Available online at www.sciencedirect.com
SCIENCE DIRECT

J. Differential Equations 187 (2003) 269–292

Journal of Differential Equations

http://www.elsevier.com/locate/jde

Symbolic dynamics and Arnold diffusion

Jacky Cresson

Equipe de Mathématiques de Besançon, Université de Franche-Comté, 16 route de Gray, 25030 Besançon, cedex, France

Received October 20, 2000; revised April 18, 2002

Abstract

We consider hyperbolic tori of three degrees of freedom initially hyperbolic Hamiltonian systems. We prove that if the stable and unstable manifold of a hyperbolic torus intersect transversaly, then there exists a hyperbolic invariant set near a homoclinic orbit on which the dynamics is conjugated to a Bernoulli shift. The proof is based on a new geometrico-dynamical feature of partially hyperbolic systems, the transversality-torsion phenomenon, which produces complete hyperbolicity from partial hyperbolicity. We deduce the existence of infinitely many hyperbolic periodic orbits near the given torus. The relevance of these results for the instability of near-integrable Hamiltonian systems is then discussed. For a given transition chain, we construct chain of hyperbolic periodic orbits. Then we easily prove the existence of periodic orbits of arbitrarily high period close to such chain using standard results on hyperbolic sets.
© 2002 Elsevier Science (USA). All rights reserved.

Keywords: Hamiltonian systems; Hyperbolic tori; Symbolic dynamics

1. Introduction

We consider Graff tori of [13] three degrees of freedom initially hyperbolic Hamiltonian systems. We extend the classical Birkhoff–Smale theorem for hyperbolic points or normally hyperbolic tori to partially hyperbolic Graff tori: if the stable and unstable manifold of a partially hyperbolic Graff tori intersect transversaly, then there exists a hyperbolic invariant set near a homoclinic orbit on which the dynamics is conjugated to a Bernouilli shift.

E-mail address: cresson@math.univ-fcomte.fr.

0022-0396/03/$ - see front matter © 2002 Elsevier Science (USA). All rights reserved.
PII: S 0 0 2 2 - 0 3 9 6 (0 2) 0 0 0 5 3 - 0

Several authors, including Treshchev [20] and Holmes and Marsden [14], have conjectured such a result. Difficulties are due to the partial hyperbolicity of the tori. A particular case is studied by Easton [10].

Our proof is based on the following key results:

(i) if the flow on the torus is with torsion, then we have a result similar to the classical λ-lemma of J. Palis for hyperbolic points [6];

(ii) transversality of the stable and unstable manifold together with torsion of the flow on the torus produce "hyperbolicity";

(iii) the minimal dynamics on the torus allows us to localize a neighbourhood of the homoclinic orbit where hyperbolicity exists (it gives the alphabet of symbolic dynamics).

An immediate application of this result is the so-called Arnold diffusion [2,11]. Assuming the existence of a transition chain, we prove the existence of a "dual" chain of periodic hyperbolic orbits, replicating the *given* chain of partially hyperbolic tori. Then, we substitute the task of tracking and studying the dynamics near a transition chain by applying the well-known facts about the dual chain of periodic hyperbolic orbits, being mostly the consequences of the λ-lemma. We then prove a statement of Holmes and Marsden [14]: there exists periodic orbits of arbitrarily high period close to the chain.

Application of these results to the computation of Arnold diffusion time is given in [8].

This paper is organized as follow: In Section 2, we recall some known results about partially hyperbolic tori and Poincaré map associated to a section of these tori. In Section 3 we state our main result: the dynamics near a homoclinic partially hyperbolic tori is conjugated to a Bernoulli shift with an infinite number of symbols. As a consequence, we prove the existence of homoclinic hyperbolic periodic orbits. We then discuss some consequences of these results for the problem of instability in Hamiltonian systems. We prove that given a transition chain there exists periodic orbits of arbitrary high period along the chain. In Section 4, we prove our main result.

2. Graff tori

2.1. Graff tori

We denote $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ and $O_n(x; y)$ a function of order $||x||^n$, parametrized by y. Let $H_\mu(J, \phi, p, q)$ be a three degrees of freedom initially hyperbolic Hamiltonian system (see [4,18]), where $(J, \phi, p, q) \in \mathbb{R}^2 \times \mathbb{T}^2 \times \mathbb{R} \times \mathbb{T}$. We denote by $\Phi(t, x)$ its flow.

Definition 2.1. A Graff torus is an invariant partially hyperbolic torus for which there exists a neighbourhood $V(\mu)$ such that the Hamiltonian takes the normal form

$$\tilde{H}_\mu(\theta, I, s, u) = \omega I + \lambda s u + f(I, su) + \mu g(\theta, I, s, u), \tag{1}$$

with $f(I,su)=O_2(I,su)$, and $g(I,\theta,s,u)=O_2(I,su;\theta,I,s,u)$, $\lambda>0$ and $\omega\in\mathbb{R}^2$ satisfies a diophantine condition

$$|\omega.k|\geqslant\frac{\gamma}{|k|^{\tau}},\quad \forall k\in\mathbb{Z}^2\backslash\{0\},\tag{2}$$

with $\tau>1$, where $V(\mu)$ is of the form

$$\tilde{V}(\mu)=\{(\theta,I,s,u)\in\mathbb{T}^2\times\mathbb{R}^2\times\mathbb{R}\times\mathbb{R}|\ |s|<\kappa,\ |u|<\kappa,\ |I|\leqslant\kappa\mu\},$$

with $\kappa>0$.

Niederman [18] and Eliasson [12] have proved that this normal form is valid for diophantine hyperbolic tori of initially hyperbolic Hamiltonian systems.

2.2. Poincaré section

As ω is non-resonant, there exists (see [16]) a section S and an analytic coordinates system (θ,ρ,s,u), $\theta\in\mathbb{T}$, $\rho\in\mathbb{R}$, $(s,u)\in\mathbb{R}^2$ in which the torus is given by $T=\{(\theta,\rho,s,u)\in\mathbb{T}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\,|\,\rho=0,s=u=0\}$, and the Poincaré map has the form

$$f(\theta,\rho,s,u)=(\theta+v(\rho),\rho,\lambda s,\lambda^{-1}u)+r(\theta,\rho,s,u),\tag{3}$$

where $\lambda<1$, v and r are analytic, and $r=O_2(\rho,s,u)$ in a domain

$$V(\mu)=\{(\theta,\rho,s,u)\in\mathbb{T}^2\times\mathbb{R}^2\times\mathbb{R}\times\mathbb{R}|\ |s|<\kappa,\ |u|<\kappa,\ |\rho|\leqslant\kappa\mu\}.$$

We have also that $v(\rho)=v+v_1\rho$, where v satisfies a diophantine condition

$$\forall k\in\mathbb{Z}\backslash\{0\},\quad |k.v|>\frac{\gamma}{|k|^{\tau}}\tag{4}$$

and $v_1\neq 0$.

3. Symbolic dynamics and Arnold diffusion

3.1. Symbolic dynamics and Graff tori

Let H_μ be a three degrees of freedom initially hyperbolic Hamiltonian systems and T a Graff torus of H_μ. We denote by $\mathscr{H}$ the energy level containing T and S the Poincaré section associated to T defined in Section 2.2.

3.1.1. Assumptions

It is assumed that:

(h_1) The stable and unstable manifolds of T intersect transversaly in $\mathscr{H}$.

Remark 3.1. This condition is generic [15].

We denote by Γ a homoclinic orbit to T. The homoclinic orbit Γ intersects the section S in $p^+ \in W^s(T)$ and $p^- \in W^u(T)$. We note that $p^+ = (\theta^+, s^+, 0, 0)$ and $p^- = (\theta^-, 0, 0, u^-)$. We define two neighbourhoods of p^+ and p^- as

$$B^+ = \{|\theta - \theta^+| \leqslant \delta_+,\ |s - s^+| \leqslant \delta_+,\ |\rho| \leqslant \delta_+ \mu,\ |u| \leqslant \delta_+\},$$

$$B^- = \{|\theta - \theta^-| \leqslant \delta_-,\ |s| \leqslant \delta_-,\ |\rho| \leqslant \delta_- \mu,\ |u - u^-| \leqslant \delta_-\},$$

where $\delta_+ > 0$, $\delta_- > 0$ sufficiently small, independent of μ.

Transversal map. Let $\tau : S \to [0, \infty[$,

$$\tau(p) = \sup\{t > 0 \mid \Phi(s,p) \in \mathscr{H} \backslash S \text{ for } 0 < s \leqslant t\},$$

where Φ is the flow associated to the Hamiltonian X_H.

We define the set $\Xi = \{p \in S \mid \tau(p) < \infty\}$. We can suppose δ_+ sufficiently small to have $B^+ \subset \Xi$ and $B^- \subset f(\Xi)$. We denote $D_n = \{q \in B^+ : f^n(q) \in B^-\}$. We also choose δ_+ and δ_- sufficiently small such that $B^- \cap B^+ = \emptyset$. Then, we have $D^n \cap D^m = \emptyset$ for $n \neq m$. We denote $D = \bigcup_{n \geqslant 1} D_n$.

The map $\mathscr{F} : D \to B^-$ defined by $\mathscr{F}(q) = f^n(q)$ for all $q \in D_n$, will be called the *transversal map*.

Homoclinic map. Let Λ be the *homoclinic map* from B^- to B^+ defined by $\Lambda(q) = \Phi(q, \sigma(q))$, where $\sigma(q) = \inf\{t > 0 \mid \Phi(t, q) \in B^+\}$. It is defined on a neighbourhood of p^- which is included in B^-. The homoclinic map $\Lambda : B^- \to B^+$ has the form

$$\Lambda(x) = p^+ + \Pi h + \Lambda_2 h,$$

where $h = x - p^-$, $\Pi = D_{p^-}(\Lambda)$ and Λ_2 is the term of order $\geqslant 2$.

(h_2) (transversality) We assume that the matrix Π, written on the basis $(e_\theta, e_s, e_\rho, e_u)$ is given by

$$\Pi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & a & 0 & b \\ \alpha & 0 & 1 & 0 \\ 0 & c & 0 & d \end{pmatrix},$$

where $(a,b,c,d) \in \mathbb{R}^4$ such that $ad - bc = 1$, $\alpha \in \mathbb{R}$ with $\alpha \neq 0$ and α of order μ^k, where $k > 0$ is a constant.

Remark 3.2. Our assumption on the form of Π is based on our computation of the homoclinic map in the Arnold's example [5, Chapter 2], where this form was obtained.

- Let $\Lambda_l(x) = p^+ + \Pi$ be the linear part of Λ. We have $T\Lambda_l = \Pi$. We denote $w_\xi^\sigma(T) = W^\xi(T) \cap B^\sigma$ and for a given manifold Δ, $\Delta^\sigma = \Delta \cap B^\sigma$, $\sigma = \pm$, $\xi = s, u$.

For $\mu = 0$, we have $\Pi(T_x w_u^-(T)) = T_{\Lambda_l(x)} w_s^+(T)$ for $x \in w_u^-(T)$. Moreover, if $\Delta \cap\!\mid W^u(T)$ in B^-, then $\Pi(T_x \Delta^-) + T_{\Delta_l(x)} w_s^+(T) = T_{\Lambda_l(x)} M$. Hence, for $\mu = 0$, the choice of the matrix Π expresses the coincidence of the stable and unstable manifold of the torus in the invariant set $I = const$.

For $\mu \neq 0$, we have $\Pi(T_x w_u^-(T)) + T_{\Lambda_l(x)} w_s^+(T) = T_{\Lambda_l(x)} M$, which expresses the transversality of the intersection between the stable and unstable manifold of the torus.

- We do not know the generic form of the matrix Π.

(h_3) (control of the remainders) We assume that

$$r = O_2(I, su) \quad \text{and} \quad \Lambda_2 = O_2(I, su). \tag{5}$$

Transition map. We define the set $\mathcal{Q}$ by $\mathcal{Q} = \{q \in B^+ \mid \Psi(q) \in B^+\}$, where the *transition* map $\Psi : \mathcal{Q} \to B^+$ is given by $\Psi(q) = \Lambda \circ F(q)$.

3.1.2. Main result

We need some notations to state our theorem.

Alphabet. We define the *alphabet* set by

$$\mathscr{A} = \{\infty > n \geqslant n_0 \mid |\theta^+ - \theta^- + n\nu| < \delta_-^{1+2\delta}\},$$

where $n_0 = \gamma \delta_-^{-(1+\delta)\tau}$ by [3], with $\delta > 0$ a constant such that $\delta_-^\delta \leqslant \frac{1}{2}$.

Window. Let $X = (\Theta, S) \in \mathbb{T} \times \mathbb{R}$, $Y = (R, U) \in \mathbb{R} \times \mathbb{R}$, $||.||_\infty$ the supnorm, and $\mathscr{B}$ a ball of centre 0 and radius 1. We denote for any $n \in \mathscr{A}$,

$$\mathscr{H}_n = \{Z = (\Theta, S, R, U) \in \mathbb{T} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \mid |U - \lambda^n \mu^{-1} u^-| < \lambda^n, \ |R| < \lambda^n\},$$

and $\mathscr{D} = \bigcup_{n \in \mathscr{A}} \mathscr{H}_n$. We define the *window* [1,10] $\mathscr{W} : \mathscr{B} \mapsto B^+$ by $\mathscr{W}(Z) = p^+ + W.Z$ and

$$W = \begin{pmatrix} \mu & 0 & \mu^{\kappa+1}\alpha^{-1} & 0 \\ 0 & \mu & 0 & b\mu d^{-1} \\ 0 & 0 & \mu^{\kappa+1} & 0 \\ 0 & 0 & 0 & \mu \end{pmatrix},$$

with $\kappa > 0$ such that $\mu^\kappa \alpha^{-1} = \mu^\delta$, where $\delta > 0$.

We denote by $\mathscr{L} : \mathscr{B} \mapsto \mathscr{B}$ the *window map* defined by $\mathscr{L} = \mathscr{W}^{-1} \circ \Psi \circ \mathscr{W}$.

Then we have the following theorem:

Theorem 3.1 (Symbolic dynamics for a Graff torus). *For μ sufficiently small, the map $\mathscr{L}$ possesses an invariant hyperbolic set*

$$I = \{ Z \in \mathscr{D} \mid \mathscr{L}^n(Z) \in \mathscr{D}, \quad \forall n \in \mathbb{Z}\}$$

such that $\mathscr{L}|_I$ is topologically conjugated to the shift on the alphabet $\mathscr{A}$. Then, there exists a homeomorphism, ϕ, such that the following diagram commutes:

$$\begin{array}{ccc} I & \mapsto^{\mathscr{L}} & I \\ \phi\downarrow & & \downarrow\phi \\ \Sigma^{\mathscr{A}} & \mapsto^{\sigma} & \Sigma^{\mathscr{A}} \end{array},$$

where $\Sigma^{\mathscr{A}} = \prod_{i=-\infty}^{\infty} \mathscr{A}$ and σ is the shift mapping on this space.

The scheme of proof of Theorem 3.1 is given in the next section (the formal proof is given in Section 4).

Remark 3.3.

- It must be emphasized that hyperbolicity of the invariant set is important. Indeed, these sets are stable under small perturbation while the persistence of the partially hyperbolic tori is subjected to constraints of arithmetical nature.
- Theorem 3.1 is also true under the weak assumption that ω in (1) is non-resonant.

3.1.3. Scheme of proof

The proof is divided in three steps. The first is to compute the window map $\mathscr{L}$. We then focus our interest on the linear part of $\mathscr{L}$. We prove that this linear map is hyperbolic. The main point is that the transversality coupled with the torsion of the linear flow on the torus create hyperbolicity. We then prove the existence of symbolic dynamics for this map using classical results on criterion for Chaos in the hyperbolic case. Third, we prove that the remainder is kept under control in a small neighbourhood of the torus. This control of the remainder is made possible via the explicit computation of the remainder of the normal form made by Eliasson [12] and Niederman [18].

(a) *Computation of $\mathscr{L}$*: Let $Z = (X, Y) \in \mathscr{B}$, the map $\mathscr{L}$ is defined by $\mathscr{L}(Z) = q + L.Z + R(Z)$, where $q = (0, \lambda^n s^+/\mu d,\ \alpha(\theta^+ - \theta^- + n\omega)/\mu,\ c\lambda^n s^+/\mu)$,

$$L = \begin{pmatrix} 0 & 0 & -\alpha^{-1}\mu^{\kappa} & 0 \\ 0 & \lambda^n d^{-1} & 0 & \lambda^n b/d^2 \\ \alpha\mu^{-\kappa} & 0 & 2 + \alpha n v_1 & 0 \\ 0 & c\lambda^n & 0 & cbd^{-1}\lambda^n + d\lambda^{-n} \end{pmatrix}$$

and R is of order 2 in Z.

This map is *hyperbolic*. Indeed, in the hyperbolic direction, we have eigenvalues which have a modulus $\neq 1$. In the angular and neutral directions, the eigenvalues β

are solution of the following equation $\beta^2 - \beta(2 + \alpha n v_1) + 1 = 0$, and are of modulus $\neq 1$ if and only if $\alpha \neq 0$ and $v_1 \neq 0$.

(b) *Criterion of Chaos for* $l(Z) = q + L.Z$: This part is essentially technical. We prove that we have the stable and unstable cones conditions (see [22]). This follows, from a geometrical point of view, from the λ-lemma for partially hyperbolic tori whose flow on the torus is with torsion (see [6]).

(c) *Control of the remainder*: Parts (a) and (b) allows us to prove that the linear map l possesses an hyperbolic invariant set $\mathscr{I}$ on which the dynamics is conjugated to a shift on an infinite number of symbols. Several difficulties arise due to the non-compactness of the set of symbols $\mathscr{A}$. Indeed, the hyperbolic invariant set $\mathscr{I}$ is then non-compact and we can not apply classical perturbation results for compact hyperbolic invariant set. The usual way to deal with this problem is [17, p. 101] to compactify the alphabet $\mathscr{A}$ by taking into account the ∞ symbol. The invariant set $\bar{\mathscr{I}}$ is then compact. But, this set is no longer hyperbolic since it contains the invariant torus T which is not hyperbolic (this is not the case for Moser's proof of the Birkhoff–Smale theorem for hyperbolic point in $\mathbb{R}^2$). However, we prove directly, via simple computations, that the remainder of the map l does not destroy the hyperbolic invariant set $\mathscr{I}$. This is due to a good behaviour of the normal form obtained by Eliasson [12] and Niederman [18] for 1-hyperbolic tori.

3.2. Transition chains

3.2.1. Dynamics around a transversal homoclinic Graff torus

For $n \in \mathscr{A}$, we denote by (n) the infinite sequence $\{\dots, n, n, n, \dots\}$. Let $p(n) = \phi^{-1}(n)$ be the associated fixed point of $\mathscr{L}$ in I by ϕ^{-1}. We obtain

$$\Psi \circ \mathscr{W}(p(n)) = \mathscr{W}(p(n)), \tag{6}$$

then $\mathscr{W}(p(n))$ is a fixed point of the transversal map in the Poincaré section.

(h_4) The Poincare map f defined on S can be extended to a neighbourhood of the homoclinic orbit Γ such that, the homoclinic map Λ is given by

$$\Lambda = f^d,$$

where d is an integer.

This assumption is already made by Moser [17]. The integer $d \in \mathbb{N}$ is related to the *homoclinic time*.

Under (h_4), we deduce that the orbit through $\mathscr{W}(p(n))$ is periodic of period $n + d$ (for the system). We denote by O_n this periodic orbit. We have $p(n) \to p^+$ when $n \to +\infty$.

We then have the following corollary of Theorem 3.1.

Corollary 3.1. *For all* $n \in \mathscr{A}$, *the periodic orbit* O_n *is hyperbolic, with a two dimensional stable (resp., unstable) manifold, denoted* $W^s(O_n)$ *(resp.,* $W^u(O_n)$*).*

Proof. This follows from the hyperbolic structure of the invariant set I. $\square$

3.2.2. Dual chain of hyperbolic periodic orbits

Let $\mathcal{T} = (T_i)_{i=1,\dots,N}$ be a family of Graff tori such that the unstable manifold $W^u(T_i)$ intersects transversaly with the stable manifold $W^s(T_{i+1})$ in $\mathcal{H}$. Moreover, we assume that for all $i \in \{1, \dots, N\}$, $W^u(T_i)$ and $W^s(T_i)$ intersect transversely in $\mathcal{H}$.

For all torus T_i of the family, we denote by S_i its section, h_i a homoclinic orbit to T_i, p_i^+ (resp., p_i^-) the intersection between h_i and $W^s(T_i) \cap S$ (resp., $W^u(T_i) \cap S$) and B_i^+ (resp., B_i^-) a neighbourhood of p_i^+ (resp., p_i^-) defined as in Section 3.1.1.

For each $i = 1, \dots, N$, we denote by O_n^i a periodic hyperbolic orbits obtained by corollary A near T_i.

Let γ_i be a heteroclinic orbit between the tori T_i and T_{i+1}, and $\Gamma_i : B_i^- \to B_{i+1}^+$ the heteroclinic map. Let A be a torus of the family $\mathcal{T}$ or a periodic orbits obtained by Corollary 3.1. We denote $w_i^{\xi,\sigma}(A) = W^{\xi}(A) \cap B_i^{\sigma}$ for $\xi = s, u$ and $\sigma = \pm$.

We assume that:

(i) the heteroclinic orbit γ_i intersects B_i^- (resp., B_{i+1}^+) in a point $q_i^- \in W^u(T_i)$ (resp., $q_i^+ \in W^s(T_{i+1})$);
(ii) there exists a diffeomorphism $\mathcal{F}_i^{\sigma} : w_i^{u,\sigma}(T_{i+1}) \to w_i^{u,\sigma}(T_i)$ for $\sigma = \pm$.

Remark 3.4. Assumptions (i) and (ii) are verified in examples using hyperbolic KAM theory.

Proposition 3.1. *Under* (i) *and* (ii), *there exists a family of periodic hyperbolic orbits $\mathcal{O} = (O_i)_{i=1,\dots,N}$ near $\mathcal{T}$ such that $W^u(O_i)$ intersects $W^s(O_{i+1})$ transversaly in $\mathcal{H}$.*

Proof. It follows from the hyperbolic structure of I that $w_i^{u,+}(O_n^i)$ is a graph of an analytic function ξ_n^+ over $w_i^{s,+}(T_i)$, for μ sufficiently small. Moreover, for n sufficiently large, there exists two constants C_1 and C_2, independent of μ and n, such that $\sup_{x \in w_i^{s,+}(T_i)} |\xi_n^+(x)| \leqslant C_1 \mu^{-1} \lambda^n$ and $\sup_{x \in w_i^{s,+}(T_i)} |D\xi_n^+(x)| \leqslant C_2 \mu^{-3} \lambda^n$, by definition of $\mathcal{H}_n$. Then, $w_i^{u,+}(O_n^i)$ and $w_i^{s,+}(T_i)$ intersect transversaly in $\mathcal{H}$ for n sufficiently large.[1] By the λ-lemma (see [6]), there exists $k \in \mathbb{N}$ such that $f^k(w_i^{u,+}(O_n^i))$ is as close as we want to $w_i^{u,-}(T_i)$ in C^1 topology, for k sufficiently large. The image of $f^k(w_i^{u,+}(O_n^i))$ and $w_i^{u,-}(T_i)$ by Γ_i are then as close as we want in C^1 topology in B_{i+1}^+. We have also $w_{i+1}^{s,+}(O_n^{i+1})$ as close as we want to $w_{i+1}^{s,+}(T_{i+1})$ in C^1 topology for n sufficiently large. As $w_{i+1}^{u,+}(T_i)$ and $w_{i+1}^{s,+}(T_{i+1})$ intersect transversaly in $\mathcal{H}$, we deduce from (i) and (ii) that $w_{i+1}^{s,+}(O_n^{i+1})$ and $\Gamma_i(f^k(w_i^{u,+}(O_n^i)))$ intersect transversaly in $\mathcal{H}$. We conclude the proof by induction. $\square$

[1] By the persistence lemma proved in [7], as the angle between $w_i^{u,+}(T_i)$ and $w_i^{s,+}(T_i)$ is of order μ^k, we must have $n \geqslant \frac{2k+3}{\log(1/\lambda)} \log\left(\frac{1}{\mu}\right)$.

The existence of orbits shadowing the chain follows easily from the standard λ-lemma of Palis (see [19]).

Corollary 3.2. *Let $p \in W^s(O_1)$ (resp. $q \in W^u(O_N)$) and $\mathcal{U}$ (resp. $\mathcal{V}$) be an arbitrary neighbourhood of p (resp. q), then there exists trajectories $\xi(t)$ and a real $T > 0$ such that $\xi(0) \in \mathcal{U}$ and $\xi(T) \in \mathcal{V}$.*

3.2.3. Holmes–Marsden conjecture

In [14], Holmes and Marsden discuss, in a heuristic paragraph about "Nonintegrability and Arnold diffusion" (Section 3, pp. 672–673), dynamical and analytical consequences of the existence of a transition chain. One of their conjecture is, among other things, that "since two-way transition chain can be chosen, we can find periodic motions of arbitrarily high period close to such chains, just as in the standard two-dimensional horseshoes example" [14, p. 673]).

In this section, we prove this conjecture using the dual chain of hyperbolic periodic orbits. We need some definitions.

We say that there exists periodic orbits of *arbitrary high period* if for any $P \in \mathbb{R}$, there exists a periodic orbits of period $\tilde{P}$ with $\tilde{P} > P$.

Let $p \in W^s(T_1)$ (resp., $q \in W^u(T_N)$), and U (resp., V) be an arbitrary neighbourhood of p (resp., q). We say that there exists periodic orbits of arbitrarily high period *close to the chain*, if there exists periodic orbits of arbitrarily high period denoted by $\xi(t)$, such that $\xi \cap U \neq \emptyset$ and $\xi \cap V \neq \emptyset$, where $\xi = \{\xi(t); t \in \mathbb{R}\}$.

We have the following corollary of Proposition 3.1:

Corollary 3.3. *There exists periodic orbits of arbitrarily high period close to the chain $\mathcal{T}$.*

This follows from Proposition 3.1 using standard results on hyperbolic set, as developed by Alekseev [1] and Easton [10].

4. Proof of Theorem 3.1

4.1. Reminder about symbolic dynamics

We recall some basic results about symbolic dynamics and criteria for Chaos as exposed by Wiggins [21, p. 108–150]).

We consider a map $f : D \to \mathbb{R}^n \times \mathbb{R}^m$ where D is a closed and bounded $(n+m)$-dimensional set contained in $\mathbb{R}^n \times \mathbb{R}^m$. We assume that f is analytic.

4.1.1. Definitions and notations

We denote D_x (resp., D_y) the set of $x \in \mathbb{R}^n$ (resp., $y \in \mathbb{R}^m$) for which there exits $y \in \mathbb{R}^m$ (resp., $x \in \mathbb{R}^n$) with $(x, y) \in D$. Let I_x (resp., I_y) be a closed, simply connected n (resp., m) dimensional set contained in D_x (resp., D_y).

Definition 4.1. A μ_+ horizontal (resp., μ_- vertical) slice, $\bar{H}$ (resp., $\bar{V}$), is defined to be the graph of a function $h : I_x \to \mathbb{R}^m$ (resp., $v : I_y \to \mathbb{R}^n$), μ_+ (resp., μ_-) Lipschitz, $0 < \mu_+ \leqslant \infty$ (resp., $0 < \mu_- < \infty$), such that $\bar{H} \subset D$ (resp., $\bar{V} \subset D$).

Fix some μ_+, $0 \leqslant \mu_- < \infty$. Let $\bar{H}$ be a μ_+ horizontal slice, and let $J^m \subset D$ be an m-dimensional topological disk intersecting $\bar{H}$ at any, but only one, point of $\bar{H}$. Let $\bar{H}^\alpha$, $\alpha \in I$, be the set of all μ_h horizontal slices that intersect the boundary of J^m and have the same domain as $\bar{H}$, where I is some index set. Let

$S_H = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid x \in I_x,\ y$ such that, $\forall x \in I_x$, given any line L through (x, y) parallel to the $x = 0$ plane, then L intersects the points $(x, h_\alpha(x))$, $(x, h_\beta(x))$ for some $\alpha, \beta \in I$ with (x, y) between these two points along $L.\}$.

Then a μ_h horizontal slab H is defined to be the closure of S_H.

Definition 4.2. The vertical boundary of a μ_+ horizontal slab H is denoted $\partial_v H$ and is defined as $\partial_v H = \{(x, y) \in H \mid x \in \partial I_x\}$. The horizontal boundary of H is denoted $\partial_h H$ and is defined as $\partial_h H = \partial H - \partial_v H$.

Definition 4.3. Let H and $\tilde{H}$ be μ_+ horizontal slabs. $\tilde{H}$ is said to intersect H fully if $\tilde{H} \subset H$ and $\partial_v \tilde{H} \subset \partial_v H$.

4.1.2. Criteria for Chaos in the hyperbolic case

Let $S = \mathbb{N}$, and let H_i, $i \geqslant 1$ be a set of disjoint μ_+-horizontal slabs with $D_H = \bigcup_{i=1}^N H_i$. We assume that f is one to one on D_H and we define

$$f(H_i) \cap H_j = V_{ji}, \quad \forall i, j \in S$$

and

$$H_i \cap f^{-1}(H_j) = f^{-1}(V_{ji}) = H_{ij}, \quad \forall i, j \in S$$

We denote $\Sigma = \prod_{i=-\infty}^{+\infty} S$. Let $s \in \Sigma$, $s = \{\dots s_{-n}, \dots, s_0, \dots s_n \dots\}$, we define the *shift map* $\sigma : \Sigma \mapsto \Sigma$ as $[\sigma(s)]_i = s_{i+1}$.

Let $\mathscr{H} = \bigcup_{i,j \in S} H_{ij}$ and $\mathscr{V} = \bigcup_{i,j \in S} V_{ij}$, then $f(\mathscr{H}) = \mathscr{V}$. Let $z_0 = (x_0, y_0) \in \mathscr{V} \cup \mathscr{H}$, the *stable sector* at z_0, denoted $S^+_{z_0}$, is defined as follows:

$$S^+_{z_0} = \{(\xi_{z_0}, \eta_{z_0}) \in \mathbb{R}^n \times \mathbb{R}^m \mid |\eta_{z_0}| \leqslant \mu_+ |\xi_{z_0}|\}.$$

The *unstable sector* at z_0, denoted $S^-_{z_0}$, is defined as

$$S^-_{z_0} = \{(\xi_{z_0}, \eta_{z_0}) \in \mathbb{R}^n \times \mathbb{R}^m \mid |\xi_{z_0}| \leqslant \mu_- |\eta_{z_0}|\}.$$

We denote $S^+_{\mathscr{H}} = \bigcup_{z_0 \in \mathscr{H}} S^+_{z_0}$, $S^+_{\mathscr{V}} = \bigcup_{z_0 \in \mathscr{V}} S^+_{z_0}$, $S^-_{\mathscr{H}} = \bigcup_{z_0 \in \mathscr{H}} S^-_{z_0}$ and $S^-_{\mathscr{V}} = \bigcup_{z_0 \in \mathscr{V}} S^-_{z_0}$.

We have the following hypothesis:

(A) $Df(S^-_{\mathscr{H}}) \subset S^-_{\mathscr{V}}$ and $Df^{-1}(S^+_{\mathscr{V}}) \subset S^+_{\mathscr{H}}$. Moreover, if $(\xi_{z_0}, \eta_{z_0}) \in S^-_{z_0}$ and $Df(z_0)(\xi_{z_0}, \eta_{z_0}) = (\xi_{f(z_0)}, \eta_{f(z_0)}) \in S^-_{f(z_0)}$, then we have

$$|\eta_{f(z_0)}| \geqslant \mu^{-1} |\eta_{z_0}|.$$

Similarly, if $(\xi_{z_0}, \eta_{z_0}) \in S^+_{z_0}$ and $Df^{-1}(z_0)(\xi_{z_0}, \eta_{z_0}) = (\xi_{f^{-1}(z_0)}, \eta_{f^{-1}(z_0)}) \in S^+_{f^{-1}(z_0)}$ then we have

$$|\xi_{f^{-1}(z_0)}| \geqslant \mu^{-1} |\xi_{z_0}|,$$

where $0 < \mu < 1 - \mu_- \mu_+$.

We denote $a_1 = ||\partial_y f_1||\, ||(\partial_y f_2)^{-1}||$, $a_2 = (1 - ||\partial_x f_1|| ||(\partial_y f_2)^{-1}||)$, $a_3 = ||\partial_x f_2||\, ||(\partial_y f_2)^{-1}||$, and

$$a_1 \mu_+^2 - a_2 \mu_+ + a_3 < 0, \tag{$*$}$$

$$\mu_+ < \frac{1 - ||(\partial_y f_2)^{-1}||}{||\partial_y f_1||\, ||(\partial_y f_2)^{-1}||}, \tag{$**$}$$

$$a_3 \mu_-^2 - a_2 \mu_- + a_1 < 0, \tag{$*_{\rm v}$}$$

$$\mu_- < \frac{1 - ||\partial_x f_1||}{||\partial_x f_2||}, \tag{$**_{\rm v}$}$$

$$0 \leqslant \mu_+ \mu_- < 1. \tag{$***$}$$

(B) H_i, $i \geqslant 1$ are μ_+ horizontal slabs with μ_+ satisfying $(*)$ and $(**)$. For all $i, j \in S$ such that V_{ji} is a μ_- vertical slab with μ_- satisfying $(*)_v$, $(**)_v$ and $(***)$. Moreover, we require $\partial_v V_{ji} \subset \partial f(H_i)$ and $f^{-1}(\partial_v V_{ji}) \subset \partial_v H_i$.

We have:

Theorem 4.1. *If f satisfies assumptions* (A) *and* (B), *then it possesses an hyperbolic invariant set I, $I \subset D_{\mathscr{H}}$, on which f is topologically conjugated to the shift on the alphabet Σ.*

4.2. Preliminaries

4.2.1. A remark on $\mathscr{W}$

We explain the choice of $\mathscr{W}$, made in Section 3.1.2. The role of the window map is to "*simplify*" the geometry of the problem, in order to have a clear understanding of the set of *horizontals* and *verticals* on which the symbolic dynamics is constructed.

Let $\mathscr{W}(Z) = p^+ + WZ$, where W is a matrix to compute. We want that W ensures the following conditions:

(i) For $Z \in \mathscr{B}$, we have $\mathscr{W}(Z) \in B^+$.
(ii) The horizontals of $\mathscr{W}$ are parallel to the tangent plane of the stable manifold in B^+.
(iii) The image of a vertical of $\mathscr{W}$ by Λ^{-1} is parallel to the tangent plane of the unstable manifold of the torus in B^-.

Conditions (ii) and (i) implies that W can be chosen as

$$W = \begin{pmatrix} p_1 & p_2 & q_1 & q_2 \\ p_3 & p_4 & q_3 & q_4 \\ 0 & 0 & \mu^{\kappa+1} & 0 \\ 0 & 0 & 0 & \mu \end{pmatrix}, \tag{7}$$

where p_i and q_i are unknown constants.

Let $e_R = (0,0,1,0)$, $e_U = (0,0,0,1)$ and for a matrix M and a vector V, we denote by $(MV)_x$ the x-component of the vector MV. Condition (ii) gives the following constraints:

$$\begin{aligned} (\Pi^{-1} W(e_R))_R &= 0, \\ (\Pi^{-1} W(e_R))_S &= 0, \\ (\Pi^{-1} W(e_U))_R &= 0, \\ (\Pi^{-1} W(e_U))_S &= 0. \end{aligned} \tag{8}$$

We then obtain

$$\begin{aligned} -\alpha q_1 + \mu^{\kappa+1} &= 0, \\ d v q_3 &= 0, \\ -\alpha q_2 &= 0, \\ d v q_4 - b v \mu &= 0. \end{aligned} \tag{9}$$

Hence, the matrix W has the form

$$W = \begin{pmatrix} p_1 & p_2 & \mu^{\kappa+1}\alpha^{-1} & 0 \\ p_3 & p_4 & 0 & b\mu d^{-1} \\ 0 & 0 & \mu^{\kappa+1} & 0 \\ 0 & 0 & 0 & \mu \end{pmatrix}. \tag{10}$$

We complete the matrix W by requiring it to be invertible. For example, we can take $p_1 = \mu$, $p_2 = 0$, $p_3 = 0$, $p_4 = \mu$. We then have the form of W used in Section 3.1.2.

4.2.2. Preliminaries

For $z \in D_n$, we have $\mathscr{F}(z) = (nv, 0, 0, 0) + Fz + r(z)$, where

$$F = \begin{pmatrix} 1 & 0 & nv_1 & 0 \\ 0 & \lambda^n & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \lambda^{-n} \end{pmatrix} \tag{11}$$

and r is of order 2 in z.

We have the following lemma:

Lemma 4.1. *Let $Z = (X, Y) \in \mathbb{R}^2 \times \mathbb{R}^2 \in \mathscr{H}_n$, $n \in \mathscr{A}$, then, we have $\mathscr{W}(Z) \in D_n$.*

Proof. For $Z \in \mathscr{H}_n$, we have $\mathscr{W}(Z) = (\theta^+, s^+, 0, 0) + W.Z$. We deduce that

$$\mathscr{W}(Z) = \begin{pmatrix} \theta^+ + \mu\Theta + \mu^{\kappa+1}\alpha^{-1}R \\ s^+ + \mu S + b\mu d^{-1}U \\ \mu^{\kappa+1}R \\ \mu U \end{pmatrix}. \tag{12}$$

We remark that $\mathscr{W}(Z) \in B^+$. Hence, we have $f^n \circ \mathscr{W}(Z)$ which is equal to

$$f^n \circ \mathscr{W}(Z) = \begin{pmatrix} \theta^+ + \mu\Theta + mu^{\kappa+1}\alpha^{-1}R + nv + nv_1\mu^{\kappa+1}R \\ \lambda^n(s^+ + \mu S + b\mu d^{-1}U) \\ \mu^{\kappa+1}R \\ \lambda^{-n}\mu U \end{pmatrix}. \tag{13}$$

We denote by z the point $f^n \circ \mathscr{W}(Z)$. As for $Z \in \mathscr{H}_n$, we have $|U - \lambda^n\mu^{-1}u^-| < \lambda^n$,we obtain that $z = (\theta, \rho, s, u)$ satisfies $|u - u^-| \leqslant \mu$. Moreover, we have $|\rho| \leq \mu^{\kappa+1}$ and $|s| \leqslant \mu$ as long as $\lambda^n|s^+| < \mu$ and $\lambda^n|bd^{-1}| \leqslant \mu$.

We have $\theta - \theta^- = (\theta^+ - \theta^- + nv) + \mu(\theta + \mu\alpha^{-1}R + nv_1R)$. As $|R| < \lambda^n$, and by definition of the alphabet $\mathscr{A}$, $|\theta - \theta^- + nv| < \delta_-^{1+\delta}$, we have $|\theta - \theta^-| < \delta_-$. $\square$

We are now ready to compute the window map $\mathscr{L}:\mathscr{B}\to\mathscr{B}$ defined by $\mathscr{L}=\mathscr{W}^{-1}\circ\Psi\circ\mathscr{W}$.

Lemma 4.2. *We have for* $Z\in\mathscr{H}_n$, $\mathscr{L}(Z)=q+L.Z+R(Z)$, *where* $q=(0,\lambda^n s^+/\mu d,\alpha(\theta^+-\theta^-+nv)/\mu,c\lambda^n s^+/\mu)$, *and*

$$L=\begin{pmatrix} 0 & 0 & -\alpha^{-1}\mu^{\kappa} & 0\\ 0 & \lambda^n d^{-1} & 0 & \lambda^n b/d^2\\ \alpha\mu^{-\kappa} & 0 & 2+\alpha n v_1 & 0\\ 0 & c\lambda^n & 0 & cbd^{-1}\lambda^n+d\lambda^{-n}\end{pmatrix} \tag{14}$$

and R is of order 2 *in* Z.

Proof. We have for $Z\in\mathscr{H}_n$,

$$F\circ W=\mu\begin{pmatrix} 1 & 0 & \mu^{\kappa}(nv_1+\alpha^{-1}) & 0\\ 0 & \lambda^n & 0 & b\lambda^n d^{-1}\\ 0 & 0 & \mu^{\kappa} & 0\\ 0 & 0 & 0 & \lambda^{-n}\end{pmatrix}. \tag{15}$$

By composition with the linear part of the homoclinic map, we have

$$\Pi\circ F\circ W=\mu^{-1}\begin{pmatrix} 1 & 0 & \mu^{\kappa}(nv_1+\alpha^{-1}) & 0\\ 0 & a\lambda^n & 0 & abd^{-1}\lambda^n+b\lambda^{-n}\\ \alpha & 0 & \mu^{\kappa}(\alpha v_1+2) & 0\\ 0 & c\lambda^n & 0 & cbd^{-1}\lambda^n+d\lambda^{-n}\end{pmatrix}. \tag{16}$$

As

$$W^{-1}=\mu^{-1}\begin{pmatrix} 1 & 0 & -\alpha^{-1} & 0\\ 0 & 1 & 0 & -bd^{-1}\\ 0 & 0 & \mu^{-\kappa} & 0\\ 0 & 0 & 0 & 1\end{pmatrix}, \tag{17}$$

we obtain that the linear part of the window map is

$$L = W^{-1} \circ \Pi \circ F \circ W = \begin{pmatrix} 0 & 0 & -\mu^{\kappa}\alpha^{-1} & 0 \\ 0 & d^{-1}\lambda^{n} & 0 & bd^{-2}\lambda^{n} \\ \alpha\mu^{-\kappa} & 0 & \alpha n v_1 + 2 & 0 \\ 0 & c\lambda^{n} & 0 & cbd^{-1}\lambda^{n} + d\lambda^{-n} \end{pmatrix}, \tag{18}$$

which concludes the proof. $\square$

The following lemma illustrates the *transversality-torsion* phenomenon (see [9]).

Lemma 4.3. *The matrix L is hyperbolic if and only if $\alpha \neq 0$ and $v_1 \neq 0$. Moreover, its eigenvalues are given by*

$$l_1 \sim d\lambda^{-n}, \quad l_2 \sim d^{-1}\lambda^{n}, \quad l_3 \sim \alpha n v_1, \quad l_4 \sim (\alpha n v_1)^{-1}$$

for n sufficiently great, and μ sufficiently small.

Proof. The characteristic polynomial of L is given by

$$P(x) = (x^2 - x(\alpha n v_1 + 2) + 1)(x^2 - xa(n) + 1), \tag{19}$$

where $a(n) = a\lambda^{2n} + d\lambda^{-n}$.

We first study the second factor of this polynomial. The eigenvalues are given by

$$x_{\pm} = \frac{(a\lambda^{n} + d\lambda^{-n})}{2}(1 \pm \sqrt{1 - 4(a\lambda^{n} + d\lambda^{-n})^{-2}}). \tag{20}$$

So, for n sufficiently large, we have

$$x_{\pm} \sim \frac{(a\lambda^{n} + d\lambda^{-n})}{2}\left(1 \pm 1 \pm \frac{-2}{(a\lambda^{n} + d\lambda^{-n})^{2}}\right), \tag{21}$$

so that we have two hyperbolic directions, whose Lyapounov exponents are (up to negligible factors), $x_{+} \sim d\lambda^{-n}$ and $x_{-} \sim d^{-1}\lambda^{n}$.

For the first factor, the eigenvalues are given by

$$x_{\pm} = 1 + \frac{\alpha n v_1}{2} \pm \frac{1}{2}(\alpha^2 (n v_1)^2 + 4\alpha n v_1)^{1/2}. \tag{22}$$

We see that if $\alpha = 0$ (or $v_1 = 0$) then, we have $x_{\pm} = 1$.

If $\alpha \neq 0$ and $v_1 \neq 0$, we obtain two hyperbolic directions whose Lyapounov exponents are of order $2 + \tau(n v_1)^{-1} + \alpha n v_1$ and $(\alpha n v_1)^{-1}$, for n sufficiently large and μ sufficiently small. This concludes the proof of lemma 4.3. $\square$

4.3. The linear case

We prove that the linear part of $\mathscr{L}$ satisfies assumptions (A) and (B). As a consequence, it is conjugated to a shift automorphism on the alphabet $\mathscr{A}$.

We denote $l(Z) = (l_1(Z), l_2(Z))$, where $l_i : \mathscr{B} \to \mathbb{R}^2$ is defined by $l_i(Z) = q_i + L_i.Z$, $i = 1, 2$, with $q_1 = (0, \lambda^n s^+ (\mu d)^{-1})$, $q_2 = (\alpha(\theta^+ - \theta^- + nv)\mu^{-1}, c\lambda^n s^+ \mu^{-1})$, and

$$L_1 = \begin{pmatrix} 0 & 0 & -\mu^\kappa \alpha^{-1} & 0 \\ 0 & \lambda^n d^{-1} & 0 & \lambda^n b d^{-2} \end{pmatrix}, \tag{23}$$

$$L_2 = \begin{pmatrix} \alpha\mu^{-\kappa} & 0 & 2 + \alpha n v_1 & 0 \\ 0 & c\lambda^n & 0 & cbd^{-1}\lambda^n + d\lambda^{-n} \end{pmatrix}. \tag{24}$$

4.3.1. Estimates

We have

$$\partial_X l_1 = \begin{pmatrix} 0 & 0 \\ 0 & \lambda^n d^{-1} \end{pmatrix}, \quad \partial_Y l_2 = \begin{pmatrix} 2 + \alpha n v_1 & 0 \\ 0 & cbd^{-1}\lambda^n + d\lambda^{-n} \end{pmatrix} \tag{25}$$

and

$$\partial_Y l_1 = \begin{pmatrix} -\mu^\kappa \alpha^{-1} & 0 \\ 0 & \lambda^n b d^{-2} \end{pmatrix}, \quad \partial_X l_2 = \begin{pmatrix} \alpha\mu^{-\kappa} & 0 \\ 0 & c\lambda^n \end{pmatrix}. \tag{26}$$

We deduce the following estimates:

$$||\partial_X l_1|| = |d^{-1}|\lambda^n,$$

$$||(\partial_Y l_2)^{-1}|| = (2 + \alpha n v_1)^{-1} \tag{27}$$

and

$$||\partial_Y l_1|| = \mu^\kappa \alpha^{-1},$$

$$||\partial_X l_2|| = \alpha\mu^\kappa. \tag{28}$$

4.3.2. Condition (A)

We denote by

$$S^+(p) = \{(v_X(p), v_Y(p)) \in \mathbb{R}^2 \times \mathbb{R}^2; |v_Y(p)| \leqslant \mu^+ |v_X(p)|\},$$

$$S^-(p) = \{(v_X(p), v_Y(p)) \in \mathbb{R}^2 \times \mathbb{R}^2; |v_X(p)| \leqslant \mu^- |v_Y(p)|\}, \tag{29}$$

with $v_X = (v_\Theta, v_S)$, $v_Y = (v_R, v_U)$, the *stable* and *unstable* sector at p, respectively (see Section 4.1).

We denote

$$(v'_X, v'_Y) = L(v_X, v_Y). \tag{30}$$

We have

$$v'_X = (v'_\Theta, v'_S) = (-\mu^\kappa \alpha^{-1} v_R, \lambda^n d^{-1} v_S + \lambda^n b d^{-2} v_U),$$

$$v'_Y = (v'_R, v'_U)$$

$$= (\alpha \mu^{-\kappa} v_\Theta + (2 + \alpha n v_1) v_R, c\lambda^n v_S + (cbd^{-1}\lambda^n + d\lambda^{-n}) v_U). \tag{31}$$

(a) $L(S^-_{\mathscr{H}}) \subset S^-_{\mathscr{V}}$. We have

$$|v'_X| = \mu^\kappa \alpha^{-1} |v_R| + |\lambda^n d^{-1} v_S + \lambda^n b d^{-2} v_U|,$$

$$|v'_Y| = |\alpha \mu^{-\kappa} v_\Theta + (2 + \alpha n v_1) v_R| + |c\lambda^n v_S + (cbd^{-1}\lambda^n + d\lambda^{-n}) v_U|. \tag{32}$$

We assume that $|v_X| \leqslant \mu^- |v_Y|$. Then, we have

$$|v'_Y| = (2 + \alpha n v_1) \alpha \mu^{-\kappa} |v_\Theta (2 + \alpha n v_1)^{-1} + \alpha^{-1} \mu^\kappa v_R|$$

$$+\ |d| \lambda^{-n} |d^{-1} c \lambda^{2n} v_S + (cbd^{-2} \lambda^{2n} + 1) v_U|$$

$$\geqslant \alpha \mu^{-\kappa} |v_\Theta| + |d| \lambda^{-n} |v_U| - \delta |v'_X|, \tag{33}$$

where $\delta = \max((2 + \alpha n v_1) \alpha \mu^{-\kappa}, |d|)$.

As

$$|v'_X| \leqslant \mu^\kappa \alpha^{-1} |v_R| + \lambda^n |bd^{-2}| |v_U| + \lambda^n |d^{-1}| |v_S|, \tag{34}$$

we deduce that for $\mu \to 0$, we have $|v'_X| \to 0$. Hence, for μ sufficiently small, we have, if $|v_U| \neq 0$,

$$|v_U| \geqslant |v'_X|. \tag{35}$$

Then, using (33), we obtain that

$$|v'_Y| \geqslant |d| \lambda^{-n} |v'_X| - \alpha |v'_X| \geqslant |v'_X| \lambda^{-n} (|d| - \alpha \lambda^n). \tag{36}$$

We deduce that for μ sufficiently small,

$$|v'_X| \leqslant 2\lambda^n |d|^{-1} |v'_Y|. \tag{37}$$

We must have

$$2\lambda^n |d|^{-1} \leqslant \mu^-, \tag{38}$$

which is the case for μ sufficiently small.

If $|v_U| = 0$, then, we have

$$\frac{|v'_X|}{|v'_Y|} \leqslant \frac{\mu^\kappa \alpha^{-1} |v_R| + \lambda^n |d^{-1}||v_S|}{c\lambda^n |v_S| + (2\alpha n v_1)|v_R| - \alpha\mu^{-\kappa}|v_\Theta} \leqslant 2\frac{\mu^\kappa \alpha^{-1}}{2 + \alpha n v_1}$$

for μ sufficiently small. Then, we must have

$$\mu^- \geqslant 2\frac{\mu^{\kappa\alpha^{-1}}}{2 + \alpha n v_1}.$$

Remark 4.1. Here, we have used the torsion in order to satisfy the unstable cone condition. This computation must be compared with the proof of the λ-lemma in [6].

(b) $L^{-1}(S^+_{\mathscr{V}}) \subset S^+_{\mathscr{H}}$.

- We assume that

$$|v_R| + |v_U| \leqslant \mu^+(|v_\Theta| + |v_S|). \tag{39}$$

We denote for $v = (v_X, v_Y) \in S^+_{\mathscr{V}}$,

$$L^{-1}v = (v'_X, v'_Y) \tag{40}$$

with

$$L^{-1} = \begin{pmatrix} \alpha n v_1 + 2 & 0 & \mu^\kappa \alpha^{-1} & 0 \\ 0 & cbd^{-1}\lambda^n + d\lambda^{-n} & 0 & -\lambda^n b d^{-2} \\ -\alpha\mu^{-\kappa} & 0 & 0 & 0 \\ 0 & -c\lambda^n & 0 & d^{-1}\lambda^n \end{pmatrix}. \tag{41}$$

We deduce that

$$\begin{aligned} v'_X &= (v'_\Theta, v'_S) \\ &= \left((\alpha n v_1 + 2)v_\Theta + \mu^\kappa \alpha^{-1} v_R, \left(c\frac{b}{d}\lambda^n + d\lambda^{-n}\right)v_S - \lambda^n \frac{b}{d^2} v_U\right), \\ v'_Y &= (v'_R, v'_U) = (-\alpha\mu^{-\kappa} v_\Theta, -c\lambda^n v_R + \frac{\lambda^n}{d} v_U). \end{aligned} \tag{42}$$

We have

$$\begin{aligned} |v'_X| &= |(\alpha n v_1 + 2)v_\Theta + \mu^\kappa \alpha^{-1} v_R| + |(cbd^{-1}\lambda^n + \lambda^{-n} d)v_S - \lambda^n b d^{-2} v_U|, \\ &\geqslant (\alpha n v_1 + 2)|v_\Theta| - \mu^\kappa \alpha^{-1}|v_R| + \lambda^{-n}|d + cbd^{-1}\lambda^{2n}||v_S| \\ &\quad - \lambda^n |b| d^{-2} |v_U|, \end{aligned} \tag{43}$$

which gives using (39),

$$\begin{aligned} |v'_X| &\geqslant (\alpha n v_1 + 2)|v_\Theta| - \mu^\kappa \alpha^{-1} + \frac{\lambda^{-n}}{\mu^+}|d + cbd^{-1}\lambda^{2n}|(|v_R| + |v_U|) \\ &\quad - \lambda^n |b| d^{-2} |v_U|, \\ &\geqslant (\alpha n v_1 + 2)|v_\Theta| + |v_R|\left(\frac{\lambda^{-n}}{\mu^+}|d + cbd^{-1}\lambda^{2n}| - \mu^\kappa \alpha^{-1}\right) \\ &\quad + \left(\frac{\lambda^{-n}}{\mu^+}|d + cbd^{-1}\lambda^{2n}| - \lambda^n |b| d^{-2}\right)|v_U|. \end{aligned} \tag{44}$$

As $\mu^\kappa \alpha^{-1} = \mu^\delta$, we have for $\mu \to 0$, $\mu^\kappa \alpha^{-1} \to 0$. Moreover, $\lambda^n \to 0$ for $\mu \to 0$. Then, for μ sufficiently small we have

$$\begin{aligned} |v'_X| &\geqslant (\alpha n v_1 + 2)|v_\Theta| + \left| v_R \left| \frac{\lambda^{-n}}{2\mu^+} \right| d + cbd^{-1}\lambda^{2n} \right| \\ &\quad + \left| v_U \left| \frac{\lambda^{-n}}{2\mu^+} \right| d + cbd^{-1}\lambda^{2n} \right|. \end{aligned} \tag{45}$$

For μ sufficiently small, we have $\alpha n v_1 + 2 \leqslant \frac{\lambda^{-n}}{2\mu^+}|d + cbd^{-1}\lambda^{2n}|$. Hence, we obtain

$$|v'_X| \geqslant (\alpha n v_1 + 2)(|v_\Theta| + |v_R| + |v_U|). \tag{46}$$

We have also that

$$|v'_Y| \leqslant \alpha\mu^{-\kappa}|v_\Theta| + |c|\lambda^n |v_R| + \lambda^n |d|^{-1} |v_U|. \tag{47}$$

As $\alpha\mu^\kappa \leqslant \alpha n v_1 + 2$ for $n \in \mathscr{A}$, and $\lambda^n \to 0$ when $\mu \to 0$, we deduce that

$$|v'_Y| \leqslant (\alpha n v_1 + 2)^{-1} |v'_Y|. \tag{48}$$

We conclude that $L^{-1}(S^+_{\mathscr{V}}) \subset S^+_{\mathscr{H}}$ if

$$(\alpha n v_1 + 2)^{-1} \leqslant \mu^+, \tag{49}$$

which is the case for μ sufficiently small.

(c) *Consequences*. We can take

$$\mu^+ = (\alpha n_0 v_1 + 2)^{-1}, \quad \mu^- = 2\frac{\mu^\kappa \alpha^{-1}}{2 + \alpha n_0 v_1}, \tag{50}$$

in order to ensure the previous conditions.

Moreover, we have for $v \in S^-$,

$$|v'_Y| \geqslant \beta^{-1} |v_X|, \tag{51}$$

and, for $v \in S^+$,

$$|v'_X| \geqslant \beta^{-1} |v'_Y|, \tag{52}$$

where β is given by

$$\beta^{-1} = \max(\alpha n v_1 + 2, \tfrac{1}{2}\mu^{-\kappa}\alpha(2 + \alpha n v_1)). \tag{53}$$

We have

$$\mu^+\mu^- = 2\mu^\kappa \alpha^{-1} (\alpha n_0 v_1 + 2)^{-2}. \tag{54}$$

We note that $1 - \mu^+\mu^- \to 1$ and $\alpha \to 0$ when $\mu \to 0$. Hence, for μ sufficiently small, we have

$$0 < \beta \leqslant 1 - \mu^+\mu^-. \tag{55}$$

This concludes the proof of condition (A).

4.3.3. *Condition (B)*

It is easy to picture the sets $\mathscr{H}_n$ and $V_{nm} = l(\mathscr{H}_n) \cap \mathscr{H}_m$ and to verify condition (B).

Of course, for μ sufficiently small and $n > m$, we have

$$\mathscr{H}_n \cap \mathscr{H}_m = \emptyset, \tag{56}$$

so that the horizontal slabs are disjoint sets.

We recall that $V_{nm} = l(\mathscr{H}_n) \cup \mathscr{H}_m$. We denote for $Z \in \mathscr{H}_n$, $l(Z) = (Z'_n)$, so that

$$
\begin{aligned}
&\Theta'_n = -\mu^{\kappa}\alpha^{-1}R,\\
&S'_n = \lambda^n s^+(\mu d)^{-1} + \lambda^n S d^{-1} + b d^{-2}\lambda^n U,\\
&R'_n = \alpha\mu^{-(\kappa+1)}(\theta^+ - \theta^- + nv) + \alpha\mu^{-\kappa}\Theta + \alpha(nv_1 + 2)R,\\
&U'_n = c\lambda^n s^+\mu^{-1} - du^-\mu^{-1} + c\lambda^n S + cb\lambda^n d^{-1}U + d\lambda^n U.
\end{aligned} \tag{57}
$$

As $Z = (\Theta, S, R, U) \in \mathscr{H}_n$, we have $|R| < \lambda^n$, and as a consequence

$$|\Theta'_n| < \mu^{\kappa}\alpha^{-1}\lambda^n. \tag{58}$$

From $|\Theta| < 1$, $|R| < \lambda^n$ and $\alpha\mu^{-\kappa} = \mu^{\delta}$, with $\delta > 0$, we deduce that the set spanned by R'_n, denoted by $I_{R'_n}$ is such that

$$I_{R'_n} \supset [0, 1] \tag{59}$$

Moreover, as $|d|\lambda^{-n}|U - \lambda^n u^- \mu^{-1}| \leqslant |d|\lambda^{-n}$, we deduce that for μ sufficiently small the set $I_{U'_n}$ spanned by U'_n when Z varies in $\mathscr{H}_n$ is such that

$$I_{U'_n} \supset [0, 1]. \tag{60}$$

Moreover, we deduce that

$$|S'_n - \lambda^n s^+(\mu d)^{-1}| = \lambda^n |d|^{-1} |bU + S|. \tag{61}$$

As $|U| < \lambda^n + \lambda^n \mu^{-1} u^-$ and $|S| \leqslant 1$, we obtain

$$|S'_n - \lambda^n s^+(\mu d)^{-1}| \leqslant \tfrac{3}{2}\lambda^n |d|^{-1}. \tag{62}$$

We then deduce easily that the vertical slab V_{nm} is defined by the constraints given by $l(\mathscr{H}_n)$ for Θ, S and by $\mathscr{H}_m$ for R and U.

This remarks allows us to prove, without any computations, that the vertical boundary of V_{nm} belongs to the boundary of $l(H_n)$, so that we have the inclusion

$$\partial_v V_{nm} \subset \partial l(\mathscr{H}_n). \tag{63}$$

The proof that $l^{-1}(\partial_v V_{nm}) \supset \partial_v \mathscr{H}_n$ follows easily in the same way.

We have

$$
\begin{aligned}
&a_1 = \mu^{\kappa}\alpha^{-1}(2 + \alpha n v_1)^{-1},\\
&a_2 = 1 - |d^{-1}|\lambda^n (2 + \alpha n v_1)^{-1},\\
&a_3 = \alpha\mu^{\kappa}(2 + \alpha n v_1)^{-1}.
\end{aligned} \tag{64}
$$

We deduce that for μ sufficiently small, in order to satisfy conditions $(*)$ and $(*)_v$ of assumption (B), we must have

$$\frac{\mu^{2\kappa}}{2}(2+\alpha n v_1)^{-2} < \mu_\sigma < 1\mu^\kappa (2+\alpha n v_1)^{-2},$$

for $\sigma = \pm$. This inequality is satisfied, by definition of μ_+ and μ_-. A simple computation proves that conditions $(**)$ and $(**)_v$ are satisfied.

4.4. Control of the remainder

We denote by $\mathscr{I}_0$ the hyperbolic set obtained in the previous section for l. By Theorem 4.1, there exists a homeomorphism h_0 such that the following diagram commutes

$$\begin{array}{ccc} \mathscr{I}_0 & \mapsto^l & \mathscr{I}_0 \\ h_0 \downarrow & & \downarrow h_0 \\ \Sigma_{\mathscr{A}} & \mapsto^\sigma & \Sigma_{\mathscr{A}} \end{array}. \tag{65}$$

It remains to prove that the hyperbolic invariant set $\mathscr{I}_0$ persists for $\mathscr{L}$. The strategy, is first to estimate the remainder of $\mathscr{L}$. Using the Eliasson–Niederman normal form, we obtain for each horizontal slab $\mathscr{H}_n$ the following control lemma (see Lemma 4.2, Section 4.2.2 for notations).

Lemma 4.4 (Control lemma). *For each point $Z \in \mathscr{H}_n$, we have*

$$||R(Z)|| < C\lambda^{2n},\ ||DR(Z)|| < \hat{C}\mu\lambda^n, \tag{66}$$

where $C > 0$, $\hat{C} > 0$ are constant.

Proof. The first step is to control the remainder in the *normal form domain*. We have, using notations from Section 1.2, $\mathscr{F}(z) = (nv, 0, 0, 0) + F.z + r(z)$:

Lemma 4.5. *For $Z \in \mathscr{H}_n$, we have*

(i) $||r(\mathscr{W}(Z))|| < c\lambda^{2n}$,
(ii) $||Dr(\mathscr{W}(Z))|| < c\lambda^{2n}$.

Proof. This follows easily from the Eliasson–Niederman normal form. Indeed, the remainder depends only on the product $(su)^2$ and I^2. So, by definition of the alphabet $\mathscr{A}$, we deduce Lemma 4.5. □

The second step is to control the remainder when the trajectory leaves the domain of the normal form. We have

$$R(Z) = \Pi(r(W.Z + p^+)) + \Lambda_2(f(W.Z) - p^-)$$

by definition. Using assumption (h_3), and point (i) of Lemma 4.5, we deduce that $||R(Z)|| \leqslant C\lambda^{2n}$ for $Z \in \mathscr{H}_n$, where $C > 0$ is a constant. As,

$$DR(Z) = \Pi W(Dr(W.Z + p^+)) + W.f(W.Z).D\Lambda_2(f(W.Z) - p^-),$$

it follows by a simple computation and point (ii) of Lemma 4.5, that $||DR(Z)|| \leqslant \hat{C}\mu\lambda^n$, where $\hat{C} > 0$ is a constant. $\square$

In order to prove the persistence of I_0, we must prove that for each $n \in \mathscr{A}$, conditions (A) and (B) are satisfied. We remark that, as long as the map f is fixed in Section 4.1, conditions (A) and (B) are stable under small perturbations. For each $n \in \mathscr{A}$, l is fixed. In this case, we can compute the maximal size of the perturbation allowed in $\mathscr{H}_n$. By classical results on perturbations of hyperbolic map, we have that the mapping $\mathscr{L}$ satisfies conditions (A) and (B) if, for each $n \in \mathscr{A}$,

$$\sup_{z \in \mathscr{H}_n} (||R(Z)||, ||DR(Z)||) < \lambda^{n+\delta}$$

with $\delta > 0$.

By the control lemma, this is the case. Then, $\mathscr{L}$ satisfies conditions (A) and (B). This concludes the proof of the theorem.

Acknowledgments

I thank J-C. Yoccoz for a discussion about a previous version of the manuscript, leading to a better understanding of the geometry of this problem.

References

[1] V. Alekseev, Quasirandom dynamical systems, Math. USSR. Sb. 5 (1968) 73–128.

[2] V. Arnold, Instability of dynamical systems with several degrees of freedom, Soviet Math. Dokl. 5 (1966) 581–585.

[3] J. Bourgain, F. Golse, B. Wennberg, On the distribution of free path lengths for the periodic Lorentz gas, Comm. Math. Phys. 190 (1998) 491–508.

[4] L. Chierchia, G. Gallavotti, Drift and diffusion in phase space, Ann. Inst. H. Poincaré 60 (1) (1994) 1–144.

[5] J. Cresson, Instability properties of near-integrable Hamiltonian systems, Thesis, Observatoire de Paris, 1997, pp. 188.

[6] J. Cresson, Un λ-lemme pour des tores partiellement hyperboliques, C. R. Acad. Sci. Paris Sér. I 331 (2000) 65–70.

[7] J. Cresson, Periodic orbits and Arnold diffusion, Discrete Contin. Dynam. Systems 9 (2) (2003) 451–469.

[8] J. Cresson, C. Guillet, Periodic orbits and Arnold diffusion, preprint, 2000.

[9] J. Cresson, C. Guillet, Création d'hyperbolicité et phénomène de transversalité-torsion, preprint, 2001.

[10] R.W. Easton, Homoclinic phenomena in Hamiltonian systems with several degrees of freedom, J. Differential Equations 29 (1978) 241–252.

[11] R.W. Easton, Orbit structure near trajectories biasymptotic to invariant tori, in: R. Devaney, Z. Nitecki (Eds.), Classical Mechanics and Dynamical Systems, Marcel Dekker, New York, 1981, pp. 55–67.
[12] L.H. Eliasson, Biasymptotic solutions of perturbed integrable Hamiltonian systems, Bol. Soc. Brasil 25 (1) (1994) 57–76.
[13] S. Graff, On the conservation of hyperbolic tori for hamiltonian systems, J. Differential Equations 15 (1974) 1–69.
[14] P.J. Holmes, J.E. Marsden, Melnikov's method and Arnold diffusion for perturbations of integrable Hamiltonian systems, J. Math. Phys. 23 (4) (1982) 669–675.
[15] P. Lochak, J.P. Marco, D. Sauzin, On the splitting of invariant manifolds in multidimensional near-integrable Hamiltonian systems, manuscript, 1999.
[16] J-P. Marco, Transition le long des chaînes de tores invariants pour les systèmes hamiltoniens analytiques, Ann. Inst. H. Poincaré 64 (2) (1996) 205–252.
[17] J. Moser, Stable and Random Motions in Dynamical Systems, in: Annals of Mathematics Studies, Vol. 77, Princeton University Press, Princeton, NJ, 1973.
[18] L. Niederman, Dynamics around a chain of simple resonant tori in nearly integrable Hamiltonian systems, J. Differential Equation 161 (1) (2000) 1–41.
[19] J. Palis, W. De Melo, Geometric Theory of Dynamical Systems: An Introduction, Springer, Berlin, 1982.
[20] D. Treshchev, Hyperbolic tori and asymptotic surfaces in Hamiltonian systems, Russian J. Math. Phys. 2 (1994) 1.
[21] S. Wiggins, Global Bifurcations and Chaos, in: Applied Mathematical Science, Vol. 73, Springer, Berlin, 1988.
[22] J.C. Yoccoz, Introduction to hyperbolic dynamics, in: Real and Complex Dynamical Systems, (Hillerod, 1993), 265–291, NATO Adv. Sci. Inst. Ser. C. Math. Phys. Sci. 464, Kluwer Academic, Dordrecht, 1995.