Ergod. Th. & Dynam. Sys. (2004), **24**, 847–871
DOI: 10.1017/S014338570400001X
© *2004 Cambridge University Press*
Printed in the United Kingdom

Averaging principle for fully coupled dynamical systems and large deviations

YURI KIFER
Institute of Mathematics, Hebrew University, Jerusalem 91904, Israel
(e-mail: kifer@math.huji.ac.il)

(*Received* 10 *January* 2003 *and accepted in revised form* 1 *March* 2003)

Abstract. In the study of systems which combine slow and fast motions, the averaging principle suggests that a good approximation of the slow motion on long time intervals can be obtained by averaging its parameters over the fast variables. When the slow and fast motions depend on each other (fully coupled), as is usually the case, for instance, in perturbations of Hamiltonian systems, the averaging prescription cannot always be applied, and when it does work, this is usually only in some averaged with respect to initial conditions sense. In this paper we first give necessary and sufficient conditions for the averaging principle to hold (in the above sense) and then, relying on some large deviations arguments, verify them in the case when the fast motions are hyperbolic (Axiom A) flows for each freezed slow variable. It turns out that in this situation the Lebesgue measure of initial conditions with bad averaging approximation tends to zero exponentially fast as the parameter tends to zero.

1. *Introduction*

The evolution of many real systems can be viewed as a combination of slow and fast motions which leads to complicated double scale equations that cannot usually be directly analyzed. In the 19th century, in applications to celestial mechanics, it was already well understood (although without rigorous justification) that a good approximation of the slow motion can be obtained by averaging its parameters over the fast variables. Later, averaging methods were applied in signal processing and, rather recently, to model climate–weather interactions (see [**13**] and [**20**]). The classical setup of averaging, justified rigorously in [**7**], presumes that the fast motion does not depend on the slow motion, and most of the work on averaging treats this case only. On the other hand, in real systems the slow and fast motions depend on each other, which leads to the more difficult fully coupled case which we study in this paper. This situation emerges, in particular, in typical perturbations of Hamiltonian systems, where the motion on constant energy manifolds is fast but the motion across them is slow.

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

Namely, we consider a system of differential equations for $X^\varepsilon = X^\varepsilon_{x,y}$ and $Y^\varepsilon = Y^\varepsilon_{x,y}$,

$$\frac{dX^\varepsilon(t)}{dt} = \varepsilon B(X^\varepsilon(t), Y^\varepsilon(t), \varepsilon), \quad \frac{dY^\varepsilon(t)}{dt} = b(X^\varepsilon(t), Y^\varepsilon(t), \varepsilon) \tag{1.1}$$

with initial conditions $X^\varepsilon(0) = x$, $Y^\varepsilon(0) = y$ on the product $\mathbb{R}^d \times \mathbf{M}$, where $\mathbf{M}$ is a compact n-dimensional Riemannian manifold and $B(x, y, \varepsilon)$, $b(x, y, \varepsilon)$ are vector fields on $\mathbb{R}^d$ and $\mathbf{M}$, respectively, which are bounded and smooth in x, y and Lipshitz continuous in ε together with their derivatives in x, y. The solutions of (1.1) determine the flow of diffeomorphisms Φ^t_ε on $\mathbb{R}^d \times \mathbf{M}$ acting by $\Phi^t_\varepsilon(x, y) = (X^\varepsilon_{x,y}(t), Y^\varepsilon_{x,y}(t))$. Taking $\varepsilon = 0$ we arrive at the flow $\Phi^t = \Phi^t_0$ acting by $\Phi^t_0(x, y) = (x, F^t_x y)$, where F^t_x is another family of flows given by $F^t_x y = Y_{x,y}(t)$, with $Y = Y_{x,y} = Y^0_{x,y}$ being the solution of

$$\frac{dY(t)}{dt} = b(x, Y(t)), \quad Y(0) = y \tag{1.2}$$

and $b(x, y) = b(x, y, 0)$. It is natural to view the flow Φ^t as describing an idealized physical system where the parameters $x = (x_1, \dots, x_d)$ are assumed to be constant, while the perturbed flow Φ^t_ε is regarded as describing a real system where evolution of these parameters is also taken into consideration.

Set $B(x, y) = B(x, y, 0)$ and assume that the limit

$$\bar B(x) = \bar B_y(x) = \lim_{T\to\infty} T^{-1} \int_0^T B(x, F^t_x y)\, dt \tag{1.3}$$

exists and is the same for 'many' y. Namely, let μ_x be an ergodic invariant measure of the flow F^t_x, then the limit (1.3) exists for μ_x-almost all y and is equal to

$$\bar B(x) = \bar B_{\mu_x}(x) = \int B(x, y)\, d\mu_x(y). \tag{1.4}$$

If $b(x, y)$ does not, in fact, depend on x, then $F^t_x = F^t$ and $\mu_x = \mu$ are also independent of x and we arrive at the classical setup. In this case, Lipschitz continuity of B implies already that $\bar B(x)$ is also Lipshitz continuous in x, and so there exists a unique solution $\bar X = \bar X_x$ of the averaged equation

$$\frac{d\bar X^\varepsilon(t)}{dt} = \varepsilon \bar B(\bar X^\varepsilon(t)), \quad \bar X^\varepsilon(0) = x. \tag{1.5}$$

Then the classical averaging principle states (see [**29**]) that

$$\lim_{\varepsilon\to 0} \sup_{0\le t\le T/\varepsilon} |X^\varepsilon_{x,y}(t) - \bar X^\varepsilon_x(t)| = 0 \tag{1.6}$$

for all x and for those y for which (1.3) holds true. In the general case (1.1), the averaged vector field $\bar B(x)$ in (1.3) and (1.4) may even not be continuous in x, let alone Lipschitz, and so (1.5) may have many solutions or none at all. Moreover, in general, there exists no natural well dependent on $x \in \mathbb{R}^d$ family of invariant measures μ_x, since dynamical systems F^t_x may have rather different properties for different x. Even when all measures μ_x are the same, the averaging principle often does not hold in the form (1.6), for instance, in the presence of resonances (see [**24**]). Thus even basic results on the approximation of

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

the slow motion by the averaged motion in the fully coupled case should be formulated in a different way and requires stronger and more specific assumptions.

If the convergence in (1.3) is uniform in x and y, then (see, for instance, [**20**]) any limit point $\bar{Z}(t) = \bar{Z}_x(t)$ as $\varepsilon \to 0$ of $Z^\varepsilon_{x,y}(t) = X^\varepsilon_{x,y}(t/\varepsilon)$ is a solution of the averaged equation

$$\frac{d\bar{Z}(t)}{dt} = \bar{B}(\bar{Z}(t)), \quad \bar{Z}(0) = x. \tag{1.7}$$

It is known that the limit in (1.3) is uniform in y if and only if the flow F^t_x on $\mathbf{M}$ is uniquely ergodic, i.e. it possesses a unique invariant measure, which occurs rather rarely. If, in addition, uniform convergence in y is also required, we arrive at very few specific models of uniquely ergodic families. Thus, the uniform convergence in (1.3) assumption is too restrictive and excludes many interesting cases. The first relatively general result on fully coupled averaging is probably due to Anosov [**1**] (see also [**24**] and [**20**]). Relying on the Liouville theorem, he showed that if each flow F^t_x preserves a probability measure μ_x on $\mathbf{M}$ having a C^1 dependent on x density with respect to the Riemannian volume m on $\mathbf{M}$, then for any $\delta > 0$,

$$\operatorname{mes}\left\{(x, y) : \sup_{0\le t\le T/\varepsilon} |X^\varepsilon_{x,y}(t) - \bar{X}^\varepsilon_x(t)| > \delta\right\} \to 0 \quad \text{as } \varepsilon \to 0, \tag{1.8}$$

where mes is the product of m and the Lebesgue measure in a relatively compact domain $\mathcal{X} \subset \mathbb{R}^d$. An improvement of (1.8) to (1.6) is not possible in a general fully coupled situation (see the discussion after Theorem 2.1).

Anosov's theorem is tailor-made for perturbations of Hamiltonian systems, which are a classical object for applications of averaging methods. In this case the whole space is fibered into manifolds of constant energy and each unperturbed flow F^t_x moves on one of these manifolds preserving the Liouville measure there. This suggests considering equation (1.1) on a (locally trivial) fiber bundle $\mathcal{M} = \{(x, y) : x \in U, y \in M_x\}$ with a base U being an open subset in a Riemannian manifold N and fibers M_x being diffeomorphic compact Riemannian manifolds (see [**30**]). On the other hand, $\mathcal{M}$ has a local product structure and if $\|B\|$ is bounded then the slow motion stays in one chart during time intervals of order Δ/ε with Δ small enough. Hence, studying the behavior of the solutions of (1.1) on each such time interval separately, we come back to the product space $\mathbb{R}^d \times \mathbf{M}$ setup and will only have to piece results together to see the picture on a larger time interval of length T/ε.

Observe that Anosov's theorem does not cover other situations where the averaging principle is known to work, for instance when the fast motion $F^t_x = F^t$ does not depend on the slow variable x, where no assumptions on invariant measures except ergodicity are required. So it is useful to have necessary and sufficient conditions which ensure that for given probability measures ν on $\mathbb{R}^d$ and μ_x on $\mathbf{M}$,

$$\int_{\mathcal{X}} \int_{\mathbf{M}} \sup_{0\le t\le T/\varepsilon} |X^\varepsilon_{x,y}(t) - \bar{X}^\varepsilon_x(t)|\, d\mu_x(y)\, d\nu(x) \to 0 \quad \text{as } \varepsilon \to 0 \tag{1.9}$$

and in §3 we derive general conditions of this kind which can be verified in specific situations. In addition to the known cases described above, relying on a large deviations argument we prove that these conditions hold true when the fast motions are hyperbolic flows.

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

This setup may also occur for Hamiltonian systems when manifolds of constant energy are compact and have negative curvature and the fast motions turn out to be geodesic flows there. More specifically we consider the situation where $b(x, y)$ is C^2 in x and y and for each x in a closure of a relatively compact domain $\mathcal{X}$ the flow F_x^t is Anosov or, more generally, Axiom A in a neighborhood of an attractor Λ_x. Let μ_x^{SRB} be the Sinai–Ruelle–Bowen (SRB) invariant measure of F_x^t on Λ (see §4) and set $\bar{B}(x) = \int B(x, y)\, d\mu_x^{\mathrm{SRB}}(y)$. It is known (see [**9**]) that the vector field $\bar{B}(x)$ is Lipschitz continuous in x, and so the averaged equations (1.5) and (1.7) (with right-hand side independent of y) have unique solutions $\bar{X}^\varepsilon(t)$ and $\bar{Z}(t) = \bar{X}^\varepsilon(t/\varepsilon)$. Still, in general, the measures μ_x^{SRB} are singular with respect to the Riemannian volume on $\mathbf{M}$ and so the method of [**1**] cannot be applied here. We check that our general conditions are satisfied in this case, and so (1.9) holds both with $\mu_x = \mu_x^{\mathrm{SRB}}$ and with μ_x being the Riemannian volume on $\mathbf{M}$. Moreover, we prove that in this situation the convergence in (1.8) to zero is exponentially fast in ε. Observe that the peculiarity of the fully coupled situation does not allow us to work with just one hyperbolic flow, but requires considering continuously changing fast motion which did not appear previously in the hyperbolic setup and calls for a special technique. In particular, the full flow Φ_ε^t on $\mathbb{R}^d \times \mathbf{M}$ defined above and viewed as a small perturbation of the partially hyperbolic system Φ_0^t plays an important role in our considerations. The corresponding problems for averaging in the discrete time case with difference equations in place of (1.1) were studied in [**21**]. Non-resonant and resonant unperturbed fast motions which are toral rotations differ from each other by having a unique and many invariant measures, respectively, and non-resonant directions occur 'much more frequently' than resonant ones. When unperturbed motions are Axiom A flows they all have an abundance of invariant measures, and so all of these flows may be viewed as 'resonant' in some sense, but the averaging principle in the form (1.9) still takes place in this case.

Once the convergence of $Z_{x,y}^\varepsilon(t) = X_{x,y}^\varepsilon(t/\varepsilon)$ to $\bar{Z}_x(t) = \bar{X}_x^\varepsilon(t/\varepsilon)$ as $\varepsilon \to 0$ is established, it is interesting to study the asymptotic behavior of the normalized error

$$V_{x,y}^{\varepsilon,\theta}(t) = \varepsilon^{\theta-1}(Z_{x,y}^\varepsilon(t) - \bar{Z}_x(t)), \quad \theta \in [\tfrac{1}{2}, 1] \tag{1.10}$$

where the cases $\theta = 1$, $\theta \in (\frac{1}{2}, 1)$ and $\theta = \frac{1}{2}$ correspond to large, moderate and Gaussian deviations, respectively. When all dynamical systems F_x^t are the same Axiom A flows, such results were obtained previously in [**18**] and [**19**], but in the fully coupled situation the proof of this kind of assertion encounters certain difficulties. A series of related results for the case when ordinary differential equations in (1.1) are replaced by fully coupled stochastic differential equations appeared in [**16, 26, 31**] and [**32**] (although it seems that [**32**] contains a serious gap in the proof). When the fast process does not depend on the slow one, such results were obtained in [**15**] and [**11**]. The methods there are quite different from ours and they are heavily based firstly on the Markov property of processes emerging there, and secondly on the uniformity and non-degeneracy of the fast diffusion term, assumptions which cannot be satisfied in our circumstances as our fast motion is very degenerate from this point of view. Recently, a series of results on Cramer's type asymptotics in fully coupled averaging for certain classes of fast motions (diffusions, Markov chains, Axion A diffeomorphisms) appeared in [**4**] and [**5**], but the methods there do not work for the Axiom A flows considered here.

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

The study of deviations from the averaged motion in the fully coupled case seems to be quite important for applications, especially from the phenomenological point of view. In addition to the perturbations of Hamiltonian systems mentioned above, there are many non-Hamiltonian systems which can naturally be considered from the beginning as a combination of fast and slow motions. For instance, Hasselmann [**13**] based his model of weather–climate interaction on the assumption that weather is a fast chaotic motion depending on climate as a slow motion which differs from the corresponding averaged motion mainly by a small diffusion term. Of course, it is hard to believe that real world chaotic systems can be described precisely by an Anosov or Axiom A flow, but one may take comfort in the Chaotic Hypothesis [**12**]: 'A chaotic mechanical system can be regarded for practical purposes as a topologically mixing Anosov system'.

2. *Main results*

We consider a system of differential equations (1.1) on the product $\bar{\mathcal{X}} \times \mathbf{M}$ where $\mathcal{X} \subset \mathbb{R}^d$ is an open set, $\bar{\mathcal{X}}$ is its closure and $\mathbf{M}$ is a compact C^2 Riemannian manifold. We assume that there exists $L > 0$ such that for all $\varepsilon \geq 0$, $x, z \in \bar{\mathcal{X}}$ and $y, v \in \mathbf{M}$,

$$\|B(x, y, \varepsilon) - B(z, v)\| + \|b(x, y, \varepsilon) - b(z, v)\| \leq L(\varepsilon + |x - z| + d_{\mathbf{M}}(y, v))$$
$$\text{and} \quad \|B(x, y)\| + \|b(x, y)\| \leq L, \tag{2.1}$$

where $d_{\mathbf{M}}$ is the distance on $\mathbf{M}$. Together with (1.1) we also consider equation (1.5) on $\bar{\mathcal{X}}$ with coefficients $\bar{B}$ such that there exists $\bar{L} > 0$ such that for all $x, z \in \bar{\mathcal{X}}$,

$$\|\bar{B}(x) - \bar{B}(z)\| \leq \bar{L}|x - z| \quad \text{and} \quad \|\bar{B}(x)\| \leq \bar{L}. \tag{2.2}$$

The Lipschitz continuity conditions (2) and (2.2) ensure the existence and uniqueness of solutions of (1.1) and (1.5), respectively. If $\bar{B}$ is defined by (1.4), then (2.2) is equivalent to the existence of $\tilde{L} > 0$ such that for all $x, z \in \mathcal{X}$,

$$\left| \int_{\mathbf{M}} B(x, y)\, d(\mu_x - \mu_z)(y) \right| \leq \tilde{L}|x - z|, \tag{2.3}$$

which is a condition of regular dependence of μ_x on x. Set $\mathcal{X}_t = \{x \in \mathcal{X} : X^{\varepsilon}_{x,y}(s) \in \mathcal{X}, \bar{X}^{\varepsilon}_x(s) \in \mathcal{X}$ for all $y \in \mathbf{M}$ and $s \in [0, t/\varepsilon]\}$. It is clear that $\mathcal{X}_t$ is an open set and by (2) and (2.2) it follows that $\mathcal{X}_t \supset \{x \in \mathcal{X} : \inf_{z \notin \mathcal{X}} |z - x| > 2t \max(L, \bar{L})\}$. Introduce

$$E_{\varepsilon}(t, \delta) = \left\{ (x, y) \in \mathcal{X}_t \times \mathbf{M} : \left| \frac{1}{t} \int_0^t B(x, Y^{\varepsilon}_{x,y}(u))\, du - \bar{B}(x) \right| > \delta \right\}.$$

In the next section we derive the following result.

THEOREM 2.1. *Suppose that (2) and (2.2) hold and let μ be a probability measure on $\mathcal{X} \times \mathbf{M}$. Then*

$$\lim_{\varepsilon \to 0} \int_{\mathcal{X}_T} \int_{\mathbf{M}} \sup_{0 \leq t \leq T/\varepsilon} |X^{\varepsilon}_{x,y}(t) - \bar{X}^{\varepsilon}_x(t)|\, d\mu(x, y) = 0 \tag{2.4}$$

if and only if there exists an integer valued function $n = n(\varepsilon) \to \infty$ as $\varepsilon \to 0$ such that for any $\delta > 0$,

$$\lim_{\varepsilon \to 0} \max_{0 \leq j < n(\varepsilon)} \mu\{(\mathcal{X}_T \times \mathbf{M}) \cap \Phi_{\varepsilon}^{-jt(\varepsilon)} E_{\varepsilon}(t(\varepsilon), \delta)\} = 0, \tag{2.5}$$

where $t(\varepsilon) = T/\varepsilon n(\varepsilon)$ and, recall, $\Phi^t_{\varepsilon}(x, y) = (X^{\varepsilon}_{x,y}(t), Y^{\varepsilon}_{x,y}(t))$.

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

Taking into account that $Y^{\varepsilon}_{x,y}(t)$ and $Y^{0}_{x,y}(t)$ stay close during the time $t \le t(\varepsilon)$ with $t(\varepsilon)$ much smaller than $\log(1/\varepsilon)$, we obtain a sufficient condition for (2.4) in the form of (2.5) with $E_0(\cdot,\cdot)$ in place of $E_\varepsilon(\cdot,\cdot)$. In addition to the proof of Theorem 2.1, we show in the next section how to check (2.5) in two situations where (2.4) was known before, namely when the fast motion $Y^{\varepsilon}_{x,y}$ does not depend on the slow motion $X^{\varepsilon}_{x,y}$ and in the Anosov theorem setup. The latter requires ν to have a bounded C^1 density with respect to the Lebesgue measure on $\mathbb{R}^d$ and μ_x, $x \in \mathcal{X}$ to be invariant measures of the corresponding unperturbed flows F^t_x so that μ_x is ergodic for ν-almost all (a.a.) x and for each $x \in \mathcal{X}$ the measure μ_x has a density $q_x = q_x(y) > 0$ with respect to the Riemannian volume on $\mathbf{M}$ which is C^1 in both x and y. We also exhibit a simple example of a system satisfying (2)–(2.3) but not (2.5), and so the convergence (2.4) does not hold there. In fact, we take fast motions there independent of the slow one but with μ having the form $d\mu(x,y) = d\mu_x(y)\,dx$.

Theorem 2.1 gives conditions for convergence in average in the averaging principle. In view of resonances (see, for instance, [**24**]) it is impossible for many interesting examples to ensure (1.6) for all $x \in \mathcal{X}$ and μ_x-a.a y, where μ_x is a reasonable family of probability measures on $\mathbf{M}$. Moreover, Neistadt explained to me that in Example 7 of [**3**], the convergence (1.6) does not hold true for any initial condition (from a large domain). There is a very restricted class of systems where (1.6) holds true for all $x \in \mathcal{X}$ and $y \in \mathbf{M}$. This happens, for instance, when Arnold's conditions for two-frequency systems are satisfied (see §3.5 in [**24**] and §5.1 in [**3**]). If the convergence in (1.3) is uniform in $x \in \mathcal{X}$ and $y \in \mathbf{M}$, then (1.6) takes place as well. It will be clear from our arguments in the next section (see (3.13)) that it suffices to assume a bit less, namely that for any $\delta > 0$ there exists ε_δ such that for any positive $\varepsilon \le \varepsilon_\delta$ one can find an integer valued function $n(\varepsilon) \to \infty$ as $\varepsilon \to \infty$ so that $E_\varepsilon(t(\varepsilon), \delta) = \emptyset$ where, again, $t(\varepsilon) = T(n(\varepsilon)\varepsilon)^{-1}$. Such conditions can only be satisfied for some families of uniquely ergodic dynamical systems such as flows on a circle and horocycle flows nicely depending on a parameter (slow variable).

Another situation where we are able to verify (2.5) is the case of fast motions being slowly changing (uniformly) hyperbolic flows where the averaging principle in the form (2.4) was not known before. First, we recall basic definitions; see [**8, 22**]. Let F^t be a C^2 flow on a compact Riemannian manifold $\mathbf{M}$ given by a differential equation

$$\frac{dF^t y}{dt} = b(F^t y), \quad F^0 y = y.$$

A compact F^t-invariant set $\Lambda \subset \mathbf{M}$ is called hyperbolic if there exists $\kappa > 0$ and a splitting $T_\Lambda \mathbf{M} = \Gamma^s \oplus \Gamma^0 \oplus \Gamma^u$ into continuous subbundles Γ^s, Γ^0, Γ^u of the tangent bundle $T\mathbf{M}$ restricted to Λ, such that the splitting is invariant with respect to the differential DF^t of F^t, Γ^0 is the one-dimensional subbundle generated by the vector field b, and there is $t_0 > 0$ such that for all $\xi \in \Gamma^s$, $\eta \in \Gamma^u$ and $t \ge t_0$,

$$\|DF^t \xi\| \le e^{-\kappa t}\|\xi\| \quad \text{and} \quad \|DF^{-t}\eta\| \le e^{-\kappa t}\|\eta\|. \tag{2.6}$$

A hyperbolic set Λ is said to be basic hyperbolic if the periodic orbits of $F^t|_\Lambda$ are dense in Λ, $F^t|_\Lambda$ is topologically transitive and there exists an open set $U \supset \Lambda$ with

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

$\Lambda = \bigcap_{-\infty<t<\infty} F^t U$. Such a Λ is called a basic hyperbolic attractor if for some open set U and $t_0 > 0$,

$$F^{t_0}\bar{U} \subset U \quad \text{and} \quad \bigcap_{t>0} F^t U = \Lambda,$$

where $\bar{U}$ denotes the closure of U. If $\Lambda = \mathbf{M}$ then F^t is called an Anosov flow.

ASSUMPTION 2.2. *The family $b(x,\cdot)$ in (1.2) consists of C^2 vector fields on an n-dimensional Riemannian manifold $\mathbf{M}$ with uniform C^2 dependence on the parameter x belonging to a relatively compact connected open set $\mathcal{X}$ and depending continuously on x in its closure $\bar{\mathcal{X}}$. Each flow F_x^t, $x \in \bar{\mathcal{X}}$ on $\mathbf{M}$ given by*

$$\frac{dF_x^t y}{dt} = b(x, F_x^t y), \qquad F_x^0 y = y \tag{2.7}$$

possesses a basic hyperbolic attractor Λ_x with a splitting $T_{\Lambda_x}\mathbf{M} = \Gamma_x^s \oplus \Gamma_x^0 \oplus \Gamma_x^u$ satisfying (2.6) with the same $\kappa > 0$, and there exists an open set $\mathcal{W} \subset \mathbf{M}$ and $t_0 > 0$ such that

$$\Lambda_x \subset \mathcal{W}, \quad F_x^t \bar{\mathcal{W}} \subset \mathcal{W}\, \forall t \geq t_0 \quad \textit{and} \bigcap_{t>0} F_x^t \mathcal{W} = \Lambda_x \quad \forall x \in \bar{\mathcal{X}}. \tag{2.8}$$

Let $J_x^u(t,y)$ be the Jacobian of the linear map $DF_x^t(y) : \Gamma_x^u(y) \to \Gamma_x^u(F_x^t y)$ with respect to the Riemannian inner products and set

$$\varphi_x^u(y) = -\frac{dJ_x^u(t,y)}{dt}\bigg|_{t=0}. \tag{2.9}$$

The function $\varphi_x^u(y)$ is known to be Hölder continuous in y, since the subbundles Γ_x^u are Hölder continuous (see [**8**] and [**22**]), and $\varphi_x^u(y)$ is C^1 in x (see [**9**]). Let $\mathcal{W}$ satisfy (2.8) and set $\mathcal{W}_x^t = \{y \in \mathcal{W} : F_x^s y \in \bar{\mathcal{W}},\ \forall s \in [0,t]\}$. A set $E \subset \mathcal{W}_x^t$ is called (δ,t)-separated for the flow F_x if $y,z \in E$, $y \neq z$ imply $d(F_x^s y, F_x^s z) > \delta$ for some $s \in [0,t]$, where $d(\cdot,\cdot)$ is the distance function on $\mathbf{M}$. For each continuous function ψ on $\mathcal{W}$ set $P_x(\psi,\delta,t) = \sup\big\{\sum_{y\in E} \exp\int_0^t \psi(F_x^s y)\,ds : E \subset \mathcal{W}_x^t \text{ is } (\delta,t)\text{-separated for } F_x\big\}$, $P_x(\psi,\delta,t) = 0$ if $\mathcal{W}_x^t = \emptyset$, and

$$P_x(\psi,\delta) = \limsup_{t\to\infty} \frac{1}{t} \log P_x(\psi,\delta,t).$$

The latter is monotone in δ, and so the limit

$$P_x(\psi) = \lim_{\delta\to 0} P_x(\psi,\delta)$$

exists and is called the topological pressure of ψ for the flow F_x^t. Let $\mathcal{M}_x$ denote the space of F_x^t-invariant probability measures on Λ_x, then (see, for instance, [**22**]) the variational principle

$$P_x(\psi) = \sup_{\mu\in\mathcal{M}_x} \left(\int \psi\, d\mu + h_\mu(F_x^1) \right) \tag{2.10}$$

holds, where $h_\mu(F_x^1)$ is the Kolmogorov–Sinai entropy of the time-one map F_x^1 with respect to μ. If q is a Hölder continuous function on Λ_x then there exists a unique F_x^t-invariant measure μ_x^q on Λ_x, called the equilibrium state for $\varphi_x^u + q$, such that

$$P_x(\varphi_x^u + q) = \int (\varphi_x^u + q)\, d\mu_x^q + h_{\mu_x^q}(F_x^1). \tag{2.11}$$

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

We denote μ_x^0 by $\mu_x^{\rm SRB}$ since it is usually called the SRB measure. Since Λ_x are attractors, we have that $P_x(\varphi_x^u) = 0$ (see [**8**]). We now replace the condition (2) by the following.

ASSUMPTION 2.3. *There exist $L, \varepsilon_0 > 0$ such that for all $x \in \bar{\mathcal{X}}$, $y \in \mathbf{M}$, and $\varepsilon \in [0, \varepsilon_0)$,*

$$\|B(x, y, \varepsilon)\|_{C^1([0,\varepsilon_0)\times\bar{\mathcal{X}}\times\mathbf{M})} + \|b(x, y, \varepsilon)\|_{C^2([0,\varepsilon_0)\times\bar{\mathcal{X}}\times\mathbf{M})} \le L, \tag{2.12}$$

where $\|\cdot\|_{C^i([0,\varepsilon_0)\times\bar{\mathcal{X}}\times\mathbf{M})}$ is the C^i norm of the corresponding vector fields on $[0, \varepsilon_0) \times \bar{\mathcal{X}} \times \mathbf{M}$ (taking only the right derivative in ε at 0).

Set

$$\bar{B}(x) = \int B(x, y)\, d\mu_x^{\rm SRB}(y). \tag{2.13}$$

Then under Assumption 2.3, $\bar{B}$ is C^1 in x (see [**9**]) and so (2.2) is automatically satisfied.

THEOREM 2.4. *Suppose that Assumptions 2.2 and 2.3 hold true. Define $\bar{B}$ by (2.13) and let μ be the product of a probability measure ν with support in $\mathcal{X}_T$ and the normalized Riemannian volume $m_{\mathcal{W}}$ on $\mathcal{W}$. Then (2.5) is satisfied with $t(\varepsilon) = T/\varepsilon n(\varepsilon)$ whenever both $t(\varepsilon) \to \infty$ and $n(\varepsilon) \to \infty$ as $\varepsilon \to 0$, and so (2.4) holds true. Moreover, for any $a > 0$ there exist $c > 0$ and ε_a such that*

$$\mu\left\{(x, y) \in \mathcal{X}_T \times \mathcal{W} : \sup_{0\le t\le T/\varepsilon} |X^\varepsilon_{x,y}(t) - \bar{X}^\varepsilon_x(t)| > a\right\} \le e^{-c/\varepsilon}, \tag{2.14}$$

provided $\varepsilon \le \varepsilon_a$. The result remains true if in place of the above we take μ defined by $d\mu(x, y) = d\nu(x)\, d\mu_x^{\rm SRB}(y)$.

Observe that we can take, in particular, ν to be the Dirac measure (unit mass) at a point $x \in \mathcal{X}$, i.e. (2.4) follows here without integration in x and (2.14) can be replaced by

$$\mu_x\left\{y \in \mathcal{W} : \sup_{0\le t\le T/\varepsilon} |X^\varepsilon_{x,y}(t) - \bar{X}^\varepsilon_x(t)| > a\right\} \le e^{-c/\varepsilon} \tag{2.15}$$

for either $\mu_x = m_{\mathcal{W}}$ or $\mu_x = \mu_x^{\rm SRB}$. It is still not clear whether in the setup of Theorem 2.4 it is possible to derive the convergence (1.6) for all (or for Lebesgue almost all) $x \in \mathcal{X}_T$ and for $m_{\mathcal{W}}$-a.a. $y \in \mathcal{W}$ and not just convergence in average (2.4) or in measure (2.14) and (2.15).

3. *General conditions of averaging*

We start proving Theorem 2.1 by first deriving that (2.4) implies (2.5). Set $\rho^\varepsilon_{x,y}(t) = X^\varepsilon_{x,y}(t) - \bar{X}^\varepsilon_x(t)$, $\rho^\varepsilon_{x,y} = \sup_{0\le t\le T/\varepsilon} |\rho^\varepsilon_{x,y}(t)|$ and so by (2.4),

$$\rho^\varepsilon = \int_{\mathcal{X}_T}\int_{\mathbf{M}} \rho^\varepsilon_{x,y}\, d\mu(x, y) \to 0 \quad \text{as } \varepsilon \to 0. \tag{3.1}$$

Take $n = n(\varepsilon) \in \mathbb{N}$ which will be specified later and set $t(\varepsilon) = T(n\varepsilon)^{-1}$. We have

$$X^\varepsilon_{x,y}(kt(\varepsilon)) - \bar{X}^\varepsilon_{x,y}(kt(\varepsilon)) = \sum_{j=0}^{k-1} R^\varepsilon_{x,y}(jt(\varepsilon), (j+1)t(\varepsilon), \varepsilon),$$

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

where

$$R^{\varepsilon}_{x,y}(s,t,\varepsilon) = \varepsilon \int_s^t (B(X^{\varepsilon}_{x,y}(u), Y^{\varepsilon}_{x,y}(u), \varepsilon) - \bar{B}(\bar{X}^{\varepsilon}_{x}(u)))\, du = \rho^{\varepsilon}_{x,y}(t) - \rho^{\varepsilon}_{x,y}(s),$$

and so

$$|R^{\varepsilon}_{x,y}(s,t,\varepsilon)| \le 2\rho^{\varepsilon}_{x,y}. \tag{3.2}$$

Set $x^{\varepsilon}_k = X^{\varepsilon}_{x,y}(kt(\varepsilon))$, $y^{\varepsilon}_k = Y^{\varepsilon}_{x,y}(kt(\varepsilon))$ and $\bar{x}^{\varepsilon}_k = \bar{X}^{\varepsilon}_{x}(kt(\varepsilon))$. Then

$$R^{\varepsilon}_{x,y}(jt(\varepsilon), (j+1)t(\varepsilon), \varepsilon) = \varepsilon \int_0^{t(\varepsilon)} (B(X^{\varepsilon}_{x^{\varepsilon}_j, y^{\varepsilon}_j}(u), Y^{\varepsilon}_{x^{\varepsilon}_j, y^{\varepsilon}_j}(u), \varepsilon) - \bar{B}(\bar{X}^{\varepsilon}_{\bar{x}^{\varepsilon}_j}(u)))\, du. \tag{3.3}$$

Furthermore, by (2)

$$\begin{aligned} &\left| \int_0^{t(\varepsilon)} (B(X^{\varepsilon}_{x^{\varepsilon}_j, y^{\varepsilon}_j}(u), Y^{\varepsilon}_{x^{\varepsilon}_j, y^{\varepsilon}_j}(u), \varepsilon) - B(x^{\varepsilon}_j, Y^{\varepsilon}_{x^{\varepsilon}_j, y^{\varepsilon}_j}(u)))\, du \right| \\ &\qquad \le L \int_0^{t(\varepsilon)} |X^{\varepsilon}_{x^{\varepsilon}_j, y^{\varepsilon}_j}(u) - x^{\varepsilon}_j|\, du \le L\varepsilon t(\varepsilon)(Lt(\varepsilon) + 1). \end{aligned} \tag{3.4}$$

Similarly, by (1.4), (2) and (2.2),

$$\left| \int_0^{t(\varepsilon)} (\bar{B}(\bar{X}^{\varepsilon}_{\bar{x}^{\varepsilon}_j}(u) - \bar{B}(\bar{x}^{\varepsilon}_j)))\, du \right| \le \bar{L}L\varepsilon t^2(\varepsilon) \tag{3.5}$$

and

$$|\bar{B}(\bar{x}^{\varepsilon}_j) - \bar{B}(x^{\varepsilon}_j)| \le \bar{L}|\bar{x}^{\varepsilon}_j - x^{\varepsilon}_j| \le \bar{L}\rho^{\varepsilon}_{x,y}. \tag{3.6}$$

It follows from (3.3)–(3.6) that

$$\begin{aligned} &\left| R^{\varepsilon}_{x,y}(jt(\varepsilon), (j+1)t(\varepsilon), \varepsilon) - \varepsilon \int_0^{t(\varepsilon)} (B(x^{\varepsilon}_j, Y^{\varepsilon}_{x^{\varepsilon}_j, y^{\varepsilon}_j}(u)) - \bar{B}(x^{\varepsilon}_j))\, du \right| \\ &\qquad \le \varepsilon t(\varepsilon)(L^2\varepsilon t(\varepsilon) + L\varepsilon + \bar{L}L\varepsilon t(\varepsilon) + \bar{L}\rho^{\varepsilon}_{x,y}), \end{aligned} \tag{3.7}$$

and so by (3.2), taking into account that $\varepsilon t(\varepsilon) = T/n$, we obtain

$$\left| \frac{1}{t(\varepsilon)} \int_0^{t(\varepsilon)} B(x^{\varepsilon}_j, Y^{\varepsilon}_{x^{\varepsilon}_j, y^{\varepsilon}_j}(u))\, du - \bar{B}(x^{\varepsilon}_j) \right| \le L^2T/n + L\varepsilon + \bar{L}LT/n + (\bar{L} + 2n/T)\rho^{\varepsilon}_{x,y}. \tag{3.8}$$

Now choose an integer valued function $n = n(\varepsilon) \to \infty$ as $\varepsilon \to 0$ so that $n(\varepsilon)\rho^{\varepsilon} \to 0$ as $\varepsilon \to 0$. Next, given $\delta > 0$ choose $\varepsilon_\delta > 0$ such that for any $\varepsilon < \varepsilon_\delta$,

$$\frac{L^2T}{n(\varepsilon)} + L\varepsilon + \frac{\bar{L}LT}{n(\varepsilon)} \le \frac{1}{2}\delta.$$

It then follows from (3.8) that for such ε,

$$\begin{aligned} &(\mathcal{X}_T \times \mathbf{M}) \cap \Phi^{-jt(\varepsilon)}_{\varepsilon} E_{\varepsilon}(t(\varepsilon), \delta) \subset A_{\varepsilon}(\delta) \\ &\qquad = \left\{ (x,y) \in \mathcal{X}_T \times \mathbf{M} : \left(\bar{L} + \frac{2n(\varepsilon)}{T} \right) \rho^{\varepsilon}_{x,y} > \frac{1}{2}\delta \right\}. \end{aligned}$$

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

By Chebyshev's inequality,

$$\mu(A_\varepsilon(\delta)) \le \frac{2}{\delta}\left(\bar L + \frac{2n(\varepsilon)}{T}\right)\rho^\varepsilon. \tag{3.9}$$

The right-hand side of (3.9) tends to zero as $\varepsilon \to 0$ by our choice of $n(\varepsilon)$, completing the proof of the first part of Theorem 2.1.

Next, we prove that (2.5) yields (2.4). By (2) and (2.2),

$$\begin{aligned}|X^\varepsilon_{x,y}(t) - \bar X^\varepsilon_x(t)| &= \varepsilon\left|\int_0^t (B(X^\varepsilon_{x,y}(s), Y^\varepsilon_{x,y}(s), \varepsilon) - \bar B(\bar X^\varepsilon_x(s)))\,ds\right| \\ &\le \varepsilon\left|\int_0^t (B(X^\varepsilon_{x,y}(s), Y^\varepsilon_{x,y}(s)) - \bar B(X^\varepsilon_{x,y}(s)))\,ds\right| \\ &\quad + L\varepsilon^2 t + \bar L\varepsilon\int_0^t |X^\varepsilon_{x,y}(s) - \bar X^\varepsilon_x(s)|\,ds. \end{aligned}\tag{3.10}$$

This together with Gronwall's inequality gives

$$\begin{aligned}&\sup_{0\le t\le T/\varepsilon} |X^\varepsilon_{x,y}(t) - \bar X^\varepsilon_x(t)| \\ &\quad \le e^{\bar L T}\varepsilon\left(\sup_{0\le t\le T/\varepsilon}\left|\int_0^t (B(X^\varepsilon_{x,y}(s), Y^\varepsilon_{x,y}(s)) - \bar B(X^\varepsilon_{x,y}(s)))\,ds\right| + LT\right).\end{aligned}\tag{3.11}$$

Now, given $n(\varepsilon)$ and $t(\varepsilon)$ appearing in (2.5), using (3.4) and (3.5) we can write that

$$\begin{aligned}&\varepsilon\sup_{0\le t\le T/\varepsilon}\left|\int_0^t (B(X^\varepsilon_{x,y}(s), Y^\varepsilon_{x,y}(s)) - \bar B(X^\varepsilon_{x,y}(s)))\,ds\right| \\ &\quad \le 2L\varepsilon t(\varepsilon) + \varepsilon\sum_{j=0}^{n(\varepsilon)-1}\left|\int_{jt(\varepsilon)}^{(j+1)t(\varepsilon)} (B(X^\varepsilon_{x,y}(s), Y^\varepsilon_{x,y}(s)) - \bar B(X^\varepsilon_{x,y}(s)))\,ds\right| \\ &\quad = 2L\varepsilon t(\varepsilon) + \varepsilon\sum_{j=0}^{n(\varepsilon)-1}\left|\int_0^{t(\varepsilon)} (B(X^\varepsilon_{x_j^\varepsilon,y_j^\varepsilon}(s), Y^\varepsilon_{x_j^\varepsilon,y_j^\varepsilon}(s)) - \bar B(X^\varepsilon_{x_j^\varepsilon,y_j^\varepsilon}(s)))\,ds\right| \\ &\quad \le 2L\varepsilon t(\varepsilon) + \varepsilon^2 n(\varepsilon)t(\varepsilon)L(L+\bar L) \\ &\qquad + \varepsilon t(\varepsilon)\sum_{j=0}^{n(\varepsilon)-1}\left|\frac{1}{t(\varepsilon)}\int_0^{t(\varepsilon)} B(x_j^\varepsilon, Y^\varepsilon_{x_j^\varepsilon,y_j^\varepsilon}(s))\,ds - \bar B(x_j^\varepsilon)\right| \\ &\quad \le 2L\varepsilon t(\varepsilon) + \varepsilon^2 n(\varepsilon)t(\varepsilon)L(L+\bar L) + \varepsilon t(\varepsilon)n(\varepsilon)\delta + 2L\varepsilon t(\varepsilon)\sum_{j=0}^{n(\varepsilon)-1}\mathbb{I}_{E_\varepsilon(t(\varepsilon),\delta)}(x_j^\varepsilon, y_j^\varepsilon)\end{aligned}\tag{3.12}$$

where $\mathbb{I}_A(z) = 1$ if $z \in A$ and 0 otherwise. Integrating against μ both parts of (3.11) and (3.12) we obtain

$$\begin{aligned}&\int_{\mathcal{X}_T}\int_{\mathbf{M}}\sup_{0\le t\le T/\varepsilon}|X^\varepsilon_{x,y}(t) - \bar X^\varepsilon_x(t)|\,d\mu(x,y) \\ &\qquad \le e^{LT}\left(\frac{2LT}{n(\varepsilon)} + \varepsilon TL(L+\bar L) + T\delta + 2L\eta_\varepsilon(\delta)\right),\end{aligned}\tag{3.13}$$

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

where

$$\eta_\varepsilon(\delta) = \max_{0\le j\le n(\varepsilon)-1} \mu\{(\mathcal{X}_T \times \mathbf{M}) \cap \Phi_\varepsilon^{-jt(\varepsilon)} E_\varepsilon(t(\varepsilon), \delta)\}.$$

By (2.5), $\eta_\varepsilon(\delta) \to 0$ as $\varepsilon \to 0$, and so letting first $\varepsilon \to 0$ and then $\delta \to 0$ in (3.13), we obtain (2.4), completing the proof of Theorem 2.1. □

Observe that by (2),

$$\begin{aligned} d(Y^\varepsilon_{x,y}(t), Y_{x,y}(t)) &\le \int_0^t \|b(X^\varepsilon_{x,y}(s), Y^\varepsilon_{x,y}(s), \varepsilon) - b(x, Y_{x,y}(s))\|\, ds \\ &\le \varepsilon t L(L+1) + L\int_0^t d(Y^\varepsilon_{x,y}(s), Y_{x,y}(s))\, ds. \end{aligned} \tag{3.14}$$

Hence, by Gronwall's inequality,

$$d(Y^\varepsilon_{x,y}(t), Y_{x,y}(t)) \le \varepsilon t L(L+1)e^{Lt}. \tag{3.15}$$

Then for $t(\varepsilon) = o(\log(1/\varepsilon))$ and all ε small enough,

$$E_0(t(\varepsilon), \delta/2) \supset E_\varepsilon(t(\varepsilon), \delta), \tag{3.16}$$

where E_0 is defined by taking $Y_{x,y}$ in place of $Y^\varepsilon_{x,y}$ in the definition of E_ε, i.e.

$$E_0(t, \delta) = \left\{(x, y) \in \mathcal{X} \times \mathbf{M} : \left|\frac{1}{t}\int_0^t B(x, F^s_x y)\, ds - \bar{B}(x)\right| > \delta\right\}.$$

Thus we obtain the following.

COROLLARY 3.1. *Suppose that there exists an integer valued function $n = n(\varepsilon) \to \infty$ as $\varepsilon \to 0$ such that $t(\varepsilon) = T/\varepsilon n(\varepsilon) = o(\log(1/\varepsilon))$ and for any $\delta > 0$,*

$$\lim_{\varepsilon\to 0} \max_{0\le j<n(\varepsilon)} \mu((\mathcal{X}_T \times \mathbf{M}) \cap \Phi_\varepsilon^{-jt(\varepsilon)} E_0(t(\varepsilon), \delta)) = 0. \tag{3.17}$$

Then (2.5) is also satisfied and so (2.4) holds.

Next we will verify (3.17) in the simplest situation when the fast motion $Y^\varepsilon_{x,y}$ does not depend on the slow one, so that we can write $Y^\varepsilon_{x,y}(t) = F^t y$ for some flow F^t on $\mathbf{M}$, and $\mu = \nu \times \mu_0$, where ν is a probability measure on $\bar{\mathcal{X}}$ and μ_0 is an F^t-invariant ergodic probability measure on $\mathbf{M}$. In this case we can write

$$\mu\{(\mathcal{X}_T \times \mathbf{M}) \cap \Phi_\varepsilon^{-t} E_0(s, \delta)\} = \int_{\mathcal{X}_T} \mu_0(D_x)\, d\nu(x), \tag{3.18}$$

where $D_x = D_x(t, s, \delta) = \{y \in \mathbf{M} : (X^\varepsilon_{x,y}(t), F^t y) \in E_0(s, \delta)\}$. For each $z \in \mathcal{X}$ set $D_z(s, \delta) = \{y \in \mathbf{M} : (z, y) \in E_0(s, \delta)\}$. By (2) and (2.2), $D_w(s, \delta) \subset D_z(s, \delta/2)$, provided $|w - z| \le \frac{1}{2}\delta(L + \bar{L})^{-1}$. Hence,

$$\begin{aligned} D_{x,z} &= D_{x,z}(t, s, \delta) \\ &= \{y \in \mathbf{M} : |X^\varepsilon_{x,y}(t) - z| \le \tfrac{1}{2}\delta(L+\bar{L})^{-1} \text{ and } F^t y \in D_{X^\varepsilon_{x,y}(t)}(s, \delta)\} \\ &\subset F^{-1} D_z(s, \delta/2). \end{aligned}$$

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

By (2), $X^\varepsilon_{x,y}(t) \in \mathcal{X}$ if $x \in \mathcal{X}_T$ and $t \le T/\varepsilon$. Hence, if $z_1, \dots, z_l$ is a minimal $\frac{1}{2}\delta(L+\bar L)^{-1}$-net in $\bar{\mathcal{X}}$, then $X^\varepsilon_{x,y}(t), t \le T/\varepsilon$ belongs to a $\frac{1}{2}\delta(L+\bar L)^{-1}$-ball around some z_i, and so for any $x \in \mathcal{X}_T$,

$$D_x(t,s,\delta) \subset \bigcup_{i=1}^{l} D_{x,z_i}(t,s,\delta) \subset \bigcup_{i=1}^{l} F^{-1} D_{z_i}(s,\delta/2).$$

Since μ_0 is F^t-invariant and ergodic, we obtain from here that

$$\mu_0(D_x(t,s,\delta)) \le \sum_{i=1}^{l} \mu_0(E_{z_i}(s,\delta/2)) \to 0 \quad \text{as } s \to \infty$$

and (3.17) follows, taking $t = jt(\varepsilon)$ and $s = t(\varepsilon)$.

Next, we establish (3.17) in the Anosov theorem setup. Namely, suppose that $\mu_x(U) = \int_U q(x,y)\, dm_{\mathbf{M}}(y)$ for each $x \in \bar{\mathcal{X}}$ and any Borel set $U \subset \mathbf{M}$, where $q(x,y) > 0$ is differentiable in x and y. Assume also now that $\bar X$ is compact and for Lebesgue almost all $x \in \bar X$ and μ_x-a.a. y,

$$\bar B(x) = \lim_{t\to\infty} \frac{1}{t} \int_0^t B(x, F^s_x y)\, ds = \int B(x,y) q(x,y)\, d\mu_x(y). \tag{3.19}$$

Then using the Liouville theorem $\operatorname{div}_y(q(x,y)b(x,y)) = 0$ we derive in the same way as in [**1**] and [**20**, p. 186] that for any Borel set $V \subset \mathcal{X}_T \times \mathbf{M}$ and $t \in [0,T]$,

$$e^{-CT}\nu(V) \le \mu(\Phi^{t/\varepsilon}_\varepsilon V) \le e^{CT}\nu(V) \tag{3.20}$$

for some $C > 0$, where $d\mu(x,y) = q(x,y)\, d\mu_x(y)\, dx$. It follows from (3.19) and (3.20) that

$$\mu((\mathcal{X}_T \times \mathbf{M}) \cap \Phi^{-jt(\varepsilon)}_\varepsilon E_0(t(\varepsilon),\delta)) \le e^{CT}\mu(E_0(t(\varepsilon),\delta)) \to 0 \quad \text{as } \varepsilon \to 0,$$

and so (3.17) holds.

Next, we describe an example where (2.4) does not hold.

Example 3.2. Let $\mathbf{M}$ be the circle $\mathbb{T}^1$ of length 1, all F^t_x be identity maps, $\mu_x = \delta_x$ be the unit mass at $x \in \mathbb{R}^1$, and $B(x,y) = B(y)$ be a C^1 function on $\mathbb{T}^1$. The measures μ_x are trivially ergodic and $\bar B(x) = \int B(y)\, d\mu_x(y) = B(x)$ extended as a periodic function to $\mathbb{R}^1$ satisfies (2.2). Then $X^\varepsilon_{x,y}(t) = x + \varepsilon t B(y)$, $Y^\varepsilon_{x,y}(t) \equiv y$ and $\bar X^\varepsilon_x(t) = x + \varepsilon \int_0^t B(\bar X^\varepsilon_x(s))\, ds$. Hence,

$$\int_0^1 \int_{\mathbb{T}^1} \sup_{0\le t\le T/\varepsilon} |X^\varepsilon_{x,y}(t) - \bar X^\varepsilon_x(t)|\, d\mu_x(y)\, dx = \int_0^1 \sup_{0\le t\le T/\varepsilon} |\varepsilon t B(x) + x - \bar X^\varepsilon_x(t)|\, dx. \tag{3.21}$$

Take, for instance, $B(x) = \cos^2(2\pi x)$, then $\bar X^\varepsilon_x(t) = (2\pi)^{-1}\arctan(2\pi\varepsilon t + \tan(2\pi x))$ and making a change of variables, we obtain that the right-hand side of (3.21) is equal to

$$\int_0^1 \sup_{0\le t\le T} |t\cos^2(2\pi x) + x - (2\pi)^{-1}\arctan(2\pi t + \tan(2\pi x))|\, dx$$

and the latter does not depend on ε and is not zero, as is easy to see (i.e. (2.4) does not hold). It is easy to directly check that (2.5) is not satisfied as well. Similar examples can be constructed with all F^t_x coinciding with a rational rotation of, say, a two-dimensional torus, with the measures μ_x being the ergodic invariant (normalized one-dimensional Lebesgue) measures supported by corresponding invariant cycles.

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

4. *Dynamics of* Φ_ε^t

Any vector $\xi \in T(\mathbb{R}^d \times \mathbf{M}) = \mathbb{R}^d \oplus T\mathbf{M}$ can be uniquely written as $\xi = \xi^{\mathcal{X}} + \xi^{\mathcal{W}}$, where $\xi^{\mathcal{X}} \in T\mathbb{R}^d$ and $\xi^{\mathcal{W}} \in T\mathbf{M}$, and it has the Riemannian norm $|||\xi||| = |\xi^{\mathcal{X}}| + \|\xi^{\mathcal{W}}\|$, where $|\cdot|$ is the usual Euclidean norm on $\mathbb{R}^d$ and $\|\cdot\|$ is the Riemannian norm on $\mathbf{M}$. The corresponding metrics on $\mathbf{M}$ and on $\mathbb{R}^d \times \mathbf{M}$ will be denoted by $d_{\mathbf{M}}$ and d, respectively, so that if $z_1 = (x_1, w_1)$, $z_2 = (x_2, w_2) \in \mathbb{R}^d \times \mathbf{M}$, then $d(z_1, z_2) = |x_1 - x_2| + d_{\mathbf{M}}(w_1, w_2)$. It is known (see [**27**]) that the hyperbolic splitting $T_{\Lambda_x}\mathbf{M} = \Gamma_x^s \oplus \Gamma_x^0 \oplus \Gamma_x^u$ over Λ_x can be continuously extended to the splitting $T_{\mathcal{W}}\mathbf{M} = \Gamma_x^s \oplus \Gamma_x^0 \oplus \Gamma_x^u$ over $\mathcal{W}$ which is forward invariant with respect to DF_x^s and satisfies exponential estimates with a uniform in $x \in \mathcal{X}$ positive exponent which we denote again by $\kappa > 0$, i.e. we assume now that

$$\|DF_x^t \xi\| \le e^{-\kappa t}\|\xi\| \quad \text{and} \quad \|DF_x^{-t}\eta\| \le e^{-\kappa t}\|\eta\|, \tag{4.1}$$

provided $\xi \in \Gamma_x^s(w)$, $\eta \in \Gamma_x^u(F_x^t w)$, $t \ge t_0$ and $w \in \mathcal{W}$. Moreover, by [**9**] (see also [**28**]) we can choose these extensions so that $\Gamma_x^s(w)$ and $\Gamma_x^u(w)$ will be Hölder continuous in w and C^1 in x in the corresponding Grassmann bundle. Actually, since $\mathcal{W}$ is contained in the basin of each attractor Λ_x, any point $w \in \mathcal{W}$ belongs to the stable manifold $W_x^s(v)$ of some point $v \in \Lambda_x$ (see [**8**]), and so we naturally choose $\Gamma_x^s(w)$ to be the tangent space to $W_x^s(v)$ at w. Now each vector $\xi \in T_{x,w}(\mathcal{X} \times \mathcal{W}) = T_x\mathcal{X} \oplus T_w\mathcal{W}$ can be represented uniquely in the form $\xi = \xi^{\mathcal{X}} + \xi^s + \xi^0 + \xi^u$ with $\xi^{\mathcal{X}} \in T_x\mathcal{X}$, $\xi^s \in \Gamma_x^s(w)$, $\xi^0 \in \Gamma_x^0(w)$ and $\xi^u \in \Gamma_x^u(w)$. We also denote $\xi^{0s} = \xi^s + \xi^0$ and $\xi^{0u} = \xi^u + \xi^0$. For each small $\varepsilon, \alpha > 0$ set $\mathcal{C}^u(\varepsilon, \alpha) = \{\xi \in T(\mathcal{X} \times \mathcal{W}) : \|\xi^{0s}\| \le \varepsilon\alpha^{-2}\|\xi^u\|$ and $\|\xi^{\mathcal{X}}\| \le \varepsilon\alpha^{-1}\|\xi^u\|\}$ and $\mathcal{C}^u_{x,w}(\varepsilon, \alpha) = \mathcal{C}^u(\varepsilon, \alpha) \cap T_{x,w}(\mathcal{X} \times \mathcal{W})$, which are cones around Γ^u and $\Gamma_x^u(w)$, respectively. Similarly, we define $\mathcal{C}^s(\varepsilon, \alpha) = \{\xi \in T(\mathcal{X} \times \mathcal{W}) : \|\xi^{0u}\| \le \varepsilon\alpha^{-2}\|\xi^s\|$ and $\|\xi^{\mathcal{X}}\| \le \varepsilon\alpha^{-1}\|\xi^s\|\}$ and $\mathcal{C}^s_{x,w}(\varepsilon, \alpha) = \mathcal{C}^s(\varepsilon, \alpha) \cap T_{x,w}(\mathcal{X} \times \mathcal{W})$, which are cones around Γ^s and $\Gamma_x^s(w)$, respectively. Put $(\mathcal{X} \times \mathcal{W})_t = \{(x, w) : \Phi_\varepsilon^\tau(x, w) \in (\mathcal{X} \times \mathcal{W})\ \forall \tau \in [0, t]\}$, where, recall, Φ_ε^t is the flow determined by (1.1).

LEMMA 4.1. *There exist* $\alpha_0, \varepsilon(\alpha), t_1 > 0$ *such that if* $z = (x, y) \in (\mathcal{X} \times \mathcal{W})_t$ *and* $t \ge t_1, \alpha \le \alpha_0, \varepsilon \le \varepsilon(\alpha)$ *then*

$$D_z\Phi_\varepsilon^t \mathcal{C}_z^u(\varepsilon, \alpha) \subset \mathcal{C}^u_{\Phi_\varepsilon^t z}(\varepsilon, \alpha), \quad \mathcal{C}_z^s(\varepsilon, \alpha) \supset D_z\Phi_\varepsilon^{-t}\mathcal{C}^s_{\Phi_\varepsilon^t z}(\varepsilon, \alpha), \tag{4.2}$$

and for any $\xi \in \mathcal{C}_z^u(\varepsilon, \alpha)$, $\eta \in \mathcal{C}^s_{\Phi_\varepsilon^t z}$,

$$|||D_z\Phi_\varepsilon^t \xi||| \ge e^{(1/2)\kappa t}|||\xi|||, \quad |||D_z\Phi_\varepsilon^{-t}\eta||| \ge e^{(1/2)\kappa t}|||\eta|||. \tag{4.3}$$

Proof. Let $\xi = \xi^{\mathcal{X}} + \xi^u + \xi^{0s} \in T_z(\mathcal{X} \times \mathbf{M})$, $D_z\Phi^t\xi^{\mathcal{X}} = \zeta = \zeta^{\mathcal{X}} + \zeta^u + \zeta^{0s} \in T_{\Phi^t z}(\mathcal{X} \times \mathbf{M})$, $z = (x, y)$, $D_y F_x^t \xi^u = \eta^u$, and $D_y F_x^t \xi^{0s} = \eta^{0s}$. Then $D_z\Phi^t\xi = \zeta^{\mathcal{X}} + (\zeta^u + \eta^u) + (\zeta^{0s} + \eta^{0s})$ and $\|\xi^{\mathcal{X}}\| = \|\zeta^{\mathcal{X}}\|$, $\|\zeta^u\| \le Ce^{Ct}\|\xi^{\mathcal{X}}\|$, $\|\zeta^{0s}\| \le Ce^{Ct}\|\xi^{\mathcal{X}}\|$, $\|\eta^{0s}\| \le C\|\xi^{0s}\|$ for some $C > 0$ independent of ξ and $\|\eta^u\| \ge e^{\kappa t}\|\xi^u\|$ if $t \ge t_0$. Hence, for $t \ge t_0$,

$$\|\xi^u + \eta^u\| \ge \|\eta^u\| - \|\zeta^u\| \ge e^{\kappa t}\|\xi^u\| - Ce^{Ct}\|\xi^{\mathcal{X}}\| \tag{4.4}$$

and

$$\|\zeta^{0s} + \eta^{0s}\| \le \|\zeta^{0s}\| + \|\eta^{0s}\| \le Ce^{Ct}\|\xi^{\mathcal{X}}\| + C\|\xi^{0s}\|. \tag{4.5}$$

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

Suppose that $\|\xi^u\| \ge \alpha\varepsilon^{-1}\|\xi^{\mathcal{X}}\|$ and $\|\xi^u\| \ge \alpha^2\varepsilon^{-1}\|\xi^{0s}\|$. Then by (4.4),

$$\|\zeta^u + \eta^u\| \ge (\tfrac{1}{2}e^{\kappa t}\alpha\varepsilon^{-1} - Ce^{Ct})\|\xi^{\mathcal{X}}\| + \tfrac{1}{2}e^{\kappa t}\alpha^2\varepsilon^{-1}\|\xi^{0s}\|. \tag{4.6}$$

Set $t_1 = \kappa^{-1}\ln(8C+5)$, choose $\alpha_0 > 0$ so that $e^{\kappa t} \ge 4\alpha Ce^{Ct}$ for any $t \le t_0$ and all $t \in [t_1, 2t_1]$, and set $\varepsilon(\alpha) = (\alpha/4)\min(\alpha, C^{-1}e^{-2Ct_1})$. Then by (4.6) we see that $D_z\Phi^t\xi \in \mathcal{C}^u_{\Phi^t z}(\varepsilon, 2\alpha)$ for all $t \in [t_1, t_2]$, and so by continuity of the splitting $\Gamma^s_x \oplus \Gamma^0_x \oplus \Gamma^u_x$ and by perturbation arguments using Assumption 2.3, it follows that $D_z\Phi^t_\varepsilon\xi \in \mathcal{C}^u_{\Phi^t_\varepsilon z}(\varepsilon, \alpha)$ for all $t \in [t_1, 2t_1]$, provided ε is small enough. Repeating this argument for $t \in [it_1, (i+1)t_1]$, $i = 2, 3, \dots$, we conclude the proof of the first part of (4.2) and its second part follows in the same way.

Next, for $t \ge t_0$,

$$\begin{aligned}
|||D_z\Phi^t\xi||| &\ge \|\eta^u\| - \|\zeta^{\mathcal{X}}\| - \|\zeta^u\| - \|\zeta^{0s}\| - \|\eta^{0s}\| \\
&\ge e^{\kappa t}\|\xi^u\| - (1 + 2Ce^{Ct})\|\xi^{\mathcal{X}}\| - C\|\xi^{0s}\| \\
&\ge (e^{\kappa t} - \alpha^{-1}\varepsilon(1 + 2Ce^{Ct}) - \alpha^{-2}\varepsilon C)\|\xi^u\| \\
&\ge (e^{\kappa t} - \alpha^2(1 + 2Ce^{Ct}) - \alpha^{-2}\varepsilon C)(1 + \varepsilon\alpha^{-1} + \varepsilon\alpha^{-2})^{-1}|||\xi|||.
\end{aligned} \tag{4.7}$$

Choose $\alpha_0, \varepsilon(\alpha)$ so small (for instance, $\varepsilon(\alpha) = \alpha^3$) that for all $\alpha \le \alpha_0$ and $\varepsilon \le \varepsilon(\alpha)$,

$$e^{\kappa t} - \varepsilon\alpha^{-1}(1 + 2Ce^{Ct}) - \varepsilon\alpha^{-2}C \ge (1 + \varepsilon\alpha^{-1} + \varepsilon\alpha^{-2})e^{(2/3)\kappa t}$$

for all $t \in [t_1, 2t_1]$. Then, $|||D_z\Phi^t\xi||| \ge e^{(2/3)\kappa t}|||\xi|||$ for all such t, and so if ε is small enough we also have $\|D_z\Phi^t_\varepsilon\xi\| \ge e^{(1/2)\kappa t}\|\xi\|$. Using (4.2) and repeating this argument for $D_z\Phi^{it_1}_\varepsilon\xi$, $i = 1, 2, \dots$ in place of ξ we derive (4.3) for all $t \ge t_1$. □

For any linear subspace Ξ of $T_z(\mathbb{R}^d \times \mathbf{M})$, denote by $J^\Xi_\varepsilon(t, z)$ the absolute value of the Jacobian of the linear map $D_z\Phi^t_\varepsilon : \Xi \to D_z\Phi^t_\varepsilon\Xi$ with respect to inner products induced by the Riemannian metric. For each $z = (x, y) \in \mathbb{R}^d \times \mathbf{M}$, also set

$$J^u_\varepsilon(t, z) = \exp\left(-\int_0^t \varphi^u_{X^\varepsilon_{x,y}(s)}(Y^\varepsilon_{x,y}(s))\,ds\right). \tag{4.8}$$

Let n_u be the dimension of $\Gamma^u_x(w)$ which does not depend on x and w by continuity considerations. If Ξ is an n^u-dimensional subspace of $T_z(\mathcal{X} \times \mathcal{W})$, $z = (x, y)$, and $\Xi \subset \mathcal{C}^u_{x,y}(\varepsilon, \alpha)$ then it follows easily from Assumptions 2.2, 2.3 and Lemma 4.1 that there exists a constant $C_1 > 0$ independent of $z \in \mathcal{X} \times \mathcal{W}$ and a small ε such that for any $t \ge 0$,

$$(1 - C_1\varepsilon)^t \le J^\Xi_\varepsilon(t, z)(J^u_\varepsilon(t, z))^{-1} \le (1 + C_1\varepsilon)^t. \tag{4.9}$$

Denote by $U(z, \rho)$ the ball in $\mathbb{R}^d \times \mathbf{M}$ centered at z and let $\mathcal{D}^u_\varepsilon(z, \alpha, \rho, C)$ be the set of all C^1 embedded n_u-dimensional closed discs $D \subset \mathcal{X} \times \mathcal{W}$ such that $z \in D$, $TD \subset \mathcal{C}^u(\varepsilon, \alpha)$ and if $v \in \partial D$, then $C\rho \le d_D(v, z) \le C^2\rho$, where TD is the tangent bundle over D, ∂ denotes the boundary and d_D is the interior metric on D. If $D \in \mathcal{D}^u_\varepsilon(z, \alpha, \rho, C)$ then, clearly, $D \subset U(z, C^2\rho)$, and if ε/α^2 and ρ are small enough then $d(v, z) \ge \rho$ for any $v \in \partial D$. Let $D \in \mathcal{D}^u_\varepsilon(z, \alpha, \rho, C)$ and $z = (x, y) \in D \subset \mathcal{X} \times \mathcal{W}$. Set $U^\varepsilon_D(t, z, \gamma) = \{\tilde{z} \in D : d_D(\Phi^s_\varepsilon z, \Phi^s_\varepsilon\tilde{z}) \le \gamma\ \forall s \in [0, t]\}$ and let $\pi_1 : \mathcal{X} \times \mathcal{W} \to \mathcal{X}$ and $\pi_2 : \mathcal{X} \times \mathcal{W} \to \mathcal{W}$ be natural projections on the first and second factors, respectively.

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

LEMMA 4.2. *Let ε, α, t, (x, y) be as in Lemma 4.1 and $T > 0$. Then there exist $\rho_0, c, c_{\rho,T}, C > 0$ such that if $\rho \le \rho_0$, $D \in \mathcal{D}^u_\varepsilon((x, y), \alpha, \rho, C)$, $z \in D$, $V_{s,t}(z) = \Phi^s_\varepsilon U^\varepsilon_D(t, z, C\rho)$, $V_t(z) = V_{t,t}(z)$, $V = V_{0,t}(z) \subset D = \emptyset$ and $t \ge 0$ then:*

(i) $d_{V_{s,t}(z)}(\Phi^s_\varepsilon v, \Phi^s_\varepsilon z) \le c^{-1} e^{-(1/2)\kappa(t-s)} d_{V_t(z)}(\Phi^t_\varepsilon v, \Phi^t_\varepsilon z) \le c^{-1} C\rho e^{-(1/2)\kappa(t-s)}$ *for any $v \in V$ and $s \in [0, t]$, where d_U is the interior distance on U;*

(ii) $TV_{s,t}(z) \subset \mathcal{C}^u(\varepsilon, \alpha)$ *and* $V_t(z) \in \mathcal{D}^u_\varepsilon(\Phi^t_\varepsilon z, \alpha, \rho, \sqrt{C})$;

(iii) *for all $v \in V$ and $0 \le s \le t$,*

$$|\pi_1 \Phi^s_\varepsilon v - \pi_1 \Phi^s_\varepsilon z| \le C c^{-1} \rho\varepsilon\alpha^{-1}(1 - \varepsilon\alpha^{-1} - \varepsilon\alpha^{-2})^{-1} e^{-(1/2)\kappa(t-s)};$$ *and*

(iv) $c_{\rho,T} \le m_D(V) J^u_\varepsilon(t, z) \le c^{-1}_{\rho,T}$ *provided $t \le T/\varepsilon$, where m_D is the induced Riemannian volume on D.*

Proof. (i) Let γ be a smooth curve on $V_t(z)$ connecting $a = \Phi^t_\varepsilon z$ and $b = \Phi^t_\varepsilon v$, then $\tilde\gamma = \Phi^{s-t}_\varepsilon \gamma$ is a smooth curve on $V_{s,t}(z)$ connecting $\Phi^s_\varepsilon z$ and $\Phi^s_\varepsilon v$. Since $T\tilde\gamma \subset TV_s(z) \subset \mathcal{C}^u(\varepsilon, \alpha)$, then by (4.3), $\text{length}(\gamma) \ge e^{\beta(t-s)} \text{length}(\tilde\gamma)$ if $t - s \ge t_1$. Then for such t and s,

$$d_{V^\varepsilon_s}(\Phi^s_\varepsilon v, \Phi^s_\varepsilon z) \le \text{length}(\tilde\gamma) \le e^{-\beta(t-s)} \text{length}(\gamma). \tag{4.10}$$

Observe that (i) is non-trivial only for large $t - s$, so minimizing in γ in (4.10) we derive the assertion (i).

Next we derive (ii). Its first part follows from (4.2). By the definition of V, $d_{V_t(z)}(w, z) \le C\rho$ for any $w \in \partial V_t(z)$. It remains to show that $d_{V_t(z)}(w, z) \ge \sqrt{C}\rho$ for any $w \in \partial V_t(z)$. Indeed, suppose $d_{V_t(z)}(w, z) < \sqrt{C}\rho$. Set

$$d_1 = \sup_{-t_1 \le s \le 0, v \in \mathcal{X} \times \mathcal{W}} \|D_v \Phi^s_\varepsilon\|.$$

By Assumption 2.3 via perturbation arguments we conclude that $d_1 > 0$ provided that ε is small enough. Let $w = \Phi^t_\varepsilon v$. It follows from Lemma 4.1 that $d_{V_s(z)}(\Phi^s_\varepsilon v, \Phi^s_\varepsilon z) < d_1\sqrt{C}\rho$ for all $s \in [0, t]$. Hence, if $\sqrt{C} \ge d_1$ then $v \notin \partial V$ and so $w \notin \partial V_t(z)$.

In order to derive (iii), take an arbitrary smooth curve γ on $V_{s,t}(z)$ connecting $\Phi^s_\varepsilon v$ and $\Phi^s_\varepsilon z$. Then $d\gamma(s)/ds = \dot\gamma(s) \in \mathcal{C}^u_{\gamma(s)}(\varepsilon, \alpha)$. It follows that if $\gamma(s) = (\gamma^{\mathcal{X}}(s), \gamma^{\mathbf{M}}(s))$ with $\gamma^{\mathcal{X}}(s) \in \mathcal{X}$ and $\gamma^{\mathbf{M}}(s) \in \mathbf{M}$, then $\dot\gamma^{\mathbf{M}}(s) = \dot\gamma^{0s}(s) + \dot\gamma^u(s)$, $\|\dot\gamma^c X(s)\| \le \varepsilon\alpha^{-1}\|\dot\gamma^u(s)\|$, $\|\dot\gamma^{0s}(s)\| \le \varepsilon\alpha^{-2}\|\dot\gamma^u(s)\|$ and so

$$|||\dot\gamma(s)||| \ge \|\dot\gamma^u(s)\| - \|\dot\gamma^{0s}(s)\| - \|\dot\gamma^c X(s)\| \ge \|\dot\gamma^u(s)\|(1 - \varepsilon\alpha^{-1} - \varepsilon\alpha^{-2}).$$

Hence

$$\|\dot\gamma^{\mathcal{X}}(s)\| \le \varepsilon\alpha^{-1}(1 - \varepsilon\alpha^{-1} - \varepsilon\alpha^{-2})^{-1}|||\dot\gamma(s)|||,$$

and so

$$|\pi_1 \Phi^s_\varepsilon v - \Phi^s_\varepsilon z| \le \varepsilon\alpha^{-1}(1 - \varepsilon\alpha^{-1} - \varepsilon\alpha^{-2})^{-1} \text{length}(\gamma). \tag{4.11}$$

Minimizing the right-hand side in (4.11) over such γ we obtain (iii) using (i).

Finally, (iv) follows from (4.9), (i), (ii) and the Hölder continuity of φ^u (as a function on $\mathcal{X} \times W$). □

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

For each $y \in \Lambda_x$ and $\gamma > 0$ small enough, set $W^s_x(y,\gamma) = \{\tilde y \in \mathcal{W} : d_{\mathbf{M}}(F^t_x y, F^t_x \tilde y) \le \gamma\ \forall t \ge 0\}$ and $W^u_x(y,\gamma) = \{\tilde y \in \mathcal{W} : d_{\mathbf{M}}(F^t_x y, F^t_x \tilde y) \le \gamma\ \forall t \le 0\}$ which are local stable and unstable manifolds for F_x at y. According to [**14**] and [**27**], these families can be included into continuous families of n^s- and n^u-dimensional stable and unstable C^1 discs $W^s_x(y,\gamma)$ and $W^u_x(y,\gamma)$, respectively, defined for all $y \in \mathcal{W}$ and such that $W^s_x(y,\gamma)$ is tangent to Γ^s_x, $W^u_x(y,\gamma)$ is tangent to Γ^u_x, $F^t_x W^s_x(y,\gamma) \subset W^s_x(F^t_x y,\gamma)$, and $W^u_x(y,\gamma) \supset F^{-t}_x W^u_x(F^t_x y,\gamma)$. Actually, as we noted above, if $y \in \mathcal{W}$ then y belongs to a stable manifold $W^s_x(\tilde y)$ of some $\tilde y \in \Lambda_x$ and we choose $W^s_x(y,\gamma)$ to be the subset of $W^s_x(\tilde y)$.

LEMMA 4.3. *For any $0 < \rho_1 < \rho_0$ small enough, each $D \in \mathcal{D}^u_\varepsilon(z,\alpha,\rho,C)$ and a continuous function g on $\mathcal{X}\times\mathcal{W}$ uniformly in $x' \in \mathcal{X}$, $z \in \mathcal{X}\times\mathcal{W}$ and $\rho \in [\rho_1,\rho_0]$,*

$$\lim_{t\to\infty} \frac{1}{t} \log \int_D \exp\left(\int_0^t g(x', F^r_{\pi_1 z}\pi_2 v)\, dr \right) dm_D(v) = P_{F_{\pi_1 z}}(g(x',\cdot) + \varphi^u_{\pi_1 z}), \quad (4.12)$$

where m_D is the induced Riemannian volume on D.

Proof. Set $\tilde W = \pi_2 D$, $x = \pi_1 z$ and $y \in \pi_2 z$. By standard transversality considerations, we can define a one-to-one map $\tilde\pi : \tilde W \to \tilde\pi\tilde W \subset W^u_x(y,\tilde C\rho)$ by $\tilde\pi(\tilde w) = w \in F^\tau_x W^s_x(\tilde w,\tilde C\rho)$, provided that $\alpha,\rho,|\tau| > 0$ are small and $\tilde C > 0$ is sufficiently large. By the absolute continuity of the stable foliation arguments (see, for instance, [**25**, §3.3]), we conclude that $\tilde\pi$ and its inverse have bounded Jacobians. It follows that it suffices to establish (4.12) for $D = \{x\}\times W$, where $W = W^u_x(y,\gamma)$ uniformly in $\gamma \in [\gamma_0,\gamma_0^{-1}]$, $\gamma_0 > 0$.

Set

$$\tilde W_\tau = F^\tau_x W, \quad W_{r,q} = \bigcup_{r\le\tau\le q} F^\tau_x W$$

and

$$I^V_{x,x'}(t) = \int_V \exp\left(\int_0^t g(x', F^\tau_x v)\, d\tau \right) dm_V(v).$$

Then

$$\begin{aligned} I^{W_\tau}_{x,x'}(t) &= \int_{W_\tau} \exp\left(\int_0^t g(x', F^\theta_x v)\, d\theta \right) dm_{W_\tau}(v) \\ &= \int_W \exp\left(\int_\tau^{t+\tau} g(x', F^\theta_x v)\, d\theta \right) J^u_x(\tau, v)\, dm_W(v), \end{aligned}$$

and so

$$re^{-2(\tau+r)\|g\|} I^W_{x,x'}(t) \le I^{W_{\tau,\tau+r}}_{x,x'}(t) \le re^{(\tau+r)(2\|g\|+\|\varphi^u\|)} I^W_{x,x'}(t), \quad (4.13)$$

where $\|\cdot\|$ is the supremum norm on $\mathcal{X}\times\mathcal{W}$. Since $\mathcal{W} \subset \bigcup_{v\in\Lambda_x} W^s_x(v)$ by [**8**], then given $\eta > 0$ there exist $\gamma(\eta), \tau(\eta) > 0$, $v \in \Lambda_x$ and $U \subset \bigcup_{-r\le\theta\le r} F^\theta_x W^u_x(v)$ such that for any $\gamma \le \gamma(\eta)$ we can define a one-to-one map $\phi : U \to W_{\tau(\eta),\tau(\eta)+r}$ by $\phi(w) \in W^s_x(w,\eta)$. By the standard absolute continuity of the stable foliation considerations (see [**25**, §3.3]), it follows that ϕ and its inverse have bounded Jacobians which together with the above arguments yield that it suffices to prove Lemma 4.3 only when $y \in \Lambda_x$, and so (see [**8**])

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

$W = W^u_x(y,\gamma) \subset \Lambda_x$. We observe also that without loss of generality we can assume γ to be sufficiently small since we can always cover $W^u_x(y,\gamma)$ by $W^u_x(y_i,\tilde\gamma)$, $i = 1,\dots,k$ with $y_1 = y$, $k = k(\gamma,\tilde\gamma)$ and small $\tilde\gamma \le \gamma$, so proving Lemma 4.3 for all such $W^u_x(y_i,\tilde\gamma)$ will imply it for $W^u_x(y,\gamma)$ itself.

So now assume that $W = W^u_x(y,\gamma) \subset \Lambda_x$ and we claim that for any $\eta > 0$ there exists $\tau(\eta,\gamma) > 0$ such that $W_{0,\tau}$ forms an η-net in Λ_x for any $\tau \ge \tau(\eta,\gamma)$. Indeed, by topological transitivity there exists $v \in \Lambda_x$ whose orbit is dense in Λ_x, and so by standard ergodicity considerations with respect to any ergodic invariant measure with full support on Λ_x (take, for instance, the SRB measure), we conclude that for any τ, $\{F^r_x v, r \ge \tau\}$ is already dense in Λ_x. Then by the transversality of W^s_x and $\bigcup_\theta F^\theta_x W^u_x$, there exists $r > 0$ such that $F^r_x v \in W^s_x(w,\gamma)$ for some $w \in W$, and so the forward orbit of w is dense in Λ_x, whence our claim holds. By compactness and structural stability considerations it follows that we can choose the same $\tau(\eta,\gamma)$ for all $y \in \mathcal{W}$ and $x \in \mathcal{X}$.

For any set $V \subset \Lambda_x$, put $U_V(t,y,\zeta) = \{v \in V : d(F^r_x y, F^r_x v) \le \zeta\ \forall r \in [0,t]\}$. Recall that a finite set $E \subset \Lambda_x$ is called (ζ,t)-separated for the flow F^t_x if $y,\tilde y \in E$, $y \ne \tilde y$ implies $\tilde y \notin U_{\Lambda_x}(t,y,\zeta)$. A set $E \subset \Lambda_x$ is called (ζ,t)-spanning if for any $y \in \Lambda_x$ there is $\tilde y \in E$ such that $y \in U_{\Lambda_x}(t,\tilde y,\zeta)$. Let $W_{0,\tau}$ be an η-net in Λ_x and E be a maximal (ζ,t)-separated subset of $W_{0,\tau}$. Then $U_{W_{0,\tau}}(t,y,\zeta/2)$, $y \in E$ are disjoint sets. By transversality of $W_{0,\tau}$ and W^s_x there exists $C_1 > 0$ such that for any $y \in \Lambda_x$ we can find $v(y) \in W_{0,\tau}$ such that $y \in W^s_x(v(y), C_1\eta)$, and so for some $w(y) \in E$, $y \in U_{\Lambda_x}(t,w(y),C_2(\zeta+\eta))$ with some $C_2 > 0$ large enough but independent of x, y, ζ, η. Hence, E is $(C_2(\zeta+\eta),t)$-spanning, and so $W_{0,\tau} \subset_{y\in E} U_{W_{0,\tau}}(t,y,C_2(\zeta+\eta))$. Assume, first, that $g = g(x',v)$ in (4.12) is Hölder continuous in v. Then by standard volume lemma arguments (see [**8**]) we obtain for $V = U_{W_{0,\tau}}(t,y,\gamma)$, $\gamma > 0$ and $y \in W_{0,\tau}$ that

$$\left|\log \int_V \exp\left(\int_0^t g(x', F^r_x v)\,dr\right) dm_V(v) - \int_0^t (g(x', F^r_x y) + \varphi^u_x(F^r_x y))\,dr\right| \le C(\gamma), \tag{4.14}$$

where $C(\gamma) > 0$ does not depend on x, x', y, t. Now (4.12) follows (cf. [**17, 18**]) from the above by (4.13), (4.14) and a uniform in (γ,t)-separated and (γ,t)-spanning sets approximation of the topological pressure (see, for instance, [**6**] and [**10**]). For a general continuous g, approximate it uniformly by Hölder continuous functions and (4.12) will follow in this case again. The limit (4.12) is uniform in x' and in z since $P_{F_{\pi_1 z}}(g(x',\cdot) + \varphi^u_{\pi_1 z})$ is uniformly continuous in x' (easy to see) and it is uniformly continuous in z (see [**9**]) and, furthermore, it follows from Lemma 5.1 from [**9**] that the family $(1/t)\log\int_D \exp(\int_0^t g(x', F^r_{\pi_1 z}\pi_2 v)\,dr)\,dm_D(v)$, $t \ge 1$ is equicontinuous in z. □

PROPOSITION 4.4. *For any $\rho, C, b > 0$ with C large and $C\rho$ small enough there exists a positive function $\zeta_{b,\rho,T}(\Delta,s,\varepsilon)$ such that*

$$\limsup_{\Delta\to 0}\ \limsup_{\varepsilon\to 0}\ \limsup_{s\to\infty}\ \zeta_{b,\rho,T}(\Delta,s,\varepsilon) = 0 \tag{4.15}$$

and for any $x, x' \in \mathcal{X}$, $y \in \mathcal{W}$, $t \ge t_1$, $\tau \le T/\varepsilon - t$, $\beta \in \mathbb{R}^d$, $|\beta| \le b$, $D \subset \mathcal{X}_t \times \mathcal{W}$,

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

$D \in \mathcal{D}^u_\varepsilon((x,y),\alpha,\rho,C)$ and $z \in D$ satisfying $V = U^\varepsilon_D(t,z,C\rho) \subset D$, we have

$$\begin{aligned}\Big|\frac{1}{\tau}\log\int_V \exp\Big\langle\beta,\int_t^{t+\tau} B(x',Y^\varepsilon_v(s))\,ds\Big\rangle\,dm_D(v)\\ +\frac{1}{\tau}\log J^u_\varepsilon(t,z) - P_{F_{\pi_1 z_t}}(\langle\beta,B(x',\cdot)\rangle+\varphi^u_{\pi_1 z_t})\Big|\\ \le \zeta_{b,\rho,T}\left(\varepsilon\tau,\min\left(\tau,\left(\log\frac{1}{\varepsilon}\right)^\lambda\right),\varepsilon\right),\end{aligned} \tag{4.16}$$

where $\langle\cdot,\cdot\rangle$ is the inner product, $z_t = \Phi^t_\varepsilon z$, $\lambda \in (0,1)$ and m_D is the induced Riemannian volume on D.

Proof. By (1.1) and (2.12) for any $w, \tilde w \in \mathcal{W}$,

$$\begin{aligned}d(\Phi^s_\varepsilon w,\Phi^s\tilde w) &\le d(w,\tilde w)+\int_0^s \|b(\Phi^u_\varepsilon w,\varepsilon)-b(\Phi^u\tilde w)\|\,du\\ &\le d(w,\tilde w)+L\varepsilon s+L\int_0^s d(\Phi^u_\varepsilon w,\Phi^u\tilde w)\,du,\end{aligned}$$

where, recall, $\Phi = \Phi_0$, and so by Gronwall's inequality

$$d(\Phi^s_\varepsilon w,\Phi^s\tilde w) \le e^{Ls}(d(w,\tilde w)+L\varepsilon s). \tag{4.17}$$

Hence,

$$d_{\mathbf{M}}(\pi_2\Phi^s_\varepsilon w, F^s_{\pi_1 z_r}\pi_2 w) \le d(\Phi^s_\varepsilon w,\Phi^s(\pi_1 z_r,\pi_2 w)) \le e^{Ls}(|\pi_1 w-\pi_1 z_r|+L\varepsilon s). \tag{4.18}$$

Recall, that by Lemma 4.2(iii) for any $w \in V_r(z)$, $r \le t$,

$$|\pi_1 z_r - \pi_1 w| \le Cc^{-1}\rho\varepsilon\alpha^{-1}(1-\varepsilon\alpha^{-1}-\varepsilon\alpha^{-2})^{-1}. \tag{4.19}$$

Set

$$I^\varepsilon_{x'}(v,r,q) = \exp\Big\langle\beta,\int_r^q B(x',\pi_2(\Phi^s_\varepsilon v))\,ds\Big\rangle.$$

Then

$$\int_{V_{0,r}(z)} I^\varepsilon_{x'}(v,r,r+s)\,dm_D(v) = \int_{V_r(z)} I^\varepsilon_{x'}(w,0,s)J^{\Xi_w}_\varepsilon(-r,w)\,dm_{V_r(z)}(w), \tag{4.20}$$

where Ξ_w is the tangent space to $V_r(z)$ at w and $m_{V_r(z)}$ is the induced Riemannian volume on $V_r(z)$. By (4.8), (4.9), Lemma 4.2(i) and the Hölder continuity of the function φ^u,

$$C_2^{-1}(1+C_1\varepsilon)^{-r} \le J^{\Xi_w}_\varepsilon(-r,w)J^u_\varepsilon(r,z) \le C_2(1-C_1\varepsilon)^{-r} \tag{4.21}$$

for some $C_2 > 0$ independent of ε, r, $z \in D$ and $w \in V_r(z)$. Since $V_r(z) \in \mathcal{D}^u_\varepsilon(z_r,\alpha,\rho,\sqrt{C})$ by Lemma 4.2(ii), it follows from (2.12) and (4.18)–(4.21) that

$$\begin{aligned}(\nu_b(\varepsilon,s))^{-1} &\le \int_{V_r(z)} I^\varepsilon_{x'}(w,0,s)\,dm_{V_r(z)}(w)\\ &\quad\times\left(\int_{V_r(z)}\exp\Big\langle\beta,\int_0^s B(x',F^\sigma_{\pi_1 z_r}\pi_2 w)\,d\sigma\Big\rangle dm_{V_r(z)}(w)\right)^{-1}\\ &\le \nu_b(\varepsilon,s),\end{aligned} \tag{4.22}$$

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

where

$$v_b(\varepsilon, s) = 2 + C_3 \exp(C_3 b e^{Ls}\varepsilon),$$

$C_3 > 0$ is a constant independent of ε, ρ, b, r, z, x' and $\beta \in \mathbb{R}^d$ with $|\beta| \le b$.

Next, choose $\lambda \in (0, 1)$ and $\theta(\varepsilon) \in [(\log(1/\varepsilon))^{\lambda}, 2(\log(1/\varepsilon))^{\lambda}]$ so that $n = \tau/\theta(\varepsilon)$ is an integer. If $n \le 1$ then (4.16) follows from (4.20)–(4.22) and Lemma 4.3. Now, let $n > 1$, $k < n$ and $v \in V_{0,t}(z)$. Then by (2.12) and Lemma 4.2(i) for any $w \in V_{0,t+k\theta(\varepsilon)}(v)$,

$$\begin{aligned} C_4^{-1} &\le I^{\varepsilon}_{x'}(w, t, t + (k+1)\theta(\varepsilon)) \\ &\quad \times (I^{\varepsilon}_{x'}(v, t, t + k\theta(\varepsilon)) I^{\varepsilon}_{x'}(w, t + k\theta(\varepsilon), t + (k+1)\theta(\varepsilon)))^{-1} \le C_4 \end{aligned} \tag{4.23}$$

and

$$C_4^{-1} \le I^{\varepsilon}_{x'}(v, t, t + k\theta(\varepsilon))(I^{\varepsilon}_{x'}(w, t, t + k\theta(\varepsilon)))^{-1} \le C_4, \tag{4.24}$$

where $C_4 = C_4(b) = e^{4bLC\rho c^{-1}\kappa^{-1}}$. Integrating (4.23), we obtain

$$\begin{aligned} C_4^{-1} &\le \int_{V_{0,t+k\theta(\varepsilon)}(v)} I^{\varepsilon}_{x'}(w, t, t + (k+1)\theta(\varepsilon))\, dm_D(w) (I^{\varepsilon}_{x'}(v, t, t + k\theta(\varepsilon)) \\ &\quad \times \int_{V_{0,t+k\theta(\varepsilon)}(v)} I^{\varepsilon}_{x'}(w, t + k\theta(\varepsilon), t + (k+1)\theta(\varepsilon))\, dm_D(w))^{-1} \\ &\le C_4. \end{aligned} \tag{4.25}$$

From the estimates (4.20)–(4.22) together with Lemma 4.3, we conclude that for some $C_5 > 0$ independent of t, k, ε, ρ, v, and x',

$$\begin{aligned} &C_5^{-1} e^{-\theta(\varepsilon)\eta_{b,\rho}(\varepsilon)} (v_b(\varepsilon, \theta(\varepsilon)))^{-1} \\ &\quad \le \int_{V_{0,t+k\theta(\varepsilon)}(v)} I^{\varepsilon}_{x'}(w, t + k\theta(\varepsilon), t + (k+1)\theta(\varepsilon))\, dm_D(w) J^u_{\varepsilon}(t + k\theta(\varepsilon), v) \\ &\qquad \times \exp(-\theta(\varepsilon) P_{F_{\pi_1 v_{t+k\theta(\varepsilon)}}}(\langle \beta, B(x', \cdot)\rangle + \varphi^u_{\pi_1 v_{t+k\theta(\varepsilon)}})) \\ &\quad \le e^{\theta(\varepsilon)\eta_{b,\rho}(\varepsilon)} v_b(\varepsilon, \theta(\varepsilon)), \end{aligned} \tag{4.26}$$

where $v_s = \Phi^s_{\varepsilon} v$ and $\eta_{b,\rho}(\varepsilon) > 0$, $\eta_{b,\rho}(\varepsilon) \to 0$ as $\varepsilon \to 0$. Observe that by (1.1), (2.12) and (4.19),

$$|\pi_1 v_{t+k\theta(\varepsilon)} - \pi_1 z_t| \le L\varepsilon\tau(1+\varepsilon) + Cc^{-1}\rho\varepsilon\alpha^{-1}(1 - \varepsilon\alpha^{-1} - \varepsilon\alpha^{-2})^{-1},$$

and so setting $P = P_{F_{\pi_1 z_t}}(\langle \beta, B(x', \cdot)\rangle + \varphi^u_{\pi_1 z_t})$, by (2.12) and [**9**] (see also [**23, 28**]) we obtain that

$$|P - P_{F_{\pi_1 v_{t+k\theta(\varepsilon)}}}(\langle \beta, B(x', \cdot)\rangle + \varphi^u_{\pi_1 v_{t+k\theta(\varepsilon)}})| \le C_6 \varepsilon\tau, \tag{4.27}$$

for some $C_6 = C_6(b) > 0$ independent of v, $k \le n$, ε, t, z, x', and $\beta \in \mathbb{R}^d$ with $|\beta| \le b$ provided that, say, $\tau \ge 1$, which we can assume without loss of generality.

A finite set $E \subset D$ will be called $(s, \gamma, \varepsilon, D)$-separated if $v_i, v_j \in E$, $v_i \ne v_j$ implies that $v_i \notin U^{\varepsilon}_D(s, v_j, \gamma)$. Let E_k be a maximal $(t + k\theta(\varepsilon), C\rho, \varepsilon, D)$-separated set in D and define

$$E_k^U = \{v \in E_k : U^{\varepsilon}_D(t + k\theta(\varepsilon), v, C\rho) \cap U \ne \emptyset\}.$$

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

Then for $k \ge 1$,

$$U^\varepsilon_D(t,z,\gamma + a_k C\rho) \supset \bigcup_{v\in E_k^{U^\varepsilon_D(t,z,\gamma)}} U^\varepsilon_D(t+k\theta(\varepsilon), v, C\rho) \supset U^\varepsilon_D(t,z,\gamma), \tag{4.28}$$

where $a_k = c^{-1}e^{-(1/2)\kappa k\theta(\varepsilon)}$ and the left-hand side of (4.28) follows from Lemma 4.2(i) assuming that ε is small enough. Observe also that $U^\varepsilon_D(t+k\theta(\varepsilon), v, C\rho/2)$ are disjoint for different $v \in E_k$. For $k = 1, 2, \dots, n-1$ set $V(k,\rho) = U^\varepsilon_D\big(t, z, C\rho\big(1 + \sum_{j=k}^{n-1} a_j\big)\big)$ and $V(-k,\rho) = U^\varepsilon_D\big(t, z, C\rho\big(1 - \sum_{j=k}^{n-1} a_j\big)\big)$ with $V(n,\rho) = V = V_{0,t}(z)$. Then by (4.24)–(4.28) and Lemma 4.2(iii),

$$\begin{aligned}
&\int_{V(k+1,\rho)} I^\varepsilon_{x'}(v, t, t+(k+1)\theta(\varepsilon))\, dm_D(w)\\
&\quad\le \sum_{v\in E_k^{V(k+1,\rho)}} \int_{V_{0,t+k\theta(\varepsilon)}(v)} I^\varepsilon_{x'}(w, t, t+(k+1)\theta(\varepsilon))\, dm_D(w)\\
&\quad\le C_4 \sum_{v\in E_k^{V(k+1,\rho)}} I^\varepsilon_{x'}(v, t, t+k\theta(\varepsilon))\\
&\qquad\times \int_{V_{0,t+k\theta(\varepsilon)}(v)} I^\varepsilon_{x'}(w, t+k\theta(\varepsilon), t+(k+1)\theta(\varepsilon))\, dm_D(w)\\
&\quad\le C_5C_4 e^{\theta(\varepsilon)(\eta_{b,\rho}(\varepsilon)+C_6\varepsilon\tau+P)}\nu_b(\varepsilon,\theta(\varepsilon))\\
&\qquad\times \sum_{v\in E_k^{V(k+1,\rho)}} (J^u_\varepsilon(t+k\theta(\varepsilon), v))^{-1} I^\varepsilon_{x'}(v, t, t+k\theta(\varepsilon))\\
&\quad\le C_5C_4^2 c^{-1}_{\rho/2,T} e^{\theta(\varepsilon)(\eta_{b,\rho}(\varepsilon)+C_6\varepsilon\tau+P)}\nu_b(\varepsilon,\theta(\varepsilon))\\
&\qquad\times \sum_{v\in E_k^{V(k+1,\rho)}} \int_{U^\varepsilon_D(t+k\theta(\varepsilon),v,C\rho/2)} I^\varepsilon_{x'}(w, t, t+k\theta(\varepsilon))\, dm_D(w)\\
&\quad\le C_5C_4^2 c^{-1}_{\rho/2,T} e^{\theta(\varepsilon)(\eta_{b,\rho}(\varepsilon)+C_6\varepsilon\tau+P)}\nu_b(\varepsilon,\theta(\varepsilon)) \int_{V(k,\rho)} I^\varepsilon_{x'}(w, t, t+k\theta(\varepsilon))\, dm_D(w).
\end{aligned} \tag{4.29}$$

Similarly, we obtain

$$\begin{aligned}
&\int_{V(-(k+1),\rho)} I^\varepsilon_{x'}(w, t, t+(k+1)\theta(\varepsilon))\, dm_D(w)\\
&\qquad\ge C_5^{-1}C_4^{-2} c_{\rho,T} e^{-\theta(\varepsilon)(\eta_{b,\rho/2}(\varepsilon)+C_6\varepsilon\tau+P)}(\nu_b(\varepsilon,\theta(\varepsilon)))^{-1}\\
&\qquad\quad\times \int_{V(-k,\rho)} I^\varepsilon_{x'}(w, t, t+k\theta(\varepsilon))\, dm_D(w).
\end{aligned} \tag{4.30}$$

Employing (4.29) and (4.30) recursively for $k = n-1, n-2, \dots, 1$ and estimating $\int_{V(\pm1,\rho)} I^\varepsilon_{x'}(w, t, t+k\theta(\varepsilon))\, dm_D(w)$ by (4.26) with $k = 0$ and with $V(\pm1,\rho)$ in place of $V_{0,t}(v)$, we derive (4.16) with $\zeta_{b,\rho,T}(\Delta, s, \varepsilon) = s^{-1}\log(C_5C_4^2(c^{-1}_{\rho,T} + c^{-1}_{\rho/2,T})) + \eta_{b,\rho}(\varepsilon) + \eta_{b,\rho/2}(\varepsilon) + 2(\theta(\varepsilon))^{-1}\log\nu_b(\varepsilon,\theta(\varepsilon)) + C_6\Delta$, provided that $n \ge 1$. □

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

5. *Large deviations arguments*

We need the following version of a general uniform in parameters, upper large deviations bound when usual assumptions hold true with errors as in Proposition 4.4. For simplicity we will formulate the result for $\mathbb{R}^d$-valued random vectors although the same arguments work for random variables with values in a Banach space.

PROPOSITION 5.1. *Let $H = H_z(\beta)$ and $\eta = \eta(\beta)$ be functions on $\mathbb{R}^d \times \mathbb{R}^d$ and on $\mathbb{R}^d$, respectively, and $\{\Xi_\tau^z, \tau \geq 1, z \in \mathbb{R}^d\}$ be a family of $\mathbb{R}^d$-valued random vectors on a probability space $(\Omega, \mathcal{F}, P)$ such that $|\Xi_\tau^z| \leq C < \infty$ with probability one for some constant C and all $\tau \geq 1, z \in \mathbb{R}^d$. Suppose also that H_z depends continuously on z belonging to a compact set $G \subset \mathbb{R}^d$. Set*

$$L_z(\alpha) = \sup_{\beta \in \mathbb{R}^d} (\langle \beta, \alpha \rangle - H_z(\beta)), \quad \alpha \in \mathbb{R}^d. \tag{5.1}$$

Then for any $\lambda > 0$, closed sets $K_z \subset \mathbb{R}^d$ and a number $A \leq \inf_{z \in G} \inf_{\alpha \in K_z} L_z(\alpha)$, there exist $b = b(\lambda, A, K)$ and $\tau_0 = \tau_0(\lambda, A, K)$ (where $K = \{(z, \alpha) : z \in G, \alpha \in K_z\}$) such that if for some $\tau \geq \tau_0$ and for all $z \in G$ and any $\beta \in \mathbb{R}^d$ with $|\beta| \leq b$,

$$\tau^{-1} \log E e^{\tau \langle \beta, \Xi_\tau^z \rangle} \leq H_z(\beta) + \eta(\beta), \tag{5.2}$$

then

$$P\{\Xi_\tau^z \in K_z\} \leq \exp(-\tau(A - \eta_b - \lambda)), \tag{5.3}$$

where $\eta_b = \sup\{\eta(\beta) : |\beta| \leq b\}$.

Proof. Set $U_{C,K} = K \cap (G \times \overline{U_C(0)})$, where $U_\gamma(\alpha) = \{\tilde{\alpha} : |\tilde{\alpha} - \alpha| < \gamma\}$ and $\bar{U}$ denotes the closure of U. Given $\lambda > 0$, by the choice of A it follows from (5.1) that for any $(z, \alpha) \in U_{C,K}$ there exists $\beta_A(z, \alpha) \in \mathbb{R}^d$ such that

$$\langle \beta_A(z, \alpha), \alpha \rangle - H_z(\beta_A(z, \alpha)) > A - \frac{\lambda}{4}. \tag{5.4}$$

Set $q_{z,\rho}(\beta) = \sup\{|H_z(\beta) - H_{\tilde{z}}(\beta)| : |z - \tilde{z}| \leq \rho\}$. By the continuity of H_z in z, $q_{z,\rho}(\beta) \to 0$ as $\rho \to 0$. Put also $\gamma_{A,\lambda}(z, \alpha) = (\lambda/4) \min(1, |\beta_A(z, \alpha)|^{-1}, \sup\{\rho : q_{z,\rho}(\beta_A(z, \alpha)) \leq (\lambda/4)\})$. Cover the compact set $U_{C,K}$ by open sets $U_{\gamma_{A,\lambda}(z,\alpha)}(z, \alpha)$, $(z, \alpha) \in U_{C,K}$ where $U_\gamma(z, \alpha) = \{(\tilde{z}, \tilde{\alpha}) : |\tilde{z} - z|^2 + |\tilde{\alpha} - \alpha|^2 < \gamma^2\}$. Let $\{U_{\gamma_{A,\lambda}(z_i,\alpha_i)}(z_i, \alpha_i),\ i = 1, \ldots, n\}$, be a finite subcover with a minimal number of elements and put $b = \max_{1 \leq i \leq n} |\beta_A(z_i, \alpha_i)|$. Then, by (5.2) and (5.4) for any $z \in G$ and $i = 1, \ldots, n$,

$$\begin{aligned} e^{\tau \eta_b} &\geq E \mathbb{I}_{(z, \Xi_\tau^z) \in U_{\gamma_{A,\lambda}(z_i,\alpha_i)}(z_i,\alpha_i)} e^{\tau(\langle \beta_A(z_i,\alpha_i), \Xi_\tau^z \rangle - H_{z_i}(\beta_A(z_i,\alpha_i)))} \\ &\geq e^{-\tau\lambda/4} E \mathbb{I}_{(z, \Xi_\tau^z) \in U_{\gamma_{A,\lambda}(z_i,\alpha_i)}(z_i,\alpha_i)} e^{\tau(\langle \beta_A(z_i,\alpha_i), \alpha_i \rangle - H_{z_i}(\beta_A(z_i,\alpha_i)))} \\ &\geq e^{\tau(A - \lambda/2)} P\{(z, \Xi_\tau^z) \in U_{\gamma_{A,\lambda}(z_i,\alpha_i)}(z_i, \alpha_i)\}. \end{aligned}$$

Summing these inequalities over $i = 1, \ldots, n$ we obtain

$$P\{\Xi_\tau^z \in K_z\} \leq e^{-\tau(A - \eta_b - (3/4)\lambda - \tau^{-1} \log n)}. \tag{5.5}$$

So given $\lambda > 0$ and K, we obtain b and n and then choose $\tau_0 = (4/\lambda) \log n$, which together with (5.5) yield (5.3). □

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

Set

$$H_{x'x''}(\beta) = P_{x'}(\langle \beta, B(x'', \cdot)\rangle + \varphi^u_{x'}) \quad \text{and} \quad L_{x'x''}(\xi) = \sup_{\beta\in\mathbb{R}^d}(\langle \beta, \xi\rangle - H_{x'y'}(\beta)). \tag{5.6}$$

As explained in [**18**], using the variational principle for the topological pressure (2.13) and the convexity and upper semicontinuity of the entropy (as a function of measure) of hyperbolic flows (see [**6**]) together with the duality theorem from convex analysis (see [**2**, p. 201]) we can express $L_{x'x''}(\xi)$ in a more explicit form. Namely, set

$$I_{x'x''}(\eta) = \begin{cases} -\int \varphi^u_{x'}\, d\eta - h_\eta(F^1_{x'}) & \text{if } \eta \in \mathcal{M}_{x'}, \\ \infty & \text{if } \eta \notin \mathcal{M}_{x'}, \end{cases} \tag{5.7}$$

where, recall, $\mathcal{M}_x$ is the space of F^t_x-invariant probability measures on Λ_x. Then

$$L_{x'x''}(\xi) = \inf\left\{ I_{x'x''}(\eta) : \int B(x'', y)\, d\eta(y) = \xi \right\} \tag{5.8}$$

if $\eta \in \mathcal{M}_{x'}$ satisfying the condition in braces exists and $L_{x'x''}(\xi) = \infty$ otherwise. By the uniqueness of the equilibrium state $\mu^{\mathrm{SRB}}_{x'}$ for $\varphi^u_{x'}$, we conclude that

$$L_{x'x''}(\xi) = 0 \quad \text{if and only if } \xi = \xi^{\mathrm{SRB}}_{x'x''} = \int B(x'', y)\, d\mu^{\mathrm{SRB}}_{x'}(y). \tag{5.9}$$

COROLLARY 5.2. *Let $G \subset \mathcal{X}$ be a compact set. Then for any $\lambda > 0$, closed sets $K_{x'x''} \subset \mathbb{R}^d$ depending on $x', x'' \in G$ and a number $A \le \inf_{x',x''\in G} \inf_{\xi\in K_{x'x''}} L_{x'x''}(\xi)$, there exist $\kappa_0, \varepsilon_0, \tau_0 > 0$ such that if $\varepsilon \in (0, \varepsilon_0]$, $\tau \ge \tau_0$, $\varepsilon\tau \le \kappa_0$, $t \in [0, T/\varepsilon - \tau]$, $V = U^\varepsilon_D(t, z, C\rho) \subset D$ is the same as in Proposition 4.4 and $|x' - \pi_1 z_t| \le \kappa_0$ (where, again, $z_t = \Phi^t_\varepsilon z$), then*

$$m_D\left\{ v \in V : \tau^{-1}\int_t^{t+\tau} B(x'', Y^\varepsilon_v(s))\, ds \in K_{x'x''} \right\} \le e^{-\tau(A-\lambda)} m_D(V). \tag{5.10}$$

Set $K^\delta_{x'x''} = \{\xi \in \mathbb{R}^d : |\xi - \xi^{\mathrm{SRB}}_{x'x''}| \ge \delta\}$, Then for any $\delta > 0$,

$$\gamma_G(\delta) = \inf_{x',x''\in G} \inf_{\xi\in K^\delta_{x'x''}} L_{x'x''}(\xi) > 0, \tag{5.11}$$

and there exist $\kappa_0, \varepsilon_0, \tau_0 > 0$ such that if $\varepsilon \in (0, \varepsilon_0]$ $\tau \ge \tau_0$, $\varepsilon\tau \le \kappa_0$, $t \in [0, T/\varepsilon - \tau]$ and $|x' - \pi_1 z_t| \le \kappa_0$, then

$$m_D\left\{ v \in V : \left| \tau^{-1}\int_t^{t+\tau} B(x'', Y^\varepsilon_v(s))\, ds - \xi^{\mathrm{SRB}}_{x'x''} \right| \ge \delta \right\} \le e^{-(1/2)\tau\gamma_G(\delta)} m_D(V) \tag{5.12}$$

and

$$m_D(V \cap \Phi^{-t}_\varepsilon E_\varepsilon(\tau, \delta)) \le e^{-(1/2)\tau\gamma_G(\delta/2)} m_D(V). \tag{5.13}$$

Proof. The first part of Corollary 5.2 follows directly from Propositions 4.4 and 5.1 taking into account Lemma 4.2(iv) and the continuous dependence of $P_{x'}(\langle \beta, B(x'', \cdot)\rangle + \varphi^u_{x'})$ on $x', x'' \in \bar{\mathcal{X}}$ (see [**9**]). The estimate (5.12) follows from (5.10) with $K_{x'x''} = K^\delta_{x'x''}$ once we derive (5.11). Since $H_{x'x''}(\beta)$ defined by (5.6) is continuous in x', x'', and β then, as is

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

easy to see, $L_{x'x''}(\xi)$ defined by (5.6) is lower semicontinuous in x', x'', and β. By (5.9), $L_{x'x''}(\xi) > 0$ for any $x', x'' \in G$ and $\xi \in K^{\delta}_{x'x''}$. Hence, by lower semicontinuity, the infinum of $L_{x'x''}(\xi)$ over compact sets $x', x'' \in G$ and $\xi \in K^{\delta}_{x'x''} \cap \{\zeta \in \mathbb{R}^d : |\zeta| \le L\}$ is positive. Taking into account that $L_{x'x''}(\xi) = \infty$ whenever $|\xi| > L$, we derive (5.11). Next, take $x' = x''$ with $|x' - \pi_1 z_1| \le \kappa_0$, then we obtain (5.12) with $\xi^{\mathrm{SRB}}_{x'x''}$ equal to $\bar{B}(x')$ given by (2.13). By [**9**], $\bar{B}$ satisfies (2.2) and taking into account (4.19) we conclude that

$$\left\{v \in V : \left|\tau^{-1}\int_t^{t+\tau} B(X^{\varepsilon}_v(s), Y^{\varepsilon}_v(s))\,ds - \bar{B}(X^{\varepsilon}_v(t))\right| \ge \delta\right\}$$
$$\subset \left\{v \in V : \left|\tau^{-1}\int_t^{t+\tau} B(x', Y^{\varepsilon}_v(s))\,ds - \bar{B}(x')\right| \ge \frac{1}{2}\delta\right\},$$

provided that ε is small enough, which together with (5.12) yield (5.13). □

Now we can complete the proof of Theorem 2.4. Let ν be a probability measure on $\mathcal{X}$ with $G = \operatorname{supp}\nu \subset \mathcal{X}_T$. For small ρ and large C so that $C\rho$ is still small, we consider for each $x \in G$ and $y \in \mathcal{W}$ the balls $D_x(y)$ and $D'_x(y)$ on $W^u_x(y,\gamma)$ centered at y and having radii $C\rho$ and $2C\rho$, respectively, in the interior metric on $W^u_x(y,\gamma)$. Then $D_x(y), D'_x(y) \in \mathcal{D}^u_{\varepsilon}((x,y),\alpha,\rho,C)$. Choose a maximal $(t, C\rho, \varepsilon, D_x(y))$-separated set E on $D_x(y)$. Then $\bigcup_{v\in E} U^{\varepsilon}_{D_x(y)}(t,v,C\rho) \supset D_x(y)$ and by Lemma 4.2(i), $U^{\varepsilon}_{D_x(y)}(t,v,C\rho) \subset D'_x(y)$ for all $v \in E$ provided that t is large enough. Moreover, $U^{\varepsilon}_{D_x(y)}(t,v,\frac{1}{2}C\rho)$ are disjoint for different $v \in E$. This together with (5.13) yields that

$$m_{D_x(y)}(D_x(y) \cap \Phi^{-t}_{\varepsilon} E_{\varepsilon}(\tau,\delta)) \le e^{-(1/3)\tau\gamma_G(\delta/2)} m_{D_x(y)}(D_x(y)), \tag{5.14}$$

provided that τ is large enough. Set

$$Q_x(y) = \cup\{D_x(F^r_x\tilde{y}) : |r| \le C\rho, \tilde{y} \in W^s_x(y, C\rho)\}$$

and $R_x(y) = \bigcup_{|\tilde{x}-x|\le C\rho} Q_{\tilde{x}}(y)$. We choose $C\rho$ sufficiently small so that $\pi_1 R_x(y) \subset \mathcal{X}_T$ whenever $x \in G$. Since conditional measures of μ^{SRB}_x on unstable manifolds are equivalent to the Riemannian volume there, we conclude from (5.14) and the above construction that for both choices of the measure μ in Theorem 2.4,

$$\mu(R_x(y) \cap \Phi^{-t}_{\varepsilon} E_{\varepsilon}(\tau,\delta)) \le e^{-(1/4)\tau\gamma_G(\delta/2)} m_{D_x(y)}(D_x(y)), \tag{5.15}$$

provided τ is large enough. Covering $G \times \bar{\mathcal{W}}$ by a finite number of sets $R_x(y)$, we arrive at

$$\mu((\mathcal{X}_T \times \mathbf{M}) \cap \Phi^{-t}_{\varepsilon} E_{\varepsilon}(\tau,\delta)) \le e^{-(1/5)\tau\gamma_G(\delta/2)}, \tag{5.16}$$

provided that τ is large enough. Choosing $\tau = t(\varepsilon) = T/\varepsilon n(\varepsilon)$ so that both $t(\varepsilon) \to \infty$ and $n(\varepsilon) \to \infty$ as $\varepsilon \to 0$, we derive (2.5). Choosing $\tau = t(\varepsilon) = \Delta\varepsilon^{-1}$ with a small $\Delta = T/n > 0$ and, respectively, with n large but independent of ε, we derive (2.14) from (3.11), (3.12) and (5.16). □

Remark 5.3. Taking $t(\varepsilon) = o(\log(1/\varepsilon))$ we can derive (3.17), and so prove the first part of Theorem 2.4, applying the uniform upper large deviations bound of Proposition 5.1 to hyperbolic flows F^s_x in place of $Y^{\varepsilon}_{x,y}(s)$. In doing so we can employ Theorem 2.1 from [**18**] in place of Proposition 4.4 above, but the former result should be made uniform in x using arguments from [**9**] (similarly to Lemma 4.3 above). Nevertheless, this approach cannot give an exponential estimate in (2.14), only an estimate of order ε^c for some $c > 0$.

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

Remark 5.4. The averaging in difference equations with fast motions being diffeomorphisms acting in a neighborhood of their hyperbolic attractors was considered in [**21**]. Some details concerning a uniform upper large deviations bound were not provided there and they should be completed similarly to the above arguments using Proposition 5.1.

Acknowledgement. The author was partially supported by the US–Israel BSF.

References

[1] D. B. Anosov. Averaging in systems of ordinary differential equations with fast oscillating solutions. *Izv. Acad. Nauk SSSR Ser. Mat.* **24** (1960), 731–742 (in Russian).

[2] J. P. Aubin and I. Ekeland. *Applied Nonlinear Analysis*. Wiley, New York, 1984.

[3] V. I. Arnold, V. V. Kozlov and A. I. Neishtadt. Mathematical aspects of classical and celestial mechanics. *Dynamical Systems III (Encyclopedia of Mathematical Sciences, 3)*. Ed. V. I. Arnold. Springer, Berlin, 1988.

[4] V. I. Bakhtin. Cramér asymptotics in a system with slow and fast Markovian motions. *Theory Prob. Appl.* **44** (1999), 1–17.

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

[6] R. Bowen. Periodic orbits for hyperbolic flows. *Am. J. Math*. **94** (1972), 1–30.

[7] N. N. Bogolyubov and Yu. A. Mitropol'skii. *Asymptotic Methods in the Theory of Nonlinear Oscillations*. Hindustan, Delhi, 1961.

[8] R. Bowen and D. Ruelle. The ergodic theory of Axiom A flows. *Invent. Math*. **29** (1975), 181–202.

[9] G. Contreras. Regularity of topological and metric entropy of hyperbolic flows. *Math. Z.* **210** (1992), 97–111.

[10] E. Franko. Flows with unique equilibrium state. *Am. J. Math*. **99** (1977), 486–514.

[11] M. I. Freidlin. The averaging principle and theorems on large deviations. *Russian Math. Surveys* **33**(5) (1978), 107–160.

[12] G. Gallavotti. Chaotic hypothesis and universal large deviations properties. *Doc. Math. J. DMV (Extra Volume ICM 1998)*, **I** (1998), 205–233.

[13] K. Hasselmann. Stochastic climate models, Part I. Theory. *Tellus* **28** (1976), 473–485.

[14] M. Hirsch, J. Palis, G. Pugh and M. Shub. Neighborhoods of hyperbolic sets. *Invent. Math*. **9** (1970), 121–134.

[15] R. Z. Khasminskii. On stochastic processes defined by differential equations with a small parameter. *Theory Prob. Appl*. **11** (1966), 211–228.

[16] R. Z. Khasminskii. On the averaging principle for Itô stochastic differential equations. *Kibernetika (Prague)* **4** (1968), 260–279 (in Russian).

[17] Yu. Kifer. Large deviations in dynamical systems and stochastic processes. *Trans. Amer. Math. Soc*. **321** (1990), 505–524.

[18] Yu. Kifer. Averaging in dynamical systems and large deviations. *Invent. Math*. **110** (1992), 337–370.

[19] Yu. Kifer. Limit theorems in averaging for dynamical systems. *Ergod. Th. & Dynam. Sys.* **15** (1995), 1143–1172.

[20] Yu. Kifer. Averaging and climate models. *Stochastic Climate Models (Progress in Probability, 49)*. Birkhäuser, Basel, 2001.

[21] Yu. Kifer. Averaging in difference equations driven by dynamical systems. Geometric methods in dynamics. *Astérisque* **287** (2003), 103–123.

[22] A. Katok and B. Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems*. Cambridge University Press, Cambridge, 1995.

[23] A. Katok, G. Knieper, M. Pollicott and H. Weiss. Differentiability and analyticity of topological entropy for Anosov and geodesic flows. *Invent. Math*. **98** (1989), 581–597.

[24] P. Lochak and C. Meunier. *Multiple Averaging for Classical Systems*. Springer, New York, 1988.

[25] R. Mañé. *Ergodic Theory of Differentiable Dynamics*. Springer, Berlin, 1987.

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press

[26] E. Pardoux and A. Yu. Veretennikov. On Poisson equation and diffusion approximation II. *Ann. Prob.* **31** (2003), 1166–1192.
[27] C. Robinson. Structural stability of vector fields. *Ann. Math*. **99** (1974), 154–175.
[28] D. Ruelle. Differentiation of SRB states. *Commun. Math. Phys*. **187** (1997), 227–241.
[29] J. A. Sanders and F. Verhurst. *Averaging Methods in Nonlinear Dynamical Systems*. Springer, Berlin, 1985.
[30] I. Tamura. *Topology of Foliations: An Introduction*. American Mathematical Society, Providence, RI, 1992.
[31] A. Yu. Veretennikov. On the averaging principle for systems of stochastic differential equations. *Math. USSR Sbornik* **69** (1991), 271–284.
[32] A. Yu. Veretennikov. On large deviations in the averaging principle for SDEs with 'full dependence'. *Ann. Prob*. **27** (1999), 284–296.

https://doi.org/10.1017/S014338570400001X Published online by Cambridge University Press