

Computer Assisted Proof of Drift Orbits Along Normally Hyperbolic Manifolds

Maciej J. Capiński^{a,1,*}, Jorge Gonzalez^{b,2}, Jean-Pierre Marco^c, J.D. Mireles James^{d,3}

^a*Faculty of Applied Mathematics, AGH University of Science and Technology, al. Mickiewicza 30, 30-059 Kraków, Poland.*

^b*School of Mathematics, Georgia Institute of Technology, 686 Cherry Street, Atlanta, GA, 30332, USA.*

^c*Institut de Mathématiques, Analyse algébrique, Université Pierre et Marie Curie, 175 rue du Chevaleret, 75013 Paris, France.*

^d*Department of Mathematical Sciences, Florida Atlantic University, 777 Glades Road, Boca Raton, FL 33431, USA.*

Abstract

Normally hyperbolic invariant manifolds theory provides an efficient tool for proving diffusion in dynamical systems. In this paper we develop a methodology for computer assisted proofs of diffusion in a-priori chaotic systems based on this approach. We devise a method, which allows us to validate the needed conditions in a finite number of steps, which can be performed by a computer by means of rigorous-interval-arithmetic computations. We apply our method to the generalized standard map, obtaining diffusion over an explicit range of actions.

Keywords: Normally hyperbolic manifold, Arnold diffusion, scattering map, topological shadowing, computer assisted proof

2010 MSC: 37J25, 37J40

1. Introduction

One of the fundamental problems of classical mechanics is to understand the dynamics of perturbations of completely integrable systems. Indeed the solar system itself can be viewed as a system of weakly coupled (completely integrable in the Liouville sense) two body problems, and the question of its stability has captivated mathematicians since the days of Newton. To formalize the discussion let $\mathbb{T} = \mathbb{R} \bmod 2\pi$, let $\mathbb{A}^n = T^*\mathbb{T}^n = \mathbb{T}^n \times \mathbb{R}^n$ denote the annulus with the angle-action coordinates (θ, r) , endowed with the symplectic⁴ form $\omega = \sum_{i=1}^n dr_i \wedge d\theta_i$. Consider a Hamiltonian of the form

$$H(\theta, r) = h(r) + f(\theta, r), \quad (1)$$

*Corresponding author.

¹Partially supported by the NCN grant 2018/29/B/ST1/00109. The work has been conducted during the visit to FAU sponsored by the Fulbright Foundation.

²Partially supported by NSF grant MSPRF DMS-2001758.

³Partially supported by NSF grant DMS 1813501.

⁴We refer to [1] for a comprehensive presentation of symplectic geometry

where f is small in some suitable function space (analytic, C^∞ , C^k , etcetera). The Hamiltonian equations of motion generated by H are

$$\begin{aligned}\dot{\theta}_i &= \partial_{r_i} H(\theta, r) = \partial_{r_i} h(r) + \partial_{r_i} f(\theta, r) \\ \dot{r}_i &= -\partial_{\theta_i} H(\theta, r) = -\partial_{\theta_i} f(\theta, r).\end{aligned}$$

We say that the system is completely integrable when $f \equiv 0$, as all the orbits move with constant velocity on invariant tori.

When f is small, the evolution of the action variables r_i are “slow”. The fact that this evolution is “extremely slow” emerged from averaging methods originally developed by Lagrange and Laplace, furthered by Poincaré and Birkhoff, culminating in the work of Littlewood [2] and in the major achievements of Nekhoroshev [3]. Thanks to the work of these and many subsequent authors, it is now well-known that if h is strictly convex and analytic, then the drift in the action variables cannot exceed a variation of $O(\varepsilon^{1/2n})$ during an $O(\exp(1/\varepsilon)^{1/2(n-2)})$ -long time. Here ε measures the size of the perturbation function f .

Examining the problem from another direction, Kolmogorov [4] proved the first results on “perpetual stability” of solutions of analytic systems (1). Kolmogorov’s approach is geometric in essence: he proves that – provided that f is analytic, that the frequency vector $\nabla h(r^0)$ is Diophantine, and that certain non-degeneracy conditions on h hold – then the integrable invariant tori of the form $\mathbb{T}^n \times \{r^0\}$ persist and are only slightly deformed when the perturbation f is added to the system. By Diophantine we mean that there are constants $\gamma > 0$ and $\tau > 0$ such that

$$|k \cdot \nabla h(r^0)| \geq \frac{\gamma}{\|k\|^\tau}, \quad \forall k \in \mathbb{Z}^n \setminus \{0\}.$$

Arnold and Moser then added their own contributions to this initial result, giving rise to what is now known as the KAM theory [5, 6, 7, 8]. See also [9] for much more complete discussion of the KAM theory and its development.

Taken together, the KAM and averaging theories provide indispensable information about perturbations of integrable Hamiltonian systems. The KAM theory says that *some* orbits remain close to orbits of the unperturbed system for all time (the KAM tori), while the averaging theory says that *all* orbits stay close to the unperturbed level sets of the Hamiltonian for exponentially long times. A natural question is to ask *do there exist orbits whose Hamiltonian increases (or decreases) by an “arbitrarily” large amount, on a long enough time scale?*

The first explicit example actually exhibiting this phenomenon was given by Arnold in [10], and had the form:

$$H_\varepsilon(\theta, r) = r_0 + \frac{1}{2}(r_1^2 + r_2^2) + \mu \cos \theta_2 + \varepsilon g(\theta, r), \quad \theta \in \mathbb{T}^3, \quad r \in \mathbb{R}^3, \quad (2)$$

where g is an explicit fixed trigonometric polynomial, and μ and ε are independent parameters. The example has several important special properties, such as:

- When $\mu = \varepsilon = 0$, the system reduces to h and is completely integrable in angle-action form.
- When $\mu > 0$ and $\varepsilon = 0$, the system H_0 is Liouville-integrable. In particular, it admits a normally hyperbolic (and symplectic) invariant annulus $\mathcal{A}_0 = \mathbb{A}^2 \times \{O\}$, where $O = (0, 0) \in \mathbb{A}$ is the hyperbolic fixed point of the pendulum $\frac{1}{2}r_2^2 + \mu \cos \theta_2$. The stable and unstable manifolds of \mathcal{A}_0 take the form $W^\pm(\mathcal{A}_0) = \mathbb{A}^2 \times W^\pm(O)$. The Hamiltonian flow in restriction to \mathcal{A}_0 is completely integrable, in the sense that it admits a foliation by the Lagrangian (for the induced structure) invariant tori $(\mathbb{T}^2 \times \{(r_0, r_1)\})_{(r_0, r_1) \in \mathbb{R}^2}$.

- For fixed μ and small enough ε (ε has to be exponentially small w.r.t. μ in Arnold's example), the annulus \mathcal{A}_0 is only slightly deformed and gives rise to a 4-dimensional normally hyperbolic (symplectic) invariant annulus \mathcal{A}_ε close to \mathcal{A}_0 , with a rich homoclinic structure, while the Hamiltonian flow on \mathcal{A}_ε is close to completely integrable.

It is important to stress that the perturbation g is carefully chosen in Arnold's example, so that the annulus \mathcal{A}_0 is still invariant when $\varepsilon > 0$ and the dynamics on \mathcal{A}_0 remains unchanged. Moreover, the perturbation does not depend on the action variable r and the transversal intersection of stable and unstable manifolds can be proved for all invariant circles on \mathcal{A}_0 . By exploiting these facts Arnold was able to show that for $\mu, \varepsilon > 0$ small enough, H_ε admits a solution $\gamma_\varepsilon(t) = (\theta(t), r(t))$ which drifts of order 1 in action for suitable (very large) $T_{\mu,\varepsilon}$. That is

$$r_1(0) < 0, \quad r_1(T_{\mu,\varepsilon}) > 1,$$

for this orbit. This provided the first explicit example where orbits of the perturbed system "diffuse" as far and as fast⁵ from the unperturbed level sets as allowed by averaging theory.

The use of two independent parameters in Arnold's example simplifies a lot the study: (2) is to be compared with (1), where the size of f is the *only* available parameter. Nevertheless, Arnold's example became a jumping off point for a large body of work. By now this is a thriving industry and it is known that diffusion occurs under a wide variety of hypotheses.

Another (deeper) question raised by Arnold is the case where the unperturbed Hamiltonian is completely integrable and in action-angle form (the famous "fundamental problem of dynamics" of Poincaré). Given a Hamiltonian system h which depends only on the actions, does there exist a large (residual) set of perturbations g such that orbits diffuse in the previous fashion - or even visit any prescribed collection of open subsets of an energy level? It turns out that this question is extremely delicate, and there are still many important open problems in this active area of research. The present discussion is by no means intended as a literature review of the field, we refer to [15] for a very nice result in any dimension, together with relevant references.

Another line of study comes from weakening the hypothesis that the unperturbed system is completely integrable (in the Liouville sense). Consider for example systems of the form (2), in which the parameter μ is fixed but not small (say $\mu = 1$). Such systems are referred to as *a priori* unstable, since they already admit hyperbolic invariant objects when $\varepsilon = 0$. The main difficulty in studying *a priori* unstable systems is their "singular character" or lack of transversality, coming from the fact that the manifolds $W^\pm(\mathcal{A}_0)$ coincide when $\varepsilon = 0$. Detecting homoclinic intersections in such systems for generic g when $\varepsilon \neq 0$ is far from trivial and requires new ingredients from variational methods, weak KAM theory (both in the convex case) or symplectic topology in the general case.

This complication motivated the introduction of a still less degenerate class of examples, for which $W^\pm(\mathcal{A}_0)$ transversely intersect even in the case $\varepsilon = 0$. This class of systems is known as *a priori* chaotic, and is the main topic of the present work (see [16] and [17] for examples in this category closely related to Arnold's). Studying such systems is simpler, which leaves open the possibility of asking new and more quantitative questions, e.g. what is the threshold in ε under which diffusion phenomena can appear, or, what is the maximal length of diffusive trajectories? These questions require new methods, and it turns out that in realistic physical systems the

⁵The fact that the speed of Arnold diffusion coincides with the prediction of averaging theory was indeed proved much later, see [11, 12, 13, 14].

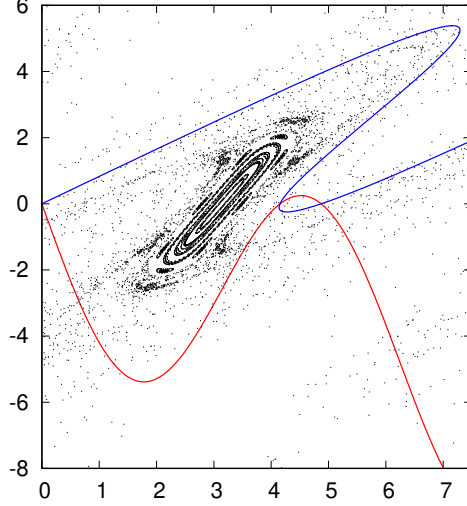


Figure 1: Phase space structure of the Chirikov Standard Map when $\alpha = 4$. Black dots indicate the dynamics of a number of “typical” orbits. The stable and unstable manifolds of the fixed point at the origin are depicted by the red and blue curves respectively.

relevant quantities to estimate are difficult to compute. Our aim is to provide an explicit example illustrating the relevance of computer-assisted methods of proof in such problems.

To simplify the construction we shift our focus to symplectic maps instead of Hamiltonian vector fields. This reduction is natural, since taking a Poincaré section in an energy manifold results in a symplectic diffeomorphism. The main example of the paper is the family of symplectic diffeomorphisms $f_\varepsilon: \mathbb{R}^2 \times \mathbb{T}^2 \rightarrow \mathbb{R}^2 \times \mathbb{T}^2$ defined by

$$f_\varepsilon(x, y, \theta, I) = \begin{pmatrix} x + y + \alpha \sin(x) \\ y + \alpha \sin(x) \\ \theta + I \\ I \end{pmatrix} + \varepsilon \begin{pmatrix} \cos(x) \sin(\theta) \\ \cos(x) \sin(\theta) \\ \sin(x) \cos(\theta) \\ \sin(x) \cos(\theta) \end{pmatrix}, \quad (3)$$

where $(x, y) \in \mathbb{R}^2$ and $(\theta, I) \in \mathbb{T}^2$. Observe that the map f_ε can be seen as a perturbation of a standard map (variables (x, y)) coupled to an I -parametrized rotation on \mathbb{T}^2 (variables (θ, I)). Indeed, when $\varepsilon = 0$ the two systems do not interact and the dynamics is a product.

Of particular interest, the standard map has a hyperbolic fixed point at the origin O in \mathbb{R}^2 . In the present work we do not treat α as a perturbation parameter, and will show that for fixed $\alpha = 4$ the stable and unstable manifolds intersect transversely at some point P ; see Figure 1 (so that the parameter α plays the role of μ in Arnold’s example). Consequently, f_0 admits an invariant torus $\{O\} \times \mathbb{T}^2$, which is readily seen to be normally hyperbolic, and whose stable and unstable manifolds intersect transversely along a homoclinic torus $\{P\} \times \mathbb{T}^2$. By the Birkhoff-Smale theorem, a large enough iterate of the standard map admits a horseshoe (homeomorphic to $\{0, 1\}^{\mathbb{Z}}$ endowed with the product topology) near the origin. Consequently, for N large enough, the coupling f_0^N admits a fibered horseshoe (hyperbolic lamination), close to $\{0\} \times \mathbb{T}^2$ and homeomorphic to $\{0, 1\}^{\mathbb{Z}} \times \mathbb{T}^2$, on which it induces a fiber-preserving dynamics. This problem was formalized

in [18, 19, 20].

When $\varepsilon > 0$, the preservation of the fibres is broken, and nothing prevents the orbits from drifting along the base \mathbb{T}^2 in the I direction. In this paper we use constructive computer assisted arguments to prove that such drift orbits do indeed exist for f_ε , and that they have lengths independent of the size ε of the perturbation. This makes the system a significant example in the *a-priori* chaotic case. Moreover, the present work provides a self contained exposition of constructive computer assisted methods for proving the existence of diffusion phenomena in explicit examples.

Our results are based on shadowing theorems for scattering maps worked out in [21]. A scattering map is a function from a normally hyperbolic invariant manifold to itself, defined through appropriate intersections of fibres of its stable and unstable manifolds. In [21] it is shown that pseudo orbits resulting from iterations of scattering maps are shadowed by true orbits of the system. We use this method in our main results, which are contained in Theorems 12, 19, 20 and 21. Theorems 12, 19, 20 establish orbits which diffuse over an explicit interval of actions. Theorem 21 establishes orbits which shadow sequences of actions, chosen from the interval. The aim of this paper is to provide tools which can be used to obtain computer assisted proofs. To check the hypotheses of our theorems one needs to compute the scattering maps of the unperturbed system, and to check certain explicit inequalities which measure the influence of the perturbation on the action. This influence is computed by considering finite fragments of homoclinic orbits. We apply our results to give a computer-assisted proof of diffusion for the system given by Equation (3). In a forthcoming paper we plan an application to the Planar Restricted Three Body Problem, with mass parameters of the Jupiter-Sun system.

An alternative approach for computer assisted proof of diffusion is given in [22]. The work just cited is based on the method of correctly aligned windows. The difference compared to the present work is that [22] requires an explicit construction of ‘connecting sequences’ of windows. These windows are then used for shadowing arguments. Here we establish transversal intersections of stable/unstable manifolds leading to scattering maps, and check our conditions along homoclinic orbits. The shadowing is automatically ensured by [21].

The remainder of the paper is organized as follows. In Section 2 we review some preliminary information about normally hyperbolic invariant manifolds, scattering maps, and the interval Newton method. In Section 3 we lay out our main theoretical results, namely the constructive hypothesis which are used to establish Arnold diffusion in explicit examples. Section 4 applies the method to the example system. Proofs of some of the theorems and lemmas are relegated to the Appendices.

2. Preliminaries

Throughout the paper, for $x \in \mathbb{R}^n$, by $\|x\|$ we shall mean the Euclidean norm. For a set A in a topological space we shall write \bar{A} to denote its closure.

2.1. Normally hyperbolic invariant manifolds

In this section we recall the notion of a normally hyperbolic invariant manifold and state the main result concerning its persistence under small perturbation. A classic reference for this material is [19, 23].

Definition 1. *Let M be a smooth n -dimensional Riemannian manifold, and let $f : M \rightarrow M$ be a C^r diffeomorphism, with $r > 1$. Let $\Lambda \subset M$ be a compact manifold without boundary,*

invariant under f , i.e., $f(\Lambda) = \Lambda$. We say that Λ is a normally hyperbolic invariant manifold (with symmetric rates) if there exist a constant $C > 0$, rates $0 < \lambda < \mu^{-1} < 1$ and a Tf invariant splitting for every $x \in \Lambda$

$$T_x M = E_x^u \oplus E_x^s \oplus T_x \Lambda,$$

such that

$$v \in E_x^u \Leftrightarrow \|Df^k(x)v\| \leq C\lambda^{-k} \|v\|, \quad k \leq 0, \quad (4)$$

$$v \in E_x^s \Leftrightarrow \|Df^k(x)v\| \leq C\lambda^k \|v\|, \quad k \geq 0, \quad (5)$$

$$v \in T_x \Lambda \Rightarrow \|Df^k(x)v\| \leq C\mu^{|k|} \|v\|, \quad k \in \mathbb{Z}. \quad (6)$$

Let $d(x, \Lambda)$ stand for the distance between a point x and the manifold Λ . Given a normally hyperbolic invariant manifold and a suitable small tubular neighbourhood $U \subset M$ of Λ one defines its local unstable and local stable manifold [19] as

$$\begin{aligned} W_\Lambda^u(f, U) &= \{y \in M \mid f^k(y) \in U, d(f^k(y), \Lambda) \leq C_y \lambda^{|k|}, k \leq 0\}, \\ W_\Lambda^s(f, U) &= \{y \in M \mid f^k(y) \in U, d(f^k(y), \Lambda) \leq C_y \lambda^k, k \geq 0\}, \end{aligned}$$

where C_y is a positive constant, which can depend on y . We define the (global) unstable and stable manifolds as

$$W_\Lambda^u(f) = \bigcup_{n \geq 0} f^n(W_\Lambda^u(f, U)), \quad W_\Lambda^s(f) = \bigcup_{n \geq 0} f^{-n}(W_\Lambda^s(f, U)).$$

The manifolds $W_\Lambda^u(f, U)$, $W_\Lambda^s(f, U)$, $W_\Lambda^u(f)$ and $W_\Lambda^s(f)$ are foliated by

$$\begin{aligned} W_x^u(f, U) &= \{y \in M \mid f^k(y) \in U, d(f^k(y), f^k(x)) \leq C_{x,y} \lambda^{|k|}, k \leq 0\}, \\ W_x^s(f, U) &= \{y \in M \mid f^k(y) \in U, d(f^k(y), f^k(x)) \leq C_{x,y} \lambda^k, k \geq 0\}, \end{aligned}$$

where $x \in \Lambda$ and $C_{x,y}$ is a positive constant, which can depend on x and y ,

$$W_x^u(f) = \bigcup_{n \geq 0} f^n(W_{f^{-n}(x)}^u(f, U)), \quad W_x^s(f) = \bigcup_{n \geq 0} f^{-n}(W_{f^n(x)}^s(f, U)).$$

Let

$$l < \min \left\{ r, \frac{|\log \lambda|}{\log \mu} \right\}. \quad (7)$$

The manifold Λ is C^l smooth, the manifolds $W_\Lambda^u(f)$, $W_\Lambda^s(f)$ are C^{l-1} and $W_x^u(f)$, $W_x^s(f)$ are C^r [24]. Normally hyperbolic manifolds, as well as their stable and unstable manifolds and their fibres persist under small perturbations [19].

2.2. Shadowing of scattering maps

Our diffusion result is based on shadowing lemmas for scattering maps found in [21], which we now summarize.

Let (M, ω) be a smooth symplectic manifold. Let us assume that $\Lambda \subset M$ is a normally hyperbolic invariant manifold for a C^r symplectic map $f : M \rightarrow M$, where $r > 1$. We assume

that Λ is symplectic with the symplectic form $\omega|_\Lambda$, then $f|_\Lambda$ is symplectic on Λ . We define two maps,

$$\begin{aligned}\Omega_+ &: W_\Lambda^s(f) \rightarrow \Lambda, \\ \Omega_- &: W_\Lambda^u(f) \rightarrow \Lambda,\end{aligned}$$

where $\Omega_+(x) = x_+$ iff $x \in W_{x_+}^s(f)$, and $\Omega_-(x) = x_-$ iff $x \in W_{x_-}^u(f)$. These are referred to as the wave maps.

Definition 2. We say that a manifold $\Gamma \subset W_\Lambda^u(f) \cap W_\Lambda^s(f)$ is a homoclinic channel for Λ if the following conditions hold:

(i) for every $x \in \Gamma$

$$T_x W_\Lambda^s(f) + T_x W_\Lambda^u(f) = T_x M, \quad (8)$$

$$T_x W_\Lambda^s(f) \cap T_x W_\Lambda^u(f) = T_x \Gamma, \quad (9)$$

(ii) the fibres of Λ intersect Γ transversally in the following sense

$$T_x \Gamma \oplus T_x W_{x_+}^s(f) = T_x W_\Lambda^s(f), \quad (10)$$

$$T_x \Gamma \oplus T_x W_{x_-}^u(f) = T_x W_\Lambda^u(f), \quad (11)$$

for every $x \in \Gamma$,

(iii) the wave maps $(\Omega_\pm)|_\Gamma : \Gamma \rightarrow \Lambda$ are diffeomorphisms onto their image.

Definition 3. Assume that Γ is a homoclinic channel for Λ and let

$$\Omega_\pm^\Gamma := (\Omega_\pm)|_\Gamma.$$

We define a scattering map σ^Γ for the homoclinic channel Γ as

$$\sigma^\Gamma := \Omega_+^\Gamma \circ (\Omega_-^\Gamma)^{-1} : \Omega_-^\Gamma(\Gamma) \rightarrow \Omega_+^\Gamma(\Gamma).$$

We have the following symplectic property of the scattering map.

Theorem 4. [21] Assume that M is endowed with a symplectic form ω and that $\omega|_\Lambda$ is also symplectic. Assume that f is symplectic. Assume that there exists a homoclinic channel Γ and so the scattering map σ^Γ is well defined. Then, the scattering map σ^Γ is symplectic.

We have the following theorem, which is the main tool which we use to obtain our results.

Theorem 5. [21] Assume that $f : M \rightarrow M$ is a sufficiently smooth map, $\Lambda \subset M$ is a normally hyperbolic invariant manifold with stable and unstable manifolds which intersect transversally along a homoclinic channel $\Gamma \subset M$, and σ is the scattering map associated to Γ .

Let $m_1, \dots, m_\ell \in \mathbb{N}$ be a fixed sequence of integers. Let $\{x_i\}_{i=0, \dots, \ell}$ be a finite pseudo-orbit in Λ , that is a sequence of points in Λ of the form

$$x_{i+1} = f^{m_i} \circ \sigma^\Gamma(x_i), \quad i = 0, \dots, \ell - 1, \ell \geq 1. \quad (12)$$

Then for every $\delta > 0$ there exists an orbit $\{z_i\}_{i=0, \dots, \ell}$ of f in M , with $z_{i+1} = f^{k_i}(z_i)$ for some $k_i > 0$, such that $d(z_i, x_i) < \delta$ for all $i = 0, \dots, \ell$.

Remark 6. The original statement of Theorem 5 from [21] does not require compactness of Λ . The assumptions are that f preserves measure absolutely continuous with respect to the Lebesgue measure on Λ , and that σ sends positive measure sets to positive measure sets. Moreover, it is assumed that the pseudo-orbit is contained in some open set $\mathcal{U} \subset \Lambda$ with almost every point of \mathcal{U} recurrent for $f|_{\Lambda}$. In our case, since Λ is compact and $f|_{\Lambda}$ is symplectic, almost every point in Λ is recurrent for $f|_{\Lambda}$, so we can take $\mathcal{U} = \Lambda$ and simplify the statement of the theorem.

Remark 7. In [21] the statement of the theorem is for pseudo-orbits of the form $x_{i+1} = \sigma^{\Gamma}(x_i)$. Here we shadow pseudo-orbits of the form (12), but this is the same result as that from [21] for the following reason.

The proof of the theorem in [21] is based on a general shadowing lemma [21, Lemma 3.1] which ensures that given a pseudo-orbit of the form $y_{i+1} = f^{k_i} \circ \sigma^{\Gamma} \circ f^{n_i}(y_i)$ where the numbers of iterates k_i, n_i are big enough, we are able to find an orbit of the form $z_{i+1} = f^{k_i+n_i}(z_i)$, δ -close to the pseudo-orbit y_i .

The shadowing of a pseudo-orbit $x_{i+1} = \sigma^{\Gamma}(x_i)$ is proven in [21] by combining [21, Lemma 3.1] with recurrence. First, by using recurrence, a pseudo-orbit of the form $y_{i+1} = f^{k_i} \circ \sigma^{\Gamma} \circ f^{n_i}(y_i)$ is constructed close to the pseudo-orbit $x_{i+1} = \sigma^{\Gamma}(x_i)$. The k_i, n_i are chosen to be big enough to apply from [21, Lemma 3.11]. The true orbit, which follows from [21, Lemma 3.11], shadows the pseudo-orbit y_i , but since this lies close to x_i one obtains the shadowing of the pseudo-orbit $x_{i+1} = \sigma^{\Gamma}(x_i)$.

The proof of the shadowing of a pseudo-orbit of the form (12) follows from the same construction: One can use recurrence to construct a pseudo-orbit of the form $y_{i+1} = f^{k_i} \circ \sigma^{\Gamma} \circ f^{n_i}(y_i)$, so that y_i are close to x_i from (12). The lemma [21, Lemma 3.1] ensures that y_i can be shadowed by a true orbit. Since y_i is close to the pseudo-orbit x_i from (12) we obtain the shadowing of (12) by a true orbit.

Remark 8. The result can be immediately extended to the case where we have a finite number of scattering maps $\sigma_1, \dots, \sigma_L$ to shadow

$$x_{i+1} = f^{m_i} \circ \sigma_{\alpha_i}(x_i), \quad i = 0, \dots, \ell - 1, \ell \geq 1,$$

for two prescribed sequences $m_1, \dots, m_{\ell} \in \mathbb{N}$ and $\alpha_1, \dots, \alpha_{\ell} \in \{1, \dots, L\}$; see [21, Theorem 3.7].

3. Main results

Let $f_{\varepsilon} : \mathbb{R}^{2d} \times \mathbb{T}^2 \rightarrow \mathbb{R}^{2d} \times \mathbb{T}^2$ be a family of smooth maps, which depends smoothly on the parameter ε and is of the form

$$f_{\varepsilon}(u, s, I, \theta) = f_0(u, s, I, \theta) + \varepsilon g(u, s, I, \theta), \quad (13)$$

where $u, s \in \mathbb{R}^d$, $\theta, I \in \mathbb{T}$ and where $g = f_{\varepsilon} - f$; here, and throughout the rest of the paper in similar instances, the subtraction is done in the covering space of $\mathbb{R}^{2d} \times \mathbb{T}^2$.

Assume that f_{ε} are smooth symplectic maps for a symplectic form $\omega = du \wedge ds + dI \wedge d\theta$, assume that for $\varepsilon = 0$

$$\Lambda_0 = \{(0, 0, I, \theta) : I, \theta \in \mathbb{T}^1\} \simeq \mathbb{T}^2,$$

is a normally hyperbolic invariant manifold, and that I is a constant of motion for the unperturbed system, i.e.

$$\pi_I f_0(x) = \pi_I x, \quad (14)$$

for any $x \in \mathbb{R}^{2d} \times \mathbb{T}^2$, where $\pi_I(u, s, I, \theta) = I$.

Remark 9. We assume smoothness of the maps since our main tool for the proof will be Theorem 5, which requires sufficient smoothness.

Remark 10. The assumption that Λ_0 is a torus simplifies the arguments, as Λ_0 is compact without boundary and the normally hyperbolic manifold theorem ensures that Λ_0 is perturbed to a nearby compact normally hyperbolic invariant manifold Λ_ε . Having compactness of Λ_ε is convenient, but not necessary if one applies non-compact versions of the normally hyperbolic theorem [25, 26, 27, 28].

Our objective is to provide conditions under which for any sufficiently small $\varepsilon > 0$ there exist a point x_ε and a number of iterates n_ε for which

$$\pi_I(f_\varepsilon^{n_\varepsilon}(x_\varepsilon) - x_\varepsilon) > 1. \quad (15)$$

The coordinates have the following roles. The u, s are the coordinates on unstable and stable bundles, respectively, of Λ_0 . The θ is an angle and I plays the role of a constant of motion for $\varepsilon = 0$. In the setting of action-angle coordinates, the I would be chosen as the action. We shall refer to I as an ‘action’, slightly abusing the terminology. In this paper we restrict to the case where the angle and action are one dimensional each. We do so for the sake of simplicity⁶.

A typical setting where our result applies is that of a time dependent perturbation of a Hamiltonian system of the form

$$x' = J\nabla_x H_\varepsilon(x, t) = J\nabla_x(H(x) + \varepsilon G(x, t)), \quad (16)$$

where $H : \mathbb{R}^{2d} \times \mathbb{T}^2 \rightarrow \mathbb{R}$, $G : \mathbb{R}^{2d} \times \mathbb{T}^2 \times \mathbb{T} \rightarrow \mathbb{R}$ and

$$J = \begin{pmatrix} 0 & Id \\ -Id & 0 \end{pmatrix}, \quad \text{for} \quad Id = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In such case we can take $f_\varepsilon(x) = \Phi_{2\pi}^\varepsilon(x, t_0)$, for some $t_0 \in [0, 2\pi)$, where $\Phi_t^\varepsilon(x_0, t_0)$ stands for the time t flow induced by (16) with the initial condition (x_0, t_0) .

Another possibility is to consider the flow induced by (16) in the extended phase space and consider a section of the form $\Sigma \times \mathbb{T}^1$ in $\mathbb{R}^{2d+2} \times \mathbb{T}^1$. Then f_ε can then be chosen as the section-to-section map along the flow in the extended phases space. The time coordinate plays the role of the angle θ and we choose I as the Hamiltonian H of the unperturbed system.

Remark 11. A typical setting is when Λ_0 is a normally hyperbolic invariant cylinder (possibly with a boundary) with $\theta \in \mathbb{T}^1$ and $I \in \mathbb{R}$. We would then have $f_\varepsilon : \mathbb{R}^{2d} \times \mathbb{R} \times \mathbb{T} \rightarrow \mathbb{R}^{2d} \times \mathbb{R} \times \mathbb{T}$. In such case we can consider $I \in [0, 1]$ and artificially ‘glue’ the system so that I is in \mathbb{T}^1 to apply our result directly, provided that certain conditions are met. Details of how this can be done are found in Appendix A.

The next theorem is our first main result. It provides conditions for the existence of orbits which diffuse in I .

⁶We believe that our methods can be generalised to the higher dimensional case. We make comments how to do so in Remarks 16, 22 after the statements of our results.

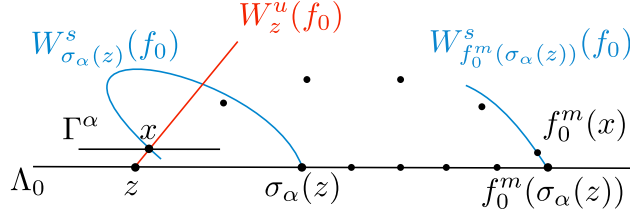


Figure 2: The setting for Theorem 12.

Theorem 12. Assume that there is a neighborhood U of Λ_0 and a positive constant L_g such that for every $z \in \Lambda_0$ and every $x_u \in W_z^u(f_0, U)$, $x_s \in W_z^s(f_0, U)$,

$$\begin{aligned} |\pi_I(g(0, x_u) - g(0, z))| &\leq L_g \|x_u - z\|, \\ |\pi_I(g(0, x_s) - g(0, z))| &\leq L_g \|x_s - z\|. \end{aligned} \quad (17)$$

Assume also that there exist positive constants C, λ , such that $\lambda \in (0, 1)$ and for every $z \in \Lambda_0$ and every $x_u \in W_z^u(f_0, U)$, $x_s \in W_z^s(f_0, U)$ we have

$$\begin{aligned} \|f_0^n(z) - f_0^n(x_u)\| &< C\lambda^{|n|} && \text{for all } n \leq 0, \\ \|f_0^n(z) - f_0^n(x_s)\| &< C\lambda^n && \text{for all } n \geq 0. \end{aligned} \quad (18)$$

Assume that for $\varepsilon = 0$ we have a sequence $\Gamma^1, \dots, \Gamma^L \subset U$ of homoclinic channels for f_0 , with corresponding wave maps $\Omega_\pm^\alpha : \Gamma^\alpha \rightarrow \Lambda_0$ and scattering maps $\sigma_\alpha : \text{dom}(\sigma_\alpha) \rightarrow \Lambda_0$ for $\alpha = 1, \dots, L$.

Assume that for every $z \in \Lambda_0$

1. There exists an $\alpha \in \{1, \dots, L\}$ such that $z \in \text{dom}(\sigma_\alpha)$.
2. There exists an $m \in \mathbb{N}$ and a point $x \in \Gamma^\alpha$, $x \in W_z^u(f_0, U) \cap W_{\sigma_\alpha(z)}^s(f_0)$ such that $f_0^m(x) \in W_{f_0^m(\sigma_\alpha(z))}^s(f_0, U)$ (see Figure 2) and

$$\sum_{j=0}^{m-1} \pi_I g(f_0^j(x)) - \frac{1+\lambda}{1-\lambda} L_g C > 0. \quad (19)$$

(The above α, m and x can depend on the choice of z .)

Then for sufficiently small $\varepsilon > 0$ there exist an x_ε and $n_\varepsilon > 0$ such that

$$\pi_I(f_\varepsilon^{n_\varepsilon}(x_\varepsilon) - x_\varepsilon) > 1. \quad (20)$$

Before giving the proof let us make a couple of comments about the assumptions.

Remark 13. Assumption (17) will readily hold since Λ_0 is compact, so we can take \bar{U} to be compact as well and existence of L_g follows from the fact that g is smooth. Conditions (18) will hold due to the contraction and expansion properties along the stable and unstable manifolds. What is important for us is to have explicit bounds L_g, C and λ which enter into the key assumption (19).

Remark 14. Condition (19) measures the influence of the perturbation term g on the coordinate I . This can be thought of as a discrete analogue of the following formula for the perturbed scattering map from [21]

$$\sigma_\varepsilon = \sigma_0 + \varepsilon J \nabla S \circ \sigma_0 + O(\varepsilon^2),$$

where

$$S(x) = \lim_{T \rightarrow +\infty} \int_{-T}^0 \left[\frac{dH_\varepsilon}{d\varepsilon} \Big|_{\varepsilon=0} \circ \Phi_t \circ (\Omega_-^{\Gamma_0})^{-1} \circ \sigma_0^{-1}(x) - \frac{dH_\varepsilon}{d\varepsilon} \Big|_{\varepsilon=0} \circ \Phi_t \circ \sigma_0^{-1}(x) \right] dt \\ + \lim_{T \rightarrow +\infty} \int_0^T \left[\frac{dH_\varepsilon}{d\varepsilon} \Big|_{\varepsilon=0} \circ \Phi_t \circ (\Omega_+^{\Gamma_0})^{-1}(x) - \frac{dH_\varepsilon}{d\varepsilon} \Big|_{\varepsilon=0} \circ \Phi_t(x) \right] dt,$$

and where Φ_t is the flow of the unperturbed system. Instead of computing an integral along a homoclinic orbit of the flow, we compute a sum along a discrete orbit. An important feature is that we are computing it along a finite fragment of the homoclinic. The second term in (19) takes into account the truncated tail. We believe that for computer assisted proofs computing a bound for a sum for a discrete dynamical system, which comes from considering a time shift map or Poincaré map, is more convenient than computing a bound on an integral over a trajectory of a flow.

Remark 15. In Theorem 12 we assume that the homoclinic channels are in U , meaning that they are close to Λ_0 . This is not a restrictive assumption, since a homoclinic channel which is far away can be propagated close to Λ_0 by using backward iterates of f_0 .

Remark 16. Theorem 12 can be generalised to the setting of higher dimensional θ and I as follows. If we have actions I_1, \dots, I_k , we can single out one of them (say $I = I_1$) for the conditions (17) and (19), to obtain diffusion towards the singled out action.

Remark 17. We have assumed that $f_\varepsilon(x) = f_0(x) + \varepsilon g(x)$. We can assume just as well that $f_\varepsilon(x) = f_0(x) + \varepsilon g(\varepsilon, x)$, with smooth $g(\varepsilon, x)$. Then in conditions (17) and (19) we can write $g(0, \cdot)$ instead of $g(\cdot)$, and the result will follow from the same arguments. Analogous modifications can be made also in subsequent theorems. We consider $g(x)$ instead of $g(\varepsilon, x)$ since it simplifies and shortens the notation.

Proof of Theorem 12. The manifold Λ_0 is perturbed to a normally hyperbolic invariant manifold Λ_ε for f_ε . Moreover, for sufficiently small ε , if $z \in \Lambda_\varepsilon$, $x_u \in W_z^u(f_\varepsilon, U)$ and $x_s \in W_z^s(f_\varepsilon, U)$, then

$$\begin{aligned} \left\| f_\varepsilon^n(z) - f_\varepsilon^n(x_u) \right\| &< C \lambda_\varepsilon^{|n|} && \text{for all } n \leq 0, \\ \left\| f_\varepsilon^n(z) - f_\varepsilon^n(x_s) \right\| &< C \lambda_\varepsilon^n && \text{for all } n \geq 0, \end{aligned} \tag{21}$$

with λ_ε converging to λ as ε tends to zero.

Since transversal intersections persist under perturbation, the homoclinic channels $\Gamma^1, \dots, \Gamma^L$ for f_0 are perturbed to homoclinic channels $\Gamma_\varepsilon^1, \dots, \Gamma_\varepsilon^L$ for f_ε , provided that $\varepsilon > 0$ is sufficiently small. This leads [24] to a scattering map $\sigma_\alpha^\varepsilon : \Omega_-^{\Gamma_\varepsilon^\alpha} \rightarrow \Omega_+^{\Gamma_\varepsilon^\alpha}$ for f_ε .

Our first objective is to show that for any $z_\varepsilon \in \Lambda_\varepsilon$ there exist an $m \in \mathbb{N}$ and $\alpha \in \{0, \dots, L\}$ (both m and α can depend on z_ε) such that

$$\pi_I(f_\varepsilon^m \circ \sigma_\alpha^\varepsilon(z_\varepsilon) - z_\varepsilon) > \varepsilon C, \tag{22}$$

where $c > 0$ is a constant, small enough so that we have

$$\sum_{j=0}^{m-1} \pi_I g(f_0^j(x)) - \frac{1+\lambda}{1-\lambda} L_g C > c, \quad (23)$$

for any $z \in \Lambda_0$ (with the same c). We can find such small c because of (19) and compactness of Λ_0 .

It turns out that (22) is the main step in our proof, since once it is established the result follows from the shadowing Theorem 5. Below we first prove (22) and then discuss how to apply the shadowing method.

Consider now a $z_\varepsilon \in \Lambda_\varepsilon$. By our assumptions, for every $z \in \Lambda_0$ we have an $\alpha \in \{1, \dots, L\}$, $m \in \mathbb{N}$ and $x \in W_z^u(f_0, U) \cap W_{\sigma_\alpha(z)}^s(f_0)$ such that $f_0^m(x) \in W_{f_0^m(\sigma_\alpha(z))}^s(f_0, U)$ and (23) holds. This means that for sufficiently small ε we shall have an $x_\varepsilon \in W_{z_\varepsilon}^u(f_\varepsilon, U)$ such that $f_\varepsilon^m(x_\varepsilon) \in W_{f_\varepsilon^m(\sigma_\alpha^{\varepsilon}(z_\varepsilon))}^s(f_\varepsilon, U)$, and by (21)

$$\|f_\varepsilon^j(z_\varepsilon) - f_\varepsilon^j(x_\varepsilon)\| < C\lambda_\varepsilon^{|j|} \quad \text{for } j \leq 0 \quad (24)$$

$$\|f_\varepsilon^{m+j}(\sigma_\alpha^{\varepsilon}(z_\varepsilon)) - f_\varepsilon^{m+j}(x_\varepsilon)\| < C\lambda_\varepsilon^j \quad \text{for } j \geq 0. \quad (25)$$

Due to (23) and the continuous dependence of $x_\varepsilon, \lambda_\varepsilon$ on ε , for sufficiently small ε we shall have

$$\sum_{j=0}^{m-1} \pi_I g(f_\varepsilon^j(x_\varepsilon)) - \frac{1+\lambda_\varepsilon}{1-\lambda_\varepsilon} L_g C > c.$$

In order to show (22) we will split our estimates into three terms

$$f_\varepsilon^m(\sigma_\alpha^{\varepsilon}(z_\varepsilon)) - z_\varepsilon = [f_\varepsilon^m(\sigma_\alpha^{\varepsilon}(z_\varepsilon)) - f_\varepsilon^m(x_\varepsilon)] + [f_\varepsilon^m(x_\varepsilon) - x_\varepsilon] + [x_\varepsilon - z_\varepsilon], \quad (26)$$

and investigate bounds on the projection π_I for each of them. We start by showing that

$$|\pi_I [f_\varepsilon^m(\sigma_\alpha^{\varepsilon}(z_\varepsilon)) - f_\varepsilon^m(x_\varepsilon)]| \leq \varepsilon \frac{1}{1-\lambda_\varepsilon} L_g C. \quad (27)$$

Indeed, since $f_\varepsilon(x) = f_0(x) + \varepsilon g(x)$ and $\pi_I f_0(x) = \pi_I x$, for any x_1, x_2 we have

$$\begin{aligned} \pi_I f_\varepsilon(x_1) - \pi_I f_\varepsilon(x_2) &= \pi_I f_0(x_1) + \varepsilon \pi_I g(x_1) - \pi_I f_0(x_2) - \varepsilon \pi_I g(x_2) \\ &= \pi_I(x_1 - x_2) + \varepsilon \pi_I(g(x_1) - g(x_2)). \end{aligned}$$

It follows by induction that

$$\pi_I (f_\varepsilon^j(x_1) - f_\varepsilon^j(x_2)) = \pi_I [x_1 - x_2] + \varepsilon \sum_{i=0}^{j-1} \pi_I (g(f_\varepsilon^i(x_1)) - g(f_\varepsilon^i(x_2))). \quad (28)$$

Taking $x_1 = f_\varepsilon^m(\sigma_\alpha^{\varepsilon}(z_\varepsilon))$ and $x_2 = f_\varepsilon^m(x_\varepsilon)$ from (28) with $\pi_I [x_1 - x_2]$ moved to the left hand

side, we have

$$\begin{aligned}
& \left| \pi_I [f_\varepsilon^m(\sigma_\alpha^\varepsilon(z_\varepsilon)) - f_\varepsilon^m(x_\varepsilon)] \right| \\
&= \left| \pi_I (f_\varepsilon^{m+j}(\sigma_\alpha^\varepsilon(z_\varepsilon)) - f_\varepsilon^{m+j}(x_\varepsilon)) - \varepsilon \sum_{i=0}^{j-1} \pi_I (g(f_\varepsilon^{m+i}(\sigma_\alpha^\varepsilon(z_\varepsilon))) - g(f_\varepsilon^{m+i}(x_\varepsilon))) \right| \\
&< C\lambda_\varepsilon^j + \varepsilon L_g \sum_{i=0}^{j-1} \|f_\varepsilon^{m+i}(\sigma_\alpha^\varepsilon(z_\varepsilon)) - f_\varepsilon^{m+i}(x_\varepsilon)\| \\
&< C\lambda_\varepsilon^j + \varepsilon L_g \sum_{i=0}^{j-1} C\lambda_\varepsilon^i,
\end{aligned}$$

where the last two inequalities follow from (25). Letting $j \rightarrow \infty$, we obtain (27).

Now consider the third term from (26). An analogous bound to (27) is obtained as follows. From (28) we have that

$$\begin{aligned}
\pi_I(x_1 - x_2) &= \pi_I [f_\varepsilon^{-j}(x_1) - f_\varepsilon^{-j}(x_2)] + \varepsilon \sum_{i=0}^{j-1} \pi_I (g(f_\varepsilon^{i-j}(x_1)) - g(f_\varepsilon^{i-j}(x_2))) \\
&= \pi_I [f_\varepsilon^{-j}(x_1) - f_\varepsilon^{-j}(x_2)] + \varepsilon \sum_{i=-j}^{-1} \pi_I (g(f_\varepsilon^i(x_1)) - g(f_\varepsilon^i(x_2))). \tag{29}
\end{aligned}$$

Taking $x_1 = x_\varepsilon$ and $x_2 = z_\varepsilon$, from (29) we obtain

$$\begin{aligned}
|\pi_I(x_\varepsilon - z_\varepsilon)| &\leq \left| \pi_I [f_\varepsilon^{-j}(x_\varepsilon) - f_\varepsilon^{-j}(z_\varepsilon)] \right| + \varepsilon \sum_{i=-j}^{-1} \left| \pi_I (g(f_\varepsilon^i(x_\varepsilon)) - g(f_\varepsilon^i(z_\varepsilon))) \right| \\
&< C\lambda_\varepsilon^j + \varepsilon L_g \sum_{i=-j}^{-1} \|f_\varepsilon^i(x_\varepsilon) - f_\varepsilon^i(z_\varepsilon)\| \\
&< C\lambda_\varepsilon^j + \varepsilon L_g \sum_{i=1}^j C\lambda_\varepsilon^i,
\end{aligned}$$

where the last two inequalities follow from (24). Taking $j \rightarrow \infty$ gives

$$|\pi_I(x_\varepsilon - z_\varepsilon)| \leq \varepsilon \frac{\lambda_\varepsilon}{1 - \lambda_\varepsilon} CL_g. \tag{30}$$

We now turn to the middle term from (26). Since $f_\varepsilon(x) = f_0(x) + \varepsilon g(x)$ and $\pi_I f_0(x) = \pi_I x$, it follows that (below we consider $x = f_\varepsilon^j(x_\varepsilon)$)

$$\begin{aligned}
\pi_I (f_\varepsilon(f_\varepsilon^j(x_\varepsilon)) - f_\varepsilon^j(x_\varepsilon)) &= \pi_I f_0(f_\varepsilon^j(x_\varepsilon)) + \varepsilon \pi_I g(f_\varepsilon^j(x_\varepsilon)) - \pi_I f_\varepsilon^j(x_\varepsilon) \\
&= \varepsilon \pi_I g(f_\varepsilon^j(x_\varepsilon)),
\end{aligned}$$

so

$$\pi_I (f_\varepsilon^m(x_\varepsilon) - x_\varepsilon) = \sum_{j=0}^{m-1} \pi_I (f_\varepsilon^{j+1}(x_\varepsilon) - f_\varepsilon^j(x_\varepsilon)) = \varepsilon \sum_{j=0}^{m-1} \pi_I g(f_\varepsilon^j(x_\varepsilon)). \tag{31}$$

Combining (26), (27), (30), (31) gives

$$\pi_I(f_\varepsilon^m(\sigma_\alpha^\varepsilon(z_\varepsilon)) - z_\varepsilon) > \varepsilon \left(\sum_{j=0}^{m-1} \pi_I g(f_\varepsilon^j(x_\varepsilon)) - \frac{1 + \lambda_\varepsilon}{1 - \lambda_\varepsilon} CL_g \right).$$

Since the right hand side of the inequality above depends continuously on ε , from (23) we obtain (22) for sufficiently small ε .

This establishes the key step (22). We now apply Theorem 5 to prove our result. Indeed, since $\omega|_{\Lambda_0}$ is nondegenerate, the same is true for $\omega|_{\Lambda_\varepsilon}$ for sufficiently small ε . Choose $x_0 \in \Lambda_\varepsilon$ having $\pi_I x_0 = 0$ and consider α_0, m_0 , (which are allowed to depend on x_0) such that for

$$x_1 := f_\varepsilon^{m_0} \circ \sigma_{\alpha_0}^\varepsilon(x_0),$$

we have $\pi_I(f_\varepsilon^{m_0} \circ \sigma_{\alpha_0}^\varepsilon(x_0) - x_0) > c\varepsilon$. This can be done due to (22). Repeating the procedure, choosing α_i, m_i for which $(\pi_I f_\varepsilon^{m_i} \circ \sigma_{\alpha_i}^\varepsilon(x_i) - x_i) > c\varepsilon$ we obtain a pseudo-orbit x_0, \dots, x_N , where $x_{i+1} := f_\varepsilon^{m_i} \circ \sigma_{\alpha_i}^\varepsilon(x_i)$, for which

$$\pi_I(x_N - x_0) > Nc\varepsilon.$$

Choosing N large enough, we obtain that $\pi_I(x_N - x_0) > 1$. By Theorem 5 the pseudo-orbit x_0, \dots, x_N is δ -shadowed by a true orbit, so by choosing

$$\delta < \frac{1}{2}(\pi_I(x_N - x_0) - 1),$$

we have the claim. ■

In Theorem 12 we assume that for any point in Λ_0 we can find a pseudo-orbit such that we have a gain in I . Note however that we do not need to have (19) for all $z \in \Lambda_0$. It is enough to have (19) for z on some smaller subset of Λ_0 , provided that we can ensure that the pseudo-orbit constructed in the proof of Theorem 12 returns to that set. Below we formulate Theorem 19, which will make this statement precise. First we introduce one notion.

Definition 18. Consider the topology on $\Lambda_0 \cap \{I \in [0, 1]\}$ induced by Λ_0 . We say that an open set $S \subset \Lambda_0 \cap \{I \in [0, 1]\}$ is a strip in Λ_0 iff

$$S \cap \{z \in \Lambda_0 : \pi_I z = \iota\} \neq \emptyset \quad \text{for any } \iota \in [0, 1].$$

(Recall that we consider $\mathbb{T} = \mathbb{R} \bmod 2\pi$; the interval $I \in [0, 1]$ is a strict subset of $[0, 2\pi)$. Since S is open in the topology induced on $\Lambda_0 \cap \{I \in [0, 1]\}$ we require that it contains points with $I = 0$ and $I = 1$.)

We refer to S as a ‘strip’ because usually we would choose it to be of the shape as in the left hand side of Figure 3. In principle though a strip might look differently, for instance as on the right hand side plot in figure 3.

In subsequent two theorems we consider two strips S^+ and S^- . The strip S^+ is used to validate diffusion in I , which increases I by order one. The strip S^- will be used to prove diffusion in which I decreases by order one.

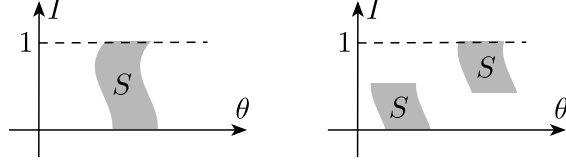


Figure 3: A typically shaped strip (left) and a ‘strip’ consisting of two connected components (right).

Theorem 19. Assume that conditions (17) and (18) are satisfied, and that for $\varepsilon = 0$ we have the sequence of scattering maps $\sigma_\alpha : \text{dom}(\sigma_\alpha) \rightarrow \Lambda_0$ for $\alpha = 1, \dots, L$. Let $S^+ \subset \Lambda_0$ be a strip⁷. Assume that for every $z \in \overline{S^+}$:

1. there exist a constant $m \in \mathbb{N}$ and an $\alpha \in \{1, \dots, L\}$ for which $z \in \text{dom}(\sigma_\alpha)$ and

$$f_0^m \circ \sigma_\alpha(z) \in S^+, \quad (32)$$

2. there exists a point $x \in W_z^u(f_0, U) \cap W_{\sigma_\alpha(z)}^s(f_0)$ such that $f_0^m(x) \in W_{f_0^m(\sigma_\alpha(z))}^s(f_0, U)$ and

$$\sum_{j=0}^{m-1} \pi_I g(f_0^j(x)) - \frac{1+\lambda}{1-\lambda} L_g C > 0. \quad (33)$$

(The m, α and x can depend on z .) Then for sufficiently small $\varepsilon > 0$ there exist an x_ε and $n_\varepsilon > 0$ such that

$$\pi_I(f_\varepsilon^{n_\varepsilon}(x_\varepsilon) - x_\varepsilon) > 1. \quad (34)$$

Moreover, we can choose x_ε to be arbitrarily close to S^+ .

Proof. The result follows by making minor adjustments to the arguments in the proof of Theorem 12. Let $S_\varepsilon^+ \subset \Lambda_\varepsilon$ be the perturbation of the strip $S^+ \subset \Lambda_0$. As in the proof of Theorem 12 we construct a pseudo orbit $x_{i+1} = f_\varepsilon^{m_i} \circ \sigma_{\alpha_i}^\varepsilon(x_i)$, starting with a point $x_0 \in S_\varepsilon^+$ with $\pi_I x_0 = 0$. Note that we assume that (32) holds for any $z \in \overline{S^+}$ (with choices of m and α depending on z). This means that for sufficiently small ε , and for any point $z_\varepsilon \in S_\varepsilon^+$, there is an $m = m(z_\varepsilon)$, $\alpha = \alpha(z_\varepsilon)$ such that $f_\varepsilon^{m(z_\varepsilon)} \circ \sigma_{\alpha(z_\varepsilon)}^\varepsilon(z_\varepsilon) \in S_\varepsilon^+$. In other words, z_ε ‘returns’ to the strip for sufficiently small ε . Due to the compactness of $\overline{S^+}$, a sufficiently small choice of ε guarantees that we have $f_\varepsilon^{m(z_\varepsilon)} \circ \sigma_{\alpha(z_\varepsilon)}^\varepsilon(z_\varepsilon) \in S_\varepsilon^+$ for all $z_\varepsilon \in S_\varepsilon^+$. In short, condition (32) ensures that the pseudo-orbit $x_{i+1} = f_\varepsilon^{m_i} \circ \sigma_{\alpha_i}^\varepsilon(x_i)$ remains within the strip S_ε^+ for sufficiently small ε . By (33) and identical arguments to those from Theorem 12 we therefore have

$$\pi_I(x_{i+1} - x_i) > \varepsilon c,$$

for some $c > 0$, and the result follows from the shadowing argument, by applying Theorem 5, just as in the proof of Theorem 12. Theorem 5 ensures that we can δ -shadow the pseudo-trajectory $\{x_i\}$, which means that the true trajectory can start from x_ε , which is δ -close to $x_0 \in S^+$, for any $\delta > 0$ of our choice. ■

A mirror result gives diffusion in the opposite direction.

⁷We add the plus in the superscript for S^+ since this strip is used to increase I . In subsequent theorem we will have another strip S^- to obtain diffusion in the opposite direction.

Theorem 20. Assume that conditions (17) and (18) are satisfied, and that for $\varepsilon = 0$ we have the sequence of scattering maps $\sigma_\alpha : \text{dom}(\sigma_\alpha) \rightarrow \Lambda_0$ for $\alpha = 1, \dots, L$. Let $S^- \subset \Lambda_0$ be a strip. Assume that for every $z \in \overline{S^-}$:

1. there exist a constant $m \in \mathbb{N}$ and an $\alpha \in \{1, \dots, L\}$ for which $z \in \text{dom}(\sigma_\alpha)$ and

$$f_0^m \circ \sigma_\alpha(z) \in S^-,$$

2. there exist a point $x \in W_z^u(f_0, U) \cap W_{\sigma_\alpha(z)}^s(f_0)$ such that $f_0^m(x) \in W_{f_0^m(\sigma_\alpha(z))}^s(f_0, U)$ and

$$\sum_{j=0}^{m-1} \pi_I g(f_0^j(x)) + \frac{1+\lambda}{1-\lambda} L_g C < 0.$$

(The m, α and x can depend on z .) Then for sufficiently small $\varepsilon > 0$ there exist an x_ε and $n_\varepsilon > 0$ such that

$$\pi_I(x_\varepsilon - f_\varepsilon^{n_\varepsilon}(x_\varepsilon)) > 1. \quad (35)$$

Moreover, we can choose x_ε to be arbitrarily close to S^- .

Proof. The proof follows as in the proof of Theorem 19. ■

By combining the two strips we obtain shadowing of any prescribed finite sequence of actions.

Theorem 21. Assume that two strips S^+ and S^- satisfy assumptions of Theorems 19 and 20, respectively. If in addition

1. for every $z \in \overline{S^+}$ there exists an n (which can depend on z) such that $f_0^n(z) \in S^-$, and
2. for every $z \in \overline{S^-}$ there exists an n (which can depend on z) such that $f_0^n(z) \in S^+$,

then there exists an $M > 0$, such that for any given finite sequence $\{I_k\}_{k=0}^N$, and for sufficiently small ε there exists an orbit of f_ε which εM -shadows the actions I_k ; i.e. there exists a point x_ε and a sequence of integers $n_1^\varepsilon \leq n_2^\varepsilon \leq \dots \leq n_N^\varepsilon$ such that

$$\left\| \pi_I f_\varepsilon^{n_k^\varepsilon}(x_\varepsilon) - I_k \right\| < \varepsilon M. \quad (36)$$

Proof. Suppose that $I_1 > I_0$. (The opposite case will be analogous.) As in the proof of Theorem 19, we construct a pseudo orbit $x_{i+1} = f_\varepsilon^{m_i} \circ \sigma_{\alpha_i}^\varepsilon(x_i)$, $x_i \in S_\varepsilon^+$, starting with a point x_0 with $\pi_I x_0 = I_0$, such that

$$\pi_I(x_{i+1} - x_i) > \varepsilon c,$$

for some $c > 0$. By smooth dependence of f_ε and $\sigma_{\alpha_i}^\varepsilon$ on ε , and by the compactness of the normally hyperbolic manifold, we can choose large enough M so that for every such x_i we have

$$|\pi_I(x_{i+1} - x_i)| < \varepsilon M/2. \quad (37)$$

(The M can be chosen to be independent from x_i .) We can therefore find a pseudo orbit for which $|\pi_I x_{i_1} - I_1| < \varepsilon M/2$, for some $i_1 > 0$. If $I_2 > I_1$, and we carry on as in the proof of Theorem 19, continuing with our pseudo-orbit along S_ε^+ , until we reach x_{i_2} such that $|\pi_I x_{i_2} - I_2| < \varepsilon M/2$. If on the other hand $I_2 < I_1$, then we take $x_{i_1+1} = f_\varepsilon^n(x_{m_{i_1}})$, where the n is the number from assumption 1. (for $z = x_{i_1}$). For sufficiently small ε we will obtain that $x_{i_1+1} \in S_\varepsilon^-$. We now

construct the subsequent points x_i along the strip S_ε^- , going down in I along each step, until we reach x_{i_2} satisfying $|\pi_I x_{i_2} - I_2| < \varepsilon M/2$. (Possibly we might need to enlarge M , so that we ensure (37) for points from the strip S_ε^- .) Depending on whether $I_{k+1} > I_k$ or $I_{k+1} < I_k$ we proceed in an analogous manner: to move up in I we construct the given fragment of the pseudo-orbit along S_ε^+ ; and to go down in I we construct the given fragment of the pseudo-orbit along S_ε^- . Assumptions 1., 2. ensure that our pseudo-orbit can be chosen to jump between the strips S_ε^+ and S_ε^- at any stage of the construction.

This way we construct a pseudo orbit for which

$$|\pi_I x_{i_k} - I_k| < \varepsilon M/2 \quad \text{for } k = 0, \dots, N.$$

By Theorem 5 the pseudo-orbit x_i can be δ -shadowed by a true orbit. For fixed ε , we can choose $\delta = \varepsilon M/2$, which concludes our proof. ■

Remark 22. Theorems 19, 20, 21 can be generalised to the setting of higher dimensional I by singling out one action, as in Remark 16. The definition of the strip is then with respect to that particular action.

4. Example of application

In this section we discuss our example, the generalized standard map, to which we apply our method. We give a computer assisted proof of the existence of diffusing orbits by applying Theorem 21. We validate the assumptions of the theorem using two independent implementations, which use different methods to obtain bounds on the stable/unstable manifolds of the NHIM. The first is based on cone conditions [29, 30, 31], and the second on the parameterization method [32, 33, 34]. The stable/unstable manifolds calculations are discussed in more detail in Section Appendix B.

4.1. The Generalized Standard Map

Let $V(q)$ be a $(2\pi\mathbb{Z})^n$ -periodic function. Consider a map $f : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ given by

$$f(q, p) = (q + p + \nabla V(q), p + \nabla V(q)).$$

Remark 23. The map f is symplectic and has the generating function

$$S(q, Q) = \frac{1}{2} \|Q - q\|^2 + V(q).$$

Remark 24. When $V = 0$ the map is completely integrable. When $n = 1$ and $V(q) = \alpha \cos(q)$ then we obtain the Chirikov Standard Map.

For our example, taking $q = (x, \theta)$, $p = (y, I)$ and

$$V_\varepsilon(x, \theta) = \alpha \cos(x) - \varepsilon \sin(x) \sin(\theta),$$

we obtain a family of maps (3). To be in line with the setup from section 3 we interpret that $f_\varepsilon : \mathbb{R}^2 \times \mathbb{T}^2 \rightarrow \mathbb{R}^2 \times \mathbb{T}^2$. (We could just as well interpret f_ε to be on \mathbb{T}^4 .)

In our example we take $\alpha = 4$. For this parameter, when $\varepsilon = 0$, on the x, y coordinates we have a hyperbolic fixed point at the origin. The reader can get a sense of the dynamics by referring to the simulation results illustrated in Figure 1.

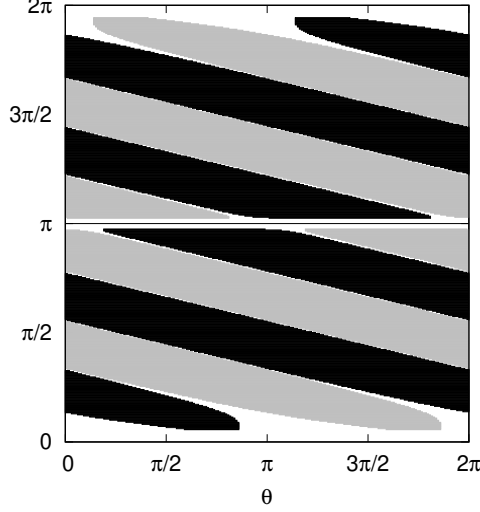


Figure 4: The strips from Theorem 21 for the map (3), validated by our computer program. The S^+ is in black and S^- in grey. The angle θ is on the horizontal axis and I on the vertical axis.

For $\varepsilon = 0$ the system consists of a pair of decoupled maps $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ and $G : \mathbb{T}^2 \rightarrow \mathbb{T}^2$

$$f_0(x, y, I, \theta) = (F(x, y), G(I, \theta)). \quad (38)$$

The origin on the x, y plane is a hyperbolic fixed point of F and $DF(0)$ has eigenvalues λ, λ^{-1} for $\lambda = 3 - 2\sqrt{2}$ (here we took $\alpha = 4$).

The torus

$$\Lambda_0 = \{(0, 0, I, \theta) : \theta \in \mathbb{T}^1, I \in \mathbb{T}^1\},$$

is a normally hyperbolic invariant manifold for f_0 with the rates λ and $\mu = (\sqrt{5}/2 + 1)/2$. (The μ is the norm of the matrix acting on I, θ in (3) for $\varepsilon = 0$.)

We consider the standard symplectic form

$$\omega = dx \wedge dy + dI \wedge d\theta.$$

The maps f_ε are ω -symplectic and $\omega|_{\Lambda_0}$ is non-degenerate.

We prove the following result.

Theorem 25 (Diffusion in the generalized standard map). *There exists an $M > 0$ such that for every finite sequence $\{I_l\}_{l=0}^L \subset [\frac{1}{5}, \pi - \frac{1}{10}]$ and for every sufficiently small $\varepsilon > 0$, there exist a sequence of integers $n_1^\varepsilon, \dots, n_L^\varepsilon$, a point z_0^ε , and a sequence of points $z_l^\varepsilon := f_\varepsilon^{n_l^\varepsilon}(z_{l-1}^\varepsilon)$ for $l = 1, \dots, L$, such that*

$$|\pi_I z_l^\varepsilon - I_l| < \varepsilon M, \quad \text{for } l = 0, \dots, L.$$

Remark 26. *The proof of this theorem is based on computer assisted validation of the assumptions of Theorem 21. The strips validated by our computer program are depicted in Figure 4.*

Remark 27. From our validation of the strips (see Figure 4) it follows also that we can take the interval $[\pi + \frac{1}{10}, 2\pi - \frac{1}{5}]$ instead of $[\frac{1}{5}, \pi - \frac{1}{10}]$ in Theorem 25. Between these two intervals though, at $I = \pi$ and $I = 0$, we have gaps, which our method is unable to overcome. In other words, we are not able to establish an orbit which would start with $I \in (0, \pi)$ and finish with $I \in (\pi, 2\pi)$ (and vice versa).

Remark 28. The diffusion is in fact established for intervals reaching in I slightly closer to 0 and π than stated in Theorem 25, where we have rounded down the intervals. Our computer assisted proof based on the parameterization method does a better job and produces higher (in I) strips than the method based on cone conditions. This is because the parametrization method leads to much higher accuracy of the bounds on the stable/unstable manifolds, which is then reflected in better accuracy of the remaining computations. Both methods though can be used to validate the I -intervals stated in Theorem 25 and Remark 27.

Remark 29. If we take the parameter α in (3) closer to zero, then the unstable eigenvalues at the origin becomes smaller and the problem becomes more challenging numerically. This is because with weak hyperbolicity it is more difficult to obtain good estimates on the manifolds; also the homoclinic excursion takes more iterates. We have found that close to $\alpha = 0.15$ the method based on cone conditions fails at establishing the bounds for the intersection of the stable/unstable manifolds, but the parametrization method can still be applied.

4.2. Interval Newton Method

In our computer assisted proofs we use the following classical result, which allows one to conclude from the existence of a “good enough” approximate solution that there exists a true solution to a nonlinear system of equations.

By an interval matrix $\mathbf{A} \subset \mathbb{R}^{k \times k}$, we mean a matrix whose elements are intervals. Let $\mathcal{F} : \mathbb{R}^k \rightarrow \mathbb{R}^k$ be a C^1 function and $U \subset \mathbb{R}^k$. We shall denote by $[D\mathcal{F}(U)]$ the interval enclosure of a Jacobian matrix on the set U . This means that $[D\mathcal{F}(U)]$ is an interval matrix defined as

$$[D\mathcal{F}(U)] = \left\{ A \in \mathbb{R}^{k \times k} \mid A_{ij} \in \left[\inf_{x \in U} \frac{d\mathcal{F}_i}{dx_j}(x), \sup_{x \in U} \frac{d\mathcal{F}_i}{dx_j}(x) \right] \text{ for } i, j = 1, \dots, k \right\}.$$

Let $\mathbf{A} \subset \mathbb{R}^{k \times k}$ be an interval matrix. We shall write \mathbf{A}^{-1} to denote an interval matrix, for which $M \in \mathbf{A}^{-1}$, if $M^{-1} \in \mathbf{A}$.

Theorem 30. [35] (Interval Newton method) Let $\mathcal{F} : \mathbb{R}^k \rightarrow \mathbb{R}^k$ be a C^1 function and $X = \prod_{i=1}^k [a_i, b_i]$ with $a_i < b_i$. If $[D\mathcal{F}(X)]$ is invertible and there exists an x_0 in X such that

$$N(x_0, X) := x_0 - [D\mathcal{F}(X)]^{-1} f(x_0) \subset X,$$

then there exists a unique point $x^* \in X$ such that $\mathcal{F}(x^*) = 0$.

4.3. Proof of Theorem 25

The proof of Theorem 25 exploits computer assisted validation methods for studying the local stable/unstable manifolds of fixed points. We apply these for the map F from (38), i.e. the unperturbed map acting on x, y . We take the origin as our fixed point of F . The methods allow us to obtain an open interval $J \subset \mathbb{R}$ and smooth functions $P_u : J \rightarrow \mathbb{R}^2$ and $P_s : J \rightarrow \mathbb{R}^2$ such that $P_u(J)$ is the local unstable manifold $W_0^u(F, U)$ of the origin for F , and $P_s(J)$ is the

local stable manifold $W_0^s(F, U)$ of the origin for F , for some neighbourhood U of the origin. We give a description of both methods in section [Appendix B](#). For the purpose of this section it is enough that we can obtain explicit bounds for such functions, as well as for their first derivatives. Moreover, the methods allow us to obtain explicit bounds $C, \lambda \in \mathbb{R}, C, \lambda > 0$ such that

$$\begin{aligned} \|F^i(P_s(x))\| &\leq C\lambda^i \\ \|F^{-i}(P_u(x))\| &\leq C\lambda^i \end{aligned} \quad \text{for all } i \in \mathbb{N} \text{ and } x \in J. \quad (39)$$

The functions P_u and P_s give only a local description of the unstable and stable manifolds. To establish their intersections we use the following parallel shooting approach. We consider an integer M such that for some $x \in J$ the point $F^M(P_u(x))$ returns back to the neighbourhood of the fixed point. Such M is chosen by performing numerical experiments. We then define $\mathcal{F} : \bar{J} \times B_1 \times \dots \times B_M \times \bar{J} \rightarrow \mathbb{R}^{2M+2}$, where $B_i \subset \mathbb{R}^2$ are cartesian products of two closed intervals, as

$$\begin{aligned} \mathcal{F}(x, v_0, \dots, v_{M-1}, y) \\ := (P_u(x) - v_0, F(v_0) - v_1, \dots, F(v_{M-2}) - v_{M-1}, F(v_{M-1}) - P_s(y)). \end{aligned}$$

If we establish the existence of a point $p^* = (x^*, v_0^*, \dots, v_{M-1}^*, y^*)$ for which

$$\mathcal{F}(p^*) = 0, \quad (40)$$

then we have established a sequence of points v_0^*, \dots, v_M^* , where $v_1^* = P_u(x^*)$ and $v_M^* = P_s(y)$, along a homoclinic orbit to zero. The bound on the solution of (40) can be established by using the interval Newton theorem⁸; see section 4.2. This way, we obtain a homoclinic orbit within a set of the form

$$v_i^* \in [x_i^0 - r, x_i^0 + r] \times [y_i^0 - r, y_i^0 + r] \quad \text{for } i = 0, \dots, M, \quad (41)$$

where x_i^0, y_i^0 are written in Table 1. (Our M is equal to 10.)

We use two methods to obtain bounds on P_u and P_s . In the case of the first method, by using cones, we obtain

$$r = r_{\text{cones}} = 1.5 \cdot 10^{-7}, \quad (42)$$

and by using the second method, i.e. the parameterization method, we obtain

$$r = r_{\text{param}} = 6.5 \cdot 10^{-15}. \quad (43)$$

(The bounds on our computer program are in fact often tighter and vary from point to point. Here we have rounded them up to write a uniform enclosure r for all considered points.)

Since we use the interval Newton method as the tool for our validation we also obtain transversality of obtained intersection of our manifolds. (Such results are well known, see for instance [36] for a similar approach. We add the proof in the appendix to keep the work self-contained.)

⁸An alternative could be to use the Newton-Krawczyk theorem or a version of the Newton-Kantorovich theorem. We use the interval Newton theorem because of its simplicity and the fact that it is sufficient for our needs in this particular example.

Table 1: Homoclinic orbit

i	x_i^0	y_i^0
0	0.003855589164542	0.003194074612644
1	0.022471982225036	0.018616393060494
2	0.130968738959384	0.108496756734347
3	0.761844080808229	0.630875341848845
4	4.153747139236954	3.391903058428725
5	4.153747139236954	0.000000000000001
6	0.761844080808229	-3.391903058428725
7	0.130968738959384	-0.630875341848845
8	0.022471982225036	-0.108496756734347
9	0.003855589164542	-0.018616393060494
10	0.000661514551898	-0.003194074612644

Lemma 31. *The manifolds $W_0^u(F)$ and $W_0^s(F)$ intersect transversally.*

Proof. The proof is given in [Appendix C](#). ■

Define the sequence

$$(x_i^*, y_i^*) := F^i(v_0^*) \quad \text{for all } i \in \mathbb{Z}.$$

Note that $(x_i^*, y_i^*) = v_i^*$, for $i = 0, \dots, M$. We now show that for $\varepsilon = 0$ the map (3) has a well defined homoclinic channel with a global scattering map.

Lemma 32. *The set*

$$\Gamma = \{(x_0^*, y_0^*, I, \theta) : I, \theta \in \mathbb{T}^1\},$$

is a homoclinic channel for f_0 and the associated scattering map σ is globally defined and is the identity on Λ_0 .

Proof. To show that Γ is a homoclinic channel for f_0 we need to prove points (i), (ii) and (iii) from Definition 2.

We start by observing that for $p \in \Gamma$

$$T_p\Gamma = \{(0, 0)\} \times \mathbb{R}^2. \quad (44)$$

Since $W_0^u(F)$, $W_0^s(F)$ intersect transversally in \mathbb{R}^2 at v_0^* we also have

$$T_{v_0^*}W_0^s(F) \oplus T_{v_0^*}W_0^u(F) = \mathbb{R}^2, \quad (45)$$

$$T_{v_0^*}W_0^s(F) \cap T_{v_0^*}W_0^u(F) = \{0\}. \quad (46)$$

Since $W_\Lambda^u(f_0) = W_0^u(F) \times \mathbb{T}^2$ and $W_\Lambda^s(f_0) = W_0^s(F) \times \mathbb{T}^2$ we see that for $p \in \Gamma$

$$T_pW_\Lambda^u(f_0) = T_{v_0^*}W_0^u(F) \times \mathbb{R}^2, \quad (47)$$

$$T_pW_\Lambda^s(f_0) = T_{v_0^*}W_0^s(F) \times \mathbb{R}^2. \quad (48)$$

From (45), (47), (48) and (46), (47), (48), (44) we obtain, respectively,

$$\begin{aligned} T_p W_\Lambda^s(f_0) + T_p W_\Lambda^u(f_0) &= \mathbb{R}^4, \\ T_p W_\Lambda^s(f_0) \cap T_p W_\Lambda^u(f_0) &= \{(0, 0)\} \times \mathbb{R}^2 = T_p \Gamma, \end{aligned}$$

which proves (i) from Definition 2.

Since any two points that converge to each other need to start with the same values on I, θ we see that for any $z \in \Lambda_0$

$$W_z^u(f_0) = W_0^u(F) \times \{\pi_{(I, \theta)} z\}, \quad (49)$$

and

$$W_z^s(f_0) = W_0^s(F) \times \{\pi_{(I, \theta)} z\}. \quad (50)$$

This means that the wave maps are of the form

$$\Omega_\pm(x, y, I, \theta) = (0, 0, I, \theta). \quad (51)$$

Clearly $(\Omega_\pm)|_\Gamma$ are diffeomorphisms as required in (iii) from Definition 2.

From (49), (50) we see that for any $p \in \Gamma$ and $z \in \Lambda_0$

$$T_p W_z^u(f_0) = T_{v_0^*} W_0^u(F) \times \{(0, 0)\}, \quad (52)$$

and

$$T_p W_z^s(f_0) = T_{v_0^*} W_0^s(F) \times \{(0, 0)\}. \quad (53)$$

Combining (44) with (52), (53) and comparing with (47), (48) gives

$$T_p \Gamma \oplus T_p W_z^u(f_0) = T_{v_0^*} W_0^u(F) \times \mathbb{R}^2 = T_p W_\Lambda^u(F), \quad (54)$$

and

$$T_p \Gamma \oplus T_p W_z^s(f_0) = T_{v_0^*} W_0^s(F) \times \mathbb{R}^2 = T_p W_\Lambda^s(F), \quad (55)$$

which means that we have (ii) from Definition 2. We have established that Γ is a homoclinic channel. From (51) we see that the associated scattering map σ is globally defined and is the identity on Λ_0 . ■

We validate the strips S^+ and S^- with the shapes as in Figure 4. These are composed of small overlapping rectangular fragments. Below we introduce a lemma which we then apply on each such rectangular part. First we introduce a notation. For $a, b \in [0, 2\pi)$ we define the interval $[a, b] \subset \mathbb{T} = \mathbb{R} \bmod 2\pi$ as

$$[a, b] = \begin{cases} \{x \in \mathbb{T} : a \leq x \leq b\} & \text{if } a \leq b, \\ \{x \in \mathbb{R} : b \leq x \leq a + 2\pi\} \bmod 2\pi & \text{if } b < a. \end{cases} \quad (56)$$

We define $(a, b) \subset \mathbb{T}$ as the interior of $[a, b]$.

Let $I_1, I_2 \in (0, 2\pi)$ satisfy $I_1 < I_2$. Let $s_1, s_2 \in \mathbb{T}$, and consider strips on Λ_0 of the form

$$\{(0, 0)\} \times [I_1, I_2] \times [s_1, s_2]. \quad (57)$$

(In (57) the interval $[s_1, s_2]$ is in the sense (56).) We now have the following lemma.

Lemma 33. *If*

$$\sum_{i=0}^{M-1} \sin(x_i^*) \cos(\theta + iI) > 3 \frac{1+\lambda}{1-\lambda} C, \quad (58)$$

and if for every $(I, \theta) \in [I_1, I_2] \times [s_1, s_2]$ there exists an $m \geq M$ (the m can depend on the choice of (I, θ)) such that

$$\theta + mI \in (s_1, s_2), \quad (59)$$

then assumptions of Theorem 19 hold true for our map (3) on the strip (57).

Proof. Condition (32) follows from (59). We need to validate (33). Since $v_M^* \in P_s(J)$, from (39) it follows that $|x_m^*| < C\lambda^{m-M}$, for $m \geq M$.

Consider an arbitrary fixed $(I, \theta) \in [I_1, I_2] \times [s_1, s_2]$ and let

$$C_m := \sum_{j=0}^{m-1} \sin(x_j^*) \cos(\theta + jI).$$

Since for $j \geq M$ we know that $|x_j^*| < C\lambda^{j-M}$, we see that for $m \geq M$

$$|C_m - C_M| \leq \sum_{j=M}^{m-1} |\sin(x_j^*)| |\cos(\theta + jI)| \leq C \frac{1 - \lambda^{m-M}}{1 - \lambda} < C \frac{1 + \lambda}{1 - \lambda}. \quad (60)$$

Observe that the map $(x, y, I, \theta) \rightarrow \sin(x) \cos(\theta)$ is Lipschitz with the constant $L_g = 2$.

For $z = (0, 0, I, \theta) \in \{(0, 0)\} \times [I_1, I_2] \times [s_1, s_2]$, consider $x = (x_0^*, y_0^*, I, \theta) \in W_z^u(f_0, U) \cap W_{\sigma_a(z)}^s(f_0)$. Since $(x_0^*, y_0^*) = v_0^* \in P_u(J)$ and $v_M^* \in P_s(J)$, for every $m \geq N$, $f_0^m(x) \in W_{f_0^m(\sigma_a(z))}^s(f_0, U)$. Also, for every $m \geq M$, by using (58) and (60), we obtain

$$\begin{aligned} \sum_{j=0}^{m-1} \pi_I g(f_0^j(x)) - \frac{1+\lambda}{1-\lambda} L_g C &= \sum_{j=0}^{m-1} \sin(x_j^*) \cos(\theta + jI) - 2 \frac{1+\lambda}{1-\lambda} C \\ &\geq C_M - |C_m - C_M| - 2 \frac{1+\lambda}{1-\lambda} C \\ &\geq C_M - 3 \frac{1+\lambda}{1-\lambda} C \\ &> 0, \end{aligned}$$

which ensures (33). This finishes our proof. ■

Remark 34. *A mirror result lets us validate assumptions of Theorem 20. The only difference is that instead of (58), we require*

$$\sum_{i=0}^{N-1} \sin(x_i^*) \cos(\theta + iI) < -3 \frac{1+\lambda}{1-\lambda} C.$$

We are now ready to prove Theorem 25.

Proof of Theorem 25. By Lemma 31 the stable and unstable manifolds of the origin for the map F intersect transversally. Moreover, we have explicit bounds for a homoclinic orbit along

this intersection, written in Table 1 and (41–43). This means that, by Lemma 32, the scattering map for the unperturbed system is well defined.

Using the bounds from Table 1 and (41–43), which give an enclosure of a finite fragment of the homoclinic orbit, and together with the aid of Lemma 33, our computer program constructs the strip S^+ from Figure 4. This is done as follows.

Both the parameterization method [37] and the method based on cones from Appendix B allow us to obtain explicit bounds on the λ and C for which we have (18). These constants are needed for the validation of (58). In the case of the cone method from Appendix B we use Lemma 40 combined with Lemmas 41, 42 and obtain

$$\begin{aligned} C &= 0.000915553935545, \\ \lambda &= 0.171573780396. \end{aligned}$$

(See Appendix B for a more detailed discussion.) These constants come out from our computer assisted validation. The parametrization method gives similar, marginally different constants.

In our computer assisted proof we choose $M = 10$, which corresponds to the length of the fragment of the homoclinic orbit from Table 1.

The computer validation of the strip S^+ from Figure 4 is obtained by subdividing the interval $\mathbf{I} := \left[\frac{1}{10}, 2\pi - \frac{1}{10}\right]$ into $\kappa_I := 400$ closed intervals of equal length, which we denote as $\mathbf{I}_1, \dots, \mathbf{I}_{\kappa_I}$; and subdividing the interval $[0, 2\pi]$ into $\kappa_\theta := 200$ closed intervals of equal length, which we denote as $\theta_1, \dots, \theta_{\kappa_\theta}$. Our computer program checks whether condition (58) holds for $(I, \theta) \in \mathbf{I}_i \times \theta_j$ for each $i \in \{1, \dots, \kappa_I\}$ and $j \in \{1, \dots, \kappa_\theta\}$. For the computation of the sum in (58) we use our computer assisted bound for the fragment of the homoclinic orbit from Table 1 obtained in (41–43). If (58) does not hold true for a given i, j , then the box $\mathbf{I}_i \times \theta_j$ is disregarded. (It does not constitute a validated fragment of the strip S^+ .) The second step is to check (59). For a given $i \in \{1, \dots, \kappa_\theta\}$ and $j \in \{1, \dots, \kappa_I\}$ the m needed for (59) with $(I, \theta) \in \mathbf{I}_i \times \theta_j$ is found automatically by our program. (The m can depend on the choice of i, j .) In the case that the program fails and can not find an m for which the condition (59) holds, the box $\mathbf{I}_i \times \theta_j$ is also disregarded. If a box $\mathbf{I}_i \times \theta_j$ passes both the validation of (58) and (59), then by Lemma 33 it is a subset of S^+ . By performing these checks for $i \in \{1, \dots, \kappa_I\}$ and $j \in \{1, \dots, \kappa_\theta\}$ we obtain the computer validated strip S^+ from Figure 4.

We use a mirror method (see Remark 34) to construct the strip S^- from Figure 4, for which assumptions of Theorem 20 are satisfied. We also validate that for these two strips conditions 1. and 2. of Theorem 21 are fulfilled.

After such validation the result follows from Theorem 21. ■

The computer assisted proof using cone conditions for the validation of intersections of the manifolds was performed with the CAPD⁹ library [38]. The parameterization method approach was implemented in Matlab. The source code is available on the web page of the corresponding author.

Acknowledgements

We would like to thank the anonymous Reviewers for their comments, suggestions and corrections, which helped us improve our paper.

⁹Computer Assisted Proofs in Dynamics: <http://capd.ii.uj.edu.pl>

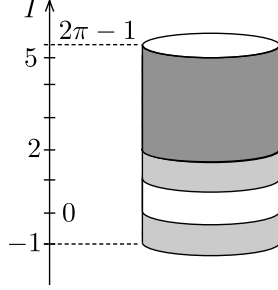


Figure A.5: For $I \in [0, 1]$ the system is not modified (white area). In the light grey regions the system is modified by the ‘bump’ function. For $I \in [2, 2\pi - 1]$ we ‘freeze’ the system to be the unperturbed map, which is represented by the dark grey area.

Appendix A. Modification of a system with a normally hyperbolic invariant cylinder to one with a normally hyperbolic invariant torus

Consider a family of maps $f_\varepsilon : \mathbb{R}^{2d} \times \mathbb{R} \times \mathbb{T}^1 \rightarrow \mathbb{R}^{2d} \times \mathbb{R} \times \mathbb{T}^1$ to be a parameter dependent family of the 2π time shift along the trajectory maps (or section-to-section maps), as described in section 3, for a Hamiltonian system

$$x' = J\nabla_x(H(x) + \varepsilon G(x, t)),$$

where $H : \mathbb{R}^{2d+2} \rightarrow \mathbb{R}$ and $G : \mathbb{R}^{2d+2} \times \mathbb{T} \rightarrow \mathbb{R}$. We consider $I(x) = H(x)$ as the preserved quantity for $\varepsilon = 0$, and our coordinates are $x = (u, s, I, \theta) \in \mathbb{R}^{2d} \times \mathbb{R} \times \mathbb{T}^1$. We assume that for $f_{\varepsilon=0}$ the manifold $\Lambda_0 = \{0\} \times \mathbb{R} \times \mathbb{T}$ is an invariant cylinder.

We will modify the system so that we obtain a map \tilde{f}_ε defined on $\mathbb{R}^{2d} \times \mathbb{T}^2$ for which

$$f_\varepsilon(u, s, I, \theta) = \tilde{f}_\varepsilon(u, s, I, \theta) \quad \text{for } I \in [0, 1].$$

We start by explaining the idea, which is depicted in Figure A.5. For $I \in [0, 1]$ we leave the system as it is. We then employ a ‘bump’ function so that at the edges of the domain $I \in [-1, 2]$, i.e. for $I = 2$ and $I = -1 = 2\pi - 1$, we have $\tilde{f}_\varepsilon = f_0$. For the remaining $I \in (2, 2\pi - 1)$ we ‘freeze’ the system taking $\tilde{f}_\varepsilon = f_0$.

In detail, we consider a smooth ‘bump’ function¹⁰ $b : \mathbb{R} \rightarrow [0, 1]$ for which

$$\begin{aligned} b(I) &= 0 & \text{for } I \in \mathbb{R} \setminus (-1, 2), \\ b(I) &= 1 & \text{for } I \in [0, 1], \end{aligned}$$

and take \tilde{f}_ε to be the 2π time shift maps (or section-to-section maps) for a modified ODE

$$x' = J\nabla_x(H(x) + b(H(x))\varepsilon G(x, t)).$$

Such modification can allow us to apply our results directly in two cases described in the lemmas below:

¹⁰For instance $b(x) = \exp(-(1-x^2)^{-1})$ for $x \in [-1, 0]$, $b(x) = 1$ for $x \in [0, 1]$, $b(x) = \exp(-(1-(1-x)^2)^{-1})$ for $x \in [1, 2]$ and zero otherwise.

Lemma 35. *If for $\tilde{f}_\varepsilon|_{\Lambda_\varepsilon}$ we have a Cantor set of KAM tori on Λ_ε , and assumptions of Theorem 12 (Theorem 19, Theorem 20, Theorem 21) are satisfied for \tilde{f}_ε , then there exists an x_ε for which the condition (20) ((34), (35), (36)) holds for f_ε .*

Proof. The region on Λ_ε between every two invariant tori constitutes an invariant set for \tilde{f}_ε . This means that we can obtain x_ε by means of Theorems 12 (Theorem 19, Theorem 20, Theorem 21) for \tilde{f}_ε which are in $\{I \in (0, 1)\}$. (This is because orbits in Λ_ε resulting from Poincaré recurrence, which are used for the shadowing construction from Theorem 5, will be contained between KAM tori, which we can choose to be in $\{I \in (0, 1)\}$.) Since in $\{I \in (0, 1)\}$ the maps \tilde{f}_ε and f_ε coincide, we obtain diffusing orbits for f_ε . ■

Lemma 36. *If assumptions of Theorem 19 are satisfied for \tilde{f}_ε , then there exist an x_ε and $n_\varepsilon \in \mathbb{N}$ for which*

$$|\pi_I f_\varepsilon^{n_\varepsilon}(x_\varepsilon) - \pi_I x_\varepsilon| > \frac{1}{3}.$$

Proof. We can use the following dichotomy:

1. There exists an orbit in Λ_ε which changes in I by more than $\frac{1}{3}$ for the map $f_\varepsilon|_{\Lambda_\varepsilon}$; then there is nothing to prove.
2. There are no orbits in Λ_ε , which diffuse in I by more than $\frac{1}{3}$ for the map $f_\varepsilon|_{\Lambda_\varepsilon}$. Then orbits of f_ε starting from $\Lambda_\varepsilon \cap \{I \in [\frac{1}{3}, \frac{2}{3}]\}$ will not leave $\Lambda_\varepsilon \cap \{I \in (0, 1)\}$. On $\{I \in (0, 1)\}$ we know that $f_\varepsilon = \tilde{f}_\varepsilon$. We can use Theorem 19 to construct orbits that change in I by more than $\frac{1}{3}$ for \tilde{f}_ε . The constructed orbits will remain in $\{I \in (0, 1)\}$. (This is because orbits in Λ_ε for Poincaré recurrence used for the construction from Theorem 5 will be contained in $\{I \in (0, 1)\}$.) On $\{I \in (0, 1)\}$ we know that $\tilde{f}_\varepsilon = f_\varepsilon$, so any orbit changing in I by $\frac{1}{3}$ for the map \tilde{f}_ε does so also for f_ε .

■

Remark 37. *We obtain a mirror result to Lemma 36 for Theorem 20.*

Remark 38. *Lemma 36 is easily generalised to higher dimensions. We either 1) diffuse along some action on Λ_ε , or 2) diffuse in the selected action which is controlled by Theorem 19.*

Lemma 35 is not easily generalised, since we can have gaps between KAM tori in higher dimensions.

Appendix B. Invariant manifolds and their intersections

We now discuss computation of the local stable and local unstable manifolds, with a focus on obtaining mathematically rigorous computer assisted error bounds on all approximations.

We focus on two methods, which we used for two independent validations of the intersection of the manifolds. The first is based on cone conditions, and the second on the parameterisation method. The first is simpler, and we provide full details, simplifying the results to our particular setting. For the parameterisation method we restrict to Remark 43 where we point our reader to the relevant references.

Below described method of cones is based on the more general results from [29]. We reformulate these for our particular setting, simplifying and giving sketches of proofs, in order to keep the paper self-contained.

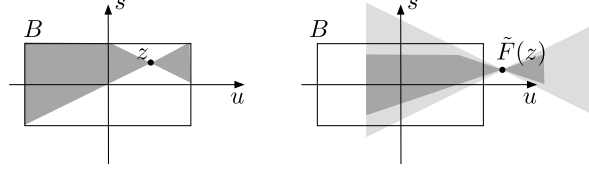


Figure B.6: The cone at z intersected with B (in dark grey) is mapped into the cone at $\tilde{F}(z)$ (in light grey).

Let F be the map (38), i.e. the unperturbed map f_0 acting on x, y ; where we recall that $\alpha = 4$. Let $\mathcal{P} \in \mathbb{R}^{2 \times 2}$ and $\tilde{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be defined as follows

$$\mathcal{P} := \begin{pmatrix} 1 + \sqrt{2} & 1 - \sqrt{2} \\ 2 & 2 \end{pmatrix}, \quad \tilde{F}(z) := \mathcal{P}^{-1}F(\mathcal{P}z).$$

The matrix \mathcal{P} is the coordinate change to Jordan form for $DF(0)$ and $\tilde{F}(z)$ is the map expressed in local coordinates, which diagonalize the stable and unstable directions at the origin, i.e. $D\tilde{F}(0) = \text{diag}((3 - 2\sqrt{2})^{-1}, 3 - 2\sqrt{2})$. We refer to these as the local coordinates, and write $z = (u, s)$. (The u stands for ‘unstable’ and s for ‘stable’.)

Let $\mathcal{L} \in \mathbb{R}$ be a fixed constant satisfying $\mathcal{L} > 0$ and define $C : \mathbb{R}^2 \rightarrow \mathbb{R}$ as

$$C(u, s) = \mathcal{L}|u| - |s|.$$

For $z \in \mathbb{R}^2$ we define the cone at z as $C^+(z) := \{v : C(z - v) \geq 0\}$ (see Figure B.6). Let $r > 0$ be fixed, $J := [-r, r] \subset \mathbb{R}$ and let $B \subset \mathbb{R}^2$ be the rectangle $B := [-r, r] \times [-\mathcal{L}r, \mathcal{L}r]$.

Definition 39. We say that \tilde{F} satisfies cone conditions in B if for every $z \in B$ we have (see Figure B.6)

$$\tilde{F}(C^+(z) \cap B) \subset C^+(\tilde{F}(z)).$$

We have the following lemma, which gives bounds on the unstable manifold in the local coordinates.

Lemma 40. If \tilde{F} satisfies cone conditions in B , and there exists a $\lambda < 1$ such that for every $z \in C^+(0)$ we have

$$|\pi_u \tilde{F}(u, s)| > \lambda^{-1} |u|, \tag{B.1}$$

then there exists a smooth function $w : J \rightarrow [-r\mathcal{L}, r\mathcal{L}]$, such that

$$W_0^u(\tilde{F}, B) = \{(u, w(u)) : u \in J\}.$$

Moreover, $|\frac{d}{du}w(u)| \leq \mathcal{L}$ and for every $u \in J$

$$\|\tilde{F}^{-n}(u, w(u))\| < \lambda^n \sqrt{1 + \mathcal{L}^2} |u|. \tag{B.2}$$

Proof. Since 0 is a hyperbolic fixed point of \tilde{F} , locally at the fixed point the unstable manifold exists, is smooth, and tangent to the horizontal axis, hence it is contained in $C^+(0)$. Cone condition together with (B.1) ensure that the unstable manifold is stretched through B to become

a graph above J . Since locally, close to zero, the unstable manifold is tangent to the horizontal axis it is a graph of a function with the Lipschitz constant smaller than \mathcal{L} . This property is preserved as the manifold is stretched throughout B thanks to the cone condition.

To show (B.2) note that for $z \in C^+(0)$, since $|\pi_s z| < \mathcal{L} |\pi_u z|$, we obtain $\|z\| \leq \sqrt{1 + \mathcal{L}^2} |\pi_u z|$. Thus, from (B.1),

$$\|z\| < \sqrt{1 + \mathcal{L}^2} |\pi_u z| < \sqrt{1 + \mathcal{L}^2} \lambda |\pi_u \tilde{F}(z)|.$$

Taking $z = \tilde{F}^{-n}(u, w(u))$ and using (B.1) we obtain

$$\|\tilde{F}^{-n}(u, w(u))\| < \sqrt{1 + \mathcal{L}^2} \lambda |\pi_u \tilde{F}^{-n+1}(u, w(u))| < \dots < \sqrt{1 + \mathcal{L}^2} \lambda^n |u|,$$

as required. ■

In practice we can validate cone conditions and (B.1) from the interval enclosure of the derivative of \tilde{F} on B .

Lemma 41. *If $[D\tilde{F}(B)](C^+(0)) \subset C^+(0)$ then \tilde{F} satisfies cone conditions.*

Proof. Let $z \in B$ and $v \in C^+(z) \cap B$. Since $v - z \in C^+(0)$, from our assumption it follows that

$$\begin{aligned} \tilde{F}(v) - \tilde{F}(z) &= \int_0^1 \frac{d}{dt} \tilde{F}(z + t(v - z)) dt \\ &= \int_0^1 D\tilde{F}(z + t(v - z)) dt (v - z) \in [DF(B)](v - z) \subset C^+(0), \end{aligned} \quad (\text{B.3})$$

hence $\tilde{F}(C^+(z)) \subset C^+(\tilde{F}(z))$, as required. ■

Above lemma is straightforward to apply in interval arithmetic by checking that

$$[D\tilde{F}(B)](\{1\} \times [-\mathcal{L}, \mathcal{L}]) \subset C^+(0).$$

Lemma 42. *Let $a_{11}, a_{12}, a_{21}, a_{22}$ be real intervals such that $[D\tilde{F}(B)] = (a_{ij})_{i,j \in \{1,2\}}$. If $a_{11} - \mathcal{L}|a_{12}| > \lambda^{-1}$ then (B.1) is fulfilled.*

Proof. Let $(u, s) \in C^+(0) \cap B$. From a mirror argument to (B.3) and since $|s| \leq \mathcal{L}|u|$,

$$|\pi_u F(u, s)| \in |\pi_u [DF(B)](u, s)| \geq a_{11}|u| - \mathcal{L}|a_{12}||u| > \lambda^{-1}|u|,$$

as required. ■

Using a computer program we compute an interval enclosure $[D\tilde{F}(B)]$. This enclosure is used to validate, via Lemmas 41 and 42, the assumptions of Lemma 40. This way we obtain $w : J \rightarrow [-r\mathcal{L}, r\mathcal{L}]$, and define $P_u : J \rightarrow \mathbb{R}^2$ by

$$P_u(x) := \mathcal{P}(x, w(x)).$$

Note that since $w(x)$ is Lipschitz with constant \mathcal{L} , $w(x) \in [-\mathcal{L}x, \mathcal{L}x]$, our method allows us to obtain the explicit bound

$$P_u(x) \subset \mathcal{P}(\{x\} \times [-\mathcal{L}x, \mathcal{L}x]), \quad \text{for every } x \in J.$$

Moreover, by Lemma 40 we know that $\frac{d}{dx}(x, w(x)) \in \{1\} \times [-\mathcal{L}, \mathcal{L}]$, which gives the bound on the derivative of P_u as

$$\frac{d}{dx} P_u(x) \subset \mathcal{P}(\{1\} \times [-\mathcal{L}, \mathcal{L}]), \quad \text{for every } x \in J.$$

From (B.2) we also see that for every $x \in J$

$$\|F^{-n}(P_u(x))\| = \|\mathcal{P}\tilde{F}^{-n}(x, w(x))\| \leq \|\mathcal{P}\| \lambda^n \sqrt{1 + \mathcal{L}^2} |x| \leq C \lambda^n,$$

for $C := \|\mathcal{P}\| \sqrt{1 + \mathcal{L}^2} r$; recall that $J = [-r, r]$. We thus see that we have all the bounds for P_u , which are required by section 25.

The function P_s and associated bounds can be obtained the same way, by considering F^{-1} instead of F .

Remark 43 (Parameterization method for invariant manifolds). *As α approaches zero in the standard map, the fixed point becomes very weakly hyperbolic (eigenvalues approach unity) and hence the dynamics near the fixed point becomes very slow. As a result, homoclinic excursions take more time, and become difficult to track using only the linear approximation of the stable/unstable manifolds.*

To overcome these difficulties we can use – for smaller α – the approach developed in [36], a work which is itself based on the parameterization method of [39, 40, 34]. Using the parameterization method we compute high order Taylor expansions of the stable/unstable manifolds, along with validated error bounds on all truncation errors. This strategy leads to a mathematically rigorous representation of the stable/unstable manifold which is valid in a large neighbourhood of the fixed point. We refer the interested reader to the book of [37] for much more complete discussion of the parameterization method.

Appendix C. Proof of Lemma 31

We will show that the tangent lines to $W_0^u(F)$ and $W_0^s(F)$ at the intersection point v_M^* span \mathbb{R}^2 . Note that $v_M^* = F^M(P_u(x^*))$. Defining inductively the sequence of vectors $w_0, \dots, w_{M-1} \in \mathbb{R}^2$ as $w_0 := DP_u(x^*)$ and $w_k := DF(v_k^*)w_{k-1}$, we see that

$$\frac{d}{dx} F^M(P_u(x))|_{x=x^*} = w_{M-1}.$$

If w_{M-1} was collinear with $\frac{d}{dy} P_s(y)|_{y=y^*}$, then there would exist an $\alpha \neq 0$ for which $\frac{d}{dy} P_s(y^*) = \alpha w_{M-1}$. Taking the vector $V = (1, w_0, \dots, w_{M-1}, 1/\alpha)$ would lead to

$$D\mathcal{F}(p^*)V = 0.$$

This is a contradiction, since if p^* is validated by the use of Theorem 30, then the matrix $D\mathcal{F}(p^*)$ must be invertible.

References

- [1] P. Libermann, C.-M. Marle, *Symplectic geometry and analytical mechanics*, Vol. 35 of Mathematics and its Applications, D. Reidel Publishing Co., Dordrecht, 1987, translated from the French by Bertram Eugene Schwarzbach. doi:10.1007/978-94-009-3807-6. URL <https://doi.org/10.1007/978-94-009-3807-6>
- [2] J. E. Littlewood, *The Lagrange Configuration in Celestial Mechanics*, Proc. London Math. Soc. (3) 9 (4) (1959) 525–543. doi:10.1112/plms/s3-9.4.525. URL <https://doi.org/10.1112/plms/s3-9.4.525>

- [3] N. N. Nehorošev, An exponential estimate of the time of stability of nearly integrable Hamiltonian systems, *Uspehi Mat. Nauk* 32 (6(198)) (1977) 5–66, 287.
- [4] A. N. Kolmogorov, On conservation of conditionally periodic motions for a small change in Hamilton's function, *Dokl. Akad. Nauk SSSR (N.S.)* 98 (1954) 527–530.
- [5] V. I. Arnold, Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian, *Uspehi Mat. Nauk* 18 (5 (113)) (1963) 13–40.
- [6] J. Moser, Convergent series expansions for quasi-periodic motions, *Math. Ann.* 169 (1967) 136–176.
- [7] V. I. Arnold, Small denominators and problems of stability of motion in classical and celestial mechanics, *Uspehi Mat. Nauk* 18 (6 (114)) (1963) 91–192.
- [8] C. L. Siegel, J. K. Moser, *Lectures on celestial mechanics*, Classics in Mathematics, Springer-Verlag, Berlin, 1995, translated from the German by C. I. Kalme, Reprint of the 1971 translation.
- [9] R. de la Llave, [A tutorial on KAM theory](#), in: *Smooth ergodic theory and its applications* (Seattle, WA, 1999), Vol. 69 of Proc. Sympos. Pure Math., Amer. Math. Soc., Providence, RI, 2001, pp. 175–292. doi:10.1090/pspum/069/1858536.
URL <https://doi.org/10.1090/pspum/069/1858536>
- [10] V. I. Arnold, Instability of dynamical systems with many degrees of freedom, *Dokl. Akad. Nauk SSSR* 156 (1964) 9–12.
- [11] J.-P. Marco, D. Sauzin, [Stability and instability for Gevrey quasi-convex near-integrable Hamiltonian systems](#), *Publ. Math. Inst. Hautes Études Sci.* (96) (2002) 199–275 (2003). doi:10.1007/s10240-003-0011-5.
URL <https://doi.org/10.1007/s10240-003-0011-5>
- [12] P. Lochak, J.-P. Marco, [Diffusion times and stability exponents for nearly integrable analytic systems](#), *Cent. Eur. J. Math.* 3 (3) (2005) 342–397. doi:10.2478/BF02475913.
URL <https://doi.org/10.2478/BF02475913>
- [13] A. Bounemoura, J.-P. Marco, [Improved exponential stability for near-integrable quasi-convex Hamiltonians](#), *Nonlinearity* 24 (1) (2011) 97–112. doi:10.1088/0951-7715/24/1/005.
URL <https://doi.org/10.1088/0951-7715/24/1/005>
- [14] J. Zhang, K. Zhang, [Improved stability for analytic quasi-convex nearly integrable systems and optimal speed of Arnold diffusion](#), *Nonlinearity* 30 (7) (2017) 2918–2929. doi:10.1088/1361-6544/aa72b7.
URL <https://doi.org/10.1088/1361-6544/aa72b7>
- [15] P. Bernard, V. Kaloshin, K. Zhang, [Arnold diffusion in arbitrary degrees of freedom and normally hyperbolic invariant cylinders](#), *Acta Math.* 217 (1) (2016) 1–79. doi:10.1007/s11511-016-0141-5.
URL <https://doi.org/10.1007/s11511-016-0141-5>
- [16] V. Gelfreich, D. Turaev, [Arnold diffusion in a priori chaotic symplectic maps](#), *Comm. Math. Phys.* 353 (2) (2017) 507–547. doi:10.1007/s00220-017-2867-0.
URL <https://doi.org/10.1007/s00220-017-2867-0>
- [17] J.-P. Marco, Twist maps and Arnold diffusion for diffeomorphisms, in: *Variational methods*, Vol. 18 of Radon Ser. Comput. Appl. Math., De Gruyter, Berlin, 2017, pp. 473–495.
- [18] J.-P. Marco, [Modèles pour les applications fibrées et les polysystèmes](#), *C. R. Math. Acad. Sci. Paris* 346 (3-4) (2008) 203–208. doi:10.1016/j.crma.2007.11.017.
URL <https://doi.org/10.1016/j.crma.2007.11.017>
- [19] M. W. Hirsch, C. C. Pugh, M. Shub, [Invariant manifolds](#), *Bull. Amer. Math. Soc.* 76 (1970) 1015–1019. doi:10.1090/S0002-9904-1970-12537-X.
URL <https://doi.org/10.1090/S0002-9904-1970-12537-X>
- [20] V. Kaloshin, J. Zhang, K. Zhang, [Normally hyperbolic invariant laminations and diffusive behaviour for the generalized Arnold example away from resonances](#), arXiv: Dynamical Systems.
- [21] M. Gidea, R. de la Llave, T. M-Seara, [A general mechanism of diffusion in Hamiltonian systems: qualitative results](#), *Comm. Pure Appl. Math.* 73 (1) (2020) 150–209. doi:10.1002/cpa.21856.
URL <https://doi.org/10.1002/cpa.21856>
- [22] M. J. Capiński, M. Gidea, [Arnold diffusion, quantitative estimates and stochastic behavior in the three-body problem](#).
URL <https://arxiv.org/abs/1812.03665>
- [23] N. Fenichel, [Persistence and smoothness of invariant manifolds for flows](#), *Indiana Univ. Math. J.* 21 (1971/72) 193–226. doi:10.1512/iumj.1971.21.21017.
URL <https://doi.org/10.1512/iumj.1971.21.21017>
- [24] A. Delshams, R. de la Llave, T. M. Seara, [Geometric properties of the scattering map of a normally hyperbolic invariant manifold](#), *Adv. Math.* 217 (3) (2008) 1096–1153.
URL <https://doi-org.ezproxy.fau.edu/10.1016/j.aim.2007.08.014>
- [25] P. W. Bates, K. Lu, C. Zeng, [Existence and persistence of invariant manifolds for semiflows in Banach space](#), *Mem. Amer. Math. Soc.* 135 (645) (1998) viii+129. doi:10.1090/memo/0645.

- URL <https://doi.org/10.1090/memo/0645>
- [26] P. W. Bates, K. Lu, C. Zeng, Persistence of overflowing manifolds for semiflow, *Comm. Pure Appl. Math.* 52 (8) (1999) 983–1046. doi:10.1002/(SICI)1097-0312(199908)52:8<983::AID-CPA4>3.3.CO;2-F
URL [https://doi.org/10.1002/\(SICI\)1097-0312\(199908\)52:8<983::AID-CPA4>3.3.CO;2-F](https://doi.org/10.1002/(SICI)1097-0312(199908)52:8<983::AID-CPA4>3.3.CO;2-F)
- [27] P. W. Bates, K. Lu, C. Zeng, Approximately invariant manifolds and global dynamics of spike states, *Invent. Math.* 174 (2) (2008) 355–433. doi:10.1007/s00222-008-0141-y
URL <https://doi.org/10.1007/s00222-008-0141-y>
- [28] J. Eldering, Normally hyperbolic invariant manifolds, Vol. 2 of *Atlantis Studies in Dynamical Systems*, Atlantis Press, Paris, 2013, the noncompact case. doi:10.2991/978-94-6239-003-4
URL <https://doi.org/10.2991/978-94-6239-003-4>
- [29] P. Zgliczyński, Covering relations, cone conditions and the stable manifold theorem, *J. Differential Equations* 246 (5) (2009) 1774–1819. doi:10.1016/j.jde.2008.12.019
URL <https://doi.org/10.1016/j.jde.2008.12.019>
- [30] M. J. Capiński, P. Zgliczyński, Cone conditions and covering relations for topologically normally hyperbolic invariant manifolds, *Discrete Contin. Dyn. Syst.* 30 (3) (2011) 641–670. doi:10.3934/dcdis.2011.30.641
URL <https://doi.org/10.3934/dcdis.2011.30.641>
- [31] M. J. Capiński, P. Zgliczyński, Geometric proof for normally hyperbolic invariant manifolds, *J. Differential Equations* 259 (11) (2015) 6215–6286. doi:10.1016/j.jde.2015.07.020
URL <https://doi.org/10.1016/j.jde.2015.07.020>
- [32] X. Cabré, E. Fontich, R. de la Llave, The parameterization method for invariant manifolds. I. Manifolds associated to non-resonant subspaces, *Indiana Univ. Math. J.* 52 (2) (2003) 283–328. doi:10.1512/iumj.2003.52.2245
URL <https://doi.org/10.1512/iumj.2003.52.2245>
- [33] X. Cabré, E. Fontich, R. de la Llave, The parameterization method for invariant manifolds. II. Regularity with respect to parameters, *Indiana Univ. Math. J.* 52 (2) (2003) 329–360. doi:10.1512/iumj.2003.52.2407
URL <https://doi.org/10.1512/iumj.2003.52.2407>
- [34] X. Cabré, E. Fontich, R. de la Llave, The parameterization method for invariant manifolds. III. Overview and applications, *J. Differential Equations* 218 (2) (2005) 444–515. doi:10.1016/j.jde.2004.12.003
URL <https://doi.org/10.1016/j.jde.2004.12.003>
- [35] G. Alefeld, Inclusion methods for systems of nonlinear equations—the interval Newton method and modifications, in: *Topics in validated computations* (Oldenburg, 1993), Vol. 5 of *Stud. Comput. Math.*, North-Holland, Amsterdam, 1994, pp. 7–26.
- [36] J. D. Mireles James, K. Mischaikow, Rigorous a-posteriori computation of (un)stable manifolds and connecting orbits for analytic maps, *SIAM J. Appl. Dyn. Syst.* 12 (2) (2013) 957–1006. doi:10.1137/12088224X
URL <http://dx.doi.org/10.1137/12088224X>
- [37] A. Haro, M. Canadell, J.-L. s. Figueras, A. Luque, J.-M. Mondelo, The parameterization method for invariant manifolds, Vol. 195 of *Applied Mathematical Sciences*, Springer, [Cham], 2016, from rigorous results to effective computations. doi:10.1007/978-3-319-29662-3
URL <http://dx.doi.org.ezproxy.fau.edu/10.1007/978-3-319-29662-3>
- [38] T. Kapela, M. Mrozek, D. Wilczak, P. Zgliczyński, *Capd:dynsys: a flexible c++ toolbox for rigorous numerical analysis of dynamical systems*, *Communications in Nonlinear Science and Numerical Simulation* (2020) 105578doi:<https://doi.org/10.1016/j.cnsns.2020.105578>
URL <https://www.sciencedirect.com/science/article/pii/S1007570420304081>
- [39] X. Cabré, E. Fontich, R. de la Llave, The parameterization method for invariant manifolds. I. Manifolds associated to non-resonant subspaces, *Indiana Univ. Math. J.* 52 (2) (2003) 283–328.
- [40] X. Cabré, E. Fontich, R. de la Llave, The parameterization method for invariant manifolds. II. Regularity with respect to parameters, *Indiana Univ. Math. J.* 52 (2) (2003) 329–360.