ANNALES DE L'I. H. P., SECTION A

LEONID A. BUNIMOVICH
JAN REHACEK

On the ergodicity of many-dimensional focusing billiards

Annales de l'I. H. P., section A, tome 68, nº 4 (1998), p. 421-448
<http://www.numdam.org/item?id=AIHPA_1998__68_4_421_0>

Ann. Inst. Henri Poincaré,

Vol. 68, n° 4, 1998, p. 421-448. *Physique théorique*

On the ergodicity of many-dimensional focusing billiards

by

Leonid A. BUNIMOVICH

Southeast Applied Analysis Center School of Mathematics
Georgia Institute of Technology Atlanta, GA 30332

and

Jan REHACEK

Center for Nonlinear Studies, MS-B258 Los Alamos
National Laboratory Los Alamos, NM87545

ABSTRACT. – The affirmative answer is given to the long standing question whether there exist nowhere dispersing ergodic billiards in dimensions greater than two. To do this we construct a class of n-dimensional domains, $n \geq 3$, generating ergodic billiards (that are also Bernoulli systems). © Elsevier, Paris.

Key words: chaos, ergodic billiards, focusing billiards, hyperbolicity.

RÉSUMÉ. – Nous répondons de façon positive à la question posée de longue date de savoir s'il existe des billiards ergodiques nulle part dispersifs en dimensions supérieures ou égales à trois. Dans ce but, nous construisons une classe de domaines en dimension $n \geq 3$ engendrant des billiards ergodiques (qui sont de plus des systèmes de Bernoulli). © Elsevier, Paris.

1. INTRODUCTION

It is well known that chaotic properties of dynamical systems of physical origin are generated by hyperbolicity. Sometimes hyperbolicity is referred to as sensitive dependence on initial conditions. It means that nearby trajectories exponentially diverge in a phase space.

Historically this type of motion was first observed and proved to exist in the geodesic flows on surfaces of negative curvature **[H]**. Then the results were essentially generalized by Anosov, Sinai and Smale **[A]**, **[S]**, **[AS]**, **[Sm]**.

However, all these papers deal with smooth dynamical systems that exhibit uniform hyperbolicity. The last means that the rates of divergence of trajectories at different points of a phase space vary smoothly and their ratios are bounded away from zero as well as from infinity. In dynamical systems that appear in physical models, however, the hyperbolicity is usually nonuniform.

The most visible and classical examples of such models are billiards. Billiards naturally appear as models in statistical mechanics, optics, acoustics etc. The theory of hyperbolic billiards was essentially developed by Sinai and his school. He introduced the class of billiards with smooth dispersing (convex inwards) boundaries **[S1]**, which were later called Sinai billiards and proved that such billiards are ergodic and B-systems in any (finite) dimension. These results were extended to dispersing 2D billiards with nonsmooth boundaries in **[BS]**.

It was the general belief for a long time, both in the mathematical and physical communities, that billiards with focusing components always demonstrate a regular, rather than chaotic behavior. This ideology has been confirmed by Lazutkin's proof **[L]** of the abundance of caustics for billiards in the convex 2D domains with smooth boundaries.

It has been discovered, however, that this ideology is wrong, because there exists another mechanism of hyperbolicity that works in some billiards with focusing boundaries **[B1]**. This mechanism was called the mechanism of defocusing. Later it has been shown that the same mechanism generates chaotic (hyperbolic) behavior in some geodesic flows on surfaces of positive curvature **[D1,2]**, **[BG]**. However, all the examples studied in these papers are two dimensional.

For more than 20 years the problem, whether the mechanism of defocusing can generate hyperbolicity in dynamical systems with dimension greater than two, was open. The affirmative answer to these questions was claimed in **[B2]**, where the construction of corresponding high-dimensional chaotic focusing billiards was outlined. In **[W1]** was constructed an example of a linearly stable periodic trajectory in some 3D region whose boundary contains semispheres. This demonstrates again that the problem under consideration is rather delicate. The rigorous proofs of the existence of high dimensional focusing billiards with nonzero Lyapunov exponents were given in the recent papers **[BR1,2]**. Here we prove ergodicity and B-property for billiards with spherical caps that are smaller than 60° (as opposed to 90° in **[BR1,2]**). This additional restriction seems to be mostly of a technical nature. However, it may have some implications for the structure of the spectrum of the corresponding quantum systems.

The main difficulty in studying high dimensional focusing billiards lies in the fact that focusing is very weak in some directions. This phenomenon is called astigmatism and the corresponding elementary formulas were discovered more than 160 years ago (see **[C]** and the formula (2.2) below). Bunimovich's claim made in **[B2]** was also made because of this formula. However, in this paper we need to impose even stronger restrictions on the size of spherical caps. Therefore the mechanism of defocusing discovered in **[B1]** works in such billiards in much more subtle manner.

2. DESCRIPTION OF THE BILLIARD AND THE MAIN RESULT

We will study billiards in some class of $n-$dimensional regions ($n \geq 2$) bounded by flat walls and spherical caps. By a spherical cap we mean a piece of sphere bounded by some plane. It is our aim to prove that for some class of such regions the billiard motion is a B-system.

Let $B \subset R^n$ be a region described above whose boundary is equipped with a field of inward unit normal vectors $n(q)$ (see Fig. 1). The billiard system is then realised by a point particle, moving with a unit velocity inside this region and being reflected according to the law "the angle of incidence equals the angle of reflection" at the boundary. In terms of the velocity of the particle this means, that the tangent component of it remains the same after the reflection, while the normal component changes the sign, according to the rule

$$v_+ = v_- - 2(n(q), v_-)n(q).$$

The phase space $\mathcal{M}$ of such a system is the restriction of the unit tangent bundle of R^n to B and we'll use the standard notation $x = (q, v)$, where x is the phase point, q is the point in the configuration space and v is the unit velocity vector. The billiard flow preserves the Liouville measure $d\nu = dq \wedge d\omega$, where dq and $d\omega$ are Lebesgue measures on B and the unit sphere respectively. This flow will be denoted by S^t, thus yielding the continuous dynamical system $(\mathcal{M}, S^t, \nu)$.

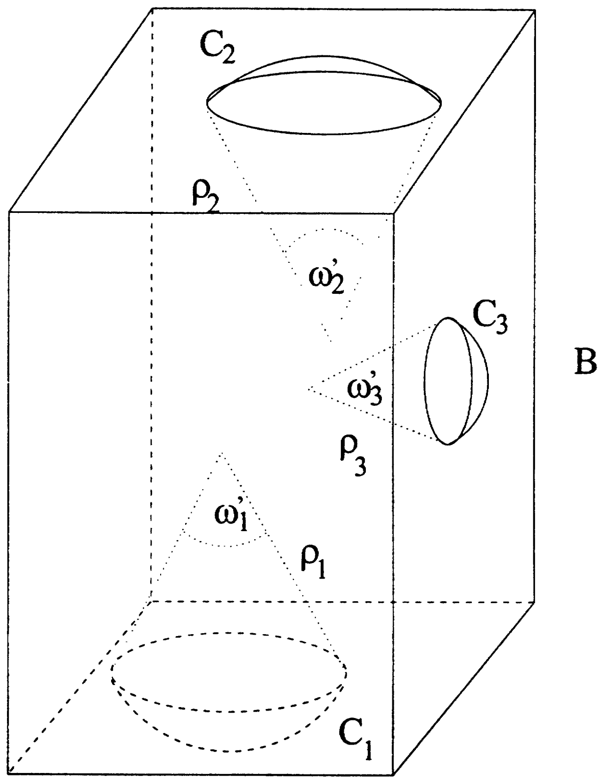

Fig. 1. – Billard region B.

As is customary for billiard systems, rather than studying the dynamics of the continuous system, we will be studying the so called billiard map. Then many properties, e.g. ergodicity, of a billiard flow follow automatically from the corresponding properties of a billiard map. Let us denote

$$M = \{x = (q, v), q \in \partial B, (v, n(q)) > 0\}.$$

The projection onto the configurational space will be denoted by π, i.e. $\pi(x) = q$. For $x = (q, v) \in M$ we define $\tau(x)$ to be the first positive moment of intersection of the billiard orbit determined by x with the boundary. Then $Tx = (q', v') = S^{\tau}x$, so that q' is the point of next reflection and v' is the outcoming velocity vector at that point.

Since very often we will be working with the $(n - 1)$-dimensional hypersurface γ, in the rest of the paper we will adopt the convention $m = n - 1$. Not to confuse the reader with dimensions, let us repeat that the boundary of our billiard region $B \subset R^n$ consists of m dimensional manifolds (flat walls and parts of the m dimensional sphere). A subspace, perpendicular to an orbit, has also the dimension $m = n - 1$. The n-dimensional billiard mapping T preserves the projection of the Liouville measure on the boundary

$$d\mu(q, v) = const.(v, n(q)).dq \wedge d\omega,$$

where dq is the m-dimensional Lebesgue measure on the boundary ∂B generated by the volume and $d\omega$ is the m-dimensional Lebesgue measure on the unit sphere. The *const* is the usual normalizing constant so that $\mu(M) = 1$.

In order to study the behavior in the vicinity of a given billiard trajectory, we introduce a concept of an m-dimensional infinitesimal control surface γ (also called a wavefront) of class C^2 perpendicular to the orbit. The rate at which the neighboring trajectories diverge is defined by the curvature operator (the operator of the second fundamental form) of the surface γ. The reflection from the focusing components of the boundary may cause that some of the principle curvatures are negative. However, during the free path the corresponding families of trajectories defocus and from then on diverge from an orbit under consideration, i.e. from a certain point on the free path (the conjugate point) the curvature operator is positive definite. To demonstrate that the defocusing mechanism works in many dimensions, we show that if the surface γ approaches a spherical cap with the positive definite curvature operator, then the whole billiard region can be configured in such a way, that after passing through the spherical cap, the surface γ will focus "relatively soon" (and thus the curvature operator becomes positive definite again). We will specify this in the proof of Lemma 1 (see also **[BR1,2]**). This property ensures that the mechanism of defocusing discovered in **[B2]** for 2-D billiard works also for some n-dimensional regions.

It is known and obvious that a part of the orbit reflecting from a given spherical cap lies in the same $2D$ plane, which contains the center of the corresponding sphere. We denote this plane by P. This plane then defines the unique direction on a hyperplane U, perpendicular to the orbit. We will call this unique direction a planar direction or a planar subspace, while its orthogonal complement will be called an orthogonal or transversal subspace. The hyperplane U thus naturally splits into

$$U = U_p \oplus U_t,$$

where $U_p = U \cap P$ and U_t is the $(m - 1)$-dimensional orthogonal complement to U_p in U. Since the control surface γ is tangent to U

the planar direction and the orthogonal directions are naturally defined on γ. Let us mention, however, that these directions are not intrinsic properties of the control surface and can be defined only with respect to a given cap.

For understanding the dynamics of the system in the vicinity of the given orbit it is important to know how the curvatures of the control surface γ evolve. In the planar direction they change just like in the 2D billiards

$$\kappa^+ = \kappa^- - \frac{2k}{cos(\phi)}. \tag{2.1}$$

In the orthogonal subspace, however, they obey (**[C]**, p.66)

$$\kappa^+ = \kappa^- - 2kcos(\phi), \tag{2.2}$$

where κ^+ (κ^-) is the curvature after (before) the reflection, ϕ is the angle of reflection and k is the curvature of the boundary. Finally, during the free path the curvature evolution is known to be

$$\kappa(t) = \frac{\kappa(0)}{1+\kappa(0)t}. \tag{2.3}$$

The main difficulty in dealing with billiard dynamics is the existence of singular points, i.e. points $x \in M$ at which the mapping T (or T^{-1}) is not defined or is discontinuous. In our case of focusing billiard there are no "tangencies" (as in dispersing billiards) and so we have to deal only with "multiple reflections".

Let us denote by S_{-1} the set of points where the mapping T is not defined (points whose trajectory hits intersections of regular boundary components) and by S_1 the points where T^{-1} is undefined (points $x = (q, v)$ with q in these intersections). Thus points from S_1 can be further iterated by T, giving rise to manifolds where higher iterates of T^{-1} are not defined.

More precisely, for a positive integer l we define the set of points where T^{-l} is not defined by

$$S_l = S_1 \cup T(S_1) \cup ... \cup T^{l-1}(S_1)$$

and similarly

$$S_{-l} = S_{-1} \cup T^{-1}(S_{-1}) \cup ... \cup T^{-l+1}(S_{-1})$$

is the set, where T^l is not defined. Further let $S_\infty = \cup_{n>0} S_n$ and $S_{-\infty} = \cup_{n<0} S_n$. Their intersection $S = S_{-\infty} \cap S_\infty$ is the set of points whose orbit terminates both in the past and in the future.

Finally, we denote the set of points for which either forward or backward trajectory can be defined by $M_0 = M \backslash S$.

We now formulate the main result of this paper. Let ω' be the angular size of a given spherical cap C and ρ its radius.

DEFINITION 1. – The *zone of focusing* of a given spherical cap is a part of the billiard region bounded by the flat wall to which the spherical cap is attached, by the transparent wall parallel to it and intersecting the center of the spherical cap and, finally, by surrounding flat walls (in Fig. 2 the transparent walls associated with the cap C_i are denoted by V_i).

DEFINITION 2. – Let $Q \subset R^{n-1}$ be a regular polyhedron, which tiles the whole space R^{n-1}. A region $B \in R^n$ will be called *type (1)* region, if it consists of a product $Q \times [0, a]$ for some $a > 0$ and all the spherical caps are attached only to the sides $Q \times \{0\}$ and $Q \times \{a\}$ (these walls will be called principal faces). Moreover, the intersection of the spherical cap with the flat face should be "inscribed" to the polyhedron Q, i.e. all the faces of the polyhedra should be tangent to this intersection. Finally, their zones of focusing have to be disjoint, i.e. $\rho_1 + \rho_2 < a$, where ρ_1 and ρ_2 are radii of the spherical caps attached to the principal faces. If one of them holds more than one cap, then we take maximum of all the possible radii.

DEFINITION 3. – A region $Q \subset R^n$ will be called *type (2)*, if it consists of a rectangular box with spherical caps attached to some of its faces. More precisely, the type (2) region must have at least one spherical cap attached to each pair of opposite sides of the rectangular box. However, any billiard trajectory going from one cap to the other must, at some point, be outside of both zones of focusing. Thus in Fig. 2 the caps C_1 and C_2 are improperly placed (the full line represents part of the billiard orbit that is inside zones of focusing during traversing from C_1 to C_2), while either of them can coexist with the spherical cap C_3 in the type (2) region. In addition to it, the type (2) region must have at least one spherical cap attached to each pair of opposite sides of the rectangular box.

DEFINITION 4. – A phase point $x = (q, v)$ will be called *essential* if its forward or backward semitrajectory has at least one reflection from a spherical cap. We denote by M_1 the set of all essential points.

The main result of this paper is the following.

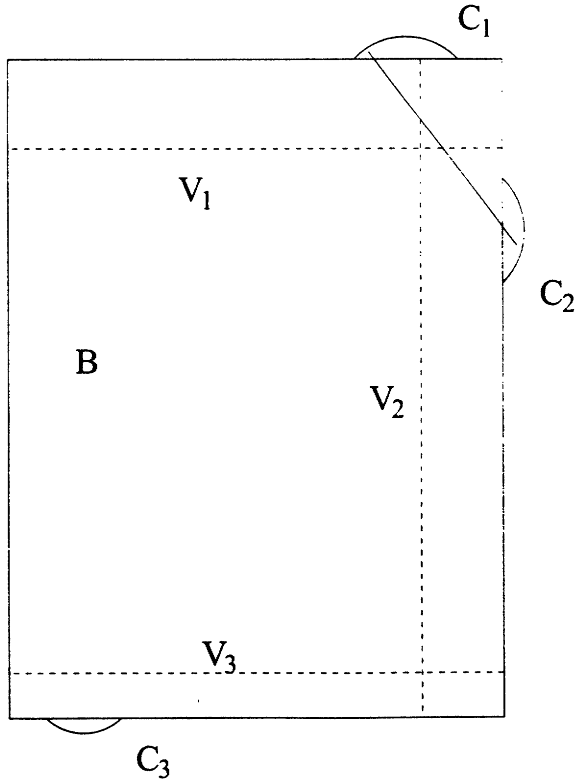

Fig. 2. – Caps on the side.

THEOREM 1. – *Let B be a billiard region of the type (1) or (2) and let all the spherical caps attached to the flat walls have angular size smaller than* $60°$. *Then billiard system in this region is ergodic.*

The full-blown proof of ergodicity for billiard systems uses techniques developed over the years in **[BS], [SCh], [KSS], [LW], [Ch], [M]**, which are based on the original Sinai's approach **[S1]**. It consists in showing that the expansion of local unstable manifolds prevails over the fractioning of these by images of singularity manifolds. This allows to deduce that in every neighborhood of a phase point x there is a sufficient number of long enough smooth stable and unstable fibers that are necessary to carry the construction of Hopf's chain of stable and unstable manifolds.

In order to establish the existence of these manifolds we use a combination of techniques of invariant cones and of monotone quadratic forms (see **[LW], [M]**). The function $Q : TM \to R$ is a quadratic form if Q_x is a quadratic form on T_xM for any $x \in M$. We define these forms in the proof of Lemma 1. Given this quadratic form we can define the following sets

$$E_x^s = \{u \in T_xM, Q(DT^n u) < 0, n \geq 0\},$$

$$E_x^u = \{u \in T_xM, Q(DT^n u) > 0, n \leq 0\},$$

which are contracting and expanding subspaces and

$$C_x^- = \{u \in T_xM, Q_x u \leq 0\},$$

$$C_x^+ = \{u \in T_xM, Q_x u \geq 0\},$$

which are called contracting and expanding cones.

We base our proof on the abstract versions of the so-called Fundamental Theorem of the billiards theory (see for instance **[Ch], [M], [LW]**) that establishes local ergodicity provided certain conditions are met. Our approach is essentially based on **[Ch]**.

Fundamental Theorem: Assume that the billiard system (M, T, μ) satisfies the Conditions (A)-(E) (listed below). Then for each essential point $x \in M_0$ (i.e. $x \in M_0 \cap M_1$) there exists a neighborhood $U(x)$ contained (mod 0) in one ergodic component.

Now we list the conditions that the billiard system has to satisfy:

Condition A: (continuity) For almost all $x \in M$ there exist local stable and unstable manifolds and their tangent subspaces E_x^s and $E_x^u \subset T_xM$ depend continuously on x.

Condition B: (Sinai-Chernov ansatz) Almost every point on the singularity manifold enters regions where the quadratic form increases, i.e. for almost all $x \in S_{-1}(S_1)$

$$\lim_{n \to -\infty} \|Q(DT_x^n s)\| = \infty, \forall s \in C_x^-,$$

$$(\lim_{n \to \infty} \|Q(DT_x^n u)\| = \infty, \forall u \in C_x^+),$$

where DT^n is the differential of the n-th iterate of the billiard map.

Condition C: (double singularities) For each integer $n > 0$, denote by Δ_n the set of $x \in M$ such that there exist positive integers $M \leq n$ and $N \leq n$ satisfying $T^{-N}x \in S_1$ and $T^M x \in S_{-1}$. Then Δ_n is a finite union of compact submanifolds of codimension greater than 1.

Condition D: (thickness of the singularity region) For $\epsilon > 0$ let U_ϵ be the $\epsilon-$ neighborhood of the singularity set $S_1 \cup S_{-1}$, that is measured in a pseudometric, described in the proof of Lemma 4. Then there exists a constant $C > 0$ such that

$$\mu(U_\epsilon) \leq C\epsilon.$$

Condition E: (transversality) At almost every point x of the singularity submanifold S_1 the stable subspace E^s_x is transversal to S_1, i.e. E^s_x is not a subspace of $T_x S_1$. Similarly, E^u_x is not a subspace of $T_x S_{-1}$ for almost every point of S_{-1}.

The Lemmas dealing with these conditions are proved in Sect. 3. Besides them, we will need some properties of essential points M_1.

PROPOSITION 1. – *Essential points have full measure, i.e.* $\mu(M_1) = 1$.

Proof. – For the sake of brevity we give the proof just for the case of a cube with attached spherical caps. It is easy to see that any velocity vector $v = (v^1, ..., v^n), ||v|| = 1$, with pairwisely incomensurable components defines a positive (as well as negative) semitrajectory which gets reflected from a spherical cap. Indeed, by the standard trick with reflecting the trajectories that never hit spherical caps from flat walls one gets a flow on a torus.

Poincare recurrence theorem now implies the following statement.

COROLLARY. – *Trajectories of almost all points of M_1 have infinitely many reflections from spherical caps.*

PROPOSITION 2. – *The set M_1 of essential points contains a subset of full measure $\hat{M}_1, \mu(\hat{M}_1) = 1$, which is arcwise connected.*

Proof. – For the sake of brevity we will discuss only the domains of type (2). It is enough to consider a cube K, which has at least one spherical cap attached to every pair of its parallel faces. (Recall that B is constructed as a rectangular box K with some attached spherical caps). The case of a rectangular prism, whose principal faces tile R^m (type (1) region) can be treated in a similar way based on the approach developed in Sect. 3. Define the base $\hat{C}_i, i = 1, ..., p, p \geq n$, of a spherical cap C_i as the intersection of the sphere containing this cap with the hyperplane (the face of K) to which C_i is attached. The base $\hat{C}_i$ is an m-dimensional disk.

Consider a velocity vector $v = (v^1, ..., v^n), ||v|| = 1$ (and recall that v is always directed inwards B). Suppose that the components of v are rationally independent and a positive semitrajectory $\{T^k x\}, x = (q, v), q \in \partial B$ is defined ($k \geq 0$). It is well-known that the ergodic components of a billiard in a rectangular parallelepiped in R^m are the one-parameter groups of shifts of an m-dimensional torus. Thus such semitrajectory $\{T^k x\}, 0 \leq k < \infty$, is everywhere dense in ∂B and eventually hits a base $\hat{C}$ of a spherical cap C.

Therefore, we must consider only such vectors v, whose coordinates are rationally dependent. Suppose that coordinates of a vector v satisfy more than one relation over the field of rational numbers. Then the union of corresponding tori, which are closures of semitrajectories $\{T^k x\}, x = (q, v), k \geq 0\,(k \leq 0)$, have at least codimension two in M. There is only a countable number of such tori whose union we denote by $W^+\,(W^-)$.

Therefore we should consider only such vectors v whose coordinates satisfy only one rational relation. There is only a countable number of such relations. Each rational relation defines an $m-1$ dimensional unit sphere on an m-dimensional unit sphere of all admissible velocities that satisfy some rational relation r.

Let $F_1, F_2, ..., F_{2n}$ be the sequence of faces of the cube K. Denote by $R_{r,l}$ a union of all vectors v whose components satisfy the relation r and whose basepoint belongs to some fixed face F_l.

Let $D_{r,l}$ be a measurable subset consisting of the points $x = (q, v) \in M$, such that $\pi(x) = q \in F_l$ and the components of the velocity vector v satisfy the rational relation r. It is important to mention that $D_{r,l}$ doesn't necessarily contain all points of M with such properties, but instead only a measurable subset consisting of such points.

We call $D_{r,l}$ a positive rational strip if its positive semitrajectory $T^k D_{r,l}, 0 \leq k < \infty$ never hits a base of any spherical cap (we assume here that some semitrajectories of a rational strip may terminate at singularities of the boundary ∂K). In the same way we define negative rational strips. Denote by $D_{r,l_{max}}$ the maximal rational strip corresponding to a relation r and having a basepoint in the face F_l.

We now invoke the Kronecker's Theorem: If $1, \alpha_1, ..., \alpha_k$ are rationally independent, then for any ϵ and any real numbers $x_1, ..., x_k$ there exists an integer n and a family of integers $p_1, ..., p_k$ such that

$$\mid n\alpha_i - p_i - x_i \mid < \epsilon, \; 1 \leq i \leq k.$$

This theorem ensures that for a type (2) domain there is only a finite number of maximal rational strips (recall that we consider only the rational strips generated by a single rational relation). Indeed, let us denote by $v^{\perp} \in R^n$ the orthogonal complement of all the velocity vectors satisfying the relation r. This is an $n-1$-dimensional subspace $V \subset R^n$, hence its orthogonal complement is one dimensional. For any rational strip, we fix a base point q and consider the subspace V. When we unfold it, and project it back to the cube K we obtain finitely many parallel hyperplanes, whose distance (in the normal direction) will be denoted by s. The maximal length of the projection of some base $\hat{C}$ onto $v^{\perp}$ will be denoted by a. Then, according to the construction of the type (2) domain, for only a finitely many rational strips $s > a$. This was the reason for us, to attach spherical caps to at least one of any pair of parallel faces of ∂K. In the same way, one may consider negative rational strips.

None of the finite number of the maximal rational strips cuts M_1 into the separated domains because any maximal rational strip $D_{r,l_{max}}$ has "holes" corresponding to velocity vectors v, satisfying r and such that a trajectory of $x = (q, v), q \in F_l$ hits some spherical cap. A union of such holes is $F_{r,l} \backslash D_{r,l_{max}}$.

We now remove from M_1 all positive and negative rational strips. According to the previous consideration the resulting set M_1' will be arcwise connected. Finally, the set $\hat{M}_1 = (M_1' \cap M_0) \backslash (W^+ \cup W^-)$ is also arcwise connected because a removal of a countable number of codimension two submanifolds doesn't change the arcwise connectivity.

Remark. – For type (1) domains there is an infinite (countable) number of rational strips. Those strips are accumulated in the vicinity of rational strips that are parallel to faces of ∂B to which the spherical caps are attached. However, by making the bases of spherical caps to be inscribed into the corresponding polyhedra, we can still connect points in a "good" subset of M_1 of full measure through the intersections of boundaries of bases of the spherical caps with another (not containing spherical caps) regular components of ∂B.

Proof of Theorem 1. – Provided that the Fundamental Theorem holds, which will be the content of the next section, ergodicity is obtained by the standard method. Indeed, each essential point of M_0 has a neighborhood which belongs mod 0 to one ergodic component. According to the above Propositions there exists an arcwise connected full measure set $\hat{M}_1 \subset M_0 \cap M_1$ consisting of points to which the Fundamental Theorem applies. We can connect any two points from that set by the path which avoids the double singularities (this is possible because the manifolds of double singularities have at least codimension 2). Finally, by virtue of compactness, we cover this path by finitely many open neighborhoods, each of them belonging to one ergodic component.

3. PROOF OF THE MAIN RESULT

In this section we prove that billiard regions described in the Theorem 1 satisfy Conditions (A)-(E). We begin by showing how to define quadratic forms that establish the existence of local stable and unstable manifolds.

LEMMA 1. – *A billiard region B described in Theorem 1 satisfies Condition A.*

Proof. – The key ingredient for establishing the existence of local stable and unstable manifolds is an invariant cone field in TM, i.e. cones (or more exactly sectors) defined in the tangent space T_xM of almost every phase point x and being mapped into each other by a differential of the billiard map (see **[W2]**). Some facts from the theory of cone-fields (from **[LW]**), as well as the characterization of tangent vectors using infinitesimal fronts and basic results about their propagation inside a spherical cap (from **[BR2]**) are summarized in the Appendix.

We begin by introducing artificial transparent walls (see Fig. 3) whose purpose is to provide enough free path for a control surface (a wavefront) γ to defocus. These transparent walls V and V' play the role of the border of the zone of focusing from Definition 1. They are parallel to the sides to which the spherical caps are attached and pass through the centers of the corresponding spheres.

It was shown in **[BR2]** (Lemma 1 and Lemma 2) that any control surface that passes the transparent wall towards a cap with positive definite curvature operator (i.e. the corresponding beam of rays is diverging) has positive definite curvature operator again when it leaves through the same transparent wall. That means that after a series of reflections from a spherical cap a diverging beam becomes diverging again. This property is called absolute focusing (see **[B2,B3]**). For spherical caps under consideration it immediately follows from (2.2) (see **[C]**, **[W1]**). Of course, during the passage through the spherical cap and immediately after it, the curvature operator may have a non-trivial eigenspace with negative eigenvalues (principal curvatures). The whole point of moving the transparent walls V and V' back so they intersect the centres of the attached spheres is to give the control surface ample time to defocus.

The above described defocusing property of control surfaces between points V_{in} and V_{out} serves as a basis for definition of cones (sectors) at the tangent spaces of these points. If the tangent spaces are parametrized by r', ϕ', then they are defined as the standard sectors $C_x^+ = C(V_1, V_2)$, where $V_1 = (r', 0)$ and $V_2 = (0, \phi')$. The boundary of this sector thus

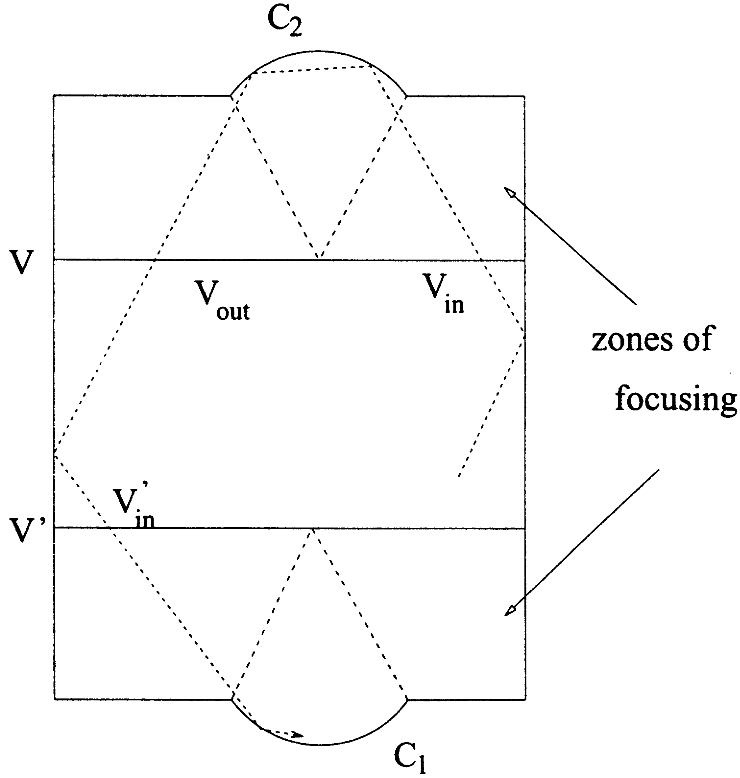

Fig. 3. – Zones of focusing.

consists of a Lagrangian subspace V_1 that corresponds to a beam of parallel trajectories and V_2 corresponding to a beam of trajectories emanating from one (configuration) point. The defocusing property then means that the image of V_1 by a billiard map is mapped strictly inside a corresponding cone at V_{out}, while the image of V_2 may not be mapped strictly inside (this is an analogy from planar billiards where a beam emanating from the center comes back and focuses at the center again). This is why we give two focusing zones a positive distance. This extra free path then causes the beam that may have focused at V_{out} to be diverging at V'_{in} and that "pushes" V_2 inside the corresponding cone at that point. Thus between V_{in} and V'_{in} the billiard map maps one cone strictly inside the other.

With the above choice of V_1 and V_2 one can define the quadratic form Q_x by (this is a special case of a general definition mentioned in the Appendix):

$$Q(x') = r'.\phi'.$$

It is clear that if $Q(x') > 0$ (or $Q(x') \geq 0$), then we can find an infinitesimal surface with positive definite (semidefinite) curvature operator, such that a vector x' lies in the Lagrangian subspace that corresponds to it (the reader can find more details in **[BR2]**). This facilitates the study of the dynamics of T.

Having defined the quadratic form at the configuration point V_{in} (see Fig. 3), we can now repeat the same construction at the next point V'_{in} and so on. Since the quadratic form Q maps the sector $C(V_1, V_2)$ at V_{in} into the analogous sector at V_{out}, the linearized billiard mapping is monotone between V_{in} and V_{out} and because of a positive distance between the two zones of focusing it is strictly monotone between V_{in} and V'_{in}. One can now define a linear mapping $\hat{T}$ between tangent spaces at different V_{in}s which is strictly monotone ($Q(\hat{T}u) > Q(u)$) or observe that the linearization of the original billiard map T is eventually strictly monotone. In either case Theorem 5.1 (from **[W2]**) now assures the existence of the expanding and contracting subspaces E^u_x and E^s_x and Theorem 2.1 from **[M]** their continuous dependence on x. Note that for regions of the type (2) the positivity of the distance between two zones of focusing is guaranteed by the placement of the spherical caps (Fig. 2).

LEMMA 2. – *A billiard region B described in Theorem 1 satisfies Condition B.*

Proof. – In order to prove the Sinai-Chernov Ansatz we need to show that almost every point on the singularity manifold enters some spherical cap. Indeed, every passage through the spherical cap causes the control surface γ to focus in all directions at or before the point V_{out}. After that the positive distance between the two zones of focusing forces the boundary of the sector at V_{in} to be mapped strictly inside of the sector at the next point V_{in}. This causes the quadratic form to increase and this increase is uniform, since it depends only on the minimal distance between the two zones of focusing.

To show that this happens, i.e. that almost every orbit starting on the singularity manifold enters a spherical cap, requires a detailed knowledge of the orbit in the polygonal part of our region. For this reason we have restricted the regions in consideration to rectangular or tiling boxes. For these, we will first show that any configurational point q in the boundary has the property, that for almost every unit velocity vector v the trajectory determined by $x = (q, v)$ enters some spherical cap. From this we then deduce the desired property of the singularity manifold. While for rectangular boxes this is trivial and essentially follows from the dynamics on the m-torus, for type (1) regions we will show this in more detail.

Suppose that the flat walls Q to which the caps are attached are parallel to the hyperplane $x_n = 0$. We then project the billiard region onto this plane and likewise we project the billiard trajectory. After we unfold the flat side Q to its symmetric images RQ, we can consider a projected billiard trajectory in the form of a straight line, going through different images, rather then a piecewise linear trajectory, bouncing within the same side Q (this is a well-known "trick" from polygonal billiards).

Now we pick a point in each "cell" (for instance in Fig. 4 we have picked the centers of each hexagon) and label it as $l_I = l_{i_1 i_2 \ldots i_m}$. The points l_I constitute a lattice that is determined by m vectors $e_i = l_{000\ldots1\ldots000} - l_{000\ldots000}$. Recall, that each cell is actually a projection of the billiard region B along the x_n axis and since the "side" flat walls are perpendicular to Q, the whole billiard region tiles R^n. Unless the n-th component of the velocity is 0, the billiard orbit hits infinitely often the principal face Q, since the reflections off the side walls do not change this component. As a matter of fact, this component can change its magnitude only by reflecting from the spherical cap in which case we would be done.

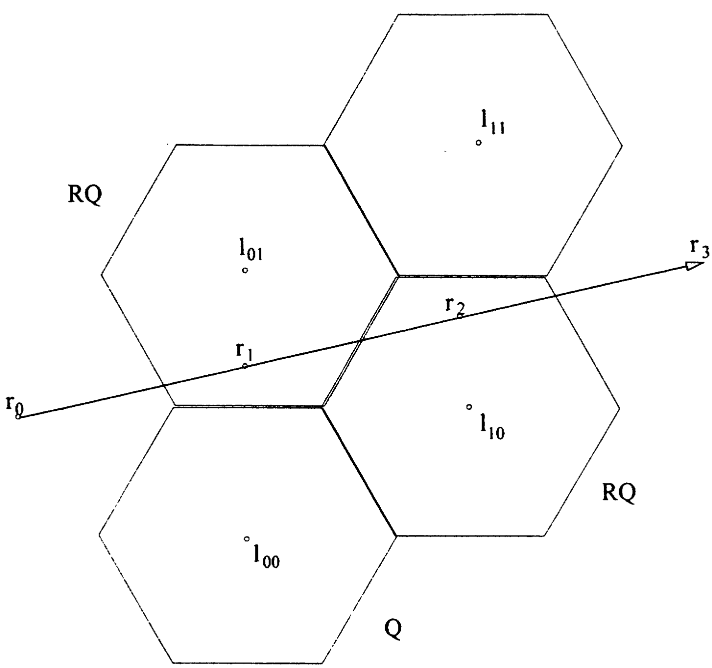

Fig. 4. – Unfolding the projected orbit.

By projecting along x_n we have effectively erased the information about the dynamics in the vertical direction, so we have to somehow keep the track of reflections from the principal faces parallel to the plane $x_n = 0$. The projections of the points where the orbit reflects from Q will be denoted by r_i. If the nth component of velocity is large (i.e. the billiard trajectory is steep), the points r_i will be close to each other, because it takes less time to traverse the billiard region from the bottom to the top. In the opposite case, of course, the spacing between r_i's will be wide.

We claim that for almost all velocity vectors the points r_i are dense in a parallelogram $E = (e_1, e_2, ..., e_m)$. More precisely, let us denote by $w = r_1 - r_0$. Then we claim that for any $\epsilon > 0$, every $r_0 \in R^{n-1}$ and almost every velocity projection w there exists a natural number n and a lattice point l_I such that

$$\|r_0 + nw - l_I\| \leq \epsilon.$$

To see this, we first make a linear transformation that transforms vectors $(e_1, e_2, ..., e_n)$ into an orthonormal basis. Since this means just introducing new coordinates in the parallelogram E (that is now a unit cube), the values of r_0 and w that yielded dense trajectories will yield dense trajectories again. Moreover, we can shift the point r_0 to the origin $l_{000..000}$. It is now clear that the velocities w for which the points r_i are dense in the unit cube are those, whose coordinates are irrational numbers. Or, to put it differently, the velocities that do not yield dense orbits are those that have rationally dependent coordinates.

The set of velocities having this property is of (m-dimensional) measure 0 and thus for almost all w the points r_i fill in the parallelogram E densely. That means that the same points will also fill in the flat side Q densely. Since the projection of the spherical cap onto Q is an open set, one of r_i's will fall into this projection. This point r_{i_0} is the intersection of the billiard orbit with the flat side Q and thus for every point on the boundary almost every velocity yields a trajectory that visits a spherical cap.

A beam of trajectories shot from a particular singular boundary point (i.e. from the "edge" of our region) represents only an m-dimensional submanifold of the full $2m - 1$ dimensional singularity manifold. However, the full manifold can be foliated by these m dimensional submanifolds. Indeed, for any point from the $m - 1$ dimensional intersection of any two regular boundary components we can shoot a similar beam of trajectories. Since each family of such trajectories gives rise to an m dimensional submanifold of the singularity manifold, we can just observe that the statement of the lemma follows from integration over this intersection.

Each leaf of the foliation has, according to the discussion above, the full measure of points visiting some cap and the same is true about the singularity manifold.

LEMMA 3. – *A billiard region B described in Theorem 1 satisfies Condition C.*

Proof. – The singularity manifolds have codimension 1 and are caused only by orbits ending at the points of intersection of regular boundary components (so called "multiple reflections"). Consider a point $x \in M$ such that there exist such integers $N > 0$ and $M > 0$ that $y^+ = T^M x \in S_{-1}$ and $y^- = T^{-N} x \in S_1$.

Denote the images of the singularity manifolds at x by $S_M(x)$ and $S_{-N}(x)$ respectively. If these two manifolds were not transversal at x the tangent spaces to these two submanifolds at x would coincide and have dimension $2m-1$. However, the control surfaces corresponding to beams of trajectories emanating from y^+ and y^- have positive and negative definite curvature operators respectively. Hence the $2m-1$ dimensional tangent space to both singularity manifolds would have to contain 2 disjoint n dimensional subspaces, which is impossible.

This argument doesn't hold only in the case when the trajectory between y^+ and y^- never leaves the given zone of focusing. That means that both y^+ and y^- have their base point in the "edge" of a spherical cap. In that case we deduce the claim from the fact that the planar direction can now be defined along the whole orbit between y^+ and y^- and is transversal to the directions determined by the $m-1$ dimensional edges.

LEMMA 4. – *A billiard region B described in Theorem 1 satisfies Condition D.*

Proof. – For simplicity we will consider a different Poincare section of the billiard flow, defined by passing through different zones of focusing, rather than by reflecting at the boundary. Let us denote the points of entrances to zones of focusing by $V_{in}^1, V_{in}^2, \ldots$ and the induced mapping by $T_{in} V_{in}^i = V_{in}^{i+1}$. For any corresponding phase point $x = (q, v)$, $q = V_{in}^i$ we coordinatize the tangent space by $(r', \phi') \in R^m \times R^m$, where r' corresponds to the configurational displacement along the m-dimensional subspace perpendicular to the velocity vector v. Due to the strict invariance of cones $C_x^{\pm}$ between different V_{in}s the local unstable manifold at x is defined by a positive definite curvature operator K. This operator defines a curved hypersurface $\mathcal{K}$ in the vicinity of x. We define $w(x)$ as the length of the shortest geodesic along $\mathcal{K}$ connecting $\pi(x)$ with $\pi(y)$, where

$y = (\pi(y), u)$ and u is a velocity vector such that a trajectory determined by y ends at a singularity (i.e. at the intersection of two regular boundary components). The length of a geodesic on $\mathcal{K}$ is defined in a standard way by a pseudonorm $\|x'\| = Q_x(x')$ for $x' \in T_xM$, which acts as a norm in the unstable subspace E_x^u. Owing to the positive distance between different zones of focusing (which increases the form Q on the vectors from the unstable cone C_x^+) the billiard map T_{in} expands vectors from the unstable space E_x^u when measured in this pseudonorm.

The distance function $w(x)$ is equivalent to the function $z(x)$ defined in **[SCh]** as long as the curvature operator K has bounded curvatures (the z function defines the distance from singularities as the maximal diameter of a cylindrical neighborhood in B of the trajectory connecting x and Tx and avoiding singularities). From strict cone invariance, the unstable space E_x^u lies in the interior of the unstable cone C_x^+. This subspace then corresponds to a control surface with positive definite curvature operator K and its eigenvalues satisfy $0 < \kappa_i < const$ for all $i = 1, 2, ..., m$ and the constant depends only on the distance between different zones of focusing. Hence there are two positive numbers A, B (depending on the κ_is) such that

$$Az(x) \leq w(x) \leq Bz(x).$$

Thus we can measure the ϵ neighborhood of the singularity set using the function z which is technically easier. That means that instead of calculating the measure of the set $U_\epsilon = \{x \in M, w(x) \leq \epsilon\}$, we can estimate the measure of $V_\epsilon = \{x \in M, z(x) \leq \epsilon\}$.

The only possibility for a phase point $x = (q, v)$ to be in V_ϵ is for q to lie in the $\epsilon/cos(\phi)$ neighborhood of S_{-1}. The m-dimensional volume of this set is less than $const.\epsilon/cos(\phi)$, where $const$ depends only on the parameters of the region B. Then the μ-measure of the set of such points will be less than $const.\epsilon$.

LEMMA 5. – *A billiard region B described in Theorem 1 satisfies Condition E.*

Proof. – Condition E (in **[Ch]** property 5') can be reformulated in terms of the cone field $C(x)$ as follows (see the "proper alignment of singularity sets" in **[LW]**, sect. 8): all the tangent subspaces to singularity submanifolds have their skew-orthogonal complements in C_x^+ (the singularity manifolds are of codimension 1 and their complements with respect to the standard symplectic form have thus dimension 1).

In order to check this condition, for each point $x \in S_1$ we have to find a skew-orthogonal complement $\hat{c} \in TM$ of the tangent space to the singularity manifold ($\hat{c} \perp T_x S_1$) and then move it by a differential of the flow to the border of a zone of focusing. Since our billiard system doesn't have tangential singularities, let us consider a point $x = (q, v)$ where q lies in the intersection of 2 regular boundary components ∂B_i and ∂B_j. The singularity manifold around x is then constituted by all $y = (p, v)$, where $p \in \partial B_i \cap \partial B_j$ and v is an arbitrary inward pointing velocity vector. The orthogonal complement of the singularity manifold at x then has the form $c = (r', 0)$, where $0 = (0, ..., 0) \in R^m$ and $r' \in R^m$ is an orthogonal complement of the intersection $\partial B_i \cap \partial B_j$ in either B_i or B_j, or to be more exact, it is a projection of this vector onto a subspace $(r_1, ..., r_m, 0, ..., 0) \in T_x M$. Hence the skew-orthogonal complement of the singularity manifold at x has the form $\hat{c} = (0, r') \in T_x M$ (it follows from the symplectic geometry that if the codimension 1 subspace has a euclidean complement (a, b), then its skew-orthogonal complement is $(b, -a)$).

In order to find, whether the skew-orthogonal complements to singularity sets are in C_x^+ we have to consider the evolution of a control surface that originated at q, i.e. the beam of trajectories that started at q with different velocities v. Strictly speaking we would need only to consider the evolution of the part of the control surface corresponding to the vector $\hat{c}$, but it is easier to look at the evolution of the whole surface. And at this point we need the restriction of the angular size of the spherical cap to $60°$. Since this computation needs to be done only in the transversal directions, we will make use of the formula (3.4) in the Appendix with $\rho = 1$.

Let us consider a part of the billiard orbit in a spherical cap with the angular size smaller than $60°$ and having $N + 1$ reflections in it, including the one at the singularity point, where this orbit begins. Since we are going to consider a beam of trajectories emanating from this point, we have to find the curvature of the corresponding wavefront at the point A_1 (Fig. 5), which will make the beam focused at the singularity point. This point can be thought of as the first reflection point and hence $\kappa' = -1/cos(\phi)$. If we denote the angle subtended at the center by the first and the last point of reflection by ω', then the angle $\omega = \omega' + \omega'/N$. Using the relation $\phi = 90° - \frac{\omega'}{2N}$ and the above expressions for ω and κ' we can now evaluate the exit curvature (i.e. at the point A_3 in Fig. 5). After a few elementary trigonometric manipulations we obtain

$$\kappa'' = cotan(\omega' + \frac{\omega'}{2N}) sec(\frac{\omega'}{2N}).$$

Since we consider orbits having at least two reflections (including the singularity point) $N > 0$ and from $\omega' < 60°$ we now easily deduce that

$\kappa'' > 0$ at the exit point and hence also at the point V_{out}, which lies farther on the orbit. This means that any beam of trajectories emanating from the singularity point x (and that includes the vector $\hat{c} \in T_x M$) will have positive definite curvature operator at V_{out}. Hence Lemma 5 follows.

Let us remark that the fact that our condition E follows from the "proper alignement of the singularity sets" mentioned at the beginning of the previous proof is implied by the following observation: First recall from lemma 1 that for a vector $u = (r', \phi') \in T_x M$ the quadratic form is defined by a scalar product of m-dimensional vectors $Q(u) = r'.\phi'$. For such vector let us also denote $\hat{u} = (-\phi', r')$ and recall again that if u is the orthogonal complement to a codimension one subspace, then $\hat{u}$ is the skew-orthogonal complement to the same subspace. From the definition of Q it is also clear that if $u \in C_x^+$ then $\hat{u} \in C_x^-$ and vice versa. The stable subspace $E_x^s \subset C_x^-$, while $E_x^u \subset C_x^+$, both being m-dimensional subspaces.

Suppose for contradiction that $E_x^s \subset TS_1$. Let u be the orthogonal complement of TS_1 such that $\hat{u}$ is in C_x^+. Then $u \in C_x^-$ is orthogonal also to $E_x^s \subset TS_1$ which is a contradiction, since the cone C_x^- cannot contain an $m + 1$ dimensional subspace spanned by E_x^s and u.

This concludes the proof of the Fundamental Theorem.

COROLLARY. – *Billiard system (M, T, μ) considered in Theorem 1 and the corresponding flow $(\mathcal{M}, S^t, \nu)$ are Bernoulli systems.*

This statement follows immediately from standard techniques and the results of **[ChH, GO, KS, S1, S2]**.

CONCLUDING REMARKS. – The mechanism of defocusing acts in higher than two dimensions in much more subtle manner than in 2D. The famous formula (2.2) from the geometric optics **[C]** immediately gives the idea that the corresponding spherical caps must have an internal angle less than 90°. However, to prove ergodicity one needs to obey some other (besides defocusing of incoming rays) conditions, which forced us to restrict an allowed spherical caps even more **[B2]**. At the moment, we are not able to prove ergodicity in the condition of **[B2]**. Actually, our spherical caps with internal angles less than 60° are exactly inscribed into the pieces of spheres considered in **[B2]**. We believe, however, that ergodicity can be proved for spherical caps with internal angles less than 90°. Furthermore, Wojtkowski's example **[W1]** with semispheres gives just linear stability of some periodic trajectories. However, on contrary to 2D, it is not true in higher dimensions that generically linearly stable periodic trajectory is, in fact, stable. The general belief is that it is not the case. Therefore, while Wojtkowski's example shows a loss of hyperbolicity, it doesn't demonstrate that spherical caps with internal angles 90° or even with bigger ones cannot belong to the boundary of ergodic billiards.

Appendix

In this appendix we review some basic facts about the tangent vectors of a billiard phase space, about their evolution and about cones in the tangent spaces.

Let us fix a point $x = (q, v)$ and consider an m-dimensional hypersurface U perpendicular to the vector v. A tangent vector $x' = (r', \phi') \in T_x M$ can naturally be related to a family of orbits generated by a displacement $r' = dr$ along a subspace perpendicular to v and the angle increment $\phi' = d\phi$ (see Fig. 5):

$$o(\sigma) = (r(\sigma), \phi(\sigma)) = (r + \sigma r', \phi + \sigma \phi').$$

Note that this family represents a curve in the phase space, satisfying $o(0) = x = (q, v)$ and $o'(0) = x'$. Hence this family is a natural representative of a class of equivalent curves from the usual definition of a tangent vector. In the configuration space this family (representing a tangent vector x') can be visualized as a perpendicular envelope of the family of orbits $o(\sigma)$ at q (a representant of this family is a dashed curve in Fig. 5).

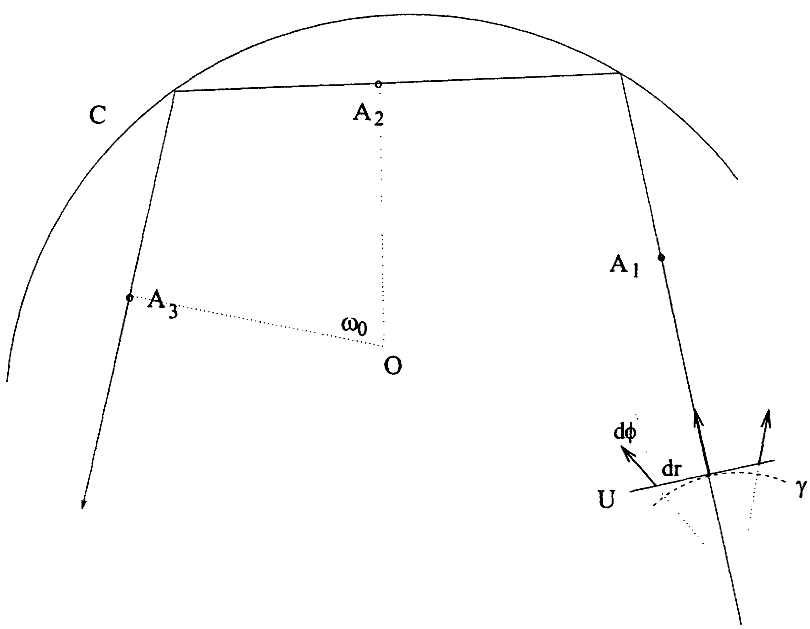

Fig. 5. – Evolution of tangent vectors.

A tangent vector $x' = (r', \phi')$ when acted on by a differential of the billiard flow evolves to a vector $x'' = (r'', \phi'')$. Even though the most natural way how to describe the dynamics in the vicinity of a given orbit is to consider evolution of the infinitesimal control surface (also caled a wavefront) that is tangent to U we will at first consider only evolution of 1-dimensional curves corresponding to planar and transversal directions on U. That means that $x' = (r', \kappa r')$, where r' is one of the above directions.

In order to obtain clear understanding of the dynamics, we consider the evolution only from the point A_i to the point A_{i+1}, where A_i is the middle point of a segment connecting the points of the ith and $(i-1)$th reflections. Between any pair of such points A_i and A_{i+n} the evolution of the tangent vectors can be expressed as

$$\begin{pmatrix} r'' \\ \phi'' \end{pmatrix} = G \begin{pmatrix} r' \\ \phi' \end{pmatrix},$$

where

$$G_p = \begin{pmatrix} 1 & 0 \\ \frac{2n}{\rho cos(\phi)} & 1 \end{pmatrix} \tag{3.1}$$

in the planar direction and

$$G_t = \begin{pmatrix} cos(\omega) & \rho sin(\omega).sin(\phi) \\ \frac{-sin(\omega)}{\rho sin(\phi)} & cos(\omega) \end{pmatrix} \tag{3.2}$$

in the transversal directions. Here $\omega = n\omega_0$ is an angle subtended by the points A_i and A_{i+n} at the center O of the spherical cap. In order to express the differences between the two cases, let us rescale the configurational coordinates differently. In the planar direction we set $r_p = r_p/\rho cos(\phi)$ while in the transversal directions $r_t = r_t/\rho sin(\phi)$. The evolution matrices, after the rescaling, assume the following form

$$G_p = \begin{pmatrix} 1 & 0 \\ 2n & 1 \end{pmatrix}, \quad G_t = \begin{pmatrix} cos(\omega) & sin(\omega) \\ -sin(\omega) & cos(\omega) \end{pmatrix}.$$

Now we see the fundamental difference between the behavior in the planar direction (which coincides with the behavior of planar billiards) and in the transversal directions (which one has to deal with in many dimensional cases). In the planar direction, the action of the evolution matrix G is a shear, while in the transversal direction it is a rotation. As a result of this, in the transversal direction a piece of a sphere does not have an absolutely focusing property (which is equivalent to focusing between any pair of reflections, see e.g. **[B1]**, **[B2]**, **[D]**).

When, instead of looking at the tangent vectors $x' = (r', \phi')$ we look at the curvature of the associated perpendicular envelope (control curve) γ, we obtain the following formulas:

$$\kappa'' = \frac{2n}{\rho cos(\phi)} + \kappa' \tag{3.3}$$

in the planar direction and

$$\kappa'' = \frac{\frac{-sin(\omega)}{\rho sin(\phi)} + cos(\omega).\kappa'}{cos(\omega) + \rho sin(\phi) sin(\omega).\kappa'} \tag{3.4}$$

in the transversal directions. Here κ' is the curvature of the control surface in the middle of the entrance cord (the point A_1 in Fig. 4) and κ'' is the curvature in the middle of the exit cord (the point A_3). Thus in the planar direction, the dynamics adds to a curvature a constant amount (as we go from A_i to A_{i+1}), while in the transversal direction the curvature evolution is in a certain sense periodic.

This difference causes very strong focusing in the planar direction, while in the orthogonal directions the focusing is much weaker. As a matter of fact, a surface that was focusing in the orthogonal direction exactly while passing through the transparent wall (on its way towards the cap) does so on its way back too. Thus the transparent walls cannot be moved any closer towards the spherical caps.

Of course, in general an incoming surface may have principle curvature directions not identical to the planar and transversal subspaces defined in the previous section. In that case the evolution of curvatures between V_{in} and V_{out} is more complicated and the evolution in the planar and transversal direction cannot be described separately since the principal curvature directions now rotate after each reflection and the quantity subtracted at the reflection in each direction is a combination of (2.1) and (2.2). The fact that one can nevertheless restrict himself to considering the evolution of planar and transversal directions separately is justified either by an abstract approach of invariant cones (**[LW]**) or more geometrically by enclosing an incoming control surface between two spherical caps (this approach is exploited in **[BR1,2]**.

We now describe the construction of the invariant cones and the quadratic form Q. To the configuration point V_{in} (see Fig. 3) there corresponds a phase point $x = (V_{in}, v)$. At this phase point we take a quotient of $T_x(Q \times S^m)$ by the velocity vector v and denote the resulting linear space by $T_x M$. We chose this notation, because this quotient space can be parametrized, just as the tangent space to the phase space, by m positional coordinates r_i' and by m angles ϕ_i'.

Let us denote a tangent vector at x by $x' = (r', \phi')$, where $r' = (r_1', ..., r_m')$ and $\phi' = (\phi_1', ..., \phi_m')$. Any control surface γ can be thought of as an $m-$dimensional subspace of the $2m-$dimensional space $T_x M$, spanned by m independent vectors $x_i' = (r_i', \phi_i')$. Then the second fundamental form K of the control surface maps the configurational vector onto the angular vector

$$\phi' = K \cdot r'.$$

Since the curvature of a surface in the direction of a unit vector u is given by (Ku, u), taking $u = r'/|r'|$ allows us to compute the curvature of the family $o(\sigma)$ as

$$\kappa = \frac{\phi' \cdot r'}{|r'|^2}, \tag{3.5}$$

and we can always rescale the vector x' so that r' is a unit vector.

However, not every m-D subspace in the $2m$-D tangent space corresponds to an infinitesimal perpendicular surface. In order that $span(x_1', ..., x_m')$ corresponds to an infinitesimal surface, it is necessary and sufficient that for all $i, j = 1, ..., m$

$$r_i' \cdot \phi_j' = r_j' \cdot \phi_i'. \tag{3.6}$$

This is just a condition for the symmetricity of the curvature matrix K. If we think of R^{2m} as a symplectic space with a standard symplectic form Ω, then the equation (3.6) becomes just $\Omega(x_i', x_j') = 0$. Hence the infinitesimal surfaces, perpendicular to the orbit can be identified with the Langrangian subspaces of R^{2m}, i.e. with planes that are skew-orthogonal to themselves ($\Omega(x', y') = 0$ for any two vectors from that plane).

In order to define the quadratic form at a point $x = (V_{in}, v)$ we will need a notion of sectors, that are multidimensional versions of cones (for more detailed treatment see **[LW]**). Let $V_1, V_2 \subset T_x M$ be two transversal Lagrangian subspaces, i.e. every vector in $w \in T_x M$ can be uniquely written as $w = v_1 + v_2$, where $v_i \in V_i$. This decomposition allows one to define a quadratic form $Q(w) = \Omega(v_1, v_2)$ in $T_x M$. Recall that

$$\Omega(x', y') = r' \cdot \psi' - s' \cdot \phi',$$

where $r', s', \phi', \psi' \in R^m$ and $x' = (r', \phi'), y' = (s', \psi') \in R^{2m} \cong T_x M$ ("$\cong$" denotes an isomorphism between the linear spaces). Given V_1 and V_2 we can define a sector (cone)

$$C = C(V_1, V_2) = (w \in T_x M, Q(w) \geq 0). \tag{3.7}$$

The interior of the sector is then defined as the set of vectors on which the quadratic form Q is strictly positive.

Since the definition (3.7) is difficult to work with, we will now evaluate the quadratic form Q explicitly for a particular choice of the Lagrangian subspaces V_1 and V_2. Namely,

$$V_1 = \{(r', 0), r' \in R^m\}, \tag{3.8}$$

$$V_2 = \{(0, r'), r' \in R^m\}. \tag{3.9}$$

It is clear that these two subspaces are Lagrangian and that they are transversal, i.e. $R^{2m} = V_1 \oplus V_2$, where "$\oplus$" stands for a direct sum of subspaces. These subspaces correspond to infinitesimal surfaces, one of which is flat and one is focusing (i.e. it has an infinite curvature) and the corresponding sector is called the standard sector. Any vector $x' = (r', \phi')$ can be decomposed into V_1 and V_2 as $x' = (r', 0) + (0, \phi')$ and that yields $Q(x') = r'.\phi'$.

Of course, different choice of Lagrangian spaces V_1 and V_2 generates different Q. For instance for $m = 2$ we can consider Lagrangian subspaces corresponding to spherical surfaces with radii $\delta < \epsilon$

$$V_1 = span[(1, 0, \delta, 0), (0, 1, 0, \delta)],$$

$$V_2 = span[(1, 0, \epsilon, 0), (0, 1, 0, \epsilon)].$$

Then

$$Q(x') = \frac{1}{\epsilon - \delta}((\epsilon + \delta) r' \cdot \phi' - \epsilon\delta \|r'\|^2 - \|\phi'\|^2) = \frac{1}{\epsilon - \delta}(\epsilon r' - \phi') \cdot (\phi' - \delta r').$$

There are many Lagrangian subspaces in the standard sector. They correspond to control surfaces with positive definite curvature operator K. The Lagrangian subspace is then described by $V = (r', K.r')$ with $r' \in R^m$. On the other hand, each vector from the standard sector (corresponding to an infinitesimal curve) can be embedded into a Lagrangian subspace contained in it. This is important because the evolution of control surfaces is easier to study than the evolution of infinitesimal vectors. Thus invariance of sectors can be fully described by means of curvature operators of the infinitesimal surfaces perpendicular to the orbit. Indeed, from (3.5) we see that for the standard sector $Q(x') = \kappa\|r'\|^2$, where κ is a curvature of a perpendicular envelope of a beam of trajectories defined by x'.

Finally, let us mention that if this invariance causes standard sectors at x to be mapped strictly inside the sectors at future iterates and this process results at one Lagrangian subspace, then this subspace correspond to an infinitesimal local unstable manifold whose curvature operator K can be formally obtained by an infinite continued fraction discussed in **[SCh]**.

ACKNOWLEDGMENTS

We thank N. Chernov and N. Simanyi for helpful suggestions and a discussion. L. Bunimovich was partially supported by NSF grant DMS-963063.

REFERENCES

[A] D. V. Anosov, Geodesic Flows on Riemannian Manifolds with Negative Curvature, *Proc. Steklov Inst.*, **90**, 1967.

[AS] D. V. Anosov, Ya G. Sinai, Certain Smooth Ergodic Systems, *Russ. Math. Surveys* **22**, 1967, pp. 103-167.

[B1] L. A. Bunimovich, On the Ergodic Properties of Nowhere Dispersing Billards, *Comm. Math. Phys.,* Vol. **65**, 1979, pp. 295-312.

[B2] L. A. Bunimovich, Many-Dimensional Nowhere Dispersing Billiards with chaotic behavior, *Physica D*, **33**, 1988, pp.58-64.

[B3] L. A. Bunimovich, On Absolutely Focusing Mirrors. In Ergodic Theory and Related Topics III. (ed. by U.Krengel), *Lect. Notes in Math.* 1514, New York: Springer-Verlag, 1992, pp. 62-82.

[BG] K. Burns, M. Gerber, Real Analytic Bernoulli Geodesic Flows on S^2. *Ergod. Th. Dyn. Sys.,* **9**, 1989, pp. 27-45.

[BR1] L. A. Bunimovich, J. Rehacek, Nowhere Dispersing 3D Billiards with Non-vanishing Lyapunov Exponents, *Comm. Math. Phys.*, **189**, 1997, pp. 729-758.

[BR2] L. A. Bunimovich, J. Rehacek, How Many Dimensional Stadia Look Like, Preprint, 1997.

[BS] L. A. Bunimovich, Ya G. Sinai, On the Fundamental Theorem of the Theory of Dispersing Billiards, *Math. Sb.*, **90**, 1973, pp. 415-431.

[C] H. Coddington, Treatise on Reflection and Retraction of Light, Simpkin & Marshall, London 1829.

[Ch] N. I. Chernov, Local Ergodicity of Hyperbolic Systems with Singularities, *Funct. Anal. Appl.,* **27**, 1993, pp. 51-54.

[ChH] N. I. Chernov, C. Haskell, Nonuniformly Hyperbolic K-Systems are Bernoulli, *Ergodic Theory and Dynamical Systems*, **16**, 1996, pp.19-44.

[ChT] N. I. Chernov, S. Troubetzkoy, Ergodicity of Billiards in Polygons with Pockets, preprint 1998.

[D1] V. J. Donnay, Geodesic flow on the two-sphere. I: Positive measure entropy. Ergod. Th. Dyn. Sys., Vol. **8**, 1988, pp. 531-553.

[D2] V. J. Donnay, Geodesic flow on the two-sphere. II: Ergodicity. In Dynamical Systems (ed. by J. C. Alexander), Lect. Notes 1342, New York: Springer-Verlag 1988, pp. 112-153.

[GO] G. Gallavotti, D. S. Ornstein, Billiards and Bernoulli Schemes, *Comm. Math. Phys.*, **38**, 1974, pp. 83-101.

[H] E. Hopf, Statistik der Geodätischen Linien in Mannigfaltigkeiten Negativer Krummung, *Ber. Verh. Sächs. Akad. Wiss.*, **91**, 1939, pp. 261-304.

[KS] A. Katok, J. M. Strelcyn, Smooth Maps with Singularities. Invariant Manifolds, Entropy and Billiards, *Lect. Notes in Math.* 1222, Berlin, Springer 1986.

[KSS] A. Krámli, N. Simányi and D. Szász, A Transversal Fundamental Theorem for Semi-Dispersing Billiards, *Comm. Math. Phys.*, **129**, 1990, pp. 535-560.

[L] V. F. Lazutkin, On the Existence of Caustics for the Billiard Ball Problem in a Convex Domain, *Math. USSR Izv.*, Vol. **37**, 1973, pp. 183-216.

[LW] C. Liverani, M. Wojtkowski, Ergodicity in Hamiltonian Systems, *Dynamics Reported*, **4**, 1993, pp. 130-302.

[M] R. Markarian, New Ergodic Billiards: Exact Results, *Nonlinearity*, Vol. **6**, 1993, pp. 819-841.

[S1] Ya G. Sinai, Dynamical Systems with Elastic Reflections, *Russ. Math. Surveys*, **25**, 1970, pp. 137-189.

[S2] Ya G. Sinai, Dynamical Systems with Countably Multiple Lebesgue Spectrum II., *Izv. Akad. Nauk. USSR. Ser. Math.*, **30**, 1966, pp. 15-68.

[SCh] Ya G. Sinai, N. I. Chernov, Ergodic Properties of Certain Systems of Two-dimensional Discs and Three-dimensional Balls, *Russ. Math. Surveys*, **42**, 1987, pp. 181-207.

[Sm] S. Smale, Differentiable Dynamical Systems, *Bull. of AMS.* **73**, 1967, pp. 747-817.

[W1] M. P. Wojtkowski, Linearly Stable Orbits in 3 Dimensional Billiards. Comm. Math. Phys., **129**, 1990, pp. 319-327.

[W2] M. P. Wojtkowski, Invariant Families of Cones and Lyapunov Exponents. *Ergod. Th. & Dyn. Sys.* **5**, 1985, pp. 145-161.

(Manuscript received on October 28th, 1997;
revised version received on February 17th, 1998.)