Ergodicity and correlation decay in billiards

PhD thesis

Imre Péter Tóth

Institute of Mathematics, Technical University of Budapest

Advisor:
Domokos Szász

2005

Ezen értekezés bírálatai és a védésről készült jegyzőkönyv a későbbiekben a Budapesti Műszaki és Gazdaságtudományi Egyetem Természettudományi Karának Dékáni Hivatalában elérhető.

Alulírott Tóth Imre Péter kijelentem, hogy ezt a doktori értekezést magam készítettem és abban csak a megadott forrásokat használtam fel. Minden olyan részt, amelyet szó szerint, vagy azonos tartalomban, de átfogalmazva más forrásból átvettem, egyértelműen, a forrás megadásával megjelöltem.

. . . .
Tóth Imre Péter

Contents

Chapter 1

Introduction

The theory of billiards is one of the fields of Dynamical Systems theory that is most directly motivated by Statistical Physics. In the Dynamical Systems literature, a billiard is the dynamical system describing the motion of a single point particle which moves freely in a container Q – which can have a complicated shape – and bounces back elastically when reaching the walls. However, this also covers the motion of several bodies with elastic collisions: instead of many bodies in a few dimensions, one considers a single point particle in many dimensions. The condition that the bodies cannot overlap is translated into the condition that the point particle cannot enter certain regions – that is, into the shape of the container Q. Actually, this is where the name comes from: the motion of billiard balls on a table can be described by such a dynamical system if we neglect friction and the loss of energy at collisions. This example also demonstrates well the key feature of the most studied billiard systems: the motion of the system is highly unpredictable if the initial conditions are not known precisely. In exactly which sense this motion is unpredictable, is the basic question of billiard theory.

One important example of billiard systems with strong physical motivation is the class of hard ball systems. They are simple mathematical models for ideal gases: atoms are modeled by balls which bounce elastically and momentarily. Their history is strongly connected to the Boltzmann Ergodic Hypothesis, a fundamental assumption of equilibrium Statistical Physics. This hypothesis states that **in any large physical system in equilibrium, the time average of any measurable quantity will tend, as time goes to infinity, to the "ensemble average", that is, the average value of the measurable on many realizations of the system, chosen according to the equilibrium distribution.** This equality of the averages is called the *ergodicity* of the system.

At the present state of science, this statement is far too general for a rigorous mathematical discussion. The first problem is with the word "large": it is unclear how large the system is excepted to be, or in what sense a limit should be taken as the size goes to infinity. The second problem is with the "equilibrium distribution": the theory of dynamical systems has many widely investigated open questions concerning in what systems, and in which sense, an equilibrium distribution exists at all. However, the main problem is with the phrase "any physical system": we don't have any unified way of handling the vast variety of physical systems one can imagine: we are forced to look at them "one at a time".

In 1963 Sinai [37] proposed to study the ergodicity of hard ball systems – systems of hard balls moving on a flat torus (thus colliding with each other only). These systems are complicated enough to show many statistical properties of the ideal gases they are used to model, yet simple enough to be accessible with tools of rigorous Mathematics. The existence of a unique equilibrium distribution is no problem. It was Sinai's great discovery that the problem of "what is a large system" is also easy to handle: he conjectured that such a system is ergodic for *every* possible number of balls greater than one. Ergodicity is, of course, understood on the subset of the phase space defined by the trivial invariants of motion (energy, momentum and centre of mass). This conjecture is the *Boltzmann-Sinai Ergodic Hypothesis*. Sinai proved the conjecture in 1970 for the simplest case of two balls in two dimensions [38]. Since then, much work has been done (by, among others, Sinai, Bunimovich, Chernov, Krámli, Simányi and Szász), and now the conjecture is near to being proven in general. This thesis also contains one contribution: in Chapter 4 we give an improved proof of the local ergodicity theorem for semi-dispersing billiards (including hard ball systems). This improvement is necessary in view of the results of Chapter 3 on the structure of the singularity set of these systems.

Another important and physically motivated class of billiard systems is the class of Lorentz processes. A Lorentz Process is a simple deterministic model of Brownian motion: a single particle moves in a periodic array of strictly convex obstacles (called scatterers), and bounces back elastically when reaching one. This could be a model e.g. of an electron moving in a crystal lattice. Many results show that the motion of such a particle resembles in many aspects a random walk (and a Wiener-process), but many questions are still open, especially if the system is high (more than two) dimensional.

Convention 1.1. *Throughout this thesis, when concerning dimensions, "high" or "multi" mean "more than two".*

Part of this thesis is motivated by one of these open questions, the decay of correlations in (high dimensional) Lorentz processes.

In Dynamical Systems, there are many notions of "chaotic" or "statistical" behaviour. In this thesis we mainly look at two of these: *ergodicity* (in Chapter 4) and *exponential decay of correlations* (EDC) (in Chapter 5). A third notion, *hyperbolicity* is used excessively as a tool, and a fourth, the *central limit theorem* (CLT) is obtained as a byproduct of EDC.

For a better understanding of the work, we give the definitions we use, of these notions here:

Definition 1.2. *The dynamical system (M, T, μ) is ergodic if every μ-measurable T-invariant set $A \subset M$ has either $\mu(A) = 1$ or $\mu(A) = 0$.*

Ergodic components of the dynamical system (M, T, μ) are the atoms of the σ-algebra of T-invariant μ-measurable subsets of M. Obviously, (M, T, μ) is ergodic if and only if it has an ergodic component with measure 1.

Definition 1.3. *We say that the dynamical system (M, T, μ) has exponential decay of correlations (EDC) , if for every pair of Hölder-continuous functions $f, g : M \to \mathbb{R}$ there exist constants $C < \infty$ and $a > 0$ such that for every $n \in \mathbb{N}$*

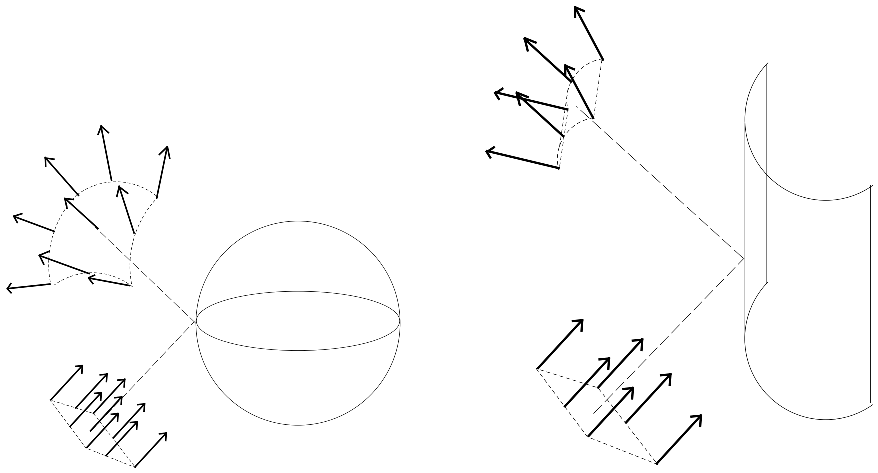
$$\left| \int_M f(x)g(T^n x)d\mu(x) - \int_M f(x)d\mu(x) \int_M g(T^n x)d\mu(x) \right| \leq Ce^{-an}.$$

Definition 1.4. *We say that (M, T, μ) satisfies the central limit theorem (CLT) (for Hölder continuous functions) if for every $\eta > 0$ and every Hölder-continuous function $f : M \to \mathbb{R}$ with $\int f \, d\mu = 0$, there exists a $\sigma_f \geq 0$ such that*

$$\frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} f \circ T^i \xrightarrow{\text{distr}} \mathcal{N}(0, \sigma_f)$$

where $\mathcal{N}(0, \sigma_f)$ is the Gaussian distribution with variance σ_f^2.

Figure 1.1 demonstrates the key feature of billiards which is responsible for their chaotic behaviour. This is the "scattering" mechanism which causes (certain) nearby trajectories to divert from each-other. We show two well-distinguished cases. One is a *dispersing* billiard, meaning that scatterers are strictly convex, and so parallel trajectories can be diverted (from each other) in all directions through a single collision. The other is only *semi-dispersing*: scatterers are not strictly convex, resulting in a "less effective" scattering.



(a) dispersing

(b) semi-dispersing

Figure 1.1: scattering in dispersing and semi-dispersing billiards

1.1 History

Mathematical studies of billiards have begun over thirty years ago. Ya. Sinai in his seminal paper of 1970 [38] described the first large class of billiards with truly chaotic behaviour – with nonzero Lyapunov exponents, positive entropy, enjoying ergodicity, mixing, and (as was later discovered by G. Gallavotti and D. Ornstein [22]) the Bernoulli property. Sinai billiards are defined in two dimensions ($d = 2$), i.e. the container $Q \subset \mathbb{R}^2$ or $Q \subset \mathbb{T}^2$, and the boundary of Q must be concave (i.e., convex outward Q), similarly to the Lorentz process (where the scatterers are convex). Due to the geometric concavity, the boundary

∂Q scatters (disperses) bundles of geodesic lines falling upon it, see Figure 1.1. For this reason, Sinai billiards are said to be *dispersing*.

Lorentz processes in two dimensions have been studied very thoroughly since 1970. Many fine ergodic and statistical properties have been established by various researchers, including P. Bleher, L. Bunimovich, N. Chernov, J. Conze, C. Dettmann, G. Gallavotti, A. Krámli, J. Lebowitz, D. Ornstein, K. Schmidt, N. Simányi, Ya. Sinai, D. Szász, and others (see the references). The latest major result for this model (the exponential decay of correlations) was obtained by L.-S. Young [43]. The success in these studies had significant impact on modern statistical mechanics. The methods and ideas originally developed for the planar Lorentz process were applied to many other classes of physical models – see recent reviews by Cohen, Gallavotti, Ruelle and Young [21, 32, 44].

On the other hand, the progress in the study of the multi-dimensional Lorentz process (where $d > 2$) has been much slower and somewhat controversial. Relatively few papers were published covering specifically the case $d > 2$, especially in contrast to the big number of works on the 2-D case. Furthermore, the arguments in the published articles were usually rather sketchy, as in Chernov's paper [11]. It was commonly assumed that the geometric properties of the multi-dimensional Lorentz process were essentially similar to those of the 2-D system, and so the basic methods of study should be extended from 2-D to any dimension at little cost. Thus, the authors rarely elaborated on details.

Recent discoveries proved that spatial dispersing billiards are *very much* different from planar ones. Bunimovich and Reháček [10] studies of *astigmatism* (the role of which was already stressed by Wojtkowski in [42]), in the somewhat different context of focusing billiards, emphasized the known fact that billiard trajectories that are initially parallel, may focus very rapidly in one plane and very slowly in the orthogonal planes. Astigmatism is unique to 3-D (and higher dimensional) billiards, it cannot occur on a plane.

In Chapter 5 we will also consider a generalization of the above billiard systems, the so-called "soft billiard". The history of soft billiards is as old as that of the hard ones. After many works of others, Donnay and Liverani in [17] have proven ergodicity of 2-dimensional systems under nearly necessary conditions. Exponential decay is discussed for some of these systems in this thesis. For the higher dimensional case, to my knowledge, the first result is the proof of hyperbolicity in [7], not yet appeared. Some more details about the history of soft billiards are given in Section 5.1.

1.2 Structure of the thesis, results

In proofs of ergodicity of billiards – especially in higher dimensions – the key tool is the "Fundamental Theorem" or "local ergodicity theorem", which was first proven by Chernov and Sinai in [13]. Later several modifications were presented by others. The version of Krámli, Simányi and Szász ([28]) is described in detail in Chapter 4.

In proofs of exponential decay of correlations, the key tool is the tower construction of Young [43]. We don't give a description of this method here, instead, we use a theorem of Chernov from [12] which lists explicit geometrical properties of the system that imply the applicability of Young's construction, and thus EDC and CLT. In Chapter 5, the conditions of Chernov's theorem are checked for a modification of the Lorentz process. A precise formulation of Chernov's theorem is given in the Appendix.

In Chapter 2 we give a precise definition of the billiard dynamical system, and introduce some basic notions of the theory.

In Chapters 3 and 4 we consider multi-dimensional *dispersing* billiards. We show that multi-dimensionality has great effect on the dynamics in the dispersing case as well – the system requires much more elaborated study than the 2D process. As already mentioned, much of this work was motivated by the problem of EDC in high dimensional Lorentz processes. This property is still *not* proven. Chapter 3 points out an important difficulty in the application of Chernov's (and, actually, also Young's) method. This is the main result of that chapter, **Theorem 3.1 about a "pathological" behaviour of singularity manifolds, meaning the blow-up of curvatures at certain submanifolds**. Actually, singularity manifolds are, in these pathologies, not even differentiable. Indeed, as it will be shown in Chapter 3, the unit normal vector to the singularity manifold has different directional limits at the pathological points – the geometry is pretty much like the classical Whitney umbrella $x^2z = y^2$ in $\mathbb{R}^3$. This phenomenon is again unique to billiards in dimension $d \geq 3$.

This discovery calls for an immediate reconsideration of the local ergodicity theorem, this is done in Chapter 4. An important new condition, the "local Lipschitz decomposability of singularities" is added, in order to counterbalance the pathological behaviour of curvatures. This property is conjectured to be true in general, however, we are only able to prove it for the case of algebraic billiards – when scatterers are defined by algebraic equations. The main results of the chapter are **theorems 4.16 and 4.18, which state local ergodicity when singularities are locally Lipschitz decomposable**, and **Corollary 4.39, which states local ergodicity for algebraic billiards**. Fortunately, almost all important examples we know of, are algebraic.

Due to the pathologies described in Chapter 3, the proof of exponential correlation decay for high dimensional billiards is still very hard. This is also indicated by the fact that in the 6 years that have elapsed by the appearence of [43], noone was able to apply the method for this case. Therefore, in Chapter 5 we return to the problem of EDC in a modified approach: Instead of high dimensions, we look at a modification of the two-dimensional billiard system, a so-called *soft billiard*, where the collision with the scatterers is not momentary, but, instead, is a motion in some potential. See the details there. The main result of that chapter is **Theorem 5.9, which states exponential decay of correlations (and the central limit theorem) for a class of soft billiard systems**. Proof of EDC for such a system requires extra work in two dimensions. However, there is hope (see Chapter 6) that in high dimensions such a modified system could be handled easier than the original: the pathological behaviour of singularities from Chapter 3 does not show up. This research is in progress: at the moment, hyperbolicity is proven for a class of multi-dimensional soft billiards [7]. Discussion of this is already out of the scope of this thesis.

It's useful to observe that the proof of EDC for hard ball systems is yet further away. That is, in some sense, hard ball systems are more difficult than the multi-dimensional Lorentz process. The main difference is that a Lorentz-process – even in high dimension – is a *dispersing* billiard, while hard ball systems (of more than two balls) are only *semi-dispersing* (see Figure 1.1).

All of this work is based on the material that has appeared in the papers [4], [3], [6] and [5]. Unfortunately, the notation used in these papers often had to be changed considerably, in order to keep notation consistent within this thesis.

Chapter 2

Preliminaries on billiards

2.1 The dynamical system

In this section we describe the traditional 'hard collision' billiard systems, which are called billiards in the literature. This notion will be used through Chapters 3 and 4. In Chapter 5 we will discuss similar, yet different systems called 'soft billiards'. There will be given new definitions of the dynamical system.

A billiard is a dynamical system describing the motion of a point particle in a connected, compact domain $Q \subset \mathbb{T}^d$. In general, the boundary of the domain, ∂Q is assumed to be piecewise C^3-smooth (see Remark 2.10). However, for the results of Section 4.4 to hold, we will need to impose the further restriction of algebraicity on the billiard. Inside Q the motion is uniform while the reflection at the boundary ∂Q is elastic. As the absolute value of the velocity is a first integral of motion, we fix this value to be 1. So the phase space of the billiard flow is the set Q of possible configuration points, each point equipped with possible unit velocity vectors. The set of possible velocity vectors is typically $\mathbf{S}^{d-1}$, however, at ∂Q, where momentary collisions take place, it is only a hemisphere – pre-collision (incoming) and post-collision (outgoing) velocities have to be identified, so that for every moment in time, the state of the particle can be uniquely defined. So, the phase space is $\mathcal{M} = Q \times \mathbf{S}^{d-1}/\sim$, where $\sim$ is the equivalence of appropriate incoming and outgoing velocities. Still we will denote a phase point x as $x = (q, v)$ with $q \in Q$ and $v \in \mathbb{R}^d$, $|v| = 1$. The dynamics S^t is the function from $\mathcal{M}$ to $\mathcal{M}$ that tells in which point $S^t(x)$ of the phase space the system will be at time t, if it is in x at time zero. The usual invariant measure for this system is the Liouville probability measure μ_0 on $\mathcal{M}$, which is essentially the product of the Lebesgue measures, i.e. $d\mu_0 = \text{const.}d\mathcal{L}_0 d\mathcal{H}$ where $\mathcal{L}_0$ is the Lebesgue measure on Q and $\mathcal{H}$ is the Riemannian measure on $\mathbf{S}^{d-1}$.

Definition 2.1. *The resulting dynamical system $(\mathcal{M}, \{S^t, t \in \mathbb{R}\}, \mu_0)$ is the billiard flow.*

Instead of the flow, one often considers a discrete time version of the system by looking only at collision moments. In other words, the boundary ∂Q defines a natural cross-section for the billiard flow. Consider namely

$$M = \{(q, v) \in \mathcal{M} \mid q \in \partial Q\}.$$

With slight abuse of notation, one can write $M = \partial\mathcal{M}$. (Strictly speaking, $\mathcal{M}$ has no boundary after the identification of incoming and outgoing velocities.)

Notation. *$n(q)$ will denote the unit normal vector of ∂Q at q, pointing inward Q.*

φ will denote the collision angle, which is the angle of $n(q)$ and the outgoing velocity v, such that $\cos\varphi = \langle v, n(q)\rangle$. When $d = 2$, it will be convenient to consider φ as a signed angle, $\varphi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$.

We will always use outgoing velocities, such that

$$M = \{(q, v) \in \mathcal{M} \mid q \in \partial Q,\ \langle n(q), v\rangle \geq 0\} = \partial Q \times \mathbf{S}_+^{d-1},$$

where $+$ means that we only take into account the hemisphere of the outgoing velocities.

The discrete time dynamics T will denote the first return map to M: For any $x \in \mathcal{M}$ we set $t^+(x) := \inf\{t > 0 \mid S^t x \in M\}$, and $T^+x := S^{t^+(x)}x$ (of course, $T^+ : \mathcal{M} \to M$). Then the Poincaré section map $T : M \to M$ is defined as $Tx := T^+x$ for $x \in M$.

The natural invariant measure for this map, which is the 'projection' of the flow-invariant measure μ_0 onto M by the dynamics, is again absolutely continuous with respect to the product of Lebesgue measures:

$$d\mu = \text{const.}\langle n(q), v\rangle d\mathcal{L}d\mathcal{H} = \text{const.}\cos\varphi d\mathcal{L}d\mathcal{H},$$

where $\mathcal{L}$ is the Riemannian measure on ∂Q and $\mathcal{H}$ is still the Riemannian measure on $\mathbf{S}^{d-1}$.

Definition 2.2. *The resulting discrete time dynamical system (M, T, μ) is the billiard map.*

Remark 2.3. *If ∂Q is only piecewise smooth (that is, there billiard has 'cornar points'), a precise definition of the phase space and the dynamics requires more care. For example, at certain points the set of possible outgoing velocities is not a hemisphere – it may be less or more, depending on the way the boundary components intersect. Also, $n(q)$ is not unique, and it is not clear what "elastic collision" should mean. Thus the dynamics is not uniquely defined (see also Remark 2.9). In this work we don't investigate these problems – the reader may see the literature, e.g. [34].*

We will work with semi-dispersing billiards, which means that the smooth components of ∂Q are convex (as seen from outside Q). If these components are strictly convex, the billiard is dispersing. In other words:

Definition 2.4. *A billiard system is semi-dispersing if for any point q of a smooth component of ∂Q, the curvature operator (or second fundamental form) of ∂Q – which is actually the derivative operator of the function $n(q)$ – is positive semi-definite. If this operator is positive definite for every $q \in \partial Q$, the billiard is dispersing.*

The name 'dispersing' comes from the effect that particles reaching ∂Q at nearby points with parallel velocities will be diverted from each-other – a planar front will be 'scattered' (see Figure 1.1 and Section 2.2). In a semi-dispersing billiard, planar fronts are scattered in some directions. In these cases, components of the complement of Q are often called 'scatterers'.

Throughout this work, unless otherwise emphasized, we will work with the discrete time dynamical system. Most results obtained can be transferred to the flow, e.g. the ergodicity of the map implies that of the flow (see [28]). In Chapter 5 we will discuss a modification of the above system, where collisions are not momentary, the walls are 'soft'. See the details there.

We finish this section by one more

Notation. *ρ will denote the Riemannian metric on M arising from the natural bundle structure (cf. [4]).*

2.2 Fronts

Our most important tools in describing hyperbolicity are local orthogonal manifolds or simply *fronts*. A front $\mathcal{W}$ is defined in the flow phase space rather than in the Poincaré section.

Definition 2.5. *Take a smooth one codimensional submanifold E of the whole configuration space Q, and add the unit normal vector $v(q)$ of this submanifold at every point q as a velocity, continuously. Consequently, at every point the velocity points to the same side of the submanifold E. The set*

$$\Sigma = \{(q, v(q)) \mid q \in E\} \subset \mathcal{M}, \tag{2.1}$$

where $v : E \to \mathbf{S}^{d-1}$ is continuous (smooth) and $v \perp E$ at every point of E, is called a front.

The derivative of this function v, called B plays a crucial role: $dv = Bdq$ for tangent vectors (dq, dv) *of the front.* B acts on the tangent plane $\mathcal{T}_q E$ of E, and takes its values from the tangent plane $\mathcal{J} = \mathcal{T}_{v(q)}\mathbf{S}^{d-1}$ of the velocity sphere. These are both naturally embedded in the configuration space Q, and can be identified through this embedding. So we just write $B : \mathcal{J} \to \mathcal{J}$. B is nothing else than the curvature operator of the submanifold E. Yet we will prefer to call it second fundamental form (SFF), in order to avoid confusion with other curvatures that are coming up. Obviously, B is symmetric.

Definition 2.6. *A front is called convex if its SFF B is positive semi-definite. A front is called strictly convex if its SFF B is positive definite.*

In billiards, fronts remain fronts during time evolution – at least locally, and apart from some singularity lines.

Remark 2.7. *We note that this will not exactly be the case in the 'soft' billiards of Chapter 5. However, fronts there can be used just as good as in 'hard' billiards: time can be rescaled locally so that slower particles of the front are 'awaited'. This will not modify the discrete time dynamics we use there.*

When we talk about a front, we sometimes think of it as the subset (submanifold) of the (whole) flow phase space $\mathcal{M}$ (for example, when we talk about time evolution under the flow), but sometimes just as the submanifold E of the configuration space Q (for example, when we talk about the tangent space or the curvature of the front). This should cause no confusion.

2.3 Hyperbolicity

Hyperbolicity is a key property of most billiard systems, used heavily in the proof of other properties representing strong statistical behaviour. In general, hyperbolicity means that all Lyapunov exponents are non-zero. [1] In Hamiltonian systems (including billiards) this means that at every point of the phase space the tangent space can be decomposed into two subspaces of equal dimension, on one of which (the *unstable* subspace) the dynamics is expanding, and on the other (the *stable* subspace) contracting. However, in billiards the notion of hyperbolicity is always connected to the geometrical picture of instability: the unstable subspace corresponds to some strictly convex front. See the outgoing front of Figure 1.1 (dispersing case) for an example. Phase points forming such a front will depart from each other, corresponding to expansion is the unstable direction. In semi-dispersing billiards, convex fronts remain convex during their time evolution, so this expansion is preserved. In other examples, including focusing billiards (e.g. the famous Bunimovich stadium) and soft billiards, the mechanism is more complicated, but the essence is the same: convex fronts remain convex. Section 5.2 describes this mechanism in detail for soft billiards. There a precise statement of (uniform) hyperbolicity is also formulated (Proposition 5.17). A further notion connected to hyperbolicity, the *sufficiency* of a phase point will be described in Section 4.1.1.

2.4 Singularities

Consider the set of tangential reflections, i.e.

$$\mathcal{S}_0 := \{(q,v) \in M \mid \langle v, n(q)\rangle = 0\} = \partial M.$$

It is easy to see that the map T is not continuous at the set $T^{-1}\mathcal{S}_0$. Further pre-images of $\mathcal{S}_0$ are

$$\mathcal{S}_k := T^{-k}\mathcal{S}_0 \quad (k > 0),$$

the so-called 'higher order' singularities. The singularity set (i.e. where it is not continuous) for a higher iterate T^n is

$$\mathcal{S}^{(n)} = \cup_{i=1}^{n} \mathcal{S}_i. \tag{2.2}$$

What is even worse, the derivative of the dynamics blows up near these singularities, as we will see. In the multi-dimensional case ($d > 2$) this is crowned by a vast anisotropy: expansion is blowing up in one direction, while disappearing in the other (both within the unstable subspace). This is the heart of the phenomenon we discuss in Chapter 3.

Generally it was assumed in the literature that the set $\mathcal{S}^{(n)}$ is a finite collection of smooth and compact submanifolds of the Poincaré phase space M. However, for multi-dimensional semi-dispersing billiards these manifolds can be treated as submanifolds of M only in a topological sense (see Chapter 3). So we view all these sets as (finite unions of) topologically embedded one codimensional compact submanifolds with boundary. They have smooth manifold structure in the interiour, however, the behaviour at the boundary is irregular (the curvature diverges).

[1] In the case of a billiard flow, one Lyapunov exponent – corresponding to the flow direction – is necessarily zero. Hyperbolicity means that all *others* are nonzero. See also Section 4.1.1.

Remark 2.8. *Above the (tangential) singularities have been introduced for the Poincaré section map T. But in some of the arguments later, they will also be needed for the flow. Remark 4.28 will discuss how this extension of the singularities is understood.*

Remark 2.9. *In case the boundary ∂Q is only piecewise smooth, further singularities – the multiple collisions – arise. As already mentioned in Remark 2.3, at such points neither $n(q)$, nor the flow dynamics is uniquely defined, thus we can speak about several "branches" of a trajectory. The singularity set must also be treated with more care. For this reason, in all cases we will denote by $\mathcal{S}^+$ the set of all singular phase points, which can be points of $\mathcal{S}_0$ or multiple collision points supplied with the possible* outgoing *velocities. (See [34] and its references for details). In the present work we consider only tangential singularities. Multiple collisions can be treated in an analogous way, although the main difficulty – the blow-up of the derivative of the dynamics – does not appear here.*

Remark 2.10. *the condition that we set in the definition of the billiard, that ∂Q should be piecewise C^3, ensures exactly that $\mathcal{S}_1$ is piecewise C^1. However, smoothness of $\mathcal{S}_2$ can not be ensured by higher regularity of ∂Q.*

2.5 Dimensions of different manifolds

In order to make the picture clearer, let us summarize the dimensions of the different spaces and manifolds that show up. Suppose that a billiard particle is moving in d dimensions, that is, on the torus $\mathbb{T}^d$. Then the dimension of the flow phase space is $2d-1$ (the energy is constant). The phase space of the Poincaré section map is $2d-2$ dimensional. Stable and unstable subspaces are $d-1$ dimensional (both for the flow and the map). Singularity manifold have $2d-2$ dimensions for the flow and $2d-3$ dimensions for the map.

2.6 Different notions of norms and metrics

In this section we summarize the different notions of metrics and distances used in the theory of billiards. This description is rather technical, and can be skipped for a first reading. It will only be essential for the understanding of technical details of chapters 4 and 5.

Let us assume that two phase points $x=(q,v)$ and $x'=(q',v')$ and a vector in the tangent plane at x, $w=(\delta q,\delta v)$ are fixed. Most of the metric notions of the theory use the *Euclidean norm* $\|w\|=\sqrt{|\delta q|^2+|\delta v|^2}$ and the generated Euclidean distance $\rho(x,x')$. The measure on M corresponding to this Riemannian metric (generated by the volume form) is simply the Lebesgue measure const. $dq\,dv$. However, in several other statements referred (see e.g. [28], especially the Erratum), two other metrics come about. For their definition we fix the notation for two $d-1$ dimensional linear subspaces in $\mathbb{R}^d$: $\mathcal{T}$, the one orthogonal to $n(q)$ and $\mathcal{J}$, the one orthogonal to v. Furthermore we introduce the linear operator $V:\mathcal{J}\to\mathcal{T}$ which is simply the projection parallel to v. (On details see [4].)

This way we may define the *invariant norm* of a vector: $\|w\|_i=\sqrt{|V^{-1}\delta q|^2+|\delta v|^2}$ and the generated invariant distance $\rho_i(x,x')$. The name 'invariant' comes from the fact that the measure corresponding to this Riemannian metric (via the volume form) is the invariant measure $d\mu = \text{const.}\,|\langle v, n(q)\rangle|\,dq\,dv$. Note that $\|w\||\langle v, n(q)\rangle| \leq \|w\|_i \leq \|w\|$, thus the two distances are equivalent if we can ensure $|\langle v, n(q)\rangle| \geq c$ for some constant c. This happens throughout the proof of the Fundamental Theorem (cf. Section 4.3) where we work in a neighbourhood of an interiour point $x \in M$ and thus the two metrics are (locally) equivalent.

The third metric-type quantity is the so-called *p-metric* $\|w\|_p = |V^{-1}\delta q|$. Even though this is a degenerate metric in general (that is the reason for the name 'p' – pseudo), it is non-degenerate when restricted to vectors w corresponding to convex fronts (cf. [28], [4]). Its importance is related to the fact that the most convenient way of handling hyperbolicity issues is in terms of the p-metric (see e.g. Lemma 4.2).

Related to the above mentioned metrics there are two ways of measuring distance of a phase point $x = (q, v)$ from the set of tangential reflections $\mathcal{S}_0$. $z(x) = \rho_i(x, \mathcal{S}_0)$ is simply the distance in terms of ρ_i. Alternatively we may consider tubular neighbourhoods U_r (of radius r) of the flow trajectory starting out of x in the configuration space Q. Then define $z_{tub}(x)$ as the supremum of the radii r for which the tube does not intersect the set of singular reflections (see [13] and [28], especially the Erratum). It is not difficult to see that $z(x) \leq z_{tub}(x)$.

Chapter 3

Geometry of singularities in multi-dimensional dispersing billiards

In several papers that appeared, singularities were assumed – either explicitly or implicitly – to consist of smooth one codimensional submanifolds of the phase space. Often, even a uniform bound on the curvature was assumed, independent of the order of the singularity. These assumptions were needed in order to estimate the measures of neighbourhoods of the singularities. Such a uniform bound on curvature does exist in 2-dimensional billiards. However, it does not exist in higher dimensions. In this chapter we present a counter example in a 3-dimensional dispersing billiard, but similar counter examples exist in every multi-dimensional semi-dispersing billiard. In correspondence with the notations introduced in Section 2.4, we will use the notation $\mathcal{S}_1$ and $\mathcal{S}_2$ for the set of those phase points, the trajectories of which have tangential first and second collisions, respectively. We will demonstrate that already the curvature of $\mathcal{S}_2$ has no upper bound, i.e. the curvature blows up near a point where the singularity manifold is not even differentiable.

To avoid confusion let us set up one further convention. As already mentioned, billiard dynamics has singularities: points where the billiard map is not continuous. These singularities occur on one codimensional submanifolds of the phase space. The development of the theory is based on considering connected and essentially smooth components of the singularity manifolds. The recently discovered phenomenon described below shows that these components are, indeed, only essentially smooth. On certain two codimensional submanifolds of them pathologies occur: singularities in the sense of algebraic singularity theory. To avoid confusion we will refer to these singular two codimensional submanifolds as *'pathologies'* – in contrast to the *'singularities'*, the singularity manifolds of the dynamics itself.

3.1 Counter example for bounded curvature

In this section we prove

Theorem 3.1. ***Generally in multi-dimensional dispersing billiards, the curvature of $\mathcal{S}_2$ is not bounded. The unit normal vector is not a continuously differentiable function of its base point.***

The proof is rather implicit. We start with a specific $3D$ dispersing billiard and the indirect assumption that the curvature is bounded, and find that the two-step singularity intersects the one-step singularity tangentially at every point of their intersection, except for a one codimensional degeneracy, where the intersection is not tangent. However – as a consequence of bounded curvature – our indirect assumption implies that the unit normal vector of $\mathcal{S}_2$ is a continuously differentiable function of its base point. Thus the set of those points where the two singularity manifolds intersect non-tangentially is open in $\mathcal{S}_1 \cap \mathcal{S}_2$. This way we get a contradiction. Then we argue that the same situation would occur in any multi-dimensional dispersing billiard.

Consider the situation demonstrated in Figure 3.1. To perform as transparent an argument as possible,

- the parameters on the figure and in the calculations below are different,
- the first scatterer, the surface where the trajectories start out, is a plane – thus it is not strictly convex.

Nevertheless the reader can easily see that these modifications have no real significance. We are in 3 dimensions, so take a standard 3D Cartesian coordinate system. Let the first 'scatterer' be the $\{z = 0\}$ plane. Let the second scatterer be the sphere with centre $O_1 = (0, -1, 1)$ and radius $R = 1$. Let the third scatterer be the sphere with centre $O_2 = (1, 0, 2)$ and radius $R = 1$. We look at the component of the phase space corresponding to the first scatterer, near the phase point $(x_0 = 0, y_0 = 0, v_{x0} = 0, v_{y0} = 0)$. Of course, $v_{z0} = 1$, and the trajectory is the z axis. We are interested in the singularity manifold belonging to a tangent *second* collision. To describe this, let $D \in \mathbb{R}^4$ be the set of those points (x, y, v_x, v_y) the trajectories of which hit the first sphere. Let $r : D \to \mathbb{R}$ be the distance of the trajectory and O_2. That is, the singularity manifold we are looking at is the set

$$\mathcal{S}_2 \;=\; \{(x, y, v_x, v_y) \in D \mid r(x, y, v_x, v_y) = 1\}.$$

So, if we want to construct the normal vector of the singularity manifold, we just need to calculate the gradient of r. We will directly calculate the partial derivatives. Since $(x_0, y_0, v_{x0}, v_{y0}) = (0, 0, 0, 0)$ is on the boundary of D, we can only hope to find one-side partial derivatives. What is even worse: $(x, y, v_x, v_y) = (x, 0, 0, 0) \in D$ only if $x = 0$, so we cannot differentiate with respect to x. The same is true for v_x. What we can do is take these partial derivatives at the points $(0, y, 0, v_y)$ and then the limits

$$\lim_{y\to 0}\lim_{v_y\to 0}\frac{\partial}{\partial x} r(x, y, v_x, v_y)\big|_{x=v_x=0}.$$

(We will see that it is important to fix $x = v_x = 0$. If we were to approach the origin via a different path, we could get a different limit.)

We start with the indirect assumption that $\mathcal{S}_2$ has bounded curvature. This implies that the unit normal vector of $\mathcal{S}_2$ is a continuously differentiable function of its base point with bounded derivative. In this way it makes sense to define the normal vector of $\mathcal{S}_2$ on the boundary points of $\mathcal{S}_2$ as the limit of (unit) normal vectors on the interiour. For us the indirect assumption will mean that the limit

$$\mathrm{grad} r(0, 0, 0, 0) := \lim_{(x,y,v_x,v_y)\to(0,0,0,0)} \mathrm{grad} r(x, y, v_x, v_y)$$

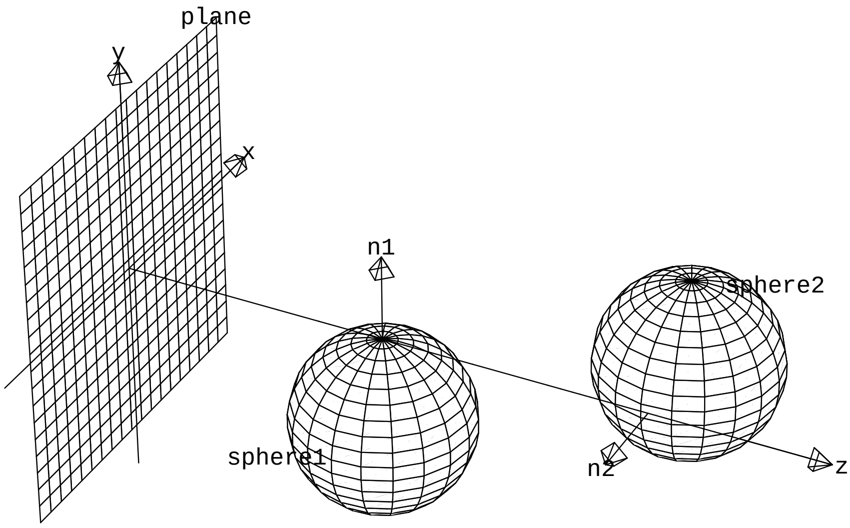


Figure 3.1: the studied billiard configuration

exists.

The closer a reflection is to tangential, the less effect it has on the 'neutral' direction. In our case, the reflection on the first sphere causes 'no scattering' in the x direction. That is, let (v'_x, v'_y, v'_z) be the velocity after the first collision. The x direction being the 'neutral' direction means that

$$\lim_{y\to 0} \frac{\partial}{\partial v_x} v'_x(0, y, 0, 0) = 1,$$

which implies that

$$\lim_{y\to 0} \frac{\partial}{\partial v_x} r(0, y, 0, 0) = -2.$$

Similarly,

$$\lim_{y\to 0} \frac{\partial}{\partial x} v'_x(0, y, 0, 0) = 0,$$

which implies that

$$\lim_{y\to 0} \frac{\partial}{\partial x} r(0, y, 0, 0) = -1.$$

According to our indirect assumption, this means that

$$\frac{\partial}{\partial x} r(0, 0, 0, 0) = -1$$

and

$$\frac{\partial}{\partial v_x} r(0, 0, 0, 0) = -2.$$

For the other two components, fix $x = v_x = 0$. So the trajectory is in the $\{x = 0\}$ plane, the scattering is just a 2D problem. We will calculate the one-side partial derivatives $\frac{\partial}{\partial y} r(0, 0, 0, 0)$ and $\frac{\partial}{\partial v_y} r(0, 0, 0, 0)$.

To find out about v'_y, let φ be the angle of the first sphere's radius at the first collision point and the $n_1 = (0,1,0)$ vector. If $v_y = 0$, then $1-\cos\varphi = -y$ ($y < 0$, of course), which, in leading order, gives $\varphi = \sqrt{-2y}$. It can be seen that after the reflection $v'_y = \sin 2\varphi$. That is, the trajectory is far from being a line. However, it is diverted in the very direction which – in the first order – does not affect its distance from O_2. Instead, in leading terms, $r^2 = 1 + (v'_y)^2$.

Putting these together, we get $r = \sqrt{1-8y}$, that is,

$$\frac{\partial}{\partial y} r(0,0,0,0) = -4.$$

If we fix $y = 0$, the exact same consideration gives $r = \sqrt{1-8v_y}$, that is,

$$\frac{\partial}{\partial v_y} r(0,0,0,0) = -4$$

as well. All together, we get

$$\mathrm{grad} r(0,0,0,0) = (-1,-4,-2,-4).$$

This is (one possible limit of) the normal vector of the singularity at the point $(x = 0, y = 0, v_x = 0, v_y = 0)$.

It is easy to see that the singularity corresponding to a tangent reflection on the first sphere has the normal vector

$$\mathrm{grad} r_0(x,y,v_x,v_y) = (0,-1,0,-1).$$

That is, the two singularities are not tangent at this point.

The previous consideration for grad$r$ also shows that this behaviour is exceptional. It is the result of the fact that in the first order r was unaffected by v'_y. If the radii n_1 and n_2 of the scatterers at the reflection points $(x,y,z) = (0,0,1)$ and $(x,y,z) = (0,0,2)$ had not been orthogonal, the result would have been

$$\frac{\partial r}{\partial y} = \infty, \frac{\partial r}{\partial v_y} = \infty,$$

corresponding to a normal vector $(0,1,0,1)$, meaning that the two singularities are tangent. Non-tangentiality of the two singularities is a one codimensional degeneracy.

As we have pointed out at the beginning of the section, this contradicts our indirect assumption on the boundedness of the curvature. In this way we have only proven that the assumption was false. However, we believe that the picture of the singularity suggested above is correct, the singularities *are* tangent almost everywhere, and their curvature only blows up near the pathological points described.

3.2 Discussion

For a rigorous proof of some finer properties (such as ergodicity and correlation decay (see Definition 1.3)) of multi-dimensional dispersing billiards, it seems essential to characterize singularities in a systematic way. Part of this characterization – which is adequate for the proof of local ergodicity – is done in Section 4.2. In the present section we do not give rigorous proofs; we would like to point out some analogies to and emphasize some interesting features of the irregularities demonstrated above.

The Whitney-umbrella. Consider the one codimensional set in $\mathbb{R}^3$ defined by the polynomial equation

$$\{(x, y, z) \in \mathbb{R}^3 \mid x^2 z = y^2\},$$

the Whitney-umbrella. 'One half' of this set (its intersection with the quadrants $\{xy \leq 0\}$) is shown in Figure 3.2. For simplicity we use the notations W_2 for this 'half-umbrella' and W_1 for the $\{z = 0\}$ plane. Clearly

- W_2 terminates on W_1 (in the points of the x axis), thus $W_1 \cap W_2 = \partial W_2$.
- at every point of the x axis where $x \neq 0$ the intersection of W_2 and W_1 is tangential.
- W_2 has smooth manifold structure in its interiour; nevertheless, near the origin its curvature is unbounded as the normal vector changes rapidly (actually, the unit normal vector does not even have a well-defined limit at the origin).

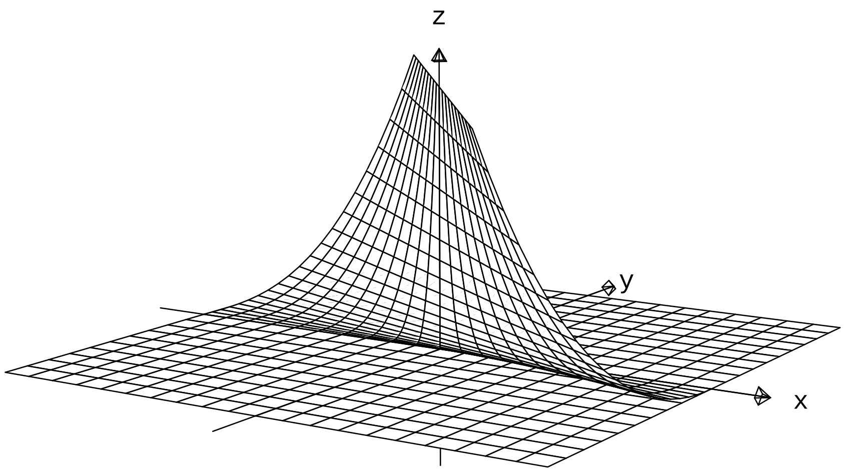


Figure 3.2: the Whitney umbrella

By these properties the geometry of singularities in the counter example is analogous to Figure 3.2. [1] W_1 corresponds to $\mathcal{S}_1$, W_2 corresponds to $\mathcal{S}_2$ while the origin corresponds to the set of those doubly tangential reflections where the two radii are orthogonal (this set is one codimensional in $\mathcal{S}_1 \cap \mathcal{S}_2$).

Lipschitz decomposability of the Whitney-umbrella

As mentioned in the beginning of this section, assumptions about the curvature of singularity manifolds were used in estimates on the size of a neighbourhood of the manifold. This analogy also shows that bounded curvature is not needed for the neighbourhood of a manifold to be small. Indeed, the 'half-umbrella' W_2 can be cut further into two pieces

[1] To be precise, the situation in Figure 3.2 has one dimension less – in contrast to W_2 the singularities are 3-dimensional manifolds – but this has little significance to the analogy.

(namely, its intersections with the quadrants $\{x \geq 0, y \leq 0\}$ and $\{x \leq 0, y \geq 0\}$), each of which is the graph of a Lipschitz function, when viewed from the appropriate direction. Indeed, easy calculations show that if we choose the direction $(1,1,1)$ or $(-1,-1,1)$ to be 'vertical' (respectively), these 'quarter-umbrellas' become graphs of Lipschitz functions with Lipschitz constant $\sqrt{2}$. So the whole Whitney-umbrella consists of four such graphs plus a one-dimensional tail. This tail (the negative z axis) has no analogue in the singularities of billiards. It's only there because the umbrella was defined in an algebraic way. However, if we consider the whole umbrella instead of the part actually corresponding to the singularity, this tail will not spoil our measure-theoretic estimates because it has one dimension less than the rest of the set.

Generalization I. First let us consider the first-step singularity $\mathcal{S}_1$. By the notations of the previous counter example we may characterize the points (x, y, v_x, v_y) belonging to $\mathcal{S}_1$ easily. These are precisely those for which $d(x, y, v_x, v_y) = 1$, where $d(.,.,.,.)$ is the distance of the point $O_1 = (0, -1, 1)$ from the line that passes through the point $(x, y, 0)$ and has direction specified by the velocity components v_x, v_y. As d is a smooth function of its variables there is no curvature blow-up for $\mathcal{S}_1$ – and, for first-step singularities in general. Thus $\mathcal{S}_2$ is a *pre-image of a smooth one codimensional compact submanifold*, however, *the map* under which the pre-image is taken *has unbounded derivatives and is highly anisotropic.* Curvature blow-up occurs only at those points of $\mathcal{S}_2$ (near its intersection with $\mathcal{S}_1$) where the map behaves irregularly.

In correspondence with the above observation we conjecture that curvature blow-up is not a feature peculiar to $\mathcal{S}_2$, it is present in the pre-images of one codimensional smooth submanifolds in general. Consider for example two-step *secondary* singularities Γ_2 – those phase points for which at the second iterate, instead of tangentiality, the collision term $(\langle n, v\rangle)$ is a given constant (see Chapter 5 for more detail). In the specific example of section 3.1 such secondary singular trajectories are precisely those that touch tangentially a sphere of radius R' ($R' < 1$) at the second iterate. It is clear that the geometry of Γ_2 is completely analogous to the geometry of $\mathcal{S}_2$.

Generalization II. Our calculations do not use any speciality of the explicitly given billiard configuration. Doubly tangential reflections for which the normal vectors of the scatterers at the consecutive collisions are orthogonal can be found in *any multi-dimensional semi-dispersing billiard.* Near such trajectories a similar calculation can be performed, and the same pathology will show up.

Generalization III. All in all, the discovered pathology is general. In addition, *the higher step singularities* $\mathcal{S}_k$; $(k \geq 3)$ may show even wilder behaviour near their intersections. Nevertheless, we strongly conjecture that a nice geometric characterization – suggested by the analogy with the Whitney-umbrella in the case of $\mathcal{S}_2$ – can be performed.

We have mentioned these generalizations to present the reader the picture of singularities we have in mind. Nevertheless, for our further discussion we do not need to verify any of these calculations or generalizations since they are completely independent.

Chapter 4

Local ergodicity theorem for multi-dimensional semi-dispersing billiards

This chapter is devoted to the ergodicity of semi-dispersing billiards. We restrict our attention to the question of *local* ergodicity. The main theorem of the chapter, Theorem 4.16 is called the "Fundamental Theorem" or "local ergodicity theorem" for semi-dispersing billiards. The essence of the statement is that for every sufficient point (see Definition 4.1) a neighbourhood of the point belongs to one ergodic component. [1] However, the statement of the theorem in its raw form is not about ergodic components, but about the existence of (many long) stable and unstable manifolds. To get local ergodicity from that, one uses the chain method of Hopf. See e.g. [13].

In contrast to smooth dynamical systems, billiards have singularities which make the application of the classical methods substantially more difficult. One reason is that in the neighbourhood of orbits tangent to the obstacles (the so-called tangent singularities) the derivative of the Poincaré section map diverges. Nevertheless, Sinai's celebrated 1970 result demonstrated that, at least for $d = 2$, the hyperbolicity caused by the strictly convex scatterers overcomes the harmful effect of singularities. In fact, he showed that $2D$ dispersing billiards, i.e. those with strictly convex obstacles, are ergodic and even K-mixing [38].

Multi-dimensional geometry is, however, essentially richer so it is not surprising that it had taken 17 years until Chernov and Sinai [13] could extend Sinai's original result to multi-dimensional dispersing billiards. This remarkable achievement was a corollary of their *local ergodicity theorem*, often called the *Fundamental Theorem*, formulated for semi-dispersing billiards, i.e. those with convex scatterers. Their theorem got slightly generalized with the clarification of some technical details and conditions by Krámli, Simányi and Szász [28] in 1990.

The considerations in the proof of the local ergodicity theorem are local. As a matter of fact, by assuming the boundedness (from above) of the curvature of all images of the 'tangent collisions' set $\mathcal{S}_0$ – which is a straightforward fact for $d = 2$ – it became possible to assume that they are linear objects, at least locally. However, as discussed in Chapter 3

[1] Having this statement, the proof of ergodicity boils down to showing that there are many sufficient phase points, which is not a topic of this thesis.

(the result was originally published in [4]), for $d \geq 3$, in the neighbourhoods of singularities the images of other singularities (and of other smooth one codimensional submanifolds of the phase space) develop a pathological behaviour contradicting the boundedness of the curvatures. Therefore, for its own interest, but also for its various important consequences it became necessary to correct the original arguments and this is the aim of this chapter. Indeed, instead of the boundedness property of the aforementioned curvatures we formulate a new condition, the so called *Lipschitz decomposability condition.* Roughly speaking it requires that the singularities can be decomposed into a finite number of graphs of locally Lipschitz functions with the boundaries of these graphs being not too wild. This assumption, together with the other requirements of the local ergodicity theorem, is already sufficient to save the old proof. The next question is, of course, when this new condition holds. Fortunately, we can verify it under one additional requirement: we assume that the scatterer boundaries are *algebraic.* Luckily enough, the main examples of multi-dimensional semi-dispersing billiards are all algebraic. Just think – first of all – of hard ball systems [36], [34], of the Lorentz process with spherical scatterers [23], [41], of general algebraic cylindrical billiards [40], [35], and of the multi-dimensional stadia designed by Bunimovich and Reháček [10].

For keeping our exposition possibly short, we rely heavily on that of [28]. In Section 4.1 we summarize the necessary notations and prerequisites from the aforementioned work. In Section 4.2 we present the Lipschitz decomposability property of the singularities. Based on this assumption, in Section 4.3 we reformulate the local ergodicity theorem and discuss in detail where and how the classical proof of [13] and [28] should be modified. Finally, in Section 4.4 it is shown that the Lipschitz decomposability property holds for algebraic billiards.

The methods, though quite elementary, come from different branches of mathematics. Throughout the arguments we try to keep the exposition self-contained. More details on the basic notions from algebra or geometric measure theory can be found in the books [2], [33], [39] and [20], [18], [19]; respectively.

4.1 Multi-dimensional semi-dispersing billiards

In this section we summarize some basic properties of semi-dispersing billiards. Our aim is to introduce the most important concepts and fix the notation in order to keep the exposition of the chapter self-contained. For a more detailed description see the literature, especially [28].

Notation. *For any $n \in \mathbb{N}$, Δ_n stands for the set of doubly singular phase points up to order n, i.e.*

$$\Delta_n = \left\{ x \in M \mid \text{ there exist indices } k_1 \neq k_2, |k_i| \leq n \text{ such that } T^{k_1}x \in \mathcal{S}_0,\ T^{k_2}x \in \mathcal{S}_0 \right\}.$$

We are mainly interested in phase points with regular or with at most once singular trajectories, thus we consider the following sets:

$$
\begin{aligned}
M^0 :=& \quad M \setminus \bigcup_{n\in\mathbb{Z}} \mathcal{S}_n \\
M^* :=& \quad M \setminus \bigcup_{n=1}^{\infty} \Delta_n \\
M^1 :=& \quad M^* \setminus M^0.
\end{aligned}
$$

As to regular and at most once singular phase points of the flow, the sets $\mathcal{M}^0$, $\mathcal{M}^*$ and $\mathcal{M}^1$ refer to flow-images of M^0, M^* and M^1, respectively.

4.1.1 Hyperbolicity

Beside the presence of singularities, the most important feature of semi-dispersing billiard dynamics is that it is – at least locally and non-uniformly – hyperbolic. A highly important consequence of this fact is the abundance of local invariant manifolds. The notion of a local invariant manifold will be used in the traditional sense, i.e. a C^1-smooth, connected submanifold $\gamma^s \subset \partial M$ is a *local stable manifold* at $x \in \partial M$ iff

$$
\begin{aligned}
&(i) \quad x \in \gamma^s \\
&(ii) \quad \exists K(\gamma^s), C(\gamma^s) > 0 \text{ such that for any } y_1, y_2 \in \gamma_s \\
&\qquad\ \rho(T^n y_1, T^n y_2) \le K \exp(-Cn)\rho(y_1, y_2).
\end{aligned}
$$

Local stable manifolds for the inverse dynamics T^{-1} will be referred to as local unstable manifolds.

Let us consider a nonsingular finite trajectory segment for the flow: $S^{[a,b]}x$, where $a < 0 < b$ and a, b, 0 are not moments of collision.
$\mathcal{N}_0(S^{[a,b]}x)$, the *neutral subspace* at time 0 for the segment $S^{[a,b]}x$ is defined as follows:

$$
\begin{aligned}
\mathcal{N}_0(S^{[a,b]}x) := \quad \{ \quad & w \in \mathbb{R}^d : \exists\delta > 0 \text{ such that } \forall\alpha \in (-\delta, \delta) \\
& v(S^a(q(x) + \alpha w, v(x))) = v(S^a x) \ \& \\
& v(S^b(q(x) + \alpha w, v(x))) = v(S^b x) \quad \}.
\end{aligned}
$$

Observe that $v(x) \in \mathcal{N}_0(S^{[a,b]}x)$ is always true, the neutral subspace is at least 1 dimensional. Neutral subspaces at time moments different from 0 are defined by $\mathcal{N}_t(S^{[a,b]}x) := \mathcal{N}_0(S^{[a-t,b-t]}(S^t x))$, thus they are naturally isomorphic to the one at 0.

Definition 4.1. *The non-singular trajectory segment $S^{[a,b]}x$ is* sufficient *if for some (and in that case for any) $t \in [a, b] : \dim(\mathcal{N}_t(S^{[a,b]}x)) = 1$. A point $x \in \mathcal{M}^0$ is said to be sufficient if its entire trajectory $S^{(-\infty,\infty)}x$ contains a finite sufficient segment. Singular points are treated by the help of trajectory branches (see [28]): a point $x \in \mathcal{M}^1$ (this means that the entire trajectory contains one singular reflection) is sufficient if both of its trajectory branches are sufficient.*

All these concepts have their natural counterparts for the billiard map phase space M. For example, a smooth piece $\Sigma \in M$ of the image of a local orthogonal manifold in $\mathcal{M}$ is referred to as a front as well.

The treatment of hyperbolicity is traditionally related to fronts through the following simple phenomena. Near sufficient phase points, hyperplanes in Q orthogonal to the flow evolve into strictly convex fronts. Convex fronts remain convex under time evolution. The importance of this is shown by the lemma below. Before formulating it we introduce one more notation: $D^n_{y,\Sigma}$ is the derivative of the (nth power of the) dynamics T^n restricted to the front Σ.

Lemma 4.2. *(Equivalent of Lemma 2.13 from [28].) For every $x \in M^0$ for which the trajectory is sufficient there exists a neighbourhood $U(x)$ and a constant $0 < \lambda(x) < 1$ such that*

- *through almost every point $y \in U(x)$ there do pass uniformly transversal local stable and unstable manifolds $\gamma^s(y)$ and $\gamma^u(y)$ of dimension $d-1$;*
- *for any $y \in U(x)$ and any convex front $\Sigma^\pm$ passing through $\pm y$:*

$$\|(D^\tau_{\pm y,\Sigma^\pm})^{-1}\|_p < \lambda(x),$$

where $\tau \in \mathbb{Z}^+$ is the first return time to $U(x)$.

More details about local hyperbolicity and semi-dispersing billiards in general can be found in [28].

4.2 Lipschitz property of singularities

We have mentioned that in several papers that appeared, singularities were assumed – either explicitly or implicitly – to consist of smooth one codimensional submanifolds of the phase space. Often, even a uniform bound on the curvature was assumed, independent of the order of the singularity. This is true for $2D$ billiards. However, it is not true in higher dimensions. In Chapter 3 we presented a counter example in a $3D$ dispersing billiard. Already the curvature of $\mathcal{S}_2$ has no upper bound, i.e. the curvature blows up near a point where the singularity manifold is not even differentiable. After this example we propose another property, which, in most applications, can replace the bounded curvature assumption. We conjecture that this property: *the Lipschitz decomposability of singularities* holds for multi-dimensional semi-dispersing billiards.

The analogy shown in Chapter 3 – the *Whitney umbrella* – not only illustrates better the pathological situation in three dimensions (rather than our counter example in dimension 4), but also suggests the way out: to substitute the condition on the boundedness of curvatures with the Lipschitz decomposability property.

When treating ergodic or stochastic properties of singular systems, we need to understand the properties of singularities in order to know that their neighbourhood is of small measure. By assuming that the singularities are smooth, e.g. they have bounded curvature, in local considerations one can treat them as planes, by choosing an appropriately small scale. This, of course, implies that the intersection of a (smooth) singularity component and a sphere of radius r has a surface-volume of order r^{m-1} where $m = 2d-2$ is the dimension of the phase space. Similarly, the δ-neighbourhood of such a singularity-piece has measure of order $r^{m-1}\delta$. These properties have been used in several papers without being checked. We now know that the curvature is in general not bounded, so a more careful investigation is essential.

Notation. *For any subset in a Riemannian manifold $H \subset \mathbb{M}$, $H^{[\delta]}$ shall denote its δ-neighbourhood:*

$$H^{[\delta]} := \{x \in \mathbb{M} \mid \rho(x, H) \leq \delta\}.$$

To ensure that the regularity properties mentioned hold, we (approximately) propose to assume that the singularities have components which are graphs of Lipschitz functions – instead of assuming they have smooth components. A subset H of $\mathbb{R}^m$ will be called a Lipschitz graph, if we can choose a Cartesian coordinate system so that H becomes the graph of a Lipschitz function: $H = \{(x, f(x)) \mid x \in D\}$ with some (measurable) $D \subset \mathbb{R}^{m-1}$ and $f : \mathbb{R}^{m-1} \to \mathbb{R}$ Lipschitz-continuous. Being a Lipschitz graph ensures that H is rectifiable, and that for its surface-volume one has $\mu(H) \leq C\mu(D)$ where the constant C depends only on the Lipschitz constant of f. The main property of Lipschitz graphs is shown by the following very basic

Lemma 4.3. *Let $D \subset \mathbb{R}^{m-1}$ arbitrary, $f : \mathbb{R}^{m-1} \to \mathbb{R}$ Lipschitz-continuous with Lipschitz-constant L. Let $H = \{(x, f(x)) \mid x \in D\} \subset \mathbb{R}^m$. Denote by $\mathcal{L}^m$ the Lebesgue-measure in $\mathbb{R}^m$, and by $\mathcal{L}^{m-1}$ the Lebesgue-measure in $\mathbb{R}^{m-1}$. Then*

$$\mathcal{L}^m(H^{[\delta]}) \leq 2\delta\sqrt{L^2+1}\mathcal{L}^{m-1}(\bar{D}) + o(\delta)$$

Proof. Just notice that

$$H^{[\delta]} \subset \{(x, y) \mid x \in D^{[\delta]}, |y - f(x)| \leq \delta\sqrt{L^2+1}\},$$

where $D^{[\delta]}$ is the δ-neighbourhood of D in $\mathbb{R}^{m-1}$. This implies

$$\mathcal{L}^m(H^{[\delta]}) \leq 2\delta\sqrt{L^2+1}\mathcal{L}^{m-1}(D^{[\delta]})$$

which gives the lemma, since $\mathcal{L}^{m-1}(D^{[\delta]}) \to \mathcal{L}^{m-1}(\bar{D})$ as $\delta \to 0$. □

To precisely formulate the property that we propose instead of smoothness of the singularities, we need the following two definitions:

Definition 4.4. *(cf. [20]) A function $f : D \subset \mathbb{R}^{m-1} \to \mathbb{R}$ will be called locally Lipschitz (with Lipschitz constant L), if for any $x \in D$ there exists a neighbourhood $U \subset \mathbb{R}^{m-1}$ of x such that the restricted function $f\Big|_{D\cap U}$ is Lipschitz (with Lipschitz constant L).*

In all our applications D will be open. Notice that in this case, f typically cannot even be extended to $\bar{D}$ in a continuous way.

Definition 4.5. *$H \subset \mathbb{R}^m$ will be called a (one codimensional) locally Lipschitz graph (with Lipschitz constant L), if we can choose an appropriate Cartesian coordinate system so that H becomes the graph of a locally Lipschitz function:*

$$H = \{(x, f(x) \mid x \in D\}$$

with some $D \subset \mathbb{R}^{m-1}$ and $f : D \to \mathbb{R}$ locally Lipschitz (with constant L).

We will be mainly interested in the case when the domain D is an open set in $\mathbb{R}^{m-1}$, then

– H will be called an open locally Lipschitz graph (even though it is not an open set in $\mathbb{R}^m$),

– and we will denote by ∂H the boundary of H as of a surface: $\partial H = \bar{H} \setminus H$.

Now we are able to define the regularity property that should replace the smoothness of singularities. This property, called 'Lipschitz decomposability' will be defined for subsets of $\mathbb{R}^m$ here. For Lipschitz decomposability of subsets of a Riemannian manifold, see Remark 4.7.

Definition 4.6. *Consider $H \subset \mathbb{R}^m$, and $L \in \mathbb{R}$. H will be called a 'Lipschitz decomposable' (one codimensional) subset with constant L if it can be decomposed into a finite number of open locally Lipschitz graphs and a small remainder set in the following way: There exist H^* and $H_1, ..., H_K$ such that:*

- $H \subset \bigcup_{i=1}^{K} \bar{H}_i \bigcup H^*$,
- $H_i \cap H_j = \emptyset$ *for any* $i \neq j$,
- *every H_i is a one codimensional open locally Lipschitz graph (with constant L),*
- $\mathcal{L}^m \left(\left((\bigcup_{i=1}^{K} \partial H_i) \cup H^* \right)^{[\delta]} \right) = o(\delta)$.

The set H^* is included in the decomposition for technical reasons: we want to allow for sets H having parts of strictly higher codimension. This occurs generically if H is an algebraic subvariety of $\mathbb{R}^n$. See Section 3.1 on the Lipschitz decomposability and the one dimensional tail of the Whitney-umbrella. We wish to substitute our singularity manifold by algebraic varieties (see Section 4.4), that's why we allow this higher codimensional tail. Nevertheless we would like to note that such higher codimensional parts are not present in the singularities of semi-dispersing billiards.

Remark 4.7. *Lipschitz decomposability in Riemannian manifolds. Throughout the chapter – and in particular in Conjecture 4.8 below – subsets of a compact Riemannian manifold $\mathbb{M}$ are considered. For $H \subset \mathbb{M}$ Lipschitz decomposition is understood in terms of coordinate charts.*

To be more precise, let us fix some convention related to the atlas $\{U_t, \psi_t\}_{t=1}^T$ for $\mathbb{M}$ first. It is important that $\mathbb{M}$ is compact thus we may consider a finite atlas. We say that the atlas is bi-Lipschitz if all charts $\psi_t : U_t \to \mathbb{R}^m$ are bi-Lipschitz maps, i.e. both ψ_t and $(\psi_t)^{-1}$ are Lipschitz with some constant $K > 1$. All atlases considered are assumed to be bi-Lipschitz with a fixed constant. This ensures that Euclidean distance on $\mathbb{R}^m$ is comparable to Riemannian metric on the manifold, and thus our metric estimates indeed apply in the arguments of Section 4.3. Note that bi-Lipschitzness – with Lipschitz constant arbitrarily close to one – can always be obtained by choosing the coordinate patches sufficiently small.

As to the problem of Lipschitz decomposition, we will say that $H \subset \mathbb{M}$ is Lipschitz decomposable whenever a finite bi-Lipschitz atlas can be chosen, such that for all charts $\psi_t(H \cap U_t)$ is Lipschitz decomposable as a subset of $\mathbb{R}^m$, in the sense of Definition 4.6.[2]

The precise property that we expect the singularities of semi-dispersing billiards to have is formulated in the form of a conjecture:

[2]The delicate question, how sensitive this notion of Lipschitz decomposition is to the choice of the atlas, needs further investigation.

Conjecture 4.8. *For any semi-dispersing billiard with a finite horizon there exists an $L \in \mathbb{R}$ such that for any integer N the set $\bigcup_{|n|\le N} \mathcal{S}_n$ of singularities of order up to N is 'Lipschitz decomposable' with constant L.*

It is worth noting that by introducing 'transparent walls' (cf. [13]) any semi-dispersing billiard can be reduced to one with a finite horizon.

The statement of this conjecture will appear word by word among the conditions of the modified version of the Fundamental Theorem for semi-dispersing billiards stated in Section 4.3.1. The conjecture will be proven for the utmost important special case of semi-dispersing billiards with algebraic scatterers in Section 4.4.

To help the reader understand why this 'Lipschitz decomposability' property is defined exactly as it is, we present two more lemmas in this section. These are the lemmas through which the decomposability of singularities will be used.

Lemma 4.9. *Let $H \in \mathbb{R}^m$ be a one codimensional locally Lipschitz graph with $H = \{(x, f(x)) \mid x \in D\}$, $D \subset \mathbb{R}^{m-1}$ open, $f : D \to \mathbb{R}$ locally Lipschitz with constant L. Assume furthermore that $\mathcal{L}^m\left((\partial H)^{[\delta]}\right) = o(\delta)$.*
Let $D' \subset D$ arbitrary, $H' = \{(x, f(x)) \mid x \in D'\}$. Then

$$\mathcal{L}^m\left(H'^{[\delta]}\right) \le 2\delta\sqrt{L^2+1}\mathcal{L}^{m-1}\left(\bar{D}'\right) + o(\delta).$$

Proof. Let $x_0 \in D$, $X_0 = (x_0, f(x_0)) \in H$. If $\operatorname{dist}(x_0, \partial D) > \delta$ then

$$B_\delta(X_0) \subset \left\{(x, y) \mid x \in D,\ \operatorname{dist}(x, x_0) \le \delta,\ |y - f(x)| \le \delta\sqrt{L^2+1}\right\}.$$

On the other hand, if $d = \operatorname{dist}(x_0, \partial D) \le \delta$, then there exists an $x_1 \in \partial D$ with $\operatorname{dist}(x_0, x_1) = d$.

With this x_1, for every $0 \le t < 1$ $x_t := tx_1 + (1-t)x_0 \in D$, otherwise $\operatorname{dist}(x_0, \partial D) < d$ would hold. The function $g : [0, 1) \to \mathbb{R}$, $g(t) := f(x_t)$ is Lipschitz with constant dL, so $g(1) := \lim_{t\nearrow 1} g(t)$ exists and $|g(1) - g(0)| \le dL$.

Obviously $X_1 := (x_1, g(1)) \in \partial H$ and $\operatorname{dist}(X_0, X_1) \le d\sqrt{L^2+1}$. That is, $B_\delta(X_0) \subset B_{(\sqrt{L^2+1}+1)\delta}(X_1)$. Putting everything together, we have

$$(H')^{[\delta]} \subset \left\{(x, y) \mid x \in (D')^{[\delta]} \cap D, |y - f(x)| \le \delta\sqrt{L^2+1}\right\} \cup (\partial H)^{[(\sqrt{L^2+1}+1)\delta]}. \tag{4.1}$$

This implies

$$\mathcal{L}^m\left((H')^{[\delta]}\right) \le 2\delta\sqrt{L^2+1}\mathcal{L}^{m-1}\left((D')^{[\delta]}\right) + \mathcal{L}^m\left((\partial H)^{[(\sqrt{L^2+1}+1)\delta]}\right).$$

This gives the statement of the lemma since $\mathcal{L}^{m-1}\left((D')^{[\delta]}\right) = \mathcal{L}^{m-1}\left(\bar{D}'\right) + o(1)$ and the second term is $o(\delta)$ because of our assumption. □

In the next lemma, π will denote the projection of $\mathbb{R}^m$ to $\mathbb{R}^{m-1}$ parallel to the last axis: $\pi((x, y)) := x$ when $x \in \mathbb{R}^{m-1}$ and $y \in \mathbb{R}$.

Lemma 4.10. *Let $H \subset \mathbb{R}^m$ be a one codimensional locally Lipschitz graph with $H = \{(x, f(x)) \mid x \in D\}$, $D \subset \mathbb{R}^{m-1}$ open, $f : D \to \mathbb{R}$ locally Lipschitz with constant L. Let $\delta > 0$ and $G \subset \mathbb{R}^m$ be such that* $\operatorname{dist}(G, \partial H) > (\sqrt{L^2+1}+1)\delta$. *Then*

$$\mathcal{L}^m\left(H^{[\delta]} \cap G\right) \le 2\delta\sqrt{L^2+1}\mathcal{L}^{m-1}\left(\pi(G)\right).$$

Proof. Let $H' = H$. (4.1) holds just like in the previous lemma. Since $\mathrm{dist}(G, \partial H) > (\sqrt{L^2+1}+1)\delta$, this means that

$$H^{[\delta]} \cap G \subset \left\{ (x,y) \mid x \in D \cap \pi(G), |y - f(x)| \leq \delta\sqrt{L^2+1} \right\},$$

which gives the statement of the lemma. □

4.3 The Fundamental Theorem

In this section we present our version of the Fundamental Theorem – or local ergodicity theorem – for semi-dispersing billiards. The first version by Chernov and Sinai [13] offered a quite involved, but in essence very transparent formulation of the theorem and a delicate proof. A self-contained exposition of the original ideas with detailed conditions and arguments was provided in [28] where a slightly more general, the so called 'transversal' version of the Fundamental Theorem was announced – mainly with its application to three-billiards in mind. Several other papers have appeared in the 90's with nice expositions of the theorem, even for classes of dynamical systems more general than the original semi-dispersing billiard setting (eg. Hamiltonian systems with singularities in [29]). However, all of these papers assumed that for all powers of the dynamics the singularity set is a finite collection of one codimensional smooth and compact submanifolds. Since, as our counter example shows, this is not the case, it became utmost necessary to replace this assumption.

Throughout the section our main reference is [28]. Actually, our aim is to demonstrate that it is possible to modify the proof presented there to the case when the singularity sets are not smooth but just finitely Lipschitz decomposable. After formulating the conditions and the statement of the theorem, we give a sketch of the proof (that goes along the lines of [28]) and work out those parts in more detail, where the original argument is to be modified. However, the original notation had to be modified at many points – in order to be consistent with the other chapters of this work.

4.3.1 Formulation of the theorem

Before its formulation, it is important to point out the **conditions** under which the modified proof of the theorem works. We use the notations introduced in Section 4.1.

Condition 4.11. (Chernov-Sinai Ansatz, Condition 3.1 from [28]). *For $\nu_{\mathcal{S}^+}$-almost every point $x \in \mathcal{S}^+$ [3] we have $x \in M^*$ and, moreover, the positive semi-trajectory of the point x is sufficient.*

What follows below is our new condition – Lipschitz decomposability – on singularities. In the original proof smoothness was assumed, even though it was only formulated as a condition for the set of double singularities – see Condition 3.3 from [28]).

Condition 4.12. *There exists an $L \in \mathbb{R}$ such that for every $N \in \mathbb{N}$ the singularity set $\bigcup_{|n|\leq N} \mathcal{S}_n$ is 'Lipschitz decomposable' with constant L (cf. Conjecture 4.8).*

[3] $\mathcal{S}^+$ is the set of all singular phase points (tangent or multiple collisions) introuced in Remark 2.9.

Remark 4.13.

- *For the set of singular reflections itself the original property remains true:* $\mathcal{S}_0$ *is a finite collection of smooth compact manifolds of codimension 1.*
- *Condition 4.12 can only be satisfied by semi-dispersing billiards with a finite horizon. However, the infinite horizon case can easily be reduced to the finite horizon case (cf. [13], [28]).*
- *As to the original exposition, one more condition was assumed on the geometry of the scatterers (the regularity of the set of degenerate tangencies - Condition 3.2 in [28]). However, the role of this condition was to guarantee that points belonging to two different smooth components of the singularity set belong to finitely many codimension 2 submanifolds. Now instead of smooth components we have locally Lipschitz graphs and it is enough to require that the* δ*-neighbourhoods of their boundaries have a volume of* $o(\delta)$*, which is a little less than being two codimensional (cf. Definition 4.6).*

To formulate the Fundamental Theorem, we introduce the notion of regular coverings. Note that $m = 2d - 2$ is the dimension of the (Poincaré) phase space M. The next definition will not be absolutely precise, for we omit some technical details for the sake of easier understanding. For a precise formulation please see Definition 3.4 in [28]).

Definition 4.14. *Let us assume that for a point* $x \in M^*$ *and its neighbourhood* $U(x)$ *a smooth foliation* $U(x) = \cup_{\alpha \in B^{d-1}} \Gamma_\alpha$ *is given. The foliae* Γ_α *are* $d-1$*-dimensional manifolds uniformly transversal to all possible local stable manifolds (*B^{d-1} *is the standard* $d-1$*-dimensional open ball).*

The parameterized family of finite coverings

$$\mathcal{G}^\delta = \{G_i^\delta \mid i = 1, ..., I(\delta)\ \} \qquad 0 < \delta < \delta_0$$

is a family of regular coverings *iff:*

1. *each* G_i^δ *is an open parallelepiped of dimension* $2d-2$*;*
2. *the* $d-1$*-dimensional faces of* G_i^δ *are all cubes with edge-length* δ*, moreover, they may belong to two different categories: the* s-faces *are 'parallel with leaves of the stable foliation' while the* Γ-faces *are 'parallel' with the leaves of the foliation* Γ*;*[4]
3. *For any point, the maximal number of parallelepipeds covering it is* 2^{2d-2}*;*
4. *if* $G_i^\delta \cap G_j^\delta \neq \emptyset$*, then*
$$\mu(G_i^\delta \cap G_j^\delta) \geq c_1 \delta^{2d-2}$$
with c_1 *independent of* δ*.*

Some further convention:

[4]More precisely if we consider the center of each parallelepiped $w_i^\delta \in G_i^\delta$, the s- and Γ- faces are parallel to the tangent planes $\mathcal{T}_{w_i^\delta}\gamma^s(w_i^\delta)$ and $\mathcal{T}_{w_i^\delta}\Gamma(w_i^\delta)$, respectively.

Definition 4.15. *Given any G_i^δ its* s-jacket, *$\partial^s(G_i^\delta)$ is the union of those $(2d-3)$-dimensional faces of G_i^δ which contain at least one s-face of it. The Γ-jacket, $\partial^\Gamma(G_i^\delta)$ is defined similarly. Clearly, $\partial(G_i^\delta) = \partial^s(G_i^\delta) \cup \partial^\Gamma(G_i^\delta)$. We say that a stable manifold* $\gamma^s(y)$ intersects G_i^δ correctly *if:*

$$\partial(G_i^\delta \cap \gamma^s(y)) \subset \partial^\Gamma(G_i^\delta).$$

Theorem 4.16. ***(The Fundamental Theorem)*** *We assume that:*
- *conditions 4.11 and 4.12 are satisfied;*
- *a sufficient phase point $x \in M^*$ is given;*
- *a smooth transversal foliation Γ in a neighbourhood U_0 of x is fixed;*
- *a constant $0 < \epsilon_1 < 1$ is chosen.*

Then there is a sufficiently small neighbourhood $U_{\epsilon_1}(x)$ such that for any $U(x) \subset U_{\epsilon_1}(x)$ and for any family of regular coverings, the covering $\mathcal{G}^\delta$ can be divided into two disjoint subsets, $\mathcal{G}_g^\delta$ and $\mathcal{G}_b^\delta$ (called 'good' and 'bad'), in such a way that:
(I) For any $G_i^\delta \in \mathcal{G}_g^\delta$ and any s-face E^s of it, the set:

$$\{y \in G_i^\delta \mid \rho(y, E^s) < \epsilon_1\delta \text{ and } \gamma^s(y) \text{ intersects correctly}\}$$

has positive relative μ-measure in G_i^δ.
(II)

$$\mu\left(\bigcup_{G_i^\delta \in \mathcal{G}_b^\delta} G_i^\delta\right) = o(\delta).$$

Remark 4.17. *With suitable modifications of the proof the theorem applies to all sufficient points $x \in M^*$ (see [28]), however, for simplicity here we restrict ourselves to regular phase points.*

As mentioned before, the statement of local ergodicity follows from this theorem through the application of Hopf's chain method, see e.g. [13]. This way we get

Theorem 4.18. ***(Local ergodicity theorem) Suppose Conditions 4.11 and 4.12 are satisfied and x is a sufficient phase point. Then, there is a neighbourhood U of x such that every point of U belongs to the same ergodic component.***

4.3.2 Proof of the Fundamental Theorem

Here we would like to give a sketch of the proof following [28]. For brevity we do not repeat the whole argument. Our aim is to emphasize the main ideas on the one hand and point out those parts where the original proof is to be modified on the other hand. Several arguments apply word by word, as to these, we do not give an exposition, just refer to the original paper. Those steps that need non-trivial modification are emphasized and worked out in detail.

Throughout the section we think of the sufficient point $x \in M^0$ and its neighbourhood U as being fixed. y usually denotes some point in U. Furthermore, a sufficiently small δ is kept fixed – thus we work with one particular covering $\mathcal{G}^\delta$. Of course, for every $G_i^\delta \in \mathcal{G}^\delta$ we have $\text{diam}(G_i^\delta) \leq m\delta$, where $m = 2d - 2$ is the dimension of the Poincaré phase space M. As a preparation for the main argument we state two important lemmas:

Lemma 4.19. *In correspondence with condition 4.12 let us denote the Lipschitz components of* $\cup_{|n|\le N}\mathcal{S}_n$ *with* $\mathcal{S}^i$ *(*$i = 1, ..., K$*), rest with* $\mathcal{S}^*$*, and the Lipschitz-constant with* L*. Consider the set*

$$\Delta_{\delta,N} := \{x \mid \exists i, j \le K, \ , i \ne j, \ \rho(x, \mathcal{S}^i) \le \delta, \ \rho(x, \mathcal{S}^j) \le \delta\} \cup \{x \mid \rho(x, \mathcal{S}^*) \le \delta\}.$$

For all N*:*

$$\mu(\Delta_{\delta,N}) = o(\delta).$$

This lemma plays essentially the role of Lemma 4.6 from [28]. However, the proof of it is the first point where the original proof of the Fundamental Theorem had to be modified.

Proof. Fix an index i and find a coordinate system so that

$$\mathcal{S}^i = \{(x, f_i(x)) | \ x \in D_i\}.$$

where $D_i \subset \mathbb{R}^{m-1}$. π_i shall denote the usual projection onto $\mathbb{R}^{m-1}$: $\pi_i((x, y)) := x$, when $x \in \mathbb{R}^{m-1}$ and $y \in \mathbb{R}$. Obviously $\partial D_i = \pi_i(\partial\mathcal{S}^i)$, so $\mathcal{L}^m\left((\partial\mathcal{S}^i)^{[\delta]}\right) \ge 2\delta\mathcal{L}^{m-1}(\partial D_i)$. So the condition $\mathcal{L}^m\left((\partial\mathcal{S}^i)^{[\delta]}\right) = o(\delta)$ implies $\mathcal{L}^{m-1}(\partial D_i) = 0$.

As a consequence, for any $\eta > 0$ it is possible to find $\eta' > 0$ such that the (closure of the) open η'-neighbourhood of ∂D_i inside D_i has $\mathcal{L}^{m-1}$-measure less than η. Let us denote this open neighbourhood by D^i_η and furthermore

$$\Delta^i_\eta = \left\{(x, f_i(x)) \mid x \in D^i_\eta\right\}.$$

Now consider the parts of the singularity far away from the borders of the singularities. For different i-s the sets $\mathcal{S}^i \backslash \Delta^i_\eta$ $(i = 1, ..., K)$ are pairwise disjoint compact sets (as they are continuous images of compact sets). Consequently, for δ small enough the sets $(\mathcal{S}^i \setminus \Delta^i_\eta)^{[\delta]}$ are pairwise disjoint as well. Now for the set mentioned in the Lemma, we can write:

$$\Delta_{\delta,N} \subset \Delta^{[\delta]}_\eta \cup (\mathcal{S}^*)^{[\delta]}$$

where $\Delta_\eta = \bigcup_{i=1}^{K} \Delta^i_\eta$.

Now apply Lemma 4.9 to get

$$\mathcal{L}^m((\Delta^i_\eta)^{[\delta]}) \le 2\sqrt{L^2 + 1}\delta\eta + o(\delta).$$

This means that

$$\mathcal{L}^m(\Delta^{[\delta]}_\eta) \le 2K\sqrt{L^2 + 1}\delta\eta + o(\delta)$$

stands for every $\eta > 0$, meaning that $\mathcal{L}^m(\Delta^{[\delta]}_\eta) = o(\delta)$. Together with $\mathcal{L}^m((\mathcal{S}^*)^{[\delta]}) = o(\delta)$ this gives the statement of the lemma. □

Before formulating the other key lemma, we would like to note that – for Lipschitz-continuous functions are differentiable almost everywhere – in almost every point of a singularity component (i.e. in one particular open locally Lipschitz graph $\hat{\mathcal{S}}$) it makes sense to talk about their (one codimensional) tangent planes $\mathcal{T}_{\hat{y}}\hat{\mathcal{S}}$. Knowing the behaviour of the tangent plane wherever it exists allows us to think about the “direction” of the whole $\hat{\mathcal{S}}$.

Lemma 4.20. *Given any $x \in M^0$ and any $\epsilon > 0$ there is a neighbourhood $U(x) \subset M$ of x such that for any two local stable manifolds γ_1^s, γ_2^s and any $(2d-3)$-dimensional Lipschitz component $\hat{\mathcal{S}}$ of some $\mathcal{S}_n$ $(n > 0)$ intersecting $U(x)$ with points y_1, y_2 and $\hat{y}$, lying on the three manifolds, respectively, so that $\mathcal{T}_{\hat{y}}\hat{\mathcal{S}}$ exists:*

$$\begin{aligned} \sphericalangle(\mathcal{T}_{y_1}\gamma_1^s, \mathcal{T}_{y_2}\gamma_2^s) &< \epsilon, \\ \sphericalangle(\mathcal{T}_{y_1}\gamma_1^s, \mathcal{T}_{\hat{y}}\hat{\mathcal{S}}) &< \epsilon. \end{aligned} \tag{4.2}$$

This Lemma is on the parallelization effect and it is exactly the same as Lemma 4.9 in [28] – the original argument applies. Nevertheless it might be useful to point out what the second inequality in (4.2) means: there is a $(d-1)$-dimensional subspace of the tangent space at almost any point of the $(2d-3)$-dimensional manifold $\hat{\mathcal{S}}$ very close to the stable subspace. Note, however, that $\hat{\mathcal{S}}$ may behave extremely widely – i.e. in a non-smooth manner – in the remaining $(d-2)$ dimensions (in case $d \geq 3$).

Notation. *The following two quantities measure the hyperbolicity near the point $y \in M^0$. Let*

$$\kappa_{n,0}(y) = \inf_{\Sigma} \|(D^n_{-T^n y, \Sigma})^{-1}\|_p^{-1},$$

where the inf *is taken over all convex local orthogonal manifolds passing through $-T^n y$. Furthermore denote*

$$\kappa_{n,\delta}(y) = \inf_{\Sigma} \inf_{w \in \Sigma} \|(D^n_{w,\Sigma})^{-1}\|_p^{-1}.$$

Here the infimum is taken for the set of convex fronts Σ passing through $-T^n y$ such that (i) T^n is continuous on Σ and (ii) $T^n \Sigma \subset B_\delta(-y)$.

Remark 4.21. *(cf. Lemma 5.3 in [28] and Lemma 4.2 in the present thesis) It is not difficult to see that $\kappa_{n,\delta}(y)$ is an increasing function of n. Furthermore, for sufficient points y clearly:*

$$\lim_{n \to \infty} \kappa_{n,0}(y) = \infty.$$

(Here we do not state in general that $\kappa_{n,0}$ grows exponentially, linear growth – which is obvious for sufficient points y – is enough.)

The following subsets of the neighbourhood $U \ni x$ depend on the constants δ and c_3.

$$\begin{aligned} U^g &:= \{y \in U \mid \forall n \in \mathbb{Z}_+,\ z_{tub}(T^n y) \geq (\kappa_{n,c_3\delta}(y))^{-1} c_3 \delta\}; \\ U^b &:= U \setminus U^g; \\ U^b_n &:= \{y \in U \mid z_{tub}(T^n y) < (\kappa_{n,c_3\delta}(y))^{-1} c_3 \delta\} \end{aligned} \tag{4.3}$$

Remark 4.22. *Note that for the points $y \in U^g$ the stable manifold extends to the boundary of $B_{c_3\delta}(y)$, the ball of radius $c_3\delta$ around y (cf. Lemma 5.4 from [28]). The constant c_3 will be chosen in an appropriate way to guarantee that for any $y \in U^g \cap G_i^\delta$ the stable leaf $\gamma^s(y)$ intersects G_i^δ correctly unless it intersects $\partial^s G_i^\delta$.*

Furthermore, we introduce the class of permitted functions.

Definition 4.23. *A function $F : \mathbb{R}_+ \to \mathbb{Z}_+$ defined in a neighbourhood of the origin is permitted whenever $F(\delta) \nearrow \infty$ as $\delta \searrow 0$. For a fixed permitted function $F(\delta)$ we define $U^b_\omega = \cup_{n > F(\delta)} U^b_n$.*

Most of the statements to come hold for any permitted function $F(\delta)$. At one point of the argument we shall fix one particular $F(\delta)$.

Lemma 4.24. (**Tail bound**; Lemma 6.1 from [28]). *For any permitted function:*

$$\mu(U^b_\omega) = o(\delta).$$

The measure estimates in the proof of the Tail Bound are related to $\mathcal{S}_0$, the set of singular reflections. As already mentioned, this set (in contrast to the higher iterates $\mathcal{S}_n$) is a finite collection of smooth and compact one codimensional submanifolds of the phase space. Consequently, there is no need for Lipschitz decomposition here, thus we do not include the proof. Essentially, the original argument from [28] applies, nevertheless, at the definition of the small set of non-sufficient points a little more care is needed. We would also like to emphasize that the proof of the Tail Bound is the point where the Chernov-Sinai Ansatz (Condition 4.11) is exploited. On more details see [28].

Remark 4.25. *In what follows we will work with distances defined by the Euclidean metric ρ. However, as the* interiour *point x in M is fixed and its neighbourhood $U(x)$ is fixed we have $|\langle v, n(q)\rangle| \geq c$ for some positive constant $c = c(x)$ in this neighbourhood. Thus the two distances ρ and ρ_i are equivalent (cf. Section 2.6).*

Now we can start proving the Fundamental Theorem by telling explicitly how the collection of parallelepipeds $\mathcal{G}^\delta$ is divided into a good and a bad part. We say $G^\delta_i \in \mathcal{G}^\delta_b$ iff
(A) either

- it intersects more than one Lipschitz component of $\mathcal{S}^{F(\delta)}$ (the singularities of $T^{F(\delta)}$),
- or it intersects only one component $\hat{\mathcal{S}}$, but $\rho(G^\delta_i, \partial\hat{\mathcal{S}}) \leq \delta$,
- or it intersects the remaining small set $\mathcal{S}^*$.

(B) or it is not of type (A), but it has an s-face E^s such that

$$\mu(G^\delta_i \cap (E^s)^{[\epsilon_1\delta]} \cap U_{ic}) \leq \frac{\epsilon_3}{4}\mu(G^\delta_i), \tag{4.4}$$

where ϵ_3 is a positive constant to be defined later and U_{ic} is the set of points in G^δ_i with correctly intersecting local stable manifolds.

Now we choose one particular permitted function $F(\delta)$: by virtue of Lemma 4.19 there definitely exists a permitted function such that:

$$\mu(\Delta_{(m+1)\delta,F(\delta)}) = o(\delta). \tag{4.5}$$

(Remember that $m = 2d - 2$ is the dimension of the Poincaré phase space.)

As a consequence the overall measure of bad parallelepipeds of type (A) is $o(\delta)$ (such parallelepipeds lie inside the set $\Delta_{(m+1)\delta,F(\delta)}$).

It is time to tell about the choice of our small constants ϵ_i as well. In the formulation of the Fundamental Theorem one particular constant ϵ_1 is given. We shall choose three further constants in the following order: $\epsilon_1 \to \epsilon_3 \to \epsilon_4 \to \epsilon_2$. It is utmost important that all of these choices are independent of δ. (they are chosen in the arguments below, ϵ_3 in

1., ϵ_4 in 2. and ϵ_2 in 3.). After all these choices are made we fix the neighbourhood $U_{\epsilon_1}(x)$ (see the formulation of the Fundamental Theorem) in such a way that for all Lipschitz components $\hat{\mathcal{S}}$ of some $\mathcal{S}_n$ ($n > 0$) that intersect $U_{\epsilon_1}(x)$:

$$\begin{aligned} \sphericalangle(\gamma_1^s, \gamma_2^s) &< \epsilon_2, \\ \sphericalangle(\gamma_1^s, \hat{\mathcal{S}}) &< \epsilon_2. \end{aligned} \tag{4.6}$$

Such a choice is clearly possible by virtue of Lemma 4.20. Here the second inequality is understood at every point of $\hat{\mathcal{S}}$ where it makes sense, that is, where $\hat{\mathcal{S}}$ is differentiable.

One more remark: having fixed the neighbourhood $U_{\epsilon_1}(x)$ and the foliation Γ uniformly transversal to the stable foliation, it is possible to uniformly compare two different measures for each product-type set inside $U_{\epsilon_1}(x)$. More precisely there is a constant $c_4 > 0$ such that given any product-type set, the ratio of its μ-measure and its measure that arises as a product of measures in the $s-$ and $\Gamma-$directions lies between c_4^{-1} and c_4.

From now on G_i^δ will always denote a bad parallelepiped of type (B). The proof of the Fundamental Theorem follows from the small arguments to come.

1. Let us first give an estimate from below on the measure of $G_i^\delta \cap (E^s)^{[\epsilon_1\delta]}$ where E^s is an s-face for the bad parallelepiped G_i^δ. By the above remark on product measures:

$$\mu(G_i^\delta \cap (E^s)^{[\epsilon_1\delta]}) \geq c_4^{-1}(\epsilon_1\delta)^{d-1}\delta^{d-1} \geq c_6\epsilon_1^{d-1}\mu(G_i^\delta) \geq \epsilon_3\mu(G_i^\delta),$$

in case $\epsilon_3(\epsilon_1)$ is chosen sufficiently small.

2. For estimates from above we fix the constant $\epsilon_4 = \epsilon_4(\epsilon_3)$ sufficiently small. The measure of points near the s-jacket (which consists of $2(d-1)$ faces of dimension $2d-3$), is:

$$\mu(G_i^\delta \cap (\partial^s G_i^\delta)^{[\epsilon_4\delta]}) \leq 2(d-1)c_4\epsilon_4\delta\delta^{2d-3} \leq \frac{\epsilon_3}{4}\mu(G_i^\delta).$$

We need one more estimate of similar type. *This is the second point where the original proof has to be modified, and the smoothness/Lipschitzness of singularity components is used.* Recall that for a bad parallelepiped of type (B) there is at most one Lipschitz component $\hat{\mathcal{S}}$ of the singularity set $\mathcal{S}^{F(\delta)}$ for $T^{F(\delta)}$ intersecting it. We are interested in estimating the measure of the $\epsilon_4\delta$-neighbourhood of this Lipschitz graph inside the parallelepiped. If $\epsilon_4 < \frac{1}{\sqrt{L^2+1}+1}$ (L is the Lipschitz-constant), then, by the construction of type (A) parallelepipeds, Lemma 4.10 can be applied, and gives

$$\mu(G_i^\delta \cap (\hat{\mathcal{S}})^{[\epsilon_4\delta]}) \leq c_5\epsilon_4\delta\delta^{2d-3} \leq \frac{\epsilon_3}{4}\mu(G_i^\delta).$$

whenever again $\epsilon_4(\epsilon_3)$ is small enough.

3. Now we choose $\epsilon_2(\epsilon_4)$ small enough, so that by (4.6) stable manifolds and singularity components are 'almost parallel'. Namely, the smallness of ϵ_2 should guarantee that for any $y \in G_i^\delta$ for which $\gamma^s(y)$ does not intersect correctly we have:

$$y \in (G_i^\delta \cap (\hat{\mathcal{S}})^{[\epsilon_4\delta]}) \cup (G_i^\delta \cap (\partial^s G_i^\delta)^{[\epsilon_4\delta]}) \cup U_\omega^b. \tag{4.7}$$

To see that, given a suitable choice of ϵ_2, the above formula is valid, we make two remarks.

- First we note that for stable manifolds and singularity components 'not to approach each other too quickly', being 'almost parallel' is enough at almost every point of $\hat{\mathcal{S}}$.
- Recalling the definitions from (4.3) and the various notions of distances from Section 2.6 what we see immediately is that the inclusion of (4.7) is valid with writing $G_i^\delta \cap (\cup_{n \le F(\delta)} U_n^b)$ instead of $G_i^\delta \cap \hat{\mathcal{S}}^{[\epsilon_4 \delta]}$ for the first set. Nevertheless, with a suitable choice of ϵ_2 we certainly have $G_i^\delta \cap (\hat{\mathcal{S}}^{[\epsilon_4 \delta]}) \subset G_i^\delta \cap (\cup_{n \le F(\delta)} U_n^b)$ as (i) $z_{tub}(x) \ge z(x)$ and (ii) the Euclidean distance ρ and the distance ρ_i (in terms of which $z(x)$ is defined) are equivalent, see Remark 4.25.

We only need some minor considerations to complete the proof. Observe first that for good parallelepipeds the statement (I) evidently holds. As for (II) we have already shown that bad parallelepipeds of type (A) are of measure $o(\delta)$ (recall (4.5)), we shall show the same for those of type (B) as well. Indeed, let us consider a G_i^δ with an s-face E^s for which (4.4) holds. By the arguments 1.-3. above:

$$\mu(G_i^\delta \cap U_\omega^b) \ge \mu(G_i^\delta \cap (E^s)^{[\epsilon_1 \delta]} \cap U_\omega^b) \ge \frac{\epsilon_3}{4} \mu(G_i^\delta).$$

Now recall that in a regular covering there are at most 2^{2d-2} parallelepipeds with a non-empty common intersection. Thus:

$$2^{2d-2} \mu(U_\omega^b) \ge \sum{}' \mu(G_i^\delta \cap U_\omega^b) \ge \frac{\epsilon_3}{4} \sum{}' \mu(G_i^\delta),$$

where $\sum'$ denotes the sum over bad parallelepipeds of type (B). By the Tail Bound (Lemma 4.24) we have $\sum' \mu(G_i^\delta) = o(\delta)$ thus the proof of Theorem 4.16 is complete. □

4.4 The case of algebraic scatterers

The main aim of this section is to show that the singularity submanifolds of algebraic semi-dispersing billiards satisfy the Lipschitz decomposability property formulated in Conjecture 4.8. Fortunately, the most important examples of semi-dispersing billiards are algebraic. Consequently, the algebraicity condition does not essentially restrict the applicability of the Fundamental Theorem.

For definiteness we will say that the zero-set of a system of polynomial equations is an *algebraic variety* (we will use these notions over the real ground field). Any (measurable) subset of a k-dimensional algebraic variety will be denoted as a k-dimensional SSAV (for 'subset of an algebraic variety'). As for the dimension of an algebraic variety, see [33]. We also use the following definition.

Definition 4.26. *A semi-dispersing billiard is algebraic if it has finitely many scatterers and the boundary of each of these scatterers is a finite union of one codimensional SSAV-s (as subsets of $\mathbb{T}^d \subset \mathbb{R}^d$).*

Remark 4.27. *Assume, in general, that we are given a Riemannian manifold $\mathbb{M} = \mathbb{M}_m$ and a subset $A \subset \mathbb{M}$. We say that A is a k-dimensional weakly algebraic subset of $\mathbb{M}$ if it is possible to find an appropriate atlas $\{U_t, \psi_t\}_{t=1}^T$ on $\mathbb{M}$ such that, for every t, $\psi_t(U_t \cap A)$ $(\subset \mathbb{R}^m)$ is a k-dimensional SSAV in $\mathbb{R}^m$. Bi-Lipschitzness of the atlas $\{U_t, \psi_t\}_{t=1}^T$ can always be assumed (cf. Remark 4.7)*

Note that being 'weakly algebraic' is really a weak notion due to the high degree of freedom in the choice of the atlas. For example, every smooth curve is 1-dim. weakly algebraic.

What follows below in three subsections is a proof of Lipschitz decomposability for the singularities $\mathcal{S}_n$ in an algebraic billiard. In Section 4.4.1 it is shown that singularities are algebraic as subsets of $\mathbb{R}^{2d}$. This implies that $\mathcal{S}_n \subset \partial M$ is algebraic in the sense of Remark 4.27 as well.[5] The proof is completed in Sections 4.4.2 and 4.4.3 where a Lipschitz decomposition is constructed for any (one codimensional) SSAV of $\mathbb{R}^m$.

4.4.1 The algebraicity of $\mathcal{S}_n$

Our approach generalizes that of section 3 of [36]. Since there a detailed exposition was given, here we are satisfied by referring to the main steps of the complexification of the dynamics. Still we completely explain those parts where our arguments are different.

In a nutshell the picture in [36] is the following.

- the authors were only considering quadratic boundaries since hard ball systems are quadratic billiards;
- and for the quadratic case they elaborated a most detailed algebraic analysis of the situation.

Here we do not need such a delicate picture. But on the other hand, we are treating the general algebraic case. The chain of field extensions of [36] relied upon the explicit solvability of the arising quadratic equations and applied the related elimination of the square roots. In the general case we rather apply the norm used in Galois theory.

We first fix some notation – slightly different from the usual conventions – at this point. According to Definition 4.26 above, $\partial Q = \cup_{j=1}^{J} \partial Q_j$, where both the components ∂Q_j and their boundaries are all appropriate dimensional SSAV-s (the decomposition is finer than the one into connected components in $\mathbb{R}^d$). In other words, for each ∂Q_j there is a (non-zero) irreducible polynomial $B_j(q)$ such that

$$\partial Q_j \subset \{q \in \mathbb{R}^d \mid B_j(q) = 0\}.$$

Note that symbolic collision sequences (4.8) are defined in terms of these algebraic boundary components as well.

From this point on it will be suitable to consider orbit segments $S^{[0,T]}x_0, T > 0$ of the billiard flow with T sufficiently large. In fact, it will be useful to also drop the condition $||v|| = 1$. Consequently, the dimension of our phase space will be $2d$ (first the phase space will be $\mathbb{T}^d \times \mathbb{R}^d$ and later just $\mathbb{R}^{2d}$).

The symbolic collision sequence of $S^{[0,T]}x_0$ will be denoted by

$$\sigma = \Sigma(S^{[0,T]}x_0) = (\sigma_1, \sigma_2, \dots, \sigma_n) \qquad (n \geq 0) \tag{4.8}$$

[5] In a small neighbourhood of $y \in M$ identify the tangent plane $\mathcal{T}_y M$ with $\mathbb{R}^m$ and restrict the orthogonal projection $\Pi : \mathbb{R}^{2d} \to \mathcal{T}_y M$ onto M to obtain coordinate charts.

Remark 4.28. *By definition, (q_0, v_0) corresponds to the initial, generally non-collision phase point x_0 of the flow. Furthermore $T^k x_0 = x_k = (q_k, v_k) \in \partial Q_{\sigma_k}$ for every $1 \le k \le n$ (we note that for a phase point $x_0 \notin M$ of the flow $Tx_0 \in M$ coincides by definition with the first point where the positive semi-orbit of x_0 reaches the boundary M; in [28] this map was denoted by T^+). By a slight abuse of notation we will keep denoting by $\mathcal{S}_n$ (introduced in Section 2.4) the n-th inverse image of $\mathcal{S}_0$ in a $2d$-dimensional neighbourhood of x_0 .*

Having fixed σ, we first explore the algebraic relationship between the consecutive x_ks. For being able to carry out arithmetic operations on our data, we lift the genuine orbit segment to the covering Euclidean space of the torus. This can be done by a straightforward generalization of the trivial Proposition 3.1 of [36].

Proposition 4.29. *Let $S^{[0,T]}x_0$ be an orbit segment of the discretized dynamics. Assume that a certain pre-image (Euclidean lifting) $\tilde{q}_0 \in \mathbb{R}^d$ of the position $q_0 \in \mathbb{T}^d$ is given. Then there is a uniquely defined Euclidean lifting $\{\tilde{q}_i \in \mathbb{R}^d | 0 \le k \le n\}$ of the given orbit segment which, when considered in continuous time, is a time-continuous extension of the original lifting $\tilde{q}_0$. Moreover, for every collision σ_k there exists a uniquely defined integer vector $a_k \in \mathbb{Z}^d$ – named the adjustment vector of σ_k – such that*

$$B_{\sigma_k}(\tilde{q}_k - a_k) = 0 \qquad (1 \le k \le n).$$

The orbit segment $\tilde{\omega} = \{\tilde{q}_k | 0 \le k \le n\}$ is called the lifted orbit segment with the system of adjustment vectors $\mathcal{A} = (a_0, \dots, a_n) \in \mathbb{Z}^{(n+1)d}$.

In the sequel, $\langle , \rangle$ denotes Euclidean inner product of d-dimensional real vectors. Our next proposition is also a straightforward extension of Proposition 3.3 of [36].

Proposition 4.30. *Between the kinetic data corresponding to σ_{k-1} and σ_k one has the following algebraic relations:*

— the linear collision equation

$$v_k = v_{k-1} - 2\langle v_{k-1}, n_k \rangle n_k \qquad (1 \le k \le n) \tag{4.9}$$

where n_k is the outer unit normal vector of the scatterer Q_{σ_k} at the point of impact;

— the linear free flight equation

$$\tilde{q}_k = \tilde{q}_{k-1} + \tau_k v_{k-1} \qquad (1 \le k \le n) \tag{4.10}$$

— where the time slot $\tau_k = t_k - t_{k-1}$ $(t_0 = 0)$ in (4.10) is determined by the polynomial equation

$$B_{\sigma_k}(\tilde{q}_{k-1} + \tau_k v_{k-1} - a_k) = 0. \tag{4.11}$$

Next we turn to the complexification of the billiard map T. Given the pair $(\Sigma, \mathcal{A}) = (\sigma_1, \sigma_2, \dots, \sigma_n; a_0, a_1, \dots, a_n)$, the equations (4.9), (4.10), (4.11) make it possible to algebraically characterize the kinetic data $(\tilde{q}_k, v_k)$ by using the preceding data $(\tilde{q}_{k-1}, v_{k-1})$. Since – at the moment – we are dealing with genuine, real orbit segments, in this situation the equations have at least one positive, real root τ_k; in case of several such roots its selection is unique by the geometry of the problem. Our further arguments, however, also use the algebraic closedness of the arising fields and therefore we complexify the dynamics. From this point on, our approach, though related but nevertheless will already be different from that of [36].

Definition 4.31. *For $n = 0$ the field $\mathbb{K}_0 = \mathbb{K}(\emptyset; \emptyset)$ is the transcendental extension $\mathbb{C}(\mathcal{B})$ of the coefficient field $\mathbb{C}$ by the algebraically independent formal variables*

$$\mathcal{B} = \{(\tilde{q}_0)_j, (v_0)_j | 1 \le j \le d\}$$

Suppose now that the commutative field $\mathbb{K}_{n-1} = \mathbb{K}(\Sigma'; \mathcal{A}')$ has already been defined, where $\Sigma' = (\sigma_1, \sigma_2, \dots, \sigma_{n-1}); \mathcal{A}' = (a_0, a_1, \dots, a_{n-1})$. Consider now the polynomial equation

$$b_l \tau^l + b_{l-1} \tau^{l-1} + \ldots + b_0 = 0 \tag{4.12}$$

arising from (4.11) with k=n. It defines a new field element τ_n to be adjoined to the field $\mathbb{K}_{n-1}$ (of course, $b_0, \dots, b_l \in \mathbb{K}_{n-1}$). At this point, however, we should be a bit cautious. If the equation 4.12 is irreducible, then all its roots are algebraically equivalent, and τ_n can denote any of them. If (4.11) is reducible, then we should select a particular irreducible factor of its. Indeed, since we are only interested in the images of $\mathcal{S}$, at each step we choose such an irreducible factor of (4.12) which, when its root τ_n gets evaluated for real values of x_0, gives us a real root of (4.12) which is actually the real root specified after (4.11). This irreducible factor defines the extension $\mathbb{K}_n = \mathbb{K}_{n-1}(\tau_n)$.

In such a way we are given a chain of extensions $\mathbb{K}_0, \mathbb{K}_1, \dots, \mathbb{K}_n$ where for every $k = 1, \dots, n$ the relation $\mathbb{K}_k = \mathbb{K}_{k-1}(\tau_k)$ holds. By our construction and by the theorem on the prime element of algebra, $\mathbb{K}_n$ can also be expressed as $\mathbb{K}_0(\tilde{\tau}_n)$ for some $\tilde{\tau}_n \in \mathbb{K}_n$ with minimal polynomial $m(\alpha)$ over $\mathbb{K}_0$.

By applying the previous construction we are going to look for an algebraic characterization of $\mathcal{S}_n$.

For every $x_0 \in M_{\Sigma,\mathcal{A}} = \{x \in M \mid \Sigma(S^{[0,T]}x) = \Sigma,\ \mathcal{A}(S^{[0,T]}x) = \mathcal{A}\}$ one has $\tilde{q}_n \in \mathbb{K}_n$. $\tilde{q}_n$ can formally be understood as a function $\tilde{q}_n(x_0, \tau_1, \dots, \tau_n)$ or (by the theorem on the prime element) simply as a function $\tilde{q}_n(x_0, \tilde{\tau}_n)$ with values in $\mathbb{K}_n$. We will be considering this function exactly in $M_{\Sigma,\mathcal{A}}$, that is, where Σ and $\mathcal{A}$ are constants. Consider Q_{σ_n} at the point $T^n x_0$. At this point the submanifold $B_{\sigma_n}(\tilde{q}_n - a_n) = 0$ has a normal vector $\vec{n}$ which can be expressed by the partial derivatives of B_{σ_n} at $\tilde{q}_n - a_n$. The condition $\tilde{x}_n \in \mathcal{S}$ just says that $\langle \vec{n}, v_n \rangle = 0$. Here both $\vec{n}$ and v_n are elements of $\mathbb{K}_n$, i. e. formal functions of x_0 and $\tilde{\tau}_n$. Consequently $\langle \vec{n}, v_n \rangle = \Phi(\tilde{\tau}_n)$ where Φ is a polynomial whose coefficients are rational functions over $\mathbb{K}_0$. Take now the (Galois-) norm (cf. [39]) of this element, i.e.

$$\|\Phi\| = \Pi \Phi(\tilde{\tau}_n^i)$$

where the product is taken for all roots $\tilde{\tau}_n^i$ of the irreducible polynomial m. This norm does not vanish since it is the product of non-zero elements in the normal hull of $\mathbb{K}_n$. Moreover, it is a symmetric polynomial of the elements $\tilde{\tau}_n^i$. As such it can be expressed as a polynomial of the elementary symmetric polynomials of the variables $\tilde{\tau}_n^i : 1 \le i \le l$. These elementary symmetric polynomials can, however, be easily expressed by the coefficients of m which are elements of $\mathbb{K}_0$. As a consequence, we obtain a non-zero element of $\mathbb{K}_0$. The construction just described generalizes the elimination of the square roots method applied in [36]. By our construction it remains also true that this polynomial has real coefficients for our real, dynamical orbit. All in all for every fixed Σ and $\mathcal{A}$ the resulting piece of $\mathcal{S}_n$ is an algebraic submanifold. From the finiteness of the horizon it is clear that in the case of our real dynamics only a finite number of Σ and $\mathcal{A}$ provide a non-empty piece of $\mathcal{S}_n$.

In this way we have established

Theorem 4.32. *$\mathcal{S}_n$ is a finite union of one codimensional SSAV-s in $\mathbb{R}^{2d}$.*

4.4.2 Dimension and measure of algebraic varieties

The motivation for this section is that we need to estimate the Lebesgue-measure (denoted here by $\mathcal{L}^m$) of the δ-neighbourhood of an algebraic variety. Actually we only need that $\mathcal{L}^m(H^{[\delta]}) = o(\delta)$ if H is (at least) two codimensional, but our results will be more general than that.

As we will see, this problem is closely related to the box dimension and the so-called Minkowski-content of H (on box dimension, Minkowski-content and their relation to Hausdorff dimension and measure, see Section 3.1 in [19]). To start, let us recall some notions and basic facts related to box dimension.

Definition 4.33. *Let H be a bounded subset of $\mathbb{R}^m$, $0 \le d \in \mathbb{R}$. Then the quantities*

$$\overline{\mathcal{M}}^d(H) := \limsup_{\delta\to 0} \frac{\mathcal{L}^m(H^{[\delta]})}{\delta^{m-d}},$$

$$\underline{\mathcal{M}}^d(H) := \liminf_{\delta\to 0} \frac{\mathcal{L}^m(H^{[\delta]})}{\delta^{m-d}}$$

are called the upper and lower d-dimensional Minkowski-content of H.

Definition 4.34. *Let H be a bounded subset of $\mathbb{R}^m$, $\epsilon > 0$. The set $I \subset H$ is called an ε-net in H if $H \subset I^{[\varepsilon]}$.*

We will always be interested in finite ε-nets I, and we will never use that $I \subset H$.

Notation. $\dim_H H$ *and* $\dim_B H$ *will denote the Hausdorff and box dimensions of H, respectively. For $0 \le d \in \mathbb{R}$ we will denote the d-dimensional Hausdorff-measure by $\mathcal{H}^d$.*

Some simple facts:

(a) $\dim_H H \le \underline{\dim}_B H \le \overline{\dim}_B H$

(b) $\mathcal{H}^d(H) \le \underline{\mathcal{M}}^d(H) \le \overline{\mathcal{M}}^d(H)$

(c) If $\overline{\mathcal{M}}^d(H) < \infty$, then $\overline{\dim}_B H \le d$.

(d) If $\dim_H H < d$ then $\mathcal{H}^d(H) = 0$.

(e) If $\overline{\mathcal{M}}^d(H) < \infty$ then $\mathcal{L}^m(H^{[\delta]}) = O(\delta^{m-d})$.

(f) If I is an ε-net (finite) in $H \in \mathbb{R}^m$ then $\mathcal{L}^m(H^{[\varepsilon]}) \le (2\varepsilon)^m |I|$, where $|I|$ is the cardinality of I.

Now we turn to the investigation of algebraic varieties. Our proposition will be an easy corollary of the following lemma.

Lemma 4.35. *Let $H = \hat{H} \cap [0,1]^m$, where $\hat{H} \subset \mathbb{R}^m$ is an algebraic variety. Let k be the maximum of the degrees of the polynomials defining $\hat{H}$. Let $\varepsilon > 0$, $0 \le d \in \mathbb{Z}$. Let $c > 1$ arbitrary. We claim that if $\mathcal{H}^{d+1}(H) = 0$ then, if ε is small enough, there exists a $(d \cdot \varepsilon)$-net I in H with*

$$|I| \le N_{m,d,k,\varepsilon} := \sum_{i=0}^{d} c^i \left(\frac{m!}{(m-i)!} \right)^{3/2} k^{m-i} \frac{1}{\varepsilon^i}.$$

Proof. The proof goes by induction on d, and the induction is based on the following **Fact:** For every $x \in [0,1]$ let $H_x = H \cap (\{x\} \times [0,1]^{m-1})$. Then $\mathcal{H}^{d+1}(H) = 0$ implies that for Lebesgue almost every $x \in [0,1]$, $\mathcal{H}^d(H_x) = 0$. This is an easy consequence of Theorem 5.8 in [18].

The same is true for subsets of H arising by fixing another (than the first) coordinate: for every $1 \le l \le m$ if $P^l_x := [0,1]^{l-1} \times \{x\} \times [0,1]^{m-l}$ and $H^l_x := P^l_x \cap H$ then we have $\mathcal{H}^d(H^l_x) = 0$ for $\mathcal{L}^1$-a.e. x. We will take advantage of this by choosing ε' arbitrary (later on we will fix $\varepsilon' = \frac{\varepsilon}{\sqrt{m}}$) and fixing $K \le \frac{c}{\varepsilon'}$ points: $0 = x_{l,1} < \ldots < x_{l,K} = 1$, such that $x_{l,j+1} - x_{l,j} \le \varepsilon'$ and $\mathcal{H}^d(H^l_{x_{l,j}}) = 0$ for every j. The $m \cdot K$ hyperplanes $P^l_{x_{l,j}} : \ l = 1, \ldots, m, \ j = 1, \ldots, K$ cut H into blocks of diameter $\le \varepsilon' \sqrt{m}$.

Notice that if H has a point A in any of these blocks, then either it also has one (B) on the surface of the block, so that $\operatorname{dist}(A,B) \le \varepsilon' \sqrt{m}$, or the entire component of H containing A is inside the block.

1. We start the induction with $d = 0$. The previous construction gives (for any ε') $\mathcal{H}^0(H^l_{x_{l,j}}) = 0$, that is, $H^l_{x_{l,j}} = \emptyset$ for every l, j. That is, the components are points, and we can certainly find the $0 \cdot \varepsilon = 0$-net $I = H$ with $|I| \le k^m = N_{m,0,k,\varepsilon}$, an upper bound for the number of components coming from Bezout's theorem.

2. Suppose we have the statement for some $d - 1 \ge 0$.

3. We prove for d. That is, $H \subset [0,1]^m$, $\mathcal{H}^{d+1}(H) = 0$. Apply the previous construction with $\varepsilon' = \frac{\varepsilon}{\sqrt{m}}$. The set $H^l_{x_{l,j}}$ is now an algebraic variety in $[0,1]^{m-1}$, the polynomials defining it can be derived from those defining H by fixing a variable. So the degrees can not grow. We can use the inductive assumption for the $mK \le mc\frac{\sqrt{m}}{\varepsilon}$ sets $H^l_{x_{l,j}}$ with $m \to m-1$ and the same k. Thus taking a $(d-1)\varepsilon$-net on every $H^l_{x_{l,j}}$ according to the inductive assumption, and choosing a point from every component that happens to be entirely inside a block, we get a $d \cdot \varepsilon$-net I in H with $|I| \le mc\frac{\sqrt{m}}{\varepsilon} N_{m-1,d-1,k,\varepsilon} + k^m = N_{m,d,k,\varepsilon}$

□

This lemma leads to the following

Proposition 4.36. *If $H = \hat{H} \cap [0,1]^m$ where $\hat{H} \subset \mathbb{R}^m$ is an algebraic variety, and k is the maximum of the degrees of the polynomials defining $\hat{H}$, then $s := \dim_H(H) = \dim_B(H) \in \mathbb{Z}$ and*

$$0 < \mathcal{H}^s(H) \le \underline{\mathcal{M}}^s(H) \le \overline{\mathcal{M}}^s(H) \le 2^m s^s \left(\frac{m!}{(m-s)!} \right)^{3/2} k^{m-s} < \infty. \tag{4.13}$$

Proof. By (b) we have $\mathcal{H}^s(H) \le \underline{\mathcal{M}}^s(H) \le \overline{\mathcal{M}}^s(H)$. On the other hand, if we choose $d \in \mathbb{Z}$ in such a way that $\mathcal{H}^{d+1}(H) = 0$ then by the Lemma for any $c > 1$:

$$\overline{\mathcal{M}}^d(H) \stackrel{\text{def}}{=} \limsup_{\varepsilon\to 0} \frac{\mathcal{L}^m(H^{[d\varepsilon]})}{(d\varepsilon)^{m-d}} \stackrel{(f)}{\le} \limsup_{\varepsilon\to 0} \frac{(2d\varepsilon)^m N_{m,d,k,\varepsilon}}{(d\varepsilon)^{m-d}} \stackrel{\text{Lemma}}{=} \tag{4.14}$$
$$= \limsup_{\varepsilon\to 0} 2^m d^d c^d \sum_{i=0}^{d} \left(\frac{m!}{(m-i)!}\right)^{3/2} k^{m-i} \varepsilon^{d-i} =$$
$$= c^d 2^m d^d \left(\frac{m!}{(m-d)!}\right)^{3/2} k^{m-d} < \infty.$$

By (c) and (d) this implies $\overline{\dim}_B(H) \le d$ if $s < d+1$ (or even if $\mathcal{H}^{d+1}(H) = 0$). This contradicts (a) unless $\overline{\dim}_B(H) = s \in \mathbb{Z}$ (or even if $\mathcal{H}^s(H) = 0$). Now with $d = s$ (4.14) implies the right end of (4.13). □

Corollary 4.37. *If H is a bounded subset of an (at least) two (algebraic) codimensional algebraic variety in $\mathbb{R}^m$, then $\mathcal{L}^m(H^{[\delta]}) = o(\delta)$.*

Proof. Knowing from [20] that the algebraic and Hausdorff dimensions coincide, the proposition actually gives $\overline{\mathcal{M}}^{m-2}(H) < \infty$ which means (by (e)) that $\mathcal{L}^m(H^{[\delta]}) = O(\delta^2)$. □

4.4.3 Lipschitz decomposability of algebraic varieties

In this subsection our aim is to establish the fact that one codimensional SSAV-s possess the finite Lipschitz decomposability property (in the sense of Definition 4.6). Having already shown the algebraic nature of $\mathcal{S}_n$ – and thus $\mathcal{S}^{(n)}$, this way we find that algebraic billiards satisfy Conjecture 4.8. The main result of the subsection is:

Theorem 4.38. *Any one codimensional algebraic variety H is Lipschitz decomposable (in the sense of Definition 4.6) with any constant $L > 0$.*

In the following, π shall denote the standard projection of $\mathbb{R}^m$ to $\mathbb{R}^{m-1}$. That is, $\pi(x, y) = x$ for any $x \in \mathbb{R}^{m-1}$, $y \in \mathbb{R}$.

Proof. We construct the decomposition of H. Fix an arbitrary $L > 0$. Let I(H) denote the ring of polynomials vanishing on H. Let H^* be the set of points in H where the gradient of every polynomial in $I(H)$ vanishes. We know from [2] that this set is at least two (algebraic) codimensional, so Corollary 4.37 ensures that H^* is good (for the purpose of Definition 4.6). For the points $x \in H \setminus H^*$, there is at least one $P \in I(H)$ for which $\mathrm{grad}P(x) \neq 0$ and the gradients of all polynomials in $I(H)$ are parallel to $\mathrm{grad}P(x)$. In the following we will assume $H = \{x | P(x) = 0\}$ for one such P, only for the sake of more transparent notation.

Fix a finite collection of unit vectors $v_1, \dots, v_N$ in $\mathbb{R}^m$, such that for any nonzero vector $v \in \mathbb{R}^d$, there is a v_i for which $\tan(\sphericalangle(v, v_i)) < L' < L$. We shall identify those components of H that are Lipschitz graphs as viewed from the direction v_i. We will omit the index i. The construction clearly depends on the vector $v = v_i$. Having fixed v it is possible to choose an orthogonal coordinate system in $\mathbb{R}^m$ such that the mth base vector points in the direction v. For $\arctan(L') < \varphi < \arctan(L)$ and $h = cos(\varphi)$, consider the following subset of the algebraic variety:

$$\begin{aligned} H^{<\varphi} &= \{x \in H \mid \sphericalangle(\mathrm{grad}P(x), v) < \varphi\} = \\ &= \left\{x \in H \mid (\frac{\partial}{\partial x_m}P(x))^2 > h^2(\mathrm{grad}P(x))^2\right\}. \end{aligned}$$

Note that $H^{<\varphi} \cap H^* = \emptyset$, because the inequality in the definition of $H^{<\varphi}$ is strict. We claim that for almost every possible φ, $\partial H^{<\varphi}$ is two codimensional. Indeed,

$$\partial H^{<\varphi} \subset H^{=\varphi} := \left\{x \in H \mid (\frac{\partial}{\partial x_m}P(x))^2 = h^2(\mathrm{grad}P(x))^2\right\}.$$

The intersection of $H^{=\varphi}$-s corresponding to different φ-s is H^*, which is two-codimensional, so its one codimensional Hausdorff-measure is zero. However, Proposition 4.36 says that the union of all $H^{=\varphi}$-s (which is part of H) has a finite one codimensional Hausdorff-measure. So apart from a countable number of φ-s, the one codimensional Hausdorff-measure of $H^{=\varphi}$ is zero. Since $H^{=\varphi}$ is algebraic, Proposition 4.36 tells us that almost every $H^{=\varphi}$ is two codimensional.

We fix $H' = H^{<\varphi}$ with one such $\arctan(L') < \varphi < \arctan(L)$.

We will cut H' into locally Lipschitz graphs. Let $k : \mathbb{R}^{m-1} \to \mathbb{N}$ be the multiplicity of $\pi(H')$. Clearly for every $x \in \pi(H')$ the restriction of P to $\pi^{-1}(x)$ is nonzero, so k is bounded by the degree of P, and the Implicit Function Theorem implies that it is lower semicontinuous. So, the set $D_1 \subset \mathbb{R}^{m-1}$ where k is maximal, is open. Here we can define the finitely many functions $f_{1,1}, \dots, f_{1,k_{max}} : D_1 \to \mathbb{R}$ taking the least, second least, ..., greatest element of $\pi^{-1}(x)$ for some $x \in D_1$. the Implicit Function Theorem implies that these functions are locally Lipschitz with constant L and that their graphs are disjoint.

Now we claim that the boundary of these graphs is two codimensional. Indeed, $H^{=\varphi}$ is two codimensional and algebraic, so $\pi(H^{=\varphi})$ is also part of a one codimensional algebraic variety in $\mathbb{R}^{m-1}$. The pre-image (by π) of this variety is one codimensional in $\mathbb{R}^m$, and the boundary of our graphs is on the intersection of this pre-image with H. This intersection is transversal (ensuring two codimensions) at points of $H' \setminus H^*$, and the rest of the boundary is in H^*.

Now erase the closure of these graphs from H'. So the argument can be repeated with k_{max} already at least one less. The procedure ends in finitely many steps, and so finitely many open locally Lipschitz graphs are constructed. Their closures cover H' by construction, and their boundary is two codimensional.

We carry out this construction for every v_i, and get a covering of the entire $H \setminus H^*$ by finitely many locally Lipschitz graphs. To get the sets H_1,...,H_N in Definition 4.6 we only need to make these graphs disjoint by subtracting the closure of one from the other. □

Putting together the statements of theorems 4.18, 4.32 and 4.38 we get

Corollary 4.39. ***(The Fundamental Theorem for algebraic semi-dispersing billiards.) Suppose Condition 4.11 is satisfied and the semi-dispersing billiard is algebraic. Let x be a sufficient phase point. Then, there is a neighbourhood U of x such that every point of U belongs to the same ergodic component.***

Chapter 5

Correlation decay in soft billiards

5.1 Introduction

Consider the motion of a point particle in a planar periodic array of circular scatterers. There are several physically relevant dynamical systems related to such a geometrical configuration. For all of these, motion outside the scatterers is uniform, i.e. the point particle proceeds along a straight (geodesic) line with constant (unit) velocity. If the particle bounces off the circles according to the laws of elastic collision (angle of incidence equals to the angle of reflection), we talk about the *hard Lorentz process*, a paradigm for strongly chaotic behaviour. Among other important properties ergodicity [38, 13] and exponential decay of correlations [43, 12] have been proven for the corresponding billiard system.

Here we consider the following natural modification. The scatterers are no longer hard disks, the point particle may enter them. The particle moves according to some rotation symmetric potential which vanishes identically outside the disks. This modification leaves the most important features of the dynamics unchanged: the systems we can describe will be hyperbolic dynamical systems with singularities. Examination of such a generalization of billiard dynamics is interesting in itself, however, we also have another motivation. As mentioned in the Introduction (and will be discussed in the Outlook, Chapter 6) we hope to be able to construct a soft billiard which is hyperbolic, but the singularities behave more regularly than those of hard billiards. This may open the way to a proof of exponential correlation decay in higher dimensional systems.

Even the issue of these softened Lorentz processes has a large literature. Results point into two different directions. On the one hand, for quite general softening of the potential, the chaotic behaviour is no longer present. Stable periodic orbits and islands appear in the phase space. This is generally the case with smooth potentials, see [31, 15, 16] and references therein.

However, in many cases, especially when the potential is not C^1, the chaotic behaviour persists.[1] The investigation of such soft billiards dates back to the pioneering works of Sinai [37] and Kubo et al. [26], [27]. There are two different approaches present in the literature to this hyperbolic case. On the one hand, under conditions on the derivatives (up to the second) of the potential the Hamiltonian flow turned out to be equivalent to

[1] In [17] there is a smooth potential example with ergodic behaviour, too. However it is unstable with respect to small perturbations like varying the full energy level.

a geodesic flow on a negatively curved manifold. This point of view is especially suitable for potentials with Coulomb type singularities, see [24] on details.

The other approach – which is actually what we follow – is to study dynamics as a hyperbolic system with singularities. [30] and, especially, [17] – which is one of our main references – are written in the spirit of this principle. Actually, in most cases it is convenient to study the discrete time dynamics, a naturally defined Poincaré section map of the Hamiltonian flow – this is the track we are going to take.

Hyperbolicity of these systems is mainly related to the properties of the so called *rotation function* that can be calculated from the potential. Being a bit technical its definition and relevant properties are discussed in Section 5.1.1. Formulation of our main theorem (Theorem 5.9) is likewise left to the next section as it is in terms of the rotation function. Nevertheless, it might be useful to point out that

- In case the rotation function (and the billiard configuration) satisfies some hyperbolicity condition (see Definition 5.5), the soft billiard system is hyperbolic and ergodic. Although a little bit otherwise stated, this fact was proved in [17]. The condition is, essentially, necessary for ergodicity (note however Remark 6.1).

- In this chapter we concentrate on decay of correlations. If – in addition to those needed for hyperbolicity – the rotation function satisfies further regularity conditions (see Definition 5.7), the rate of mixing is proved to be exponential.

In most of the chapter we think of the rotation function as being fixed with the desired properties. It is only Section 5.4 when we turn to some specific potentials. Nevertheless, three technical conditions supposed to hold throughout the chapter are:

- In order to be able to define a rotation function at all, we introduce $h(r) = r^2(1 - 2V(r))$ (cf. Section 5.4) and require $h'(r) > 0$ for all but finitely many r (this condition ensures the lack of trapping zones, cf. [17])

- The scattering occurs on rotation symmetric potentials of finite range – that is, the potential for every scatterer is concentrated on a circle and depends only on the distance from the center. (Note that this is the case in our references like [17], too.)

- The horizon is finite (i.e. the maximum time between two enterings of consecutive potential disks is uniformly bounded from above for any trajectory).

The proof of our main theorem is based on our second main reference, on [12]. In this paper, by implementing the techniques of L. S. Young from [43], N. Chernov showed that given any hyperbolic system with singularities for which one can show the validity of certain technical properties, correlations decay exponentially fast.

What we perform below is the proof of these technical properties for our 'soft' billiard system. Even though the existence of invariant cone fields is established in [17], the *uniformity* of hyperbolicity (Section 5.2.2) needs detailed investigation. An even more important new difficulty that we have to overcome is the treatment of quantities connected to the second derivative of the dynamics, especially while traveling through the potential. An analysis finer than before – in this sense – of the evolution of fronts is needed. This applies especially to the self-contained proof of curvature and distortion bounds (Section 5.2.3).

It is a key aspect of our method that arguments related to expansion and distortion can be carried out by considering motion inside and outside the potential disks separately. Actually, our choice of the outgoing phase space and the Euclidean metric (see Section 5.1.1) is related to this point of view and not to the tradition of [12]. (Using the Euclidean metric with the phase space of *incoming* particles instead of outgoing, our distortion bounds would no longer hold.) The splitting of motion into 'potential' and 'free' intervals is, however, slightly restrictive. Namely, certain soft billiard systems that seem ergodic and exponentially mixing are covered neither in this work nor in [17] (see also Chapter 6, especially Remark 6.1).

We note that it is not clear how sharp our results are. On the one hand, the conditions for ergodicity – which are part of our conditions – formulated by Donnay and Liverani are more or less sharp (see [16]). On the other hand, the conditions formulated for EDC by Chernov are sufficient, but most probably not necessary. So, although we know that Chernov's conditions (eg. the bounded curvature assumption and the distortion bounds) are *not* satisfied when our regularity conditions are not met, it is well possible that EDC still occurs. At some points of the discussion we will point out why our regularity conditions are necessary for Chernov's method to work.

As a beginning, we introduce the dynamical system, together with some necessary notions, and state the main result of the chapter, Theorem 5.9. The proof is briefly sketched in Section 5.1.6. Details of the proof are in Sections 5.2 and 5.3. Finally, specific examples where the theorem is valid are discussed in Section 5.4.

5.1.1 The dynamical system and the rotation function

Consider finitely many disjoint circles of radius R on the unit two-dimensional flat torus $\mathbb{T}^2$. (Thinking of a periodic array of circular disks on the Euclidean plane $\mathbb{R}^2$ would not be very much different.) We require that the configuration has finite horizon: there is a certain constant $\tau_{\max}$ such that any straight segment longer than $\tau_{\max}$ on $\mathbb{R}^2$ intersects at least one of the scatterers.

Remark 5.1. *As the circles are disjoint, the minimum distance between two scatterers is bigger than some positive constant* $\tau_{\min}$.

Let the Hamiltonian motion of our point particle be described by a potential which is identically zero outside and is some rotation symmetric function $V(r)$ inside the circular scatterers (here r is the distance from the center of the scatterer). For simplicity we fix the mass and the full energy of our point particle as

$$m = 1, \qquad E = \frac{1}{2}.$$

This way the free flight velocity has unit length, $|v| = 1$ (in other words $v \in \mathbf{S}^1$, the unit circle in $\mathbb{R}^2$).

We know that as a result of the assumptions we will make, the Hamiltonian flow restricted to this surface of constant full energy is ergodic with respect to Liouville measure (cf. Definition 5.5 and the remarks following it.) Equivalently one can say that the map corresponding to the naturally defined Poincaré section of the flow (see below) is ergodic. Our aim is *to study the rate of mixing for* this *map.*

We work with the Poincaré section of outgoing velocities (particles that have just left one of the scatterers).

Notation. *Denote by M the Poincaré section of outgoing particles. Sometimes we will also use the notation $M_+ = M$ to stress that this is the outgoing phase space, to avoid confusion.*

The phase points are the boundary points of the scatterers, equipped with unit velocities pointing outwards. The phase space M is a finite union of cylinders (each corresponding to one of the circular scatterers). Coordinates for the cylinders are:

Notation. *s denotes the arclength parameter along the scatterer (starting from a point arbitrarily fixed), describing position of the outgoing particle.*

φ denotes the collision angle, the angle that the outgoing velocity makes with the normal vector of the scatterer in the point s. Clearly $\varphi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$.

The position can be equivalently described by another angle parameter $\Theta \in [0, 2\pi]$, for which $s = R\Theta$ (here R is the radius of the scatterer).

Note that M defined this way is a (finite union of) Riemannian manifold(s).

Let

$$|dx|_e = \sqrt{ds^2 + d\varphi^2} \tag{5.1}$$

denote the Riemannian metric on M, which will be referred to as the Euclidean metric (e-metric).

Later on, just like in the hard billiard case, we will introduce the p-metric for soft systems as well.

The dynamics T is the first return map onto M.

Notation for the Lebesgue measure on M is m, i.e. $dm = ds\, d\varphi$. Furthermore, given a curve γ in M we denote the Lebesgue measure on γ with m_γ (this is simply the length on γ).

Denote by μ the natural invariant probability measure on M. μ is absolutely continuous w.r.t. Lebesgue, and the density is of the form

$$d\mu = \ const.\ \cos(\varphi)\, dm = \ const.\ \cos(\varphi)\, ds\, d\varphi. \tag{5.2}$$

It is this latter measure for which T is shown to be ergodic and K-mixing in [17] under the appropriate conditions, and this is the one we work with as well.

Definition 5.2. *The dynamical system (M, T, μ) is the soft billiard map.*

Remark 5.3. *In a completely similar manner we could consider the Poincaré section M_- of incoming particles. The two coordinates would be the point of income and the angle the incoming velocity makes with the (opposite) normal vector. However, in some key steps of the proof – eg. the distortion bounds of Section 5.2.3.2 – we heavily use that our phase space is the outgoing, and* not *the incoming Poincaré section.*

With slight abuse of notation we often refer to the incoming Poincaré coordinates with the same symbols s and φ. That should cause no confusion.

To describe the first return map T we decompose the motion into two parts: free flight among the scatterers and flight in the potential of the scatterers. Free flight can be treated completely analogously to the billiard case. The particle leaves one of the scatterers in the point s_0 with velocity φ_0 and reaches some other scatterer in point s (or equivalently, Θ) with unit incoming velocity that makes an angle φ with the (opposite) normal vector $n(s)$ at the point of income. After some inter-potential motion the particle leaves the circle in some point $s_1(= R\Theta_1)$ with outgoing velocity specified by φ_1. Out of symmetry reasons $\varphi_1 = \varphi$, thus the only nontrivial quantity is the angle difference $\Delta\Theta = \Theta_1 - \Theta$. Again out of symmetry reasons $\Delta\Theta$ depends only on the angle φ.

The role in the map T played by the potential is completely described by the function $\Delta\Theta(\varphi)$.

Definition 5.4. *From here on we will refer to this function $\Delta\Theta(\varphi)$ as* ***the rotation function****.*

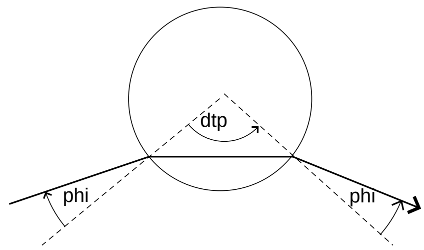


Figure 5.1: meaning of the rotation function

Being mainly interested in the differential aspects of T we introduce one more

Notation.

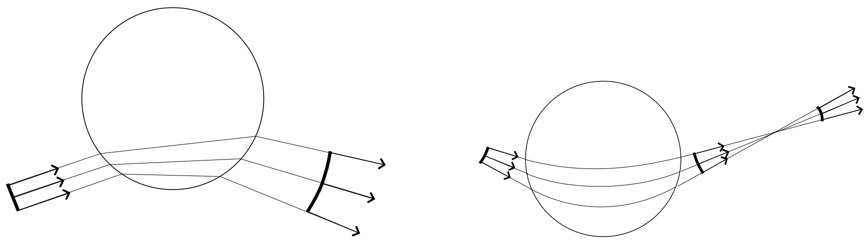
$$\kappa(\varphi) = \frac{d\Delta\Theta(\varphi)}{d\varphi}.$$

5.1.2 Mechanisms of hyperbolicity

There is a rather large class of potentials for which [17] obtained ergodicity and hyperbolicity, both repelling and attracting. Reason for their success is that V. Donnay and C. Liverani have found conditions on κ sufficient for hyperbolic behaviour that later on also turned out to be essentially necessary. Before formulating these we turn to a simple, though slightly artificial example. Fix the potential inside the scatterers as identically zero. Thus, when entering the circle, the particle proceeds along its straight trajectory without changing its direction or velocity magnitude. This system is, of course, not hyperbolic: neutral fronts remain neutral and convex ones loose more and more convexity as they evolve. It is straightforward to calculate $\Delta\Theta(\varphi) = \pi - 2\varphi$ and thus $\kappa(\varphi) = -2$ identically.

In view of the above example it is not surprising that the value $\kappa = -2$ is to be avoided. Actually, [17] obtained results in two different cases.

Dispersing case. Assume $\kappa < -2$ or $\kappa \geq 0$ for all φ. The soft billiard is ergodic and hyperbolic. Mechanism of hyperbolicity is, just like in the hard case, related to convex fronts. I.e. incoming convex fronts (those reaching the potential disk) turn into outgoing convex fronts (convex when leaving the disk). See Figure 5.2.



(a) dispersing case (b) defocusing case

Figure 5.2: mechanisms of hyperbolicity

Defocusing case. Assume there is some $\delta > 0$ such that $0 \geq \kappa > -2 + \delta$. In addition, the configuration is such that there is some lower bound $t_{\min}$ on the free flight between two disks which satisfies $t_{\min} > \frac{2R(2-\delta)}{\delta}$. The soft billiard is ergodic and hyperbolic. Nevertheless, the mechanism of hyperbolicity is different from the one in the dispersing case. Incoming convex fronts may turn into outgoing concave fronts. These concave fronts, as there is enough time until the next disk is reached, defocus and turn into convex fronts during free flight. Thus, when entering a potential region for the next time, they are convex again. See Figure 5.2. The mechanism is discussed in full detail in Section 5.2.2.

The above brief discussion sheds some light on the fact why these conditions are, essentially, necessary for chaotic behaviour. Assume there is some φ for which κ approaches -2 from above and the suitable bound on $t_{\min}$ is missing. The outgoing concave fronts do not have enough time to defocus, thus they remain concave for all positive times. In case one can construct a periodic orbit in the vicinity of these persistent concave fronts, the periodic orbit turns out to be stable. [16] shows obstructions for ergodicity roughly along these lines.

5.1.3 Regularity of the rotation function

Our aim is to prove exponential decay of correlations in certain ergodic soft billiards. Based on these results we conjecture that the rate of mixing is exponential in essentially all the (finite horizon) cases for which [17] obtained hyperbolic ergodicity. Nevertheless, in contrast to [17], we do not have many – actually, we do have only two classes of – specific potentials for which exponential decay of correlations can be explicitly shown.

The small number of specific examples is related to the necessity of understanding the 'second derivative of the dynamics'. To obtain curvature and distortion bounds – which is inevitable for the application of methods from [12] and [43] – we need to check the derivative of $\kappa(\varphi)$.

Below two important properties are defined and discussed, in terms of which our main theorem is formulated.

Definition 5.5. *The soft billiard system satisfies property H (H stands for 'hyperbolicity') in case*

1. *there is some positive constant c such that $|2 + \kappa(\varphi)| > c$ for all φ;*

2. *the configuration of scatterers is such that the distance of any two circles is bounded below by $\tau_{\min}$ where*
$$\tau_{\min} \gneqq \max_{\varphi} \left\{ -2R\kappa(\varphi)\frac{\cos\varphi}{2+\kappa(\varphi)} \right\}.$$

Remark 5.6.

- *Although a bit otherwise formulated, it was essentially proven in [17] that soft billiard systems with property H are hyperbolic and ergodic. The mechanism of hyperbolicity is explained in Section 5.2.2. (Actually, the hyperbolicity shown in [17] is not enough for our purposes, so we prove uniform hyperbolicity.)*
- *Note that in case $\kappa > 0$ or $\kappa < -2$ for all φ, the lower bound for $\tau_{\min}$ turns out to be negative. Thus the second assumption is only restrictive in the opposite case, and the closer κ may get to -2 from above, the more restrictive it is.*
- *In case there is some φ for which $0 > \kappa > -2$, a positive lower bound on the free path is to be assumed. As already mentioned, we will also require that the horizon is finite – that is, the length of free flight is bounded. Thus a planar periodic configuration of circles is needed which has finite horizon and (a possibly great) given $\tau_{\min}$ simultaneously. At first sight it seems questionable whether such configurations exist at all, nevertheless, as proven in [9], this happens with positive probability in a random construction.*

Definition 5.7. *The rotation function is termed regular in case the following properties hold.*

1. $\Delta\Theta(\varphi)$ *is piecewise* uniformly *Hölder continuous. I.e. there are constants $C < \infty$ and $\alpha > 0$, and furthermore, $[-\frac{\pi}{2}, \frac{\pi}{2}]$ can be partitioned into finitely many intervals, such that for any φ_1 and φ_2 (from the interiour of one of the intervals):*
$$|\Delta\Theta(\varphi_1) - \Delta\Theta(\varphi_2)| \le C|\varphi_1 - \varphi_2|^{\alpha}.$$
2. $\Delta\Theta(\varphi)$ *is a piecewise C^2 function of φ on the closed interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$, in the above sense. (Note, however, that κ, in contrast to $\Delta\Theta$, can happen to have no finite one-sided limits at discontinuity points.)*
3. *There is some finite constant C such that*
$$|\kappa'(\varphi)| \le C|(2+\kappa(\varphi))^3|$$
where $\kappa'(\varphi)$ is the derivative of κ with respect to φ.
4. *For the final property consider any discontinuity point φ_0 where $\kappa(\varphi)$ (in contrast to $\Delta\Theta(\varphi)$) has no finite limit from the left. Of course, in case there is no finite limit from the right, the analogous property is similarly assumed.*

   Restricted to some interval $[\varphi_0 - \epsilon, \varphi_0)$; $\omega(\varphi) = \frac{2+\kappa(\varphi)}{\cos\varphi}$ is a monotonic function of φ.

Remark 5.8. *Note that in case κ is C^1 (or piecewise C^1 with boundedness of itself and of κ') regularity is automatic. In case the asymptotics of κ near some discontinuity is some power law $(\varphi_0 - \varphi)^{-\xi}$ (with $\xi > 0$), regularity means $\frac{1}{2} \le \xi < 1$.*

5.1.4 Singularities

Just like in billiards the dynamics T is not smooth at certain one codimensional submanifolds (curves) of M. Consider the set of tangential reflections:

$$\mathcal{S}_0 = \left\{ (s,\varphi) \in M \mid \varphi = \pm\frac{\pi}{2} \right\}.$$

Actually $\mathcal{S}_0 = \partial M$ (the boundary of the phase space). We know that T is not continuous at $\mathcal{S}_1 = T^{-1}\mathcal{S}_0$, i.e. at the preimages of tangential reflections. However, additional singularities appear at

$$\mathcal{Z}_0 = \{ (s,\varphi) \in M \mid \varphi = \varphi_0 \}$$

in case φ_0 is some discontinuity point for $\Delta\Theta(\varphi)$, $\kappa(\varphi)$ or $\kappa'(\varphi)$. In such a case we will consider the phase space as if it were cut into two regions, more precisely $\mathcal{Z}_0$ is treated as part of the boundary. As κ is not differentiable at $\mathcal{Z}_0$, T is not C^1 at the preimage of this set, at $\mathcal{Z}_1 = T^{-1}\mathcal{Z}_0$.

Similarly to the definition of $\mathcal{S}^{(n)}$ in (2.2), we introduce

$$\mathcal{Z}^{(n)} = \mathcal{Z}_1 \cup T^{-1}\mathcal{Z}_1 \cup \cdots \cup T^{-n+1}\mathcal{Z}_1.$$

The n-th iterate of the dynamics is not smooth precisely at $\mathcal{Z}^{(n)} \cup \mathcal{S}^{(n)}$.

The geometrical structure of $\mathcal{Z}^{(n)}$ is much similar to that of $\mathcal{S}^{(n)}$. Indeed, one can think of $\mathcal{Z}_1$ as the set of those trajectories that would touch tangentially a smaller disk (one of radius $R\sin(|\varphi_0|)$) at the next collision. The following properties of the singularity set are of crucial importance:

- $\mathcal{Z}^{(n)} \cup \mathcal{S}^{(n)}$ is a finite union of C^2 curves.
- *Continuation property.* Each endpoint, x_0, of every unextendable smooth curve $\gamma \subset \mathcal{Z}^{(n)} \cup \mathcal{S}^{(n)}$, lies either on the extended boundary $\mathcal{Z}_0 \cup \mathcal{S}_0$ or on another smooth curve $\gamma' \subset \mathcal{Z}^{(n)} \cup \mathcal{S}^{(n)}$ that itself does not terminate at x_0.
- *Complexity property.* Let us denote by K_n the complexity of $\mathcal{Z}^{(n)} \cup \mathcal{S}^{(n)}$, i.e. the maximal number of smooth curves in $\mathcal{Z}^{(n)} \cup \mathcal{S}^{(n)}$ that intersect or terminate at any point of $\mathcal{Z}^{(n)} \cup \mathcal{S}^{(n)}$. K_n grows sub-exponentially with n.

For the proof of these properties in the billiard setting see the literature, especially [14], our case is analogous.

One more similarity to 'hard' billiards is that for technical reasons later on we will introduce countably many secondary singularities parallel to the lines of $\mathcal{S}^{(n)}$. Such secondary singularities are to be introduced parallel to $\mathcal{Z}^{(n)}$ as well, in case $|\kappa|$ is unbounded as $\varphi \to \varphi_0$, at least from one side. We will turn back to this question in Section 5.2.2.6.

5.1.5 Formulation of the theorem

Now we are ready to formulate our main theorem, which uses definitions 1.3 and 1.4.

Theorem 5.9. ***Suppose that the soft billiard system*** (M,T,μ) ***satisfies property H and the rotation function is regular. Suppose furthermore that there are no corner points and the horizon is finite (***$0 < \tau_{min}$, $\tau_{max} < \infty$***).***

Then, the dynamics enjoys, in addition to ergodicity and hyperbolicity, exponential decay of correlations and the central limit theorem for Hölder-continuous functions.

5.1.6 Sketch of the proof

The proof of Theorem 5.9 is rather long, thus for a better understanding, we first give a brief sketch of the main argument. As already discussed in the Introduction, the main new difficulty is the treatment of quantities connected to the second derivative of the dynamics. This is needed for the proof of curvature bounds, and, especially, distortion bounds (described below). It is also worth mentioning that the choice of the right phase space and metric is not trivial. For example, using the Euclidean metric (as we do) with the phase space of *incoming* particles instead of outgoing, our distortion bounds would no longer hold. The complete discussion of the proof is given in Sections 5.2 and 5.3.

Actually, our theorem is an application of [12]. In this paper N. Chernov showed that given a hyperbolic system with singularities for which certain properties can be shown, correlations decay exponentially. Our proof establishes these properties for the investigated soft billiard systems. More precisely, our theorems are the consequences of Theorem 2.1 from [12] (see the Appendix, Theorem A.1 for a precise formulation) and the properties with a **bold** typeface below. Arguments are to be presented at three different levels.

1. ***Uniform hyperbolicity*** *(Proposition 5.17) and geometric properties.* We define stable/unstable manifolds as those curves in M that correspond to fronts that remain concave/convex for all future/past iterates of T. Unlike in [17], we need to show that contraction/expansion (in the natural Riemannian metric of M) is indeed uniform. To achieve this, one needs to introduce an auxiliary metric quantity (the p-metric of billiard literature, cf. Section 2.6 for a general definition or (5.26) for the specific form in two dimensions) and furthermore, needs to discuss the cases of $\kappa \geq 0$, $0 > \kappa \geq -1$, $-1 > \kappa > -2 + \delta$ and $-2 - \delta > \kappa$ separately. In addition, uniform *transversality* of stable and unstable manifolds is to be shown. Besides hyperbolicity, it is to be shown that singularities and stable/unstable manifolds, even though not necessarily transversal to each other, can have tangential intersections of at most some polynomial rate. This property, often termed as *alignment*, is the point where Hölder continuity of the rotation function (the first property from Definition 5.7) is applied.

2. *Technicalities on stable/unstable manifolds.* To apply the result of [12] (or more generally, the ideas of L.-S. Young from [43]) one needs to show that stable/unstable manifolds enjoy certain regularity properties. The proof of **uniform curvature bounds** (Proposition 5.23) relies on the fact that curvature of a convex front (when viewed as a submanifold of the flow phase space) cannot blow up during time evolution as distances on the front always grow faster than the inhomogeneities in its shape. This behaviour is why we need the third regularity property in Definition 5.7. It might be worth mentioning that the proof of curvature bounds seemed much more difficult for the case of $-1 < \kappa < 0$ at first sight as in this case one needs to handle fronts that defocus even within the potential, nevertheless, finally, based on Definition 5.7, we could find an argument that applies to all allowed κ values. As to **distortion bounds** (Proposition 5.24) and the **absolute continuity** (Proposition 5.28) of holonomy maps, following the idea already applied in 'hard' billiards (cf. [12] and references), the phase space is partitioned into 'homogeneity layers' by

introducing further, 'artificial' singularities. This is to be done with special care, for – in addition to the phenomena near tangential singularities common to all billiards – further unbounded derivatives may appear at discontinuities of κ. This is the point where the fourth regularity property from Definition 5.7 is exploited.

3. ***Growth properties of unstable manifolds*** (Proposition 5.29). The main idea of the papers [43] and [12] is that expansion of unstable manifolds is uniformly stronger than their fractioning caused by the presence of singularities. This is quantified by three rather technical growth formulas in [12]. However, establishing the validity of these formulas is the part of our proof which is closest to the analogous discussion from [12]. We note that this is the point where the *alignment* property of singularities and unstable manifolds is used.

The proof consists of checking all these properties described.

Ingredients for the proof of Theorem 5.9 are in Sections 5.2 and 5.3. Actually, following tradition (eg. [12]) we modify the dynamical system in several steps (Conventions 5.16 and 5.19). We will use a phase space $\bar{M}$, which is the original M cut into (countably many) connected components by singularities and so called 'secondary singularities'. We will also use a higher iterate of the dynamics $T_1 = T^{m_0}$ with some m_0 to be found later.

It is the modified dynamical system $(\bar{M}, T_1, \mu)$ for which the conditions for EDC and CLT given in [12] are checked. Precisely, EDC and CLT for $(\bar{M}, T_1, \mu)$ are the consequence of propositions 5.17, 5.23, 5.24, 5.28 and 5.29 and Theorem 2.1 from [12].

Exponential decay of correlations and the central limit theorem for (M, T, μ) follow easily from EDC and CLT for $(\bar{M}, T_1, \mu)$.

For the reader's convenience, we give a formulation of Theorem 2.1 from [12] in the Appendix.

In the two corollaries in Section 5.4 we present two important classes of potentials for which calculation of the rotation function is relatively straightforward. In these two cases the assumptions of Theorem 5.9 are satisfied.[2]

Before turning to the details of the above proof, we fix some conventions.

Convention 5.10. *Constants that depend only on the map T itself (like $\tau_{\min}, \tau_{\max}, \dots$) will be called* global constants*.*

Positive and finite global constants, whose value is otherwise not important, will be often denoted by just c or C (typically c for lower bounds and C for upper). That is, in two different lines of the same section, C can mean two different numbers.

Two quantities f and g defined on (the tangent bundle of) M (or on some subset like the unstable cone field, see Section 5.2.1) will be called equivalent *($f \sim g$) if there are some global positive constants c and C such that $cf \le g \le Cf$.*

[2]It might be worth mentioning that these potentials, as functions on $\mathbb{R}^2$, are not C^1, thus the equations of motion are to be integrated with care. One needs to integrate inside and outside the disks separately and apply plausible boundary conditions: the magnitude of the velocity at R_- can be obtained from the kinetic energy, and the tangential velocity component is continuous at R.

5.2 Fronts, u-manifolds and unstable manifolds

5.2.1 Fronts and their geometric properties

Analysis of the time evolution of fronts (Section 2.2) is the key to almost all the geometric properties of the system that we need. For this reason, we first discuss time evolution of an arbitrary front. Later subsections will deal with special cases. In two dimensions, a front is just a curve in configuration space.

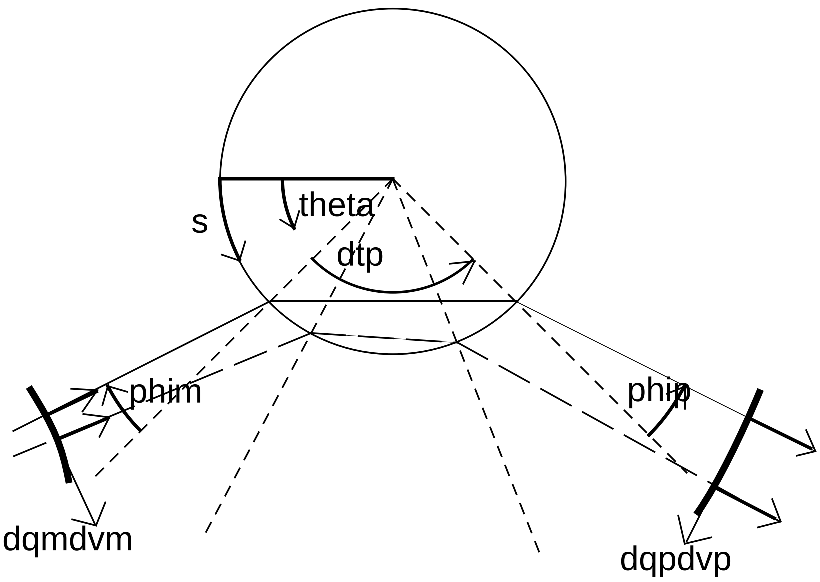


Figure 5.3: conventions for notation and signs for fronts

Consider a front with a reference point just before reaching a scatterer, and another ‘perturbed’ point nearby. With the notations of Section 5.1.1 (see also Figure 5.3), the perturbation bringing the reference trajectory into the perturbed one is (dq_-, dv_-) just before collision, $(ds_-, d\varphi_-)$ in the incoming Poincaré section, $(ds_+, d\varphi_+)$ in the outgoing Poincaré section, (dq_+, dv_+) just after collision, and (dq'_-, dv'_-) just before the next collision. In general:

Convention 5.11. *with any x, x' denotes the equivalent of the quantity x after one collision – that is, one collision ahead in time.*

The evolution of the perturbations is:

$$\begin{aligned}
ds_- &= \frac{dq_-}{\cos\varphi_-} \\
d\Theta_- &= \frac{ds_-}{R} \\
d\varphi_- &= dv_- + d\Theta_- \\
d\Theta_+ &= d\Theta_- + \kappa d\varphi_- \\
d\varphi &:= d\varphi_+ = d\varphi_- \\
ds_+ &= Rd\Theta_+ \\
dq_+ &= -\cos\varphi ds_+ \\
dv_+ &= -d\Theta_+ - d\varphi_+
\end{aligned} \tag{5.3}$$

while crossing the potential. For the evolution equations of free flight, we introduce the

Notation. $\tau = \tau(x)$ *will denote the length of free flight of the particle before reaching the next scatterer.*

So, during free flight we have

$$\begin{aligned}
dq'_- &= dq_+ + \tau dv_+ \\
dv'_- &= dv_+.
\end{aligned} \tag{5.4}$$

Note that the angles of incidence and reflection are measured in different directions – in order to keep them equal, as they traditionally are, – but dq_- and dq_+ (just like dv_- and dv_+) are measured in the same direction, unlike usually in billiards.

Based on these, we can find out about the evolution of the derivative $B = \frac{dv}{dq}$.

In this 2D case, the SFF B of a front is just a number – the derivative of the unit normal vector (velocity) $v(q)$ of a front: $dv = Bdq$ for tangent vectors (dq, dv) *of the front.*

Notation. $m = \frac{d\varphi}{ds}$ *will denote the slope of the (trace of the) front in the Poincaré section.*

Remember, B is the curvature of the submanifold E, yet we prefer to call it second fundamental form (SFF), in order to avoid confusion with other curvatures. The term 'form' refers to higher dimensional cases when B is a symmetric operator.

(5.3) gives

$$\begin{aligned}
m_- &= \cos\varphi B_- + \frac{1}{R} \\
\frac{1}{m_+} &= \frac{1}{m_-} + R\kappa \\
\cos\varphi B_+ &= m_+ + \frac{1}{R}
\end{aligned} \tag{5.5}$$

while crossing the potential, which can be summarized in

$$B_+ = \frac{2 + \kappa(\varphi) + (1 + \kappa(\varphi))R\cos\varphi B_-}{R\cos\varphi(1 + \kappa(\varphi) + \kappa(\varphi)R\cos\varphi B_-)}, \tag{5.6}$$

and (5.4) gives

$$\frac{1}{B'_-} = \frac{1}{B_+} + \tau \tag{5.7}$$

during free flight.

Notation.

$$\lambda_1 := \frac{dq_+}{dq_-}, \tag{5.8}$$

$$\lambda_2 := \frac{dq'_-}{dq_+}, \tag{5.9}$$

$$\lambda := \lambda_1 \lambda_2.$$

These are exactly the expansion factors along the front, for the respective 'pieces' of the dynamics. (They are also expansion factors in the Poincaré section, but in the p-metric – as we will see later.) We have

$$\lambda_1 = 1 + \kappa + \kappa R \cos\varphi B_- = 1 + \kappa R m_- = \frac{m_-}{m_+}, \tag{5.10}$$

$$\lambda_2 = 1 + \tau B_+ = \frac{B_+}{B'_-}. \tag{5.11}$$

To study decay of correlations, we need one more derivative.

Notation. $D = \frac{dB}{dq}$.

This is exactly the *curvature* of the *front* as of a *subset of the flow phase space* (and not as of a subset of the configuration space – unlike B, cf. (2.1)).

To study the evolution of D we need to consider two small pieces of the front, one around the reference point, and one around the perturbed one. Let the change in the SFF be

$$dB_{--} = D_- dq.$$

before scattering, and

$$dB_{++} = D_+ dq.$$

after scattering. dB_{--} is *not* the difference of SFF-s at the points of incidence, because the perturbed point has to travel another $d\tau_- = \tan\varphi_- dq_-$ to reach the scatterer ($d\tau$ can be negative), which changes its SFF according to the rules (5.7) of free flight. Taking that into account, we have

$$dB_- = dB_{--} - B_-^2 d\tau_- = dB_{--} - B_-^2 \tan\varphi_- dq_-. \tag{5.12}$$

Similarly, for the fronts leaving the potential,

$$dB_{++} = dB_+ - B_+^2 d\tau_+ = dB_+ - B_+^2 \tan\varphi_+ dq_+. \tag{5.13}$$

(Note our convention on the signs of dq_-, dq_+, φ_- and φ_+.)

To follow the evolution of curvature we introduce

Notation. $D_1 = \frac{dB_-}{dq_-} (\neq D_-)$, $K_- = \frac{dm_-}{ds_-}$, $K_+ = \frac{dm_+}{ds_+}$, $D_2 = \frac{dB_+}{dq_+}$ *and* $\eta(\varphi) = \frac{d\kappa(\varphi)}{d\varphi}$.

With these we get from (5.3), (5.5), (5.8), (5.10), (5.12) and (5.13)

$$\begin{aligned} D_1 &= D_- - \tan\varphi B_-^2 \\ K_- &= \cos^2\varphi D_1 - \sin\varphi B_- m_- \\ K_+ &= \frac{1}{\lambda_1^3}K_1 - R\left(\frac{m_-}{\lambda_1}\right)^3\eta \\ \cos^2\varphi D_2 &= -K_+ - \sin\varphi B_+ m_+ \\ D_+ &= D_2 - \tan\varphi B_+^2 \end{aligned} \tag{5.14}$$

while crossing the potential, and, from (5.4), (5.7) and (5.9)

$$D'_- = \frac{1}{\lambda_2^3}D_+ \tag{5.15}$$

during free flight.

5.2.2 Invariance of convex fronts, u-fronts and u-manifolds

In [17] it is shown – although not explicitly stated in this integrated form – that if Property H (defined in Definition 5.5) is satisfied, then convex fronts with suitably small SFF-s (the upper bound may be ∞) either remain convex, or focus before reaching the next scatterer, and become convex again, with suitably small SFF. This property is called the 'invariance of convex fronts'. In the present work we also require (see Theorem 5.9) that τ be bounded from below by some $\tau_{\min} > 0$ even in the case when [17] did not (the 'no corner points' assumption), and an upper bound $\tau_{\max}$ (the 'finite horizon' assumption). In order to establish estimates that we will need later, we must repeat some steps of the argument in [17]. We omit details of the calculations, these can be done by the reader or can be found in the above paper.

Notation.

$$\begin{aligned} \tau_1 &= \max\left\{0, \max_{\varphi}\left\{-R\kappa(\varphi)\frac{\cos\varphi}{2+\kappa(\varphi)}\right\}\right\}, \\ B^* &= \frac{1}{\tau_1} \quad (\infty \text{ if } \tau_1 = 0). \end{aligned} \tag{5.16}$$

Let us consider the three different possible cases for the value of κ.

- If $\kappa < -2-\delta$ and $B_- > 0$, then from (5.5) we get $B_+ > \frac{\delta}{R(1+\delta)\cos\varphi} \geq \frac{\delta}{R(1+\delta)}$.
- If $\kappa \geq 0$ and $B_- > 0$, then from (5.5) we get $B_+ > \frac{2+\kappa}{R(1+\kappa)\cos\varphi} > \frac{1}{R}$.
- If $-2 < \kappa < 0$ and $0 < B_- < B^*$, then (5.16) gives $B_- < -\frac{2+\kappa}{R\kappa\cos\varphi}$. Putting that in (5.5), we get that either $m_+ > 0$ and thus $B_+ > \frac{1}{R\cos\varphi} \geq \frac{1}{R}$ or $m_+ < 0$ and thus $B_+ < \frac{2+\kappa}{R\kappa\cos\varphi} \leq -B^*$.

In any of these cases, (5.7) implies that $B_* < B'_- < B^{**}$ with some global constants $B_* > 0$ and $B^{**} < B^*$, whenever $\tau_{min} > 2\tau_1$, which is exactly Property H. All in all,

$$B_* < B_- < B^{**} \text{ implies } B_* < B'_- < B^{**}. \tag{5.17}$$

This motivates our

Definition 5.12. *A u-front is a front with $B_* < B_- < B^{**}$. A u-manifold is the trace of a u-front on the Poincaré phase space.*

and

Definition 5.13. *An s-front is a front with $B_* < -B_+ < B^{**}$. An s-manifold is the trace of an s-front on the Poincaré phase space.*

As we have seen, u-manifolds remain u-manifolds under time evolution. s-fronts are exactly the u-fronts of the inverse dynamics.

The aim of this subsection is to show important properties of u-fronts and u-manifolds, which are stronger than those shown for an arbitrary front in the previous subsection. In Section 5.2.3 we further restrict to the case of unstable manifolds, which are special kinds of u-manifolds.

5.2.2.1 Expansion estimates along u-fronts

First we work out estimates for the expansion along a front from one moment of incidence to the next. We will use these estimates later to estimate expansion of our dynamics T in our outgoing Poincaré phase space M.

Consider a u-front with the earlier notations. We start with an easy observation we will often use: from (5.5) and (5.17) we get $\frac{1}{R} < m_- < \frac{1}{R} + B^{**}$, which implies

$$m_- \sim 1. \tag{5.18}$$

To get the order of magnitude for the expansion factor λ, put the formulas in (5.6) and (5.10) together, and get that

$$\frac{R\cos\varphi B_+\lambda_1}{2+\kappa(\varphi)} = 1 + (1+\kappa(\varphi))RB_-\frac{\cos\varphi}{2+\kappa(\varphi)}.$$

The right hand side is trivially bounded from above since B_- is bounded, and so is $\frac{1+\kappa(\varphi)}{2+\kappa(\varphi)} = 1 - \frac{1}{2+\kappa(\varphi)}$. On the other hand,

- It is greater than 1 if $\frac{1+\kappa(\varphi)}{2+\kappa(\varphi)} > 0$.
- If $\frac{1+\kappa(\varphi)}{2+\kappa(\varphi)} \le 0$ (that is, $-2 < \kappa(\varphi) \le -1$), then

$$1 + (1+\kappa(\varphi))RB_-\frac{\cos\varphi}{2+\kappa(\varphi)} \ge 1 + (1+\kappa(\varphi))R\frac{\cos\varphi}{2+\kappa(\varphi)}B^* \ge$$

$$\ge 1 + (1+\kappa(\varphi))R\frac{\cos\varphi}{2+\kappa(\varphi)}\,\frac{2+\kappa(\varphi)}{-R\kappa(\varphi)\cos\varphi} = \frac{-1}{\kappa(\varphi)} \ge \frac{1}{2}.$$

All in all, using $\lambda_2 = \frac{B_+}{B'_-} \sim B_+$ (see (5.11) and (5.17)) we have

$$\lambda \sim B_+\lambda_1 \sim \frac{2+\kappa(\varphi)}{\cos\varphi}, \tag{5.19}$$

which is one of our key estimates. Notice that the right hand side cannot be too small due to Property H (Definition 5.5).

We can also get the order of magnitude for λ_1 and λ_2 separately: (5.10) and (5.18) gives

$$|\lambda_1|\sqrt{1+m_+^2} = \sqrt{\lambda_1^2+m_-^2} \sim |2+\kappa(\varphi)|. \tag{5.20}$$

(The last equivalence is true because both sides are bounded away from zero, and can only be big when they grow linearly with κ.) Notice that λ_1 can be very small (even zero), and can even change signs while $2+\kappa(\varphi)$ remains positive. Of course, m_+ has to be infinity (and change signs) simultaneously.

Putting (5.19) and (5.20) together, we get

$$\frac{|\lambda_2|}{\sqrt{1+m_+^2}} \sim \frac{1}{\cos\varphi}. \tag{5.21}$$

This last line can be rewritten as

$$1 \sim \frac{|\lambda_2|\cos\varphi}{\sqrt{1+m_+^2}} \sim \frac{|B_+|\cos\varphi}{\sqrt{1+m_+^2}} = \frac{|m_+ + \frac{1}{R}|}{\sqrt{1+m_+^2}},$$

which implies that there is a global constant c such that

$$\left|m_+ + \frac{1}{R}\right| > c. \tag{5.22}$$

5.2.2.2 Expansivity

To obtain hyperbolicity, we must see that u-manifolds are expanded by the dynamics. In the first round we prove a lemma about the expansion on u-fronts from collision to collision.

Lemma 5.14. *There exists a global constant $\Lambda > 1$, such that for every u-front, $|\lambda| \geq \Lambda$.*

Proof. Besides $\tau > 0$ and $B_- > 0$ we will use that $\tau \geq 2\tau_1 + d$ where $d := \tau_{min} - 2\tau_1 > 0$, and $\tau_1 \geq -R\kappa(\varphi)\frac{\cos\varphi}{2+\kappa(\varphi)}$ for every φ (see Definition 5.5 and (5.16)). Altogether:

$$\tau \geq d - 2R\kappa(\varphi)\frac{\cos\varphi}{2+\kappa(\varphi)}. \tag{5.23}$$

We will also use from (5.16) and Definition 5.12 that

$$0 < B_- \leq \frac{2+\kappa(\varphi)}{-R\kappa(\varphi)\cos\varphi} \tag{5.24}$$

whenever the right hand side is positive, which is the $-2 < \kappa(\varphi) < 0$ case. Now we start by putting together (5.10), (5.11) and (5.6) to get

$$\lambda = (1+\kappa(\varphi)Rm_-)(1+\tau B_+) = 1+\kappa(\varphi)+\kappa(\varphi)R\cos\varphi B_- + \tau\left(\frac{2+\kappa(\varphi)}{R\cos\varphi} + (1+\kappa(\varphi))B_-\right)$$

We estimate this taking care of the signs of the particular terms.

- If $\kappa(\varphi) \leq -2-\delta$, then
$$\lambda \leq 1+\kappa(\varphi) \leq -1-\delta.$$

- If $-2+\delta \le \kappa(\varphi) \le -1$ then both coefficients of B_- are negative, so we can use (5.24) to estimate the right hand side from below. In the next step we find the coefficient of τ positive, so we can use (5.23). What we get is

$$\begin{aligned}
\lambda &\ge 1+\kappa(\varphi)+\kappa(\varphi)R\cos\varphi\frac{2+\kappa(\varphi)}{-R\kappa(\varphi)\cos\varphi}+\\
&\quad +\tau\left(\frac{2+\kappa(\varphi)}{R\cos\varphi}+(1+\kappa(\varphi))\frac{2+\kappa(\varphi)}{-R\kappa(\varphi)\cos\varphi}\right)=\\
&= -1+\tau\,\frac{2+\kappa(\varphi)}{-\kappa(\varphi)R\cos\varphi}\\
&\ge -1+\left(d-2R\kappa(\varphi)\frac{\cos\varphi}{2+\kappa(\varphi)}\right)\frac{2+\kappa(\varphi)}{-\kappa(\varphi)R\cos\varphi}\\
&\ge 1+\frac{d\delta}{2R}.
\end{aligned}$$

- If $-1 \le \kappa(\varphi) \le 0$, then the coefficient of τ is positive, so we first use (5.23) to estimate the right hand side from below. In the next step we find one coefficient of B_- positive, so we just use $B_- > 0$, and one coefficient of B_- negative, so we can use (5.24). What we get is

$$\begin{aligned}
\lambda &\ge 1+\kappa(\varphi)+\kappa(\varphi)R\cos\varphi B_- +\\
&\quad +\left(d-2R\kappa(\varphi)\frac{\cos\varphi}{2+\kappa(\varphi)}\right)\left(\frac{2+\kappa(\varphi)}{R\cos\varphi}+(1+\kappa(\varphi))B_-\right)=\\
&= 1+d\frac{2+\kappa(\varphi)}{R\cos\varphi}+d(1+\kappa(\varphi))B_- -\kappa(\varphi)-\frac{\kappa^2(\varphi)R\cos\varphi}{2+\kappa(\varphi)}B_-\\
&\ge 1+d\frac{2+\kappa(\varphi)}{R\cos\varphi}-\kappa(\varphi)-\frac{\kappa^2(\varphi)R\cos\varphi}{2+\kappa(\varphi)}\,\frac{2+\kappa(\varphi)}{-R\kappa(\varphi)\cos\varphi}\\
&\ge 1+\frac{d}{R}.
\end{aligned}$$

- If $0 < \kappa(\varphi)$, then

$$\lambda \ge 1+\frac{2d}{R}.$$

□

5.2.2.3 Transversality

Lemma 5.15. *We will see that u- and s-manifolds are uniformly transversal. I.e. there is some global constant $\alpha_0 > 0$ such that given any two tangent vectors (in the outgoing Poincaré phase space) dx_s and dx_u of an s- and a u-manifold, respectively, we have*

$$\sphericalangle(dx_u, dx_s) > \alpha_0.$$

Proof. To see this, use Definition 5.13 and (5.5) to get $m_+^s \sim -1$ for the slope of any s-manifold. This way, it is enough to see that the slopes of u- and s-manifolds are bounded away, that is, $|m^u_+ - m^s_+| > c$. To get this, use – again – Definition 5.13, Definition 5.12, the estimates before them and (5.5) to get

$$-\frac{1}{R} > m^s_+ > -\frac{1}{R} - \cos\varphi B^{**}$$

$$m^u_+ > 0 \text{ or } m^u_+ < -\frac{1}{R} - \cos\varphi B^* \tag{5.25}$$

so either $m^u_+ - m^s_+ > \frac{1}{R}$ or $m^s_+ - m^u_+ > \cos\varphi(B^* - B^{**})$. This implies the statement when $\cos\varphi$ is not too small. However, when $\cos\varphi$ is small, we have to use the estimate (5.22) and (5.25) to see also that

$$m^u_+ > 0 \text{ or } m^u_+ < -\frac{1}{R} - c$$

which completes the proof. □

5.2.2.4 Hyperbolicity

In what follows we will consider time evolution of vectors tangent to u-manifolds. Notation both in the incoming and the outgoing phase space will be of the type $dx = (ds, d\varphi)$. In addition to the e-metric (5.1) we will use the p-metric introduced in Section 2.6, which now has the form

$$|dx|_p = |ds|\cos(\varphi). \tag{5.26}$$

The p-metric measures distances along the corresponding u-front. It is degenerate on the whole tangent bundle. However, when restricted to a u-manifold *in the incoming phase space*, by (5.18) we have:

$$|dx|_p \sim |dx|_e \cos(\varphi).$$

According to Lemma 5.14, u-vectors are expanded uniformly *(from collision to collision, that is, in the incoming phase space)* in the p-metric:

$$|DT|_p = \lambda \geq \Lambda > 1$$

To obtain expansion in the e-metric and the outgoing phase space, we look at the n-th iterate of the outgoing phase space dynamics the following way:

- switch to p-metric
- reach the next scatterer
- do $n-1$ steps in the incoming phase space
- cross the potential
- switch back to Euclidean metric.

This way we get

$$|DT^n\, dx|_e = \frac{\sqrt{1 + m^2_{+(n)}}}{\cos\varphi_{(n)}} \lambda_{1(n)}\lambda_{(n-1)}\lambda_{(n-2)} \dots \lambda_{(1)}\lambda_2 \frac{\cos\varphi}{\sqrt{1+m^2_+}} |dx|_e$$

where symbols with ()-ed subscripts mean values at the appropriate iterate of the phase point. Using (5.20), (5.21) and Lemma 5.14 we get

$$|DT^n\, dx|_e \sim \lambda_{(n-1)}\lambda_{(n-2)}\dots\lambda_{(1)}\frac{2+\kappa(\varphi_{(n)})}{\cos\varphi_{(n)}}|dx|_e. \tag{5.27}$$

This way we have

$$|DT^n\, dx|_e > c_1\Lambda^n|dx|_e \tag{5.28}$$

with some global constant c_1. Again, this is for u-vectors in the outgoing phase space.

The transversality of s- and u- vectors, stated in Proposition 5.17 implies that the product of (length) expansion factors for s- and u- vectors is equivalent to the n-step (Lebesgue) volume expansion factor. Using (5.2), and the T-invariance of μ, we get that if dx is a u-vector and dy is an s-vector, then

$$\frac{|DT^n\, dx|_e}{|dx|_e}\,\frac{|DT^n\, dy|_e}{|dy|_e} \sim \frac{\cos\varphi}{\cos\varphi_{(n)}}.$$

Combining this with (5.27) we get

$$|DT^n\, dy|_e \sim \frac{\cos\varphi}{2+\kappa(\varphi_{(n)})}\frac{1}{\lambda_{(n-1)}\lambda_{(n-2)}\dots\lambda_{(1)}}|dy|_e,$$

which implies

$$|DT^n\, dy|_e < \frac{C_1}{\Lambda^n}|dx|_e \tag{5.29}$$

with some global constant C_1. Again, this is for s-vectors in the outgoing phase space.

Convention 5.16. *We choose a positive integer m_0 the following way. First take m_1 such that $c_1\Lambda^{m_1} > 1$ and $\frac{C_1}{\Lambda^{m_1}} < 1$. This way any enough high power of the dynamics, T^m with $m > m_1$ is uniformly expanding along u-manifolds and uniformly contracting along s-manifolds with $\Lambda_1 = \Lambda^{m-m_1}$. Now recall the notion and the basic properties of complexity K_n from Section 5.1.4. As K_n grows subexponentially we may choose m_2 for which we have $K_m < \Lambda^{m-m_1}$ whenever $m > m_2$. We fix $m_0 = \min(m_1, m_2) + 1$.*

The advantage of this choice is that the iterate $T_1 = T^{m_0}$ is uniformly hyperbolic (see the proposition to come) with constant Λ_1 for which $\Lambda_1 > K_{m_0} + 1$. This later fact we only use in Section 5.3.

Let us summarize what we have seen so far from the hyperbolic properties in the following

Proposition 5.17. *There exist two families of cones $C_s(x)$ and $C_u(x)$ – called stable and unstable cones – in the tangent space of M such that*

$$DT(C_u(x)) \subset C_u(Tx) \text{ and } C_s(Tx) \subset DT(C_s(x)).$$

The stable/unstable cone is uniformly contracting/expanding:

$$\begin{aligned} |DT_1^{-1}(dx)| &\geq \Lambda_1|dx| \qquad \forall dx \in C_s(x), \\ |DT_1(dx)| &\geq \Lambda_1|dx| \qquad \forall dx \in C_u(x). \end{aligned}$$

Furthermore, the two cone fields are uniformly transversal in the sense above.

Vectors of the stable/unstable cone are often called s- and u-vectors.

Proof. The two cones are formed by the tangent vectors of s- and u-manifolds, respectively. Invariance is the implication (5.17), recalling Definition 5.12 and 5.13. Expansion and contraction are (5.28), (5.29) and Convention 5.16. Transversality is Lemma 5.15. □

We note that so far we have only used that our billiard satisfies property H, which is a property already formulated in [17], and which is known from [16] to be essentially necessary for ergodicity.

5.2.2.5 Alignment

We need to investigate the relative position of u-manifolds and singularities in order to find out how much of a u-manifold can be 'close' to a singularity. Our aim is to prove the following

Lemma 5.18. *Take any smooth component $\mathcal{Z}$ of $T^{-k}\mathcal{Z}_0$ with $k \geq 0$, where*

$$\mathcal{Z}_0 = \{ \ (s, \varphi) \in M \mid \varphi = \varphi_0 \ \}$$

with any $\varphi_0 \in [-\frac{\pi}{2}, \frac{\pi}{2}]$. *Given some small positive δ let us denote the δ-neighbourhood of $\mathcal{Z}$ by $\mathcal{Z}^{[\delta]}$. There are global constants $C < \infty$ and $\alpha > 0$ such that for any u-manifold W we have*

$$m_W \left(\mathcal{Z}^{[\delta]} \cap W \right) \leq C\delta^{\alpha}, \tag{5.30}$$

where m_W is the Lebesgue measure – the length – on the u-manifold W.

Proof. If $k > 0$, then $\mathcal{Z}$ is an s-manifold, and is transversal to our u-manifold W according to Lemma 5.15, so the statement holds even with $\alpha = 1$.

So take $k = 0$, then $\mathcal{Z}$ is described by $m_{\mathcal{Z}} = 0$. If $\kappa(\varphi)$ remains bounded near φ_0, then for our u-manifold W,

$$\frac{1}{m_+} = \frac{1}{m_-} + R\kappa(\varphi)$$

is bounded (see (5.5) and (5.18)), so the two curves are transversal again, we can choose $\alpha = 1$.

The interesting case is $k = 0$, $\kappa(\varphi) \to \infty$ as $\varphi \to \varphi_0$. In this case (5.18) ensures that $\frac{1}{m_-}$ is negligible – say, less than ε portion – compared to $R\kappa(\varphi)$. This – through (5.5) and the definition $m_+ = \frac{d\varphi_+}{ds_+} = \frac{d\varphi}{ds}$ – implies that for u-manifolds

$$(1 - \varepsilon)R\kappa(\varphi) \leq \frac{ds}{d\varphi} \leq (1 + \varepsilon)R\kappa(\varphi)$$

Integrating this with respect to φ and using the definition of $\kappa(\varphi)$, we get

$$(1 - \varepsilon)R(\Delta\Theta(\varphi) - \Delta\Theta(\bar{\varphi}) \leq s - \bar{s} \leq (1 + \varepsilon)R(\Delta\Theta(\varphi)) - \Delta\Theta(\bar{\varphi}))$$

which means that, close (enough) to a $\kappa(\varphi) \to \infty$ singularity, a u-manifold is (arbitrarily) similar to the graph of the rotation function $\Delta\Theta(\varphi)$. Now the Hölder-continuity of $\Delta\Theta(\varphi)$ required in the regularity condition (Definition 5.7) implies the statement of the Lemma. □

We note that the proof of alignment is the only place where we use our assumption that the rotation function is Hölder-continuous. The above proof shows that Hölder-continuity is indeed a necessary condition for alignment. Alignment is *not* among the conditions of Chernov's theorem which our proof is based on, but we will use it in the proof of the growth properties (Proposition 5.29). At that place it seems to be unavoidable, so we think that Hölder-continuity of the rotation function is needed for Chernov's method to work. On the other hand, as already pointed out earlier, we do not claim that it is a necessary condition for EDC.

5.2.2.6 Homogeneity strips, secondary singularities and homogeneous u-manifolds

Notation.

$$\omega(\varphi) := \frac{2+\kappa(\varphi)}{\cos\varphi} \tag{5.31}$$

We will see that expansion in the e-metric is unbounded as $|\omega(\varphi)| \to \infty$. This certainly happens in the vicinity of $\pm\frac{\pi}{2}$, nevertheless, there can exist other discontinuity values φ_0 with the same property. Big expansion comes together with big variations of expansion (i.e. distortion) rates along u-manifolds. For that reason we need to partition the phase space into homogeneity layers in which $\omega(\varphi)$ is nearly constant. We fix a large integer k_0 (to be specified in Section 5.3) and define for $k > k_0$ the I-strips as

$$I_k = \left\{ (s,\varphi) \mid k^2 \leq |\omega(\varphi)| < (k+1)^2 \right\} \tag{5.32}$$

Recall from Definition 5.7 that whenever $\lim_{\varphi\to\varphi_0} |\omega(\varphi)| = \infty$, there exists an interval $[\varphi_0 - \epsilon, \varphi_0)$ restricted to which $|\omega(\varphi)|$ is a monotonic function of φ. We partition a subinterval of this interval into I-strips, thus k_0 is chosen accordingly large. In case there are several discontinuity points of $\omega(\varphi)$ (with unbounded one-sided limits) we may construct further I-strips, $I_k^{(s)}$, analogously. Here the index s labels the finitely many discontinuities of this kind.

Furthermore take $I_0^{(u)}$; $u = 1 \cdots U$ where the index u labels the finitely many connected components of the complement of all the above layers (that is, the 'remaining part' of the phase space).

We will use the notations Γ_0 for the countably many boundary components of I-strips.

Convention 5.19. *From now on, Γ_0 – just like $\mathcal{S}_0$ and $\mathcal{Z}_0$ before – is considered as part of the boundary of the phase space. That is, we will use a modified phase space $\bar{M}$, whose connected components are the homogeneity strips I_k (and $I_0^{(u)}$).*

In complete analogy with primary singularities we introduce furthermore the notations Γ_1 and $\Gamma^{(n)}$ for the corresponding preimages. The geometric properties of these secondary singularity lines are analogous to those of primary ones (for example, (5.30) applies).

Definition 5.20. *We will say that a* u-manifold is homogeneous *whenever it is contained in one of the homogeneity strips I_k (or $I_0^{(u)}$).*

In sections 5.2.3.2 and 5.3 we will be concerned with u-manifolds that remain homogeneous for several steps of the dynamics.

5.2.3 Regularity properties of unstable manifolds

Definition 5.21. *An unstable manifold is a u-manifold for which all past iterates are u-manifolds as well.*

Analogously, a stable manifold is an s-manifold for which all future iterates are s-manifolds as well.

From the theory of hyperbolic systems (see [12] and references therein) we know that there is a unique unextendable unstable (and similarly a unique unextendable stable) manifold through ($\mu-$)almost every point of $\bar{M}$. Thus it makes sense to talk about the *(un)stable manifold through the point.*

We will also refer to unstable manifolds as 'local unstable manifolds' (LUMs), stressing the fact that they are (and all their past iterates as well are) contained in some homogeneity layer I_k. (Remember that our phase space ends on the boundary of I_k, so I_{k+1} is already another connected component.)

In this subsection we deal with properties of unstable manifolds which are stronger than those proved before for arbitrary u-manifolds in Section 5.2.2.

5.2.3.1 Curvature bounds

In what follows we obtain bounds on unstable manifolds that will guarantee that their curvature is uniformly bounded from above.

First we look at u-fronts as submanifolds of the flow phase space.

Putting the formulas in (5.14) and (5.15) together, we get

$$D'_- = -\frac{D_-}{\lambda^3} + \frac{2\sin\varphi B_-^2}{\lambda^3\cos\varphi} - \frac{2\sin\varphi {B'_-}^2}{\lambda_2\cos\varphi} + \frac{\sin\varphi B_-}{R\lambda^3\cos^2\varphi} + \frac{\sin\varphi B'_-}{R\lambda_2^2\cos^2\varphi} - Rm_-^3\frac{\eta}{\lambda^3\cos^2\varphi}$$

Our key estimate (5.21) implies

$$\cos\varphi|\lambda_2| \sim \sqrt{1+m_+^2}$$

which is bounded from below. So, in the above sum, terms number 2,3,4 and 5 are all bounded in absolute value. The last term is bounded due to our assumption

$$\left|\frac{\eta(\varphi)}{(2+\kappa(\varphi))^3}\right| < C.$$

As a consequence, we have

$$|D'_-| \le \frac{|D_-|}{\Lambda^3} + C_2, \tag{5.33}$$

with some global constant C_2 and can state

Lemma 5.22. *There is a global constant $\hat{D}$ such that for almost any point of the phase space, the front corresponding to the LUM has*

$$|D_-| \le \hat{D}$$

Proof. Choose $\hat{D} = \frac{C_2\Lambda^3}{\Lambda^3-1}$. Now suppose indirectly that there is a set $H \subset M$ of positive measure, for the points of which $|D_-| > \hat{D}+\varepsilon$. Then (5.33) implies that there is a $c(\varepsilon) > 0$ such that $|D_-| > \hat{D}+\varepsilon+c$ on $T^{-1}H$. This implies that $|D_-| > \hat{D}+\varepsilon+2c$ on $T^{-2}H$, and so on: $|D_-| > \hat{D}+\varepsilon+kc$ on $T^{-k}H$ for all $k > 0$. But the $T^{-k}H$-s are all sets of equal positive measure, which contradicts the finiteness of the phase space. □

As a consequence, we can give curvature bounds for local unstable manifolds in the incoming and outgoing phase spaces. Since an unstable manifold in the Poincaré section is the graph of a function $\varphi = \varphi(s)$, its curvature is given by

$$g = \frac{\varphi''(s)}{\sqrt{1+(\varphi'(s))^2}^3} = \frac{K}{\sqrt{1+m^2}^3}.$$

We have reached

Proposition 5.23. *There is a global constant C such that for almost any point of the phase space, the front corresponding to the LUM has*

$$|g_+| \leq C.$$

Proof. It can be read from (5.14) that

$$|K_-| < C, \tag{5.34}$$

thus

$$|g_-| < C.$$

To find out about g_+, we write

$$g_+ = \frac{K_+}{\sqrt{1+m_+^2}^3} = \left(\frac{m_+}{\sqrt{1+m_+^2}}\right)^3 \frac{K_-}{m_-^3} - R\frac{\eta}{\sqrt{1+(\frac{1}{m_-}+R\kappa(\varphi))^2}^3}.$$

This is also bounded in absolute value due to our assumption

$$\left|\frac{\eta(\varphi)}{(2+\kappa(\varphi))^3}\right| < C$$

(see Definition 5.7). □

We note that this proof suggests that our condition

$$|\kappa'(\varphi)| \leq C|(2+\kappa(\varphi))^3|$$

is necessary for bounded curvature, and consequently for Chernov's method to work.

5.2.3.2 Distortion bounds

Length of a u-manifold W is expanded by T^n locally with a factor

$$J_{W,n}(x) = \frac{|DT^n dx|_e}{|dx|_e},$$

where dx is the vector tangent to the curve of W at x. The aim of this subsection is to prove

Proposition 5.24. *Let W be an unstable manifold on which T^n is smooth. Assume that $W_i = T^i W$ is a homogeneous unstable manifold for each $1 \leq i \leq n$. Then for all $x, \bar{x} \in W$*

$$|\ln J_{W,n}(x) - \ln J_{W,n}(\bar{x})| \leq C \left[\mathrm{dist}_{W_n}(T^n x, T^n \bar{x})\right]^{\frac{1}{5}}.$$

Proof. Note that $J_{W,n}(x) = \prod_{i=0}^{n-1} J_{W_i,1}(T^i x)$. Hence, it is enough to prove the proposition for $n = 1$, because $\mathrm{dist}(T^i x, T^i \bar{x})$ grows uniformly exponentially in i due to (5.28). So we put $n = 1$.

Denote $x' = Tx$ and, we will use a ' to denote quantities related to the point x'.

Recall from Section 5.2.2 that the expansion factor is easily calculated in the p-metric. To obtain $J := J_{W,1}(x)$ we transform $|dx|_e$ to $|dx|_p$, take the p-expansion factor from (5.8) and (5.9) and transform back. This way:

$$J = \frac{\sqrt{1+m'^2}}{\cos\varphi'} \lambda_1' \lambda_2 \frac{\cos\varphi}{\sqrt{1+m^2}}.$$

In order to calculate the change in the logarithm of J as we move from x to $\bar{x}$, it is best to write it with the help of (5.31) in the form

$$J = \omega(\varphi') J_1' J_2 \tag{5.35}$$

with

$$J_1 = \frac{\sqrt{1+m_+^2}}{2+\kappa(\varphi)} \lambda_1$$

and

$$J_2 = \frac{\cos\varphi}{\sqrt{1+m_+^2}} \lambda_2.$$

(5.20) and (5.21) imply

$$|J_1| \sim |J_2| \sim 1. \tag{5.36}$$

The change in logarithm of the three factors can be calculated independently, moreover, J_1 and J_2 are expected to change moderately, while $\omega(\varphi')$ can be kept under good control, because it depends only on φ'. The three terms are investigated in three lemmas. Thus Proposition 5.24 is the direct consequence of the three lemmas 5.25, 5.26 and 5.27. Of course, the first and third (concerning J_1 and $\omega(\varphi)$) have to be applied with $'$-es. When applying Lemma 5.27, we use the trivial fact $|\varphi - \bar{\varphi}| \leq \mathrm{dist}(x, \bar{x})$. □

In the arguments below, as usual, quantities with neither $+$ nor $-$ in their index are meant to have a $+$, that is, to be in the outgoing phase space.

Lemma 5.25. *There exists a global constant C such that when a perturbation of size dx is performed on the base point, we have*

$$|d \ln J_1| \leq C|dx|.$$

Proof. In many estimates, we will use – without further mention – that m_- and K_- are bounded (see (5.18) and (5.34)).

With the help of (5.10) we choose the form

$$J_1 = \frac{\sqrt{(1+\kappa(\varphi)Rm_-)^2 + m_-^2}}{2+\kappa(\varphi)}.$$

When calculating the differential, we use

$$d\kappa(\varphi) = \eta(\varphi)d\varphi = \frac{\eta(\varphi)m_+}{\sqrt{1+m_+^2}}dx$$

and

$$dm_- = K_- ds_- = K_- \frac{ds_+}{\lambda_1} = \frac{K_-}{\lambda_1}\frac{dx}{\sqrt{1+m_+^2}} = \frac{K_-}{\sqrt{\lambda_1^2+m_-^2}}dx$$

Calculating the differential, we get

$$\begin{aligned} d \ln J_1 &= \frac{m_- + \kappa(\varphi)R + \kappa^2(\varphi)R^2 m_-}{((1+\kappa(\varphi)Rm_-)^2 + m_-^2)^{3/2}} K_- dx + \\ &+ \frac{2Rm_- - 1 - m_-^2 + (2Rm_- - 1)Rm_-\kappa(\varphi)}{(2+\kappa(\varphi))((1+\kappa(\varphi)Rm_-)^2 + m_-^2)} \frac{\eta(\varphi)m_+}{\sqrt{1+m_+^2}}dx. \end{aligned}$$

The coefficient of dx in the first term is obviously bounded since the denominator is one degree higher in $\kappa(\varphi)$ and is bounded away from zero. In the second term, we use (5.10) and (5.20) to get

$$\left|\frac{m_+}{\sqrt{1+m_+^2}}\right| = \left|\frac{m_-}{\sqrt{1+m_+^2}\lambda_1}\right| \sim \left|\frac{1}{2+\kappa(\varphi)}\right| \tag{5.37}$$

so, looking again at the degrees of polynomials (in κ) in the numerator and denominator of the second term, we have

$$\frac{2Rm_- - 1 - m_-^2 + (2Rm_- - 1)Rm_-\kappa(\varphi)}{(2+\kappa(\varphi))(1+\kappa(\varphi)Rm_-)^2 + m_-^2)} \frac{\eta(\varphi)m_+}{\sqrt{1+m_+^2}} \leq C\left|\frac{\eta(\varphi)}{(2+\kappa(\varphi))^3}\right| \leq C_1.$$

□

Lemma 5.26. *There exists a global constant C such that when a perturbation of size dx is performed on the base point, we have*

$$|d \ln J_2| \leq C|dx'|$$

(note the ' on the right hand side).

Proof. With the help of (5.5) and (5.11) we choose the form

$$J_2 = \frac{\cos\varphi + \tau m_+ + \frac{\tau}{R}}{\sqrt{1+m_+^2}}.$$

When calculating the differential, we use

$$d\varphi = \frac{m_+}{\sqrt{1+m_+^2}} dx$$

and

$$dm_+ = K_+ ds_+ = K_+ \frac{dx}{\sqrt{1+m_+^2}} = (1+m_+^2) g_+ dx.$$

This way we get

$$d\ln J_2 = \frac{-\sin\varphi m_+}{\cos\varphi \lambda_2 \sqrt{1+m_+^2}} dx + B'_- d\tau - \frac{m_+(\frac{\tau}{R} + \cos\varphi) - \tau}{\cos\varphi\lambda_2} g_+ dx.$$

Due to (5.21), the coefficient of dx in the first term is equivalent to $\frac{-\sin\varphi m_+}{1+m_+^2}$, and in the third term to $-\frac{m_+(\frac{\tau}{R}+\cos\varphi)-\tau}{\sqrt{1+m_+^2}} g_+$, both of which are bounded (cf. (5.37)).

We finish by estimating dx and $d\tau$ with dx'. First,

$$dx' = J dx \sim \left| \frac{2+\kappa(\varphi')}{\cos\varphi'} \right| dx \geq c dx. \tag{5.38}$$

Second, the triangle inequality implies $|d\tau| \leq |ds| + |ds'_-|$. On the one hand, (5.38) implies $|ds| \leq |dx| \leq C|dx'|$. On the other hand (5.20) implies,

$$dx' = \sqrt{1 + {m'_+}^2} |\lambda'_1| ds'_- \sim |2+\kappa(\varphi')| ds'_- \geq c ds'_-.$$

These give

$$|d\tau| \leq C|dx'|.$$

□

Lemma 5.27. *There exists a global constant C such that if $x = (s,\varphi)$ and $\bar{x} = (\bar{s}, \bar{\varphi})$ are in the same homogeneity layer*

$$I_k = \left\{ (s,\varphi) \mid k^2 \leq |\omega(\varphi)| < (k+1)^2 \right\},$$

then

$$|\ln|\omega(\varphi)| - \ln|\omega(\bar{\varphi})|| \leq C|\varphi - \bar{\varphi}|^{1/5}.$$

Proof. We use the notation $\omega'(\varphi) = \frac{d}{d\varphi}\omega(\varphi)$. It is easy to see that the regularity of $\kappa(\varphi)$ implies

$$\left| \frac{\omega'(\varphi)}{\omega^3(\varphi)} \right| \leq C.$$

That is, everywhere inside J_k,

$$\left| \frac{d|\ln\omega(\varphi)|}{d\varphi} \right| = \left| \frac{\omega'(\varphi)}{\omega(\varphi)} \right| \leq C|\omega(\varphi)|^2 \leq 2Ck^4.$$

This, together with the obvious $k^2 \le |\omega(\varphi)|, |\omega(\bar\varphi)| < (k+1)^2$, implies

$$|\ln|\omega(\varphi)| - \ln|\omega(\bar\varphi)|| \le \min\left\{2Ck^4|\varphi-\bar\varphi|, \ln(k+1)^2 - \ln k^2\right\} \le \min\left\{2Ck^4|\varphi-\bar\varphi|, \frac{2}{k}\right\}.$$

It is easy to check that for every k and every ξ

$$\min\left\{2Ck^4|\xi|, \frac{2}{k}\right\} \le 2C^{1/5}\xi^{1/5},$$

which completes the proof. □

5.2.3.3 Absolute continuity

After proving that the expansion factors vary nicely between nearby points on the same u-manifold, we now investigate their behaviour at points of different u-manifolds that lie on the same s-manifold. This is the absolute continuity property. Just like it was with the distortion bounds, it is important to consider homogeneous manifolds.

We introduce the simplified notation $J^u_k(x)$ and $J^s_k(x)$ for the k-step length expansion factor at x along the unstable and the stable manifold, respectively.

Proposition 5.28. *Let W_s be a small s-manifold, $x, \bar x \in W_s$, and $W_u, \bar W_u$ two u-manifolds crossing W_s at x and $\bar x$, respectively. Assume that T^k is smooth on W_s and T^iW_s is a homogeneous s-manifold for each $0 \le i \le k$. Then*

$$|\ln J^u_k(x) - \ln J^u_k(\bar x)| \le C$$

where C is a global constant.

Proof. We have bounds on the change in expansion as we move along unstable manifolds. In order to have such bounds as we move along stable manifolds, we wish to use the fact that stable manifolds are turned into unstable ones when we revert time. However, this time reflection symmetry is not complete: we always work in the outgoing Poincaré section, and reverting time turns this into the incoming one. To deal with the problem, we introduce the map P which is the dynamics through the potential, and which maps from the incoming to the outgoing Poincaré section. That is,

$$P((s_-, \varphi)) := (s_+, \varphi) = (s_- + R\Delta\Theta(\varphi), \varphi).$$

We can see from (5.5) that if $dx_- = (ds_-, d\varphi)$ is a tangent vector of the incoming phase space, then

$$|DP(dx_-)|_e = \sqrt{1+m_+^2}|\lambda_1|\frac{1}{\sqrt{1+m_-^2}}|dx_-|_e.$$

Denote by $\nu(x)$ the expansion factor of DP along the unstable manifold at x, that is $\nu(x) = \frac{|DP(dx_-)|_e}{|dx_-|_e}$ where dx is an unstable vector at x. We can use (5.20) and (5.18) to get

$$\nu(x) \sim |2 + \kappa(\varphi)|.$$

We also introduce the 'turn back' operator, which we will denote by a '$-$' sign: this turns incoming phase points into outgoing phase points which corresponds to reverting the velocity. '$-$' is almost the identity function from M_- to M_+, only the collision angle is reverted (see our sign convention in figure 5.3):

$$- \;:\; M_- \to M_+$$

$$-(s,\varphi_-) := (s_+, \varphi_+) = (s_-, -\varphi_-).$$

With these notations, if $x = P(y)$, the time reflection symmetry implies

$$J_k^s(x) = \frac{\nu(-x)}{J_k^u(-T^k y)\nu(-T^k x)} \sim \frac{1}{J_k^u(-T^k y)} \frac{|2+\kappa(\varphi)|}{|2+\kappa(\varphi_k)|} \tag{5.39}$$

The transversality of stable and unstable vectors, stated in Proposition 5.17 implies that $J_k^u(x)J_k^s(x)$ is equivalent to the k-step (Lebesgue) volume expansion factor. Using (5.2), and the T-invariance of μ, we get

$$J_k^u(x)J_k^s(x) \sim \frac{\cos\varphi}{\cos\varphi_k} \tag{5.40}$$

Putting together (5.39) and (5.40) we get

$$J_k^u(x) \sim J_k^u(-T^k y)\frac{2+\kappa(\varphi_k)}{2+\kappa(\varphi)}\frac{\cos\varphi}{\cos\varphi_k} = J_k^u(-T^k y)\frac{\omega(\varphi_k)}{\omega(\varphi)}$$

The same is true for $\bar{x} = P(\bar{y})$, so we have

$$|\ln J_k^u(x) - \ln J_k^u(\bar{x})| \le |\ln J_k^u(-T^k y) - \ln J_k^u(-T^k \bar{y})| + \left|\ln\left|\frac{\omega(\varphi_k)}{\omega(\bar{\varphi}_k)}\right|\right| + \left|\ln\left|\frac{\omega(\varphi)}{\omega(\bar{\varphi})}\right|\right| + C.$$

To see the boundedness of the first term of the right hand size we can apply Proposition (5.24), because $-T^k y$ and $-T^k\bar{y}$ are on the same local *un*stable manifold. The second and third term is bounded because W_s and $T^k W_s$ are homogeneous, see Definition 5.20. Now the proof of Proposition 5.28 is complete. $\square$

5.3 Growth properties of unstable manifolds

This last section og the proof is concerned with the growth properties of LUMs. Our aim is to show that LUMs "grow large and round, on the average". This is expressed in the formulas of Proposition 5.29 below.

Recall Convention 5.16. Throughout the section we use the higher iterate of the dynamics, $T_1 = T^{m_0}$. This has singularity set (secondary and primary) $\Xi = \Gamma^{(m_0)}$. For the higher iterates of T_1 the singularity set is $\Xi^{(n)} = \Xi \cup {T_1}^{-1}\Xi \cup \ldots \cup T^{-n+1}\Xi$.

δ_0**-LUM's**. To formulate and prove further important conditions on growth of LUMs we need to recall several notions and notations from [12]. Let $\delta_0 > 0$. We call W a δ_0-LUM if it is a LUM and $\mathrm{diam} W \le \delta_0$. For an open subset $V \subset W$ and $x \in V$ denote by $V(x)$ the connected component of V containing the point x. Let $n \ge 0$. We call an open subset $V \subset W$ a (δ_0, n)-subset if $V \cap (\Xi^{(n)}) = \emptyset$ (i.e., the map T_1^n is smooth and homogeneous on V) and $\mathrm{diam} T_1^n V(x) \le \delta_0$ for every $x \in V$. Note that $T_1^n V$ is then a union of δ_0-LUM's. Define a function $r_{V,n}$ on V by

$$r_{V,n}(x) = d_{T_1^n V(x)}(T_1^n x, \partial T_1^n V(x)).$$

Note that $r_{V,n}(x)$ is the radius of the largest open ball in $T_1^n V(x)$ centered at $T_1^n x$. In particular, $r_{W,0}(x) = d_W(x, \partial W)$.

One further notation we introduce is $\mathcal{U}_\delta$ (for any $\delta > 0$), the δ-neighbourhood of the closed set $\Xi \cup S_0 \cup Z_0$.

The aim of this section is to prove the Proposition below.

Proposition 5.29. *There are constants $\alpha_0 \in (0,1)$ and $\beta_0, D_0, \eta, \chi, \zeta > 0$ with the following property. For any sufficiently small $\delta_0, \delta > 0$ and any δ_0-LUM W there is an open $(\delta_0, 0)$-subset $V_\delta^0 \subset W \cap \mathcal{U}_\delta$ and an open $(\delta_0, 1)$-subset $V_\delta^1 \subset W \setminus \mathcal{U}_\delta$ (one of these may be empty) such that $m_W(W \setminus (V_\delta^0 \cup V_\delta^1)) = 0$ and that $\forall \varepsilon > 0$*

$$m_W(r_{V_\delta^1,1} < \varepsilon) \leq \alpha_0 \Lambda_1 \cdot m_W(r_{W,0} < \varepsilon/\Lambda_1) + \varepsilon \beta_0 \delta_0^{-1} m_W(W), \tag{5.41}$$

$$m_W(r_{V_\delta^0,0} < \varepsilon) \leq D_0 \delta^{-\eta}\, m_W(r_{W,0} < \varepsilon) \tag{5.42}$$

and

$$m_W(V_\delta^0) \leq D_0\, m_W(r_{W,0} < \zeta \delta^\chi). \tag{5.43}$$

Proof of this Proposition goes along the lines of the arguments from [12]. First let us consider

Accumulation of singularity lines. There are two sources of accumulation of the components of the set Ξ that can cut LUM's into arbitrary many pieces.

First, the set Γ_1 consists of countably many curves stretching approximately parallel to some curves in $\mathcal{S}_1$ (or $\mathcal{Z}_1$) and approaching them. So, each set $T^{-1} I_k$ and $k \neq 0$, is a narrow strip with curvilinear boundaries. The expansion of unstable fibers in these strips can be estimated using (5.35), (5.36) and (5.32). More precisely, let $W \subset T^{-1} I_k$ be a LUM, for some $k \neq 0$. Then the expansion factor, $J^u(x)$, on W satisfies

$$J^u(x) \sim \omega(\varphi) \sim k^2 \qquad \forall x \in W. \tag{5.44}$$

Second, there might be multiple intersections of the curves in $\mathcal{S}_1 \cup \mathcal{Z}_1$. Recall K_n, the complexity of $\mathcal{S}^{(n)} \cup \mathcal{Z}^{(n)}$ and its properties from Section 5.1.4. Specifically important for us is the choice of the higher iterate $T_1 = T^{m_0}$ with its relevant properties, see Convention 5.16.

Indexing system. Before proving the proposition we introduce a handy indexing system, cf. [12]. Let $\delta_0 > 0$ and W be a δ_0-LUM. If δ_0 is small enough, then W crosses at most K_{m_0} curves of the set $\mathcal{S}^{(m_0)} \cup \mathcal{Z}^{(m_0)}$, so the set $W \setminus (\mathcal{S}^{(m_0)} \cup \mathcal{Z}^{(m_0)})$ consists of at most $K_{m_0} + 1$ connected curves, let us call them $W_1, \ldots, W_p$ with $p \leq K_{m_0} + 1$.

On each W_j the map T_1 (as a map on M) is smooth, but any W_j may be cut into arbitrary many (countably many) pieces by other curves in Ξ, which are the preimages of the boundaries of I_k. Let $\Delta \subset W$ be a connected component of the set $W \setminus \Xi$. It can be identified with the (m_0+1)-tuple $(k_1, \ldots, k_{m_0}; j)$ such that $\Delta \subset W_j$ and $T^i \Delta \subset I_{k_i}$ for $1 \leq i \leq m_0$. Note that this identification is almost unique. Indeed, given j, $(T^i \Delta \subset) T^i W_j$ is contained in a strip of the phase space that lies between two horizontal lines: two components of $\mathcal{S}_0 \cup \mathcal{Z}_0$. It might happen that expansion factors diverge – and consequently, homogeneity strips have been constructed – at both sides of the strip. Thus given the index k_i, we have $T^i \Delta \subset I_{k_i}$, where I_{k_i} can be the k_i-th layer from one of the two homogeneity structures. In such a case we use the following convention; the homogeneity layers at the 'upper' and 'lower' ends of the phase space strip (corresponding to j) are labeled by odd and even numbers, respectively. This way the indexing system is made unique and (5.44) remains true.

All in all, we will write $\Delta = \Delta(k_1, \ldots, k_{m_0}; j)$. Of course, some strings $(k_1, \ldots, k_{m_0}; j)$ may not correspond to any piece of W, for such strings $\Delta(k_1, \ldots, k_{m_0}; j) = \emptyset$.

Denote by $J_1^u(x) = J^u(x) \cdot \ldots \cdot J^u(T^{m_0-1}x)$ the expansion factor of unstable vectors under DT_1. Let $|\Delta| = m_\Delta(\Delta)$ be the Euclidean length of a LUM Δ. We record two important facts:

(a) For every point $x \in \Delta(k_1, \ldots, k_{m_0}; j)$ we have

$$J_1^u(x) \geq L_{k_1,\ldots,k_{m_0}} := \max\left\{\Lambda_1,\, C_{20} \prod_{k_i \neq 0} k_i^2\right\}$$

where C_{20} is some positive global constant. This follows from (5.44).

(b) For each $\Delta(k_1, \ldots, k_{m_0}; j)$ we have

$$|\Delta(k_1, \ldots, k_{m_0}; j)| \leq M_{k_1,\ldots,k_{m_0}} := \min\left\{|W|,\, C_{21} \prod_{k_i \neq 0} k_i^{-2}\right\}$$

where $C_{21} = C_{20}^{-1}|W|_{\max}$ and $|W|_{\max}$ is the maximal length of LUMs in M. This follows from the previous fact.

Next, put

$$\theta_0 := 2\sum_{k=k_0}^{\infty} k^{-2} \leq 4/k_0$$

and let us turn to the proof of our growth formulas.

Let W be a δ_0-LUM and $\delta > 0$ be small. For each connected component $\Delta \subset W \setminus \Xi$ put $\Delta^0 = \Delta \cap \mathcal{U}_\delta$ and $\Delta^1 = \mathrm{int}(\Delta \setminus \mathcal{U}_\delta)$ (recall $\mathcal{U}_\delta$ is the δ-neighbourhood of $\Xi \cup \mathcal{S}_0 \cup \mathcal{Z}_0$). Due to the Continuation property (cf. Section 5.1.4) and to Alignment (cf. Lemma 5.18), the set Δ^0 consists of two subintervals adjacent to the endpoints of Δ (they may overlap and cover Δ, of course). The set Δ^1 is either empty or a subinterval of Δ. We put $W^1 = \cup_{\Delta \subset W \setminus \Xi} \Delta^1$.

Proof of (5.41). For each Δ^1 the set $T_1(\Delta_1 \cap \{r_{W^1,1} < \varepsilon\})$ is the union of two subintervals of $T_1\Delta^1$ of length ε adjacent to the endpoint of $T_1\Delta^1$. Using the above indexing system we get

$$\begin{aligned} m_W(r_{W^1,1} < \varepsilon) &\leq \sum_{k_1,\ldots,k_{m_0},j} 2\varepsilon L^{-1}_{k_1,\ldots,k_{m_0}} \\ &\leq 2\varepsilon p\left[\Lambda_1^{-1} + C_{20}{}^{-1}(\theta_0 + \theta_0^2 + \ldots + \theta_0^{m_0})\right] \\ &\leq 2\varepsilon(K_{m_0}+1)\left(\Lambda_1^{-1} + C_{20}{}^{-1} m_0\theta_0\right). \end{aligned}$$

We now assume that k_0 is large enough so that

$$\alpha_0 := (K_{m_0}+1)(\Lambda_1^{-1} + C_{20}{}^{-1} m_0\theta_0) < 1$$

and thus get

$$m_W(r_{W^1,1} < \varepsilon) \leq \min\{|W|, 2\alpha_0\varepsilon\}.$$

The first term on the right hand side of (5.41) is equal to

$$\alpha_0\Lambda_1 \min\{|W|, 2\varepsilon/\Lambda_1\} = \min\{\alpha_0\Lambda_1|W|, 2\alpha_0\varepsilon\}.$$

Since $\alpha_0\Lambda_1 > 1$, we get

$$m_W(r_{W^1,1} < \varepsilon) \le \alpha_0\Lambda_1 \cdot m_W\left(r_{W,0} < \varepsilon/\Lambda_1\right). \tag{5.45}$$

Next, to obtain an open $(\delta_0, 1)$-subset V_δ^1 of W^1, one needs to further subdivide the intervals $\Delta^1 \subset W$ such that $|T_1\Delta^1| > \delta_0$. Each such LUM $T_1\Delta^1$ we divide into s_Δ equal subintervals of length $\le \delta_0$, with $s_\Delta \le |T_1\Delta^1|/\delta_0$. If $|T_1\Delta^1| < \delta_0$, then we set $s_\Delta = 0$ and leave Δ^1 unchanged. Then the union of the preimages under T_1 of the above intervals will make V_δ^1. Now we must estimate the measure of the ε-neighbourhood of the additional endpoints of the subintervals of $T_1\Delta^1$. This gives

$$\begin{aligned} m_W(r_{V_\delta^1,1} < \varepsilon) - m_W(r_{W^1,1} < \varepsilon) &\le \sum_{\Delta\subset W\setminus\Xi} 2s_\Delta\varepsilon|C_{22}\Delta^1|/|T_1\Delta^1| \\ &\le \sum_{\Delta\subset W\setminus\Xi} 2C_{22}\varepsilon|\Delta^1|/\delta_0 \\ &\le 2C_{22}\varepsilon\delta_0^{-1}|W|. \end{aligned}$$

Here $C_{22} = \exp(\mathrm{const}\cdot|W|_{\max}^{\frac{1}{5}})$ is an upper bound on distortions on LUM's, see Proposition 5.24. Combining the above bound with (5.45) completes the proof of (5.41) with $\beta_0 = 2C_{22}$.

We now *prove (5.42)*. It is enough to consider $\varepsilon < |W|/2$, so that the right hand side of (5.42) equals $2D_0\delta^{-\eta}\varepsilon$. We can put $V_\delta^0 = W \setminus \overline{V_\delta^1}$. Then the left hand side of (5.42) does not exceed $2J_\delta\varepsilon$, where J_δ is the number of nonempty connected components of the set $\overline{V_\delta^0}$, which is at most the number of connected components of $W \setminus \Xi$ of length $> 2\delta$. Hence, clearly $J_\delta \le |W|/\delta \le \delta_0/\delta$. This proves (5.42) with $\eta = 1$.

Finally, we *prove the inequality (5.43)*. Again, let Δ be a connected component of $W \setminus \Xi$ and Δ^0, Δ^1 be defined as above, with the set Δ^0 consisting of two subintervals adjacent to the endpoints of Δ. By (5.30) – and the analogous property for the secondary singularities, see Sections 5.2.2.5 and 5.2.2.6 – each of these subintervals has length smaller than $C\delta^\alpha$.

Now, the right hand side of (5.43) equals $D_0 \min\{|W|, 2\zeta\delta^\chi\}$. So, it is enough to show that $m_W(V_\delta^0) \le B\delta^\chi$ for some $B, \chi > 0$. We have

$$\begin{aligned} m_W(V_\delta^0) &\le \sum_{\Delta\subset W\setminus\Xi} \min\{2C\delta^\alpha, |\Delta|\} \\ &\le \sum_{k_1,\dots,k_{m_0},j} \min\{2C\delta^\alpha, M_{k_1,\dots,k_{m_0}}\} \\ &\le \mathrm{const}\cdot\delta^\alpha + \mathrm{const}\cdot \sum_{k_1,\dots,k_{m_0}}^{*} \min\left\{\delta^\alpha, \prod_{k_i\neq 0} k_i^{-2}\right\} \end{aligned}$$

where $\sum^*$ is taken over m_0-tuples that contain at least one nonzero index $k_i \neq 0$. The following Lemma – Lemma 7.2 from [12], which was proved in the Appendix of that paper – completes the proof of (5.43) with $\chi = \frac{\alpha}{2m_0}$.

Lemma 5.30. *Let $\epsilon > 0$ and $m \geq 1$. Then*

$$\sum_{k_1,\ldots,k_m \geq 2} \min\left\{\epsilon, (k_1 \cdot \ldots \cdot k_m)^{-2}\right\} \leq B(m) \cdot \epsilon^{1/2m}.$$

With the help of this lemma Proposition 5.29, and, consequently, Theorem 5.9 is proved. □. □.

5.4 Specific potentials

In this section we would like to show that, as important corollaries of Theorem 5.9, exponential decay of correlations can be established for certain specific potentials. To prove such corollaries we need to calculate the rotation function $\Delta\Theta(\varphi)$ from the potential $V(r)$.

As to the detailed description of the Hamiltonian flow in a circularly symmetric potential, we refer to the literature, e.g. [17] and references therein. Most important is that besides the full energy there is an additional integral of motion, the angular momentum l, that can be calculated for a specific trajectory as

$$l = R \sin\varphi$$

where φ is the collision angle at income. For brevity of notation it is worth introducing the function

$$h(r) = (1 - 2V(r))r^2.$$

By the presence of the angular momentum motion is completely integrable and is described by the pair of differential equations (recall our convention that the full energy is $E = \frac{1}{2}$):

$$\begin{aligned} \dot{r}^2 &= r^{-2}(h(r) - l^2) \\ r^2\dot{\Theta} &= l. \end{aligned}$$

Combining these we get

$$\frac{d\Theta}{dr} = \pm \frac{l}{r\sqrt{h(r) - l^2}} \tag{5.46}$$

where the sign depends on whether r is increasing or decreasing. More precisely, there is a minimum radius

$$\hat{r} = \hat{r}(\varphi): \qquad h(\hat{r}) = l^2 = R^2 \sin^2\varphi,$$

down to which r decreases (with negative sign in (5.46)) and from which r increases (with positive sign in (5.46)). This results in

$$\Delta\Theta(\varphi) = 2\int_{\hat{r}}^{R} \frac{l}{r\sqrt{h(r) - l^2}} dr. \tag{5.47}$$

For a generic potential, the dependence of (5.47) on φ is rather implicit: φ is present both in the integrand (via l) and in the limits (via $\hat{r}$). One possible strategy to follow is to obtain some even more complicated formulas for the derivatives in the general case, and based on those perform estimates that guarantee the desired dynamical properties.

This is possible as long as only hyperbolicity and ergodicity is treated – like in [17] – and thus only the first derivative, $\kappa(\varphi) = \Delta\Theta'(\varphi)$ is needed. However, for rate of mixing you need one more derivative, $\kappa'(\varphi) = \Delta\Theta''(\varphi)$, cf. Definition 5.7. Finding good sufficient conditions on the potential $V(r)$ that guarantee the regularity of κ seems to be a very hard task, if possible at all. Thus we have chosen instead to investigate some specific cases where $\Delta\Theta$ is directly computable from (5.47). Of course, this way we could handle a much narrower class of potentials than [17], nevertheless, the established dynamical property is stronger.

Corollary 5.31. *Consider the case of a constant potential,*

$$V(r) = V_0 \text{ for any } r \in [0, R).$$

Correlations decay with an exponential rate in case

- $V_0 > 0$ *and the configuration is arbitrary,*
- $V_0 < 0$ *and the configuration is such that* $\tau_{\min} > \frac{2R}{\sqrt{1-2V_0}-1}$.

Remark 5.32. *Actually, the analysis of this constant potential case from the point of ergodicity dates back to the late eighties, to [25] and [1]. Rate of mixing is, to our knowledge, discussed for the first time. For potential values* $V_0 > \frac{1}{2}$ *the particle cannot enter the disks, the system is equivalent to the traditional dispersing billiard, thus we consider the opposite case,* $V_0 < \frac{1}{2}$.

Proof. Let us introduce the quantity

$$\nu = \sqrt{1 - 2V_0} \tag{5.48}$$

which is less or greater than 1 depending on the sign of V_0. Let us consider the case of positive V_0 first and introduce furthermore the angle φ_0 for which:

$$\nu = \sin\varphi_0.$$

In case $|\varphi| > \varphi_0$, $|l|$ is greater than the maximum value $h(r)$ can take, which indicates that the particle has too large angular momentum to enter the potential, thus $\Delta\Theta = 0$. In the opposite case of $|\varphi| < \varphi_0$ it is easy to obtain $\hat{r} = \frac{R|\sin\varphi|}{\nu}$ and perform the integration of (5.47). All in all

$$\Delta\Theta(\varphi) = \begin{cases} 2\arccos\left(\frac{\sin(\varphi)}{n}\right) & \text{if } |\varphi| < \varphi_0, \\ 0 & \text{if } |\varphi| \geq \varphi_0. \end{cases}$$

See the left half of Figure 5.4.

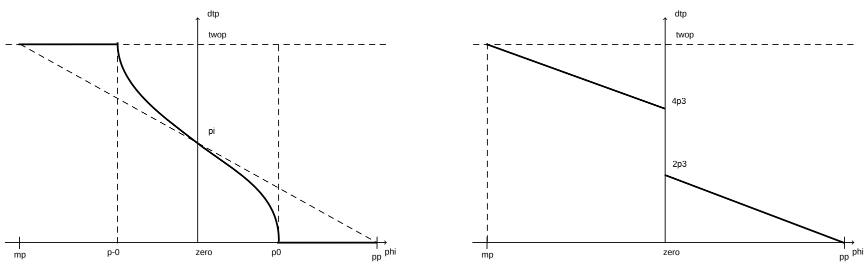
On the one hand, whatever a configuration we have, the system satisfies property H (cf. Definition 5.5), as either $\kappa = 0$ or $\kappa \leq \frac{-2}{\nu} < -2$. On the other hand, κ is a piecewise C^1 function of φ and it behaves as $(\varphi_0 - \varphi)^{-\frac{1}{2}}$ near the discontinuity point φ_0. Thus κ is regular (cf. Definition 5.7 and the remarks following it). This means that the first statement of our Corollary follows from Theorem 5.9.

Now let us turn to the case of $V_0 < 0$ (i.e. $\nu > 1$). It is even simpler to calculate the rotation function (5.47):

$$\Delta\Theta(\varphi) = 2\arccos\left(\frac{\sin\varphi}{\nu}\right)$$

for all φ. As $\nu > 1$, this is a C^2 function on the interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$, thus κ is definitely regular. As to property H, we have $0 > \kappa \geq -\frac{2}{\nu}$ where the minimum is obtained at $\varphi = 0$. Thus the assumption on the configuration from Definition 5.5 reads as $\tau_{\min} > \frac{2R}{\nu-1}$ and the second statement of the Corollary follows from Theorem 5.9. $\square$

Remark 5.33. *Note that motion in the constant potential is equivalent to the problem of diffraction form geometric optics. More precisely, we can think of the disks as if they were made of a material optically different from their neighbourhood, where the relative diffraction coefficient is ν from (5.48). In case the disks are optically less dense than their neighbourhood (i.e. $\nu < 1$, $V_0 > 0$), we may observe the phenomenon of complete reflection that corresponds to the limiting angle φ_0.*



(a) $V(r) = V_0 (> 0)$ (b) $V(r) = \frac{1}{2}\left(1 - \left(\frac{r}{R}\right)^{\beta}\right)$ $(\beta > 0)$

Figure 5.4: rotation function for two examples

Corollary 5.34. *Given constants $A > 0$ and $\beta > -2$, consider the potential*

$$V(r) = A\left(1 - \left(\frac{r}{R}\right)^{\beta}\right).$$

Correlations decay at an exponential rate in case:

- $A = \frac{1}{2}$, $0 > \beta (> -2)$ *and the configuration is arbitrary,*
- $A = \frac{1}{2}$, $\beta > 0$ *and the configuration is such that* $\tau_{\min} > \frac{2R}{\beta}$.

Remark 5.35. *Note that according to our construction the chosen value for the constant A, $A = \frac{1}{2}$ is exactly the full energy. If we had a different value for A, the integration in (5.47) would be much more complicated. In other words, Corollary 5.34, in contrast to Corollary 5.31 is unstable with respect to variations of the full energy (see also the discussion below, following the proof). Nevertheless it is nice to have at least one potential with exponential mixing for any kind of power law behaviour (if $\beta \leq -2$, a positive measure set of trajectories is pulled into the center of the disk, cf. [17]).*

Proof. By straightforward calculation

$$h(r) = \frac{r^{2+\beta}}{R^{\beta}}; \quad \text{and} \quad \hat{r} = R\,|\sin\varphi|^{\frac{2}{2+\beta}}.$$

Then it is not hard to integrate in (5.47):

$$\Delta\Theta(\varphi) = \frac{4}{2+\beta}\left(\frac{\pi}{2} - \varphi\right)$$

for all $\varphi \neq 0$. See the right half of Figure 5.4. Thus $\Delta\Theta$ is piecewise linear (in the general case with one discontinuity of the first kind at $\varphi = 0$) and thus

$$\kappa = -\frac{4}{2+\beta}$$

identically. Regularity (in terms of Definition 5.7) is automatic.

Let us consider the attracting potentials, $\beta < 0$ first. In such a case the potential has a singularity at the center of the disk, resulting in the discontinuity at $\varphi = 0$. [3] Nevertheless, $\kappa < -2$, thus property H (cf. Definition 5.5) and consequently the first statement of the Corollary follows.

Now if $\beta > 0$, as $A = \frac{1}{2}$, the 'top' of the potential is equal to the energy. As a consequence, for the initial value $\varphi = 0$ the flow is not uniquely defined, resulting in the discontinuity for the rotation function. However, in accordance with Definition 5.5, property H is satisfied if $\tau_{\min} > \frac{2R}{\beta}$. Thus the second statement of the Corollary holds. □

5.4.1 Discussion

As already mentioned, Corollary 5.34 is much sensitive to the convention $E = \frac{1}{2}$. Though very difficult to calculate, it is interesting to guess what happens if one perturbs the constant A (or equivalently, the full energy level).

Let us consider the case $\beta > 0$ first. With A either increased or decreased from the value $\frac{1}{2}$, the physical reason for the discontinuity at $\varphi = 0$ disappears and we expect smooth rotation functions. By continuity of the potential at R, $\Delta\Theta(\frac{\pi}{2}) = 0$ seems also reasonable. As to the initial value $\varphi = 0$ let us have a look at the case $A < \frac{1}{2}$ first. There is no reason for the trajectory to deviate in direction: it slows down, reaches the center and then speeds up following a linear track. Thus $\Delta\Theta(0) = \pi$. This altogether implies on basis of Lagrange's mean value theorem that there definitely exists at least one $\varphi \in (0, \frac{\pi}{2})$ for which $\kappa(\varphi) = -2$. In such a case, however, stable periodic orbits tend to appear and the system is most likely not even ergodic, cf. [16]. One can suspect that a typical repelling potential which has a maximum less than the total energy, leads to non-ergodic soft billiards in a similar fashion.

In the opposite case of $A > \frac{1}{2}$ the behaviour of trajectories in the vicinity of $\varphi = 0$ is completely different. As the top of the potential is higher then the full energy, the particle cannot 'climb' it thus it should 'turn back'. We expect $\Delta\Theta(0) = 0$ and a smooth rotation function with $\kappa > -2$ for all φ. That would mean ergodicity and possibly exponential mixing in case of a suitable configuration (cf. Definition 5.5). All in all, ergodic and statistical behaviour is much sensitive to perturbation of the full energy level.

In case of $\beta < 0$ it is no so easy to guess. Nevertheless, we can say something rather surprising in one particular case that indicates similar sensitivity. Choose $\beta = -1$ and

[3] However, in case $\beta = -2(1 - \frac{1}{n})$, the left and right limits coincide, this corresponds to the possibility of regularizing the flow, cf. [17] and [24].

$A = 1$. It is not difficult to obtain $h(r) = 2r - r^2$. The integral in (5.47) is a bit more complicated now, nevertheless, its is possible to evaluate:

$$\Delta\Theta(\varphi) = 2\pi - 2\varphi \tag{5.49}$$

which means $\kappa = -2$ identically. This corresponds to the least ergodic behaviour we can have. It is straightforward to obtain that an identically zero potential ($V(r) = 0$ for all r) would result in $\Delta\Theta(\varphi) = \pi - 2\varphi$. Thus by (5.49) in this particular case of $A = 1, \beta = -1$ trajectories evolve as if they passed on freely and were reflected when leaving the disc.

Thus if $\beta = -1$, we may have exponential mixing ($A = \frac{1}{2}$) and stability ($A = 1$). As to other values of A it is worth mentioning that ergodicity follows from [17] in case $A < \frac{1}{2}$.

Chapter 6

Outlook

As described in the Introduction, much of this work is aimed at exponential decay of correlations in high-dimensional systems. Moving in that direction, a first step is already taken: hyperbolicity is proven for a class of multi-dimensional soft billiards [7]. To proceed towards decay of correlations, we plan to follow the way described here.

1. As to the possibly most direct challenge, we conjecture that there exist rapidly mixing potentials for which the condition $|\kappa + 2| > c$ (i.e. property H from Definition 5.5) is not satisfied for nearly tangential trajectories. Thus these systems are not covered by Theorem 5.9, even more, at least to our knowledge, there is no result in the literature on the ergodicity or hyperbolicity of such soft billiards either. Thus we make the following

   Remark 6.1. *Note that it is possible that κ tends to -2 as $\varphi \to \frac{\pi}{2}$, nevertheless, $|\frac{2+\kappa}{\cos\varphi}| > c$ and the system can be hyperbolic (possibly ergodic or exponentially mixing). This question is under investigation.*

   The difficulty with the treatment of this case is, as already mentioned in Section 5.1, that the separate investigation of motion inside and outside the disks seems not to work at several arguments.

   For such a system, even a proof of hyperbolicity in two dimensions would be interesting in itself.

2. In the higher dimensional case, that is, softenings of multi-dimensional dispersing billiards (motivated e.g. by the three dimensional Lorentz process with spherical scatterers), even ergodicity seems to be difficult. That is because the algebraic approach to handle singularities does not work for soft systems – so no fundamental theorem is known at the moment. That is why a soft system where $|\frac{2+\kappa}{\cos\varphi}|$ is bounded away from both zero and infinity may be easier to handle. In this case the pathology in the singularities does not appear.

At the same time, it would be very nice to get rid of the algebraicity assumption in the Fundamental Theorem. As mentioned before, while all the important hard billiard examples are algebraic, soft billiards are *not*, so proof of their ergodicity in high dimensions is presently blocked by the weakness of the Fundamental Theorem.

Another direction of future research, motivated mainly by applications to Physics, could be the further investigation of those systems for which rapid mixing is already established. For example, as mathematical evidence on the existence of diffusion and other transport coefficients is given, it would be interesting to understand the dependence of these on certain parameters like the full energy level.

Last but not least, in contrast to the generality of Theorem 5.9, it is striking how narrow the class of specific potentials is for which we could apply the result in Section 5.4. It would be desirable to establish – at least numerically – our reasonable regularity properties for as wide a class of potentials as possible.

Acknowledgements

I first thank all my co-authors that I could work with them and learn from them, and we could solve these problems together. Péter Bálint deserves special thanks for the careful reading of the draft and all his valuable suggestions. But most of all, I thank my advisor, Domokos Szász for his help in every detail of my PhD work.

I should thank Nándor Simányi for many valuable discussions and suggestions, and the Department of Stochastics at the Technical University of Budapest for providing excellent conditions for the work. Last but not least, I thank my family for their support and patience.

Appendix

Here we provide, for the reader's convenience, a very short, yet mainly self-contained formulation of Theorem 2.1 from [12]. For self-containedness, many notions and notations are repeatedly introduced. First we give the conditions **P0 . . . P6** which are required, and then the statement of the theorem.

P0. The dynamical system *is a map* $T : M \setminus \Gamma \to M$*, where* M *is an open subset in a* C^∞ *Riemannian manifold,* $\bar{M}$ *is compact.* Γ *is a closed subset in* $\bar{M}$*, and* T *is a* C^2 *diffeomorphism of its range onto its image.* Γ *is called the singularity set.*

P1. Hyperbolicity. *We assume there are two families of cone fields* C^u_x *and* C^s_x *in the tangent planes* $\mathcal{T}_x M$*,* $x \in \bar{M}$ *and there exists a constant* $\Lambda > 1$ *with the following properties:*

- $DT(C^u_x) \subset C^u_{Tx}$ *and* $DT(C^s_x) \supset C^s_{Tx}$ *whenever* DT *exists;*
- $|DT(v)| \geq \Lambda |v| \quad \forall v \in C^u_x$*;*
- $|DT^{-1}(v)| \geq \Lambda |v| \quad \forall v \in C^s_x$*;*
- *these families of cones are continuous on* $\bar{M}$*, their axes have the same dimensions across the entire* $\bar{M}$ *which we denote by* d_u *and* d_s*, respectively;*
- $d_u + d_s = \dim M$*;*
- *the angles between* C^u_x *and* C^s_x *are uniformly bounded away from zero:*

$$\exists\ \alpha > 0 \text{ such that } \forall x \in M \text{ and for any } dw_1 \in C^u_x \text{ and } dw_2 \in C^s_x \text{ one has}$$

$$\sphericalangle(dw_1, dw_2) \geq \alpha$$

The C^u_x *are called the unstable cones whereas* C^s_x *are called the stable ones.*

The property that the angle between stable and unstable cones is uniformly bounded away from zero is called **transversality**.

Some notation and definitions. For any $\delta > 0$ denote by $\mathcal{U}_\delta$ the δ-neighbourhood of the closed set $\Gamma \cup \partial M$. We denote by ρ the Riemannian metric in M and by m the Lebesgue measure (volume) in M. For any submanifold $W \subset M$ we denote by ρ_W the metric on W induced by the Riemannian metric in M, by m_W the Lebesgue measure on W generated by ρ_W and by diam$W$, the diameter of W in the ρ_W metric.

LUM-s. To be able to formulate the further properties to be checked the reader is kindly reminded of the notion of local unstable manifolds. We call a ball-like submanifold $W^u \subset M$ a local unstable manifold (LUM) if (i) $\dim W^u = d_u$, (ii) T^{-n} is defined and

smooth on W^u for all $n \geq 0$, (iii) $\forall x, y \in W^u$ we have $\rho(T^{-n}x, T^{-n}y) \to 0$ exponentially fast as $n \to \infty$.
We denote by $W^u(x)$ (or just $W(x)$) a local unstable manifold containing x. Similarly, local stable manifolds (LSM) are defined.

P2. SRB measure. *The dynamics T has to have an invariant ergodic Sinai-Ruelle-Bowen (SRB) measure μ. That is, there should be an ergodic probability measure μ on M such that for μ-a.e. $x \in M$ a LUM $W(x)$ exists, and the conditional measure on $W(x)$ induced by μ is absolutely continuous with respect to $m_{W(x)}$.*
Furthermore, the SRB-measure should have nice mixing properties: the system (T^n, μ) is ergodic for all finite $n \geq 0$.

In our case the SRB measure is simply the Liouville-measure defined by (5.2) in Section 5.1.1. Absolute continuity of μ is straightforward, while the other above required properties (invariance, ergodicity, mixing) are proved in [17].

P3. Bounded curvature. *The tangent plane of an unstable manifold at the point x should be a Lipschitz function of x. By this we mean that a base can be chosen in every tangent plane so that every base vector is a Lipschitz function of x.*

Some notation. Denote by $J^u(x) = |\det(DT|E^u_x)|$ the Jacobian of the map T restricted to $W(x)$ at x, i.e. the factor of the volume expansion on the LUM $W(x)$ at the point x.

P4. Distortion bounds. *Let x, y be in one connected component of $W \setminus \Gamma^{(n-1)}$, which we denote by V. Then*

$$\log \prod_{i=0}^{n-1} \frac{J^u(T^i x)}{J^u(T^i y)} \leq \varphi\left(\rho_{T^n V}(T^n x, T^n y)\right)$$

where $\varphi(\cdot)$ is some function, independent of W, such that $\varphi(s) \to 0$ as $s \to 0$.

P5. Absolute continuity. *Let W_1, W_2 be two sufficiently small LUM-s, such that any LSM W^s intersects each of W_1 and W_2 in at most one point. Let $W_1' = \{x \in W_1 : W^s(x) \cap W_2 \neq \emptyset\}$. Then we define a map $h : W_1' \to W_2$ by sliding along stable manifolds. This map is often called a holonomy map. This has to be absolutely continuous with respect to the Lebesgue measures m_{W_1} and m_{W_2}, and its Jacobian (at any density point of W_1') should be bounded, i.e.*

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

with some $C' = C'(T) > 0$.

A few words are in order to discuss how our Proposition 5.28 implies property (P5). Let us consider the unique ergodic SRB-measure μ for the dynamical system (in our billiard dynamics this is precisely the Liouville measure defined by (5.2)). We know that the conditional measure on any LUM induced by μ is absolutely continuous with respect to the Lebesgue measure on the unstable manifold. These conditional measures are often referred to as u-SRB measures and their density w.r.t. the Lebesgue measure, $\rho_W(x)$ is given by the following equation (cf. [12]):

$$\frac{\rho_W(x)}{\rho_W(y)} = \lim_{n \to \infty} \prod_{i=1}^{n} \frac{J^u(T^{-i}x)}{J^u(T^{-i}y)}.$$

Actually, what directly follows from Proposition 5.28 is that if we consider two nearby LUM-s W and $\bar{W}$ and points $x, \bar{x}$ on them joint by the holonomy map along an s-manifold, then the ratio of $\rho_W(x)$ and $\rho_{\bar{W}}(\bar{x})$, the densities for the two u-SRB measures is uniformly bounded. However, taking into account the invariance of μ and the uniform contraction along s-manifolds, we may get the uniform bound on the distortion of Lebesgue measures, i.e. the property we assumed in (P5).

Some further notation. Let $\delta_0 > 0$. We call W a $\boldsymbol{\delta_0}$**-LUM** if it is a LUM and $\mathrm{diam} W \leq \delta_0$. For an open subset $V \subset W$ and $x \in V$ denote by $V(x)$ the connected component of V containing the point x. Let $n \geq 0$. We call an open subset $V \subset W$ a (δ_0, n)-subset if $V \cap \Gamma^{(n)} = \emptyset$ (i.e., the map T^n is smoothly defined on V) and $\mathrm{diam} T^n V(x) \leq \delta_0$ for every $x \in V$. Note that $T^n V$ is then a union of δ_0-LUM-s. Define a function $r_{V,n}$ on V by

$$r_{V,n}(x) = \rho_{T^n V(x)}(T^n x, \partial T^n V(x))$$

Note that $r_{V,n}(x)$ is the radius of the largest open ball in $T^n V(x)$ centered at $T^n x$. In particular, $r_{W,0}(x) = \rho_W(x, \partial W)$.

Now we are able to give the last group of technical properties that have to be verified:

P6. Growth of unstable manifolds *Let us assume there is a fixed $\delta_0 > 0$. Furthermore, there exist constants $\alpha_0 \in (0,1)$ and $\beta_0, D_0, \kappa, \sigma, \zeta > 0$ with the following property. For any sufficiently small $\delta > 0$ and any δ_0-LUM W there is an open $(\delta_0, 0)$-subset $V_\delta^0 \subset W \cap \mathcal{U}_\delta$ and an open $(\delta_0, 1)$-subset $V_\delta^1 \subset W \setminus \mathcal{U}_\delta$ (one of these may be empty) such that the two sets are disjoint, $m_W(W \setminus (V_\delta^0 \cup V_\delta^1)) = 0$ and $\forall \varepsilon > 0$*

$$m_W(r_{V_\delta^1,1} < \varepsilon) \leq \alpha_0 \Lambda \cdot m_W(r_{W,0} < \varepsilon / \Lambda) + \varepsilon \beta_0 \delta_0^{-1} m_W(W)$$

$$m_W(r_{V_\delta^0,0} < \varepsilon) \leq D_0 \delta^{-\kappa}\, m_W(r_{W,0} < \varepsilon)$$

and

$$m_W(V_\delta^0) \leq D_0\, m_W(r_{W,0} < \zeta \delta^\sigma)$$

Now we can formulate Theorem 2.1 from [12].

Theorem A.1. ***(Chernov, 1999)*** *Under the conditions **P0 ... P6**, the dynamical system enjoys exponential decay of correlations and the central limit theorem for Hölder-continuous functions.*

The properties stated in the theorem are defined in definitions 1.3 and 1.4.

Bibliography

[1] P. R. Baldwin, *Soft billiard systems*, Physica D **29** (1988), 321–342.

[2] J. Bochnak, M. Coste and M-F. Roy, *Géometrie algébrique réelle*, Springer, 1987

[3] P. Bálint, N. I. Chernov, D. Szász and I. P. Tóth, *Multi-dimensional Semi-dispersing Billiards: Singularities and the Fundamental Theorem*, Annales Henri Poincaré **3** (2002) 451–482.

[4] P. Bálint, N. I. Chernov, D. Szász and I. P. Tóth, *Geometry of Multi-dimensional Dispersing Billiards*, Astérisque **286** (2003) 119–150.

[5] P. Bálint and I. P. Tóth, *Correlation decay in certain soft billiards*, Commun. Math. Phys. **243** (2003) 55–91.

[6] P. Bálint and I. P. Tóth, *Mixing and its rate in 'soft' and 'hard' billiards motivated by the Lorentz process*, Physica D **187** (2004) 128–135.

[7] P. Bálint and I. P. Tóth, *Hyperbolicity of multi-dimensional soft billiards*, submitted to Discrete and Cont. Dyn. Syst.

[8] L. Boltzmann, *Über die Eigenschaften monocyklischer und anderer damit verwandter Systeme*, Kronecker J. = J. Reine Angew. Math. **98** (1884) 68–94, Jahrbuch Fortschritte Math. **16** 863

[9] K. Böröczky, Jr. and G. Tardos, *The longest segment in the complement of a packing*, Mathematika **49** (2002) 45–49.

[10] L. A. Bunimovich and J. Reháček, *How high dimensional stadia look like*, Commun. Math. Phys. **197** (1998) 277–301.

[11] N. Chernov, *Statistical properties of the periodic Lorentz gas. Multidimensional case.* J. Stat. Phys. **74** (1994), 11–53.

[12] N. Chernov, *Decay of correlations and dispersing billiards*, J. Stat. Phys. **94** (1999), 513–556.

[13] N. Chernov and Ya. G. Sinai, *Ergodic Properties of Certain Systems of 2–D Discs and 3–D Balls.*, Russ. Math. Surv. **(3) 42** (1987), 181–201.

[14] N. Chernov and L.-S. Young, *Decay of Correlations for Lorentz gases and hard balls*, 89–120. in *Hard ball systems and the Lorentz gas*, Encyclopedia Math. Sci. **101**, Springer (2000), ed. D. Szász

[15] V. Donnay, *Elliptic islands in generalized Sinai billiards*, Ergodic Theory and Dyn. Syst. **16** (1997) no. 5, 975–1010.

[16] V. Donnay, *Non-ergodicity of two particles interacting via a smooth potential*, J. Stat. Phys. **96** (1999) no. 5-6, 1021–1048.

[17] V. Donnay and C. Liverani, *Potentials on the two-torus for which the Hamiltonian flow is ergodic*, Commun. Math. Phys. **135** (1991), 267–302.

[18] K. Falconer, *The Geometry of Fractal Sets*, Cambridge University Press, 1985

[19] K. Falconer, *Fractal Geometry, Mathematical Foundations and Applications*, John Wiley & Sons, 1990

[20] H. Federer, *Geometric Measure Theory*, Springer, 1969

[21] G. Gallavotti and E.D.G. Cohen, *Dynamical ensembles in stationary states*, J. Stat. Phys. **80** (1995), 931–970.

[22] G. Gallavotti and D. Ornstein, *Billiards and Bernoulli schemes*, Commun. Math. Phys. **38** (1974), 83–101.

[23] E. H. Hauge, *What can we learn from Lorentz models?*, Transport Phenomena, Lecture Notes in Physics, Springer **31** (1974), 377.

[24] A. Knauf, *Ergodic and topological properties of Coulombic periodic potentials*, Commun. Math. Phys. **110** (1987), 89–112.

[25] A. Knauf, *On soft billiard systems*, Physica D **36** (1989), 259–262.

[26] I. Kubo, *Perturbed billiard systems, I.*, Nagoya Math. J. **61** (1976), 1–57.

[27] I. Kubo and H. Murata, *Perturbed billiard systems II, Bernoulli properties*, Nagoya Math. J. **81** (1981), 1–25.

[28] A. Krámli, N. Simányi and D. Szász, *A "Transversal" Fundamental Theorem for Semi-Dispersing Billiards*, Commun. Math. Phys. **129** (1990), 535–560.

[29] C. Liverani and M. P. Wojtkowski, *Ergodicity in Hamiltonian Systems*, Dynamics Reported (New Series) **4** (1995), 130–202.

[30] R. Markarian, *Ergodic properties of plane billiards with symmetric potentials*, Commun. Math. Phys. **145** (1992), 435–446.

[31] V. Rom-Kedar and D. Turaev, *Big islands in dispersing billiard-like potentials*, Physica D **130** (1999) no. 3-4, 187–210.

[32] D. Ruelle, *Smooth dynamics and new theoretical ideas in nonequilibrium statistical mechanics*, J. Stat. Phys. **95** (1999) 393–468.

[33] I. R. Shafarevich, *Basic Algebraic Geometry*, Springer, 1974

[34] N. Simányi, *Proof of the Boltzmann-Sinai Ergodic Hypothesis for Typical Hard Disk Systems*, Inventiones Mathematicae **154(1)** (2003) 123–178. arXiv:math.DS/0008241

[35] N. Simányi, *The Complete Hyperbolicity of Cylindric Billiards*, Ergodic Theory and Dynamical Systems **22** (2002) 281–302. arXiv:math.DS/9906139

[36] N. Simányi and D. Szász, *Hard Ball Systems are Completely Hyperbolic*, Annals of Math. **149** (1999) 35–96.

[37] Ya. G. Sinai, *On the foundations of the ergodic hypothesis for a dynamical system of statistical mechanics*, Dokl. Akad. Nauk SSSR **153** (1963), 1262–1264.

[38] Ya. G. Sinai, *Dynamical systems with elastic reflections. Ergodic properties of dispersing billiards*, Russ. Math. Surv. **25** (1970), 137–189.

[39] I. Stewart, *Galois Theory*, Chapman and Hill, London, 1973

[40] D. Szász, *The K-Property of "Orthogonal" Cylindric Billiards*, Commun. Math. Phys. **160** (1994) 581–597.

[41] D. Szász (ed.), *Hard Ball Systems and the Lorentz Gas*, Encyclopedia of Mathematical Sciences, **101**, Springer, 2000

[42] M. P. Wojtkowski, *Linearly Stable Orbits in 3 Dimensional Billiards*, Commun. Math. Phys. **105** (1986) 391–414.

[43] L.-S. Young, *Statistical properties of dynamical systems with some hyperbolicity*, Annals of Math. **147** (1998), 585–650.

[44] L.-S. Young, *Ergodic theory of chaotic dynamical systems*, XIIth International Congress of Mathematical Physics (ICMP'97) (Brisbane), 131–143, Internat. Press, Cambridge, MA, 1999