Commun. Math. Phys. 146, 357–396 (1992)

Communications in
Mathematical
Physics
© Springer-Verlag 1992

# Ergodic Systems of $n$ Balls in a Billiard Table

**Leonid Bunimovich**[1,2], **Carlangelo Liverani**[3,4], **Alessandro Pellegrinotti**[5] **and Yurii Suhov**[6,7]

[1] Shirshov Institute of Oceanology, Russian Academy of Sciences, Moscow, Russia
[2] Universität Bielefeld, Fakultät für Physik, W-4800 Bielefeld 1, FRG
[3] Mathematics Department, University of Rome II, Rome, Italy
[4] Partially supported by CNR grant n. 203.01.52
[5] Mathematics Department, University of Rome I, Rome, Italy
[6] Institute for Problems of Information Transmission, Russian Academy of Sciences, Moscow, Russia
[7] Statistical Laboratory, DPMMS, University of Cambridge, Cambridge CB2 1SB, England, UK

Received September 12, 1991

**Abstract.** We consider the motion of $n$ balls in billiard tables of a special form and we prove that the resulting dynamical systems are ergodic on a constant energy surface; in fact, they enjoy the $K$-property. These are the first systems of interacting particles proven to be ergodic for an arbitrary number of particles.

## Table of Contents



## 0. Introduction

Consider the motion of $n$ identical balls of radius $R$ in a cube (or, more generally, in an appropriate domain) $Q \subset \mathbb{R}^d$ ($d \geqq 2$) that interact elastically among themselves and with the (piece-wise smooth) boundary $\partial Q$. The Boltzmann hypothesis claims that the restriction of this dynamical system on a manifold of constant energy is ergodic. In fact this hypothesis stimulated the initial development of the notions of ergodic theory itself in the works of L. Boltzmann [B] and J.W. Gibbs [Gi]. The outstanding contribution in the approach to this problem was made by Ya.G. Sinai in his papers, [S1, S2] where some powerful methods were developed that are used now, not only for this, but for many problems of the theory of dynamical systems. The method proposed recently in [SC] is also very important in proving ergodicity for hyperbolic dynamical systems with singularities. Such methods allowed one to prove ergodicity of the system of three and four balls on a torus (i.e. on a cube with identified opposite faces) [KSS1, KSS2]. Unfortunately, new and serious technical problems, which require the development of some specific methods, appear at each step from $n$ to $n+1$ balls. So, the problem of ergodicity for a system of an arbitrary number of elastically interacting balls is still open. The only result, to date, for a system of hard balls are about Lyapunov exponents and the existence of ergodic components of positive measure [SC]. Such results, while falling short of the original goal, are important, not only for the specific problem at hand, but also in the more general context of the ergodic problem for systems of an arbitrary number of interacting particles.

In fact, only one other model is known so far in which it is possible to obtain comparable results. That is, for a system of one dimensional particles falling under gravity and colliding among themselves and with a floor, it is known that the Lyapunov exponents are positive almost everywhere in the phase space [W2, W3] (provided the masses of the particles are not all equal and the lighter particles are above the heavier ones), moreover the ergodicity of the system is proven when only two particles are present [C]. For a review of the systems of many particles for which one can prove that the Lyapunov exponents are non-zero almost everywhere see [W5].

In this paper we solve the above-mentioned problem of ergodicity for a system of $n$ billiard balls, when the balls (we also call them particles) move in boxes of special type. We will see that some of these boxes are generated by a periodic Lorentz gas with a kind of a bounded free path (finite horizon) condition, see [BS, Bu1]. This allows us to introduce a class of models of statistical mechanics that, to our knowledge, was never considered before. These models are intermediate ones between the gas of hard balls and the Lorentz gas model.

We discuss in detail only the two dimensional case (where balls are actually discs moving in a plain domain). It turns out that the higher dimensional cases can be treated in a similar way (in fact, they are much easier); in due time we will outline the changes necessary in higher dimensions.

The paper is organized as follows:

Section 1 discusses the techniques available to tackle the problem of ergodicity for a general system of $n$ balls. We recall some of the literature and we describe in more detail the structure of the argument developed in the following sections. We also present an explicit model to which the rest of the discussion will mainly refer. In Sect. 2 we derive some explicit results on the evolution of the tangent vectors under the flow. These are well known facts [S3, W4, W5] but we present them here to help the reader. Section 3 deals with the Lyapunov exponents. We produce explicit conditions, for the systems of $m < n$ balls, under which the Lyapunov exponents of the $n$-balls system are non-zero almost everywhere. In Sect. 4 we show that our system decomposes in, at most, countably many mod 0 open ergodic components. Section 5 is devoted to the proof that the system has only one ergodic component. In Sect. 6 we discuss other models, in particular a periodic Lorentz gas, to which our strategy can be applied. Finally, there are two technical appendices.

Appendix I deals with the transversality of some manifolds. Appendix II reminds the reader of the construction of the Poincaré section and of how to translate our results for the Poincaré map to results for the billiard flow.

## 1. General Facts and a Model

As already mentioned in the introduction, the contributions of many different people have crystallized, through the years, into a standard strategy to deal with the problem of ergodicity for systems of elastically interacting balls. The argument developed follows, ideally, the path outlined by Hopf [H], but addresses in particular two difficulties typical of these systems. The first is the lack of uniform hyperbolicity (e.g. for trajectories through which the balls never collide among themselves). The second is a violation of smoothness of the dynamics. Even if we consider the flow generated by the dynamics at times when no collisions occur there are discontinuities, in the derivatives, at points that experience, in the time interval under consideration, a tangent or a multiple collision. More precisely, a trajectory may experience a tangence of a ball to the boundary of the billiard table or grazing of several balls. Besides, multiple collisions may occur, when two or more balls collide with each other or with the boundary at the same time. We call such events "singularities" or qualify them as "non-smoothness" in the behavior of a given trajectory.

The initial step, in the above-mentioned strategy, is to show that none of the Lyapunov exponents is zero [Kr, S1]. A typical technique, in this context, is to find a cone field, in the tangent bundle, which is eventually strictly invariant for the dynamics [W1].

Let $\phi^t$ be the billiard flow (for $n$ balls in two dimensions) on the phase space $\tilde{\mathcal{M}} \subset \mathbb{R}^{2n} \times \mathbb{R}^{2n}$. We call $\mathcal{M}$ a submanifold of $\tilde{\mathcal{M}}$ with constant (kinetic) energy $E = \frac{1}{2}\langle p, p\rangle$; owing to the law of the conservation of energy, $\phi^t \mathcal{M} = \mathcal{M}$. For that reason we also call sometimes $\mathcal{M}$ the phase space of the system under consideration. Given $x \in \mathcal{M}$, we introduce in the tangent space $\mathcal{T}_x\mathcal{M}$ the basis induced by the coordinates $q$ and $p$, so that a tangent vector will have coordinates $\xi = (\delta q, \delta p)$. The free flow direction is given by $(p, 0)$; clearly this tangent vector is preserved by the dynamics. More importantly, the dynamics preserves the property of being perpendicular to the flow direction (see Lemma 2.3). Since to the flow direction corresponds a zero Lyapunov exponent, it is natural to define our cone at $x$ in the subspace of $\mathcal{T}_x\mathcal{M}$ perpendicular to the flow. An invariant cone field is then a collection of cones $C(x) \subset \hat{\mathcal{T}}_x\mathcal{M} = \{(\delta q, \delta p) \in \mathcal{T}_x\tilde{\mathcal{M}} \mid \langle p, \delta p\rangle = 0;\ \langle(\delta q, \delta p),\ (p, 0)\rangle = 0\}$ such that $d\phi^t\, C(x) \subseteqq C(\phi^t x)\ \forall t \geqq 0$. Notice that an element of $\hat{\mathcal{T}}_x\mathcal{M}$ can be interpreted as an equivalence classs of curves (also called variations) $\gamma$: $[-\varepsilon, \varepsilon] \to \mathcal{M}$, $\gamma(0) = x$. Any such curve can be pictured as a collection of $n$ curves in the two dimensional space describing variation of positions for $n$ particles in the system, each curve endowed with its own family of velocity vectors; see Fig. 1.

For the systems under consideration here, the first results on the Lyapunov exponents where found in the classical Sinai papers [S1, S3], where he used the language of continuous fractions. Yet, we obtain Sinai's results by using an invariant cone family introduced by Wojtkowski in [W4, W5]. Accordingly, we define $C(x) = \{(\delta q, \delta p \in \hat{\mathcal{T}}_x\mathcal{M} \mid \langle \delta q, \delta p\rangle \geqq 0\}$ (and $C_-(x)$

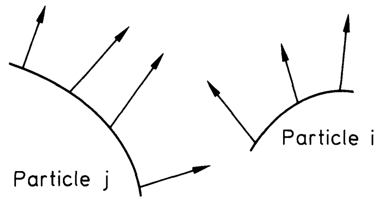


**Fig. 1.** Representation of variations

$= \{(\delta q, \delta p) \in \hat{\mathscr{T}}_x\mathscr{M} \mid \langle \delta q, \delta p\rangle \leqq 0\}$ for the backward dynamics); for a single ball on a table it is quite easy to see that the preceding cone corresponds to a family of diverging trajectories (note that the vector spaces $(\delta q, B\delta q)$, where $B \geqq 0$, belongs to $C(x)$).

Between two collisions a tangent vector $\xi = (\delta q, \delta p)$ evolves according to the following equations:

$$d\phi^t \xi = (\delta q + t\delta p, \delta p) \tag{1.1}$$

(we assume that all the particles have mass 1). It is therefore clear that, if a family of trajectories is divergent ($\langle \delta q, \delta p\rangle \geqq 0$), the free dynamics preserves such a property. A more complex computation shows that the same is true when a collision, between two particles or with the boundary, occurs (see Sect. 2).

What we need in order to have the desired cone property is the "strict" invariance. This means that any family of trajectories on the boundary of the cone ($\langle \delta q, \delta p\rangle = 0$) will be, after some time, strictly divergent, i.e. strictly contained in the cone ($\langle \delta q, \delta p\rangle > 0$). In fact a theorem from [W1], applied to our situation, states that, if we have an invariant measurable cone family such that almost every cone has an image strictly contained in the cone at the image point, then the Lyapunov exponents are different from zero almost everywhere.

As noted already in [KSS4], the above-mentioned property is determined by a finite piece of trajectory; in the future we use the symbol $(x, [\tau_1, \tau_2])$, where $x \in \mathscr{M}$ and $\tau_1, \tau_2 \in \mathbb{R}$, $\tau_1 < \tau_2$, to designate the piece of trajectory $\{\phi^t(x)\}_{t \in [\tau_1, \tau_2]}$.

**Definition 1.1.** *A piece of trajectory* $(x, [\tau_1, \tau_2])$ *is called sufficient for the vector* $(\delta q, \delta p) \neq 0$, *where* $(\delta q, \delta p) \in \partial C(\phi^{\tau_1}(x))$ $(\langle \delta q, \delta p\rangle = 0)$, iff $d\phi^{\tau_2 - \tau_1}(\delta q, \delta p) \in \operatorname{int} C(\phi^{\tau_2}(x))$.

*Moreover,* $(x, [\tau_1, \tau_2])$ *is called sufficient* iff *it is sufficient for each non-zero vector in* $\partial C(\phi^{\tau_1}(x))$.

*Finally, a point* $x$ *is called sufficient forward (backward)* iff *there exists* $\tau \in \mathbb{R}^+$ ($\tau \in \mathbb{R}^-$) *such that* $(x, [0, \tau])$ $((x, [\tau, 0]))$ *is sufficient.*

To clarify the relations between forward and backward sufficiency, that is sufficiency for $\phi^t$ versus sufficiency for $\phi^{-t}$, see Lemma 4.1.

Consequently, to obtain all the Lyapunov exponents different from zero, it is enough to show that almost all the points are sufficient. Nevertheless, in order to prove ergodicity, it is necessary to have a more detailed knowledge of the set of non-sufficient points.

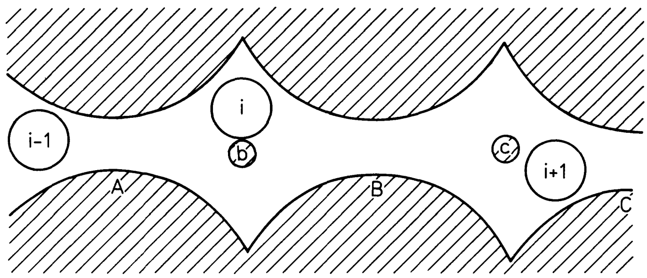


**Fig. 2.** Billiard table containing $n$ balls

For our systems (which are a particular case of a semi-dispersing billiard) the so-called "Transversal Fundamental Theorem for Semi-Dispersing Billiards" enables us to prove that, given a sufficient point with a smooth history, there exists a neighborhood of the point that belongs mod 0 to a single ergodic component. In other words, only one ergodic component has an intersection of positive measure with each sufficiently small neighborhood of the point. Moreover, such a component is open, apart from a set of zero measure. On the other hand, the Fundamental Theorem requires that almost all points (with respect to the induced measure) on the singularity manifolds of the Poincaré map and of its inverse are sufficient [SC, KSS3, LW]. This is called the Sinai-Chernov Ansatz. We check in Sect. 5 that such a condition is satisfied for our examples. In addition, some knowledge of the structure of the singularity manifolds is required; we discuss this in the first part of Appendix I.

To prove global ergodicity it is necessary to show that the set of non-sufficient points does not separate the phase space, and, moreover, it is well related to the structure of the singularities (the Sinai-Chernov Ansatz again). Our proof of these facts, inspired by [SC] and [KSS1], is carried out in Sect. 5.

Next, we introduce an example for which it is possible to establish the stated results, so that the reader has something concrete to refer to in the next sections of the paper (the general class of examples to which our techniques apply is discussed in Sect. 6).

*A Special Billiard Table.* Figure 2 shows the billiard we suggest for consideration. The boundary $\partial Q$ of the billiard table $Q$ is strictly inwardly-convex. There is one ball in each "cell" between consecutive "bottlenecks" A, B, C, ... The balls are of radius $R$ and mass 1; they cannot cross the bottlenecks, but can collide with their neighbors and move in their cells around corresponding obstacles a, b, c, ... The obstacles make it impossible for a ball to have a collision with its neighbor on the left followed by a collision with the one on the right, or vice versa, without having a collision with the boundary. Some of the obstacles may not be necessary in the case of a billiard table having dimension $d$ greater than 2: it is generally sufficient to have an obstacle every $d-1$ consecutive balls.

We call a system of $n$ balls the dynamical system generated by the motion of $n$ adjacent balls in the above-described region with the invariant measure on $\mathscr{M}$ induced by the Lebesgue volume in $\mathbb{R}^{2n} \times \mathbb{R}^{2n}$ (this measure is also called the Lebesgue measure on $\mathscr{M}$).

We conclude this section by stating explicitly the theorem proved in this paper.

**Main Theorem.** *The dynamical system generated by the motion of any number n of adjacent balls in the billiard table desribed above (or in the ones described in Sect. 6) is a K-flow on each connected component of a constant energy manifold (with strictly positive energy).*

## 2. Cone Fields and Dynamics

Here we examine the behavior of the cone field with respect to the dynamics. To study the evolution of the tangent vectors, it is necessary to use the explicit form of the laws of reflection.

*(I) Reflection by the Boundary $\partial Q$.* Let us study the collision of particle $k$ with the boundary. Calling $\eta$ the unit vector inwardly normal to the boundary, at the point of collision, and $v$ the unit tangent vector, we have

$$\begin{aligned} q_k' &= q_k, \\ p_k' &= \langle p_k, v(x)\rangle\, v(x) - \langle p_k, \eta(x)\rangle\, \eta(x) = p_k - 2\langle p_k, \eta(x)\rangle\, \eta(x), \end{aligned} \tag{2.1}$$

where $(q_k, p_k)$ and $(q_k', p_k')$ are the coordinates of the particle $k$ before and after the collision, respectively. (Similar notations are used below for all types of collisions.) If we consider a variation of trajectory of our system corresponding to the tangent vector $(\delta q, \delta p)$ at the collision point and suppose that $\delta q$ is parallel to $v(x)$ (in which case all the trajectories experience a collision of the particle $k$ with $\partial Q$ at the same time), we have:

$$\begin{aligned} \delta q_k' &= \delta q_k, \\ \delta p_k' &= \delta p_k - 2\langle \delta p_k, \eta\rangle \eta - 2\langle p_k, \delta q_k\rangle K\eta - 2\langle p_k, \eta\rangle K\delta q_k, \end{aligned} \tag{2.2}$$

where $K > 0$ is the (inward) curvature of the boundary at the collision point. Using (2.2) yields

$$\langle \delta q_k', \delta p_k'\rangle = \langle \delta q_k, \delta p_k\rangle - 2\langle p_k, \eta\rangle \langle \delta q_k, K\delta q_k\rangle \geqq \langle \delta q_k, \delta p_k\rangle \tag{2.3}$$

for $\langle p_k, \eta\rangle \leqq 0$, in order for the collision to happen. Hence, we have the following lemma.

**Lemma 2.1.** *A collision of particle k with the boundary $\partial Q$ is sufficient for $(\delta q, 0)$ ($\langle \delta q, \delta p\rangle$ increases strictly) unless*

(i) $$\langle p_k, \eta\rangle = 0 \quad \textit{(tangent collision)}$$

*or*

(ii) $$\delta q_k = \lambda p_k \quad \textit{for some } \lambda \in \mathbb{R}.$$

*Proof.* Notice that (2.3) was derived under the assumption that $\delta q_k$ was parallel to $\partial Q$. This can be achieved by letting different trajectories flow different amounts of times. To be more accurate, we define $\delta \tilde{q} = \delta q + v p$ – observe that $\langle \delta \tilde{q}, \delta p\rangle = \langle \delta q, \delta p\rangle$ (since the conservation of the energy $E = \frac{1}{2}\langle p, p\rangle$ implies

$\langle \delta p, p \rangle = 0$) – and we choose $\nu$ such that $\langle \eta, \delta \tilde{q}_k \rangle = 0$. We can then apply (2.3) to the variation $(\delta \tilde{q}, \delta p)$. The previous condition reads

$$-\langle \delta q_k, \eta \rangle = \nu \langle \eta, p_k \rangle$$

and can be satisfied only if (i) does not hold. Thus, we will have strict increase provided that $\delta \tilde{q} \neq 0$, or $\delta q_k \neq -\nu p_k$ (remember that $\langle \eta, p_k \rangle < 0$). Observe that after the collision $\delta q'$ is given by $\delta q' = \delta \tilde{q} - \nu p'$. □

*(II) Collision Between Particles k and* $k+1$. In this case, let $\tilde{\eta} = q_{k+1} - q_k$ be the vector that joins the center of $k$ with the center of $k+1$. If we set $\eta = \tilde{\eta} / \| \tilde{\eta} \|$ and we call $v$ the perpendicular to $\eta$, then

$$\begin{aligned} q_k' &= q_k, \\ q_{k+1}' &= q_{k+1}, \\ p_k' &= \langle p_k, v \rangle v + \langle p_{k+1}, \eta \rangle \eta = p_k + \langle p_{k+1} - p_k, \eta \rangle \eta, \\ p_{k+1}' &= \langle p_{k+1}, v \rangle v + \langle p_k, \eta \rangle \eta = p_{k+1} - \langle p_{k+1} - p_k, \eta \rangle \eta. \end{aligned} \tag{2.4}$$

A variation $(\delta q, \delta p)$ corresponds to trajectories that collide at the same time if

$$\| (q_{k+1} + \delta q_{k+1}) - (q_k + \delta q_k) \| = 2R + \mathcal{O}(\delta q^2), \tag{2.5}$$

where $R$ is the radius of the balls; consequently, (2.5) yields

$$\langle \delta q_{k+1} - \delta q_k, \eta \rangle = 0. \tag{2.6}$$

For a variation satisfying (2.6), differentiating (2.4) gives

$$\begin{aligned} \delta q_k' &= \delta q_k, \\ \delta q_{k+1}' &= \delta q_{k+1}, \\ \delta p_k' &= \delta p_k + \pi_k, \\ \delta p_{k+1}' &= \delta p_{k+1} - \pi_k, \\ \pi_k &= \langle \delta p_{k+1} - \delta p_k, \eta \rangle \eta + (2R)^{-1} \langle p_{k+1} - p_k, \delta q_{k+1} - \delta q_k \rangle \eta \\ &\quad + (2R)^{-1} \langle p_{k+1} - p_k, \eta \rangle (\delta q_{k+1} - \delta q_k). \end{aligned} \tag{2.7}$$

**Lemma 2.2.** *A collision involving particles k and* $k+1$ *is sufficient for* $(\delta q, 0)$ *unless*

(i) $\quad \langle p_{k+1} - p_k, \eta \rangle = 0 \quad$ *(tangent collision)*

*or*

(ii) $\quad \delta q_k + \lambda p_k = \delta q_{k+1} + \lambda p_{k+1} \quad$ *for some* $\lambda \in \mathbb{R}$.

*Proof.* Similarly to the preceding case, we first modify our variation and then apply (2.7). Consider $\delta \tilde{q} = \delta q + \nu p$, then $\delta \tilde{q}$ satisfies (2.6) if

$$-\langle \delta q_{k+1} - \delta q_k, \eta \rangle = \nu \langle \eta, p_{k+1} - p_k \rangle.$$

The previous condition can always be fulfilled, provided (i) is false. Next, using (2.7), we obtain

$$\begin{aligned} &\langle \delta \tilde{q}_k', \delta p_k' \rangle + \langle \delta \tilde{q}_{k+1}', \delta p_{k+1}' \rangle \\ &\quad = \langle \delta \tilde{q}_k, \delta p_k \rangle + \langle \delta \tilde{q}_{k+1}, \delta p_{k+1} \rangle - \langle \pi_k, \delta \tilde{q}_{k+1} - \delta \tilde{q}_k \rangle \\ &\quad = \langle \delta \tilde{q}_k, \delta p_k \rangle + \langle \delta \tilde{q}_{k+1}, \delta p_{k+1} \rangle - \frac{1}{2R} \langle p_{k+1} - p_k, \eta \rangle \, \| \delta \tilde{q}_{k+1} - \delta \tilde{q}_k \|^2. \end{aligned}$$

Since $\langle p_{k+1} - p_k, \eta \rangle$ has to be negative, for the collision to happen, the only contingency in which $\langle \delta q, \delta p \rangle$ does not increase is $\delta \tilde{q}_{k+1} = \delta \tilde{q}_k$. □

**Lemma 2.3.** *The collisions preserve the properties:*

(1) $\langle \xi, (0, p) \rangle = 0$ *(conservation of energy)*,
(2) $\langle \xi, (p, 0) \rangle = 0$ *(geodesic-like)*.

*Moreover, the forward dynamics increases the configuration norm* $\| \xi \|_q = \| \delta q \|$ *for each* $\xi = (\delta q, \delta p) \in C$, *while the backward dynamic increases the configuration norm of each vector* $\xi \in C_-$.

*Proof.* Lemma 2.3 follows by direct computation, using (2.2), (2.7). □

## 3. Sufficiency and Lyapunov Exponents

The aim of this section is to find a large set of sufficient points (see Definition 1.1). We consider explicitly only the sufficiency for the forward trajectory; the discussion of the backward sufficiency is completely analogous (with the only proviso that the cone is now given by $\langle \delta q, \delta p \rangle \leqq 0$).

Let us start by considering a tangent vector $(\delta q, \delta p)$ such that $\langle \delta q, \delta p \rangle = 0$ and with $\delta p_k \neq 0$ *for some* $k$; then, given (1.1), it will follow, for $t > 0$, that $\langle \delta q, \delta p \rangle > 0$ at time $t$. Hence, sufficiency will be verified, with respect to the given vector, after any arbitrarily small time. Consequently, we need to discuss only the variations of the form $(\delta q, 0)$.

Before going any further we need a few definitions:

**Definition 3.1.**

$\mathscr{R} = \{(q, p) \in \mathscr{M} \mid$ *the balls experience a multiple or tangent collision*$\}$,
$\mathscr{M}_0 = \{(q, p) \in \mathscr{M} \mid$ *the forward trajectory never intersects* $\mathscr{R}\}$,
$C_i = \{(q, p) \in \mathscr{M}_0 \mid$ *the next collision will be* $i$ *with* $i + 1$ *(non-tangent)*$\}$,
$\tilde{C}_i = \{(q, p) \in \mathscr{M}_0 \mid$ *the next collision will be* $i$ *with* $\partial Q$ *(non-tangent)*$\}$,
$\Delta_i = \{x = (q, p) \in \mathscr{M}_0 \mid$ *there exists* $t^* \in \mathbb{R}^+$ *such that* $i$, $i + 1$ *never collide, for* $t > t^*\}$,

$$\tilde{\Delta} = \bigcup_{i=1}^{n-1} \Delta_i,$$

$$\Omega = \mathscr{M}_0 \backslash \tilde{\Delta},$$

$\sum_i^1 = \{x \in C_i \mid$ *the vectors* $p_i$ *and* $p_{i+1}$ *are parallel*$\}$,

$\sum_i^2 = \{x \in \tilde{C}_i \mid$ *the vector* $p_i$ *is parallel to* $p_i$ *after the next collision*$\}$,

$\sum_i^3 = \{x \in C_i \mid$ *the vector* $p_{i+1}$ *is parallel to* $p_i$ *after the next collision*$\}$,

$\sum_i^4 = \{x \in C_i \mid$ *the vector* $p_i$ *is parallel to* $p_i$ *after the next collision*$\}$.

The set $\Omega$ is roughly the set of points whose trajectories, after any given time, contain all the possible types of collisions. With the next theorem, we start to heavily use the concept of sub-sets of (Lebesgue) measure zero and codimension two in the phase space $\mathcal{M}$; a few clarifications are called for. The relevance of these sets, for the problem of ergodicity, was first noticed in [S3], and is due to the fact that they cannot separate the phase space [E]. In Sects. 3, 4, by codimension two we will mean "codimension two in the Poincaré section," which really means codimension three in $\mathcal{M}$. In fact, in Appendix I, we discuss manifolds with the property of being perpendicular to the flow direction. For these manifolds we talk sometimes of codimension three since the direction of the flow is counted. This may be confusing, but it is convenient, since the theorems to which we refer [SC, KSS3, LW] are stated for maps (here the Poincaré map from collision to collision: see Appendix II) while many properties are most readily checked for the flow.

**Theorem 3.2.** *$\Omega$ is a set of sufficient points, apart from a sub-set of measure zero and codimension two.*

*Proof.* Unfortunately, the proof is an analysis of many different cases. First of all, for a particle to have a velocity zero is a codimension two condition. We can therefore restrict ourselves to the situation in which all the particles have, and keep, a non-zero velocity. We then start our analysis by concentrating on two generic particles $i$ and $i+1$.

**Lemma 3.3.** *If $x \in \Omega$ and $d\phi^t(\delta q, 0) = (\delta q(t), 0)\ \forall t > 0$, then either $\delta q_i(t) = \lambda p_i(t)$ and $\delta q_{i+1}(t) = \lambda p_{i+1}(t)$ for some $\lambda \in \mathbb{R}$, $t \in \mathbb{R}^+$, or $x$ belongs to a set of measure zero and codimension two.*

*Proof.* We recall that $\Omega \subset \mathcal{M}_0$, this means that the trajectory does not experience a singular collision; in particular it does not experience any tangent collision. We can then use Lemmas 2.1, 2.2 to characterize the vectors that are not sufficient after a give collision.

Among all the collisions occurring in the system, we distinguish and call relevant, the following: $i$ with the boundary $\partial Q$, $i+1$ with $\partial Q$ and $i$ with $i+1$.

*Case A. The First Relevant Collisions are $i$ with $\partial Q$ and $i+1$ with $\partial Q$.* In this case let $t_1 \in \mathbb{R}^+$ be the first time the collision $i$ with $i+1$ occurs ($t_1$ exists because of the definition of $\Omega$). Owing to the geometry of the billiard table $Q$ and the hypotheses under consideration, the preceding collisions involving $i$, $i+1$ must have been with the boundary $\partial Q$. According to Lemma 2.1, a variation is non-sufficient, after such collisions, only if it is of the form $\delta q_i(t_1^-) = \lambda p_i(t_1^-)$, $\delta q_{i+1}(t_1^-) = \lambda' p_{i+1}(t_1^-)$, for some $\lambda$, $\lambda' \in \mathbb{R}$ (we have used the relations that describe the evolution in time of the variations) where by $t_1^-$ we mean the instant before the collision. Moreover, Lemma 2.2 forces

$$\lambda p_i(t_1^-) + \nu p_i(t_1^-) = \lambda' p_{i+1}(t_1^-) + \nu p_{i+1}(t_1^-). \tag{3.1}$$

If $p_i(t_1^-)$ is not parallel to $p_{i+1}(t_1^-)$, then the only solution of (3.1) is $\lambda + \nu = 0 = \lambda' + \nu$ which implies $\lambda = \lambda'$ and the lemma. If $p_i(t_1^-)$ is parallel to $p_{i+1}(t_1^-)$, then a direct computation yields

$$\begin{aligned}\delta q_i(t_1^+) &= \delta q_i(t_1^-) + \nu(p_i(t_1^-) - p_i(t_1^+)) = (\lambda + \nu)\, p_i(t_1^-) - \nu p_i(t_1^+),\\ \delta q_{i+1}(t_1^+) &= (\lambda + \nu)\, p_i(t_1^-) - \nu p_{i+1}(t_1^+).\end{aligned}$$

So far we have established the desired result apart from points belonging to the manifold $\phi^{-t_1}(\sum_i^1)$. Note that in general, $\sum_i^1$ has codimension $d-1$, where $d$ is the dimension of the billiard table $Q$; thus, in the higher dimensional cases, the result is already established, apart from a set of measure zero and codimension two. However, in the two dimensional case (the one at hand) $\sum_i^1$ has only codimension one and, therefore, we need further analysis. Such simplifications, in higher dimension, also occur in all the cases discussed later. In order not to interrupt the exposition, we will no longer comment on them explicitly.

For points in $\sum_i^1$, due to the geometry of the table, either $i$ or $i+1$ collide with $\partial Q$ after the time $t_1$. We consider the first eventuality; the second one can be treated in a similar fashion.

In order not to be sufficient, after the collision of $i$ with the boundary, it is necessary that $\delta q_i(t_1^+) = \sigma p_i(t_1^+)$; this forces $\lambda + \nu = 0 = \sigma + \nu$, unless $p_i(t_1^-)$ is parallel to $p_i(t_i^+)$ ($x$ belongs to a pre-image of $\sum_i^2$). Given that the points, for which $p_i(t_1^-)$, $p_{i+1}(t_1^-)$ and $p_i(t_1^+)$ are simultaneously parallel, form a manifold of codimension two (see Appendix I), we can dismiss them. Therefore, after this last collision, we have

$$\begin{aligned}\delta q_i(t) &= \lambda p_i(t),\\ \delta q_{i+1}(t) &= \lambda p_{i+1}(t)\end{aligned}$$

as desired.

*Case B. The First Relevant Collisions are: $i+1$ with $\partial Q$ Followed by $i$ with $i+1$.* The simplest possibility consists of the following sequence of collisions: $i+1$ with $\partial Q$, $i$ with $i+1$, $i$ with $\partial Q$ and $i+1$ with $\partial Q$. It follows, from an analysis similar to the preceding one, that the expected result is determined outside the following codimension two manifold: $p_{i+1}(t_1^-)$, $p_{i+1}(t_1^+)$ and $p_i(t_1^+)$ are simultaneously parallel ($t_1$ is the time at which the collision $i$ and $i+1$ takes place). Next, suppose that one of the two particles (e.g. $i+1$) will not hit the boundary before the next $i$, $i+1$ collision occurs (notice that, since the particle has non-zero velocity, it will definitely collide with $\partial Q$, if nothing else happens). Accordingly, we consider the following sequence of collisions: $i+1$ with $\partial Q$, $i$ with $i+1$ and $i$ with $\partial Q$. Such a trajectory has the desired property unless $x \in \phi^{-t_1}(\sum_i^3)$ ($t_1$ is, again, the time of the collision $i$, $i+1$). If $x$ belongs to the previous codimension one manifold, then, after the above sequence of collisions, we have

$$\begin{aligned}\delta q_i(t) &= \lambda p_i(t),\\ \delta q_{i+1}(t) &= (\lambda' + \nu)\, p_{i+1}(t_1^-) - \nu p_{i+1}(t_1^+).\end{aligned} \tag{3.2}$$

A crucial point, for the success of our discussion, is the possibility of controlling the evolution of such variations until the next $i$, $i+1$ collision.

**Sub-Lemma 3.4.** *If $\phi^{t_1^+}(x)$ does not belong to a pre-image of the manifolds $\sum_{i-1}^3$, $\sum_i^4$ or to a set of measure zero and codimension two, then $\delta q_i$ will still satisfy (3.2) before the next $i$, $i+1$ collision.*

*Proof.* If only collisions with the boundary are involved, then it is clear that $\delta q_i(t) = \lambda p_i(t)$. Let us see what happens when $i$ collides with $i-1$. If such a collision takes place at time $t_2$, we have

$$\delta q_{i-1}(t_2^+) = (\lambda + \alpha)\, p_i(t_2^-) - \alpha q_{i-1}(t_2^+),$$
$$\delta q_i(t_2^+) = (\lambda + \alpha)\, p_i(t_2^-) - \alpha q_i(t_2^+). \tag{3.3}$$

Next, one of the two particles must collide with $\partial Q$, and the Sub-Lemma follows. □

The usefulness of the previous Sub-Lemma is emphasized by the following

**Sub-Lemma 3.5.** *The pre-image of the manifolds $\sum_j^k$ intersect transversally the manifolds $\sum_m^l$ for each $j$, $k$, $l$, $m$.*

*Proof.* See Appendix I. □

According to the previous discussion, when the next $i$, $i+1$ collision takes place, (3.2) will still hold, again out of a set of measure zero and codimension two. It follows, using Sub-Lemma 3.5, that this last collision ensures $\delta q_i = \lambda p_i$, $\delta q_{i+1} = \lambda p_{i+1}$, out of a set of measure zero and codimension two.

*Case C. The First Relevant Collision is $i$ with $i+1$.* In this case the definition of $\Omega$ implies that another $i$, $i+1$ collision takes place. We can then skip the first collision and apply the preceding arguments to the subsequent trajectory. This concludes the proof of Lemma 3.3. □

Let us review our situation. Lemma 3.3 implies that, given $x \in \Omega$ and a pair of neighboring particles $i$ and $i+1$, a time $t_i$ and a number $\lambda_i$ exist such that $\delta q_i(t_i) = \lambda_i p_i(t_i)$, $\delta q_{i+1}(t_i) = \lambda_i p_{i+1}(t_i)$. Moreover, Sub-Lemma 3.4 tells us that, for $t > t_i$, such a situation can change, for $i+1$ or $i$, only if $x$ belongs to the pre-image of one of the codimension one manifolds $\sum_{i-1}^k$, $\sum_i^k$, $\sum_{i+1}^k$, $\sum_{i+2}^k$. Also Sub-Lemma 3.5 ensures that only one such instance can occur, since all the previous mentioned manifolds intersect transversally (so that the set of their intersections has measure zero and codimension two). Then, after some time $t^*$, the worst possible situation will be

$$\delta q_i(t^*) = \lambda_1 p_i(t^*) \qquad i \leqq k,$$
$$\delta q_i(t^*) = \lambda_2 p_i(t^*) \qquad i > k$$

for some $k \in \{1, \ldots, n-1\}$. In addition, any collision between $k$ and $k+1$ would force $\lambda_1 = \lambda_2$ (always ignoring a set of measure zero and codimension two). Since the definition of $\Omega$ ensures that such a collision happens, we have $\delta q(t) = \lambda p(t)$, for $t$ large enough. Finally, remembering that $(\delta q, 0)$ is perpendicular to the flow direction, it follows that $\lambda = 0$. Hence all the variations are sufficient. □

A question remains about the size of the set $\Omega$. To this end we have the following lemma.

**Lemma 3.6.** *If, for any $m < n$, the dynamics generated by m balls is mixing, then the set $\mathcal{M}\backslash\Omega$ is of measure zero for the system of n balls.*

*Proof.* Looking back at Definition 3.1, we see that it is sufficient to show that the set $\tilde{\Delta}_0 = \{x \in \mathcal{M}_0 \mid i \text{ never collides with } i+1 \text{ for all } t \geqq 0 \text{ and for some } i \in \{1, \ldots, n\}\}$, has measure zero. In fact, the set $\mathcal{M}\backslash\mathcal{M}_0$ is of measure zero since it is composed of countably many codimension one manifolds. In addition, given $\tau \in \mathbb{R}^+$, we have $\phi^\tau(\tilde{\Delta}_0) \subset \tilde{\Delta}_0$ and $\phi^\tau(\tilde{\Delta}) = \tilde{\Delta}$. Since for each $x \in \tilde{\Delta}$ a $\tau \in \mathbb{R}^+$ exists such that $\phi^\tau(x) \in \tilde{\Delta}_0$, this implies that the sets $\tilde{\Delta}_0$ and $\tilde{\Delta}$ have the same measure.

Let $\Gamma_1$, $\Gamma_2$ be two sets in $Q \times \mathbb{R}^2$ such that, if $(q_i, p_i) \in \Gamma_1$ and $(q_{i+1}, p_{i+1}) \in \Gamma_2$, the next collision involving $i$ or $i+1$ is a collision between the two balls. Clearly it is possible to choose the sets $\Gamma_1$ and $\Gamma_2$ so that they have a positive Lebesgue volume. Consequently, the sets $\tilde{\Gamma}_2(\tilde{E}) = \left\{(q, p) \in \mathcal{M}_0 \mid (q_{i+1}, p_{i+1}) \in \Gamma_2, \sum_{j=i+1}^{n} \|p_j\|^2 = \tilde{E}\right\}$, for the value of $\tilde{E}$ in some interval, are of strictly positive measure. (We have in mind here the measure $\tilde{\mu}_{\tilde{E}}^{(n-i)}$ that is the projection of the invariant measure on $\mathcal{M}$ to the manifold $\mathcal{M}_{\tilde{E}}^{(n-i)}$ where the collection of the last $n - i$ particles has total energy $\tilde{E}$). It is then a consequence of the ergodicity of the sub-dynamics generated by the first $i$ particles and of Fubini's Theorem that, for almost all $x \in \mathcal{M}_0$, there exists a monotonically increasing sequence $\{t_k(x)\}$ with $\lim_{k\to\infty} t_k(x) = +\infty$ such that $(\bar{q}_i(t_k(x)), \bar{p}_i(t_k(x))) \in \Gamma_1$, where $(\bar{q}_i(t), \bar{p}_i(t))$ is determined by the dynamics generated by the first $i$ particles only (the system obtained by erasing the last $n - i$ particles). Calling $\chi_{\tilde{\Gamma}_2(\tilde{E})}$ the characteristic function of $\tilde{\Gamma}_2(\tilde{E})$, it follows that, for almost all $x \in \tilde{\Delta}_0$, $\chi_{\tilde{\Gamma}_2(\tilde{E}(x))}(\phi^{t_k(x)}(x)) = 0$ for all $k > 0$. ($\tilde{E}(x)$ denotes here the total energy of the last $n - i$ balls in a point $x$ of the phase space $\mathcal{M}$.) Notice that, for such $x$, the dynamics is the product of two independent dynamics: the one generated by the first $i$ balls and the one generated by the remaining $n - i$ balls. The result is then a consequence of the mixing property of the dynamics generated by the last $n - i$ balls. In fact, if $\chi_{\tilde{\Delta}_0}$ is the characteristic function of $\tilde{\Delta}_0$, we establish that $\tilde{\Delta}_0$ is of measure zero thanks to the relation

$$0 = \lim_{k\to\infty} \int_{\mathcal{M}} \chi_{\tilde{\Gamma}_2(\tilde{E})}(\phi^{t_k(x)}(x))\, \chi_{\tilde{\Delta}_0}(x)\, d\tilde{\mu}_{\tilde{E}}^{(n-i)}(x)\, d\tilde{E}$$

$$= \int \tilde{\mu}_{\tilde{E}}^{(n-i)}(\tilde{\Gamma}_2(\tilde{E}))\, \tilde{\mu}_{\tilde{E}}^{(n-i)}(\tilde{\Delta}_0 \cap \mathcal{M}_{\tilde{E}}^{(n-i)})\, d\tilde{E}$$

for the right-hand side of this relation implies $\tilde{\mu}_{\tilde{E}}^{(n-i)}(\tilde{\Delta}_0 \cap \mathcal{M}_{\tilde{E}}^{(n-i)}) = 0$ for each $\tilde{E}$, $i$. □

*Remark 3.7.* Lemma 3.6 is a version of the "weak lemma on avoiding balls" (see [KSS1]). Its proof is based, in essence, on the following fact: in an ergodic system the set of points that avoids a region of positive measure has zero measure. In the following we will need a slight generalization of Lemma 3.6. More precisely, it is convenient to introduce the following model: a system of $n$ balls divided into two sub-systems; the first consisting of the first $i$ balls and the second by the last $n-i$. These two sub-systems are almost independent: we allow the ball $i$ to collide with the ball $i+1$ only if the distance of the centers of the two balls would became smaller than $2R-\delta$ ($R$ being the radius of the balls and $\delta < R$ is a given positive number) under the evolution of the two independent sub-systems (meaning that the balls $i$ and $i+1$ can freely cross each other). In this situation we quote Lemma 3.6 and this Remark to claim that the set of points for which $i$ and $i+1$ never interacts has zero measure. It is clear that this statement can be proven in complete analogy with the proof of Lemma 3.6.

The above lemma suggests that our strategy will be a proof by induction on the number of balls. Note that it is sufficient to prove the ergodicity since, for our examples, the mixing (and the $K$-property) on the ergodic components is insured by the general theory (see [P, KS]). Also, in the same papers, it is proven that, if our system has Lyapunov exponents different from zero almost everywhere, then it decomposes mod-0 into, at most, countably many different ergodic components.

## 4. Local Ergodicity

The aim of this section is to introduce an invariant set $\tilde{\Omega} \supset \Omega$, such that for each $x \in \tilde{\Omega}$ there exists a neighborhood of $x$ that belongs, apart from a set of measure zero, to only one ergodic component (one says in this situation that the system is locally ergodic near $x$).

As was noted before, our system can be naturally treated as a billiard in $2n$ dimensions. Owing to the fact that the boundary $\partial Q$ of the original region is dispersing (its curvature $K$ is strictly positive), the boundary of the corresponding $2n$ dimensional billiard is semi-dispersing (has non-negative curvature). This allows us to use the material elaborated for general semi-dispersing billiards.

More precisely, to prove local ergodicity we apply the aforementioned Hopf argument using the results from the general theory of semi-dispersing billiards (see [SC, KSS3]) which allow us to produce an abundance of manifolds of size $\delta$, for $\delta$ sufficiently small, in some neighborhood of a sufficient point. Since this type of argument is explained, in a more or less detailed way, in various paper [SC, KSS3, Bu2, LW] we will refer to them for all the technical points. The key observation in the construction of $\tilde{\Omega}$ is provided by the following.

**Lemma 4.1.** *If the piece of trajectory $(x, [t_1, t_2])$ of $\phi^t$ is sufficient (see Definition 1.1), also the reverse trajectory $(x, [-t_2, -t_1])$, of $\phi^{-t}$, is sufficient.*

*Proof.* Assuming this is not the case, it would mean that there exists $(\delta q, \delta p) \in \mathcal{T}_{\phi^{t_2}(x)}\mathcal{M}$ such that $\langle \delta q, \delta p\rangle = 0$ and, for $(\delta q', \delta p') = d\phi^{t_2-t_1}(\delta q, \delta p)$, $\langle \delta q', \delta p'\rangle = 0$.

But the latter would imply

$$\langle \delta q, \delta p\rangle > 0$$

because of the sufficiency of the forward trajectory. □

The previous lemma, remembering that $\phi^t\Omega \subset \Omega$ for $t \geqq 0$, readily suggests

**Definition 4.2.**

$$\tilde{\Omega} = \left( \bigcup_{t \leqq 0} \phi^t(\Omega) \right) \cup \left( \bigcup_{t \geqq 0} \phi^t(-\Omega) \right).$$

*Here* $-(q, p)$ *stands for the time reversal point* $(q, -p)$.

The next assertion, Theorem 4.3, is a specific version of the Transversal Fundamental Theorem for semi-dispersing billiards. It is stated without a proof; we refer to the above-mentioned papers for the latter. Some basic definitions and facts related to the formulation of this theorem are discussed in Appendix II.

We denote by $\partial_+ \mathcal{M}$ the "outgoing" part of the boundary $\partial \mathcal{M}$ of the phase space $\mathcal{M}$ and by $T$ the Poincaré map between consecutive collisions induced by the flow $\phi^t$. Let $\mu$ denote the standard $T$-invariant measure on $\partial_+ \mathcal{M}$. Furthermore, $\mathcal{R}^+ \subset \partial_+ \mathcal{M}$ denotes the singularity set for $T$ and $\mathcal{R}^- \subset \partial_+ \mathcal{M}$ that for $T^{-1}$. Finally, let $\mu_+$ be the measure on $\mathcal{R}^+$ induced by the measure $\mu$ and $\mu_-$ be that on $\mathcal{R}^-$.

The definitions of the local stable and unstable manifolds and related objects figuring in the formulation of the Fundamental Theorem (such as parallelograms with smooth faces (sides) parallel or transversal to those manifolds) may be found in the same papers as before. Speaking of a diameter (or size) of a given set we have in mind the standard Riemannian metric on $\partial_+ \mathcal{M}$. We also use, in a repeated manner, various intermediate geometric constructions such as a natural identification of the tangent spaces at different points; this is done by using the Euclidean structure of the phase space.

**Theorem 4.3** (Unstable Version of the Transversal Fundamental Theorem for Semi-Dispersing Billiards). *Suppose that* $\mathcal{R}^+$ *is a finite union of* $4n - 3$ *dimensional* $\mathcal{C}^2$*-manifold (apart from the boundary that is supposed to consist of a finite union of* $4n - 4$ *dimensional manifolds) and so is* $\mathcal{R}^-$. *Suppose furthermore that the following assumptions are valid.*

i) *For each* $z \in \mathcal{R}^-$, *the tangent space* $\mathcal{T}_{Tz}(T\mathcal{R}^-)$ *contains a* $2n - 1$ *dimensional subspace that belongs strictly to the forward cone* $C(Tx)$, *and for each* $z \in \mathcal{R}^+$, *the tangent space* $\mathcal{T}_{T^{-1}z}(T^{-1}\mathcal{R}^+)$ *contains a* $2n - 1$ *dimensional subspace that belongs strictly to the backward cone* $C_-(T^{-1}z)$.

ii) *(The Sinai-Chernov Ansatz) For* $\mu_-$*-almost all* $z \in \mathcal{R}^-$, *the tangent maps* $D_z T^m$ *obey*

$$\lim_{m \to \infty} \| D_z T^m \xi \| = \infty \qquad \text{for all } \xi \in C(z)\{0\}.$$

*Let* $x \in \partial_+ \mathcal{M}$ *be a point with a smooth and sufficient trajectory for* $\phi^t$, $t \leqq 0$ *(backward smoothness and sufficiency). Then:*

a) $E^u = \bigcap_{m > 0} DT^m(C(T^{-m}(x)))$ *is a* $2n - 1$ *dimensional subspace of* $\mathcal{T}_x \partial_+ \mathcal{M}$.

b) *For each* $c_1 \in (1/2, 1)$ *there exist: a neighborhood of* $x$ *(in* $\partial_+ \mathcal{M}$*),* $\mathcal{U}(x)$, *a constant* $c_2 \in (1 - c_1, 1)$ *and a natural number* $k_*$, *such that the following holds. Suppose a one-parameter family* $\{\mathcal{G}_\delta, \delta > 0\}$, *of coverings of* $\mathcal{U}(x)$, *is given, with the following properties:*

1. *The elements of covering $\mathscr{G}_\delta$ are topological open parallelograms with smooth sides, of size $\delta$, either parallel to $E^u$ (that is, with tangent space, at each point, containing a $2n-1$ dimensional sub-space parallel to $E^u$) or uniformly transversal to $E^u$ (more precisely, with a $2n-1$ dimensional subspace, of the tangent space at each point, contained in the backward cone $C_-(x)$).*

2. *Each element $G \in \mathscr{G}_\delta$ intersects at most $k_*$ other elements. Moreover, for each $G \in \mathscr{G}_\delta$ there exists a collection of neighboring elements $N(G) \subset \mathscr{G}_\delta$ such that $\bigcup_{G' \in N(G)} G' \supset G$ and $\mu(G' \cap G) > c_2 \mu(G)$ for each $G' \in N(G)$ (note that the volume of an element of the cover will be, roughly, $\delta^{4n-2}$).*

*Let us divide the whole set of the elements of the covering $\mathscr{G}_\delta$ into two disjoint collections $\mathscr{G}_\delta^{(0)}$ and $\mathscr{G}_\delta^{(1)}$ as follows. An element $G \in \mathscr{G}_\delta$ belongs to $\mathscr{G}_\delta^{(0)}$* iff *the total measure of the collections of unstable manifolds $W^u$ in $G$ such that $W^u \cap \partial G = W^u \cap (\partial G \setminus \{\text{sides parallel to } E^u\})$ (the manifolds run almost parallel to $E^u$ from side to side in $G$) is larger than $c_1 \mu(G)$. Otherwise, $G$ belongs to $\mathscr{G}_\delta^{(1)}$.*

*Then the total measure of the elements in $\mathscr{G}_\delta$, but not in $\mathscr{G}_\delta^{(0)}$, is small in a strong sense; namely:*

$$\lim_{\delta \to 0} \delta^{-1} \mu\left( \bigcup_{G \in \mathscr{G}_\delta^{(1)}} G \right) = 0.$$

*Remark 4.4.* In the literature mentioned above the Fundamental Theorem is stated in a slightly different fashion. The main difference consists in the definition of $\mathscr{G}_\delta^{(0)}$: where it is required an abundance of unstable manifolds only near the boundary of an element. Nevertheless, the theorem holds also in our version without any significant change to the proof. In addition, hypothesis i) is sometimes not mentioned explicitly, although it can be found, in a form similar to the one used here, in [Bu2].

Theorem 4.3 has the obvious analog for the local stable manifolds (provided the Sinai-Chernov Ansatz, stated in (ii), is replaced with its converse for $\mathscr{R}^+$ and the point $x$ has a forward smooth and sufficient trajectory). We prove in Appendix I that condition (i) is satisfied for our cases. Furthermore, in Theorem 5.14 we prove that, for the system of $n$ balls, the set of non-sufficient points is of $\mu_\pm$-measure zero in $\mathscr{R}^\pm$ under the condition that, for any $m < n$, the same holds for the system of $m$ balls and, in addition, the set of non-sufficient points has codimension two in the phase space of the $m$ balls system. Moreover, it is known that, in a more general situation, for any semi-dispersing billiard, if a point $z \in \partial_+ \mathscr{M}$ is sufficient, then $\lim_{m \to \infty} \| D_z T^m \xi \| = \infty$ for all $\xi \in C(z) \setminus \{0\}$ (see [S3, LW]). Here it is crucial that hypothesis ii) does not require exponential growth of the tangent vector (points that satisfy (ii) may have zero Lyapunov exponents).

Summing up, we have seen that Theorem 4.3 applies to the system of $n$ balls if the set of non-sufficient points has zero measure and codimension two in $\mathscr{M}$ and zero $\mu_\pm$-measure in $\mathscr{R}^\pm$ for any system of $m < n$ balls.

*Remark 4.5.* To verify conditions (ii) we need to know properties of the trajectory at all times and to possess information about sets of zero measure. This makes the above condition hard to check (at least in general), but it is probably unavoidable and reflects the fact that the approach considered here cannot be purely measure theoretic.

Our next assertion, Theorem 4.6, can also be found in the literature ([SC, KSS1, LW]). We will sketch the proof, so that the reader can see how Theorem 4.3 is used to obtain some knowledge about ergodic components.

**Theorem 4.6** (Local Ergodicity). *Assuming conditions* (i) *and* (ii) *of Theorem 4.3, it follows that, given $x \in \tilde{\Omega}$ out of a sub-set of measure zero and codimension two, there exists a neigborhood of $x$ that belongs, apart from a set of zero measure, to only one ergodic component.*

*Proof* (a Sketch). From Theorem 3.2, Definition 4.2 and the hypotheses at hand follows that we can apply Theorem 4.3 to $x$. We start by sketching the proof of the theorem under the assumption that the point $x$ has a smooth trajectory and is sufficient in both directions. Since sufficiency is a property of a segment of trajectory, this last assumption has the only purpose to simplify the discussion: in general one can consider $x$ and one of its images to obtain a point sufficient backward and one sufficient forward on the same orbit. Therefore, the following argument can easily be adapted to the situation without the last assumption.

The proof uses, as we said before, the standard Hopf argument [H]. In our context this means that, given a sufficient point with a smooth trajectory, we can employ Theorem 4.3 to produce an abundance of stable and unstable manifolds of diameter $\delta$. Those manifolds can then be used to construct chains that connect different ergodic components in the neighborhood of $x$, showing that this neighborhood intersects only one ergodic component (see [KSS3, Zig-Zag Theorem], also [Bu2]).

More precisely, one observes that, e.g., a stable manifold belongs, morally, to only one ergodic component, since the forward ergodic average for any continuous function is the same for all the points in the manifold. In addition, the stable and unstable manifolds form absolutely continuous foliations (see [KS]) and the backward and forward ergodic averages are equal almost everywhere. These last two facts imply that, when a set of positive measure of stable manifolds intersects a set of positive measure of unstable manifolds, almost all the intersection points lie in the set in which the forward and backward ergodic average are the same.

Now take a family of parallelograms coverings $\mathscr{G}_\delta$, $\delta > 0$, with the properties listed in conditions b1) and b2) of Theorem 4.3; the construction of those covering is rather standard (see, e.g., [BS, KSS3]). From the previous considerations it follows easily that, according to Theorem 4.3 and its stable version, we can construct, near the sides of the parallelograms that are from the collection $\mathscr{G}_\delta^{(0)}$, thick chains of stable and unstable manifolds that belong to the same ergodic component. Moreover, property b) of Theorem 4.3 implies that a connected component of the union of parallelograms from $\mathscr{G}_\delta^{(0)}$ has a measure arbitrarily near to the total measure of $\mathscr{G}_\delta$, when $\delta$ goes to zero. In fact, for $\mathscr{G}_\delta^{(0)}$ to be disconnected uniformly in $\delta$, it is necessary that it is divided by a boundary composed of elements of $\mathscr{G}_\delta^{(1)}$ that enclose a volume of order one. But this is possible only if the measure of the union of the elements from $\mathscr{G}_\delta^{(1)}$ is proportional to $\delta$ (that is, approximately the area of the dividing boundary times $\delta$), contrary to b) (see [KSS3] for more details).

A little more careful argument is needed for points that have a smooth trajectory in one direction only, e.g., forward, but the same conclusion holds (see

[SC, KSS3, Bu2, LW]). The key observation, in this case, is that the construction of Theorem 4.3 can still be carried out, although only in one direction (forward), producing an abundance of stable manifolds. Moreover such manifolds are transversal to the singularity manifolds. Here "singularity manifold" is used, loosely, to mean points at which either a tangent or a multiple collision occurs, (and images, under the dynamics, of such points), see Appendix I. Theorem 4.3 can also be applied to each side of the singularity manifold; the previous considerations show that a neighborhood of $x$ contains at most two ergodic components separated by this manifold. Finally, it is possible to use the stable manifolds that cross the singularity manifold to show that the two sides of the singularity manifolds actually belong to the same ergodic component. □

## 5. Global Ergodicity

In the previous section we were able to prove that, if the system with $m$ balls is ergodic and mixing for any $m < n$, then the phase space of the system generated by $n$ balls decomposes mod 0 in, at most, countably many ergodic components, each component being open. The following stage is to prove that there is only one ergodic component. Our proof is by induction on the number of balls. Since the statement for one ball is well known [S1, S2, G], we need only to prove the inductive step from $m < n$ to $n$. The subsequent theorem clarifies the relevant properties that we need to study.

**Theorem 5.1.** *If the set $\mathcal{M}\backslash\tilde{\Omega}$ is of measure zero and codimension two and condition* ii) *of Theorem 4.3 holds, then the system is ergodic, and has the K-property (which implies mixing).*

*Proof.* On the one hand, according to Theorem 4.6, if $A \subset \tilde{\Omega}$ is connected, then it belongs to only one ergodic component. This result is stated there for the Poincaré map but it can be translated easily in the corresponding statement for the flow $\phi^t$. On the other hand, $\mathcal{M}\backslash\tilde{\Omega}$ of codimension two implies that $\tilde{\Omega}$ is connected [E], hence the result. The $K$-property, as already mentioned, is a consequence of the general theory of hyperbolic systems [P, KS, S2]. □

We approach the conclusion of our discussion with Theorem 5.2 below.

**Theorem 5.2.** *If, for each $m < n$, the set $\mathcal{M}\backslash\tilde{\Omega}$ is of measure zero and codimension two, and condition* (ii) *of Theorem 4.3 holds for the m and n balls system, then the set $\mathcal{M}\backslash\tilde{\Omega}$ is of measure zero and codimension two also for the system of n balls.*

*Proof.* By definition, $\tilde{\Omega} \supset \Omega$. Moreover, in our hypothesis, Theorem 3.6, together with Theorem 5.1, claims that $\mu(\mathcal{M}\backslash\Omega) = 0$, so $\mu(\mathcal{M}\backslash\tilde{\Omega}) = 0$. Notice that this implies that the Lyapunov exponents are different from zero almost everywhere and that almost every point has local stable and unstable manifolds (see end of Sect. 3).

To prove the second, and harder, part of the statement we note that the complement of $\tilde{\Omega}$ decomposes naturally in three disjoint sets:

1. double singularity points,
2. points, with smooth trajectories, for which the system splits, both in the past and in the future, into non-interacting sub-systems,
3. points for which the trajectory meets a singularity manifold in one time direction and behaves like two non-interacting sub-systems in the other.

We discuss the three cases separately, showing that each one of the three sets has codimension two.

*Double Singularities*. These are points for which the trajectory meets a singularity manifold both in the past and in the future. By definition this set is given by

$$\left(\bigcup_{t>0}\phi^t\mathscr{R}^-\right)\cap\left(\bigcup_{t<0}\phi^t\mathscr{R}^+\right). \tag{5.1}$$

Here, with a slight change of notation with respect to the preceding section, $\mathscr{R}^\pm$ are the singularity sets in $\mathscr{M}$ (corresponding to tangent and multiple collisions) which are related in a natural way to the previous ones (which were subsets of $\partial_+\mathscr{M}$). It is shown in Appendix I that the pre-images of $\mathscr{R}^+$ are transversal to the images of $\mathscr{R}^-$ (their tangent spaces contain $2n-1$ dimensional subspaces that belong strictly the complementary cones). Therefore this set has codimension two, being the countable union of sets of codimension two $\left(\text{note that the union } \bigcup_{t\in[0,1]}\phi^t\mathscr{R}^\pm \text{ is of codimension one in } \mathscr{M}\right)$.

*Non-Interacting Sub-Systems*. Here we are looking at the set of points with smooth trajectory for which there exist $k_-$, $k_+\in\{1,\dots,n\}$ and $t_-$, $t_+\in\mathbb{R}$, $t_-\leqq t_+$, such that, on the one hand, there is no $k_-$, $k_-+1$ collision for each $t\leqq t_-$ and, on the other hand, the particles $k_+$ and $k_++1$ never collide for $t\geqq t_+$.

It is convenient to call $\Delta_+(k_+,t_+)$ and $\Delta_-(k_-,t_-)$ the sets of points, with smooth trajectory, which do not experience a $k_+$, $k_++1$ collision for $t\geqq t_+$, and a $k_-$, $k_-+1$ collision for $t\leqq t_-$, respectively. Then the present goal is to prove that

$$\Delta=\bigcup_{\substack{k_+,k_-\in\{1,\dots,n\}\\ t_-\leqq t_+\in\mathbb{R}}}(\Delta_+(k_+,t_+)\cap\Delta_-(k_-,t_-))$$

is a set of codimension two.

Since $\Delta_+(k,t)\subset\Delta_+(k,j)$ if $j\geqq t$ and $\Delta_-(k,t)\subset\Delta_-(k,j)$ if $j\leqq t$, we have

$$\Delta=\bigcup_{\substack{k_+,k_-\in\{1,\dots,n\}\\ j_-\leqq j_+\in\mathbb{Z}}}(\Delta_+(k_+,j_+)\cap\Delta_-(k_-,j_-)).$$

It is therefore sufficient to show that each one of the sets $\Delta_+(k_+,j_+)\cap\Delta_-(k_-,j_-)$ is of codimension two (since a countable union of codimension two sets is again of codimension two, see [KSS1]). To pursue the argument, a few definitions are necessary.

**Definition 5.3.** *Given $m \in \mathbb{N}$ and $E > 0$ we put:*

$\mathcal{M}_{m,E}$ = *phase space for the m balls system with energy E,*
$\tilde{\Omega}_{m,E}$ = *the set $\tilde{\Omega}$ (see Definition 4.2) for the system of m balls with energy E,*
$\Theta_{m,E} = \mathcal{M}_{m,E} \backslash \tilde{\Omega}_{m,E}$,
$\phi^t_{m,E}$ = *dynamics for the system of m balls with energy E,*
$W^s_{m,E}(x)$ = *stable manifold at x for the system of m balls with energy E, provided it exists and has dimension $2m-1$,*
$W^u_{m,E}(x)$ = *the same for the unstable manifold,*
$E(x)$ = *(kinetic) energy of the point $x = (q, p)$.*

During the rest of the proof we often consider the sets and sub-dynamics relative to $m < n$ balls as objects embedded in the system of $n$ balls. We will not state this explicitly: the context is self-explanatory. In addition, we suppress the subscript $E$ unless its omission creates ambiguites.

**Lemma 5.4.** *The sets $\tilde{\Theta}_i = \bigcup_{E=E_1+E_2} (\Theta_{i,E_1} \times \mathcal{M}_{n-i,E_2}) \cup (\mathcal{M}_{i,E_1} \times \Theta_{n-i,E_2})$, for any* $E > 0$ and $i \in \{1, \ldots, n$, *are of codimension two in $\mathcal{M}_{n,E}$.*

*Proof.* We start by noticing that $\mathcal{M}_{i,E_1} \times \mathcal{M}_{n-i,E_2}$, for any $E_1 > 0$, and $E_2 > 0$, is a codimension one manifold in $\mathcal{M}_{n,E}$ where $E = E_1 + E_2$. At the same time, the invariant measure with respect to the relative dynamics on the latter manifold decomposes naturally into the product of the invariant measures on the previous ones. Furthermore, by the hypotheses of Theorem 5.2, $\Theta_{i,E_1}$ has codimension two in $\mathcal{M}_{i,E_1}$. Consequently $\Theta_{i,E_1} \times \mathcal{M}_{n-i,E_2}$ has codimension two in $\mathcal{M}_{i,E_1} \times \mathcal{M}_{n-i,E_2}$. The proof of this last statement can be found in [E, 1.5.16] or it can be carried out similarly to the proof of Property 4 in [KSS4, Sect. 4.1]. The ideas used in the above-mentioned papers are the same as those we will use to conclude the proof of Lemma 5.4, that is, to prove the following: if $x \in \mathcal{M}_{n,E}$, $E > 0$, and $\mathcal{U}(x)$ is a neigborhood of $x$, then $\mathcal{U}(x) \backslash \tilde{\Theta}_i$ is connected.

Let $\mathcal{U}_{E_1}(x) = \mathcal{U}(x) \cap (\mathcal{M}_{i,E_1} \times \mathcal{M}_{n-i,E_2})$, $E_2 = E - E_1$, then $\mathcal{U}(x)$ can be represented as the union $\bigcup_{E_1 \in [0,E]} \mathcal{U}_{E_1}(x)$. We restrict ourselves to the case in which all the sets $\mathcal{U}_{E_1}(x)$ are connected. Indeed, it is clear that any neighborhood $\mathcal{U}(x)$ contains another neighborhood that satisfies the previous assumption. Hence, the general case can always be reduced to the present one.

If $\mathcal{U}(x) \backslash \tilde{\Theta}_i$ is not connected, then there exist open disjoint sets $V_1$, $V_2$ such that $V_1 \cup V_2 \supset \mathcal{U}(x) \backslash \tilde{\Theta}_i$, $V_j \cap (\mathcal{U}(x) \backslash \tilde{\Theta}_i) \neq \emptyset$, $j \in \{1,2\}$. However, we will show that this is impossible.

Let us define, for $E_1 \in (0, E)$ and $E_2 = E - E_1$,

$$\pi(E_1) = \begin{cases} 0 & \text{if } \mathcal{U}_{E_1}(x) = \emptyset \\ 1 & \text{if } \mathcal{U}_{E_1}(x) \backslash (\Theta_{i,E_1} \times \mathcal{M}_{n-i,E_2} \cup \mathcal{M}_{i,E_1} \times \Theta_{n-i,E_2}) \subset V_1 \, . \\ 2 & \text{if } \mathcal{U}_{E_1}(x) \backslash (\Theta_{i,E_1} \times \mathcal{M}_{n-i,E_2} \cup \mathcal{M}_{i,E_1} \times \Theta_{n-i,E_2}) \subset V_2 \end{cases}$$

The extreme values ($E_1 = 0$ and $E_1 = E$) correspond to cases in which $\mathcal{U}_{E_1}(x)$ is a manifold of codimension larger than or equal to two and consequently can be ignored. The function $\pi$ is well defined since $\mathcal{U}_{E_1}(x) \backslash (\Theta_{i,E_1} \times \mathcal{M}_{n-i,E_2} \cup \mathcal{M}_{i,E_1}$

$\times \Theta_{n-i,E_2})$ is connected (in fact, it is a connected set from which a codimension two sub-set has been removed).

Moreover, $\pi$ is a continuous function when different from zero. To see this, it is sufficient to consider a point $y \in \mathscr{U}_{E_1}(x) \backslash (\Theta_{i,E_1} \times \mathscr{M}_{n-i,E_2} \cup \mathscr{M}_{i,E_1} \times \Theta_{n-i,E_2})$ for which $\pi(E_1) \neq 0$. It is clear that there exists a non-empty neighborhood $B(y) \subset V_{\pi(E_1)} \cap \mathscr{U}(x)$. Accordingly, since $\mathscr{M}_{i,E_1} \times \mathscr{M}_{n-i,E_2}$ forms a continuous foliation of $\mathscr{M}$, $\mathscr{U}_{\mathscr{E}}(x)$ will have a non-empty intersection with $B(y)$, and therefore with $V_{\pi(E_1)}$, for values of $\mathscr{E}$ closed to $E_1$. It then follows that $\mathscr{U}_{\mathscr{E}}(x) \backslash ((\Theta_{i,E_1} \times \mathscr{M}_{n-i,E_2}) \cup (\mathscr{M}_{i,E_1} \times \Theta_{n-i,E_2})) \subset V_{\pi(E_1)}$. Obviously, this means that $\pi(E_1)$, for $\pi(E_1) \neq 0$, can assume only one value, contrary to the hypothesis that $\mathscr{U}(x)$ is disconnected. $\square$

As a consequence of Lemma 5.4, we can restrict ourselves to the study of the sets $\Delta_+^{(1)}(k_+, j_+) \cap \Delta_-^{(1)}(k_-, j_-)$, where

$$\Delta_+^{(1)}(k_+, j_+) = \Delta_+(k_+, j_+) \backslash \tilde{\Theta}_{k_+},$$
$$\Delta_-^{(1)}(k_-, j_-) = \Delta_-(k_-, j_-) \backslash \tilde{\Theta}_{k_-}.$$

Given $x \in \Delta_+^{(1)}(k_+, j_+) \cap \Delta_-^{(1)}(k_-, j_-)$, it follows from the definition that, setting $x^+ = \phi^{j_+}(x) = (x_1^+, x_2^+)$ and $x^- = \phi^{j_-}(x) = (x_1^-, x_2^-)$, we have $x_1^+ \in \tilde{\Omega}_{k_+}$, $x_2^+ \in \tilde{\Omega}_{n-k_+}$ and $x_1^- \in \tilde{\Omega}_{k_-}$, $x_2^- \in \tilde{\Omega}_{n-k_-}$.

Our strategy will be to prove the following (which implies the codimension two of the intersection $\Delta_+^{(1)}(k_+, j_+) \cap \Delta_-^{(1)}(k_-, j_-)$; see [E]): there exists a neighborhood $\mathscr{U}(x)$ of $x$ such that any open sub-set in $\mathscr{U}_r(x) = \mathscr{U}(x) \backslash (\Delta_+^{(1)}(k_+, j_+) \cap \Delta_-^{(1)}(k_-, j_-))$ is connected.

To prove the above statement we proceed in a fashion similar to the one employed in Sect. 4. In other words, we construct: (i) a pair of foliations induced by local stable and unstable manifolds for our reduced sub-systems, and (ii) a family of parallelogram coverings $\mathscr{G}_\delta$, $\delta > 0$, with properties analogous to the ones indicated in conditions b1) and b2) of Theorem 4.3. We show that, precisely as in Theorem 4.3, a "growing majority" of the elements of the covering has, near the sides, a "large" collection of manifolds that belong to $\mathscr{U}_r(x)$. Finally, as in Theorem 4.6, chains of these manifolds are used to prove that $\mathscr{U}_r(x)$ is connected.

First of all we choose $\mathscr{U}(x)$ so small that $\phi^t(y)$ is smooth for each $t \in [j_-, j_+]$ and $y \in \mathscr{U}(x)$.

The second step is to discard some unwanted directions; we do so by defining, near the point $x$, some local manifolds of codimension one and two. We see in the following how to extend the results obtained for points in such manifolds to the full neighborhood of $x$.

By $B_\delta(\cdot)$ we denote the ball of radius $\delta$ around a given point in Euclidean space of the appropriate dimensionality.

**Definition 5.5.** *Given $z \in \mathscr{M}$, $z = (z_1, z_2) \in \mathbb{R}^{4k} \times \mathbb{R}^{4(n-k)}$, $z_i = (q^{(i)}(z), p^{(i)}(z))$ and $q(z) = (q^{(1)}(z), q^{(2)}(z))$, we define the local manifolds $\hat{\mathscr{U}}(z, \delta_0)$ and $\tilde{\mathscr{U}}(z, \delta_0)$ of codimension one and two, respectively, by*

$$\hat{\mathscr{U}}(z, \delta_0) = \{(q, p) \in \mathscr{M} \cap B_{\delta_0}(z) \mid \langle q - q(z), p \rangle = 0\},$$
$$\tilde{\mathscr{U}}(z, \delta_0) = \{(q^{(1)}, q^{(2)}, p^{(1)}, p^{(2)}) \in \mathscr{M} \cap B_{\delta_0}(z) \mid \langle q^{(1)} - q^{(1)}(z), p^{(1)} \rangle = 0$$
$$\langle q^{(2)} - q^{(2)}(z), p^{(2)} \rangle = 0\}.$$

Our strategy is based on a study of the sub-systems. It is therefore necessary to decompose our manifolds according to the foliation $\bigcup_{\mathscr{E}} \mathscr{M}_{k,\mathscr{E}} \times \mathscr{M}_{n-k,E-\mathscr{E}}$.

**Definition 5.6.** *Given* $z \in \mathscr{M}$, $z = (z_1, z_2) \in \mathbb{R}^{4k} \times \mathbb{R}^{4(n-k)}$, $z_i = (q^{(i)}(z), p^{(i)}(z))$, *we define local manifolds* $\hat{\mathscr{U}}_k(z,\delta) \subset \mathscr{M}_k$ *and* $\hat{\mathscr{U}}_{n-k}(z,\delta) \subset \mathscr{M}_{n-k}$ *by*

$$\mathscr{U}_k(z,\delta) = \{(q^{(1)}, p^{(1)}) \in B_\delta(z_1) \mid \|p^{(1)}\| = \|p^{(1)}(z)\|;\ \langle q^{(1)} - q^{(1)}(z), p^{(1)}\rangle = 0\},$$

$$\hat{\mathscr{U}}_{n-k}(z,\delta) = \{(q^{(2)}, p^{(2)}) \in B_\delta(z_2) \mid \|p^{(2)}\| = \|p^{(2)}(z)\|;\ \langle q^{(2)} - q^{(2)}(z), p^{(2)}\rangle = 0\}.$$

It follows from the definitions that there exists $\delta_1 > \delta_0$ such that, for $w \in \tilde{\mathscr{U}}(z,\delta_0)$, the intersection $(\hat{\mathscr{U}}_k(w,\delta_1) \times \hat{\mathscr{U}}_{n-k}(w,\delta_1)) \cap \tilde{\mathscr{U}}(z,\delta_0)$ is a manifold of codimension one in $\tilde{\mathscr{U}}(z,\delta_0)$ with the boundary contained in the boundary of $\tilde{\mathscr{U}}(z,\delta_0)$. The final step in our analysis of the neighborhood $\mathscr{U}(x)$ is the construction of the aforementioned foliations. First we define a direction transversal to the energy foliation.

**Definition 5.7.** *Given* $z \in \mathscr{M}$, $z = (z_1, z_2) \in \mathbb{R}^{4k} \times \mathbb{R}^{4(n-k)}$, $z_i = (q^{(i)}(z), p^{(i)}(z))$, *we define the curve* $l_z\colon [-a, a] \to \mathscr{M}$, *a being chosen small enough depending on* $z$, *by* $l_z(s) = (l_z^{(1)}(s), l_z^{(2)}(s))$ *with*

$$l_z^{(1)}(s) = \left(q^{(1)}(z), \left(1 + \frac{s}{\|p^{(1)}(z)\|^2}\right)^{\frac{1}{2}} p^{(1)}(z)\right),$$

$$l_z^{(2)}(s) = \left(q^{(2)}(z), \left(1 - \frac{s}{\|p^{(2)}(z)\|^2}\right)^{\frac{1}{2}} p^{(2)}(z)\right).$$

It is easy to check that $E(l_z(s)) = E(z)$ (which means that $l_z$ is a curve in $\mathscr{M}$, i.e. at constant total energy). In addition, if $\delta_0$ is chosen small enough, there exists $\delta_2 > \delta_1$ such that

$$\bigcup_{s \in [-\delta_2, \delta_2]} (\hat{\mathscr{U}}_k(l_z(s), \delta_2) \times \hat{\mathscr{U}}_{n-k}(l_z(s), \delta_2)) \supset \tilde{\mathscr{U}}(z, \delta_0).$$

*Remark 5.8.* Note that sufficiency and the related properties are of a geometric nature. That is, they depend only on the geometry of the trajectory and not on the value of the total energy. It is then an important remark that the points on the curves $l_z^{(1)}$ and $l_z^{(2)}$ have the same trajectory, apart from the time and velocity scaling, under the flows $\phi_k^t$ and $\phi_{n-k}^t$, respectively.

This concludes our discussion on the decomposition of the quantities of interest according to the structure of the sub-systems. We are now ready to use it in the problem at hand.

It is convenient to perform our constructions around the points $x^+$ and $x^-$ separately. By hypothesis, $x^+ \in \Delta_+^{(1)}(k_+, 0)$ and $x^- \in \Delta_-^{(1)}(k_-, 0)$. According to Remark 5.8, this means that it is possible to choose $\delta_2$, and consequently $\delta_0$, so small that $l_{x^+}^{(1)}(s) \notin \tilde{\Theta}_{k_+}$ and $l_{x^-}^{(2)}(s) \notin \tilde{\Theta}_{k_-}$ for $s \in [-\delta_2, \delta_2]$. We can then apply, inside each one of the sets $\hat{\mathscr{U}}_{k_+}(l_{x^+}(s), \delta)$, $\hat{\mathscr{U}}_{n-k_+}(l_{x^+}(s), \delta)$, $\hat{\mathscr{U}}_{k_-}(l_{x^-}(s), \delta)$ and $\hat{\mathscr{U}}_{n-k_-}(l_{x^-}(s), \delta)$, the results of Sect. 4 (Theorem 4.3 in particular) to the flows $\phi_{k_+}^t$, $\phi_{n-k_+}^t$, $\phi_{k_-}^t$, $\phi_{n-k_-}^t$, respectively. See Appendix II for details on how to apply Theorem 4.3 to flows, instead of maps, in our present situation. In doing so, we try to obtain coverings of the sets $\tilde{\mathscr{U}}(x^+, \delta_0)$ and $\tilde{\mathscr{U}}(x^-, \delta_0)$ and abundances of local stable and unstable manifolds in these sets. More precisely, we have in mind manifolds $\tilde{W}^u = W_{k_-}^u \times W_{n-k_-}^u$ near $x^-$ and $\tilde{W}^s = W_{k_+}^s \times W_{n-k_+}^s$ near $x^+$. Note that these manifolds are related to the local stable and unstable manifolds of the complete system only in the case where the latter never experiences a collision of type $k_+$, $k_+ + 1$ in the future (or $k_-$, $k_- + 1$ in the past). Indeed, in this case the dynamics of the complete system would be precisely the direct product of the two sub-dynamics. In order to avoid confusion, let us remark again that, for example, $W^u_{k_-}$ is obtained by erasing the balls $\{k_- + 1, \dots, n\}$ and using the dynamics of the remaining ones.

**Definition 5.9.**

$$C^+_\sigma(i) = \{(q,p) \in \mathcal{M} \mid \|(q',p') - (q,p)\| < \sigma \textit{ implies that the next collision, for } \phi^t(q',p')\textit{, will be between } i \textit{ and } i+1\},$$

$$C^-_\sigma(i) = \{(q,p) \in \mathcal{M} \mid \|(q',p') - (q,p)\| < \sigma \textit{ implies that the previous collision, for } \phi^t(q',p')\textit{, has been between } i \textit{ and } i+1\},$$

$$t^-_\sigma(y) = \sup\{t \leqq 0 \mid \phi^t_{k_-} \times \phi^t_{n-k_-}(y) \in C^-_\sigma(k_-)\},$$

$$t^+_\sigma(y) = \inf\{t \geqq 0 \mid \phi^t_{k_+} \times \phi^t_{n-k_+}(y) \in C^+_\sigma(k_+)\}.$$

**Lemma 5.10.** *If, given $y \in \tilde{\mathcal{U}}(x_-, \delta_0)$ and $\delta > \delta_0$, $t^-_\delta(y) > -\infty$, then all the points of the smooth manifold $\tilde{W}^u(y) \cap B_\delta(y)$ (provided it exists) have a $k_-$, $k_- + 1$ collision in the past.*

*Proof.* On the one hand, if some point of $\tilde{W}^u(y)$ has a $k_-$, $k_- + 1$ collision in the time interval $[t^-_\delta(y), 0]$, we have what we are looking for. On the other hand, the points of $\tilde{W}^u(y) \cap B_\delta(y)$ have a configuration distance less than $\delta$ from $y$ (i.e. if $w \in \tilde{W}^u(y) \cap B_\delta(y)$ then $\|q(y) - q(w)\| < \delta$) and Lemma 2.3 implies that such distance cannot grow under the action of $\phi^t_{k_-} \times \phi^t_{n-k_-}$, $t \leqq 0$. Therefore, since the points that do not experience a $k_-$, $k_- + 1$ collision will evolve according to the two independent sub-dynamics, their distance from $\phi^t_{k_-} \times \phi^t_{n-k_-}(y)$ cannot be more then $\delta$. Consequently, remembering that the definition of $t^-_\delta(y)$ implies that, following the trajectory of $y$ under the product dynamics, the two center of the balls $k_-$ and $k_- + 1$ would pass at a distance less than $R - \delta$ from each other, we can claim that the points under consideration are bound to have a collision of type $k_-$, $k_- + 1$ after the time $t^-_\delta(y)$. □

From the proof it is immediately apparent that the same argument proves the analogous statement for $\tilde{W}^s$. Lemma 5.10 goes in the right direction insofar as it shows that a simple condition can ensure that the local manifolds $\phi^{-j_-}(\tilde{W}^u)$, $\phi^{-j_+}(\tilde{W}^s)$ belong to $\mathcal{U}_r(x)$. Unfortunately, there is one drawback: the fact that $\tilde{W}^u$, $\tilde{W}^s$ are $2n-2$ dimensional manifolds; this means that, generically, they do not intersect each other.

It is then clear that we need $2n-1$ dimensional manifolds with the property stated in Lemma 5.10, if we want to be able to construct chains of intersecting manifolds. We will achieve this by adding an extra dimension to the manifolds $\tilde{W}^u$, $\tilde{W}^s$.

**Definition 5.11.** *Given a point $z = (z_1, z_2) \in \mathcal{M}_{k_-} \times \mathcal{M}_{n-k_-}$, where $z_i = (q^{(i)}(z), p^{(i)}(z))$, if a local manifold $\tilde{W}^u(z)$ exists, we define*

$$W^u_0(z) = B_{\delta_0}(z) \cap \{(w_1 + s(p^{(1)}(w), 0),\ w_2 + \alpha(w)\, s(p^{(2)}(w), 0))\}_{\substack{w \in \tilde{W}^u(z) \\ s \in [-1,1]}},$$

*where*

$$\alpha(w) = -\frac{\langle p^{(1)}(w), p^{(1)}(w)\rangle}{\langle p^{(2)}(w), p^{(2)}(w)\rangle}.$$

*The corresponding definition holds for* $W_0^s(z)$.

**Lemma 5.12.** *Given* $y \in \tilde{\mathscr{U}}(x^-, \delta_0)$ *and* $y' \in \tilde{\mathscr{U}}(x^+, \delta_0)$, *the following facts hold:*

i) $W_0^u(y)$ *is a smooth* $2n-1$ *dimensional manifold (if* $\tilde{W}^u(y)$ *exists).*

ii) $W_0^u(y)$ *is in the stable direction and satisfies the same perpendicularity conditions, with respect to the flow, as does the unstable manifold* $W^u(y)$.

iii) $W_0^u(y)$ *exists, unless* $y$ *belongs to a subset of* $\tilde{\mathscr{U}}(x^-, \delta_0)$ *of measure zero.*

iv) $W_0^u(y) \cap \Delta_-^{(1)}(k_-, 0) = \emptyset$ *unless* $y$ *belongs to a set of measure zero.*

v) *The properties corresponding to* (i)–(iv) *hold for* $W_0^s$.

vi) *If the manifolds* $\phi^{-j_-}(W_0^u(y))$, $\phi^{-j_+}(W_0^s(y'))$ *intersect, then they intersect transversally (provided* $k_- \neq k_+$ *and the intersection point does not belong to a set of measure zero).*

*Proof.* Property i) follows from the regularity of the line bundle that generates $W_0^u$.

The tangent space $\mathscr{T}_w W_0^u(y)$ at a point $w = (w_1, w_2) \in W_0^u(y)$ consists of vectors of the form $(\delta q, \delta p) = (\delta q^{(1)}, \delta q^{(2)}, \delta p^{(1)}, \delta p^{(2)}) + \lambda(v(w,k), 0, 0)$. Here $(\delta q^{(1)}, \delta p^{(1)}) \in \mathscr{T}_{w_1} W_{k_-}^u$, $(\delta q^{(2)}, \delta p^{(2)}) \in \mathscr{T}_{w_2} W_{n-k_-}^u$ and $\lambda \in \mathbb{R}$. Furthermore, setting $w = (q, p)$, we defined

$$v(w,k)_i = \begin{cases} p_i & \text{for } i \leqq k \\ \alpha(w)\, p_i & \text{for } i > k \end{cases}$$

($\alpha(w)$ is determined in Definition 5.11). Since

$$\langle \delta q, p\rangle = \lambda(\|p^{(1)}\|^2 + \alpha\,\|p^{(2)}\|^2) = 0$$

and

$$\langle \delta q, \delta p\rangle = \langle \delta q^{(1)}, \delta p^{(1)}\rangle + \langle \delta q^{(2)}, \delta p^{(2)}\rangle \geqq 0$$

we have that $\mathscr{T}_w W_0^u(y)$ is contained in the unstable cone [note that if one of the previous vectors lies on the boundary of the cone, then it must be of the form $\lambda(v(w,k_-), 0, 0)$].

The third statement of the lemma follows then by noticing that the stable and unstable manifolds of the sub-systems exist almost everywhere and that the construction of Definition 5.11 can be carried out at every point, apart from the manifold of codimension $\geqq 2$ defined by $p^{(2)} = 0$.

A simple variation of Lemma 3.6 (see Remark 3.7) shows that, in our hypotheses, for almost all points $y \in \tilde{\mathscr{U}}(x^-, \delta_0)$, $t_\varepsilon^-(y) < \infty$ for each $\varepsilon \in (0, R)$. The proof of iv) can then be concluded with an argument completely similar to the one used in the proof of Lemma 5.10. To see this we notice that, letting $z = (q^{(1)}, q^{(2)}, p^{(1)}, p^{(2)})$ and $\phi^t(z) = (\tilde{q}^{(1)}, \tilde{q}^{(2)}, \tilde{p}^{(1)}, \tilde{p}^{(2)})$, we get the following. If particles $k_-$ and $k_- + 1$ do not collide, then $\phi^t(z + s(p^{(1)}, \beta p^{(2)}, 0, 0)) = \phi^t(z) + s(\tilde{p}^{(1)}, \beta\tilde{p}^{(2)}, 0, 0)$ for each $\beta \in \mathbb{R}$.

Property v) is easily verified with the same argument as that used above. This leaves us with the only task being to prove (vi).

Let $w \in \phi^{-j_-}(W_0^u(y)) \cap \phi^{-j_+}(W_0^s(y'))$. We recall that $\phi^{-j_-}(\tilde{W}^u(y))$ and $\phi^{-j_+}(\tilde{W}^s(y'))$ are strictly contained in complementary cones. On the contrary, $\phi^{-j_-}(W_0^u(y))$ and $\phi^{-j_+}(W_0^s(y'))$, although contained in complementary cones, may intersect the boundaries of their cones and therefore fail to be transversal. Thus, we need only to check that the images of the added directions are transversal. This last property is certainly true, out of a manifold of codimension one, because of the condition $k_+ \neq k_-$. To verify this statement we consider two different possibilities. First, there is the case in which neither the balls $k_-$, $k_- + 1$ nor the balls $k_+$, $k_+ + 1$ collide for $t \in [j_-, j_+]$; a direct check shows then that the two added directions are linearly independent. Second, there is the case in which at least one of the two abovementioned collisions occurs for $j \in [j_-, j_+]$; suppose it is the $k_+$, $k_+ + 1$ collision. According to Lemma 2.2 the added direction becomes sufficient after the collision, out of a codimension one manifold (defined by the condition that the two colliding particles have parallel velocities). Since the union of such manifolds forms a set of measure zero, this ends the proof. □

To conclude the argument we must provide a covering to which we can apply Theorem 4.3. A covering with the necessary properties can be constructed by starting with two transversal $2n-1$ dimensional foliations: the $2n-1$ dimensional sides of the elements of the covering can be chosen from those foliations (one can convince oneself that this is the case thinking of $2n-1$ dimensional linear subspaces of $\mathbb{R}^{4n-2}$). From our hypotheses and point (a) of Theorem 4.3 it follows that there exist subspaces $E_{k_-}^u$, $E_{n-k_-}^u$, $E_{k_+}^s$ and $E_{n-k_+}^s$ defined by

$$\begin{aligned}
E_{k_-}^u &= \bigcap_{t>0} d\phi_{k_-}^t (C^{(k_-)}(\phi_{k_-}^{-t} x_1^-)),\\
E_{n-k_-}^u &= \bigcap_{t>0} d\phi_{n-k_-}^t (C^{(n-k_-)}(\phi_{n-k_-}^{-t} x_2^-)),\\
E_{k_+}^s &= \bigcap_{t>0} d\phi_{k_+}^{-t} (C_-^{(k_+)}(\phi_{k_+}^{t} x_1^+)),\\
E_{n-k_+}^s &= \bigcap_{t>0} d\phi_{n-k_+}^{-t} (C_-^{(n-k_+)}(\phi_{n-k_+}^{t} x_2^+)),
\end{aligned}$$

where $C^{(j)}$, $C_-^{(j)}$ stand for the unstable and stable cone for the system of $j$ particles.

From now on we will identify any tangent spaces with a standard Euclidean space of the corresponding dimensionality. Accordingly, for $y = (y_1, y_2) \in \hat{\mathscr{U}}(x^+, \delta_0)$ and $y' = (y_1', y_2') \in \hat{\mathscr{U}}(x^-, \delta_0)$, and for $\delta_2$ sufficiently small, we define the manifolds

$$\begin{aligned}
\tilde{E}_s(y) &= \{l_w(t) \mid w \in y + E_{k_+}^s \times E_{n-k_+}^s,\ t \in [-\delta_2, \delta_2]\},\\
\tilde{E}_u(y') &= \{l_w(t) \mid w \in y' + E_{k_-}^u \times E_{n-k_-}^u,\ t \in [-\delta_2, \delta_2]\},
\end{aligned}$$

and, through them, the $2n-1$ dimensional foliations $\mathscr{F}_+ = \{\tilde{E}^s(y)\}_{y \in \hat{\mathscr{U}}(x^+, \delta_0)}$ and $\mathscr{F}_- = \{\tilde{E}^u(y')\}_{y' \in \hat{\mathscr{U}}(x^-, \delta_0)}$. Finally, we are able to specify the coverings to which we apply Theorem 4.3. The covering of $\hat{\mathscr{U}}(x^+, \delta_0)$, $\mathscr{G}_\delta^+$, is constructed by using the foliations $\mathscr{F}_+$ and $\phi^{j_+ - j_-}\mathscr{F}_-$ while the covering of $\hat{\mathscr{U}}(x^-, \delta_0)$, $\mathscr{G}_\delta^-$, is constructed by using the foliations $\mathscr{F}_-$ and $\phi^{j_- - j_+}\mathscr{F}_+$. After thinking a while it appears clear that the two coverings can be chosen so that $\phi^{j_-}\mathscr{G}_\delta^- = \phi^{j_+}\mathscr{G}_\delta^+ = \mathscr{G}_\delta$, and $\{\mathscr{G}_\delta\}$ is a family of coverings of $\hat{\mathscr{U}}(x, \delta_0)$ that satisfy conditions b1) and b2) of Theorem 4.3.

In addition, the intersections

$$\mathscr{G}_\delta^-(k_-) = \mathscr{G}_\delta^- \cap \hat{\mathscr{U}}_{k_-}(x^-, \delta_2),$$
$$\mathscr{G}_\delta^-(n-k_-) = \mathscr{G}_\delta^- \cap \hat{\mathscr{U}}_{n-k_-}(x^-, \delta_2),$$
$$\mathscr{G}_\delta^+(k_+) = \mathscr{G}_\delta^+ \cap \hat{\mathscr{U}}_{k_+}(x^+, \delta_2),$$

and

$$\mathscr{G}_\delta^+(n-k_+) = \mathscr{G}_\delta^+ \cap \hat{\mathscr{U}}_{n-k_+}(x^+, \delta_2)$$

form coverings of neighborhoods $\hat{\mathscr{U}}_{k_-}(x^-, \delta_2)$, $\hat{\mathscr{U}}_{n-k_-}(x^-, \delta_2)$, $\hat{\mathscr{U}}_{k_+}(x^+, \delta_2)$ and $\hat{\mathscr{U}}_{n-k_+}(x^+, \delta_2)$, respectively, with the properties required to apply Theorem 4.3. This means that any element $G \in \mathscr{G}_\delta^\pm(\cdot)$, apart from a set of total measure $o(\delta)$, has an abundance of $\delta$-long $\tilde{W}^u$- and $\tilde{W}^s$-manifolds (hence of $W_0^u$- and $W_0^s$-manifolds) near its sides (see assertion b) of Theorem 4.3).

Moreover, according to Remark 5.8, the construction obtained through Theorem 4.3 in the above-mentioned neighborhoods is completely consistent with the same construction carried out in the neighborhoods of $l_{x^-}^{(1)}(s)$, $l_{x^-}^{(2)}(s)$, $l_{x^+}^{(1)}(s)$, and $l_{x^+}^{(2)}(s)$, $s \in [-\delta_2, \delta_2]$, so that, if the intersection of $G$ with the neighborhoods of $x^-$ and $x^+$ has an abundance of stable and unstable manifolds then, all the elements $G$ along the curves $l_{x^-}$ and $l_{x^+}$ contain a large measure of manifolds $W_0^u$, $W_0^s$. We consider the set

$$\begin{aligned}\hat{\mathscr{U}}_r(x, \delta_0) = \{y \in \hat{\mathscr{U}}(x, \delta_0) \backslash (\varDelta_+^{(1)}(k_+, j_+) \cap \varDelta_-^{(1)}(k_-, j_-) \mid \\ W_0^u(y) \text{ or } W_0^s(y) \text{ exist and their intersection with} \\ \varDelta_+^{(1)}(k_+, j_+) \cap \varDelta_-^{(1)}(k_-, j_-) \text{ is empty}\}.\end{aligned}$$

We know already that $\mu(\hat{\mathscr{U}}_r(x, \delta_0)) = \mu(\hat{\mathscr{U}}(x, \delta_0))$ (it follows from the equality $\mu(\mathscr{M} \backslash \tilde{\Omega}) = 0$ and Lemma 5.12). Now, given any two points $w$, $w'$ in $\hat{\mathscr{U}}_r(x, \delta_0)$, we can connect neighborhoods of $w$ and $w'$ by chains of $W_0^u$- and $W_0^s$-manifolds, with intersections out of the set of zero measure where the transversality of the manifolds may fail (see Lemma 5.12(vi)). Indeed, this can be done by using a simplified version of the argument sketched in the proof of Theorem 4.6.

As anticipated, this shows that the set $\hat{\mathscr{U}}(x, \delta_0) \backslash \hat{\mathscr{U}}_r(x, \delta_0)$ which contains the intersection $\varDelta_+^{(1)}(k_+, j_+) \cap \varDelta_-^{(1)}(k_-, j_-) \cap \hat{\mathscr{U}}(x, \delta_0)$ has codimension two. By using the flow direction we obtain codimension two in $\mathscr{U}(x)$.

The result is proven so far for $k_+ \neq k_-$. If $k_+ = k_- = k$, there are two possibilities. In the first case there is a $k$, $k+1$ collision for $t \in (j_-, j_+)$. When this possibility takes place it is easy to see that assertion (vi) of Lemma 5.12 holds and the argument can be completed as before. In the second case the $k$, $k+1$ collision never takes place; this case corresponds to $x \in \varDelta_+^{(1)}(k, 0) \cap \varDelta_-^{(1)}(k, 0)$ for some $k \in \{1, \ldots, n\}$. To discuss this last possibility it is useful to introduce again the explicit dependency from the energy. Let $E_1$ be the energy of the first $k$ particles and $E$ the total energy of the system. In this case, the manifolds $\tilde{W}^u$, $\tilde{W}^s$ are not in a generic position with respect to each other: they both belong to $\mathscr{M}_{k, E_1} \times \mathscr{M}_{n-k, E-E_1}$. This means that it is possible to use, for example, chains of manifolds $W_0^u$, $\tilde{W}^s$ in order to carry out our argument and prove that the intersection $(\varDelta_+^{(1)}(k, 0) \cap \varDelta_-^{(1)}(k, 0)) \cap \mathscr{U}_{E_1}(x)$, where $\mathscr{U}_{E_1}(x) = \mathscr{U}(x) \cap (\mathscr{M}_{k, E_1} \times \mathscr{M}_{n-k, E-E_1})$ has codimension two. The desired result is then proven in the same fashion as in the case of Lemma 5.4.

*Singularity and Non-Interacting Sub-Systems.* The statement that the corresponding set is of codimension two, although of a purely topological nature, is a consequence of the Sinai-Chernov Ansatz that will be proven in Theorem 5.14. Indeed, here we are assuming the Ansatz to prove the desired result.

Let us concentrate on a point $x$ whose trajectory experiences a singular collision in the past and separates into non-interacting sub-systems in the future, the other possibility being left to the reader. If $\mathscr{U}(x)$ is a sufficiently small neighborhood of $x$, then we call $\tilde{\mathscr{R}}$ the image in $\mathscr{U}(x)$ of the singularity manifold that intersects the trajectory of $x$. It then follows that $\tilde{\mathscr{R}}$ divides the neighborhood into two disjoint connected components.

The sets we are considering are $\tilde{\mathscr{R}} \cap \Delta_+(k,j)$ where $k \in \{1, \ldots, n\}$ and $j > 0$ and our goal is to show that $\mathscr{U}(x) \backslash (\tilde{\mathscr{R}} \cap \Delta_+(k,j))$ is connected. We note that $\mathscr{U}(x) \cap \tilde{\mathscr{R}}$ has positive measure with respect to the induced measure on $\tilde{\mathscr{R}}$. Accordingly, by the Sinai-Chernov Ansatz, to almost all the points $y \in \mathscr{U}(x) \cap \tilde{\mathscr{R}}$ we can apply Theorem 4.3.

Remembering that $\tilde{\mathscr{R}}$ is transversal to the foliations by the local manifolds $W_0^s$, we obtain a family of local manifolds $W_0^s(y)$, of positive total measure, which cross $\tilde{\mathscr{R}}$ and therefore connect the two components of $\mathscr{U}(x)$. This concludes the proof, since almost all the manifolds $W_0^s$ belong to the complement of $\Delta_+(k,j)$ by properties (iv)–(v) stated in Lemma 5.12. □

We still need to prove that the Sinai-Chernov Ansatz holds for the $n$-ball system. We will obtain this last result by studying a stronger property.

**Definition 5.13** (Generalized Ansatz). *For each $\mathscr{C}^2$-manifold $W \subset \mathscr{M}$ in the stable (unstable) direction (see Definition I.1 in Appendix I) the following equality holds:*

$$\mu_W(W \cap \{\textit{sufficient points}\}) = \mu_W(W),$$

*where $\mu_W$ is the Lebesgue measure restricted to $W$.*

The property stated in Definition 5.13 implies the Sinai-Chernov Ansatz since, apart from manifolds of codimension three, the singularity set $\mathscr{R}$ is composed of the finite union of smooth manifolds (see Appendix I) and since, for semi-dispersing billiards, a sufficient point with a smooth trajectory will automatically have the unbounded property required in the Ansatz (see [S3, LW]).

The Generalized Ansatz holds for the one-ball system where every point on $W$ is sufficient (see [S, G]). We can therefore continue in our strategy and the prove the validity of the Generalized Ansatz by induction on the number of balls.

**Theorem 5.14.** *If, for any $m < n$, the Generalized Ansatz holds and the set $\mathscr{M} \backslash \tilde{\Omega}$ is of measure zero and codimension two for the system of m balls, then the Generalized Ansatz holds also for the systems of n balls.*

*Proof.* We discuss only manifolds in the stable direction; the other possibility can be treated by exactly the same arguments and is left to the reader. For each $z \in W$ we will prove that there exists $\delta > 0$ such that

$$\mu_W(W \cap (-\Omega) \cap B_\delta(z)) = \mu_W(B_\delta(z) \cap W).$$

As explained in Sect. 3 and Definition 4.2, $-\Omega$ consists of points that are sufficient in the past so that the above statement implies Theorem 5.14. Choosing $\delta$ small enough we can suppose, without any loss of generality, that there exists a $\mathscr{C}^2$-function $\tilde{W}$: $B_\delta(z) \to \mathbb{R}$ such that $W_\delta(z) = W \cap B_\delta(z) = \{w \in B_\delta(z) \subset \mathscr{M} \mid \tilde{W}(w) = 0\}$.

The first thing to notice is that the points in $W_\delta(z)$ that have a singularity in the past have $\mu_W$-measure zero. In fact, a point has a non-smooth trajectory in the past only if it belongs to the image of some singularity manifold. But those images are all in the unstable direction (see Appendix I) and therefore transversal to $W$. Consequently, the set of non-smooth points, in the past, belonging to $W$ is contained in a countable union of smooth sub-manifolds of codimension one (in $W$) and hence is of $\mu_W$-measure zero. Accordingly, the set $W \cap (\mathscr{M} \backslash (-\Omega))$ is contained, apart from a set of $\mu_W$-measure zero, in the union

$$\Delta_-(W) = \bigcup_{\substack{k \in \{1, \ldots, n\} \\ j \in \mathbb{N}}} (\Delta_-(k, -j) \cap W).$$

For a definition of the sets $\Delta_-(k, -j)$ see the beginning of "Non-interacting sub-systems" in Theorem 5.2. The lemma is then equivalent to

$$\mu_W(\Delta_-(k, -j) \cap W_\delta(z)) = 0 \qquad \forall k \in \{1, \ldots, n-1\} \text{ and } j \in \mathbb{N}. \tag{5.2}$$

Moreover, since $T^{-j}W$ is a finite collection of smooth manifolds in the stable direction, it suffices to discuss the case $j = 0$.

The proof will be by contradiction: we suppose that $\mu_W(\Delta_-(k,0) \cap W_\delta(z)) > 0$, for some $k$, and we will derive a contradiction.

We use, in part, the same notation as that used in Theorem 5.2 but, for the convenience of the reader, we again introduce most of it explicitly. The discussion at the beginning of Appendix I (Lemma I.3) shows that $W$ in the stable direction is equivalent to $\nabla \tilde{W}$ in the unstable cone. Given $w \in W_\delta(z)$, we write $w = (w_1, w_2)$, $w_i = (q^{(i)}, p^{(i)})$, where $w_1 \in \mathscr{M}_k$ gives the positions and the velocities of the first $k$ balls and $w_2 \in \mathscr{M}_{n-k}$ of the last $n-k$; as before we call $E(w)$ the energy of the point $w$. Analogously, for each $w \in W_\delta(z)$ we will set $\nabla_w \tilde{W} = \pi(w) = (\pi_1(w), \pi_2(w))$ with $\pi_i(w) = (\xi_i(w), \eta_i(w))$, $\pi_1(w) \in \mathscr{M}_k$ and $\pi_2(w) \in \mathscr{M}_{n-k}$.

We start by defining

$$W_* = \{w \in W_\delta(z) \mid \langle \pi(w), (p^{(1)}(w), \alpha(w) p^{(2)}(w), 0, 0)\rangle = 0\},$$

where $p^{(1)}$, $p^{(2)}$ and $\alpha$ are defined in Definition 5.11 (where one chooses $k_- = k$). Essentially, $W_*$ is the part of $W_\delta(z)$ that contains the neutral direction for the dynamics of points belonging to $\Delta_-(k,0)$. Our first claim is the following:

$$\mu_W((W_\delta(z) \backslash W_*) \cap \Delta_-(k,0)) = 0.$$

Suppose it is not true. It is then possible to construct the set

$$B = \{w + s(p^{(1)}(w), \alpha(w) p^{(2)}(w), 0, 0) \mid w \in (W_\delta(z) \backslash W_*) \cap \Delta_-(k,0),\; s \in [-\delta_1, \delta_1]\}.$$

Clearly $\mu(B) > 0$, but the points in $B$, by construction, have the property that, under that product dynamics $\phi_k^t \times \phi_{n-k}^t$, the balls $k$, $k+1$ never get closer than $2R - 2\delta_1$. Since $\delta_1$ can be arbitrarily small, Lemma 3.6 (see also Remark 3.7) implies that $\mu(B) = 0$, in contradiction with our assumption.

We are left with the possibility

$$\mu_W(W_* \cap \Delta_-(k, 0)) > 0.$$

We will show that this is impossible by using an argument similar to the one just employed; the only difference will be in the construction of the transversal fibers. Using the previous notations the property of being in the stable direction reads:

$$\langle \xi_1(w), \eta_1(w) \rangle \geqq -\langle \xi_2(w), \eta_2(w) \rangle. \tag{5.3}$$

It is therefore natural to define $W_1 = \{w \in W_* \mid \langle \xi_1(w), \eta_1(w) \rangle \geqq 0\}$ and $W_2 = \{w \in W_* \mid \langle \xi_2(w), \eta_2(w) \rangle \geqq 0\}$; it then follows from (5.3) that $W_1 \cup W_2 = W_*$. In addition, since the manifold $W$ is of class $\mathscr{C}^2$, so that $w \mapsto \pi(w)$ is a $\mathscr{C}^1$-function, the boundaries of $W_i$ are $\mathscr{C}^1$-manifolds of codimension one in $W$. Therefore, calling $W_i^0 = \operatorname{int} W_i$, we have

$$\mu_W(W_* \setminus (W_1^0 \cup W_2^0)) = 0.$$

Note that this is the only place in which we use the $\mathscr{C}^2$-smoothness of $W$. In the following we will assume that $z \in W_1^0$ and $\delta$ is so small that $W_* = W_1^0$; the other possibilities are completely analogous.

Let $\mathscr{U}(z)$ be a sufficiently small neighborhood of $z_2$ in $\mathbb{R}^{4(n-k)}$. For each $w_2 \in \mathscr{U}(z)$ we define $W_z^{(1)}(w_2) = \{w_1 \in \mathscr{M}_{k, E(z) - E(w_2)} \mid w = (w_1, w_2) \in \tilde{W}_\delta(z)\}$. From the previous discussion it follows that $W_z^{(1)}(w_2)$ is contained in $\mathscr{M}_{k, E(z) - E(w_2)}$ and it is a smooth, codimension two manifold in the stable direction. Indeed, if $\delta v = (\delta v_1, \delta v_2) \in \mathscr{T}_w \mathscr{M}$ with $\delta v_i = (\delta q^{(i)}, \delta p^{(i)})$, then the property $\delta v \in \mathscr{T}_w W_1^0$ implies

$$\begin{aligned} \langle (p, 0), \delta v \rangle &= 0, \\ \langle (0, p), \delta v \rangle &= 0, \\ \langle \pi(w), \delta v \rangle &= 0. \end{aligned} \tag{5.4}$$

Moreover, if $\delta v_1 \in \mathscr{T}_{w_1} W_z^{(1)}(w_2)$, then by definition $(\delta v_1, 0) \in \mathscr{T}_w W_1^0$. Accordingly, (5.4) becomes

$$\begin{aligned} \langle (p_1, 0), \delta v_1 \rangle &= 0, \\ \langle (0, p_1), \delta v_1 \rangle &= 0, \\ \langle \pi_1(w), \delta v_1 \rangle &= 0. \end{aligned} \tag{5.5}$$

That is, $W_z^{(1)}(w_2)$ is a manifold in $\mathscr{M}_{k, E(z) - E(w_2)}$ perpendicular to the flow direction and to $\pi_1(w)$. Next, we want to check that this manifold is in the stable direction. Since $\pi_1(w)$ it is not necessarily in $\mathscr{M}_{k, E(z) - E(w_2)}$ we need to find another perpendicular vector to $W_z^{(1)}(w_2)$; to this end we define $\pi_1^*(w) = (\xi_1^*, \eta_1^*) = \pi_1(w) + \lambda(p_1, 0) + \sigma(0, \pi_1(w))$ where $\lambda$ and $\sigma$ are chosen so that

$$\begin{aligned} \langle \pi_1^*(w), (p_1, 0) \rangle &= 0, \\ \langle \pi_1^*(w), (0, p_1) \rangle &= 0. \end{aligned} \tag{5.6}$$

In fact, remembering that $w \in W_*$ and that $\pi(w)$ is perpendicular to the flow direction, it follows that $\lambda = 0$. We can then compute

$$\langle \xi_1^*, \eta_1^* \rangle = \langle \xi_1, \eta_1 \rangle + \sigma \langle \xi_1, p_1 \rangle \geqq 0, \tag{5.7}$$

where we have used (5.3) and (5.6). The inequality (5.7), as discussed in Appendix I, implies that $W_z^{(1)}(w_2)$ is in the stable direction. From the definition it is also clear that the union $\bigcup_{w_2 \in \mathscr{U}(z_2)} W_z^{(1)}(w_2)$ forms a neighborhood of $z$ in $W$. This means that, if we apply all the previous construction to a point $z \in W$ which is a density point of $\Delta_-(k,0)$, with respect to $\mu_W$, then the hypotheses at hand imply, by Fubini's theorem, the existence of a set $A(z_2) \subset \mathscr{U}(z_2)$ of positive Lebesgue measure such that

$$\mu_{W_z^{(1)}(w_2)}(W_z^{(1)}(w_2) \cap \Delta_-(k,0)) > 0 \qquad \forall w_2 \in A(z_2).$$

We are now ready to produce the contradiction.

We will use again the basic idea of constructing a set of positive measure in which the trajectory, under the dynamics defined by the product of the dynamics of the first $k$ and last $n-k$ balls (i.e. dynamics in which the balls $k$, $k+1$ can, in principle. cross each other without interacting), has the property that the centers of the balls $k$, $k+1$ are never closer that $2R$ minus some fixed, arbitrarily small, amount (almost no collision possible for the true dynamics). The proof is then concluded since the existence of such a set of positive measure is in contradiction with Lemma 3.6 (see also Remark 3.7).

Let us go ahead with the construction. By hypothesis we have that, for each $w_2 \in \mathscr{U}(z_2)$, $\mu_{W_z^{(1)}(w_2)}$-almost all points of $W_z^{(1)}(w_2)$ are sufficient. For each $w_2 \in A(z_2)$ we can choose $w_1 \in W_z^{(1)}(w_2)$ to be both a $\mu_{W_z^{(1)}(w_2)}$-density point of $\Delta_-(k,0)$ and a sufficient point for the dynamics of the first $k$ balls. We can then use Theorem 4.3 (and the comments in Appendix II) to see that there exists a set $\tilde{B}(w_2) \subset W_z^{(1)}(w_2) \cap \Delta_-(k,0)$ and a $\delta > 0$ such that $\mu_{W_z^{(1)}(w_2)}(\tilde{B}(w_2)) > 0$ and each point in $\tilde{B}(w_2)$ has an unstable manifold of size $\delta$. In fact, choosing $c_1$ close enough to one in Theorem 4.3(b) and remembering that $W_z^{(1)}(w_2)$ is in the stable direction (that is, transversal to the unstable manifolds), it follows that $\mu_{W_z^{(1)}(w_2)}(W_z^{(1)}(w_2) \cap \mathscr{G}_\delta^{(1)}) = o(\delta)$. Choosing $\delta$ small enough we can then construct the set $B(w_2) = \bigcup_{w \in \tilde{B}(w_2)} W^u_{k, E(z)-E(w_2)}(w)$. Next, by the absolute continuity of the unstable foliation (see [KS]), we have $\mu_{k, E(z)-E(w_2)}(B(w_2)) > 0$ where $\mu_{k, E(z)-E(w_2)}$ is the Lebesgue measure restricted to $\mathscr{M}_{k, E(z)-E(w_2)}$.

Moreover, since $\tilde{B}(w_2)$ belongs to $\Delta_-(k,0)$ and the configuration space size of the unstable manifolds cannot expand in the past (see Lemma 2.3 and Lemma 5.10), for each point $w_1 \in B(w_2)$ we have that the trajectory of the point $(w_1, w_2)$ has the desired behaviour in the past. More precisely, if we consider the product dynamics $\phi_k^t \times \phi_{n-k}^t$ on $\mathscr{M}_{k, E(z)-E(w_2)} \times \mathscr{M}_{n-k, E(w_2)}$, then the trajectory $(\phi_k^t(w_1), \phi_{n-k}^t(w_2))$, $t < 0$, has the property that the particles $k$, $k+1$ are never closer than $2R - \delta$. Finally, we construct the set $B_0(z) = \bigcup_{w_2 \in A(z_2)} B(w_2)$; by Fubini's Theorem, $\mu(B_0(z)) > 0$ and this set has the same property as the sets $B(w_2)$. But this is in contradiction to Lemma 3.6 (since our hypotheses and Theorem 5.1 ensure the necessary ergodic properties for the systems of $m < n$ balls). $\square$

*Review*. Since the argument explained here may seem intricate, let us review the logic of our induction. From the literature is known that the one-ball system has no non-sufficient points and it is then mixing (it is actually a $K$-system). Next, we suppose that, for any $m < n$, the set of non-sufficient points for the system of $m$ balls is of measure zero and codimension two and the Generalized Ansatz (see Definition 5.13) holds for this system. This allows us to use Theorem 5.1 (based on Theorems 4.3, 4.6) to prove the ergodicity and mixing property of the systems of $m$ balls. Lemma 3.6 then implies that the set of non-sufficient points for the $n$ balls system is of measure zero. A more careful analysis, carried out in Theorem 5.2, shows that the set of non-sufficient points for the $n$ balls system has codimension two. Finally, Theorem 5.14 shows that the Generalized Ansatz holds for the system of $n$ balls, therefore closing the induction.

## 6. Other Boxes (Periodic Lorentz Gas)

In this section we discuss other examples that can be treated with the same technique as developed in the previous sections.

We start by pointing out the parts of our argument that are model-dependent.

In Sect. 2 we made explicit use of the fact that the curvature $K$ of the boundary $\partial Q$ of the billiard table in which the motion of the balls takes place, is strictly positive. This is an essential feature in our discussion: a flat boundary would not provide any hyperbolicity, thus making much more difficult the discussion of the sufficiency of the trajectories. In general it could be possible to allow flat pieces of $\partial Q$, and even special convex ones (see [Bu3]), provided that the invariance of the cone structure is preserved and that between two ball-ball collisions involving any given particle $k$ there are enough collisions with $\partial Q$ to ensure that all the non-sufficient vectors satisfy $\delta q_k = \lambda p_k$ (the equivalent of Lemma 2.1).

In Sect. 3 other model-dependent hypotheses were used in an essential way: the fact that the geometry of the boundary $\partial Q$ constrains the region which each ball can explore and the property that, when two neighboring balls collide, the next collision involving one of them will be with the boundary. The previous two conditions are all that is needed to prove Lemma 3.3, while Lemma 3.6 is a modification of the weak lemma on avoiding balls from [KSS1], which holds in a more general situation.

Section 4 contains result that are model independent. In particular Theorem 4.3, as stated, applies to any semi-dispersing billiard. Moreover, after adding some hypothesis (see [LW]), Theorem 4.3 holds for a quite general class of symplectic maps.

Section 5 contains the proof that the hypotheses of Theorem 4.3 are in general satisfied and that the set where this theorem cannot be applied does not separate the phase space. The proof of hypothesis i) is referred to in Appendix I which contains quite general arguments that hold for any semi-dispersing billiard. Hypothesis ii) is proven at the end of Sect. 5 by a reasoning which relies on quite general arguments. The analysis of the set of the non-sufficient points is contained in Theorem 5.2. Looking at the proof it is clear that it can be generalized to show the following. Given a system of $n$

balls, for which all the $m < n$ sub-systems are mixing, the set of points with trajectories along which the system splits into two independent sub-systems, both in the past and in the future, is of measure zero and codimension two. In fact, the mixing condition can be further weakened: the ergodicity of the direct product of any two sub-dynamics suffices. In Sect. 3 we have proven that the points with the above-mentioned property are sufficient. Obviously, since Sect. 3 is model-dependent, Theorem 5.2 implies sufficiency on a large set only for the billiards to which Sect. 3 applies.

In conclusion, given a system of $n$ balls in a region $Q$ with dispersing boundary ($K > 0$), the only requirements needed in order to aply our techniques are the ones on which depends Sect. 3. From this consideration it is clear that a generlization of Sect. 3 is the missing ingredient to treat the general case.

Of course, the example introduced at the end of Sect. 1 is by no means the unique one which possesses the properties that we require. We shall describe three different models that share such properties and, therefore, can be studied using our strategy.

The first model is a straight channel containing obstacles. There is one ball between each pair of obstacles. The size and shape of the obstacles confine the balls; yet, at the sometime, allow them to collide against each other. In addition, the narrow regions between the flat and the curved boundary, where the particles can interact with each other, are shaped in a way that prevents two consecutive collisions between particles – see Fig. 3.

The second model is a cylinder surface with two types of obstacles. The obstacles that intersect the dashed lines have the same function that the obstacles in the previous model, while the internal ones prevent consecutive collisions between particles; see Fig. 4. Note that this last example can be lifted from the cylinder to the plane providing a two-dimensional array of scatterers.

The third example is an even simpler two-dimensional array of scatterers. It can be constructed starting from a triangular lattice on the plane and putting a scatterers around any vertex (shadowed discs in Fig. 5). If the radius of the scatterers and the radius of the moving balls are chosen properly, then each ball is confined among three neighboring scatterers. In addition, the geometry of the lattice is such that it is impossible to have two consecutive ball-ball collisions without having a collision with the boundary in between. The last model is in fact a periodic Lorentz gas where instead of point particles we have hard spheres. As was noted before, the condition that no ball can have two consecutive collisions with the other balls without some intermediate collision(s) with the scatterers corresponds, in a sense, to the finiteness of the horizon in the Lorentz gas [BS, Bu1]. In general, any periodic array of obstacles that enjoy the above-mentioned properties will provide a model to which the techniques described in this paper apply. It would be interesting to investigate kinetic properties of these models, e.g. diffusion of the energy.

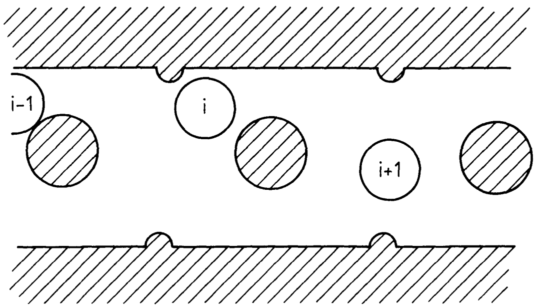


**Fig. 3.** Straight channel with obstacles

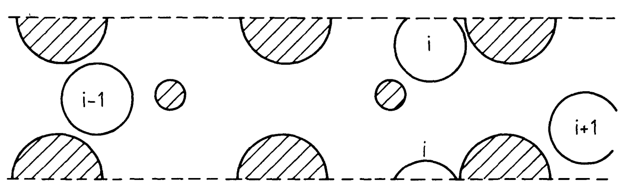


**Fig. 4.** Cylindrical table with obstacles

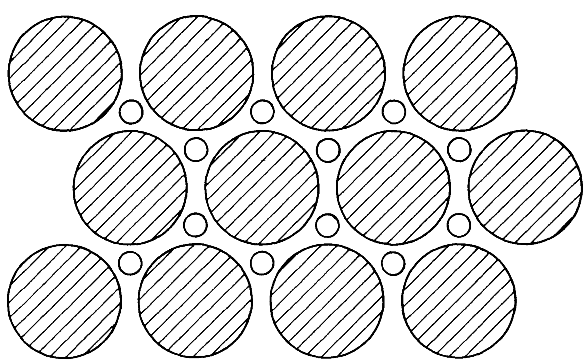

**Fig. 5.** Two dimensional array of scatterers

## Appendix I (Transversality)

Let us start with some general facts. Recall that the dimensionality of the phase space $\mathscr{M}$ of our system is $4n-1$.

**Definition I.1.** *We say that a sub-manifold in $\mathscr{M}$ is (strictly) in the stable (unstable) direction if its tangent space is perpendicular to the flow direction and contains a $2n-1$ dimensional subspace belonging (strictly) to the stable (unstable) cone.*

Notice that, since the stable and unstable cone are not disjoint, a manifold can be both in the stable and in the unstable direction; however, if a manifold is strictly in one of the two directions, it cannot be simultaneously in the other. This is emphasized by the following lemma.

**Lemma I.2.** *If a manifold $\Sigma_1$, strictly in the stable direction, intersects a manifold $\Sigma_2$, in the unstable direction, then the intersection is transversal.*

*Proof.* We will show that, for any $z \in \Sigma_1 \cap \Sigma_2$, $\dim(\mathscr{T}_z\Sigma_1 \oplus \mathscr{T}_z\Sigma_2) = \dim(\mathscr{M})-1$, where the $-1$ is a consequence of the fact that both manifolds do not contain the flow direction. To see this, let $A_i \subset \mathscr{T}_z\Sigma_i$ be $2n-1$ subspace such that $A_1 \subset C_-(z)$, strictly and $A_2 \subset C(z)$. We will check that $\dim(A_1 \oplus A_2) = \dim(A_1)+\dim(A_2)$. The only manner in which the previous statement can fail is if there exist linearly dependent vectors $\xi_1 = (\delta q_1, \delta p_1) \in A_1$ and $\xi_2 = (\delta q_2, \delta p_2) \in A_2$ (let us say, $\xi_1 + \xi_2 = 0$). But, given that $\langle \delta q_1, \delta p_1 \rangle < 0$ and $\langle \delta q_2, \delta p_2 \rangle \geqq 0$ for each $\xi_1, \xi_2 \neq 0$, we would have then that $\xi_1 = \xi_2 = 0$. □

In view of Lemma I.2, it is useful to have a simple criterion to check if a manifold is in the stable or in the unstable direction. To this end, let us consider a codimension two manifold $\Sigma \subset \mathscr{M}$ whose tangent space is perpendicular to the flow direction (as already noticed, this corresponds to codimension one in the Poincaré section – see Appendix II). We will study its tangent space $\mathscr{T}_z\Sigma$ at a point $z = (q,p) \in \Sigma$ (for simplicity, the lower index $z$ is omitted wherever possible). Observe that $\mathscr{T}\Sigma$ is a $4n-3$ dimensional space. Since $\mathscr{T}\Sigma$ is perpendicular to $(p,0)$, in $\mathscr{T}\mathscr{M}$ there is only one vector $\eta = (\eta_q, \eta_p)$ perpendicular to both $(p,0)$ and $\mathscr{T}\Sigma$. We restrict ourselves to $2n-1$ dimensional subspaces $W \subset \mathscr{T}\Sigma$, with the property that for each $\delta q$, $\langle \delta q, p \rangle = 0$ there exists $\delta p$ such that $(\delta q, \delta p) \in W$. The vectors of such a subspace can be represented as $(\delta q, A\delta q)$, where $A$ is a $2n \times 2n$ matrix such that $\langle A\delta q, p \rangle = 0$ for each $\delta q$ (the case of a more general subspace can be treated similarly). For these vectors we have

$$\langle \eta, (\delta q, A\delta q) \rangle = 0$$

or, equivalently,

$$\langle \eta_q + A^T \eta_p, \delta q \rangle = 0.$$

The last equation implies $\eta_q + A^T \eta_p = \lambda p$ for some $\lambda \in \mathbb{R}$.

**Lemma I.3.** *A codimension two manifold in $\mathscr{M}$, whose tangent space is perpendicular to the flow direction, is in the stable (unstable) direction* iff *its normal vector lies in the unstable (stable) cone.*

*Proof.* We start by proving the necessity of our condition. Let $\{(\delta q, A\delta q)\} \subset \mathscr{T}\Sigma$ belong to the unstable cone. Then

$$\langle \delta q, A\delta q\rangle \geqq 0$$

for each $\delta q$ such that $\langle \delta q, p\rangle = 0$. According to our previous discussion, and remembering that $\eta \in \mathscr{T}\mathscr{M}$ (i.e. $\langle \eta_p, p\rangle = 0$), we have

$$\langle \eta_p, \eta_q\rangle = -\langle \eta_p, A^T\eta_p\rangle + \lambda\langle \eta_p, p\rangle = -\langle \eta_p, A\eta_p\rangle \leqq 0. \tag{I.1}$$

To prove that (I.1) is a sufficient condition, consider first the case $\langle \eta_q, \eta_p\rangle \neq 0$. We can then define the symmetric matrix

$$Av = -(\langle \eta_q, \eta_p\rangle)^{-1}\langle \eta_q, v\rangle\eta_q.$$

Clearly, $A$ satisfies $\langle A\delta q, p\rangle = 0$ and the space $\{(\delta q, A\delta q)\}$ belongs to $\mathscr{T}\Sigma$. Indeed,

$$\langle(\eta_q, \eta_p), (\delta q, A\delta q)\rangle = 0.$$

Moreover,

$$\langle \delta q, A\delta q\rangle = -\frac{\langle \eta_q, \delta q\rangle^2}{\langle \eta_q, \eta_p\rangle} \geqq 0.$$

In the case $\langle \eta_q, \eta_p\rangle = 0$, consider the subspace generated by the vector $(0, \eta_q)$ and the vectors $(\delta q, 0)$, with $\langle \delta q, \eta_q\rangle = 0$. The generic vector $\xi$, in such space, will then have the form $\xi = \lambda(\delta q, 0) + \mu(0, \eta_q)$. A direct computation shows that the above subspace belongs to $\mathscr{T}\Sigma$ and lies in the boundary of the unstable cone. □

We can now use the above criterion to study the manifolds that appear in the paper.

Let us start with the codimension two manifolds that form the singularity set $\mathscr{R}^+$. By this we mean any manifold, perpendicular to the flow direction, composed of points for which one of future collision will be a singular one. For the sake of brevity, in this Appendix we denote such a manifold again by $\mathscr{R}^+$ and consider the case where the next collision is a singular one. In our system we have three types of singularities:

(1) tangent collision with $\partial Q$,
(2) tangent collision between two balls,
(3) multiple collisions.

We will consider them one by one.

(1). Suppose that the next collision for $z \in \mathscr{R}^+$ is a tangent collision of particle $k$ with the boundary $\partial Q$, and let $\eta$ be the unit normal vector to $\partial Q$ at the collision point. Given a vector $\xi \in \mathscr{T}_z\mathscr{R}^+$, we transport it along the flow direction until the collision point. Let us call $\xi = (\delta q, \delta p)$ the resulting tangent vector. Recall that the manifold of tangent collision at the collision point is defined by

$$\langle p_k, \eta\rangle = 0. \tag{I.2}$$

Suppose $\langle \delta q_k, \eta \rangle = 0$, so that $\delta q_k$ is tangent to $\partial Q$. Then differentiating (I.2) yields

$$\langle \delta p_k, \eta \rangle + K \langle p_k, \delta q_k \rangle = 0, \tag{I.3}$$

where $K > 0$ is the curvature of the boundary $\partial Q$ at the collision point. If $\delta q_k$ is perpendicular to the boundary, and therefore to $p_k$, we obtain that

$$\langle \delta p_k, \eta \rangle = \langle \delta p_k, \eta \rangle + K \langle p_k, \delta q_k \rangle = 0.$$

By taking linear combination we get that all the tangent vectors, belonging to the image of $\mathscr{T}_z \mathscr{R}^+$ just before the collision, are perpendicular to the vector $(\tilde{\eta}_q, \tilde{\eta}_p) = (0, \ldots, -Kp_k, 0, \ldots, n, 0, \ldots, 0)$; such a vector is not perpendicular to $(p, 0)$, but it can be used to construct the vector $(\eta_q, \eta_p) = (Kp_1, \ldots, Kp_{k-1}, 0, Kp_{k+1}, \ldots, Kp_n, 0, \ldots, n, 0, \ldots, 0)$ that has all the necessary properties for applying Lemma I.3. Since $\langle \eta_q, \eta_p \rangle = 0$, the manifold $\mathscr{R}^+$, at the collision point, is in the stable direction. Moreover, if we consider a manifold $\mathscr{R}^+$ before the collision (which is a pre-image of the above one), it is strictly in the stable direction. To see this, we define vectors $\{\beta_i\}_{i=1}^{2n-2}$ by

$$\begin{aligned} \langle \beta_i, \beta_j \rangle &= \delta_{ij}, \\ \langle \beta_i, p \rangle &= 0, \\ \langle \beta_i, \eta_q \rangle &= 0. \end{aligned} \tag{I.4}$$

It follows that the vectors $\{\beta_i, p, \eta_q\}$ form an orthonormal basis in $\mathbb{R}^{2n}$. Moreover, setting

$$Av = -\sum_{i=1}^{2n-2} \langle \beta_i, v \rangle \beta_i \tag{I.5}$$

the vectors $\{(\delta q, A\delta q) \mid \langle \delta q, p \rangle = 0; \langle \delta q, \eta_q \rangle = 0\}$ together with $\{(0, \eta_q)\}$ span a $2n-1$ dimensional subspace perpendicular to $\eta$. In addition, this space is in the stable direction. In fact, the only vector not strictly contained in the stable cone is $(0, \eta_q)$. The image of this vector, at any epoch before the collision is $(t\eta_q, \eta_q)$, where $t > 0$ is the time that remains until the collision. Thus $\mathscr{R}^+$ is strictly in the stable direction at any time preceding the collision.

(2). Suppose that the next collision for $z \in \mathscr{R}^+$ is a tangent collision between particles $k$ and $k+1$. Let $\eta$ be the unit vector in the direction of the line connecting the centers of the two balls, at the moment of collision. In this case the collision manifold at the collision point is defined by

$$\langle p_{k+1} - p_k, \eta \rangle = 0. \tag{I.6}$$

As before, we differentiate (I.6):

$$\langle \delta p_{k+1} - \delta p_k, \eta \rangle + R^{-1} \langle p_{k+1} - p_k, \delta q_{k+1} - \delta q_k \rangle = 0.$$

The conclusion then follows in complete analogy with the previous discussion.

(3). This case splits naturally into three sub-cases: a) a particle collides with one of its neighbors and with the boundary simultaneously; b) two particles collide at the same time (not with each other) and c) a particle collides simultaneously with two different parts of the boundary.

a) If we consider a variation at $z$, for which, along the infinitesimal family of trajectories, the collisions of particle $k$ with $\partial Q$ happen simultaneously ($\langle \delta q_k, \eta \rangle = 0$), the condition for $z$ to be a point of $\mathscr{R}^+$, owing to a contemporaneous collision with particle $k+1$, reads

$$\| q_{k+1} - q_k \| = 2R.$$

Differentiating now yields

$$\langle \delta q_{k+1} - \delta q_k, q_{k+1} - q_k \rangle = 0.$$

So, the vector $\tilde{\eta} = (0, \ldots, q_k - q_{k+1}, q_{k+1} - q_k, 0, \ldots, 0)$ is normal to $\mathscr{T}_z \mathscr{R}^+$. It is then clear that $\mathscr{R}^+$ is in the stable direction at the collision time, and that it is strictly in the stable direction at any preceding time.

b) This case is dealt in a fashion similar to the preceding one.

c) This case happens when a ball is colliding at the last reachable point in a corner or a bottleneck. It is easy to convince oneself that, in our examples, there is only a discrete number of positions in which a given ball, say $i$, can touch two different parts of the boundary $\partial Q$. Let $\tilde{q}$ be one such position, then the tangent vectors to the manifolds of the points that experience a double collision of the particle $i$ with $\partial Q$ will be given by $(\delta q, \delta p)$ with $\delta q_i = 0$. Accordingly, if $v$ is perpendicular to $p_i$, the vector $\eta = (0, \ldots, v, \ldots, 0)$ is perpendicular to the signularity manifold and to the flow direction. The same reasoning as that used before shows that $\mathscr{R}^+$ is strictly in the stable cone at any time preceding the collision.

It is simple to observe that the above discussion yields analogous results for $\mathscr{R}^-$ (the manifolds of points that experience a singular collision in the past); namely, $\mathscr{R}^-$ is strictly in the unstable direction at any time following the singular collision.

We still need to discuss the manifolds $\Sigma_i^k$ and their images.

Since the discussion is completely standard, we will limit ourselves to $\Sigma_i^1$, and will leave the other possibilities to the reader. We study first $\Sigma_i^1$ and then its images after the $i$, $i+1$ collision (see Definition 3.1). Let us define the $2 \times 2$ matrix $J$ by

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Obviously, $\langle Jv, v \rangle = 0$ for all $v \in \mathbb{R}^2$. Using the matrix $J$, we can write the condition $z = (q, p) \in \Sigma_i^1$ as

$$\langle Jp_i, p_{i+1} \rangle = 0. \tag{I.7}$$

Hence, the vector $(\delta q, \delta p) \in \mathscr{T}_z \Sigma_i^1$ iff

$$\langle Jp_i, \delta p_{i+1} \rangle - \langle Jp_{i+1}, \delta p_i \rangle = 0.$$

Thus, the vector $\eta = (0, \ldots, -Jp_{i+1}, Jp_i, 0, \ldots, 0)$ is orthogonal to $\mathcal{T}_z\Sigma_i^1$. According to Lemma I.3, $\Sigma_i^1$ is then a manifold in the stable direction. Next, we will prove that the image of $\Sigma_i^1$, after the $i$, $i+1$ collision, is strictly in the unstable direction. In analogy with (I.4), we consider vectors $\beta_i$ such that

$$\langle \beta_i, \beta_j \rangle = 0,$$
$$\langle \beta_i, p \rangle = 0,$$
$$\langle \beta_i, \eta_p \rangle = 0,$$

and a matrix $A$ defined as in (I.5) (apart from the minus sign). It can be easily verified that $V \subset \mathcal{T}_z\Sigma_i^1$, where $V = \{(\delta q, A\delta q) \mid \langle \delta q, p \rangle = 0\}$. Moreover, $V$ is contained in the unstable cone. The reason why the inclusion is not strict is the vector $(\eta_p, A\eta_p) = (\eta_p, 0)$: this vector lies on the boundary of the cone. If $(\eta_p, 0)$ becomes strictly contained in the cone after the collision, all the space $V$ will be strictly unstable (remember that $\langle \eta_p, p \rangle = 0$). According to Lemma 2.2, if $\langle p_{i+1} - p_i, \eta \rangle \neq 0$ (which is warranted by the definition of $\Sigma_i^1$), the above-mentioned vector may fail to become sufficient only if

$$-Jp_{i+1} + \lambda p_i = Jp_i + \lambda p_{i+1}$$

for some $\lambda$. But this is clearly impossible, out of the measure zero and codimension two set $\{p_i = p_{i+1} = 0\}$, since

$$\langle J(p_{i+1} + p_i), p_{i+1} - p_i \rangle = 0$$

by definition.

## Appendix II (Flows and Maps)

We have already noticed that, while the results of Sect. 4 are stated for maps, many properties are most readily checked for flows. In this Appendix we remind the reader of the concept of the Poincaré section and we show how to translate results or the map into results for the flow and vice versa. These things are either simple or well known so the Appendix is added only for the sake of clarity.

The phase space $\mathcal{M}$ of our examples is a smooth manifold with a piecewise smooth boundary $\partial\mathcal{M}$. Physically, $\partial\mathcal{M}$ is obtained by the boundary $\partial Q$ of the table in which the balls move and by the boundary of the cylinders that describe the forbidden positions of the balls (the centers of two balls cannot get closer than $2R$). All these conditions are formulated in terms of the vector $q \in \mathbb{R}^{2n}$ describing the positions of the balls, or, equivalently, the position of a particle in the corresponding $2n$ dimensional billiard. Therefore, we can speak about a configurational boundary of this billiard. The flow $\phi^t$ in $\mathcal{M}$ generates, in the $2n$ dimensional billiard picture, linear trajectories that are elastically reflected by the configurational boundary. As we repeatedly noted before, in the examples considered in this paper $\mathcal{M}$ has the configurational boundary with non-negative curvature (semi-dispersing billiards).

As to the boundary of $\mathcal{M}$, it splits naturally into two parts: $\partial_+\mathcal{M} = \{(q,p) \in \partial\mathcal{M} \mid \langle p, \eta\rangle \geqq 0\}$ and $\partial_-\mathcal{M} = \{(q,p) \in \partial\mathcal{M} \mid \langle p,\eta\rangle \leqq 0\}$, where $\eta \in \mathbb{R}^{2n}$ is the inward normal vector to the configurational boundary at the point $(q,p)$. Physically speaking, $\partial_-\mathcal{M}$ consists of the points just before a collision while $\partial_+\mathcal{M}$ consists of the points just after (remember that the flow is not defined at the collision times).

Let us call $\mathcal{B}$ the non-singular part of $\partial_+\mathcal{M}$; the boundary of $\mathcal{B}$ is constituted by points that correspond to tangent or a multiple collisions. The flow $\phi^t$ induces a map $T\colon \mathcal{B} \to \mathcal{B}$. This map is symplectic with respect to the induced symplectic structure. In particular, the measure $\mu$ on $\mathcal{B}$ induced by the projection of the Lebesgue measure on $\mathcal{M}$ along the flow direction is an invariant measure for $T$. The transformation $T$ is called the Poincaré map or Poincaré section asssociated to $\phi^t$. Sometimes in the literature the name of the Poincaré section is used for the pair $(\mathcal{B}, T)$ or for the corresponding dynamical system with the invariant measure $\mu$.

The set contains two subsets: $\mathcal{R}^+$, the set of points for which a tangent or multiple collision occurs in a future, but no such collision occurs in the past, and $\mathcal{R}^-$, the set of points for which a tangent or multiple collision occurs in a past, but no such collision occurs in the future. Both sets have $\mu$-measure zero, but carry their own measures $\mu_\pm$ induced by $\mu$. The map $T$ and its positive iterates are defined on $\mathcal{R}^-$ while the inverse map $T^{-1}$ and its positive iterates (i.e. the negative iterates of $T$) are defined on $\mathcal{R}^+$.

From the above construction it is clear that the properties of $T$ can be lifted to properties of $\phi^t$. Let us consider, for example, Theorem 4.3. If $x$ is a sufficient point with a smooth trajectory in the past and $\pi(x)$ is the first point on $\partial\mathcal{M}$ reached by $\phi^t(x)$, for $t \leqq 0$, then Theorem 4.3 can be applied to $\pi(x)$. Since, according to Lemma 2.3, the billiard flow preserves the property of being perpendicular to the flow direction, it is natural to define orthogonal manifolds like the ones used in Sect. 5.

Let $\hat{\mathcal{U}}(x)$ be a local manifold containing $x$ such that $\mathcal{T}_y\hat{\mathcal{U}}(x)$ is uniformly transversal to the flow direction at $y$ for each $y \in \hat{\mathcal{U}}(x)$ (for example an hyperplane in $\mathcal{M}$). It is then obvious that it is possible to transfer the construction carried out in Theorem 4.3 to $\hat{\mathcal{U}}(x)$ using the flow. If we want to produce a covering in a full neighborhood of $x$ we can just use the covering induced on the manifolds $\hat{\mathcal{U}}(\phi^\tau(x))$ for $\tau \in \left\{\frac{n}{\|p\|}\delta\right\}_{n \in I \subset \mathbb{Z}}$ and complete them with edges in the flow direction. We can then construct chains of, e.g., weak-unstable manifolds (the unstable manifold supplemented with the flow direction) and stable manifolds and perform the same construction as discussed in Theorem 4.6. The general theory of dynamical systems also ensures that, for semi-dispersing billiards, the mixing and $K$-property of the Poincaré map imply, respectively, that the billiard flow is mixing and $K$ [KS, S2].

*Acknowledgements.* We would like to thank N. Chernov, L. Chierchia, V. Donnay, S. Olla, N. Simányi, D. Szász and M. Wojtkowski for helpful and enlightening discussions. Particular thanks go to J. Guckenheimer, Director of the Center for Applied Mathematics at Cornell University where one of us (C. Liverani) was visiting during part of this work; in addition L. Bunimovich and C. Liverani thank the I.S.I. Foundation for the kind hospitality in I.S.I. where this paper was finished. Finally, we would like to thank D. Szász and S. Troubetzkoy for carefully reading a draft of the paper.

## References


[B] Boltzmann, L.: Lectures on gas theory (translation from german). Berkeley, CA: University of California Press, 1964

[BG1] Burns, K., Gerber, M.: Continuous invariant cone families and ergodicity of flows in dimension three. Ergodic Th. Dynam. Syst. **9**, 19–25 (1989)

[BG2] Burns, K., Gerber, M.: Real analytic Bernoulli geodesic flows on $S^2$. Ergodic Theoret. Dynam. Syst. **9**, 27–45 (1989)

[BS] Bunimovich, L. A., Sinai, Ya. G.: Statistical properties of Lorentz gas with periodic configuration of scatterers. Commun. Math. Phys. **78**, 479–497 (1981)

[Bu1] Bunimovich, L. A.: Decay of correlation in dynamical systems with chaotic behaviour. Sov. Phys. JETPh **62**, 1452–1471 (1985)

[Bu2] Bunimovich, L. A.: A theorem on ergodicity of two-dimensional hyperbolic billiards. Commun. Math. Phys. **130**, 599–621 (1990)

[Bu3] Bunimovich, L. A.: On billiards close to dispersing. Matem. Sbornik **95**, 40–73 (1974)

[C] Chernov, N. I.: Ergodic Hamiltonian system of two particles in an external field. Preprint

[DL] Donnay, V., Liverani, C.: Potentials on the two-torus for which the Hamiltonian flow is ergodic. Commun. Math. Phys. **135**, 267–302 (1991)

[E] Engelking, R.: Dimension theory. Amsterdam: North Holland 1978

[Gi] Gibbs, J. W.: Elementary principles in statistical mechanics. New York, 1902

[G] Gallavotti, G.: Lectures on the billiard. Lect. Notes in Phys. Vol. **38**, pp. 236–295. Berlin, Heidelberg, New York: Springer 1975

[H] Hopf, E.: Statistik der Geodatischen Linien in Mannigfaltigkeiten negativer Krümmung. Ber. Verh. Sächs. akad. wiss., Leipzig **91**, 261–304 (1939)

[K] Katok, A.: Invariant cone families and stochastic properties of a smooth dynamical systems. Preprint

[Kr] Krylov, N. S.: Works on the foundation of statistical physics. Princeton, NJ: Princeton University Press 1979

[KS] Katok, A., Strelcyn, J. M. with collaboration of Ledrappier F. and Przytycki F.: Invariant manifolds, entropy and billiards, smooth maps with singularities. Lect. Notes in Math. Vol. **1222**. Berlin, Heidelberg, New York: Springer 1986

[KSS1] Krámli, A., Simányi, N., Szász, D.: Three billiard balls on the $\nu$-dimensional torus is a $K$-flow. Ann. Math. **133**, 37–72 (1991)

[KSS2] Krámli, A., Simányi, N., Szász, D.: The $K$-property of four billiard balls. Commun. Math. Phys. **144**, 107–148 (1992)

[KSS3] Krámli, A., Simányi, N., Szász, D.: A "transversal" fundamental theorem for semi-dispersing billiards. Commun. Math. Phys. **129**, 535–560 (1990); Erratum. Commun. Math. Phys. **138**, 207–208 (1991)

[KSS4] Krámli, A., Simányi, N., Szász, D.: Ergodic properties of semi-dispersing billiards. I two cylindric scatterers in the 3-$D$ torus. Nonlinearity **2**, 311–326 (1989)

[LW] Liverani, C., Wojtkowski, M. P.: Ergodicity of Hamiltonian systems. Preprint

[O] Osceledec, V. I.: A multiplicative ergodic theorem: characteristic Lyapunov exponents of dynamical systems. Trans. Mosc. Math. Soc. **19**, 197–231 (1968)

[P] Besin, Ya. B.: Lyapunov characteristic exponents and smooth ergodic theory. Russ. Math. Surv. **32**(4), 55–114 (1977)

[SC] Sinai, Ya. G., Chernov, N. I.: Ergodic properties of certain systems of two-dimensional discs and three-dimensional balls. Usp. Math. Nauk **42**, 153–174 (1987)

[S1] Sinai, Ya. G.: On the foundations of the ergodic hypothesis for a dynamical system of statistical mechanics. Dokl. Akad. Nauk. SSSR **153**, 1261–1264 (1963)

[S2] Sinai, Ya. G.: Dynamical systems with elastic reflections. Russ. Math. Surv. **25**(1), 137–189 (1970)

[S3] Sinai, Ya. G.: The development of Krylov's ideas. In: Works on the foundations of statistical physics, by Krylov. Princeton, NJ: Princeton University Press 1979, pp. 239–281

[W1] Wojtkowski, M.: Invariant families of cones and Lyapunov exponents. Ergodic Theoret. Dynam. Syst. **5**, 145–161 (1985)

[W2] Wojtkowski, M.: A system of one dimensional balls with gravity. Commun. Math. Phys. **126**, 507–533 (1990)
[W3] Wojtkowski, M.: The system of one dimensional balls in an external field II. Commun. Math. Phys. **127**, 425–432 (1990)
[W4] Wojtkowski, M.: Measure theoretic entropy of the system of hard spheres. Ergodic Theoret. Dynam. Syst. **8**, 133–153 (1988)
[W5] Wojtkowski, M.: Systems of classical interacting particles with nonvanishing Lyapunov exponents. Preprint
[W6] Wojtkowski, M.: Private communication

Communicated by Ya. G. Sinai