NEW DIRECTIONS IN FULLY COUPLED AVERAGING FOR DYNAMICAL SYSTEMS.

YURI KIFER

INSTITUTE OF MATHEMATICS
HEBREW UNIVERSITY
JERUSALEM, ISRAEL

ABSTRACT. We describe some results and formulate few problems concerning dynamical systems which combine fast and slow motions, both depending on each other. The heuristic averaging principle which prescribes to approximate the slow motion by averaging its parameters in fast variables does not always work in this setup and if it does work then usually only in some average with respect to initial conditions sense. We exhibit also results which rely on stochastic properties of fast motions such as large deviations and stochastic resonances.

1. INTRODUCTION

Evolution of many real systems can be viewed as a combination of motions taking place with significantly different velocities which leads to complicated multiscale equations. We can arrive also at this setup viewing a physical system as a perturbation of an ideal one, the latter depending on parameters which at the first approximation are considered as constants of motion. In the real system these parameters start moving slowly (may be also with significantly different speeds) which lead to a multiscale motion. Such problems arised first in celestial mechanics in 18th century considering a multibody planet motion as a perturbation of certain two body problem which can be integrated exactly.

Most of the time we will consider classical two scale systems

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

$X^\varepsilon = X^\varepsilon_{x,y}$, $Y^\varepsilon = Y^\varepsilon_{x,y}$ with initial conditions $X^\varepsilon(0) = x$ and $Y^\varepsilon(0) = y$. At the end we will describe also a stochastic resonance phenomenon which emerges in the

setup of three scale systems

$$\frac{dW^{\delta,\varepsilon}(t)}{dt} = \delta\varepsilon A(W^{\delta,\varepsilon}(t), X^{\delta,\varepsilon}(t), Y^{\delta,\varepsilon}(t))$$

(1.2)
$$\frac{dX^{\delta,\varepsilon}(t)}{dt} = \varepsilon B(W^{\delta,\varepsilon}(t), X^{\delta,\varepsilon}(t), Y^{\delta,\varepsilon}(t))$$

$$\frac{dY^{\delta,\varepsilon}(t)}{dt} = b(W^{\delta,\varepsilon}(t), X^{\delta,\varepsilon}(t), Y^{\delta,\varepsilon}(t)),$$

$W^{\delta,\varepsilon} = W^{\delta,\varepsilon}_{w,x,y}$, $X^{\delta,\varepsilon} = X^{\delta,\varepsilon}_{w,x,y}$, $Y^{\delta,\varepsilon} = Y^{\delta,\varepsilon}_{w,x,y}$ with initial conditions $W^{\delta,\varepsilon}(0) = w$, $X^{\delta,\varepsilon}(0) = x$ and $Y^{\delta,\varepsilon}(0) = y$. In general, right hand sides in (1.1) and (1.2) may explicitly depend on ε and δ but, usually, this does not lead to qualitatively new effects, so in order to avoid unnecessary technicalities we do not consider this case here. We assume that $W^{\delta,\varepsilon} \in \mathbb{R}^l$, $X^{\delta,\varepsilon} \in \mathbb{R}^d$ while $Y^{\delta,\varepsilon}$ evolves on a compact n-dimensional Riemannian manifold M and the coefficients A, B, b are bounded smooth vector fields on $\mathbb{R}^l$, $\mathbb{R}^d$ and M, respectively, depending on other variables as parameters. The solution of (1.2) determine the flow of diffeomorphisms $\Phi^t_{\delta,\varepsilon}$ on $\mathbb{R}^l \times \mathbb{R}^d \times M$ acting by $\Phi^t_{\delta,\varepsilon}(w,x,y) = (W^{\delta,\varepsilon}_{w,x,y}(t), X^{\delta,\varepsilon}_{w,x,y}(t), Y^{\delta,\varepsilon}_{w,x,y}(t))$. Taking $\varepsilon = \delta = 0$ we arrive at the (unperturbed) flow $\Phi^t = \Phi^t_{0,0}$ acting by $\Phi^t_{0,0}(w,x,y) = (w, x, F^t_{w,x}y)$ where $F^t_{w,x}$ is another family of flows given by $F^t_{w,x}y = Y_{w,x,y}(t)$ with $Y = Y_{w,x,y} = Y^{0,0}_{w,x,y}$ which are solutions of

(1.3)
$$\frac{dY(t)}{dt} = b(w, x, Y(t)), \quad Y(0) = y.$$

It is natural to view the flow $\Phi^t_{0,0}$ as describing an idealized physical system where parameters $w = (w_1, ..., w_l)$, $x = (x_1, ..., x_d)$ are assumed to be constants of motion while the perturbed flow $\Phi^t_{\delta,\varepsilon}$ is regarded as describing a real system where evolution of these parameters is also taken into consideration.

Consider (1.1) and assume that the limit

(1.4)
$$\bar{B}(x) = \bar{B}_y(x) = \lim_{T\to\infty} T^{-1} \int_0^T B(x, F^t_x y) dt$$

(where $F^t_x y = Y^0_{x,y}(t)$) exists and it is the same for "many" y's, for instance, for almost all y's with respect to some measure(s). Namely, let μ_x be an ergodic invariant measure of the flow F^t_x. Then the limit (1.4) exists for μ_x−almost all y and it is equal to

(1.5)
$$\bar{B}(x) = \bar{B}_{\mu_x}(x) = \int B(x, y) d\mu_x(y).$$

If $\bar{B}(x)$ is Lipschitz continuous then we can speak about a unique solution $\bar{X}^\varepsilon = \bar{X}^\varepsilon_x$ of the averaged equation

(1.6)
$$\frac{d\bar{X}^\varepsilon(t)}{dt} = \varepsilon\bar{B}(\bar{X}^\varepsilon(t)), \quad \bar{X}^\varepsilon(0) = x.$$

Averaging

The averaging principle suggests to approximate X^ε by $\bar{X}^\varepsilon$ on time intervals of order $1/\varepsilon$. This approach works well when the vector field b in (1.1) does not depend on the slow variables, i.e. when $b(x, y) = b(y)$, and so all F_x^t's coincide with some flow F^t. In this case for any ergodic F^t-invariant measure μ the limit (1.4) exists for μ-almost all (a.a.) y's and it coincides with $\int B(x, y) d\mu(y)$. It is well known (see, for instance, [27]) that for such y's,

$$(1.7) \qquad \sup_{0 \le t \le T/\varepsilon} |X_{x,y}^\varepsilon(t) - \bar{X}_x^\varepsilon(t)| \longrightarrow 0 \quad \text{as} \quad \varepsilon \to 0.$$

An example due to Neishtadt which will be described in the next section shows that in the fully coupled case, i.e. when the coefficients in (1.1) depend both on x and y, the convergence (1.7) for fixed initial conditions , in general, does not hold true and it is possible to speak about this convergence only in some average with respect to initial conditions sense. This example is based on the phenomenon called the "capture into resonance" which is well known in perturbations of integrable Hamiltonian systems. It would be interesting to understand whether such nonconvergence examples can be constructed in another important setup which will be discussed here where fast motions are hyperbolic dynamical systems.

We will consider also the discrete time case where (1.1) and (1.2) are replaced by difference equations for sequences $W^{\delta,\varepsilon}(n) = W_{w,x,y}^{\delta,\varepsilon}(n)$, $X^{\delta,\varepsilon}(n) = X_{w,x,y}^{\delta,\varepsilon}(n)$, and $Y^{\delta,\varepsilon}(n) = Y_{w,x,y}^{\delta,\varepsilon}(n)$, $n = 0, 1, 2, ...$, so that

$$(1.8) \qquad X^\varepsilon(n+1) - X^\varepsilon(n) = \varepsilon\Psi(X^\varepsilon(n), Y^\varepsilon(n)), \ X^\varepsilon(0) = x,$$
$$Y^\varepsilon(n+1) = \Phi(X^\varepsilon(n), Y^\varepsilon(n)), \ Y^\varepsilon(0) = y,$$

or

$$W^{\delta,\varepsilon}(n+1) - W^{\delta,\varepsilon}(n) = \varepsilon\delta\Xi(W^{\delta,\varepsilon}(n), X^{\delta,\varepsilon}(n), Y^{\delta,\varepsilon}(n)), \ W^{\delta,\varepsilon}(0) = w,$$
$$(1.9) \ X^{\delta,\varepsilon}(n+1) - X^{\delta,\varepsilon}(n) = \varepsilon\Psi(W^{\delta,\varepsilon}(n), X^{\delta,\varepsilon}(n), Y^{\delta,\varepsilon}(n)), \ X^{\delta,\varepsilon}(0) = x,$$
$$Y^{\delta,\varepsilon}(n+1) = \Phi(W^{\delta,\varepsilon}(n), X^{\delta,\varepsilon}(n), Y^{\delta,\varepsilon}(n)), \ Y^{\delta,\varepsilon}(0) = y$$

where Ξ and Ψ are smooth vector functions and $F_{w,x} = \Phi(w, x, \cdot) : M \to M$ (or $F_x = \Phi(x, \cdot) : M \to M$ in the case of (1.8)) is a smooth map (a diffeomorphism or an endomorphism). As usual in dynamical systems, it is quite useful to consider the discrete time setup whenever possible since it provides a richer source of examples than the continuous time (flow) case and it may better clarify the situation (see, for instance, the example at the end of Section 3). In the case of the system (1.8) set $F_x = \Phi(x, \cdot) : M \to M$. Then, if the limit

$$(1.10) \qquad \bar{\Psi}(x) = \bar{\Psi}_y(x) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \Psi(x, F_x^n y)$$

exists, it is the same for "many" y's and $\bar{\Psi}(x)$ is Lipschitz continuous then we can speak again about the averaged equation

$$(1.11) \qquad \frac{d\bar{X}^\varepsilon(t)}{dt} = \varepsilon\bar{\Psi}(\bar{X}^\varepsilon(t)), \quad \bar{X}^\varepsilon(0) = x$$

and study the approximation of $X_{x,y}^\varepsilon(n)$ by $\bar{X}_x^\varepsilon(n)$ for $n \in [0, T/\varepsilon]$.

In the next section we will discuss convergence in (1.7) for fixed initial conditions and exhibit Neishtadt's nonconvergence example. In Section 3 we formulate a general result which provide the convergence in (1.7) in some averaged in the initial conditions sense. In Sections 4 and 5 we discuss results which rely on stochastic (chaotic) properties of the fast motion. Namely, we will deal with the case where F_x^t (or F_x), $x \in \mathbb{R}^d$ is a family of Axiom A flows (or diffeomorphisms) in a vicinity of a hyperbolic attractor, in particular, they could be Anosov systems. Such situation can arise in perturbations of nonintegrable Hamiltonian systems which are geodesic flows on manifolds of constant energy which are supposed to be negatively curved. In the discrete time case we can also have F_x, $\mathbb{R}^d$ to be a family of expanding transformations which yields a wealth of explicit examples. This setup enables us to obtain probabilistic descriptions of the error $X_{x,y}^\varepsilon(t) - \bar{X}_x^\varepsilon(t)$ in the averaging approximation in the form of large deviations and stochastic resonance type results. This series of results seems to be important as a fenomenological justification of models of weather–climate interactions where weather is considered as a fast chaotic motion and climate as a slow one (see [14], [15], [11], and [20]). Several assertions which are known in the uncoupled case (all F_x are the same) have not been proved yet in the fully coupled situation and we formulate them as problems in Section 6.

2. Neishtadt's nonconvergence example

For any probability measure μ on M set

$$(2.1) \qquad \bar{B}_\mu(x) = \bar{B}(x) = \int B(x,y)d\mu(y).$$

In the uncoupled case we have only one flow F^t and then for any ergodic F^t-invariant measure μ and for μ-a.a. y's the slow motion $X_{x,y}^\varepsilon(t)$ is close on time intervals of order T/ε to the averaged motion $\bar{X}_x^\varepsilon(t)$ which solves (1.6) with $\bar{B} = \bar{B}_\mu$, i.e. for such y's (1.7) holds true. In the fully coupled case the situation is more complicated. Change the time and set $Z_{x,y}^\varepsilon(t) = X_{x,y}^\varepsilon(t/\varepsilon)$. Since we assume that the vector field is bounded then for fixed x and y $\{Z_{x,y}^\varepsilon(t), \varepsilon > 0, t \in [0,T]\}$ is a compact family of Lipschitz continuous curves in $\mathbb{R}^d$ with respect to the uniform norm $\|\varphi\| = \sup_{0 \leq t \leq T} |\varphi(t)|$, $\varphi : [0,T] \to \mathbb{R}^d$. A general result from [3] sais,

essentially, that any limit point $Z^0 = Z_x^0$ of this family is a solution of an equation of the form

$$(2.2) \qquad \frac{dZ^0(t)}{dt} = B_{\mu_{Z^0(t)}}(Z^0(t)), \quad Z^0(0) = x$$

where μ_z is some F_z^t-invariant probability measure and B_μ is defined by (2.1). In particular, if all flows F_z^t are uniquely ergodic (which happens very rarely) then we obtain the convergence (1.7) for all initial conditions.

The following example due to Neishtadt (which appeared previously in [2] by different reasons) shows that, in general, we cannot obtain results more precise than the above assertion concerning the convergence (1.7) for individual initial conditions. Consider the system of equations

$$(2.3) \qquad \dot{I} = \varepsilon(4 + 8\sin\gamma - I), \quad \dot{\gamma} = I$$

with the corresponding averaged equation

$$(2.4) \qquad \dot{J} = \varepsilon(4 - J).$$

Here γ belongs to the circle $\mathbb{T}^1$ parametrized by the interval $[-2\pi, 0]$ with the end points glued together. Denote by $(I_{I_0,\gamma_0}^\varepsilon(t), \gamma_{I_0,\gamma_0}^\varepsilon(t))$ and by $J_{I_0}^\varepsilon(t)$ the solution of (2.3) and of (2.4), respectively, with the initial conditions $I_{I_0,\gamma_0}^\varepsilon(0) = I_0$, $\gamma_{I_0,\gamma_0}^\varepsilon(0) = \gamma_0$ and $J_{I_0}^\varepsilon(0) = I_0$.

2.1. Proposition. *For any initial condition* (I_0, γ_0) *with* $-2 < I_0 < -1$ *there is a sequence* $\varepsilon_n \to 0$ *as* $n \to \infty$ *such that* $I_{I_0,\gamma_0}^{\varepsilon_n}(t) < 0$ *for all* $t \geq 0$ *and* $J_{I_0}^{\varepsilon_n}(1/\varepsilon_n) > 3/2$, *so, in particular,*

$$(2.5) \qquad \sup_{0 \leq t \leq 1/\varepsilon_n} |I_{I_0,\gamma_0}^{\varepsilon_n}(t) - J_{I_0}^{\varepsilon_n}(t)| > 3/2.$$

A full proof of this assertion can be found in [24]. In fact, (2.5) holds true for any (I_0, γ_0) belonging to certain strip S_ε having width of order $\varepsilon^{3/2}$ which winds around the lower half $\{I < 0, \varphi \in [0.2\pi]\}$ of the phase cylinder so that the distance between subsequent coils of S_ε is of order ε. So when $\varepsilon \to 0$ the strip S_ε passes trough all points, say, of the domain $\{-2 < I_0 < -1\}$.

The phenomenon above is due to the resonance $I = 0$. When $I \neq 0$ then the equation $\dot{\gamma} = I$ defines a circle rotation which preserves only the Lebesgue measure on it and the time average of any continuous function coincides with its space average with respect to the Lebesgue measure. On the other hand, $\dot{\gamma} = 0$ defines the identity transformation which preserves, of course, all probability measures but more importantly, the time average of a continuous function will be just its value at the initial point and not its space average with respect to the Lebesgue measure which is usually different unless we have a constant function. More generally, non-resonant and resonant unperturbed fast motions which are toral rotations differ from

each other by having a unique and many invariant measures, respectively, with resonant directions occuring "very rarely". It seems that a more important reason for problems in averaging due to resonances is connected with the fact that the reference (Lebesgue) measure becomes nonergodic for some parameters and the time averaging there has nothing to do with the space averaging with respect to this measure. We will discuss again this problem in the next section considering fast motions being Axiom A systems and expanding transformations which have abundance of invariant measures but there are natural families of ergodic invariant measures so that time and space averages coincide for almost all initial conditions with respect to appropriate measures.

3. GENERAL CONVERGENCE RESULTS

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

$$(3.1) \quad \|B(x, y) - B(z, v)\| + \|b(x, y) - b(z, v)\| \leq L(|x - z| + d_M(y, v))$$

$$\text{and } \|B(x, y)\| + \|b(x, y)\| \leq L$$

where d_M is the distance on M. Together with (1.1) we consider also the equation (1.6) on $\bar{\mathcal{X}}$ with coefficients $\bar{B}$ for which there exists $\bar{L} > 0$ such that for all $x, z \in \bar{\mathcal{X}}$,

$$(3.2) \quad \|\bar{B}(x) - \bar{B}(z)\| \leq \bar{L}|x - z| \text{ and } \|\bar{B}(x)\| \leq \bar{L}.$$

The Lipschitz continuity conditions (3.1) and (3.2) ensure existence and uniqueness of solutions of (1.1) and (1.6), respectively. If $\bar{B}$ is defined by (1.5) with $\mu = \mu_x$ then (3.2) is equivalent to the existence of $\bar{L} > 0$ such that for all $x, z \in \mathcal{X}$,

$$(3.3) \quad |\int_M B(x, y) d(\mu_x - \mu_z)(y)| \leq \bar{L}|x - z|,$$

which is a condition of regular dependence of μ_x on x. Set $\mathcal{X}_t = \{x \in \mathcal{X} : X^\varepsilon_{x,y}(s) \in \mathcal{X}, \bar{X}^\varepsilon_x(s) \in \mathcal{X} \text{ for all } y \in M \text{ and } s \in [0, t/\varepsilon]\}$. It is clear that $\mathcal{X}_t$ is an open set and by (3.1) and (3.2) it follows that $\mathcal{X}_t \supset \{x \in \mathcal{X} : \inf_{z \notin \mathcal{X}} |z - x| > 2t \max(L, \bar{L})\}$. Introduce

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

The following result is proved in [23] (and the same result for the discrete time setup (1.8) is obtained in [21]).

3.1. Theorem. *Suppose that (3.1) and (3.2) hold true and let μ be a probability measure on $\mathcal{X} \times M$. Then*

$$(3.4) \qquad \lim_{\varepsilon \to 0} \int_{\mathcal{X}_T} \int_M \sup_{0 \le t \le T/\varepsilon} |X_{x,y}^\varepsilon(t) - \bar{X}_x^\varepsilon(t)| d\mu(x,y) = 0$$

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

$$(3.5) \qquad \lim_{\varepsilon \to 0} \max_{0 \le j < n(\varepsilon)} \mu\{(\mathcal{X}_T \times M) \cap \Phi_\varepsilon^{-jt(\varepsilon)} E_\varepsilon(t(\varepsilon), \delta)\} = 0,$$

where $t(\varepsilon) = \frac{T}{\varepsilon n(\varepsilon)}$ and $\Phi_\varepsilon^t(x, y) = (X_{x,y}^\varepsilon(t), Y_{x,y}^\varepsilon(t))$.

Taking into account that $Y_{x,y}^\varepsilon(t)$ and $Y_{x,y}^0(t)$ stay close during the time $t \le t(\varepsilon)$ with $t(\varepsilon)$ much smaller than $\log(1/\varepsilon)$, we obtain a sufficient condition for (3.4) in the form of (3.5) with $E_0(\cdot, \cdot)$ in place of $E_\varepsilon(\cdot, \cdot)$. It is not difficult (see [23]) to check (3.5) in two situations where (3.4) was known before, namely, when the fast motion $Y_{x,y}^\varepsilon$ does not depend on the slow motion $X_{x,y}^\varepsilon$ and in the situation of the Anosov theorem (see [1]). The latter requires that $d\mu(x, y) = d\mu_x(y)d\nu(x)$ with ν having a bounded C^1 density with respect to the Lebesgue measure on $\mathbb{R}^d$ and μ_x, $x \in \mathcal{X}$ being invariant measures of the corresponding unperturbed flows F_x^t 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 M that is C^1 in both x and y.

Theorem 3.1 gives conditions for convergence in average in the averaging principle. In view of resonances (see, for instance, [25]) it is impossible for many interesting examples to ensure (1.7) for all $x \in \mathcal{X}$ and y. One still could hope that the convergence in average (3.4) could be improved to convergence almost everywhere but somehow this question has not been touched upon until recently in the literature. In the example of the previous section the convergence (1.7) does not hold true for any initial condition from a large open domain. Thus the type of convergence to the averaged motion described in Theorem 3.1 cannot be improved, in general, in the fully coupled averaging setup.

There is a very restricted class of systems where (1.7) holds true for all $x \in \mathcal{X}$ and $y \in M$. This happens, for instance, when Arnold's conditions for two-frequency systems are satisfied (see Section 3.5 in [25] and Section 5.1 in [2]). If the convergence in (1.4) is uniform in $x \in \mathcal{X}$ and $y \in M$ then (1.7) takes place, as well. In fact, 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 (3.5) is the case of fast motions being slowly changing Axiom A flows where the averaging principle in the form (3.4) has been established first in [23] using this approach.

3.2. Assumption. The family $b(x, \cdot)$ in (1.1) consists of C^2 vector fields on an n-dimensional Riemannian manifold 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 M given by

$$(3.6) \qquad \frac{dF_x^t y}{dt} = b(x, F_x^t y), \qquad F_x^0 y = y$$

possesses a basic hyperbolic attractor Λ_x (see [17]) with a hyperbolic splitting $T_{\Lambda_x} M = \Gamma_x^s \oplus \Gamma_x^0 \oplus \Gamma_x^u$, where Γ_x^s, Γ_x^u, and Γ_x^0 are the stable, unstable, and flow directions, respectively, and there exists an open set $W \subset M$ with the closure $\bar{W}$ and $t_0 > 0$ such that

$$(3.7) \qquad \Lambda_x \subset W, \quad F_x^t \bar{W} \subset W \ \forall t \geq t_0, \text{ and } \bigcap_{t>0} F_x^t W = \Lambda_x \ \forall x \in \bar{\mathcal{X}}.$$

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

$$(3.8) \qquad \varphi_x^u(y) = -\frac{dJ_x^u(t, y)}{dt}\Big|_{t=0}.$$

The function $\varphi_x^u(y)$ is known to be Hölder continuous in y, since the subbundles Γ_x^u are Hölder continuous (see [17]), and $\varphi_x^u(y)$ is C^1 in x (see [10]). The Sinai-Ruelle-Bowen measure μ_x^{SRB} of F_x^t is the unique equilibrium state of F_x^t for the function φ_x^u (see [9]), i.e. it is the only F_x^t-invariant probability measure on Λ_x whose topological pressure is zero (since Λ_x is an attractor). We replace now the condition (3.1) by the following stronger one:

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

$$(3.9) \qquad \|B(x, y)\|_{C^1(\bar{\mathcal{X}} \times M)} + \|b(x, y, \varepsilon)\|_{C^2(\bar{\mathcal{X}} \times M)} \leq L$$

where $\| \cdot \|_{\bar{\mathcal{X}} \times M)}$ is the C^i norm of the corresponding vector fields on $\bar{\mathcal{X}} \times M$.

Set

$$(3.10) \qquad \bar{B}(x) = \int B(x, y) d\mu_x^{\text{SRB}}(y)$$

then under Assumption 3.3 $\bar{B}$ is C^1 in x (see [10]), and so (3.2) is automatically satisfied. The following result was proved in [23] and its discrete time counterpart (with $F_x = \Phi(x, \cdot)$ in (1.8) being C^2 expanding endomorphisms or Axiom A diffeomorphisms in a vicinity of a hyperbolic attractor) was derived in [21].

3.4. Theorem. *Suppose that Assumptions 3.2 and 3.3 hold true. Define $\bar{B}$ by (3.10) 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 (3.5) is satisfied with $t(\varepsilon) = \frac{T}{\varepsilon n(\varepsilon)}$ whenever both $t(\varepsilon) \to \infty$ and $n(\varepsilon) \to \infty$ as $\varepsilon \to 0$, and so (3.4) holds true. Moreover, for any $a > 0$ there exist $c > 0$ and ε_a such that*

$$(3.11) \qquad \mu\{(x,y) \in \mathcal{X}_T \times \mathcal{W} : \sup_{0 \le t \le T/\varepsilon} |X^\varepsilon_{x,y}(t) - \bar{X}^\varepsilon_x(t)| > a\} \le e^{-c/\varepsilon},$$

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

Observe that we can take, in particular, ν to be the Dirac measure (unit mass) at a point $x \in \mathcal{X}$, i.e. (3.4) follows here without integration in x and (3.11) can be replaced by

$$(3.12) \qquad \mu_x\{y \in \mathcal{W} : \sup_{0 \le t \le T/\varepsilon} |X^\varepsilon_{x,y}(t) - \bar{X}^\varepsilon_x(t)| > a\} \le e^{-c/\varepsilon}$$

for either $\mu_x = m_\mathcal{W}$ or $\mu_x = \mu^{SRB}_x$. It is clear that in this setup Φ^t_ε is a small perturbation of the partially hyperbolic dynamical system Φ^t_0 but this observation dies not help in the analysis.

Note that Neishtadt's example discussed in the previous section is constructed in the standard resonance framework where for some x the measures μ_x (which all coincide with the Lebesgue measure there) become nonergodic. In the setup of Theorem 3.4 all measures μ_x are ergodic, and so it is still not clear whether it is possible in these circumstances to derive the convergence (1.7) for all (or for Lebesgue almost all) $x \in \mathcal{X}_T$ and for $m_\mathcal{W}$-almost all $y \in \mathcal{W}$ and not just convergence in average (3.4) or in measure (3.11) and (3.12). One difficulty in understanding this problem is related to the fact that the relevant F^t_x-invariant ergodic measures μ^{SRB}_x are, in general, singular with respect to each other and with respect to $m_\mathcal{W}$. Still, even in the case when all μ^{SRB}_x are equivalent to (or even coincide with) the Riemannian volume on M (for instance, when F^t_x, $x \in \mathbb{R}^d$ are geodesic flows with respect to slowly varying metrics or F^t_x, $x \in \mathbb{R}^d$ are all conjugated to a flow preserving the Riemannian volume by means of a family of diffeomorphisms) the answer is not clear. Consider, for instance, the following explicit discrete time example which manifests, in particular, usefulness of the discrete time setup as a rich source of examples.

3.5. Example. Let M be the unit circle $\mathbb{T}^1$ in $\mathbb{R}^2$ centered at $(0,0)$ and define $F_x :$ $\mathbb{T}^1 \to \mathbb{T}^1$ by $F_x e^{i\varphi} = e^{i(2\varphi + x)}$ where $\varphi, x \in \mathbb{R}^1$. Let B be a continuous 2π-periodic function on $\mathbb{R}^1$ such that $\int_0^{2\pi} B(\varphi) d\varphi = 0$ (take, for instance, $B(\varphi) = \sin \varphi$ or $B(\varphi) = \cos \varphi$. Consider $X^\varepsilon(n) = X^\varepsilon_{x,y}(n)$, $Y^\varepsilon(n) = Y^\varepsilon_{x,y}(n)$ given by

$$(3.13) \qquad X^\varepsilon(n+1) - X^\varepsilon(n) = \varepsilon B(\varphi^\varepsilon(n)), \ X^\varepsilon(0) = x$$

$$Y^\varepsilon(n+1) = F_{X^\varepsilon(n)} e^{i\varphi^\varepsilon(n)}, \ \varphi^\varepsilon(n+1) = 2\varphi^\varepsilon(n) + X^\varepsilon(n),$$

$$Y^\varepsilon(0) = y = e^{i\varphi}, \ \varphi^\varepsilon(0) = \varphi.$$

Here all F_x's preserve the Lebesgue measure Leb on $\mathbb{T}^1$ and the averaged (with respect to Leb) equation is

$$(3.14) \qquad \frac{d\bar{X}^\varepsilon(t)}{dt} = 0.$$

It follows from both the Anosov type theorem from [21] (see Corollary 2.2 there) and from the discrete time version of Theorem 3.4 above which was proved in [23] that

$$(3.15) \qquad \lim_{\varepsilon \to 0} \int_{\mathbb{T}^1} \sup_{0 \le n \le T/\varepsilon} |X^\varepsilon_{x,y}(n) - x| dy = 0$$

and for any $\delta > 0$ there exist $c = c_\delta > 0$ such that for all sufficiently small $\varepsilon > 0$,

$$(3.16) \qquad \text{Leb}\{y \in \mathbb{T}^1 : \sup_{0 \le n \le T/\varepsilon} |X^\varepsilon_{x,y}(n) - x| > \delta\} \le e^{-c/\varepsilon}.$$

It is still not clear whether for Leb-a.a. y's and all or Leb-a.a. x's,

$$(3.17) \qquad \sup_{0 \le n \le T/\varepsilon} |X^\varepsilon_{x,y}(n) - x| \longrightarrow 0 \text{ as } \varepsilon \to 0.$$

4. LARGE DEVIATIONS

In this section we exhibit precise large deviations bounds which both will improve the estimate (3.12) and will enable us to study the behavior of X^ε on much longer, exponential large in $1/\varepsilon$, time intervals and employ these results in the next section for deriving stochastic resonance type assertions. Assume that Assumptions 3.2 and 3.3 hold true. Recall, that the topological pressure of a continuous function ψ for the flow F_x^t satisfies the following variational principle (see, for instance, [17]),

$$(4.1) \qquad P_x(\psi) = \sup_{\mu \in \mathcal{M}_x} \left(\int \psi d\mu + h_\mu(F_x^1) \right)$$

where $\mathcal{M}_x$ denotes the space of F_x^t-invariant probability measures on Λ_x and $h_\mu(F_x^1)$ 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

$$(4.2) \qquad P_x(\varphi_x^u + q) = \int (\varphi_x^u + q)d\mu_x^q + h_{\mu_x^q}(F_x^1).$$

We denote μ_x^0 by μ_x^{SRB} since it is the Sinai–Ruelle–Bowen (SRB) measure for F_x^t. Since Λ_x are attractors we have that $P_x(\varphi_x^u) = 0$ (see [9]).

For any probability measure ν on W define

$$(4.3) \qquad I_x(\nu) = \begin{cases} -\int \varphi_x^u d\nu - h_\nu(F_x^1) & \text{if } \nu \in \mathcal{M}_x \\ \infty & \text{otherwise.} \end{cases}$$

It is known that $h_\nu(F_x^1)$ is upper semicontinuous in ν since hyperbolic flows are entropy expansive (see [8]). Thus $I_x(\nu)$ is a lower semicontinuous functional in ν and it is also convex since entropy h_ν is affine in ν.

Denote by C_{0T} the space of continuous curves γ_t, $t \in [0,T]$ in $\mathbb{R}^d$ which is the space of continuous maps of $[0,T]$ into $\mathbb{R}^d$. For each absolutely continuous $\gamma \in C_{0T}$ set

$$(4.4) \qquad S_{0T}(\gamma) = \int_0^T \inf\{I_{\gamma_t}(\nu) : \dot{\gamma}_t = \bar{B}_\nu(\gamma_t), \nu \in \mathcal{M}_{\gamma_t}\}dt,$$

(where $\bar{B}_\nu(x) = \int B(x,y)d\nu(y)$), provided for Lebesgue almost all $t \in [0,T]$ there exists $\nu_t \in \mathcal{M}_{\gamma_t}$ for which $\dot{\gamma}_t = \bar{B}_{\nu_t}(\gamma_t)$, and $S_{0T}(\gamma) = \infty$ otherwise. It follows from [9] and [10] that

$$S_{0T}(\gamma) \geq S_{0T}(\gamma^u) = -\int_0^T P_{\gamma_t^u}(\varphi_{\gamma_t^u}^u)dt = 0$$

where γ_t^u is the unique solution of the equation

$$(4.5) \qquad \dot{\gamma}_t^u = \bar{B}(\gamma_t^u), \quad \gamma_0^u = x,$$

where $\bar{B}(z) = \bar{B}_{\mu_z^{SRB}}(z)$, and the equality $S_{0T}(\gamma) = 0$ holds true if and only if $\gamma = \gamma^u$.

Define the uniform metric on C_{0T} by

$$\rho_{0T}(\gamma, \eta) = \sup_{0 \leq t \leq T} |\gamma_t - \eta_t|$$

for any $\gamma, \eta \in C_{0T}$. Set $\Psi_{0T}^a(x) = \{\gamma \in C_{0T} : \gamma_0 = x, S_{0T}(\gamma) \leq a\}$. It follows from [10] and Section 9.1 of [16] that S_{0T} is a lower semicontinuous functional on C_{0T}, and so $\Psi_{0T}^a(x)$ is a closed set. The following result can be derived employing the machinery of [18], [23] and a certain modification of [28].

4.1. Theorem. *Suppose that $X_{x,y}^\varepsilon$ and $Y_{x,y}^\varepsilon$ are solutions of (1.1) with coefficients satisfying Assumptions 3.2 and 3.3. Set $Z_{x,y}^\varepsilon(t) = X_{x,y}^\varepsilon(t/\varepsilon)$ then for any $a, \delta, \lambda > 0$*

and every $\gamma \in C_{0T}$, $\gamma_0 = x$ there exists $\varepsilon_0 = \varepsilon_0(x, \gamma, a, \delta, \lambda) > 0$ such that for $\varepsilon < \varepsilon_0$,

$$(4.6) \qquad \mu_x \left\{ y \in \mathcal{W} : \rho_{0T}(Z^\varepsilon_{x,y}, \gamma) < \delta \right\} \geq \exp \left\{ -\frac{1}{\varepsilon}(S_{0T}(\gamma) + \lambda) \right\}$$

and

$$(4.7) \qquad \mu_x \left\{ y \in \mathcal{W} : \rho_{0T}(Z^\varepsilon_{x,y}, \Psi^a_{0T}(x)) \geq \delta \right\} \leq \exp \left\{ -\frac{1}{\varepsilon}(a - \lambda) \right\}$$

where either $\mu_x = m$ or $\mu_x = \mu^{SRB}_x$ and m is the normalized Riemannian volume on M. The functional $S_{0T}(\gamma)$ for $\gamma \in C_{0T}$ is finite if and only if $\dot\gamma_t = \bar{B}_{\nu_t}(\gamma_t)$ for $\nu_t \in \mathcal{M}_{\gamma_t}$ and Lebesgue almost all $t \in [0,T]$. Furthermore, $S_{0T}(\gamma)$ achieves its minimum 0 only on $\gamma \in C_{0T}$ satisfying $\dot\gamma_t = \bar{B}(\gamma_t)$ for all $t \in [0,T]$. In particular, for any $\delta > 0$ there exist $c(\delta) > 0$ and $\varepsilon_0 > 0$ such that for all $\varepsilon < \varepsilon_0$,

$$(4.8) \qquad m \left\{ y \in \mathcal{W} : \rho_{0T}(Z^\varepsilon_{x,y}, \bar{Z}_x) \geq \delta \right\} \leq \exp \left(-\frac{c(\delta)}{\varepsilon} \right)$$

where $\bar{Z}_x = \gamma^u$ is the unique solution of (4.5) which is another form of (3.12).

Emloying the machinery of [21] we obtain a similar result for the discrete time setup where $F_x = \Phi(x, \cdot)$ in (1.8) are C^2 expanding endomorphisms or Axiom A diffeomorphisms in a vicinity of a hyperbolic attractor.

Next, let $V \subset \mathcal{X}$ be a connected open set and set $\tau^\varepsilon_{x,y}(V) = \inf\{t \geq 0 : Z^\varepsilon_{x,y}(t) \notin V\}$ where we take $\tau^\varepsilon_{x,y}(V) = \infty$ if $X^\varepsilon_{x,y}(t) \in V$ for all $t \geq 0$.

4.2. Corollary. *Under the conditions of Theorem 4.1 for any $T > 0$ and $x \in V$,*

$$(4.9) \qquad \lim_{\varepsilon \to 0} \varepsilon \log m \left\{ y \in \mathcal{W} : \tau^\varepsilon_{x,y}(V) < T \right\}$$
$$= -\inf \left\{ S_{0t}(\gamma) : \gamma \in C_{0T}, t \in [0,T], \gamma_0 = x, \gamma_t \neq V \right\}.$$

Next, we will study exponentially long in $1/\varepsilon$ time behavior of Z^ε under certain assumptions on the averaged motion $\bar{Z}_x$. The arguments here rely on Theorem 4.1 and they follow the strategy of [13] and [18] but the fully coupled case is more involved though its main difficulties lie already in the proof of Theorem 4.1. Let $V \subset \mathbb{R}^d$ be a connected open set with a compact closure $\bar{V}$. Put

$$R_V(x, z) = \inf\{S_{0,T}(\gamma) : T \geq 0, \gamma \in C_{0,T}, \gamma \subset V, \gamma_0 = x, \gamma_T = z\}.$$

Let $\bar{F}^t x = \bar{Z}_x(t) = \bar{X}_x(t/\varepsilon)$ be the flow determined by the averaged system (1.5) with $\bar{B}(x) = \bar{B}_{\mu^{SRB}_x}(x)$ and assume that

$$(4.10) \qquad \bar{F}^t \bar{V} \subset V \quad \text{for all} \quad t > 0.$$

Suppose that the ω-limit set of the flow $\bar{F}^t$ in V is contained in a disjoint union of a finite number of compacts $K_1, ..., K_l \subset V$ such that $x, z \in K_i$ if and only if

$R(x, z) = R(z, x) = 0$. Among these compacts we specify the attractors $K_1, ..., K_k$ of the flow $\bar{F}^t$ which are characterized by the property that $R(x, z) > 0$ for any $x \in K_j$ and $z \notin K_j$, $j = 1, ..., k$. Suppose that transitions between each pair K_i and K_j are possible in the sense that there exist $T > 0$ and $\gamma \in C_{0T}$, $\gamma \subset V$ such that $\gamma_0 \in K_i$, $\gamma_T \in K_j$, and $\dot{\gamma}_t = \bar{B}_{\nu_t}(\gamma_t)$ for some $\nu_t \in \mathcal{M}_\Lambda^f$ and almost all $t \in [0, T]$. In this case $R_{ij} = R_V(x, z) < \infty$ for all $x \in K_i$, $z \in K_j$, $i, j = 1, ..., k$ and these numbers describe transitions between the compacts K_i, $i = 1, ..., k$ in the following way (introduced in [13]). Set $L = \{1, ..., k\}$. Given $i \in L \cup \{*\}$, a graph consisting of arrows $(m \to n)$ $(m \neq i, m, n \in L, n \neq m)$ is called an i-graph if every point $m \neq i$ is the origin of exactly one arrow and the graph has no circles. Let $\Gamma_i(Q)$ be the set of all i-graphs over $Q \subset L \cup \{*\}$. Denote by V_i the domains of attraction of K_i, $i = 1, ..., k$ and choose

$$\delta < \min\{\text{dist}(V_i, K_j) : i, j = 1, ..., k, i \neq j\}.$$

Put $\tau_{x,y}^\varepsilon(i, Q) = \inf\{t : \text{dist}(X_{x,y}^\varepsilon(t), \cup_{j \in L \setminus Q} K_j) < \delta\}$ where $x \in V_i$. Relying on Theorem 4.1 and employing the machinery of [18] enhanced for the fully coupled case it is possible to derive the following result.

4.3. Theorem. *Fix an arbitrary $Q \subset L$. Set $R_{i*} = \min_{j \notin Q} R_{ij}$ and assume that this minimum is achieved only at one point $j_Q(i) \notin Q$. Let γ^* be the unique $*$-graph for which*

$$(4.11) \qquad \min_{\gamma \in \Gamma_*(Q \cup \{*\})} \sum_{(m \to n) \in \gamma} R_{mn} = \sum_{(m \to n) \in \gamma^*} R_{mn} < \infty.$$

Then for any $x \in V_i$, $i \in Q$,

$$(4.12) \qquad \lim_{\varepsilon \to 0} m\{y \in \mathcal{W} : X_{x,y}^\varepsilon(\tau_{x,y}^\varepsilon(i, Q)) \notin V_{j_Q(k)}\} = 0,$$

where k is such that the arrow $(k \to *)$ is the last in the path going from i to $*$ in the $*$-graph γ^*. If $Q = \{i\}$ then $j = j_Q(i)$ satisfies $R_{ij} = \min_{l \in L \setminus \{i\}} R_{il}$ and

$$(4.13) \qquad \lim_{\varepsilon \to 0} \varepsilon \log \int_{\mathcal{W}} \tau_{x,y}^\varepsilon(i, Q) dm(y) = R_{ij},$$

$$(4.14) \qquad \lim_{\varepsilon \to 0} (m(\mathcal{W}))^{-1} m\left\{y \in \mathcal{W} : e^{\varepsilon^{-1}(R_{ij} - \alpha)} < \tau_{x,y}^\varepsilon(i, Q) < e^{\varepsilon^{-1}(R_{ij} + \alpha)}\right\} = 1$$

for any $\alpha > 0$.

Again, a discrete time version of this result with expanding or Axiom A maps F_x can be derived using the technique of [21].

Observe, that rare transitions between attractors of the averaged system were discussed in the framework of climate models in [11] and [15]. Time estimates for such transitions given in Theorem 4.3 play an important role in the stochastic resonance setup which we consider in the next section.

5. STOCHASTIC RESONANCE

Next, we will describe certain stochastic resonance type phenomenon where the slowest motion $W^{\varepsilon,\delta}$ in (1.2) becomes periodic. The scheme of this construction was suggested by M.Freidlin (cf. [12]) and the corresponding proofs are supposed to appear in our joint paper.

Set $\tilde{W}^{\varepsilon,\delta}(t) = W^{\varepsilon,\delta}(\frac{t}{\delta\varepsilon})$, $\tilde{X}^{\varepsilon,\delta}(t) = X^{\varepsilon,\delta}(\frac{t}{\delta\varepsilon})$, $\tilde{Y}^{\varepsilon,\delta}(t) = Y^{\varepsilon,\delta}(\frac{t}{\delta\varepsilon})$, and pass from (1.2) to the equations in the new time

$$
\begin{aligned}
\frac{d\tilde{W}^{\delta,\varepsilon}(t)}{dt} &= A(\tilde{W}^{\delta,\varepsilon}(t), \tilde{X}^{\delta,\varepsilon}(t), \tilde{Y}^{\delta,\varepsilon}(t)) \\
\frac{d\tilde{X}^{\delta,\varepsilon}(t)}{dt} &= \delta^{-1}B(\tilde{W}^{\delta,\varepsilon}(t), \tilde{X}^{\delta,\varepsilon}(t), \tilde{Y}^{\delta,\varepsilon}(t)) \\
\frac{d\tilde{Y}^{\delta,\varepsilon}(t)}{dt} &= (\delta\varepsilon)^{-1}b(\tilde{W}^{\delta,\varepsilon}(t), \tilde{X}^{\delta,\varepsilon}(t), \tilde{Y}^{\delta,\varepsilon}(t)),
\end{aligned}
\tag{5.1}
$$

Assume that the equation (1.2) satisfy the assumptions similar to Assumptions 3.2 and 3.3 (with $\mathbb{R}^l \times \mathbb{R}^d$ in place of $\mathbb{R}^d$), in particular, that $F^t_{w,x}y = Y^{0,0}_{w,x,y}(t)$ have a C^2 dependence on w, x, for all w, x they are Axiom A flows in a neighborhood $\mathcal{W}$ which contains a basic hyperbolic attractor $\Lambda_{w,x}$ for $F^t_{w,x}$ and $\mathcal{W}$ itself is contained in the basin of $\Lambda_{w,x}$. Set

$$
\bar{B}_w(x) = \bar{B}(w, x) = \int B(w, x, y) d\mu^{SRB}_{w,x}(y)
\tag{5.2}
$$

where $\mu^{SRB}_{w,x}$ is the SRB measure for $F^t_{w,x}$ and let $\bar{X}^{(w)}$ be the solution of the averaged equation

$$
\frac{d\bar{X}^{(w)}(t)}{dt} = \bar{B}_w(\bar{X}^{(w)}(t)).
\tag{5.3}
$$

First, we apply averaging and large deviations estimates in averaging from the previous section to two last equations in (5.1) freezing the slowest variable w (i.e. taking for a moment $\delta = 0$). Namely, set $\hat{X}(t) = X^{\varepsilon,0}_{w,x,y}(t/\varepsilon)$ and $\hat{Y}(t) = Y^{\varepsilon,0}_{w,x,y}(t/\varepsilon)$ so that

$$
\begin{aligned}
\frac{d\hat{X}(t)}{dt} &= B(w, \hat{X}(t), \hat{Y}(t)) \\
\frac{d\hat{Y}(t)}{dt} &= \varepsilon^{-1}b(w, \hat{X}(t), \hat{Y}(t)).
\end{aligned}
\tag{5.4}
$$

Suppose that $l = 1$, $d = 2$ (i.e. $W^{\varepsilon,\delta}$ is one dimensional and $X^{\varepsilon,\delta}$ is two dimensional) and that the solution $X^{(w)}(t)$ of (5.3) has the limit set consisting of two attracting points K^w_1 and K^w_1 and a separatrix (separating between their basins). Let $S^w_{0T}(\gamma)$, $\gamma \subset \mathbb{R}^d$ be the large deviations rate functional for the system (5.4) defined in the previous section (see (4.4) and set for $i, j = 1, 2$,

$$
R_{ij}(w) = \inf\{S^w_{0T}(\gamma) : \gamma \in C_{0T}, \gamma_0 = K^w_i, \gamma_T = K^w_j, T \geq 0\}
\tag{5.5}
$$

(cf. with R_{ij} in Theorem 4.3). Put $x_i = K_i^w$,

$$(5.6) \qquad \bar{B}_i(w) = \int B(w, x_i, y) d\mu_{w,x_i}^{SRB}(y)$$

and assume that for all w,

$$(5.7) \qquad \bar{B}_1(w) < 0 \quad \text{and} \quad \bar{B}_2(w) > 0$$

which means that $W_{w,x,y}^{\varepsilon,\delta}(t)$ decreases (increases) while $X_{w,x,y}^{\varepsilon,\delta}(t)$ stays close to K_1^w (to K_2^w) for "most" y's with respect to $\mu_{w,x}^{SRB}$ and also with respect to the Riemannian volume on M restricted to $\mathcal{W}$.

The proof of the following statement is not written yet with all details and so it is called here an assertion rather than a theorem.

5.1. Assertion. *Suppose that there exist strictly increasing and decreasing functions* $w_-(r)$ *and* $w_+(r)$, *respectively, so that*

$$R_{12}(w_-(r)) = R_{21}(w_+(r)) = r$$

and $w_-(\lambda) = w_+(\lambda) = w^*$ *for some* $\lambda > 0$ *while* $w_-(r) < w^* < w_+(r)$ *for* $r < \lambda$. *Assume that* $\delta \to 0$ *and* $\varepsilon \to 0$ *in such a way that*

$$(5.8) \qquad \lim_{\varepsilon, \delta \to 0} \delta \varepsilon \ln \varepsilon^{-1} = \rho < \Lambda.$$

Then for any w, x *there exists* $t_0 > 0$ *so that the slowest motion* $W_{w,x,y}^{\varepsilon,\delta}(t + t_0)$, $t \geq 0$ *converges weakly (as* $\varepsilon, \delta \to 0$ *so that (5.8) holds true) as a random process on the probability space* $(M, \mu_{w,x}^{SRB})$ *(or on* $(\mathcal{W}, m_{\mathcal{W}})$ *where* $m_{\mathcal{W}}$ *is the normalized Riemannian volume on* $\mathcal{W}$) *to a periodic function* $\Psi(t)$, $\Psi(t + T) = \Psi(t)$ *with*

$$T = T(\rho) = \int_{w_-(\rho)}^{w_+(\rho)} \frac{dw}{|\bar{B}_1(w)|} + \int_{w_-(\rho)}^{w_+(\rho)} \frac{dw}{|\bar{B}_2(w)|}.$$

A heuristic explanation of this result is the following. When the intermediate motion $\tilde{X}^{\varepsilon,\delta}$ is close to K_1^w the slowest motion $W^{\varepsilon,\delta}$ decreases until $w = w_-(\rho)$ where $R_{12}(w) = \rho$. In view of (4.14) and the scaling (5.8) between ε and δ, a moment later $R_{12}(w)$ becomes less than ρ and $\tilde{X}^{\varepsilon,\rho}$ jumps immediately close to K_2^w. There $\bar{B}_2(w) > 0$, and so $W^{\varepsilon,\delta}$ starts to grow until it reaches $w = w_+(\rho)$ where $R_{21}(w) = \rho$. A moment later $R_{21}(w)$ becomes smaller than ρ and in view of (4.14) $\tilde{X}^{\varepsilon,\delta}$ jumps immediately close to K_1^w. This leads to a close to periodic behavior of $W^{\varepsilon,\delta}$.

6. LIMIT THEOREMS

In this section we return to the system (1.1) under Assumptions 3.2 and 3.3 and will discuss limit theorems type results such as a Gaussian and a diffusion approximations of the slow motion and moderate deviations for the error in averaging.

These results have been proved for the uncoupled case but in the fully coupled case they mostly remain as conjectures.

It follows from [19] that under Assumptions 3.2 and 3.3 for each $x \in \mathbb{R}^d$ and $i, j = 1, ..., d$ the limit

$$(6.1) \qquad a_{ij}(x) = \lim_{t \to \infty} \tfrac{1}{t} \int_{\Lambda_x} d\mu_x^{SRB}(y) \left(\int_0^t (B_i(x, F_x^u y) - \bar{B}_i(x)) du \right.$$

$$\left. \int_0^t (B_j(x, F_x^v y) - \bar{B}_j(x)) dv \right)$$

exists and the matrix $a(x) = \left(a_{ij}(x) \right)_{i,j=1,...,d}$ is nonnegative definite. Moreover, combining [19] and [10] we conclude that $a(x)$ is C^2 in x and there exists a Lipschitz continuous symmetric matrix $\sigma(x)$ such that $\sigma^2(x) = a(x)$ (cf. [22]). Set $Z_{x,y}^\varepsilon(t) = X_{x,y}^\varepsilon(t/\varepsilon)$ and $\bar{Z}_x(t) = \bar{X}_x^\varepsilon(t/\varepsilon)$ where $\bar{X}^\varepsilon$ satisfies (1.5) with $\bar{B}(x) = \int B(x, y) d\mu_x^{SRB}(y)$ and notice that $\bar{Z}$ does not depend on ε. For each fixed x define the stochastic process

$$(6.2) \qquad \xi_x^\varepsilon(t, y) = \varepsilon^{-1/2}\left(Z_{x,y}^\varepsilon(t) - \bar{Z}_x(t) \right), \quad t \in [0, T], \quad y \in W$$

on the probability space (W, m_W) with W introduced in Assumption 3.2 and m_W being the normalized Riemannian volume there.

6.1. Conjecture. *For each fixed x the process $\xi_x^\varepsilon(t, \cdot)$, $t \in [0, T]$ weakly converges as $\varepsilon \to 0$ to a $\mathbb{R}^d$-valued Gaussian Markov process $\xi_x^0(t)$ on (W, m) satisfying the equation*

$$(6.3) \qquad \xi_x^0(t) = G_x^0(t) + \int_0^t \nabla\bar{B}(\bar{Z}_x(s))\xi_x^0(s) ds$$

where $\left(\nabla\bar{B}(x) \right)_{ij} = \frac{\partial \bar{B}_i(x)}{\partial x_j}$ and $G_x^0(t)$ is a Gaussian process with independent increments, zero expectation and the covariance matrix $\int_0^t a(\bar{Z}_x(s)) ds$.

In the uncoupled case this result was proved in [19] with μ^{SRB} in place of the Riemannian volume but it can be obtained for the latter, as well. In the probabilistic setup when fast motions are nondegenerate diffusion processes in place of Axiom A flows this result follows from [4], [5] and [6]. In the discrete time fully coupled setup (1.8) (with $F_x = \Phi(x, \cdot)$ being Axiom A diffeomorphisms or expanding transformations) the corresponding counterpart of the assertion of Conjecture 6.1 can be derived by a slight extension of arguments from [4] and [5]. In the continuous time fully coupled setup the assertion is highly plausible but there are substantial technical difficulties to justify it rigorously.

Next we will discuss Hasselmann's diffusion approximation of the time changed slow motion $Z_{x,y}^\varepsilon(t)$ which was suggested in [14].

6.2. Conjecture. *For each* $\varepsilon > 0$ *and* x *there exists a Brownian motion* $W(t)$ *(depending on* ε *and* x*) defined on the product probability space*

(6.4) $$(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}, Leb) \times (\mathcal{W}, \mathcal{B}, m_W)$$

(where $\mathcal{B}$ *is the corresponding Borel* σ*-field) such that if* $S^\varepsilon(t) = S_x^\varepsilon(t)$ *is the solution of the stochastic differential equation*

(6.5) $$dS^\varepsilon(t) = \bar{B}(S^\varepsilon(t))dt + \sqrt{\varepsilon}\sigma(S^\varepsilon(t))dW(t), \quad S^\varepsilon(0) = x$$

and $Z_{x,\cdot}^\varepsilon(t)$ *is extended to the product (6.4) in the trivial way so that it does not depend on the first factor then*

(6.6) $$E \sup_{0 \le t \le T} |Z_{x,\cdot}^\varepsilon(t) - S_x^\varepsilon|^2 \le C_{\delta,T}\varepsilon^{1+\delta}$$

for any sufficiently small $\delta > 0$ *where* $C_{\delta,T} > 0$ *does not depend on* $\varepsilon > 0$.

Observe, that the diffusion S^ε provides a better approximation of the slow motion Z^ε than the averaged motion $\bar{Z}$ but, in fact, Hasselmann suggested this approximation in [14] hoping to employ it in the study of rare transitions of Z^ε (which represented climate in his model) between attractors of $\bar{Z}$ as described in Section 4. Alas, it turns out that Z^ε and S^ε have different large deviations rate functionals and the latter cannot describe the former on exponentially large in $1/\varepsilon$ time intervals (see [22] and [24]). This becomes especially clear if we observe that for each $t > 0$ the diffusion $S^\varepsilon(t)$ can be arbitrarily far away with a small but relevant for large deviations probability though $Z^\varepsilon(t)$ cannot be farther away from the initial point than Lt with L taken from (3.1). In the uncoupled case the above assertion was proved in [24] (see also [22]). In the fully coupled probabilistic setup with fast motions being nondegenerate diffusions the assertion of the last conjecture was proved in [6]. Combining the machinery of [4], [5], and [6] this assertion can be derived also in the fully coupled discrete time setup (with $F_x = \Phi(x, \cdot)$ in (1.8) being Axiom A diffeomorphisms or expanding transformations). Again, in the continuous time fully coupled setup formulated above the assertion is highly plausible but its proof is not known yet.

Finally, we will discuss moderate deviations. Namely, consider

(6.7) $$\xi_{x,y}^{\varepsilon,\kappa}(t) = \varepsilon^{\kappa-1}(Z_{x,y}^\varepsilon(t) - \bar{Z}_x(t)).$$

The case $\kappa = 1/2$ is considered in Conjecture 6.1. The study of the case $\kappa = 1$ leads to the large deviations setup. The intermediate case $1/2 < \kappa < 1$ corresponds to moderate deviations asymptotics. Assume that the matrix $a(z)$ defined by (6.1) is invertible for all z (it suffices to take z with $|z| < LT$ where L comes from Assumption 3.3). Set

(6.8) $$S_{0T}^x(\gamma) = \frac{1}{2}\int_0^T \left(a(\bar{Z}_x(t))^{-1}(\dot{\gamma}_t - \nabla\bar{B}(\bar{Z}_x(t))\gamma_t), (\dot{\gamma}_t - \nabla\bar{B}(\bar{Z}_x(t))\gamma_t)\right)$$

(where $(\cdot, \cdot)$ denotes the inner product) if γ_t is absolutely continuous in t and we put $S_{0T}^x(\gamma) = \infty$ for other γ's from the space C_{0T}^0 of continuous curves γ in $\mathbb{R}^d$ defined on $[0, T]$ with $\gamma_0 = 0$. For each $a \geq 0$ set

$$\Gamma_{0T}^{a,x} = \{\gamma \in C_{0T}^0 : S_{0T}^x(\gamma) \leq a\}$$

and let $\rho_{0T}(\gamma, \varphi) = \sup_{0 \leq t \leq T} |\gamma_t - \varphi_t|$.

6.3. Conjecture. *For any* $\kappa \in (1/2, 1)$, $a, \delta, \lambda > 0$, $x \in \mathbb{R}^d$ *and* $\gamma \in C_{0T}^0$ *there exists* $\varepsilon_0 > 0$ *such that for* $\varepsilon < \varepsilon_0$,

$$(6.9) \qquad m_W\{y : \rho_{0T}(\xi_{x,y}^{\varepsilon,\kappa}, \gamma) \leq \delta\} \geq \exp\left(-\varepsilon^{1-2\kappa}(S_{0T}^x(\gamma) + \lambda)\right)$$

and

$$(6.10) \qquad m_W\{y : \rho_{0T}(\xi_{x,y}^{\varepsilon,\kappa}, \Gamma_{0T}^{a,x}) \geq \delta\} \leq \exp\left(-\varepsilon^{1-2\kappa}(a - \lambda)\right).$$

In the uncoupled case this assertion was proved in [19]. In the fully coupled discrete time case for κ sufficiently close to $1/2$ the assertion can be derived employing the Cramer type asymptotics from [4] and [5]. Again, in the fully coupled continuous time case the assertion has not been justified rigorously as yet.

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] V.I. Arnold, V.V. Kozlov, A.I. Neishtadt, *Mathematical Aspects of Classical and Celestial Mechanics (Dynamical Systems III, V.I. Arnold ed., Encyclopedia Math. Sci., 3)*, (1988) Springer-Verlag, Berlin.

[3] Z. Artstein and A. Vigodner, *Singularly perturbed ordinary differential equations with dynamic limits*, Proc. Royal Soc. Edinburgh (Sec. A), **126** (1996), 541–569.

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

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

[6] V.I. Bakhtin and Yu. Kifer, *Diffusion approximation for slow motion in fully coupled averaging*, Prob. Theory and Rel. Fields, **129** (2004), 157–181.

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

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

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

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

[11] S. Corti, F. Molteni, and T.N. Palmer, *Signature of recent climate change in frequencies of natural atmospheric circulation regimes*, Nature **398** (1999), 799–802.

[12] M. Freidlin, *Quasi-deterministic approximation, metastability and stochastic resonance*, Physica D **137** (2000), 333–352.

[13] M.I. Freidlin and A.D. Wentzell, *Random Perturbations of Dynamical Systems*, 2nd ed., (1998), Springer–Verlag, New York.

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

[15] K. Hasselmann, *Linear and nonlinear signatures*, Nature **398** (1999), 755–756.

[16] A.D. Ioffe and V.M. Tikhomirov, *Theory of Extremal Problems*, (1979) North-Holland, Amsterdam.

[17] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, (1995), Cambridge Univ. Press, Cambridge.

[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*, in: *Stochastic Climate Models*, (eds. P.Imkeller and J.-S. von Storch), Progress in Probability **49** (2001), 171–178, Birkhäuser.

[21] Yu. Kifer, *Averaging in difference equations driven by dynamical systems*, in: *Geometric Methods in Dynamics*, (eds. W.de Melo, M.Viana, J.-C. Yoccoz), Asterisque **287** (2003), 103–123.

[22] Yu. Kifer, L^2 *Diffusion approximation for slow motion in averaging*, Stochastics and Dynamics, **3** (2003), 213–246.

[23] Yu. Kifer, *Averaging principle for fully coupled dynamical systems and large deviations*, Ergod. Th.& Dynam. Sys., **24** (2004), 847–871.

[24] Yu. Kifer, *Some recent advances in averaging*, in: *Modern Dynamical Systems and Applications*, (eds. M.Brin, B.Hasselblatt, Y.Pesin,) (2004) Cambridge Univ. Press, Cambridge.

[25] P. Lochak and C. Meunier, *Multiple Averaging for Classical Systems*, (1988), Springer-Verlag, New York.

[26] D. Monrad and W. Philipp, *Nearby variables with nearby conditional laws and a strong approximation theorem for Hilbert space valued martingales*, Probab. Th. Rel. Fields **88** (1991), 381–404.

[27] J.A. Sanders and F.Verhurst, *Averaging Methods in Nonlinear Dynamical Systems*, (1985), Springer-Verlag, Berlin.

[28] A.Yu. Veretennikov, *On large deviations in the averaging principle for SDEs with "full dependence"*, Ann. Probab., **27** (1999), 284–296.

INSTITUTE OF MATHEMATICS, THE HEBREW UNIVERSITY, JERUSALEM 91904, ISRAEL
E-mail address: kifer@math.huji.ac.il