Commun. Math. Phys. 129, 535–560 (1990)

Communications in
Mathematical
Physics
© Springer-Verlag 1990

A "Transversal" Fundamental Theorem for Semi-Dispersing Billiards

A. Krámli[1], N. Simányi[2,★], and D. Szász[2,3,★]

[1] Computer and Automation Institute of the Hungarian Academy of Sciences
[2] Mathematical Institute of the Hungarian Academy of Sciences, H-1364 Budapest, P.O.B. 127, Hungary
[3] Also: Princeton University and the Institute for Advanced Study, Princeton, NJ, USA

Dedicated to Joel L. Lebowitz on the occasion of his 60th birthday

Abstract. For billiards with a hyperbolic behavior, Fundamental Theorems ensure an abundance of geometrically nicely situated and sufficiently large stable and unstable invariant manifolds. A "Transversal" Fundamental Theorem has recently been suggested by the present authors to prove *global ergodicity* (and then, as an easy consequence, the K-property) of semi-dispersing billiards, in particular, the global ergodicity of systems of $N \geqq 3$ elastic hard balls conjectured by the celebrated *Boltzmann-Sinai ergodic hypothesis.* (In fact, the suggested "Transversal" Fundamental Theorem has been successfully applied by the authors in the cases $N=3$ and 4.) The theorem generalizes the Fundamental Theorem of Chernov and Sinai that was really the fundamental tool to obtain *local ergodicity* of semi-dispersing billiards. Our theorem, however, is stronger even in their case, too, since its conditions are simpler and weaker. Moreover, a complete set of conditions is formulated under which the Fundamental Theorem and its consequences like the Zig-zag theorem are valid for general semi-dispersing billiards beyond the utmost interesting case of systems of elastic hard balls. As an application, we also give conditions for the ergodicity (and, consequently, the K-property) of dispersing-billiards. "Transversality" means the following: instead of the stable and unstable foliations occurring in the Chernov-Sinai formulation of the stable version of the Fundamental Theorem, we use the stable foliation and an arbitrary nice one transversal to the stable one.

1. Introduction

Smooth dynamical systems with singularities satisfying a hyperbolicity condition play an utmost important role in the theory of dynamical systems for (i) they contain such interesting classes as systems of elastic hard balls or – more generally

★ Research partially supported by the Hungarian National Foundation for Scientific Research, grant No. 819/1

– semi-dispersing billiards (and also a lot of non-dispersing ones, like the Bunimovich stadium) and maps like the Lozi-map; (ii) their theory is closely related to the theory of non-uniformly hyperbolic smooth dynamical systems. Since Sinai's celebrated paper (S-1970) it has been well understood that, once the positivity of the ergodic components has been proven, the main tool for obtaining additional, topological inference on the ergodic components of the system is a suitable form of the so-called fundamental theorem.

As mentioned before, the class of semi-dispersing billiards (i.e. those with – not necessarily strictly – convex scatterers) contains hard-ball-systems as well, and it was this class of billiards for which, in 1987, Chernov and Sinai (S-Ch, 1987) proved a strong fundamental theorem immediately in the multidimensional setting. The main consequence of the Chernov-Sinai theory is, in general, that a suitable open neighborhood of a phase point possessing a sufficiently rich trajectory belongs to one ergodic component (local ergodicity).

In fact, the system of two elastic balls on the ν-torus $\mathbf{T}^\nu : \nu \geqq 2$ is isomorphic to a dispersing billiard (i.e. to one with strictly convex scatterers) and the main application of their theory was that this system was a K-flow on the submanifold of the phase space specified by the trivial conservation laws. For completeness, this result will be derived in Sect. 7 and we shall see that, in this case, global ergodicity follows relatively easily from local ergodicity.

If, however, one aims at proving global ergodicity for general semi-dispersing billiards, then essentially new problems arise in going beyond local ergodicity (that already assumes that the Ansatz, a special global condition of the Chernov-Sinai theory has been verified, which in itself is not an easy task, in general). In fact, the method should also be complemented by geometric-algebraic and topological (dimension-theoretic) tools (see, for instance, K–S–Sz (1989-B) for a billiard in $\mathbf{T}^3$ with two cylindric scatterers or K–S–Sz (1989-C) for the system of three balls on $\mathbf{T}^\nu : \nu \geqq 2$). This latter part of the argument where – surprisingly enough – the fundamental theorem gets applied again, and this is exactly the place where its "transversal" form is, in general, needed. This means that, when ensuring an abundance of not too short local stable invariant manifolds, instead of the unstable foliation we assume an arbitrary smooth foliation (in a small neighborhood of a sufficient point) about which we only require its transversality to the stable one.

In the applications to semi-dispersing billiards one also finds that the conditions of the fundamental theorem should also be relaxed. To this end we reshape the Ansatz to make its verification simpler (see Condition 3.1 in this paper) and use the sufficiency assumption for a point in its minimal form that can imply local hyperbolicity at all.

As to the proof one can observe that the method of Chernov and Sinai is so robust that, apart from simpler additional ideas, it also works for proving the necessary stronger form of the theorem (cf. K–S–Sz (1989-C), Theorem 2.4). Since the original proof, we think a gem of the theory in itself, was formulated in an extremely concise and occasionally sketchy way and, moreover, because of the very importance of the fundamental theorems, the main aim of the present paper is to give an elaborated proof of our "transversal" version of the fundamental theorem.

Our theorem, however, contains two additional assumptions: one on the subset of degenerate tangencies and further one on the subset of double singularities. These conditions seem to be necessary for the method and, implicitly, are also used in S–Ch (1987). As a matter of fact, in the applications to date, their verification is not too hard and, for instance, for systems of hard balls they hold obviously.

The careful reader will notice the difference between the definitions of bad parallelograms given in S–Ch (1987) and in Sect. 5 of the present paper. Our definition helps to resolve a small gap in the original proof.

Finally, we note that the definition of the function $z(x)$ for the distance from the singularities is also different here. It is based on our metric ϱ introduced in the phase space of the Poincaré section map, which makes the whole discussion more natural and simple.

The paper is organized as follows. Section 2 prepares the formulation, done in Sect. 3, of the fundamental theorem and of its main corollaries (the Zig-zag Theorem and the one on local ergodicity) by collecting the necessary notions and facts about semi-dispersing billiards. Section 4 contains elementary geometric facts about semi-dispersing billiards to be used in the proofs. In Sects. 5 and 6 we prove the fundamental theorem. Section 5 gives the main body of the proof while Sect. 6 separates the proof of the Tail Bound, the only place where the Ansatz is used. Finally as an application, in Sect. 7 we also give a detailed proof of the ergodicity of dispersing billiards including systems of two balls.

2. Semi-Dispersing Billiards and Invariant Manifolds

Billiards. A billiard is a dynamical system describing the motion of a point particle in a connected, compact domain $Q \subset \mathbf{R}^d$ or $Q \subset \mathbf{T}^d = \mathrm{Tor}^d$, $d \geqq 2$ with a piecewise C^2-smooth boundary. Inside Q the motion is uniform while the reflection at the boundary ∂Q is elastic (the angle of reflection equals the angle of incidence). Since the absolute value of the velocity is a first integral of motion, the phase space of our system can be identified with the unit tangent bundle over Q. Namely, the configuration space is Q while the phase space is $M = Q \times S_{d-1}$, where S_{d-1} is the surface of the unit d-ball. In other words, every phase point x is of the form (q, v), where $q \in Q$ and $v \in S_{d-1}$. The natural projections $\pi: M \to Q$ and $p: M \to S_{d-1}$ are defined by $\pi(q, v) = q$ and by $p(q, v) = v$, respectively.

Suppose that $\partial Q = \cup_1^k \partial Q_i$, where ∂Q_i are the smooth components of the boundary. Denote $\partial M = \partial Q \times S_{d-1}$ and let $n(q)$ be the unit normal vector of the boundary component ∂Q_i at $q \in \partial Q_i$ directed inwards to Q.

The flow $\{S^t\}: t \in \mathbf{R}$ is determined for the subset $M' \subset M$ of phase points whose trajectories never cross the intersections of the smooth pieces of ∂Q and do not contain an infinite number of reflections in a finite time interval. If μ denotes the Liouville (probability) measure on M, i.e. $d\mu(q, v) = \text{constant} \times dq \cdot dv$, where dq and dv are the Lebesgue measures on Q and on S_{d-1} respectively, then under certain conditions $\mu(M') = 1$ and μ is invariant [cf. K–S–F (1980)]. The interior of the phase space M can be endowed with the natural Riemannian metric. For our present purpose it is sufficient to pose the following assumption.

Condition 2.1. (Residuality of trajectories with a finite accumulation.) *The set of phase points in whose trajectory the moments of reflections accumulate in a finite time interval form a residual set.*

We remind the reader that a subset is residual if it is contained in a countable union of closed, codimension 2 *0-sets* [*cf.* K–S–Sz (1989-C)].

We remark that a strong form of Condition 2.1 holds for billiard systems isomorphic to systems of elastic hard balls in $\mathbf{T}^\nu$, the ν-dimensional torus. As a matter of fact, for these systems the aforementioned set is empty [cf. G (1981) and I (1988)].

If at every point $q \in \partial Q$ the normal vectors $n(q)$ (oriented inwards Q) of smooth pieces of ∂Q are linearly independent, then, by K–S–F (1980), then the set in question has measure 0, but its topological smallness is not treated.

The dynamical system $(M, \{S^t\}, \mu)$ is said to be a *billiard*. Notice, that $(M, \{S^t\}, \mu)$ is neither everywhere defined nor smooth.

The main object of the present paper is a particularly interesting class of billiards: that of *semidispersing* ones where, for every $q \in \partial Q$, *the* second fundamental form $K(q)$ of the boundary is non-negative (if, moreover, for every $q \in \partial Q$, $K(q)$ is positive, then the billiard is called a *dispersing* one).

It will be convenient to denote $(q, -v)$ by $-x$ if $x=(q,v)$; then, of course, for $y=S^t x$ we have $-x=S^t(-y)$.

Invariant Manifolds. We recall that a C^1-smooth, connected submanifold $\gamma^s \subset M$ without boundary is called a *local stable (invariant) manifold* for $\{S^t\}$ at $x \in M$ iff

(i) $x \in \gamma^s$.

(ii) $\exists C_i = C_i(\gamma^s) > 0$ $(i=1,2)$ such that, for any $y_1, y_2 \in \gamma^s$, $t>0$,

$$\operatorname{dist}(S^t y_1, S^t y_2) \leqq C_1 \exp\{-C_2 t\} \operatorname{dist}(y_1, y_2).$$

A local stable manifold for $\{S^{-t}\}$ is called a *local unstable manifold* for $\{S^t\}$.

In what follows we summarize some facts about the existence of invariant manifolds and their dimensions.

In the construction of invariant manifolds a crucial role is played by the time evolution equation for the second fundamental form of codimension 1 submanifolds in Q orthogonal to a given vector x. Let $x=(q,v) \in M \backslash \partial M$ and choose a C^2-smooth codimension 1 submanifold $\tilde{\Sigma} \subset Q \backslash \partial Q$ such that $q \in \tilde{\Sigma}$ and $v = v(q)$ is a unit normal vector to $\tilde{\Sigma}$ at q. Denote by Σ the normal section of the unit tangent bundle of Q restricted to $\tilde{\Sigma}$ (Σ is uniquely defined by the orientation $(q,v) \in \Sigma$). We call Σ a *local orthogonal manifold* with support $\tilde{\Sigma}$.

Recall that *the second fundamental form* $B_\Sigma(x)$ of Σ (or $\tilde{\Sigma}$) at x is defined through

$$v(q+dq) - v(q) = B_\Sigma(x) dq + o(\|dq\|)$$

and is a self-adjoint operator acting in the $(d-1)$-dimensional tangent hyperplane $\mathscr{I}(x)$ of $\tilde{\Sigma}$ at x.

A local orthogonal manifold Σ is called *convex* if $B_\Sigma(y) \geqq 0$ for every $y \in \Sigma$.

Consider a trajectory $\{x^t = S^t x : t \in \mathbf{R}\}$. Between consecutive reflections the hyperplanes $\mathscr{I}(x^t)$ can simply be identified by a projection parallel with v. Further, this identification can be extended along the whole trajectory by determining it in points of ∂M. Let $x \in \partial M$. Then, first of all, $S^{t-}x$ and $S^{t+}x$ can be identified by gluing $y, y' \in \partial M$ iff $\pi(y) = \pi(y') = q$ and, moreover, $p(y') = p(y) - 2(n(q), p(y))n(q)$. We shall use the following convention: if, for $x^t \in \partial M$, we write x^{t-} or x^{t+}, then we mean (q^t, v^{t-}) and (q^t, v^{t+}) respectively, while, if we write x^t, then we mean the glued object as represented by x^{t+}. In addition, for $x^t \in \partial M$, (i) it makes sense to define $\mathscr{I}(x^{t-})$ and $\mathscr{I}(x^{t+})$ and the identification between them is determined by a projection parallel to $n(q)$ and (ii) we denote by $V(x^t)$ the projection of $\mathscr{I}(x^{t+})$ onto $\mathscr{T}_{q^t}\partial Q$ parallel to v^{t+} and by $V^*(x^t)$ its adjoint which, in fact, projects $\mathscr{T}_{q^t}\partial Q$ onto $\mathscr{I}(x^{t+})$ parallel to $n(q^t)$, where $\mathscr{T}_{q^t}\partial Q$ is the tangent hyperplane to ∂Q in $q \in \partial Q$.

Poincaré Map. Traditionally one reduces the study of the ergodic properties of the flow $\{S^t\}$ to that of a discrete time dynamical system. This system is given by an automorphism T, the so-called Poincaré section of $\{S^t\}$, where the section is defined by using $\partial M^+ = \{x = (q, v) \in \partial M, (v, n(q)) \geqq 0$ and $\exists \varepsilon$ such that $\forall 0 < \delta < \varepsilon\ S^\delta x \notin \partial M\}$. However, if necessary we can, in addition, introduce "virtual" hyperplanar boundaries of codimension 1 in the configuration space (and, correspondingly, in the phase space, too) and thus we can obtain a system with a bounded free path (lack of infinite horizon!). These virtual walls are transparent [cf. S–Ch (1987)], i.e. when hitting these walls the velocity remains unchanged. If the union of these virtual boundaries in Q is $\widetilde{\partial Q}$, then the phase space of this new discrete time system will be $\widehat{\partial M} = \partial M \cup \widetilde{\partial M}$, where $\widetilde{\partial M} = \widetilde{\partial Q} \times S_{d-1}$.

Let $\tau: M \to \mathbf{R}^+$ be defined as follows:

$$\tau(x) := \inf\{t > 0 : S^t x \in \partial M^+\}.$$

Introduce the mapping $T^+ : M \to \widehat{\partial M}$ as $T^+ x := S^{\tau(x)+0}x$. The restriction T of the mapping T^+ to $\widehat{\partial M}$ is the Poincaré section of the flow $\{S^t\}$. Then $(\widehat{\partial M}, T, \mu_1)$ is a dynamical system with $d\mu_1(q, v) = \text{const} \cdot |(v, n(q))| dq \cdot dv$, where dq is the Riemannian measure on $\widehat{\partial Q} := \partial Q \cup \widetilde{\partial Q}$. An additional remark is that $-x = T^n(-y)$ if $y = T^n x$.

A smooth piece $\Sigma \subset \widehat{\partial M}$ of the T^+-image of a local orthogonal manifold in M will also be called a *local orthogonal manifold* (in $\widehat{\partial M}$). Finally, we say that a local orthogonal manifold $\Sigma \subset \widehat{\partial M}$ is *convex* if it is the T^+-image of a convex orthogonal submanifold in M.

Important Convention. For brevity, by slightly abusing the notation, we will throughout this paper denote $\widehat{\partial M}$ (and $\widehat{\partial Q}$) by simply ∂M (and ∂Q)!

The natural Riemannian metric ϱ on ∂M inducing the invariant measure μ_1 can be defined via the equation

$$(d\varrho)^2 = (n(q), v)^2 \cdot (dq_1)^2 + (dq_2)^2 + (dv)^2, \tag{2.2}$$

$$(x = (q, v) \in \partial M, \quad dq = (dq_1, dq_2) \in \mathscr{T}_q \partial Q, \quad dv \in \mathscr{I}(x)).$$

In this formula the component $dq_2 \in \mathscr{T}_q \partial Q$ of dq is orthogonal to v while dq_1 is a scalar multiple of $v - (n(q), v) \cdot n(q)$.

Semidispersing Billiards. From now on, throughout this paper, we restrict our attention to semi-dispersing billiards. For a local orthogonal manifold Σ to $x \in M \setminus \partial M$ let $\Sigma^t = S^t \Sigma$. The invariant manifolds are the solutions of a differential-geometric problem: at every point their second fundamental form should be equal to an operator-valued continued fraction. To obtain these continued fractions we have to recall the evolution laws for B_{Σ^t}. For the sake of definiteness, we fix a point $x \in M$ and understand all operators mapping some $\mathcal{I}(x^t)$ into itself as operators mapping $\mathcal{I}(x)$ into itself by using the unitary identification introduced above.

Proposition 2.3. (i) *If* $x \in M$ *and for* $0 \leqq t' \leqq t\, x^{t'} \notin \partial M$, *then*

$$B_{\Sigma^t}(x^t) = B_\Sigma(x)(1 + tB_\Sigma(x))^{-1}.$$

(ii) *If* $x^t \in \partial M$, *then*

$$B_{\Sigma^{t+}}(x^{t+}) = B_{\Sigma^{t-}}(x^{t-}) + 2(v^{t+}, n(q^t))\, V^*(x^t) K(q^t) V(x^t).$$

The proposition and the semi-dispersing property imply that, if $B_\Sigma(x) \geqq 0$, $x \in \Sigma$, then $B_{\Sigma^t}(x^t) \geqq 0$ for every $t \geqq 0$, i.e. convexity of $\tilde{\Sigma}$ is preserved under S^t, $t \geqq 0$.

The billiard flow $\{S^t\}$ has singularities of different types. Since we shall always work with the Poincaré section map T it is sufficient to give a more detailed description of the set $\mathcal{R} \subset \partial M$ of the singularities of T. $\mathcal{R} = \mathcal{R}' \cup \mathcal{R}''$ is a $2d-3$ dimensional CW-complex which consists of two types of CW-complexes:

a) $\mathcal{R}' := \pi^{-1}(\cup_j \partial(\partial Q)_j) \cup T^{-1}[\pi^{-1}(\cup_j \partial(\partial Q)_j))]$ (the "double reflections" and their preimages under T);
b) $\mathcal{R}'' := \{x = (q, v) \in \partial M : (v, n(q)) = 0\} \cup T^{-1}[\{x = (q, v) \in \partial M : (v, n(q)) = 0\}]$ ("tangencies" and their pre-images under T).
For $n \in \mathbf{Z}$ let

$$\mathcal{R}^n = T^{-n}\mathcal{R}.$$

Denote for arbitrary $n \in N$ by Δ_n the set of double singularities of maximal order n. Δ_n consists of points $x \in \partial M$ for which there exist two different integers k_1 and k_2($|k_1| \leqq n$, $|k_2| \leqq n$) such that $T^{k_1}x$ and $T^{k_2}x$ belong to the set of singular reflections ("double" or "tangential" ones). The set of singular reflections is denoted by $\mathcal{SR}$ ($\mathcal{R} = \mathcal{SR} \cup T^{-1}(\mathcal{SR})$). Introduce the following notations:

$$\partial M^0 := \partial M \Big\backslash \bigcup_{n \in Z} \mathcal{R}^n,$$

$$\partial M^* := \partial M \Big\backslash \bigcup_{n=1}^{\infty} \Delta_n,$$

$$\partial M^1 := \partial M^* \backslash \partial M^0.$$

Let, moreover, $M^* := (T^+)^{-1} \partial M^*$.

Next, for every $x \in M''$ ($M'' = \{x : S^t x \notin \mathcal{R}$ for every $t \geqq 0\}$) we introduce the operator-valued continued fractions $B(x)$ mentioned before. Let $t_0 = 0$ and $0 < t_1 < t_2 < \dots$ be the sequence of reflection moments of the positive semi-trajectory $\{x^t : t \geqq 0\}$. One can show [cf. S (1979), Ch (1982)] that, for semi-dispersing billiards, the operator-valued continued fraction

$$B(x) = \cfrac{1}{s_1 + \cfrac{1}{2\cos\varphi_1 V_1^* K_1 V_1 + \cfrac{1}{s_2 + \cfrac{1}{2\cos\varphi_2 V_2^* K_2 V_2 + \cfrac{1}{\dots}}}}} \tag{2.4}$$

exists whenever $x \in M''$ and is a non-negative linear operator for every $x \in M''$, where $s_n = t_n - t_{n-1}$, $K_n = K(q^{t_n})$, $V_n^{(*)} = V^{(*)}(x^{t_n})$ and $\cos\varphi_n = (v^{t_n+}, n(q^{t_n}))$.

Remark 2.5. It is very important to observe that a finite analog of the present continued fraction formula is also valid for the second fundamental form

$$B_{S^t(\Sigma)}(-x)$$

of the S^t-image of any convex local orthogonal manifold Σ containing the phase-point $-S^t(x)$ ($x \in M''\backslash\partial M$, $S^t(x) \notin \partial M$). This formula, which is a consequence of Proposition 2.3, looks as follows:

$$B_{S^t(\Sigma)}(-x) = \cfrac{1}{s_1 + \cfrac{1}{2\cos\varphi_1 V_1^* K_1 V_1 + \cfrac{1}{s_2 + \cfrac{1}{2\cos\varphi_2 V_2^* K_2 V_2 + \ldots + \cfrac{1}{s_n + \cfrac{1}{B_\Sigma(-S^t x)}}}}}} \tag{2.6}$$

(Here s_n is the length of the trajectory segment between $S^t(x)$ and the last reflection before it.) The semi-dispersing property (non-negativity of $V_k^* K_k V_k$) implies the non-negativity of every partial continued fraction of the right-hand side in (2.6) and, moreover, the formula in (2.6) is monotone in the self-adjoint operator variable $B_\Sigma(-S^t x)$. These properties together with the continuity of the operator coefficients $s_k = s_k \cdot 1$ and $V_k^* K_k V_k$ and also with the convergence of the right-hand side in (2.6) yield the following proposition:

Proposition 2.7. *Using the notions and notations above, for every $\varepsilon > 0$ there exist a large positive integer n_ε and a suitably small neighborhood U_ε of x in M such that for every $y \in U_\varepsilon \cap M''$ and for every convex local orthogonal manifold Σ containing $-S^t y$ such that $S^t y$ is after the n_ε^{th} reflection of y (in the positive semi-trajectory) we have that*

$$\|\tilde{B}(x) - \tilde{B}_{S^t\Sigma}(-y)\| < \varepsilon .$$

Here the operators $\tilde{B}(x)$ and $\tilde{B}_{S^t\Sigma}(-y)$ are extended from $\mathscr{J}(x)$ and $\mathscr{J}(y)$ to the bigger vector space $\mathscr{T}Q$, taking the value zero in the orthogonal complement of $\mathscr{J}(x)$ and $\mathscr{J}(y)$, respectively.

The dilation effect of the dynamics on an orthogonal manifold Σ is described by the linearization of S^t, i.e. by

$$D^t_{x,\Sigma} : \mathscr{T}_x\Sigma \to \mathscr{T}_{S^t x} S^t\Sigma . \tag{2.8}$$

The linearization of T^n is

$$D^n_{x,\Sigma} : \mathscr{T}_x\Sigma \to \mathscr{T}_{T^n x}(T^n\Sigma) , \tag{2.9}$$

where Σ is a local orthogonal manifold in ∂M through $x \in \partial M$. We will be consequent in using $D^t_{x,\Sigma}$ for the flow and $D^n_{x,\Sigma}$ for the map. Observe that, for Σ convex $\|(D^t_{x,\Sigma})^{-1}\| \leqq 1$ and $\|(D^n_{x,\Sigma})^{-1}\| \leqq 1$ [see Ch (1982)].

It is of special importance to characterize the positive subspace of $B(x)$. To this end we decompose

$$\mathscr{I}(x) = \mathscr{I}_+(x) \oplus \mathscr{I}_0(x),$$

where

$$B(x)|_{\mathscr{I}_+(x)} > 0, \qquad B(x)|_{\mathscr{I}_0(x)} = 0.$$

For $x \in M''$ let $j(x) = \dim \mathscr{I}_+(x)$, and

$$\Omega = \{x \in M: \text{for some neighborhood } U(x) \subset M, \; j(y) \text{ is a positive constant in } U(x) \cap M''\}$$

being, of course, an open set. For $x = (q, v) \in \Omega$, consider the tangent space $\mathscr{T}_x M = \mathscr{T}_q Q \oplus \mathscr{T}_v S_{d-1}$, where there is a natural isomorphism between $\mathscr{T}_v S_{d-1}$ and $\mathscr{I}(x)$. The set $E(x) = \{(e, f)) : e \in \mathscr{I}_+(x),\ f = -B(x)e\}$ is a linear subspace of $\mathscr{T}_x M$ and $\dim E(x) = j(x)$.

Theorem 2.10 [Ch (1982)]. *If for some $q \in \partial Q$, $K(q) \neq 0$, then $\Omega \neq \emptyset$ and, for a.e. $x \in \Omega$, there exists a local stable manifold $\gamma^s(x)$ of dimension $j(x)$ at x and $\mathscr{T}_x \gamma^s(x) = E(x)$. The* *proof* of this theorem and further steps of the analysis are based on a simple but important observation: the positive subspace $\mathscr{I}_+(x)$ of $B(x)$ is finitely defined. Namely, there exists a function $l_0 : M'' \to \mathbf{Z}_+$ such that

$$\mathscr{I}_0(x) = \{e \in \mathscr{I}(x): V_l^* K_l V_l e = 0 \text{ for every } l = 1, 2, \ldots, l_0(x)\}.$$

Denote by $l(x)$ the minimal such function $l_0(x)$. It is easy to see that $l(x)$ is upper semicontinuous on Ω. The following characterisation of $\mathscr{I}_+(x)$ is trivial:

$$\mathscr{I}_+(x) = \mathscr{L}\{\mathscr{I}_{m,+}(x): 1 \leqq m \leqq l(x)\},$$

where $\mathscr{L}$ is the linear subspace spanned by those indicated in the brackets and $\mathscr{I}_{m,+}(x)$ denotes the positive subspace of $\mathscr{I}(x)$ for the operator $V_m^* K_m V_m$.

Remark. The most important special case $j(x) = d-1$ of the last theorem (more exactly, the a.e. existence of invariant manifolds in a neighborhood of each sufficient point) can easily be obtained from the Fundamental Theorem (Theorem 3.6).

The following simple lemma is an easy consequence of the definitions.

Lemma 2.11. *If $j(x^{0-}) = d-1$, then $l(x^{0-}) \geqq l((-T^{l(x)-1}x)^{0-})$.*

If $T^+ x \in \mathscr{R}^n \cap M^1$, $n \geqq 0$, then $B(y)$ has two limiting values as $y \to x$ [cf. S–Ch (1987). p. 170] and they will be denoted by $B^1(y)$, $B^2(y)$. In this case all the notions and notations introduced above for $B(x)$ will have two values and will be distinguished by superscripts 1 and 2, e.g. $\mathscr{I}^i_+(x), l^i(x), j^i(x); \mathrm{i} = 1, 2$. For these points, too, we introduce $j(x) = \min\{j^1(x), j^2(x)\}$ and $l(x) = \max\{l^1(x), l^2(x))\}$.

Sufficiency of Trajectories. For any interval (a,b), $-\infty \leqq a < b \leqq \infty$ we will throughout denote by $S^{(a,b)}x$ the trajectory segment $\{S^t x : a < t < b\}$. As already said the notion of sufficiency is important because it is the weakest requirement that can ensure local hyperbolicity of a semi-dispersing billiard.

Definition 2.12. A trajectory segment $S^{(a,b)}x$, $-\infty \leqq a < b \leqq \infty$ satisfying $S^{(a,b)}x \cap \mathcal{SR} = \emptyset$ is said to be sufficient if

$$\dim \mathcal{L}\{\mathcal{I}_{t,+}(x) : t \in (a,b), S^t x \in \partial M\} = d-1,$$

where, for any t such that $S^t x \in \partial M$, $\mathcal{I}_{t,+}(x)$ denotes the positive subspace of $\mathcal{I}(x)$ for the operator V^*KV taken at $S^t x$. If $S^{(a,b)}x$ intersects $\mathcal{SR}$ just once, then $S^{(a,b)}x$ is called *sufficient* if both branches of the trajectory are sufficient in the time interval (a,b).

Finally a point $x \in M^*$ is said to be *sufficient* if its trajectory $S^{(-\infty,\infty)}$ is sufficient.

Note that, by virtue of Lemma 2.11, x is sufficient if and only if $-x$ is sufficient.

The last lemma of this section prepares our improvement for the strong fundamental theorem of Chernov and Sinai. But first we should again introduce some notations. For any $y \in U(x) = U$ let

$$\tau_U(y) = \inf\{t > 0 : S^t(y) \in U(x) \text{ and for some } s \in (0,t), S^s(y) \notin U(x)\}$$

and, for any convex orthogonal manifold $\Sigma \ni y$, let

$$D^{\tau}_{y,\Sigma} =: D^{\tau_U(y)}_{y,\Sigma}.$$

Lemma 2.13. *For every $x \in M^0$ such that $S^{(0,\infty)}$ is sufficient, there exist a neighborhood $U(x)$ and a positive constant $\lambda = \lambda(x) < 1$ such that*

(i) *through almost every point $y \in U(x)$ there do pass uniformly transversal local stable and unstable manifolds $\gamma^s(y)$ and $\gamma^u(y)$ of dimension $d-1$;*
(ii) *for any $y \in U(x)$ and for any convex orthogonal manifold $\Sigma^{\pm} \ni \pm y$,*

$$\|(D^{\tau}_{\pm y, \Sigma^{\pm}})^{-1}\| < \lambda.$$

Due to the continuity of $B(y)$, if $U(x)$ is sufficiently small, then, for every $y \in U(x)$, $l(y) \leqq l(x)$ and $\{V_m^* K_m V_m : 1 \leqq m \leqq l(x)\}$ is a y-dependent, sufficient sequence of operators uniformly close to the analogous sequence defined for x. It follows from the Poincaré recurrence theorem that for every neighborhood $U(x)$ and for a.e. $y \in U(x) \cap M^0$ the recurrence time $\tau_{-U}(-y)$ is finite and $S^{\tau_{-U}(-y)}(-y) := y' \in (-U(x)) \cap M^0$. Then, in view of Lemma 2.11, for a.e. $y \in U(x) \cap M^0$, the trajectory $\{S^t(-y) : 0 \leqq t \leqq \tau_{-U}(-y)\}$ contains a sufficient sequence of reflection operators thus implying $B(-y) > 0$ and $j(-y) = d-1$. Now the existence Theorem 2.10 can be applied providing the a.e. existence in $U(x)$ of local stable and unstable manifolds of dimension $(d-1)$.

The validity of (ii) follows easily from Theorem 2.10, from the positivity and continuity of B and, finally, from the existence of a sufficient sequence of reflections.

Remark 2.14. The assertion of the previous lemma also holds if $x \in M^1$.

Further Convention. The essential constants of the proof are denoted by $\lambda = \Lambda^{-1}$, $0 < \lambda < 1$, $c_i > 0$, and $\varepsilon_j > 0$. The difference between the c_i's and ε_j's is that the latter ones can be chosen arbitrarily small by appropriately shrinking the neighborhood we are working in. All the constants will be independent of the parameter δ and about the c_i's we will always say what they depend on.

3. Formulation of the Fundamental Theorem

The main aim of this section is to formulate the Fundamental Theorem in its most general and applicable form. We note that there are two dual forms of the Fundamental Theorem: the first one providing an ample set of not too short local stable invariant manifolds and the other one stating the same property for the local unstable manifolds. Now we are going to draw up the first (stable) version of the Fundamental Theorem; the dualization, being an easy exercise, is left to the reader.

In order to phrase the theorem we need three important preliminary conditions.

Condition 3.1 (Chernov-Sinai Ansatz). *For $\nu_{S\mathcal{R}}$-almost every point $x \in S\mathcal{R}$ we have $x \in \partial M^*$ and the positive semitrajectory of the point x is sufficient, where $\nu_{S\mathcal{R}}$ denotes the measure on the codimension 1* CW-*subcomplex $S\mathcal{R}$ of ∂M induced by the Riemannian metric ϱ.*

The other important regularity condition needed for the proof of the Fundamental Theorem is:

Condition 3.2 (Regularity of the set of degenerate tangencies). *The set*

$$\{x = (q, v) \in \partial Q \times S_{d-1} : (v, n(q)) = 0 \text{ and } K(q)v = 0\}$$

is a finite union of compact smooth submanifolds of ∂M (usually with boundary), i.e. this set is a CW-*subcomplex of $\mathcal{R}$.*

We remark that Condition 3.2 trivially holds for semi-dispersing billiards with solely cylindric scatterers.

Our last regularity condition concerns the sets Δ_n of double singularities:

Condition 3.3 (Regularity of double singularities). *For every $n \in \mathbf{N}$ the set Δ_n is a finite union of compact smooth submanifolds of ∂M.*

Definition 3.4. Let $x \in \partial M^*$ and let $U(x)$ be an open neighborhood of x in ∂M *diffeomorphic to* $\mathbf{R}^{2d-2}$ and $U(X) = \bigcup_{\alpha \in B^{d-1}} \Gamma_\alpha$ a smooth foliation of $U(x)$ with $(d-1)$-dimensional smooth submanifolds Γ_α which are uniformly transversal to all possible local stable invariant manifolds in $U(x)$ (B^{d-1} is the standard $(d-1)$-dimensional open ball, i.e. the factor of $U(x)$ by the foliation).

The parametrized family

$$\mathcal{G}^\delta = \{G_i^\delta : i = 1, 2, \ldots, I(\delta)\} \qquad (0 < \delta < \delta_0)$$

of finite coverings of $U(x)$ is called a *family of regular coverings* iff the following five requirements are fulfilled:

(a) all the sets G_i^δ are *open* parallelepipeds of dimension $2d-2$, i.e. they are images of the standard unit cube $[0,1]^{2d-2} \subset \mathbf{R}^{2d-2}$ under inhomogeneous linear mappings $\mathbf{R}^{2d-2} \to U(x)$ where linearity is defined in terms of a fixed coordinate system in $U(x)$, say the exponential coordinates using the mapping $\exp_x$;
(b) for the centers $w_i^\delta \in M^0$ of G_i^δ (according to this coordinate system) the tangent spaces $E(w_i^\delta)$ and $\mathscr{T}_{w_i^\delta}\Gamma(w_i^\delta)$ are parallel (according to the coordinate system) to some $(d-1)$-dimensional faces of G_i^δ. The faces of G_i^δ parallel to $E(w_i^\delta)$ are called s-faces (there are 2^{d-1} of them), while those faces of G_i^δ parallel to $\mathscr{T}_{w_i^\delta}\Gamma(w_i^\delta)$ are called Γ-faces (there are also 2^{d-1} of them) and they are supposed to be cubes with edge-length δ;
(c) if $G_i^\delta \cap G_j^\delta \neq \emptyset$, then

$$\mu_1(G_i^\delta \cap G_j^\delta) \geqq c_1 \cdot \delta^{2d-2},$$

where $c_1 > 0$ is a fixed number (not depending on δ);
(d) for every $\delta < \delta_0$ there are at most 2^{2d-2} different indices $1 \leqq i_1 < i_2 < \ldots < i_k \leqq I(\delta)$ such that

$$\bigcap_{j=1}^{k} G_{i_j}^\delta \neq \emptyset;$$

(e) the system of the centers

$$\{w_i^\delta : i = 1, 2, \ldots, I(\delta)\}$$

constitutes a $(2d-2)$-dimensional linear lattice with edge-length e.g. $(1-0.01)\delta$ such that the stable- and Γ-faces of the elementary parallelepipeds of this lattice are cubes. (In the notion of this linear lattice again the fixed coordinate system in $U(x)$ is used.)

The following lemma, stating the existence of regular coverings, can be obtained using simple geometric arguments, see S–Ch (1987).

Lemma 3.5. *Let $x \in \partial M^*$ be a sufficient point and let U_0 be an open neighborhood of x in ∂M with the smooth foliation $U_0 = \bigcup_{\alpha \in B^{d-1}} \Gamma_\alpha$ as above (recall that the manifolds Γ_α are uniformly transversal to all possible local stable invariant manifolds). Then there are arbitrarily small neighborhoods U of x in ∂M having families of regular coverings with respect to the foliation $U = \bigcup_{\alpha \in B^{d-1}} (\Gamma_\alpha \cap U)$.*

Now we are in the position of formulating the "Transversal" Fundamental Theorem for semi-dispersing billiards generalizing the Fundamental Theorem of S–Ch (1987). As said in the Introduction, the present version is stronger because our form of the Ansatz is simpler and our condition of sufficiency is weaker and any transversal foliation $U_0 = \bigcup_{\alpha \in B^{d-1}} \Gamma_\alpha$ can be used instead of the partition into local unstable invariant manifolds. These improvements are important in applications, e.g. in the case of three billiard balls on tori [see K–S–Sz (1989-C)].

We introduce the following notation: $\partial^\Gamma(G_i^\delta)$ is the union of those $(2d-3)$-dimensional faces of G_i^δ which contain at least one Γ-face of G_i^δ. We call $\partial^\Gamma(G_i^\delta)$ the Γ-jacket of G_i^δ. The notion of the s-jacket $\partial^s(G_i^\delta)$ of G_i^δ is quite similar: It is the union of the remaining $(2d-3)$-dimensional faces of G_i^δ. (They are just those $(2d-3)$-

dimensional faces which contain at least one s-face of G_i^δ.) It is clear that $\partial(G_i^\delta) = \partial^\Gamma(G_i^\delta) \cup \partial^s(G_i^\delta)$. If the condition

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

is fulfilled for a point y then we say that the invariant manifold $\gamma^s(y)$ intersects the parellelepiped G_i^δ *correctly.*

Theorem 3.6 ("Transversal" Fundamental Theorem). *Suppose that*

(i) *Conditions 3.1–3.3 are fulfilled for the semi-dispersing billiard flow* $(M, \{S^t\}, \mu)$;
(ii) *the point* $x \in \partial M^* \setminus \bigcup_{n \geqq 0} \mathscr{R}^n$ *is sufficient;*
(iii) ε_1 *is a fixed constant between zero and one;*
(iv) *a smooth foliation* Γ *uniformly transversal to the local stable manifolds is given in a neighborhood* U_0 *of* x.

Then there exists a small neighborhood $U_{\varepsilon_1}(x)$ *of* x *in* ∂M *such that for every neighborhood* $U(x) \subset U_{\varepsilon_1}(x)$ *of* x *and for every family*

$$\mathscr{G}^\delta = \{G_i^\delta : i = 1, 2, \ldots, I(\delta)\} \quad (0 < \delta < \delta_0)$$

of regular coverings of $U(x)$ *the covering* $\mathscr{G}^\delta$ *can be divided into two disjoint subsets* $\mathscr{G}_g^\delta$ *and* $\mathscr{G}_b^\delta$ *such that*

(I) *for every parallelepiped* $G_i^\delta \in \mathscr{G}_g^\delta$ *and for every s-face* E^s *of* G_i^δ *the set*

$$\{y \in G_i^\delta : \varrho(y, E^s) < \varepsilon_1 \delta \text{ and } \exists\ (d-1)\text{-dimensional } \gamma^s(y) \text{ such that } \partial(G_i^\delta \cap \gamma^s(y)) \subset \partial^\Gamma(G_i^\delta)\}$$

has positive μ_1*-measure;*
(II)

$$\frac{\mu_1(\bigcup \mathscr{G}_b^\delta)}{\delta} \to 0$$

i.e.

$$\mu_1\Big(\bigcup_{G_i^\delta \in \mathscr{G}_b^\delta} G_i^\delta\Big) = o(\delta).$$

Remark 3.7. This theorem guarantees the existence of an ample set of points $y \in U(x) \subset \partial M$ with suitably long local stable invariant manifolds in arbitrarily small neighborhoods $U(x)$ of x in ∂M if only $x \in \partial M^*$ is a sufficient point. The following simple generalization is important in applications [see K–S–Sz (1989-C)]: instead of assuming that $x \in \partial M^*$ is sufficient we can assume that $x \in M^* \setminus \partial M$ is sufficient and $\mathscr{E} \subset M$ is a small ball-like C^2-smooth codimension one submanifold of the phase space M containing the point x in its interior and being transversal to the flow. In this (more general) setup every sufficiently small manifold $\mathscr{E}$ can be mapped to a neighborhood $U(T^+x) \subset \partial M$ of T^+x via the correspondence $y \to S^{\tau(y)}y$ $(y \in \mathscr{E})$, where the positive numbers $\tau(y)$ are close to $\tau(x)$ and $S^{\tau(y)}y = T^+y$. Using this mapping $\mathscr{E} \to U(T^+x)$ every geometric object in $U(T^+x)$ playing role in the Fundamental Theorem has its counterpart in $\mathscr{E}$ and the theorem itself remains true in $\mathscr{E}$ without any modification.

Remark 3.8. Assume that $x \in \partial M^* \cap \mathscr{R}^n$ $(n \geqq 0)$. The statement of the Fundamental Theorem also remains true in this case. We only need the following modification in (II):

$$\mu_1(\bigcup \{G_i^\delta : G_i^\delta \in \mathscr{G}_b^\delta,\ G_i^\delta \cap \mathscr{R}^n = \emptyset\}) = o(\delta) \qquad (\delta \to 0).$$

Remark 3.9. The grid condition (e) in the definition of regular coverings (Definition 3.4) is not necessary in the Fundamental Theorem, nonetheless in all applications it is enough to know that the theorem is true for every family of regular coverings satisfying the condition (e) too. The other reason for retaining the condition (e) is of didactics: the best way for constructing families of regular coverings (Lemma 3.5) is to begin with a linear lattice $\{w_i^\delta : i = 1, 2, \ldots, I(\delta)\}$ of centers of the parallelepipeds to be constructed. In this way the geometric and combinatorial structure of the covering $\mathscr{G}^\delta = \{G_i^\delta : i = 1, 2, \ldots, I(\delta)\}$ will be more transparent than it would be without condition (e).

In the applications one often uses two corollaries of the fundamental theorem whose formulations are less technical. The first one is called the Zig-zag Theorem in K–S–Sz (1989-B), and is also derived there from the fundamental theorem phrased for the case when the foliation Γ is $\Gamma^u = \{\gamma^u\}$. As a matter of fact, Γ^u is not a smooth foliation as Γ was supposed to be in Definition 3.4, but it is clear from all our proofs that about Γ its transversality to $\Gamma^s = \{\gamma^s\}$ and the continuity of $\mathscr{T}_{\gamma(y)}$ were only exploited.

Corollary 3.10 (Zig-zag Theorem). *Assume the conditions of the Fundamental Theorem for the flow $(M, \{S^t\}, \mu)$ and let the base-point $x \in \partial M^*$ be sufficient. Then there exist arbitrarily small neighborhoods $U(x)$ of x in ∂M such that for every null-set $N \subset U(x)$ there exists a set $A = A(N) \subset U(x)$ of full measure (i.e. $\mu_1(A) = \mu_1(U(x))$) such that we have:*

(i) $A \cap N = \emptyset$,
(ii) *for every pair of points $y, z \in A$ there exist two finite sequences $\gamma_1^s, \gamma_2^s, \ldots, \gamma_k^s$, and $\gamma_1^u, \gamma_2^u, \ldots, \gamma_k^u$ of local stable and unstable invariant manifolds in $U(x)$ such that*

(a) $y \in \gamma_1^s,\ z \in \gamma_k^u$;
(b) $\emptyset \neq \gamma_i^s \cap \gamma_i^u \subset A$ $(i = 1, 2, \ldots, k)$, $\emptyset \neq \gamma_i^u \cap \gamma_{i+1}^s \subset A$ $(i = 1, 2, \ldots, k-1)$.

We note that, because of transversality, the non-empty sets $\gamma_i^s \cap \gamma_i^u$ and $\gamma_i^u \cap \gamma_{i+1}^s$ must contain exactly one point.

Proof. First we prove the corollary assuming $x \in \partial M^0$, i.e. x does not belong to any manifold of singularity. At the end of the proof we shall briefly discuss the small modifications in the proof needed for the general case $x \in \partial M^*$.

Set $U(x)$ a suitably small neighborhood of x in ∂M having a family $\{\mathscr{G}^\delta : \delta < \delta_0\}$ of regular coverings, cf. Lemma 3.5. Let the null-set $N \subset U(x)$ be given. For every $\delta < \delta_0$ we consider the set G_g^δ of good parallelepipeds of the covering $\mathscr{G}^\delta$. Here G_g^δ is chosen in such a way that its elements are "good parallelepipeds" with respect to both the stable and the unstable versions of the Fundamental Theorem, $\mu_1\left(\bigcup_{G_i^\delta \in G_b^\delta} G_i^\delta\right) = o(\delta)$ and, moreover, in the application of the stable (unstable) version of the theorem the role of the transversal foliation Γ is played by the partition $\{\gamma^u\}$ (and $\{\gamma^s\}$, respectively). The actual choice of the *small* parameter

$\varepsilon_1 > 0$ is arbitrary. Two elements G_i^δ and G_j^δ of $\mathscr{G}_g^\delta$ are connected with an edge, by definition, if and only if $G_i^\delta \cap G_j^\delta \neq \emptyset$. Thus $\mathscr{G}_g^\delta$ turns out to be a graph. Its connected components are called clusters. One cluster of $\mathscr{G}_g^\delta$ with a maximal number of elements is denoted by $\mathscr{M}^\delta$.

Sublemma 3.11.

$$\mu_1(U(x)) - \mu_1\Big(\bigcup_{G_i^\delta \in \mathscr{M}^\delta} G_i^\delta\Big) \to 0 \quad (\text{as } \delta \to 0).$$

Proof. The statement of the sublemma is an easy consequence of a fact from combinatorics which we are going to formulate as follows. Let the unit cube $[0,1]^m \subset \mathbf{R}^m$ be divided into N^m small cubes

$$C_{k_1,\ldots,k_m} = \left\{(x_1, x_2, \ldots, x_m) \in \mathbf{R}^m : \frac{k_j}{N} \leqq x_j \leqq \frac{k_j+1}{N}, j = 1, 2, \ldots, m\right\}$$
$$(k_j = 0, 1, \ldots, N-1).$$

Let, moreover, $\mathscr{C}_g^N$ be a family of such cubes with the property

$$\lambda\Big(\bigcup_{C_{k_1,\ldots,k_m} \notin \mathscr{C}_g^N} C_{k_1,\ldots,k_m}\Big) = o(N^{-1}).$$

(Here λ is the Lebesgue measure in $\mathbf{R}^m$.) Two elements C and C' of C_g^N are connected with an edge iff $C \cap C' \neq \emptyset$. (The cubes are closed!) In this way C_g^N is also a graph. One of its connected components with a maximal number of vertices is denoted by H^N. The well-known combinatorial fact is that

$$\lambda\Big(\bigcup_{C \in H^N} C\Big) \to 1 \qquad (\text{as } N \to \infty).$$

This can be proved using an easy induction on the dimension m.

Now we return to the proof of the Zig-zag theorem. We denote by J the set of points $y \in U(x) \setminus \bigcup_{m=-\infty}^{+\infty} \mathscr{R}^m$ having $(d-1)$-dimensional invariant manifolds $\gamma^s(y)$ and $\gamma^u(y)$ such that y is their interior point. It is known from Ch (1982) that $\mu_1(J) = \mu_1(U(x))$.

By Sublemma 3.11 there exists a sequence $\delta_1, \delta_2, \delta_3, \ldots$ tending to zero such that the set $\liminf_{n\to\infty} [\cup \mathscr{M}^{\delta_n}]$ has full measure in $U(x)$.

Let

$$A_1 := (J \setminus N) \cap \Big[\liminf_{n\to\infty} \Big(\bigcup_{G_i^{\delta_n} \in \mathscr{M}^{\delta_n}} G_i^{\delta_n}\Big)\Big].$$

We obviously have that $\mu_1(A_1) = \mu_1(U(x))$. The essential statement (ii) of the Zig-zag Theorem without the inclusions $(\gamma_i^s \cap \gamma_i^u) \subset A$ and $(\gamma_i^u \cap \gamma_{i+1}^s) \subset A$ is a consequence of the goodness of the parallelepipeds from $\mathscr{M}^\delta$ (in the sense of both the stable and the unstable versions of the Fundamental Theorem), of the existence of the invariant manifolds for the points of A and, finally, of the connectedness of $\mathscr{M}^\delta$. (In the explanation of this statement some topology must be applied concerning the intersections of C^1-small perturbations of s-faces and u-faces of parallelepipeds.) In order to prove the existence of a subset $A \subset A_1$ with $\mu_1(A_1)$

$=\mu_1(A)$ such that the relations $\gamma_i^s \cap \gamma_i^u \subset A$ and $\gamma_i^u \cap \gamma_{i+1}^s \subset A$ hold for appropriate sequences $\gamma_1^s, \gamma_2^s, \ldots, \gamma_k^s$; $\gamma_1^u, \gamma_2^u, \ldots, \gamma_k^u$ we can consider an ample family of perturbed sequences of the original one and we can also use the absolute continuity (see K–S (1986)) of the foliations $\{\gamma^s\}$, $\{\gamma^u\}$ to gain the inclusion relations demanded in (ii) (in more detail see K–S–Sz (1989-A)).

Finally, we discuss the case $x \in \partial M^1$, say $x \in \mathscr{R}^n$, $n>0$. In this situation the smooth component $\mathscr{R}^n(x)$ of $\mathscr{R}^n$ containing x splits the neighborhood $U(x)$ into two semi-balls. Statement (II) of the stable version of Theorem 3.6 is only true for bad parallelepipeds, not intersecting $\mathscr{R}^n(x)$. The previous proof of the Zig-zag Theorem applies for both open semi-balls of $U(x)$. The only remaining task is to connect general pairs of points lying in different semi-balls of $U(x)$. This can be done by passing through $\mathscr{R}^n(x)$ along local *unstable* invariant manifolds, because the *unstable* version of Theorem 3.6 holds in $U(x)$ without restrictions ($n>0$).

Corollary 3.12. *Assume the conditions of the Fundamental Theorem for the flow* $(M, \{S^t\}, \mu)$ *and for the base point* $x \in (\partial M) \cap M^* = \partial M^*$. *Then there exists a neighborhood* $U(x)$ *of* x *in* ∂M *contained in one ergodic component of the system* $(\partial M, T, \mu_1)$.

Proof. Using Hopf's classical method and the Zig-zag Theorem we get the statement of this corollary in a straightforward way.

4. Elementary Geometric Facts about Semi-Dispersing Billiards

In this section we summarize some facts concerning the geometry of the invariant manifolds and the set of singularities of semi-dispersing billiards. All the statements will be formulated in terms of the Poincaré section map.

Lemma 4.1. *If the regularity of the set of degenerate tangencies (Condition 3.2) and of double singularities (Condition 3.3) hold, then, for any* $n \in \mathbf{N}$, *the set* Δ_n *of double singularities is the union of a finite number of compact codimension 2 smooth manifolds.*

Proof. It is sufficient to prove the statement of the lemma for the intersection of the set $S\mathscr{R}$ of singular reflections with its image under T^k ($1 \leq k \leq 2n$); applying an appropriate power of T we get the lemma. The proof is based on the following simple geometric fact:

Sublemma 4.2. *Every* $d-1$*-dimensional convex local orthogonal manifold* $\Sigma \subset \partial M$ *intersects the set of singular reflections in at most a finite number of* $(d-2)$*-dimensional smooth manifolds.*

The statement of the sublemma follows from the definition of the convexity of $\Sigma \subset \partial M$: it is the T^+ image of a convex local orthogonal manifold in M.

To prove Lemma 4.1 observe first that the set of singular reflections consists of a finite number of at most $(2d-3)$-dimensional open disks D_i. The regularity of the set of degenerate tangencies provides that D_i can be chosen in such a way that if D_i belongs to the set of "tangential reflections," then either $K(q)v=0$ for every $(q,v) \in D_i$ or $K(q)v \neq 0$ for every $(q,v) \in D_i$.

If D_i belongs to the set of "double reflections," then there exist two smooth pieces $\partial Q'$ and $\partial Q''$ of ∂Q such that $\pi(D_i) \subset \partial Q' \cap \partial Q''$. Let us fix a phase point $(q, v) \in D_i$. The set

$$\Sigma_q^k := \{T^k(q, v') : v' \in S_{d-1}, \|v' - v\| < \varepsilon_q\} \tag{4.3}$$

is a convex orthogonal manifold.

Letting q run over $\pi(D_i)$ and applying Sublemma 4.2 we get the statement of the lemma for D_i.

If D_i belongs to the set of "tangential reflections" then there is no such natural way of representing $T^k D_i$ as a union of convex orthogonal manifolds. Assume first that $K(q)v \neq 0$ for every $(q, v) \in D_i$.

For a fixed $(q, v) \in D_i$ consider the geodesic $g(s) \subset \partial Q : g(0) = q$, $g'(0) = v$ and s is the arclength-parameter. Since $K(q)v \neq 0$, $g''(s) = -\alpha(s)n(g(s))$, where $\alpha(s) > 0$ in a small interval $-s_0 \leqq s \leqq s_0$. This means that $g(s)$ is, up to the second order, a circle with radius $(\alpha(0))^{-1}$ which has its center on the half-line $\{q - \beta n(q), \beta \in \mathbf{R}_+\}$. It is well known from the theory of 2-dimensional dispersing billiards (see V (1982)) that there exists a "synchronizing" function $\tau(s) > 0$ $(-s_0 \leqq s \leqq s_0)$ such that the projection of the curve $\{\pi(S^{\tau(s)}(g(s), g'(s))), -s_0 \leqq s \leqq s_0\}$ on the plane spanned by the vectors v and $n(q)$ will be a convex planar curve. Moreover, if, for $|s| \leqq s_0$, the function $\tau(s)$ is sufficiently small then the curvature of the planar curve defined above is large enough. These facts imply

Sublemma 4.4. *The T^+-image of the manifold*

$$\Sigma_g := \{S^{\tau(s)}(g(s), v') : |s| \leqq s_0, \|v(s) - v'\| < \varepsilon_0, v' \in \mathscr{T}_{g(s)}\partial Q\} \tag{4.5}$$

is a convex local orthogonal manifold containing $T^+(S^{\tau(s)}(g(s), g'(s)))$ and, moreover, $T^-\Sigma_g \subset D_i$, where $T^-x := -T^+(-x)$.

Sublemma 4.4 provides that Sublemma 4.2 can be applied to $T^{k-1}T^+\Sigma_g$, so Sublemma 4.2 together with Condition 3.3 imply the statement of Lemma 4.1 for D_i.

If $K(q)v = 0$ for every $(q, v) \in D_i$, then TD_i is automatically of codimension 2. This statement is obvious if D_i itself is of codimension 2. In the case $\dim D_i = 2d - 3$ we can observe that D_i collapses along the trajectory of the flow therefore its image under T^+ must be of codimension 2. Hence the lemma.

Now we can estimate the measure of the phase points being simultaneously close to two smooth pieces of $\mathscr{R}^n$.

Lemma 4.6. *Denote by $\mathscr{R}_i$ the smooth components of $\bigcup_{|n| \leqq N} \mathscr{R}^n$ $(i = 1, \ldots, K)$. Then*

$$\mu_1\{x : \exists i, j \leqq K, i \neq j, \varrho(x, \mathscr{R}_i) < \delta, \varrho(x, \mathscr{R}_j) < \delta\} = o(\delta)$$

if N is fixed.

Proof. Lemma 4.1 implies that for every $\mathscr{R}_i$ and for $\nu_{\mathscr{R}_i}$-almost every $y \in \mathscr{R}_i$, $\varrho(y, \Delta_N) > 0$. Thus for every $\eta > 0$ there exists an open subset $\Delta_{\eta,i} \subset \mathscr{R}_i$ containing $\Delta_N \cap \mathscr{R}_i$ such that $\nu_{\mathscr{R}_i}(\Delta_{\eta,i}) < \eta$. Further, for every $\eta > 0$, there exists a $\delta > 0$ such that the neighborhoods $(R_i \setminus \Delta_{\eta,i})^{[\delta]} : 1 \leqq i \leqq K$ are disjoint, where $A^{[\eta]} := \{y : \exists x \in A$ such that $\varrho(x, y) < \eta\}$.

Therefore

$$\{x: \exists i,j \leq K, i \neq j, \varrho(x,\mathscr{R}_i)<\delta, \varrho(x,\mathscr{R}_j)<\delta\} \subset (\Delta_\eta)^{[\delta]} \subset \Delta_\eta^{[\delta]},$$

where $\Delta_\eta = \bigcup_{i=1}^{K} \Delta_{\eta,i}$. From the fact that the surface measure of Δ_η in $\bigcup_{|n|\leq N} \mathscr{R}^n$ is less than $K\eta$ we obtain that the μ_1-measure of the right-hand side of the previous inclusion is of order $K\eta\delta$. By choosing η sufficiently small the lemma follows.

In Sect. 3 (in Definition 3.4/(a)) we used the geodesic coordinate system in a sufficiently small neighborhood $U(x)$ of x in ∂M. Thus the tangent spaces of the submanifolds of $U(x)$ can be regarded as linear subspaces of $\mathbf{R}^{2d-2}$. Let us recall the definition of the angle of two linear subspaces $\mathscr{L}_1$ and $\mathscr{L}_2(\subset\mathbf{R}^{2d-2})$.

$$\operatorname{angle}(\mathscr{L}_1,\mathscr{L}_2) := \sup_{v_2\in\mathscr{L}_2} \inf_{v_1\in\mathscr{L}_1} \operatorname{angle}(v_1,v_2),$$

where angle (v_1, v_2) denotes the angle of two vectors v_1 and $v_2 \in \mathbf{R}^{2d-2}$. We restrict our considerations to angles between the tangent hyperplanes of two convex local orthogonal submanifolds of dimension $d-1$ and to those between a $(d-1)$-dimensional convex local orthogonal submanifold and a $(2d-3)$-dimensional smooth component of $\mathscr{R}^n$. In fact, we want to prove that the angles in question are sufficiently small; to this end we are looking for upper estimates of them.

Remark 4.8. Recall that the tangent space $\mathscr{T}_y\Sigma$ $(y\in\Sigma,\ y\notin\partial M)$ of a convex local orthogonal manifold is uniquely determined by $\mathscr{I}(y)$ and the second fundamental form $B_\Sigma(y)$:

$$\mathscr{T}_y\Sigma = \{(e,f): e\in\mathscr{I}(x), f=B_\Sigma(y)e\}$$

(see Sect. 2). Using this formula we see that angle $(\mathscr{T}_y\Sigma, \mathscr{T}_{y'}\Sigma')\to 0$ if $\|\tilde{B}(y)-\tilde{B}(y')\| \to 0$ and $\varrho(y,y')\to 0$ and a similar statement is true for the Poincaré map, too. (For the definition of $\tilde{B}$ cf. Proposition 2.7.)

Lemma 4.9. *For every $x\in\partial M^0$ and for every ε there exists a neighborhood $U(x)$ $\subset\partial M$ of x such that for every γ_1^s, γ_2^s and for every smooth $(2d-3)$-dimensional piece $\hat{\mathscr{R}}$ of $\mathscr{R}^n$ $(n>0)$ intersecting $U(x)$ and for every y_1, y_2, and $y_{\hat{\mathscr{R}}}$ $(\in U(x))$ lying on γ_1^s, γ_2^s, and $\hat{\mathscr{R}}$, respectively, we have*

(i) $\operatorname{angle}(\mathscr{T}_{y_1}\gamma_1^s, \mathscr{T}_{y_2}\gamma_2^s)<\varepsilon$,
(ii) $\operatorname{angle}(\mathscr{T}_{y_1}\gamma_1^s, \mathscr{T}_{y_{\hat{\mathscr{R}}}}\hat{\mathscr{R}})<\varepsilon$.

Proof. By virtue of the continuity of $\tilde{B}(y)$ (cf. Proposition 2.7) for every $\eta>0$ one can choose $\hat{U}(x)$ in such a way that for every $y_1, y_2\in U(x)$ $\|\tilde{B}(y_1)-\tilde{B}(y_2)\|<\eta$. Remark 4.8 implies inequality (i), if η is chosen small enough.

In order to prove (ii) it is sufficient to find a $(d-1)$-dimensional subspace $\mathscr{T}_{y_{\hat{\mathscr{R}}}}\Sigma$ of $\mathscr{T}_{y_{\hat{\mathscr{R}}}}\hat{\mathscr{R}}$, such that

$$\operatorname{angle}(\mathscr{T}_{y_1}\gamma_1^s, \mathscr{T}_{y_{\hat{\mathscr{R}}}}\Sigma)<\varepsilon.$$

This can be attained by constructing a convex local orthogonal manifold $-y_{\hat{\mathscr{R}}}\in\Sigma$ $\subset -\hat{\mathscr{R}}$, having a second fundamental form $B_\Sigma(-y_{\hat{\mathscr{R}}})$ close to $B(y_1)$. The construction of the convex orthogonal manifold Σ is the same as that of the manifolds Σ_q^k

and Σ_g in (4.3) and (4.5), respectively, for $-T^k\mu_{\hat{\mathscr{R}}}$, where $T^k y_{\hat{\mathscr{R}}}$ belongs to the set of singular reflections ($k=n$ or $k=n+1$).

(In the third case when $T^k y_{\hat{\mathscr{R}}}$ belongs to a tangential reflection and $K(q')\equiv 0$ in a neighborhood of $\pi(T^k y_{\hat{\mathscr{R}}})$ in ∂Q, then $\dim\hat{\mathscr{R}}<2d-3$ because of the collapse and the lemma does not refer to this case.)

For every point $x\in\partial M$ we denote by $z(x)$ the ϱ-distance of x from the compact set $\mathscr{R}\subset\partial M$, i.e.

$$z(x)=\min\{\varrho(x,y):y\in\mathscr{R}\}.$$

Lemma 4.10. *There exists a constant c_2 such that for every $\eta>0$,*

$$\mu_1\{x\in\partial M:z(x)<\eta\}\leqq c_2\eta.$$

Proof. The statement of the lemma is an obvious consequence of the definition of the Riemannian metric ϱ.

5. Proof of the Theorem Using the Tail Bound

The reader is reminded of our convention to denote by ∂M the extended boundary also including the transparent walls. Moreover, we will assume $x\in\partial M^0$ since the case ∂M^1 is analogous.

In the forthcoming proofs, a fundamental role will be played by the minimal expansion rate $\kappa_{n,\delta}(y)$ for the finite trajectory segment $T^{[0,n]}(-T^n y)$ in a small neighborhood of a phase point.

Definition 5.1. Consider now $y\in\partial M^0$. Denote first

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

where $D^n_{y,\Sigma}$ was defined in Sect. 2 (cf. Definition 2.9) and the inf is taken for convex local orthogonal manifolds Σ through $-T^n y$ in ∂M. (It is easy to see that the inf is attained for a $(d-1)$-dimensional hyperplanar orthogonal manifold in Q supplied with parallel velocities.) Denote, moreover,

$$\kappa_{n,\delta}(y)=\inf_{\Sigma}\inf_{w\in\Sigma}\|(D^n_{w,\Sigma})^{-1}\|^{-1}, \tag{5.2}$$

where the lower bound in Σ is taken for convex local orthogonal manifolds Σ through $-T^n y$ such that

(i) T^n is continuous on Σ,
(ii) $T^n\Sigma\subset B_\delta(-y)$.

For later use we formulate a simple property of $\kappa_{n,\delta}$.

Lemma 5.3. *$\kappa_{n,\delta}(y)$ is an increasing function of n.*

Proof. The monotonicity arises from two circumstances: from the trivial monotonicity in n of $\kappa_{n,0}(y)$ and from the observation that every submanifold Σ^{n+1} to be considered in the lower bound in (5.2) at the definition of $\kappa_{n+1,\delta}(y)$ is of the form $T^{-1}\Sigma^n$, where Σ^n is a submanifold to be considered in the lower bound at the definition of $\kappa_{n,\delta}(y)$. Hence the lemma.

We think of $U \ni x$ as being fixed though in course of the proof we shall put several conditions for its smallness. Then it makes sense to define the sets

$$U^g := \{y \in U : \forall n \in \mathbf{Z}_+, z(T^n y) \geqq (\kappa_{n,c_3\delta}(y))^{-1} c_3 \delta\},$$
$$U^b := U \backslash U^g,$$
$$U_n^b := \{y \in U : z(T^n y) < (\kappa_{n,c_3\delta}(y))^{-1} c_3 \delta\},$$

where c_3 is a suitable constant to be described later. Denote also

$$r(y) := \sup\{r : \partial(\gamma^s(y) \cap B_r(y)) \subset \partial B_r(y)\}.$$

The reason for introducing U^g (g for "good") is shown by our first lemma.

Lemma 5.4. *If $y \in U^g$, then $r(y) \geqq c_3 \delta$.*

Proof. The argument relies upon Sinai's classical philosophy: "expansion prevails over fractioning." We want to use the fact that $\gamma^s(y)$, if it exists at all and its dimension is $d-1$, is a superset of the $n \to \infty$ limit of the submanifolds $-\Sigma_0^n(y) := -T^n \Sigma_n^n(y)$, where $\Sigma_n^n(y)$ is a convex local orthogonal submanifold of ∂M through $-T^n y$ which, for definiteness, can be taken as the T^+-image of a $(d-1)$-dimensional hyperplanar manifold in Q of radius $c_3\delta$ supplied with parallel orthogonal and suitably oriented velocities [see P (1977) and Ch (1982)].

Now Lemma 5.4 is the consequence of the definitions of $\kappa_{n,\delta}(y)$ and U^g. Indeed, denote by $\tilde{\Sigma}_0^n$ the connected piece of $(-\Sigma_0^n(y)) \cap B_{c_3\delta}(y)$ containing y. Assume on the contrary that there exists a $w \in \partial \tilde{\Sigma}_0^n$ such that $w \notin \partial B_{c_3\delta}(y)$. This is only possible if $\varrho(y, w) < c_3\delta$ and, moreover, for some $k : 0 \leqq k \leqq n$, $w \in T^{-k}(S\mathscr{R})$. But on the other hand, for every $k : 0 \leqq k \leqq n$, $z(T^k y) \geqq (\kappa_{k,c_3\delta}(y))^{-1} c_3 \delta$, i.e. $\varrho(T^k w, T^k y) \geqq (\kappa_{k,c_3\delta}(y))^{-1} c_3 \delta$ implying $\varrho(w, y) \geqq c_3 \delta$, a contradiction. The convergence of manifolds $\tilde{\Sigma}_0^n$ follows from Proposition 2.7. Lemma 5.4 is proved.

Remark 5.5. An utmost important remark is that, if we choose c_3 depending on d and on x sufficiently large, then the property $r(y) \geqq c_3 \delta$, $y \in G_i^\delta$ will ensure that $\gamma^s(y)$ intersects G_i^δ correctly unless $\gamma^s(y)$ intersects $\partial^s G_i^\delta$.

In the proof of the theorem $F = F(\delta)$ will always denote a function $F : \mathbf{R}_+ \to \mathbf{Z}_+$ defined in a neighborhood of the origin such that $F(\delta) \nearrow \infty$ as $\delta \searrow 0$. For such a fixed function F we denote

$$U_\omega^b := \bigcup_{n > F(\delta)} U_n^b. \tag{5.6}$$

Tail Bound (Lemma 6.1). *For any fixed permitted function F*

$$\mu_1(U_\omega^b) = o(\delta). \tag{5.7}$$

While Sect. 6 is devoted to proving this statement we now prove the fundamental theorem by also using the Tail Bound.

Proof of Theorem 3.6. Fix $\varepsilon_1 > 0$. Consider a family $\mathscr{G}^\delta = \{G_i^\delta : 1 \leqq i \leqq I(\delta)\}$ of coverings satisfying the conditions of the theorem in a sufficiently small neighborhood $U(\subset \partial M)$ of x and fix a function F. Then we will say that G_i^δ is *bad*, i.e. $G_i^\delta \in \mathscr{G}_b^\delta$ iff

(i) either it intersects more than one smooth piece of the singularities of $T^{F(\delta)}$,
(ii) or it intersects not more than one smooth piece of the singularities of $T^{F(\delta)}$ and

there exists an s-face E^s of G_i^δ such that

$$\mu_1(G_i^\delta \cap (E^s)^{[\varepsilon_1 \delta]} \cap U_{ic}) \leqq \frac{\varepsilon_3}{4} \cdot \mu_1(G_i^\delta), \tag{5.8}$$

where ε_3 will be specified later and U_{ic} denotes the set of points in G_i^δ whose local stable invariant manifold intersects G_i^δ *correctly*, see the definition just before Theorem 3.6.

Otherwise we say that $G_i^\delta \in \mathscr{G}_g^\delta$ (g stands for *good*, again), thus $\mathscr{G}^\delta = \mathscr{G}_g^\delta \cup \mathscr{G}_b^\delta$.

We note that, in view of the transversality of the s-faces and the Γ-faces of the parallelepipeds, which is uniform in U, there exist two constants $0 < c_4 < c_5$ such that the ratio of the μ_1-measure of each product-type set arising in the forthcoming arguments and its "volume" calculated by multiplying the "volumes" of its factors lies between c_4 and c_5.

Now, in view of Lemma 4.6, it is easy to bound the total measure of parallelepipeds from $\mathscr{G}_b^\delta$ intersecting more than one smooth piece of the singularities of $T^{F(\delta)}$. Indeed, their total measure is, of course, $o(\delta)$ if $F(\delta)$ is increasing sufficiently slowly.

From now on we shall fix such a function F and will *only* consider parallelepipeds G_i^δ intersecting *not more than one* regular piece of singularity of $T^{F(\delta)}$.

In this case, by choosing U sufficiently small, we can assume that the order of this piece of singularity, if it exists at all, is sufficiently large ($\geqq n_0$, say), a fact that, by Lemma 4.9, also implies that the angle of this piece of singularity, denoted by $\hat{\mathscr{R}}$, and of any $\gamma^s(y) : y \in G_i^\delta$ is smaller than ε_2!

Give first a lower bound for $\mu_1(G_i^\delta \cap (E^s)^{[\varepsilon_1 \delta]})$, where E^s is an s-face of G_i^δ:

$$\mu_1(G_i^\delta \cap (E^s)^{[\varepsilon_1 \delta]}) \geqq c_4 (\varepsilon_1 \delta)^{d-1} \delta^{d-1} \geqq c_6 \varepsilon_1^{d-1} \mu_1(G_i^\delta), \tag{5.9}$$

where $c_6 > 0$. Choose $\varepsilon_3 = \varepsilon_3(\varepsilon_1)$ to satisfy $\varepsilon_3 < c_6 \varepsilon_1^{d-1}$.

The following observation is fundamental for the proof: by choosing $\varepsilon_2 = \varepsilon_2(\varepsilon_4)$ sufficiently small, it will be true that

if, G_i^δ intersects at most one regular piece of singularity of $T^{F(\delta)}$ and for a $y \in G_i^\delta$ with $G_i^\delta \in \mathscr{G}_g^\delta$, $\gamma^s(y)$ does not intersect G_i^δ correctly, then necessarily

$$y \in (G_i^\delta \cap (\hat{\mathscr{R}})^{[\varepsilon_4 \delta]}) \cup (G_i^\delta \cap (\partial^s G_i^\delta)^{[\varepsilon_4 \delta]}) \cup U_\omega^b. \tag{5.10}$$

For the proof of (5.10) see the method of constructing local invariant manifolds in Lemma 5.4 and the statement of Lemma 4.9.

As we did above, it is simple to give upper bounds for the measures of the first two sets appearing on the right-hand side of the previous inclusion. In fact,

$$\mu_1(G_i^\delta \cap (\hat{\mathscr{R}})^{[\varepsilon_4 \delta]}) \leqq c_6 c_5 \varepsilon_4 \delta \delta^{2d-3} \leqq \frac{\varepsilon_3}{4} \mu_1(G_i^\delta) \tag{5.11}$$

and

$$\mu_1(G_i^\delta \cap (\partial^s G_i^\delta)^{[\varepsilon_4 \delta]}) \leqq 2(d-1) c_5 \varepsilon_4 \delta \delta^{2d-3} \leqq \frac{\varepsilon_3}{4} \mu_1(G_i^\delta) \tag{5.12}$$

if only $\varepsilon_4 = \varepsilon_4(\varepsilon_3)$ is sufficiently small.

Observe first that, by (5.8), for $G_i^\delta \in \mathscr{G}_g^\delta$ statement (I) of the Theorem is evident.

Consider next a bad parallelepiped of type (ii) assuming that for some s-face E^s of G_i^δ (5.8) holds. Then from (5.9–12) we obtain

$$\mu_1(G_i^\delta \cap U_\omega^b) \geqq \mu_1(G_i^\delta \cap (E^s)^{[\varepsilon_1\delta]} \cap U_\omega^b) \geqq \frac{\varepsilon_3}{4} \cdot \mu_1(G_i^\delta).$$

But then

$$2^{2d-2}\mu_1(U_\omega^b) \geqq \Sigma' \mu_1(G_i^\delta \cap U_\omega^b) \geqq \frac{\varepsilon_3}{4} \Sigma' \mu_1(G_i^\delta),$$

where the summation Σ' is taken for bad parallelepipeds of type (ii). This chain of inequalities, however, combined with the Tail Bound immediately implies that for our F, too, $\Sigma'\mu_1(G_i^\delta) = o(\delta)$, thus providing the necessary bound for $\sum_{G_i^\delta \in \mathscr{G}_b^\delta} \mu_1(G_i^\delta)$.

Finally, by rethinking the proof, one can see that the parameters can really be chosen in the way indicated there. Indeed, ε_2 should be chosen as a function of ε_4, ε_4 as a function of ε_3, and, conclusively, ε_3 as a function of ε_1, the parameter figuring the theorem. Also, one can see that the neighborhood $U(x)$ should be small as a function of ε_1 (to intersect singularity lines of sufficiently high order, only). Hence Theorem 3.6.

6. Proof of the Tail Bound

The aim of this section is to prove

Lemma 6.1 (Tail Bound). *For any function $F(\delta)$ such that $F(\delta) \nearrow \infty$ as $\delta \searrow 0$,*

$$\mu_1(U_\omega^b) = o(\delta)$$

(as to the definition of U_ω^b see (5.6)*).*

Proof of Lemma 6.1. The proof is based on the following simple

Lemma 6.2. *Assume that, for any $\delta > 0$, $a_{n,m}^\delta \geqq 0$, $n, m \in \mathbf{Z}_+$ is a double array of numbers satisfying the following conditions:*

(i) *There exist numbers A_m such that for every $m \in \mathbf{Z}_+$ and $\delta > 0$,*

$$\sum_n a_{n,m}^\delta < A_m.$$

(ii) *For every $m \in \mathbf{Z}_+$,*

$$\lim_{\delta \to 0, N \to \infty} \sum_{n \geqq N} a_{n,m}^\delta = 0.$$

(iii)

$$\sum_m A_m < \infty.$$

Then

$$\lim_{\delta \to 0, N \to \infty} \sum_{n \geqq N} \sum_m a_{n,m}^\delta = 0.$$

Proof of Lemma 6.2. Let $\eta>0$. By choosing M sufficiently large, we have $\sum_{m>M} A_m<\eta$ and, consequently,

$$\sum_{n\geqq N}\sum_m a^\delta_{n,m}\leqq \sum_{m=0}^{M}\sum_{n\geqq N} a^\delta_{n,m}+\eta$$

uniformly in δ and N. Fix M large and apply (ii): then there exist δ_0 and N_0 such that for every $m\leqq M$, $\delta\leqq\delta_0$, and $N\geqq N_0$,

$$\sum_{n\geqq N} a^\delta_{n,m}<\frac{\eta}{M+1}.$$

Thus, if $\delta\leqq\delta_0$ and $N\geqq N_0$, then

$$\sum_{n\geqq N}\sum_m a^\delta_{n,m}<2\eta.$$

Hence the lemma.

Choose U and $\Lambda=\lambda^{-1}$ according to Lemma 2.13. Then, by Poincaré recurrence, almost every point of U infinitely often returns to U and, on U, we have the uniform lower bound Λ for the expansion rate. Lemma 6.1 will be established if we check the conditions of Lemma 6.2 for $a^\delta_{n,m}=\delta^{-1}\mu_1(U^b_{n,m})$, where

$$U^b_{n,m}:=\{y\in U: z(T^n y)<(\kappa_{n,c_3\delta}(y))^{-1}c_3\delta \text{ and } \kappa_{n,c_3\delta}(y)\in[\Lambda^m,\Lambda^{m+1})\}.$$

(The reader is reminded that each of the sets U^g, U^b, U^b_n, $U^b_{n,m}$ depends on the parameter δ.)

In fact, by Lemma 6.2,

$$\mu_1(U^b_\omega)\leqq \sum_{n\geqq F(\delta)}\sum_m \mu_1(U^b_{n,m})=o(\delta).$$

In checking the conditions of Lemma 6.2 the following lemma will play a decisive role.

Lemma 6.3. *For every $m\in\mathbf{Z}_+$, $0\leqq n_1<n_2$,*

$$T^{n_1}U^b_{n_1,m}\cap T^{n_2}U^b_{n_2,m}=\emptyset. \tag{6.4}$$

Proof of Lemma 6.3. Assume that, on the contrary, there exists $w\in T^{n_1}U^b_{n_1,m}\cap T^{n_2}U^b_{n_2,m}$. Then, from the definition of $U^b_{n,m}$, we have necessarily

$$T^{-n_1}w,\ T^{-n_2}w\in U, \tag{6.5}$$

and also

$$\kappa_{n_1,c_3\delta}(T^{-n_1}w),\ \kappa_{n_2,c_3\delta}(T^{-n_2}w)\in[\Lambda^m,\Lambda^{m+1}). \tag{6.6}$$

But any Σ permitted in the definition of $\kappa_{n_2,c_3\delta}(T^{-n_2}w)$ [cf. (5.2)] is also permitted in the definition of $\kappa_{n_1,c_3\delta}(T^{-n_1}w)$ and, moreover, because of (6.5) the minimal expansion rate of $T^{(n_2-n_1)}$ on the manifold $T^{n_1}\Sigma$ is at least Λ. This contradicts to (6.6). Lemma 6.3 is proved.

Now by Lemmas 6.3 and 4.10,

$$\begin{aligned}\delta^{-1}\sum_n \mu_1(U^b_{n,m})&=\delta^{-1}\sum_n \mu_1(T^n U^b_{n,m})=\delta^{-1}\mu_1(\cup_n T^n U^b_{n,m})\\ &\leqq \delta^{-1}\mu_1\{y:\exists n, z(y)<(\kappa_{n,c_3\delta}(T^{-n}y))^{-1}c_3\delta,\\ &\qquad \kappa_{n,c_3\delta}(T^{-n}y)\in[\Lambda^m,\Lambda^{m+1})\}\\ &\leqq \delta^{-1}\mu_1\{y: z(y)<\Lambda^{-m}c_3\delta\}\leqq (c_2\cdot c_3)\Lambda^{-m},\end{aligned}$$

thus immediately implying the fulfilment of conditions (i) and (iii) of Lemma 6.2. Its condition (ii) will follow from our following lemma:

Lemma 6.7. *For every* $m \in \mathbf{Z}_+$ *and every function* $F(\delta) \nearrow \infty$ $(\delta \searrow 0)$ *we have*

$$\sum_{n \geqq F(\delta)} \mu_1(U^b_{n,m}) = o(\delta).$$

Proof of Lemma 6.7. We use the following simple consequence of the Ansatz (Condition 3.1): For almost every $y \in \mathscr{R}$ (with respect to the Riemannian measure in $\mathscr{R}$)

$$\lim_{n \to \infty} \kappa_{n,0}(T^{-n}y) = \infty. \tag{6.8}$$

Denote by $\tilde{\mathscr{R}}$ the subset of those points of $\mathscr{R}$ where (6.8) holds. Since the subset of those points of $\mathscr{R}$ whose trajectory enters once more to $\mathscr{R}$ has $\nu_{\mathscr{R}}$-measure 0, we can also assume that for points of $\tilde{\mathscr{R}}$ both semi-trajectories never enter $\mathscr{R}$ again (cf. Lemma 4.1.)

Let now $m \in \mathbf{Z}_+$ be fixed. For every $y \in \tilde{\mathscr{R}}$ there exists an $n_0 := n_0(y, m)$ such that

$$\kappa_{n_0,0}(T^{-n_0}y) > \Lambda^{m+2}.$$

Then, because of the continuity of $\kappa_{n,0}(T^{-n}y)$, for every $y \in \tilde{\mathscr{R}}$ there exists a neighborhood $V_1 := V_1(y, m) \subset \partial M$ of y such that, for every $w \in V_1$,

$$\kappa_{n_0,0}(T^{-n_0}w) > \Lambda^{m+2},$$

where $n_0 = n_0(y, m)$. Now, also by continuity, and by the definition of $\kappa_{n,\delta}(y)$, for every $y \in \tilde{\mathscr{R}}$, there exists a neighborhood $V_2 := V_2(y, m)\ (\subset \partial M)$ of y and there exists a positive number $\delta_1 := \delta_1(y, m)$ such that for every $\delta < \delta_1$ and for every $w \in V_2$,

$$\kappa_{n_0,\delta}(T^{-n_0}w) > \Lambda^{m+2}.$$

Finally, since $\kappa_{n,\delta}(T^{-n}w)$ increases in n, we also have that for every $\delta < \delta_1$, every $w \in V_2$ and every $n \geqq n_0$,

$$\kappa_{n,\delta}(T^{-n}w) > \Lambda^{m+2}.$$

Next we claim that for every $y \in \tilde{\mathscr{R}}$ and every $\delta < (c_3)^{-1}\delta_1$,

$$\left(\bigcup_{n \geqq n_0} T^n U^b_{n,m}\right) \cap V_2 = \emptyset.$$

Indeed,

$$T^n U^b_{n,m} \subset \{w : \kappa_{n,c_3\delta}(T^{-n}w) \in [\Lambda^m, \Lambda^{m+1})\},$$

thus the claim holds true if $n \geqq n_0$, and $c_3\delta < \delta_1$.

Because of regularity, for every $\eta > 0$, there exists a compact subset K_η of $\tilde{\mathscr{R}}$ such that $\nu_{\mathscr{R}}(\tilde{\mathscr{R}} \setminus K_\eta) < \eta$. Then, because of the compactness of K_η, one can choose a finite subset $\{y_1, \ldots, y_l\} \subset K_\eta$ such that

$$\bigcup_{i=1}^{l} V_2(y_i) \supset K_\eta,$$

implying the existence of a $\delta_\eta > 0$ such that

$$\bigcup_{i=1}^{l} V_2(y_i) \supset K_\eta^{[\delta_\eta]}.$$

We can assume $\delta_\eta \leqq \min_{1 \leqq i \leqq l} \delta_1(y_i, m)$. Let $N_\eta := \max_{1 \leqq i \leqq l} n_0(y_i, m)$.

Then, for every $\delta < \frac{\delta_\eta}{c_3}$,

$$\bigcup_{n \geqq N_\eta} T^n U^b_{n,m} \cap \bigcup_{i=1}^{l} V_2(y_i, m) = \emptyset,$$

and hence

$$\bigcup_{n \geqq N_\eta} T^n U^b_{n,m} \cap K_\eta^{[\delta_\eta]} = \emptyset.$$

But, on the other hand

$$T^n U^b_{n,m} \subset \{w \in \partial M : z(w) < c_3 \delta\} \subset \mathscr{R}^{[c_3\delta]}. \tag{6.9}$$

Consequently, whenever $c_3\delta < \delta_\eta$, we have

$$\bigcup_{n \geqq N_\eta} T^n U^b_{n,m} \subset (\mathscr{R} \setminus K_\eta)^{[c_3\delta]}. \tag{6.10}$$

Now, by Lemma 6.3

$$\sum_{n \geqq N_\eta} \mu_1(T^n U^b_{n,m}) = \mu_1\left(\bigcup_{n \geqq N_\eta} T^n U^b_{n,m}\right),$$

and, moreover, by (6.10) the right-hand side is at most of the order $\delta \cdot v_{\mathscr{R}}(\mathscr{R} \setminus K_\eta) < \delta \cdot \eta$ whenever $c_3\delta < \delta_\eta$. Hence Lemma 6.7 and simultaneously the Tail Bound are proved.

7. An Application: Ergodicity of Two Billiards

Originally, various forms of the fundamental theorem were proved for two-dimensional dispersing billiards with a finite horizon [see S (1970), B–S (1973), G (1975) and also K–S–Sz (1989-A)]. The efficiency of Chernov-Sinai type fundamental theorems is partially explained by the fact that they also cover the multidimensional case without assuming the finiteness of the horizon. This is why the Chernov-Sinai theory implies important new results for dispersing billiards, too, namely their ergodicity (and, moreover, their K-property).

Theorem 7.1. *Every dispersing billiard with $Q \subset \mathbf{T}^d$ $(d \geqq 2)$ and satisfying Condition 2.1 is a K-system.*

We notice that the case $Q \subset \mathbf{R}^d$, Q compact is also contained in this theorem.

Proof. Denote by ∂M the extended boundary of the phase space M as introduced in Sect. 2. Then the horizon is, of course, finite. It will be convenient to write ∂M as $\bigcup_{i=1}^{l} \partial M_i$, where ∂M_i is a smooth cell of dimension $2d-2$. Since the billiard is dispersing, every orbit which ever enters $\partial M \backslash \partial(\partial M)$ is sufficient. On the other hand, the subset of points in ∂M whose trajectories hit singularities ($\mathcal{SR}$) more than once form a codimension 2 subset [see Lemma 4.15 in K–S–Sz (1989-C)]. For dispersing billiards the regularity Conditions 3.2 and 3.3 are obvious. Assume the Ansatz for a moment. Then Corollary 3.12 of the fundamental theorem combined with the two aforementioned facts immediately implies that each ∂M_i lies in one ergodic component. Thus global ergodicity follows by connecting neighboring ∂M_i-*s* with beams of trajectories of positive measure step by step. This can be done since Q is connected.

Next we verify the Ansatz. We have to show that the orbit of $\nu_{\mathcal{R}}$-almost every point of $\mathcal{SR}$ enters $(\partial M \backslash \mathcal{SR}) \cap \{(q, v) \in \partial M : K_q > 0\}$ (virtual walls are excluded!) at least once in the past and once in the future. If the positive semitrajectory of a point (q, v) never enters the previous set, then necessarily $\{q + tv(\mathrm{mod}\, \mathbf{T}^d) : t \geqq 0\}$ is not dense in $\mathbf{T}^d$. But it is well known (see e.g. Lemma 3.1.1 of K–S–F (1980) that then $v_1, \ldots, v_d$ are rationally dependent, i.e. there exist integers $n_1, \ldots, n_d$ such that

$$n_1 v_1 + \ldots + n_d v_d = 0 .$$

Consequently, for any fixed q, the velocities for which $\{q + tv(\mathrm{mod}\, \mathbf{T}^d) : t \geqq 0\}$ is not dense in $\mathbf{T}^d$ form a countable union of codimension 1 submanifolds in S_{d-1} and thus their subset $(\subset S_{d-1})$ has measure 0. Therefore, by Fubini's theorem, the Ansatz follows. Hence the theorem.

An interesting particular case of the previous theorem gives the K-property of two-billiards. Consider the dynamical system of two elastic balls of radius R given on the ν-torus $\mathbf{T}^\nu : \nu \geqq 2$ whose motion is uniform with elastic collisions. Assume the conservation laws have the form

$$\begin{gathered} q_1 + q_2 = 0 , \\ v_1 + v_2 = 0 , \quad v_1^2 + v_2^2 = 1 , \end{gathered} \tag{7.2}$$

where (q_i, v_i) are the phase points of the balls $(i = 1, 2)$. Then it is easy to see that the system of two balls is isomorphic to a dispersing billiard whose phase point is (q_1, v_1) and whose phase space is $Q \times \frac{1}{\sqrt{2}} S_{\nu-1}$, where $Q \subset \mathbf{T}^\nu$ can be obtained as the subset of $\mathbf{T}^\nu$ satisfying

$$\mathrm{dist}_{\mathbf{T}^\nu}(q_1, -q_1) \geqq 2R \tag{7.3}$$

(hard core condition!).

Equivalently saying we obtain Q by cutting out from $\mathbf{T}^\nu$ 2^ν spheres of radius R. If $\nu \geqq 5$ then Q is connected provided $R < 1/2$; if $2 \leqq \nu \leqq 4$, then Q is connected for $R < \frac{\sqrt{\nu - 1}}{4}$ but Q contains 2^ν connected components for $R > \frac{\sqrt{\nu - 1}}{4}$. In any case, the corresponding billiard system is a dispersing one, satisfying Conditions 2.1 and 3.1–3, and we have

Theorem 7.4. *The system of two billiard balls of radius R on $\mathbf{T}^\nu$ restricted to the phase space defined by the conservation laws* (7.2) *and the hard core condition* (7.3) *is*

(i) *a K-system if $\nu \geqq 5$ or $2 \leqq \nu \leqq 4$ and $R < \frac{\sqrt{\nu - 1}}{4}$;*

(ii) *a K-system on each of the* 2^ν *connected components of the phase space if* $2 \leqq \nu \leqq 4$ *and* $\frac{\sqrt{\nu-1}}{4} < R < \frac{1}{2}$.

Acknowledgements. D. Szász expresses his sincere gratitude to John Mather, Tom Spencer and Arthur Wightman for their kind hospitality during his visit at the Department of Mathematics of Princeton University and in the School of Mathematics of the Institute for Advanced Study in the spring of 1989. Special thanks are due to Tom Spencer for encouragement to write this paper.

The authors are deeply indebted to Kolya Chernov and Carlangelo Liverani for their careful reading of the manuscript and for stimulating discussions.

References

B–S (1973) Bunimovich, L.A., Sinai, Ya.G.: On the fundamental theorem of dispersing billiards. Math. Sb. **90**, 415–431 (1973)

C (1982) Chernov, N.I.: Construction of transversal fibers for multidimensional semi-dispersing billiards. Funkt. Anal. i. Pril. **16**, 35–46 (1982)

G (1975) Gallavotti, G.: Lectures on the billiard. Lecture Notes in Physics, vol. 38. Moser, J. (ed.), pp. 236–295. Berlin, Heidelberg, New York: Springer 1975

G (1981) Galperin, G.: On systems of locally interacting and repelling particles moving in space. Trudy MMO **43**, 142–196 (1981)

I (1988) Illner, R.: On the number of collisions in a hard sphere particle system in all space. Technical Report (1988)

K–S (1986) Katok, A., Strelcyn, J.-M.: Invariant manifolds, entropy and billiards, smooth maps with singularities. Lecture Notes in Mathematics, vol. 1222. Berlin, Heidelberg, New York: Springer 1986

K–S–F (1980) Kornfeld, I.P., Sinai, Ya.G., Fomin, S.V.: Ergodic theory. Moscow: Nauka 1980

K–S–Sz (1989-A) Krámli, A., Simányi, N., Szász, D.: Dispersing billiards without focal points on surfaces are ergodic. Commun. Math. Phys. **125**, 439–457 (1989)

K–S–Sz (1989-B) 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)

K–S–Sz (1989-C) Krámli, A., Simányi, N., Szász, D.: Three billiard balls on the ν-dimensional torus is a K-flow (submitted to Ann. Math.)

P (1977) Pesin, Ya.B.: Lyapunov characteristic exponents and smooth ergodic theory. Usp. Mat. Nauk. **32**, 55–112 (1977)

S (1970) Sinai, Ya.G.: Dynamical systems with elastic reflections. Usp. Mat. Nauk. **25**, 141–192 (1970)

S (1979) Sinai, Ya.G.: Ergodic properties of the Lorentz gas. Funkcional. Anal. i. Pril. **13**, 46–59 (1979)

S–Ch (1987) Sinai, Ya.G., Chernov, N.I.: Ergodic properties of some systems of 2-D discs and 3-D spheres. Usp. Mat. Nauk. **42**, 153–174 (1987)

V (1982) Vetier, A.: Sinai billiard in a potential field (construction of stable and unstable fibers). Colloquia Soc. Math. J. Bolyai **36**, 1079–1146 (1982)

Communicated by J. N. Mather

Received August 8, 1989