Geometric shadowing in slow–fast Hamiltonian systems

Niklas Brännström1, Emiliano de Simone1 and Vassili Gelfreich2

1 Department of Mathematics, PO Box 68, FIN-00014 University of Helsinki, Finland
2 Mathematics Institute, University of Warwick, CV4 7AL Coventry, UK

Received 16 July 2009, in final form 18 March 2010
Published 13 April 2010
Online at stacks.iop.org/Non/23/1169

Abstract
We study a class of slow–fast Hamiltonian systems with any finite number of degrees of freedom, but with at least one slow one and two fast ones. At \( \epsilon = 0 \) the slow dynamics is frozen. We assume that the frozen system (i.e. the unperturbed fast dynamics) has families of hyperbolic periodic orbits with transversal heteroclinics.

For each periodic orbit we define an action \( J \). This action may be viewed as an action Hamiltonian (in the slow variables). It has been shown in Brännström and Gelfreich (2008 Physica D 237 2913–21) that there are orbits of the full dynamics which shadow any finite combination of forward orbits of \( J \) for a time \( t = O(\epsilon^{-1}) \).

We introduce an assumption on the actions of periodic orbits which enables us to shadow any continuous curve (of arbitrary length) in the slow phase space for any time. The slow dynamics shadows the curve as a purely geometrical object, thus the time on the slow dynamics has to be reparametrized.

Mathematics Subject Classification: 37C50, 37J45, 70H14

1. Introduction

Let us consider the Hamiltonian system defined by the Hamiltonian function

\[
H = H(x, y, u, v; \epsilon)
\]

and the symplectic form

\[
\Omega = dy \wedge dx + \frac{1}{\epsilon} dv \wedge du,
\]

where \( x, y \in \mathbb{R}^{2m} \) and \( u, v \in \mathbb{R}^{2d} \). We assume that the parameter \( \epsilon \) is small, hence the Hamiltonian system is ’slow–fast’ with \( (x, y) \) being fast variables and \( (u, v) \) being slow ones.
This can readily be seen from the equations of motion
\[
\begin{align*}
\dot{x} &= \frac{\partial H}{\partial y}, & \dot{u} &= \varepsilon \frac{\partial H}{\partial v}, \\
\dot{y} &= -\frac{\partial H}{\partial x}, & \dot{v} &= -\varepsilon \frac{\partial H}{\partial u}.
\end{align*}
\] (1)

We note that the form of (1) is not unusual, it appears in many applications.

If we set \( \varepsilon = 0 \) the equations of motion are reduced to
\[
\begin{align*}
\dot{x} &= \frac{\partial H}{\partial y}, & \dot{u} &= 0, \\
\dot{y} &= -\frac{\partial H}{\partial x}, & \dot{v} &= 0.
\end{align*}
\] (2)

We refer to (2) as the frozen system.

We are interested in describing the slow dynamics of system (1). In general it is a difficult task to exactly describe the slow dynamics and therefore often only approximations, or averaged solutions, are sought. The procedure to obtain the averaged solutions, i.e. an averaging method, is typically tailored for the particular class of systems it is applied to. Averaging has been used under various assumptions on the fast dynamics, see, e.g., [1, 3, 5, 8, 9, 11]. Regardless of which averaging method is employed the aim is to derive effective equations for the slow dynamics which are independent of the fast variables \((x, y)\). We may write such averaged slow dynamics as
\[
\dot{u} = \varepsilon \left\langle \frac{\partial H}{\partial v} \right\rangle, & \dot{v} = -\varepsilon \left\langle \frac{\partial H}{\partial u} \right\rangle,
\] (3)

where \(\langle \cdot \rangle\) denotes averaging with respect to \((x, y)\). A crucial point in the justification of all averaging methods is to verify that time averages can be approximated by space averages. This property holds for most but not all trajectories. For example, if the fast dynamics spends most of its time close to a periodic orbit there is, in general, no reason to believe that the time average would converge to the space average.

In fact when the fast dynamics spends most of its time close to periodic orbits the slow dynamics of (1) may behave very differently from the averaged dynamics (3). In [4] it was shown that if the system satisfies assumptions [A1] and [A2] (defined later) then there exist trajectories of the full dynamics whose slow component deviates significantly from (3). In particular it was shown that there are orbits shadowing accessible paths of finite length composed of forward trajectories of the auxiliary systems
\[
\dot{u} = \varepsilon \left\langle \frac{\partial J_c}{\partial v} \right\rangle, & \dot{v} = -\varepsilon \left\langle \frac{\partial J_c}{\partial u} \right\rangle,
\] where \(J_c\) is an action defined on a periodic orbit (labelled by \(c\)) in the fast phase space (see definition 1). This shadowing result is valid on time-scales of order \(O(\varepsilon^{-1})\). We note that this method yields trajectories which deviate at the rate \(O(\varepsilon)\) from (3).

This approach is a generalization of the mechanism proposed in [6] for studying drift of the energy in a Hamiltonian system which depends on time explicitly and slowly. In this set-up it was shown that switching between fast periodic orbits does indeed provide the fastest possible rate of energy growth in several situations (see [6]). An interesting direct application of this theory is a rigorous proof of Fermi acceleration for a class of billiards with slowly moving boundary [7].

In this paper we describe a special class of systems where the upper bound on the time for which the shadowing result holds can be lifted, i.e. we provide a description of the shadowing
orbits for all times. Moreover, we show that for any continuous curve in the slow phase space there is a trajectory of the full dynamics whose slow component shadows it. We achieve this by refining the mechanism in [4, 6] such that for any two $O(\varepsilon)$ close points in the slow phase space there is a trajectory which starts in a neighbourhood of the first point and ends up in a neighbourhood of the second one. To prove this we have to introduce an assumption on the mutual relationship between the actions $J_c$. This assumption ensures that any vector in the tangent space of a point in the slow phase space can be written as a linear combination of gradients of the actions $J_c$ where all coefficients are positive. We identify this linear combination with a path in the slow phase space and denote it guiding path (which is a generalization of accessible path). We only consider guiding paths of length $O(\varepsilon)$ and show that for any such guiding path there exists a trajectory of the full dynamics whose slow component shadows it. Thus by approximating any given curve by a set of points which are $O(\varepsilon)$ apart we can find trajectories of the full dynamics whose slow component shadows a guiding path between these points. Then [6] implies that there exists a trajectory which lies close to the union of trajectories shadowing the guiding paths, hence there is a trajectory which shadows the entire curve. The idea of the proof is similar to that of [4], that is, using that the full dynamics takes place on normally hyperbolic manifolds where certain action Hamiltonians are preserved for long times.

The paper is structured as follows. In the next section we present the assumptions on the frozen system (2) and state our main result. Before proving this result we present an example where the assumptions on the frozen system are verified to hold. In section 3 we provide a short summary of the results from [4, 6] on normal hyperbolicity which we will require to prove our main theorem. In section 4 we state two lemmas which we then combine to prove theorem 1. The proofs of the two lemmas are of a more technical nature and have been postponed to section 5 to increase the readability of section 4.

2. Set-up and statement of the result

We impose a number of assumptions on the frozen system (2). Let $D \subset \mathbb{R}^{2d}$ be an open and bounded subset.

[A1] We assume that the frozen system (2) has $n$ families of hyperbolic periodic orbits $L_c(u, v), c \in \{c_1, \ldots, c_n\}$, defined for all $(u, v) \in D$.

[A2] We assume that each of the periodic orbits has a family of heteroclinic orbits to every other periodic orbit, i.e. for all $c_i, c_j \in \{c_1, \ldots, c_n\}$ and $(u, v) \in D$ there are a pair of transversal heteroclinic orbits

\[
\Gamma_{c_i c_j} (u, v) \subset W^u(L_{c_i} (u, v)) \cap W^s(L_{c_j} (u, v)),
\]

\[
\Gamma_{c_j c_i} (u, v) \subset W^u(L_{c_j} (u, v)) \cap W^s(L_{c_i} (u, v)).
\]

We note that under assumptions [A1] and [A2] the frozen system has a family of uniformly hyperbolic invariant transitive sets $\Lambda_{(u,v)}$, also known as Smale horseshoes (see for instance [10]). The dynamics on the Smale horseshoe can be described by symbolic dynamics. Let $\Lambda := \bigcup_{(u,v) \in D} \Lambda_{(u,v)}$. For each family of hyperbolic periodic orbits $L_c(u, v), c \in \{c_1, \ldots, c_n\}$, we define a family of actions.

**Definition 1.** The action $J_c$ of a periodic orbit $L_c$ is defined by the integral

\[
J_c(u, v) := \oint_{L_c(u,v)} y \, dx.
\]
The function $J_c(u, v)$ is independent of the fast variables and can be considered as a Hamiltonian function which generates some dynamics in the slow variables

\begin{equation}
\dot{u} = \frac{1}{T_c(u, v)} \frac{\partial J_c}{\partial v}, \quad \dot{v} = -\frac{1}{T_c(u, v)} \frac{\partial J_c}{\partial u},
\end{equation}

where $T_c$ is the period of the periodic orbit $L_c$. Following the notation of [12] we refer to (4) as the \textit{guiding system} and its equations of motion can be written in the concise form

\begin{equation}
\dot{z} = (\dot{u}, \dot{v}) = : X_c(z),
\end{equation}

where $X_c$ is the \textit{guiding vector field}.

[A3] We assume that there is a closed subset $\overline{D} \subset D \subset \mathbb{R}^{2d}$ such that for all $(u, v) \in \overline{D}$, $0$ is inside the convex envelope (or convex hull) of the vectors from the set $\{\nabla_{(u, v)} J_{c_i} : i = 1, \ldots, n\}$.

**Remark 1.** Assumption [A3] has a simple geometrical interpretation: any vector $\vec{w} \in \mathbb{R}^{2d}$ can be written as a linear combination

\begin{equation}
\vec{w} = \sum_{k=1}^{n} a_k \nabla_{(u, v)} J_{c_i}
\end{equation}

with $a_k \geqslant 0$ for all $k$,

where the coefficients $a_k$ may depend on the point $(u, v) \in \overline{D}$.

It is obvious that at least $n = 2d + 1$ actions are required to satisfy this assumption. Allowing $n \geqslant 2d + 1$ we extend the class of systems that satisfy assumption [A3] without introducing any complications to the proof.

We note that assumptions [A1] and [A2] are quite commonly satisfied in Hamiltonian dynamics. In particular, this is true if the fast system is uniformly hyperbolic. In some perturbative situations the standard Melnikov method can be used to verify these assumptions. On the other hand assumption [A3] substantially restricts our class of systems. In the next section we provide an example of a system which satisfies this assumption. Finally we note that assumptions [A1]–[A3] are stated in the form suitable for numerical verification as they involve periodic and transversal heteroclinic orbits only.

We can now state the main result. Let $\pi : \mathbb{R}^{2m+2d} \to \mathbb{R}^{2d}$ be the projection on the slow variables.

**Theorem 1.** Let $\gamma : \mathbb{R}^+ \to \overline{D}$ be any continuous curve. Assume that the frozen system (2) satisfies assumptions [A1], [A2] and [A3]. Then there exist positive constants $C$ and $\varepsilon_0$ such that for any $0 < \varepsilon < \varepsilon_0$ there exists a trajectory $\psi(t)$ of the full system (1) and a continuous monotone reparametrization of time $T(t)$ such that the slow component $z(t) := \pi \psi(t)$ of $\psi(t)$ satisfies $z(0) = \gamma(0)$ and

\begin{equation}
\|z(T(t)) - \gamma(t)\| \leqslant C \varepsilon,
\end{equation}

for all $t \geqslant 0$.

Before proving the theorem we construct an example where assumptions [A1], [A2] and [A3] are shown to be satisfied.

2.1. Example
Let $\varepsilon = 0$ and let $\mu$ be a small parameter. Consider a Hamiltonian function of the form

\begin{equation}
H(x, y) = H_0(x, y) - \mu H_1(x, y),
\end{equation}
where \((x, y) \in \mathbb{R}^4\). Assume that the dynamics of the ‘unperturbed’ Hamiltonian \(H_0\) has three hyperbolic periodic orbits \(L_c, c \in \{c_1, c_2, c_3\}\) with transversal heteroclinic connections (hence assumptions [A1] and [A2] are satisfied). The implicit function theorem implies that the hyperbolic periodic orbits persist for \(\mu\) sufficiently small, and furthermore the periodic orbits depend smoothly on \(\mu\). Therefore we can expand the periodic orbit \(L_c\) as well as the action \(J_c\) in a power series in \(\mu\). For the action we write

\[
J_c = J_c^0 + \mu J_c^1 + O(\mu^2),
\]

where \(J_0\) is the action of the periodic orbits of the ‘unperturbed’ Hamiltonian \(H_0\), \(J_0\) and \(J_1\) are constants as the Hamiltonian is independent of \((u, v)\). The action \(J_0\) and its first order correction \(J_1\) are generically non-zero. Now, by abusing the notation slightly we let \(\mu\) depend on \((u, v)\) in the following way

\[
\mu = \mu(x, y, u, v) = \mu \sum_{i=1}^{3} \chi_{L_c+i\delta}(x, y)\phi_i(u, v) + \mu^3 \sum_{i=1}^{3} \chi_{L_c+i\delta}(x, y)f(x, y),
\]

where \(\mu\) is a small parameter, \(\chi\) is the indicator function, \(L_c+i\delta\) is a \(\delta\)-neighbourhood of \(L_c\), i.e.,

\[
L_c+i\delta := \{(x, y) \text{ s.t. dist}((x, y), L_c) \leq \delta\}
\]

and \(f\) is so chosen as to interpolate \(H_1\) into a \(C^\infty\) function. Then the action of the periodic orbit \(L_c\) becomes

\[
J_c(u, v) = J_c^0 + \mu \phi_i(u, v)J_i^1 + O(\mu^2).
\]

For example, by choosing

\[
\phi_1 = \mathrm{sgn}(J_1^1)v, \\
\phi_2 = -\mathrm{sgn}(J_2^1)u, \\
\phi_3 = \mathrm{sgn}(J_3^1)u - \mathrm{sgn}(J_1^1)v,
\]

where \(\mathrm{sgn}(A) := 1\) if \(A > 0\) and \(\mathrm{sgn}(A) := -1\) if \(A < 0\), will generate

\[
\nabla_{(u,v)}J_1^1 = \mu \vec{e}_u, \\
\nabla_{(u,v)}J_2^1 = \mu \vec{e}_v, \\
\nabla_{(u,v)}J_3^1 = -\mu \vec{e}_u - \mu \vec{e}_v.
\]

This choice ensures that assumption [A3] is satisfied.

3. Actions, normal hyperbolicity and symbolic dynamics

For the frozen system (2) the families of hyperbolic periodic orbits \(L_c\) form normally hyperbolic invariant manifolds. By analysing the dynamics in a vicinity of the periodic orbits and the heteroclinic connections using Poincaré sections and maps it was shown in [4, 6] that the dynamics in this part of the phase space can be described using symbolic dynamics. That is, there is a one-to-one correspondence between codes \(\xi\) and trajectories of the frozen system. It has also been shown that these normally hyperbolic manifolds persist for \(\epsilon\) small allowing for the dynamics of the full system (1) to be described using symbolic dynamics. The theory of normal hyperbolicity gives us existence of trajectories in the vicinity of the hyperbolic periodic
orbits (and the heteroclinic connections) and the symbolic dynamics gives us a means to study them.

These arguments help us to find trajectories which stay close to the families $L_c$ and switch between different families in a prescribed order. When a trajectory of the full system stays $\varepsilon$-close to $L_c$ for a long time its slow component evolves like a solution of the guiding equation (4).

The following arguments closely follow the paper [4].

In order to find the change in slow variable after one rotation near $L_c$ we integrate the slow component of the vector field along the exact trajectory and use (1) to conclude

$$\bar{u} - u = \int_0^{\Delta t} \dot{u} \, dt = \varepsilon \int_0^{T_c(u,v)} \frac{\partial H}{\partial v} \bigg|_{x_c(t,v,u),y_c(t,v,u),u,v} \, dt + O(\varepsilon^2),$$

$$\bar{v} - v = \int_0^{\Delta t} \dot{v} \, dt = -\varepsilon \int_0^{T_c(u,v)} \frac{\partial H}{\partial u} \bigg|_{x_c(t,v,u),y_c(t,v,u),u,v} \, dt + O(\varepsilon^2),$$

(6)

where the error terms come from replacing the exact trajectory by the frozen periodic trajectory $(x_c, y_c)$ and from the difference between the return time $\Delta t$ and the period of the frozen fast trajectory $T_c$. The integrals in the right-hand side can be expressed in terms of the action $J_c$

$$J_c = \int_{L_c} y \, dx = \int_0^{T_c} y_c \, \dot{x}_c \, dt.$$

Indeed, differentiating $J_c$ with respect to $u$, integrating by parts and taking into account periodicity and the equations of motion, we get

$$\frac{\partial J_c}{\partial u} = \int_0^{T_c} \left( \frac{\partial y_c}{\partial u} \frac{\partial x_c}{\partial t} - \frac{\partial x_c}{\partial u} \frac{\partial y_c}{\partial t} \right) dt = \int_0^{T_c} \left( \frac{\partial y_c}{\partial u} \frac{\partial H}{\partial y} + \frac{\partial x_c}{\partial u} \frac{\partial H}{\partial x} \right) dt.$$

Then differentiating identity $H(x_c, y_c, u, v) = \text{const}$ we get

$$\frac{\partial J_c}{\partial u} = \int_0^{T_c} \frac{\partial H}{\partial u} \bigg|_{x_c(t,v,u),y_c(t,v,u),u,v} \, dt.$$

Repeating these arguments with $u$ replaced by $v$ we also get

$$\frac{\partial J_c}{\partial v} = \int_0^{T_c} \frac{\partial H}{\partial v} \bigg|_{x_c(t,v,u),y_c(t,v,u),u,v} \, dt.$$

Substituting the last two equalities into (6) we arrive at

$$\bar{u} = u - \varepsilon \frac{\partial J_c}{\partial v} + O(\varepsilon^2), \quad \bar{v} = v + \varepsilon \frac{\partial J_c}{\partial u} + O(\varepsilon^2).$$

(7)

Taking into account that the time taken by one rotation is approximately $T_c(u,v)$ we see that the slow dynamics is approximated by solutions of the guiding equation.

Now let us summarize the results of [4, 6] which we will use to prove theorem 1. This will also introduce the notation which we will use in the subsequent sections. For each periodic orbit $L_c$ we denote by $\Sigma_c$ a Poincaré section and by $x_i, y_i, z_i$ (where $z_i = (u_i, v_i)$) we denote the $i$th intersection of a solution of (1) with one of these Poincaré sections. Let $\xi = \{\xi_i\}_{i=-\infty}^{\infty}$ be a bi-infinite sequence of letters $\xi_i \in \{c_1, \ldots, c_n\}$ where $c_i$ is the index of the $n$ actions. The sequence $\xi$ is called the code of the dynamics. If, for example, $\xi_i = c_2$ and $\xi_{i+1} = c_2$ then for some time $t_i$ the fast dynamics will hit the Poincaré surface $\Sigma_{c_2}$ and then make one round close to $L_{c_2}$ before hitting $\Sigma_{c_2}$ again. It has been shown that by specifying an initial condition for the slow dynamics the code $\xi$ generates a trajectory of (1) whose slow dynamics on the Poincaré sections is given by

$$z_{i+1} = z_i + \varepsilon \phi_{\xi_i,\xi_{i+1}}(x_i(z_i, \xi, \varepsilon), y_i(z_i, \xi, \varepsilon), z_i, \xi),$$

(8)
where $\phi_{\xi,\xi_i} \in C^1$ is the displacement of the slow variable $z = (u, v)$ between two consecutive hits of the fast dynamics with the Poincaré surfaces, that is $z_i \in \Sigma_1$. Note that the functions $x_i$ and $y_i$ depend on the entire code $\xi$ not just the current element $\xi_i$ and the next element $\xi_{i+1}$.

Lemma 2 in [4] implies that for $\epsilon = 0$ and any two codes $\xi^{(1)}$ and $\xi^{(2)}$ that satisfy $\xi^{(1)}_i = \xi^{(2)}_i$ for $|i| \leq n$ for any $n$ the following estimate holds:

$$\max \{ \|x_i(z, \xi^{(1)}) - x_i(z, \xi^{(2)})\|, \|y_i(z, \xi^{(1)}) - y_i(z, \xi^{(2)})\| \} \leq 2r\lambda^{-|i|},$$

where the constants $r > 0$ and $0 < \lambda < 1$ do not depend on the sequences $\xi^{(1)}, \xi^{(2)}$. For $\epsilon$ small lemma 3 in [4] implies the functions $(x_i, y_i)$ are defined for all small $\epsilon$ and all $z \in T$, that they are uniformly bounded along with their first derivatives with respect to $z$ and satisfy (9). Moreover, by the lemma, there is a constant $C_0 > 0$, independent of the code $\xi$, such that

$$\|x_i(z, \xi, \epsilon) - x_i(z, \xi, 0), y_i(z, \xi, \epsilon) - y_i(z, \xi, 0)\| < C_0\epsilon,$$

for all $i \in \mathbb{Z}$.

4. Proof of theorem 1

In this section we construct a code $\xi^{(\ast)}$ that generates a trajectory $z^{(\ast)}$ of the full dynamics (8) which satisfies theorem 1. The proof relies on two lemmas. The first one states that two trajectories of (8) which have a slow iterate in common stay uniformly close to each other as long as their respective codes coincide (lemma 1). The second one states that given any two points in the slow phase space which are $O(\epsilon)$ close we can find a trajectory which goes from the neighbourhood of the first point to the neighbourhood of the second point. This trajectory is essentially obtained by updating the code of another trajectory. We then combine these two lemmas to give an inductive proof of theorem 1.

4.1. Uniform closeness of trajectories

As mentioned in the previous paragraph, finding a trajectory that satisfies theorem 1 is an iterative process which involves updating the code. The crucial point here is that when we update the code we do not only alter the future of the trajectory but we switch to another trajectory. Differently put, as we update the code the entire slow dynamics changes (not only the ‘future’ iterates of the trajectory); in fact if one considers the description of the slow component of the full dynamics given in equation (8) one sees that the first two arguments of the function $\phi_{\xi,\xi_i}$ depend on the entire code $\xi$ (not just on the current and next to current code elements). The following lemma gives us a uniform estimate on how much the ‘past’ of the trajectory changes when the code is updated and the initial condition is kept fixed, i.e. $z_0^{(a)} = z_0^{(b)}$. The estimate appears in [6] (see the paragraph between equation (56) and (57) therein) but since the proof was only sketched and the result is crucial to our theory we state it together with a full proof.

Lemma 1. There exists $\epsilon_0 > 0$ such that for any $0 < \epsilon < \epsilon_0$ the following holds: if $\xi^{(a)}$ and $\xi^{(b)}$ are two codes such that for some $N > 0$

$$\xi_i^{(a)} = \xi_i^{(b)} \quad \text{for all } |i| \leq N$$

and $z^{(a)}$ and $z^{(b)}$ are two slow trajectories of (8) which correspond to the codes $\xi^{(a)}$ and $\xi^{(b)}$, respectively, and $z_0^{(a)} = z_0^{(b)}$ then

$$\|z_i^{(a)} - z_i^{(b)}\| \leq K\epsilon \quad \text{for all } 0 \leq i \leq N,$$

(11)
where

$$K := \frac{8 \|\phi\|_{C^1} r \lambda}{1 - \lambda}$$

(12)

and $\|\phi\|_{C^1} := \max_{\xi, c_1} \|\phi_{\xi, c_1}\|_{C^1}$.

Note that the constant $K$ is independent of $N$. We postpone the proof of this lemma to section 5.1.

With $K$ given by (12) we let

$$A_1 = \max_{c} \sup_{(u, v) \in \mathcal{D}} 3\varepsilon (K + \|\phi_{c, c}\|_{C^1} \frac{4r}{1 - \lambda} + 1 + |T_e(u, v)| |X_e(u, v)|),$$

where by $\phi_{c, c}$ we denote $\phi_{\xi, c_1}$ with $\xi_i = c$ for all $i$. We define recursively

$$A_i := \max_{c} \sup_{(u, v) \in \mathcal{D}} 3\varepsilon (A_{i-1} + \|\phi_{c, c}\|_{C^1} \frac{4r}{1 - \lambda} + 1 + |T_e(u, v)| |X_e(u, v)|).$$

Let

$$A := A_{2d},$$

(13)

where $d$ is the number of slow degrees of freedom.

**Lemma 2.** Let $L > 0$ be any fixed constant and $A$ given by (13). There exists $\varepsilon_0 > 0$ such that for any $0 < \varepsilon < \varepsilon_0$ and any $z \in \mathcal{D}$ and any code $\xi^{(a)}$ with a corresponding trajectory $z^{(a)}$ satisfying

$$\|z^{(a)}_0 - z\| \leq \epsilon (L + A)$$

the following holds: There exists another code $\xi^{(b)}$ with

$$\xi^{(b)}_i = \xi^{(a)}_i \quad i < 0$$

such that for any $p \leq 0$ there exists $N \in \mathbb{N}$ and a corresponding trajectory $z^{(b)}$ which satisfies

$$z^{(a)}_p = z^{(b)}_p,$$

$$\|z^{(b)}_i - z\| \leq \epsilon A.$$

Moreover there exists a uniformly bounded constant $C_1 > 0$ such that

$$\|z^{(b)}_i - z\| \leq \epsilon (C_1 + A) \quad \text{for } 0 \leq i \leq N.$$

(14)

Next we combine these two lemmas to show that we can shadow any curve $\gamma : \mathbb{R}^+ \rightarrow \overline{D}$.

### 4.2. Combining the results

**Proof of theorem 1.** Let $\gamma : \mathbb{R}^+ \rightarrow \overline{D}$ be any curve. Pick a constant $L > 0$ arbitrarily. Take the smallest $\varepsilon_0$ of lemmas 1 and 2. The following analysis is valid for any $0 < \varepsilon < \varepsilon_0$. Using $L$ we define a sequence $t_i$, $i \in \mathbb{N}$, as follows:

$$\begin{align*}
t_0 &= 0, \\
t_{i+1} &= \min_{t > t_i} \{ t : \|\gamma(t) - \gamma(t_i)\| = \varepsilon L \}
\end{align*}$$

(15)

and if for some $k$

$$\|\gamma(t) - \gamma(t_k)\| \leq \varepsilon L \quad \text{for all } t > t_k$$

then

$$t_{i+1} = t_i + 1 \quad \text{for all } i \geq k.$$

The sequence $t_i$ divides the curve $\gamma$ into points $\gamma(t_i)$ which are at most $\varepsilon L$ apart. Next we take any code $\xi^{(0)}$ and a corresponding trajectory $z^{(0)}$ such that $z^{(0)}_0 = \gamma(t_0) = \gamma(0)$. We also define $P(0) = 0$. 

Inductive assumption. There exists a code \( \xi^{(l)} \) and a monotone sequence \( P(l) \) such that for all \( 1 \leq l \leq k \)
\[
\xi_j^{(l)} = \xi_j^{(l-1)} \quad \text{for } j < P(l - 1).
\]
Moreover, there is a trajectory \( z^{(l)} \) corresponding to \( \xi^{(l)} \) such that
\[
z_0^{(l)} = z_0^{(0)}
\]
and
\[
\left\| z_j^{(l)} - \gamma(t_j) \right\| \leq \varepsilon A. \tag{16}
\]
Furthermore
\[
\left\| z_j^{(l)} - \gamma(t_j) \right\| \leq \varepsilon(C_1 + A) \quad \text{for } P(l - 1) \leq j \leq P(l), \tag{17}
\]
where \( C_1 \) is given by lemma 2.

Inductive step. We will use lemma 2 to verify the inductive assumption. Set \( p = -P(k) \),
\( z = \gamma(t_{k+1}) \) and \( z_i^{(a)} = z_{P(k)+i}^{(b)} \) then using (16) and (15) we get
\[
\left\| z_0^{(a)} - z \right\| \leq \varepsilon(A + L).
\]
Then applying lemma 2 implies that there exists a code \( \xi_i^{(k+1)} := \xi_i^{(b)} \) which satisfies
\[
\xi_i^{(k+1)} = \xi_i^{(k)} \quad \text{for } i < P(k)
\]
and generates a trajectory \( z_i^{(k+1)} \) which satisfies
\[
z_0^{(k+1)} = z_0^{(k)} = \gamma(0).
\]
Furthermore there exists \( N_{k+1} \) such that
\[
\left\| z_{P(k)+N_{k+1}}^{(k+1)} - z \right\| \leq \varepsilon A
\]
and
\[
\left\| z_j^{(k+1)} - \gamma(t_{k+1}) \right\| \leq \varepsilon(C_1 + A) \quad \text{for } P(k) \leq j \leq P(k + 1).
\]
The monotone sequence \( P(k) \) is defined inductively by
\[
P(0) = 0,
P(k) = P(k - 1) + N_k.
\]
This concludes the inductive step.

By induction there exists a unique code \( \xi^{(*)} \) which for all \( k > 0 \) satisfies
\[
\xi_i^{(*)} = \xi_i^{(k)} \quad \text{for } i \leq P(k).
\]
Denote by \( z^{(*)} \) the trajectory of (8) which corresponds to the code \( \xi^{(*)} \) and satisfies \( z_0^{(*)} = \gamma(0) \).

Next we want to apply lemma 1 to show that for all \( k \) the trajectory \( z^{(*)} \) lies close to the points \( \gamma(t_i) \). Set \( \xi_i^{(a)} = \xi_i^{(*)}, \xi_i^{(b)} = \xi_i^{(k)} \) and \( N = P(k) \). Then lemma 1 implies
\[
\left\| z_i^{(*)} - z_i^{(k)} \right\| \leq \varepsilon K \quad \text{for } 0 \leq i \leq P(k).
\]
Combining this estimate with (17) gives
\[
\forall k > 0 \quad \left\| z_i^{(*)} - \gamma(t_i) \right\| \leq \varepsilon(K + C_1 + A) \quad \text{for } P(k - 1) \leq i \leq P(k). \tag{18}
\]
To conclude the proof we have to define the time reparametrization $T(t)$ of theorem 1 and pass from the discrete solution $z^{(*)}(t)$ to the time continuous $z^{(*)}(t)$. Let us begin with the time reparametrization. For $t_k \leq t \leq t_{k+1}$ we define

$$T_k(t) = \tau_{P(k)} + \frac{\tau_{P(k+1)} - \tau_{P(k)}}{t_{k+1} - t_k}(t - t_k),$$

where $\tau_i$ is the time $z^{(*)}(t)$ intersects the Poincaré surface $\Sigma_{z^{(*)}}$, i.e. $z^{(*)}(\tau_i) = z_i^{(*)}$. The complete $T(t)$ is obtained by gluing together all $T_k(t)$.

Next for $t_k \leq t \leq t_{k+1}$, and consequently $P(k) \leq i \leq P(k+1)$, we estimate the distance between $z^{(*)}(t)$ and $\gamma(t)$ as follows:

$$\|z^{(*)}(T(t)) - \gamma(t)\| \leq \|z^{(*)}(T(t)) - z^{(*)}(T(t_{k+1}))\| + \|z^{(*)}(T(t_{k+1})) - z^{(*)}(P_{(k+1)})\| + \|\gamma(t_{k+1}) - \gamma(t)\|.$$

Consider the right-hand side. The second term is 0 by definition, the third term is bounded by (18) and the fourth one by (15). The first term in the right-hand side we estimate as

$$\|z^{(*)}(T(t)) - z^{(*)}(T(t_{k+1}))\| \leq \sup_{(u,v) \in \mathcal{D}} \| \dot{z} \|_{C^1}(1 + \max(\|\frac{\partial x}{\partial z}\|, \|\frac{\partial y}{\partial z}\|)),$$

(19)

We note that the set of periodic orbits and transversal heteroclinics is compact. The full trajectory, whose projection on the slow phase space is $z^{(*)}$, lies in a compact neighbourhood of this set, therefore $\dot{z}$ given by (1) is uniformly bounded on this set: there exists a constant $C_2 > 0$, independent of $k$, such that

$$\|z^{(*)}(T(t)) - z^{(*)}(T(t_{k+1}))\| \leq \varepsilon C_2.$$

By choosing $C := C_2 + L + K + C_1 + A$ we have

$$\|z^{(*)}(T(t)) - \gamma(t)\| \leq \varepsilon C \quad \text{for } t > 0.$$

5. Technical results

In this section we include the proofs of lemmas 1 and 2 that we used to prove theorem 1.

5.1. Proof of lemma 1

Let $\varepsilon_0$ be sufficiently small for the normal hyperbolicity estimates to be valid and smaller than $\frac{1-\lambda}{\pi \xi^2}$, where

$$C = \sup_{(x,v) \in \mathcal{D}} \| \phi \|_{C^1} \left(1 + \max \left(\left\| \frac{\partial x}{\partial z} \right\|, \left\| \frac{\partial y}{\partial z} \right\| \right) \right).$$

(19)

Consider the two codes $\xi^{(a)}$ and $\xi^{(b)}$. By assumption we have

$$\xi^{(a)}_i = \xi^{(b)}_i \quad |i| \leq N,$$

and as a consequence of normal hyperbolicity that

$$\|x_i(z, \xi^{(a)}) - x_i(z, \xi^{(b)}), y_i(z, \xi^{(a)}) - y_i(z, \xi^{(b)})\| \leq 2\lambda |z^{N-|i|},$$

(20)

for all $|i| \leq N$, see (9). For convenience we repeat equation (8) which gives the slow component of the full dynamics generated by the code $\xi$

$$z_{i+1} = z_i + \varepsilon \phi_{\xi,xi}(z_i, \varepsilon; \xi), y_{i+1}(z_i, \varepsilon; \xi, z_i, \varepsilon).$$
We note that the functions $x_i$ and $y_i$ depend on the full code $\xi$ (not only on the current element in the code). Taking the difference of the $z$ components of the two trajectories coded by $\xi^{(a)}$ and $\xi^{(b)}$ and using that they have the same initial condition $\xi^{(a)}_0 = \xi^{(b)}_0$ we get

$$
\|x_k^{(a)} - x_k^{(b)}\| = \epsilon \left\| \sum_{i=0}^{k-1} \phi_{\xi^{(a)}_{i+1}}(x_i^{(a)}, \epsilon; \xi^{(a)}_i, y_{i+1}(\xi^{(a)}_i, \epsilon; \xi^{(a)}_i), \xi^{(a)}_i, \epsilon) - \phi_{\xi^{(b)}_{i+1}}(x_i^{(b)}, \epsilon; \xi^{(b)}_i, y_{i+1}(\xi^{(b)}_i, \epsilon; \xi^{(b)}_i), \xi^{(b)}_i, \epsilon) \right\|,
$$

Now, $\xi^{(a)}_i = \xi^{(b)}_i$ for all $|i| < N$, therefore $\phi_{\xi^{(a)}_{i+1}}$ and $\phi_{\xi^{(b)}_{i+1}}$ are the same functions in the interval we are studying. Using the mean value inequality and suppressing the dependence of $\epsilon$ in the notation we get

$$
\|x_k^{(a)} - x_k^{(b)}\| \leq \epsilon \|\phi\|_{C^1} \left\| \sum_{i=0}^{k-1} \left( x_i^{(a)}(\xi^{(a)}_i, \xi^{(a)}_i) - x_i^{(b)}(\xi^{(b)}_i, \xi^{(b)}_i), y_{i+1}(\xi^{(a)}_i, \xi^{(a)}_i) - y_{i+1}(\xi^{(b)}_i, \xi^{(b)}_i) \right) \right\|
$$

By adding and subtracting terms and using the triangle inequality we rewrite this expression in a form where inequality (20) can be used

$$
\|x_k^{(a)} - x_k^{(b)}\| \leq \epsilon \|\phi\|_{C^1} \sum_{i=0}^{k-1} \left\| \left( x_i^{(a)}(\xi^{(a)}_i, \xi^{(a)}_i) - x_i^{(b)}(\xi^{(b)}_i, \xi^{(b)}_i), y_{i+1}(\xi^{(a)}_i, \xi^{(a)}_i) - y_{i+1}(\xi^{(b)}_i, \xi^{(b)}_i) \right) \right\|
$$

Using inequality (20) to estimate the first term and the mean value inequality to estimate the second term we get

$$
\|x_k^{(a)} - x_k^{(b)}\| \leq \epsilon \|\phi\|_{C^1} \sum_{i=0}^{k-1} 2r \lambda^{N-i} + \epsilon \|\phi\|_{C^1} \sum_{i=0}^{k-1} \left\| z_i^{(a)} - z_i^{(b)} \right\|
$$

which we rewrite as

$$
\|x_k^{(a)} - x_k^{(b)}\| \leq 2\epsilon \|\phi\|_{C^1} \frac{\lambda^{N-k+1}}{1-\lambda} + \epsilon C \sum_{i=0}^{k-1} \left\| z_i^{(a)} - z_i^{(b)} \right\|, \quad (21)
$$

where $C$ is given by (19). Now, in order to prove lemma 1 we will show that

$$
\|z_k^{(a)} - z_k^{(b)}\| \leq \epsilon K \lambda^{N-k}, \quad (22)
$$

for all $0 \leq k < N$ where $K$ is given by (11). The Gronwall type of estimate follows from inequality (21) by finite induction. Since $z_0^{(a)} = z_0^{(b)}$ the statement is valid for $k = 0$. Let us assume that (22) is true for all $k \leq m$. Consider the case $k = m + 1$. By (21) we get

$$
\|z_{m+1}^{(a)} - z_{m+1}^{(b)}\| \leq 2\epsilon \|\phi\|_{C^1} \frac{\lambda^{N-m}}{1-\lambda} + \epsilon C \sum_{i=0}^{m} \left\| z_i^{(a)} - z_i^{(b)} \right\|.
$$
The induction assumption implies that
\[
\|z_{m+1}^{(a)} - z_{m+1}^{(b)}\| \leq 2\varepsilon \|\phi\|_{C^1} r_{\varepsilon} \frac{\lambda^{N-m}}{1 - \lambda} + \varepsilon C \sum_{i=0}^{m} \varepsilon K \lambda^{N-i} \leq 2\varepsilon \|\phi\|_{C^1} r_{\varepsilon} \frac{\lambda^{N-m}}{1 - \lambda} + \varepsilon^2 C K \lambda^{N-m}.
\]
For
\[
K := \frac{8 \|\phi\|_{C^1} r_{\varepsilon} \lambda}{1 - \lambda}
\]
and \(0 < \varepsilon < \frac{1 - \lambda}{2C}\lambda\) we have
\[
K > 2 \|\phi\|_{C^1} r_{\varepsilon} \lambda = \frac{1 - \lambda}{1 - \lambda} - \varepsilon C \lambda.
\]
Thus it follows that
\[
\|z_{m+1}^{(a)} - z_{m+1}^{(b)}\| \leq \varepsilon K \lambda^{N-(m+1)}. \tag{23}
\]
Since \(0 < \lambda < 1\) and \(0 < m + 1 \leq N\) the lemma follows from (23). We note that \(K\) is independent of \(N\) and \(\varepsilon\). □

5.2. Proof of lemma 2

Fix \(L > 0\) arbitrarily and let \(A\) be given by (13). Then we choose \(\varepsilon_0\) as the smallest of that required by lemma 1 and
\[
C = \frac{1}{\sup_{(u,v) \in \mathcal{D}_L} \max_{(u,v) \in [0,1]} \sum_{j=1}^{L+A} \|T_{(u,v)} \|_{\mathcal{D}_L}^{1/2}}
\]
where \(C\) is given by (19). Then for any \(\varepsilon < \varepsilon_0\) let \(\xi^{(a)}\) be a code generating a trajectory of (8) such that
\[
\|z_0^{(a)} - z\| \leq \varepsilon (L + A).
\]
To show that there exists a code \(\xi^{(b)}\) that satisfies the lemma we introduce the notion of guiding path. Denoting by \(T_{z_0}^{(a)}(\mathbb{R}^{2d})\) the tangent space of \(\mathbb{R}^{2d}\) at the point \(z_0^{(a)}\), consider the vector \(\vec{v} \in T_{z_0}^{(a)}(\mathbb{R}^{2d})\) pointing towards \(z\) with length \(\varepsilon (L + A)\). Condition [A3] implies that there exists (a possibly non-unique) injective map \(\sigma : \{1, \ldots, 2d\} \rightarrow \{c_1, \ldots, c_n\}\) (where \(c_i\) is the index of the periodic orbits), with \(\sigma = \sigma(\vec{v})\) and non-negative coefficients \(a_1, \ldots, a_{2d}\) such that
\[
\vec{v} = \sum_{i=1}^{2d} \varepsilon a_i \vec{X}_{\sigma(i)}(z_0^{(a)}). \tag{24}
\]
We refer to the coefficients \(a_i\) as guiding times. We denote\(^3\) the guiding path between \(z_0^{(a)}\) and \(z\) as
\[
G(t, z_0^{(a)}, z) = z_0^{(a)} \left( t - \sum_{j=1}^{i} a_j \right) + \varepsilon \left( \sum_{j=1}^{i} a_j \right) \vec{X}_{\sigma(i)}(z_0^{(a)}) \tag{25}
\]
for \(\varepsilon \sum_{j=1}^{i} a_j < t < \varepsilon \sum_{j=1}^{i+1} a_j\). Let us denote by \(W_{\vec{v}} := \{\vec{X}_{\sigma(1)}(z_0^{(a)}), \ldots, \vec{X}_{\sigma(2d)}(z_0^{(a)})\}\) the \(2d \times 2d\) matrix whose columns are the \(2d\) Hamiltonian vectors needed to represent \(\vec{v}\) as a linear
\(^3\) At this point we abuse the notation by identifying vector fields and points in \(\mathbb{R}^{2d}\).
combination of the type (24). Now, condition [A3] implies that \{\vec{X}_{\sigma(1)}(z_0^{(a)}), \ldots, \vec{X}_{\sigma(2d)}(z_0^{(a)})\} is a basis in \(\mathbb{R}^{2d}\), thus \(W_\varepsilon\) is invertible and from (24), denoting \(\vec{a} = (a_1, \ldots, a_{2d})\), we get
\[
\|\vec{a}\| \leq \left\|W_{\varepsilon}^{-1}\right\| \frac{\|\vec{b}\|}{\varepsilon} \leq D,
\]
(26)
where \(D = \sup_{\varepsilon \in T} \sup_{\varepsilon \in T_{\omega}(\mathbb{R}^{2d})} \|W_{\varepsilon}^{-1}\|(L + A) < \infty\) depends only on \(L + A\). Hence the guiding times \(a_i\) are uniformly bounded.

Without loss of generality\(^4\) we will prove the lemma for \(d = 1\), and we write \(v = \varepsilon a_1 X_1 + \varepsilon a_2 X_2\). Then the guiding path consists of two segments. We will show that these two segments can be shadowed one at a time, hence the code \(\xi^{(b)}\) will be obtained by updating the code \(\xi^{(b)}\) twice. The rules for updating the code are
\[
\begin{align*}
\bar{\xi}_k^{(b)} &= \xi_k^{(a)} & \text{for } k \leq 0, \\
\bar{\xi}_k^{(b)} &= c_1 & \text{for all } k > 0
\end{align*}
\]
(27)
and
\[
\begin{align*}
\bar{\xi}_k^{(b)} &= \xi_k^{(b)} & \text{for } k \leq \left\lfloor \frac{a_1}{T_{c_1}} \right\rfloor, \\
\bar{\xi}_k^{(b)} &= c_2 & \text{for all } k > \left\lfloor \frac{a_1}{T_{c_1}} \right\rfloor,
\end{align*}
\]
(28)
where \(\left\lfloor \cdot \right\rfloor\) denotes rounding up to the next integer and \(T_{c_1}\) denotes the period of the periodic orbit \(L_{c_1}\) (also note that by (26) \(a_i\) are of order \(O(\varepsilon)\)). We define \(\bar{N} = \left\lfloor \frac{\bar{a}}{T_{c_1}} \right\rfloor\) and \(N^b = \left\lfloor \frac{\bar{a}}{T_{c_1}} \right\rfloor\) and let \(N = \bar{N} + N^b\).

We begin by updating the code according to (27) which gives us the code \(\vec{\xi}^{(b)}\). For any \(p \leq 0\) we pick the trajectory of (8) corresponding to \(\xi^{(b)}\) such that \(\vec{\xi}^{(b)} = z^{(a)}_{\bar{c}}\). Let us also consider two auxiliary sequences, \(z_{k}^{(b)}\) which corresponds to the code \(\xi^{(b)}\) and \(\bar{z}_{k}^{(b)}\) which corresponds to the constant code \((c_1)^{N^b}\). Both sequences satisfy the initial condition \(z_0^{(b)} = z_{0}^{(a)} = z_{0}^{(a)}\).

By definition (27) we have \(\bar{z}_{k}^{(b)} = c_1\) for all \(k \geq 0\). Then the following lemma holds.

**Lemma 3 (lemma 5, [4]).** For any \(K_0 > 0\), \(l_0 > 0\), there is \(\varepsilon_0 > 0\) such that for any \(|\varepsilon| < \varepsilon_0\) and any two codes \(\xi^1\) and \(\xi^2\) such that for some index \(j\)
\[
\xi^1_{j+i} = \xi^2_{j+i} = c \quad 0 \leq i \leq N_0(\varepsilon) \equiv \left\lfloor \frac{l_0}{\varepsilon} \right\rfloor
\]
the inequality \(\|z_j^{(1)} - z_j^{(2)}\| \leq \varepsilon K_0\) implies
\[
\|z_j^{1} + N - z_j^{2} + N\| \leq C_1 e^{CN_{C_2}} \quad 0 \leq N \leq N_0(\varepsilon),
\]
where
\[
C_1 = \left\|\frac{\partial \Phi_{z}}{\partial (x, y)}\right\| \frac{4r}{1 - \lambda} + K_0
\]
and
\[
C_2 = \left\|\frac{\partial \Phi_{z}}{\partial z}\right\| + \left\|\frac{\partial \Phi_{z}}{\partial (x, y)}\right\| \max \left\{\left\|\frac{\partial x_{z}}{\partial z}\right\|, \left\|\frac{\partial y_{z}}{\partial z}\right\|\right\}.
\]
In the compact set \(\overline{D}\) we have
\[
\|z_j^{1} + N - z_j^{2} + N\| \leq 3\varepsilon C_1 \quad 0 \leq N \leq \frac{1}{\varepsilon C_2}.
\]

\(^4\) The proof for \(d > 1\) is analogous to the case \(d = 1\). The only difference is that the guiding path will consist of more segments. Consequently the constant \(A\) depends on \(d\).
which implies that
\[ \left\| \tilde{z}_k^{(b)} - z_k^{(c)} \right\| \leq 3\varepsilon \left( \left\| \phi_{c1,1} \right\|_{C1} \frac{4r}{1 - \lambda} \right) \]  
(29)
for all \( 0 \leq k \leq N \). By lemma 1 we have the following bound:
\[ \left\| \tilde{z}_0^{(b)} - z_0^{(a)} \right\| = \| z_0^{(b)} - z_0^{(a)} \| \leq \varepsilon K, \]  
(30)
since \( z_k^{(b)} \) and \( z_k^{(a)} \) share the same code for \( k \leq 0 \) and \( \tilde{z}_0^{(b)} = z_0^{(a)} \). Now, since \( \tilde{z}_k^{(b)} \) and \( z_k^{(b)} \) are generated from the identically same code, lemma 3 implies that
\[ \left\| \tilde{z}_k^{(b)} - z_k^{(c)} \right\| \leq 3\varepsilon \left( \left\| \phi_{c1,1} \right\|_{C1} \frac{4r}{1 - \lambda} + K \right) \]  
(31)
for all \( 0 \leq k \leq N \). Combining estimates (29) and (31), we obtain
\[ \left\| \tilde{z}_k^{(b)} - z_k^{(c)} \right\| \leq \left\| \tilde{z}_k^{(b)} - z_k^{(c)} \right\| + \left\| z_k^{(c)} - z_k^{(c)} \right\| \leq 3\varepsilon \left( \left\| \phi_{c1,1} \right\|_{C1} \frac{4r}{1 - \lambda} + K \right), \]  
(32)
for all \( 0 \leq k \leq \tilde{N} \). Now it remains to prove that \( z_1^{(c)} \) stays close to the guiding path \( G \). By definition we have
\[ G(t, z_0^{(a)}, z) = z_0^{(a)} - t\Omega^{-1} \nabla J_c(z_0^{(a)}) \]  
for \( 0 \leq t \leq \varepsilon a_1 \),
where \( \Omega^{-1} \) is the inverse of the symplectic matrix. The map (8) takes a simple form for the constant code \( (c_1) \), see [4] and the proof of lemma 1 therein,
\[ z^{(c)}_1 = z^{(a)}_0 - \varepsilon k T_{c1} \Omega^{-1} \nabla J_c(z^{(a)}_0) + O(\varepsilon^2). \]
Therefore
\[ \left\| \tilde{z}_N^{(b)} - G(\varepsilon a_1, z_0^{(a)}, z) \right\| \leq O(\varepsilon^2) + \varepsilon \max_{(u,v)\in\Omega} |T_c(u,v)||X_c(u,v)| \]  
(34)
\[ \leq \varepsilon (1 + \max_{(u,v)\in\Omega} |T_c(u,v)||X_c(u,v)|), \]  
(35)
where the \( O(\varepsilon^2) \) term is the difference of the two maps and the second term compensates for the fact that \( a_1 \) in general is not a multiple of \( T_{c1} \), this round off error is bounded by the strength of the vector field times the largest period. Collecting the estimates, we obtain
\[ \left\| \tilde{z}_N^{(b)} - G(\varepsilon a_1, z_0^{(a)}, z) \right\| \leq 3\varepsilon \left( \left\| \phi_{c1,1} \right\|_{C1} \frac{4r}{1 - \lambda} + K \right) + \varepsilon \]  
(36)
\[ + \varepsilon \max_{(u,v)\in\Omega} |T_c(u,v)||X_c(u,v)|, \]  
(37)
which is the accuracy with which we have shadowed the first segment of the guiding path.
Let us continue by shadowing the second segment of the guiding path. We begin by updating the code to \( \varepsilon^{(b)} \), using the rule given by equation (28). We pick the trajectory \( z^{(b)} \) which corresponds to \( \varepsilon^{(b)} \) and satisfies \( z_p^{(b)} = z_p^{(a)} \). We repeat the arguments above to shadow the second segment of the guiding path. Indeed, we consider two auxiliary sequences \( \tilde{z}_k^{(b)} \), which corresponds to the code \( \varepsilon^{(b)} \), and \( z_k^{(b)} \), which corresponds to the constant code \( (c_2) \).
Both sequences satisfy the initial condition
\[ \tilde{z}_k^{(b)} = z_k^{(b)} = G(\varepsilon a_1, z_0^{(a)}, z). \]
By definition (27) we have \( \varepsilon_k^{(b)} = c_2 \) for all \( k > \tilde{N} \). Then by lemma 3
\[ \left\| \tilde{z}_k^{(b)} - z_k^{(b)} \right\| \leq 3\varepsilon \left( \left\| \phi_{c2,2} \right\|_{C1} \frac{4r}{1 - \lambda} \right) \]  
(38)
for all $0 \leq k \leq N^b$. Lemma 1 together with (36) yields the following bound:
\[
\|z_N^{(b)} - \tilde{z}_N^{(b)}\| = \|z_N^{(b)} - G(\varepsilon a_1, z_0^{(a)}, z)\| \leq \max_c \sup_{(u,v) \in \overline{D}} (6 \varepsilon \|\phi_c\| C 1^2 4r 1 - \lambda + 2 \varepsilon K + \varepsilon + \varepsilon |T_c(u,v)| |X_c(u,v)|) .
\]

Then lemma 3 implies that
\[
\|z_N^{(b)} + k - \tilde{z}_N^{(b)} + k\| \leq 3 \varepsilon \max_c \sup_{(u,v) \in \overline{D}} \left( 7 \varepsilon \|\phi_c\| C 1^2 4r 1 - \lambda + 4 K + 1 + |T_c(u,v)| |X_c(u,v)| \right) .
\]

for all $0 \leq k \leq N^b$. Combining estimates (38) and (39) gives
\[
\|z_N^{(b)} + k - \tilde{z}_N^{(b)} + k\| \leq \left( 8 \varepsilon \|\phi_c\| C 1^2 4r 1 - \lambda + 4 K + 1 + |T_c(u,v)| |X_c(u,v)| \right) .
\]

for all $0 \leq k \leq N^b$. Using that $\tilde{z}_N^{(b)} = G(t_1, z(a)_0, z)$ the guiding path is shadowed by $\tilde{z}_N^{(b)}$ exactly analogously to the first segment. The result is
\[
\|\tilde{z}_N^{(b)} - G(\varepsilon a_2, z_0^{(a)}, z)\| \leq \varepsilon + \varepsilon \max_c \sup_{(u,v) \in \overline{D}} |T_c(u,v)| |X_c(u,v)| .
\]

Using that $z_N^{(b)} = z_N^{(b)}$ by definition and combining (41) and (43) gives
\[
\|z_N^{(b)} - G(\varepsilon a_2, z_0^{(a)}, z)\| =
\leq 4 \varepsilon \max_c \sup_{(u,v) \in \overline{D}} \left( 6 \varepsilon \|\phi_c\| C 1^2 4r 1 - \lambda + 3 K + 1 + |T_c(u,v)| |X_c(u,v)| \right)
\]

Thus, using definition (13) of $A$
\[
A = 4 \varepsilon \max_c \sup_{(u,v) \in \overline{D}} \left( 6 \varepsilon \|\phi_c\| C 1^2 4r 1 - \lambda + 3 K + 1 + |T_c(u,v)| |X_c(u,v)| \right) ,
\]

and the definition of guiding path gives
\[
\|z_N^{(b)} - z\| \leq \varepsilon A .
\]

The constant $A$ is uniformly bounded on the compact space $\overline{D}$. Lastly, from the fact that the guiding times $a_i$ are uniformly bounded it follows that the length of the guiding path is bounded by $\varepsilon C_1$, where $C_1$ is a uniform constant. This and the estimates above imply that
\[
\|z_k^{(b)} - z\| \leq \varepsilon (C_1 + A) \quad \text{for} \quad 0 \leq k \leq N .
\]

6. Conclusions

In this paper we study a class of slow–fast Hamiltonian systems where the fast system has chaotic (horseshoe type) dynamics. Under an additional assumption we establish that the slow component of the dynamics behaves in a way which resembles a random walk and, in particular, can ‘shadow’ any given continuous curve. If the additional assumption is violated
we expect the slow dynamics to behave differently, more precisely, to resemble superposition of a systematic drift (described by some kind of an averaged system) and a diffusion-like process.

The results described in this paper and in the previous works [4, 6] have been partially motivated by the study of the Arnold diffusion and mechanisms of instability in a priori unstable Hamiltonian systems.

Acknowledgments

The research of NB was funded by the Academy of Finland and the EU research training network CODY. The research of the second author was funded by the Academy of Finland.

References

[1] Anosov D 1960 Averaging in systems of ODEs with rapidly oscillating solutions Izv. Akad. Nauk. SSSR 24 721–42
[2] Arnold V I 1964 Instability of dynamical systems with several degrees of freedom Sov. Math. Dokl. 5 581–5
[3] Bogoliubov N N and Mitropolskii Yu A 1961 Asymptotic Methods in the Theory of Nonlinear Oscillations (London: Gordon and Breach)
[4] Brännström N and Gelfreich V 2008 Drift of slow variables in slow–fast Hamiltonian systems Physica D 237 2913–21
[5] Dumas H S and Golse F 1997 The averaging method for perturbations of mixing flows Ergod. Theory Dyn. Syst. 17 1339–58
[6] Gelfreich V and Turaev D 2008 Unbounded energy growth in Hamiltonian systems with a slowly varying parameter Commun. Math. Phys. 283 769–94
[7] Gelfreich V and Turaev D 2008 Fermi acceleration in non-autonomous billiards J. Phys. A: Math. Theor. 41 212003
[8] Kifer Y 2004 Averaging principle for fully coupled dynamical systems and large deviations Ergod. Theory Dyn. Syst. 24 847–71
[9] Lochak P and Meunier C 1988 Multiphase Averaging for Classical Systems with Applications to Adiabatic Theorems (New York: Springer)
[10] Moser J 1973 Stable and Random Motions in Dynamical Systems (Annals of Mathematical Studies vol 77) (Princeton: Princeton University Press)
[11] Neishtadt A 1976 Averaging in multi-frequency systems: II Sov. Phys.—Dokl 21 80–2 (Engl. Transl.)
Neishtadt A 1976 Dokl. Akad. Nauk. SSSR 226 (6) 1295–8
[12] Sanders J, Verhulst F and Murdock J 2007 Averaging Methods in Nonlinear Dynamical Systems (New York: Springer) 2 Rev edn