On Local Ergodicity in Hyperbolic Systems with Singularities

N.I.Chernov

September 14, 2006

Abstract

We consider hyperbolic dynamical systems with singularities such as billiards and similar Hamiltonian systems. For this class of systems we formulate the theorem on local ergodicity in a rather general form. Besides, the existing version of this theorem is strengthened here by discarding one of its assumptions.

The ergodic properties of billiard dynamical systems and similar Hamiltonian systems with hyperbolic behavior are being intensively studied, see [1-10]. One of key results in this theory is so called theorem on local ergodicity. It was formulated in a rather general form in [5,6] for the systems with two degrees of freedom. Here we strengthen this theorem by discarding of one of its assumptions. It should be also noted that we work with systems of arbitrary dimension.

Let M be a smooth d-dimensional manifold with boundary ∂M and $T : M \to M$ be a smooth transformation with singularities at the set $S_- = (T^{-1}\partial M) \cup \partial M$. Denote $S_n = T^n(\partial M)$, and for any $m \le n$ set $S_{m,n} = S_{n,m} = S_m \cup \ldots \cup S_n$. For any finite n the set $S_{0,n}$ (note that it is the set of singular points of T^{-n}) is supposed to be a finite union of smooth compact submanifolds with boundary and of dimension $\le d-1$. It is convenient to assume that all the iterates of T are defined on the subset $M_0 = M \setminus S_{-\infty,\infty}$.

Let ν be an absolutely continuous probability measure on M with a bounded density and invariant under T.

We impose certain quite general conditions on our system, which are formulated in [11]. They are related to the singularity set S_- and usually hold in applications [1-9].

Due to the Oseledec' theorem [12] there exists a limit

$$\lim_{n\to\pm\infty} \frac{1}{n} \ln \|DT^n v\| = \chi(v) \tag{1}$$

for almost every point $x \in M$ and any vector $v \in \mathcal{T}_\S\mathcal{M}$. This limit is called the Lyapunov exponent (for x and v). Thus, a DT-invariant decomposition $\mathcal{T}_\S\mathcal{M} = \mathcal{E}_\S^\sqcap \oplus \mathcal{E}_\S^\int \oplus \mathcal{E}_\S^\prime$ is defined such that $\chi(v)$ is positive in E_x^u, negative in E_x^s and zero in E_x^0. We suppose that T is hyperbolic, which means that $\dim E_x^0 = 0$ for almost every point x. For simplicity we also suppose that $d^+ \stackrel{\text{def}}{=} \dim E_x^u$ and $d^- \stackrel{\text{def}}{=} \dim E_x^s$ are constants in M (and $d^+ + d^- = d$). The subspace E_x^u and its vectors are called unstable, while E_x^s – stable.

The relation (1) guarantees an asymptotically exponential rate of expansion and contraction of vectors $v \in E_x^{u,s}$ under the iterates of DT. However, we need also a certain monotonicity of this expansion (and contraction). To this end we suppose that in $\mathcal{TM}$ there exists a metric (or at least a pseudometric, which is a metric if restricted to the subspaces E_x^u and E_x^s), such that $\|DTv\| \geq \|v\|$ for all $v \in E_x^u$ and $\|DTv\| \leq \|v\|$ for all $v \in E_x^s$ and $x \in M$. In what follows only this "monotone" metric is used. In applications [1-10] this metric can be always constructed in some way.

According to the Pesin theory, for almost every point $x \in M$ there exist d^+-dimensional local unstable manifold (abbr., LUM) $\gamma^u(x) \subset M$ and d^--dimensional local stable manifold (LSM) $\gamma^s(x) \subset M$, both contain the point x. The former is contracted exponentially fast under T^n as $n \to -\infty$ (in the past), and the latter – as $n \to \infty$ (i.e. in the future). For any point $y \in \gamma^{u,s}(x)$ the subspace $E_y^{u,s}$ is a tangent space to $\gamma^{u,s}(x)$. For hyperbolic systems with singularities (when $S_- \neq \emptyset$) the manifolds $\gamma^{u,s}$ have been constructed in [11]. In this case they are of finite size and have the boundary. So, we can define two functions $r^{u,s}(x) = \{$the distance of the point x to the boundary $\partial\gamma^{u,s}(x)$ in the internal metric of the submanifold $\gamma^{u,s}(x)$, induced by our "monotone" metric $\|\cdot\|\}$. Besides, the boundary $\partial\gamma^u(x)$ lies on the singularity manifold S_n for $n \geq 0$, and the boundary $\partial\gamma^s(x)$ lies on S_n for $n \leq 0$. Then it is reasonable to denote $k^u(x) = \min\{k \geq 0 : \partial\gamma^u(x) \bigcap S_k \neq \emptyset\}$ and $k^s(x) = \min\{k \geq 0 : \partial\gamma^s(x) \bigcap S_{-k} \neq \emptyset\}$.

Definition. A point $x \in M$ is called u-essential if for any $\Lambda > 1$ there exist $n \geq 0$ and a neighborhood $V(x)$ such that for any point y in $V(x)$ and any vector $v \in E_y^u$ (when this subspace exists) one has $\|DT^n v\| \geq \Lambda \|v\|$.

In other words, in the neighborhood $V(x)$ any LUM is expanded in at least Λ times in all directions after n iterates of T. An s-essential point is defined in the same way (but one should replace E^u by E^s and set $n \leq 0$).

Definition. A point $x \in M$ is called sufficient if there exist $\Lambda > 1$, two integers $n_1 < n_2$ and a neighborhood V of the point $T^{n_1}x$ such that for any point $y \in V$ and any vectors $v_1 \in E_y^u$ and $v_2 \in E_y^s$ (when these subspaces exist) one has $\|DT^{n_2 - n_1} v_1\| \geq \Lambda \|v_1\|$ and $\|DT^{n_2 - n_1} v_2\| \leq \Lambda^{-1} \|v_2\|$.

In other words, there is a finite time interval $[n_1, n_2]$ during which any LUM starting near the point $T^{n_1}x$ is expanded in at least Λ times in all directions, and any LSM starting near the same point is contracted in at least Λ times.

Our definitions of essential and sufficient points differ from those introduced in [2,3,7], but ours are formulated in the weakest possible form. In applications the sufficient and essential points are usually defined in more explicit terms such as invariant cone fields [12], or special quadratic forms [6] or geometrical properties like the convexity or concavity [2,7] etc., which are easier to check.

Below we formulate all the conventional assumptions for the local ergodic theorem.

Property 1 (double singularities) *For any finite $n \geq 1$ the intersection $S_{1,n} \cap S_{-1,-n}$ (this consists of the singular points of both T^n and T^{-n}) is a finite union of compact submanifolds of dimension $\leq d - 2$.*

Property 2 (continuity) *The fields of subspaces E_x^u and E_x^s are continuous at each sufficient point. Morover, the limit spaces $\lim_{y \to x} E_y^u$ and $\lim_{y \to x} E_y^s$ should be transversal at a sufficient point x.*

Property 3 (thickness of singularity region) *For $\varepsilon > 0$ denote U_ε the ε- neighborhood of the set $S_{-1,1}$. Then, $\nu(U_\varepsilon) < const \cdot \varepsilon$.*

Property 4 (ansatz) *Almost every point of the submanifold S_1 (wrt the internal Riemannian metric) is u-essential, and almost every point of S_{-1} is s-essential.*

Property 5 (parallelization) *The angles between the submanifold S_n and the adjacent LUM's (LSM's) uniformly tend to zero as $n \to \infty$ (resp., as $n \to -\infty$).*

Note that the property 3 is far from being trivial because the ε-neighborhood is defined through the "monotone" metric, not through the regular one. If this monotone metric is actually a pseudo-metric, then the neighborhood U_ε may be quite large. In this case the choice of the pseudo-metric becomes a very delicate matter – see, for instance, [10].

The property 3 is needed for proving the following two estimates on the distribution of the lengths of LUM's and LSM's. Let x be a sufficient point and $V(x)$ be a sufficiently small neighborhood of x. Then [2,7] for any $\delta > 0$

$$\nu\{y \in V(x) : r^{u,s}(y) < \delta\} \leq C(x)\delta. \tag{2}$$

Moreover, if $F(\delta)$ is a function such that $F(\delta) \to \infty$ as $\delta \to 0$, then

$$\nu\{y \in V(x) : r^{u,s}(y) < \delta \text{ and } k^{u,s} > F(\delta)\} \leq C(x)\delta\varphi(\delta), \tag{3}$$

where $\varphi(\delta) \to 0$ as $\delta \to 0$. The estimate (2) states that the "short" LUM's and LSM's occupy not so large subset of the neighborhood $V(x)$, while the estimate (3) tells us that most of these "short" LUM and LSM end ("are torn") in the images of the singularity set in a not so distant future (or past), i.e. in S_n for not so large values of n.

Theorem 1 (on local ergodicity) *If a hyperbolic system (M, T, ν) possesses the above properties 1-5, then each sufficient point $x \in M_0$ has a neighborhood $V(x)$ belonging (almost surely) in one ergodic component of the transformation T.*

The proof of Theorem 1 first appeared in [2] for the billiard systems with a semidispersing boundary. Then it was modified and generalized in [7,6,5]. In applications [1-10] the properties 1-5 usually hold, although their verification (especially that of the properties 4 and 5) is not so simple. Only one of the cited examples – the system of falling balls introduced by M.Wojtkowski in [8,9] (when the number of balls is three or more) – does not possess the property 5. This means that the submanifolds S_n can deviate from the directions of the LUM's very far for arbitrary large $n > 0$. These manifolds can even approach the directions of the LSM's!

The aim of this paper is to strengthen Theorem 1 by replacing the condition 5 with a weaker one:

Property 6 (transversality) *At almost every point x of the submanifold S_1 (wrt the internal metric in S_1) the subspace E_x^s is transversal to S_1, i.e. $E_x^s \not\subset \mathcal{T}_{\S}\mathcal{S}_{\infty}$. Analogously, $E_x^u \not\subset \mathcal{T}_{\S}\mathcal{S}_{-\infty}$ for almost every point $x \in S_{-1}$.*

One should remember that $\mathcal{T}_{\S}\mathcal{S}_{\infty}$ is a subspace of codimension one in the tangent space $\mathcal{T}_{\S}\mathcal{M}$, so either the intersection $E_x^s \cap \mathcal{T}_{\S}\mathcal{S}_{-\infty}$ is a subspace of codimension one in E_x^s (in the transversal case) or E_x^s itself is just a subspace of $\mathcal{T}_{\S}\mathcal{S}_{\infty}$. Furthermore, in the case $E_x^s \not\subset \mathcal{T}_{\S}\mathcal{S}_{\infty}$ we define an angle between the subspaces E_x^s and $\mathcal{T}_{\S}\mathcal{S}_{\infty}$ as the angle between the vector $v \in E_x^s$ orthogonal to the intersection $E_x^s \cap \mathcal{T}_{\S}\mathcal{S}_{\infty}$ and the subspace $\mathcal{T}_{\S}\mathcal{S}_{\infty}$.

The property 6 allows the singularity manifolds $S_1, S_2, \ldots, S_n, \ldots$ to deviate from the LUM's and even to approach the LSM's, but it forbids them to become parallel to the LSM's. Note that if the property 6 fails, then some open region on the submanifold S_1 foliates with a smooth family of LSM's. In the latter case this region on S_1 can actually separate two ergodic components in the phase space M. At least, the main idea of the proof of Theorem 1 (the construction of so called "Hopf chain", joining two arbitrary points in $V(x)$) does not work, because there are obstacles ("walls") through which none of the Hopf chains can go. These obstacles are obtained by the images of the above open region in S_1 in the future. So, one cannot weaken the condition 6 any more, it seems to be the maximal possible relaxation of the condition 5.

Let us point out another example of a hyperbolic system possessing the property 6 but not the property 5, in addition to the above falling balls. Consider a hyperbolic toral automorphism [14] and then cut the underlying torus into several parts along a finite number of smooth hypersurfaces. Suppose one can displace and permute these parts so that they will cover again the entire torus. If after this permutation one applies the linear hyperbolic automorphism, then one obtains a piecewise linear hyperbolic transformation. The cut surfaces may be chosen arbitrary but the conditions 1 and 6 must hold. In particular, no open part of the cut surfaces should foliate with a smooth family of stable or unstable subspaces. As an illustrative example, one can cut two identical but disjoint balls out of the torus and then interchange the balls keeping the outside area unmovable.

The rest of this paper contains the proof of Theorem 1 under the assumptions 1-4 and 6 (instead of 5). We suppose that the reader is familiar with the proof of Theorem 1 due to the papers [2,7]. The main tool of the proof is a parametrized family of coverings of the neighborhood $V(x)$ with special systems of subsets called parallelograms $G_1^{(\delta)}, \dots, G_{I(\delta)}^{(\delta)}$, where $\delta > 0$ is a small parameter of the family (note, that this construction has first appeared in [1]). In the proof these parallelograms are classified into "good" ones which contain an ample set of points with LUM's and LSM's of size $> \delta$, and "bad" ones – the rest of them. The total measure of "bad" parallelograms is estimated as $o(\delta)$, i.e. it is equal to $\delta\varphi(\delta)$ where $\varphi(\delta) \to 0$ as $\delta \to 0$. The replacement of the condition 5 with 6 causes difficulties with only parallelograms intersecting exactly one of the smooth components of the set $S_{-N,N}$, where $N = F(\delta)$ is a large number specified in the proof, see [2,7].

Consider a parallelogram $G = G_i^{(\delta)}$ intersecting a smooth submanifold $R \subset S_n$, where $|n| \le N$. Let also $n > 0$ for the sake of definiteness. Suppose first that the angle between the subspace E_y^s and the tangent space $\mathcal{T}_\dagger\mathcal{R}$ at any point $y \in R \cap G$ is greater than ε_1 (where $\varepsilon_1 = \varepsilon_1(\delta)$ is to be specified below). Show then that such a parallelogram is "good". Indeed, since $n > 0$ the parallelogram G contains an ample set of points with "long" (of size $> \delta$) LSM's. More precisely, these LSM's fill a subset $G' \subset G$ of measure $\nu(G') > (1 - \varepsilon_2)\nu(G)$ where $\varepsilon_2(\delta) \to 0$ as $\delta \to 0$ (the proof of this statement goes the same way as in [2,7]). Besides, it is easy to choose ε_1 such that $\varepsilon_1 \gg \varepsilon_2$ but $\varepsilon_1(\delta) \to 0$ as $\delta \to 0$ (e.g., let $\varepsilon_1 = \sqrt{\varepsilon_2}$). One gets then that there are "long" LSM's in G which go through the hypersurface R and so connect two parts of G separated by the hypersurface R. Both these parts also contain ample sets of points with LUM's which are either adjacent to R or extended to the boundary ∂G. This is enough for a parallelogram to be used in the construction of the Hopf chain, i.e. to be "good" (cf. [7], lemma 3.11).

Now suppose the contrary: the angles between E_y^s and $\mathcal{T}_\dagger\mathcal{R}$ for $y \in R \cap G$ are less than ε_1. We add such a parallelogram to the collection of the "bad" ones, and then our aim is to estimate their total measure. Let G be one of such parallelograms intersecting a smooth manifold $R \subset S_n$ for some n, $0 < n \le N$. Then in virtue of the property 2 the angles between the subspaces E_y^u for $y \in G$ and $\mathcal{T}_\dagger\mathcal{R}$ for $y \in R \cap G$ exceed a positive constant

$c(x_0) > 0$. This implies that $r^u(y) < c_1(x)\delta$ for all $y \in G$ and another constant $c_1(x) > 0$. Combining this estimate with the inequalities (2), (3) and then applying the property 6 together with the fact that $\varepsilon_1(\delta) \to 0$ as $\delta \to 0$ we conclude that the total measure of parallelograms under consideration does not exceed $o(\delta)$. Hence, the statement of Theorem 1 is proven under the assumptions 1-4 and 6.

As in [2-7], our theorem admits the following modification:

Theorem 2 *Let a hyperbolic system (M, T, ν) possess the properties 1-4 and 6. Then each point x belonging either in M_0 or in just one smooth component of the set $S_{-\infty,\infty}$ has a neighborhood $V(x)$ which is contained (almost surely) in one ergodic component of T.*

Finally, let us briefly describe how Theorem 2 should be used for proving the ergodicity in applications. If the set of points x indicated in Theorem 2 is of full measure and arcwise connected, then certain standard simple reasonings show that the transformation T has the only ergodic component, so it is ergodic and even has a K-property, see [2-5]. Furthermore, the points of intersections of two or more smooth components of the set $S_{-\infty,\infty}$ cannot prevent the above set of points from being arcwise connected due to the property 1. On the other hand, the set of insufficient points may be either empty (as a result of strong hyperbolic properties of T, see for instance [10]) or rather complicated and intricate. In the latter case it may require a lot of work to overcome its influence, as was the case in [3,4]. Certain experience on this subject is gained in studying concrete systems in [1-10].

References

[1] Ya.G.Sinai, Ergodic properties of the Lorentz gas, Funkt. analysis and appl. **13** (1979), 192-202.

[2] Ya.G.Sinai, N.I.Chernov, Ergodic properties of some systems of 2-dimensional discs and 3-dimensional spheres, Russ. Math. Surv. **42** (1987), 181-207.

[3] A.Krámli, N.Simányi, D.Szász, Three Billiard Balls on the ν-Dimensional Torus is a K-Flow, Annals of Mathematics, **133** (1991), 37-72.

[4] L.A.Bunimovich, C.Liverani, A.Pellegrinotti, Yu.M.Sukhov, Ergodic Systems of N Balls in a Billiard Table, preprint 1991.

[5] V.Donnay, C.Liverani, Potentials on the Two-Torus for Which the Hamiltonian Flow is Ergodic, Comm. Math. Phys. **135** (1991), 267-302.

[6] R.Markarian, The Fundamental Theorem of Sinai-Chernov for Dynamical Systems with Singularities, preprint 1991.

[7] A.Krámli, N.Simányi, D.Szász, A "Transversal" Fundamental Theorem for Semi-Dispersing Billiards, Comm. Math. Phys. **129** (1990), 535-560.

[8] M.Wojtkowski, A System of One Dimensional Balls With Gravity, Comm. Math. Phys. **126** (1990), 507-533.

[9] M.Wojtkowski, The System of One Dimensional Balls in An External Field, Comm. Math. Phys. **127** (1990), 425-432.

[10] N.I.Chernov, The Ergodicity of a Hamiltonian System of Two Particles in an External Field, Physica D (in press).

[11] A.Katok, J.-M.Strelcyn, Invariant manifolds, entropy and billiards; smooth maps with singularities, Lecture Notes in Mathematics vol 1222, Springer, New York, 1986.

[12] V.I.Oseledec, A Multiplicative Ergodic Theorem: Characteristic Lyapunov Exponents of Dynamical Systems, Trans. Moscow Math. Soc. **19** (1968), 197-231.

[13] M.Wojtkowski, Invariant Families of Cones and Lyapunov Exponents, Erg. Th. Dyn. Syst. **5** (1985), 145-161.

[14] Dynamical Systems II, Encyclopaedia of Mathematical Sciences, 2, Springer, Berlin, 1989.