TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 356, Number 4, Pages 1637–1689
S 0002-9947(03)03335-X
Article electronically published on September 22, 2003

LIMIT THEOREMS FOR PARTIALLY HYPERBOLIC SYSTEMS

DMITRY DOLGOPYAT

Abstract. We consider a large class of partially hyperbolic systems containing, among others, affine maps, frame flows on negatively curved manifolds, and mostly contracting diffeomorphisms. If the rate of mixing is sufficiently high, the system satisfies many classical limit theorems of probability theory.

1. Introduction

The study of the statistical properties of deterministic systems constitutes an important branch of smooth ergodic theory. According to a modern view, a chaotic behavior of deterministic systems is caused by the exponential instability of nearby trajectories. The best illustration of this statement is provided by Axiom A diffeomorphisms, where the expansion of some directions and the contraction of complementary ones are uniform. Both qualitative [2, 3, 8] and quantitative [64, 88, 91] properties of such systems are well understood.

Much less is known in other cases, in spite of significant advances in the recent years. There are two main ways of weakening the uniform hyperbolicity conditions [68]. The first one is the theory of nonuniformly hyperbolic systems of Pesin [66, 67]. (Some refinements of this theory are given in [48, 74, 18, 63].) Now the qualitative behavior of such systems is quite well understood. Interesting results concerning the quantitative theory are obtained in [17, 58, 91, 92].

The second direction of research is the theory of partially hyperbolic systems. Here hyperbolicity should be uniform, but only in some directions. The attraction of this theory is that the question about ergodic properties of a single diffeomorphism is reduced to understanding the ergodic behavior of a usually large holonomy group [13], and the larger the group, the fewer invariant sets it has. Even though currently there are significant technical difficulties in justifying this reduction, the conditions of the theorems obtained this way are relatively easy to check (see [75, 76, 90, 46, 14, 15]) without the formidable analytic work common in nonuniformly hyperbolic theory.

In any case, the results of [35, 75, 76, 7] show that there is a non-trivial theory applicable to a large class of partially hyperbolic systems. Our paper concerns limit theorems for partially hyperbolic systems. More precisely, similarly to the nonuniformly hyperbolic situation, we study the relation between mixing properties

Received by the editors April 17, 2002 and, in revised form, March 19, 2003.
2000 *Mathematics Subject Classification.* Primary 37D30; Secondary 60Fxx.
Key words and phrases. Partial hyperbolicity, central limit theorem, Gibbs measure, absolute continuity.
This work was partly supported by an Elisabeth Proctor Fellowship at Princeton, a Miller Fellowship at Berkeley, and a Sloan Fellowship at PennState.

of the system and the limit theorems it satisfies. The paper [58] shows that it is more convenient to work with a qualitative version of the K-property.

Central to this approach is a notion of an almost Markov family. This is a slight generalization of a Markov family, but its construction is much simpler. An example of an almost Markov family is given by the set of all domains with bounded geometry of the boundary.

Following [58], we assume that for some almost Markov family the images of all elements under the iterations of our system become uniformly distributed. The rate of convergence is essentially independent of the choice of the almost Markov family, and so it is a natural measure of the speed of K–mixing.

Remark. We note that almost sure convergence suffices for the K–property; we require uniform convergence, so there are K–systems with zero convergence rate [45]. In principal, in many places it should be possible to replace uniform estimates by L^1-bounds, but the proofs would become much more complicated. Also there are many simple systems enjoying the K–property yet not satisfying the central limit theorem and other limit theorems of probability theory. Thus in this paper we restrict ourselves to uniform convergence.

The result of our study is the generalization of many limit theorems which were previously known in the Anosov or Axiom A context ([78, 23, 70, 40, 55]) to a large class of partially hyperbolic systems. Some of our results were known before (see Section 6). However, our results seem to be the most general ones currently available for partially hyperbolic systems, implying all that was known before and presenting a unified proof for many seemingly different systems.

In the next section we define the class of the systems we consider. We also recall the notion of u-Gibbs state introduced in [69] and playing a central role in our analysis. Section 3 describes some simple properties of systems with unique u-Gibbs state. The statements of our main results are given in Sections 4 and 5. They are based on the assumption that the system under consideration has a unique u-Gibbs state with good mixing properties (mixing is understood in the sense described above). Section 4 contains various versions of the central limit theorem, and Section 5 presents various other results. In Section 6 we apply our results to classical partially hyperbolic systems. The proofs of the statements of Section 3 are given in Sections 7 and 8. The statements of Section 4 are proved in Sections 9-15. The statements of Section 5 are proved in Sections 16-18.

The appendix collects various results related to the absolute continuity of the unstable foliations for which the author could not find convenient references.

To conclude this section, let us briefly describe possible extensions of our results. First there are some natural classes of non-uniformly partially hyperbolic systems or partially hyperbolic systems with singularities (e.g., some weakly interacting particle systems) where our methods seem to be useful. However, specific features of each particular example seem to be very important in the proofs, so we do not pursue this subject here. Second, a pleasant feature of our approach is that in most cases it is not required that the initial distribution is invariant with respect to dynamics; we only ask that it has smooth conditional measures on unstable leaves. Since we do not assume stationarity, our methods seem to be useful in the study of the time-dependent ([4, 5]) and, in particular, random case (cf. [29]). Third, probably, most of out results are valid for flows with assumption of K-mixing for the flow being replaced by a weaker condition of K-mixing for a suitable Poincaré map as in [78, 23, 50, 51], etc. Also some of our results admit generalizations to the case where instead of one diffeomorphism a family of partially hyperbolic systems is considered.

2. Partial hyperbolicity

Let M be a compact Riemannian manifold and $f : M \to M$ a C^2– diffeomorphism. f is called *partially hyperbolic* if there are an f–invariant splitting

$$T_xM = E_u \oplus E_c \oplus E_s$$

and constants $\lambda_1 \le \lambda_2 < \lambda_3 \le \lambda_4 < \lambda_5 \le \lambda_6$, $\lambda_2 < 1$, $\lambda_5 > 1$, such that

$$\begin{aligned} \forall v \in E_s \quad & \lambda_1||v|| \le ||df(v)|| \le \lambda_2||v||, \\ \forall v \in E_c \quad & \lambda_3||v|| \le ||df(v)|| \le \lambda_4||v||, \\ \forall v \in E_u \quad & \lambda_5||v|| \le ||df(v)|| \le \lambda_6||v||. \end{aligned}$$

We assume that $E_u \neq 0$. On the other hand, the reader can assume in what follows that $E_s = 0$, replacing E_c by $E_c \oplus E_s$. We denote by W^u the foliation tangent to E_u. We say that F is a *u-set* if F belongs to a single leaf of W^u. By volume, diameter and so on of a u-set we mean the volume, diameter, etc. induced by the Riemann structure on W^u.

The important property of W^u is its *absolute continuity.* Call a set A *u-negligible* if it intersects each W^u-leaf at a set of zero leaf volume. We say that some property holds *u-almost surely* if it fails on a u-negligible set. A measure ν is called *u-absolutely continuous* if it assigns zero measure to u-negligible sets. Absolute continuity of W^u means that the Lebesgue measure is u-absolutely continuous. Absolute continuity is the most basic property for the study of statistical properties of Lebesgue–almost every point. Thus it is useful to consider all u-absolutely continuous measures. (Since in this paper we are dealing with u-absolutely continuous measures only, we consider two sets equal if they differ by a u-negligible set. In particular, we do not distinguish between two u-sets if their difference has zero leaf measure.) Among the absolutely continuous measures, the special role belongs to f-invariant ones. u-absolutely continuous f-invariant measures are called *u-Gibbs states.* u-Gibbs states were studied in [69]. Among other things, they show that if F is a nice u-set and μ is the normalized Lebesgue measure on F, then any limit point of $\frac{1}{n}\sum_{j=0}^{n-1} f_*^j\mu$ is a u-Gibbs state. In this paper we study partially hyperbolic systems satisfying two requirements. First, they have the unique u-Gibbs state ν. Second, not only the Birkhoff averages of μ but $f_*^j\mu$ itself converges to ν. To give the precise formulation we need to define a collection of nice u-sets.

A collection $\mathcal{P}$ of u-sets is called *an almost Markov family* if there are constants r_1, r_2, v, C, γ such that $\forall P \in \mathcal{P}$

(a) $\operatorname{diam}(P) \le r_1$;

(b) $\operatorname{Vol}(P) \ge v$;

(c) $P = \overline{\operatorname{Int}(P)}$, and $\operatorname{Vol}\{p : d(p, \partial P) \le \varepsilon\} \le C\varepsilon^\gamma$;

(d) for any u-set F there are disjoint sets $P_i \in \mathcal{P}$ such that $\bigcup_i P_i \subset F$ and $F \setminus \bigcup_i P_i \subset \{p : d(p, \partial F) \le r_2\}$;

(e) $\bigcup_{\mathcal{P}} P = M$.

An almost Markov family is called *Markov* if

(f) for any $P \in \mathcal{P}$ there are $P_i \in \mathcal{P}$ such that $fP = \bigcup_i P_i$.

Proposition 1. *Any f has a Markov family.*

(In [81] a family of sets satisfying (f) but not (a)–(e) was constructed. The family satisfying (a)–(e) as well is obtained in [82]. Formally, [82] proves the existence of Markov *partitions* for Anosov diffeomorphisms (i.e. when $E_c = 0$). However, this is done by constructing the Markov families for f and f^{-1} and showing that they can be fitted together nicely. It can be seen that the construction of the Markov family for f never uses the assumption that $E_c = 0$, so it is valid for arbitrary partially hyperbolic f.)

Examples of almost Markov families.

(I) If $r_1$and C are large and v is small, then the collection of all sets satisfying (a)–(c) is an almost Markov family.

(II) If $\dim E_u = 1$, then the set of all curves of length between 1 and 2 is a Markov family.

(III) If $\mathcal{P}$ is an almost Markov family and F is a domain in some leaf of W^u with piecewise smooth boundary, then $\mathcal{P} \cup \{F\}$ is an almost Markov family.

We can associate to each u-set F a probability density as follows. For $x_1, x_2 \in F$ let

$$\rho(x_1, x_2) = \prod_{j=0}^{\infty} \frac{\det(df^{-1}|E_u)(f^{-j}x_1)}{\det(df^{-1}|E_u)(f^{-j}x_2)}.$$

Choose $x_0 \in F$ and let $\rho_F(x) = C\rho(x, x_0)$, where $C = \left(\int_F \rho(x, x_0)dx\right)^{-1}$. (Here '$\int_F dx$' means the integration over the leaf of W^u containing F with the induced volume form.) Since $\rho(x, x_0') = \rho(x, x_0)\rho(x_0, x_0')$, this definition does not depend on the choice of x_0. If $A \in C(X)$, then $\int_F A(fx)\rho_F(x)dx = \int A(y)\rho_F(y)dy$.

Let $\mathcal{P}$ be an almost Markov family, P a u-set satisfying (a)–(c), and n a natural number. By (d) $\exists P_j \in \mathcal{P}$ such that

$$f^n P = (\bigcup_j P_j) \cup Z, \tag{1}$$

where $Z \subset \{x : d(x, \partial f^n P) \le r_2\}$. We call (1) *an almost Markov decomposition of $f^n P$ (with respect to $\mathcal{P}$)*. Let $c_j = \int_{f^{-n}P_j} \rho_P(x)dx$, $c = \int_{f^{-n}Z} \rho_P(x)dx$. Then

$$c \le C_1 \operatorname{meas}(f^{-n}Z) \le C_1 \operatorname{meas}(\{x : d(x, \partial P) \le \frac{r_2}{\lambda_5^n}\}) \le C_2(\frac{r_2}{\lambda_5^n})^\gamma \le C_3\zeta^n$$

for some $\zeta < 1$.

Now let us introduce the measures we consider. Choose an almost Markov family $\mathcal{P}$. Fix some constants R, α. Let $E_1(\mathcal{P}, R, \alpha)$ be the set of the measures given by the following expression: for $A \in C(M)$

$$\ell(A) = \int_P A(x)e^{G(x)}\rho_P(x)dx,$$

where $P \in \mathcal{P}$, $|G(x_1) - G(x_2)| \le Rd(x_1, x_2)^\alpha$ and $\ell(1) = 1$. We will refer to the above functional as $\ell(P, G)$ and write $\ell(P)$ for $\ell(P, 0)$. Let $E_2(\mathcal{P}, R, \alpha)$ be the convex hall of $E_1(\mathcal{P}, R, \alpha)$ and $E(\mathcal{P}, R, \alpha) = \overline{E_2(\mathcal{P}, R, \alpha)}$. Usually we will drop some of the parameters $\mathcal{P}, R, \alpha$ if it does not cause confusion.

Examples of admissible measures.

(a) Probably the most important example is the following.

Proposition 2. *Let $\mathcal{P}$ be a maximal family from Example I of Section 2. If R is large enough and α is small enough, then the Lebesgue measure belongs to $E(R, \alpha)$.*

This follows from the Hölder continuity of E_u and the Hölder continuity of the unstable holonomy Jacobian. See Appendix A.

(b) It is not difficult to see by a standard Kukutani-Markov argument that there is always a u-Gibbs state in $E(0,0)$. Conversely, [69] show that any u-Gibbs state belongs to $E(0,0)$. Below we prove that several sets have full ℓ-measure for any $\ell \in E$. The following statement is useful.

Proposition 3. *The set $Y \subset X$ has zero ℓ-measure for any $\ell \in E$ if and only if it is u-negligible.*

See Appendix A for more details.

3. Formulation of results. uunique ergodicity and strong u-transitivity

Our first assumption throughout this paper is that f has unique u-Gibbs state. We will call such systems *uuniquely ergodic*, and write $f \in$ UuEe. By [69] any limit point of the measures of the form

$$\mu_n(A) = \frac{1}{n}\sum_{j=0}^{n-1} \ell^{(n)}(A \circ f^j), \tag{2}$$

where $\ell^{(n)} \in E$, is a u-Gibbs state. Conversely, any u-Gibbs state ν is a limit point of measures μ_n as above with $\ell^n \equiv \nu$. Thus an equivalent way to define uunique ergodicity is the following.

Definition. f is uuniquely ergodic if $\forall A \in C(M)$, uniformly in $\ell \in E$,

$$\frac{1}{n}\sum_{j=0}^{n-1} \ell(A \circ f^j) \to \nu(A).$$

If $f \in$ UuEe, we have a bound on the rate of convergence for Hölder functions.

Given $A \in C(M)$, let $\mathcal{S}_n(A)(x) = \sum_{j=0}^{n-1} A(f^j x)$. Sometimes we will write simply $\mathcal{S}$ if A is clear.

Theorem 1. *If $f \in UuEe$, then $\forall A \in C^\gamma(M)$ with $\nu(A) = 0$, $\forall \varepsilon\ \exists C_\varepsilon, c_\varepsilon$ such that $\forall \ell \in E$*

$$\ell(|\mathcal{S}_n(A)| > \varepsilon n) \le C_\varepsilon e^{-nc_\varepsilon}.$$

The proof is given in Section 7. Since $C^\gamma(M)$ is dense in $C(M)$, we get

Corollary 1 (Law of Large Numbers)**.** $\forall A \in C(M)$ $\frac{\mathcal{S}_n}{n} \to \nu(A)$ *u-almost surely.*

In dynamical systems language this statement can be reformulated as follows. Let μ be an f-invariant measure. Define the *basin of* μ, $B(\mu)$, to be the set of forward μ-regular points

$$B(\mu) = \{x : \forall A \in C(X)\ \frac{1}{n}\mathcal{S}_n(A) \to \nu(A)\}.$$

μ is called an *SRB measure* if its basin has positive Lebesgue measure. Thus the previous corollary can be restated as follows.

Corollary 2. *If f has a unique u-Gibbs state ν, then ν is also an SRB measure and $B(\nu)$ has whole Lebesgue measure.*

In order to get quantitative results about the behavior of $\mathcal{S}_n$, we need to impose stronger restrictions on f. We say that f is *strongly u-transitive* if, for some almost Markov collection $\mathcal{P}$, $\forall A \in C(M)$ $\forall P \in \mathcal{P}$

$$\int_P A(f^n x)\rho_P(x)dx \to \nu(A), \tag{3}$$

where ν is some probability measure on M. (The argument below shows that this definition is independent of the choice of $\mathcal{P}$.)

Starting from this point, we will assume that f is strongly u-transitive. We need a qualitative bound for the rate of convergence in (3). To formulate this more precisely, let us discuss the space of observables we consider. Let $\mathbb{B}$ be a Banach function algebra such that there is a continuous embedding $i : \mathbb{B} \to C^\gamma(M)$. We assume that there exists a measure ν such that $\forall \ell \in E$ $\forall A \in \mathbb{B}$

$$|\ell(A \circ f^n) - \nu(A)| \le a(n)||A||_{\mathbb{B}}, \tag{4}$$

where $a(n) \to 0$ as $n \to \infty$.

$a(n)$ is essentially independent of the choice of a Markov family. More precisely, we have

Proposition 4. *If $\mathcal{P}'$ is another almost Markov family, then $\forall \ell \in \mathcal{P}'$*

$$|\ell(A \circ f^n) - \nu(A)| \le a'(n)||A||_{\mathbb{B}},$$

where $a'(n) \le C_1 a(\frac{n}{C_2}) + C_3\theta^n$.

Remark. The reader can check that the conditions of all theorems we formulate are stable with respect to replacing $a(n)$ by $C_1 a(\frac{n}{C_2}) + C_3\theta^n$.

Proof. Here and below, θ denotes a constant less than 1 which can change from entry to entry.

Take any $Q \in \mathcal{P}'$. Let $f^{\frac{n}{2}}Q = (\bigcup_j P_j) \cup Z$ be its almost Markov decomposition with respect to $\mathcal{P}$. Take $A \in \mathbb{B}$ with $||A||_{\mathbb{B}} \le 1$. We have

$$I = \int_Q e^{G(x)}\rho_Q(x)A(f^n x)dx = \sum_j c_j \int_{P_j} e^{G(f^{-\frac{n}{2}}y)}\rho_{P_j}(y)A(f^{\frac{n}{2}}y)dy + O(\theta^{\frac{n}{2}}).$$

Choose $y_j \in f^{-\frac{n}{2}}P_j$. Then

$$\begin{aligned} I &= \sum_j c_j e^{G(f^{-\frac{n}{2}}y_j)} \int_{P_j} \rho_{P_j}(y)A(f^{\frac{n}{2}}y)dy + O(\theta^{\frac{n}{2}}) \\ &= \sum_j c_j e^{G(f^{-\frac{n}{2}}y_j)} \left[\nu(A) + O(a(\frac{n}{2}))\right] + O(\theta^{\frac{n}{2}}). \end{aligned}$$

In particular, letting $A \equiv 1$, we get

$$1 = \int_Q e^{G(x)}\rho_Q(x)dx = \sum_j c_j e^{G(f^{-\frac{n}{2}}y_j)} + O(\theta^{\frac{n}{2}}).$$

The last two identities prove the proposition. □

Plugging $\ell = A\nu$ into (3), we see that (f, ν) is mixing. In fact it is also mixing of all orders, as the next statement shows.

Theorem 2 (Multiple mixing). *Fix k. There are constants C_1 and C_2 such that $\forall A_1, A_2 \dots A_k \in \mathbb{B}\ \forall \ell \in E$*

$$\left|\ell\left(\prod_{j=1}^{k} A(f^{n_j}x)\right) - \prod_{j=1}^{k}\nu(A_j)\right| \leq C_1\left[a\left(\frac{m}{C_2}\right)+\theta^m\right]\prod_{j=1}^{k}||A_j||_{\mathbb{B}},$$

where $m = \min(n_j - n_{j-1}), n_0 = 0$.

The proof is given in Section 8.

4. Formulation of results. Central limit theorem

Here we formulate various versions of the central limit theorems for the systems under consideration. Most of the proofs use methods of moments [44].

Throughout this section we assume that $\sum_n a(n) < \infty$.

Theorem 3 (Invariance Principle). *There is a constant $s > 0$ such that the following holds. Let $A \in \mathbb{B}$ be a function such that $\nu(A) = 0$, $\sigma(A) = \sum_{j=-\infty}^{\infty} \nu(A(A\circ f^j)) \neq 0$. Let $\mathcal{P}$ be a Markov family and let $P \in \mathcal{P}$. Then there are a probability space (Ω, μ), and a Brownian motion $w(t)$ and a sequence ξ_n, both defined on Ω, such that*

(a) *the distribution of ξ_n is the same as the distribution of $\mathcal{S}_n(x)$ with respect to $\ell(P)$, and*

(b) *$\exists \sigma_n$ such that $\frac{\sigma_n}{n} \to \sigma(A)$ and $|\xi_n - w(\sigma_n)| \leq C(\omega)n^{\frac{1}{2}-s}$.*

Corollary 3 (Law of the Iterated Logarithm).

$$\limsup \frac{\mathcal{S}_n(x)}{\sqrt{2\sigma(A)n\ln\ln n}} = 1, \quad \liminf \frac{\mathcal{S}_n(x)}{\sqrt{2\sigma(A)n\ln\ln n}} = -1$$

u-almost surely.

Corollary 4 (Central Limit Theorem). *$\forall \mathcal{P}, R, \alpha\ \forall \ell \in E(\mathcal{P}, R, \alpha)$ the random process $X_n(t) = \frac{\sum_{j\leq nt} A(f^j x)}{\sqrt{n}}$ converges weakly to the Brownian motion with average zero and variance $\sigma(A)$.*

Let $\mathbb{B}^d$ denote the space of functions $M \to \mathbb{R}^d$ such that each coordinate belongs to $\mathbb{B}$. Consider the sequence $z_n \in \mathbb{R}^d$ given by

$$z_{n+1} - z_n = \varepsilon A(z_n, f^n x), \qquad z_0 = a, \tag{5}$$

where the function $A(z, x)$ is three times differentiable with respect to z and the norms $||\frac{\partial^\alpha A(z,\cdot)}{\partial^\alpha z}||_{\mathbb{B}^d}$ are uniformly bounded for $0 \leq |\alpha| \leq 3$. Let q_n be the solution of the averaged equation

$$q_{n+1} - q_n = \varepsilon \bar{A}(q_n), \qquad q_0 = a,$$

where

$$\bar{A}(q) = \int A(q, x)d\nu(x).$$

Let $DA(z, x)$ denote the partial derivative of A with respect to z. Let $\Delta_n = z_n - q_n$. Denote $\Delta_t^\varepsilon = \frac{\Delta_{[\frac{t}{\varepsilon}]}}{\sqrt{\varepsilon}}$.

Theorem 4 (Short time fluctuations in averaging). *If $a(n) \leq \frac{\text{Const}}{n^2}$, then $\forall \mathcal{P}, R, \alpha$ $\forall \ell \in E(\mathcal{P}, R, \alpha)$, as $\varepsilon \to 0$ the function Δ_t^ε converges weakly to the solution of*

$$d\Delta(t) = D\bar{A}(q(t))\Delta dt + dB,$$

where B is a Gaussian process with independent increments, zero mean and covariance matrix

$$\langle B, B\rangle(t) = \int_0^t \sigma(A(q(s), \cdot))ds. \tag{6}$$

Theorem 5 (Long time fluctuations in averaging). *Suppose that A in* (5) *has zero mean,*

$$\bar{A}(z) = \int A(z, x)d\nu(x) \equiv 0,$$

and that $a(n)n^2 \to 0$ as $n \to \infty$. Let $Z_t^\varepsilon = Z_{[\frac{t}{\varepsilon^2}]}$. Then, as $\varepsilon \to 0$, Z_t^ε converges weakly to the diffusion process $Z(t)$ with drift

$$a(z) = \sum_{n=1}^{\infty} \int DA(z, f^n x)A(z, x)d\nu(x)$$

and diffusion matrix $\sigma(A(z, \cdot))$.

Remark. As usual, after Theorem 5 is proved for smooth bounded functions, the stopping time argument can be used to extend it to the more general framework where the limiting diffusion process has no explosions.

The proofs of the results of this section are given in Sections 9–15. Sections 9 and 10 contain some auxiliary estimates. Theorem 3 is proven in Section 11, Theorem 4 in Sections 12–13 and Theorem 5 in Section 15.

Note. Surveys on central limit theorems for dynamical systems can be found in [21, 30, 17].

5. Formulation of results. Other limit theorems

Theorem 6 (Three Series Theorem). *If $\sum_n a(n) \leq \infty$, $A_n \in \mathbb{B}$, $||A_n||_{\mathbb{B}} \leq 1$ and c_n is a sequence such that $\sum_n c_n \nu(A_n) < \infty$, $\sum_n c_n^2 < \infty$, then $\sum_n c_n A_n(f^n x)$ converges u-almost surely.*

The proof is given in Section 16.

To formulate our next results, we suppose that ν has a smooth density. We assume also that for any ball B of radius r, and for any $\ell \in E$,

$$|\ell(1_B(f^n x)) - \nu(B)| \leq \text{Const}\, r^{-\alpha} \left(\frac{1}{n}\right)^k,$$

where $1_B(x)$ stands for the indicator function of B. Denote $d = \dim(X)$, $d_u = \dim E_u$.

Theorem 7 (Borel-Cantelli Lemma). *Assume that $\frac{k}{\alpha+1} > \frac{d}{d_u}$. If $\{B_n\}$ is a sequence of balls, then $\sum_n 1_{B_n}(f^n x)$ converges ν–almost surely $\Leftrightarrow \sum_n r(B_n)^d < \infty$, and $\sum_n 1_{B_n}(f^n x)$ diverges ν–almost surely $\Leftrightarrow \sum_n r(B_n)^d = \infty$.*

The proof is given in Section 17.

Theorem 8 (Poisson Law). *Assume that* $\frac{k}{\alpha+1} > \frac{1}{d_u}$. *Let* x_0 *be a non-periodic point and* $B_n = B(x_0, r)$. *Denote* $X_n(\Delta) = \sum_{j\nu(B_n)\in\Delta} 1_{B_n}(f^j(x))$. *Then,* $\forall \ell \in E$, *as* $n \to \infty$, $X_n(\Delta)$ *converges to a Poisson process with density* 1.

The proof is given in Section 18.

6. Applications

Here we give some examples to which our theorems apply. The main examples of strongly transitive systems belong to the class of Anosov actions. (See [73, 12, 31, 47] for general discussions of the Anosov actions.) In this case E_c is the tangent space to the orbits of some Lie group G and $f(x) = g_f x, g \in G$. We hope, however, that more examples of systems satisfying our assumptions will appear with the further development of the theory of partially hyperbolic systems (cf. [1, 7, 80, 27]).

Throughout this section, we say that f *is strongly u-transitive with exponential rate* if (4) is satisfied with $\mathbb{B} = C^\gamma(M)$ and $a(n) = C\theta^n$ for some $\theta < 1$. We say that f *is strongly u-transitive with superpolynomial rate* if for each r there is $k = k(r)$ such that (4) is satisfied with $\mathbb{B} = C^k(M)$, and $a(n) = C_r n^{-r}$.

(a) Anosov diffeomorphisms. These are defined by the condition that $E_c = 0$. This is perhaps the most studied class of partially hyperbolic systems (see [2, 3, 8]), and most of our results are well known for Anosov diffeomorphisms.

Proposition 5 (see e.g. [8]). *Topologically transitive Anosov diffeomorphisms are strongly u-transitive with exponential rate.*

Corollary 5. *All theorems of Sections 4 and 5 hold true for topologically transitive Anosov diffeomorphisms.*

(b) Time one maps of Anosov flows. These are Anosov actions with $G = R$.

Proposition 6. (a) ([24, 25]) *Suppose that* f *is a time one map of a topologically transitive Anosov flow whose stable and unstable foliations are jointly non-integrable. Then* f *is strongly u-transitive with superpolynomial rate. If in addition* E_u *and* E_s *are* C^1 *then* f *is strongly u-transitive with exponential rate.*

(b) ([59]) *Time one maps of contact Anosov flows are strongly u-transitive with exponential rate.*

Corollary 6. *Time one maps of topologically transitive Anosov flows with jointly non-integrable stable and unstable foliations satisfy the conclusions of Theorems 3–6 and their corollaries. If in addition* E_u *and* E_s *are* C^1 *or the flow preserves a contact structure, then all the results of Sections 4 and 5 apply.*

Remark. It is easy to see that strong u-transitivity with exponential rate implies exponential convergence in (4) for piecewise Hölder functions such as indicators of balls. On the other hand, strong u-transitivity with superpolynomial rate gives only power decay for indicators. For this reason it is unclear if Theorems 7 and 8 hold for time one maps of arbitrary Anosov flows.

(c) Compact skew extensions of Anosov diffeomorphisms. Let $h : N \to N$ be a topologically transitive Anosov diffeomorphism, K a compact connected Lie group, $M = N \times K$, and $\tau : N \to K$ a smooth map. Let $f(x, y) = (hx, \tau(x)y)$. Thus here $G = \mathbb{Z} \times K$. Compact skew extensions are studied in [11, 12, 15, 26].

Proposition 7 ([26]). *Generic skew extension is strongly u-transitive with superpolynomial rate. In particular, if K is semisimple then all ergodic extensions are strongly u-transitive with superpolynomial rate. Also, if N is an infranilmanifold, then all stably ergodic extensions are strongly u-transitive with superpolynomial rate.*

Corollary 7. *Generic compact skew extensions of Anosov diffeomorphisms satisfy the conclusions of Theorems 3–6 and their corollaries.*

(d) Quasihyperbolic toral automorphisms. Here $M = \mathbb{T}^d$ and $f(x) = Qx \pmod 1$, where $Q \in SL_d(\mathbb{Z})$, $sp(Q) \not\subset S^1$.

Proposition 8 ([45]). *Quasi-hyperbolic toral automorphisms are strongly u-transitive with exponential rate.*

Corollary 8. *All theorems of Sections 4 and 5 hold for quasihyperbolic toral automorphisms.*

(e) Translations on homogeneous spaces. Let $M = G/\Gamma$, where G is a connected semisimle group without compact factors and Γ is an irreducible compact lattice in G. Let $f(x) = gx$, $g = \exp(X)$.

Proposition 9 ([53]). *Suppose that there is a factor G' of G which is not locally isomorphic to $SO(n,1)$ or $SU(n,1)$ and such that the projection g' of g to G is not quasiunipotent (i.e. $sp(\mathrm{ad}(g')) \not\subset (S^1)$). Then f is strongly u-transitive with exponential rate.*

Corollary 9. *All theorems of Sections 4 and 5 hold true for translations of homogeneous spaces satisfying the conditions of the last proposition.*

(f) Mostly contracting diffeomorphisms. Let $f : M \to M$ be partially hyperbolic. f is called *mostly contracting* if $\exists \epsilon > 0$ such that for any u-Gibbs state ν

$$\lim_{n\to\infty} \frac{\nu(\ln ||df^n|E_c||)}{n} \leq -\epsilon.$$

See [7, 16, 27] for examples of mostly contracting diffeomorphisms.

Proposition 10 ([27]). *Suppose that $f : M \to M$ is a mostly contracting topologically mixing diffeomorphism,* $\dim(M) = 3$, $\dim(E_c) = 1$. *Then f strongly u-transitive with exponential rate.*

Remark. It is likely that the restrictions on dimensions given here are unnecessary (cf. [16]).

Corollary 10. *All theorems of Sections 4 and 5 hold true for mostly contracting topologically mixing diffeomorphism on three-dimensional manifolds.*

Remark. The set of mostly contracting diffeomorphisms is open. The simplest examples of mostly contracting diffeomorphisms can be constructed by perturbing Anosov actions. Thus this result is the first step in extending our results beyond Anosov actions.

Other examples of diffeomorphisms satisfying our conditions could be constructed using following observations. Let $M = M_1 \times M_2$ and $f = f_1 \times f_2$, where the f_j are partially hyperbolic. If both f_1 and f_2 are strongly u-transitive with either exponential or superpolynomial rate, then the same is true for f.

Notes. As we mentioned before, not all of these results are new. Below we list the results which were known before:

- Anosov diffeomorphisms: Theorem 3 and Corollary 3 ([23]), Theorem 4 ([50]), Corollary 4 ([78]), Theorem 6 ([55]), Theorem 8 ([40]). These articles also consider Anosov flows, but instead of time one maps they deal with
$$\mathcal{S}^{(t)}(A) = \int_0^t A(g_s x) ds,$$
where g_s is the flow in question. Our results are therefore slightly stronger. Let us remark, by the way, that our formulations might be more appropriate from the point of view of applications, because in practise it is possible to measure $\mathcal{S}_n(A, g_1)$ rather than $\mathcal{S}^{(t)}(A)$. On the other hand, the results for $S^{(t)}$ are usually proven under a weaker assumption than that of Proposition 6 (and Corollary 6 is false under these weaker assumptions). However, it seems possible to extend our results to treat the case when (4) holds not for a time one map of a flow but for a suitably chosen Poincaré map.
- Quasihyperbolic toral automorphsims: Theorem 3, Corollaries 3 and 4 ([33, 34]), Theorem 4 ([65]).
- Translations on homogeneous spaces: Theorem 3, Corollaries 3 and 4 ([56, 57]), Theorem 4 ([65]). Also, [54, 87] contain results quite similar in spirit to our Theorems 6 and 7, even though Theorems 6 and 7 are not explicitly stated there. ([56, 57, 54, 87] do not suppose that M is compact, requiring only that $\mathrm{Vol}(M) < \infty$.)

However, the advantage of our method is that we give a unified proof for all these different classes of dynamical systems, and this proof would seem to be of interest even in the known cases.

7. Large deviations

Here we prove Theorem 1. First we verify our claim for $\ell = \ell(P) \in E(\mathcal{P}, 0, 0)$, where $\mathcal{P}$ is a Markov family. It is enough to estimate $\ell(\mathcal{S}_n(A) > \varepsilon n)$; the case $\ell(\mathcal{S}(A) < -\varepsilon n)$ is dealt with similarly. Denote $B(x) = A(x) - \frac{\varepsilon}{2}$. By our assumption there exists n such that $\forall P \in \mathcal{P}$
$$\int_P \mathcal{S}_n(B)(x)\rho_P(x)dx \leq -\frac{n\varepsilon}{4}.$$
Also there exists some C such that $\forall P \in \mathcal{P}$ $\forall n$
$$\mathrm{Osc}_P(\mathcal{S}_n(B) \circ f^{-n}) \leq C, \tag{7}$$
where $\mathrm{Osc}_P(A) = \max_P(A) - \min_P(A)$.

Hence $\exists n, \alpha < 0$ such that $\forall P \in \mathcal{P}$ for any decomposition $f^n P = \bigcup_j P_j$, $P_j \in \mathcal{P}$,
$$\sum_j c_j \max_{f^{-n}P_j} \mathcal{S}_n(B) \leq \alpha,$$
where $c_j = \int_{f^{-n}P_j} \rho_P(x)dx$.

Corollary 11. $\exists \gamma > 0, \theta < 1$ *such that*
$$\sum_j c_j \exp\left(\gamma \max_{f^{-n}P_j} \mathcal{S}_n(B)\right) < \theta.$$

Proof. Let

$$\phi(\gamma)=\sum_j c_j \exp\left(\gamma \max_{f^{-n}P_j} \mathcal{S}_n(B)\right).$$

Then $\phi(0)=1$, $\phi'(0)\le\alpha$. □

Corollary 12. *$\forall m>0$ there is a decomposition $f^{nm}P=\bigcup_j P_j$ such that*

$$\sum_j c_j \exp\left(\gamma \max_{f^{-nm}P-j} \mathcal{S}_{nm}(B)\right)\le\theta^m.$$

Proof (By induction). Decompose $f^nP=\bigcup_j Q_j$ and let $f^{n(m-1)}Q_j=\bigcup_k P_{jk}$ be a decomposition such that

$$\sum_k c_{jk} \exp\left(\gamma \max_{f^{-n(m-1)}P_{jk}} \mathcal{S}_{n(m-1)}(B)\right)\le\theta^{m-1}.$$

We have

$$\max_{f^{-nm}P_{jk}} \mathcal{S}_{nm}(B)\le \max_{f^{-n}Q_j}\mathcal{S}_n(B)+\max_{f^{-n(m-1)}P_{jk}}\mathcal{S}_{n(m-1)}(B).$$

Therefore

$$\begin{aligned}&\sum_{jk} c_jc_{jk}\exp\left(\gamma\max_{f^{nm}P_{jk}}\mathcal{S}_{nm}(B)\right)\\ &\le\sum_j c_j\exp\left(\gamma\max_{f^{-n}Q_j}\mathcal{S}_n(B)\right)\sum_k c_{jk}\exp\left(\gamma\max_{f^{-n(m-1)}P_{jk}}\mathcal{S}_{n(m-1)}(B)\right)\\ &\le\theta^{m-1}\sum_j c_j\exp\left(\gamma\max_{f^{-n}Q_j}\mathcal{S}_n(B)\right)\le\theta^m.\end{aligned}$$

□

Combining this with (7) and using $|\mathcal{S}_N(B)-\mathcal{S}_{N+k}(B)|\le Kk$, we get

Corollary 13. *$\exists C_1,\gamma,\rho_1<1$ such that $\forall\ell\in E$*

$$\ell(\exp(\gamma(\mathcal{S}_N(A)-\frac{N\varepsilon}{2})))\le C_1\rho_1^N.$$

Proof of Theorem 1. By the above corollary, $\forall\ell\in E(\mathcal{P},0,0)$

$$\ell(\mathcal{S}_N(A)\ge\frac{\varepsilon N}{2})\le C_1\rho_1^N.$$

Using the same argument for bounding $\mathcal{S}_N(A)$ from below, we get $\forall\ell\in E(\mathcal{P},0,0)$

$$\ell(|\mathcal{S}_N(A)|\ge\frac{\varepsilon N}{2})\le C_2\rho_2^N.$$

Now, given $\mathcal{P}',R,\alpha$, consider $\ell\in E_1(\mathcal{P},R,\alpha)$, say $\ell=\ell(Q,G)$. Decompose $N=N_1+N_2$, where $N_1=\delta N$, $N_2=(1-\delta)N$. Then

$$\ell(|\mathcal{S}_N(A)|\ge\varepsilon N)\le\ell\left(|\mathcal{S}_{N_2}(A)\circ f^{N_1}|\ge\frac{\varepsilon N}{2}\right)+\ell\left(|\mathcal{S}_{N_1}(A)|\ge\frac{\varepsilon N}{2}\right).$$

The second term is void if δ is small enough. Consider an almost Markov decomposition $f^{N_1}Q = (\bigcup_j P_j) \cup Z$ with respect to $\mathcal{P}$. Then

$$\ell\left(|\mathcal{S}_{N_2}(A)\circ f^{N_1}| \geq \frac{\varepsilon N}{2}\right)$$
$$\leq \text{Const}\left(c + \sum_j c_j \ell_j\left(|\mathcal{S}_{N_2}(A)| \geq \frac{\varepsilon N}{2}\right)\right)$$
$$\leq \text{Const}\, C_2 \rho_2^N.$$

(Here $\ell_j = \ell(P_j)$.) □

Notes. (1) Many results in smooth ergodic theory have partially hyperbolic versions. For example, Corollary 2 corresponds to the statement that a homeomorphism $h : F \to F$ of a compact F is uniquely ergodic if and only if $\frac{1}{n}\sum_{j=0}^{n} A(h^j x) \to \nu(A)$ for all x. However, for partially hyperbolic systems convergence does not hold for all x. The papers [20, 53] produce many non-negative C^∞ functions for which $A_n \equiv 0$ on a set of large Hausdorff dimension.

(2) For Anosov diffeomorphisms one can get quite precise asymptotics for $\ln \ell(|\mathcal{S}_n| > \varepsilon n)$. See [50, 51]. It is unlikely that the similar results could be obtained under our assumptions, because the asymptotics involve integrals of A with respect to Gibbs states other than SRB measure, and here we only assume good behavior with respect to SRB measures. On the other hand, the asymptotics for moderate deviations (see [52]) involve only integrals with respect to the SRB measure itself, and so it is likely to be generalizable to the settings of u-transitive systems. We do not pursue this topic here, however.

(3) In case $f \notin$ UuEe we can obtain the following generalization of Corollary 1.

Proposition 11. *$\forall A \in C(M)$, u-almost surely,*

$$\liminf \frac{\mathcal{S}_n(A)}{n}, \limsup \frac{\mathcal{S}_n(A)}{n} \in [\inf(\mu(A)), \sup(\mu(A))],$$

where the infimum and the supremum are taken over the set of u-Gibbs measures.

The proof is a verbatim repetition of the proof of Corollary 1.

8. Multiple mixing

Proof of Theorem 2. We argue by induction on k. We can assume that $||A_j|| \leq 1$.

(I) $k = 1$. It is enough to consider the case $\ell = \ell(P, G) \in E_1$. We have

$$I = \int_P e^{G(x)} \rho_P(x) A(f^n x) dx$$
$$= \int_{f^{\frac{n}{2}}P} e^{G(f^{-\frac{n}{2}}y)} \rho_{f^{\frac{n}{2}}P}(y) A(f^{\frac{n}{2}}y) dy.$$

Let $f^{\frac{n}{2}}P = (\bigcup P_j) \cup Z$ be an almost Markov decomposition. Choose $y_j \in P_j$; then

$$\begin{aligned} I &= \sum_j c_j \int_{P_j} \rho_{P_j}(y) e^{G(f^{-\frac{n}{2}}y)} A(f^{\frac{n}{2}}y)dy + O(\theta^n) \\ &= \sum_j c_j e^{G(f^{-\frac{n}{2}}y_j)} \int_{P_j} \rho_{P_j}(y) A(f^{\frac{n}{2}}y)dy + O(\theta^n) \\ &= \sum_j c_j e^{G(f^{-\frac{n}{2}}y_j)} \nu(A) + O\left(\theta^n + a\left(\frac{n}{2}\right)\right). \end{aligned}$$

Finally,

$$\sum_j c_j e^{G(f^{-\frac{n}{2}}y_j)} = \ell(1) + O(\theta^n) = 1 + O(\theta^n);$$

(II) From k to $k+1$. Denote $N = \frac{n_1+n_2}{2}$. Again consider an almost Markov decomposition $f^N P = (\bigcup P_j) \cup Z$. Similarly to (I),

$$\int_P e^{G(x)} \rho_P(x) \prod_{j=1}^{k+1} A(f^{n_j}x)dx$$

$$= \sum_j c_j e^{G(f^{-N}y_j)} A_1(f^{-(N-n_1)}y_j) \int_{P_j} \rho_{P_j}(y) \prod_{j=2}^{k+1} A(f^{n_j-N}y)dy + O(\theta^m).$$

The first term is

$$\begin{aligned} &\sum_j c_j e^{G(f^{-N}y_j)} A_1(f^{-(N-n_1)}y_j) \\ &= \int_P e^{G(x)} \rho_P(x) A_1(f^{n_1}x)dx + O(\theta^m) \\ &= \nu(A_1) + O(\theta^m), \end{aligned}$$

and the second one equals

$$\prod_{j=2}^{k+2} \nu(A_j) + O\left(a\left(\frac{m}{C_2(k)}\right) + \theta^m\right)$$

by induction. □

9. Moment estimates

Starting from this section, we suppose that the $a(m)$ satisfy

$$\sum_m a(m) < +\infty.$$

Let $A_j \in \mathbb{B}$ be a sequence of functions such that $||A_j||_{\mathbb{B}} \le K$, $\nu(A_j) = 0$. Let $S_n = \sum_{j=0}^{n-1} A_j(f^j x)$.

Lemma 1.

(a) $|\ell(S_n) \le \text{Const}$;
(b) $\ell(S_n^2) \le \text{Const}\, n$;
(c) $|\ell(S_n^3)| \le \text{Const}\, n^{\frac{3}{2}}$;
(d) $\ell(S_n^4) \le \text{Const}\, n^2$,

where the constants in (a)–(d) *depend only on* K *but not on sequence* A_j.

(e) *Let $A(t,x)$ be a function defined on $[0,\mathbf{T}]\times M$ such that for all $t\in[0,\mathbf{T}]$ we have $A(t,\cdot)\in\mathbb{B}$, $||A(t,\cdot)||_{\mathbb{B}}\leq K$ and $\int A(t,x)d\mu(x)=0$. Let*

$$S_\varepsilon(t)=\sum_{j=0}^{[\frac{t}{\varepsilon}]}A(\varepsilon j, f^j x). \tag{8}$$

Then, as $\varepsilon\to 0$,

$$\varepsilon\ell(S_\varepsilon(t)^2)\to\int_0^t\sigma(A(s,\cdot))ds,$$

where

$$\sigma(A)=\sum_{j=-\infty}^{\infty}\nu(A(A\circ f^j)).$$

Proof. (a) We have

$$|\ell(S_n)|=|\sum_{j=0}^{n-1}\ell(A_j(f^jx))|\leq \mathrm{Const}\sum_j a(j)\leq\mathrm{Const}\,.$$

(b) We have

$$\ell(S_n^2)=\sum_{j,k}\ell(A_j(f^jx)A_k(f^kx))\leq\mathrm{Const}\sum_{j,k}a\left(\frac{|j-k|}{C}\right).$$

Now for fixed m there are less than $2n$ pairs (j,k) with $|j-k|=m$. So

$$\ell(S_n^2)\leq\mathrm{Const}\,n\sum_m a\left(\frac{m}{C}\right)\leq\mathrm{Const}\,.$$

(e) Fix some large M. We have

$$\begin{aligned}\ell(S_\varepsilon(t)^2)&=\sum_{j,k=0}^{n-1}\ell(A(\varepsilon j,f^jx)A(\varepsilon k,f^kx)\\&=\sum_{|j-k|<M}\ell(A(\varepsilon j,f^jx)A(\varepsilon k,f^kx)\\&=\sum_{|j-k|\geq M}\ell(A(\varepsilon j,f^jx)A(\varepsilon k,f^kx)=I+I\!I.\end{aligned}$$

By the argument of (b), $|\varepsilon I_\varepsilon|\leq\mathrm{Const}\sum_{m>M}a(m)\to 0$ as $M\to\infty$. On the other hand, for fixed M the following holds. Let $\varepsilon j\to s$; then

$$\begin{aligned}\sum_{|k-j|<M}\ell(A(\varepsilon j,f^jx)A(\varepsilon k,f^kx))&\to\sum_{|q|<M}\nu(A(s,x)A(s,f^qx))\\&=\sigma(A(s,\cdot))+o_{M\to\infty}(1).\end{aligned}$$

Thus

$$\varepsilon\ell(S_\varepsilon^2(t))\to\int_0^t\sigma(A(s,\cdot))ds+o(1).$$

Letting $M\to\infty$, we obtain (e).

(c) follows from (b) and (d), so it suffices to establish (d). We have

$$\ell(S_n^4)=\sum_{j_1,j_2,j_3,j_4}\ell((A_{j_1}(f^{j_1}x)A_{j_2}(f^{j_2}x)A_{j_3}(f^{j_3}x)A_{j_4}(f^{j_4}x)).$$

First, let us estimate the terms where not all indices j_p are different. The sum over terms with at most two different indices is bounded by Const $\times$(the number of terms), hence by Const n^2. Also,

$$J = \sum \ell(A_{j_1}(f^{j_1}x)A_{j_2}(f^{j_2}x)A^2_{j_3}(f^{j_3}x)) \le \text{Const} \sum a\left(\frac{\min j_p - j_{p-1}}{C}\right).$$

For fixed m, the number of terms with $\min(n_j - n_{j-1}) = m$ equals Const n^2. Thus

$$J \le \text{Const}\, n^2 \sum_m a(m).$$

Now, up to the terms of order n^2,

$$\begin{aligned}\ell(S_n^4) &= 12 \sum_{j_3} \sum_{j_1,j_2=1}^{j_3} \sum_{j_4=j_3}^{n} \ell(A_{j_1}(f^{j_1}x)A_{j_2}(f^{j_2}x)A_{j_3}(f^{j_3}x)A_{j_4}(f^{j_4}x)) + O(n^2)\\ &= 12 \sum_{j_3} \sum_{j_4=j_3}^{n} \ell(S^2_{j_3}A_{j_3}(f^{j_3}x)A_{j_4}(f^{j_4}x)) + O(n^2).\end{aligned}$$

Proposition 12. $\forall l\ \forall j_3$

$$\ell\left(\sum_{j_4=j_3}^{n} S^2_{j_3}A_{j_3}(f^{j_3}x)A_{j_4}(f^{j_4}x)\right) \le \text{Const}\, j_3.$$

Proof. Again it suffices to verify this for $l \in E_1$, say $\ell = \ell(P,G)$. Consider an almost Markov decomposition $f^{j_3}P = (\bigcup_q P_q) \cup Z$. Choose $y_q \in P_q$; then

$$\begin{aligned}&\int_P e^{G(x)}\rho_P(x)S^2_{n_3}(x)A_{j_3}(f_{j_3}x)A_{j_4}(f_{n_4}x)dx\\ &= O(\theta^{j_3}) + \sum_q c_q S^2_{n_3}(y_q) \sum_{j_4=j_3}^{n} \int_P e^{G(f^{-j_3}y)}\rho_{P_q}(y)A_{j_3}(y)A_{j_4}(f^{j_4-j_3}y)dy\\ &+ \sum_q c_q \sum_{j_4=j_3}^{n} \int_P e^{G(f^{-j_3}y)}\rho_{P_q}(y)[S^2_{n_3}(f^{-j_3}y) - S^2_{n_3}(y_q)]A_{j_3}(y)A_{j_4}(f^{j_4-j_3}y)dy\\ &= I + I\!I.\end{aligned}$$

By Theorem 2, $I \le \text{Const} \sum_q c_q S^2_{j_3}(y_q)$. Now $\text{Osc}_{f^{-j_3}P_q} S^2_{j_3} \le \text{Const}\, j_3$, so

$$\sum_q c_q S^2_{j_3}(y_q) \le \text{Const}\, j_3 + \ell(S^2_{j_3}) \le \text{Const}\, j_3.$$

Moreover,

$$\begin{aligned}I\!I &= \sum_q c_q \sum_{j_4=j_3}^{n} \int_P e^{G(f^{-j_3}y)}\rho_{P_q}(y)[S_{j_3}(f^{-j_3}y) - S_{j_3}(y_q)][S_{j_3}(f^{-j_3}y) + S_{j_3}(y_q)]\\ &\qquad \times A_{j_3}(y)A_{j_4}(f^{j_4-j_3}y)dy\\ &= \sum_q c_q \sum_{k=0}^{j_3-1} \sum_{j_4=j_3}^{n} \int_P \Big\{e^{G(f^{-j_3}y)}\rho_{P_q}(y)[S_{j_3}(f^{-j_3}y) - S_{j_3}(y_q)]\\ &\qquad \times[A_k(f^{k-j_3}y) + A_k(f^k y_q)]\Big\} A_{j_3}(y)A_{j_4}(f^{j_4-j_3}y)dy.\end{aligned}$$

The part in brackets is uniformly bounded and uniformly Hölder continuous. Thus by Theorem 2 the sum over j_4 is uniformly bounded for any q, k. Hence

$$I\!I \le \text{Const} \sum_q c_q \sum_k 1 = \text{Const}\, j_3 \sum_q c_q \le \text{Const}\, j_3.$$

□

Now

$$\ell(S_n^4) \le \text{Const} \sum_{j<n} j + O(n^2) = O(n^2).$$

This concludes the proof of Lemma 1. □

10. Tightness

In all theorems of Section 4 it suffices by the definition of weak convergence in $C[0,\infty[$ to show that for each $\mathbf{T} > 0$ the corresponding processes converge in $C[0,\mathbf{T}]$. So let $\mathbf{T}$ be fixed from now on, until the end of Section 15.

Lemma 2. *Let $S_\varepsilon(t)$ be defined by* (8). *Then the family $\{\sqrt{\varepsilon} S_\varepsilon(t)\}$ is tight.*

Proof. Let $Y(N) = \{X(t) : \forall m > N\ \forall k\ |X(\frac{k+1}{2^m}) - X(\frac{k}{2^m})| < \frac{1}{2^{\frac{m}{8}}}\}$. Then $Y(N)$ is compact in $C[0,\mathbf{T}]$ for all N. Let us estimate $\ell(\sqrt{\varepsilon} S_\varepsilon(t) \notin Y(N))$. We have

$$\ell\left(\left[\sqrt{\varepsilon}\left|S_\varepsilon\left(\frac{k+1}{2^m}\right) - S_\varepsilon\left(\frac{k}{2^m}\right)\right|\right]^4\right) \le C\varepsilon^2 \left(\frac{1}{2^m\varepsilon}\right)^2 = C2^{-2m}.$$

So, for given k,

$$\ell\left(\left[\sqrt{\varepsilon}\left|S_\varepsilon\left(\frac{k+1}{2^m}\right) - S_\varepsilon\left(\frac{k}{2^m}\right)\right|\right] < \frac{1}{2^{\frac{m}{8}}}\right) \le C2^{-2m}(2^{\frac{m}{8}})^4 = C2^{-\frac{3m}{2}}.$$

Hence

$$\ell\left(\exists k \quad \left[\sqrt{\varepsilon}\left|S_\varepsilon\left(\frac{k+1}{2^m}\right) - S_\varepsilon\left(\frac{k}{2^m}\right)\right|\right] < \frac{1}{2^{\frac{m}{8}}}\right)$$
$$\le \text{Const}\, \mathbf{T} 2^m 2^{-\frac{3m}{2}} = \text{Const}\, \mathbf{T} 2^{-\frac{m}{2}}.$$

Thus $\ell(\sqrt{\varepsilon} S_\varepsilon(t) \notin Y(N)) \le \text{Const}\, 2^{-\frac{N}{2}}$. □

The next statement is used in Section 11. Take some α between 1 and 2. Denote $n_k = \sum_{j=1}^k j^\alpha$, $\eta_k = \mathcal{S}_{n_k}(A)$. Choose θ such that $\frac{1}{6} + \frac{1}{6\alpha} < \theta < \frac{1}{2\alpha}$.

Lemma 3. *Almost surely*

$$\max_{j\le k} \max_{n_{j-1}\le l\le n_j} |\mathcal{S}_l(A) - \eta_{n_{j-1}}| \le Ck^{\alpha(\frac{1}{2}+\theta)}.$$

Proof. Let $[l_1, l_2]$ be an interval of the form

$$l_1 = n_{j-1} + \frac{pj^\alpha}{2^m}, \quad l_2 = n_{j-1} + \frac{(p+1)j^\alpha}{2^m}.$$

We claim that almost surely

$$|\mathcal{S}_{l_2} - \mathcal{S}_{l_1}| \ge \sqrt{l_2 - l_1}\, j^{\alpha\theta} \tag{9}$$

only finitely many times. Indeed the probability of such an event is less than

$$\frac{\mathbb{E}(|\mathcal{S}_{l_2} - \mathcal{S}_{l_1}|^4)}{j^{4\alpha\theta}(l_2 - l_1)^2} \le \frac{C}{j^{4\alpha\theta}}.$$

(9) can happen only if $l_2 - l_1 \geq \sqrt{l_2 - l_1} j^{\alpha\theta}$; that is, $l_2 - l_1 \geq j^{2\alpha\theta}$. Thus for fixed j we have $O(j^{\alpha(1-2\theta)})$ events, and so

$$\text{Prob}(\exists l_1, l_2 \quad \text{satisfying (9) with given } j) = O(j^{\alpha(1-6\theta)}).$$

By assumption $\alpha(1 - 6\theta) < -1$. This completes the proof. □

11. Invariance principle

Proof of Theorem 3. We keep the notation of the previous section. Let us begin by recalling the facts about martingales we will use in this and the following sections. Proofs can be found for example in [38]. Let $(Z_n, \mathcal{G}_n)$ be a martingale pair. Then $Y_n = Z_n - Z_{n-1}$ is called a *martingale difference sequence*. We consider only martingales satisfying $Z_0 = 0$ and $\mathbb{E}(Z_n^2) < \infty$.

Proposition 13. (a) (Doob convergence theorem) *If $\mathbb{E}(Z_n^2)$ is bounded then Z_n converges almost surely.*

There are constants C_1 and C_2 such that for any martingale $(Z_n, \mathcal{G}_n)$ as above the following holds:

(b) *Let $Z^* = \max_n Z_n$, $\Delta Z = \sum_n Y_n^2$. Then*

$$\frac{1}{C_1}\mathbb{E}((\Delta Z)^2) \leq \mathbb{E}(Z^{*4}) \leq C_1 \mathbb{E}((\Delta Z)^2).$$

(c) (Skorohod representation theorem) *After possibly enlarging the probability space, we can find a Brownian motion $w(t)$ and stopping times T_j such that if $\tau_k = \sum_{j=1}^k T_j$, then $Z_k = w(\tau_k)$, $\mathbb{E}(T_k|\mathcal{F}_k) = \mathbb{E}(Y_k^2)$ and $\mathbb{E}(T_k^2) \leq C_2 \mathbb{E}(Y_k^4)$.*

(d) *Let η_n be a $\mathcal{G}_n$-measurable sequence such that*

$$\beta_n = \sum_{j=1}^{\infty} \mathbb{E}(\eta_{n+1-j}|\mathcal{F}_{n-1}) \leq \text{Const}.$$

Then

$$\eta_n = Y_n + \beta_{n+1} - \beta_n, \tag{10}$$

where Y_n is a martingale difference sequence.

Let $\ell = \ell(P)$. First we define an increasing sequence of sigma–algebras $\mathcal{F}_n$ on P. Let $\mathcal{F}_0 = \{\emptyset, P\}$. Suppose that $\mathcal{F}_n$ is generated by $\{P_{j,n}\}$ such that $f^n P_{j,n} \in \mathcal{P}$. Decompose $f^{n+1} P_{j,n} = \bigcup_k P_{jk,n}$ and let $\mathcal{F}_{n+1}$ be generated by $f^{-n-1} P_{jk,n}$. Write $\mathcal{G}_k = \mathcal{F}_{n_k}$, $\tilde{\eta}_k = \mathbb{E}(\eta_k|\mathcal{G}_k)$. Note that $|\eta_k - \tilde{\eta}_k| \leq \text{Const}$.

Lemma 4. $\exists C$ *such that* $\forall j$ $\sum_k |\mathbb{E}(\tilde{\eta}_{j+k}|\mathcal{G})| \leq C$.

Proof. Let Q be an element of $\mathcal{G}_j$. Then

$$\mathbb{E}(\tilde{\eta}_{j+k}|\mathcal{G}_j) = \mathbb{E}(\eta_{j+k}|\mathcal{G}_j) = \int_Q \rho_Q(y) \sum_{l=n_{j+k-1}+1}^{n_{j+k}} A(f^{l-n_j} y) dy.$$

Thus $\sum_k |\mathbb{E}(\tilde{\eta}_{j+k}|\mathcal{G})| \leq \sum_{l=1}^{\infty} a(l)$. □

Write $\tilde{\eta}_k = \zeta_k + \beta_k - \beta_{k+1}$, where $\beta_k = \sum_{l=0}^{\infty} \mathbb{E}(\tilde{\eta}_{k+l}|\mathcal{G}_{k-1})$. Let $S_k = \sum_{l=1}^k \zeta_l$. Then $(S_k, \mathcal{G}_k)$ is a martingale, and $|S_k - \mathcal{S}_{n_k}| \leq \text{Const}\, k$. Given N, define M_N by the condition that $n_{M_N} \leq N < n_{M_N+1}$.

Proposition 14. *$\exists s_1$ such that almost surely*

$$\mathcal{S}_N - S_{M_N} = O(N^{\frac{1}{2}-s_1}).$$

Proof. We have $\mathcal{S}_N - S_{M_N} = (\mathcal{S}_N - \mathcal{S}_{M_N}) + (\mathcal{S}_{M_N} - S_{M_N}) = I + II$. For I we have

$$I = O(M^{(\frac{1}{2}+\theta)\alpha}) = O(N^{(\frac{1}{2}+\theta)\frac{\alpha}{\alpha+1}})$$

by Lemma 3, and $(\frac{1}{2}+\theta)\frac{\alpha}{\alpha+1}) \leq \frac{1}{2}$ since $\theta < \frac{1}{2\alpha}$. On the other hand, $II = O(M) = O(N^{\frac{1}{1+\alpha}})$, and $\frac{1}{\alpha+1} < \frac{1}{2}$ as $\alpha > 1$. □

Let w, T_j and τ_k be as in Proposition 13(c).

Proposition 15. *$\exists \sigma_N$ such that $\frac{\sigma_N}{N} \to \sigma(A)$ and $\sum_{j=1}^{M_N} T_j - \sigma_N = O(N^{1-s_2})$.*

Proof. We have

$$\begin{aligned}\sum_{j=1}^{M_N} T_j &= \sum_{j=1}^{M_N}[T_j - \mathbb{E}(T_j|\mathcal{G}_{j-1})] + \sum_{j=1}^{M_N}[\mathbb{E}(\zeta_j^2|\mathcal{G}_{j-1}) - \zeta_j^2] \\ &\quad + \sum_{j=1}^{M^N}[\zeta_j^2 - \mathbb{E}(\zeta_j^2)] + \sum_{j=1}^{M^N}\mathbb{E}(\zeta_j^2) \\ &= I + II + III + IV.\end{aligned}$$

To estimate I, write $R_j = T_j - \mathbb{E}(T_j|\mathcal{G}_{j-1})$, $\mathbb{E}(T_j|\mathcal{G}_{j-1}) = Dj^\alpha + r_j$, where r_j is uniformly bounded. Thus

$$\mathbb{E}(R_j^2) = \mathbb{E}(T_j^2) - 2\mathbb{E}(T_j r_j) + D^2 j^{2\alpha} \leq Cj^{2\alpha}.$$

Since R_j is a martingale difference sequence, $\sum_j \frac{R_j}{j^{\alpha+\frac{1}{2}+\epsilon}}$ converges almost surely by Proposition 13(b). Writing

$$R_j = \left(\frac{R_j}{j^{\alpha+\frac{1}{2}+\varepsilon}}\right) j^{\alpha+\frac{1}{2}+\varepsilon}$$

and summing by parts, we obtain

$$\sum_{j=1}^{M} R_j \leq \text{Const}(\omega) M^{\alpha+\frac{1}{2}+\epsilon} = O(n_M^{1-s_3}).$$

II can be bounded the same way as I. Namely let $L_j = \zeta_j^2 - \mathbb{E}(\zeta_j^2|\mathcal{G}_{j-1})$; then

$$\mathbb{E}(L_j^2) = \mathbb{E}([\zeta_j^2 - \mathbb{E}(\zeta_j^2|\mathcal{G}_{j-1})]^2) \leq \mathbb{E}([A_{j^\alpha} \circ f^{n_j}]^2) + O(j^{2\alpha}) = O(j^{2\alpha}),$$

so as before $II = \sum_{j=1}^{M} L_j = O(n^{1-s_3})$. Also, similarly to Lemma 1,

$$\mathbb{E}\left(\left[\sum_{j=1}^{M}\tilde{\zeta}^2 - \mathbb{E}(\tilde{\zeta}^2)\right]^2\right) = \mathbb{E}\left(\left[\sum_{j=1}^{M}\zeta^2 - \mathbb{E}(\zeta^2)\right]^2\right) \leq \text{Const}\, n_M^2,$$

so by Borel-Cantelli $III = O(n_M^{\frac{7}{8}})$ almost surely. Therefore, $\sum_{j=1}^{M_N} T_j = \sum_{j=1}^{M_N}\mathbb{E}(\zeta_j^2) + O(n^{1-s_3})$. By Section 9, $\sum_{j=1}^{M}\mathbb{E}(\zeta_j^2) \sim \sum_{j=1}^{M}\sigma(A)j^\alpha = \sigma(A)n_M$. □

Thus we have

$$S_k = w(\tau_k) = w(\sigma_{n_k}) + [w(\tau_k) - w(\sigma_{n_k})] = w(\sigma n_k) + O(n_k^{\frac{1-s_4}{2}})$$

almost surely.

This identity together with Proposition 14 proves Theorem 3. □

Note. Our exposition mostly follows [71].

12. Convergence to the Gaussian process

Theorem 9. *Let $S_\varepsilon(t)$ be defined as in* (8)*. Then as $\varepsilon \to 0$ the process $\sqrt{\varepsilon}S_\varepsilon(t)$ converges weakly to a Gaussian random process $\mathbf{S}(t)$ with zero mean and covariance matrix*

$$\langle \mathbf{S}(t), \mathbf{S}(t) \rangle = \int_0^t \sigma(A(s, \cdot))ds.$$

Remark. Clearly this theorem implies Corollary 4.

Proof. By Lemma 2 $\{S_\varepsilon(t)\}$ is a tight family, so we need only to verify convergence of finite dimensional distributions. Let us start with one-dimensional distributions. Denote $n = \frac{1}{\varepsilon}$. Define

$$\hat{S}_k = \sum_{j=(k-1)n^{\frac{3}{5}}}^{kn^{\frac{3}{5}} - n^{\frac{1}{10}}} A(\varepsilon j, f^j x),$$

$$\bar{S}_k = \sum_{j=kn^{\frac{3}{5}} - n^{\frac{1}{10}}}^{kn^{\frac{3}{5}} - 1} A(\varepsilon j, f^j x),$$

$$S^*(t) = \sum_{k=0}^{\left[\frac{t}{n^{\frac{3}{5}}}\right] - 1} \hat{S}_k,$$

$$S^{**}(t) = \sum_{k=0}^{\left[\frac{t}{n^{\frac{3}{5}}}\right] - 1} \bar{S}_k.$$

Then by Lemma 1 $S^{**}(t) \to 0$ in $L^2(l)$ and, in particular, $S^{**}(t) \to 0$ in probability. Let $\psi_k(\xi) = \ell(e^{i\sqrt{\varepsilon}\hat{S}_k\xi})$.

Proposition 16.

$$\psi_k(\xi) = 1 - \varepsilon^{\frac{2}{5}}\sigma(A(k\varepsilon^{\frac{2}{5}}, \cdot))(1 + o(1)).$$

Proof. We have

$$\psi_k(\xi) = \mathbb{E}_l\left(1 + i\sqrt{\varepsilon}\hat{S}_k\xi - \frac{\varepsilon\hat{S}_k^2}{2}\xi^2 - i\varepsilon^{\frac{3}{2}}\frac{\hat{S}_k^3}{6}\xi^3 + O\left(\varepsilon^2 h S_k^2\xi^4\right)\right).$$

Using Lemma 1, we get

$$\psi_k(\xi) = 1 - \varepsilon^{\frac{2}{5}}\sigma(A(s, \cdot))(1 + o(1)) + O(\varepsilon^{\frac{1}{2}} + \varepsilon^{\frac{3}{5}} + \varepsilon^{\frac{4}{5}}),$$

where the main term comes from $\varepsilon\frac{\hat{S}_k^2}{2}\xi^2$. This proves the proposition. □

Let $\phi_k(\xi) = \ell(e^{i\sqrt{\varepsilon}S_k^*\xi})$.

Proposition 17.

$$\ln \phi_{k+1}(\xi) = \ln \phi_k(\xi) - \varepsilon^{\frac{2}{5}}\sigma\left(A(k\varepsilon^{\frac{2}{5}}, \cdot)\right)\frac{\xi^2}{2} + o\left(\varepsilon^{\frac{2}{5}}\right). \tag{11}$$

Proof. It suffices to verify this for $\ell \in E_1$.

(I) The case $k = 0$ constitutes Proposition 16.

(II) $k > 0$. Decompose $f^{kn^{\frac{3}{5}}}P = (\bigcup_j P_j) \cup Z$. Let $q = kn^{\frac{3}{5}}$. Choose $y_j \in P_j$. Then

$$\ell\left(\exp(i\sqrt{\varepsilon}S_{k+1}^*\xi)\right)$$
$$= \sum_j c_j \exp(i\sqrt{\varepsilon}S_k^*(f^{-q}y_j)\xi)\exp(G(f^{-q}y_j))\int_{P_j} e^{i\sqrt{\varepsilon}S_1^*(y)\xi}\rho_{P_j}(y)dy + O(\theta^{n^{\frac{1}{10}}}).$$

By Proposition 16

$$\int_{P_j} e^{i\sqrt{\varepsilon}S_1^*(y)\xi}\rho_{P_j}(y)dy = (1 - \varepsilon^{\frac{2}{5}}\sigma(A(k\varepsilon^{\frac{2}{5}}, \cdot))(1 + o(1)).$$

Hence

$$\phi_{k+1}(\xi) = \sum_j c_j \exp(i\sqrt{\varepsilon}S_k^*(f^{-q}y_j))\exp(G(f^{-q}y_j)))(1 - \varepsilon^{\frac{2}{5}}\sigma(A(k\varepsilon^{\frac{2}{5}}, \cdot))(1 + o(1)))$$
$$= \phi_k(\xi)(1 - \varepsilon^{\frac{2}{5}}\sigma(A(k\varepsilon^{\frac{2}{5}}, \cdot))(1 + o(1))) + O(\theta^{-n^{\frac{1}{10}}}).$$

Taking logarithms of both sides, we obtain the required statement. □

Now, summing (11) for $k = 0, \ldots, [tn^{\frac{2}{5}}]$, we get

$$\ln \ell(e^{i\sqrt{\varepsilon}S^*(t)\xi}) \sim -\frac{\xi^2}{2}\int_0^t \sigma(A(s, \cdot))ds.$$

Since $\sqrt{\varepsilon}[S_\varepsilon(t) - S_\varepsilon^*(t)] \to 0$ in probability, we see that one-dimensional distributions of $\sqrt{\varepsilon}S_\varepsilon(t)$ converge to those of $\mathbf{S}(t)$. To consider the general case, let $t_1, \ldots, t_r$, $\xi_1, \ldots, \xi_r$ be numbers. Denote $\eta_j = \sum_{m=1}^j \xi_m$. We have

$$\sum_j \xi_j S_\varepsilon(t_j) = \sum_j \eta_j [S_\varepsilon(t_j) - S_\varepsilon(t_{j-1})].$$

By the same argument as in the proof of Proposition 11, we obtain

$$\ln \ell\left(\exp[i\sqrt{\varepsilon}\sum_j \xi_j S_\varepsilon(t_j)]\right) \sim -\frac{1}{2}\sum_j \eta_j^2 \int_{t_{j-1}}^{t_j} \sigma(A(s, \cdot))ds.$$

This implies convergence of multidimensional distributions, and so proves Theorem 9. □

Note. By the same argument one can obtain versions of the central limit theorem for families of diffeomorphisms. One only has to check the uniformity of the estimates of the previous sections. The following statement is used in [28].

Proposition 18. *Let f_ε be a family of partially hyperbolic systems such that $\exists C, r, v,$ a function space $\mathbb{B}$, and a sequence $\{a(n)\}$ such that*

$$\sum_{n=1}^{\infty} a(n) < \infty,$$

and a linear functional $\omega : \mathbb{B} \to \mathbb{R}$ such that for any P_ε belonging to the (C, r_1, v)-universal family from example I of Section 2 and $\forall \rho$ such that $||\rho||_{C^\gamma(P_\varepsilon)} \leq 1$ the following estimate holds:

$$\left| \int_{P_\varepsilon} A(f_\varepsilon^n x)\rho(x)dx - \nu(A) - \varepsilon\omega(A) \right| \leq ||A||(a(n) + o(\varepsilon)).$$

Let n_ε be a sequence such that $n_\varepsilon \to \infty$ and $n_\varepsilon \varepsilon^2 \to c$, where $c \geq 0$. Then, if x is chosen according to Lebesgue measure,

$$\frac{\sum_{j=0}^{n-1} \left[A(f_\varepsilon^j x) - \nu(A) \right]}{\sqrt{n_\varepsilon}} \to \mathcal{N}(c\omega(A), D(A)).$$

13. Short time fluctuations in averaging. Moments of slowly changing quantities

To simplify the notation we present the proofs of Theorems 4 and 5 only for the case $d = 1$. The reader will have no difficulty in establishing multidimensional analogies of our results, but the notation in higher-dimensional settings becomes much more complicated.

Here we prove Theorem 4. We have

$$\begin{aligned} \Delta_{n+1} - \Delta_n &= \varepsilon \left[A(z_n, f^n x) - \bar{A}(q_n) \right] \\ &= \varepsilon \left[A(q_n, f^n x) - \bar{A}(q_n) \right] + \varepsilon \left[A(z_n, f^n x) - \bar{A}(q_n, f^n x) \right]. \end{aligned}$$

Using the Hadamard lemma, we rewrite the second term as

$$A(z_n, f^n x) - \bar{A}(q_n, f^n x) = \left[DA(q_n, f^n x) + \zeta(q_n, f^n x, \Delta_n) \right] \Delta_n,$$

where ζ is a smooth function of its arguments, $\zeta(q, x, 0) = 0$. Denote

$$\begin{aligned} \mathcal{Q}_n &= D\bar{A}(q_n) + \bar{\zeta}(q_n, \Delta_n), \\ \beta_n &= \left[DA(q_n, f^n x) + \zeta(q_n, f^n x, \Delta_n) - \mathcal{Q}_n \right] \Delta_n, \\ \gamma_n &= A(q_n, f^n x) - \bar{A}(q_n). \end{aligned}$$

Then our equation can be rewritten as

$$\Delta_{n+1} - \Delta_n = \varepsilon \left[\mathcal{Q}_n \Delta_n + \beta_n + \gamma_n \right].$$

Let L_n be the solution of

$$L_{n+1} - L_n = \varepsilon \mathcal{Q}_n L_n. \tag{12}$$

Substitute $\Delta_n = L_n \rho_n$; then we have

$$L_{n+1}(\rho_{n+1} - \rho_n) = \varepsilon(\beta_n + \gamma_n), \tag{13}$$

so

$$\rho_{n+1} = \varepsilon \sum_{j=0}^{n} L_{j+1}^{-1}(\beta_j + \gamma_j). \tag{14}$$

The next is a special case of Theorem 9.

Proposition 19. *Let* $\gamma_t^\varepsilon = \frac{\gamma_{[\frac{t}{\varepsilon}]}}{\sqrt{\varepsilon}}$. *Then, as* $\varepsilon \to 0$, γ_t^ε *converges to* B, *which is the Gaussian process defined by* (6).

In order to estimate the moments of $\sum_n L_{n+1}^{-1}\beta_n$ and $\sum_n L_{n+1}^{-1}\gamma_n$ we need the following statement, the proof of which occupies the most of this section.

Proposition 20. *Let* $A(\delta, x)$ *satisfy* $\int A(\delta,x)d\nu(x) = 0$ *for all* δ. *Let* $\theta_p(\delta) = ||A^p(\delta,\cdot)||_{\mathbb{B}}$. *Suppose that* θ_p *are smooth functions of* δ. *Let* $\kappa_p(\delta) = \frac{d\theta_p}{d\delta}$, $\tilde{\kappa}_p(\delta) = ||\frac{d}{d\delta}A^p(\delta,\cdot)||_{\mathbb{B}}$. *Suppose that* $|\kappa_p(\delta)| < \text{Const}$, $|\tilde{\kappa}_p(\delta)| < \text{Const}$ *for* $p \le 4$. *Let* $\{\delta_n(x)\}$ *satisfy*

$$\delta_{n+1} - \delta_n = \varepsilon B(\delta_n, f^n x, \varepsilon), \tag{15}$$

where for all m $||\frac{d^m}{d\delta^m}B(\delta,\cdot)||_{\mathbb{B}}$ *is uniformly bounded and*

$$B(\delta, x, \varepsilon) = B(\delta, x) + O(\varepsilon).$$

Let $T = \sum_{j=m}^{m+\frac{1}{\sqrt{\varepsilon}}} A(\delta_j, f^j)$. *Then*

$$|\ell(T)| \le \text{Const}\left[\ell(\theta_1(\delta_m)) + \sqrt{\varepsilon}\right], \tag{a}$$

$$\left|\ell(T^2)\right| \le \text{Const}\left[\ell(\theta_2(\delta_m)) + \varepsilon\right]\frac{1}{\sqrt{\varepsilon}}, \tag{b}$$

$$\left|\ell(T^4)\right| \le \text{Const}\left[\frac{\ell(\theta_4(\delta_m))}{\varepsilon} + \ell(|\kappa_4(\delta_m)|) + \ell(|\tilde{\kappa}_1(\delta_m)\theta_3(\delta_m)|)\right. \tag{c}$$

$$\left. + \ell(|\theta_3(\delta_m)|) + \ell(|\tilde{\kappa}_1(\delta_m)\theta_2(\delta_m)|) + \varepsilon\ell(|\theta_2(\delta_m)|) + \varepsilon^{\frac{3}{2}}\ell(|\tilde{\kappa}_1(\delta_m)) + \varepsilon^2\right].$$

Proof. Let

$$T' = \sum_j A(\delta_m, f^j x),$$

$$T'' = \varepsilon \sum_{j>k} \frac{dA}{d\delta}(\delta_m, f^j x) B(\delta_m, f^k x).$$

Lemma 5.

$$T = T' + T'' + O(\sqrt{\varepsilon}).$$

Proof. We have

$$T = \sum_j A(\delta_j, f^j x) = \sum_j A(\delta_m, f^j x) + \sum_j [A(\delta_j, f^j x) - A(\delta_m, f^j x)].$$

The first term is equal to T'. The second term can be estimated as follows:

$$\sum_j [A(\delta_j, f^j x) - A(\delta_m, f^j x)] = \sum_j \frac{dA}{d\delta}(\delta_m, f^j x)(\delta_j - \delta_m) + \sum_j O\left((\delta_j - \delta_m)^2\right).$$

Now

$$\delta_m - \delta_j \le \text{Const}\,|m-j|\varepsilon \le \text{Const}\,\sqrt{\varepsilon}.$$

Hence

$$\sum_j (\delta_j - \delta_m)^2 \le \text{Const}\,\frac{1}{\sqrt{\varepsilon}}\varepsilon \le \text{Const}\,\sqrt{\varepsilon}.$$

Now

$$\delta_j - \delta_m = \sum_{k=m}^{j} \varepsilon B(\delta_k, f^k x) + \sum_{k=m}^{j} O(\varepsilon^2) = \sum_{k=m}^{j} \varepsilon B(\delta_k, f^k x) + O(\varepsilon^{\frac{3}{2}})$$

and

$$B(\delta_k, f^k x) = B(\delta_m, f^k x) + O(|\delta_k - \delta_m|) = B(\delta_m, f^k x) + O(\sqrt{\varepsilon}).$$

Hence

$$\sum_j \left[A(\delta_j, f^j x) - A(\delta_m, f^j x)\right] = \varepsilon \sum \sum_{k>j} \frac{dA}{d\delta}(\delta_j, f^j x) B(\delta_m, f^k x) + O(\sqrt{\varepsilon}),$$

as claimed. □

To estimate T'', rewrite

$$\ell\left(\sum_{j>k}\sum \frac{dA}{d\delta}(\delta_m, f^j x) B(\delta_m, f^k x)\right) = \ell\left(\sum_k B(\delta_m, f^k x) \sum_{j>k} \frac{dA}{d\delta}(\delta_m, f^j x)\right).$$

Now $\int \frac{dA}{d\delta}(\delta, x) d\mu(x) = 0$, so similarly to Lemma 1 we obtain that for any fixed k

$$\left|\ell\left(B(\delta_m, f^k x) \sum_{j>k} \frac{dA}{d\delta}(\delta_m, f^j x)\right)\right| \leq \text{Const}.$$

Hence

$$\varepsilon\left|\ell\left(\sum_{j>k}\sum \frac{dA}{d\delta}(\delta_m, f^j x) B(\delta_m, f^k x)\right)\right| \leq \varepsilon \times \text{Const}\,\frac{1}{\sqrt{\varepsilon}} = \text{Const}\,\sqrt{\varepsilon}.$$

To estimate $\ell(\sum_j A(\delta_m, f^j x))$, it is enough to treat the case $\ell = \ell(P, G)$. In this case we consider an almost Markov decomposition $f^m P = (\bigcup_q P_q) \cup Z$. Choose $y_q \in P_q$. We have

$$\int_P e^{G(x)} \sum_j A(\delta_m, f^j x) \rho_P(x) dx$$

$$= \sum_q c_q \int_{P_q} e^{G(f^{-m}(y)} \sum_j A(\delta_m(f^{-m} y), f^{j-m} y) \rho_{P_q}(y) dy + O(\theta^m).$$

For fixed q,

$$\int_{P_q} e^{G(f^{-m} y)} \sum_j A(\delta_m(y), f^{j-m} y) \rho_{P_q(y)} dy \tag{16}$$

$$= \int_{P_q} e^{G(f^{-m} y)} \sum_j A(\delta_m(y_q), f^{j-m} y) \rho_{P_q(y)} dy$$

$$+ \int_{P_q} e^{G(f^{-m} y)} \sum_j [A(\delta_m(y), f^{j-m} y) - A(\delta_m(y_q), f^{j-m} y)] \rho_{P_q(y)} dy.$$

Lemma 6. *$\exists C$ such that for small ε*

$$|\delta_j(x_1) - \delta_j(x_2)| \leq C \varepsilon d^\gamma(f^j x_1, f^j x_2), \tag{17}$$

where γ is such that $\mathbb{B} \subset C^\gamma(M)$.

Proof. Let C_k be a constant such that for $j<k$

$$|\delta_j(x_1)-\delta_j(x_2)|\le C_k\varepsilon d^\gamma(f^jx_1,f^jx_2).$$

Then

$$\begin{aligned}|\delta_{k+1}(x_1)-\delta_{k+1}(x_2)| &\le |\delta_k(x_1)-\delta_k(x_2)|+\varepsilon|B(\delta_k(x_1),f^kx_1)-B(\delta_k(x_2),f^kx_2)|\\ &\le C_k\varepsilon d^\gamma(f^jx_1,f^jx_2)+\varepsilon|B(\delta_k(x_1),f^kx_1)-B(\delta_k(x_1),f^kx_2)|\\ &\quad+\varepsilon|B(\delta_k(x_1),f^kx_2)-B(\delta_k(x_2),f^kx_2)|\\ &\le C_k\varepsilon d(f^jx_1,f^jx_2)+\varepsilon C(B)|\delta_k(x_1)-\delta_k(x_2)|+\varepsilon||B||d^\gamma(f^kx_1,f^k(x_2))\\ &\le [C_k\varepsilon+C(B)\varepsilon^2+\varepsilon||B||]d^\gamma(f^kx_1,f^kx_2).\end{aligned}$$

Since f is partially hyperbolic, $\exists\theta<1$ such that

$$d(f^kx_1,f^kx_2)\le\theta d(f^{k+1}x_1,f^{k+1}x_2).$$

Thus

$$|\delta_{k+1}(x_1)-\delta_{k+1}(x_2)|\le\varepsilon[C_k+C(B)\varepsilon+||B||]\theta^\gamma d^\gamma(f^{k+1}x_1,f^{k+1}x_2).$$

If ε is small enough, then $\varepsilon C(B)\le 1$, so

$$C_{k+1}\le(C_k+1+||B||)\theta^\gamma.$$

Thus if

$$C_{k+1}\le\frac{(1+||B||)\theta^\gamma}{1-\theta^\gamma},$$

then (17) holds. □

Thus the second term in the RHS of (16) is less than

$$\sum_{j=m}^{m+\frac{1}{\sqrt{\varepsilon}}}\text{Const}\,\varepsilon=\text{Const}\,\sqrt{\varepsilon}.$$

Now

$$\begin{aligned}&\int_{P_q}e^{G(f^{-m}y)}\sum_j A(\delta_m(y_q),f^{j-m}y)\rho_{P_q(y)}dy\\ &=||A(\delta_m(y_q),\cdot)||_{\mathbb{B}}\int_{P_q}e^{G(f^{-q}y)}\sum_j\frac{A(\delta_m(y_q),f^{j-m}y)}{||\delta_m(y_q),\cdot)||}\rho_{P_q}(y)dy\\ &\le||A(\delta_m(y_q),\cdot)||\sum_j a(\frac{j-m}{C})\le\text{Const}\,||A(\delta_m(y_q),\cdot)||.\end{aligned}$$

So,

$$\left|\ell\left(\sum_j A(\delta_m,f^jx)\right)\right|\le\sum_q c_q\theta_1(\delta_m(y_q))+\text{Const}\,\sqrt{\varepsilon}.$$

Using again Lemma 6 and the assumption that θ_1 depends smoothly on δ, we get

$$\sum_q c_q\theta_1(\delta(y_q))=\sum_q\int\theta_1(\delta_m(f^{-m}y))\rho_{P_q}(y)dy+O(\varepsilon)=\ell(\theta_1(\delta_m))+O(\varepsilon).$$

This completes the proof of (a).

(b) By Lemma 5, $T=T'+T''+O(\sqrt{\varepsilon})$. Hence

$$\ell(T^2)\le\text{Const}[\ell((T')^2)+\ell((T'')^2)+\varepsilon].$$

Lemma 7.

$$\ell((T')^2) \leq \text{Const}\left(\frac{\ell(\theta_2(\delta_m))}{\sqrt{\varepsilon}} + \sqrt{\varepsilon}\right).$$

Proof. It suffices to give a proof in the case $\ell = \ell(P, G)$. Let $f^m P = (\bigcup_q P_q) \cup Z$ be an almost Markov decomposition. Choose $y_q \in P$. Then

$$\int_P e^{G(x)}(T')^2 \rho_P(x)dx = \sum_q c_q \int_{P_q} e^{G(f^{-m}y)}(T'(f^{-m}t))^2 \rho_{P_q}(y)dy + O(\theta^m).$$

Now

$$\int_{P_q} e^{G(f^{-m}y)}(T'(f^{-m}t))^2 \rho_{P_q}(y)dy$$
$$= \int_{P_q} e^{G(f^{-m}y)}\left[\sum_j A(f^j x, \delta_m(y_q))\right]^2 \rho_{P_q}(y)dy$$
$$+ \int_{P_q} e^{G(f^{-m}y)}\left(\left[\sum_j A(\delta_m(y), f^j x)\right]^2 - \left[\sum_j A(\delta_m(y_q), f^j x)\right]^2\right)\rho_{P_q}(y)dy$$
$$= I_q + I\!I_q.$$

Now

$$I_q = ||A(\delta_m(y_q), \cdot)||^2 \int_{P_q} e^{G(f^{-m}y)} \frac{\left[\sum_j A(\delta_m(y_q), f^j x)\right]^2}{||A(\delta_m(y_q), \cdot)||^2}\rho_{P_q}(y)dy,$$

and by the argument of Lemma 1 the last integral is $O(\frac{1}{\sqrt{\varepsilon}})$. Hence

$$I_q \leq \text{Const}\,\frac{\theta_2(y_q)}{\sqrt{\varepsilon}}.$$

By Lemma 6,

$$\theta_2(y_q) = \int_{P_q} \theta_2(y)\rho_{P_q}(y)dy + O(\varepsilon).$$

Summation over q gives

$$\sum_q c_q I_q \leq \text{Const}\left[\frac{\ell(\theta_2(\delta_m))}{\sqrt{\varepsilon}} + \sqrt{\varepsilon}\right]. \tag{18}$$

Now

$$I\!I_q = \int_{P_q} e^{G(f^{-m}y)}\left[\sum_j (A(\delta_m(y), f^j x) - A(\delta_m(y_q), f^j x))\right]$$
$$\times\left[\sum_j (A(\delta_m(y), f^j x) + A(\delta_m(y_q), f^j x))\right]\rho_{P_q}(y)dy.$$

By Lemma 6,

$$\sum_j \left(\delta_m(y), f^j x) - A(\delta_m(y_q), f^j x)\right) \leq \sum_j O(\varepsilon) = O(\sqrt{\varepsilon}).$$

On the other hand,

$$\sum_j \left(A(\delta_m(y), f^j x) + A(\delta_m(y_q), f^j x)\right)$$
$$\leq 2\sum_j ||A(\delta_m(y), \cdot)|| + O(\varepsilon)$$
$$\leq \text{Const}\left(\frac{||A(\delta_m, \cdot)||}{\sqrt{\varepsilon}} + \sqrt{\varepsilon}\right).$$

Thus

$$\sum_q c_q I_q \leq \text{Const}\, \ell(|\theta_1(\delta_m)|). \tag{19}$$

But

$$\ell(|\theta_1(\delta_m)|) = \ell\left(\frac{|\theta_1(\delta_m)|}{\sqrt[4]{\varepsilon}}\sqrt[4]{\varepsilon}\right) \leq \frac{1}{2}\left(\frac{\ell(\theta_2(\delta_m))}{\sqrt{\varepsilon}} + \sqrt{\varepsilon}\right) \tag{20}$$

Combining (18), (19) and (20), we obtain the lemma. □

Now

$$\ell((T'')^2)$$
$$= \varepsilon^2 \sum_{k_1<j_1, k_2<j_2, j_1<j_2} \ell\left(\frac{dA}{d\delta}(\delta_m, f^{j_1}x)\frac{dA}{d\delta}(\delta_m, f^{2_1}x)B(\delta_{k_1}, f^{k_1}x)B(\delta_{k_2}, f^{k_2}x)\right).$$

By the argument of Lemma 1 we see that for fixed k_1, j_1, k_2

$$\left|\sum_{j_2} \ell\left(\frac{dA}{d\delta}(\delta_m, f^{j_1}x)\frac{dA}{d\delta}(\delta_m, f^{2_1}x)B(\delta_m, f^{k_1}x)B(\delta_m, f^{k_2}x)\right)\right| \leq \text{Const},$$

so the whole sum is bounded by

$$\ell((T'')^2) \leq \text{Const}\, \varepsilon^2 \left(\frac{1}{\sqrt{\varepsilon}}\right)^3 = \text{Const}\, \sqrt{\varepsilon}. \tag{21}$$

Lemma 7 and (21) prove (b).

To prove (c) we again use

$$\ell(T^4) \leq \text{Const}(\ell((T')^4) + \ell((T'')^4) + \varepsilon^2).$$

Lemma 8.

$$\ell((T')^4) \leq \text{Const}\left[\varepsilon + \ell(|\tilde{\kappa}_1\theta_3|) + \ell(|\theta_3|) + \ell(|\tilde{\kappa}_1\theta_2|) + \varepsilon\ell(|\theta_2|) + \ell(\theta_4(\delta_m)) + \varepsilon\ell(\kappa_4(\delta_m))\right].$$

Proof. It suffices to consider the case $\ell = \ell(P, G)$. We argue as in the proof of (a). Let $f^m P = (\bigcup_q P_q) \cup Z$ be an almost Markov decomposition. Choose $y_q \in P_q$. Then

$$\int e^{G(x)}(T')^4 \rho_P(x)dx$$
$$= \sum_q c_q \int_{P_q} e^{G(f^{-m}y)}(T'(f^{-m}y))^4 \rho_{P_q}(y)dy + O(\theta^m).$$

Now

$$\int_{P_q} e^{G(f^{-m}y)}(T'(f^{-m}y))^4\rho_{P_q}(y)dy$$
$$= \int_{P_q} e^{G(f^{-m}y)}\left[\sum_j A(\delta_m(y_q), f^j x)\right]^4 \rho_{P_q}(y)dy$$
$$+ \int_{P_q} e^{G(f^{-m}y)}\left[(\sum_j A(\delta_m(y), f^j x))^4 - (\sum_j A(\delta_m(y_q), f^j x))^4\right]\rho_{P_q}(y)dy = I_q + I\!I_q.$$

Reasoning as in Lemma 1(d), we obtain

$$|I_q| \le \mathrm{Const}\left(\frac{1}{\varepsilon}\right)^2 \theta_4(\delta_m(y_q)) = \mathrm{Const}\,\frac{\theta_4(\delta_m(y_q))}{\varepsilon}.$$

On the other hand,

$$I\!I_q = \int_{P_q} e^{G(f^{-m}y)}\left[(\sum_j A(\delta_m(y), f^j x))^4 - (\sum_j A(\delta_m(y_q), f^j x))^4\right]\rho_{P_q}(y)dy$$
$$= \int_{P_q} e^{G(f^{-m}y)}\left[(\sum_j A(\delta_m(y), f^j x) - A(\delta_m(y_q), f^j x))\right]$$
$$\times \left[(\sum_j A(\delta_m(y)), f^j x)^3 + (\sum_j A(\delta_m(y), f^j x))^2(\sum_j A(\delta_m(y_q), f^j x))\right.$$
$$\left. + (\sum_j A(\delta_m(y), f^j x))(\sum_j A(\delta_m(y_q), f^j x))^2 + (\sum_j A(\delta_m(y_q), f^j x))^3\right]\rho_{P_q}(y)dy.$$

Now

$$\sum_j \left[A(\delta_m(y), f^j x) - A(\delta_m(y_q), f^j x)\right] \tag{22}$$
$$= \left[\sum_j DA(\delta_q(y), f^j x)(\delta_m(y) - \delta_m(y_q))\right] + O(\sum_j [\delta_m - \delta - m(y_q)]^2)$$
$$= \left[\sum_j DA(\delta_q(y), f^j x)(\delta_m(y) - \delta_m(y_q))\right] + O(\sqrt{\varepsilon^3}),$$

since by Lemma 6 each term in the second sum is $O(\varepsilon^2)$.

Lemma 9.

$$(\sum_j A(\delta_m(y), f^j x))^3 + (\sum_j A(\delta_m(y), f^j x))^2 (\sum_j A(\delta_m(y_q), f^j x)) + (\sum_j A(\delta_m(y), f^j x))(\sum_j A(\delta_m(y_q), f^j x))^2 + (\sum_j A(\delta_m(y_q), f^j x))^3 = 4\left(\sum_j A(\delta_m(y_q), f^j x)\right)^3 + O\left[\left(\frac{\theta_2(\delta_m(y))}{\sqrt{\varepsilon}}\right) + \sqrt{\varepsilon}\right]. \tag{23}$$

Proof. Consider, for example, the first term. Other terms can be handled similarly. We have

$$\left(\sum_j A(\delta_m(y), f^j x)\right)^3 = \left[\left(\sum_j A(\delta_m(y), f^j x)\right)^3 - \left(\sum_j A(\delta_m(y_q), f^j x)\right)^3\right] + \left(\sum_j A(\delta_m(y_q), f^j x)\right)^3.$$

The first term here equals

$$\sum_j \left[A(\delta_m(y), f^j x) - A(\delta_m(y_q), f^j x)\right] \times \left[\left(\sum_k A(\delta_m(y), f^k x)\right)^2 + \left(\sum_k A(\delta_m(y), f^k x)\right)\left(\sum_r A(\delta_m(y_q), f^r x)\right) + \left(\sum_k A(\delta_m(y_q), f^k x)\right)^2\right].$$

By Lemma 6 the first factor is

$$\frac{1}{\sqrt{\varepsilon}} O(\varepsilon) = O(\sqrt{\varepsilon}).$$

On the other hand, using the formula for difference of squares the same way we did for cubes, we obtain that the first factor is $O(\frac{1}{\varepsilon}\theta_2(\delta_m) + 1)$. □

Multiplying (22) and (23), we get

$$\begin{aligned}
I\!I_q = 4\int_{P_q} e^{G(f^{-m}y)} \left[\sum_j DA(\delta_q(y), f^j x)(\delta_m(y)-\delta_m(y_q))\right] \\
\times \left(\sum_j A(\delta_m(y_q), f^j x)\right)^3 \rho_{P_q}(y)dy \\
+O\left(\sqrt{\varepsilon^3}\int_{P_q} e^{G(f^{-m}y)} \left|\sum_j A(\delta_m(y_q), f^j x)\right|^3 \rho_{P_q}(y)dy\right) \\
+\frac{1}{\sqrt{\varepsilon}}O\left(\int_{P_q} e^{G(f^{-m}y)} \left[\sum_j DA(\delta_q(y), f^j x)(\delta_m(y)-\delta_m(y_q))\right] |\theta_2(\delta_m(y))|\rho_{P_q}(y)dy\right) \\
+O\left(\varepsilon\int_{P_q} e^{G(f^{-m}y)}|\theta_2(\delta_m(y)|\rho(y)dy\right) \\
+O\left(\sqrt{\varepsilon}\int_{P_q} e^{G(f^{-m}y)} \sum_j DA(\delta_m(y_q), f^j x)(\delta_m(y)-\delta_m(y_q))\rho_{P_q}(y)dy\right) + O(\varepsilon^2) \\
= I\!I_q^{(1)} + I\!I_q^{(2)} + I\!I_q^{(3)} + I\!I_q^{(4)} + I\!I_q^{(5)} + I\!I_q^{(6)}.
\end{aligned}$$

By the argument of Lemma 1(d) we obtain

$$\begin{aligned}
I\!I_q^{(1)} \leq \text{Const} \sup_{P_q} \varepsilon\kappa_1(\delta_m(y))|\theta_3(\delta_m(y))| \left(\frac{1}{\sqrt{\varepsilon}}\right)^2 \\
= \text{Const} \sup_{P_q} \tilde{\kappa}_1(\delta_m(y))|\theta_3(\delta_m(y))| \\
= \text{Const}\left[\int_{P_q} e^{G(f^{-m}y)}\tilde{\kappa}_1(\delta_m(y))|\theta_3(\delta_m(y))|\rho_{P_q}(y)dy + \varepsilon\right].
\end{aligned}$$

Also

$$I\!I_q^{(2)} \leq \text{Const}\int_{P_q} e^{G(f^{-m}y)}|\theta_3(\delta_m(y))|\rho_{P_q}(y)dy,$$

$$I\!I_q^{(3)} \leq \text{Const}\int_{P_q} e^{G(f^{-m}y)}|\tilde{\kappa}_1(\delta_m(y))||\theta_2(\delta_m(y))|\rho_{P_q}(y)dy,$$

and, since $|\delta_m(y)-\delta_m(y_q)| = O(\varepsilon)$,

$$I\!I^{(5)} \leq \text{Const}\,\varepsilon^{\frac{3}{2}}\int_P e^{G(f^{-m}y)}|\tilde{\kappa}_1(\delta_m)|\rho_{P_q}(y)dy.$$

Thus

$$\sum_q c_q|I\!I_q| \leq \text{Const}\left[\varepsilon + \ell(|\kappa_1\theta_3|) + \ell(|\theta_3|) + \ell(|\kappa_1\theta_2|) + \varepsilon\ell(|\theta_2|) + \varepsilon^{\frac{3}{2}}\ell(|\kappa_1|) + \varepsilon^2\right].$$

Also,

$$\theta_4(\delta_m(y_q)) = \int_{P_q} e^{G(f^{-m}y)}\theta_4(\delta_m(y))\rho_{P_q}(y)dy$$
$$+O(\varepsilon \int_{P_q} e^{G(f^{-m}y)}|\kappa_4(\delta_m(y))|\rho_{P_q}(y)dy + \varepsilon^2).$$

Hence

$$\sum_q c_q|I_q| \le \mathrm{Const}(\ell(\theta_4(\delta_m)) + \varepsilon\ell(\kappa_4(\delta_m)) + \varepsilon^2).$$

This completes the proof of Lemma 8. □

On the other hand, the inequality

$$\ell((T'')^4) \le \mathrm{Const}\,\varepsilon$$

can be proven similarly to Lemma 1(d). This together with Lemma 8 completes the proof of (c). The proof of Proposition 20 is complete. □

Proposition 21.

$$\left|\ell\left(\sum_{j=1}^{\frac{T}{\varepsilon}} L_{j+1}^{-1}\gamma_j\right)\right| \le \mathrm{Const}\,.$$

Proof. We have

$$L_{j+1}^{-1}\gamma_j = L_{j-\sqrt{n}}^{-1}\gamma_j + \left[L_{j+1}^{-1} - L_{j-\sqrt{n}}^{-1}\right]\gamma_j.$$

Now

$$\left|l\left(\sum_j L_{j-\sqrt{n}}^{-1}\gamma_j\right)\right| \le n\,\mathrm{Const}\,a\left(\frac{\sqrt{n}}{C}\right) \le n\frac{\mathrm{Const}}{(\sqrt{n})^2} = \mathrm{Const}\,.$$

Also,

$$\left[L_{j+1}^{-1} - L_{j-\sqrt{n}}^{-1}\right]\gamma_j = \varepsilon L_{j-\sqrt{n}}^{-1}\sum_k \mathcal{Q}_k L_{j-\sqrt{n}}^{-1}\gamma_j + O\left(\left\|\left[\sum_{k=j-\sqrt{n}}^{j}\varepsilon\mathcal{Q}_k\right]\right\|^2\right)$$
$$= I_j + I\!I_j.$$

But we have $I\!I_j = O((\sqrt{\varepsilon})^2) = O(\varepsilon)$. Thus

$$\sum_j \ell(I\!I_j) = \sum_j O\left(\left\|\left[\sum_{k=j-\sqrt{n}}^{j}\varepsilon Q_k\right]\right\|^2\right) = O(1).$$

Also, similarly to the proof of Lemma 1,

$$\sum_j \ell(I_j) = \ell\left(\varepsilon\sum_{j>k} L_{j-\sqrt{n}}^{-1}\sum_k Q_k L_{j-\sqrt{n}}^{-1}\gamma_j\right) \le \mathrm{Const}\,\varepsilon\sum_{j,k} a\left(\frac{j-k}{C}\right).$$

Now, for fixed k,

$$\sum_{j>k} a\left(\frac{j-k}{C}\right) \le \mathrm{Const}\,.$$

So

$$\sum_j \ell(I_j) \leq \text{Const}\,\varepsilon \sum_{j=1}^{\frac{1}{\varepsilon}} O(1) = O(1).$$

□

14. Short time fluctuations in averaging. Recursive bounds

Here we complete the proof of Theorem 13. Let

$$a_{m,p} = \sup_\ell \left| \ell\left(\left(\sum_{j=0}^{\frac{m}{\sqrt{\varepsilon}}} L_j^{-1}\gamma_j \right)^p \right) \right|.$$

Lemma 10. (a) $a_{m,2} \leq \text{Const}\, m\sqrt{n}$.
(b) $a_{m,4} \leq \text{Const}\, m^2 n$.

Proof. We want to relate $a_{m+1,p}$ to $a_{m,p}$. Let $\bar{S} = \sum_{j=0}^{\frac{m}{\sqrt{\varepsilon}}} L_j^{-1}\gamma_j$, $\hat{S} = \sum_{\frac{m}{\sqrt{\varepsilon}}}^{\frac{m+1}{\sqrt{\varepsilon}}} L_j^{-1}\gamma_j$. We have

$$\ell((\bar{S}+\hat{S})^2) = \ell(\bar{S}^2) + 2\ell(\bar{S}\hat{S}) + \ell(\hat{S}^2).$$

Applying Proposition 20 to the last term, we get

$$\ell(\hat{S}^2) \leq \text{Const}\,\sqrt{n}.$$

To estimate the second term, we write

$$\bar{S}\hat{S} = \sqrt{a_{m,2}}\left(\frac{\bar{S}}{\sqrt{a_{m,2}}}\hat{S}\right)$$

and apply Proposition 20 with

$$A(x,\bar{S},L) = \frac{1}{\sqrt{a_m,2}}\bar{S}L^{-1}A(x).$$

Then $|\theta_1(\bar{S},L)| \leq \text{Const}$, so

$$|\ell(\bar{S}\hat{S})| \leq \text{Const}\,\sqrt{a_{m,2}},$$

and hence

$$\ell((\bar{S}+\hat{S})^2) \leq \ell(\bar{S}^2) + \text{Const}\left[\sqrt{n} + \sqrt{a_{m,2}}\right].$$

Taking the supremum over l, we obtain

$$a_{m+1,2} \leq a_{m,2} + \text{Const}(\sqrt{n} + \sqrt{a_{m,2}}).$$

Let $a_{m,2} = K_m m\sqrt{n}$. Then we get

$$K_{m+1}(m+1)\sqrt{n} \leq K_m m\sqrt{n} + \text{Const}(\sqrt{n} + \sqrt{K_m m\sqrt{n}})$$
$$\leq K_m m\sqrt{n} + \text{Const}(1+\sqrt{K_m})\sqrt{n}.$$

(The last inequality follows from the fact that $m \leq \mathbf{T}\sqrt{n}$.) Dividing by $m+1$, we get

$$K_{m+1} \leq K_m - \frac{K_m - (\sqrt{K_m}+1)\,\text{Const}}{m+1}.$$

If K is such that $K \geq (\sqrt{K}+1)\,\text{Const}$, then $K_m \leq K$ implies that $K_{m+1} \leq K$, and so (a) is proved by induction.

To prove (b), write

$$\ell((\bar{S}+\hat{S})^4) \le \ell(\bar{S}^4) + \text{Const}\left[\ell(\bar{S}^3\hat{S}) + \ell(\bar{S}^2\hat{S}^2) + \ell(\bar{S}\hat{S}^3) + \ell(\hat{S}^4)\right].$$

To estimate $\ell(\bar{S}^3\hat{S})$ we write

$$\bar{S}^3\hat{S} = (a_{m,4})^{\frac{3}{4}}\left(\frac{\bar{S}^3}{a_{m,4}^{\frac{3}{4}}}\hat{S}\right)$$

and apply Proposition 20 with

$$A(x,\bar{S},L) = \frac{1}{a_{m,4}^{\frac{3}{4}}}\bar{S}^3 L^{-1}A(x).$$

Then $|\theta_1(\bar{S},L)| \le \text{Const}$, so

$$|\ell(\bar{S}^3\hat{S})| \le \text{Const}\, a_{m,4}^{\frac{3}{4}}.$$

Also by Proposition 20

$$\ell(\hat{S}^4) \le \text{Const}\, n.$$

To estimate the other terms, we apply the Hölder inequality to get

$$\ell((\bar{S}+\hat{S})^4) \le \ell(\bar{S}^4) + \text{Const}\left[a_{m,4}^{\frac{3}{4}} + n + \ell(\bar{S}^4)^{\frac{1}{4}}\ell(\hat{S}^4)^{\frac{3}{4}} + \sqrt{\ell(\bar{S}^4)\ell(\hat{S}^4)}\right].$$

Taking the supremum, we get

$$a_{m+1,4} - a_{m,4} \le \text{Const}\left[a_{m,4}^{\frac{3}{4}} + \sqrt{a_{m,4}n} + a_{m,4}^{\frac{1}{4}}n^{\frac{3}{4}} + n\right].$$

Let $a_{m,2} = K_m m^2 n$; then we get

$$\begin{aligned} &K_{m+1}(m+1)^2 n - K_m m^2 n \\ &\le \text{Const}\left[K_m^{\frac{3}{4}} m^{\frac{3}{2}} n^{\frac{3}{4}} + \sqrt{K_m} m\sqrt{n}\sqrt{n} + K_m^{\frac{1}{4}}\sqrt{m} n^{\frac{1}{4}} n^{\frac{3}{4}} + n\right] \\ &\le \text{Const}\left[K_m^{\frac{3}{4}} mn + \sqrt{K_m} mn + K_m^{\frac{1}{4}}\sqrt{m} n + n\right]. \end{aligned}$$

(In the last inequality we are using the fact that $m \le \mathbf{T}\sqrt{n}$.) So if K is large enough, then $K_m \le K$ implies that $K_{m+1} \le K$. This proves (b). □

Now let

$$b_{m,p} = \sup_{\ell}\left|\left(\left(\sum_{j=0}^{\frac{m}{\sqrt{\varepsilon}}}\beta_j\right)^p\right)\right|,$$

$$d_{m,p} = \sup_{\ell}\left|\left(\Delta^p_{\frac{m}{\sqrt{\varepsilon}}}\right)\right|.$$

Using equation (14) and Lemma 10, we get

$$d_{m,p} \le \text{Const}(a_{m,p} + b_{m,p})\varepsilon^p \le \text{Const}(b_{m,p} + (m\sqrt{n})^{\frac{p}{2}})\varepsilon^p. \tag{24}$$

The next step gives recursive relations for $b_{m,p}$.

Proposition 22. (a)

$$b_{m+1,2} - b_{m,2} \leq \text{Const}(\sqrt{b_{m,2}d_{m,2}} + \sqrt{n}d_{m,2}).$$

(b) *Let* $D_m = \frac{d_{m,4}}{\varepsilon} + d_{m,4}^{\frac{3}{4}} + d_{m,2} + \varepsilon^{\frac{3}{2}}$. *Then*

$$b_{m+1,4} - b_{m,4} \leq \text{Const}\left[b_{m,4}^{\frac{3}{4}}\sqrt{d_{m,2}} + \sqrt{b_{m,4}D_m} + b_{m,4}^{\frac{1}{4}}D_m^{\frac{3}{4}} + D_m\right].$$

Proof. (a) Let $R' = \sum_{j=1}^{\frac{m}{\sqrt{\varepsilon}}} \beta_j$, $R'' = \sum_{\frac{m}{\sqrt{\varepsilon}}}^{\frac{m+1}{\sqrt{\varepsilon}}} \beta_j$. We have

$$\ell((R'+R'')^2) = \ell((R')^2) + 2\ell(R'R'') + \ell((R'')^2).$$

Thus

$$b_{m+1,2} - b_{m,2} \leq \left[\sup_\ell \ell(R'R'') + \ell((R'')^2)\right].$$

Applying Proposition 20 with $\delta = (q, \Delta)$ and

$$A(q,\Delta,x) = \frac{[DA(q,x) + \zeta(q,x,\Delta) - \mathcal{Q}]\,\Delta}{d_{m,2}},$$

we get

$$\ell((R'')^2) \leq \text{Const}\, d_{m,2}\sqrt{n}.$$

Applying Proposition 20 with $\delta = (q, \Delta, R')$ and

$$A(q,\Delta,x) = \frac{[DA(q,x) + \zeta(q,x,\Delta) - \mathcal{Q}]\,\Delta R'}{\sqrt{d_{m,2}b_{m,2}}},$$

we obtain

$$|\ell(R'R'')| \leq \ell(\Delta_{m\sqrt{n}}R') \leq \text{Const}\,\sqrt{d_{m,2}b_{m,2}}.$$

(b) First we estimate $\ell((R'')^4)$. To this end we apply Proposition 20 and note that

$$|\theta_4(\delta_m)| \leq \text{Const}\,\Delta_{m\sqrt{n}}^4,$$
$$|\kappa_4(\delta_m)| \leq \text{Const}\,|\Delta_{m\sqrt{n}}|^3,$$
$$|\theta_3(\delta_m)| \leq \text{Const}\,|\Delta_{m\sqrt{n}}|^3,$$
$$|\theta_2(\delta_m)| \leq \text{Const}\,|\Delta_{m\sqrt{n}}|^2,$$
$$|\tilde{\kappa}_1(\delta_m)\theta_3(\delta_m)| \leq \text{Const}\,|\Delta_{m\sqrt{n}}|^3,$$
$$|\tilde{\kappa}_1(\delta_m)\theta_2(\delta_m)| \leq \text{Const}\,|\Delta_{m\sqrt{n}}|^2,$$
$$|\tilde{\kappa}_1(\delta_m)| \leq \text{Const},$$

and, that, by the Hölder inequality

$$\ell\left(|\Delta_{m\sqrt{n}}|^3\right) \leq d_{m,4}^{\frac{3}{4}}.$$

Thus

$$\ell((R'')^4) \leq \text{Const}\, D_m.$$

To estimate $\ell((R')^3R'')$ we apply Proposition 20 with

$$A = \frac{(R')^3\,[DA(q,x) + \zeta(q,x,,\Delta) - \mathcal{Q}_n]\,\Delta}{\sqrt{d_{m,2}}b_{m,4}^{\frac{3}{4}}}.$$

This gives

$$\ell((R')^3R'') \leq \text{Const}\,\sqrt{d_{m,2}}b_{m,4}^{\frac{3}{4}}.$$

To estimate the remaining terms we use the Hölder inequality. This completes the proof of (b) . □

Now, using the *a priori* bound $|\Delta_m| \leq \text{Const}$, we see that the contribution of $b_{m,p}$ to (24) is not larger that the contribution of $a_{m,p}$. Thus

$$d_{m,p} \leq \text{Const}(m\sqrt{n})^{\frac{p}{2}}\varepsilon^p. \tag{25}$$

Plugging this bound into Proposition 22(a), we get

$$b_{m+1,2} - b_{m,2} \leq \text{Const}(\sqrt{b_{m,2}} + 1)\sqrt{\varepsilon}.$$

From this we obtain by induction that, for $m \leq \frac{1}{\sqrt{\varepsilon}}$,

$$b_{m,2} \leq \text{Const}\, m\sqrt{\varepsilon}. \tag{26}$$

Also, (25) implies that

$$d_{m,p} \leq \text{Const}\, \varepsilon^{\frac{p}{2}}.$$

Hence $D_m \leq \text{Const}\, \varepsilon$. The inequality of Proposition 22(b) becomes

$$b_{m+1,4} - b_{m,4} \leq \text{Const}\left[b_{m,4}^{\frac{3}{4}}\sqrt{\varepsilon} + \sqrt{b_{m,4}\varepsilon} + b_{m,4}^{\frac{1}{4}}\varepsilon^{\frac{3}{4}} + \varepsilon\right].$$

Now, repeating the argument of Lemma 10(b), we get

$$b_{m,4} \leq \text{Const}\, m^2 n\varepsilon^2. \tag{27}$$

Proposition 23. (a) $\{\Delta_t^\varepsilon\}$ *is a tight family.*

(b) *Let* $\beta_t^\varepsilon = \frac{\beta_{[\frac{t}{\varepsilon}]}}{\sqrt{\varepsilon}}$. *Then* $\beta_t^\varepsilon \to 0$ *in probability as* $\varepsilon \to 0$.

Proof. In view of the inequalities (26)–(27), the proof of (a) is similar to the proof of Lemma 2. The similar argument implies that $\{\frac{\beta_t^\varepsilon}{\sqrt{\varepsilon}}\}$ is tight, and so $\beta_t^\varepsilon \to 0$. □

Proposition 24. *Let* $L_t^\varepsilon = L_{[\frac{t}{\varepsilon}]}$. *Then, as* $\varepsilon \to 0$, L_t^ε *converges to the solution of the ODE*

$$\frac{dL}{dt} = DA(q(t))L.$$

Proof. This follows immediately from the equation (12), the bound

$$||\bar{\zeta}(q_n, \Delta_n)|| \leq \text{Const}\, ||\Delta_n||$$

and the fact that $\Delta_{[\frac{t}{\varepsilon}]} \to 0$ weakly. □

Proof of Theorem 4. If Δ_t is some limit of Δ_t^ε, then it follows from (13) and Propositions 19, 23 and 24 that Δ_t satisfies the equation

$$\Delta(t) = L(t)\int_0^t L^{-1}(s)dB(s), \tag{28}$$

where $\frac{dL}{dt} = D\bar{A}(q(t))L$. Differentiating (28), we get

$$d\Delta = D\bar{A}(q(t))\Delta dt + dB(t).$$

This completes the proof of Theorem 4. □

15. Long time fluctuations in averaging

Here we prove Theorem 5. Recall from Section 13 that for the sake of notational simplicity we give the proof only for the case $z \in \mathbb{R}^1$, the general case being completely similar.

Lemma 11. $\left|\ell\left(\sum_{j=0}^n A(z_j, f^j x)\right)\right| \le \mathrm{Const}(1+\varepsilon n).$

Proof. Let $r = \beta_\varepsilon\sqrt{\frac{1}{\varepsilon}}$, where β_ε is chosen so that $\beta_\varepsilon \to 0$ but $\frac{a_r}{\varepsilon} \to 0$ as $\varepsilon \to 0$. For $j > r$ we write

$$z_j = z_{j-r} + \sum_{m=j-r}^{j-1} \varepsilon A(z_m, f^m x).$$

From this and Lemma 5 we obtain

$$A(z_j, f^j x) = A(z_{j-r}, f^j x) + \varepsilon \sum_{m=j-r}^{j-1} DA(z_{j-r}, f^{j-r}x)A(z_m, f^m x) + O(\varepsilon^2 r^2).$$

Similarly to the proof of Theorem 2, we get

$$\left|\ell(A(z_{j-r}, f^j x))\right| \le \mathrm{Const}\, a_r.$$

Thus

$$\ell\left(\sum_{j=r}^n A(z_j, f^j x)\right) \tag{29}$$
$$= O(\varepsilon^2 n r^2) + O(n a_r) + \varepsilon \ell\left(\sum_{m=0}^n A(z_m, f^m x) \sum_{k=m+1}^{m+r} DA(z_{k-r}, f^k x)\right).$$

By the argument of Lemma 1 the contribution to the last term of each fixed m can be bounded by

$$\mathrm{Const}\, \varepsilon \sum_{k=m+1}^{m+r} a\left(\frac{k-m}{C}\right) \le \mathrm{Const}\, \varepsilon \sum_k a\left(\frac{k}{C}\right).$$

Thus $\ell(\sum_{j=r}^n A(z_j, f^j x)) = O(\varepsilon n)$. Similarly

$$\sum_{j=0}^r \ell(A(z_j, f^j x))$$
$$= \sum_{j=0}^r \ell(A(z_0, f^j x)) + \varepsilon \sum\sum_{j>k} \ell\left(A(z_k, f^k x) DA(z_0, f^j x)\right) + O(\varepsilon^2 r^3).$$

Similarly to Lemma 1(a) and (b), we can estimate the first term here by Const and the second by Const εr. □

Corollary 14. (a)

$$\ell\left(\left[\sum_{j=0}^{\frac{\Delta}{\varepsilon}} A(z_j, f^j x)\right]^2\right) \le \mathrm{Const}\,\frac{\Delta}{\varepsilon}.$$

(b) *As* $\Delta \to 0$,

$$\varepsilon \ell\left(\left[\sum_{j=0}^{\frac{\Delta}{\varepsilon}} A(z_j, f^j x)\right]^2\right) \sim \Delta\sigma(z_0).$$

Proof. (a) follows from Lemma 10(a).

(b) We have

$$\ell\left(\left[\sum_{j=0}^{\frac{\Delta}{\varepsilon}} A(z_j, f^j x)\right]^2\right) = \sum_{j,k} \ell\left(A(z_j, f^j x) A(z_k, f^k x)\right).$$

Break this sum into two parts:

$$\sum_{|j-k|\leq K} \ell(A(z_j, f^j x) A(z_k, f^k x)) \sim \frac{\Delta}{\varepsilon}\nu\left(\sum_{k=-K}^{K} A(z_0, x) A(z_0, f^k x)\right).$$

On the other hand,

$$\left|\sum_{|j-k|>K} \ell(A(z_j, f^j x) A(z_k, f^k x))\right| = o_{K\to\infty}(1) + O\left(\frac{\Delta^2}{\varepsilon}\right).$$

Letting $K \to \infty$, we obtain the required statement. □

Lemma 12. (a)

$$\ell\left(\left[\sum_{j=0}^{\frac{\Delta}{\varepsilon^2}} A(z_j, f^j x)\right]^2\right) \leq \text{Const}\,\frac{\Delta}{\varepsilon^2}.$$

(b) *As* $\Delta \to 0$,

$$\ell\left(\left[\sum_{j=0}^{\frac{\Delta}{\varepsilon^2}} A(z_j, f^j x)\right]^2\right) \sim \sigma(z_0)\frac{\Delta}{\varepsilon^2}.$$

Proof. (a) We proceed by induction. Namely, we will show that for each k there is a constant R_k such that

$$\ell\left(\left[\sum_{j=0}^{\frac{2^k}{\varepsilon}} A(z_j, f^j x)\right]^2\right) \leq R_k \frac{2^k}{\varepsilon^2}. \tag{30}$$

Corollary 14 show that this is true for $k=1$. Let us see haw to pass from k to $k+1$. We have

$$\ell\left(\left[\sum_{j=0}^{\frac{2^{k+1}}{\varepsilon}} A(z_j, f^j x)\right]^2\right) = \ell\left(\left\{\left[\sum_{j=0}^{\frac{2^k}{\varepsilon}} A(z_j, f^j x)\right] + \left[\sum_{\frac{2^k}{\varepsilon}+1}^{\frac{2^{k+1}}{\varepsilon}} A(z_j, f^j x)\right]\right\}^2\right)$$
$$= \ell\left(\left[\sum_{j=0}^{\frac{2^k}{\varepsilon}} A(z_j, f^j x)\right]^2 + \left[\sum_{\frac{2^k}{\varepsilon}+1}^{\frac{2^{k+1}}{\varepsilon}} A(z_j, f^j x)\right]^2\right)$$
$$+2\ell\left(\sum_{0\leq j\leq \frac{2^k}{\varepsilon}<m\leq \frac{2^{k+1}}{\varepsilon}} A(z_j, f^j x)A(z_m, f^m x)\right).$$

The sum of the first two terms is bounded by $R_k\frac{2^{k+1}}{\varepsilon}$, by the induction hypothesis. By the argument of Lemma 11 the last term is less than

$$\ell\left(\left|\sum_{j=0}^{\frac{2^k}{\varepsilon}} A(z_j, f^j x)\right|\right) \text{Const } 2^k.$$

By the induction hypothesis the first factor here is at most $\sqrt{R_k}\frac{2^{\frac{k}{2}}}{\sqrt{\varepsilon}}$. Thus

$$\frac{R_{k+1}2^{k+1}}{\varepsilon} \leq 2\frac{R_k 2^k}{\varepsilon} + O\left(\sqrt{\frac{R_k 2^k}{\varepsilon}}2^k\right).$$

In other words,

$$R_{k+1} \leq R_k + O(\sqrt{R_k 2^k \varepsilon}).$$

Let $R_k^* = \max(R_k, 1)$; then

$$R_{k+1}^* \leq R_k^*\left(1 + O\left(\sqrt{2^k\varepsilon}\right)\right).$$

Hence

$$R_k^* \leq R_0^* \prod_{j=0}^{k}\left(1 + \text{Const}\sqrt{2^j\varepsilon}\right).$$

Now

$$2^j\varepsilon = 2^k\varepsilon 2^{j-k} \leq \frac{\Delta}{2^{k-j}}.$$

The second term is less than $\prod_{m=0}^{\infty}(1 + \text{Const}\sqrt{\Delta 2^{-j}})$. Hence the R_k^* and so the N_k are uniformly bounded. This proves (a).

(b): (a) implies that as $\Delta \to 0$, $z_n \to z_0$ in probability uniformly for $n < \frac{\Delta}{\varepsilon^2}$. Hence we can repeat the computation of (a), replacing (30) by the assumption that $\forall N = \frac{2^k}{\varepsilon}$,

$$\ell\left(\left[\sum_{j=0}^{N} A(z_j, f^j x)\right]^2\right) = (\sigma(z_0) + \rho_{l,k})N,$$

where $|\rho_{l,k}| < \delta_k$. We then get

$$\delta_{k+1} \leq \delta_k \left(1 + O\left(\sqrt{\frac{2^k \varepsilon}{\delta_k}}\right)\right).$$

We want to show that, given $\delta > 0$, there exists $\bar{\Delta}$ such that $|\delta_k| < \delta$ for $\Delta < \bar{\Delta}$. Let $k'(k)$ be the largest number less than k such that $|\delta_{k'}| < \delta^2$. Reasoning as in (a), we get

$$\delta_k < \delta^2 \prod_{j=k'}^{k} \left(1 + \text{Const}\, \frac{\sqrt{2^j \varepsilon}}{\delta}\right) \leq \delta^2 \prod_{l=1}^{\infty} \left(1 + \text{Const}\, \frac{\sqrt{2^{-l}\Delta}}{\delta}\right).$$

The second term converges to 1 as $\Delta \to 0$. This proves (b). □

Corollary 15. *As* $\Delta \to 0$,

$$\ell\left(\sum_{j=0}^{\frac{\Delta}{\varepsilon^2}} A(z_j, f^j x)\right) \sim \frac{\Delta}{\varepsilon} a(z_0).$$

Proof. By (29) we have

$$\ell\left(\sum_{j=0}^{\frac{\Delta}{\varepsilon^2}} A(z_j, f^j x)\right) = \varepsilon \sum_{j=0}^{\frac{\Delta}{\varepsilon^2}} \sum_k \ell\left(A(z_j, f^j x) DA(z_{k-r}, f^k x)\right) + o\left(\frac{\Delta}{\varepsilon}\right)$$

$$= \varepsilon \sum_{j=0}^{\frac{\Delta}{\varepsilon^2}} \sum_{k=1}^{K} \ell\left(A(z_j, f^j x) DA(z_{j+k-r}, f^k x)\right) + o_{K\to\infty}\left(\frac{\Delta}{\varepsilon}\right).$$

But for fixed j

$$\ell\left(A(z_j, f^j x) \sum_{k=1}^{K} DA(z_{j+k-r}, f^{j+k} x)\right) \sim \nu\left(A(z_j, f^j x) \sum_{k=1}^{K} DA(z_j, f^{j+k} x)\right).$$

Also, $z_j \to z_0$ in probability by Lemma 12; thus

$$\sum_{j=0}^{\frac{\Delta}{\varepsilon^2}} \ell\left(A(z_j, f^j x) \sum_{k=1}^{K} DA(z_{j+k-r}, f^{j+k} x)\right) \sim \frac{\Delta}{\varepsilon} \sum_{k=1}^{K} \nu\left(A(z_0, f^j x) DA(z_0, f^k x)\right).$$

Letting $K \to \infty$, we obtain the required statement. □

Lemma 13.

$$\ell\left(\left[\sum_{j=0}^{\frac{\Delta}{\varepsilon^2}} A(z_j, f^j x)\right]^4\right) \leq \text{Const}\, \frac{\Delta^2}{\varepsilon^4}.$$

Proof. We proceed as in Lemma 12. The inequality

$$\ell\left(\left[\sum_{j=0}^{\frac{\Delta}{\varepsilon}} A(z_j, f^j x)\right]^4\right) \leq \text{Const}\, \frac{\Delta^2}{\varepsilon^2}$$

follows from Lemma 10. Let M_k be the number such that

$$\ell\left(\left[\sum_{j=0}^{\frac{2^k}{\varepsilon}} A(z_j, f^j x)\right]^4\right) \leq M_k \left(\frac{2^k}{\varepsilon}\right)^2.$$

Let

$$\bar{T} = \sum_{j=0}^{\frac{2^k}{\varepsilon}} A(z_j, f^j x),$$

$$\hat{T} = \sum_{j=\frac{2^k}{\varepsilon}+1}^{\frac{2^{k+1}}{\varepsilon}} A(z_j, f^j x).$$

We have

$$\ell\left(\left[\sum_{j=0}^{\frac{2^{k+1}}{\varepsilon}} A(z_j, f^j x)\right]^4\right)$$

$$= \ell((\bar{T}+\hat{T})^4) = \ell(\bar{T}^4) + \ell(\hat{T}^4) + 4\ell(\bar{T}^3\hat{T}) + 4\ell(\bar{T}\hat{T}^3) + 6\ell(\bar{T}^2\hat{T}^2).$$

Using the argument of Proposition 20, Corollary 14 and Lemma 12, we obtain

$$\ell\left(\bar{T}^4\right) \leq M_k \frac{2^{2k}}{\varepsilon^2}, \tag{31}$$

$$\ell\left(\hat{T}^4\right) \leq M_k \frac{2^{2k}}{\varepsilon^2}, \tag{32}$$

$$|\ell(\hat{T}^3\bar{T})| \leq \text{Const}\, \ell(|\hat{T}|^3)2^k \leq \text{Const}\, M_k^{\frac{3}{4}} \frac{2^{\frac{5k}{2}}}{\varepsilon^{\frac{3}{2}}} \leq \text{Const}\, M_k^{\frac{3}{4}} \frac{2^{2k}}{\varepsilon^2} \tag{33}$$

(the last inequality is true because $2^k \leq \frac{1}{\varepsilon}$), and

$$\ell(\hat{T}^2\bar{T}^2) \leq \text{Const}\, \frac{2^{2k}}{\varepsilon^2}. \tag{34}$$

Lemma 14.

$$\left|\ell(\hat{T}\bar{T}^3)\right| \leq \text{Const}\, M_k^{\frac{3}{4}} \frac{2^{2k}}{\varepsilon^2}.$$

Proof. It is enough to prove this for $\ell = \ell(P, G)$. Denote $k^* = \frac{2^k}{\varepsilon}$. Consider an almost Markov decomposition $f^{k^*}P = \bigcup P_j \cup Z$. Denote $\xi_j = \sup_{f^{-k^*}P_j} |\hat{T}| + 1$. Then

$$\left|\ell(\hat{T}\bar{T}^3)\right|$$

$$\leq \sum_j c_j \xi_j \left| \int_{P_j} e^{G(f^{-k^*}y)} \frac{(\hat{T}(f^{-k^*}y)}{\xi_j} \left[\sum_{j=0}^{k^*-1} A(z_{k^*+j}, f^j y)\right]^3 \rho_{P_j}(y) dy \right| + O\left(\theta^{k^*}\right).$$

Now the Hölder norm of $\hat{T} \circ f^{-k^*}$ is $O(1)$. Now, any bounded function can be decomposed as a difference of two positive functions as follows:

$$A = 2||A||_{L^\infty} - (2||A||_{L^\infty} - A).$$

This implies that

$$\ell_j^*(A) = \int_{P_j} e^{G(f^{-k^*}y)} \frac{(\hat{T}(f^{-k^*}y)}{\xi_j} A(y)\rho_{P_j}(y)dy$$

can be written as $\ell_j^* = a_1\ell_j' - a_2\ell_j''$, where $\ell_j', \ell_j'' \in E(\mathcal{P}, R, \alpha)$ and $|a_1| < \text{Const}$, $|a_2| < \text{Const}$. Thus

$$\left| \ell_j \left(\left[\sum_{j=0}^{k^*-1} A(z_{k^*+j}, f^j y) \right]^3 \right) \right|$$
$$\leq (\text{Hölder}) \left(\ell_j \left(\left[\sum_{j=0}^{k^*-1} A(z_{k^*+j}, f^j y) \right]^4 \right) \right)^{\frac{3}{4}}$$
$$\leq (\text{inductive hypothesis}) \frac{M_k^{\frac{3}{4}} 2^{\frac{3k}{2}}}{\varepsilon^{\frac{3}{2}}}.$$

Thus

$$\left| \ell \left(\hat{T}\bar{T}^3 \right) \right| = O\left(\frac{M_k^{\frac{3}{4}} 2^{\frac{3k}{2}}}{\varepsilon^{\frac{3}{2}}} \sum_j c_j \xi_j \right) + O(\theta^{k^*}) = O\left(\frac{M_k^{\frac{3}{4}} 2^{\frac{3k}{2}}}{\varepsilon^{\frac{3}{2}}} \sum_j c_j \xi_j \right).$$

Using the argument of Proposition 20, we get

$$\sum_j c_j \xi_j \leq \text{Const}(\ell(|\hat{T}|) + 1) \leq \text{Const}(\sqrt{\ell(\hat{T}^2)} + 1) \leq \text{Const} \frac{2^{\frac{k}{2}}}{\sqrt{\varepsilon}},$$

where the last inequality follows by (30). Thus

$$|\ell(\hat{T}\bar{T}^3)| \leq \text{Const}\, M_k^{\frac{3}{4}} \frac{2^{2k}}{\varepsilon^2},$$

as claimed. □

Combining (31)–(34) and Lemma 14, we get

$$\frac{M_{k+1} 2^{2k+2}}{\varepsilon^2} = 2\frac{M_k 2^{2k}}{\varepsilon^2} + \text{Const}\left(M_k^{\frac{3}{4}} \frac{2^{2k}}{\varepsilon^2} + 1 \right).$$

Thus

$$M_{k+1} \leq \frac{M_k}{2} + K\left(M_k^{\frac{3}{4}} + 1 \right).$$

Hence if M is so large that

$$\frac{M}{2} \geq K\left(M_k^{\frac{3}{4}} + 1 \right),$$

then $M_k \leq M$ implies $M_{k+1} \leq M$. This completes the proof of Lemma 13. □

Corollary 16. $\{Z_t^\varepsilon\}$ *is tight.*

Proof. In view of Lemma 13, the proof is the same as the proof of Lemma 2. □

To prove Theorem 5 we need the following characterization of diffusion processes. (See, for example, [86], Exercise 4.6.6.)

Proposition 25. *Let $(\xi_t, \mathcal{F}_t)$ be a random process with continuous paths such that*

$$\xi_t - \int_0^t a(\xi_s)ds \quad \text{and} \quad \left(\xi_t - \int_0^t a(\xi_s)ds\right)^2 - \int_0^t \sigma(\xi_s)ds$$

are martingales. Then ξ_t is diffusion with drift $a(x)$ and diffusion coefficient $\sigma(x)$.

Proof of Theorem 5. By Corollary 16, $\{Z_t^\varepsilon\}$ is a tight family. Let Z be some limit of Z_t^ε. We need to show that if $Q(z_1 \dots z_m)$ is any smooth bounded function and $t_1, \dots, t_m$ are any numbers, $t_j \leq t$, then

$$\mathbb{E}\left(Q(Z_{t_1} \dots Z_{t_s})\left[Z_{t+\Delta} - Z_t - \Delta a(Z_t)\right]\right) = o(\Delta) \tag{35}$$

and

$$\mathbb{E}\left(Q(Z_{t_1} \dots Z_{t_s})\left[(Z_{t+\Delta} - Z_t)^2 - \Delta\sigma(Z_t)\right]\right) = o(\Delta). \tag{36}$$

Let us consider (35) ((36) is similar). In terms of the original family Z_t^ε we have to show that $\forall \ell \in E(\mathcal{P}, R, \alpha)$

$$\ell\left(Q(, Z_{t_1/\varepsilon^2}, \dots, Z_{t_m/\varepsilon^2})\left(Z_{(t+\Delta)/\varepsilon^2} - Z_{t/\varepsilon^2} - \Delta a(Z_{t/\varepsilon^2})\right)\right) \to 0$$

uniformly as $\varepsilon \to 0$. It suffices to verify this for $\ell = \ell(P, G)$. However, in this case the proof proceeds as before by considering an almost Markov decomposition $f^{t/\varepsilon^2} P = \left(\bigcup_j P_j\right) \cup Z$ and applying Corollary 15 to each P_j. The details are left to the reader. □

16. Three series theorem

Here we prove Theorem 6. Consider a series

$$(S) = \sum_n c_n A_n(f^n x).$$

It is enough to assume that $\nu(A_n) = 0$, $\sum_n c_n^2 < \infty$, $\sum a(m) < \infty$, $\|A_n\|_B \leq 1$. Take a Markov family $\mathcal{P}$. Let $P \in \mathcal{P}$.

Proposition 26. *(S) converges in $L^2(\ell(P))$.*

Proof. We have

$$\begin{aligned}
&\ell([\sum_{n=N}^{\infty} c_n A_n \circ f^n]^2) \\
&= \text{Const} \sum_{m,n=N}^{\infty} c_n c_m a\left(\frac{n-m}{C}\right) \\
&\leq \text{Const} \sum_{m,n=N}^{\infty} (c_n^2 + c_m^2) a\left(\frac{n-m}{C}\right) \\
&\leq \text{Const} \sum_{m,n=N}^{\infty} c_n^2 a\left(\frac{n-m}{C}\right) \\
&\leq \text{Const} \left(\sum_{m=1}^{\infty} a(m)\right)\left(\sum_{n=N}^{\infty} c_n^2\right).
\end{aligned}$$

□

Let $\mathcal{F}_N$ be as in Section 11.

Proposition 27. *$\forall Q \in \mathcal{F}_N$, $\forall A \in \mathbb{B}$ such that $\nu(A) = 0$, $||A|| \le 1$,*

$$\left|\int_Q A(f^n x)\rho_P(x)dx\right| \le \mathrm{Const}\, a(n-N)\,\mathrm{Vol}(Q).$$

Proof. Indeed,

$$\left|\int_Q A(f^n x)\rho_Q(x)dx\right| = \int_{f^N Q} A_n(f^{n-N}y)\rho_{f^N Q}(y)dy \le a(n-N),$$

but $\rho_P = c_{P,Q}\rho_Q$, where $c_{P,Q} \sim \mathrm{Vol}(Q)$. □

Proof of the theorem. Denote by B the L^2-sum $B = \sum c_n A_n \circ f^n$. We have, $\forall Q \in \mathcal{F}_r$,

$$\int_Q B(x)\rho_P(x)dx = \left(\sum_{n=1}^{r} + \sum_{n=r+1}^{\infty}\right) c_n A_n(f^n x)\rho_P(x)dx = I + II.$$

But

$$|II| \le \mathrm{Const} \sum_{n=r+1}^{\infty} c_n a(n-r)\,\mathrm{Vol}(Q) \le \mathrm{Const}(\max_{n\ge r} c_n)\,\mathrm{Vol}(Q).$$

Let y be any point in Q Then

$$I = \sum_{n=1}^{r} c_n A(f^n y)\,\mathrm{Vol}(Q) + O(\sum_{n=1}^{r} c_n \theta^{r-n})\,\mathrm{Vol}(Q).$$

The second term can be bounded as follows:

$$\sum_{n=1}^{r} c_n\theta^{r-n} \le (\max_n c_n)\sum_{n=1}^{\frac{r}{2}}\theta^{r-n} + \max_{n>\frac{r}{2}} c_n(\sum_{n>\frac{r}{2}}\theta^{r-n}) \le \mathrm{Const}(\theta^{\frac{r}{2}} + \max_{n>\frac{r}{2}} c_n).$$

Hence

$$\frac{\int_Q B(x)\rho_P(x)dx}{\mathrm{Vol}(Q)} = \sum_{n=1}^{r} c_n A(f^n y) + o(1),$$

so the theorem follows by Doob's martingale convergence theorem. □

Note. In this section we followed [55] quite closely.

17. Borel–Cantelli lemma

Here we prove Theorem 7. Let r_n be the radius of B_n, $p_n = \nu(B_n) \sim r_n^d$, $\mu_{mn} = \nu(1_{B_m}(f^m x)1_{B_n}(f^n x))$. Set $S_N = \sum_{j=1}^N 1_{B_j}(f^j x)$, $E_N = \mathbb{E}(S_N) = \sum_{n=1}^N p_n$, and

$$V_N = \mathbb{E}(S_N^2) = \sum_{m,n=1}^{N} \mu_{mn}. \tag{37}$$

Lemma 15. *$V_N \le E_N^2 + CE_N$.*

Proof. To estimate V_N we break the sum (37) into five parts. Below, ε is such that, for $k < \varepsilon\ln(\frac{1}{p_m})$, $f^k B_m \cap B_n$ has at most one component, c_1 is an arbitrary constant and c_2 is a constant whose value will be chosen at the end of this section.

(I) $m = n$. Then $I = \sum_{n=1}^N \mu_{nn} = \sum_{n=1}^N p_n = E_N$.

(II) $m < n < m + \varepsilon \ln(\frac{1}{p_m})$. Consider the set $\tilde{B}_m$ obtained as follows. For any leaf $W^u(x)$ such that $W^u_{loc}(x) \cap B_m \neq \emptyset$, choose a ball $W_{m,x}$ of radius r_m containing $W^u_{loc}(x) \cap B_m$. Let $\tilde{B}_m = \bigcup_x W_{m,x}$.

Proposition 28. $\nu(\tilde{B}_m) \leq \text{Const}\, p_m$.

Proof. Let dist_u denote the distance in the induced W^u metric. Then locally we have $\text{dist}_u(\cdot,\cdot) \leq \text{Const}\,\text{dist}(\cdot,\cdot)$. Hence $\tilde{B}_m$ is contained in a ball with the same center as B_m and of radius $\text{Const}\, r_m$. □

Proposition 29. *If $m < n < m + \varepsilon \ln(\frac{1}{p_m})$, then*

$$\mu_{mn} \leq C\theta^{n-m}(p_n + p_m).$$

Proof. Choose δ such that $f^k W_{m,x}$ contains a ball of radius $(1+\delta)^k r_m$. Consider two cases:

(a) $r_n \geq r_m(1+\frac{\delta}{2})^{n-m}$. Then $\mu_{mn} \leq p_m \leq \text{Const}(1+\delta)^{d(n-m)} p_n$.

(b) $r_n < r_m(1+\frac{\delta}{2})^{n-m}$. Let ℓ_x denote $\ell(W_{m,x})$. Then $\ell_x(f^{n-m}W_{m,x} \cap B_n) \leq \text{Const}(1+\delta)^{(n-m)}\,\text{Vol}(W_{m,x})$, and hence $\mu_{mn} \leq \text{Const}(1+\frac{\delta}{2})^{n-m} p_m$. □

(III) $m + \varepsilon \ln(\frac{1}{p_m}) \leq n \leq m + c_1 \ln(\frac{1}{p_m})$. Then $f^{n-m}W_{m,x}$ contains a ball of radius $r_m^{1-\gamma}$, $\gamma = \gamma(\varepsilon)$. Again there are two cases.

(a) If $r_n \leq r_m^{1-\frac{\gamma}{2}}$, then any component of $f^{n-m}W_{m,x} \cap B_n$ can be surrounded by an annulus of width $r_m^{1-\gamma} - r_m^{1-\frac{\gamma}{2}}$ disjoint from B_n. Thus $\exists \delta_1$ such that

$$\ell_x(f^{n-m}W_{m,x} \cap B_n) \leq \text{Const}\, p_m^{\delta_1}.$$

Thus $\mu_{mn} \leq p_m^{1+\delta_1}$.

(b) If $r_n > r_m^{1-\frac{\gamma}{2}}$, then $\mu_{mn} \leq p_m$.

(IV) $m + c_1 \ln(\frac{1}{p_m}) < n < m + (\frac{1}{p_m})^{c_2}$.

Proposition 30. $\mu_{mn} \leq \text{Const}\, p_m^{\frac{d}{d-d_u}}$.

Proof. Now any component of $f^{n-m}W_{m,x} \cap B_n$ can be surrounded by an annulus of constant width disjoint from B_n. Hence

$$\ell_x(f^{n-m}W_{m,x} \cap B_n) \leq \text{Const}\, r_n^{d_u}.$$

On the other hand, $\mu_{mn} \leq p_n$. So

$$\mu_{m,n} \leq C \sup_r(\min(r^d, r^{d_u} p_m)) = C p_m^{\frac{d}{d-d_u}}.$$

□

(V) $n > m + (\frac{1}{p_m})^{c_2}$. The following is analogous to Theorem 2.

Proposition 31. *Let B^1 and B^2 be two balls, of radii r^1 and r^2 respectively. Then, given n_0, $\exists C(n_0)$ such that*

$$|\nu(1_{B^1}(x)1_{B^2}(f^m x)) - \nu(B^1)\nu(B^2)| \leq \text{Const}\left[\left(\frac{1}{m - C\ln r^1}\right)^k \left(\frac{1}{r^2}\right)^\alpha + r^{n_0}\right].$$

So, $\mu_{mn} \le p_m p_n + \delta_{mn}$, where for δ_{mn} we have two bounds:

$$\delta_{mn} \le C[(\frac{1}{p_n})^\alpha (\frac{1}{n-m})^k + p_m^{n_0}] \quad \text{and} \quad \delta_{mn} \le p_n.$$

Hence

$$\delta_{mn} \le \sup_p \left(\min \left(C \left(\frac{1}{p}\right)^\alpha \left(\frac{1}{n-m}\right)^k + p_m^{n_0}, p \right)\right) = C \left(\frac{1}{n-m}\right)^{\frac{k}{\alpha+1}}.$$

(Here we have used that $\frac{1}{n-m} \gg p_m^{n_0}$.)

Let us sum up these terms. Direct calculation shows that

$$(I) = E_N;$$

$$(II) \le \operatorname{Const} E_N;$$

$$(III)(a) \le \operatorname{Const} \sum_m p_m^{1+\delta} \ln \left(\frac{1}{p_m}\right) \le \operatorname{Const} \sum_m p_m \le \operatorname{Const} E_N;$$

$$(IV) \le \sum_m \left(\frac{1}{p_m}\right)^{c_2} p_m^{\frac{d}{d-d_u}} \le E_N$$

if $c_2 < \frac{d_u}{d-d_u}$. To estimate (III)(b), observe that we have two lower bounds for p_m. First, $p_m \le p_n^{1+\delta}$, and second, $p_m \le e^{-\frac{n-m}{c_1}}$. Thus

$$\begin{aligned}(III)(b) &\le \sum_n \left(\sum_{n-m<(\frac{1}{p_n})^{\frac{\delta}{2}}} p_n^{1+\delta} + \sum_{m-n\ge(\frac{1}{p_n})^{\frac{\delta}{2}}} e^{-\frac{n-m}{c_1}} \right) \\ &\le \sum_n \operatorname{Const} p_n \le \operatorname{Const} E_N.\end{aligned}$$

Finally,

$$\begin{aligned}(V) &\le E_N^2 + \sum_m \sum_{j>(\frac{1}{p_m})^{c_2}} \left(\frac{1}{j}\right)^{\frac{k}{\alpha+1}} \\ &\le E_N^2 + C \sum_m p_m^{c_2(\frac{k}{\alpha+1}-1)} \le E_N^2 + CE_N\end{aligned}$$

if $c_2(\frac{k}{\alpha+1} - 1) \ge 1$, i.e., $c_2 \ge \frac{\alpha+1}{k-(\alpha+1)}$. So for c_2 we have two inequalities:

$$\frac{\alpha+1}{k-(\alpha+1)} \le c_2 < \frac{d_u}{d-d_u}.$$

They are compatible, since $\frac{k}{\alpha+1} > \frac{d}{d_u}$. Combining these bounds, we get $V_N \le E_N^2 + \operatorname{Const} E_N$, as claimed. This completes the proof of Lemma 15. □

Proof of Theorem 7. By Lemma 15, $\mathbb{E}([\frac{S_N}{E_N} - 1]^2) \le \frac{\operatorname{Const}}{E_N}$. Choose N_j so that $E_{N_j} \ge 2^j$. Then, by the Borel-Cantelli lemma, $\frac{E_{N_j}}{S_{N_j}} \to 1$ ν–almost surely. Thus $S_{N_j} \to \infty$ ν–almost surely. Since S_N is non-decreasing, $S_N \to \infty$. □

Notes. The first Borel-Cantelli lemma for a dynamical system was proved in [70]. [87] and [54] prove Borel-Cantelli for some partially hyperbolic dynamical systems on non-compact manifolds and present several applications to geometry and number theory. [19] deals with Anosov diffeomorphisms and establishes Borel-Cantelli under various assumptions on the shapes of the B_n.

18. Poisson law

Here we prove Theorem 8. Let $B_n = B(x_0, \frac{1}{n})$, $X_{n,\theta} = \sum_{j=1}^{n^\theta} 1_{B_n}(f^j x)$.

Lemma 16. *If $\theta > \frac{1+\alpha}{k}$, then $\ell(X) = n^\theta \nu(B_n)(1 + o(1))$.*

Proof. We have

$$\ell(X) = n^\theta \nu(B_n) + O\left(\sum_{j=1}^{n^\theta} \min\left((\frac{1}{n})^{d_u}, n^\alpha \left(\frac{1}{j}\right)^k\right)\right).$$

The second term is $O((\frac{1}{n})^{d_u(1-\frac{1}{k})-\frac{\alpha}{k}})$. If $\theta > \frac{1+\alpha}{k}$, then the main term here is the first one. □

Let us estimate $\ell(X \geq 2)$. Denote $W_{n,x} = B_n \cap W^u_{loc}(x)$ and $\ell_x = \ell(W_{n,x})$. Fix K. Put $\hat{B}_n(K) = \bigcup_{W_{n,x} \supset B^u(\bar{y}, \frac{1}{nK})} W_{n,x}$.

Proposition 32.

$$\frac{\nu(\hat{B}_n(K))}{\nu(B_n)} \to 1$$

as $K \to \infty$ uniformly over n.

Proof. Similarly to Proposition 28, $B_n \backslash B_n(K) \subset B(x_0, \frac{1}{n}(1 - \frac{\text{Const}}{K}))$. □

We have

$$\ell(X \geq 2) \leq \ell(\exists j \leq n^\theta : f^j x \in B_n \backslash \hat{B}_n(K))$$
$$+ \sum_m \ell(\exists j \leq n^\theta : 1_{B_n}(f^{j+m} x) = 1 | f^n x \in \hat{B}_n(K)) \ell(1_{\hat{B}_n(K)}(f^m x)).$$

By Proposition 32 the first term is less than $\varepsilon n^\theta \nu(B_n)$ if K is large enough. To bound the second term, break it into four parts.

(I) $j \leq M_0$. This term vanishes, since x_0 is not periodic.

(II) $M_0 < j \leq \epsilon \ln n$.

Proposition 33. *$\forall \epsilon\ \exists M_0$ such that $II \leq \varepsilon \ell(X)$.*

Proof. The intersection $f^j(W_{n,x}) \cap B_n$ has at most one component. Hence

$$\ell_x(f^j(W_{n,x}) \cap B_n) \leq C \frac{r_n^{d_u}}{\text{Vol}(f^j W_{n,x})} \leq C \xi^j,$$

$\xi < 1$. So

$$II \leq \sum_m 1_{\hat{B}_n(K)}(f^m x) \sum_{j=M_0}^{\infty} C \xi^j \leq \sum_{j=M_0}^{\infty} C \xi^j \ell(X) \leq C \frac{\xi^{M_0}}{1-\xi} \ell(X),$$

and the last expression goes to 0 as M_0 tends to infinity. □

(III) $\epsilon \ln n < j \leq C_1 \ln n$.

Proposition 34. *For fixed C_1, $\exists \tilde{\varepsilon}$ such that $III \leq \text{Const}(\ln n) n^{\tilde{\varepsilon}} \ell(X)$.*

Proof. Here for any component of $f^jW_{n,x}\cap B_n$ there is an annulus of width at least $(\frac{1}{n})^{1-\tilde\varepsilon}$ disjoint from B_n. Hence

$$\sum_j \ell_x(f^jW_{n,x}\cap B_n)\le \text{Const}(\ln n)\left(\frac{1}{n}\right)^{\tilde\varepsilon}.$$

□

(IV) $C_1\ln n<j\le n^\theta$.

Proposition 35. *If C_1 is large enough, then $IV\le \text{Const}(\frac{1}{n})^{d_u}n^\theta\ell(X)$.*

Proof. Here for any component of $f^jW_{n,x}\cap B_n$ there is an annulus of width of order 1 disjoint from B_n. So $\ell_x(f^jW_{n,x}\cap B_n)\le\text{Const}(\frac{1}{n})^{d_u}$. Hence

$$\sum_j \ell_x(f^jW_{n,x}\cap B_n)\le n^\theta\left(\frac{1}{n}\right)^{d_u}.$$

□

Thus $IV\le\varepsilon\ell(X)$ if $\theta<d_u$. So we have for θ the inequalities $\frac{1+\alpha}{k}<\theta<d_u$. They are compatible if $\frac{k}{\alpha+1}>\frac{1}{d_u}$. So we have

Proposition 36. *Let $\frac{1+\alpha}{k}<\theta<d_u$. Then*

$$\ell(e^{itX})=1-n^\theta\nu(B_n)(1-e^{it})+o(n^\theta\nu(B_n)).$$

Now introduce

$$X_{n,k}=\sum_{j=1}^{k}\sum_{l=jn^\theta}^{(j+1)n^\theta-n^{\frac{\theta}{2}}}1_{B_n}(f^lx).$$

Then $X_{n,(\nu(B_n)n^\theta)^{-1}}-\sum_{l=1}^{\nu(B_n)^{-1}}1_{B_n}(f^lx)$ converges to 0 in probability. Let $\phi_{n,k}(\ell,t)=\mathbb{E}_\ell(e^{itX_{n,k}})$.

Proposition 37.

$$\phi_{n,k}(\ell,t)=\left[1-n^\theta\nu(B_n)(1-e^{it})\right]^k+o(kn^\theta\nu(B_n)).$$

Proof (Induction on k). For $k=1$ this is the subject of Proposition 36. Assume that we have established our claim for k. Take $\ell\in E_1$, $\ell=\ell(P)$. Consider an almost Markov decomposition $f^{(k+1)n^\theta}P=(\bigcup_j P_j)\cup Z$. Choose $y_j\in f^{-(k+1)n^\theta}P_j$. Then

$$\begin{aligned}\phi_{n,k+1}(\ell,t)&=\sum_j c_je^{itX_{n,k}(y_j)}\phi_{n,1}(\ell(P_j),t)+O(\zeta^n)\\&=\sum_j c_je^{itX_{n,k}(y_j)}[(1-n^\theta\nu(B_n)(1-e^{it}))+O(\varepsilon n^\theta\nu(B_n))]\\&=[\phi_{n,k}(\ell,t)+O(\zeta^{n^{\frac{\theta}{2}}})][(1-n^\theta\nu(B_n)(1-e^{it}))+O(\varepsilon n^\theta\nu(B_n))]\\&=[(1-n^\theta\nu(B_n)(1-e^{it}))^k+O(\delta_k+\zeta^{n^{\frac{\theta}{2}}})][(1-n^\theta\nu(B_n)(1-e^{it}))+O(\varepsilon n^\theta\nu(B_n))]\\&=(1-n^\theta\nu(B_n)(1-e^{it}))^{k+1}+\delta_{k+1},\end{aligned}$$

where

$$\delta_k\le\delta_k+\varepsilon n^\theta\nu(B_n)+\text{Const}\,\zeta^{\frac{n}{2}}.$$

□

Proof of Theorem 8. Since $X_n(\Delta)$ is a point process, we only need to establish the convergence of finite-dimensional distributions. Let $\Delta_1, \dots, \Delta_m$ be disjoint intervals. By Proposition 37,

$$\ell(X_n(\Delta_1) = n_1) \sim \frac{\Delta_1^{n_1}}{n_1!} e^{-\Delta_1}.$$

Repeating the argument of Proposition 37, we obtain

$$\ell(\bigcap_j \{X_n(\Delta_j) = n_j\}) \sim \prod_j \left(\frac{\Delta_j^{n_j}}{n_j!}\right) e^{-\Delta_j}.$$

□

Notes. (1) There are two useful extensions of Theorem 8. The first says that if x_0 is periodic of least period T, then $X_n(\Delta)$ is asymptotically distributed as $\sum_{j \in N_\Delta} \xi_j$, where N_Δ is the Poisson process with the unit density and the ξ_j are mutually independent, independent of N_Δ and identically distributed. Their distribution can be obtained as follows. Let M be a linear transformation of a d-dimensional Euclidean space such that at least one eigenvalue of M has absolute value greater then 1. Let η be uniformly distributed in the unit ball $\mathcal{B}$. Define $\xi(M) = \sum_{k=1}^{\infty} 1_{\mathcal{B}}(M^k \eta)$. Then the ξ_j have the same distribution as $\xi(df^T(x_0))$. (The proof is the same as before, but now (I) is not zero.) Second, one can consider the pair $(j, n \operatorname{dist}(f^j x, x_0))$, where j is such that $f^j x \in B_n$, and prove the Poisson limit for this pair. (Again the proofs are very similar, but now balls need to be replaced by annuli.) One application of this generalization of the Poisson law is the following.

Corollary 17. *Let $m_n = \min_{j \le n} \operatorname{dist}(f^j x, x_0)$. If x_0 is aperiodic, then*

$$\nu(n^{\frac{1}{d}} m_n < t) \sim \exp(-K(x_0) t^d).$$

Thus, for a typical point, $(\frac{1}{n})^{\frac{1}{d}}$ is a correct normalization for $m_n(x)$. [39] studies the set of points with different asymptotic behavior of m_n.

(2) Other classes of dynamical systems satisfying the Poisson law are described in [40, 41, 42, 22]. The method of proof we use is similar to that of [79] (cf. also [72, 22]).

Appendix A. Absolute continuity

Proof of Proposition 2. We will use the following fact (see [13]). Let D_1 and D_2 be smooth $(d - d_u)$-dimensional discs transversal to E_u. Let $x_j \in D_j$ be points such that $x_2 \in W^u(x_1)$ and $\operatorname{dist}_u(x_1, x_2) \le 1$. Then locally near x_1 we can define a continuous map $p : D_1 \to D_2$ such that $px_1 = x_2$ and $px = W^u_{loc}(x)$. Then p is absolutely continuous and its Jacobian $J_p(x)$ is Hölder continuous, where the Hölder constant depends only on the angle between TD_j and E_u and the norms of the embeddings $D_j = i_j D$, D being the standard disc in $\mathbb{R}^{d-d_u}$. (In fact,

$$J_p(x) = \lim_{n \to \infty} \frac{\det(df^{-n}|TD_2)(x)}{\det(df^{-n}|TD_1)(x)}.)$$

Now let U be a parallelogram obtained as follows. Take $x_0 \in X$. Locally near x_0, chose a foliation $\mathcal{V}$ transversal to E_u. Then near x_0 we have a local product structure; that is, for $x, y \in X$ there is a unique point $z = W^u_{loc}(x) \cap V(y)$, where $V(y)$ is the leaf of $\mathcal{V}$ containing y. Write $z = [x, y]$. Consider the set U of the form $U = [V_0, W^u_{loc}(x_0)]$, where V_0 is a small disc in $V(x_0)$. We first show that the restriction of the Lebesgue measures to U belongs to $E(R, \alpha)$, where the constants R and α do not depend on the choice of V_0. Decompose $V_0 = \bigcup V_j$, where the V_j are small discs in V_0. Take $x_j \in V_j$ and let $W_j = [x_j, W^u_{loc}(x_0)]$, $U_j = [V_j, W^u_{loc}(x_0)]$. Then

$$\int_{U_j} A(x)dx = \left[\int_{W_j} dy \left(\int_{V_j(y)} A(v)dv\right)\left(\frac{dx}{dydv}\right)(y)\right](1 + o(1)),$$

where $V_j(y) = [V_j, y]$ is the slice of $\mathcal{V}$ inside U_j. By the Hölder continuity of E_u, $\frac{dx}{dydv}(y)$ is Hölder continuous. Also $\int_{V_j(y)} \sim A(y)\,\mathrm{Vol}(V_j(y))$ and $\mathrm{Vol}(V_j(y)) \sim \mathrm{Vol}(V_j)J_{p_y}(y)$, where p_y is the projection $p_y : V_j \to V_j(y)$. This verifies our claim. Now the same remains true if instead of requiring U to be a parallelogram we only ask that unstable slices of U satisfy conditions (a)–(d) of the definition of an almost Markov family and that they depend continuously on the point in the sense that if π is the projection along $\mathcal{V}$ leaves, then $\pi W_U(y) \to W_U(x)$ in the Hausdorff topology as $y \to x$. (Indeed, such sets can be approximated by parallelograms.) Now decomposing $X = \bigcup_j \hat{U}_j$, where the $\hat{U}_j$ are the sets as above, completes the proof of the proposition. □

Proof of Proposition 3. This proposition does not use the absolute continuity of W^u. In fact, it remains valid if we replace W^u by any continuous foliation with smooth leaves. We only have to show that any $\ell \in E$ assigns zero measure to u-negligible sets. Choose a small r. Let D be a $(d - d_u)$-dimensional disc. Denote by U a union of unstable balls of radii r centered at D. For $x \in D$ let ℓ_x denote $\ell(W^u_r(x))$. Then $x \to \ell_x$ is continuous (see, e.g., [75]). Thus the map $A \to \bar{A}(x) = \ell_x(A)$ is continuous from $C(U) \to C(D)$. Therefore the set $M(U)$ of measures of the form $\int_D \mu(x)\ell_x$ is weakly closed in $C(X)^*$. Now take $\ell \in E(\mathcal{P}, R, \alpha)$. By the definition it is a limit of some $\ell_j \in E_2(\mathcal{P}, R, \alpha)$. Let $\ell_j = \sum_k c_{jk}\ell(P_{jk}, G_{jk})$. If $\partial P_{jk} \cap U \neq \emptyset$, enlarge P_{jk} slightly so that the boundary of the resulting sets P'_{jk} is disjoint from U. By property (b) of an almost Markov family, this can be done in such a way that $\mathrm{meas}(P'_{jk}) \leq \mathrm{Const}\,\mathrm{meas}(P_{jk})$. Let $\ell'_j = \frac{1}{c_j}\sum c_{jk}\ell(P'_{jk})$, where c_j is the normalization constant. Then $\ell_j|_U \leq \mathrm{Const}\,\ell'_j|_U$. Thus it is enough to show that any limit point of ℓ'_j assigns zero measure to u-negligible sets. But $\ell'_j \in M(U)$. Thus if $\ell'_j \to \ell'$, then $\ell' \in M(U)$. So the statement follows by Fubini's theorem. □

Acknowledgment

A part of this paper constitutes a part of my Ph.D. thesis at Princeton University. I thank my thesis advisor Ya. G. Sinai for encouragement during my work and many helpful discussions. I am grateful to Yu. Kifer, D. Kleinbock, M. Pollicott and R. Sharp for useful conversations.

References

[1] Alves J., Bonatti C. & Viana M. *SRB measures for partially hyperbolic systems whose central direction is mostly expanding,* Inv. Math. **140** (2000) 35—398. MR **2001j:**37063b

[2] Anosov D. V. *Geodesic flows on closed Riemannian manifolds with negative curvature* Proc. Steklov Inst. Math. **90** (1967). MR **39:**3527

[3] Anosov D. V. & Sinai Ya. G. *Certain smooth ergodic systems*, Uspekhi Mat. Nauk **22** (1967), no. 5, 107–172; English transl., Russian Math. Surveys **22** (1967), no. 5, 103–167. MR **37:**370

[4] Bakhtin, V. I. *Random processes generated by a hyperbolic sequence of mappings. I,* Izv. Ross. Akad. Nauk Ser. Mat. 58 (1994), no. 2, 40–72; English transl., Russian Acad. Sci. Izv. Math. **44** (1995) 247–279.

[5] Bakhtin, V. I. *On the averaging method in a system with fast hyperbolic motions,* Proc. Belorussian Math. Inst. **6** (2000) 23–26.

[6] Barreira L. & Schmeling J. *Sets of non-typical points have full topological entropy and full Haussdorff dimension*, Israel J. Math **116** (2000) 29–70. MR **2002d:**37040

[7] Bonnatti C.& Viana M. *SRB measures for partially hyperbolic systems: mostly contracting case*, Israel J. Math **115** (2000) 157–193. MR **2001j:**37063a

[8] Bowen R. *'Equilibrium states and ergodic theory of Anosov diffeomorphisms'* Lect. Notes in Math. **470** (1975) Springer New York. MR **56:**1364

[9] Bowen R. *Weak mixing and unique ergodicity on homogeneous spaces* Isr. J. Math. **23** (1976) 267-273. MR **53:**11016

[10] Brezin J. & Shub M. *Stable ergodicity in homogeneous spaces* Bol. Soc. Brasil. Mat. (N.S.) **28** (1997) 197–210. MR **99a:**58103

[11] Brin M. *Topological transitivity of one class of dynamical systems and flows of frames on manifolds of negative curvature* Func. An. & Appl. **9** (1975) 8-16. MR **51:**6886

[12] Brin M. *Topology of group extensions of Anosov systems* Math. Notes **18** (1975) 858-864. MR **52:**15563

[13] Brin M. & Pesin Ya. B. *Partially hyperbolic dynamical systems* Math. USSR-Izvestiya **8** (1974) 177-218. MR **49:**8058

[14] Burns K., Pugh C. C. & Wilkinson A. *Stable ergodicity and Anosov flows*, Topology **39** (2000) 149–159. MR **2000i:**37031

[15] Burns K. & Wilkinson A. *Stable ergodicity of skew products*, Ann. Sci. Esc. Norm. Sup. **32** (1999) 859–889. MR **2000g:**37030

[16] Castro A. *Backward inducing and statistical properties of some partially hyperbolic attractors,* Israel J. Math. **130** (2002) 29–75.

[17] Chernov N. I. *Limit theorems and Markov approximations for chaotic dynamical systems* Prob. Th., Rel. Fields **101** (1995) 321-363. MR **96m:**28016

[18] Chernov N. I. & Haskell C. *Nonuniformly hyperbolic K-systems are Bernoulli* Erg. Th. & Dyn. Sys. **16** (1996) 19–44. MR **97k:**28031

[19] Chernov N. I. & Kleinbock D. Y. *Dynamical Borel–Cantelli lemmas for Gibbs measures,* Israel J. Math. **122** (2001), 1–27. MR **2002h:**37003

[20] Dani S. G. *On orbits of endomorphisms of tori and the Schmidt game,* Erg. Th. & Dyn. Sys. **8** (1988) 523-529. MR **90b:**58145

[21] Denker M. *The Central Limit Theorem for dynamical systems,* Banach Center Publ. **23** Warsaw, Polish Sci. Publ. (1989) 33-62. MR **92d:**28007

[22] Denker M. *Remarks on weak limit laws for fractal sets,* In Fractal geometry and stochastics (C. Bandt, S. Graf & M. Zahle editors) Progr. Prob. **37** (1995) Birkhauser, Basel. MR **97k:**28032

[23] Denker M. & Philipp W. *Approximation by Brownian motion for Gibbs measures and flows under a function,* Erg. Th. & Dyn. Sys. **4** (1984) 541-552. MR **86g:**28025

[24] Dolgopyat D. *On decay of correlations in Anosov flows,* Ann. Math. **147** (1998) 357-390. MR **99g:**58073

[25] Dolgopyat D. *Prevalence of rapid mixing in hyperbolic flows,* Erg. Th. & Dyn. Sys. **18** (1998) 1097-1114. MR **2000a:**37014

[26] Dolgopyat D. *On mixing properties of compact group extensions of hyperbolic systems*, Israel Math. J. **130** (2002) 157–205.

[27] Dolgopyat D. *On dynamics of mostly contracting diffeomorphisms,* Comm. Math. Phys. **213** (2000) 181–201. MR **2001h:**37056

[28] Dolgopyat D. *On differentiability of SRB states for partially hyperbolic systems*, to appear in Inv. Math.

[29] Dolgopyat D., Kaloshin V. & Koralov L. *Sample path properties of stochastic flows,* to appear in Ann. Prob.

[30] Eberlein E. & Taqqu M. S. (editors) *Dependence in probability and steatistics. A survey of recent results,* Progress in prob. and Stat. **11**, Birkhauser, Boston, 1986. MR **88b:**60006

[31] Ellis R. & Perrizo W. *Unique ergodicity of flows on homogeneous spaces,* Isr. J. Math. **29** (1978) 276-284. MR **57:**12774

[32] Gordin M. I. *On Central Limit Theorem for stationary processes,* Doklany AN SSSR **188** (1969) 739–741; English transl., Soviet Math. Dokl. **10** (1969), 1174–1176. MR **40:**5012
[33] Gordin M. I. *Double extensions of dynamical systems and construction of the mixing filtrations,* J. Math. Sci. (New York) **99** (2000), 1053–1060. MR **2000e:**37003
[34] Gordin M. I. *Double extensions of dynamical systems and construction of the mixing filtrations-2: quasi-hyperbolic toral automorphisms,* J. Math. Sci. (New York) **109** (2002), 2103–2114. MR **2001e:**37008
[35] Grayson M., Pugh C. C. & Shub M. *Stably ergodic diffeomorphisms,* Ann. Math. **140** (1994) 295-329. MR **95g:**58128
[36] Guivarch Y. & Hardy J. *Theoremes limites pour une classe de chaines de Markov et applications aux diffeomorphisms d'Anosov,* Annales de Inst. H. Poincare (Ser. Prob & Stat) **24** (1988) 73-98. MR **89m:**60080
[37] Guivarch Y. & Le Borgne S. *Methode de martingales et flot geodesique sur une surface de courbure constante negative*, preprint.
[38] Hall P. & Heyde C. C. *Martingale limit theory and its application,* Academic Press, New York, 1980. MR **83a:**60001
[39] Hill R. & Velani S. L. *The ergodic theory of shrinking targets,* Invent. Math. **119** (1995) 175–198. MR **96e:**58088
[40] Hirata M. *Poisson law for Axiom A diffeomorphisms,* Erg. Th. & Dyn. Sys. **13** (1993) 533-556. MR **94m:**58137
[41] Hirata M. *Poisson law for the dynamical systems with self-mixing conditions,* In Dynamical systems and chaos, Ed. N. Aoki, K. Shiraiwa & Y. Takahashi, World Scientific, Singapore (1995) v. 1, 87-96. MR **98h:**81032
[42] Hirata M., Saussol B. & Vaienti S. *Statistics of return times: a general framework and new applications,* Comm. Math. Phys. **206** (1999) 33–55. MR **2001c:**37007
[43] Hirsch M., Pugh C. C. & Shub M. *Invariant manifolds,* Lect. Notes in Math. **583** Springer-Verlag, Berlin, 1977. MR **58:**18595
[44] Ibragimov I. A. & Linnik Yu. V. *Independent and stationary sequences of random variables,* Wolters-Noordhoff Publishing, Groningen, 1971. MR **48:**1287
[45] Katok, A. *Smooth non-Bernoulli K-automorphisms,* Invent. Math. **61** (1980) 291–299. MR **84e:**58063a, MR **84e:**58063b
[46] Katok A.& Kononenko A.*Cocycle stability for partially hyperbolic systems,* Math. Res. Lett. **3** (1996) 191-210. MR **97d:**58152
[47] Katok A. & Spatzier R. J. *First cohomology of Anosov actions of higher rank abelian groups and applications to rigidity,* Inst. Hautes Études Sci. Publ. Math. **79** (1994) 131–156. MR **96c:**58132
[48] Katok A. & Strelcyn J.–M. *Invariant manifolds, entropy and billiards; smooth maps with singularities,* Lect. Notes Math. **1222** Springer, Berlin, 1986. MR **88k:**58075
[49] Katznelson Y. *Ergodic automorphisms of* $\mathbb{T}^n$ *are Bernoulli shifts*, Israel J. Math. **10** (1971) 186-195. MR **45:**3672
[50] Kifer Yu. *Large deviations in dynamical systems and stochastic processes,* Transactions AMS **321** (1990) 505–524. MR **91e:**60091
[51] Kifer Yu. *Averaging in dynamical systems and large deviations,* Inv. Math. **110** (1992) 337–370. MR **93m:**60118
[52] Kifer Yu. *Limit theorems in averaging for dynamical systems,* Erg. Th. & Dyn. Sys. **15** (1995) 1143–1172. MR **97c:**58087
[53] Kleinbock D. Y. & Margulis G. A. *Bounded orbits of nonquasiunipotent flows on homogeneous spaces,* Amer. Math. Soc. Transl. **171** (1996) 141-172. MR **96k:**22022
[54] Kleinbock D. Y. & Margulis G. A. *Logarithm laws for flows on homogeneous spaces*, Inv. Math. **138** (1999) 451–494. MR **2001i:**37046
[55] Krzyzewski K. *On convergence of certain series related to Axiom A diffeomorphisms,* Bull. Acad. Polon. Ser. Math. **30** (1982) 25-30. MR **84b:**58086
[56] Le Borgne S. *Limit theprems for non-hyperbolic automorphisms of the torus,* Israel J. Math. **109** (1999) 61–73. MR **2000a:**37001
[57] Le Borgne S. *Principes d'invariance pour les flots diagonaux sur* $\mathrm{SL}(d,\mathbb{R})/\mathrm{SL}(d,\mathbb{Z})$, Ann. Inst. H. Poincare Probab. Stat. **38** (2002), 581–612. MR **2003c:**37008

[58] Liverani C. *Central limit theorem for deterministic systems,* in International conference on dynamical systems (Eds. F. Ledrappier, J. Lewowics & S. Newhouse), Pitman Res. Notes **363** (1996) 56-75. MR **98k:**28025

[59] Liverani C. *On Contact Anosov flows,* preprint.

[60] Margulis G. A. *Certain measures that are connected with U-flows on compact manifolds,* Func. Anal. Appl. **4** (1970) 62–76. MR **42:**7865

[61] Moore C. C. *Ergodicity of flows on homogeneous spaces,* Amer. J. Math. **88** (1966) 154-178. MR **33:**1409

[62] Nitica V. & Torok A. *An open dense set of stably ergodic diffeomorphisms in a neighbourhood of a non-ergodic one,* Topology **40** (2001), 259–278. MR **2001m:**37058

[63] Ornstein D. & Weiss B. *On Bernoulli nature of systems with some hyperbolic structure,* Erg. Th. & Dyn. Sys. **18** (1998) 441–456. MR **99c:**58096

[64] Parry W. & Pollicott M. *Zeta Functions and Periodic Orbit Structure of Hyperbolic Dynamics,* Asterisque v. **187-188** (1990). MR **92f:**58141

[65] Pene F. *Averaging method for differential equations perturbed by dynamical systems,* ESAIM Probab. Stat. **6** (2002), 33–88 (electronic). MR **2003i:**34118

[66] Pesin Ya. B. *Families of invariant manifolds that correspond to nonzero characteristic exponents,* Izv. Akad. Nauk SSSR **40** (1976) 1332–1379; English tranl., Math. USSR Izv. **10** (1976), 1261–1305. MR **56:**16690

[67] Pesin Ya. B. *Characteristic Lyapunov exponents, and smooth ergodic theory,* Uspehi Mat. Nauk **32** (1977) 55–112; English transl., Russian Math. Surveys **32** (1977), no. 4, 55–114. MR **57:**6687

[68] Pesin Ya. B. *General theory of smooth hyperbolic dynamical systems,* In Encyclopedia Math. Sci., Dynamical Systems–2. Ed. Ya. G. Sinai, Springer-Verlag, Berlin, 1989. MR **91i:**58079

[69] Pesin Ya. B. & Sinai Ya. G. *Gibbs measures for partially hyperbolic attractors,* Erg. Th. & Dyn. Sys. **2** (1982) 417–438. MR **85f:**58071

[70] Philipp W. *Some metric theorems in number theory,* Pacific J. Math **20** (1967) 109-127. MR **34:**5755

[71] Philipp W. & Stout W. *Almost sure invariance principles for partial sums of weakly dependent random variables,* Mem. AMS **161** (1975). MR **55:**6570

[72] Pitskel B. S. *Poisson Limit Law for Markov chains,* Erg. Th. & Dyn. Sys. **11** (1991) 501–513. MR **92j:**60081

[73] Pugh C. C.& Shub M *Ergodicity of Anosov actions,* Inv. Math **15** (1972) 1–23. MR **45:**4456

[74] Pugh C. C. & Shub M. *Ergodic attractors,* Trans. AMS **312** (1989) 1-54. MR **90h:**58057

[75] Pugh C. C. & Shub M.*Stably ergodic dynamical systems and partial hyperbolicity,* J. Complexity **13** (1997) 125-179. MR **98e:**58110

[76] Pugh C. C. & Shub M. *Stable ergodicity and Julienne quasi-conformality*, J. European Math. Soc. **2** (2000) 1–52. MR **2003j:**37046

[77] Pugh C. C., Shub M. & Wilkinson A. *Hölder foliations,* Duke Math. J. **86** (1997) 517-546. MR **97m:**58155

[78] Ratner M. *The central limit theorem for geodesic flows on n–dimensional manifolds of negative curvature,* Israel J. Math. **15** (1973) 181-197. MR **48:**11446

[79] Sevastjanov B. A. *A Poisson Limit Law in a scheme of sums of dependant random variables,* Teor. Ver., Primenen. **17** (1972) 733–738; English transl. in Theory Probab. Appl. **17** (1972). MR **46:**10041

[80] Shub M. & Wilkinson A. *Pathological foliations and removable zero exponents,* Inv. Math. **139** (2000) 495–508. MR **2001c:**37030

[81] Sinai Ya. G. *Dynamic systems with countably-multiple Lebesgue spectrum. II,* Izv. Akad. Nauk SSSR Ser. Mat. **30** (1966) 15-68; English transl., Amer. Math. Soc. Transl. (2) **68** (1968), 34–88. MR **33:**5847

[82] Sinai Ya. G. *Markov partitions and U-diffeomorphisms,* Funkts. Anal. i Prilozhen. **2** (1968) no. 1 64–89; English transl. in Funct. Anal. Appl. **2** (1968). MR **38:**1361

[83] Sinai Ya. G. *Gibbs measures in ergodic theory,* Uspehi Mat. Nauk **27** (1972) 21-64; English transl., Russian Math. Surveys **27** (1972), no. 4, 21–70. MR **53:**3265

[84] Sinai Ya. G. *Stochasticity of dynamical systems,* Nonlinear Waves (Materials All-Union School, Gorki, 1977; A. V. Gaponov-Grekhov, editor), "Nauka", Moscow, 1979, pp. 192–212; English transl., Selecta Math. Sovietica **1** (1981), 101–119. MR **83e:**58058

[85] Starkov A. N. *New progress in the theory of homogeneous flows,* Russian Math. Surveys **52** (1997) 721–818. MR **98k:**22047
[86] Stroock D. W. & Varadhan S. *Multidimensional diffusion processes,* (1979) Springer, Berlin–New York. MR **91f:**60108
[87] Sullivan D. *Disjoint spheres, approximation by imaginary quadratic numbers, and the logarithm law for geodesics,* Acta Math. **149** (1982) 215–237. MR **84j:**58097
[88] Viana M. *Stochastic dynamics of deterministic systems,* Lecture Notes XXI Bras. Math. Colloq. IMPA, Rio de Janeiro, 1997.
[89] Viana M. *Multidimensional nonhyperbolic attractors,* Publ. IHES **85** (1997), 63–96. MR **98j:**58073
[90] Wilkinson A. *Stable ergodicity of the time–one map of a geodesic flow,* Erg. Th. & Dyn. Sys. **18** (1998) 1545–1587. MR **99m:**58129
[91] Young L.–S. *Statistical properties of dynamical systems with some hyperbolicity*, Ann. Math. **147** (1998) 585–650. MR **99h:**58140
[92] Young L.–S. *Recurrence times and rates of mixing*, Israel J. Math. **110** (1999) 153–188. MR **2001j:**37062

Department of Mathematics and Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742
E-mail address: dmitry@math.umd.edu