On statistical properties of dynamical systems

Minicourse at ESI, Vienna, June 2008

N. Chernov, University of Alabama at Birmingham

Let (X, T, μ) be a dynamical system, i.e. $T: X \to X$ is a map preserving a probability measure μ.

The dynamics is seen as a sequence $x_0, x_1, \dots, x_n, \dots$ of points in X such that $x_n = T^n(x_0)$.

Alternatively, we may only 'see' the values $F_0, F_1, \dots, F_n, \dots$ of some function $F: X \to \mathbb{R}$, where $F_n = F(x_n)$. We call F an observable.

If the sequence $\{X_n\}$, or $\{F_n\}$, was independent (relative to the measure μ), then we could easily apply all major results of classical probability theory... But this is almost never the case in deterministic systems.

So let us assume the simplest situation with dependence: the function F only takes finitely many values, say $\{1, 2, \dots, I\}$, and the sequence $\{F_n\}$ is a Markov chain; i.e. F_n depends only on F_{n-1} but not on the previous values F_{n-m}, $m \geq 2$. This Markov chain has a stationary distribution P with components $p_i = \mu(F_0 = i)$ and its transition probability matrix Π has components $\pi_{ij} = \mu(F_1 = j / F_0 = i)$.

If $\pi_{ij} = p_j$ for all i, j we would have an independent sequence. Suppose

$$\pi_{ij} \geq \gamma p_j \qquad \text{for some} \quad \gamma > 0$$

and all i, j. Then the n-step transition probabilities $\pi_{ij}^{(n)}$ (the components of the matrix Π^n) converge to the stationary distribution exponentially fast in the following sense:

$$\mathsf{Var}(\Pi_i^{(n)}, P) \leq (1 - \gamma)^n. \tag{1}$$

Here $\Pi_i^{(n)} = (\pi_{i1}^{(n)}, \dots, \pi_{iI}^{(n)})$ is the 'image' of the ith state at time n and $P = (p_1, \dots, p_I)$ is the stationary vector; we denote by Var the distance in variation between probability vectors:

$$\mathsf{Var}(P, Q) = \frac{1}{2} \sum |p_i - q_i|.$$

The condition $\pi_{ij} \geq \gamma p_j$ may be too rigid and hard to check. A weaker condition is:

$$\frac{1}{2} \sum_k |\pi_{ik} - \pi_{jk}| \leq 1 - \gamma$$

for some $\gamma > 0$ and all i, j. This is known as Doeblin's condition, or D-condition, or Dobrushin's coefficient of ergodicity. Under this condition the convergence to the stationary vector is exponentially fast in the following sense:

$$\mathsf{Var}(\Pi_i^{(n)}, P) \leq (1-\gamma/2)^n, \tag{2}$$

which is only slightly slower than (1).

Intuitively, Doeblin's condition means that for any pair of states i and j, with probability γ the trajectories jump into the same state, and then their future evolutions become identical. That is, the two trajectories are 'coupled' and become indistinguishable in the future.

We note that if, say, 3% of the state i jumps into k and 1% of the state j jumps into k, then only 1% of each state can be 'coupled'; on the other hand, 2% of the image of i in k has to move further on, until it can be coupled with some other images of j.

Coupling method. The dynamical version of this coupling procedure was developed by Young [Y99], and then modified by Bressaud and Liverani [BL02], Dolgopyat [CD08], etc. I will describe one of its simplest versions, assuming that X is a 2D manifold and $T\colon X \to X$ is a hyperbolic map (possibly, with singularities), and μ is an SRB measure.

Let W_1 and W_2 be two unstable curves in X which carry smooth probability measures ν_1 and ν_2, respectively. Their images $T^n(W_1)$ and $T^n(W_2)$ are longer curves, possibly broken into pieces by singularities. Generally, $T^n(W_1)$ and $T^n(W_2)$ are finite or countable unions of unstable curves.

Now if two pieces $V_1 \subset T^n(W_1)$ and $V_2 \subset T^n(W_2)$ are close enough to be connected by some stable manifolds, then the points $x \in V_1$ and $y \in V_2 \cap W^s(x)$ (observe that y is the image of x on V_2 under the standard holonomy map) can be 'coupled', because their further trajectories converge exponentially fast, thus their images become practically indistinguishable.

Note that we need to 'couple' the same amount of measure, thus only a portion of the measure $T^n\nu_1$ carried by V_1 can be coupled with a portion of the measure $T^n\nu_2$ carried by V_2.

To make sense of 'portions' of measures carried by curves and individual points, we have to replace each W_i, $i = 1, 2$, with a direct product $\hat{W}_i = W_i \times [0, 1]$, and the measure ν_i with $\hat{\nu}_i = \nu_i \times m$, where m is the Lebesgue measure on $[0, 1]$.

Coupling lemma. The coupling lemma precisely says that if W_1 and W_2 are not too short, then there exists a measure preserving bijection $\Theta\colon \hat{W}_1 \to \hat{W}_2$ (the 'coupling map') and a 'coupling time' function $R\colon \hat{W}_1 \to \mathbb{N}$ such that

(i) if $\Theta(x,t) = (y,s)$, then the points $T^R(x)$ and $T^R(y)$, where $R = R(x,t)$, belong to one stable manifold, i.e. $T^R(y) \in W^s\big(T^R(x)\big)$;

(ii) we have an exponential tail bound: $\hat{\nu}_1(R > A) < C\theta^A$ for some constants $C > 0$ and $\theta < 1$, and for all $A > 1$.

The coupling lemma can be extended to families of unstable curves with probability measures defined on the entire family (with smooth conditional densities on individual curves). In this way the coupling lemma also covers the SRB measure μ, as one an foliate X by smooth unstable manifolds (on which μ has smooth densities).

The coupling lemma allows us to derive sharp bounds on correlations and multiple correlations relatively easily [CM06, Chapter 7]. Those bounds are strong enough to imply the Central Limit Theorem and the ASIP without any extra properties of the dynamical system. (That is, the CLT and ASIP are logical consequences of sharp bounds on multiple correlations.)

Young's tower. A more traditional approach to statistical properties of hyperbolic maps (with singularities) is Young's tower [Y98]. It consists of a hyperbolic structure $\Omega \subset X$ ('horseshoe') which returns onto itself (in a Markov manner) under the iterations of the map, i.e. for every $x \in \Omega$ there is a return time $R = R(x) \in \Omega$ such that $T^R(x) \in \Omega$. The Markovness of returns means that

$$T^R\big(W^s(x) \cap \Omega\big) \subset W^s\big(T^R(x)\big) \cap \Omega$$

and

$$T^R\big(W^u(x) \cap \Omega\big) \supset W^u\big(T^R(x)\big) \cap \Omega.$$

(Due to these restrictions, $R = R(x)$ is not necessarily the first return of the trajectory of the point x to Ω.)

If the return time satisfies an exponential tail bound, i.e.

$$\mu(R > A) < C\theta^A$$

for some constants $C > 0$ and $\theta < 1$, and for all $A > 1$, then the map has exponential mixing rates [Y98]. If the return time satisfies a polynomial tail bound, i.e.

$$\mu(R > A) < CA^{-a}$$

for some constants $C, a > 0$, and for all $A > 1$ then the map has polynomial mixing rates [Y99].

It is hard to compare the two approaches, the tower and the coupling method. Both are very powerful and have been successfully used in the recent studies of various hyperbolic models. Arguably, the tower is a more complicated and 'rigid' object, it is rather difficult to construct; the coupling process is somewhat more flexible and more elementary (if this word can be applied at all...). But the tower works for systems with either exponential or polynomial mixing rates. The coupling method currently applies to systems with exponential rates only.

Flows. Next we describe a modification of the coupling method for hyperbolic flows. For simplicity, let Φ^t be a hyperbolic flow on a 3D manifold X, so that (strongly) stable and unstable manifolds are one-dimensional curves.

Given two unstable curves W_1 and W_2, their images $\Phi^t(W_1)$ and $\Phi^t(W_2)$ are unions of unstable curves. However the situation is more complicated than that we have for maps: if two pieces $V_1 \subset T^n(W_1)$ and $V_2 \subset T^n(W_2)$ are close enough somewhere, they usually are not close 'all the way', as they are not parallel (they look like skew lines in space). Thus coupling will only take place in a small neighborhood where those two pieces come close to each other.

To formalize coupling, let us construct a Markov approximation. First, we represent Φ^t by a suspension flow over a 2D cross-section Y. Then we partition Y into small rectangles (hyperbolic horseshoes). Then over each rectangle $R \subset Y$ we approximate the ceiling function τ by a constant $\bar{\tau}_Y$ and divide the resulting cylinder $Y \times [0, \bar{\tau}_Y]$ into boxes of a small fixed height δ (a quantum of time). Now the flow Φ^t is conveniently represented by the time δ map, $T = \Phi^\delta$, on the space X, and the system is conveniently represented by a Markov chain.

Now consider two small boxes $B_i, B_j \subset X$ that represent two states of our Markov chain. Their images will consist of long strips that stretch along unstable manifolds and have a constant height δ. When they can be connected by a stable manifold, that manifold will shrink in the future and its images will be entirely in some box B_k. That means parts of the images of B_i and B_j will come into one box B_k, at which time they are 'coupled'.

We note that only a small portion of the images of B_i and B_j are actually coupled. But this is just enough. A direct analysis shows that the transition probabilities satisfy condition

$$\sum_k \frac{\pi_{ik}\pi_{jk}}{p_k} \geq \gamma$$

for some $\gamma > 0$ and all i, j. Even though this is much weaker than Doeblin's condition, it still implies the following control on the convergence to the stationary distribution:

$$\mathsf{Var}(\Pi_i^{(n)}, P) \leq \frac{(1-\gamma/2)^{n/3}}{\sqrt{\gamma}\, p_{\min}},$$

where $p_{\min} = \min_i p_i$, see [C98]. This is slower than the convergence (2) under Doeblin's condition, but it is sufficient, for example, to prove a stretched exponential bound on correlations for dispersing billiard flows [C07].

It would be interesting to develop a dynamical (rather than Markov-chain-style) coupling method for hyperbolic flows.

Growth lemmas. While the construction of Young's tower, and proving the coupling lemma alike, are fairly complex tasks, they have one key element in common: the fast (exponential) growth of small unstable curves (or unstable manifolds, in high dimensions). All statements that describe such growth in precise terms are called *growth lemmas*.

Let W be an unstable manifold in X. Its image $T(W)$ expands locally but may be cut into pieces by singularities. In other words, $T(W)$ may consist of several unstable manifolds of rather different shapes and sizes. All too often $T(W)$ contains small (even arbitrarily small) components. How would we specify the 'growth' of W then?

One way is to assess the ϵ-neighborhood of the boundary of $T(W)$, because it consists of points passing nearby singularities where all the troubles occur. Let W carry a probability measure ν. Define

$$Z(W) = \sup_{\varepsilon > 0} \frac{\nu\big(\text{dist}(x, \partial W) < \varepsilon\big)}{\varepsilon},$$

see [C99a]. This value characterizes the size and shape of W: higher values of $Z(W)$ indicate either a small size or an irregular shape of W (or both).

Now for $n \geq 0$ define

$$Z_n(W) = \sup_{\varepsilon > 0} \frac{\nu\big(\mathrm{dist}\big(T^n(x), \partial T^n(W)\big) < \varepsilon\big)}{\varepsilon}.$$

This value characterizes the size and shape of the components of $T^n(W)$: higher values of $Z_n(W)$ indicate either a small size or an irregular shape of typical components. We would like to see the sequence $\{Z_n(W)\}$ decreasing fast. Ideally, we would like to have

$$Z_n(W) \leq \alpha^n Z_0(W) + \beta \tag{3}$$

for some constants $\alpha \in (0,1)$ and $\beta > 0$. That would mean that $Z_n(W)$ essentially decreases exponentially fast. Such a claim would be a precise version of 'exponential growth of unstable manifolds'.

Now consider the action of T^m on a small unstable manifold W: it expands W by a factor $\geq \Lambda^m$ in all directions, where $\Lambda > 1$ is the hyperbolicity constant. In addition, singularities of T^m cut $T^m(W)$ into, say, K_m pieces. It is easy to see (cf. [C99a]) that

$$Z_m(W) \leq \Lambda^{-m} K_m Z_0(W).$$

Thus it is important that

$$K_m < \Lambda^m,$$

at least for some m. Such inequalities are called *complexity bounds*, they describe the complexity of singularities in the system. If we have a complexity bound $K_m < \Lambda^m$ for some $m \geq 1$, then

$$Z_m(W) \leq \alpha\, Z_0(W) \qquad \text{for} \qquad \alpha = \Lambda^{-m} K_m < 1.$$

In addition, we may have to (artificially) divide $T^m(W)$ into smaller pieces (for using induction), which increases $Z_m(W)$ by a constant, so we get

$$Z_m(W) \leq \alpha\, Z_0(W) + \beta$$

Then, inductively the sequence $\{Z_n(W)\}$ will decrease exponentially fast as n growth. This is one (simple) version of the growth lemma.

One-step expansion estimate. Unfortunately, things may get worse. In many hyperbolic systems, such as billiards, an unstable manifold can be cut into countably many pieces by a single iteration of T. We examine this situation for 2D maps, where W is just a curve.

Let W be an unstable curve, divided by singularities of T into pieces $W = \cup_i W_i$, which are then mapped by T into unstable curves $V_i = T(W_i)$. If we ignore distortions (to simplify the analysis), then each W_i is expanded by a factor $\Lambda_i > 1$, so that $|V_i| = \Lambda_i |W_i|$ for each $i \geq 1$. The ε-neighborhood of the endpoints of V_i is mapped by T^{-1} back onto the $\Lambda^{-1}\varepsilon$-neighborhood of the endpoints of W_i, thus in the previous terms we have

$$m\big(\operatorname{dist}\big(T(x), \partial T(W)\big) < \varepsilon\big) \simeq \sum_i 2\Lambda_i^{-1}\varepsilon,$$

where m denotes the Lebesgue measure on the curve W. On the other hand,

$$m\big(\operatorname{dist}\big(x, \partial W\big) < \varepsilon\big) \simeq 2\varepsilon.$$

Thus $Z_1(W) < Z_0(W)$ if and only if

$$\sum_i \Lambda_i^{-1} < 1.$$

Let us assume that for all sufficiently short unstable curves

$$\sum_i \Lambda_i^{-1} < 1 - \gamma \tag{4}$$

for some (small) constant $\gamma > 0$. This assumption guarantees exponential growth of unstable curves under T, because it implies (3), see [CZ05].

The condition (4) holds for many planar billiards, including Sinai's dispersing billiards (with finite and infinite horizon), see [CZ05, CM06]. We call it a *one-step expansion estimate*. We will discuss its further implications next.

Linear tail bounds. The condition (4) guarantees that for any short unstable curve W its images will become long, on the average, after $n \sim -\log|W|$ iterations of T. That is,

$$Z_n(W) \leq \text{ Const} \qquad \forall n \geq \text{ Const} \cdot \big|\log|W|\big|.$$

This is yet another version of the growth lemma. The condition $Z_n(W) <$ Const means that the ε-neighborhood of the endpoints of the components of $T^n(W)$ has measure $\mathcal{O}(\varepsilon)$.

More generally, if we pick a point $x \in X$ at random with respect to the SRB measure μ, then it lies in the ε-neighborhood of an endpoint of its unstable manifold $W^u(x)$ with probability $\mathcal{O}(\varepsilon)$, i.e.

$$\mu\big(x\colon \operatorname{dist}(x, \partial W^u(x)) < \varepsilon\big) < C\varepsilon. \tag{5}$$

The same, of course, remains true for stable manifolds $W^s(x)$. We call (5) a *linear tail bound* for stable and unstable manifolds.

The linear tail bound also holds in the following 'local' sense: given an unstable curve W, which is long enough (say, of length one), pick a point $x \in W$ with respect to the Lebesgue measure ν on W; then x lies in the ε-neighborhood of an endpoint of its stable manifold $W^s(x)$ with probability $\mathcal{O}(\varepsilon)$, i.e.

$$\nu\big(x\colon \operatorname{dist}(x, \partial W^s(x)) < \varepsilon\big) < C\varepsilon. \tag{6}$$

This follows from the growth lemmas, see [CM06, Chapter 5].

The linear tail bounds (5)–(6), i.e. in both local and global senses, are used not only in the studies of statistical properties – they are essential in the classical (Sinai's) proof of ergodicity for dispersing billiards [S70], which was later extended to many other maps with singularities.

Power-law tail bounds. On the other hand, the general Katok-Strelcyn theory [KS86] of hyperbolic maps with singularities assumes that the ε-neighborhood of singularities of the map T has measure $\mathcal{O}(\varepsilon^a)$ for some $a > 0$. This is enough for the existence of stable and unstable manifolds (and their absolute continuity).

In the Katok-Strelcyn theory, if one picks a point $x \in X$ at random, then it lies in the ε-neighborhood of an endpoint of its stable or unstable manifold with probability $\mathcal{O}(\varepsilon^a)$, i.e.

$$\mu\big(x\colon \operatorname{dist}(x, \partial W^u(x)) < \varepsilon\big) < C\varepsilon^a. \tag{7}$$

We call (7) a *power-law tail bound* for stable and unstable manifolds. It is more general than the linear tail bound (5), which corresponds to $a = 1$.

To summarize, most existing techniques for investigating statistical properties (or proving ergodicity for that matter) are developed for systems with linear tail bounds (5)–(6). There are no general methods for hyperbolic maps with power-law tail bounds... We are trying to design one such method next.

Improved one-step expansion estimate. First let us better understand the linear tail bound (5). We begin with the 2D case, i.e. $\dim X = 2$. In that case X can be foliated by unstable manifolds (curves) with smooth conditional densities on them.

More generally, suppose $\{W_\alpha\}$ is a (countable or uncountable) family of unstable curves in X with a probability measure ν on that family, which has smooth conditional density on each W_α. For each x we denote by $W(x) = W_\alpha$ the curve in that family that contains x.

A general version of the (global) linear tail bound would say that

$$\nu\big(\mathrm{dist}(x, \partial W(x)) < \varepsilon\big) \leq C\varepsilon. \tag{8}$$

It clearly implies that the total mass of unstable curves shorter than ε in that family is $\mathcal{O}(\varepsilon)$, i.e.

$$\nu\big(|W_\alpha| < \varepsilon\big) \leq C\varepsilon, \tag{9}$$

but (8) is stronger than (9). The following simple lemma shows what (8) really is:

Lemma 1. *The linear tail bound* (8) *is equivalent to*

$$\int \frac{1}{|W_\alpha|}\, d\nu < \infty.$$

Thus, the average reciprocal length of unstable curves (in the given family) plays a crucial role. This motivates the following definitions.

For an unstable curve $W \subset X$ with a probability measure ν on it, we put

$$Z^{(1)}(W) = Z_0^{(1)}(W) = |W|^{-1}$$

and

$$Z_n^{(1)}(W) = \int |V_n(x)|^{-1}\, d\nu(x),$$

where $V_n(x)$ denotes the component of $T^n(W)$ that contains the point $T^n(x)$. Now it is easy to see that our one-step expansion estimate (4) exactly means that

$$Z_1^{(1)}(W) < (1-\gamma)\, Z_0^{(1)}(W)$$

for some constant $\gamma > 0$ for all sufficiently short unstable curves. Inductively this implies $\forall n \geq 1$

$$Z_n^{(1)}(W) < \alpha^n Z_0^{(1)}(W) + \beta$$

for some constants $\alpha < 1$ and $\beta > 0$, i.e. again we see an exponential decrease of $Z_n^{(1)}(W)$, which translates into an exponential growth of unstable curves.

Now suppose we want to deal with a family of unstable curves with a power-law tail bound, i.e. satisfying

$$\nu\big(\mathrm{dist}(x, \partial W(x)) < \varepsilon\big) \le C\varepsilon^a \tag{10}$$

for some $0 < a < 1$. It clearly implies that the total mass of unstable curves shorter than ε is $\mathcal{O}(\varepsilon^a)$, i.e.

$$\nu\big(|W_\alpha| < \varepsilon\big) \le C\varepsilon^a, \tag{11}$$

but again (10) is stronger than (11). The following simple lemma clarifies what (10) really means:

Lemma 2. *The power-law tail bound* (8) *means that for any $p < a$*

$$\int \frac{1}{|W_\alpha|^p}\, d\nu < \infty.$$

(Note: this is not necessarily true for $p = a$.)

It is now natural to modify our previous definitions of Z's as

$$Z^{(p)}(W) = Z_0^{(p)}(W) = |W|^{-p}$$

and

$$Z_n^{(p)}(W) = \int |V_n(x)|^{-p}\, d\nu(x).$$

Then the one-step expansion estimate can be modified as follows:

$$Z_1^{(p)}(W) < (1-\gamma)\, Z_0^{(p)}(W)$$

for some $\gamma > 0$ for all sufficiently short unstable curves W. Again, inductively this implies $\forall n \ge 1$

$$Z_n^{(p)}(W) < \alpha^n Z_0^{(p)}(W) + \beta \tag{12}$$

for some constants $0 < \alpha < 1$ and $\beta > 0$, i.e. again we see an exponential growth of unstable curves, but in a new sense, with $p < 1$ instead of $p = 1$. This is our new version of the growth lemma.

Our new one-step expansion estimate can be also stated in terms of expansion factors Λ_i, in a way similar to (4): for every short unstable curve divided by singularities into pieces $W_i \subset W$, which are then expanded under the map T by factors $\Lambda_i > 1$, we must have

$$\sum_i (V_i/W)^{1-p} \Lambda_i^{-1} < 1 - \gamma \tag{13}$$

for some small constant $\gamma > 0$. Compare this to (4): the new estimate involves an extra "weight" factor $(V_i/W)^{1-p}$ which may help reduce the sum of the series when V_i's are small, i.e. exactly in 'troublesome' situations. We present an example below.

We emphasize that our new one-step expansion estimate (13) and our new growth lemma (12) are, on the one hand, consistent with power-law tail bounds (they do not require linear tail bounds), and, on the other hand, they still can be used to prove the coupling lemma, to construct Young's tower with an exponential tail bound, and ultimately – establish exponential decay of correlations.

In fact we found two classes of hyperbolic billiards to which the standard one-step expansion estimate (4) does not apply but the new estimate (13) applies. These are Bunimovich flowers with arcs greater than semicircle and a modified stadium with a non-C^1 boundary (bounded by two parallel line segments and two arcs shorter than semicircle), see below.

Higher-dimensional one-step expansion estimate. The above theory, including the estimate (13), can be easily generalized to higher-dimensional maps. We just redefine our Z-values as follows:

$$Z^{(p)}(W) = Z_0^{(p)}(W) = \int_W |\mathrm{dist}(x, \partial W)|^{-p} \, d\nu$$

and

$$Z_n^{(p)}(W) = \int |\mathrm{dist}(T^n(x), \partial V_n(x)|^{-p} \, d\nu(x),$$

where $V_n(x)$ denotes the component of $T^n(W)$ containing the point $T^n(x)$. Then the one-step expansion estimate can be modified as follows:

$$Z_1^{(p)}(W) < (1-\gamma)\, Z_0^{(p)}(W) \tag{14}$$

for some $\gamma > 0$ for all sufficiently short unstable manifolds W. The rest of the theory will not change.

It would be interesting to find examples of multidimensional hyperbolic billiards where the linear tail bound fails, but the new one-step expansion estimate (14) holds. It appears that for the 3D Lorentz gas with finite horizon the linear tail bound still holds [BT08], but even in that case it may be easier to use (14) instead.

Systems with slow mixing rates. Suppose we have a non-uniformly hyperbolic map (with singularities). This basically means that for any $n \geq 1$ and $\delta > 0$

$$\inf_{W\colon |W|<\delta} \frac{|T^n(W)|}{|W|} = 1,$$

where $|\cdot|$ denotes the length of unstable curves and their images. In mechanical models, it is usually easy to find (and 'localize') the spots in the phase space X where the expansion of unstable curves slows down.

As a first example, consider a *pinball machine in a box*: a billiard in a rectangle with an immovable round obstacle inside. This is a semi-dispersing billiard table, known to be hyperbolic and ergodic. It is easy to see that collisions with the flat sides of the box ('polygonal part' of the boundary) do not cause expansion of unstable curves, while collisions with the round obstacle ('dispersing part' of the boundary) cause strong expansion. Thus the collision map $T\colon X \to X$ acting on the entire collision space X is non-uniformly hyperbolic.

Let X_* be the *reduced* collision space consisting of collisions with the round obstacle only, and $T_*\colon X_* \to X_*$ be the induced first return map. In fact, T_* is the collision map for the corresponding dispersing billiard on a torus (i.e., a Lorentz gas) obtained by a standard unfolding procedure, thus T_* has exponential mixing rates.

The analysis of the mixing rates of the original map $T\colon X \to X$ proceeds in three major steps:

Step 1: Construct Young's tower $\Omega \subset X_*$ with exponential tail bounds for the reduced, strongly hyperbolic map T_*. At this step a crucial part is to verify that unstable curves grow exponentially fast.

In our example T_* is the collision map for the Lorentz gas without horizon. There, the worst case for a short unstable curve W is to be cut into countably many pieces W_i which are then expanded by factors $\Lambda_i \sim i^{3/2}$ where $i \geq i_0$, and i_0 depends on $|W|$ so that $i_0 \to \infty$ as $|W| \to 0$. Thus we have

$$\sum_{\imath \geq i_0} \Lambda_i^{-1} = \sum_{\imath \geq i_0} i^{-3/2} \sim i_0^{-1/2}$$

and for sufficiently short unstable curves $i_0^{-1/2} \ll 1$, and we obtain our one-step expansion estimate (4).

Step 2: Find a 'rough' tail bound for the returns to Ω under the original map T. Points $x \in \Omega$ return to Ω slower under T than under T_* because they spend some extra time outside X_*.

First we need to find a tail bound for the return times to X_*, which is not hard as it does not involve the tower Ω. Let $S(x)$ for $x \in X_*$ be the return time to X_*. For the Lorentz gas example,

$$\mu(S = k) \sim k^{-3}, \qquad \text{hence} \qquad \mu(S > k) \sim k^{-2}. \tag{15}$$

Now suppose for $x \in \Omega$ we have $R(x) > n$, i.e. its trajectory fails to return to Ω within n iterations of T. Let m denote the number of times its trajectory visits X_* during the first n iterations (of course, $0 \leq m \leq n$). Let $L_1, \dots, L_{m+1}$ denote the intervals between successive returns to X_* (of course, $\sum L_i = n$), and let $L_{\max} = \max\{L_1, \dots, L_{m+1}\}$ denote the maximal interval between the returns to X_*.

Now we estimate the measure of such points x. From the exponential tail bound for the map T_* we have $\mu(\text{such } x\text{'s}) \leq C\theta^m$. On the other hand, $\mu(\text{such } x\text{'s}) \leq CnL_{\max}^{-2}$ from the estimate (15) on S.

The worst case scenario is $m \sim \log n$ and $L_{\max} \sim n/(\log n)$, which readily gives the following 'rough' estimate:

$$\mu(R > n) \sim C(\log n)^2/n. \tag{16}$$

Due to Young's general result [Y99] the correlations for the map T are bounded by $C(\log n)^2/n$.

But this is only a crude estimate. In the next step we will get rid of the logarithmic factor.

Step 3: Refine the tail bound (16) and eliminate the logarithmic factor. The estimate (16) would be sharp only if all the intervals $L_1, \dots, L_{m+1}$ had length of the same order as $L_{\max}$, i.e. of order $n/(\log n)$. This is however very unlikely due to the following features of the dynamics.

If a trajectory spends k iterations of T outside X_*, then after its return to X_* it will typically spend $\sim \sqrt{k}$ iterations outside X_*, then after the next return $\sim k^{1/4}$ iterations, etc. Thus typical points escape from the process of long returns to X_* very fast, we call this feature *fast escape*.

It is now a technical task to estimate the measure of 'non-typical' points that escape slowly, we omit details, see [CZ08, SV08]. This allows us to eliminate the logarithmic factor:

$$\mu(R > n) \leq C/n. \tag{17}$$

Due to Young's general result [Y99] the correlations for the map T are bounded by C/n.

Stadium. The next example is the famous Bunimovich stadium – a billiard table bounded by two semicircles connected by two parallel line segments. This domain is convex and has a C^1 boundary.

There are two types of trajectories here that cause slow expansion in the unstable direction. One is trajectories bouncing between the parallel sides, they are similar to those in the previously discussed pinball machine. The other is trajectories moving slowly ('sliding') along a semicircle.

The reduced space X_* is made by the *first* collisions with the semicircles; i.e. we remove from X all the collisions with the parallel sides, and from each series of successive collisions with one of the two semicircles we remove all but the very first one. Next we follow the three major steps described above.

At Step 1 we need to verify the one-step expansion estimate (4) for the return map $T\colon X \to X$. It may happen now that an unstable curve W is cut into pieces W_i which are then expanded by factors $\Lambda_i \sim i$, which is slower than $i^{3/2}$ in the previous example. If there were infinitely many such pieces, the series $\sum \Lambda_i^{-1}$ would diverge and the crucial estimate (4) would fail. Fortunately, the number of pieces is always finite, and a detailed calculation (see [M04, CZ05]) shows that

$$\sum \Lambda_i^{-1} \sim \frac{3}{8} \log 9 < 1,$$

which is just enough for (4) to be true. It seemed like a miracle at the time when the mixing rates for the stadium were derived first by Markarian [M04], but we know now that the new estimate (13) would work anyway – it is more flexible and less sensitive to the details of the particular system, see below.

Step 2 is rather simple and leads to the same 'rough' tail bound as in (16).

The stadium presents a new difficulty at Step 3. Its trajectories escape from long returns to X_* *slowly*. Precisely, if a trajectory spends $k = k_1$ iterations of T outside X_*, then after its return to X_* it will spend k_2 iterations outside X_*, where k_2 is between $k_1/3$ and $3k_1$ (that is, k_2 is of order k_1). The sequence $k_1, k_2, \ldots$ behaves as a Markov chain, with transition probabilities given by $\pi_{k_1 k_2} \approx \frac{3k_1}{8k_2^2}$, see [BG06]. In order to eliminate the logarithmic factor we have to carefully analyze that Markov chain and employ the large deviation theorem [CZ08].

The resulting estimate for the decay of correlations in the stadium is $\sim 1/n$. Bálint and Gouëzel [BG06] proved that this estimate is sharp, i.e. they showed that correlations for some observable do decay as $\sim 1/n$. It seems like similar results could be obtained for the other billiards with slow mixing rates described here, but this is yet to be done.

The decay rate $\sim 1/n$ is too slow for the central limit theorem to hold. In fact Bálint and Gouëzel [BG06] show that for generic observables the classical CLT fails; instead they prove a non-classical version of the CLT, with the scaling factor $\sqrt{n \log n}$, instead of $\sqrt{n}$. It would be interesting to obtain similar results for the other billiards with slow mixing rates described here.

Other models. In a similar way, correlations have been estimated for several more classes of hyperbolic billiards:

- Dispersing tables with cusps (first introduced by Machta [M83]);
- Skewed stadia (or drivebelt regions, aka squashes);
- Flower-like tables.

This was done in [CZ05, CZ08]. Studies of multidimensional dispersing billiards (Lorentz gases with finite horizon) have been conducted by Bálint and Toth [BT08].

For two particular models we have to use the new estimate (13). These are Bunimovich flowers with arcs greater than semicircle and a modified stadium with a non-C^1 boundary (bounded by two parallel line segments and two arcs shorter than semicircle).

In both cases, a short unstable curve W can be cut into infinitely many pieces W_i which are then expanded by factors $\Lambda_i \sim i$. Thus the series $\sum \Lambda_i^{-1}$ diverges, hence the standard one-step expansion estimate (4) fails.

The new estimate (13) holds because $V_i \sim 1/i$ hence

$$\sum (V_i/W)^{1-p} \Lambda_i^{-1} \sim \sum i^{-2+p} < \infty$$

because $p < 1$. A more careful calculation shows that the sum of the above series is $\ll 1$ for short enough unstable curves W, which completes the verification of (13).

References

[BG06] P. Bálint and S. Gouëzel, *Limit theorems in the stadium billiard*, Comm. Math. Phys. **263** (2006), 461–512.

[BT08] Bálint, P., and Toth, I. P., *Exponential decay of correlations in multi-dimensional dispersing billiards*, preprint.

[BL02] Bressaud, X., and Liverani, C., *Anosov diffeomorphism and coupling*, Ergod. Th. Dynam. Syst., **22** (2002), 129–152.

[C99a] Chernov, N., *Statistical properties of piecewise smooth hyperbolic systems in high dimensions*, Discr. Cont. Dynam. Syst. **5** (1999), 425–448.

[C99b] Chernov, N., *Decay of correlations in dispersing billiards*, J. Statist. Phys. **94** (1999), 513–556.

[C98] Chernov, N., *Markov approximations and decay of correlations for Anosov flows,* Ann. Math. **147** (1998), 269–324.

[C07] Chernov, N., *A stretched exponential bound on time correlations for billiard flows,* J. Statist. Phys. **127** (2007), 21–50.

[CD08] Chernov, N., and Dolgopyat, D., *Brownian Motion - I*, Memoirs of AMS, to appear.

[CM06] Chernov, N., and Markarian, R., *Chaotic Billiards*, Mathematical Surveys and Monographs, **127**, AMS, Providence, RI, 2006.

[CM07] Chernov, N., and Markarian, R., *Dispersing billiards with cusps: slow decay of correlations*, Commun. Math. Phys., **270** (2007), 727-758.

[CZ05] Chernov, N., and Zhang, H.-K., *Billiards with polynomial mixing rates.* Nonlineartity, **4** (2005), 1527–1553.

[CZ08] Chernov, N., and Zhang, H.-K., *Improved estimates for correlations in billiards*, Commun. Math. Phys., **277** (2008), 305–321.

[KS86] Katok, A., and Strelcyn, J.-M. (with the collaboration of F. Ledrappier & F. Przytycki), *Invariant manifolds, entropy and billiards; smooth maps with singularities*, Lect. Notes Math., **1222**, Springer, New York (1986).

[M83] J. Machta, *Power law decay of correlations in a billiard problem*, J. Statist. Phys. **32** (1983), 555–564.

[M04] Markarian, R., *Billiards with polynomial decay of correlations*, Erg. Th. Dynam. Syst. **24** (2004), 177–197.

[S70] Sinai, Ya. G., *Dynamical systems with elastic reflections.* Ergodic properties of diepersing billiards, Russian Math. Surveys, **25** (1970), 137–189.

[SV08] D. Szász and T. Varjú, *Limit laws and recurrence for the planar Lorentz process with infinite horizon*, J. Statist. Phys.

[Y98] Young, L.-S., *Statistical properties of systems with some hyperbolicity including certain billiards.* Ann. Math. **147** (1998), 585–650.

[Y99] Young, L.-S., *Recurrence times and rates of mixing*, Israel J. Math. **110** (1999), 153–188.