# 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 algxébrique, 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

The influence of Celestial Mechanics on the evolution of dynamical systems theory cannot be overstated. The latter originated in Newton's *Principia*, where were formulated and solved the differential equations governing the motion of two planets. In modern terminology, Newton showed that the two body problem is "completely integrable": that is, there are enough conserved quantities so that all solutions are obtained by taking intersections of their level sets. In the two body case this results in the conic sections.

The Newtonian $n$-body problem for $n \geq 3$ is far more complicated, and for several generations it was unclear whether the notion of integrability was sufficient for describing the behavior of more general gravitating systems. Interest in the problem led to the development of perturbation theory in the works of prominent mathematicians like Euler, Lagrange, Laplace, Gauss, and Hamilton during the next two centuries, and to the introduction of both the Lagrangian and

*Corresponding author.
[1]Partially 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.
[2]Partially supported by NSF grant MSPRF DMS-2001758.
[3]Partially supported by NSF grant DMS 1813501.

Hamiltonian formulations of mechanics. However, the "only" problem studied during this period was to compute with extreme precision the trajectories of (some of) the planets in the solar system, taking into account mutual interactions between them.

The work of Poincaré at the end of the $19^{th}$ century revolutionized the theory, and the fundamental questions changed in several dramatic ways.

- Instead of looking at individual solutions of a given system, Poincaré realized that considering the evolution of *all* initial conditions enables one to use (or create new) geometric tools adapted to this setting.
- Instead of looking at the evolution of a trajectory over a finite time interval, he realized that understanding the asymptotic behavior of certain special orbits as time goes to infinity yields useful information.

As an example, recall that Poincaré used the geometry of certain infinitely long homoclinic orbits to establish that the (restricted) three body problem is not integrable, simultaneously shattering the notion that integrability was sufficient for the study of all dynamics and establishing the existence of complex phenomena never before imagined. Another stunning example where both ideas are fully exploited is the celebrated "Poincaré recurrence theorem", which uses measure (probability) theory on the geometric side, and which would have been totally unreachable by any study only focusing on the evolution of single trajectories. The impressive ensemble of ideas introduced and developed by Poincaré is nowadays considered to be the very cornerstone of dynamical systems theory.

Stability remained a major concern in the new theory, and a central problem was to understand perturbations of completely integrable systems. Indeed the solar system itself can be viewed as a system of weakly coupled (completely integrable) two body problems, and the question of it's stability has captivated mathematicians since the days of Newton. To formalize the discussion[4], let $\mathbb{A}^n = T^*\mathbb{T}^n = \mathbb{T}^n \times \mathbb{R}^n$ denote the annulus with the angle-action coordinates $(\theta, r)$, endowed with the symplectic form $\Omega = \sum_{i=1}^n dr_i \wedge d\theta_i$. Consider a Hamiltonian of the form

$$H(\theta, r) = h(r) + f(\theta, r), \tag{1}$$

where $f$ is small in some suitable function space (analytic, $C^\infty$, $C^\kappa$, etcetera). The Hamiltonian differential equations generated by $H$ read

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

and we note that when $f \equiv 0$, all the orbits move with constant velocity on invariant tori.

When $f$ is small it is apparent that the evolution of the action variables $r_i$ are "slow". The fact that this evolution is "extremely slow" emerged from the averaging methods originally developed by Lagrange and Laplace, furthered by Poincaré and Birkhoff, and which culminated 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 $f$ 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, where $\varepsilon$ measures the size of the perturbation function $f$.

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

Following another line of research dating back to Poincaré, 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 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, provided that the frequency vector $\nabla h(r^0)$ is Diophantine, meaning that there are constants $\gamma > 0$ and $\tau > 0$ such that

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

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 and it's development.

Taken together, the KAM and averaging theories provide indispensable information about the dynamics of perturbations of integrable Hamiltonian systems. The KAM theory tells us that *some* orbits remain close to the unperturbed dynamics for all time (the KAM tori), while the averaging theory says that *all* orbits stay close to the unperturbed dynamics for exponentially long times. A natural question is to ask *do there exist orbits which move "arbitrarily" far from the integrable dynamics on a long enough time scale?* Indeed, when $n \geq 3$, a KAM torus has a connected complement in a constant energy level, and the existence of the full family of KAM tori (whose complement has an $O(\sqrt{\varepsilon})$ relative measure in this level) does not prevent trajectories from drifting away from the integrable dynamics on very long timescales.

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

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

where $g$ is an explicit fixed trigonometric polynomial, with $\mu$ and $\varepsilon$ independent parameters. The example has several important special properties, which we state here for a general analytic function $g$.

- 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 $\mu$ and small enough ($\varepsilon$ has to be exponentially small w.r.t. $\mu$ 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 $\mathcal{A}_\varepsilon$ close to $\mathcal{A}_0$, with a rich homoclinic structure, while the Hamiltonian flow on $\mathcal{A}_\varepsilon$ 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 $\varepsilon > 0$. By exploiting this fact Arnold was able to show that for $\mu, \varepsilon > 0$ small enough $H_\varepsilon$ 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

$$r_1(0) < 0, \qquad r_1(T_\varepsilon) > 1,$$

for this orbit. This provided an explicit example where orbits of the perturbed system "diffuse" as far and as fast[5] from the integrable dynamics as allowed by averaging theory.

The use of two independent parameters (a method originally introduced by Poincaré) 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 parturbed 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.

We consider another line of study, which comes from weakening the hypothesis that the unperturbed system is completely integrable. Consider for example systems of the form (2), in which the parameter $\mu$ 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 (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 $\varepsilon$ under which diffusion phenomenons can appear, or, what is the maximal length of diffusive trajectories? These questions require new methods, and this is the main concern of the present work. It turns out that in realistic physical systems the relevant quantities to estimate are difficult to compute, and our aim is to provide an explicit example illustrating the relevance of computer-assisted methods of proof in such problems.

In order to simplify the construction we shift our focus to symplectic maps instead of Hamiltonian vector fields. Such a 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 $f_\varepsilon \colon \mathbb{R}^2 \times \mathbb{T}^2 \to \mathbb{R}^2 \times \mathbb{T}^2$ defined by

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

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

where $(x,y) \in \mathbb{R}^2$ and $(\theta, I) \in \mathbb{T}^2$. Observe that the map $f_\varepsilon$ can be seen as a perturbation of a standard map (variables $(x,y)$) coupled to an $I$-parametrized rotation on $\mathbb{T}^2$ (variables $(\theta, I)$). 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 $\alpha$ as a perturbation parameter, and will show that for some fixed $\alpha$ the stable and unstable manifolds intersect transversely at some point $P$ (so that the parameter $\alpha$ plays the role of $\mu$ 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, 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].

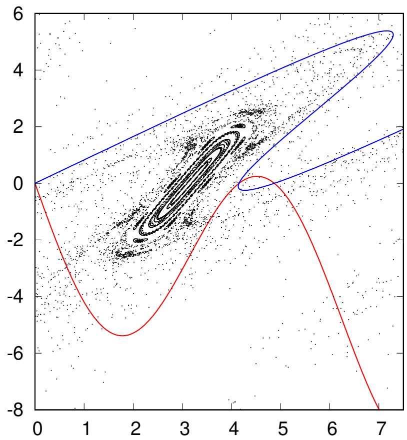


Figure 1: Phase space structure of the Chirikov Standard Map when $\alpha = 4$. Black dots indicate the dynamics of a number of "typical" orbits. The stable and unstable manifolds of the fixed point at the origin are depicted by the red and blue curves respectively.

When $\varepsilon > 0$, the preservation of the fibers 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 $f_\varepsilon$, and that they have lengths independent of the size $\varepsilon$ 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 [19]. A scattering map is a function from a normally hyperbolic invariant manifold to itself, defined through appropriate intersections of fibers of its stable and unstable manifolds. In [19] 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 11, 17, 18 and 19. Theorems 11, 17, 18 establish orbits which diffuse over an explicit interval of actions. Theorem 19 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 show two computer-assisted methods with which the scattering map can be computed: by using cones or the parameterization method. 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 [20]. This work is based on the method of correctly aligned windows. The difference compared to this paper is that [20] 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 [19].

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. Section 5 is a technical treatment of constructive methods for studying the stable/unstable manifolds fo fixed points of maps. 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. We use $\mathbb{T} = \mathbb{R}/\mathrm{mod}\ 2\pi$, to stand for a one dimensional torus and $\mathbb{T}^k$ to stand for a $k$-dimensional torus. For a set $A$ in a topological space we shall write $\overline{A}$ 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 [21].

**Definition 1.** *Let $\Lambda \subset \mathbb{R}^n$ be a compact manifold without boundary, invariant under $f : \mathbb{R}^n \to \mathbb{R}^n$, i.e., $f(\Lambda) = \Lambda$, where $f$ is a $C^r$-diffeomorphism, $r \geq 1$. We say that $\Lambda$ is a normally hyperbolic invariant manifold (with symmetric rates) if there exists a constant $C > 0$, rates $0 < \lambda < \mu^{-1} < 1$ and a $Tf$ invariant splitting for every $x \in \Lambda$*

$$\mathbb{R}^n = E_x^u \oplus E_x^s \oplus T_x\Lambda$$

*such that*

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

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

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

Let $d(x,\Lambda)$ stand for the distance between a point $x$ and the manifold $\Lambda$, induced by the Euclidean norm. Given a normally hyperbolic invariant manifold and a suitable small tubular neighbourhood $U \subset \mathbb{R}^n$ of $\Lambda$ one defines its local unstable and local stable manifold [21] as

$$W^u_\Lambda(f,U) = \left\{ y \in \mathbb{R}^n \mid f^k(y) \in U, d\left(f^k(y),\Lambda\right) \le C_y \lambda^{|k|},\ k \le 0 \right\},$$
$$W^s_\Lambda(f,U) = \left\{ y \in \mathbb{R}^n \mid f^k(y) \in U, d\left(f^k(y),\Lambda\right) \le C_y \lambda^{k},\ k \ge 0 \right\},$$

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

$$W^u_\Lambda(f) = \bigcup_{n\ge 0} f^n\left(W^u_\Lambda(f,U)\right), \qquad W^s_\Lambda(f) = \bigcup_{n\ge 0} f^{-n}\left(W^s_\Lambda(f,U)\right).$$

The manifolds $W^u_\Lambda(f,U)$, $W^s_\Lambda(f,U)$, $W^u_\Lambda(f)$ and $W^s_\Lambda(f)$ are foliated by

$$W^u_x(f,U) = \left\{ y \in \mathbb{R}^n \mid f^k(y) \in U, d(f^k(y),f^k(x)) \le C_{x,y} \lambda^{|k|},\ k \le 0 \right\},$$
$$W^s_x(f,U) = \left\{ y \in \mathbb{R}^n \mid f^k(y) \in U, d(f^k(y),f^k(x)) \le C_{x,y} \lambda^{k},\ k \ge 0 \right\},$$

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

$$W^u_x(f) = \bigcup_{n\ge 0} f^n\left(W^u_{f^{-n}(x)}(f,U)\right), \qquad W^s_x(f) = \bigcup_{n\ge 0} f^{-n}\left(W^s_{f^n(x)}(f,U)\right).$$

Let

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

The manifold $\Lambda$ is $C^l$ smooth, the manifolds $W^u_\Lambda(f)$, $W^s_\Lambda(f)$ are $C^{l-1}$ and $W^u_x(f)$, $W^s_x(f)$ are $C^r$ [22]. Normally hyperbolic manifolds, as well as their stable and unstable manifolds and their fibres persist under small perturbations [21].

**Lemma 2.** *[22] In the case that the map $f$ preserves a symplectic form $\omega$, the induced form $\omega|_\Lambda$ is a symplectic and $f|_\Lambda$ preserves $\omega|_\Lambda$.*

*2.2. Shadowing of scattering maps*

Our diffusion result is based on shadowing lemmas for scattering maps found in [19], which we now summarize.

Let us assume that $\Lambda$ is a normally hyperbolic invariant manifold for $f$, and define two maps,

$$\Omega_+ : W^s_\Lambda(f) \to \Lambda,$$
$$\Omega_- : W^u_\Lambda(f) \to \Lambda,$$

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

**Definition 3.** *We say that a manifold $\Gamma \subset W^u_\Lambda(f) \cap W^s_\Lambda(f)$ is a homoclinic channel for $\Lambda$ if the following conditions hold:*

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

$$T_x W^s_\Lambda(f) \oplus T_x W^u_\Lambda(f) = \mathbb{R}^n, \tag{8}$$
$$T_x W^s_\Lambda(f) \cap T_x W^u_\Lambda(f) = T_x\Gamma, \tag{9}$$

*(ii) the fibres of $\Lambda$ intersect $\Gamma$ transversally in the following sense*

$$T_x\Gamma \oplus T_x W^s_{x_+}(f) = T_x W^s_\Lambda(f), \tag{10}$$
$$T_x\Gamma \oplus T_x W^u_{x_-}(f) = T_x W^u_\Lambda(f), \tag{11}$$

*for every $x \in \Gamma$,*

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

**Definition 4.** *Assume that $\Gamma$ is a homoclinic channel for $\Lambda$ and let*

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

*We define a scattering map $\sigma^\Gamma$ for the homoclinic channel $\Gamma$ as*

$$\sigma^\Gamma := \Omega^\Gamma_+ \circ \left(\Omega^\Gamma_-\right)^{-1} : \Omega^\Gamma_-(\Gamma) \to \Omega^\Gamma_+(\Gamma).$$

We have the following theorem.

**Theorem 5.** *[19] Assume that $f : \mathbb{R}^n \to \mathbb{R}^n$ is a sufficiently smooth map, $\Lambda \subset \mathbb{R}^n$ is a normally hyperbolic invariant manifold with stable and unstable manifolds which intersect transversally along a homoclinic channel $\Gamma \subset \mathbb{R}^n$, and $\sigma$ is the scattering map associated to $\Gamma$.*

*Assume that $f$ preserves measure absolutely continuous with respect to the Lebesgue measure on $\Lambda$, and that $\sigma$ sends positive measure sets to positive measure sets.*

*Let $m_1, \ldots, m_n \in \mathbb{N}$ be a fixed sequence of integers. Let $\{x_i\}_{i=0,\ldots,n}$ be a finite pseudo-orbit in $\Lambda$, that is a sequence of points in $\Lambda$ of the form*

$$x_{i+1} = f^{m_i} \circ \sigma^\Gamma(x_i), \qquad i = 0, \ldots, n-1,\ n \geq 1, \tag{12}$$

*that is contained in some open set $\mathcal{U} \subset \Lambda$ with almost every point of $\mathcal{U}$ recurrent for $f|_\Lambda$. (The points $\{x_i\}_{i=0,\ldots,n}$ do not have to be themselves recurrent.)*

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

**Remark 6.** *In [19] 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 [19] for the following reason.*

*The proof of the theorem in [19] is based on a general shadowing lemma [19, Lemma 3.1] which ensures that given a pseudo-orbits 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 $z_{i+1} = f^{k_i+n_i}(z_i)$, $\delta$-close to the pseudo-orbit $y_i$.*

*The shadowing of a pseudo-orbit $x_{i+1} = \sigma^\Gamma(x_i)$ is proven in [19] by combining [19, 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 [19, Lemma 3.11]. The true orbit, which follows from [19, 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 [19, Lemma 3.1] ensures that $y_i$ can be shadowed by a true orbit. Since $y_i$ is close to the pseudo-orbit $x_i$ form (12) we obtain the shadowing of (12) by a true orbit.*

**Remark 7.** *The result can be immediately extended to the case where we have a finite number of scattering maps $\sigma_1, \ldots, \sigma_L$ to shadow*

$$x_{i+1} = f^{m_i} \circ \sigma_{\alpha_i}(x_i), \qquad i = 0, \ldots, n-1,\ n \geq 1,$$

*for two prescribed sequences $m_1, \ldots, m_n \in N$ and $\alpha_1, \ldots, \alpha_n \in \{1, \ldots, L\}$; see [19, Theorem 3.7].*

**Remark 8.** *If $f$ is symplectic for a symplectic form $\omega$, $\Lambda$ is compact and $\omega|_{\Lambda}$ is not degenerate on $\Lambda$ then $f|_{\Lambda}$ is measure-preserving. Hence, by the Poincaré recurrence theorem almost every point of $\Lambda$ is recurrent. In such setting in Theorem 5 we can take $\mathcal{U} = \Lambda$.*

### *2.3. 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.

Let $\mathcal{F} : \mathbb{R}^k \to \mathbb{R}^k$ 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

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

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 if $A \in \mathbf{A}$ then $A^{-1} \in \mathbf{A}^{-1}$.

**Theorem 9.** *[23] (Interval Newton method) Let $\mathcal{F} : \mathbb{R}^k \to \mathbb{R}^k$ be a $C^1$ function and $X = \Pi_{i=1}^k [a_i, b_i]$ 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 $\mathcal{F}(x^*) = 0$.*

## 3. Main results

Let $f_0, g : \mathbb{R}^{2d} \times \mathbb{T}^2 \to \mathbb{R}^{2d} \times \mathbb{T}^2$ and consider the following system

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

where $u, s \in \mathbb{R}^d$, $\theta, I \in \mathbb{T}$. Assume that $f_{\varepsilon}$ are symplectic maps for a symplectic form $\omega$, assume that for $\varepsilon = 0$

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

is a normally hyperbolic invariant manifold for which $\omega|_{\Lambda_0}$ is non degenerate, and that $I$ is a constant of motion for the unperturbed system, i.e.

$$\pi_I f_0(x) = \pi_I x, \tag{13}$$

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

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

$$\pi_I \left( f_\varepsilon^{n_\varepsilon}(x_\varepsilon) - x_\varepsilon \right) > 1. \tag{14}$$

The coordinates have the following roles. The $u, s$ are the coordinates on unstable and stable bundles, respectively, of $\Lambda_0$. The $\theta$ is an angle and $I$ plays the role of an 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. We do so to achieve simplicity[6].

**Remark 10.** *The assumption that $\Lambda_0$ is a torus to simplifies the arguments, as $\Lambda_0$ is compact without boundary and the normally hyperbolic manifold theorem ensures that $\Lambda_0$ is perturbed to a nearby compact normally hyperblic invariant manifold $\Lambda_\varepsilon$. Having compactess of $\Lambda_\varepsilon$ is convenient, but does not appear to be necessary.*

*A more typical setting is when $\Lambda_0$ 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 $f_\varepsilon : \mathbb{R}^{2d} \times \mathbb{R} \times \mathbb{T} \to \mathbb{R}^{2d} \times \mathbb{R} \times \mathbb{T}$. In such case consider $I \in [0, 1]$ and artificially 'glue' the system so that I is in $\mathbb{T}^1$ to apply our result. Details of how this can be done are found in section Appendix A.*

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

$$x' = J\nabla \left( H(x) + \varepsilon G(x, t) \right), \tag{15}$$

where $H : \mathbb{R}^{2d+2} \to \mathbb{R}^{2d+2}$, $G : \mathbb{R}^{2d+2} \times \mathbb{T}^1 \to \mathbb{R}^{2d+2}$ and

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

In such case we can take $f_\varepsilon(x) = \Phi^\varepsilon_{2\pi}(x, t_0)$, for some $t_0 \in [0, 2\pi)$, where $\Phi^\varepsilon_t(x_0, t_0)$ stands for the time $t$ flow induced by (16) with the initial condition $(x_0, t_0)$. If the unperturbed system admits a normally hyperbolic invariant cylinder, then we are in the setting from Remark 10.

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 $f_\varepsilon$ 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 $\theta$ and we choose $I$ as the Hamiltonian $H_0$ (energy) of the unperturbed system.

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

[6] We believe that our methods can be generalised to the higher dimensional case. We make comments how to do so in Remarks 14, 20 after the statements of our results. We also make a cautionary Remark 44 regarding potential problems while extending the higher dimensional setting to the case of normally hyperbolic cylinders in Appendix A.

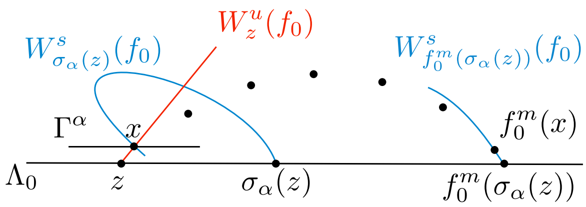


Figure 2: The setting for Theorem 11.

**Theorem 11.** *Assume that there is a neighborhood $U$ of $\Lambda_0$ and positive constants $L_g, C, \lambda$, where $\lambda \in (0,1)$, such that for every $x_1, x_2 \in U$, and every $z \in \Lambda_0$, $x_u \in W^u_z(f_0, U)$ and $x_s \in W^s_z(f_0, U)$ we have*

$$|\pi_I(g(x_1) - g(x_2))| \leq L_g \|x_1 - x_2\|, \tag{16}$$

$$\begin{aligned} \left\| f_0^n(z) - f_0^n(x_u) \right\| &< C\lambda^{|n|} && \text{for all } n \leq 0, \\ \left\| f_0^n(z) - f_0^n(x_s) \right\| &< C\lambda^{n} && \text{for all } n \geq 0. \end{aligned} \tag{17}$$

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

*Assume that for every $z \in \Lambda_0$*

1. *There exists an $\alpha \in \{1, \ldots, L\}$ such that $z \in \mathrm{dom}(\sigma_\alpha)$.*
2. *There exists an $m \in \mathbb{N}$ and a point $x \in \Gamma_\alpha$, $x \in W^u_z(f_0, U) \cap W^s_{\sigma_\alpha(z)}(f_0)$ such that $f_0^m(x) \in W^s_{f_0^m(\sigma_\alpha(z))}(f_0, U)$ (see Figure 2) and*

$$\sum_{j=0}^{m-1} \pi_I g\left(f_0^j(x)\right) - \frac{1+\lambda}{1-\lambda} L_g C > 0. \tag{18}$$

*(The above $\alpha, m$ and $x$ can depend on the choice of $z$.)*

*Then for sufficiently small $\varepsilon > 0$ there exists an $x_\varepsilon$ and $n_\varepsilon > 0$ such that*

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

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

**Remark 12.** Assumption (17) will readily hold when $g$ is $C^1$ since $\Lambda_0$ is compact, so we can take $\overline{U}$ to be compact as well. 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 $\lambda$ which enter into the key assumption (22). Condition (22) measures the influence of the perturbation term $g$ on the coordinate $I$. This can be thought of as a discrete version of a Melnikov integral. (Instead of an integral we have a sum, since we are working with a discrete system.) An important feature is that we are computing the sum along a finite fragment of a homoclinic orbit, and not along the full orbit as is the case in Melnikov theory. The second term in (22) takes into account the truncated tail.

**Remark 13.** In Theorem 11 we assume that the homoclinic channels are in $U$, meaning that they are close to $\Lambda_0$. This is not a restrictive assumption, since a homoclinic channel which is far away can be propagated close to $\Lambda_0$ by using backward iterates of $f_0$.

**Remark 14.** Theorem 11 can be generalised to the setting of higher dimensional $\theta$ and $I$ as follows. If we have actions $I_1, \ldots, I_k$, we can single out one of them (say $I = I_1$) for the conditions (17) and (22), to obtain diffusion towards the singled out action.

**Remark 15.** 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 11.** The manifold $\Lambda_0$ is perturbed to a normally hyperbolic invariant manifold $\Lambda_\varepsilon$ for $f_\varepsilon$. Moreover, for sufficiently small $\varepsilon$ if $z \in \Lambda_\varepsilon$, $x_u \in W^u_z(f_\varepsilon, U)$ and $x_s \in W^s_z(f_\varepsilon, U)$ so that

$$\begin{aligned} \left\|f_\varepsilon^n(z) - f_\varepsilon^n(x_u)\right\| &< C\lambda_\varepsilon^{|n|} && \text{for all } n \le 0, \\ \left\|f_\varepsilon^n(z) - f_\varepsilon^n(x_s)\right\| &< C\lambda_\varepsilon^{n} && \text{for all } n \ge 0, \end{aligned} \tag{19}$$

with $\lambda_\varepsilon$ converging to $\lambda$ as $\varepsilon$ tends to zero.

Since transversal intersections persist under perturbation, the homoclinic channels $\Gamma^1, \ldots, \Gamma^L$ for $f_0$ are perturbed to homoclinic channels $\Gamma^1_\varepsilon, \ldots, \Gamma^l_\varepsilon$ for $f_\varepsilon$, provided that $\varepsilon > 0$ is sufficiently small. This leads [22] to a scattering map $\sigma^\varepsilon_\alpha : \Omega^{\Gamma^\alpha_\varepsilon}_-(\Gamma^\alpha_\varepsilon) \to \Omega^{\Gamma^\alpha_\varepsilon}_+(\Gamma^\alpha_\varepsilon)$ for $f_\varepsilon$.

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

$$\pi_I\left(f_\varepsilon^m \circ \sigma^\varepsilon_\alpha(z_\varepsilon) - z_\varepsilon\right) > \varepsilon c, \tag{20}$$

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

$$\sum_{j=0}^{m-1} \pi_I g\left(f_0^j(x)\right) - \frac{1+\lambda}{1-\lambda} L_g C > c \tag{21}$$

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

It turns out that (21) is the main step in our proof, since once it is established the result follows from the shadowing Theorem 5. Below we first prove (21) and then discuss how to apply the shadowing method.

Consider now a $z_\varepsilon \in \Lambda_\varepsilon$. By our assumptions, for every $x \in \Lambda$ we have an $\alpha \in \{1, \ldots, L\}$, $m \in \mathbb{N}$ and $x \in W^u_z(f_0, U) \cap W^s_{\sigma_\alpha(z)}(f_0)$ such that $f_0^m(x) \in W^s_{f_0^m(\sigma_\alpha(z))}(f_0, U)$ and (22) holds. This means that for sufficiently small $\varepsilon$, for some $\alpha \in \{1, \ldots, L\}$ and some $m \in \mathbb{N}$ we shall have $x_\varepsilon \in W^u_{z_\varepsilon}(f_\varepsilon, U)$ and $f_\varepsilon^m(x_\varepsilon) \in W^s_{f_\varepsilon^m(\sigma^\varepsilon_\alpha(z_\varepsilon))}(f_\varepsilon, U)$, hence by (20)

$$\left\|f_\varepsilon^j(z_\varepsilon) - f_\varepsilon^j(x_\varepsilon)\right\| < C\lambda_\varepsilon^{|j|} \qquad \text{for } j \le 0 \tag{22}$$

$$\left\|f_\varepsilon^{m+j}(\sigma^\varepsilon_\alpha(z_\varepsilon)) - f_\varepsilon^{m+j}(x_\varepsilon)\right\| < C\lambda_\varepsilon^{j} \qquad \text{for } j \ge 0. \tag{23}$$

Due to (22) and the continuous dependence of $x_\varepsilon$, $\lambda_\varepsilon$, for sufficiently small $\varepsilon$ we shall have

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

In order to show (21) we will split our estimates into three terms

$$f_\varepsilon^m(\sigma_\alpha^\varepsilon(z_\varepsilon)) - z_\varepsilon = [f_\varepsilon^m(\sigma_\alpha^\varepsilon(z_\varepsilon)) - f_\varepsilon^m(x_\varepsilon)] + [f_\varepsilon^m(x_\varepsilon) - x_\varepsilon] + [x_\varepsilon - z_\varepsilon], \tag{24}$$

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

$$\left|\pi_I [f_\varepsilon^m(\sigma_\alpha^\varepsilon(z_\varepsilon)) - f_\varepsilon^m(x_\varepsilon)]\right| \le \varepsilon \frac{1}{1-\lambda_\varepsilon} L_g C. \tag{25}$$

Indeed, since $f_\varepsilon(x) = f_0(x) + \varepsilon g(x_1)$ 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

$$\pi_I\left(f_\varepsilon^j(x_1) - f_\varepsilon^j(x_2)\right) = \pi_I[x_1 - x_2] + \varepsilon\sum_{i=0}^{j-1}\pi_I\left(g\left(f_\varepsilon^i(x_1)\right) - g\left(f_\varepsilon^i(x_2)\right)\right). \tag{26}$$

Taking $x_1 = f_\varepsilon^m(\sigma_\alpha^\varepsilon(z_\varepsilon))$ and $x_2 = f_\varepsilon^m(x_\varepsilon)$ from (27) with $\pi_I[x_1 - x_2]$ moved to the left hand side, we have

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

where the last two inequalities follow from (24). Letting $j \to \infty$, we obtain (26).

Now consider the third term from (25). An analogous bound to (24) is obtained as follows. From (27) we have that

$$\begin{aligned} \pi_I(x_1 - x_2) &= \pi_I\left[f_\varepsilon^{-j}(x_1) - f_\varepsilon^{-j}(x_2)\right] + \varepsilon\sum_{i=0}^{j-1}\pi_I\left(g\left(f_\varepsilon^{i-j}(x_1)\right) - g\left(f_\varepsilon^{i-j}(x_2)\right)\right) \\ &= \pi_I\left[f_\varepsilon^{-j}(x_1) - f_\varepsilon^{-j}(x_2)\right] + \varepsilon\sum_{i=-j}^{-1}\pi_I\left(g\left(f_\varepsilon^{i}(x_1)\right) - g\left(f_\varepsilon^{i}(x_2)\right)\right). \end{aligned} \tag{27}$$

Taking $x_1 = x_\varepsilon$ and $x_2 = z_\varepsilon$, from (28) we obtain

$$
\begin{aligned}
&|\pi_I (x_\varepsilon - z_\varepsilon)| \\
&\leq \left|\pi_I \left[f_\varepsilon^{-j}(x_\varepsilon) - f_\varepsilon^{-j}(z_\varepsilon)\right]\right| + \varepsilon \sum_{i=-j}^{-1} \left|\pi_I \left(g\left(f_\varepsilon^i(x_\varepsilon)\right) - g\left(f_\varepsilon^i(z_\varepsilon)\right)\right)\right| \\
&< C\lambda_\varepsilon^j + \varepsilon L_g \sum_{i=-j}^{-1} \left\|f_\varepsilon^i(x_1) - f_\varepsilon^i(x_2)\right\| \\
&< C\lambda_\varepsilon^j + \varepsilon L_g \sum_{i=1}^{j} C\lambda_\varepsilon^i,
\end{aligned}
$$

where the last two inequalities follow from (23). Taking $j \to \infty$ gives

$$
|\pi_I (x_\varepsilon - z_\varepsilon)| \leq \varepsilon \frac{\lambda_\varepsilon}{1-\lambda_\varepsilon} C L_g. \tag{28}
$$

We now turn to the middle term from (25). Since $f_\varepsilon(x) = f_0(x) + \varepsilon g(x)$ and $\pi_I f_0(x) = x$, it follows that (below we consider $x = f_\varepsilon^j(x_\varepsilon)$)

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

so

$$
\pi_I\left(f_\varepsilon^m(x_\varepsilon) - x_\varepsilon\right) = \sum_{j=0}^{m-1} \pi_I\left(f_\varepsilon^{j+1}(x_\varepsilon) - f_\varepsilon^j(x_\varepsilon)\right) = \varepsilon \sum_{j=0}^{m-1} \pi_I g\left(f_\varepsilon^j(x_\varepsilon)\right). \tag{29}
$$

Combining (25), (26), (29), (30) gives

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

Since the right hand side of the inequality above depends continuously on $\varepsilon$, from (22) we obtain (21) for sufficiently small $\varepsilon$.

This establishes the key step (21). 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 $\varepsilon$. Since $f_\varepsilon$ is symplectic, by Remark 8 almost every point of $\Lambda_\varepsilon$ is recurrent for $f_\varepsilon|_{\Lambda_\varepsilon}$. Choose $x_0 \in \Lambda_\varepsilon$ having $\pi_I x_0 = 0$ and consider $\alpha_0, m_0$, (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 $\pi_I\left(f_\varepsilon^{m_0} \circ \sigma_{\alpha_0}^\varepsilon(x_0) - x_0\right) > c\varepsilon$. This can be done due to (21). Repeating the procedure, choosing $\alpha_i, m_i$ for which $\left(\pi_I f_\varepsilon^{m_i} \circ \sigma_{\alpha_i}^\varepsilon(x_i) - x_i\right) > c\varepsilon$ we obtain a pseudo-orbit $x_0, \ldots, x_N$, where $x_{i+1} := f_\varepsilon^{m_i} \circ \sigma_{\alpha_i}^\varepsilon(x_i)$, for which

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

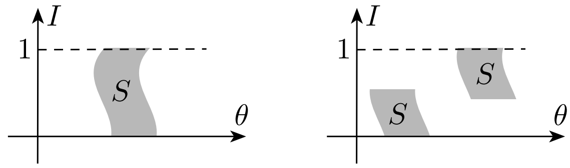


Figure 3: A typically shaped strip (left) and a 'strip' consisting of two connected components (right).

Choosing $N$ large enough, we obtain that $\pi_I(x_N - x_0) > 1$. By Theorem 5 the pseudo-orbit $x_0, \ldots, x_N$ is $\delta$-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 11 we assume that for any point in $\Lambda_0$ we can find a pseudo-orbit such that we have a gain in $I$. Note however that we do not need to have (22) for all $z \in \Lambda_0$. It is enough to have (22) for $z$ on some smaller subset of $\Lambda_0$, provided that we can ensure that the pseudo-orbit constructed in the proof of Theorem 11 returns to that set. Below we formulate Theorem 17, which will make this statement precise. First we introduce one notion.

**Definition 16.** *Consider the topology on $\Lambda_0 \cap \{I \in [0,1]\}$ induced by $\Lambda_0$. We say that an open set $S \subset \Lambda_0 \cap \{I \in [0,1]\}$ is a strip in $\Lambda_0$ iff*

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

*(Recall that we consider $\mathbb{T} = \mathbb{R}/\mathrm{mod}\,2\pi$; 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 migh look differently, for instance as on the right hand side plot in figure 3. (Provided that conditions from below corollary are fulfilled, the result holds regardless from the shape of the 'strip'.)

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.

**Theorem 17.** *Assume that conditions (17) and (18) are satisfied, and that for $\varepsilon = 0$ we have the sequence scattering maps $\sigma_\alpha : \mathrm{dom}(\sigma_\alpha) \to \Lambda_0$ for $\alpha = 1, \ldots, L$. Let $S^+ \subset \Lambda_0$ be a strip*[7]. *Assume that for every $z \in \overline{S^+}$*

1. *there exists an $\alpha \in \{1, \ldots, L\}$ for which $z \in \mathrm{dom}(\sigma_\alpha)$,*

$$f_0^m \circ \sigma_\alpha(z) \in S^+, \tag{30}$$

[7] We 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 oposite direction.

2. *there exists a constant* $m \in \mathbb{N}$ *and a point* $x \in W^u_z(f_0, U) \cap W^s_{\sigma_\alpha(z)}(f_0)$ *such that* $f_0^m(x) \in W^s_{f_0^m(\sigma_\alpha(z))}(f_0, U)$ *and*

$$\sum_{j=0}^{m-1} \pi_I g\left(f_0^j(x)\right) - \frac{1+\lambda}{1-\lambda} L_g C > 0. \tag{31}$$

*Then for sufficiently small* $\varepsilon > 0$ *there exists an* $x_\varepsilon$ *and* $n_\varepsilon > 0$ *such that*

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

**Proof.** The result follows by making minor adjustments to the arguments in the proof of Theorem 11. So, let $S_\varepsilon^+ \subset \Lambda_\varepsilon$ be the perturbation of the strip $S^+ \subset \Lambda_0$. As in the proof of Theorem 11 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 we assume that (31) holds for any $z \in \overline{S^+}$ (with choices of $m$ and $\alpha$ depending on $z$). This means that for sufficiently small $\varepsilon$, 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, $z_\varepsilon$ 'returns' to the strip for sufficiently small $\varepsilon$. Due to the compactness of $\overline{S^+}$, a sufficiently small choice of $\varepsilon$ 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 (31) ensures that the pseudo-orbit $x_{i+1} = f_\varepsilon^{m_i} \circ \sigma_{\alpha_i}^\varepsilon(x_i)$ remains within the strip $S_\varepsilon^+$ for sufficiently small $\varepsilon$. By (32) and identical arguments to those from Theorem 11 we therefore have

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

for some $c > 0$, and the result follows from the shadowing argument just as in the proof of Theorem 11. ∎

A mirror result gives diffusion in the opposite direction.

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

1. *there exists an* $\alpha \in \{1, \ldots, L\}$ *for which* $z \in \mathrm{dom}(\sigma_\alpha)$*,*

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

2. *there exists a constant* $m \in \mathbb{N}$ *and a point* $x \in W^u_z(f_0, U) \cap W^s_{\sigma_\alpha(z)}(f_0)$ *such that* $f_0^m(x) \in W^s_{f_0^m(\sigma_\alpha(z))}(f_0, U)$ *and*

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

*Then for sufficiently small* $\varepsilon > 0$ *there exists an* $x_\varepsilon$ *and* $n_\varepsilon > 0$ *such that*

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

**Proof.** The proof follows as in the proof of Theorem 17. ∎

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

**Theorem 19.** *Assume that two strips* $S^+$ *and* $S^-$ *satisfy assumptions of Theorems 17 and 18, 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 for any given finite sequence $\{I_k\}_{k=0}^N$ and any given $\delta > 0$, for sufficiently small $\varepsilon$ there exists an orbit of $f_\varepsilon$ which $\delta$-shadows the actions $I_k$; i.e. there exists a point $z_0^\varepsilon$ and a sequence of integers $n_1^\varepsilon \leq n_2^\varepsilon \leq \ldots \leq n_N^\varepsilon$ such that*

$$\left\| \pi_I f_\varepsilon^{n_k^\varepsilon}(z_0^\varepsilon) - I_k \right\| < \delta.$$

**Proof.** Suppose that $I_1 > I_0$. (The opposite case will be analogous.) As in the proof of Theorem 17, 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 $c > 0$. We can therefore find a pseudo orbit for which $\left|\pi_I x_{i_1} - I_1\right| < \delta/2$, for some $i_1 > 0$. If $I_2 > I_1$, and we carry on as in the proof of Theorem 17, continuing with our pseudo-orbit along $S_\varepsilon^+$, until we reach $x_{i_2}$ such that $\left|\pi_I x_{i_2} - I_2\right| < \delta/2$. 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 $\varepsilon$ we will obtain that $x_{i_1+1} \in S_\varepsilon^-$. We now construct the subsequent points $x_i$ along the strip $S_\varepsilon^-$, going down in $I$ along each step, until we reach $x_{i_2}$ satisfying $\left|\pi_I x_{i_2} - I_2\right| < \delta/2$. Depending on whether $I_{k+1} > I_k$ or $I_{k+1} < I_k$ we procede in an analogous manner: to move up in $I$ we construct the given fragment of the pseudo-orbit along $S_\varepsilon^+$; and to go down in $I$ we construct the given fragment of the pseudo-orbit along $S_\varepsilon^-$. Assumptions 1., 2. ensure that our pseudo-orbit can be chosen to jump between the strips $S_\varepsilon^+$ and $S_\varepsilon^-$ at any stage of the construction.

This way we construct a pseudo orbit for which

$$\left|\pi_I x_{i_k} - I_k\right| < \delta/2 \qquad \text{for } k = 0, \ldots, N.$$

By Theorem 5 the pseudo-orbit $x_i$ can be $\delta/2$-shadowed by a true orbit, which concludes our proof. ∎

**Remark 20.** Theorems 17, 18, 19 can be generalised to the setting of higher dimensional $I$ by singling out one action, as in Remark 14. 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 19. 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 [24, 25, 26], and the second on the parameterization method [27, 28, 29].

### *4.1. The Generalized Standard Map*

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

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

**Remark 21.** *The map $f$ is symplectic and has the generating function*

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

**Remark 22.** 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 $f_\varepsilon : \mathbb{R}^2\times\mathbb{T}^2 \to \mathbb{R}^2\times\mathbb{T}^2$. (We could just as well interpret $f_\varepsilon$ 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.

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

$$f_0(x,y,\theta,I) = (F(x,y), G(\theta,I)). \tag{32}$$

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

The torus

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

is a normally hyperbolic invariant manifold for $f_0$ with the rates $\lambda$ and $\mu = \sqrt{\frac{1}{2}\sqrt{5}+\frac{3}{2}}$. (The $\mu$ is the norm of the matrix acting on $\theta, I$ in (3) for $\varepsilon = 0$. In fact, $\mu$ is the famous golden ratio.)

We consider the standard symplectic form

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

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

**Remark 23.** Note that with the coupling considered in (3), for $\varepsilon > 0$ the manifold $\Lambda_0$ remains invariant, and the dynamics on it remains unchanged. We remark that our method does not depend on this property. Rather, we have chosen such a coupling so that it is evident that it is impossible to diffuse in $I$ by using the 'inner dynamics' on the perturbed manifold. Our diffusion is driven by the 'outer dynamics' along the homoclinic connections, which is clearly visible here.

We prove the following result.

**Theorem 24 (Diffusion in the generalized standard map).** *Let $\delta$ be an arbitrary, fixed, strictly positive number. Then for every finite sequence $\left\{I^l\right\}_{l=0}^{L} \subset \left[\frac{1}{5}, \pi-\frac{1}{10}\right]$, there exists an $\varepsilon > 0$, a sequence of integers $n_1^\varepsilon,\ldots,n_L^\varepsilon$ and an orbit $z_0^\varepsilon,\ldots,z_L^\varepsilon$, $z_l^\varepsilon = f_\varepsilon^{n_l^\varepsilon}\left(z_{l-1}^\varepsilon\right)$, for $l = 1,\ldots,L$, such that*

$$\left|\pi_I z_l^\varepsilon - I_l\right| < \delta, \qquad \textit{for } l = 0,\ldots,L.$$

**Remark 25.** The proof of this theorem is based on computer assisted validation of the assumptions of Theorem 19. The strips validated by our computer program are depicted in Figure 4.

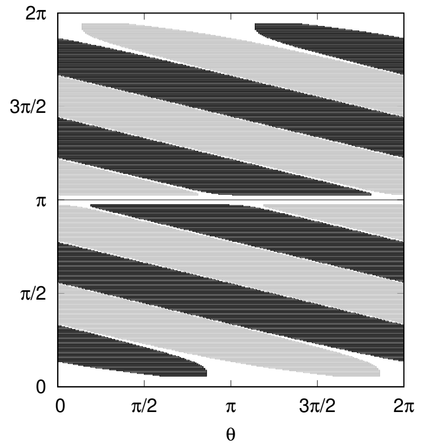


Figure 4: The strips from Theorem 19 for the map (3), validated by our computer program. The $S^+$ is in black and $S^-$ in gray. The angle $\theta$ is on the horizontal axis and $I$ on the vertical axis.

**Remark 26.** From our validation of the strips (see Figure 4) it follows also that we can take the interval $\left[\pi + \frac{1}{10}, 2\pi - \frac{1}{5}\right]$ instead of $\left[\frac{1}{5}, \pi - \frac{1}{10}\right]$ in Theorem 24. Between these two intervals though, at $I = \pi$, we have a gap, 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 27.** The diffusion is in fact be established for intervals reaching in $I$ slightly closer to 0 and $\pi$ than stated in Theorem 24, 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 24 and Remark 26.

**Remark 28.** If we take the parameter $\alpha$ in (3) closer to zero, then the unstable eigenvalues at the origin becomes smaller and the example 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, but the parametrization method can still be applied.

### *4.2. Proof of Theorem 24*

The proof of Theorem 24 exploits computer assisted validation methods for studying the local stable/unstable manifolds of fixed points. We apply these for the map $F$ from (33), 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 $P_u : J \to \mathbb{R}^2$ and $P_s : J \to \mathbb{R}^2$ 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 5. For the purpose of this section it is enough that we

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 |

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

$$\begin{aligned} \left\| F^{i}\left(P_s(x)\right)\right\| &\leq C\lambda^{i} \\ \left\| F^{-i}\left(P_u(x)\right)\right\| &\leq C\lambda^{i} \end{aligned} \qquad \text{for all } i \in \mathbb{N} \text{ and } x \in J. \tag{33}$$

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. Define $\mathcal{F}$ : $\overline{J} \times B_1 \times \ldots \times B_M \times \overline{J} \to \mathbb{R}^{2M+2}$, where $B_i \subset \mathbb{R}^2$ are cartesian products of two closed intervals, as

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

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

$$\mathcal{F}\left(p^*\right) = 0, \tag{34}$$

then we have established a sequence of points $v_0^*, \ldots, v_M^*$, where $v_1^* = P_u(x^*)$ and $v_M^* = P_s(y)$, along a homoclinic to zero. The bound on the solution of (35) can be established by using the interval Newton theorem[8]; see section 2.3. This way, we obtain a homoclinic orbit within a set of the form

$$v_i^* \in \left[x_i^0 - r, x_i^0 + r\right] \times \left[y_i^0 - r, y_i^0 + r\right] \qquad \text{for } i = 0, \ldots, M \tag{35}$$

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

$$r = r_{\text{cones}} = 1.5 \cdot 10^{-7}, \tag{36}$$

[8] An 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.

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

$$r = r_{\text{param}} = 6.5 \cdot 10^{-15}. \tag{37}$$

(The bounds on our computer program are in fact often tighter and vary from pint 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 [30] for a similar approach. We add the proof in the appendix to keep the work self-contained.)

**Lemma 29.** *The manifolds $W_0^u(F)$ and $W_0^s(F)$ intersect transversally.*

**Proof.** The proof is given in Appendix B. ■

Define the sequence

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

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

**Lemma 30.** *The set*

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

*is a homoclinic channel for $f_0$ and the associated scattering map $\sigma$ is globally defined and is the identity on $\Lambda_0$.*

**Proof.** To show that $\Gamma$ is a homoclinic channel for $f_0$ we need to prove points (i), (ii) and (iii) from Definition 3.

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

$$T_p\Gamma = \{(0,0)\} \times \mathbb{R}^2. \tag{38}$$

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

$$T_{v_0^*}W_0^s(F) \oplus T_{v_0^*}W_0^u(F) = \mathbb{R}^2, \tag{39}$$

$$T_{v_0^*}W_0^s(F) \cap T_{v_0^*}W_0^u(F) = \{0\}. \tag{40}$$

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 $p \in \Gamma$

$$T_pW_\Lambda^u(f_0) = T_{v_0^*}W_0^u(F) \times \mathbb{R}^2, \tag{41}$$

$$T_pW_\Lambda^s(f_0) = T_{v_0^*}W_0^s(F) \times \mathbb{R}^2. \tag{42}$$

From (40), (42), (43) and (41), (42), (43), (39) we obtain, respectively,

$$T_pW_\Lambda^s(f_0) \oplus T_pW_\Lambda^u(f_0) = \mathbb{R}^4,$$
$$T_pW_\Lambda^s(f_0) \cap T_pW_\Lambda^u(f_0) = \{0\} \times \mathbb{R}^2 = T_p\Gamma,$$

which proves (i) from Definition 3.

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

$$W_z^u(f_0) = W_0^u(F) \times \{\pi_{(\theta,I)}z\}, \tag{43}$$
$$W_z^s(f_0) = W_0^s(F) \times \{\pi_{(\theta,I)}z\}. \tag{44}$$

This means that the wave maps are of the form

$$\Omega_{\pm}\left(x_0^*, y_0^*, \theta, I\right) = (0, 0, \theta, I). \tag{45}$$

Clearly $\Omega_{\pm}$ are diffeomorphisms as required in (iii) from Definition 3.

From (44), (45) we see that for any $p \in \Gamma$ and $z \in \Lambda_0$

$$T_p W_z^u(f_0) = T_{v_0^*} W_0^u(F) \times \{(0,0)\}, \tag{46}$$
$$T_p W_z^s(f_0) = T_{v_0^*} W_0^s(F) \times \{(0,0)\}, \tag{47}$$

Combining (39) with (47), (48) and comparing with (42), (43) gives

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

which means that we have (ii) from Definition 3. We have established that $\Gamma$ is a homoclinic chanel. From (46) we see that the associated scattering map $\sigma$ is globally defined and is the identity on $\Lambda_0$. ■

We validate strips $S^+$ and $S^-$ with 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 $[a, b] \subset \mathbb{T}^1 = \mathbb{R}/\bmod 2\pi$ as

$$[a,b] = \begin{cases} \left\{x \in \mathbb{T}^1 : a \le x \le b\right\} & \text{if } a \le b, \\ \{x \in \mathbb{R} : b \le x \le a + 2\pi\} \bmod 2\pi & \text{if } b < a. \end{cases} \tag{48}$$

We define $(a, b) \subset \mathbb{T}^1$ as the interior of $[a, b]$.

Let $I_1, I_2 \in (0, 2\pi)$ satisfy $I_1 < I_2$. Let $s_1, s_2 \in \mathbb{T}^1$, and consider strips on $\Lambda_0$ of the form

$$\{(0,0)\} \times [s_1, s_2] \times [I_1, I_2]. \tag{49}$$

(In (50) the interval $[s_1, s_2]$ is in the sense (49).) We now have the following lemma.

**Lemma 31.** *If*

$$\sum_{i=0}^{M-1} \sin(x_i^*) \cos(\theta + iI) > 3\frac{1+\lambda}{1-\lambda}C \tag{50}$$

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

$$\theta + mI \in (s_1, s_2), \tag{51}$$

*then assumptions of Theorem 17 hold true for our map (3) on the strip (50).*

**Proof.** Condition (31) follows from (52). We need to validate (32). Since $v_M^* \in P_s(J)$, from (34) it follows that $|x_m^*| < C\lambda^{m-M}$, for $m \geq M$.

Consider an arbitrary fixed $(\theta, I) \in [s_1, s_2] \times [I_1, I_2]$ 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 $m \geq M$

$$|C_m - C_M| \leq \sum_{j=M}^{m-1} \left|\sin(x_j^*)\right| |\cos(\theta + jI)| \leq C\frac{1-\lambda^{m-M}}{1-\lambda} < C\frac{1+\lambda}{1-\lambda}. \tag{52}$$

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

For $z = (0, 0, \theta, I) \in \{(0, 0)\} \times [s_1, s_2] \times [I_1, I_2]$, consider $x = \left(x_0^*, y_0^*, \theta, I\right) \in W_z^u(f_0, U) \cap W_{\sigma_\alpha(z)}^s(f_0)$. Since $\left(x_0^*, y_0^*\right) = v_0^* \in P_u(J)$ and $v_M^* \in P_s(J)$, for every $m \geq N$, $f_0^m(x) \in W_{f_0^m(\sigma_\alpha(z))}^s(f_0, U)$. Also, for every $m \geq M$, by using (51) and (53), we obtain

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

which ensures (32). This finishes our proof. ■

**Remark 32.** A mirror result lets us validate assumptions of Theorem 18. The only difference is that instead of (51), 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 24.

**Proof of Theorem 24.** By Lemma 29 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 (36–38). This means that, by Lemma 30, the scattering map for the unperturbed system is well defined.

Using the bounds from Table 1 and (36–38), which give an enclosure of a finite fragment of the homoclinic orbit, and together with the aid of Lemma 31, our computer program constructs the strip $S^+$ from Figure 4. This strip is a union of overlapping rectangles, for which assumptions of Theorem 17 are satisfied. We use a mirror result to Lemma 31 (see Remark 32), to construct the strip $S^-$ from Figure 4, for which assumptions of Theorem 18 are satisfied. We also validate that for these two strips conditions 1. and 2. of Theorem 19 are fullfiled.

After such validation the result follows from Theorem 19. ■

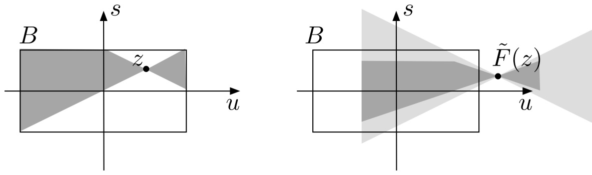


Figure 5: The cone at $z$ intersected with $B$ (in dark grey) is mapped into the cone at $\tilde{F}(z)$ (in light grey).

The computer assisted proof using cone conditions for the validation of intersections of the manifolds was performed with the CAPD[9] library [31]. The parametrisation method approach was implemented in Matlab. The source code is available on the web page of the corresponding author.

## 5. Invariant manifolds and their intersections

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

### 5.1. Cone conditions for the (un)stable manifold

The ideas presented here are based on the more general method from [24]. We reformulate the results for our particular setting, giving sketches of proofs, in order to keep the paper self-contained.

Let $F$ be the map (33), i.e. the unperturbed map $f_0$ acting on $x, y$. Let $\mathcal{P} \in \mathbb{R}^{2\times 2}$ and $\tilde{F} : \mathbb{R}^2 \to \mathbb{R}^2$ be defined as follows

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

where we recall that $\alpha = 4$. 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 diagonalizes the stable and unstable directions at the origin, i.e. $D\tilde{F}(0) = \operatorname{diag}\left((3-2\sqrt{2})^{-1}, 3-2\sqrt{2}\right)$. We refer to these as the local coordinates, as $\tilde{F}$ is expressed as $z = (u, s)$. (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 $C : \mathbb{R}^2 \to \mathbb{R}^2$ 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 5). 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 33.** *We say that $\tilde{F}$ satisfies cone conditions in $B$ if for every $z \in B$ we have (see Figure 5)*

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

[9] Computer Assisted Proofs in Dynamics: http://capd.ii.uj.edu.pl

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

**Lemma 34.** *If $\tilde{F}$ satisfies cone conditions in $B$, and there exists a $\lambda < 1$ such that for every $z \in C^{+}(0)$ we have*

$$|\pi_u \tilde{F}(u,s)| > \lambda^{-1} |u|, \tag{53}$$

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

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

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

$$\left\|\tilde{F}^{-n}(u, w(u))\right\| < \lambda^n \sqrt{1+\mathcal{L}^2}\,|u|. \tag{54}$$

**Proof.** This lemma in a slightly more general form was proven in [32]. We therefore limit ourselves to a sketch of the proof, which is given in Appendix C. ■

In practice we can validate cone conditions and (54) from the interval enclosure of the derivative of $\tilde{F}$ on $B$. We give proofs of below lemmas in the appendix.

**Lemma 35.** *If $[D\tilde{F}(B)](C^{+}(0)) \subset C^{+}(0)$ then $\tilde{F}$ satisfies cone conditions.*

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

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

**Lemma 36.** *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 (54) is fullfiled.*

Using a computer program we compute an interval enclosure $[D\tilde{F}(B)]$. This enclosure is used to validate, via Lemmas 35 and 36, the assumptions of Lemma 34. This way we obtain $w : J \to [-r\mathcal{L}, r\mathcal{L}]$, and define $P_u : J \to \mathbb{R}^2$ 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

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

Moreover, by Lemma 34 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

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

From (55) we also see that for every $x \in J$

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

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 24.

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

*5.2. Parameterization method for the (un)stable manifold with validated error bounds*

We now review a method for computing high order polynomial expansions of chart maps for the local stable/unstable manifolds, providing accurate approximations further from the fixed point. Employing these expansions improves error bounds in computer assisted proofs for connecting orbits, as it shortens the orbit segment between the local stable/unstable manifold segments. This in turn improves the condition number of the matrix appearing in the interval Newton method.

The idea behind the parameterization method is to find a chart map conjugating the given dynamics to the linear dynamics at the fixed point. The conjugacy is then used to accurately track orbits as they approach the fixed point. Since the chart map expansions are used in computer assisted proofs, it is necessary to develop explicit and mathematically rigorous bounds for the truncation errors. We also need to be able to compute rigorous enclosure on derivatives.

Our approach is adapted from more general results of [30], a work which is itself based on the parameterization method of [33, 34, 29]. We refer to the book of [35] for much more complete discussion of the parameterization method. The key is that the desired conjugacy relation is viewed nonlinear functional equation, and error bounds are obtained via fixed point arguments in appropriate function spaces. Rather than proceeding in full generality, we focus for the sake of simplicity only on the details needed in the present work.

So, suppose that $f\colon \mathbb{C}^2 \to \mathbb{C}^2$ is a function of two complex variables, analytic in an open set about the fixed point $\mathbf{z}_0$. Let $\lambda \in \mathbb{C}$ denote an unstable eigenvalue of $Df(\mathbf{z}_0)$, so that $|\lambda| > 1$, and let $\xi \in \mathbb{C}^2$ be an associated eigenvector of $\lambda$.

With $D \subset \mathbb{C}$ denoteing the unit disk in the complex plane, we look for an analytic function $P\colon D \to \mathbb{C}^2$ having that $P(0) = \mathbf{z}_0$, $P'(0) = \xi$, and that $P$ solves the invariance equation

$$f(P(\sigma)) = P(\lambda\sigma), \qquad \sigma \in D. \tag{55}$$

Such a $P$ parameterizes a local unstable manifold attached to $\mathbf{z}_0$. Indeed, the equation requires that applying the linear dynamics in $D$ is the same as applying the full dynamics on the image of $P$: that is, $P$ is a conjugacy as desired.

Since $\lambda \neq 0$ we are able to compose both sides of the invariance equation with the mapping $\sigma \mapsto \lambda^{-1}\sigma$ and obtain the equivalent fixed point problem

$$P(\sigma) = f(P(\lambda^{-1}\sigma)), \qquad \text{for } \sigma \in D. \tag{56}$$

This problem has exactly the same solution, but is better suited for a-posteriori error analysis.

Writing

$$P(\sigma) = \sum_{n=0}^{\infty} \begin{pmatrix} a_n \\ b_n \end{pmatrix} \sigma^n = \sum_{n=0}^{\infty} p_n \sigma^n,$$

we impose that $p_0 = \mathbf{z}_0$ and $p_1 = \xi$, so that $P$ satisfies the first order constraints, i.e. $P(0) = \mathbf{z}_0$ and $DP(0) = \xi$. The coefficients $p_2, p_3, p_4, \ldots$ are computed via a power matching argument, which depends strongly on the nonlinearity of $f$. See Appendix F for the derivation when $f$ is the standard map.

Let

$$P^N(\sigma) = \sum_{n=0}^{N} \begin{pmatrix} a_n \\ b_n \end{pmatrix} \sigma^n,$$

denote the approximate parameterization obtained by truncating $P$ to order $N$. That is, we suppose that the coefficients $p_0, \ldots, p_N$ are exactly the Taylor coefficients of $P$. In practice these must be computed using validated numerical methods and are known only up to interval enclosures.

Our goal is to understand the truncation error on $D$. Define the *defect* function

$$E^N(\sigma) = f[P^N(\lambda^{-1}\sigma)] - P^N(\sigma).$$

The quantity

$$\epsilon_N = \left\|E^N\right\|_0,$$

is an a-posteriori error indicator on $D$ associated with the approximation $P^N$. We note that $\epsilon_N$ is made small either by taking $N$ large or by taking $\|\xi\|$ small. In practice this is a delicate balancing act, see Remark 41 below.

Small defects do not necessarily imply small errors, and further hypotheses are needed to bound the truncation error associated with $P^N$ in terms of $\epsilon_N$. We now formulae an a-posteriori theorem, whose proof is given in Appendix H for the sake of completeness. The statement requires a little notation and a few additional definitions having to do with the Taylor remainder at $\mathbf{z}_0$.

So, for $\sigma \in \mathbb{C}$ let $|\sigma|$ denote the usual complex absolute value. We write $\mathbf{z} = (z_1, z_2) \in \mathbb{C}^2$ and endow $\mathbb{C}^2$ with the norm

$$\|\mathbf{z}\| = \max(|z_1|, |z_2|).$$

This induces a norm on the set of all $2 \times 2$ matrices with complex entries given by

$$\|A\| = \max(|a_{11}| + |a_{12}|, |a_{21}| + |a_{22}|).$$

With this norm we have that

$$\|A\mathbf{z}\| \leq \|A\|\|\mathbf{z}\|.$$

For a function $P\colon D \to \mathbb{C}^2$ we write

$$\|P\|_0 = \sup_{\sigma\in D} \|P(\sigma)\|,$$

to denote the usual supremum norm on $C^0(D, \mathbb{C}^2)$ norm. Recall that the set

$$\mathcal{D} = \left\{P\colon D \to \mathbb{C}^2 \ : \ P \text{ is analytic on } D \text{ and } \|P\|_0 < \infty\right\},$$

is a Banach space.

Fix $\mathbf{z}_0 \in \mathbb{C}^2$ and $r_*, R \in \mathbb{R}$ with $0 < r_* < R$. For $\|\mathbf{z} - \mathbf{z}_0\| < R$ and $\|\mathbf{w}\| < r_*$ write the first order Taylor expansion of $f$ as

$$f(\mathbf{z} + \mathbf{w}) = f(\mathbf{z}) + Df(\mathbf{z})\mathbf{w} + \mathcal{R}_{\mathbf{z}}(\mathbf{w}).$$

Here $\mathcal{R}_{\mathbf{z}}(\cdot)$ – the first order Taylor remainder at $\mathbf{z}$ – is analytic in both $z$ and $w$.

By the Taylor remainder theorem there are constants $0 < C_1, C_2$ with

$$\sup_{\|\mathbf{z}-\mathbf{z}_0\|\leq R} \|\mathcal{R}_{\mathbf{z}}(\mathbf{w})\| \leq C_1\|\mathbf{w}\|^2, \qquad \text{for } \|\mathbf{w}\| \leq r_*, \tag{57}$$

and

$$\sup_{\|\mathbf{z}-\mathbf{z}_0\|\leq R} \|D\mathcal{R}_{\mathbf{z}}(\mathbf{w})\| \leq C_2\|\mathbf{w}\|, \qquad \text{for } \|\mathbf{w}\| \leq r_*. \tag{58}$$

There is also a $C_3 > 1$ having

$$\sup_{\|\mathbf{z}-\mathbf{z}_0\|\leq R} \|Df(\mathbf{z})\| \leq C_3. \tag{59}$$

If $f$ is entire then $\mathcal{R}_z(w)$ is entire in both $\mathbf{z}$ and $\mathbf{w}$, and explicit constants $C_1, C_2, C_3$ for the standard map are derived in Appendix G. In general what is needed is that $f$ is analytic on a ball about $\mathbf{z}_0$ of radius $R + r_*$.

**Theorem 37 (A-posteriori error bounds for Equation** (56)**).** *Suppose that $f\colon \mathbb{C}^2 \to \mathbb{C}^2$ fixes the point $\mathbf{z}_0 \in \mathbb{C}^2$. Let $0 < r_* < R$ and suppose that and that $f$ is analytic for all $\mathbf{z} \in \mathbb{C}^2$ with*

$$\|\mathbf{z} - \mathbf{z}_0\| < R + r_*.$$

*Assume that $Df(\mathbf{z}_0)$ has a single unstable eigenvalue, denoted by $\lambda$, and let $\mu = \lambda^{-1}$ and $\xi \in \mathbb{C}^2$ be an eigenvector associated with $\lambda$. Let $P$ denote the solution of Equation* (56) *on the unit disk $D \subset \mathbb{C}$, and let $p_0, \ldots, p_N \in \mathbb{C}^2$ denote the zeroth through $N$-th order power series coefficients of $P$ subject to the constraints $p_0 = \mathbf{z}_0$ and $p_1 = \xi$.*

*Let $C_1$, $C_2$, and $C_3$ be as defined in Equations* (58)*,* (59)*, and* (60)*, and assume tha they satisfy*

$$\sum_{n=1}^{N} \mu^n \|p_n\| < R, \tag{60}$$

*and*

$$4C_2|\mu|^{2(N+1)} < \left(1 - C_3|\mu|^{N+1}\right)^2. \tag{61}$$

*If $r > 0$ has*

$$r \leq r_*, \tag{62}$$

*and*

$$C_2|\mu|^{2(N+1)}r^2 - \left(1 - C_3|\mu|^{N+1}\right) r + \epsilon_N < 0, \tag{63}$$

*then*

$$\sup_{|\sigma|\leq 1} \left\|P(\sigma) - P^N(\sigma)\right\| \leq r.$$

**Remark 38 (Existence of an $r > 0$ satisfying the hypotheses of Equation** (64).**).** Observe that

$$p(r) = C_2|\mu|^{2(N+1)}r^2 - \left(1 - C_3|\mu|^{N+1}\right) r + \epsilon_N,$$

is a quadratic polynomial with $p(0) = \epsilon_N > 0$ and $p'(0) = -\left(1 - C_3|\mu|^{N+1}\right) < 0$. If the discriminant condition hypothesized in Equation (62) is met, then $p(r)$ has two positive real roots $r_\pm$, and moreover $p(r)$ is negative for all $r \in (r_-, r_+)$. Given the problem data $C_2$, $C_3$, $|\mu|$, $N$, and $\epsilon_N$, finding an appropriate value of $r$ is a matter of solving a quadratic equation.

**Remark 39 (Stable manifold parameterization).** The standard map is symplectic, so that if $|\lambda| < 1$ is a stable eigenvalue of $Df(z_0)$ then $1/\lambda$ is an unstable eigenvalue. The standard map is entire with entire inverse, and the theorem above applies to the unstable manifold of $f^{-1}$. A more general alternative, which does not require invertibility of $f$, is to study separately the equation

$$P(\lambda\sigma) = f(P(\sigma)),$$

and develop analogous a-posteriori analysis for this equation. See [30].

**Remark 40 (Real invariant manifolds).** Suppose that $f$ is real valued for real inputs, as is the case for the standard map. Then $p_0$, is real, and we are interested in the real image of $P$. In the case considered in the present work – that of a real hyperbolic saddle – the eigenvalues and eigenvectors are real also. It follows that solutions $a_n$ and $b_n$ of Equation (F.5) are real at all orders. Taking real values of $\sigma$ provides the parameterization of the real stable manifold for $f$ at $p_0$, and treating $\sigma$ as a complex variable is a convenience which facilitates the use of analytic function theory in the error analysis.

**Remark 41 (Scaling the eigenvector).** The coefficients $p_2$, $p_3$, $p_4$, . . ., and hence the solution $P(\sigma)$, are only unique up to the choice of the scaling of the eigenvector. This is seen explicitly in Appendix F, where the formal series solution of Equation (57) is derived. This lack of uniqueness is exploited in numerical calculations, providing control over the growth rate of the coefficient sequence. Algorithms for determining optimal scalings are discussed in [36]. In the present work we determine good scalings through numerical experimentation. More precisely, we fix at the start of the calculation the order of approximation $N$. Then we adjust the scaling of the eigenvector so that coefficients of order $N$ are roughly of size machine epsilon, as the magnitude of the $N$-th power series coefficient of $P$ serves as a good heuristic indicator of the size of the truncation error.

**Remark 42 (Bounds on derivatives).** Suppose that $P(\sigma) = P^N(\sigma) + H(\sigma)$ for $|\sigma| < 1$, with $\|H\|_0 \leq r$. Then $P$ is differentiable on $D$ with

$$\frac{d}{d\sigma} P(\sigma) = \frac{d}{d\sigma} P^N(\sigma) + \frac{d}{d\sigma} H(\sigma),$$

for all $\sigma \in D$. Here $P^N(\sigma)$ is a polynomial whose derivative is given by the standard formula. The derivative of $H$ is bounded on any smaller disk thanks to the Cauchy bound

$$\sup_{|\sigma| \leq e^{-\nu}} \left\| \frac{d}{d\sigma} H(\sigma) \right\| \leq \frac{2\pi}{\nu} \|H\|_0 \leq \frac{2\pi}{\nu} r, \tag{64}$$

where $\nu > 0$. A proof is in [30].

In practice finding $\epsilon_N$ requires a bound on the tail of $f(P^N(\sigma))$, and this will of course depend on the explicit form of the map $f$. For example, if $f$ is the Chirikov Standard Map

$$f(z_1, z_2) = \begin{pmatrix} z_1 + z_2 + \alpha \sin(z_1) \\ z_2 + \alpha \sin(z_1) \end{pmatrix}, \tag{65}$$

then we write

$$f\left(P^N\left(\lambda^{-1}\sigma\right)\right) = \sum_{n=0}^{\infty} f_n \sigma^n,$$

and observe that we have to study the term $[\sin(P_1^N(\lambda^{-1}\sigma))]_n$ – the Taylor coefficients of the composition of $P^N$ with the sine function. Indeed, we have that

$$\begin{aligned} E^N(\sigma) &= f(P^N(\lambda^{-1}\sigma)) - P^N(\sigma) \\ &= \sum_{n=N+1}^{\infty} f_n \sigma^n = \alpha \sum_{n=N+1}^{\infty} [\sin(P_1^N(\lambda^{-1}\sigma))]_n \begin{bmatrix} 1 \\ 1 \end{bmatrix} \sigma^n, \end{aligned}$$

as the lower order terms cancel exactly by hypothesis, and the linear operations on $P^N$ do not contribute to the tail of the series. Then by taking

$$\epsilon_N = \sum_{n=N+1}^{\infty} \|f_n\| = \sum_{n=N+1}^{\infty} |\alpha| \left|[\sin(P_1^N(\lambda^{-1}\sigma))]_n\right|$$

we obtain

$$\|E^N\|_0 \leq \epsilon_N,$$

as needed.

However the Taylor series expansion of $\sin(P_1^N(\lambda^{-1}\sigma))$ is an infinite series, even though $P^N$ is a polynomial, and bounding the tail is not a finite calculation. The following Lemma, whose proof is found in Appendix I, exploits the fact that $\sin(P_1^N(\lambda^{-1}\sigma))$ is the solution of a certain linear differential equation involving only the known data $P^N$. This analysis reduces the necessary bound to a finite sum.

**Lemma 43.** *Suppose that $g^N : \mathbb{C} \to \mathbb{C}$ is an $N$-th order polynomial denoted by*

$$g^N(\sigma) = \sum_{n=0}^{N} \beta_n \sigma^n.$$

*We write*

$$c(\sigma) = \cos\left(g^N(\sigma)\right) = \sum_{n=0}^{\infty} c_n \sigma^n,$$

*and*

$$s(\sigma) = \sin\left(g^N(\sigma)\right) = \sum_{n=0}^{\infty} s_n \sigma^n,$$

*to denote the power series of the compositions with sine and cosine. Let*

$$s^N(\sigma) = \sum_{n=0}^{N} s_n \sigma^n,$$

*and*

$$c^N(\sigma) = \sum_{n=0}^{N} c_n \sigma^n,$$

*be the Taylor polynomials to $N$-th order, where recursion relations for the coefficients $s_n$ and $c_n$ are worked out via power matching in Appendix F. Let*

$$\hat{K} = \sum_{n=0}^{N-1} (n+1)|\beta_{n+1}|,$$

*and*

$$e_N = \max\left( \sum_{n=N+1}^{2N} \left| \sum_{k=0}^{n-1} \frac{k+1}{n} s_{n-k-1}\beta_{k+1} \right|, \sum_{n=N+1}^{2N} \left| \sum_{k=0}^{n-1} \frac{k+1}{n} c_{n-k-1}\beta_{k+1} \right| \right).$$

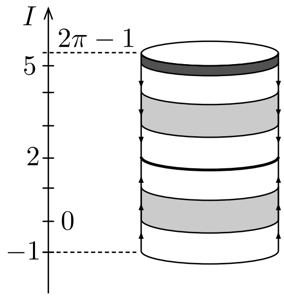


Figure A.6: For $I \in [0, 1]$ the system is not modified (bottom grey area). In the white regions the system is modified by the 'bump' function to allow for gluing at $I = 2$ and $I = -1$. The system on $I \in [2, 5]$ is a 'flipped copy' of the system on $I \in [-1, 2]$. For $I \in [5, 2\pi - 1]$ we 'freeze' $I = -1 = 2\pi - 1$.

*Assume that*

$$\frac{\hat{K}}{N+2} < 1.$$

*Then the truncation error on the unit disk D satisfies*

$$\sup_{|\sigma|\le 1} \| \sin(g^N(\sigma)) - s^N(\sigma)\| \le \frac{e_N}{1 - \frac{\hat{K}}{N+2}}, \tag{66}$$

*and similarly*

$$\sup_{|\sigma|\le 1} \| \cos(g^N(\sigma)) - c^N(\sigma)\| \le \frac{e_N}{1 - \frac{\hat{K}}{N+2}}. \tag{67}$$

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

Consider a family of maps $f_\varepsilon : \mathbb{R}^{2d} \times \mathbb{R} \times \mathbb{T}^1 \to \mathbb{R}^{2d} \times \mathbb{R} \times \mathbb{T}^1$ for which $\pi_I f_0(x) = \pi_I x$. We now modify the family to $\tilde{f}_\varepsilon : \mathbb{R}^{2d} \times \mathbb{T}^2 \to \mathbb{R}^{2d} \times \mathbb{T}^2$ so that

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

Before writing out the slightly technical formulae we first explain the idea, which is depicted in Figure A.6. 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 $\tilde{f}_\varepsilon = f_0$. Then for $I \in [2, 5]$ we 'flip' the system and glue at $I = 2$. For the remaining $I \in [5, 2\pi - 1]$ we 'freeze' the system taking $f_0$ with $I = -1 = 2\pi - 1$.

The technical details of how $\tilde{f}_\varepsilon$ is chosen are of secondary importance. The important issue is that if we prove diffusion in $I$ over the range $I \in [0, 1]$ for $\tilde{f}_\varepsilon$, then this implies diffusion for $f_\varepsilon$, as for $I \in [0, 1]$ the systems are the same. Since all the assumptions of Theorem 11 are for $\varepsilon = 0$, the discussion here is of an abstract nature. The assumptions of Theorems 11 need to be validated for $f_0$ over $I \in [0, 1]$, and then we just need to keep in mind that the construction mentioned here is valid, but it does not need to be performed in practice.

We now write out the details. Consider a smooth 'bump' function[10] $b : \mathbb{R} \to [0, 1]$ for which

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

let $g(u, s, I, \theta) := (u, s, 4 - I, \theta)$, $h(u, s, I, \theta) := (u, s, -1, \theta)$ and take

$$\tilde{f}_\varepsilon(x) = \begin{cases} f_0(x) + b(\pi_I x)(f_\varepsilon(x) - f_0(x)) & \text{for } \pi_I x \in [-1, 2], \\ f_0(g(x)) + b(4 - \pi_I x)(f_\varepsilon(g(x)) - f_0(g(x))) & \text{for } \pi_I x \in [2, 5], \\ f_0(h(x)) & \text{for } \pi_I x \in [5, 2\pi - 1]. \end{cases}$$

**Remark 44.** A similar construction works in the case when $I$ is higher dimensional; say $I \in [0, 1]^k$. In this case we control one action as mentioned in Remarks 14, 20, but we need to make sure that the strips for Theorems 17, 18, 19 do not intersect the boundary of $\{I \in [0, 1]^k\}$ on any action coordinate, except the one action which we control. Otherwise the dynamics could escape $\{I \in [0, 1]^k\}$ through the remaining actions, and the results obtained for the artificial 'glued' system would not need to be realized by the true system.

## Appendix B. Proof of Lemma 29

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 $v_M^* = F^M(P_u(x^*))$. Taking $w_0 = DP_u(x^*) \in \mathbb{R}^2$ and $w_k = DF\left(v_k^*\right) w_{k-1} \in \mathbb{R}^2$ we see that

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

If $w_{M-1}$ was colinear 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 $V = (1, w_0, \ldots, w_{M-1}, 1/\alpha)$ would lead to

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

This is a contradiction, since if $p^*$ is validated by the use of Theorem 9, so the matrix $D\mathcal{F}(p^*)$ must be invertible.

## Appendix C. Proof of Lemma 34

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 (54) ensure that the unstable manifold is streched 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 (55) 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 (54),

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

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

Taking $z = \tilde{F}^{-n}(w(u))$ and using (54) we obtain

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

as required.

## Appendix D. Proof of Lemma 35

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

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

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

## Appendix E. Proof of Lemma 36

Let $(u,s) \in C^+(0) \cap B$. From a mirror argument to (D.1) and since $|s| \le \mathcal{L}|u|$,

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

as required.

## Appendix F. Formal series calculations

Suppose that

$$p_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

is a hyperbolic fixed point of the Standard map $f$ (defined in Equation (66)), that $\lambda \in \mathbb{R}$ is an eigenvalue of $Df(p_0)$, and that $\xi \in \mathbb{R}^2$ is an associated eigenvector.

Let

$$P(\sigma) = \begin{pmatrix} X(\sigma) \\ Y(\sigma) \end{pmatrix} = \sum_{n=0}^{\infty} \begin{pmatrix} a_n \\ b_n \end{pmatrix} \sigma^n.$$

and note that

$$\begin{pmatrix} a_0 \\ b_0 \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = p_0,$$

and

$$\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = \xi.$$

While the equations

$$f(P(\sigma)) = P(\lambda\sigma) \qquad \text{or} \qquad f(P(\lambda^{-1}\sigma)) = P(\sigma),$$

have the same solutions, the form of the second equaiton is better suited to error analysis while the first is easier to work with from the perspective of formal series calculations.

Observe that while

$$P(\lambda\sigma) = \sum_{n=0}^{\infty} \lambda^n \begin{pmatrix} a_n \\ b_n \end{pmatrix} \sigma^n,$$

computationg the power series coefficients of $f(P(\sigma))$ is more delicate, due to the appearance of the composition term $\sin(X(s))$. To work out the power series of the composition let

$$s(\sigma) = \sin(X(\sigma)) = \sum_{n=0}^{\infty} s_n \sigma^n,$$

and

$$c(\sigma) = \cos(X(\sigma)) = \sum_{n=0}^{\infty} c_n \sigma^n,$$

where we note that $s_n$ and $c_n$ depend on the $a_n$ and $b_n$. Indeed, to first order we have that

$$s_0 = \sin(X(0)) = \sin(a_0), \qquad \text{and} \qquad c_0 = \cos(a_0).$$

Differentiating $s(\sigma)$ and $c(\sigma)$ leads to

$$s'(\sigma) = \cos(X(\sigma))X'(\sigma) = c(\sigma)X'(\sigma), \tag{F.1}$$

and

$$c'(\sigma) = -\sin(X(\sigma))X'(\sigma) = -s(\sigma)X'(\sigma), \tag{F.2}$$

and evaluating at $\sigma = 0$ gives

$$s_1 = c_0 a_1, \qquad \text{and} \qquad c_1 = -s_0 a_1.$$

To work out the higher order terms, we expand Equations (F.1) and (F.2) as power series and obtain that

$$\begin{aligned} \sum_{n=0}^{\infty}(n+1)s_{n+1}\sigma^n &= \left(\sum_{n=0}^{\infty} c_n\sigma^n\right)\left(\sum_{n=0}^{\infty}(n+1)a_{n+1}\sigma^n\right) \\ &= \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}(k+1)c_{n-k}a_{k+1}\right)\sigma^n, \end{aligned}$$

and

$$\begin{aligned} \sum_{n=0}^{\infty}(n+1)c_{n+1}\sigma^n &= -\left(\sum_{n=0}^{\infty} s_n\sigma^n\right)\left(\sum_{n=0}^{\infty}(n+1)a_{n+1}\sigma^n\right) \\ &= -\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}(k+1)s_{n-k}a_{k+1}\right)\sigma^n, \end{aligned}$$

Matching like powers (and reindexing) leads to

$$s_n = \frac{1}{n}\sum_{k=0}^{n-1}(k+1)c_{n-k-1}a_{k+1} = c_0 a_n + \frac{1}{n}\sum_{k=0}^{n-2}(k+1)c_{n-k-1}a_{k+1}, \tag{F.3}$$

and

$$c_n = \frac{-1}{n}\sum_{k=0}^{n-1}(k+1)s_{n-k-1}a_{k+1} = -s_0 a_n + \frac{-1}{n}\sum_{k=0}^{n-2}(k+1)s_{n-k-1}a_{k+1}, \tag{F.4}$$

for $n \geq 2$. Note that the sums on the right hand sides depend only on terms of order less than $n$.

Then

$$\begin{aligned}
f(P(\sigma)) &= f(X(\sigma), Y(\sigma)) \\
&= \begin{pmatrix} X(\sigma) + Y(\sigma) + \alpha\sin(X(\sigma)) \\ Y(\sigma) + \alpha\sin(X(\sigma)) \end{pmatrix} \\
&= \sum_{n=0}^{\infty}\begin{pmatrix} a_n + b_n + \alpha s_n \\ b_n + \alpha s_n \end{pmatrix}\sigma^n \\
&= \sum_{n=0}^{\infty}\begin{pmatrix} a_n + b_n + \alpha\left(c_0 a_n + \frac{1}{n}\sum_{k=0}^{n-2}(k+1)c_{n-k-1}a_{k+1}\right) \\ b_n + \alpha\left(c_0 a_n + \frac{1}{n}\sum_{k=0}^{n-2}(k+1)c_{n-k-1}a_{k+1}\right) \end{pmatrix}\sigma^n
\end{aligned}$$

Setting this last sum equal to $P(\lambda\sigma)$ and matching like powers leads to

$$\begin{pmatrix} a_n + b_n + \alpha c_0 a_n + \frac{\alpha}{n}\sum_{k=0}^{n-2}(k+1)c_{n-k-1}a_{k+1} \\ b_n + \alpha c_0 a_n + \frac{\alpha}{n}\sum_{k=0}^{n-2}(k+1)c_{n-k-1}a_{k+1} \end{pmatrix} = \lambda^n\begin{pmatrix} a_n \\ b_n \end{pmatrix},$$

or, upon rearranging

$$\begin{bmatrix} 1 + \alpha c_0 - \lambda^n & 1 \\ \alpha c_0 & 1 - \lambda^n \end{bmatrix}\begin{pmatrix} a_n \\ b_n \end{pmatrix} = \frac{-\alpha}{n}\sum_{k=0}^{n-2}(k+1)c_{n-k-1}a_{k+1}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tag{F.5}$$

for $n \geq 2$. Observe that the right hand side of the equation does not depend on $a_n$. Indeed, the right hand side depends only on $c_1, \ldots, c_{n-1}$, and $a_1, \ldots, a_{n-1}$.

Moreover, noting that

$$\begin{bmatrix} 1 + \alpha c_0 - \lambda^n & 1 \\ \alpha c_0 & 1 - \lambda^n \end{bmatrix} = \begin{bmatrix} 1 + \alpha\cos(x_0) - \lambda^n & 1 \\ \alpha\cos(x_0) & 1 - \lambda^n \end{bmatrix} = Df(p_0) - \lambda^n \mathrm{Id},$$

and observing that for $n \geq 2$, $\lambda^n$ is never an eigenvalue of $Df(p_0)$ (as $|\lambda| \neq 1$ by hyperbolicity), we see that Equation (F.5) is always uniquely solvable. Hence the coefficients of $P$ are formally well defined to all orders. Observe also that once we solve Equation (F.5) for $a_n$ and $b_n$ we compute $s_n$, and $c_n$ using Equations (F.3) and (F.4), as this information is needed to solve the homological equations at order $n+1$.

## Appendix G. Explicit constants for parameterization of the standard map

The next lemmas provide constants $C_1, C_2, C_3$ when $f$ is the standard map defined in Equation (66).

**Lemma 45 (Explicit constants for the Standard Map).** *Consider $f\colon \mathbb{C}^2 \to \mathbb{C}^2$ as given in Equation (66). Recall that $\lambda$ is an unstable eigenvalue and that $\mu = \lambda^{-1}$. Let $R$ and $r_*$ be* *constants satisfying* $0 < r_* < R$. *If we choose*

$$\begin{aligned} C_1 &= \frac{|\alpha|\, e^R\, (e^{r_*} + 1)}{2}, \\ C_2 &= |\alpha|\, e^R\, (e^{r_*} + 1)\,, \\ C_3 &= 2 + |\alpha|\, e^R, \end{aligned}$$

*then (58–60) are satisfied.*

**Proof.** To prove the lemma, start by expanding $f(z_1 + w_1, z_2 + w_2)$ to find that

$$\mathcal{R}_{(z_1,z_2)}(w_1, w_2) = \alpha\, [\sin(z_1)(\cos(w_1) - 1) + \cos(z_2)(\sin(w_1) - w_1)] \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

is the first order Taylor remainder for the standard map as a function of the base point $(z_1, z_2)$. Recalling that $z_1, z_2, w_1, w_2 \in \mathbb{C}$, the bounds $C_1$ and $C_2$ follow immediately from the identities $\sin(z) = (e^z - e^{-z})/2i$ and $\cos(z) = (e^z + e^{-z})/2$. Indeed we have that

$$\begin{aligned} |\cos(z)| &\leq \frac{e^{|z|} + 1}{2} \leq e^{|z|}, \\ |\sin(z)| &\leq \frac{e^{|z|} + 1}{2} \leq e^{|z|}, \\ |\cos(z) - 1| &\leq \frac{|z|^2}{4}(e^{|z|} + 1), \end{aligned}$$

and

$$|\sin(z) - z| \leq \frac{|z|^3}{12}(e^{|z|} + 1).$$

The form of $C_1$ follows directly.

Differentiating with respect to $\mathbf{w} = (w_1, w_2)$ gives that

$$D\mathcal{R}_{z_1,z_2}(w_1, w_2) = \alpha(-\sin(z_1)\sin(w_1) + \cos(z_1)(\cos(w_1) - 1)) \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix},$$

from which follows the form of $C_2$.

Finally, for the standard map we have that

$$Df(z_1, z_2) = \begin{bmatrix} 1 + \alpha\cos(z_1) & 1 \\ \alpha\cos(z_1) & 1 \end{bmatrix},$$

and the bound on $C_3$ follows from the formula for the matrix norm. ■

Observing that

$$f^{-1}(z_1 + w_1, z_2 + w_2) = f^{-1}(z_1, z_2) + Df^{-1}(z_1, z_2) \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} + \tilde{\mathcal{R}}_{(z_1,z_2)}(w_1, w_2),$$

with

$$\tilde{\mathcal{R}}_{(z_1,z_2)}(w_1, w_2) =$$
$$\begin{pmatrix} 0 \\ -\alpha\sin(z_1 - z_2)\,(\cos(w_1 - w_2) - 1) - \alpha\cos(z_1 - z_2)\,(\sin(w_1 - w_2) - (w_1 - w_2)) \end{pmatrix},$$

and that

$$Df^{-1}(z_1, z_2) = \begin{pmatrix} 1 & -1 \\ -\alpha\cos(z_1 - z_2) & 1 + \alpha\cos(z_1 - z_2) \end{pmatrix},$$

gives the following bounds.

**Lemma 46 (Explicit constants for the inverse Standard Map).** *Let $f^{-1}\colon \mathbb{C}^2 \to \mathbb{C}^2$ denote the inverse of the standard map. Choosing*

$$\begin{aligned} C_1 &= \frac{|\alpha| e^{2\tilde{M}}(e^{2r_*} + 1)}{2}, \\ C_2 &= |\alpha| e^{2\tilde{M}}(e^{2r_*} + 1), \\ C_3 &= \max\left(2, 1 + 2|\alpha| e^{\tilde{M}}\right), \end{aligned}$$

*gives that*

$$\begin{aligned} \sup_{\|\mathbf{z}-\mathbf{z}_0\| \le R} \|\tilde{\mathcal{R}}_{\mathbf{z}}(\mathbf{w})\| &\le C_1 \|\mathbf{w}\|^2, && \textit{for} \quad \|\mathbf{w}\| \le r_*, \\ \sup_{\|\mathbf{z}-\mathbf{z}_0\| \le R} \|D\tilde{\mathcal{R}}_{\mathbf{z}}(\mathbf{w})\| &\le C_2 \|\mathbf{w}\|, && \textit{for} \quad \|\mathbf{w}\| \le r_*, \end{aligned}$$

*and*

$$\sup_{\|\mathbf{z}-\mathbf{z}_0\| \le R} \|Df^{-1}(\mathbf{z})\| \le C_3.$$

## Appendix H. Proof of Theorem 37

The idea behind the proof is to write $P(\sigma) = P^N(\sigma) + H(\sigma)$ where $H$ is analytic on $D$, and to rewrite (57) as a fixed point problem for $H$.

Since the coefficients of $P^N$ are exactly the Taylor coefficients of $P$, we have that

$$H(0) = \frac{d}{d\sigma} H(0) = \ldots = \frac{d^N}{d\sigma^N} H(0) = 0.$$

That is, the truncation error function $H$ is zero to order $N$ at $\sigma = 0$. We refer to $H$ as an *analytic $N$-tail*, and let

$$\mathcal{X} = \left\{ H\colon D \to \mathbb{C}^2 :\ H \text{ is analytic}, H(0) = \ldots = H^N(0) = 0, \|H\|_0 < \infty \right\}, \tag{H.1}$$

denote the Banach space of all bounded analytic $N$-tails endowed with the $C^0$ norm. For a linear operator $\mathcal{M}\colon \mathcal{X} \to \mathcal{X}$ we write

$$\|\mathcal{M}\|_{B(\mathcal{X})} = \sup_{\|H\|_0 = 1} \|\mathcal{M}(H)\|_0,$$

to denote the operator norm. The collection of all bounded linear operators (bounded in the $\|\cdot\|_{B(\mathcal{D})}$) is denoted $B(\mathcal{D})$, and is a Banach algebra.

Using the first order Taylor expansion we rewrite the invariance equation (57) as

$$\begin{aligned} P^N(\sigma) + H(\sigma) &= f(P^N(\mu\sigma) + H(\mu\sigma)) \\ &= f(P^N(\mu\sigma)) + Df(P^N(\mu\sigma))H(\mu\sigma) + \mathcal{R}_{P^N(\mu\sigma)}(H(\mu\sigma)). \end{aligned}$$

Define

$$E^N(\sigma) = f(P^N(\mu\sigma)) - P^N(\sigma),$$

and note that $E^N$ is an analytic $N$-tail. Moreover, $E^N$ does not depend on $H$. Rearranging leads to the fixed point problem

$$H(\sigma) = E^N(\sigma) + Df(P^N(\mu\sigma))H(\mu\sigma) + \mathcal{R}_{P^N(\mu\sigma)}(H(\mu\sigma)), \tag{H.2}$$

for the truncation error. Observe that if $H \in \mathcal{X}$ is a solution of Equation (H.2) with $\|H\|_0 \leq r$ then $P = P^N + H$ solves Equation (57) and has

$$\left\|P - P^N\right\|_0 = \|H\|_0 \leq r.$$

Since $H \in \mathcal{X}$ and $P^N$ is entire, $P$ is analytic on $D$.

Writing $P^N(\mu\sigma) = \mathbf{z}$, we see that the condition given in Equation (61) gives

$$\|\mathbf{z} - \mathbf{z}_0\| = \|P^N(\mu\sigma) - p_0\| \leq \sum_{n=1}^{N} \mu^n \|p_n\| \leq R,$$

so that for all $H \in \mathcal{X}$ with $\|H\| \leq r_*$ the estimates of Equations (58), (59), and (60) give us

$$\|\mathcal{R}_{P_N(\mu\sigma)}(H)\|_0 \leq C_1 \|H\|_0^2,$$

and

$$\|D\mathcal{R}_{P^N(\mu\sigma)}(H)\|_{B(\mathcal{X})} \leq C_2 \|H\|_0.$$

Define the linear operator $\mu \colon \mathcal{X} \to \mathcal{X}$ by

$$\mu(H)(\sigma) = H(\mu\sigma).$$

The main estimate is that

$$\|\mu\|_{B(\mathcal{X})} \leq |\mu|^{N+1}.$$

To see this note that for any $H \in \mathcal{X}$ we have that

$$\sup_{\sigma \in D} \|H(\mu\sigma)\| \leq |\mu|^{N+1} \sup_{\sigma \in D} \|H(\sigma)\|, \tag{H.3}$$

by the maximum modulus principle.

Now define the fixed point operator $\Psi \colon \mathcal{X} \to \mathcal{X}$ by

$$\Psi[H](\sigma) = E^N(\sigma) + Df(P^N(\mu\sigma))H(\mu\sigma) + \mathcal{R}_{P^N(\mu\sigma)}(H(\mu\sigma)).$$

Let

$$B_r = \{H \in \mathcal{D} \,:\, \|H\|_0 \leq r\}.$$

We will show that $\Psi$ has a unique fixed point in $B_r \subset \mathcal{X}$ using the contraction mapping theorem.

Observe first that $\Psi$ is Fréchet differentiable with

$$D\Psi[H_1]H_2 = Df(P^N(\mu\sigma))H_2(\mu\sigma) + D\mathcal{R}_{P^N(\mu\sigma)}(H_1(\mu\sigma))H_2(\mu\sigma),$$

for all $H_1, H_2 \in \mathcal{X}$. From the inequality hypothesized in Equation (64) we obtain that

$$C_2|\mu|^{2(N+1)} r^2 + C_3 |\mu|^{N+1} r + \epsilon_N < r, \tag{H.4}$$

by adding $r$ to both sides of Equation (64). Dividing the inequality in Equation (H.4) by $r > 0$ leads to

$$C_2|\mu|^{2(N+1)}r + C_3|\mu|^{N+1} + \frac{\epsilon_N}{r} < 1,$$

and since each term in the sum is positive we have that

$$C_2|\mu|^{2(N+1)}r^2 + C_3|\mu|^{N+1}r < 1. \tag{H.5}$$

Now for $H \in \overline{B_r}$ we have that

$$\begin{aligned}
\|D\Psi[H]\|_{B(X)} &\leq \|Df(P^N(\mu\sigma))\mu + D\mathcal{R}_{P^N(\mu\sigma)}(H(\mu\sigma))\mu\|_{B(X)} \\
&\leq C_3\|\mu\|_{B(X)} + C_2|\mu|^{N+1}\|H\|_0\|\mu\|_{B(X)} \\
&\leq C_3|\mu|^{N+1} + C_2|\mu|^{2(N+1)}r \\
&< 1,
\end{aligned}$$

by the inequality given in Equation (H.5). It follows from the Mean Value Inequality that $\Psi$ is a contraction on $\overline{B_r}$.

Finally, to verify that $\Psi$ maps $\overline{B_r}$ into itself choose $H \in \overline{B_r}$ and note that

$$\begin{aligned}
\|\Psi[H]\|_0 &\leq \|\Psi[H] - \Psi[0]\|_0 + \|\Psi[0]\|_0 \\
&\leq \sup_{\|V\|_0 \leq r} \|D\Psi[V]\|\|H\|_0 + \|E^N\|_0 \\
&\leq \left(C_3|\mu|^{N+1} + C_2|\mu|^{2(N+1)}r\right) r + \epsilon_N \\
&\leq C_2|\mu|^{2(N+1)}r^2 + C_3|\mu|^{N+1}r + \epsilon_N \\
&\leq r,
\end{aligned}$$

again by the mean value inequality and the bound given in Equation (H.5). Then $\Psi$ is a contraction mapping from $\overline{B_r}$ into itself, and the proof is complete.

## Appendix I. Proof of Lemma 43

Define the function $Q\colon D \to \mathbb{C}$ by

$$Q(\sigma) = \frac{d}{d\sigma} g^N(\sigma) = \sum_{n=0}^{N-1} (n+1)\beta_{n+1}\sigma^n.$$

The main observation is that $s(\sigma), c(\sigma)$ solve the system of differential equations

$$\begin{aligned}
s' &= cQ \\
c' &= -sQ
\end{aligned}$$

subject to the initial conditions

$$s(0) = s_0 = \sin(g^N(0)) = \sin(\beta_0), \qquad \text{and} \qquad c(0) = c_0 = \cos(g^N(0)) = \cos(\beta_0).$$

Integrating leads to the system of equations

$$
\begin{aligned}
s(\sigma) &= s_0 + \int_0^\sigma c(z)Q(z)\,dz,\\
c(\sigma) &= c_0 - \int_0^\sigma s(z)Q(z)\,dz.
\end{aligned}
\tag{I.1}
$$

We now write $s = s^N + s^\infty$ and $c = c^N + c^\infty$ with $(s^\infty, c^\infty) \in \mathcal{X}$ (the space of analytic $N$-tails – see Equation (H.1)).Then

$$
\begin{aligned}
s^N(\sigma) + s^\infty(\sigma) &= s_0 + \int_0^\sigma c^N(z)Q(z)\,dz + \int_0^\sigma c^\infty(z)Q(z)\,dz,\\
c^N(\sigma) + c^\infty(\sigma) &= c_0 - \int_0^\sigma s^N(z)Q(z)\,dz - \int_0^\sigma s^\infty(z)Q(z)\,dz,
\end{aligned}
$$

or

$$
\begin{aligned}
s^\infty(\sigma) - \int_0^\sigma c^\infty(z)Q(z)\,dz &= -s^N(z) + s_0 + \int_0^\sigma c^N(z)Q(z)\,dz,\\
c^\infty(\sigma) + \int_0^\sigma s^\infty(z)Q(z)\,dz, &= -c^N(z) + c_0 - \int_0^\sigma s^N(z)Q(z)\,dz,
\end{aligned}
\tag{I.2}
$$

Since $c^N, s^N$ solve the system of Equations (I.1) exactly to $N$-th order we have that

$$
\begin{aligned}
s^N(\sigma) &= s_0 + \left[\int_0^\sigma c^N(z)Q(z)\,dz\right]^N,\\
c^N(\sigma) &= c_0 - \left[\int_0^\sigma s(z)Q(z)\,dz\right]^N,
\end{aligned}
$$

so that Equation (I.2)becomes

$$
\begin{aligned}
s^\infty(\sigma) - \int_0^\sigma c^\infty(z)Q(z)\,dz &= -s_0 - \left[\int_0^\sigma c^N(z)Q(z)\,dz\right]^N + s_0 + \int_0^\sigma c^N(z)Q(z)\,dz,\\
c^\infty(\sigma) + \int_0^\sigma s^\infty(z)Q(z)\,dz, &= -c_0 + \left[\int_0^\sigma s(z)Q(z)\,dz\right]^N + c_0 - \int_0^\sigma s^N(z)Q(z)\,dz,
\end{aligned}
$$

which, after cancelation is

$$
\begin{aligned}
s^\infty(\sigma) - \int_0^\sigma c^\infty(z)Q(z)\,dz &= \left[\int_0^\sigma c^N(z)Q(z)\,dz\right]^\infty,\\
c^\infty(\sigma) + \int_0^\sigma s^\infty(z)Q(z)\,dz, &= -\left[\int_0^\sigma s^N(z)Q(z)\,dz\right]^\infty,
\end{aligned}
\tag{I.3}
$$

a linear system of equations for $s^\infty, c^\infty \in \mathcal{X}$, amenable to a straight forward Neumann series analysis.

To this end define

$$
H(\sigma) = \begin{pmatrix} s^\infty(\sigma) \\ c^\infty(\sigma) \end{pmatrix}, \qquad \mathcal{E}^N(\sigma) = \begin{pmatrix} \left[\int_0^\sigma c^N(z)Q(z)\,dz\right]^\infty \\ -\left[\int_0^\sigma [s^N(z)Q(z)\,dz\right]^\infty \end{pmatrix}.
$$

Note that $\sin(g_N(\sigma)) - s^N(\sigma) = s^\infty(\sigma)$ and $\cos(g_N(\sigma)) - c^N(\sigma) = c^\infty(\sigma)$, so obtaining bounds on $H(\sigma)$ will lead to (67–68). Observe that,

$$c^N(z)Q(z) = \sum_{n=0}^{2N-1}\left(\sum_{k=0}^{n}(k+1)c_{n-k}\beta_{k+1}\right)z^n$$

as $c^N$ and $Q$ are $N$-th and $N-1$-th order polynomials respectively. Then for any $\sigma \in D$ we have that

$$\begin{aligned}
\int_0^\sigma c^N(z)Q(z)\,dz &= \int_0^\sigma \left(\sum_{n=0}^{2N-1}\sum_{k=0}^{n}(k+1)c_{n-k}\beta_{k+1}z^n\right)dz \\
&= \sum_{n=0}^{2N-1}\sum_{k=0}^{n}(k+1)c_{n-k}\beta_{k+1}\left(\int_0^\sigma z^n\,dz\right) \\
&= \sum_{n=0}^{2N-1}\sum_{k=0}^{n}(k+1)c_{n-k}\beta_{k+1}\frac{1}{n+1}\sigma^{n+1} \\
&= \sum_{n=1}^{2N}\left(\sum_{k=0}^{n-1}\frac{k+1}{n}c_{n-k-1}\beta_{k+1}\right)\sigma^n.
\end{aligned}$$

Then

$$\left[\int_0^\sigma c^N(z)Q(z)\,dz\right]^\infty = \sum_{n=N+1}^{2N}\left(\sum_{k=0}^{n-1}\frac{k+1}{n}c_{n-k-1}\beta_{k+1}\right)\sigma^n.$$

A similar calculation shows that

$$\left[\int_0^\sigma s^N(z)Q(z)\,dz\right]^\infty = \sum_{n=N+1}^{2N}\left(\sum_{k=0}^{n-1}\frac{k+1}{n}s_{n-k-1}\beta_{k+1}\right)\sigma^n.$$

Combining these with the maximum modulus principle and the triangle inequality gives

$$\|\mathcal{E}^N\|_0 \le e_N,$$

where $e_N$ is as defined in the hypothesis of the Lemma. We seek a solution $H \in \mathcal{X}$ of (I.3)

Define the linear operator $\mathfrak{M}\colon \mathcal{X} \to \mathcal{X}$ by

$$\mathfrak{M}(H)(\sigma) = \begin{pmatrix} -\int_0^\sigma c^\infty(z)Q(z)\,dz \\ \int_0^\sigma s^\infty(z)Q(z)\,dz \end{pmatrix}.$$

The linear equation for $H$ is now

$$(\mathrm{Id} + \mathfrak{M})\,H = \mathcal{E}^N.$$

We will show that

$$\|\mathfrak{M}\|_{B(\mathcal{X})} \le \frac{\hat{K}}{N+2} < 1.$$

To see this note that for any analytic $N$-tail $H$ with $\|H\|_0 < \infty$ there exists an analytic function $\hat{H}\colon D \to \mathbb{C}^2$ so that

$$H(\sigma) = \hat{H}(\sigma)\sigma^{N+1},$$

and

$$\|H\|_0 = \|\hat{H}\|_0.$$

Here equality of the $C^0$ norms is a consequence of the maximum modulus principle and the observation that $|\sigma|^{N+1} = 1$ on the boundary of the disk.

Then for any $\sigma \in D$ we have that

$$\begin{aligned}
\left|\int_0^\sigma H(z)Q(z)\,dz\right| &= \left|\int_0^\sigma z^{N+1}\hat{H}(z)Q(z)\,dz\right| \\
&= \left|\int_0^1 (t\sigma)^{N+1}\hat{H}(t\sigma)Q(t\sigma)\sigma\,dt\right| \\
&\leq \int_0^1 |\sigma|^{N+2}|\hat{H}(t\sigma)||Q(t\sigma)|t^{N+1}\,dt \\
&\leq \|\hat{H}\|_0\|Q\|_0|\sigma|^{N+2}\int_0^1 t^{N+1}\,dt \\
&\leq \|H\|_0\hat{K}\left.\frac{t^{N+2}}{N+2}\right|_0^1 \\
&\leq \frac{\hat{K}}{N+2}\|H\|_0,
\end{aligned}$$

as $|\sigma| \leq 1$. Taking the sup over all $H$ with norm one yields the result.

It now follows from the assumption that $\hat{K}/(N+2) < 1$ and the Neumann theorem that $\mathrm{Id}+\mathfrak{M}$ is invertible with

$$\|(\mathrm{Id}+\mathfrak{M})^{-1}\|_{B(\mathcal{X})} \leq \frac{1}{1-\frac{\hat{K}}{N+2}}.$$

Then

$$H = (\mathrm{Id}+\mathfrak{M})^{-1}\mathcal{E}_N,$$

and

$$\|H\|_0 \leq \frac{e_N}{1-\frac{\hat{K}}{N+2}},$$

as required.

## 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. Arnolđ, 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. Arnolḋ, 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. Arnolḋ, 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. 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
[20] M. J. Capiński, M. Gidea, Arnold diffusion, quantitative estimates and stochastic behavior in the three-body problem.
URL https://arxiv.org/abs/1812.03665
[21] 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
[22] 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
[23] 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.
[24] 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
[25] 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/dcds.2011.30.641.
URL https://doi.org/10.3934/dcds.2011.30.641
[26] 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
[27] 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

[28] 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

[29] 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

[30] 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

[31] T. Kapela, M. Mrozek, D. Wilczak, Z. P., Capd::dynsys: a flexible c++ toolbox for rigorous numerical analysis of dynamical systems, Commun. Nonlinear Sci. Numer. Simul.
URL https://arxiv.org/abs/1812.03665

[32] M. J. Capiński, Computer assisted existence proofs of Lyapunov orbits at $L_2$ and transversal intersections of invariant manifolds in the Jupiter-Sun PCR3BP, SIAM J. Appl. Dyn. Syst. 11 (4) (2012) 1723–1753. doi:10.1137/110847366.
URL https://doi.org/10.1137/110847366

[33] 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.

[34] 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.

[35] 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

[36] M. Breden, J.-P. Lessard, J. D. Mireles James, Computation of maximal local (un)stable manifold patches by the parameterization method, Indag. Math. (N.S.) 27 (1) (2016) 340–367. doi:10.1016/j.indag.2015.11.001.
URL http://dx.doi.org/10.1016/j.indag.2015.11.001