Instability for \textit{a priori} unstable Hamiltonian systems: a dynamical approach

Abed Bounemoura * and Edouard Pennamen †

March 10, 2011

Abstract

In this article, we consider an \textit{a priori} unstable Hamiltonian system with three degrees of freedom, for which we construct a drifting solution with an optimal time of instability. Such a result has been already proved by Berti, Bolle and Biasco using variational arguments, and by Treschev using his separatrix map theory. Our approach is new: it is based on a special type of symbolic dynamics corresponding to the random iteration of a family of twist maps of the annulus, and it gives the first concrete application of this idea introduced by Moeckel in an abstract setting and further studied by Marco. Our method should also be useful in obtaining the optimal time of instability in the more difficult context of \textit{a priori} stable Hamiltonian systems.

1 Introduction

The theory of perturbations of Hamiltonian systems is essentially the study of near-integrable Hamiltonian systems, generated by functions of the form

\[ H(\theta,I) = h(I) + f(\theta,I), \quad (\theta,I) \in \mathbb{A}^n = \mathbb{T}^n \times \mathbb{R}^n, \]

where \( H \) is sufficiently smooth and \( f \) sufficiently small.

1. For \( f = 0, H = h \) is an integrable system and the situation is well-understood. In this case, the variables \((\theta,I)\) are called angle-action coordinates for \( h \) and the Hamiltonian depends only on the action variables. The phase space \( \mathbb{A}^n \) is trivially foliated by invariant Lagrangian tori parametrized by the action variables: indeed the equations of motion of \( H \) read

\[
\begin{align*}
\dot{\theta} &= \nabla h(I) \\
\dot{I} &= 0,
\end{align*}
\]
so for each $I_0 \in \mathbb{R}^n$, the torus $T_0 = \mathbb{T}^n \times \{ I_0 \}$ is Lagrangian and invariant under the Hamiltonian flow. The latter is complete and restricts to a linear flow on $T_0$ with frequency $\omega_0 = \nabla h(I_0)$, that is
\[
\Phi_t^H(\theta_0, I_0) = (\theta_0 + t\omega_0 \cdot \mathbb{Z}^n, I_0) \in T_0, \quad t \in \mathbb{R}.
\]
The action variables remain fixed for all times, and all solutions are quasi-periodic.

2. Now if we let $f$ to be non zero but small, that is
\[
|f| = \varepsilon << 1,
\]
with respect to some norm $|.|$, the situation is dramatically different. For such a near-integrable system, the action variables are no longer first integrals and we may find “unstable” solutions $(\theta(t), I(t))$ that experience a substantial drift in the action direction. Such an instability phenomenon is usually referred to Arnold diffusion, after the original work of Arnold ([Arn64]). In this paper, the author devised a beautiful mechanism to construct an example of an analytic near-integrable system with three degrees of freedom $(n = 3)$, possessing an orbit $(\theta(t), I(t))$ drifting in the space of actions, that is
\[
|I(\tau) - I(0)| \geq 1
\]
for some time $\tau = \tau(\varepsilon) > 0$, which is called the time of instability (or time of diffusion).

3. However, showing that such drifting solutions exist in some large class of near-integrable Hamiltonian systems is a very difficult task. Indeed, on the one hand KAM theory ([Kol54], [Mos62] and [Arn63]) gives, under some non-degeneracy condition on $h$ and provided $\varepsilon$ is small enough, the existence of a set of positive measure of quasi-periodic solutions. These solutions are perpetually stable, in the sense that the variation of their action components is of order at most $\sqrt{\varepsilon}$ for all time. Moreover, for $n = 2$ and if the integrable Hamiltonian is isospectrally non degenerate ([Arn63]), one can even show that all solutions are stable for all time. On the other hand, for $n \geq 3$ we know from Nekhoroshev theory ([Nek77], [Nek79]) that all solutions are stable, not for all time, but for an interval of time which is exponentially long with respect to $\varepsilon^{-1}$, provided that $h$ meets some quantitative transversality condition and the perturbation is sufficiently small.

But even though the existence of drifting orbits for near-integrable Hamiltonian systems seems to be quite exceptional, Arnold conjectured that this topological instability is in fact a “typical” phenomenon (see [Arn63] or [AKN06]). We refer to [Loc99] for a very lucid and enlightening discussion on Arnold mechanism (and on Arnold diffusion in general). Under some “generic” conditions and for $n = 3$, Mather ([Mat04]) announced that this
conjecture holds if the unperturbed Hamiltonian is convex (the convexity is required by the use of his variational methods). This is so far the best result in this direction, however, his proof is highly technical and still incomplete. Another connected question, which is even harder, is to find orbits that densely fill the energy level (this is the so-called quasi-ergodic hypothesis): for instance, in [Her98] it is asked whether there exists a Hamiltonian system, $C^\infty$-smooth and for $r \geq 2$, $C^r$-close to the integrable Hamiltonian $h(I) = \frac{1}{2} |I|^2$, which has a dense orbit on an energy level (progress towards this question have been made recently in [KLS10] and [KZZ09]).

4. Yet there is another case, derived from Arnold’s original example, which is much simpler and hence more studied and understood, in which the unperturbed Hamiltonian is completely integrable (in the sense of symplectic geometry) but possesses some “hyperbolicity”. The prototype for such a system is given by an uncoupled product of rotators with a pendulum. This case is usually referred to as *a priori* unstable (following the terminology introduced in [CG94]), in contrast with the *a priori* stable case where the unperturbed system is integrable in angle-action coordinates and therefore has no hyperbolic feature (one can also say that *a priori* stable systems are “fully elliptic”). The interest in studying such *a priori* unstable systems is in fact double. First, one can use the original strategy of Arnold to find unstable orbits, by constructing and following a “transition chain” made of sets with suitable hyperbolic properties (or minimizing properties when using variational arguments). The second interest lies in the fact that, due to normal form theory, the study of *a priori* stable systems in a neighbourhood of a simple resonance reduces to the *a priori* unstable case (see [LMS03] for example). Results on the topological instability of *a priori* unstable systems under fairly general conditions can be found in [DdILS06] and [Tre04] by means of geometrical methods and in [CY04], [Ber08] and [CY09] by variational methods. However there are many difficulties in using these results to tackle Arnold’s conjecture which concerns *a priori* stable systems.

5. Another feature of *a priori* unstable systems is that one avoids all exponentially small phenomena which are typical for analytic (or Gevrey) *a priori* stable systems, as the lower bound on the time of instability imposed by Nekhoroshev’s theorem. Indeed, if $\mu$ denotes the small parameter in the *a priori* unstable case, then it was realized by Lochak that the time of instability should be polynomial, and then Bernard ([Ber96]), adapting Bessi’s work ([Bes96]), showed that one can obtain an upper bound on the time of instability which is of order $\mu^{-2}$. In [Loc99], it was conjectured that the optimal time of instability should be $\mu^{-1} \ln \mu^{-1}$, and this was proved by Berti, Bolle and Biasco [BBB03], using Bessi’s ideas and suitable *a priori* estimates, obtained by convexity arguments, to localize the minima of the action functional. This time of instability was also found by Treschev ([Tre04]) using fairly different geometric methods.
6. The goal of this article is to give yet another method to construct an example with this optimal time of instability, which should be useful to obtain the optimal time of instability for a priori stable systems. Our approach is dynamical: it uses the notion of polysystem introduced by Moeckel ([Moe02]) in the context of Arnold diffusion, and which corresponds to the random iteration of a family of maps (the term “polysystem” is due to Marco, such systems are just locally constant skew-products over a Bernoulli shift and they are also known as iterated function systems). More precisely, our example uses an explicit construction of a polysystem, which can be seen as a concrete realisation of an abstract mechanism introduced by Marco in [Mar08]. Finally, let us also note that polysystems are a crucial ingredient in the recent work of Marco towards the instability for general a priori stable systems with 3 degrees of freedom ([Mar10b], [Mar10c], [Mar10d] and [Mar10a]).

2 Main results

7. For \( n \geq 1 \), let us recall that a function \( f \in C^\infty(\mathbb{A}^n) \) is \( \alpha \)-Gevrey, for \( \alpha \geq 1 \), if for any compact subset \( K \subseteq \mathbb{A}^n \) there exist two positive constants \( A_K, B_K \) such that

\[
|\partial^k f|_{C^\infty(K)} \leq A_K B_K^{|k|} (|k|!)^\alpha, \quad k \in \mathbb{N}^{2n},
\]

with the standard multi-index notation. We shall denote by \( G^\alpha(\mathbb{A}^n) \) the space of such functions. For \( \alpha = 1 \), these are exactly the analytic functions, but for \( \alpha > 1 \), the space \( G^\alpha(\mathbb{A}^n) \) contains non zero compactly-supported functions (we refer to [MS02], Appendix A, for more details). Our perturbation will be \( \alpha \)-Gevrey for \( \alpha > 1 \).

In all this paper, we will consider a Hamiltonian system \( h \) with two degrees of freedom defined by

\[
h(\theta_1, \theta_2, I_1, I_2) = \frac{1}{2} (I_1^2 + I_2^2) + \cos 2\pi \theta_1, \quad (\theta_1, \theta_2, I_1, I_2) \in \mathbb{A}^2,
\]

which is the direct product of a pendulum \( P(\theta_1, I_1) = \frac{1}{2} I_1^2 + \cos 2\pi \theta_1 \) on the first factor and a standard rotator \( S(\theta_2, I_2) = \frac{1}{2} I_2^2 \) on the second factor. It is the simplest example of an a priori unstable integrable Hamiltonian system.

Below we shall state two versions of our theorem, one for Hamiltonian diffeomorphisms (Theorem 2.2) and one for Hamiltonian flows (Theorem 2.1). For the discrete case, our unperturbed diffeomorphism is the time-one map of the Hamiltonian flow generated by \( h \), \( \Phi^h : \mathbb{A}^2 \to \mathbb{A}^2 \). In the continuous case, our unperturbed system is the a priori unstable Hamiltonian with three degrees of freedom defined by

\[
h(\theta, I) = h(\theta_1, \theta_2, I_1, I_2) + I_3 = \frac{1}{2} (I_1^2 + I_2^2) + I_3 + \cos 2\pi \theta_1.
\]
8. Let us state our main result.

**Theorem 2.1.** For $\alpha > 1$, there exist positive constants $C$, $\mu_0$ and a function $f \in G^\alpha(\mathbb{A}^3)$ such that if $0 < \mu \leq \mu_0$, the Hamiltonian system

$$H(\theta, I) = \hat{h}(\theta, I) + \mu f(\theta, I), \quad (\theta, I) \in \mathbb{A}^3,$$

has an orbit $(\theta(t), I(t))_{t \in \mathbb{R}}$ such that

$$\lim_{t \to \pm \infty} I_2(t) = \pm \infty,$$

with the estimates

$$|I_2(\tau) - I_2(0)| \geq 1, \quad \tau \leq C\mu^{-1}\ln\mu^{-1}.$$
9. Let us recall that one of our main motivation in using a new mechanism to construct an unstable orbit with an optimal time of instability in this a priori unstable case, is to tackle the corresponding problem for the a priori stable case, which is a much more challenging problem. In fact, in this case the optimal time of instability \( \tau(\varepsilon) \) is of order \( \exp(\varepsilon^{-a}) \), for some positive exponent \( a \), and it was shown recently in [BM11] that one has the lower bound

\[
a \geq (2a(n-1))^{-1},
\]

for \( \alpha \)-Gevrey perturbations, with \( \alpha \geq 1 \) (this include the analytic case \( \alpha = 1 \)). In the same paper, the authors conjectured that this value is indeed optimal, in the sense that the upper bound \( a \leq (2a(n-1))^{-1} \) should also hold (for the moment, one only knows that \( a \leq (2a(n-2))^{-1} \), this is due to Herman, Marco and Sauzin ([MS02]) for the case \( \alpha > 1 \), and to Ke Zhang ([Zha09]) for the more difficult analytic case \( \alpha = 1 \)). We believe that using some technical tools introduced in the present work, in particular the construction of our polysystem, one should finally be able to reach the value \( a = (2a(n-1))^{-1} \) (first in the easiest case \( \alpha > 1 \)).

Indeed, one can show that polysystems do appear in the a priori stable case, as in the examples of [Mar05] and [LM05], our example being just an a priori unstable version of these. But of course, whereas in the a priori unstable case, the splitting (which is the “distance” between stable and unstable manifolds of partially hyperbolic tori, or equivalently, the size of the drift in the action after one homoclinic excursion in the annulus) is simply of order \( \mu \), in the a priori stable case it is exponentially small (with respect to \( \varepsilon \)) and therefore much more difficult to deal with. Nevertheless, we believe that one can “embed” our polysystem into an a priori stable system, either by refining the splitting estimates of [Mar05] and [LM05], or by using an embedding mechanism analogous to [MS02], in which the authors use a very clever lemma to estimate the time of instability without any splitting estimates (they easily recover the latter as a consequence of their construction).

10. We shall not prove Theorem 2.1 directly, but the following equivalent version in terms of diffeomorphisms. Given a function \( H \) on \( \mathbb{A}^n \), we shall denote by \( \Phi^H_t : \mathbb{A}^n \to \mathbb{A}^n \) the time-\( t \) map of its Hamiltonian flow and by \( \Phi^H = \Phi^H_1 \) the time-one map.

**Theorem 2.2.** For \( \alpha > 1 \), there exist positive constants \( C, \mu_0 \) and a function \( f \in G^{\alpha}(\mathbb{A}^2) \) such that if \( 0 < \mu \leq \mu_0 \), the diffeomorphism

\[
\Phi^h \circ \Phi^{\mu f} : \mathbb{A}^2 \to \mathbb{A}^2
\]

has an orbit \((\theta^k, I^k)_{k \in \mathbb{Z}}\) such that

\[
\lim_{k \to \pm \infty} I^k = \pm \infty,
\]
with the estimates
\[ |I_N^N - I_N^0| \geq 1, \quad N \leq C \mu^{-1} \ln \mu^{-1}. \]

Theorem 2.1 follows easily from Theorem 2.2 by a classical suspension argument (see [MS02] for a simple method in the Gevrey case using compactly-supported functions), so we shall not repeat the details. Of course, the constants \( \mu_0, C \) and the function \( f \) are not the same in both theorems, but we have kept the same notation for simplicity. Moreover, in the sequel we will not give explicit values for these constants, in fact sometimes it will be more convenient to use asymptotic notations: given \( u(\mu) \) and \( v(\mu) \) defined for \( \mu \geq 0 \), we shall write \( u(\mu) = O(v(\mu)) \) if there exist positive constants \( \mu' \) and \( c' \) independent of \( \mu \) such that the inequality \( u(\mu) \leq c' v(\mu) \) holds true for \( 0 \leq \mu \leq \mu' \).

11. The plan of this paper is the following. In section 3, we shall describe the perturbation. We will show that the perturbed system has an invariant normally hyperbolic manifold, the stable and unstable manifolds of which intersect transversely along a homoclinic annulus.

Then, this will be used in section 4 to show the following generalisation of the Birkhoff-Smale theorem to the normally hyperbolic case: near the homoclinic manifold, there exist an invariant set on which a suitable iterate of the system is conjugated to a symbolic dynamic, more precisely to a skew-product on the annulus \( A \) over a Bernoulli shift. Let us give a precise definition.

Given an alphabet \( A \subseteq \mathbb{N} \), we let \( \Sigma_A = A^\mathbb{Z} \) be the Cantor set of bi-infinite sequence of elements in \( A \). We will denote by \( \sigma = \sigma_A \) the left Bernoulli shift on \( \Sigma_A \), that is if \( \bar{n} = (n_k)_{k \in \mathbb{Z}} \) belongs to \( \Sigma_A \), then \( \sigma(\bar{n}) = (n'_k)_{k \in \mathbb{Z}} \) is defined by \( n'_k = n_{k+1}, \quad k \in \mathbb{Z} \).

**Definition 2.3.** Let \( M \) be a manifold. A skew-product on \( M \) over \( \sigma \) is a map \( G : \Sigma_A \times M \rightarrow \Sigma_A \times M \) of the form
\[
G(\bar{n}, x) = (\sigma(\bar{n}), F_{\bar{n}}(x)), \quad x \in M,
\]
where \( F_{\bar{n}} : M \rightarrow M \) is an arbitrary map, for \( \bar{n} \in \Sigma_A \).

However we are not able to prove the existence of our orbit using directly this symbolic dynamic. Hence in section 5, we will first construct this orbit for a simplified model called polysystem, which appears as a random iteration of “standard” maps of the annulus, and which is close to our skew-product. Here is the definition.

**Definition 2.4.** A polysystem on \( M \) is a skew-product over a Bernoulli shift of the form
\[
F_{\bar{n}}(x) = f_{n_0}(x), \quad \bar{n} \in \Sigma_A, \quad x \in M,
\]
with \( f_n : M \rightarrow M \) for \( n \in A \).
In other terms, a polysystem is just a locally constant skew-product over the shift, and this special property will greatly help us in the construction of our orbit. Moreover, polsysystems are equivalent to iterated functions systems on the annulus, and these dynamical systems are known to have drifting orbits under generic assumptions (see [Moe02]). Then, in section 6, the orbit for the polysystem will be considered as a pseudo-orbit for the skew-product and, using some hyperbolicity of our orbit, we will conclude the existence of an orbit for the general skew-product using shadowing arguments.

Finally we have gathered in an appendix some estimates on the so-called time-energy coordinates for the simple pendulum that are used in section 4.

3 Construction of the perturbation

This section is devoted to a description of our system, which is similar to the one introduced in [Mar05]. We will first consider the integrable case, and then explains the construction of the perturbation.

3.1 Integrable system

Our integrable diffeomorphism $\mathcal{F}_0 : \mathbb{A}^2 \to \mathbb{A}^2$ is the time-one map of the Hamiltonian flow generated by $h$, that is

$$\mathcal{F}_0 = \Phi^h = \Phi^{I_1^1 + I_2^1 + \cos 2\pi \theta_1},$$

which is the product of the pendulum map $\Phi^P = \Phi^{I_2^1 + \cos 2\pi \theta_1}$ and the integrable twist map $\Phi^S = \Phi^{I_2^1}$ (see figure 1).

It is an “a priori unstable” map in the sense that it possesses an invariant normally hyperbolic annulus. Indeed, let $O = (0,0)$ be the hyperbolic fixed
point of the pendulum map $\Phi^P$, its stable and unstable manifolds obviously coincide, and if $S$ is the upper part of the separatrix, then

$$W^\pm(O, \Phi^P) = S = \{(\theta_1, I_1) \in \mathbb{A} \mid I_1 = 2 \sin \pi \theta_1\}.$$  

Due to the product structure of the map $F_0$, it is easy to see that the annulus $A = O \times \mathbb{A}$ is invariant by $F_0$, and one can check that it is symplectic for the canonical structure of $\mathbb{A}^2$. But the most important feature is that this annulus is normally hyperbolic, more precisely it is $r$-normally hyperbolic for any $r \in \mathbb{N}$, in the sense of [HPS97]. To see this, just decompose the tangent bundle of $\mathbb{A}^2$ along $A$ as

$$TA\mathbb{A}^2 = TA \oplus E^s \oplus E^u,$$

where $E^s$ (resp. $E^u$) is the one-dimensional contracting (reps. expanding) direction associated to the hyperbolic fixed point. Then note that the restriction of $F_0$ to $A$ coincides with the integrable twist map $\Phi^S$, hence there is zero contraction and expansion in the tangent direction $TA$.

The 3-dimensional stable and unstable manifolds of $A$ also coincide, and are given by the product

$$W^\pm(A, F_0) = S \times A.$$  

### 3.2 Perturbed system

Now we will describe the geometric properties of our perturbed system. It will be of the form

$$F_\mu = \Phi^f \circ F_0,$$

where $\mu > 0$ is the small parameter, and $f : \mathbb{A}^2 \to \mathbb{R}$ a function of the form

$$f(\theta_1, \theta_2, I_1, I_2) = \chi(\bar{\theta}) f_1(\theta_1, I_1) f_2(\theta_2)$$

to be defined precisely below.

1. First $\chi : \mathbb{T} \to \mathbb{R}$ will be a bump function. Of course such a function can be chosen to be $\alpha$-Gevrey for $\alpha > 1$ but not analytic. More precisely, let us choose $\bar{\theta} \in \mathbb{T}$ such that

$$(\bar{\theta}, I) \in S \implies \Phi^P(\bar{\theta}, I) = (1 - \bar{\theta}, I).$$

Since the separatrix $S$ is symmetric about the section $\{\theta_1 = 1/2\}$, $\bar{\theta}$ is well-defined. Then choose any $\tilde{\theta} \in [\bar{\theta}, 1/2]$, and a function $\chi \in G^\alpha(\mathbb{T})$, $\alpha > 1$, such that

$$\chi(\theta) = \begin{cases} 
1 & \text{if } \theta \in [\bar{\theta}, 1 - \bar{\theta}], \\
0 & \text{if } \theta \notin [\bar{\theta}, 1 - \bar{\theta}]. 
\end{cases}$$

9
Since $\chi$ is identically zero at 0, the perturbation $\Phi^{u,f}$ is the identity on the normally hyperbolic annulus $A$, and therefore the latter remains invariant and normally hyperbolic for the map $F_\mu$. Moreover, as $\chi$ vanishes in a neighbourhood of 0, some pieces of the stable and unstable manifolds for $F_\mu$ will coincide with the ones of $F_0$. The use of a bump function here is only to simplify the construction, it can be relaxed but at the expense of a more involved analysis (as in [LM05]).

2. Now to define the function $f_1 : A \to \mathbb{R}$, we will use time-energy coordinates on the first factor, so we introduce the symplectic diffeomorphism

$$
\Psi : \mathcal{E} \times \mathbb{A} \longrightarrow \mathcal{E}_s \times \mathbb{A}
$$

where

$$
\mathcal{E} = \{(\theta_1, I_1) \in \mathbb{A} \mid 0 < \theta_1 < 1, \ I_1 > 0\}
$$

and

$$
\mathcal{E}_s = \{(\tau, e) \in \mathbb{R}^2 \mid e > -2, |\tau| < \frac{1}{2}T(e)\}.
$$

We refer to Appendix A for more details on those coordinates. The function $f_1$ is simply defined by

$$
f_1(\tau, e) = 1 - \frac{1}{2}\tau^2, \quad (\tau, e) \in \mathcal{E}_s,
$$

and it will create a transverse intersection between the stable and unstable manifolds $W^+(A, F_\mu)$ and $W^-(A, F_\mu)$.

3. Finally the function $f_2$ will be defined by

$$
f_2(\theta_2) = -\pi^{-1}(2 + \sin 2\pi \theta_2), \quad \theta_2 \in \mathbb{T}.
$$

The fact the $f_2$ is nowhere zero on $\mathbb{T}$ (for any $\theta_2 \in \mathbb{T}$, $|f_2(\theta_2)| \geq \pi^{-1}$) will imply that $W^+(A, F_\mu)$ and $W^-(A, F_\mu)$ intersect transversely along an annulus, and the explicit form of $f_2$ will be responsible for the non trivial dynamics along this homoclinic annulus. The existence of this homoclinic annulus is a crucial ingredient in our construction but it is also non-generic, and this explains why our method cannot handle more general perturbations.

4. Let us define the domains

$$
D = ([\bar{\theta}, 1 - \bar{\theta}] \times \mathbb{R}_+^+) \times \mathbb{A} \subseteq \mathbb{A}^2, \quad D_s = \Psi(D) \subseteq \mathcal{E}_s \times \mathbb{A},
$$

on which the perturbation (either in the original coordinates or in time-energy coordinates) is non zero, and

$$
\tilde{D} = ([\bar{\theta}, 1 - \bar{\theta}] \times \mathbb{R}_+^+) \times \mathbb{A} \subseteq \mathbb{A}^2, \quad \tilde{D}_s = \Psi(\tilde{D}) \subseteq \mathcal{E}_s \times \mathbb{A}.
$$
For \((\tau, e, \theta, I_2) \in \tilde{D}_s\), our perturbation can be written explicitly as
\[
\Phi_{\mu f}(\tau, e, \theta, I_2) = (\tau, e - \mu f_1'(\tau)f_2(\theta_2), \theta_2 - \mu f_1(\tau)f_2'(\theta_2))
\]
\[
= (\tau, e + \mu \tau f_2(\theta_2), \theta_2, I_2 + 2\mu f_1(\tau)\cos(2\pi \theta_2)),
\]
and otherwise the diffeomorphism \(\Phi_{\mu f}\) is the identity.

5. In the sequel we shall need the following property.

**Proposition 3.1.** The immersed manifolds \(W^+(A, F_\mu)\) and \(W^-(A, F_\mu)\) intersect transversely along the annulus
\[
\mathcal{I}_\mu = \{ (\tau, e, \theta, I_2) \in \mathcal{E} \times \mathbb{A} \mid \tau = e = 0 \}.
\]

Let us remark that this annulus is also given by
\[
\mathcal{I}_\mu = \{ (\theta_1, I_1, \theta_2, I_2) \in \mathbb{A}^2 \mid \theta_1 = 1/2, I_1 = 2 \}
\]
in the original coordinates.

**Proof.** By definition of \(D\), one easily sees that the sets
\[
W^+(A, F_0) \cap F_0(\tilde{D}), \quad W^-(A, F_0) \cap F_0^{-1}(\tilde{D}),
\]
are disjoint from \(D\). Since \(F_\mu\) and \(F_0\) coincide outside \(D\), we can define pieces of stable and unstable manifolds by
\[
W^+_\theta(A, F_\mu) = W^+(A, F_0) \cap F_0(\tilde{D}) \subseteq W^+(A, F_\mu)
\]
and
\[
W^-_\theta(A, F_\mu) = W^-(A, F_0) \cap F_0^{-1}(\tilde{D}) \subseteq W^-(A, F_\mu).
\]

Therefore
\[
F^{-1}_\mu(W^+_\theta(A, F_\mu)) \subseteq W^+(A, F_\mu), \quad F_\mu(W^-_\theta(A, F_\mu)) \subseteq W^-(A, F_\mu). \quad (3)
\]

Then on the one hand,
\[
F^{-1}_\mu(W^+_\theta(A, F_\mu)) = F^{-1}_0 \circ \Phi^{-\mu f}(W^+_\theta(A, F_\mu))
\]
\[
= F^{-1}_0(W^+_\theta(A, F_\mu))
\]
\[
= W^+(A, F_0) \cap \tilde{D},
\]
and on the other hand,
\[
F_\mu(W^-_\theta(A, F_\mu)) = \Phi^\mu f \circ F^{-1}_0(W^-_\theta(A, F_\mu))
\]
\[
= \Phi^\mu f(W^-(A, F_0) \cap \tilde{D}).
\]
Now in time-energy coordinates, one simply has
\[ W^{\pm}(A, F_0) = \{ (\tau, e, \theta, I) \in E \times A \mid e = 0 \}, \]
and therefore
\[ W^+(A, F_0) \cap \tilde{D} = \{ (\tau, e, \theta, I) \in \tilde{D} \mid e = 0 \}, \]
while, using the expression of the perturbation (2),
\[ \Phi^{\mu f}(W^- (A, F_0) \cap \tilde{D}) = \{ (\tau, e, \theta, I) \in \tilde{D} \mid e = \mu f(\theta) \}. \]
Since \( f \) is nowhere zero, one easily sees that the manifolds \( F^{-1}(W^+ (A, F_\mu)) \)
and \( F_\mu(W^- (A, F_\mu)) \) intersect transversely along the annulus
\[ I_\mu = \{ (\tau, e, I) \in \tilde{D} \mid e = 0 \}, \]
and the conclusion follows by (3).

4 Construction of a symbolic dynamic

6. In this section, we will take advantage of the fact that our diffeomorphism
\[ F_\mu : \mathbb{R}^2 \to \mathbb{R}^2 \]
possesses an invariant normally hyperbolic annulus \( A \), whose stable and unstable manifolds intersect transversely along a homoclinic annulus \( I_\mu \).

In the case where the normally hyperbolic manifold is a point, under such a transverse homoclinic intersection it is well-known that chaotic dynamics arise: the system has an invariant Cantor set on which a suitable iterate is conjugated to a shift map (this is the Horseshoe theorem, due to Birkhoff, Smale and Alexeev). In our more general situation, the symbolic dynamic is more complicated.

7. Recall from definition 2.3 that a skew-product (over \( \sigma \)) is completely defined by the family of maps \( (F_n)_{n \in \Sigma_A} : M \to M \), and the skew-product will be denoted by \( [[F_n]]_{n \in \Sigma_A} \). Then one can easily see that a sequence \( (n_k, x_k)_{k \in \mathbb{Z}} \in A \times M \) gives rise to an orbit \( (\sigma^k(\bar{n}), x_k)_{k \in \mathbb{Z}} \in \Sigma_A \times M \), where \( \bar{n} = (n_k)_{k \in \mathbb{Z}} \), for the skew-product \( [[F_n]]_{n \in \Sigma_A} \) if and only if
\[ F_{\sigma^k(\bar{n})}(x_k) = x_{k+1}, \quad k \in \mathbb{Z}. \]

In this section, we will need our alphabet \( A \) to contain integers of order \( \ln \mu^{-1} \), but for subsequent arguments, we will also need it to contain integers \( n \in \mathbb{N} \) as large as \( \mu^{-\frac{2}{3}} \ln \mu^{-1} \), so we may already fix
\[ A = A_\mu = \{ [\ln \mu^{-1}], \ldots, 2[\mu^{-\frac{2}{3}} \ln \mu^{-1}] \}. \]

12
For simplicity, we shall get rid of the subscript $\mu$.

8. Let $O_\mu \subseteq \mathcal{D}_s$ be a neighbourhood (in time-energy coordinates) of the homoclinic annulus $I_{\mu}$. For $x \in O_\mu$, let us set
$$n_0^x = \inf \{n \in \mathbb{N}^* \mid F_{\mu}^n(x) \in O_\mu \} \in \mathbb{N}^* \cup \{+\infty\}$$
and assuming $n_0^x < +\infty$, we can define inductively
$$n_k^x = \inf \{n \in \mathbb{N}^* \mid F_{\mu}^n \left(F_{\mu}^{n_{k-1}^x}(x)\right) \in O_\mu \} \in \mathbb{N}^* \cup \{+\infty\},$$
for $k \geq 1$ provided $n_{k-1}^x < +\infty$. Similarly we can define
$$n_{-1}^x = \inf \{n \in \mathbb{N}^* \mid F_{\mu}^{-n}(x) \in O_\mu \} \in \mathbb{N}^* \cup \{+\infty\},$$
and by induction
$$n_{-k}^x = \inf \{n \in \mathbb{N}^* \mid F_{\mu}^n \left(F_{\mu}^{-k+1}(x)\right) \in O_\mu \} \in \mathbb{N}^* \cup \{+\infty\}$$
if $n_{-k}^x < +\infty$ for $k \geq 1$.

We will show below that there exists $\Lambda_\mu \subseteq O_\mu$ such that for any $x \in \Lambda_\mu$, the doubly infinite sequence $\bar{n}^x = (n_k^x)_{k \in \mathbb{Z}}$ is a well-defined element of $\Sigma_\Lambda$. In fact, $\Lambda_\mu$ is homeomorphic to a Cantor set of annuli, and we will be interested in the dynamics restricted to this set. For that, following Moser ([Mos73]) we define the transversal map
$$\bar{F}_{\mu}(x) = F_{\mu}^{n_\bar{x}}(x), \quad x \in O_\mu, \ n_\bar{x} < +\infty,$$
which is the first return map to the neighbourhood $O_\mu$, and if the sequence $\bar{n}^x$ is well-defined, then so is $\bar{F}_{\mu}(x)$ for $n \in \mathbb{Z}$.

9. In the sequel, we will consider the discrete topology on the alphabet $A$, and the sets $A^\mathbb{N}$ and $\Sigma_\Lambda = A^\mathbb{Z}$ will be endowed with the product topology, for which they are compact and metrizable. The goal of this section is to prove the following proposition.

**Proposition 4.1.** For $\mu$ small enough, there exist a neighbourhood $O_\mu$, a set $\Lambda_\mu \subseteq O_\mu$ invariant by the transversal map $F_{\mu}$, a homeomorphism
$$\Upsilon_{\mu} : \Sigma_\Lambda \times \mathbb{A} \longrightarrow \Lambda_\mu$$
with $\psi_{\bar{n}} : \mathbb{A} \rightarrow \mathcal{E}_s$, where $\mathcal{E}_s$ is defined in (1), and a family of maps
$$F_{\mu}(\theta, I) = (\theta + n_0 I, I + 2\mu f_1(\phi_{\mu}(\theta, I)) \cos 2\pi (\theta + n_0 I)), \quad \bar{n} \in \Sigma_\Lambda, \ (\theta, I) \in \mathbb{A},$$
such that for $x = (\tau_x, \varepsilon_x, \theta_x, I_x) \in \Lambda_{\mu}$, then $\bar{n}^x \in \Sigma_\Lambda$ and
$$\bar{F}_{\mu} \circ \Upsilon_{\mu}(\bar{n}^x, \theta_x, I_x) = \Upsilon_{\mu}(\sigma(\bar{n}^x), F_{\mu}(\theta_x, I_x)).$$
Moreover the map $\phi_{\mu} : \mathbb{A} \rightarrow \mathbb{R}$ is Lipschitz and satisfies
$$|\phi_{\mu}|_{C^0(\mathbb{A})} = O \left(\mu^{2\pi-1}\right), \quad \text{Lip}(\phi_{\mu}) = O \left(\mu^{2\pi-\frac{3}{2}} \ln \mu^{-1}\right). \quad (4)$$
The statement of the above proposition seems complicated, but in fact it simply means that the restriction of \( \tilde{F} \) to the invariant set \( \Lambda \) is conjugated to the skew-product on \( A \) given by the maps

\[
F_n(\theta, I) = (\theta + n_0 I, I + 2\mu f_1(\phi_n(\theta, I)) \cos 2\pi(\theta + n_0 I)), \quad n \in \Sigma_A, \ (\theta, I) \in A.
\]

The proof of this result is long and technical. We will first prove an abstract result in section 4.1, which is contained in Proposition 4.4, and then we will apply this proposition to our example in section 4.2.

### 4.1 Symbolic dynamic

1. Here we will use the framework developed by Chaperon ([Cha04], [Cha08]), which was designed to obtain rather general invariant manifold theorems, in particular in the normally hyperbolic case.

Consider a complete metric space \( X \) and a complete subspace \( Y \) of a metric space \( F \). We endow the product space \( X \times F \) with the product metric, that is for \( (x, y), (x', y') \in X \times F \) we define

\[
d((x, y), (x', y')) = \sup \{d(x, x'), d(y, y')\}.
\]

Let us set \( Z = X \times Y \). Then we consider a Lipschitz map

\[
h = (f, g) : Z \rightarrow X \times F,
\]

and we make the following two assumptions.

- **(H1)** There exists a constant \( \rho_0^{-1} > 0 \) such that for all \( x \in X \), \( y, y' \in Y \)

\[
d(g(x, y), g(x, y')) \geq \rho_0^{-1} d(y, y').
\]

Hence, for all \( x \in X \), the map

\[
g_x : Y \rightarrow g_x(Y)
\]

is a bijection.

- **(H2)** For all \( x \in X \), we have \( Y \subseteq g_x(Y) \). Hence the map

\[
g_x^{-1} : Y \rightarrow Y
\]

is well-defined, and if

\[
G : Z \rightarrow Y
\]

is defined by \( G(x, y) = g_x^{-1}(y) \), then \( G \) is Lipschitz.

We will say that the map \( h \) satisfies hypothesis \((H)\) if it satisfies both hypotheses \((H1)\) and \((H2)\). Under these assumptions, we will consider three
positive constants $\nu_0$, $\sigma_0$ and $\kappa_0$ where $\nu_0$ is the Lipschitz constant of $G$ with respect to $Y$, that is
\[
d(G(x, y), G(x, y')) \leq \nu_0 d(y, y'), \quad \forall x \in X, \forall y, y' \in Y,
\]
which from (H1) is smaller than $\rho_0$, $\sigma_0$ is the Lipschitz constant of $G$ with respect to $X$, that is
\[
d(G(x, y), G(x', y)) \leq \sigma_0 d(x, x'), \quad \forall x \in X, \forall y, y' \in Y,
\]
and $\kappa_0$ the Lipschitz constant of $f$. Let us now formulate another hypothesis.

(L) Assume that $\sigma_0 + \nu_0 \max\{\kappa_0, 1\} < 1$. Hence for any $x \in X$, $y \in Y$ the maps
\[
G(., y) : X \longrightarrow Y, \quad G(x, .) : Y \longrightarrow Y,
\]
are contractions.

The following result is due to Chaperon ([Cha04]).

**Theorem 4.2** (Chaperon). With the previous notations, assume that $Y$ is bounded and the map
\[
h : Z \longrightarrow X \times F
\]
satisfies hypotheses (H) and (L). Then the set
\[
\Lambda = \bigcap_{n \in \mathbb{N}} h^{-n}(Z)
\]
is the graph of a contracting map $\Phi : X \rightarrow Y$, with
\[
\text{Lip}(\Phi) \leq \sigma_0(1 - \nu_0\kappa_0)^{-1} < 1.
\]

2. The above theorem is concerned with the iteration of a single map $h$, but with this formalism we obviously have an analogous result for the iteration of a family of maps. More precisely, consider a family of maps
\[
h_n = (f_n, g_n) : Z \longrightarrow X \times F, \quad n \in A,
\]
where $A$ is the given alphabet. We will assume that each map satisfies hypothesis (H) and a uniform version of hypothesis (L), that is:
\[
(L') \sup_{n \in A} \{\sigma_n + \nu_n \max\{\kappa_n, 1\}\} < 1.
\]

Then one can state the following proposition.

**Proposition 4.3.** With the previous notations, assume that $Y$ is bounded and that for any $n \in A$ the maps
\[
h_n : Z \longrightarrow X \times F
\]
satisfy hypotheses (H) and (L'). Then for any sequence \( \bar{n} = (n_k)_{k \in \mathbb{N}} \in A^\mathbb{N} \), the set

\[
\Lambda^+(\bar{n}) = \bigcap_{k \in \mathbb{N}} (h_{n_k} \circ \cdots \circ h_{n_0})^{-1}(Z)
\]

is the graph of a contracting map \( \Phi_{\bar{n}} : X \to Y \), with

\[
\text{Lip}(\Phi_{\bar{n}}) \leq \sup_{n \in A} \{ \sigma_n (1 - \nu_n \kappa_n)^{-1} \} < 1.
\]

Moreover, if we endow the space of continuous function \( C(X,Y) \) with the topology of pointwise convergence, then the map

\[
\bar{n} \in A^\mathbb{N} \mapsto \Phi_{\bar{n}} \in C(X,Y)
\]

is continuous.

The proof is the same as in Theorem 4.2, with some obvious modifications.

3. Now we consider two metric spaces \( F^+ \) and \( F^- \), and two complete subspaces \( X \subseteq F^- \) and \( Y \subseteq F^+ \). Let \( V \) another complete metric space, and \( Z = X \times Y \times V \).

One has to think of \( V \) as a central direction and \( F^- \) (resp. \( F^+ \)) as a contracting (resp. expanding) direction. We consider two families of maps \( (h_n^+)_{n \in A} \) and \( (h_n^-)_{n \in A} \) of the form

\[
h_n^+ : Z \to X \times F^+ \times V, \quad h_n^- : Z \to F^- \times Y \times V,
\]

and we decompose them as \( h_n^+ = (f_n^+, g_n^+) \), where

\[
f_n^+ : Z \to X \times V, \quad f_n^- : Z \to Y \times V
\]

and

\[
g_n^+ : Z \to F^+, \quad g_n^- : Z \to F^-.
\]

The hypotheses (H) and (L') for \( h_n^\pm \) refer to these decompositions. Let us also denote by

\[
F_n^+ : Z \to V
\]

the second component of the map \( f_n^+ \), and consider sets \( Z_n^+ \subseteq Z \) and \( Z_n^- \subseteq Z \) such that

\[
Z \cap (h_n^+)^{-1}(Z) \subseteq Z_n^+, \quad Z \cap (h_n^-)^{-1}(Z) \subseteq Z_n^-,
\]

for any \( n \in A \).

4. The aim of this section is to prove the following result.

16
Proposition 4.4. With the previous notations, assume that $X,Y$ are bounded, $V$ is locally compact and that for any $n \in A$ the maps

$$h_n^+ : Z \longrightarrow X \times F^+ \times V, \quad h_n^- : Z \longrightarrow F^- \times Y \times V$$

satisfy hypotheses (H) and (L'). Let us also assume that for any $n \in A$, (i) the maps $(h_n^+|_{Z_n^+})$ are invertible and

$$\left( (h_n^+|_{Z_n^+})^{-1} = (h_n^-|_{Z_n^-}), \quad \left( (h_n^-|_{Z_n^-})^{-1} = (h_n^+|_{Z_n^+}) \right)$$

(ii) the sets $Z_n^+$ (resp. $Z_n^-$) are pairwise disjoint.

Then for any sequence $\bar{n} \in \Sigma_A$, the map $h_{\bar{n}0}^+$ has an invariant set $\Lambda$ and there exist a homeomorphism

$$\Upsilon : \Sigma_A \times V \longrightarrow \Lambda$$

$$(\bar{n}, v) \longmapsto (\Phi_{\bar{n}}(v), v)$$

which conjugates $h_{n0}^+|_{\Lambda}$ to the skew-product on $V$ given by

$$F_{\bar{n}}(v) = F_{\bar{n}0}/\Phi_{\bar{n}}(v), \quad \bar{n} \in \Sigma_A, \ v \in V,$$

where $\Phi_{\bar{n}} : V \rightarrow X \times Y$ is a contracting map, with

$$\text{Lip}(\Phi_{\bar{n}}) \leq \sup_{n \in A}\{((\sigma_n^+(1 - \nu_n^+ \kappa_n^+))^{-1}, (\sigma_n^-(1 - \nu_n^- \kappa_n^-))^{-1})\} < 1.$$

As a consequence, the functions $F_{\bar{n}} : V \rightarrow V$, for $\bar{n} \in \Sigma_A$, are also defined by the following equation

$$h_{n0}^+(\Phi_{\bar{n}}(v), v) = (\Phi_{\bar{n}}(F_{\bar{n}}(v)), F_{\bar{n}}(v)), \quad v \in V. \quad (5)$$

This remark will be useful later on to obtain the estimates (4) in Proposition 4.4.

Proof. For $\bar{n} \in \Sigma_A$, let us define

$$\bar{n}_+ = (n_k)_{k \in \mathbb{N}} \in A^\mathbb{N}, \quad \bar{n}_- = (n_{-k})_{k \in \mathbb{N}^*} \in A^{\mathbb{N}^*}.$$

Since each family of maps $(h_n^+|_{\Lambda})_{n \in A}$ and $(h_n^-|_{\Lambda})_{n \in A}$ satisfies hypotheses (H) and (L'), we can apply Proposition 4.3 so both sets

$$\Lambda^+(\bar{n}_+) = \bigcap_{k \in \mathbb{N}} (h_{n_k}^+ \circ \cdots \circ h_{n_{0}}^+)^{-1}(Z)$$

and

$$\Lambda^-(\bar{n}_-) = \bigcap_{k \in \mathbb{N}^*} (h_{n_{-k}}^- \circ \cdots \circ h_{n_{-1}}^-)^{-1}(Z)$$
are graphs of contracting maps
\[ \Phi_{\bar{n}^+}^+ : X \times V \longrightarrow Y, \quad \Phi_{\bar{n}^-}^- : Y \times V \longrightarrow X, \]
that is
\[ \Lambda^+(\bar{n}^+) = \{(x, \Phi_{\bar{n}^+}^+(x, v), v) \mid x \in X, v \in V\} \]
and
\[ \Lambda^-(\bar{n}^-) = \{(\Phi_{\bar{n}^-}^-(y, v), y, v) \mid y \in Y, v \in V\}. \]
Moreover, one has
\[ \text{Lip}(\Phi_{\bar{n}^+}^+) \leq \sup_{n \in A} \{\sigma^+_n(1 - \nu^+_n \kappa^+_n)^{-1}\} < 1, \quad (6) \]
\[ \text{Lip}(\Phi_{\bar{n}^-}^-) \leq \sup_{n \in A} \{\sigma^-_n(1 - \nu^-_n \kappa^-_n)^{-1}\} < 1. \quad (7) \]
Therefore for each \( v \in V \), the maps
\[ \Phi_{\bar{n}^+}^+(., v) : X \longrightarrow Y, \quad \Phi_{\bar{n}^-}^-(., v) : Y \longrightarrow X, \]
are also contracting, and so are the maps
\[ \Phi_{\bar{n}^+}^+ \circ \Phi_{\bar{n}^+}^+ : X \longrightarrow X, \quad \Phi_{\bar{n}^-}^- \circ \Phi_{\bar{n}^-}^- : Y \longrightarrow Y. \]
Since \( X \) and \( Y \) are complete, these maps have fixed points \( x(\bar{n}, v) \in X \) and \( y(\bar{n}, v) \in Y \) from the contraction principle, and by uniqueness they satisfy
\[ \Phi_{\bar{n}^+}^+(x(\bar{n}, v), v) = y(\bar{n}, v), \quad \Phi_{\bar{n}^-}^-(y(\bar{n}, v), v) = x(\bar{n}, v). \]
Now let us define
\[ \Lambda(\bar{n}) = \Lambda^+(\bar{n}^+) \cap \Lambda^-(\bar{n}^-). \]
Then by the previous relation this set is non-empty since it is the graph of the contraction
\[ \Phi_{\bar{n}} : V \longrightarrow X \times Y \]
\[ v \longmapsto (x(\bar{n}, v), y(\bar{n}, v)), \]
and, from (6) and (7),
\[ \text{Lip}(\Phi_{\bar{n}}) \leq \sup_{n \in A} \{\sigma^+_n(1 - \nu^+_n \kappa^+_n)^{-1}, \sigma^-_n(1 - \nu^-_n \kappa^-_n)^{-1}\} < 1. \]
Moreover, as the maps
\[ \bar{n}^+ \in A^N \mapsto \Phi_{\bar{n}^+}^+ \in C(X \times V, Y), \quad \bar{n}^- \in A^N \mapsto \Phi_{\bar{n}^-}^- \in C(Y \times V, X), \]
are continuous, from the contraction principle one also has the continuity of the map
\[ \bar{n} \in \Sigma_A \mapsto \Phi_{\bar{n}}(v) \in X \times Y. \]
Now set
\[ \Lambda = \bigcup_{\bar{n} \in \Sigma_A} \Lambda(\bar{n}). \]

The fact that this set is invariant under \( h_{n_0}^+ \) will follow from our condition (i). Indeed, if \( z \in \Lambda \), then \( z \in \Lambda(\bar{n}) \) for some \( \bar{n} \in \Sigma_A \) and so by definition
\[ z \in \bigcap_{k \in \mathbb{N}} (h_{n_k}^+ \circ \cdots \circ h_{n_0}^+)^{-1}(Z), \quad z \in \bigcap_{k \in \mathbb{N}^*} (h_{n_0}^- \circ \cdots \circ h_{n_{-1}}^-)^{-1}(Z). \]

From the first relation we get
\[ h_{n_0}^+(z) \in \bigcap_{k \in \mathbb{N}^*} (h_{n_k}^+ \circ \cdots \circ h_{n_0}^+)^{-1}(Z) = \Lambda^+(\sigma(\bar{n})_+), \quad (8) \]
and since by hypothesis
\[ \left. (h_{n_0}^+ \big| \mathbb{Z}_{n_0}^+)^{-1} \right| = (h_{n_0}^- \big| \mathbb{Z}_{n_0}^-)^{-1} \]
and \( \Lambda(\bar{n}) \subseteq \mathbb{Z}_{n_0}^+ \), we get from the second relation
\[ h_{n_0}^+(z) \in \bigcap_{k \in \mathbb{N}} (h_{n_0}^- \circ \cdots \circ h_{n_{-1}}^-)^{-1}(Z) = \Lambda^-(\sigma(\bar{n})_-). \quad (9) \]

Now (8) and (9) means exactly that \( h_{n_0}^+(z) \in \Lambda(\sigma(\bar{n})) \), so \( \Lambda \) is positively invariant under \( h_{n_0}^+ \). In fact, a completely similar argument using condition (i) shows that
\[ h_{n_0}^+(\Lambda(\bar{n})) = \Lambda(\sigma(\bar{n})), \]
and hence \( \Lambda \) is totally invariant under \( h_{n_0}^+ \). More generally, one obtains
\[ h_{n_{k-1}}^+ \circ \cdots \circ h_{n_0}^+(\Lambda(\bar{n})) = \Lambda(\sigma^{\pm k}(\bar{n})), \quad k \in \mathbb{Z}. \]

Next let us prove that as a consequence of (ii) the union
\[ \Lambda = \bigcup_{\bar{n} \in \Sigma_A} \Lambda(\bar{n}) \]
is disjoint. Let \( \bar{n} \neq \bar{n}' \), so there exists \( l \in \mathbb{Z} \) such that \( n_l \neq n'_l \). First suppose that \( l = 0 \), then on the one hand
\[ \Lambda(\bar{n}) \subseteq \Lambda(\bar{n}_+)^+ \subseteq \mathbb{Z}_{n_0}^+ \]
and on the other hand
\[ \Lambda(\bar{n}') \subseteq \Lambda(\bar{n}'_+) \subseteq \mathbb{Z}_{n_0}^+. \]
As $Z^+_{n_0}$ and $Z^+_0$ are disjoint by hypothesis, then so are $\Lambda(\bar{n})$ and $\Lambda(\bar{n'})$. Now if $l \geq 1$, we can assume without loss of generality that $n_k = n_k'$ for $0 \leq k \leq l - 1$, and as before

$$h^+_{n_{l-1}} \circ \cdots \circ h^+_{n_0}(\Lambda(\bar{n})) = \Lambda(\sigma^l(\bar{n})) \subseteq \Lambda(\sigma^l(\bar{n})_+) \subseteq Z^+_{nl}$$

and

$$h^+_{n_{l-1}} \circ \cdots \circ h^+_{n_0}(\Lambda(\bar{n}')) = \Lambda(\sigma^l(\bar{n}')) \subseteq \Lambda(\sigma^l(\bar{n}')_+) \subseteq Z^+_{n'l},$$

so $\Lambda(\bar{n})$ and $\Lambda(\bar{n}')$ have to be disjoint. Finally, the case $l \leq -1$ is completely similar using the hypothesis that $Z^-_n$ are pairwise disjoint for $n \in A$.

To conclude, as $\Lambda$ is a disjoint union, every point $z \in \Lambda$ can be uniquely written as $z = (\Phi_{\bar{n}}(v), v)$ for $v \in V$ and $\bar{n} \in \Sigma_A$, so the map

$$\Upsilon : \Sigma_A \times V \longrightarrow \Lambda$$

$$(\bar{n}, v) \longmapsto (\Phi_{\bar{n}}(v), v)$$

is a well-defined continuous bijection, and as $V$ is locally compact, one can check that it is a homeomorphism with respect to the product topology on $\Sigma_A \times V$. Then for $z = (\Phi_{\bar{n}}(v), v) \in \Lambda(\bar{n})$, as $h^+_{n_0}(z) \in \Lambda(\sigma(\bar{n}))$ we have from the definition of $F^+_n$,

$$h^+_{n_0}(z) = \left(\Phi_{\sigma(\bar{n})}(F^+_{n_0}(z)), F^+_{n_0}(z)\right) = \left(\Phi_{\sigma(\bar{n})}(F^+_{n_0}(\Phi_{\bar{n}}(v), v)), F^+_{n_0}(\Phi_{\bar{n}}(v), v)\right).$$

The last equality exactly means that the map

$$h^+_{n_0} : \Lambda \longrightarrow \Lambda$$

is conjugated by $\Upsilon$ to the skew-product

$$G : \Sigma_A \times V \longrightarrow \Sigma_A \times V$$

$$(\bar{n}, v) \longmapsto (\sigma(\bar{n}), F(\bar{n}, v)).$$

where $F(\bar{n}, v) = F^+_{n_0}(\Phi_{\bar{n}}(v), v)$. This ends the proof.

### 4.2 Proof of Proposition 4.1

This section is entirely devoted to the proof of Proposition 4.1, this will be done in several steps. We will have to use notations and estimates on time-energy coordinates, contained in Appendix A, and this will require to choose $\mu$ sufficiently small. Moreover, we shall use various coordinates but we shall keep the same notation for the diffeomorphisms expressed in different coordinates.

**Step 1. Straightening of the invariant manifolds.**
Using time-energy coordinates on the first factor, our homoclinic annulus is given by
\[ \mathcal{I}_\mu = \{ (\tau, e, \theta, I_2) \in \tilde{D}_s | \tau = e = 0 \} \]
and in a neighbourhood of it, the stable manifold of \( F_\mu \) is given by
\[ \{ (\tau, e, \theta, I_2) \in \tilde{D}_s | e = 0 \} \]
while the unstable manifold is
\[ \{ (\tau, e, \theta, I_2) \in \tilde{D}_s | e = \mu f_2(\theta_2)\tau \}. \]
Now we introduce the change of coordinates
\[ \Theta : \mathcal{E}_s \times \mathbb{A} \rightarrow \mathcal{E}_s' \times \mathbb{A} \]
\[ (\tau, e, \theta, I_2) \mapsto (\tau', e', \theta, I_2) \]
where
\[ \tau' = \tau - (\mu f_2(\theta_2))^{-1} e, \quad e' = (\mu f_2(\theta_2))^{-1} e. \]
Its inverse is given by
\[ \Theta^{-1} : \mathcal{E}_s' \times \mathbb{A} \rightarrow \mathcal{E}_s \times \mathbb{A} \]
\[ (\tau', e', \theta, I_2) \mapsto (\tau + e', \mu f_2(\theta_2)e', \theta, I_2). \]
It follows that in these new coordinates, denoting by \( \tilde{D}'_s = \Theta(\tilde{D}_s) \), the stable manifold
\[ \{ (\tau', e', \theta, I_2) \in \tilde{D}'_s | e' = 0 \} \]
and the unstable manifold
\[ \{ (\tau', e', \theta, I_2) \in \tilde{D}'_s | \tau' = 0 \} \]
are straightened out.

**Step 2. Choice of the box \( Z \).**

Our goal is to use Proposition 4.4, so we will explain how to choose the domain \( Z \). Our central direction \( V = \mathbb{A} \) will be the annulus given by the coordinates \( (\theta_2, I_2) \), our contracting and expanding directions will be one-dimensional, so \( X \subseteq F^- = \mathbb{R} \) and \( Y \subseteq F^+ = \mathbb{R} \). We will choose
\[ X = [-\tau_0', \tau_0'], \quad Y = [0, e_0'], \]
with \( \tau_0' = c_1 \mu^{2\pi - 1}, e_0' = c_2 \mu^{2\pi - 1} \), and \( c_1 > 2\pi, c_2 > 2\pi \). Eventually
\[ Z = [-\tau_0', \tau_0'] \times [0, e_0'] \times \mathbb{A}. \]
For \( (\theta_2, I_2) \in \mathbb{A} \), let us also define the section
\[ Z(\theta_2, I_2) = [-\tau_0', \tau_0'] \times [0, e_0'] \times \{ \theta_2, I_2 \}, \]
where all the constructions will take place (see figure 2).

This domain $Z$ is located above the homoclinic annulus $I_\mu$, and for $\mu$ small enough it is contained in the domain $\tilde{D}'$, where the manifolds are straightened.

For $n \in A = \{\lfloor \ln \mu^{-1} \rfloor, \ldots, 2 \lceil \ln \mu^{-1} \rceil \}$, the point $a_n = (0, e_n) \in \tilde{D}'$, or $a'_n = (-e'_n, e'_n) \in \tilde{D}'$, is by definition $n$-periodic for the pendulum map $\Phi^P$. We want the annulus $\{a'_n\} \times A$, which is $n$-periodic for the unperturbed map $F_0$, to be included in our domain $Z$, and this requires that $e'_n < e'_0$ for any $n \in A$. First note that

$$e_0 = \mu f_2(\theta_2) e'_0 = c_2 f_2(\theta_2) \mu^{2\pi},$$

and as $f_2(\theta_2) \geq \pi^{-1}$, $c_2 f_2(\theta_2) > 2$ so $e_0 > 2\mu^{2\pi}$. Then using (21) and the fact that $A = \{\lfloor \ln \mu^{-1} \rfloor, \ldots, 2 \lceil \ln \mu^{-1} \rceil \}$, for $\mu$ is small enough (and hence $n$ is large enough) one obtains

$$\sup_{n \in A} \{e_n\} < 2\mu^{2\pi},$$

and therefore $e_n < e_0$ is satisfied for any $n \in A$, which gives $e'_n < e'_0$. Similarly, we obtain

$$-\tau'_0 < -2\pi \mu^{2\pi - 1} < -e'_n.$$

**Step 3. Construction of the domains $Z^+_n$.**

For $n \in A$, the domains $Z^+_n \cap Z(\theta_2, I_2)$ will be rectangles (see figure 3, they are parallelograms in the coordinates $(\tau, e)$), with vertices

$$A^1_n = \left( -\tau'_0, -e'_0 + \frac{\delta e}{\mu f_2(\theta_2)}, \theta_2, I_2 \right), \quad A^2_n = \left( -\tau'_0, -e'_0 - \frac{\delta e}{\mu f_2(\theta_2)}, \theta_2, I_2 \right),$$

where $\delta e = e'_0 - e'_n$. (end of page)
\[ A_n^3 = \left( \tau_0, \frac{e_n + \delta_n^e}{\mu f_2(\theta_2)}, \theta_2, I_2 \right), \quad A_n^4 = \left( \tau_0, \frac{e_n - \delta_n^e}{\mu f_2(\theta_2)}, \theta_2, I_2 \right), \]

where the parameter \( \delta_n^e \) is defined by

\[
\delta_n^e = \frac{c_3 \mu^{2\pi - 1}}{|T_n'|}, \quad n \in A
\]

for some constant \( c_3 > \max\{c_1 + c_2, c_1 + 2\pi\} \), and \( T_n' = T'(e_n) \) (see Appendix A and (22)). Hence the domains \( Z_n^+ \) are small neighbourhoods of the annuli \( \{a_n' \} \times A_n \), for \( a_n = (-e_n', e_n') \in \mathcal{D}'_n \).

For \( \mu \) small enough, one can easily check from the definition of \( \delta_n^e \) and the mean value theorem that

\[
n + c_3 \mu^{2\pi - 1} \leq T(e_n - \delta_n^e) \leq n + 2c_3 \mu^{2\pi - 1} \tag{10}
\]

and

\[
n - 2c_3 \mu^{2\pi - 1} \leq T(e_n + \delta_n^e) \leq n - c_3 \mu^{2\pi - 1}. \tag{11}
\]

Let us show that these domains \( Z_n^+ \) are pairwise disjoint. By construction it is enough to prove that \( e'(A_n^1) > e'(A_n^{2+1}) \). Choosing \( \mu \) small so

\[
n + 2c_3 \mu^{2\pi - 1} < n + 1 - 2c_3 \mu^{2\pi - 1},
\]

from (10) and (11) we get

\[
T(e_n - \delta_n^e) < T(e_n + \delta_n^{e+1}).
\]

But as the period function \( T \) is decreasing, this gives

\[
e_n - \delta_n^e > e_{n+1} + \delta_{n+1}^e
\]

which implies that \( e'(A_n^1) > e'(A_n^{2+1}) \).

Let us also remark that by definition of our domain \( \mathcal{D}'_n \), one can ensure that for \( \mu \) small enough (and so \( n - T(e) \) is small enough)

\[
\mathcal{F}_0^k(Z_n^+) \cap \mathcal{D}'_n = \emptyset, \quad 0 < k < n. \tag{12}
\]

**Step 4. Construction of the domains \( Z_n^- \).**

The construction is similar (see figure 4), namely \( Z_n^- \cap Z(\theta_2, I_2) \) is a rectangle with vertices

\[
B_n^1 = \left( -\frac{e_n}{\mu f_2(\theta_2)} - \delta_n^\pi, e_0', \theta_2, I_2 \right), \quad B_n^2 = \left( -\frac{e_n}{\mu f_2(\theta_2)} + \delta_n^\pi, e_0', \theta_2, I_2 \right),
\]

\[
B_n^3 = \left( -\frac{e_n}{\mu f_2(\theta_2)} + \delta_n^\pi, 0, \theta_2, I_2 \right), \quad B_n^4 = \left( -\frac{e_n}{\mu f_2(\theta_2)} - \delta_n^\pi, 0, \theta_2, I_2 \right),
\]

23
where $\delta_n^\tau$ is defined by

$$
\delta_n^\tau = \frac{c_4 \mu \pi^2 - 2 |T_n'|}{|T_n'|},
$$

for a constant $c_4 > \pi \max\{c_1+4\pi, c_1+c_2\}$, and $T_n' = T'(e_n)$ (see Appendix A and (22)). As in the previous step, one can check that $\tau'(B_{n+1}^1) > \tau'(B_{n}^2)$ so the domains $Z_n^-$ are pairwise disjoint.

**Step 5. Expressions of the maps $F_{\mu}^{\pm n}$ restricted to $Z_n^\pm$.**

For any $(\tau, e, \theta_2, I_2) \in \Theta^{-1}(Z_n^\pm)$, one has the following explicit expression for the unperturbed map

$$
F_0^n(\tau, e, \theta_2, I_2) = (\tau + n - T(e), e, \theta_2 + nI_2, I_2).
$$

In those coordinates, our perturbation is given by

$$
\Phi_{\mu}^{\pm f}(\tau, e, \theta_2, I_2) = (\tau, e + \mu\tau f_2(\theta_2), \theta_2, I_2 - \mu f_1(\tau) f_2'(\theta_2)).
$$

Then, from (12) we know that $F_{\mu}^n = \Phi_{\mu f} \circ F_0^n$ when restricted to $\Theta^{-1}(Z_n^\pm)$, so

$$
F_{\mu}^n(\tau, e, \theta_2, I_2) = (\tau + n - T(e), e + \mu(\tau + n - T(e)) f_2(\theta_2 + nI_2),
$$

$$
\theta_2 + nI_2, I_2 - \mu f_1(\tau + n - T(e)) f_2'(\theta_2 + nI_2)).
$$

Using the expression for $\Theta^{-1}$, for $(\tau', e', \theta_2, I_2) \in Z_n^\pm$ we compute

$$
F_{\mu}^n(\tau', e', \theta_2, I_2) = (-f_2(\theta_2 + nI_2))^{-1} f_2(\theta_2) e',
$$

$$
(1 + f_2(\theta_2 + nI_2))^{-1} f_2(\theta_2) e' + \tau' + n - T(\mu f_2(\theta_2) e'),
$$

$$
\theta_2 + nI_2,
$$

$$
I_2 - \mu f_1(\tau' + e' + n - T(\mu f_2(\theta_2) e')) f_2'(\theta_2 + nI_2)).
$$

24
If we fix \((\theta_2, I_2) \in A\), then the first two components of this map are linear with respect to \(\tau'\) and \(e'\), therefore the image of the parallelogram \(Z_+^n\) by the map \(\mathcal{F}_\mu^n\) is still a parallelogram (in a different section), with vertices \(\tilde{A}_i^n = \mathcal{F}_\mu^n(A_i^n)\) for \(i = 1, 2, 3, 4\) which can be explicitly computed.

Similarly, for \((\tau', e', \theta_2, I_2) \in Z_-^n\) we compute

\[
\mathcal{F}_\mu^{-n}(\tau', e', \theta_2, I_2) = \left((1 + (f_2(\theta_2))^{-1}f_2(\theta_2))\tau' + e' - k + T(-\mu f_2(\theta_2)\tau'), \right.
\]

\[
\left. (f_2(\tilde{\theta}_2))^{-1}f_2(\theta_2)\tau', \theta_2 + nI_2 - n\mu f_1(\tau' + e')f_2'(\theta_2), I_2 + \mu f_1(\tau' + e')f_2'(\theta_2)\right),
\]

so the image of \(Z_-^n\) by the map \(\mathcal{F}_\mu^{-n}\) is a parallelogram with vertices \(\tilde{B}_i^n = \mathcal{F}_\mu^{-n}(B_i^n)\) for \(i = 1, 2, 3, 4\).

**Step 6. Relative position of \(Z_\pm^n\) and \(\mathcal{F}_\mu^n(Z_\pm^n)\).**

Here we will prove that the figures (5) and (6) make sense, that is we will show that the horizontal (resp. vertical) edges of \(\mathcal{F}_\mu^n(Z_+^n)\) (resp. \(\mathcal{F}_\mu^{-n}(Z_-^n)\)) are not contained in \(Z\). The upper horizontal edge of \(Z_+^n\) is the segment joining \(A_2^n\) to \(A_3^n\), therefore the upper horizontal edge of \(\mathcal{F}_\mu^n(Z_+^n)\) is the segment joining \(\tilde{A}_2^n\) to \(\tilde{A}_3^n\). We can compute

\[
e'(\tilde{A}_2^n) = \left(\frac{1}{f_2(\theta_2 + nI_2)} + \frac{1}{f_2'(\theta_2)}\right) \mu^{-1}e_n - \tau_0 + n - T(e_n + \delta_n^e)
\]

and

\[
e'(\tilde{A}_3^n) = \left(\frac{1}{f_2(\theta_2 + nI_2)} + \frac{1}{f_2'(\theta_2)}\right) \mu^{-1}e_n + \tau_0 + n - T(e_n + \delta_n^e),
\]

Figure 4: Section \(Z_-^n(\theta_2, I_2)\) of the domain \(Z_-^n\)
From (11) we have
\[ T(e_n + \delta_n) \leq n - c_3\mu^{2\pi - 1}, \]
and as \( \tau'_0 = c_1\mu^{2\pi - 1} \) and \( c_3 > c_1 + c_2 \) this gives
\[ e'(\tilde{A}_n^2) > -\tau'_0 + n - T(e_n + \delta_n) \]
\[ > (c_3 - c_1)\mu^{2\pi - 1} > c_2\mu^{2\pi - 1} > \epsilon'_0. \]

We also have
\[ e'(\tilde{A}_n^3) = e'(\tilde{A}_n^2) + 2\tau'_0 > \epsilon'_0, \]
and so the upper horizontal edge of \( Z_n^+ \) is not contained in \( Z \).

For the lower horizontal edge of \( F_{\mu}^{-n}(Z_n^-) \), which is the segment joining \( \tilde{A}_n^4 \) to \( \tilde{A}_n^1 \), one has to prove that \( e'(\tilde{A}_n^4) < 0 \) and \( e'(\tilde{A}_n^1) < 0 \). We compute
\[ e'(\tilde{A}_n^4) = \left( \frac{1}{f_2(\theta_2 + nI_2)} + \frac{1}{f_2(\theta_2)} \right) \mu^{-1}e_n + \tau'_0 + n - T(e_n - \delta_n) \]
\[ \leq 2\pi\mu^{-1}e_n + (c_1 - 2c_3)\mu^{2\pi - 1} \]
\[ \leq (c_1 - 2c_3 + 2\pi)\mu^{2\pi - 1}, \]
and as \( 2c_3 > c_3 > c_1 + 2\pi \), this gives \( e'(\tilde{A}_n^4) < 0 \). We also have
\[ e'(\tilde{A}_n^1) = e'(\tilde{A}_n^4) - 2\tau'_0 < 0, \]
and so the lower horizontal edge of \( Z_n^+ \) is not contained in \( Z \).

Similarly, one can check that the vertical edges of \( F_{\mu}^{-n}(Z_n^-) \) are not contained in \( Z \), and this follows from the choice of the constant \( c_4 \).

**Step 7. Definition of the maps** \( h_n^+ \) and \( h_n^- \).

Our maps \( h_n^+ \) and \( h_n^- \) will be suitable extensions of the maps \( F_{\mu}^n \) and \( F_{\mu}^{-n} \). First we set
\[ (h_n^+)|_{Z_n^+} = (F_{\mu}^n)|_{Z_n^+}, \quad (h_n^-)|_{Z_n^-} = (F_{\mu}^{-n})|_{Z_n^-}, \]

Figure 5: Position of \( Z_n^+ \) and \( F_{\mu}^n(Z_n^+) \)
and we want to define Lipschitz extensions of $h_n^\pm$ to $Z$ in order to have

$$Z \cap (h_n^\pm)^{-1}(Z) \subseteq Z_n^\pm.$$  

Let us begin with the maps $h_n^+$. Take $z = (\tau', e', \theta_2, I_2) \in Z \setminus Z_n^+$, then we can find a unique $\tilde{z} = (\tau', e'(\tilde{z}), \theta_2, I_2) \in Z_n^+$: indeed, either

$$\frac{e_n + \delta_n^e}{\mu f_2(\theta_2)} < e' < e'_0$$

in which case we choose

$$e'(\tilde{z}) = \frac{e_n + \delta_n^e}{\mu f_2(\theta_2)},$$

or

$$0 < e' < \frac{e_n - \delta_n^e}{\mu f_2(\theta_2)}$$

and then we choose

$$e'(\tilde{z}) = \frac{e_n - \delta_n^e}{\mu f_2(\theta_2)}.$$  

Then we set

$$h_n^+(z) = F_n^\mu(\tilde{z}) + (0, \alpha_n^+(e'(\tilde{z}) - e'(z)), 0, 0),$$

for some positive constant $\alpha_n^+$ yet to be chosen. Hence the map

$$h_n^+ : Z \rightarrow [-\tau'_0, \tau'_0] \times F^+ \times A$$

is a well-defined Lipschitz extension of $F_n^\mu$ and, using the form of the extension, one can check that $Z \cap (h_n^+)^{-1}(Z) \subseteq Z_n^+$. 

27
For the maps \( h^-_{n} \), this is completely analogous. For \( z = (\tau', e', \theta_2, I_2) \in Z \setminus Z^-_{n} \), then we can find a unique \( \tilde{z} = (\tau'(\tilde{z}), e', \theta_2, I_2) \in Z^-_{n} \) where

\[
\tau'(\tilde{z}) = -\frac{e_n}{\mu f_2(\theta_2)} + \delta_n^\tau
\]

if

\[
-\frac{e_n}{\mu f_2(\theta_2)} + \delta_n^\tau < \tau' < \tau'_0,
\]

or

\[
\tau'(\tilde{z}) = -\frac{e_n}{\mu f_2(\theta_2)} - \delta_n^\tau
\]

if

\[
-\tau'_0 < \tau' < -\frac{e_n}{\mu f_2(\theta_2)} + \delta_n^\tau.
\]

Then we define

\[
h^-_{n}(z) = F^-_{\mu}(\tilde{z}) + (\alpha^-_{n}(\tau'(z) - \tau'(\tilde{z})), 0, 0, 0)
\]

for some constant positive \( \alpha^-_{n} \). The map

\[
h^+_{n}: Z \to F^- \times [0, e'_0] \times A
\]

is a well-defined Lipschitz extension of \( F^-_{\mu} \), with \( Z \cap (h^-_{n})^{-1}(Z) \subseteq Z^-_{n} \).

**Step 8. Verification of Hypotheses (H) and (L').**

Now let us show that the maps \( h^+_n \) and \( h^-_{n} \) satisfy hypotheses (H) and (L'). Recall that

\[
X = [-\tau'_0, \tau'_0] \subseteq F^- = \mathbb{R}, \quad Y = [0, e'_0] \subseteq F^+ = \mathbb{R},
\]

and let us write \( h^+_n = (f^+_n, g^+_n) \) where

\[
f^+_n: X \times Y \times A \to X \times A
\]

and

\[
g^+_n: X \times Y \times A \to F^+.
\]

For \( x \in X \times A \), consider the partial map

\[
g^+_{n,x}: Y \to F^+.
\]

By step 6 (see the figure 5), the image under \( (h^+_n)|_{Z^+_{n}} = (F^+_n)|_{Z^+_{n}} \) of the horizontal edges of \( Z^+_{n} \) do not belong to \( Z \), and this implies that

\[
Y = [0, e'_0] \subseteq g^+_{n,x}(Y \cap Z^+_{n}) \subseteq g^+_{n,x}(Y),
\]

so the hypothesis (H2) is satisfied.
Now it remains to show hypothesis (H1) and (L'). This follows from the choice of \( \alpha^+ \) and lengthy calculations of the various partial derivatives of \( F_n^\mu \), using both the explicit expression obtained in step 5 and the estimates of Appendix A. We find the following: if we define

\[
\alpha^+_n = \pi^{-1}(\mu|T'_n|),
\]

then (H1) is satisfied with

\[
\rho^+_n = O\left(\mu|T'_n|\right),
\]

and we estimate the Lipschitz constants

\[
\nu^+_n = O\left(\mu^{-1}|T'_n|^{-1}\right), \quad \sigma^+_n = O\left(\mu^{2\pi-1}\right), \quad \kappa^+_n = O\left(\mu^{2\pi-1}|T'_n|\right),
\]

so (L') is satisfied for \( \mu \) small enough as

\[
\sigma^+_n + \nu^+_n \max\{1, \kappa^+_n\} = \sigma^+_n + \nu^+_n \kappa^+_n = O\left(\mu^{2\pi-1}\right).
\]

The situation for \( h^-_n \) is of course similar.

\textbf{Step 9. Conclusions.}

All the hypotheses of Proposition 4.4 are satisfied, so we can apply it to any sequence \( \bar{n} \in A \): there exist a set \( \Lambda \subseteq Z \), invariant by \( h^+_n |_\Lambda \), and a homeomorphism

\[
\Upsilon : \Sigma_A \times \Lambda \to \Lambda
\]

\[
(n, (\theta, I)) \mapsto (\Phi_{\bar{n}}(\theta, I), (\theta, I))
\]

which conjugates \( h^+_n |_\Lambda \) to the skew-product on \( \Lambda \) given by

\[
F_{\bar{n}}(\theta, I) = F^+_n(\Phi_{\bar{n}}(\theta, I), \theta, I), \quad (\theta, I) \in \Lambda,
\]

where \( \Phi_{\bar{n}} : \Lambda \to X \times Y \) is a contracting map, with

\[
\text{Lip}(\Phi_{\bar{n}}) \leq \sup_{n \in A} \left\{ (\sigma^+_n(1 - \nu^+_n \kappa^+_n)^{-1}), (\sigma^-_n(1 - \nu^-_n \kappa^-_n)^{-1}) \right\} = O\left(\mu^{2\pi-1}\right). \quad (13)
\]

Recall also from (5) that

\[
h^+_n(\Phi_{\bar{n}}(\theta, I), \theta, I) = (\Phi_{\bar{n}}(F_{\bar{n}}(\theta, I), \theta, I), F_{\bar{n}}(\theta, I), \quad (\theta, I) \in \Lambda. \quad (14)
\]

So we let

\[
\Omega_\mu = \Theta^{-1}(Z), \quad \Lambda_\mu = \Theta^{-1}(\Lambda) \subseteq \Omega_\mu, \quad \Upsilon_\mu = \Theta^{-1} \circ \Upsilon.
\]

For \( x \in \Lambda_\mu \),

\[
h^+_n(x) = \mathcal{F}_\mu^\bar{n}(x) = \mathcal{F}_\mu(x),
\]

29
where \( \tilde{F}_\mu \) is the transversal map of \( F_\mu \) associated to \( O_\mu \). Then by definition of \( \Lambda_\mu \), the sequence \( \tilde{n}^x \) is a well-defined element of \( \Sigma_A \). Setting \( n = (\tau_x, e_x, \theta, I_x) \in \Lambda_\mu \), the above conjugacy gives

\[
\tilde{F}_\mu \circ \Upsilon_\mu(\tilde{n}^x, \theta_x, I_x) = \Upsilon_\mu(\sigma(\tilde{n}^x), F_{\tilde{n}^x}(\theta_x, I_x)).
\]

Now it remains to explain the form of \( F_{\tilde{n}} \) and the estimates (4). We can write

\[
\Upsilon_\mu : \Sigma_A \times \mathbb{A} \to \Lambda_\mu
\]
as

\[
\Upsilon_\mu(\tilde{n}, (\theta, I)) = (\psi_{\tilde{n}}(\theta, I), (\theta, I)) = (\tau_{\tilde{n}}(\theta, I), e_{\tilde{n}}(\theta, I), \theta, I),
\]
where \( \tau_{\tilde{n}} : \mathbb{A} \to \mathbb{R} \) and \( e_{\tilde{n}} : \mathbb{A} \to \mathbb{R} \) are easily deduced from

\[
\Phi_{\tilde{n}} = (\Phi_{\tilde{n}}^1, \Phi_{\tilde{n}}^2) : \mathbb{A} \to X \times Y
\]
and from \( \Theta^{-1} \). Now by definition of \( X \) and \( Y \),

\[
|\Phi_{\tilde{n}}^1|_{C^0(\mathbb{A})} = O\left(\mu^{2\pi - 1}\right), \quad |\Phi_{\tilde{n}}^2|_{C^0(\mathbb{A})} = O\left(\mu^{2\pi - 1}\right),
\]
and this gives

\[
|\tau_{\tilde{n}}|_{C^0(\mathbb{A})} = O\left(\mu^{2\pi - 1}\right), \quad |e_{\tilde{n}}|_{C^0(\mathbb{A})} = O\left(\mu^{2\pi}\right).
\]
Moreover using (13) these maps are Lipschitz and

\[
\text{Lip}(\tau_{\tilde{n}}) = O\left(\mu^{2\pi - 1}\right), \quad \text{Lip}(e_{\tilde{n}}) = O\left(\mu^{2\pi - 1}\right).
\]

Now from (14) we can write

\[
F_{\tilde{n}}(\theta, I) = (\theta + n_0 I, I + 2\mu f_1(\phi_{\tilde{n}}(\theta, I)) \cos 2\pi(\theta + n_0 I)), \quad n_0 \in \Sigma_A, \ (\theta, I) \in \mathbb{A},
\]
where the Lipschitz function \( \phi_{\tilde{n}} : \mathbb{A} \to \mathbb{R} \) satisfies the equation

\[
\phi_{\tilde{n}}(\theta, I) = \tau_{\sigma(\tilde{n})}(\theta + n_0 I, I + 2\mu f_1(\phi_{\tilde{n}}(\theta, I)) \cos 2\pi(\theta + n_0 I)). \quad (15)
\]
The estimate

\[
|\phi_{\tilde{n}}|_{C^0(\mathbb{A})} = O\left(\mu^{2\pi - 1}\right)
\]
is obvious, but for the Lipschitz constant this is more subtle. Let us define

\[
g_{\tilde{n}}(\theta, I) = (\theta + n_0 I, I + 2\mu f_1(\phi_{\tilde{n}}(\theta, I)) \cos 2\pi(\theta + n_0 I)).
\]

Then on the one hand, using the definition of \( f_1 \) we can estimate

\[
\text{Lip}(g_{\tilde{n}}) \leq \sup \{1 + \mu^{-\frac{1}{2}} \ln \mu^{-1}, \text{Lip}(\phi_{\tilde{n}})\},
\]
but on the other hand, from (15)

\[
\text{Lip}(\phi_{\tilde{n}}) \leq \text{Lip}(\tau_{\tilde{n}}) \text{Lip}(g_{\tilde{n}}) < \text{Lip}(g_{\tilde{n}}),
\]
so

\[
\text{Lip}(g_{\tilde{n}}) = O\left(\mu^{-\frac{1}{2}} \ln \mu^{-1}\right)
\]
and therefore

\[
\text{Lip}(\phi_{\tilde{n}}) = O\left(\mu^{2\pi - \frac{1}{2}} \ln \mu^{-1}\right).
\]
This concludes the proof.
5 Construction of a pseudo-orbit

1. In this section, we will restrict our study to a special type of skew-product over a Bernoulli shift, which is called “polysystem” (see [Mar08], it is also known as Iterated Function System).

Recall from definition 2.4 that a polysystem is a skew-product (over $\sigma$) such that for any $\vec{n} = (n_k)_{k \in \mathbb{Z}} \in \Sigma_A$, one has $F_{\vec{n}} = f_{n_0}$. So a polysystem does not depend on the whole sequence $\vec{n} \in \Sigma_A$ but only on its first component $n_0 \in A$, hence instead of being defined by a family of maps indexed by $\Sigma_A$, it is defined by a family of maps indexed by $A$. Then one can easily see that a sequence $(n_k, x_k)_{k \in \mathbb{Z}} \in A \times M$ gives rise to an orbit $(\sigma^k(\vec{n}), x_k)_{k \in \mathbb{Z}} \in \Sigma_A \times M$, where $\vec{n} = (n_k)_{k \in \mathbb{Z}}$, for the polysystem $[[f_n]]_{n \in A}$ if and only if

$$f_n(x_k) = x_{k+1}, \quad k \in \mathbb{Z},$$

and its projection onto $M$ corresponds to the iteration of the maps $(f_n)_{n \in A}$ in the order prescribed by the sequence $\vec{n}$.

2. In the previous section, we showed that the dynamics of the first return map of our diffeomorphism $F_\mu$ in a neighbourhood of the homoclinic annulus is conjugated to the skew-product map

$$F_{\vec{n}}(\theta, I) = (\theta + n_0 I, I + 2\mu f_1(\phi_{\vec{n}}(\theta, I)) \cos 2\pi(\theta + n_0 I)), \quad \vec{n} \in \Sigma_A, (\theta, I) \in A.$$  

This family of maps depend on the whole sequence $\vec{n} \in \Sigma_A$, but as $\phi_{\vec{n}}$ is small, $f_1(\phi_{\vec{n}}(\theta, I))$ is close to one, and so $F_{\vec{n}}$ is close to

$$f_{n_0}(\theta, I) = (\theta + n_0 I, I + 2\mu \cos 2\pi(\theta + n_0 I)), \quad n_0 \in A, (\theta, I) \in A.$$  

These “standard” maps can be seen as perturbations of iterates of the integrable twist map $T(\theta, I) = (\theta + I, I)$: more precisely we can write $f_n = V \circ T^n$ where

$$V(\theta, I) = (\theta, I + 2\mu \cos 2\pi\theta)$$

is a “vertical” map which is close to identity if $\mu$ is small.

We will first construct an orbit for the polysystem $[[f_n]]_{n \in A}$ defined by the maps $f_n$, and in the next section, we will use it as a pseudo-orbit for the skew-product $[[F_{\vec{n}}]]_{\vec{n} \in \Sigma_A}$ defined by the maps $F_{\vec{n}}$.

3. The goal of this section is to prove the following proposition.

**Proposition 5.1.** There exists a positive constant $C$ such that for $\mu$ small enough, there exists an orbit $(\sigma^k(\vec{n}), x_k)_{k \in \mathbb{Z}} \in \Sigma_A \times A$ for the polysystem $[[f_n]]_{n \in A}$ defined by

$$f_n(\theta, I) = (\theta + n I, I + 2\mu \cos 2\pi(\theta + n I)), \quad n \in A, (\theta, I) \in A.$$
such that
\[ \lim_{k \to \pm\infty} I_k = \pm \infty, \]
and the estimates
\[ |I_N - I_0| \geq 2, \quad \sum_{k=0}^{N} n_k \leq C \mu^{-1} \ln \mu^{-1}, \]
hold true.

We will see that this proposition holds only if we can choose the integers \( n_k, k \in \mathbb{Z} \), as large as \( \mu^{-\frac{1}{2}} \ln \mu^{-1} \), so this explains the choice of
\[ A = \{ [\ln \mu^{-1}], \ldots, 2\mu^{-\frac{1}{2}} \ln \mu^{-1} \}. \]
The upper bound on the sum of the integers \( n_k \) needed to produce a drift of order one is basically the “time of instability” in this context.

In the sequel, we shall only explain the construction of the positive sequence \((n_k, \theta_k, I_k)_{k \geq 0}\), since, of course, the construction of the negative sequence \((n_k, \theta_k, I_k)_{k \leq -1}\) is completely similar.

4. Let us now describe the construction of our orbit. In all this section we will need to introduce two real numbers \( 0 < K < K' < \sqrt{\frac{2}{\pi}} \). First, we fix \( K \) such that \( 0 < K < \sqrt{\frac{2}{\pi}} \) and we define the non-empty domain
\[ B_K = \{ (\theta, I) \in \mathbb{A} | \cos 2\pi \theta \geq K, \sin 2\pi \theta \leq -K \}. \]
We shall only need the first condition \( \cos 2\pi \theta \geq K \) in this section, the other condition \( \sin 2\pi \theta \leq -K \) will be used in the next section.

One can also write
\[ B_K = \{ (\theta, I) \in \mathbb{A} | -\frac{\arccos K}{2\pi} \leq \theta \leq -\frac{\arcsin K}{2\pi} \ (\mathbb{Z}) \}, \]
that is \( B_K = J_\delta \times \mathbb{R} \) (see figure 7), where \( J_\delta \) is an interval of \( \mathbb{T} \) of length
\[ \delta = \frac{\arccos K}{2\pi} - \frac{\arcsin K}{2\pi}. \]
Now \( K \) being fixed, we shall also consider \( K' \in ]K, \sqrt{\frac{2}{\pi}}[ \) such that the domain \( B_{K'} = J_{\delta/2} \times \mathbb{R} \), where \( J_{\delta/2} \) is an interval of length \( \delta/2 \). Now recall that we have \( f_n = V \circ T^n \) with
\[ V(\theta, I) = (\theta, I + 2\mu \cos 2\pi \theta), \quad T(\theta, I) = (\theta + I, I), \quad (\theta, I) \in \mathbb{A} \]
and by definition, the map \( V \) leaves the set \( B_K \) invariant and produce in \( B_K \) a drift in the \( I \)-direction which is at least equal to \( 2\mu K \) by the condition \( \cos 2\pi \theta \geq K \). Therefore, to prove the first part of our proposition, we
will construct a sequence of points \((\theta_k, I_k)_{k \in \mathbb{Z}} \in \mathbb{A}\) for which we can find a sequence of integers \((n_k)_{k \in \mathbb{Z}} \in \mathbb{A}\) such that

\[
T^{n_k}(\theta_k, I_k) \in B_K.
\] (16)

Indeed, in this case \((\theta_{k+1}, I_{k+1}) = f_{n_k}(\theta_k, I_k) = V \circ T^{n_k}(\theta_k, I_k)\) satisfies

\[
I_{k+1} - I_k \geq 2\mu K.
\]

Then to prove our second part, we shall need estimates on these integers \(n_k, k \in \mathbb{Z}\). In fact the relation (16) can be written as

\[
R_{I_k}^{n_k}(\theta_k) = \theta_k + n_k I_k \in J_\delta,
\]

where \(R_I\) is the rotation on the circle \(\mathbb{T}\) of angle \(I\). Most of the time, that is if either \(I_k\) is irrational or if \(I_k = p/q\) with \(|q| \geq \delta^{-1}\), this will be realized, and the integers \(n_k, k \in \mathbb{Z}\), can be estimated. This is an easy consequence of the “ergodization” theorem recalled below, that we shall use crucially in our construction.

5. For \(I \in \mathbb{R}\), consider the rotation \(R_I(\theta) = \theta + I\) defined on the circle \(\mathbb{T}\). Given \(0 < \delta < 1\), let \(\mathcal{J}_\delta\) be the set of intervals of \(\mathbb{T}\) of length \(\delta\). We define the \(\delta\)-ergodization time \(N(I, \delta) \leq +\infty\) by

\[
N(I, \delta) = \inf\{n \in \mathbb{N} \mid \{\theta, \ldots, R^n_I(\theta)\} \cap J \neq \emptyset, \forall \theta \in \mathbb{T}, \forall J \in \mathcal{J}_\delta\},
\]

or equivalently

\[
N(I, \delta) = \inf\{n \in \mathbb{N} \mid \bigcup_{0 \leq k \leq n} R^{-k}_I(J) = \mathbb{T}, \forall J \in \mathcal{J}_\delta\}.
\]
One can easily see that $N(I,\delta) < +\infty$ except if $I$ is a rational number with a denominator smaller than $\delta^{-1}$, but it is more difficult to prove that when it is defined, this number is essentially given by the inverse of the distance to these “bad” rationals. This is the content of the theorem below, which is due to Berti, Biasco and Bolle ([BBB03]).

**Theorem 5.2** (Berti-Bolle-Biasco). There exist a positive constant $M$ such that if

$$\mathcal{R}_\delta = \{ p/q \in \mathbb{Q} \mid |q| \leq M\delta^{-1} \},$$

then for $I \in \mathbb{R} \setminus \mathcal{R}_\delta$,

$$N(I,\delta/2) \leq d(I,\mathcal{R}_\delta)^{-1}.$$  

This is a consequence of Theorem 4.2 in [BBB03] (see also the estimate (5.3) in [BBB03]), where the above proposition is proved both for the continuous and multi-dimensional case. Of course one can give an explicit value for the numerical constant $M$ but this will not be useful.

In fact, most of the time we shall use this result in the following form.

**Lemma 5.3.** Let $I \in \mathbb{R} \setminus \mathcal{R}_\delta$, $\theta \in \mathbb{T}$ and $J \subseteq \mathbb{T}$ any interval of length $\delta/2$. Then for any $m \in \mathbb{N}$, one can find an integer

$$n \in [m, m + d(I,\mathcal{R}_\delta)^{-1}]$$

such that $\theta + nI \in J$.

6. In order to construct our orbit, now we know that we have to consider the distance to this set of “bad rationals” $\mathcal{R}_\delta$.

So given $\mu > 0$, we first define the domain of “fast drift”

$$D_F(\mu) = \left\{ (\theta, I) \in \mathbb{A} \mid d(I,\mathcal{R}_\delta) \geq \frac{1}{\ln \mu^{-1}} \right\},$$

the domain of “slow drift”

$$D_S(\mu) = \left\{ (\theta, I) \in \mathbb{A} \mid \frac{\mu^{\frac{1}{2}}}{\ln \mu^{-1}} < d(I,\mathcal{R}_\delta) < \frac{1}{\ln \mu^{-1}} \right\},$$

and finally the domain of “resonances”

$$D_R(\mu) = \left\{ (\theta, I) \in \mathbb{A} \mid d(I,\mathcal{R}_\delta) \leq \frac{\mu^{\frac{1}{2}}}{\ln \mu^{-1}} \right\}.$$  

Obviously we have

$$\mathbb{A} = D_F(\mu) \cup D_S(\mu) \cup D_R(\mu).$$
But in the sequel, it will be more convenient to further decompose these sets. Indeed, for $p/q \in \mathbb{R}_\delta$, if we define

$$D_S(\mu, p/q) = \left\{ (\theta, I) \in \mathbb{A} \mid \frac{\mu^2}{\ln \mu - 1} < |I - p/q| < \frac{1}{\ln \mu - 1} \right\}$$

and

$$D_R(\mu, p/q) = \left\{ (\theta, I) \in \mathbb{A} \mid |I - p/q| \leq \frac{\mu^2}{\ln \mu - 1} \right\},$$

then, for $\mu$ small enough, one has

$$D_S(\mu) = \bigcup_{p/q \in \mathbb{R}_\delta} D_S(\mu, p/q), \quad D_R(\mu) = \bigcup_{p/q \in \mathbb{R}_\delta} D_R(\mu, p/q)$$

and

$$\mathbb{A} = D_F(\mu) \bigcup \left( \bigcup_{p/q \in \mathbb{R}_\delta} D_S(\mu, p/q) \right) \bigcup \left( \bigcup_{p/q \in \mathbb{R}_\delta} D_R(\mu, p/q) \right). \tag{17}$$

From now on we will assume that $\mu$ is sufficiently small, with respect to $\delta$ and $M$ which are fixed, so the above decomposition (17) is in fact a partition of $\mathbb{A}$ (since the set $\mathbb{R}_\delta$ is discrete) and moreover the following properties hold true: for any $(\theta, I) \in D_S(\mu)$, there exist a unique $p/q \in \mathbb{R}_\delta$ such that $(\theta, I) \in D_S(\mu, p/q)$, and if

$$m = \inf \{ n \in \mathbb{N}^* \mid f_n(\theta, I) \notin D_S(\mu, p/q) \},$$

then $m$ is well defined and necessarily the point $(\theta_m, I_m) = f_m(\theta, I)$ does not belong to $D_S(\mu)$ (either it is in $D_R(\mu)$ or $D_F(\mu)$, since the size of the jump, which is of order $\mu$, is much smaller than the width of the connected components of $D_R(\mu)$ or $D_F(\mu)$). We also ask the same for a point $(\theta, I) \in D_R(\mu)$: under the iteration of a map $f_n$ the first time it escapes the domain $D_R(\mu, p/q)$, it also escapes the domain $D_R(\mu)$ (in fact it enters into the domain $D_S(\mu)$). This situation is depicted in figure 8.

7. The construction of our orbit will be inductive, and we will start with a point $(\theta_0, I_0) \in D_F(\mu)$. Then we have the following easy application of the ergodization theorem.

**Lemma 5.4.** Let $(\theta, I) \in D_F(\mu)$. There exists an integer

$$n \in \{ \lfloor \ln \mu^{-1} \rfloor, \ldots, 2\lfloor \ln \mu^{-1} \rfloor \}$$

such that $T^n(\theta, I) \in B_{K'}$.  

35
Proof. Recall that $T^n(\theta, I) \in B_{K'}$ if and only if $\theta + nI \in J_{\delta/2}$. Now by definition of $D_F(\mu)$,
\[ d(I, R_{\delta}) \geq \frac{1}{\ln \mu} \]
so the conclusion follows from Lemma 5.3 by choosing $m = \lfloor \ln \mu^{-1} \rfloor$.

8. As long as the orbit stays in $D_F(\mu)$, we can use the previous lemma. Then it will enter into the domain of slow drift, and this is where the ergodization theorem gives integers as large as $\mu^{-\frac{7}{2}} \ln \mu^{-1}$. However in the lemma below we will see how, after iterating a finite number of maps $f_{n_k}$, $n_k \in A$, with $\sum n_k$ of order $\mu^{-1} \ln \mu^{-1}$, one can actually cross through this domain of slow drift.

Lemma 5.5. Let $(\theta, I) \in D_S(\mu)$. There exist an index $j \in \mathbb{N}$ and integers $n_0, \ldots, n_j$ satisfying:

(i) $n_k \in A$, for $0 \leq k \leq j$;
(ii) $f_{n_k} \circ \cdots \circ f_{n_0}(\theta, I) \in B_{K'}$, for $0 \leq k \leq j$;
(iii) $f_{n_j} \circ \cdots \circ f_{n_0}(\theta, I) \notin D_S(\mu)$.

Moreover, if $j \in \mathbb{N}^*$, then

(iv) $\sum_{k=0}^{j} n_k \leq (j + 1) \ln \mu^{-1} + (2K')^{-1} \mu^{-1} \ln \mu^{-1}$.

Let us point out that in the proof of Proposition 5.1 below, the above lemma will always be used with $j \geq 1$ so item (iv) will be available.
Proof. There exists a unique \( p/q \in \mathbb{R} \) such that \( (\theta, I) \in D_S(\mu, p/q) \). By Lemma 5.3 (with \( m = \lfloor \ln \mu^{-1} \rfloor \)) and since

\[
d(I, \mathbb{R}_\delta) = |I - p/q|^{-1} \leq \mu^{-\frac{1}{2}} \ln \mu^{-1},
\]

we can find an integer \( n_0 \in A \) such that \( f_{n_0}(\theta, I) \in B_{K'} \). Now if \( f_{n_0}(\theta, I) \notin D_S(\mu) \), then we can take \( j = 0 \) in the statement and assertions \((i), (ii)\) and \((iii)\) are proven.

Otherwise, we construct inductively \((n_k)_{1 \leq k \leq j}\) where \( j \in \mathbb{N} \) is defined by

\[
j = \inf \{ k \in \mathbb{N}^* \mid f_{n_k} \circ \cdots \circ f_{n_0}(\theta, I) \notin D_S(\mu, p/q) \}.
\]

Since our orbit always stays in \( B_{K'} \), then \( j \) is obviously well-defined, and at each step we have used Lemma 5.3 so conditions \((i)\) and \((ii)\) are satisfied. Moreover by a previous remark we can also write

\[
j = \inf \{ k \in \mathbb{N}^* \mid f_{n_k} \circ \cdots \circ f_{n_0}(\theta, I) \notin D_S(\mu) \},
\]

so \( f_{n_j} \circ \cdots \circ f_{n_0}(\theta, I) \notin D_S(\mu) \), and condition \((iii)\) is also satisfied.

Finally, let us write \((\theta_0, I_0) = (\theta, I)\) and \((\theta_{k+1}, I_{k+1}) = f_{n_k} \circ \cdots \circ f_{n_0}(\theta_0, I_0), \) for \( 0 \leq k \leq j \). Since these points belong to \( B_{K'} \), we have \( I_{k+1} - I_k \geq 2 \mu K' \) for \( 0 \leq k \leq j \) and hence

\[
\sum_{k=0}^{j} n_k \leq \sum_{k=0}^{j} \ln \mu^{-1} + \sum_{k=0}^{j} \frac{1}{|I_k - p/q|} \leq (j + 1) \ln \mu^{-1} + (2K')^{-1} \mu^{-1} \sum_{k=0}^{j} \frac{I_{k+1} - I_k}{|I_k - p/q|}.
\]

The second term on the right-hand side is a Riemann sum, hence for \( \mu \) small enough it can be estimated by an integral, namely

\[
\sum_{k=0}^{j} \frac{I_{k+1} - I_k}{|I_k - p/q|} \leq 2 \int_{\frac{1}{\ln \mu^{-1}}}^{\frac{1}{\ln \mu^{-1}}} \frac{dI}{|I - p/q|} \leq 2 \int_{\frac{1}{\ln \mu^{-1}}}^{\frac{1}{\ln \mu^{-1}}} \frac{dI}{I - p/q} = 2 \int_{\frac{1}{\ln \mu^{-1}}}^{\frac{1}{\ln \mu^{-1}}} \frac{dI}{T},
\]

and as

\[
2 \int_{\frac{1}{\ln \mu^{-1}}}^{\frac{1}{\ln \mu^{-1}}} \frac{dI}{T} = 2 \ln \mu^{-\frac{1}{2}} \leq \ln \mu^{-1},
\]

this eventually gives

\[
\sum_{k=0}^{j} n_k \leq (j + 1) \ln \mu^{-1} + (2K')^{-1} \mu^{-1} \ln \mu^{-1}.
\]

This is exactly \((iv)\), and so this ends the proof.
9. Now that we have escaped the domain of slow drift, we are in the resonant domain. Here we cannot use any ergodization result. However our point belongs $B_K'$, and in the lemma below we will show that by iterating $j$ times a map $f_n$, with $n$ of order $\ln \mu^{-1}$, it can cross the resonant domain while staying in the larger domain $B_K$.

Lemma 5.6. Let $(\theta, I) \in D_R(\mu) \cap B_K'$. If $\mu$ is small enough, there exist integers
\[ n \in \{[\ln \mu^{-1}], \ldots, 2[\ln \mu^{-1}]\} \]
and $j \in \mathbb{N}^*$ such that:

(i) $f_n^k(\theta, I) \in D_R(\mu) \cap B_K$, for $0 \leq k \leq j - 1$;

(ii) $f_n^j(\theta, I) \notin D_R(\mu)$.

Note that it is possible that the point $f_n^j(\theta, I)$ does not belong to $B_K'$ either, but then we are back in the domain of slow drift and we can find an integer $n' \in A$ such that $T^{n'}(f_n^j(\theta, I)) \in B_K' \subseteq B_K$.

Proof. There exists a unique $p/q \in \mathbb{Q}$ such that $(\theta, I) \in D_R(\mu, p/q) \cap B_K'$.

We choose $\mu$ small enough so
\[ M\delta^{-1} \leq \ln \mu^{-1}, \]
and as $q \leq M\delta^{-1}$, we can find an integer $d$ such that $n = dq$ satisfies
\[ n \in \{[\ln \mu^{-1}], \ldots, 2[\ln \mu^{-1}]\}. \]

We will also add a further smallness condition on $\mu$ by requiring that
\[ \frac{2}{K \ln \mu^{-1}} < \frac{1}{2\pi} \max\{\arccos K - \arccos K', \arcsin K' - \arcsin K\}. \]

Let us define inductively
\[ f_n^k(\theta, I) = (\theta_k, I_k), \quad 0 \leq k \leq j, \]
where $j$ is defined as follows: $j = \inf\{j_1, j_2\}$ with
\[ j_1 = \inf\{k \in \mathbb{N}^* \mid (\theta_k, I_k) \notin D_R(\mu, p/q)\} \]
and
\[ j_2 = \left[\left(K \mu^{\frac{1}{2}} \ln \mu^{-1}\right)^{-1}\right] + 1. \]

It follows from the definition of $j$ that $(\theta_k, I_k) \in D_R(\mu)$, for $0 \leq k \leq j - 1$, so now let us show that $(\theta_k, I_k) \in B_K$. 38
In fact we will prove below by induction on $k$ that
\[
-\frac{\arccos K'}{2\pi} - \frac{kn\mu^\frac{1}{2}}{\ln \mu^{-1}} \leq \theta_k \leq -\frac{\arcsin K'}{2\pi} + \frac{kn\mu^\frac{1}{2}}{\ln \mu^{-1}}, \quad 0 \leq k \leq j - 1. \tag{18}
\]
Since $n \leq 2 \ln \mu^{-1}$ this implies
\[
-\frac{\arccos K'}{2\pi} - 2k\mu^\frac{1}{2} \leq \theta_k \leq -\frac{\arcsin K'}{2\pi} + 2k\mu^\frac{1}{2},
\]
and as $k \leq j$, then
\[
-\frac{\arccos K'}{2\pi} - \frac{2}{K\ln \mu^{-1}} \leq \theta_k \leq -\frac{\arcsin K'}{2\pi} + \frac{2}{K\ln \mu^{-1}},
\]
\[-\frac{\arccos K}{2\pi} \leq \theta_k \leq \frac{\arcsin K}{2\pi}.
\]
This means that $\cos 2\pi\theta_k \geq K$ and $\sin 2\pi\theta_k \leq -K$, and therefore this gives $(\theta_k, I_k) \in B_K$ for $0 \leq k \leq j - 1$.

So now let us go through the induction, that is let us prove (18). This is obviously true for $k = 0$, so let us assume it is satisfied for some $0 \leq k \leq j - 2$.

We have
\[
\theta_{k+1} = \theta_k + nr_k,
\]
but since $I_k \geq p/q - \mu^\frac{1}{2}/\ln \mu^{-1}$ and $n = dq$ then
\[
\theta_{k+1} \geq \theta_k + dp - n\mu^\frac{1}{2}/\ln \mu^{-1} = \theta_k - n\mu^\frac{1}{2}/\ln \mu^{-1} \quad [\mathbb{Z}],
\]
as $dp$ is an integer. Then, using the hypothesis of induction this gives
\[
\theta_{k+1} \geq -\frac{\arccos K'}{2\pi} - \frac{kn\mu^\frac{1}{2}}{\ln \mu^{-1}} - \frac{n\mu^\frac{1}{2}}{\ln \mu^{-1}} \geq -\frac{\arccos K'}{2\pi} - \frac{(k+1)n\mu^\frac{1}{2}}{\ln \mu^{-1}}.
\]
Similarly using the fact that $I_k \leq p/q + \mu^\frac{1}{2}/\ln \mu^{-1}$ one obtains
\[
\theta_{k+1} \leq -\frac{\arcsin K'}{2\pi} + \frac{(k+1)n\mu^\frac{1}{2}}{\ln \mu^{-1}},
\]
therefore (18) is proven and so is (i).

Now assume that $j = j_2$, then since $(\theta_k, I_k) \in B_K$ for $0 \leq k \leq j - 1$,
\[
I_{j_2} \geq I_0 + 2j_2K\mu > I_0 + 2\mu^\frac{1}{2}/\ln \mu^{-1} > p/q + \mu^\frac{1}{2}/\ln \mu^{-1},
\]
which means that $j_2 \geq j_1$. This is absurd, therefore $j = j_1$, so $(\theta_j, I_j) \notin D_R(\mu, p/q)$ and this means that $(\theta_j, I_j) \notin D_R(\mu)$. This proves (ii).
10. Now we can finally conclude the proof of Proposition 5.1.

**Proof of Proposition 5.1.** We choose any point \((\theta_0, I_0) \in D_F(\mu)\). Applying successively Lemma 5.4, Lemma 5.5, Lemma 5.6 and Lemma 5.5 once again, in this precise order, one obtains a positive sequence \((n_k, \theta_k, I_k)_{k \in \mathbb{N}} \in A \times A\) which gives a positive orbit for our polysystem, that is

\[
 f_{n_k}(\theta_k, I_k) = (\theta_{k+1}, I_{k+1}), \quad k \in \mathbb{N}.
\]

Note that since we have started in \(D_F(\mu)\), at each time Lemma 5.5 is applied with an integer \(j \geq 1\).

Now by construction, \(T^{n_k}(\theta_k, I_k) \in B_K\) for any \(k \in \mathbb{N}\), and since

\[
 (\theta_{k+1}, I_{k+1}) = V \circ T^{n_k}(\theta_k, I_k),
\]

by definition of \(B_K\) one obtains

\[
 I_{k+1} \geq I_k + 2\mu K.
\]

This clearly shows that

\[
 \lim_{k \to +\infty} I_k = +\infty,
\]

which proves the first part of the statement.

Then as

\[
 I_k - I_0 \geq 2k\mu K,
\]

if we set \(N = [(\mu K)^{-1}] + 1\),

\[
 I_N - I_0 \geq 2.
\]

It remains to estimate the sum of integers

\[
 S = \sum_{k=0}^N n_k.
\]

To do that, we will write

\[
 \sigma_1 = \{k \in [0, N] \mid (\theta_k, I_k) \in D_F(\mu) \cup D_R(\mu)\}
\]

and

\[
 \sigma_2 = \{k \in [0, N] \mid (\theta_k, I_k) \in D_S(\mu)\},
\]

so

\[
 S = \sum_{k \in \sigma_1} n_k + \sum_{k \in \sigma_2} n_k = S_1 + S_2.
\]

For each \(k \in \sigma_1\), we know that \(n_k \leq 2\ln \mu^{-1}\) and hence

\[
 S_1 \leq 2 \sum_{k \in \sigma_1} \ln \mu^{-1} \leq 2N \ln \mu^{-1} \leq 4K^{-1} \mu^{-1} \ln \mu^{-1}.
\]
To estimate $S_2$, first observe that the set $\mathcal{R}_\delta$ is discrete, so its intersection with the compact set $\{I_0 \leq I \leq I_N\}$ is finite. But the latter set is included in $\{I_0 \leq I \leq I_0 + 3\}$, so the constant

$$M = \text{card}\{p/q \in \mathcal{R}_\delta \mid I_0 \leq p/q \leq I_0 + 3\},$$

is independent of $\mu$. Then setting

$$\sigma_2(p/q) = \{k \in [0, N] \mid (\theta_k, I_k) \in D_S(\mu, p/q)\},$$

we obtain

$$S_2 = \sum_{k \in \sigma_2} n_k \leq M \sum_{k \in \sigma_2(p/q)} n_k.$$

Now each $\sigma_2(p/q)$ is not reduced to a point, so by Lemma 5.5

$$\sum_{k \in \sigma_2(p/q)} n_k \leq N \ln \mu^{-1} + (2K')^{-1} \mu^{-1} \ln \mu^{-1} \leq 2K^{-1} \mu^{-1} \ln \mu^{-1} + (2K')^{-1} \mu^{-1} \ln \mu^{-1} \leq (2K^{-1} + (2K')^{-1}) \mu^{-1} \ln \mu^{-1}.$$

This finally gives

$$S \leq (4K^{-1}) \mu^{-1} \ln \mu^{-1} + M (2K^{-1} + (2K')^{-1}) \mu^{-1} \ln \mu^{-1} \leq C \mu^{-1} \ln \mu^{-1},$$

with $C = 2 \sup \{4K^{-1}, M (2K^{-1} + (2K')^{-1})\}$, and this proves the proposition. \qed

6 Proof of Theorem 2.2

11. In this section we will prove Theorem 2.2, and in view of Proposition 4.1, this will follow easily from the next result.

**Proposition 6.1.** There exists a positive constant $C$ such that for $\mu$ small enough, there exists an orbit $(x^k(\bar{n}), x_k)_{k \in \mathbb{Z}} \in \Sigma_A \times \mathcal{A}$ for the skew-product $[[F_{\bar{n}}]]_{n \in A}$ defined by

$$F_{\bar{n}}(\theta, I) = (\theta + n_0 I, I + 2\mu f_1(\phi_{\bar{n}}(\theta, I)) \cos 2\pi(\theta + n_0 I)), \quad \bar{n} \in \Sigma_A, (\theta, I) \in \mathcal{A},$$

such that

$$\lim_{k \to \pm \infty} I'_k = \pm \infty$$

and the estimates

$$|I'_N - I'_0| \geq 1, \quad \sum_{k=0}^{N} n_k \leq C \mu^{-1} \ln \mu^{-1},$$

hold true.
The strategy will be to consider the orbit constructed in Proposition 5.1 as a pseudo-orbit for the skew-product. So in order to find a true orbit nearby, we shall need some hyperbolicity and this will be described below.

12. Recall that

\[ B_K = \{ (\theta, I) \in \mathbb{A} \mid \cos 2\pi \theta \geq K, \sin 2\pi \theta \leq -K \} \]

The condition \( \sin 2\pi \theta \leq -K \) will be used here to prove the following lemma.

**Lemma 6.2.** Let \( x \in \mathbb{A} \) such that \( T^n(x) \in B_K \), for \( n \in \mathbb{A} \), and \( \tilde{x} \in \mathbb{R}^2 \) a lift of \( x \). Then the eigenvalues \( \lambda_\pm \) of \( df_n(\tilde{x}) \) are real and for \( \mu \) small enough, they satisfy

\[ \lambda_+ > 1 + 2\sqrt{n\pi \mu K} > 1, \quad \lambda_- < 1 - 2\sqrt{n\pi \mu K} + 4n\pi \mu K < 1. \]

Moreover, if \( e_\pm \in \mathbb{R}^2 \) are eigenvectors associated to \( \lambda_\pm \), and for \( v \in \mathbb{R}^2 \),

\[ v = v_+ e_+ + v_- e_-, \]

then

\[ \frac{1}{2} |v| \leq \sup \{|v_+|, |v_-|\} \leq a|v|, \]

with \( a = O \left( \mu^{-\frac{3}{2}} (\ln \mu)^{-\frac{1}{2}} \right) \), and where \( |.| \) is the supremum norm on \( \mathbb{R}^2 \).

**Proof.** Let us write \( x = (\theta, I) \in \mathbb{A} \) and \( \tilde{x} = (\tilde{\theta}, I) \in \mathbb{R}^2 \). If we define

\[ s = -2\pi \mu \sin 2\pi (\theta + nI), \]

then \( s > 2\pi \mu K \) since \( T^n(\theta, I) = (\theta + nI, I) \in B_K \). As

\[ f_n(\theta, I) = (\theta + nI, I + 2\mu \cos 2\pi (\theta + nI)), \quad n \in \mathbb{A}, \quad (\theta, I) \in \mathbb{A}, \]

we have

\[ df_n(\tilde{\theta}, I) = \begin{pmatrix} 1 & n \\ 2s & 1 + 2ns \end{pmatrix} \in M_2(\mathbb{R}). \]

Then the eigenvalues are real and given by

\[ \lambda_\pm = 1 + ns \pm \sqrt{ns(2 + ns)}. \]

Therefore we easily obtain

\[ \lambda_+ > 1 + \sqrt{2ns} > 1 + 2\sqrt{n\pi \mu K}, \]

and then using the equality \( \lambda_+ \lambda_- = 1 \), for \( \mu \) small enough one finds

\[ \lambda_- < 1 - 2\sqrt{n\pi \mu K} + 4n\pi \mu K. \]

This proves the first part of the statement.
Then if we define
\[ \alpha_\pm = n^{-1}(\lambda_\pm - 1) = s \pm \sqrt{n^{-1}s(2+ns)}, \]
one can easily check that the vectors
\[ e_\pm = \left( \frac{1}{\alpha_\pm} \right) \in \mathbb{R}^2, \]
are eigenvectors associated to \( \lambda_\pm \). Since \(|e_+| = |e_-| = 1\), for \( v \in \mathbb{R}^2 \) if we write
\[ ||v|| = \sup\{|v_+|, |v_-|\}, \quad v = v_+ e_+ + v_- e_- , \]
it is trivial that
\[ |v| \leq |v_+| + |v_-| \leq 2||v||. \]
As for the other inequality, note that
\[ |v| = \sup\{|v_+ + v_-|, |\alpha_+ v_+ + \alpha_- v_-|\}. \]
If we define
\[ r_\pm = \frac{\alpha_+ \alpha_- - 1 \pm (\alpha_+ - \alpha_-)}{1 - \alpha_+^2}, \]
then one can check that
\[ |\alpha_+ v_+ + \alpha_- v_-| \leq |v_+ + v_-| \]
if and only if \( v_- = 0 \), or \((v_-)^{-1}v_+ \leq r_-\), or \((v_-)^{-1}v_+ \geq r_+\). Hence
\[ |v| = \begin{cases} |v_+ + v_-| & \text{if } v_- = 0, \text{ or } (v_-)^{-1}v_+ \leq r_-\text{, or } (v_-)^{-1}v_+ \geq r_+\text{,} \\ |\alpha_+ v_+ + \alpha_- v_-| & \text{if } (v_-)^{-1}v_+ \geq r_-\text{, or } (v_-)^{-1}v_+ \leq r_+. \end{cases} \]
If we study all the cases, one finds
\[ ||v|| \leq \sup \{1, |1+r_+|^{-1}, |1+r_-|^{-1}, |1+r_+|^{-1}, |1+r_-|^{-1}\} |v|. \]
Then we can estimate
\[ r_\pm = -1 \pm \sqrt{2n^{-1}s} + o\left(\sqrt{n^{-1}\mu}\right), \]
and using the fact that \( n \in A \), one finds
\[ ||v|| \leq a|v|, \]
with \( a = O\left(\mu^{-\frac{3}{4}}(\ln\mu^{-1})^{\frac{1}{2}}\right). \)
13. Let us now describe an abstract fixed point theorem that will be used to find an orbit close to our pseudo-orbit.

Consider a Banach space \((E, |.|)\) and \(T : E \to E\) a continuous linear map. Recall that the spectrum \(\text{Sp}(T)\) of \(T\) is the set of complex numbers \(\lambda\) such that \(T_C - \lambda I_{d_C}\) is not an automorphism of \(E_C\), where \(E_C\) and \(T_C\) are the complexifications of \(E\) and \(T\).

Given two real numbers \(\kappa_s, \kappa_u\) satisfying \(0 < \kappa_s < 1 < \kappa_u\), we say that \(T\) is \((\kappa_s, \kappa_u)\)-hyperbolic if

\[
\text{Sp}(T) \cap \{\kappa_s \leq |z| \leq \kappa_u\} = \emptyset.
\]

In such a case, there exists a \(T\)-invariant decomposition

\[
E = E_s \oplus E_u, \quad T(E_s) \subseteq E_s, \quad T(E_u) \subseteq E_u,
\]

and a constant \(c > 0\) such that

\[
|(T|_{E_s})^n| \leq cn_s^n, \quad |(T|_{E_u})^{-n}| \leq cn_u^n,
\]

for any \(n \in \mathbb{N}\), \(\eta_s < \kappa_s, \eta_u > \kappa_u\) and where \(|.|\) is the induced norm on linear operators.

In fact, one can always find a norm \(\|.|\) on \(E\) which is adapted to \(T\) in the following sense: \(\|.|\) is equivalent to \(|.|\) and satisfies

\[1.\] \(|x_s + x_u| = \sup\{\|x_s\|, \|x_u\|\}, \quad x_s \in E_s, \quad x_u \in E_u;\]

\[2.\] \(|T|_{E_s}| \leq \kappa_s, \quad \|(T|_{E_u})^{-1}\| \leq \kappa_u^{-1}.
\]

The following theorem is proved in [Yoc95], section 2.1.

**Theorem 6.3.** Let \(T : E \to E\) be \((\kappa_s, \kappa_u)\)-hyperbolic and let \(U : E \to E\) be a Lipschitz map such that

\[\varepsilon = \text{Lip}(U - T) < \varepsilon_0 = \inf\{1 - \kappa_s, 1 - \kappa_u^{-1}\}.
\]

Then \(U\) has a unique fixed point \(p \in E\), and if \(\|.|\) is a norm adapted to \(T\),

\[\|p\| < (\varepsilon_0 - \varepsilon)^{-1}\|U(0)\|.
\]

14. Now we can prove Proposition 6.1.

**Proof of Proposition 6.1.** Consider the orbit \((\sigma^k(\bar{n}), \theta_k, I_k)_{k \in \mathbb{Z}} \in \Sigma_A \times \mathbb{A}\) given by Proposition 5.1. Then the proof is an immediate consequence of the following claim: there exist a sequence \((\theta'_k, I'_k)_{k \in \mathbb{Z}} \in \mathbb{A}\) such that \((\sigma^k(\bar{n}), \theta'_k, I'_k)_{k \in \mathbb{Z}} \in \Sigma_A \times \mathbb{A}\) is an orbit for the skew product and

\[
|I_k - I'_k| \leq \mu^2, \quad k \in \mathbb{Z}.
\]
So let us construct this orbit.

Let \( x_k = (\theta_k, I_k)_{k \in \mathbb{Z}} \in \mathcal{A} \), and \( \tilde{x}_k = (\tilde{\theta}_k, I_k)_{k \in \mathbb{Z}} \) one of its lift in \( \mathbb{R}^2 \). First we define a linear map 

\[
T : (\mathbb{R}^2)^\mathbb{Z} \rightarrow (\mathbb{R}^2)^\mathbb{Z}
\]

by

\[
(T(v))_k = df_{n_{k-1}}(\tilde{x}_{k-1}).v_{k-1}, \quad k \in \mathbb{Z}.
\]

Using the supremum norm on \( \mathbb{R}^2 \) let us define

\[
E = \{ v = (v_k)_{k \in \mathbb{Z}} \in (\mathbb{R}^2)^\mathbb{Z} \mid \sup_{k \in \mathbb{Z}} |v_k| < \infty \}.
\]

Then \( E \) is obviously a Banach space with the norm

\[
|v| = \sup_{k \in \mathbb{Z}} |v_k|, \quad v \in E.
\]

Now recall that for any \( \tilde{x} = (\tilde{\theta}, I) \in \mathbb{R}^2 \), if \( s = -2\pi \mu \sin 2\pi (\tilde{\theta} + nI) \), then

\[
df_n(\tilde{x}) = \begin{pmatrix} \frac{1}{2s} & \frac{n}{1 + 2ns} \end{pmatrix} \in M_2(\mathbb{R}).
\]

Since the norm on \( M_2(\mathbb{R}) \) induced by the supremum norm on \( \mathbb{R}^2 \) is given by the maximum of the sums of the absolute values of the elements in each row, we obtain

\[
\sup_{k \in \mathbb{Z}} \{|df_{n_k}|_{C^0(\mathbb{R}^2)}\} = \sup_{k \in \mathbb{Z}} \{1 + n_k\} = 1 + \mu^{-\frac{1}{2}} \ln \mu^{-1} < \infty,
\]

then \( T(F) \subseteq F \) and therefore \( T \) is continuous. Moreover, by construction the sequence \( (x_k)_{k \in \mathbb{Z}} \) satisfies \( T^{n_k}(x_k) \in B_K \), so we can apply Lemma 6.2 and each map

\[
T_{k-1} : v_{k-1} \mapsto df_{n_{k-1}}(\tilde{x}_{k-1}).v_{k-1}
\]

is a \((\kappa_s^{k-1}, \kappa_u^{k-1})\)-hyperbolic linear map of \( \mathbb{R}^2 \), with

\[
\kappa_s^k = 1 - 2\sqrt{n_k \pi \mu K} + 4n_k \pi \mu K < 1, \quad k \in \mathbb{Z},
\]

and

\[
\kappa_u^k = 1 + 2\sqrt{n_k \pi \mu K} > 1, \quad k \in \mathbb{Z}.
\]

This implies that \( T \) is \((\kappa_s, \kappa_u)\)-hyperbolic, with

\[
\kappa_s = \sup_{k \in \mathbb{Z}} \{\kappa_s^k\}, \quad \kappa_u = \inf_{k \in \mathbb{Z}} \{\kappa_u^k\}.
\]

Since \( n_k \in A \), one finds

\[
\kappa_s = 1 - 2\sqrt{\pi K (\ln \mu^{-1})^{\frac{3}{2}} \mu^{\frac{3}{2}} + 4\pi K \ln \mu^{-1} \mu^{\frac{1}{2}}} < 1,
\]

and

\[
\kappa_u = 1 + 2\sqrt{\pi K (\ln \mu^{-1})^{\frac{3}{2}} \mu^{\frac{3}{2}} + 4\pi K \ln \mu^{-1} \mu^{\frac{1}{2}}} > 1.
\]
and
\[ \kappa_u = 1 + 2 \sqrt{\pi K (\ln \mu^{-1})^{3/2}} \mu^{3/2} > 1. \]

Let us set
\[ \varepsilon_0 = \inf\{1 - \kappa_s, 1 - \kappa_u^{-1}\} = O \left( \mu^{3/4} (\ln \mu)^{3/2} \right). \]

Now if
\[ \|v\| = \sup_{k \in \mathbb{Z}} \|v_k\|, \]
where \( \|\cdot\|_k \) is a norm in \( \mathbb{R}^2 \) adapted to \( T_k, k \in \mathbb{Z} \), then \( \|\cdot\| \) is adapted to \( T \), and from Lemma 6.2
\[
\frac{1}{2} |v| \leq \|v\| \leq a |v|, \quad v \in E,
\]
with \( a = O \left( \mu^{-4/3} (\ln \mu)^{-1} \right) \).

Next we define
\[ U : (\mathbb{R}^2)^\mathbb{Z} \rightarrow (\mathbb{R}^2)^\mathbb{Z} \]
by
\[ (U(v))_k = F_{\sigma^{-1}(\tilde{n})}(\tilde{x}_{k-1} + v_{k-1}) - \tilde{x}_k, \quad k \in \mathbb{Z}. \]

Note that if \( v \) is a fixed point of \( U \), then
\[ F_{\sigma^{-1}(\tilde{n})}(\tilde{x}_{k-1} + v_{k-1}) = \tilde{x}_k + v_k, \]
so \( (\sigma^k(\tilde{n}), x'_k)_{k \in \mathbb{Z}} \in \Sigma_A \times \mathbb{A} \), where \( x'_k \in \mathbb{A} \) is the projection onto \( \mathbb{A} \) of \( \tilde{x}'_k = \tilde{x}_k + v_k \in \mathbb{R}^2 \), is an orbit for the skew-product. To prove that \( U \) has a fixed point, we will use Theorem 6.3 and for that we need to estimate the Lipschitz constant of \( \Delta = U - T \) with respect to the adapted norm \( \|\cdot\| \).

First if \( v, v' \in E \) satisfy \( |v| \leq \mu^2, |v| \leq \mu^2 \), then from Taylor formula one can compute
\[ |\Delta(v) - \Delta(v')| \leq (L + 2M\mu^2)|v - v'|, \]
with
\[ L = \sup_{k \in \mathbb{Z}} \left\{ \text{Lip} \left( f_{n_k} - F_{\sigma^{k}(\tilde{n})} \right) \right\}, \quad M = \sup_{k \in \mathbb{Z}} \left\{ |d^2 f_{n_k}|_{C^0(\mathbb{R}^2)} \right\}. \]

Using the estimates (4) obtained in Proposition 4.1, one finds
\[ L = O \left( \mu^{4\pi - 3/2} \ln \mu^{-1} \right), \]
whereas
\[ M = O \left( (\ln \mu^{-1})^2 \right) \]
is obvious from the definition of the maps \( (f_n)_{n \in \mathbb{A}} \).
However, far away from 0, the maps $U$ and $T$ are not necessarily close, so for $v \in \mathbb{R}^2$ we define $\overrightarrow{U}(v)$ by

$$\overrightarrow{U}(v) = \begin{cases} U(v) & \text{if } |v| \leq \mu^2, \\ U \left( \mu^2 |v|^{-1} v \right) + (1 - \mu^2 |v|^{-1}) T(v) & \text{if } |v| \geq \mu^2. \end{cases}$$

Then setting $\Delta = \overrightarrow{U} - T$, for any $v, v' \in E$ one easily obtains

$$|\Delta(v) - \Delta(v')| \leq 2(L + 2\mu^2)|v - v'|,$$

and this gives

$$\|\Delta(v) - \Delta(v')\| \leq 4a(L + 2\mu^2)\|v - v'\|,$$

In particular, this shows $\overrightarrow{U}(F) \subseteq F$, and the Lipschitz constant of $\overrightarrow{U} - T$ with respect to the adapted norm $\| \cdot \|$ is

$$\varepsilon = \text{Lip}(\overrightarrow{U} - T) \leq 4a(L + 2\mu^2) = O \left( \mu^2 \left( \ln \mu \right)^{\frac{5}{2}} \right).$$

As

$$\varepsilon = O \left( \mu^2 \left( \ln \mu \right)^{\frac{5}{2}} \right) < \varepsilon_0 = O \left( \mu^2 \left( \ln \mu \right)^{\frac{5}{2}} \right)$$

we can finally apply Theorem 6.3: $\overrightarrow{U}$ has a unique fixed point $v' \in E$ such that

$$\|v'\| \leq (\varepsilon_0 - \varepsilon)^{-1} \|\overrightarrow{U}(0)\|,$$

and hence

$$|v'| \leq 2a(\varepsilon_0 - \varepsilon)^{-1} \|\overrightarrow{U}(0)\|.$$

Note that

$$|\overrightarrow{U}(0)| = \sup_{k \in \mathbb{Z}} \{|F_{\sigma^k(\bar{n})}(\ddot{x}_k) - \ddot{x}_{k+1}|\} = \sup_{k \in \mathbb{Z}} \{|F_{\sigma^k(\bar{n})}(\ddot{x}_k) - f_{n_k}(\ddot{x}_k)|\}.$$

Using the estimates (4) in Proposition 4.1,

$$\sup_{k \in \mathbb{Z}} |F_{\sigma^k(\bar{n})}(\ddot{x}_k) - f_{n_k}(\ddot{x}_k)| = O(\mu^{4\gamma - 1})$$

and as $2a(\varepsilon_0 - \varepsilon)^{-1} = O(\mu^{-5/4})$, we have in particular

$$|v'| \leq \mu^2. \quad (19)$$

By definition of $\overrightarrow{U}$, this shows that $v'$ is in fact a fixed point of $U$, and therefore the sequence $(\sigma^k(\bar{n}), x'_k)_{k \in \mathbb{Z}} \in \Sigma_A \times A$, where $x'_k$ is the projection onto $A$ of

$$\ddot{x}'_k = \ddot{x}_k + v'_k = (\ddot{\theta}'_k, I'_k), \quad k \in \mathbb{Z},$$

is an orbit for the skew-product. If we set $x'_k = (\theta'_k, I'_k) \in A$, then from (19) we obtain

$$|I_k - I'_k| \leq \mu^2, \quad k \in \mathbb{Z}.$$

This concludes the proof. □
Now we can eventually prove Theorem 2.2.

Proof of Theorem 2.2. Let \((\sigma_k(n), \theta'_k, I'_k)_{k \in \mathbb{Z}} \in \Sigma_A \times A\) be the orbit for the skew-product obtained in Proposition 6.1. Then, by Proposition 4.1,

\[ \Upsilon_\mu(n_k, \theta'_k, I'_k) = (\tau_k, e_k, \theta'_k, I'_k) \in E \times A, \quad k \in \mathbb{Z}, \]

and in the original coordinates

\[ \Psi^{-1}(\tau_k, e_k, \theta'_k, I'_k) = (\theta^k_1, I^k_1, \theta^k_2, I^k_2) \in A^2, \quad k \in \mathbb{Z}, \]

is an orbit for the transversal map \(\tilde{F}_\mu\). By definition of the latter one has

\[ \mathcal{F}^*_\mu(\theta^k_1, I^k_1, \theta^k_2, I^k_2) = (\theta^{k+1}_1, I^{k+1}_1, \theta^{k+1}_2, I^{k+1}_2), \quad k \in \mathbb{Z}, \]

and since \(I'_k = I^2_k\) for \(k \in \mathbb{Z}\), the theorem follows. \(\square\)

A Time-energy coordinates for the pendulum

In this appendix, we recall some elementary facts about the time-energy coordinates for the simple pendulum. We refer to [MS04], [Mar05] and [LM05] for more details.

Consider the simple pendulum defined by the Hamiltonian

\[ P(\theta, I) = \frac{1}{2} I^2 + \cos 2\pi \theta, \]

and the open domain

\[ \mathcal{E} = \{ (\theta, I) \in A \mid 0 < \theta < 1, \ I > 0 \}. \]

We define the energy function

\[ e(\theta, I) = P(\theta, I) - 1 = \frac{1}{2} I^2 + \cos 2\pi \theta - 1 \]

and, using \(\{\theta = 1/2\}\) as a reference section, the time function

\[ \tau(\theta, I) = \int_{1/2}^{\theta} \frac{d\theta}{\sqrt{2(e(\theta, I) - V(\theta))}} \]

where \(V(\theta) = \cos 2\pi \theta - 1\). For positive energy \(e > 0\), the period of motion is given by

\[ T(e) = \int_{1/2}^{1/2} \frac{d\theta}{\sqrt{2(e - V(\theta))}} \]

and it is a decreasing function. Moreover, we have the equivalent

\[ T(e) \sim 0 - (2\pi)^{-1} \ln e. \quad (20) \]

48
For negative energy \( e < 0 \), if \( \theta(e) \) is such that \( \cos 2\pi \theta(e) = 1 + e \), then

\[
T(e) = \int_{-\theta(e)}^{\theta(e)} \frac{d\theta}{\sqrt{2(e - V(\theta))}}.
\]

Therefore if we define

\[
\mathcal{E}_* = \{ (\tau, e) \in \mathbb{R}^2 \mid e > -2, |\tau| < \frac{1}{2} T(e) \}
\]

we have a diffeomorphism

\[
\Psi : \mathcal{E} \rightarrow \mathcal{E}_*
\]

\[
(\theta, I) \mapsto (\tau, e)
\]

and one can check that it is symplectic. Moreover, in those coordinates \((\tau, e)\) the flow of the pendulum is straightened out, that is

\[
\Phi^{tP}(\tau, e) = (\tau + t, e).
\]

for \( t \) small enough.

17. To conclude, for the proof of Proposition 4.1 in section 4.2 we shall need some estimates on the energy and the period of the periodic orbits (of positive energy) for the pendulum. For \( n \in \mathbb{N}^* \), we let \( e_n \) be the energy of the \( n \)-periodic orbit for \( \Phi^P \), that is \( T(e_n) = n \), we have

\[
e_n \sim_{+\infty} \exp(-2\pi n).
\]

Then if we define \( T'_n = T'(e_n) \), \( T''_n = T''(e_n) \) where \( T' \) and \( T'' \) are the first and second derivatives of the period function, we can deduce from (20) and (21) that

\[
T'_n \sim_{+\infty} -(2\pi)^{-1} \exp(2\pi n)
\]

and

\[
T''_n \sim_{+\infty} 2\pi (T'_n)^2 = (2\pi)^{-1} \exp(4\pi n).
\]

Acknowledgments.

The authors are thankful to their advisor Jean-Pierre Marco for his guidance, for many helpful discussions, comments and corrections. The first author thanks the University of Warwick, the Marie Curie training network “Conformal Structures and Dynamics” and the Instituto Nacional de Matemática Pura e Aplicada (IMPA) for their support.

References

[AKN06] V.I. Arnold, V.V. Kozlov, and A.I. Neishtadt, Mathematical aspects of classical and celestial mechanics, [Dynamical Systems III], Transl. from the Russian original by E. Khukhro, Third edition, Encyclopedia of Mathematical Sciences, 3 ed., Springer-Verlag, Berlin, 2006.
[Arn63] V.I. Arnol’d, *Proof of a theorem of A.N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations*, Russ. Math. Surv. 18 (1963), no. 5, 9–36.

[Arn64] V.I. Arnold, *Instability of dynamical systems with several degrees of freedom*, Sov. Math. Doklady 5 (1964), 581–585.

[BBB03] M. Berti, L. Biasco, and P. Bolle, *Drift in phase space: a new variational mechanism with optimal diffusion time*, J. Math. Pures Appl. (9) 82 (2003), no. 6, 613–664.

[Ber96] P. Bernard, *Perturbation of a partially hyperbolic hamiltonian system (Perturbation d’un hamiltonien partiellement hyperbolique)*, C. R. Math. Acad. Sci. Paris 323 (1996), no. 2, 189–194.

[Ber08] ——, *The dynamics of pseudographs in convex Hamiltonian systems*, no. 3, 615–669.

[Bes96] U. Bessi, *An approach to Arnold’s diffusion through the calculus of variations*, Nonlinear Anal., Theory Methods Appl. 26 (1996), no. 6, 1115–1135.

[BM11] A. Bounemoura and J.-P. Marco, *Improved exponential stability for quasi-convex Hamiltonians*, Nonlinearity 24 (2011), no. 1, 97–112.

[Bou10a] A. Bounemoura, *Effective stability for Gevrey and finitely differentiable prevalent Hamiltonians*, Submitted (2010).

[Bou10b] ——, *Nekhoroshev theory for finitely differentiable quasi-convex Hamiltonians*, Journal of Differential Equations 249 (2010), no. 11, 2905–2920.

[CG94] L. Chierchia and G. Gallavotti, *Drift and diffusion in phase space*, Ann. Inst. Henri Poincaré, Phys. Théor. 60 (1994), no. 1, 1–144.

[CG03] J. Cresson and C. Guillet, *Periodic orbits and Arnold diffusion*, Discrete Contin. Dyn. Syst. 9 (2003), no. 2, 451–470 (English).

[Cha04] M. Chaperon, *Stable manifolds and the Perron-Irwin method*, Erg. Th. Dyn. Sys. 24 (2004), no. 5, 1359–1394.

[Cha08] ——, *The lipschitzian core of some invariant manifold theorems*, Erg. Th. Dyn. Sys. 28 (2008), no. 5, 1419–1441.

[CY04] C.-Q. Cheng and J. Yan, *Existence of diffusion orbits in a priori unstable Hamiltonian systems*, J. Differ. Geom. 67 (2004), no. 3, 457–517.
[CY09] ——, Arnold diffusion in Hamiltonian systems: a priori un-
stable case, J. Differ. Geom. 82 (2009), no. 2, 229–277.

[DdILS06] A. Delshams, R. de la Llave, and T.M. Seara, A geometric
mechanism for diffusion in Hamiltonian systems overcoming in
the large gap problem: Heuristics and rigorous verification on a
model, Mem. Am. Math. Soc. 844 (2006), 141 p. (English).

[Her98] M. Herman, Some open problems in dynamical systems, Doc.
Math., J. DMV, Extra Vol. ICM Berlin 1998, vol. II, 1998,
pp. 797–808.

[HPS97] M.W. Hirsch, C.C. Pugh, and M. Shub, Invariant Manifolds,
Springer, Berlin, 1997.

[KLS10] V. Kaloshin, M. Levi, and M. Saprykina, An example of nearly
integrable Hamiltonian system with a trajectory dense in a set of
maximal Hausdorff dimension, Preprint (2010).

[Kol54] A.N. Kolmogorov, On the preservation of conditionally periodic
motions for a small change in Hamilton’s function, Dokl. Akad.
Nauk. SSSR 98 (1954), 527–530.

[KZZ09] V. Kaloshin, K. Zhang, and Y. Zheng, Almost dense orbit on
energy surface, Proceedings of the XVI-th ICMP, Prague, 314-
322, 2009.

[LM05] P. Lochak and J.P. Marco, Diffusion times and stability exponents
for nearly integrable analytic systems, Central European Journal
of Mathematics 3 (2005), no. 3, 342–397.

[LMS03] P. Lochak, J.-P. Marco, and D. Sauzin, On the splitting of invari-
ant manifolds in multidimensional near-integrable Hamiltonian
systems, Mem. Am. Math. Soc. 775 (2003), 145 p.

[Loc99] Pierre Lochak, Arnold diffusion ; a compendium of remarks and
questions., Simó, Carles (ed.), Hamiltonian systems with three
or more degrees of freedom. Proceedings of the NATO Advanced
Study Institute, 1995. Dordrecht: Kluwer Academic Publishers.,
1999.

[Mar05] J.-P. Marco, Uniform lower bounds of the splitting for analytic
symplectic systems, preprint (2005).

[Mar08] ——, Models for skew-products and polysystems, C. R. Math.
Acad. Sci. Paris (2008), no. 3-4, 203–208.

[Mar10a] ——, Arnold diffusion in a priori stable systems on \&\&^3, In
preparation (2010).
[Mar10b] ______, Generic properties of classical systems on the torus $\mathbb{T}^2$, Preprint (2010).
[Mar10c] ______, Nets of hyperbolic annuli in generic nearly integrable systems on $\mathbb{A}^3$, Preprint (2010).
[Mar10d] ______, Skew-products and polsystems in the neighborhood of hyperbolic annuli, In preparation (2010).
[Mat04] J. Mather, Arnold diffusion I: Announcement of results, J. of Math. Sciences 124 (2004), no. 5, 5275–5289.
[Moe02] R. Moeckel, Generic drift on Cantor sets of annuli, Chenciner, Alain (ed.) et al., Celestial mechanics. Dedicated to Donald Saari for his 60th birthday. Proceedings of an international conference, Northwestern Univ., Evanston, IL, USA, December 15–19, 1999. Providence, RI: AMS, American Mathematical Society. Contemp. Math. 292, 163-171, 2002.
[Mos62] J. Moser, On Invariant curves of Area-Preserving Mappings of an Annulus, Nachr. Akad. Wiss. Göttingen II (1962), 1–20.
[Mos73] ______, Stable and random motions in dynamical systems. With special emphasis on celestial mechanics. Hermann Weyl Lectures. The Institute for Advanced Study, Annals of Mathematics Studies. No.77. Princeton, N. J.: Princeton University Press and University of Tokyo Press. VIII, 199 pp., 1973.
[MS02] J.-P. Marco and D. Sauzin, Stability and instability for Gevrey quasi-convex near-integrable Hamiltonian systems, Publ. Math. Inst. Hautes Études Sci. 96 (2002), 199–275.
[MS04] ______, Wandering domains and random walks in Gevrey near-integrable systems, Erg. Th. Dyn. Sys. 5 (2004), 1619–1666.
[Nek77] N.N. Nekhoroshev, An exponential estimate of the time of stability of nearly integrable Hamiltonian systems, Russian Math. Surveys 32 (1977), 1–65.
[Nek79] ______, An exponential estimate of the time of stability of nearly integrable Hamiltonian systems II, Trudy Sem. Petrovs 5 (1979), 5–50.
[Tre04] D. Treschev, Evolution of slow variables in a priori unstable Hamiltonian systems, Nonlinearity 17 (2004), no. 5, 1803–1841.
[Yoc95] J.-C. Yoccoz, *Introduction to hyperbolic dynamics*, Branner, Bodil (ed.) et al., Real and complex dynamical systems. Proceedings of the NATO Advanced Study Institute held in Hillerød, Denmark, June 20-July 2, 1993. Dordrecht: Kluwer Academic Publishers. NATO ASI Ser., Ser. C, Math. Phys. Sci. 464, 265-291, 1995.

[Zha09] Ke Zhang, *Speed of Arnold diffusion for analytic Hamiltonian systems*, Preprint (2009).