SIAM J. Applied Dynamical Systems
Vol. 12, No. 2, pp. 957–1006

Rigorous A Posteriori Computation of (Un)Stable Manifolds and Connecting Orbits for Analytic Maps*

J. D. Mireles James† and Konstantin Mischaikow‡

Abstract. This work is concerned with high order polynomial approximation of stable and unstable manifolds for analytic discrete time dynamical systems. We develop a posteriori theorems for these polynomial approximations which allow us to obtain rigorous bounds on the truncation errors via a computer assisted argument. Moreover, we represent the truncation error as an analytic function, so that the derivatives of the truncation error can be bounded using classical estimates of complex analysis. As an application of these ideas we combine the approximate manifolds and rigorous bounds with a standard Newton–Kantorovich argument in order to obtain a kind of "analytic-shadowing" result for connecting orbits between fixed points of discrete time dynamical systems. A feature of this method is that we obtain the transversality of the connecting orbit automatically. Examples of the manifold computation are given for invariant manifolds which have dimension between two and ten. Examples of the a posteriori error bounds and the analytic-shadowing argument for connecting orbits are given for dynamical systems in dimension three and six.

Key words. stable manifolds, computer assisted proof, validated computation, connecting orbits, parameterization method, high order methods

AMS subject classifications. 37D10, 37C29, 65G20, 65P20

DOI. 10.1137/12088224X

1. Introduction.

Knowledge of the existence and intersection of invariant manifolds, particularly stable and unstable manifolds, plays a central role in our description and understanding of the global dynamics of nonlinear systems. However, because they are typically global nonlinear objects, an analytic description of invariant manifolds for specific dynamical systems is typically impossible. Thus it is not surprising that there is a large literature devoted to the numerical approximation of these objects [30]. Current computing technology makes it practical to go a step further and rigorously bound these numerical approximations. The focus of this article is twofold: one, to obtain bounds on approximations of finite dimensional stable and unstable manifolds of hyperbolic fixed points for maps and, two, to demonstrate that these bounds can efficiently be used to rigorously approximate homoclinic and heteroclinic orbits. As such this paper is an extension of our earlier work with van den Berg and Lessard [7], the distinction being that in this paper we provide sharper bounds, and work with maps as opposed to flows.

Consider a function $f\colon \mathbb{R}^n \to \mathbb{R}^n$ that is real analytic in some neighborhood of the hyper-

*Received by the editors June 25, 2012; accepted for publication (in revised form) by H. Kokubu April 1, 2013; published electronically June 19, 2013. This work was partially supported by NSF grants DMS-0915019, DMS-1125174, and CBI-0835621 and by contracts from DARPA and AFOSR.
http://www.siam.org/journals/siads/12-2/88224.html
†Department of Mathematics, Rutgers University, Piscataway, NJ 08854 (jmireles@math.rutgers.edu).
‡Department of Mathematics & BioMaPS, Rutgers University, Piscataway, NJ 08854 (mischaik@math.rutgers.edu).

bolic fixed points p and q. Let $n_s, n_u \in \mathbb{N}$ denote the dimensions of the stable and unstable eigenspaces of $Df(p)$ and $Df(q)$, respectively. It follows from the stable manifold theorem [28] that there are neighborhoods $U_s \subset \mathbb{R}^{n_s}$, $U_u \subset \mathbb{R}^{n_u}$ and analytic chart maps

$$P\colon U_s \subset \mathbb{R}^{n_s} \to \mathbb{R}^n \quad \text{and} \quad Q\colon U_u \subset \mathbb{R}^{n_u} \to \mathbb{R}^n$$

for the local unstable and stable manifolds at p and q, so that

$$P[U_s] = W^s_{\text{loc}}(p) \quad \text{and} \quad Q[U_u] = W^u_{\text{loc}}(q).$$

In general one cannot expect to compute the chart maps P and Q exactly. However, one can hope to compute the power series approximations P_N and Q_N.

The primary goal of this paper is to provide better bounds on the truncation errors associated with P_N and Q_N. We make the following technical assumptions[1] (see section 2 for a detailed description of the notation):

A1. Assume that $f\colon \mathbb{C}^n \to \mathbb{C}^n$ is an analytic function and $f(p) = p$ for some $p \in \mathbb{C}^n$.

A2. Assume that $Df(p)$ is nonsingular, diagonalizable, and hyperbolic. Let $\{\lambda^s_1, \dots, \lambda^s_{n_s}\}$ denote the stable eigenvalues, and let $\{\xi^s_1, \dots, \xi^s_{n_s}\}$ denote a particular choice of corresponding stable eigenvectors. Let Λ_s denote the $n_s \times n_s$ diagonal matrix of stable eigenvalues, and $A_s = [\xi^s_1 | \dots | \xi^s_{n_s}]$ denote the matrix whose columns are the stable eigenvectors.

A3. Assume that $P_N\colon \mathbb{C}^{n_s} \to \mathbb{C}^n$ is an Nth order polynomial, with $N \geq 2$. Moreover, assume that P_N provides an exact formal solution of the equation

$$f[P_N(\theta)] = P_N(\Lambda_s \theta), \tag{1.1}$$

up to order N, that is, that the power series coefficients of the function on the left-hand side are exactly equal to the power series coefficients of the function on the right-hand side to Nth order.

One can make corresponding assumptions A2′ and A3′ for the Q_N parameterization of the unstable manifold at q. We remark that it is possible that some of the coefficients of P_N are zero (it sometimes happens, for example, that the linear approximation given by the eigenspace satisfies (1.1) to high order). Given f and Λ, there exists a formal procedure for computing P_N as discussed in greater detail below.

Assumptions A1 and A2 are standard assumptions associated with hyperbolic fixed points. Assumption A3 arises because our approach is based on the *parameterization method* developed by Cabré, de la Llave, and Fontich in [10, 11, 12]. This provides an efficient technique for computing to Nth order a power series approximation P_N for the chart map P, as well as a general framework for establishing the convergence of the series. The key observation for this method is the fact that the chart map P satisfies the functional equation

$$f[P(\theta)] = P(\Lambda_s \theta), \tag{1.2}$$

[1] From a theoretical perspective it is easier to work with holomorphic functions. The applications presented in this paper are restricted to real analytic functions.

for $\theta \in U_s$ some neighborhood of the origin in $\mathbb{C}^{n_s}$, while also satisfying the first order constraints

$$P(0) = p, \quad DP(0) = A_s. \tag{1.3}$$

Similar remarks hold for the chart map Q of the unstable manifold by observing that Q parameterizes the stable manifold of f^{-1}. (Note that f^{-1} is defined at least locally, as the fact that p is a hyperbolic fixed point implies that f is a local diffeomorphism.)

Our contribution is to provide rigorous bounds on the truncation error $P-P_N$ and $Q-Q_N$, given certain hypotheses which can be checked on the computer using interval arithmetic. We hasten to add that we do not know P_N or Q_N exactly; the best we can do computationally is to obtain a rigorous enclosure of these polynomials using interval arithmetic. However, the bounds we present take this extra level of error into consideration and therefore are valid given the particular interval enclosure of P_N and Q_N.

From now on we present the analysis for P, the parameterization of the stable manifold. The arguments and bounds for Q, the parameterization of the unstable manifold, follow mutatis mutandis. The main result of the present work is Theorem 3.2, which makes precise the following claims.

Theorem 1.1 ("meta-theorem" paraphrasing of Theorem 3.2). *Let U_s be a neighborhood of the origin in $\mathbb{C}^{n_s}$, and suppose that P_N is a "good enough" and "properly constructed" Nth order polynomial approximately solving (1.2) on U_s. Then there exist a constant $C(N)$ and a unique analytic function $h\colon U_s \to \mathbb{C}^n$ so that $P = P_N + h$ exactly solves (1.2) on U_s. Moreover, we have*

$$\sup_{\theta \in U_s} |h(\theta)| \leq C(N) \sup_{\theta \in U_s} |f[P_N(\theta)] - P_N(\Lambda_s \theta)|\,. \tag{1.4}$$

Since $h = P - P_N$ is the truncation error associated with the approximation P_N on U_s and since h is a bounded analytic function, we can use classical estimates of complex analysis in order to bound derivatives of h as well.

Much of the technical work in the following pages goes into making precise the terms properly constructed and good enough. We must also specify the neighborhood U_s and determine precisely the constant $C(N)$. Note that the error estimate given by (1.4) is a posteriori in the sense that all terms on the right-hand side of (1.4) can be explicitly determined once a particular approximation P_N is fixed. Hence the supremum on the right-hand side of the inequality can be estimated using rigorous numerical methods, and we can obtain mathematically rigorous computer assisted bounds on the truncation error associated with the polynomial approximations P_N (and similarly for Q_N).

As might be expected, the constant $C(N)$ depends in a nontrivial way on the domain of P, the interval approximation of P_N, the nonlinearity f, and the chosen/available precision of computation. These are captured via a set of related constants that we refer to as *validation values* (see Definition 3.1 for a precise description). Note that assumption A2 fixes the magnitudes of the eigenvectors of $Df(p)$. The uniqueness conclusion in Theorem 3.2 is predicated on this assumption; different scalings of the eigenvectors lead to reparameterizations of the chart map P which satisfy (1.2).

In order to demonstrate the utility of Theorem 3.2 we use it to develop a technique to rigorously prove the existence of heteroclinic and homoclinic orbits. Recall that Beyn and Kleinkauf [8, 9] developed an approach for obtaining numerical approximations of these types of orbits via *projected boundary conditions*. More precisely, they use Newton's method to solve the system of equations

$$\begin{aligned} x_{n+1} &= f(x_n), \quad n = n_-, \dots, n_+ - 1, \\ b_-(x_{n_-}) &= b_+(x_{n_+}) = 0, \end{aligned}$$

where the zero sets of b_- and b_+ are linear approximations to the local unstable and stable manifolds at the fixed point.

We consider a slightly modified problem. Define the *connecting orbit operator equation* $F\colon \mathbb{R}^{n(k+1)} \to \mathbb{R}^{n(k+1)}$ by

$$F(\phi, x, \theta) = \begin{bmatrix} f^{-1}(x_1) - Q(\phi) \\ \vdots \\ f(x_j) - x_{j+1} \\ \vdots \\ f(x_k) - P(\theta) \end{bmatrix}, \quad j = 1, \dots, k-1, \tag{1.5}$$

where $\phi \in U_u \subset \mathbb{R}^{n_u}$, $\theta \in U_s \subset \mathbb{R}^{n_s}$, and $n_u + n_s = n$. Observe that $F(\hat{\phi}, \hat{x}, \hat{\theta}) = 0$ implies that, for any $j = 1, \dots, k$, $\hat{x}_j \in \mathbb{R}^n$ lies on a homoclinic or heteroclinic orbit depending on whether $p = q$ or $p \neq q$.

If we let F_N be the map obtained by replacing the chart maps P and Q in the definition of F by P_N and Q_N, then we are in the setting of projected boundary conditions, and hence by [8, 9] we can find an approximate numerical homoclinic orbit. As we demonstrate via Theorem 5.2, the error bounds of Theorem 3.2 are sufficient for us to apply a Newton–Kantorovich argument to conclude the existence of a true homoclinic orbit lying within a given bound of the numerical orbit. A feature of our method is that the homoclinic or heteroclinic orbits obtained are automatically transverse.

As in the setting of the stable and unstable manifold representations, it is useful to think of our approach as an a posteriori validation technique for the method of projected boundary conditions. To make this clear and to demonstrate the practicality of our approach we consider a variety of applications. We focus on two models: the *delayed Hénon map* $f\colon \mathbb{R}^n \to \mathbb{R}^n$, introduced in [49] and defined by

$$f(x_1, \dots, x_n) = \begin{pmatrix} 1 - a\, x_1^2 + b\, x_n \\ x_1 \\ \vdots \\ x_{n-1} \end{pmatrix}, \tag{1.6}$$

and the *Lomelí map* $f\colon \mathbb{R}^3 \to \mathbb{R}^3$, introduced in [35] and given by

$$f(x, y, z) = f_{\alpha,\tau,a,b,c}(x, y, z) = \begin{pmatrix} z + Q_{\alpha,\tau,a,b,c}(x, y) \\ x \\ y \end{pmatrix}, \tag{1.7}$$

where Q is the quadratic function

$$Q_{\alpha,\tau,a,b,c}(x,y) = \alpha + \tau x + ax^2 + bxy + cy^2 \quad \text{subject to} \quad a+b+c=1. \tag{1.8}$$

The delayed Hénon map has the advantage that it allows us to easily vary the dimension of the stable manifold. We make use of the Lomelí map to study the existence of homoclinic orbits which make "long" excursions, and we couple together a pair of Lomelí maps in order to obtain examples of homoclinic orbits in a higher dimensional example.

A few additional comments concerning the generality of this method are in order.

Remark 1.2. If f is not invertible, but p is a hyperbolic fixed point, then the local stable and unstable sets can still be defined, and the stable manifold theorem generalizes as in [46]. Since p is hyperbolic, f is a local analyticomorphism, and the parameterization method can still be used to compute the local manifolds. The usual care must be taken to globalize the local stable manifold due to the nonexistence of a unique inverse map (see [19, 20] for more complete discussion).

Similarly, the results associated with solutions to the connecting orbit operator (1.5) also extend to the noninvertible case under mild assumptions. For example, the choice of inverse maps is dictated by the specific problem at hand (i.e., the approximate orbit whose existence is to be validated). For a more complete discussion of connecting orbits for noninvertible maps we refer the reader to [47].

Remark 1.3. To apply the methods presented in this paper the requirement that the map be real analytic on all of $\mathbb{C}^n$ can be significantly weakened. It is sufficient to assume that f is analytic and uniformly bounded in a neighborhood of p and piecewise real analytic (linear, polynomial, etc.) as long as all of the points $\hat{x}_i$, $\tilde{x}_i$, $1 \le i \le k-1$, are bounded away from the singularity set of Df.

It is important to note that there are other successful numerical methods for obtaining rigorous enclosures of stable and unstable manifolds, as well as computer assisted proofs of the existence of connecting orbits. The first numerical implementation of a computer assisted proof for stable and unstable manifolds that we know of was developed for the one dimensional case in [39]. That reference also provides a method for establishing the existence of transverse intersections of the manifolds. More general methods for obtaining rigorous enclosures of normally hyperbolic invariant manifolds (including stable and unstable manifolds) are implemented in [54, 15, 16, 53, 52] (see [28] for the general development of normally hyperbolic invariant manifolds).

The methods just mentioned are based on a topological rather than analytic approach to studying invariant manifolds in phase space and can also be applied to proving the existence of connecting orbits and chaotic motions [40, 2, 3, 4, 23, 24, 51, 22]. We also mention the work of [29, 7, 34], which implement computer assisted methods for the rigorous study of stable and unstable manifolds based on analytic methods similar to those presented here for differential equations instead of local diffeomorphisms. The later two references also implement computer assisted proofs of the existence of symmetric connecting orbits for second order differential equations, and saddle to saddle connecting orbits for first order vector fields. In the work of [17, 42, 50], another method for validated shadowing of numerical orbits is developed which can also be adapted to proving the existence of connecting orbits for differential equations

and maps. Rather than computing stable and unstable manifolds, these methods exploit the theory of exponential dichotomies in order to obtain a posteriori validation theorems.

The parameterization method has wide application outside the scope of the present work; see [10, 11, 12, 6, 26, 27, 21, 31, 33] for theoretical developments, as well as [13, 14, 25, 7, 37] for additional numerical applications.

An outline of this paper is as follows. In section 2 we develop all the notation and basic analytical results used in the remainder of the paper. Section 3 is devoted to the proof of Theorem 3.2.

In section 4 we begin an example-driven discussion of the computational aspects of the parameterization method. Section 4.1 discusses the numerical computation of the coefficients for the approximate parameterization and examines computational performance for a family of example systems. We compute polynomial approximations of invariant manifolds whose dimension ranges from 3 to 10, and examine computational complexity. In section 4.2 we discuss the determination of the a posteriori error. Section 4.3 presents a theorem which shows that, under mild "nonresonance" hypotheses between the eigenvalues of like stability, it is possible to compute the approximate parameterization to any order and find validation values satisfying Theorem 3.2. We also develop certain a priori indicators which are useful in practice. Section 4.4 examines several explicit validated computations for invariant manifolds and explains how the validation values are chosen in practice.

In section 5 we state and prove an a posteriori validation theorem for connecting orbits. This is a kind of "shadowing theorem" in the C^ω category, and its proof requires only the finite dimensional Newton–Kantorovich theorem (our Theorem 2.7). Finally, in section 6 we present a number of computer assisted proofs of the existence of transverse homoclinic orbits in three and six dimensional examples.

2. Analytic preliminaries.

For the sake of readability and completeness this section introduces the notation used in this paper and reviews a variety of standard estimates that are essential for the proofs of Theorems 3.2 and 5.2.

2.1. Analytic functions, norms, and notation.

If $x \in \mathbb{R}$ or $z = a + ib \in \mathbb{C}$, then $|x| = \text{sign}(x)x$ and $|z| := \sqrt{a^2 + b^2}$ denote the usual absolute values. For the complex vector spaces $\mathbb{C}^n$ we always use the sup norm; so if $z = (z_1, \ldots, z_n) \in \mathbb{C}^n$, then

$$|z| := \max_{1 \le i \le n} |z_i|. \tag{2.1}$$

If the real vector space $\mathbb{R}^n$ is playing the role of the phase space of a real dynamical system, then we use again the sup norm. In this case if $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, then

$$|x| := \max_{1 \le i \le n} |x_i|. \tag{2.2}$$

The max-over-components norm is well suited for numerical work, as it is easy to evaluate and (unlike Euclidean norms) introduce no additional rounding errors. The max norm also simplifies certain analytic estimates.

On the other hand, when we consider chart maps from $\mathbb{R}^k$ to $\mathbb{R}^n$ it is sometimes useful to define the following *poly-cylinder* norms on $\mathbb{R}^k$. Let $l, m \in \mathbb{N}$, and suppose that $l + 2m = k$.

Then for any $\theta = (\theta_1, \ldots, \theta_k) \in \mathbb{R}^k$ we define

$$|\theta|_{(l,m)} = \max_{1\le i\le l,\ 1\le j\le m} \left(|\theta_i|, \sqrt{\theta_{l+2j-1}^2 + \theta_{l+2j}^2} \right). \tag{2.3}$$

In practice l is the number of real stable eigenvalues, and m is the number of complex conjugate pairs of stable eigenvalues associated with a k dimensional stable manifold.

Let $\hat{z} \in \mathbb{C}^n$, $\hat{x} \in \mathbb{R}^n$, $\hat{\theta} \in \mathbb{R}^k$, and $\nu > 0$. Suppose that $l, m \in \mathbb{N}$ satisfy $l + 2m = k$. The neighborhoods induced by the above norms are as follows:

- The *complex poly-disk* $B_\nu(\hat{z}) \subset \mathbb{C}^n$ is defined by
$$B_\nu(\hat{z}) := \{ z \in \mathbb{C}^n : |\hat{z} - z| < \nu \}.$$
- The *real poly-disk* (or cube) $D_\nu(\hat{x}) \subset \mathbb{R}^n$ is defined by
$$D_\nu(\hat{x}) := \{ x \in \mathbb{R}^n : |\hat{x} - x| < \nu \}.$$
- The (l, m) *poly-cylinder* $U_\nu^{(l,m)}(\hat{\theta}) \subset \mathbb{R}^k$ is defined by
$$U_\nu^{(l,m)}(\hat{\theta}) = \{ \theta \in \mathbb{R}^k : |\hat{\theta} - \theta|_{(l,m)} < \nu \}.$$

If the neighborhoods are centered at the origin, then to simplify notation we write $B_\nu(0) = B_\nu$, $D_\nu(0) = D_\nu$, and $U_\nu^{(l,m)}(0) = U_\nu^{(l,m)}$. When l, m are clear from context we simply write $U^{(l,m)} = U_\nu$. Note that U_ν is the product of l intervals of length 2ν and m planar Euclidean disks of radius ν. (So, for example, if $k = 3$ and $l = m = 1$, then U_ν is just the usual cylinder.)

A function $f\colon B_\nu(\hat{z}) \subset \mathbb{C}^m \to \mathbb{C}$ is analytic on the poly-disk $B_\nu(\hat{z})$ if f has a power series expansion

$$f(z) = \sum_{|\alpha|\ge 0} a_\alpha (\hat{z} - z)^\alpha, \quad a_\alpha \in \mathbb{C},\ \alpha \in \mathbb{N}^m,$$

which converges for all $z \in B_\nu(\hat{z})$. Here we use the usual *multi-index* notation, so that if $\alpha = (\alpha_1, \ldots, \alpha_m) \in \mathbb{N}^m$ and $z \in \mathbb{C}^m$, then $|\alpha| = \alpha_1 + \cdots + \alpha_m$ and $z^\alpha = z_1^{\alpha_1} \cdot \cdots \cdot z_m^{\alpha_m}$. More generally, $f\colon B_\nu(\hat{z}) \subset \mathbb{C}^m \to \mathbb{C}^n$ is analytic on $B_\nu(\hat{z})$ if $f = (f_1, \ldots, f_n)$, and each $f_j : B_\nu(\hat{z}) \subset \mathbb{C}^m \to \mathbb{C}$, $1 \le j \le n$, is analytic. Again, such an f can be expressed in power series form as

$$f(z) = \sum_{|\beta|\ge 0} b_\beta (\hat{z} - z)^\beta, \quad b_\beta \in \mathbb{C}^n,\ \beta \in \mathbb{N}^m, \tag{2.4}$$

which converges for all $z \in B_\nu(\hat{z})$.

Our focus is on analytic functions that arise as model dynamical systems. In this setting it is often the case that not all variables are directly related to one another. Since we are interested in using numerics to extract rigorous information, bounds which are computationally easy to access are useful. In particular, given an analytic function $f\colon B_\nu(\hat{z}) \subset \mathbb{C}^m \to \mathbb{C}^n$, we will make use of the second derivatives to obtain bounds which can often be improved if we can ignore any zero terms. With this in mind, using the notation of (2.4), we define the number of second order partial derivatives of f which are not identically zero on $B_\nu(\hat{z})$; i.e.,

$$N_f := \#\{\beta \in \mathbb{N}^m : |\beta| = 2 \ \text{ and } \ \partial_\beta f_i(z) \not\equiv 0 \ \text{ for } \ z \in B_\nu(\hat{z}),\ 1 \le i \le n\}. \tag{2.5}$$

See (3.7) of Theorem 3.2 for an example of its use.

We make use of two norms on analytic functions, the standard C^0 norm,

$$\|f\|_{B_\nu(\hat{z})} := \sup_{|z-\hat{z}|<\nu} |f(z)|,$$

which is useful for analytic arguments, and the weighted Wiener space (or weighted ℓ^1) norm,

$$\|f\|_{B_\nu(\hat{z}),\Sigma} := \sum_{|\alpha|\geq 0} |b_\alpha| \nu^{|\alpha|},$$

which is computationally more tractable. We have the inequality $\|f\|_{B_\nu(\hat{z})} \leq \|f\|_{B_\nu(\hat{z}),\Sigma}$ and note that both norms induce a Banach space structure on the set of bounded holomorphic maps of $B_\nu(\hat{z})$. Moreover, the maximum modulus principle implies that if f is uniformly bounded and analytic on (the open poly-disk) $B_\nu(\hat{z}) \subset \mathbb{C}^m$, then

$$\|f\|_{B_\nu(\hat{z})} = \sup_{|z-\hat{z}|=\nu} |f(z)|\,,$$

and hence f is bounded on the closed poly-disk. Finally, we have that if $\|f\|_{B_\nu(\hat{z}),\Sigma} < \infty$, then f is not only bounded but also continuous on $\partial B_\nu(\hat{z})$. Again, if the poly-disk in question is centered at the origin, then we simplify notation by letting

$$\|\cdot\|_{\nu,\Sigma} = \|\cdot\|_{B_\nu(0),\Sigma} \quad \text{and} \quad \|\cdot\|_\nu = \|\cdot\|_{B_\nu(0)}.$$

Let X and Y be Banach spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively. Given a linear map $A\colon X \to Y$, the standard operator norm is

$$\|A\|_{\mathbb{L}(X)} := \sup_{x\in X, \|x\|_X=1} \|Ax\|_Y.$$

We denote the set of all bounded linear operators on X by $\mathbb{L}(X)$ and remark that this is a Banach space. In the special case in which $X = \mathbb{C}^m$ and $Y = \mathbb{C}^n$ we denote the norm by $\|A\|_M$, and we sometimes simply write $\|A\|$ when it is understood that A is a matrix. Since we are using the sup norm (see 2.1), we have the standard estimate [5]

$$\|A\|_M \leq \sup_{1\leq i\leq n} \sum_{j=1}^{m} |a_{ij}| \leq m \sup_{1\leq i\leq n} \sup_{1\leq j\leq m} |a_{ij}|. \tag{2.6}$$

We make use of these norms in a slightly more general setting. Consider a fixed $\hat{z} \in \mathbb{C}^k$, $\nu > 0$, and an $n \times m$ matrix A whose entries are themselves analytic functions, $a_{ij} : B_\nu(\hat{z}) \subset \mathbb{C}^k \to \mathbb{C}$. Then the norm of the nonconstant matrix A is

$$\|A\|_{M,B_\nu(\hat{z})} = \max_{1\leq i\leq n} \sum_{j=1}^{m} \|a_{ij}\|_{B_\nu(\hat{z})}.$$

If, in addition, $g\colon B_\nu(\hat{z}) \subset \mathbb{C}^k \to \mathbb{C}^m$ is an analytic function, then the nonconstant matrix vector product $A \cdot g\colon B_\nu(\hat{z}) \subset \mathbb{C}^k \to \mathbb{C}^n$ is an analytic function, and we have the bounds

$$\|A \cdot g\|_{B_\nu(\hat{z})} \leq \|A\|_{M,B_\nu(\hat{z})} \|g\|_{B_\nu(\hat{z})} \leq m \|g\|_{B_\nu(\hat{z}),\Sigma} \max_{1\leq i\leq n} \max_{1\leq j\leq m} \|a_{ij}\|_{B_\nu(\hat{z}),\Sigma}.$$

The last bound is particularly useful for numerical applications.

As is indicated in the introduction, we are interested in bounds on $P - P_N$ when $P_N = P$ to Nth order in the sense of power series. Thus the following family of analytic functions plays an important role.

Definition 2.1. *An analytic function* $h\colon B_\nu(0) \subset \mathbb{C}^m \to \mathbb{C}^n$ *is an* analytic N-tail *if* h *is analytic on* $B_\nu(0)$ *and*

$$h(0) = 0, \quad Dh(0) = 0, \ldots, D^\alpha h(0) = 0 \quad \text{for } |\alpha| \le N.$$

Observe that an analytic N-tail h has a power series representation of the form

$$h(z) = \sum_{|\beta|>N} b_\beta z^\beta, \qquad b_\beta \in \mathbb{C}^n,\ \beta \in \mathbb{N}^m,$$

converging for each $|z| < \nu$. With m, n, and $\nu > 0$ fixed, we define $\mathbb{H}_N = \mathbb{H}_N(n, m, \nu)$ to be the set of bounded analytic N-tails on $B_\nu(0) \subset \mathbb{C}^m$ taking values in $\mathbb{C}^n$.

The following lemmas are well-known or easily established facts about analytic functions and N-tails.

Lemma 2.2. *If* $\hat z \in \mathbb{C}^m$, $\nu > 0$, $f\colon B_\nu(\hat z) \to \mathbb{C}^n$ *is analytic, and* $\|f\|_\nu \le M$, *then for each* $\beta \in \mathbb{N}^m$

$$|b_\beta| \le \frac{M}{\nu^{|\beta|}}.$$

The proof follows from the Cauchy estimate [1].

Lemma 2.3. *Let* h *be a bounded analytic* N*-tail on* $B_\nu(0) \subset \mathbb{C}^m$. *Let* Λ *be an* $m \times m$ *diagonal matrix with diagonal entries* $\lambda_1, \ldots, \lambda_m \in \mathbb{C}$ *such that* $0 < |\lambda_j| < 1$ *for* $1 \le j \le m$. *If* $\mu^* := \sup_j |\lambda_j|$, *then* $h \circ \Lambda$ *is a bounded analytic* N*-tail on* $B_\nu(0)$ *and*

$$\|h \circ \Lambda\|_\nu \le (\mu^*)^{N+1} \|h\|_\nu.$$

See [7, Lemma 3.2] for an elementary proof.

Lemma 2.4. *Fix* $p \in \mathbb{C}^m$, $0 < \nu' < \nu$, *and suppose that* $f\colon B_\nu(p) \subset \mathbb{C}^m \to \mathbb{C}$ *has uniformly bounded second derivatives in* $B_\nu(p)$, *so that there is a* $K > 0$ *satisfying*

$$\sup_{|\beta|=2} \|\partial_\beta f\|_{B_\nu(p)} \le K.$$

Given any points $\hat z \in B_{\nu'}(p)$ *and* $\eta \in B_{\nu-\nu'}(0)$, f *can be expanded as*

$$f(\hat z + \eta) = f(\hat z) + Df(\hat z) \cdot \eta + R_{\hat z}(\eta),$$

where $|R_{\hat z}(\eta)|/|\eta| \to 0$ *as* $|\eta| \to 0$. *In fact,* $R\colon B_{\nu'}(p) \times B_{\nu-\nu'}(0) \subset \mathbb{C}^{2m} \to \mathbb{C}$ *is an analytic function with respect to both* $\hat z$ *and* η, *and we have the bound*

$$\|R_{\hat z}\|_{\nu-\nu'} \le N_f K (\nu - \nu')^2,$$

uniformly in the variable $\hat z \in B_{\nu'}(p)$. *Here* N_f *is as defined in* (2.5).

This is the Lagrange form of the Taylor remainder theorem [1] applied on the disk $B_{\nu'}(p)$ when we have uniform bounds on the second derivatives of f on the larger disk $B_\nu(p)$. If $f\colon \mathbb{C}^m \to \mathbb{C}^n$, then Lemma 2.4 can be applied to each individual component of f.

Lemma 2.5. *If* $f\colon B_\nu(\hat z) \subset \mathbb{C}^m \to \mathbb{C}^n$ *is analytic and* $z_1, z_2 \in B_\nu(\hat z)$, *then*

$$|f(z_1) - f(z_2)| \le \|Df\|_{M, B_\nu(\hat z)} |z_1 - z_2|.$$

This is the mean value theorem in the context of the norms we are using.

2.2. Classical results from nonlinear analysis. We make extensive use of three standard theorems from nonlinear analysis. The first is the contraction mapping theorem [43]. We state the next two for specificity of exposition in the proofs.

Theorem 2.6 (Neumann series). *Let X be a Banach space. Suppose that $I : X \to X$ is the identity map, $A : X \to X$ is a bounded linear, and $M > 0$ is a real constant with $\|A\|_{\mathbb{L}(X)} \leq M < 1$. Then $I - A$ is boundedly invertible and*

$$[I - A]^{-1} = \sum_{k=0}^{\infty} A^k,$$

from which it follows that

$$\|(I - A)^{-1}\|_{\mathbb{L}(X)} \leq \sum_{k=0}^{\infty} \|A\|^k_{\mathbb{L}(X)} \leq \frac{1}{1 - M}.$$

For a proof see [43].

Theorem 2.7 (Newton–Kantorovich). *Let X and Y be Banach spaces and $F\colon U \subset X \to Y$. Assume the existence of $\hat{x} \in U$ and three positive constants r, ϵ_{NK}, and κ such that $B_r(\hat{x}) \subset U$ and F is Fréchet differentiable on $B_r(\hat{x})$. Furthermore, assume that*

1. *$DF(\hat{x})$ has bounded inverse, and*
2. *$\|DF(x) - DF(y)\|_{B(X,Y)} \leq \kappa\|x - y\|$ for all $x, y \in B_r(\hat{x})$.*

If

$$\epsilon_{NK} \geq \|DF(\hat{x})^{-1} F(\hat{x})\|_X, \tag{2.7}$$

$$\epsilon_{NK} \leq \frac{r}{2}, \tag{2.8}$$

$$4\epsilon_{NK}\,\kappa\,\|DF(\hat{x})^{-1}\|_{B(Y,X)} \leq 1, \tag{2.9}$$

then the equation $F(x) = 0$ has a unique solution in $B_r(\hat{x})$.

For a proof see [41].

The following lemma follows trivially from the preceding theorem and gives conditions under which the differential of F is boundedly at the Newton–Kantorovich root of F. We include the proof for the sake of completeness.

Lemma 2.8. *Assume the hypotheses of Theorem 2.7, and let $x_* \in B_r(\hat{x})$ be the unique root of F in $B_r(\hat{x})$. Assume in addition that $r \leq 4\epsilon_{NK}$ and that the strict inequality*

$$4\epsilon_{NK}\,\kappa\|DF(\hat{x})^{-1}\|_{B(Y,X)} < 1$$

holds. Then $DF(x_*) \in B(X, Y)$ is boundedly invertible at x_*, and moreover if $M > 0$ is any constant with $4\epsilon_{NK}\,\kappa\|DF(\hat{x})^{-1}\|_{B(Y,X)} \leq M < 1$, then we have the explicit bound

$$\|DF(x_*)^{-1}\|_{B(Y,X)} \leq \frac{\|DF(\hat{x})^{-1}\|_{B(Y,X)}}{1 - M}.$$

Proof. Note that

$$DF(x_*) = DF(\hat{x})\left[I - DF^{-1}(\hat{x})(DF(\hat{x}) - DF(x_*))\right]$$

and that we have the bound

$$\begin{aligned}\left\|DF^{-1}(\hat{x})(DF(\hat{x})-DF(x_*))\right\|_{B(Y,X)} &\le \|DF(\hat{x})^{-1}\|_{B(Y,X)}\|DF(\hat{x})-DF(x_*)\|_{B(X,Y)}\\ &\le \|DF(\hat{x})^{-1}\|_{B(Y,X)}\ \kappa\,\|\hat{x}-x_*\|_X\\ &\le \|DF(\hat{x})^{-1}\|_{B(Y,X)}\ \kappa\, r\\ &\le \|DF(\hat{x})^{-1}\|_{B(Y,X)}\ \kappa\,(4\epsilon_{NK})\\ &< 1.\end{aligned}$$

Then $DF(x_*)$ is invertible and satisfies the desired bound by Theorem 2.6. ∎

2.3. Two specific estimates. We present two specific estimates in the form of Lemma 2.9 and an inequality (2.18), which are central to the proof of Theorem 3.2.

The first provides bounds on the derivatives of an analytic function in terms of a bound on the function, but at the price of reducing the size of the domain on which the bounds are valid. The estimates themselves are not new (see, for example, [32, Lemma 4.6]); however, we include a proof because we require explicit bounds on the constants in the inequalities. Similar, but less optimal, estimates can be found in [7]. While the constants we obtain are not sharp, we do take care to obtain the optimal order in the *loss of domain parameter* σ.

Lemma 2.9 (Cauchy bounds). *Suppose that $f\colon B_\nu(0)\subset\mathbb{C}^m\to\mathbb{C}^n$ is bounded and analytic. Then for any $0<\sigma\le 1$ we have that*

$$\|\partial_i f\|_{\nu e^{-\sigma}} \le \frac{2\pi}{\nu\sigma}\|f\|_\nu, \quad \textit{which implies that} \quad \|Df\|_{\nu e^{-\sigma}} \le \frac{2\pi m}{\nu\sigma}\|f\|_\nu \tag{2.10}$$

as well as

$$\|\partial_i\partial_j f\|_{\nu e^{-\sigma}} \le \frac{4\pi^2}{\nu^2\sigma^2}\|f\|_\nu, \quad \textit{and hence} \quad \|D^2 f\|_{\nu e^{-\sigma}} \le \frac{4\pi^2 m^2}{\nu^2\sigma^2}\|f\|_\nu. \tag{2.11}$$

Proof. Consider first the one dimensional case, where $\nu>0$ and $f\colon B_\nu(0)\subset\mathbb{C}\to\mathbb{C}$ is analytic. Let $0<\sigma\le 1$. Cauchy's formula [1] states that, for any $z\in B_{\nu e^{-\sigma}}(0)$,

$$f'(z)=\frac{1}{2\pi i}\int_{|\xi|=\nu}\frac{f(\xi)}{(\xi-z)^2}\,d\xi.$$

Note that the denominator is bounded because $|z|\le\nu e^{-\sigma}$. Parameterize the path $|\xi|=\nu$ by $\xi(\theta)=\nu e^{i\theta}$, and take norms to obtain

$$\begin{aligned}|f'(z)| &= \frac{1}{2\pi}\left|\int_0^{2\pi}\frac{f[\nu e^{i\theta}]i\nu e^{i\theta}}{(\nu e^{i\theta}-z)^2}\,d\theta\right|\\ &\le \frac{1}{2\pi}\int_0^{2\pi}\frac{\nu\|f\|_\nu}{|\nu e^{i\theta}-z|^2}\,d\theta\\ &\le \frac{\|f\|_\nu}{2\pi\nu}\int_0^{2\pi}\frac{1}{|e^{i\theta}-e^{-\sigma}|^2}\,d\theta \qquad (2.12)\\ &= \frac{\|f\|_\nu}{2\pi\nu}\left(\int_{-\frac{\sigma}{2}}^{\frac{\sigma}{2}}\frac{1}{|e^{i\theta}-e^{-\sigma}|^2}\,d\theta+\int_{\frac{\sigma}{2}}^{2\pi-\frac{\sigma}{2}}\frac{1}{|e^{i\theta}-e^{-\sigma}|^2}\,d\theta\right). \qquad (2.13)\end{aligned}$$

Observe that inequality (2.12) is due to the fact that $|z| \leq \nu e^{-\sigma}$, so that the denominator is minimized when $|z| = \nu e^{-\sigma}$. Since the integrand is radially symmetric once we take the norm of f, we are free to take $z = \nu e^{-\sigma}$ and then factor a ν^2 out of the denominator of the integrand.

To obtain bounds in (2.13) note that $e^{\sigma} \geq 1+\sigma$ for all real σ and hence $\sigma/(1+\sigma) \leq 1-e^{-\sigma}$ for all $\sigma > -1$. Thus for all $0 < \sigma \leq 1$ and $0 \leq \theta \leq 2\pi$,

$$\sigma/2 \leq \frac{\sigma}{1+\sigma} \leq 1 - e^{-\sigma} \leq |e^{i\theta} - e^{-\sigma}|. \tag{2.14}$$

This implies that

$$\int_{-\frac{\sigma}{2}}^{\frac{\sigma}{2}} \frac{1}{|e^{i\theta} - e^{-\sigma}|^2}\, d\theta \leq \int_{-\frac{\sigma}{2}}^{\frac{\sigma}{2}} \frac{1}{|\frac{\sigma}{2}|^2}\, d\theta \leq \frac{4}{\sigma}. \tag{2.15}$$

In addition, $|e^{i\theta} - e^{-\sigma}| \geq \sin\theta \geq 2\theta/\pi$ for $0 \leq \theta \leq \pi/2$, and hence

$$\int_{\frac{\sigma}{2}}^{2\pi-\frac{\sigma}{2}} \frac{1}{|e^{i\theta} - e^{-\sigma}|^2}\, d\theta \leq 4 \int_{\frac{\sigma}{2}}^{\frac{\pi}{2}} \frac{\pi^2}{4\theta^2} \leq \frac{2\pi^2}{\sigma}. \tag{2.16}$$

Applying (2.15) and (2.16) to (2.13) produces the desired result:

$$\|f'\|_{\nu e^{-\sigma}} \leq \frac{1}{2\pi\nu} \left(\frac{4}{\sigma} + \frac{2\pi^2}{\sigma} \right) \|f\|_{\nu} \leq \frac{2\pi}{\nu\sigma} \|f\|_{\nu}. \tag{2.17}$$

If $f\colon B_{\nu}(0) \subset \mathbb{C}^m \to \mathbb{C}^n$, then each $f_k(z_1, \ldots, z_i, \ldots, z_m)$, $1 \leq i \leq m$, $1 \leq k \leq n$, is analytic in the ith variable (with the other variables held fixed). Thus (2.17) implies the first part of (2.10), by applying the same argument to the Cauchy integral of $\partial/\partial z_i f_k(z)$. Since the estimate is uniform in i, k, and z, we can apply (2.6) to obtain

$$\|Df\|_{\nu e^{-\sigma}} \leq \frac{2\pi m}{\nu\sigma} \|f\|_{\nu}.$$

The same estimates can be applied to the Cauchy-type integral,

$$\frac{\partial}{\partial z_i} \frac{\partial}{\partial z_j} f(z) = \frac{1}{(2\pi i)^2} \int_{|\xi_i|=\nu} \int_{|\xi_j|=\nu} \frac{f(z_1, \ldots, \xi_i, \ldots, \xi_j, \ldots, z_m)}{(\xi_i - z_i)^2\, (\xi_j - z_j)^2}\, d\xi_i\, d\xi_j,$$

to obtain in a similar fashion that

$$\|D^2 f\|_{\nu e^{-\sigma}} \leq \frac{4\pi^2 m^2}{\nu^2 \sigma^2} \|f\|_{\nu}. \qquad \blacksquare$$

While the estimates of Lemma 2.9 are applicable to any bounded analytic function, the next estimate is specific to the class of application problems being addressed in this paper. Given assumptions A1–A3, the definition of a set of *validation values* (Definition 3.1) requires a bound on the nonconstant matrix $(Df[P_N(\theta)])^{-1}$ for $|\theta| \leq \nu$. For the sake of simplicity and since it suffices for the particular examples considered in sections 4 and 6 we restrict our attention to the case where f and f^{-1} take the form of a polynomial. In the more general

setting one can modify the arguments presented here in one of several different ways, for example by either considering Taylor expansions of f and f^{-1} and rigorously bounding the remainders, or by computing by hand coarse bounds on $\|[Df]^{-1}(x)\|_M$ with the phase space variable $|p-x| \le \rho'$ (and then using the hypothesis that $\|P_N - p\|_\nu \le \rho'$), or even by using the Cauchy bounds of Lemma 2.9.

Under the present assumption that f^{-1} is a polynomial of degree M,

$$[Df]^{-1}(z) = \sum_{|\beta|\ge 0}^{M-1} B_\beta z^\beta,$$

where each B_β is an $n \times n$ matrix. As remarked in the introduction, assumptions A2 and A3 allow us to explicitly determine the coefficients of $P_N(\theta) = \sum_{0\le|\alpha|\le N} a_\alpha \theta^\alpha$. Then we can determine the matrix coefficients of the polynomial expression

$$(Df[P_N(\theta)])^{-1} = \sum_{0\le|\alpha|\le \bar{M}} C_\alpha \theta^\alpha.$$

Here $\bar{M} = N(M-1)$, and the coefficients C_α depend on B_β and a_α. This immediately leads to

$$\|Df[P_N]^{-1}\|_{\Sigma,\nu} \le \sum_{0\le|\alpha|\le\bar{M}} \|C_\alpha\|\theta^{|\alpha|}. \tag{2.18}$$

However, in the next section we maintain generality, and in the definition of validation values we simply ask for any bound on $\|[Df]^{-1}(P_N)\|_{\Sigma,\nu}$. Obtaining such a bound in practice is a problem-dependent issue.

2.4. Real manifolds associated with complex conjugate pairs of eigenvalues. All of our theoretical considerations are formulated in terms of holomorphic functions, yet all of our applications are real dynamical systems. Nevertheless several of the examples we consider admit complex conjugate eigenvalues. We now explain the method used to pass from a parameterization of a complex invariant manifold to a real one, so that complex conjugate eigenvalues present no obstruction to our methods.

Consider a real analytic map $f\colon D_\nu(p) \subset \mathbb{R}^n \to \mathbb{R}^n$ with $p \in \mathbb{R}^n$ a fixed point having that $Df(p)$ is diagonalizable with n_s stable eigenvalues. Suppose that l_s of these are real and that there are m_s complex conjugate pairs, so that $l_s + 2m_s = n_s$. We order these so that the real eigenvalues come before the complex conjugate pairs. Then λ_i for $1 \le i \le l_s$ are real, and $\lambda_{l_s+2j-1}, \lambda_{l_s+2j}$ for $1 \le j \le m_s$ are complex conjugate pairs. In other words ,

$$\lambda_{l_s+2j} = \overline{\lambda_{l_s+2j-1}}$$

for each $1 \le j \le m_s$.

We extend to a complex dynamical system $f : B_\nu(p) \subset \mathbb{C}^n \to \mathbb{C}^n$ in the usual way (namely, replacing the real variables in the power series expansion of f with complex variables). Let $\hat{P}\colon B_\nu(0) \subset \mathbb{C}^{n_s} \to \mathbb{C}^n$ denote the parameterization of the complex stable manifold. Let

$\alpha \in \mathbb{N}^{n_s}$ be a multi-index, and a_α be the corresponding coefficient of $\hat{P}$. For $1 \leq j \leq m_s$ define

$$\text{conj}_j(\alpha_1, \ldots, \alpha_{l_s}, \ldots, \alpha_{l_s+2j-1}, \alpha_{l_s+2j}, \ldots) = (\alpha_1, \ldots, \alpha_{l_s}, \ldots, \alpha_{l_s+2j}, \alpha_{l_s+2j-1}, \ldots)$$

to be the function which switches the indices associated with the jth complex conjugate pair of eigenvalues. In section 4.1 we see that the coefficients of $\hat{P}$ solve the linear system given by (4.1). One can check that from the symmetry properties of (4.1) we have

$$a_{\text{conj}(\alpha)} = \overline{a_\alpha}.$$

Then for each complex coefficient appearing in the series expansion of $\hat{P}$ we have that the complex conjugate is a coefficient as well. It follows that the function P defined by the complex conjugate change of variables

$$P(\theta_1, \ldots, \theta_{l_s}, \theta_{l_s+1}, \theta_{l_s+2}, \ldots, \theta_{n_s-1}, \theta_{n_s})$$

$$= \hat{P}(\theta_1, \ldots, \theta_{l_s}, \theta_{l_s+1} + i\theta_{l_s+2}, \theta_{l_s+1} - i\theta_{l_s+2}, \ldots, \theta_{n_s-1} + i\theta_{n_s}, \theta_{n_s-1} - i\theta_{n_s})$$

has real image; hence it parameterizes the real stable manifold of f in real phase space. The domain for the real chart map P is defined by requiring that $|\theta_i| \leq \nu$ for $1 \leq i \leq l_s$ and $|\theta_{l_s+2j-1} \pm i\theta_{l_s+2j}| \leq \nu$ for $1 \leq j \leq m_s$. This explains why we take the poly-cylinders $U_\nu^{(l_s,m_s)}(0) = U_\nu$ as defined in section 2.1 to be the domain of P. We remark that $P \circ \Lambda$ is actually shorthand for

$$P(\Lambda\theta) = \hat{P}(\lambda_1\theta_1, \ldots, \lambda_{n_s-1}(\theta_{n_s-1} + i\theta_{n_s}), \overline{\lambda_{n_s-1}}(\theta_{n_s-1} - i\theta_{n_s})),$$

in which case $P(\Lambda\theta) \subset U_\nu$ as desired.

Now suppose that $\hat{P} = \hat{P}_N + \hat{h}$ on the poly-disk $B_\nu(0) \subset \mathbb{C}^{n_s}$ and $\|\hat{h}\|_\nu \leq \delta$. Define the real polynomial approximation and truncation error P_N and h by the same complex conjugate change of variables as above. In order to bound derivatives of h in terms of the bounds on $\hat{h}$ we apply the Cauchy bounds of Lemma 2.9 in conjunction with the chain rule and obtain

$$\|Dh\|_{\nu e^{-\sigma}} \leq \frac{\pi(2l_s + 8m_s)}{\nu\sigma}\delta \tag{2.19}$$

and

$$\|D^2h\|_{\nu e^{-\sigma}} \leq \frac{\pi^2(2l_s + 8m_s)^2}{\nu^2\sigma^2}\delta \tag{2.20}$$

for any $0 < \sigma \leq 1$. The estimates hold on the reduced poly-cylinder $U_{\nu e^{-\sigma}}$.

3. A posteriori validation theorem for stable manifolds.

We begin by defining the validation values which are used in the hypotheses of Theorem 3.2.

Definition 3.1. *Let $f\colon B_\rho(p) \subset \mathbb{C}^n \to \mathbb{C}^n$ satisfy* A1 *and* A2, *and let $P_N\colon \mathbb{C}^{n_s} \to \mathbb{C}^n$ be an Nth order polynomial with $N \geq 2$ that satisfies the linear constraints of* (1.3) *as well as*

assumption A3. *A collection of positive constants* ν, ϵ_{tol}, ρ, ρ', μ^*, $\bar{K}$, *and* K_1 *are validation values for* P_N *if*

$$\|f \circ P_N - P_N \circ \Lambda\|_{\Sigma,\nu} \leq \epsilon_{\text{tol}}, \tag{3.1}$$

$$\|P_N - p\|_{\Sigma,\nu} \leq \rho' < \rho, \tag{3.2}$$

$$0 < \max_{1 \leq i \leq n_s} |\lambda_i^s| \leq \mu^* < 1, \tag{3.3}$$

$$\left\|[Df]^{-1}(P_N)\right\|_{\Sigma,\nu} \leq \bar{K}, \tag{3.4}$$

$$\max_{\substack{\beta \in \mathbb{Z}^n \\ |\beta| = 2}} \max_{1 \leq j \leq n} \left\|\partial^\beta f_j\right\|_{B_\rho(p)} \leq K_1. \tag{3.5}$$

Note that condition (3.1) is the a posteriori error, (3.2) says that the image of P_N is contained in the interior of the poly-disk of radius ρ about $p \in \mathbb{C}^n$, (3.3) bounds the spectral radius of the differential away from the unit circle, (3.4) bounds the inverse of the differential of f on the image of P_N, and (3.5) bounds the second derivatives of f on the disk of radius ρ about p. Strategies for choosing validation values are discussed in sections 4.3 and 4.4, where computations based on Theorem 3.2 are presented. For the moment we assume that a set of validation values are given. Unstable manifolds are validated by exploiting the fact that they are stable manifolds for the inverse map f^{-1}. We now come to the main result of the paper.

Theorem 3.2 (a posteriori manifold validation). *Given assumptions* A1–A3, *suppose that* ν, ϵ_{tol} , K_1, $\bar{K}$, ρ, ρ', *and* μ^* *are validation values for the polynomial* P_N. *Let* N_f *be as in* (2.5), *and suppose that* $N \in \mathbb{N}$ *and* $\delta > 0$ *satisfy the three inequalities*

$$N + 1 > \frac{-\ln(\bar{K})}{\ln(\mu^*)}, \tag{3.6}$$

$$\delta < e^{-1} \cdot \min\left(\frac{1 - \bar{K}(\mu^*)^{N+1}}{2n\pi N_f \bar{K} K_1}, \rho - \rho'\right), \tag{3.7}$$

$$\delta > \frac{2\bar{K}\epsilon_{\text{tol}}}{1 - \bar{K}(\mu^*)^{N+1}}. \tag{3.8}$$

Then there is a unique analytic N*-tail* $h\colon B_\nu(0) \subset \mathbb{C}^{n_s} \to \mathbb{C}^n$ *having that*

$$\|h\|_\nu \leq \delta \tag{3.9}$$

and that

$$P(\theta) = P_N(\theta) + h(\theta)$$

is the exact solution of (1.2). *Since* P_N *equals* P *exactly to* N*th order by* A3 *and* h *is an* N*-tail,* h *is the truncation error associated with the* N*th order Taylor expansion of* P. *Moreover, since* h *is analytic and uniformly bounded on* $B_\nu(0)$ *we use the Cauchy estimates of Lemma* 2.9 *to bound derivatives of the truncation error. The coefficients* $a_\alpha \in \mathbb{C}^n$ *for the power series expansion of* P *satisfy the bounds*

$$|a_\alpha| \leq \frac{\delta}{\nu^{|\alpha|}} \quad \text{for } |\alpha| > N. \tag{3.10}$$

We discuss the strategy of the proof of Theorem 3.2 before turning to the details. Let P_N satisfy A3, and assume that the appropriate validation values have been obtained. Define

$$h := P - P_N.$$

Observe that to prove Theorem 3.2 it is sufficient to prove that $h \in \mathbb{H}_N$ and $\|h\|_\nu < \delta$.

Key to the latter proof is the fact that h solves the functional equation that we derive as follows. Consider the Taylor expansion about $P_N(\theta)$ for any $\theta \in B_\nu(0)$ as given by

$$f[P_N(\theta) + h(\theta)] = f[P_N(\theta)] + (Df)[P_N(\theta)]\, h(\theta) + R_{P_N(\theta)}[h(\theta)]. \tag{3.11}$$

Note that (3.2) of Definition 3.1 gives that $\|P_N - p\|_\nu \le \|P_N - p\|_{\Sigma,\nu} \le \rho'$, so that for each $\theta \in B_\nu$ we have that $P_N(\theta) \in B_{\rho'}(p)$. In other words, image$(P_N) \subset B_{\rho'}(p)$. Also, we have from (3.7) of Theorem 3.2 that $\|h\|_\nu \le \delta < (\rho - \rho')e^{-1} < \rho - \rho'$, so that image$(h) \subset B_{\rho-\rho'}(0)$. Then the remainder term $R_{P_N(\theta)}[h(\theta)]$ is analytic in the variable θ on $B_\nu(0)$ by Lemma 2.4.

We rewrite the invariance equation (1.2) as

$$f[P(\theta)] = f[P_N(\theta) + h(\theta)] = [P_N + h]\Lambda_s(\theta),$$

and in turn rewrite (3.11) as

$$P_N(\Lambda_s(\theta)) + h(\Lambda_s(\theta)) = f[P_N(\theta)] + (Df)[P_N(\theta)]\, h(\theta) + R_{P_N(\theta)}(h(\theta)). \tag{3.12}$$

Define

$$E(\theta) = f[P_N(\theta)] - P_N(\Lambda_s\theta), \tag{3.13}$$

and note that E is an analytic N-tail by assumption A3. This allows us to reduce (3.12) to the functional equation

$$h[\Lambda_s\theta] - (Df)[P_N(\theta)]h(\theta) = E(\theta) + R_{P_N(\theta)}(h(\theta)). \tag{3.14}$$

As indicated above, our goal is to show that this equation has a solution in $\mathbb{H}_N$. This is done by casting (3.14) as a fixed point problem.

Consider $\mathfrak{L} : \mathbb{H}_N \to \mathbb{H}_N$ defined by the left-hand side of (3.14); that is, for any $q \in \mathbb{H}_N$ let

$$\mathfrak{L}[q](\theta) := q[\Lambda_s\theta] - (Df)[P_N(\theta)]q(\theta).$$

The following lemma (proved below) provides conditions under which we can find solutions to the equation $\mathfrak{L}[q] = p$.

Lemma 3.3. *Let $\bar{K}$ and μ^* be validation values as in Definition 3.1. Suppose that N satisfies assumption (3.6) of Theorem 3.2. Then the linear operator $\mathfrak{L}$ is boundedly invertible on $\mathbb{H}_N$, so that for any $p \in \mathbb{H}_N$ there exists a unique solution to the equation*

$$\mathfrak{L}[q] = p. \tag{3.15}$$

Moreover, we have the bound

$$\|\mathfrak{L}^{-1}\|_{\mathbb{H}_N} \le \frac{\bar{K}}{1 - \bar{K}(\mu^*)^N}.$$

Lemma 3.3 indicates that $\mathfrak{L}$ is invertible and thus, with the functional equation (3.14) in mind, we can define the nonlinear operator $\Phi : \mathbb{H}_N \to \mathbb{H}_N$ by

$$\Phi(h) = \mathfrak{L}^{-1}\left[E(\theta) + R_{P_N(\theta)}(h(\theta))\right]. \tag{3.16}$$

By the preceding discussion $P = P_N + h$ is an exact solution of (1.2) if and only if h is a fixed point of (3.16). As the following lemma (proved below) indicates, we employ the contraction mapping theorem to verify the latter condition.

Lemma 3.4. *Under the hypotheses of Theorem 3.2, Φ is a contraction on the closed ball $V_\delta = \{h \in \mathbb{H}_N : \|h\|_\nu \le \delta\}$.*

A consequence of Lemma 3.4 is that h is analytic on $B_\nu(0)$ and hence (3.10) is satisfied, which completes the proof of Theorem 3.2.

Proof of Lemma 3.3. We need to find a unique solution to (3.15), where p and q are bounded analytic N-tails on B_ν. Define the linear operator

$$A[q](\theta) := ((Df)[P_N(\theta)])^{-1} q(\Lambda\theta).$$

Using the operator norm,

$$\begin{aligned}
\|A\|_{\mathbb{H}_N} &= \sup_{\|\eta\|_\nu = 1} \|(Df[P_N])^{-1}(\eta \circ \Lambda_s)\|_\nu \\
&\le \sup_{\|\eta\|_\nu = 1} \bar{K}|\mu^*|^{N+1}\|\eta\|_\nu \\
&= \bar{K}|\mu^*|^{N+1} \\
&< 1,
\end{aligned} \tag{3.17}$$

where the first inequality is obtain using (2.18) and Lemma 2.3 and the last inequality follows from assumption (3.6) of Theorem 3.2.

Using the definition of $\mathfrak{L}$, (3.15) can be rewritten as

$$(I - A)[q](\theta) = -((Df)[P_N(\theta)])^{-1} p(\theta). \tag{3.18}$$

Thus it follows from (3.17) and Theorem 2.6 that

$$\|(I - A)^{-1}\|_{\mathbb{H}_N} \le \sum_{i=0}^{\infty} \|A\|^i_{\mathbb{H}_N} = \frac{1}{1 - \bar{K}(\mu^*)^{N+1}}.$$

The bounded invertability of $(I - A)$ implies that there exists a unique solution to (3.18) of the form

$$q(\theta) = -(I - A)^{-1}((Df)[P_N(\theta)])^{-1} p(\theta).$$

Since p and q are arbitrary,

$$\begin{aligned}
\|\mathfrak{L}^{-1}\|_{\mathbb{H}_N} &\le \|(I - A)^{-1}\|_{\mathbb{H}_N} \|(Df[P_N])^{-1}\|_{\Sigma,\nu} \\
&\le \frac{\bar{K}}{1 - \bar{K}(\mu^*)^{N+1}},
\end{aligned}$$

as desired. ∎

Proof of Lemma 3.4. To prove that Φ is a contraction on V_δ we must show that

(i) $\Phi(V_\delta) \subset V_\delta$, and
(ii) there exists a $0 < \kappa < 1$ such that for any $h_1, h_2 \in V_\delta$

$$\|\Phi(h_1) - \Phi(h_2)\|_\nu \leq \kappa \|h_1 - h_2\|_\nu.$$

We begin by establishing (i). By Lemma 2.4 (see also [7, equation 75]) we have that for any $z, \eta \in \mathbb{C}^n$ with $|z| \leq \rho'$ and $|\eta| \leq \rho - \rho'$, $R_z(\eta)$ is analytic and

$$|R_z^j(\eta)| \leq N_f K_1 (\rho - \rho')^2, \quad 1 \leq j \leq n.$$

Here $R_z^j(\eta)$ denotes the jth component of $R_z(\eta)$. From the previous equation and the fact that image$(P_N) \subset B_{\rho'}(p)$ and $\delta < \rho - \rho'$ we have, as in the proof of [7, Lemma 4.5], that for each $\theta \in B_\nu$ the remainder term $R_{P_N(\theta)}(h(\theta))$ is analytic and

$$|R_{P_N(\theta)}(h(\theta))| \leq |R_z(h(\theta))| \leq \|R_z\|_\delta \leq \frac{\delta^2}{(\rho - \rho')^2} \|R_z\|_{\rho - \rho'} \leq N_f K_1 \delta^2. \tag{3.19}$$

Applying Lemma 3.3, (3.1), and (3.19) to the definition of Φ (see (3.16)) leads to

$$\|\Phi(h)\|_\nu \leq \frac{\bar{K}}{1 - \bar{K}(\mu^*)^{N+1}} \left(\epsilon_{\text{tol}} + N_f K_1 \delta^2\right).$$

By (3.7) and (3.8), respectively,

$$\frac{\bar{K}}{1 - \bar{K}(\mu^*)^{N+1}} \epsilon_{\text{tol}} \leq \frac{\delta}{2},$$
$$\frac{\bar{K}}{1 - \bar{K}(\mu^*)^{N+1}} N_f K_1 \delta^2 \leq \frac{\delta}{2},$$

and therefore the image of Φ lies in V_δ, as desired.

To establish (ii) we begin by considering the differential of the remainder term. Let $\theta \in B_\nu$ and $z = P_N(\theta)$. Recall that condition (3.2) for the validation values implies that $|z| \leq \rho'$. Again, by hypothesis (3.7) of Theorem 3.2, $\delta < (\rho - \rho')e^{-1} < \rho - \rho'$. Choose $0 < \sigma \leq 1$, and let $\omega := \delta / (\rho - \rho')e^{-\sigma}$. Then for any $h \in V_\delta$

$$\begin{aligned}
\|DR_z(h(\theta))\|_\delta &= \|DR_z \circ \omega\|_{(\rho - \rho')e^{-\sigma}} \\
&\leq \omega \|DR_z\|_{(\rho - \rho')e^{-\sigma}} \\
&\leq \frac{\delta}{(\rho - \rho')e^{-\sigma}} \frac{2\pi n}{(\rho - \rho')\sigma} \|R_z\|_{\rho - \rho'} \\
&\leq \frac{2\pi n e^\sigma \delta}{(\rho - \rho')^2 \sigma} N_f K_1 (\rho - \rho')^2 \\
&\leq 2\pi n e N_f K_1 \delta,
\end{aligned}$$

where the second inequality follows from Lemma 2.9, the third inequality from Lemma 2.4 and (3.5), and the last inequality from the fact that $\sigma^{-1} e^\sigma$ is minimized at $\sigma = 1$. Thus, by the mean value theorem,

$$|R_z^j(h_1(\theta)) - R_z^j(h_2(\theta))| \leq 2ne\pi N_f K_1 \delta \|h_1 - h_2\|_\nu \tag{3.20}$$

for all $h_1, h_2 \in V_\delta$ and all $j \in \{1, \dots, n\}$.

Using the definition of Φ,

$$\begin{aligned}\|\Phi(h_1) - \Phi(h_2)\|_\nu &= \left\|\mathfrak{L}^{-1}[E - R_{P_N(\theta)}(h_1)] - \mathfrak{L}^{-1}[E - R_{P_N(\theta)}(h_2)]\right\|_\nu \\ &= \|\mathfrak{L}^{-1}[R_{P_N(\theta)}(h_1) - R_{P_N(\theta)}(h_2)]\|_\nu \\ &\le \frac{\bar{K}}{1 - \bar{K}(\mu^*)^{N+1}} 2ne\pi N_f K_1 \delta \|h_1 - h_2\|_\nu,\end{aligned}$$

where the last inequality follows from Lemma 3.3 and (3.20). Observe that by hypothesis (3.7) of Theorem 3.2,

$$\frac{\bar{K}}{1 - \bar{K}(\mu^*)^{N+1}} 2ne\pi N_f K_1 \delta < 1,$$

and hence (ii) is established. ∎

4. Computing stable manifolds.

4.1. Computation of the power series coefficients: Theory and practice.

Suppose that $P : U_\nu \subset \mathbb{R}^k \to \mathbb{R}^n$ parameterizes the k dimensional stable manifold of $f : \mathbb{R}^n \to \mathbb{R}^n$ at p_0, and denote the power series expansion of P by

$$P(\theta) = \sum_{|\alpha|>0} a_\alpha \theta^\alpha.$$

A formal computation shows that for any multi-index $|\alpha| \ge 2$ the power series coefficient a_α satisfies the so-called *homological equation*,

$$\left[Df(p_0) - \left(\lambda_1^{\alpha_1} \cdot \cdots \cdot \lambda_k^{\alpha_k}\right) \mathrm{Id}\right] a_\alpha = s_\alpha. \tag{4.1}$$

Here s_α is a function of all of the lower order coefficients $a_{\alpha'}$ with $|\alpha'| < |\alpha|$. The form of the function s_α depends only on the nonlinearity of the function f. Then computing the coefficient a_α requires solving a linear system of equations involving only the local information p_0, $Df(p_0)$, and the stable eigenvalues $\lambda_1, \dots, \lambda_k$, assuming the lower order coefficients have been computed recursively.

The homological equation for the parameterization method is developed abstractly in [10] and is derived for the concrete example of the Lomelí map in [37, 36]. For example, we recall from the latter reference that for a one dimensional manifold of the Lomelí map the right-hand side of the homological equation is given by

$$s_n = -\sum_{k=1}^{n-1} \begin{pmatrix} a a_k^1 a_{n-k}^1 + b a_k^1 a_{n-k}^2 + c a_k^2 a_{n-k}^2 \\ 0 \\ 0 \end{pmatrix}. \tag{4.2}$$

(We will use this formula explicitly in section 4.4.) The function s_α for a two dimensional manifold in the Lomelí system is derived in the same reference, and similar computations give the form of the function s_α for all of the examples studied in this paper. See also [7, 34] for similar developments for the Gray–Scott and Lorenz differential equations, respectively.

Now we consider the question of whether or not the homological equations have solutions. In other words, when is the chart map P with coefficients defined by (4.1) formally well defined? Throughout the discussion recall that μ^* denotes the magnitude of the stable eigenvalue which is closest to the unit circle, and let $\mu_* > 0$ denote the magnitude of the stable eigenvalue with magnitude closest to zero. (There could be several eigenvalues which share the minimum and maximum moduli, but the moduli themselves are unique.)

The homological equation (4.1) shows that the coefficient a_α is well defined as long as

$$\lambda_1^{\alpha_1} \cdot \cdots \cdot \lambda_k^{\alpha_k} \neq \lambda_i$$

for any $1 \leq i \leq k$. We say that there is a *resonance at* α whenever there is equality in the expression above. Note that since

$$|\lambda_1^{\alpha_1} \cdot \cdots \cdot \lambda_k^{\alpha_k}| \leq (\mu^*)^{|\alpha|},$$

there can be no resonances for any α with

$$(\mu^*)^{|\alpha|} < \mu_*.$$

Then it is sufficient to check directly that there are no resonances for each of the multi-indices α with

$$2 \leq |\alpha| \leq \frac{\ln(\mu_*)}{\ln(\mu^*)}, \tag{4.3}$$

in order to conclude that there are no resonances at all orders. Since this gives a finite number of conditions we conclude that a generic set of parameters leads to eigenvalues which are nonresonant.

The preceding discussion gives an algorithm for computing the coefficients of a generic chart map P to any desired finite order which we now make explicit. In what follows, a *self-validating algorithm* is an interval arithmetic algorithm whose output is mathematically rigorous in the following sense: If the algorithm executes successfully, then there exists a unique solution to the problem, and the true solution is contained in the outputted interval. If the algorithm cannot validate, then it throws an error message and terminates without output.

Let $\alpha_0 = p_0$ be an interval arithmetic enclosure of the fixed point $p_0 \in \mathbb{R}^n$. Such an enclosure is trivial to obtain when there are explicit formulas for the fixed point. Otherwise the fixed point can be found using a self-validating interval Newton method such as Algorithm 13.4 of [44]. Denote by $e_i = (0, \ldots, 1, \ldots, 0) \in \mathbb{Z}^k$ the multi-index with 1 in the ith position and 0's elsewhere, and let $\alpha_{e_i} = \xi_i$ be an interval arithmetic enclosure of the eigenvector associated with the ith stable eigenvalue of $Df(p_0)$. We use the self-validating IntLab algorithm "verifyeig" (discussed in section 13.4 of [44]) for computing interval enclosures of eigenvalues and eigenvectors.

Next we fix some $N \in \mathbb{N}$ and use the self-validating IntLab algorithm "verifylss" (discussed in section 10.5 of [44]) to solve the homological equations defining the coefficients a_α for $2 \leq |\alpha| \leq N$. Any time the algorithm executes successfully, we conclude that there is no

Table 1

Coefficient computation performance: Three dimensional manifold in four dimensional phase space. The number of coefficient refers to the number of coefficients for a polynomial of degree N, but each of these coefficients is a vector of length 4. The relative error is computed by dividing the radius of the interval enclosure of the absolute value of each component of each coefficient by the right end point of its absolute value. This gives a measure of the round-off error as a proportion of the magnitude (or the number of significant digits). We consider the relative error only of coefficients whose magnitude is greater than 10^{-50}, as we expect that coefficients which are almost zero may lose significance.

Order	Number coeff	Comp time	Worst relative error
2	10	0.17 sec	8.4×10^{-15}
3	20	0.38 sec	1.8×10^{-14}
4	35	0.63 sec	2.65×10^{-14}
5	56	1.16 sec	3.75×10^{-14}
10	286	3.9 sec	9.8×10^{-14}
15	816	11.9 sec	1.35×10^{-13}
20	1771	24.5 sec	1.4×10^{-13}
30	21824	97.2 sec	1.5×10^{-13}

resonance at order α. If the algorithm executes successfully for all $1 < |\alpha| < \ln(\mu_*)/\ln(\mu^*)$, then we have ruled out resonances to all orders, and we conclude that P is formally well defined. If P_N is the polynomial with interval coefficients resulting from such a computation, then we know that the interval coefficients of P_N enclose the true coefficients (which exist uniquely).

We now illustrate the computational costs associated with computing the coefficients for the parameterization polynomial P_N. We consider the computation time as a function of both the polynomial order N and the dimension of the invariant manifold. The example system we consider is the delayed Hénon family, as this map is well defined on $\mathbb{R}^n$ for any value of n. In all computations we take $a = 1.6$ and $b = 0.1$, in order to conform with [49]. In particular, for this choice of parameters the map has a fixed point $p_1 = (x_1, \ldots, x_1) \in \mathbb{R}^n$, with $x_1 \in \mathbb{R}$ enclosed in the interval $x_1 \subset B(0.557857598881097, 2.221 \times 10^{-16})$, and a codimension one stable manifold. We remark that all the computations in this subsection are carried out by making small modifications to the program henonLikeStableMan_intval.m [38].

First we consider the cost of computing parameterization coefficients of the three dimensional stable manifold with varying polynomial order when the dimension of the phase space is $n = 4$. We use the IntLab software [45] to obtain interval enclosures of the stable eigenvalues with radii all smaller than 3×10^{-16}, and enclosures of the eigenvalues in intervals of radii smaller than 3.5×10^{-15} in each component.

Table 1 reports the performance data for computations of the coefficients with orders between $N = 2$ and $N = 30$. Each coefficient is a vector in $\mathbb{R}^4$ (the solution of the homological equation, which is a 4×4 linear system), so each nonzero coefficient is a vector of four intervals. We also give an indication of the worst case relative error for any of the coefficients. We note that the increase in time as a function of N is not due only to the increase in the number of polynomial coefficients. It is also important that the right-hand side of the homological equation depend recursively on all lower order coefficients, so that computing the right-hand sides becomes more time consuming as N increases.

Next we compute the parameterization to third order for phase space dimensions 4 through

Table 2

Coefficient computation performance: Third order approximation of codimension one stable manifold for the Hénon-like map. The dimension of the phase space is recorded in the leftmost column. The dimension of the manifold is always one less. The run time and worst relative error of a coefficient are shown as well.

Phase space dim	Number nonzero coeff	Comp time	Worst relative error
4 (3-D manifold)	20	0.36 sec	1.71×10^{-14}
5 (4-D manifold)	35	0.59 sec	2.31×10^{-14}
6 (5-D manifold)	56	1.04 sec	2.27×10^{-14}
7 (6-D manifold)	84	1.28 sec	2.75×10^{-14}
8 (7-D manifold)	120	3.10 sec	2.65×10^{-14}
9 (8-D manifold)	165	9.45 sec	3.36×10^{-14}
10 (9-D manifold)	220	45.4 sec	3.35×10^{-14}
11 (10-D manifold)	286	275 sec	3.32×10^{-14}

11 (manifold dimensions 3 through 10). The results are given in Table 2. We illustrate the computations at order $N = 3$ in each dimension because computing the ten dimensional manifold to order $N = 4$ in eleven dimensions takes 2.39 hours. We see that each increase in dimension leads to an increase in computation time by roughly a factor of four.

4.2. Numerical radius of validity for formal solutions. Suppose that we have computed

$$P_N(\theta) = \sum_{0 \le |\alpha| \le N} a_\alpha \theta^\alpha,$$

as discussed in the previous section. While P_N is a polynomial and hence an entire function, we do not expect that P_N is a good approximation to the chart map P for all θ. Instead, we would like to determine a fixed domain on which the approximation is "good." The following definition makes this precise.

Definition 4.1. *Let $\epsilon > 0$ be a prescribed error tolerance, $\nu > 0$, and $P_N \colon U_\nu \subset \mathbb{R}^k \to \mathbb{R}^n$. We call the number ν* an ϵ-numerical radius of validity *for the approximation P_N if*

$$\|f \circ P_N - P_N \circ \Lambda\|_{\Sigma,\nu} \le \epsilon. \tag{4.4}$$

This is the residual, or a posteriori, error associated with an approximate solution of (1.2). We are asking for a ν so that the residual is less than a proscribed ϵ. As discussed in section 2, we have that

$$\sup_{|\theta| \le \nu} \|f \circ P_N - P_N \circ \Lambda\| \le \|f \circ P_N - P_N \circ \Lambda\|_{\Sigma,\nu},$$

which is how we are able to pass from numerical computations in the Σ norm to conclusions in the C_0 norm (for example, the estimate given by (3.9) in the conclusion of Theorem 3.2). In practice, numerical experimentation is enough to select a good ν.

In the applications considered in the present work, f is a polynomial map. Thus, if we let A_α and B_α denote the nonzero power series (polynomial) coefficients of $f \circ P_N$ and $P_N \circ \Lambda$, then there is an $\bar{N}$ so that

$$\|f \circ P_N - P_N \circ \Lambda\|_{\Sigma,\nu} = \sum_{|\alpha|=0}^{\bar{N}} |A_\alpha - B_\alpha| \, \nu^{|\alpha|}. \tag{4.5}$$

In particular, the evaluations of the necessary Σ norm reduce to finite sums. If f is not a polynomial, then there is a Taylor remainder associated with the A_α. This additional complication presents no serious impediments.

4.3. Choice of validation values. The following result illustrates that validation values for the problem exist as long as (a) the eigenvalues are nonresonant and (b) we can solve the homological equations with infinite numerical precision. The proof will also illuminate the point where the numerical obstructions concerning (b) arise. (See also the discussion following the proof.)

Theorem 4.2. *Suppose* A1–A3 *and that there are no resonances for each multi-index* $\alpha \in \mathbb{N}^{n_s}$ *with* $2 \leq |\alpha| \leq \ln(\mu_*)/\ln(\mu^*)$. *Then there are validation values satisfying the hypotheses of Theorem* 3.2.

Proof. By A1, f is uniformly bounded and analytic in some neighborhood of p. Thus we can choose, using for example the Cauchy estimates, a $\rho_1 > 0$ so that the second partial derivatives of f are uniformly bounded on the poly-disk $B_{\rho_1}(p)$. Let K_1 denote an explicit uniform bound on the second partial derivatives of f on $B_{\rho_1}(p)$. Furthermore, let N_f count the number of not identically zero second derivatives of f. Note that $N_f \leq n^2$.

By A2, $Df(p)$ is diagonalizable; hence each stable eigenvalue has an eigenvector of norm one. Let A_s denote the matrix whose columns are these normalized stable eigenvectors. The hypothesis that $Df(p)$ is invertible implies that f is a local diffeomorphism. Thus there is a neighborhood of p so that f^{-1} and its first derivatives are uniformly bounded. By the inverse function theorem there exists a $\rho_2 > 0$ so that $[Df]^{-1}(z)$ is invertible and uniformly bounded for $|z| \leq \rho_2$. Since f and f^{-1} are analytic, $[Df]^{-1}(z)$ has a convergent power series on $B(p, \rho_2)$. Choose $\rho = \min(\rho_1, \rho_2)$.

Set $\bar{K} = 2\|[Df]^{-1}(p)\|$, and choose $N \in \mathbb{N}$ such that $N + 1 > \frac{-\ln(2\bar{K})}{\ln(\mu^*)}$. This choice of N satisfies the hypothesis of Theorem 3.2 given by (3.6). Furthermore,

$$\bar{K}(\mu^*)^{N+1} < \frac{1}{2}. \tag{4.6}$$

Define $P_N(\theta)$ to be the polynomial with $P_N(0) = p$, $DP_N(0) = A_s$, and whose coefficients solve the homological equation to order N. The coefficients exist and are unique by the nonresonance hypothesis of Theorem 4.2 and the discussion in section 4.1 (recall especially (4.3) and the surrounding discussion).

We expand

$$[Df]^{-1}(P_N(\theta)) = \sum_{|\alpha| \geq 0} C_\alpha \theta^\alpha$$

and choose a $\nu_1 > 0$ so that

$$\sum_{|\alpha| \geq 1} \|C_\alpha\| \nu_1^{|\alpha|} \leq \left\|[Df]^{-1}(p)\right\|.$$

Then

$$\left\|[Df]^{-1}(P_N)\right\|_{\Sigma,\nu_1} \leq 2\left\|[Df(p)]^{-1}\right\|_M = \bar{K}, \tag{4.7}$$

by the triangle inequality, as $C_0 = [Df]^{-1}(P_N(0)) = [Df]^{-1}(p)$. The preceding discussion shows that P_N so defined satisfies A3. In particular the coefficients of P_N solve (1.1) term by term up to order N.

If

$$\frac{1}{4n\pi N_f K_1 \bar{K}} < \rho,$$

then choose any $0 < \rho' < \rho$ so that

$$\frac{1}{4n\pi N_f K_1 \bar{K}} < \rho - \rho'.$$

Otherwise we choose (somewhat arbitrarily) $\rho' = 0.9\rho$ and define

$$\delta \equiv e^{-1} \min\left(\frac{1}{4n\pi N_f K_1 \bar{K}}, \rho - \rho'\right).$$

Since $f \circ P_N - P_N \circ \Lambda$ is an analytic N-tail, for $N \geq 1$ we can choose $\nu_2 > 0$ so that

$$\|f \circ P_N - P_N \circ \Lambda\|_{\Sigma,\nu_2} < \epsilon_{\text{tol}} \equiv \frac{1}{4\bar{K}}\delta.$$

Then we have the bounds

$$\frac{2\bar{K}\epsilon_{\text{tol}}}{1 - \bar{K}(\mu^*)^{N+1}} < 4\bar{K}\epsilon_{\text{tol}} = \delta \leq \frac{1}{4n\pi e N_f \bar{K} K_1} < e^{-1} \min\left(\frac{1 - \bar{K}(\mu^*)^{N+1}}{2n\pi N_f \bar{K} K_1}, \rho - \rho'\right), \tag{4.8}$$

where (4.6) is used to obtain the leftmost and rightmost inequalities. The inequalities (4.8) show that hypotheses (3.7) and (3.8) of Theorem 3.2 are satisfied.

Finally, since P_N is a polynomial with $P_N(0) = p$, there is a ν_3 so that

$$\|P_N - p\|_{\Sigma,\nu_3} \leq \rho'.$$

Choosing $\nu = \min(\nu_1, \nu_2, \nu_3)$, we see that collection of validation values ν, ϵ_{tol}, K_1, $\bar{K}$, ρ, ρ', and μ^* satisfy all the hypotheses of Theorem 3.2 for P_N and N. ■

Remarks 4.3.

(I) As an interesting aside observe that this gives a proof of the analytic stable manifold theorem under the assumptions of Theorem 4.2.

(II) A technical point can be extracted from the proof, that the nonresonance assumptions of Theorem 4.2 imply our standing assumption A3. In fact in the more theoretical exposition of [10, 11, 12] the nonresonance assumption is made from the beginning. However, due to the focus of the present work on rigorous a posteriori numerics, we prefer the assumptions as stated.

(III) We remark that in practice we are usually able to choose better bounds for $\bar{K}$ and δ. Nevertheless we can obtain some useful a priori information from the argument above. Note that we made the choice to define $\bar{K} = (1 + \gamma)\|Df(p)^{-1}\|_M$ and ν_1 in (4.7) so that

$$\sum_{|\alpha|\geq 1} \|C_\alpha\| v_1^{|\alpha|} \leq \gamma \|Df(p)^{-1}\|_M,$$

with $\gamma = 1$. However, the same argument holds with any $0 < \gamma \leq 1$. This leads us to define $\bar{K}_{\text{Test}} = \|Df(p)^{-1}\|_M$ and the quantities

$$\epsilon_{\text{Test}} = \frac{1}{4n\pi e N_f K_1 \bar{K}_{\text{Test}}^2} \tag{4.9}$$

and

$$N_{\text{Test}} = \left\lceil \frac{\ln\left(\bar{K}_{\text{Test}}\right)}{|\ln \mu^*|} \right\rceil - 1. \tag{4.10}$$

Then these quantities represent a theoretical best case scenario, in the sense that in any actual computer assisted proof we will have that $\bar{K} > \bar{K}_{\text{Test}}$, that the numerical a posteriori error must satisfy $\epsilon_{\text{tol}} < \epsilon_{\text{Test}}$, and that the polynomial order must have $N \geq N_{\text{Test}}$. This is seen by independently taking the limiting cases of $0 = \alpha$ and $\alpha = 1$. Computing these test values requires only local information about the map f near p and gives an indication of the feasibility of the validated parameterization computation (in terms of both required accuracy and polynomial order) with almost no initial effort.

We also note that in applications the finite precision of the digital computer introduces two obstructions for the preceding theorem. First, even if f is not at a resonance, it could be indistinguishably close to one from a numerical standpoint, and the self-validating IntLab algorithm "verifylss" [45, 44] could fail for one or more of the homological equations. In this case we may not able to compute a P_N which satisfies the hypotheses of Theorem 3.2.

Second, even if we can compute P_N to the desired order, the resulting coefficients are intervals. So while in theory $f \circ P_N - P_N \circ \Lambda$ is an N-tail, in practice we can have access only to interval enclosures of polynomials $f \circ P_N$ and $P_N \circ \Lambda$. Then our representation of $f \circ P_N - P_N \circ \Lambda$ is a polynomial whose coefficients up to order N have size comparable to the size of the interval enclosures of the polynomial coefficients a_α, plus a tail which is order ν^{N+1}.

More precisely, let r_α denote the radius of the interval enclosure of the term $A_\alpha - B_\alpha$, where A_α and B_α are as in (4.5). Then in practice the order of ϵ_{tol} behaves like $\sum_{0 \leq |\alpha| \leq N} r_\alpha \nu^{|\alpha|} + C\nu^{N+1}$ for some constant C. In typical computations r_α is at largest on the order of machine epsilon (and often much smaller). The constant C is determined by the decay rate of the coefficients of $f \circ P_N$ and is controlled either by making ν^{N+1} small or by appropriate choice of the scalings of the eigenvectors. In applications the quantity given by (4.9) is usually much larger than machine epsilon, and it is usually the case that this sum over r_α is very small.

4.4. Validated computation of P_N: A numerical example. In this section we sometimes use $B(x, r)$ to denote $D_r(x) \subset \mathbb{R}^n$, as the former is easier to read when r is a floating point number expressed in scientific notation.

Consider the Lomelí map with the following parameters:

$$\begin{aligned} a &= 0.5, \\ b &= -0.5, \\ c &= 1.0, \\ \alpha &= -0.15999999999999, \\ \tau &= 0.8. \end{aligned}$$

These are taken as intervals with zero radius. (We note that these parameters correspond to "Dullin–Meiss parameters" of $\bar{a} = 1$, $\bar{b} = 0.5$, $\bar{c} = 0.5$, $\epsilon = 0.32$, and $\mu = -2.4$. The interested reader can see [18, 36] for the meaning of these parameters.)

At these parameters the map has fixed points $p_0 = (x_0, x_0, x_0)$ and $p_1 = (x_1, x_1, x_1)$ with

$$x_0 \in B(-0.96568542494923, 1.112 \times 10^{-16}) \text{ and } x_1 \in B(0.16568542494923, 0.556 \times 10^{-16}).$$

We focus on p_0. We find that $Df(p_0)$ has eigenvalues

$$\lambda^u_{1,2} \in B(-0.14812334137550 \pm 1.26819024777211i, 4.648 \times 10^{-16}) \subset \mathbb{C}$$

and

$$\lambda^s \in B(0.61340397027639, 2.221 \times 10^{-16}) \subset \mathbb{R},$$

with associated eigenvectors

$$\xi^u_{1,2} \in B\left(\begin{bmatrix} 0.70894027376725 \\ -0.06441392030205 \pm 0.55149380772300i \\ -0.42316155243089 \mp 0.10021684334273i \end{bmatrix}, 2.788 \times 10^{-16}\right) \subset \mathbb{C}^3$$

and

$$\xi^s \in B\left(\begin{bmatrix} 0.30540789596717 \\ 0.49789031497392 \\ 0.81168420665679 \end{bmatrix}, 2.221 \times 10^{-16}\right) \subset \mathbb{R}^3.$$

We consider the validation of the one dimensional stable manifold of p_0. We can compute several of the necessary quantities immediately. For example, we see that $N_f = 4$ and $K_1 = \max(2|a|, 2|c|, |b|) = 2$, as the nonzero second partials of f are constant. We also have that the explicit formula for the inverse of the differential is

$$[Df(x, y, z)]^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -\tau - 2ax - by & -bx - 2cy \end{pmatrix}.$$

We take $\bar{K}$ to be a number bounding the sum on the right-hand side in (2.18). Since $C_0 = [Df(p_0)]^{-1}$ is the first term in the expansion $[Df]^{-1}(P_N)(\theta) = \sum_{n=0}^{N} C_\alpha \theta^n$, we now use interval arithmetic to obtain the numerical test quantity

$$\|[Df]^{-1}(p_0)\| \leq 2.766 = \bar{K}_{\text{Test}}.$$

Then for any $\nu > 0$ the estimate on $\|[Df]^{-1}(P_N)(\theta)\|_{\nu,\Sigma}$ given by (2.18) will be larger than 2.766, and we expect that $\bar{K}$ must be larger than 2.766 no matter how small ν is.

Again using interval arithmetic, we check the a priori indicator given by (4.10) and have that

$$\frac{-\ln(2.766)}{\ln(|\lambda^s|)} \leq 2.1.$$

This suggests that we will need at least polynomial order $N_{\text{Test}} = 2$ in order to obtain a validated computation. So we try to compute validated bounds for the minimal case of $N = 2$.

We solve the homological equation for P_N to order $N = 2$ and obtain the approximation

$$P_N(\theta) = p_0 + \xi^s \theta + a_2 \theta^2,$$

where

$$a_2 \in B\left(\begin{bmatrix} -0.06926609073821 \\ -0.18408886165507 \\ -0.48925395708471 \end{bmatrix}, 8.605 \times 10^{-16} \right).$$

For the sake of completeness we observe that a_2 is the solution of the explicit equation

$$[Df(p_0) - (\lambda^s)^2 I]\, a_2 = -\begin{pmatrix} a\,(\xi_1^s)^2 + b\,(\xi_1^s)(\xi_2^s) + c\,(\xi_2^s)^2 \\ 0 \\ 0 \end{pmatrix},$$

where we have let $\xi^s = (\xi_1^s, \xi_2^s, \xi_3^s)^T$ and used the form of s_n from (4.2).

We expect that P_N is a cubically good approximation to P in the neighborhood of $\theta = 0$; i.e., for $|\theta| \leq 5 \times 10^{-6}$ we expect errors on the order of 2.21×10^{-16} (double precision machine epsilon). We see that this expectation is met by computing

$$\|f \circ P_N - P_N \circ \lambda^s\|_{\Sigma,\nu} \leq 1.25 \times 10^{-15} \equiv \epsilon_{\text{tol}},$$

using interval arithmetic with $\nu = 5 \times 10^{-6}$. We also see that

$$\|P_N - p_0\|_\nu \leq 4.059 \times 10^{-6}.$$

Let $\rho' = 4.059 \times 10^{-6}$. Taking $\rho = 4.463 \times 10^{-6}$ guarantees that $\rho > \rho'$. Now we let $a_n = (a_n^1, a_n^2, a_n^3)^T$ and expand $[Df]^{-1}(P_N)$ to see that

$$[Df]^{-1}(P_N)(\theta) = [Df(p_0)]^{-1} + \sum_{n=1}^{N} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -2a\,a_n^1 - b a_n^2 & -b\,a_n^1 - 2c\,a_n^2 \end{pmatrix} \theta^n. \tag{4.11}$$

(This expansion is correct for any N, a fact we will make use of below.) Then we check that

$$\|[Df]^{-1}(P_N)\|_{\Sigma,\nu} \leq 2.77.$$

Taking $\bar{K} = 2.77$, we see that we have the inequality

$$N + 1 = 3 > 2.082 > \frac{-\ln(\bar{K})}{\ln(\lambda^s)},$$

and in fact

$$1 - \bar{K}|\lambda^s|^3 > 0.362.$$

Now we simply check (as always using interval arithmetic) that

$$\frac{2\bar{K}\epsilon_{\text{tol}}}{1 - \bar{K}(|\lambda^s|^{N+1})} \leq 1.899 \times 10^{-14} \equiv \delta,$$

that

$$\frac{1 - \bar{K}|\lambda^s|^{N+1}}{2n\pi e N_f \bar{K} K_1} > 8.67 \times 10^{-4} > \delta,$$

and that

$$\frac{\rho - \rho'}{e} > 4.05 \times 10^{-7} > \delta.$$

Then these validation values satisfy the hypotheses of Theorem 3.2, and we conclude that the truncation error $h(\theta) = P(\theta) - P_N(\theta)$ is an analytic function on $B(0, 5 \times 10^{-6}) \subset \mathbb{C}$ having $h(0) = h'(0) = h''(0)$ and

$$\|h\|_\nu \leq \delta = 1.9 \times 10^{-14}.$$

We note that the previous example could be done more or less "by hand," as all that was required was the computation of the eigenvalues and eigenvectors of a known matrix, the solution of a single linear system of equations, and then checking certain inequalities. Checking these inequalities involves at most adding up a long list of known numbers. The role of the computer in this kind of proof is only to speed up these tedious computations.

So if we raise the order of the polynomial approximation to $N = 25$, and use the same parameters for a, b, c, τ, and α, then we have the same fixed points, eigenvalues, and eigenvectors. This time, however, there are 24 linear systems of equations to solve, and the assistance of the computer is greatly appreciated. After a little experimentation we find that taking $\nu = 0.8$ gives

$$\|f \circ P_N - P_N \circ \lambda^s\|_{\Sigma,\nu} \leq 1.04 \times 10^{-14} \equiv \epsilon_{\text{tol}},$$

which is less than one hundred multiples of double precision machine epsilon. We find that

$$\|P_N - p_0\|_{\Sigma,\nu} < 1.139,$$

so that the image of P_N lies in a poly-disk in phase space of roughly size 1 about the fixed point p_0. We take $\rho = 1.252$ and $\rho' = 1.139$ and compute that

$$\|[Df]^{-1}(P_N)\|_{\Sigma,\nu} \leq 3.79 \equiv \bar{K},$$

again using the expansion given by (4.11).

As before we have that $K_1 = 2$, $N_f = 4$, and $n = 3$. In order to define δ we compute

$$\frac{2\bar{K}}{1 - \bar{K}|\lambda^s|^{N+1}}\epsilon_{\text{tol}} < 7.85 \times 10^{-14} \equiv \delta$$

and check that

$$\frac{-\ln(\bar{K})}{\ln(\lambda^s)} < 2.726 < N + 1 = 26,$$

$$\frac{1-\bar{K}|\lambda^s|^{N+1}}{2n\pi N_f \bar{K} K_1} > 6.44 \times 10^{-4} > \delta,$$

and

$$\frac{\rho - \rho'}{e} > 0.022 > \delta.$$

Again we have that the conditions of Theorem 3.2 are satisfied and conclude that there exists a unique analytic N-tail h defined on $B(0, 0.8) \subset \mathbb{C}$ having $h(\theta) = P(\theta) - P_N(\theta)$ and

$$\|h\|_{0.8} \leq 7.85 \times 10^{-14},$$

which is a little less than 400 multiples of double precision machine epsilon. Of course, the coefficients of P_N are real, so the image of P_N is real for real θ. The interested reader can find the computations discussed in this section implemented in the file a_tanglePaperExample1.m at [38].

We hope to draw the reader's attention to the constructive nature of these arguments. We compute a specific polynomial P_N (with interval coefficients) to a chosen fixed finite order. We know that the interval coefficients of our approximation P_N enclose the actual power series coefficients of P due to the self-validating properties of the IntLab algorithms. We then check that P_N has certain properties (satisfies some inequalities) and are able to conclude that the truncation error satisfies some bounds. The computations in the remainder of the paper are similar to those just discussed.

5. Rigorous computation of transverse heteroclinic and homoclinic orbits. The computations of section 4 indicate the practicality of computing high precision Taylor approximations of even reasonably high dimensional stable manifolds. As is indicated in the introduction, one of our motivations for developing the computational aspects of the theory is to be able to rigorously compute the existence of connecting orbits. In this section we state and prove a theorem that guarantees that a numerically computed connecting orbit provides a good approximation of a true connecting orbit. We begin by stating specific, but natural, assumptions that must be satisfied.

P1. Assume that $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is real analytic with hyperbolic fixed points $q, p \in \mathbb{R}^n$. Assume that $Df(q)$ and $Df(p)$ are diagonalizable, and that $n_u + n_s = n$, where $n_u = n_u(q)$ and $n_s = n_s(p)$ are the number of stable and unstable eigenvalues at q and p. Moreover, let l_s, l_u be the number of real distinct eigenvalues, and m_s, m_u be the number of complex conjugate pairs of eigenvalues at q and p, respectively. Note that $l_s + 2m_s = n_s$ and $l_u + 2m_u = n_u$.

P2. Assume that P_{N_s} is a parameterization of $W^s(p)$ which satisfies Theorem 3.2 with validation values ν_s, ϵ_s, K_{1s}, $\bar{K}_s$, ρ_s, ρ'_s, and μ^*_s. Suppose that P_{N_s} has truncation error bound δ_s on the poly-cylinder $U_{\nu_s}^{(l_s,m_s)}$.

P3. Assume that Q_{N_u} is a parameterization of $W^u(p)$ which satisfies Theorem 3.2 applied to f^{-1}. The associated validation values are ν_u, ϵ_u, K_{1u}, $\bar{K}_u$, ρ_u, ρ'_u, and μ^*_u. Suppose that Q_{N_u} has truncation error bound δ_u on the poly-cylinder $U_{\nu_u}^{(l_u,m_u)}$.

As suggested by assumptions P1–P3, we are interested in proving the existence of a connecting orbit from q to p. If $q \neq p$, then this is a heteroclinic orbit, and if $q = p$, this is a

homoclinic orbit. As is indicated in the introduction this is done by finding a fixed point of the connecting orbit operator equation that we present in a slightly more general form:

$$
\begin{aligned}
F(\theta, x, \phi) &= \begin{bmatrix} f^{-1}(x_1) - Q_{N_u}(\theta) - h_u(\theta) \\ f^{-1}(x_2) - x_1 \\ \vdots \\ f^{-1}(x_j) - x_{j-1} \\ f(x_j) - x_{j+1} \\ \vdots \\ f(x_{k-1}) - x_k \\ f(x_k) - P_{N_s}(\phi) - h_s(\phi) \end{bmatrix} \\
&= \begin{bmatrix} f^{-1}(x_1) - Q_{N_u}(\theta) \\ \vdots \\ f^{-1}(x_j) - x_{j-1} \\ f(x_j) - x_{j+1} \\ \vdots \\ f(x_k) - P_{N_s}(\phi) \end{bmatrix} + \begin{bmatrix} -h_u(\theta) \\ \vdots \\ 0 \\ 0 \\ \vdots \\ -h_s(\phi) \end{bmatrix} \\
&:= F_N(\theta, x, \phi) + H(\theta, \phi), \qquad (5.1)
\end{aligned}
$$

where $x = (x_1, \ldots, x_k) \in \mathbb{R}^{nk}$. To once again emphasize the a posteriori nature of our approach, recall that the strategy is to begin with an approximate zero of (5.1), i.e., $\hat{x} = (\hat{\theta}, \hat{x_1}, \ldots, \hat{x}_{k-1}, \hat{\phi}) \in \mathbb{R}^{n(k+1)}$ such that

$$\|F_N(\hat{x})\| \approx 0,$$

and then prove the existence of $x^* = (\theta^*, x_1^*, \ldots, x_k^*, \phi^*)$ close to $\hat{x}$ such that

$$F(x^*) = 0.$$

As in the case of the parameterization of the stable and unstable manifolds, we need to determine local bounds associated with and determined by the choice of approximations.

Definition 5.1. *Given assumptions* P1–P3 *and* F *as in* (5.1), *let* $\hat{x} = (\hat{\theta}, \hat{x_1}, \ldots, \hat{x}_k, \hat{\phi}) \in \mathbb{R}^{n(k+1)}$ *be a vector with* $|\hat{\theta}|_{(l_u,m_u)} < \nu_u$, $|\hat{\phi}|_{(l_s,m_s)} < \nu_s$, *and* $DF_N(\hat{x})$ *nonsingular. A collection of positive constants* A_N, A_u, A_s, $\hat{\epsilon}$, r, C_β, *and* C_P *are* connecting validation values *(CVV) if the following conditions are satisfied:*

$$A_N \geq \|DF_N(\hat{x})^{-1}\|_M, \qquad (5.2)$$

$$A_u \geq \max_{1 \leq i \leq n(k+1)} \sum_{j=1}^{n} [DF_N^{-1}(\hat{x})]_{ij}, \qquad (5.3)$$

$$A_s \geq \max_{1 \leq i \leq n(k+1)} \sum_{j=n(k+1)-n+1}^{n(k+1)} [DF_N^{-1}(\hat{x})]_{ij}, \qquad (5.4)$$

$$\hat{\epsilon} \geq |DF_N(\hat{x})^{-1} F_N(\hat{x})|, \qquad (5.5)$$

$$r < \min\{\nu_u - |\theta|_{(l_u,m_u)}, \nu_s - |\phi|_{(l_s,m_s)}\}. \qquad (5.6)$$

Given (5.6), *set*

$$\sigma_s \le \min\left\{-\ln\left(\frac{|\phi|_{(l_s,m_s)}+r}{\nu_s}\right),1\right\} \quad \textit{and} \quad \sigma_u \le \min\left\{-\ln\left(\frac{|\theta|_{(l_u,m_u)}+r}{\nu_u}\right),1\right\}, \tag{5.7}$$

and require that

$$C_\beta \ge \max_{\substack{1\le j\le k\\ 1\le i\le n\\ |\beta|=2}} \left\{\|\partial^\beta f_i\|_{D_r(\hat{x}_j)}, \|\partial^\beta f_i^{-1}\|_{D_r(\hat{x}_j)}\right\}, \tag{5.8}$$

$$C_P \ge \max\left(\|D^2 P_{N_s}\|_{U_r^{(l_s,m_s)}(\hat{\theta})} + \frac{\pi^2(2l_s+8m_s)^2}{\nu_s^2\sigma_s^2}\delta_s,\ \|D^2 Q_{N_u}\|_{U_r^{(l_u,m_u)}(\hat{\phi})} + \frac{\pi^2(2l_u+8m_u)^2}{\nu_u^2\sigma_u^2}\delta_u\right). \tag{5.9}$$

Observe that there is an implicit nonlinear relation between r, C_β, and C_P. Thus care must be taken in choosing CVV. We discuss this issue in greater detail in section 6. For the moment we take the CVV as given.

The following result is the basis for our proofs of the existence of connecting orbits.

Theorem 5.2. *Assume* P1–P3. *Let* A_N, A_u, A_s, $\hat{\epsilon}$, r, C_β, *and* C_P *be CVV for a vector* $\hat{x} = (\hat{\theta}, \hat{x}_1, \dots, \hat{x}_{k-1}, \hat{\phi}) \in \mathbb{R}^{n(k+1)}$ *for which* $|\hat{\theta}|_{(l_u,m_u)} < \nu_u$, $|\hat{\phi}|_{(l_s,m_s)} < \nu_s$, *and for which* $DF_N(\hat{x})$ *is nonsingular. Let*

$$\kappa \ge \bar{N}_f C_\beta + C_P, \tag{5.10}$$

where $\bar{N}_f$ *is the maximum of the number of nonzero second partials of* f *and* f^{-1},

$$\hat{\delta} \ge A_u\delta_u + A_s\delta_s, \tag{5.11}$$

$$M_N \ge A_u\frac{\pi(2l_u+8m_u)}{\nu_u\hat{\sigma}_u}\delta_u + A_s\frac{\pi(2l_s+8m_s)}{\nu_s\hat{\sigma}_s}\delta_s, \tag{5.12}$$

where

$$\hat{\sigma}_s = \min\left(-\ln\left(\frac{|\hat{\theta}|_{(l_s,m_s)}}{\nu_s}\right),1\right) \quad \textit{and} \quad \hat{\sigma}_u = \min\left(-\ln\left(\frac{|\hat{\phi}|_{(l_u,m_u)}}{\nu_u}\right),1\right).$$

Define

$$B_r(\hat{x}) = U_r^{(l_u,m_u)}(\hat{\theta}) \times D_r(\hat{x}_1) \times \cdots \times D_r(\hat{x}_k) \times U_r^{(l_s,m_s)}(\hat{\phi}).$$

If

$$M_N < 1, \tag{5.13}$$

$$2\frac{\hat{\epsilon}+\hat{\delta}}{1-M_N} \le r, \tag{5.14}$$

$$4\kappa A_N\frac{\hat{\epsilon}+\hat{\delta}}{(1-M_N)^2} \le 1, \tag{5.15}$$

then there exists a unique $x_* \in B_r(\hat{x}) \subset \mathbb{R}^{n(k+1)}$ *satisfying* $F(x_*) = 0$.

Proof. We use the Newton–Kantorovich approach to prove the existence and uniqueness of x_*. The subtle point is that we assume knowledge about $F_N(\hat{x})$ and the existence of $DF_N(\hat{x})^{-1}$ but want conclusions about the zero of F. Thus we must check that the hypotheses of Theorem 2.7 are satisfied with respect to F and DF.

Observe that

$$\begin{aligned} DF(\hat{x})^{-1} &= [DF_N(\hat{x}) + DH(\hat{x})]^{-1} \\ &= \left[DF_N(\hat{x}) \ \left(I + DF_N(\hat{x})^{-1} DH(\hat{x})\right) \ \right]^{-1} \\ &= [I + DF_N(\hat{x})^{-1} DH(\hat{x})]^{-1} \ DF_N(\hat{x})^{-1}, \end{aligned} \tag{5.16}$$

assuming the term on the right-hand side exists. Thus, to prove that $DF(\hat{x})$ is nonsingular, and hence that condition 1 of Theorem 2.7 is satisfied, it is sufficient to show that $I + DF_N(\hat{x})^{-1} DH(\hat{x})$ is invertible.

With this in mind, let $[DF_N^{-1}(\hat{x})]_{(a:b)}$, with $a < b \in \mathbb{N}$, denote the submatrix of $DF_N^{-1}(\hat{x})$ composed of columns a through b. Note that

$$\begin{aligned} DF_N^{-1}(\hat{x}) DH(\hat{x}) &= DF_N^{-1}(\hat{x}) \begin{bmatrix} D_\theta h_u(\hat{\theta}) & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & D_\phi h_s(\hat{\phi}) \end{bmatrix} \\ &= \left[[DF_N^{-1}(\hat{x})]_{(1:n)} Dh_u(\hat{\theta}) \mid 0 \mid \cdots \mid 0 \mid [DF_N^{-1}(\hat{x})]_{(n(k+1)-n+1:n(k+1))} Dh_s(\hat{\phi}) \right], \end{aligned}$$

and therefore

$$\begin{aligned} \|DF_N^{-1}(\hat{x}) DH(\hat{x})\|_M &\leq \left(n_u \max_{1 \leq i \leq n(k+1)} \sum_{j=1}^{n} [DF_N^{-1}(\hat{x})]_{ij} \right) \|Dh_u\|_{\nu_u e^{-\hat{\sigma}_u}} \\ &\quad + \left(n_s \max_{1 \leq i \leq n(k+1)} \sum_{j=n(k+1)-n+1}^{n(k+1)} [DF_N^{-1}(\hat{x})]_{ij} \right) \|Dh_s\|_{\nu_s e^{-\hat{\sigma}_s}} \\ &\leq M_N \\ &< 1, \end{aligned}$$

where the second inequality follows from (5.3), (5.4), (2.10), and (5.12) and the last inequality from (5.13). Note that the estimates of (5.12) take into account the effect of the complex and real eigenvalues when the Cauchy bounds are applied to h_s and h_u.

Now by Theorem 2.6 the matrix $I + DF_N^{-1}(\hat{x}) DH(\hat{x})$ is invertible and

$$\left\| [I + DF_N^{-1}(\hat{x}) DH(\hat{x})]^{-1} \right\|_M \leq \frac{1}{1 - M_N}. \tag{5.17}$$

It follows from (5.16) and (5.2) that

$$\|DF(\hat{x})^{-1}\|_M \leq \frac{A_N}{1 - M_N}. \tag{5.18}$$

We now turn our attention to condition 2 of Theorem 2.7, which is a Lipschitz bound on the differential DF. Define $g_{ij}\colon B_r(\hat{x}) \subset \mathbb{R}^{n(k+1)} \to \mathbb{R}$, where $1 \le i, j \le n(k+1)$, by

$$g_{ij}(z) = \partial_j F_i(z).$$

For $x, y \in B_r(\hat{x})$,

$$\begin{aligned} |g_{ij}(x) - g_{ij}(y)| &\le \|\nabla g_{ij}\|_{M,B_r(\hat{x})}|x-y| \\ &\le \sum_{\ell=1}^{n(k+1)} \|\partial_\ell g_{ij}\|_{B_r(\hat{x})}|x-y| \\ &\le \left(\sum_{\ell=1}^{n(k+1)} \|\partial_\ell \partial_j F_i\|_{B_r(\hat{x})}\right)|x-y|, \end{aligned} \tag{5.19}$$

where the last inequality follows from the mean value theorem. Observe that

$$\begin{aligned} \|DF(x) - DF(y)\|_M &\equiv \sup_{\substack{v \in \mathbb{R}^{n(k+1)} \\ |v|=1}} |[DF(x) - DF(y)]v| \\ &\le \max_{1\le i\le n(k+1)} \sum_{1\le j\le n(k+1)} |[DF(x) - DF(y)]_{ij}| \\ &= \max_{1\le i\le n(k+1)} \sum_{1\le j\le n(k+1)} |\partial_j F_i(x) - \partial_j F_i(y)| \\ &\le \left(\max_{1\le i\le n(k+1)} \sum_{j=1}^{n(k+1)} \sum_{\ell=1}^{n(k+1)} \|\partial_\ell \partial_j F_i\|_{B_r(\hat{x})}\right)|x-y|, \end{aligned} \tag{5.20}$$

where the last inequality follows from (5.19). We need an explicit Lipschitz bound. Note that

$$\begin{aligned} \|\partial_\ell \partial_j (h_u)_i\|_{B_r(\hat{x})} &= \|\partial_\ell \partial_j (h_u)_i\|_{U_r(\hat{\theta})} \\ &\le \|\partial_\ell \partial_j (h_u)_i\|_{\nu_u e^{-\sigma_u}} \\ &\le \frac{K^2\pi^2}{\nu_u^2 \sigma_u^2}\delta_u, \end{aligned} \tag{5.21}$$

where σ_u is defined by (5.7), the last inequality follows from (2.20), and $K = 2, 8$ whether or not the indices i, j are associated with complex conjugate variables.

Similarly,

$$\|\partial_\ell \partial_j (h_s)_i\|_{B_r(\hat{x})} \le \frac{K^2\pi^2}{\nu_s^2 \sigma_s^2}\delta_s, \tag{5.22}$$

where we have again used (2.20) and $K = 2, 8$ as determined by the real or complex conjugate variables. Considering the second partial derivatives of F one component at a time and

combining (5.21), (5.22), (5.8), and (5.9), we obtain

$$\max_{1\le i\le n(k+1)} \sum_{j=1}^{n(k+1)} \sum_{\ell=1}^{n(k+1)} \|\partial_\ell \partial_j F_i\|_{B_r(\hat{x})} \le \bar{N}_f C_\beta + C_P \le \kappa,$$

where the last inequality follows from (5.10). Combining this with (5.20) gives condition 2 of Theorem 2.7.

Now we are in a position to apply Theorem 2.7, and it is important to observe that at this point r and κ are already defined in Definition 5.1 and the statement of Theorem 5.2, respectively. Define

$$\epsilon_{NK} := \frac{\hat{\epsilon} + \hat{\delta}}{1 - M_N}. \tag{5.23}$$

To complete the proof we need to show that (2.7), (2.8), and (2.9) are satisfied simultaneously.

Observe that

$$\begin{aligned} \left|DF(\hat{x})^{-1}F(\hat{x})\right| &\le \left|[I + DF_N(\hat{x})^{-1}DH(\hat{x})]^{-1}\ DF_N(\hat{x})^{-1}F(\hat{x})\right| \\ &= \left|[I + DF_N(\hat{x})^{-1}DH(\hat{x})]^{-1}\ DF_N(\hat{x})^{-1}\left(F_N(\hat{x}) + H(\hat{x})\right)\right| \\ &\le \frac{1}{1 - M_N}\left(\left|DF_N^{-1}(\hat{x})F_N(\hat{x})\right| + \left|DF_N^{-1}(\hat{x})H(\hat{x})\right|\right), \end{aligned} \tag{5.24}$$

where the last inequality follows from (5.17). A bound for the last term in (5.24) can be obtained as follows:

$$\begin{aligned} \left|DF_N^{-1}(\hat{x})H(\hat{x})\right| &= \left|DF_N^{-1}(\hat{x}) \begin{bmatrix} h_u(\hat{\theta}) \\ 0 \\ \vdots \\ 0 \\ h_s(\hat{\phi}) \end{bmatrix}\right| \\ &= \left|[DF_N^{-1}(\hat{x})]_{(1:n)}h_u(\hat{\theta}) + [DF_N^{-1}(\hat{x})]_{(n(k+1)-n+1:n(k+1))}h_s(\hat{\phi})\right| \\ &\le \left(\max_{1\le i\le n(k+1)} \sum_{j=1}^{n} \left|[DF_N^{-1}(\hat{x})]_{ij}\right|\right) \|h_u\|_{\nu_u} \\ &\quad + \left(\max_{1\le i\le n(k+1)} \sum_{j=n(k+1)-n+1}^{n(k+1)} \left|[DF_N^{-1}(\hat{x})]_{ij}\right|\right) \|h_s\|_{\nu_s} \\ &\le \hat{\delta}, \end{aligned} \tag{5.25}$$

where the last inequality follows from (5.3), (5.4), and (5.11) and where δ_u and δ_s are as indicated in assumptions P2 and P3, respectively. Applying (5.5) and (5.25) to (5.24) gives

$$\left|DF(\hat{x})^{-1}F(\hat{x})\right| \le \epsilon_{NK},$$

and hence (2.7) is satisfied. Finally, observe that (5.14) implies (2.8), and (5.15) combined with (5.18) implies (2.9). Therefore the hypotheses of Theorem 2.7 have been validated, and thus the existence and uniqueness of x_* is guaranteed. ∎

Assuming that the hypotheses of the previous theorem are satisfied, the following corollary gives almost trivial additional conditions which imply the transversality of the intersection.

Corollary 5.3. *Supposing the hypotheses of Theorem 5.2, assume in addition that*

$$r \leq 4\frac{\hat{\epsilon} + \hat{\delta}}{1 - M_N} \tag{5.26}$$

and

$$4\kappa A_N \frac{\hat{\epsilon} + \hat{\delta}}{(1 - M_N)^2} < 1. \tag{5.27}$$

Then $W^U(q)$ and $W^s(p)$ intersect transversally along the connecting orbit associated with x_*.

Proof. We must show that $W^u(q)$ and $W^s(p)$ intersect transversally at each point on the connecting orbit. Observe that (5.16) holds for all $x \in B_r(\hat{x})$. Thus the additional assumptions that r satisfies the inequality of (5.26) and that the strict inequality of (5.27) holds is enough to apply Lemma 2.8. This gives that $DF(x_*)$ is nonsingular. What remains to be shown is that the invertibility of $DF(x_*)$ gives the transversality of the connecting orbit.

Assume for the moment that the local manifolds given by the images of the parameterizations $W^u_{\text{loc}}(p) = Q[U_{\nu_u}(0)]$ and $W^s_{\text{loc}}(p) = P[U_{\nu_s}(0)]$ intersect at x_*. In this case a connecting orbit is simply a zero of the equation

$$F(\theta, \phi) = Q(\theta) - P(\phi),$$

whose root is $x_* = (\theta_*, \phi_*)$. We are assuming that $DF(x_*)$ is nonsingular, and it follows that the columns of

$$DF(x_*) = [D_\theta Q(\theta_*) | - D_\phi P(\phi_*)]$$

span $\mathbb{R}^n$. But the columns of $D_\theta P(\phi_*)$ and $D_\phi Q(\theta_*)$ span $T_{P(\phi_*)}W^s(p)$ and $T_{Q(\theta_*)}W^u(p)$, respectively. It follows that $T_{P(\phi_*)}W^s(p)$ and $T_{Q(\theta_*)}W^u(p)$ span $\mathbb{R}^n$, which is to say that x_* is a point of transverse intersection.

Now suppose that F is the connecting orbit operator defined by (5.1) with $k \geq 1$, and let $x_* \in \mathbb{R}^{n(k+1)}$ be a solution of $F = 0$. Then $x_* = (\theta_*, x_1^*, \ldots, x_k^*, \phi_*)$, and we have that

$$\begin{aligned} f[Q(\theta_*)] &= x_1^*, \\ f[x_1^*] &= f^2[Q(\theta_*)] = x_2^*, \\ &\vdots \\ f[x_i^*] &= f^{i+1}[Q(\theta_*)] = x_{i+1}^*, \\ f[x_k^*] &= f^{k+1}[Q(\theta_*)] = P(\phi_*). \end{aligned}$$

Since any f-iterate of a local unstable manifold is again a local unstable manifold, and any f-iterate of a heteroclinic point is another heteroclinic point, the fact that $Q(\theta_*)$ is a heteroclinic

point tells us that the local unstable manifold $f^{k+1}[W^u_{\text{loc}}(q)] = f^{k+1}[Q(U_{n_u}(0))]$ intersects $W^s_{\text{loc}}(p) = P[U_{\nu_s}(0)]$ at the phase space point $Q(\theta_*) = f^{k+1}[P(\phi_*)]$. Then as above (when the local stable and unstable manifolds intersected without iteration) we have that the intersection is transverse if and only if the matrix

$$[D_\theta[f^{k+1}(Q(\theta_*))]\,|\,-D_\phi P(\phi_*)] = [D_x[f^{k+1}(Q(\theta_*))]D_\theta Q(\theta_*)|\,-D_\phi P(\phi_*)]$$

is nonsingular. Note that $D_x f^i(x)$ is nonsingular for any $x \in \mathbb{R}^n$ and $i \in \mathbb{Z}$, as f is a diffeomorphism.

By hypothesis the matrix

$$DF(x_*) =$$

$$\begin{pmatrix}
-D_\theta Q(\theta_*) & Df^{-1}(x_1^*) & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\
0 & -I & Df^{-1}(x_2^*) & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \dots & -I & Df^{-1}(x_j^*) & 0 & \dots & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \dots & 0 & Df(x_j^*) & -I & \dots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & Df(x_{k-1}^*) & -I & 0 \\
0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & Df(x_k^*) & -DP(\phi_*)
\end{pmatrix}$$

is invertible. We argue by way of the contrapositive and suppose that the intersection of $W^u(q)$ and $W^s(p)$ is not transverse at $P(\phi_*)$.

If the intersection is not transverse, then there is a nonzero vector $v \in \mathbb{R}^n$ so that

$$[-D_x[f^{k+1}(Q(\theta_*))]D_\theta Q(\theta_*)|D_\phi P(\phi_*)]v = 0$$

or

$$D_x[f^{k+1}(Q(\theta_*))]D_\theta Q(\theta_*)v_1 = D_\phi P(\phi_*)v_2, \tag{5.28}$$

with $v_1 \in \mathbb{R}^{n_u}$ and $v_2 \in \mathbb{R}^{n_s}$ not both identically zero. Note that if we repeatedly apply the chain rule to the left-hand side of (5.28), we obtain the identity

$$D_x[f^{k+1}(Q(\theta_*))]D_\theta Q(\theta_*)v_1 = Df[x_k^*]Df[x_{k-1}^*]\cdot\dots\cdot Df[x_1^*]Df[Q(\theta_*)]DQ(\theta_*)v_1. \tag{5.29}$$

Now we will use v_1 and v_2 to define a vector $h \in \mathbb{R}^{n(k+1)}$, which we will apply to $DF(x_*)$. We write

$$h = (h_1, h_2, \dots, h_{j-1}, h_j, h_{j+1}, \dots, h_{k-1}, h_k, h_{k+1}, h_{k+2}),$$

where $h_1 \in \mathbb{R}^{n_u}$, $h_{k+2} \in \mathbb{R}^{n_s}$, and $h_i \in \mathbb{R}^n$ for $1 < i < k+2$. Then we define

$$\begin{aligned}
h_1 &= v_1, \\
h_2 &= Df[Q(\theta_*)]DQ(\theta_*)h_1, \\
h_3 &= Df[x_1]h_2, \\
&\vdots \quad \vdots \\
h_j &= Df[x_{j-2}]h_{j-1}, \\
&\vdots \quad \vdots \\
h_k &= Df[x_{k-2}]h_{k-1}, \\
h_{k+1} &= Df[x_{k-1}]h_k, \\
h_{k+2} &= v_2.
\end{aligned}$$

Here note that since $n_u + n_s = n$ there are $k+2$ "columns" in the matrix $DF(x_*)$; hence the definition of $k+2$ components of h.

Now we consider $DF(x_*)h$ component by component (here by "component" we mean an n dimensional subvector). The first component is

$$\begin{aligned}
-D_\theta Q(\theta_*)h_1 + Df^{-1}(x_1^*)h_2 &= -D_\theta Q(\theta_*)h_1 + Df^{-1}(x_1^*)Df[Q(\theta_*)]DQ(\theta_*)h_1 \\
&= -D_\theta Q(\theta_*)v_1 + Df^{-1}(f(Q(\theta_*)))Df[Q(\theta_*)]DQ(\theta_*)v_1 \\
&= -D_\theta Q(\theta_*)v_1 + [Df(Q(\theta_*))]^{-1}Df[Q(\theta_*)]DQ(\theta_*)v_1 \\
&= D_\theta Q(\theta_*)(v_1 - v_1) \\
&= 0,
\end{aligned}$$

where we have used the inverse function theorem and the definition of h_2.

Similarly if $2 \le i \le j$, then the ith component of $DF(x_*)h$ is

$$\begin{aligned}
-h_{i-1} + Df^{-1}(x_{i-1}^*)h_i &= -h_{i-1} + Df^{-1}(x_{i-1}^*)Df(x_{i-2}^*)h_{i-1} \\
&= -h_{i-1} + Df^{-1}(f(x_{i-2}^*))Df(x_{i-2}^*)h_{i-1} \\
&= -h_{i-1} + [Df(x_{i-2}^*)]^{-1}Df(x_{i-2}^*)h_{i-1} \\
&= -h_{i-1} + h_{i-1} \\
&= 0,
\end{aligned}$$

again by the inverse function theorem and the definition of h_i, h_{i-1}. On the other hand, for $j < i \le k$ we have that the i-component of $DF(x_*)h$ is

$$\begin{aligned}
Df(x_{i-1}^*)h_i - h_{i+1} &= Df(x_{i-1}^*)h_i - Df(x_{i-1}^*)h_i \\
&= 0,
\end{aligned}$$

just by definition of h_{i+1}. Finally, consider the $(k+1)$th component of $DF(x_*)h$. Using the definition of each of the components of h, we see that

$$\begin{aligned}
&Df(x_k^*)h_{k+1} - DP(\phi_*)h_{k+2} \\
&\quad = Df(x_k^*)h_{k+1} - DP(\phi_*)v_2 \\
&\quad = Df(x_k^*)Df[x_{k-1}^*]h_{k-1} - DP(\phi_*)v_2 \\
&\quad \vdots \\
&\quad = Df(x_k^*)Df[x_{k-1}^*]\dots Df[x_1^*]Df[Q(\theta_*)]DQ(\theta_*)v_1 - DP(\phi_*)v_2 \\
&\quad = D_x[f^{k+1}(Q(\theta_*))]D_\theta Q(\theta_*)v_1 - DP(\phi_*)v_2 \\
&\quad = 0,
\end{aligned}$$

where we pass from the third-to-last to the second-to-last line using the identity given by (5.29), and from the second-to-last to the last line by (5.28). Since h_1 and h_2 are not both zero we now have that h is a nonzero vector in the kernel of $DF(x_*)$, a matrix which is invertible by hypothesis. Then the manifolds are in fact transverse at $P(\phi_*)$. Since f is a diffeomorphism the transversality is preserved under iteration, and we have the theorem. ■

As remarked at the beginning of this section, if $q \neq p$, then Theorem 5.2 provides a proof of the existence of a heteroclinic orbit. The case of $q = p$ and homoclinic orbits is slightly more subtle. Observe that in this case if $x_* = (0, p, \dots, p, 0)$, then $F(x_*) = 0$. To avoid this trivial solution we provide the following reformulation of Theorem 5.2 for the case of homoclinic orbits.

Corollary 5.4. *Assume the same hypotheses as Theorem 5.2, where $q = p$ with the added condition that*

$$x_0 = (0_{n_u}, p, \dots, p, 0_{n_s}) \notin B_r(\hat{x}) \subset \mathbb{R}^{n(k+1)}.$$

Then there exists a unique $x_* \in B_r(\hat{x}) \subset \mathbb{R}^{n(k+1)}$ satisfying $F(x_*) = 0$, the orbit of which is not trivial (not simply the fixed point p). The resulting orbit is transverse under the same conditions given in Corollary 5.3.

6. Computer assisted proofs of transverse homoclinic orbits and chaos. In this section we use $B(x, r)$ to denote $D_r(x) \subset \mathbb{R}^n$, as the former notation is easier to read when x and r are floating point numbers and vectors. The use of the B for a ball in this section also lets us reserve D to stand for the differential.

We discuss several computer assisted proofs of the existence of homoclinic tangles. We examine one computation in detail in order to illustrate how the various parameters are chosen, and then present the results of a number of additional computations. We note that, as in section 4.4, the MATLAB software package IntLab is used throughout in order to carry out interval arithmetic computations.

Take the Lomelí map, this time with parameters

$$a = 0.5, \quad b = -0.5, \quad c = 1, \quad \alpha = -5.34, \quad \text{and} \quad \tau = 0.8.$$

(Again we remark that these correspond to Dullin–Meiss parameters of $\bar{a} = 1$, $\bar{b} = 0.5$, $\bar{c} = 0.5$, $\mu = -2.4$, and $\epsilon = 5.5$.) For these parameter values there is a fixed point at $p = (x_-, x_-, x_-)$

with

$$x_- \in B(-2.745207879911715, 4.441 \times 10^{-16}).$$

We compute that $Df(p)$ has an unstable complex conjugate pair of eigenvalues,

$$\lambda^u_{1,2} \in B(-0.402451645443971 \pm i2.035392592347574, 1.061 \times 10^{-15}),$$

and a stable eigenvalue

$$\lambda^s \in B(0.232299350932085, 2.221 \times 10^{-16}),$$

with eigenvectors

$$\xi^u_{1,2} \in B\left(\begin{bmatrix} 0.88172919902084 \\ -0.08243220682596 \pm 0.41689953325783i \\ -0.18941206925421 \mp 0.07795125439720i \end{bmatrix}, 2.381 \times 10^{-16}\right)$$

and

$$\xi^s \in B\left(\begin{bmatrix} 0.05249091964826 \\ 0.22596240341460 \\ 0.97272076959298 \end{bmatrix}, 1.388 \times 10^{-16}\right).$$

In particular we see that the matrix $Df(p)$ is hyperbolic and diagonalizable. We want to prove the existence of a homoclinic orbit for p. In fact, we will obtain the existence of a transverse homoclinic orbit and hence the existence of chaotic motions for the Lomelí map by the Smale tangle theorem [48].

We begin by computing and validating parameterizations P_{N_s} and Q_{N_u} for the stable and unstable manifolds, as in section 4.4. The results are tabulated in Table 3. Note that we compute the stable manifold to order $N_s = 50$ and the unstable manifold to order $N_u = 25$ and that we obtain truncation errors of less than 6×10^{-12} for both manifolds.

Table 3

Manifold validation performance: ($\epsilon = 5.5, \mu = -2.4$). The last line of the table is a bound on the size of the image of the parameterizations in phase space.

Dim	Order	Coeff time	Validation time	δ	ν	$\lvert\xi\rvert$	$\lVert\cdot\rVert_{\nu,\Sigma}$
1	50	0.89 sec	0.65 sec	7.92×10^{-13}	0.9	2	1.96
2	25	5.64 sec	2.30 sec	3.54×10^{-12}	0.4	1.5	1.21

A numerical Newton scheme is used to find an approximate numerical solution to the discretized homoclinic functional equation $F_N(x) = 0$ with $k = 6$ and $n = 3$. Recall that this means that the homoclinic orbit requires six iterates to make the transition from the local unstable to the local stable manifold, and that the phase space dimension is three. Then F_N is a map from $\mathbb{R}^{18}$ into itself.

As an aside we remark that the problem of locating initial guesses/approximate zero from which to begin the Newton iteration is classical and is discussed quite generally in [8] and in the literature review therein. Locating approximate solutions of the homoclinic operator

equation for the Lomelí map is discussed in [36]. We use the methods discussed in the latter reference to find the approximate zero,

$$\hat{x} = \begin{bmatrix} \hat{\theta} \\ \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \hat{x}_4 \\ \hat{x}_5 \\ \phi \end{bmatrix} = \begin{bmatrix} (-0.337379322019076, 0.088431234641040) \\ (-1.648314148155201, -3.605864990373435, -2.750773367689280) \\ (1.979508268106647, -1.648314148155201, -3.605864990373435) \\ (-1.054666610773029, 1.979508268106647, -1.648314148155201) \\ (-2.313572985270695, -1.054666610773029, 1.979508268106647) \\ (-2.642742570718999, -2.313572985270695, -1.054666610773029) \\ 0.228218016117584 \end{bmatrix}.$$

(These are taken as intervals with radius zero.) We use Theorem 5.2 to prove that there is a true solution of the homoclinic functional equation nearby.

We begin by noting that $0 < |\hat{\phi}| < 0.22822 < 0.9 = \nu_s$, so that $\hat{\phi}$ is well inside the validated domain of P_{N_u} but is also isolated well away from 0 in parameter space. Similarly $|\hat{\theta}| = \sqrt{\theta_1^2 + \theta_2^2} < 0.34878 < 0.4$, so that $\hat{\theta}$ is also well inside the validated domain of Q_{N_s} and also isolated away from zero. These conditions are essential so that there is some portion of the parameterization domain that we can "give up" or "trade in" in order to bound derivatives of the truncation errors h_s and h_u. It is also important that we are not validating the trivial solution (see Corollary 5.4), which is why we note that the parameters are bounded away from zero.

We define the "loss of domain" parameters

$$\hat{\sigma}_u = \min\left(-\ln\left(\frac{|\hat{\theta}|}{\nu_u}\right), 1\right) \in B(0.13703387555508, 1.111 \times 10^{-16})$$

and

$$\hat{\sigma}_s = \min\left(-\ln\left(\frac{|\hat{\phi}|}{\nu_s}\right), 1\right) = 1,$$

as per Definition 5.1 (CVV), and compute $F_N(\hat{x})$, $DF_N(\hat{x})$, and $DF_N(\hat{x})^{-1}$. We find that

$$\|DF_N(\hat{x})^{-1}\|_M < A_N \equiv 6.061,$$

and note that while $\hat{x}$ is taken as an interval with zero radius, the evaluations of F_N and DF_N result in interval output.

In particular we must invert and bound an interval matrix. However, since the inputs $\hat{x}$ are points, the interval entries of $DF_N(\hat{x})$ typically have small radii. In this example the

entries of $DF_N(\hat{x})$ have radii of less than 8×10^{-14}, and inverting produces a matrix with interval entries enclosing the matrix $DF_N(\hat{x})^{-1}$. In this example the entries of this inverted matrix have radii of less than 3×10^{-14}.

Next we compute that

$$|DF_N(\hat{x})^{-1} F_N(\hat{x})| < \hat{\epsilon} \equiv 9.088 \times 10^{-14}.$$

Here we actually compute the matrix vector product and then bound the resulting interval vector, as in general this will produce a better bound than $A_N|F_N(\hat{x})|$ even if the latter is cheaper to compute once A_N is known.

From the computation of $DF_N(\hat{x})^{-1}$ we also obtain the constants

$$A_u = 0.845 \quad \text{and} \quad A_s = 0.482,$$

as required by Definition 5.1. Combining these with the information from the validated parameterization computations, we have that

$$\hat{\delta} \equiv 3.3 \times 10^{-12} \qquad \text{and} \qquad M_N \equiv 2.02 \times 10^{-9} < 1.$$

We are ready to compute

$$\frac{\hat{\epsilon} + \hat{\delta}}{1 - M_N} = \epsilon_{NK} < 3.39 \times 10^{-12}$$

and define

$$2\epsilon_{NK} < 6.78 \times 10^{-12} \equiv r.$$

Now consider $B = B_r(\hat{x})$, which is an interval vector about our approximate solution. We have to bound the second derivative of F (not F_N) on B. So we check that $|\hat{\theta}| + r < 0.3488 < 0.4 = \nu_u$ and $|\hat{\phi}| + r < 0.228 < 0.8 = \nu_s$, which lets us define the loss of domain parameters σ_s and σ_u as needed in Theorem 5.2.

We observe that while the derivatives of the parameterizations need to be bounded only locally, we have no local information about the truncation errors. So we apply the Cauchy bounds of Lemma 2.9 and obtain global bounds on the truncations. In this example we obtain that the second derivatives of the truncation errors satisfy the bounds

$$\|D^2 h_u\|_{B(0,|\hat{\theta}|+r)} < 1.64 \times 10^{-6} \quad \text{and} \quad \|D^2 h_s\|_{B(0,|\hat{\phi}|+r)} < 4 \times 10^{-10}.$$

Now we use interval arithmetic to bound $D^2 P_{N_s}$ and $D^2 Q_{N_u}$ on $U_r(\hat{\phi})$ and $U_r(\hat{\theta})$, respectively. This is a purely numerical computation. In fact, since r is small, this is just polynomial evaluation with interval inputs, which IntLab has built-in features to handle efficiently.

Combining all of this, we numerically bound the expression given by (5.9) in Definition 5.1 and find that it is sufficient to take

$$C_P \equiv 2.230.$$

Similarly we evaluate C_β on B, but since the map is quadratic we just have that

$$C_\beta = \max(2|a|, 2|c|, |b|) = 2,$$

Table 4

Primary intersection validation ($\epsilon = 5.5, \mu = -2.4$): 1.89 (sec) for proof of both orbits. This does not count the time taken to compute the manifolds. Chaos confirmed in both cases.

K	$\hat{x}_1$	r
6	$(-1.648314148155201, -3.605864990373435, -2.750773367689280)$	6.92×10^{-12}
6	$(-1.692334813290302, -3.652591337627915, -2.718741184627647)$	7.12×10^{-12}

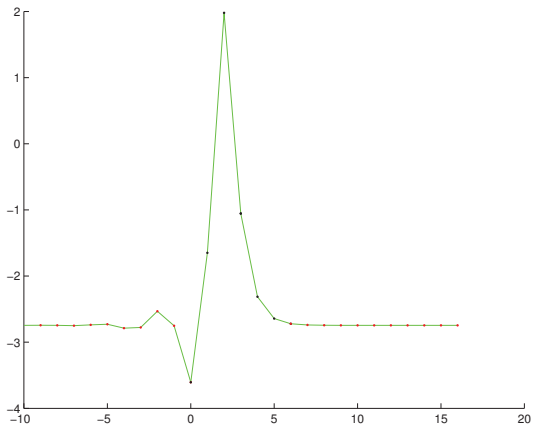

Figure 1. *x-axis projection of the validated homoclinic; $k = 6$, $\epsilon = 5.5$.*

independent of B, and similarly $\bar{N}_f = 4$.

Now we are ready to compute

$$\kappa = 10.3 \geq \bar{N}_f C_\beta + C_P$$

and check that

$$4\kappa A_N \frac{\hat{\epsilon} + \hat{\delta}}{(1 - M_N)^2} < 8.41 \times 10^{-10} < 1. \tag{6.1}$$

From this we conclude that there exists a unique x_* in $B_r(\hat{x})$ which is a root of F; i.e., there is a homoclinic orbit of p through $B_r(\hat{x})$. Moreover, since the left-hand side of (6.1) is strictly less than 1 and $r = 2\epsilon_{\mathrm{NK}} < 4\epsilon_{\mathrm{NK}}$, we have the transversality of the intersection and incidentally the existence of a homoclinic tangle/chaotic motions by Smale's theorem [48].

Table 4 summarizes the results of the computation/proof just discussed and also gives results for a similar proof of the existence of a second distinct solution of the homoclinic operator equation for $k = 6$. In these cases only the $\hat{x}_1$ data is recorded. Figure 1 shows the time series data for the x-component of the first of these two orbits. Black dots represent points in $\hat{x}$. Red points represent iterates on the local manifolds. The computation time is given in both cases as well.

We note that in both of the proofs discussed above, the time taken to compute both the rigorous interval enclosures of the coefficients as well as the validated bounds for both

Table 5

Secondary homoclinic orbits ($\epsilon = 5.5, \mu = -2.4$): Total time for proofs: 36 sec. Transversality confirmed in all cases.

K	$\hat{x}_1$	r	Time
8	$(-1.878269557294666 - 3.704360821688669 - 2.644177124855255)$	6.96×10^{-12}	2.43 sec
·	$(-1.598486534326447 - 3.712394711133192 - 2.715338895232408)$	7.15×10^{-12}	·
9	$(-1.693365888596068 - 3.516449414154529 - 2.776271298390562)$	6.98×10^{-12}	2.86 sec
·	$(-2.033965491062911 - 3.691036738831221 - 2.607784382423848)$	7.00×10^{-12}	·
·	$(-3.649752275192224 - 2.876479215542708 - 2.487231377447373)$	6.68×10^{-12}	·
11	$(-1.724921906236488 - 3.503098391735548 - 2.773685700840596)$	7.17×10^{-12}	5.79 sec
·	$(-2.089900084565888 - 3.686144425839955 - 2.594568562106802)$	7.09×10^{-12}	·
·	$(-3.634873256589227 - 2.906134859387798 - 2.482573305537549)$	6.78×10^{-12}	·
·	$(-3.620917995724487 - 2.915222901577827 - 2.483676082433866)$	6.63×10^{-12}	·
·	$(-2.114585182128023 - 3.679401143907701 - 2.591096188024756)$	6.94×10^{-12}	·
·	$(-1.768297176557754 - 3.496683655844906 - 2.765421447288818)$	7.05×10^{-12}	·
12	$(-1.613946132963925 - 3.601054205346514 - 2.761528716808955)$	7.01×10^{-12}	4.2 sec
·	$(-1.672093712060165 - 3.527103879468739 - 2.777334962197874)$	7.12×10^{-12}	·
·	$(-2.122097145983667 - 3.674130503709248 - 2.591708802528494)$	6.98×10^{-12}	·
·	$(-1.822510500455057 - 3.571768208555173 - 2.720873332899369)$	7.06×10^{-12}	·
13	$(-3.644121861531430 - 2.872709464400592 - 2.489856336907984)$	6.64×10^{-12}	6.11 sec
·	$(-1.720320939862523 - 3.656391590596805 - 2.709687450800172)$	7.05×10^{-12}	·
·	$(-1.972320520664557 - 3.693712699582179 - 2.623618282117915)$	7.10×10^{-12}	·
·	$(-3.647170226591191 - 2.867372482172479 - 2.490553201321108)$	6.76×10^{-12}	·
·	$(-1.582489566947040 - 3.527839146851471 - 2.799122415227350)$	7.26×10^{-12}	·
·	$(-1.574224069064366 - 3.529792848481951 - 2.800384605320380)$	7.10×10^{-12}	·
20	$(-1.931148725862011 - 3.707646666557216 - 2.627909579872393)$	6.90×10^{-12}	2.1 sec
·	$(-3.638326627639060 - 2.901176380906034 - 2.483107270172258)$	6.61×10^{-12}	
21	$(-3.690719490424216 - 2.823690393936636 - 2.490880594199791)$	6.64×10^{-12}	10.35 sec
·	$(-1.957194765763665 - 3.705297800511473 - 2.621878341459779)$	6.98×10^{-12}	·
·	$(-1.729640666364290 - 3.510087223951199 - 2.769773329564600)$	7.05×10^{-12}	·
·	$(-1.690639165363386 - 3.669844178437995 - 2.711227287037635)$	7.06×10^{-12}	·
·	$(-1.950380561442004 - 3.705777860019821 - 2.623527619587280)$	7.02×10^{-12}	·
·	$(-3.702924845715120 - 2.791265552326865 - 2.496202529149488)$	6.67×10^{-12}	·
·	$(-3.708117158393551 - 2.774277602263187 - 2.499173703371658)$	6.50×10^{-12}	·
·	$(-1.932029291989042 - 3.707691973193318 - 2.627641932935536)$	6.88×10^{-12}	·
·	$(-3.616786394029812 - 2.918973341522530 - 2.483676503055390)$	6.55×10^{-12}	·

manifolds is 9.22 seconds, while the validation of the two homoclinic orbits takes only 1.89 seconds. Since we can use the same polynomial approximations P_{N_s} and Q_{N_u} in any homoclinic functional equation, regardless of the size of k, we compute 32 more distinct homoclinic orbits with k varying. The results are tabulated in Table 5, and again only $\hat{x}_1$-components are recorded. Note that the time required to validate all 34 orbits is 35.8 seconds, or roughly three times the amount of time needed to rigorously compute and validate the manifolds. This suggests that high order approximation of the manifolds is especially useful when computing many distinct homoclinic orbits at a given parameter set. Figures 1 and 2 show time series data for the x-component of the shortest and longest homoclinic orbits validated. The reader interested in the details of these computations can find the implementation in the file validatedChaos_proofsOne.m [38].

We note that in the previous example the dynamics is "fast" in the sense that as few as 6 iterates are needed in order to transition from the local unstable to the local stable manifold.

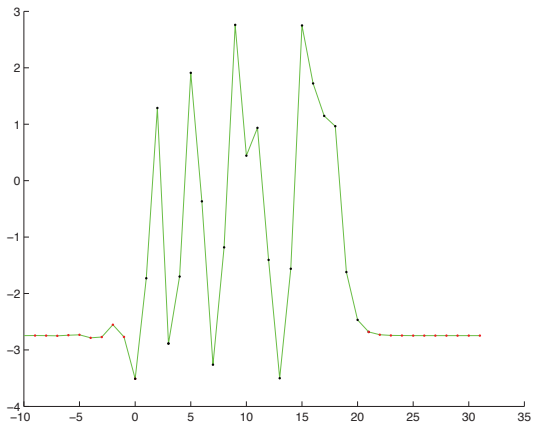

Figure 2. *x-axis projection of the validated homoclinic; $k = 21$, $\epsilon = 5.5$.*

Table 6
Manifold validation performance: ($\epsilon = 0.2, \mu = -2.4$).

K	$\hat{x}_1$	r	Time
75	$(-0.717248519714197 - 1.043252947479510 - 0.860812112677259)$	1.06×10^{-7}	3.01 sec
76	$(-1.107394504655081 - 0.745731963636135 - 0.642025567084575)$	1.43×10^{-7}	2.16 sec
111	$(-1.104148108665029 - 0.729631044649217 - 0.648872760710501)$	1.08×10^{-7}	8.62 sec
121	$(-0.995810895350469 - 0.972045779061998 - 0.671276957464922)$	1.06×10^{-7}	11.2 sec

Table 7
Homoclinic orbits ($\epsilon = 0.2, \mu = -2.4$): Transversality confirmed in all cases.

Dim	Order	Approx time	Proof time	Validated error	Radius	$\lvert\xi\rvert$	$\lVert \cdot \rVert_\nu$
1	50	0.69 sec	0.51 sec	2.76×10^{-11}	0.9	1.5	5.63
2	25	5.6 sec	2.49 sec	3.60×10^{-13}	0.4	0.5	0.32

In order to compute orbits with longer "time of flight" (higher k) we consider a Lomelí map with parameters a, b, c, and τ as before, but with $\alpha = -0.04$ (corresponding to a Dullin–Meiss value of $\epsilon = 0.2$). We do not record the fixed point, eigenvalue, and eigenvector data for this example. We just point out that since ϵ is smaller, the dynamics are "slower" in the sense that the eigenvalues are closer to the unit circle. Also, the one dimensional stable and unstable manifolds of the two distinct fixed points pass close to one another, so that orbits homoclinic to p get "captured" in a neighborhood of the other fixed point for some amount of time. (See Figures 4.6 and 4.7 in [36]. The idea is that as $\epsilon \to 0$, the system converges to a singular integrable limit. Again we refer the reader to [18, 36] for a fuller discussion.) The point at present is that we can find orbits which spend a "longer and longer time" in transition from one manifold to another by making the value of the Dullin–Meiss ϵ parameter small.

The data for the new parameterization computations is given in Table 6, with format identical to before. Table 7 gives data for the results of the homoclinic validation computations for four different orbits with values of k varying between 75 and 121. Figures 3 and 4 show

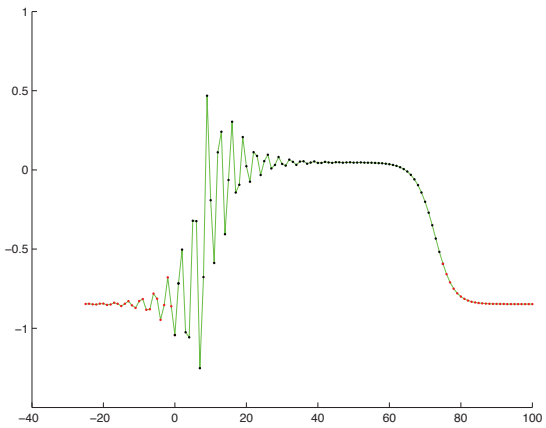

Figure 3. *x-axis projection of the validated homoclinic;* $k = 75$, $\epsilon = 0.2$.

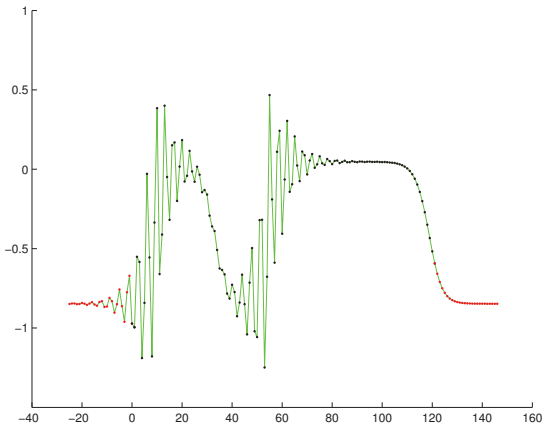

Figure 4. *x-axis projection of the validated homoclinic;* $k = 121$, $\epsilon = 0.2$.

time series data for the shortest and longest of these homoclinic orbits (x-component in both cases). We make a few additional remarks about the breakdown of the proof. These computer assisted proofs are implemented in the file validatedChaos_proofsTwo.m, available from [38].

For the successful case of $k = 121$ we see that the norm of the inverse matrix has the bound $\|DF_N(\hat{x})^{-1}\| < A_N \equiv 3.9 \times 10^5$, which is quite large. Then while we obtain the reasonable bounds $\hat{\epsilon} = 1 \times 10^{-10}$, $\hat{\delta} = 5.27 \times 10^{-8}$, $r = 1.08 \times 10^{-7}$, $M_N = 3.433 \times 10^{-6}$, and $\kappa = 8.97$, the large value of A_M gives

$$4\kappa A_N \frac{\hat{\epsilon} + \hat{\delta}}{(1 - M_N)^2} < 0.74 < 1,$$

only a little less than 1. While the computation goes through, the value is only marginally smaller than 1 (i.e., is the same order of magnitude as 1).

When we repeat the computation for $k = 122$ the proof fails. In this case we are able to find an approximate orbit which still has $\hat{\epsilon} < 1 \times 10^{-9}$. However, we have $A_N = 3.3 \times 10^5$, $M_N = 1.1 \times 10^{-5}$, $\hat{\delta} < 2.5 \times 10^{-7}$, and $\kappa = 8.7$, so that

$$4\kappa A_N \frac{\hat{\epsilon} + \hat{\delta}}{(1 - M_N)^2} > 2.78,$$

and the hypotheses of the Newton–Kantorovich theorem are not satisfied. Note that in the $k = 122$ computation the value of A_N is actually a little better then in the $k = 121$ cases. However, we lose an order of magnitude in both the values of $\hat{\delta}$ and M_N, and these push the bound above over 1.

We have made no real effort to fine tune the computation and push the proof further (mainly because these computations are intended to give a broad picture of the use of the method rather than exhaustively explore the dynamics of the Lomelí map). It is likely that by adjusting the parameterizations we could decrease $\hat{\delta}$ enough to improve the bounds. More importantly, we are using "off the shelf" codes to obtain the inverse of $DF_N(\hat{x})$ and the bound A_N. However, the matrix itself has considerable structure which could be exploited. It is likely that taking advantage of the sparse-banded structure of the differential would lead to better numerics. Finally, it is often the case that some gain (as much as an additional order of magnitude) can be obtained in interval arithmetic computations by switching to a C^{++} implementation. The example computations of the present section are presented in the spirit of a proof of concept and do not necessarily represent definitive limitations of the method.

The last two computations also show the connection between the time of flight k and the size of the interval matrices in a particular proof. This highlights another reason to prefer high order approximations to the invariant manifold, namely, that it prevents us from studying orbit segments which begin and end near the fixed point. In a sense we are "saving up" our k, spending it only on orbit segments which truly exhibit global dynamics. This is helpful when we are trying to follow "long excursions."

As a final example we consider a computer assisted proof involving higher dimensional manifolds. We begin by coupling a pair of three dimensional Lomelí maps,

$$\begin{array}{rcl} f_1(x_1, y_1, z_1) & \equiv & f_{\alpha_1,\tau_1,a_1,b_1,c_1}(x_1, y_1, z_1) \\ & \text{and} & \\ f_2(x_2, y_2, z_2) & \equiv & f_{\alpha_2,\tau_2,a_2,b_2,c_2}(x_2, y_2, z_2), \end{array}$$

in order to obtain the six dimensional dynamical system $G : \mathbb{R}^6 \to \mathbb{R}^6$ given by

$$G(x_1, y_1, z_1, x_2, y_2, z_2) \equiv \begin{bmatrix} f_1(x_1, y_1, z_1) + \varepsilon g_2(y_2,) \\ f_2(x_2, y_2, z_2) + \varepsilon g_1(y_1) \end{bmatrix}, \tag{6.2}$$

where

$$g_1(y_1) \equiv (y_1 - x_1^+)(y_1 - x_1^-) \quad \text{and} \quad g_2(y_2) \equiv (y_2 - x_2^+)(y_2 - x_2^-).$$

Here $x_{1,2}^{\pm}$ denotes the first coordinate of each of the two fixed points in the $f_{1,2}$ systems. (Since the fixed points of the uncoupled systems are on the $x = y = z$ line, it is enough to specify

only the x-coordinate of each of the uncoupled fixed points. See [35].) Note that this coupling does not move the fixed points in the $f_{1,2}$ systems but does perturb the eigenspaces. In the discussion that follows we will work with only floating point approximations for the sake of brevity. However, we stress that in the actual computations interval arithmetic is used.

We take parameters $a_1 = a_2 = 0.5$, $b_1 = b_2 = -0.5$, $c_1 = c_2 = 1$, $\tau_1 = \tau_2 = 0.8$, $\alpha_1 = -5.339$, and $\alpha_2 = -5.939$ (corresponding to Dullin–Meiss parameters of $\epsilon_1 = 5.5$ and $\epsilon_2 = 6.1$). The maps are coupled with a strength of $\varepsilon = 5 \times 10^{-7}$.

The reason for the small coupling strength is that in order to obtain a numerical starting point for the Newton iteration scheme we continue away from the product system having $\varepsilon = 0$. Since we allow the coupling function to perturb the eigenspaces of the individual systems, the coupled system is quite sensitive to this parameter, and a tangency develops quickly as ε increases. However, our proof exploits the small parameters only in the sense that it is helpful for locating an initial guess for a homoclinic in the coupled system. We have made no attempt at an exhaustive study of the six dimensional system. The coupled system only serves to illustrate that this kind of computer assisted proof can be made to function in higher dimensions.

The coupled map has a fixed point at $p = (x_1, x_1, x_1, x_2, x_2, x_2)$ with $x_1 \approx -2.7452078799$ and $x_2 \approx -2.8698178070$. We find that the differential $DG(p)$ has two pairs of unstable complex conjugate eigenvalues, $\lambda^u_{1,2} \approx -0.429 \pm i2.076$ and $\lambda^u_{3,4} \approx -0.402 \pm i2.035$, and also a pair of real distinct stable eigenvalues, $\lambda^s_1 \approx 0.232$ and $\lambda^s_2 = 0.222$. Then the fixed point has a four dimensional unstable manifold and a two dimensional stable manifold. We show that these manifolds intersect transversally, using the arguments developed above.

We begin then by considering the two dimensional stable manifold at p. Note that for this map, and at this fixed point, we have $n = 6$, $N_f = 10$, $\mu^*_s \approx 0.232$, and

$$K_1 = \sup(2|a_1|, 2|c_1|, |b_1|, 2|a_2|, 2|c_2|, b_2) + 2\varepsilon$$

(where we recall that ε is the coupling parameter). At the fixed point we find that $\|[DG]^{-1}(p)\|_M \approx 4.57$. Then we compute the a priori indicators given by (4.9) and (4.10) and see that

$$\frac{\ln(\|[DG]^{-1}(p)\|)}{|\ln(\mu^*_s)|} < 1.9 \quad \text{and} \quad \frac{1}{4n\pi e N_f K_1 \|[DG]^{-1}(p)\|^2} \approx 5 \times 10^{-7}.$$

The first number suggests that we will be able to validate the approximate parameterization of the stable manifold for as long as $N_s+1 \geq 2$. Note that, while the second number suggests that the residual error will need to be small, 10^{-7} is still a long way from machine epsilon. Then the validated manifold computation should be reasonable, and we are confident in proceeding. While it seems that we are free to parameterize the manifolds to any order, a little numerical experimentation shows that taking $N_s = 9$ gives a good trade-off between computational speed and numerical accuracy, while still giving a parameterization domain whose image moves us away from the fixed point. Table 8 summarizes the results.

For the unstable manifold we have the same values of n, K_1, and N_f as before. However, this time we must interpret the a priori indicators in terms of stable manifolds of G^{-1}. Then

$$\mu^*_u = \frac{1}{|\lambda^u_{3,4}|} \approx 0.482$$

Table 8

Manifold validation performance: Six dimensional example.

Dim	Order	Approx time	Proof time	Validated error	Radius	$\|\xi\|$	$\|\cdot\|_\nu$
2	9	1.09 sec	0.97 sec	4.66×10^{-13}	0.001	1.4	1.46×10^{-3}
4	5	2.02 sec	1.75 sec	6.59×10^{-9}	0.001	0.4	7.1×10^{-4}

Table 9

Primary intersection validation: Six dimensional example. Chaos confirmed. The similarity in the coordinates of x_1 is due to the fact that the orbit begins close to the fixed point.

K	$\hat{x}_1$	r	Time
20	$\begin{pmatrix} -2.743916182731272 \\ -2.745285611304841 \\ -2.745493386071762 \\ -2.868750115142552 \\ -2.869921042268439 \\ -2.870035612699910 \end{pmatrix}$	3.07×10^{-8}	2.97 sec

and

$$\|[DG^{-1}]^{-1}(p)\| = \|DG(p)\| \approx 5.94.$$

In this case computing the quantities given by (4.9) and (4.10) leads to

$$\frac{\ln(\|[DG](p)\|)}{|\ln(\mu_u^*)|} \approx 4 \quad \text{and} \quad \frac{1}{4n\pi e N_f K_1 \|[DG]^{-1}(p)\|^2} \approx 5 \times 10^{-6},$$

and the suggestion is that the computation requires a parameterization order of at least $N_u + 1 \geq 5$, or $N_u = 4$. We actually find that good numerical/timing results are obtained with $N_u = 5$, and we can in fact validate the computations at this order. The validated computations summarized in Table 8 also give performance data for the unstable manifold computation.

With these manifolds we are able to find an approximate solution of the homoclinic operator equation with $k = 20$. The validation of the connecting orbit proceeds as before. The results are summarized in Table 9. We note that since we also validate the transversality of the orbit, the computation actually gives the existence of chaotic motions in the six dimensional dynamical system by Smale's tangle theorem [48].

Acknowledgments. The authors would like to thank two anonymous referees for carefully reading the manuscript and making several very helpful comments and corrections; the final version of the paper was greatly improved by their suggestions. We also thank Rafael de la Llave, Jean-Philippe Lessard, Jan Bouwe van den Berg, Maciej Capinski, and Christian Reinhardt for helpful discussions which improved the present work.

REFERENCES

[1] L. V. Ahlfors, *Complex Analysis. An Introduction to the Theory of Analytic Functions of One Complex Variable*, 3rd ed., Int. Ser. Pure Appl. Math., McGraw–Hill, New York, 1978.

[2] G. Arioli, *Periodic orbits, symbolic dynamics and topological entropy for the restricted 3-body problem*, Comm. Math. Phys., 231 (2002), pp. 1–24.

[3] G. Arioli and P. Zgliczyński, *Symbolic dynamics for the Hénon–Heiles Hamiltonian on the critical level*, J. Differential Equations, 171 (2001), pp. 173–202.
[4] G. Arioli and P. Zgliczyński, *The Hénon–Helies Hamiltonian near the critical energy level—some rigorous results*, Nonlinearity, 16 (2003), pp. 1833–1852.
[5] K. Atkinson, *An Introduction to Numerical Analysis*, Wiley, New York, 1987.
[6] I. Baldomá, E. Fontich, and R. de la Llave, *The parameterization method for one-dimensional invariant manifolds of higher dimensional parabolic fixed points*, Discrete Contin. Dynam. Systems, 17 (2007), pp. 835–865.
[7] J. B. van den Berg, J. D. Mireles-James, J.-P. Lessard, and K. Mischaikow, *Rigorous numerics for symmetric connecting orbits: Even homoclinics of the Gray–Scott equation*, SIAM J. Math. Anal., 43 (2011), pp. 1557–1594.
[8] W.-J. Beyn and J.-M. Kleinkauf, *The numerical computation of homoclinic orbits for maps*, SIAM J. Numer. Anal., 34 (1997), pp. 1207–1236.
[9] W. Beyn and J. Kleinkauf, *Numerical approximation of homoclinic chaos*, Numer. Algorithms, 14 (1997), pp. 25–53.
[10] X. Cabré, E. Fontich, and R. de la Llave, *The parameterization method for invariant manifolds.* I. *Manifolds associated to non-resonant subspaces*, Indiana Univ. Math. J., 52 (2003), pp. 283–328.
[11] X. Cabré, E. Fontich, and R. de la Llave, *The parameterization method for invariant manifolds.* II. *Regularity with respect to parameters*, Indiana Univ. Math. J., 52 (2003), pp. 329–360.
[12] X. Cabré, E. Fontich, and R. de la Llave, *The parameterization method for invariant manifolds.* III. *Overview and applications*, J. Differential Equations, 218 (2005), pp. 444–515.
[13] R. Calleja and R. de la Llave, *Fast numerical computation of quasi-periodic equilibrium states in* $1D$ *statistical mechanics, including twist maps*, Nonlinearity, 22 (2009), pp. 1311–1336.
[14] R. Calleja and R. de la Llave, *A numerically accessible criterion for the breakdown of quasi-periodic solutions and its rigorous justification*, Nonlinearity, 23 (2010), pp. 2029–2058.
[15] M. Capinski, *Covering relations and the existence of topologically normally hyperbolic invariant sets*, Discrete Contin. Dynam. Systems, 23 (2009), pp. 705–725.
[16] M. J. Capiński and P. Roldán, *Existence of a center manifold in a practical domain around* L_1 *in the restricted three-body problem*, SIAM J. Appl. Dyn. Syst., 11 (2012), pp. 285–318.
[17] B. Coomes, H. Koçak, and K. Palmer, *Homoclinic shadowing*, J. Dynam. Differential Equations, 17 (2005), pp. 175–215.
[18] H. R. Dullin and J. D. Meiss, *Quadratic volume-preserving maps: Invariant circles and bifurcations*, SIAM J. Appl. Dyn. Syst., 8 (2009), pp. 76–128.
[19] J. P. England, B. Krauskopf, and H. M. Osinga, *Computing one-dimensional stable manifolds and stable sets of planar maps without the inverse*, SIAM J. Appl. Dyn. Syst., 3 (2004), pp. 161–190.
[20] J. P. England, B. Krauskopf, and H. M. Osinga, *Bifurcations of stable sets in noninvertible planar maps*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 15 (2005), pp. 891–904.
[21] E. Fontich, R. de la Llave, and Y. Sire, *A method for the study of whiskered quasi-periodic and almost-periodic solutions in finite and infinite dimensional Hamiltonian systems*, Electron. Res. Announc. Math. Sci., 16 (2009), pp. 9–22.
[22] Z. Galias and P. Zgliczyński, *Abundance of homoclinic and heteroclinic orbits and rigorous bounds for the topological entropy for the Hénon map*, Nonlinearity, 14 (2001), pp. 909–932.
[23] M. Gidea and P. Zgliczyński, *Covering relations for multidimensional dynamical systems*, J. Differential Equations, 202 (2004), pp. 59–80.
[24] M. Gidea and P. Zgliczyński, *Covering relations for multidimensional dynamical systems.* II, J. Differential Equations, 202 (2004), pp. 59–80.
[25] A. Guillamon and G. Huguet, *A computational and geometric approach to phase resetting curves and surfaces*, SIAM J. Appl. Dyn. Syst., 8 (2009), pp. 1005–1042.
[26] Á. Haro and R. de la Llave, *A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: Numerical algorithms*, Discrete Contin. Dyn. Syst. Ser. B, 6 (2006), pp. 1261–1300.
[27] Á. Haro and R. de la Llave, *A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: Rigorous results*, J. Differential Equations, 288 (2006), pp. 530–579.
[28] W. M. Hirsch and C. C. Pugh, *Stable manifolds and hyperbolic sets*, in Global Analysis (Berkeley 1960), Proc. Sympos. Pure Math. 14, AMS, Providence, RI, 1970, pp. 133–163.

[29] T. Johnson and W. Tucker, *A note on the convergence of parametrised non-resonant invariant manifolds*, Qual. Theory Dyn. Syst., 10 (2011), pp. 107–121.
[30] B. Krauskopf, H. M. Osinga, E. J. Doedel, M. E. Henderson, J. Guckenheimer, A. Vladimirsky, M. Dellnitz, and O. Junger, *A survey of methods for computing (un)stable manifolds of vector fields*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 15 (2005), pp. 763–791.
[31] R. de la Llave, A. González, A. Jorba, and J. Villanueva, *KAM theory without action-angle variables*, Nonlinearity, 18 (2005), pp. 855–895.
[32] R. de la Llave, *A tutorial on KAM theory*, in Smooth Ergodic Theory and Its Applications (Seattle, WA, 1999), Proc. Sympos. Pure Math. 69, AMS, Providence, RI, 2001, pp. 175–292.
[33] R. de la Llave and J. D. Mireles James, *Parameterization of invariant manifolds by reducibility for volume preserving and symplectic maps*, Discrete Contin. Dynam. Systems, 32 (2012), pp. 4321–4360.
[34] J. P. Lessard, J. D. Mireles James, and C. Reinhardt, *Computer assisted proof of transverse saddle-to-saddle connecting orbits for first order vector fields*, J. Dynam. Differential Equations, submitted; preprint and code available online at http://www.math.rutgers.edu/∼jmireles/saddleToSaddle.html.
[35] H. Lomelí and J. Meiss, *Quadratic volume-preserving maps*, Nonlinearity, 11 (1998), pp. 557–574.
[36] J. D. Mireles James, *Quadratic volume-preserving maps: (Un)stable manifolds, hyperbolic dynamics, and vortex bubble bifurcations*, J. Nonlinear Sci., to appear; DOI 10.1007/s00332-012-9162-1.
[37] J. D. Mireles James and H. Lomelí, *Computation of heteroclinic arcs with application to the volume preserving Hénon family*, SIAM J. Appl. Dyn. Syst., 9 (2010), pp. 919–953.
[38] J. D. Mireles James and K. Mischaikow, `cap_ImCoLab`: *Computer Assisted Proof for Invariant Manifolds and Connecting Orbits Library in MatLab/IntLab* (2012), software, online at http://www.math.rutgers.edu/∼jmireles/tanglePaperPage.html.
[39] A. Neumaier and T. Rage, *Rigorous chaos verification in discrete dynamical systems*, Phys. D, 67 (1994), pp. 327–346.
[40] S. Newhouse, M. Berz, J. Grote, and K. Makino, *On the estimation of topological entropy on surfaces*, in Geometric and Probabilistic Structures in Dynamics, Contemp. Math. 469, AMS, Providence, RI, 2008, pp. 243–270.
[41] J. M. Ortega, *The Newton–Kantorovich theorem*, Amer. Math. Monthly, 75 (1968), pp. 658–660.
[42] K. J. Palmer, *Exponential dichotomies, The shadowing lemma and transversal homoclinic points*, in Dynamics Reporten, Vol. 1, Dynam. Report. Ser. Dynam. Systems Appl. 1, Wiley, Chichester, UK, 1988, pp. 265–306.
[43] W. Rudin, *Functional Analysis*, 2nd ed., Internat. Ser. Pure Appl. Math., McGraw–Hill, New York, 1991.
[44] S. Rump, *Verification methods: Rigorous results using floating-point arithmetic*, Acta Numer., 19 (2010), pp. 287–449.
[45] S. Rump, *INTLAB—INTerval LABoratory*, in Developments in Reliable Computing, Tibor Csendes, ed., Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999, pp. 77–104; see also http://www.ti3.tuhh.de/rump/.
[46] E. Sander, *Hyperbolic sets for noninvertible maps and relations*, Discrete Contin. Dynam. Systems, 5 (1999), pp. 339–358.
[47] E. Sander, *Homoclinic tangles for noninvertible maps*, Nonlinear Anal., 41 (2000), pp. 259–276.
[48] S. Smale, *Diffeomorphisms with many periodic points*, in Differential and Combinatorial Topology (A Symposium in Honor of Marston Morse), Princeton University Press, Princeton, NJ, 1965, pp. 63–80.
[49] J. C. Sprott, *High-dimensional dynamics in the delayed Hénon map*, Electronic J. Theoret. Phys., 3 (2006), pp. 19–35.
[50] D. Stoffer and K. Palmer, *Rigorous verification of chaotic behavior of maps using validated shadowing*, Nonlinearity, 12 (1999), pp. 1683–1698.
[51] D. Wilczak, *Abundance of heteroclinic and homoclinic orbits for the hyperchaotic Rössler system*, Discrete Contin. Dyn. Syst. Ser. B, 11 (2009), pp. 1039–1055.
[52] A. Wittig, *Rigorous High-Precision Enclosures of Fixed Points and Their Invariant Manifolds*, Ph.D. Dissertation, Department of Mathematics, Michigan State University, Lansing, MI, 2011.
[53] A. Wittig, M. Berz, J. Grote, K. Makino, and S. Newhouse, *Rigorous and accurate enclosure of invariant manifolds on surfaces*, Regul. Chaotic Dyn., 15 (2010), pp. 107–126.
[54] P. Zgliczyński, *Covering relations, cone conditions and the stable manifold theorem*, J. Differential Equations, 246 (2009), pp. 1774–1819.