COMPUTER ASSISTED ERROR BOUNDS FOR LINEAR APPROXIMATION OF (UN)STABLE MANIFOLDS AND RIGOROUS VALIDATION OF HIGHER DIMENSIONAL TRANSVERSE CONNECTING ORBITS

J. D. MIRELES JAMES*

Abstract. This paper presents a method for computing validated error bounds on the truncation error associated with the linear approximation of an (un)stable invariant manifold by its eigenspace. The method is based on studying a certain functional equation which describes a chart map for the invariant manifold. The truncation error is represented as a bounded analytic function on an explicitly given neighborhood. Moreover studying this functional equation leads to a method for bounding derivatives of the truncation error as well. The methods developed in the present work are well suited for studying higher dimensional invariant manifolds, and validated numerical results are provided for manifolds of dimension up to one hundred. As an application of these ideas we present some computer assisted proofs of transverse homoclinic connecting orbits.

1. Introduction. Invariant manifolds play a central role in the qualitative theory of dynamical systems. Over the last several decades substantial effort has gone into developing numerical methods for studying invariant manifolds associated with rest points of a dynamical system. A thorough review of the literature is beyond the scope of the present work and we refer the interested reader to [43, 16] for excellent overviews of the literature. At present we remark that many numerical methods for globalizing stable manifolds exploit the fact that the manifold is tangent to its stable eigenspace. The present work is concerned with obtaining mathematically rigorous bounds on the quality of this linear approximation via computer assisted arguments.

When the dynamical system under consideration is smooth enough (say C^1) the linear approximation of the manifold is quadratically good and in a "small enough" neighborhood of the rest point we expect that the truncation error is proportional to the square of the radius of the neighborhood. Our goal is to replace this qualitative statement with precise numerical bounds on the truncation error in a fixed neighborhood of the rest point. The computer is used in order to obtain precisely several system and dimension dependent constants needed in the argument.

Our approach is based on the *Parameterization Method* of Cabré, Fontich, and de la Llave [14, 15, 16]. The Parameterization Method provides a general functional analytic framework for studying invariant manifolds. We review the essentials in Section 2.1. For the moment we recall only the most visible feature of the method, which is to look for a chart map for the local invariant manifold which satisfies a certain invariance equation. In the present work we ask what precisely is the error made when we solve this invariance equation only to first order? Answers to this question are provided by Theorem 3.2 for analytic discrete time dynamical systems and by Theorem 4.2 for analytic differential equation.

Remark 1.1 (Derivatives). One useful feature of the Parameterization Method is that it provides several mechanisms for studying derivatives of the invariant manifolds. When the truncation error is represented as a bounded analytic function it is possible to bound the derivatives using the classical Cauchy bounds (see for example Lemma 2.2). The Cauchy bounds work best when the domain of the function is large relative to the bound on the size of the function. For the truncation error associated with

*Rutgers University, Department of Mathematics, 110 Frelinghuysen Rd, Piscataway, NJ 08854, USA.

the linear approximation this is not the case. A more quantitative discussion of this matter is taken up in Remark 5.1 parts (III) and (IV). The problem of obtaining good derivative bounds on the derivatives of the truncation error when the linear approximation is employed is addressed in Section 6. There we develop an alternative, computer assisted method for bounding derivatives which again exploits the operator equation for the chart map given by the Parameterization Method.

Remark 1.2 (Connecting Orbits). The intersection of stable and unstable manifolds give rise to connecting orbits: i.e. orbits which accumulate in forward/backward time to some specified rest points. Connecting orbits are widely studied in nonlinear analysis as they are low dimensional objects which carry global information about the system. Canonical illustrations of this principle in dynamical systems theory are the tangle theorems of Smale and Silnikov [68, 67]. An example from nonlinear analysis is the version of Morse theory developed by Witten and later on used by Floer [73, 65]. In this version of Morse theory the homology groups of a compact n-dimensional manifold are recovered by studying only zero dimensional equilibria and one dimensional connecting orbits between them. (See [66] for a fuller introduction to this branch of nonlinear analysis).

A classical numerical method for computing connecting orbits is the method of projected boundary conditions developed by Doedel, and Friedman for differential equations [25, 26] and by Beyn and Kleinkauf for diffeomorphisms [12, 13, 11]. Here the connecting orbit is reformulated as the solution of a certain boundary value problem, where the boundary conditions are taken to be on the local unstable and stable manifolds. This reformulation transforms the qualitative dynamical systems problem of finding connecting orbits into the more standard and quantitative problem of numerically solving a boundary value problem.

The references [7, 50, 54] develop a-posteriori, computer assisted arguments for the method of projected boundary conditions which facilitate rigorous computation of connecting orbits for differential equations and maps. These methods employ the Parameterization Method (with rigorous computer assisted error bounds) in order to control the boundary conditions needed in the problem set up. However the computer assisted arguments developed in the references just mentioned have up to now been implemented for only low dimensional problems (dimensions between two and six), as the high order methods used to compute the parameterizations are not practical for representing manifolds with a large number of stable or unstable dimensions.

As an application of the error bounds developed in the present work we give some computer assisted proofs for a higher dimensional problem in Section 7. We use a modification of the a-posteriori method developed in [54], extending to higher dimensions the results of that work.

Remark 1.3 (High Order Versus Low Order Parameterization). The references [54, 7, 51] consist of an effort to implement a high order approach to rigorous computation of local stable/unstable manifolds based on the Parameterization Method of [14, 15, 16]. The methods of [54, 7, 51] are “high order”, in the sense that in order to apply them it is necessary to approximate the manifolds to at least a minimum polynomial order determined by some problem dependent spectral gap conditions. (See for Example the condition given by Equation (3.6) in Theorem 3.2 of [54] and the condition given by Equation (64) in Theorem 4.2 of [7]). The present work is by contrast a “low order” offering, in the sense that here we are concerned with only linear approximation of a (often much higher dimensional) invariant manifold. Our goal is to avoid hypotheses which require the polynomial approximation be computed to some pre-specified order.

The key to both the high and the low order validation schemes is to start by obtaining some approximate solution of the invariance equation. Then a new fixed point operator for the associated truncation error is derived, and we show that this new fixed point operator is in fact a contraction on the Banach space of all candidates for the truncation error. Before beginning the nonlinear analysis it is first necessary to invert a particular linear operator in order to show that the fixed point operator is well defined. In the high order case one assumes that an approximation is known exactly to N-th order in the sense of power series. This makes the truncation error zero to N-th order, and if N is large enough it is possible to use some nice properties of N-tails in order to invert the linear operator.

The low order argument proceeds in a different way. A truncation error operator equation is derived via local considerations (local with respect to the rest point of the dynamical system). The local argument reduces the linear operator to constant coefficient and is solved directly. In other words explicit formulas for the action of the linear operator on Taylor coefficients are obtained. A benefit of this local approach is that it reveals a clear relationship between the spectrum of the differential at the rest point and the invertibility of the linear operator. This relationship is codified in a set of *non-resonance conditions* between the eigenvalues which, as we will see in Sections 3 and 4, can be checked *a-priori*. Considering the resonance conditions leads not only to the invertibility, but also to bounds on the linear operator.

In addition, the fact that the linear operator is inverted directly suggests that in the low order case it is more natural to work with Banach spaces of Taylor coefficient sequences. The desire to work with a sequence space also informs the choice of norms. In the present work we use Wiener norms, or weighted sums of the Taylor coefficients, rather than using the supremum norms as in [7, 54, 51]. This change in the choice of norm effects substantially the details of the nonlinear analysis which must be carried out in order to prove that the fixed point operator for the truncation error is a contraction, so that the present work has a distinctly different flavor than previous work of [7, 54, 51]. Nevertheless the present work is closely related to these previous works in the sense that taken together they constitute an effort to implement rigorous numerical methods for studying local stable/unstable manifolds based on the functional analytic approach of the Parameterization Method.

Remark 1.4 (Analytic Versus C^k Category). It should be noted that while the results of the present work apply to analytic maps and flows, the stable manifold theorem requires only C^k regularity. Indeed the original theoretical development of the Parameterization Method given in [14, 15, 16] is formulated for C^k mappings. The present work focuses on computer assisted validation methods for invariant manifolds and our arguments are streamlined by the use of complex analysis. This simplification is not unreasonable as many dynamical systems encountered in applied mathematics are given by analytic maps or vector fields.

That said, a great deal of work has gone into developing tools for studying nonlinear operator equations in the C^k category by computer assisted arguments. The interested reader might consult for example [47, 48, 49, 24] and the references discussed therein. The C^k approach discussed in the previous references could be applied to the study of solutions of Equations (2.1) and (2.2), however the technical details would be substantially different from those presented in the present work. The topological methods mentioned in Section 1.1 are also used in order to study stable/unstable manifolds. This analysis is carried out directly in phase space rather than in function space. Topological computer assisted arguments are often made in the C^k rather than the analytic category, through the use of isolating neighborhoods and cone conditions.

The remainder of the paper is organized as follows. In Section 1.1 we give a brief overview of the literature related to rigorous computation of local stable/unstable manifolds and computer assisted proof of connecting orbits.

The technical core of the paper is in Section 2. In Section 2.1 we review the basic ideas of the Parameterization Method, while in Section 2.2 we establish basic notations and recall some facts about analytic functions of several complex variables. In Section 2.3 we define the function spaces used throughout the paper and review some of their basic properties. Section 2.4 is where we develop the nonlinear analysis needed in order to obtain the validated error bounds which are the main goal of the paper. We solve a certain operator equation which arises in the study of both continuous and discrete time dynamical systems. Finally in Section 2.5 we recall some facts from Newton-Kantorovich analysis on $\mathbb{R}^n$.

In Section 3 we develop the necessary non-resonance conditions for maps, and state and prove our main theorem for maps. In Section 4 we repeat this analysis for differential equations. The main technical results of Sections 3 and 4 are Lemmas 3.3 and 4.3, which provide conditions under which certain linear operators are boundedly invertible.

In Section 5 we present the results of some numerical computations. The familiar Lorenz system is considered in Section 5.1, while in Section 5.2 we give results for the delayed Hénon map in a number of dimensions.

Section 6 is about bounding partial derivatives of the parameterization. In Section 6.1 we derive the functional equation which describes the partial derivatives of the parameterization, while in Section 6.2 we present the results of some numerical example calculations. In Section 7 we review the main ideas behind a-posteriori proof of the existence of connecting orbits based on the method of projected boundary conditions. Finally in Section 8 we present the results of some computer assisted proofs of connecting orbits for the delayed Hénon mapping in dimension up to twenty.

1.1. Related Work. The present work is by no means the first attempt to consider mathematically rigorous computation of stable and unstable manifolds. By now the topic has a long history as well as a substantial literature. We pause for a brief overview but caution that the following discussion is by no means exhaustive. The aim of the present section is to give only a rapid survey which directs the interested reader to the appropriate literature.

The first mathematically rigorous computer assisted study of stable and unstable manifolds which we are aware of is [56]. Here we find a method for rigorously enclosing the stable and unstable manifolds of planar diffeomorphisms. This paper also illustrates the close connection between connecting orbits and stable and unstable manifolds. Indeed we find here an early example of computer assisted proof of chaos based on studying intersection of the invariant manifolds. These results have been extended and generalized by many authors.

General purpose tools for computer assisted study of invariant manifolds based on topological expansion/contraction combined with some cone conditions are developed in [75, 31, 5, 20, 21], and the references discussed therein. These tools are implemented in the C++ library known as CAPD [8] In fact the second reference deals more generally with normally hyperbolic invariant sets, of which the stable/unstable manifolds of fixed points are only an example. These computations are based on studying expansion and contraction directly in the phase space and have a distinctly topological flavor. Moreover they are combined with a topological form of shadowing in order to study connecting orbits, as seen for example in the works [29, 6, 71, 19, 22]. We refer the reader directly to these papers for a more thorough discussion of the use of topological methods for computer assisted proof in dynamical systems. See also the work of [37, 38] for studies of symmetric connecting orbits of differential equations based on a rigorous computer assisted Melnikov analysis.

Another collection of general methods for studying stable and unstable manifolds based on the theory of Taylor Models are found in [9, 57, 72, 74], and the references discussed therein. These tools are implemented in the high level language called COSY Infinity [10]. With these methods the invariant manifolds are represented using high order polynomials. A somewhat subtle, but quite important remark is that the coefficients of the polynomials are not required to be the Taylor coefficients of a parameterization of the manifold, nor are they required to be interval enclosures of these coefficients. (Note that in general it is unlikely that it is even possible to represent the Taylor coefficients exactly as floating point numbers). The flexibility gained by suppressing the requirement that the polynomial coefficients are the Taylor coefficients has profound implication when iterates of the manifold are considered. Numerical round off errors associated with the polynomial coefficients are reallocated into the tail bounds by a process called *shrink wrapping* [40]. Shrink wrapping diminishes the "the wrapping effect" greatly, and the methods of the works just cited are able to propagate the invariant manifold for impressively long times. It should be noted that this flexibility is bought at the cost of loss of control of derivatives.

The rigorous validation of the error bounds for the Taylor Model approximations of the invariant manifold developed in [72, 74] are based on topological arguments in phase space. The approximate manifold is enclosed by a nonlinear polygon and the expansion/contraction of the nonlinear polygon is studied (in conjunction with some cone type conditions) in order to obtain tight C^0 bounds on the approximation error. In addition the Taylor Models of the invariant manifolds provide powerful tools for studying connecting dynamics and topological entropy, as seen for example in [57, 72]. Again, some topological crossing conditions must be checked in phase space in order to verify transversality of the connections.

Another vein of research which exploits high order polynomial approximation of the invariant manifolds is found in [41, 3, 7, 54]. Here the coefficients of the polynomial are typically required to be interval enclosures of the Taylor coefficients, and some control of derivatives is maintained. Another trait of these methods is that the error bounds are obtained through analytic and functional analytic methods in a Banach space of error functions, rather than topological methods in phase space. The methods just cited are most closely related to the present work and are discussed more extensively in the introduction remarks of the previous section.

A set of methods for computer assisted study of connecting dynamics which bypass direct study of the invariant manifolds are found in [70, 59, 23]. Here one studies the connecting orbits by imposing boundary conditions "at infinity" which are controlled using the theory of exponential dichotomies. Systems with exponential dichotomies have excellent "shadowing" or a-posteriori properties and these are used in order to validate the results of good numerics. The exponential dichotomies are also used to verify transversality for the orbits, so that these techniques are also used to prove the existence of chaos. See also the work of [42] for a method which studies the infinite time boundary value problem by projecting onto an appropriate Hermite spectral basis.

Another family of methods which provide computer assisted verification of the existence of connecting orbits without explicitly computing invariant manifolds are those based on the Conley Index Theory. These methods are implemented on the computer using methods of computational algebraic topology (cubical homology is especially well suited for computer implementation). We suggest [2, 24] for fuller discussion of these methods and the surrounding literature. The first reference treats maps while the second treats the Swift-Hohenberg partial differential equation.

2. Background. Suppose that $f: U \subset \mathbb{R}^n \to \mathbb{R}^n$ is a C^k diffeomorphism. (Here $k = \infty, \omega$ are also allowed and C^ω denotes analytic mappings). Let $p \in \mathbb{R}^n$ be a fixed point of f so that $f(p) = p$. The fixed point p is called a hyperbolic fixed point if $Df(p)$ is invertible and its spectrum does not intersect the unit circle. The point sets

$$W^s(p) := \{x \in U \,:\, f^n(x) \to p \text{ as } n \to \infty\},$$

$$W^u(p) := \{x \in U \,:\, f^{-n}(x) \to p \text{ as } n \to \infty\},$$

are called the stable and unstable sets of p respectively. The stable manifold theorem (see for example [60, 62]) states that the stable and unstable sets are in fact C^k manifolds tangent at p to the linearly stable (respectively linearly unstable) subspace of $Df(p)$. The dimension of these manifolds is equal to the number of stable (respectively unstable) eigenvalues of $Df(p)$. For this reason we call $W^{s,u}(p)$ the stable and unstable manifolds of p.

The stable/unstable manifolds may be immersed in $\mathbb{R}^n$ in complicated ways. Given an open neighborhood V of p the local stable set of p is

$$W^s_{\text{loc}}(p, V) = \{x \in U \,:\, f^n(x) \in V \text{ for all } n \geq 0\},$$

and the local unstable set of p is

$$W^u_{\text{loc}}(p, V) = \{x \in U \,:\, f^{-n}(x) \in V \text{ for all } n \geq 0\}.$$

Another consequence of the stable manifold theorem is that there exists a $V \subset U$ so that $W^{s,u}_{\text{loc}}(p, V)$ are C^k embedded disks, tangent at p to the stable/unstable linear subspace of $Df(p)$. The neighborhood V is not unique, so in practice we write $W^{s,u}_{\text{loc}}(p)$ to mean a local stable/unstable manifold for p relative to some neighborhood $V \subset U$. The sequel is largely concerned with parametric representations of the local stable/unstable manifolds. These definitions and theorems extend naturally to flows generated by vector fields by considering time tau maps. The interested reader and find precise definitions for example in [14, 16, 43]. In the sequel we employ these notions with no further comment.

2.1. Overview of The Parameterization Method for (Un)Stable Manifolds. The Parameterization Method, as introduced in [14, 15, 16], is a general functional analytic framework for studying stable and unstable manifolds associated with fixed points and equilibria of discrete and continuous time dynamical systems. The basic idea is to try to find a normal form for the stable (or unstable) dynamics in the chart space as well as a chart map for the invariant manifold (a parameterization) which conjugates the dynamics on the invariant manifold to the dynamics in the chart space given by the normal form.

For example suppose that $p \in \mathbb{R}^n$ and $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is a diffeomorphism with $f(p) = p$. Assume p is a hyperbolic fixed point in the sense of diffeomorphisms, and that $Df(p)$ has $k > 0$ stable eigenvalues $\lambda_1, \ldots, \lambda_n$ arranged by magnitude so that

$$0 < |\lambda_1| \leq |\lambda_2| \leq \ldots \leq |\lambda_k| < 1 < |\lambda_{k+1}| \leq \ldots \leq |\lambda_n|.$$

Let

$$\Lambda = \left(\begin{array}{ccc} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_k \end{array} \right),$$

denote the $k \times k$ diagonal matrix of stable eigenvalues.

For $\theta \in \mathbb{R}^k$ we define a normal form for the dynamics in parameter space by the mapping $\theta \to \Lambda\theta$. In other words we hope to "model" the dynamics on the stable manifold by the linear dynamics given by the diagonal matrix Λ. Then we seek a chart map $P\colon \mathbb{R}^k \to \mathbb{R}^n$ which conjugates the dynamics on the local stable manifold to the "model" or normal form dynamics given by the diagonal matrix Λ. More formally, we look for a chart map P which satisfies the conjugacy equation

$$f[P(\theta)] = P[\Lambda\theta], \tag{2.1}$$

for all θ in some neighborhood $U \subset \mathbb{R}^k$ of the origin.

This gives a functional equation for the chart map P. In fact is is natural to require that $P(0) = p$ and $DP(0) = A$ where A is an $n \times k$ matrix whose columns span the stable eigenspace associated with $\lambda_1, \ldots, \lambda_k$, making Equation (2.1) a functional equation with prescribed first order data. Observe that if $x \in \text{image}(P)$ then iterates of x converge to p, as is seen by repeatedly applying the conjugacy given by Equation (2.1). Then x is on the stable manifold by definition, and image(P) is a local stable manifold for p.

Similarly if $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is a smooth vector field, $p \in \mathbb{R}^n$ is a hyperbolic equilibrium, and Λ and A are as before then we look for a parameterization satisfying the equation

$$f[P(\theta)] = DP(\theta)\Lambda\theta, \tag{2.2}$$

which says that the tangent space of P is everywhere tangent to the vector field. Again it is natural to take $P(0) = p$ and $DP(0) = A$. To see that this is the correct conjugacy relation, suppose that ϕ is the flow generated by f. Equation (2.2) is derived from

$$\phi(P(\theta), t) = P\left(e^{\Lambda t}\theta\right),$$

by differentiating with respect to t and then taking the limit as $t \to 0$. But the previous equation makes it clear that if $x \in \text{imag}(P)$ then $\phi(x,t) \to p$ as $t \to \infty$, so that P is a parameterization of $W^s_{\text{loc}}(p)$.

In general we cannot solve the functional equation for P in closed form. Rather we compute an approximation to P, and then try to bound the associated error. In the present work we are interested in the linear approximation

$$P_1(\theta) = p + A\theta \approx P(\theta),$$

for $\theta \in U \subset \mathbb{R}^k$. The question is, on a given fixed domain U, how good is the approximation of P by P_1? This question is studied in detail in the sequel, under some additional (fairly weak) *non-resonance* assumptions on the eigenvalues.

We refer to [14, 15, 16] for a much more general treatment of the Parameterization Method, and remark that validation schemes for higher order approximations to P are developed in [7, 54]. For numerical implementation of these higher order approximation schemes we refer to [7].

2.2. Analytic Functions of Several Complex Variables: Definitions and Properties. Throughout the sequel it is convenient to work with complex analytic functions. Then we make some definitions. Let $z = a + ib \in \mathbb{C}$ and $|z| := \sqrt{a^2 + b^2}$ denote the usual absolute value on $\mathbb{C}$. We endow $\mathbb{C}^n$ the norm

$$\|z\| := \max_{1 \leq i \leq n} |z_i|.$$

For $z_0 \in \mathbb{C}^n$, $\rho > 0$ this norm induces the *polydisk* of radius ρ centered at z_0, i.e.

$$\mathbb{D}^n_{z_0}(\rho) := \{z \in \mathbb{C}^n \; : \; \|z - z_0\| < \rho\}.$$

When n is understood from context we write $\mathbb{D}^n_{z_0}(\nu) = \mathbb{D}_{z_0}(\nu)$. When $z_0 = 0$ we write $\mathbb{D}_0(\nu) = \mathbb{D}(\nu)$.

An $n \times k$ matrix A of complex numbers induces a linear map $A \colon \mathbb{C}^k \to \mathbb{C}^n$ under the usual matrix-vector product. We employ the matrix norm

$$\|A\| := \max_{1 \leq m \leq n} \sum_{i=1}^{k} |a_{mi}|$$

and have that $\|Az\| \leq \|A\| \|z\|$ for any $z \in \mathbb{C}^k$.

If $U \subset \mathbb{C}^n$ is an open set then the function $f \colon U \to \mathbb{C}$ is analytic in U if

$$\frac{\partial}{\partial z_i} f(z_1, \ldots, z_n) := \lim_{h \to 0} \frac{f(z_1, \ldots, z_i + h, \ldots, z_n) - f(z_1, \ldots, z_i, \ldots, z_n)}{h},$$

exists for each $1 \leq i \leq n$ and $(z_1, \ldots, z_n) \in U$. If $f \colon U \to \mathbb{C}$ is an analytic function on the open set U and $z_0 \in U$ with $\mathbb{D}_{z_0}(\rho) \subset U$, then f has convergent power series representation

$$f(z) = \sum_{|\beta|=0}^{\infty} a_\beta (z - z_0)^\beta.$$

Here $\beta := (\beta_1, \ldots, \beta_n) \in \mathbb{N}^n$ is an n-dimensional multi-index, $|\beta| := \beta_1 + \ldots + \beta_n$, $a_\beta \in \mathbb{C}$, and $z^\beta := z_1^{\beta_1} \cdot \ldots \cdot z_n^{\beta_n}$. The series converges absolutely for each $z \in \mathbb{D}_{z_0}(\rho)$ and uniformly on compact subsets of $\mathbb{D}_{z_0}(\rho)$. The elements of the sequence $\{a_\beta\}_{\beta \in \mathbb{N}^n}$ are called the Taylor coefficients of f. If f is analytic and $|f(z)| \leq M$ for all $z \in \mathbb{D}_{z_0}(\rho)$ then the Taylor coefficients satisfy

$$|a_\beta| \leq \frac{M}{\rho^{|\beta|}}. \tag{2.3}$$

Equation (2.3) is referred to as the Cauchy estimate. See any standard text on complex analysis (for example [1]) for the proof.

If f and g are analytic on U and $c \in \mathbb{C}$ then $f+g$, cf, ∂f, and $f \cdot g$ are all analytic on U. It follows by induction that $\partial^{|\beta|}/\partial z^{\beta} f$ is analytic on U for every multi-index $\beta \in \mathbb{N}^n$. For $f \colon \mathbb{D}_{z_0}(\nu) \to \mathbb{C}$ given by $f(z) = \sum_{|\beta|=0}^{\infty} a_{\beta}(z - z_0)^{\beta}$ we define the norm

$$\|f\|_{\nu} := \sum_{|\beta|=0}^{\infty} |a_{\beta}| \nu^{|\beta|}. \tag{2.4}$$

$\|f\|_{\nu} < \infty$ implies that the power series converges absolutely for each $\|z - z_0\| \leq \nu$, and that f is analytic on $\mathbb{D}_{z_0}(\nu)$. In addition we have that

$$\sup_{\|z - z_0\| \leq \nu} |f(z)| \leq \|f\|_{\nu}.$$

Remark 2.1 (Domain of analyticity versus domain of continuity). If $\|f\|_{\nu} < \infty$ it follows from Abel's Theorem that f extends to a continuous function on $\overline{\mathbb{D}_{z_0}(\nu)}$. Then we could in fact take the domain of f to be $\overline{\mathbb{D}_{z_0}(\nu)}$. Nevertheless we will continue to say that $\mathbb{D}_{z_0}(\nu)$ is the domain of f as this is the set on which f is an analytic function. However it should be understood that when desirable we can always take advantage of the fact that f is defined and continuous on the boundary of the polydisk.

The following estimate bounds the derivatives of an analytic function in terms of a bound on the supremum of the function over a polydisk. The cost is that the bound on the derivative holds only on a smaller polydisk. A proof which provides the explicit constants given here can be found in [54].

Lemma 2.2 (Cauchy Bounds). *Suppose that $f \colon \mathbb{D}_0(\nu) \subset \mathbb{C}^m \to \mathbb{C}^n$ is bounded and analytic. Then for any $0 < \sigma \leq 1$ we have that*

$$\sup_{\|z\| \leq \nu e^{-\sigma}} \|\partial_i f(z)\| \leq \frac{2\pi}{\nu\sigma} \sup_{\|z\| \leq \nu} \|f(z)\| \quad \text{and} \quad \sup_{\|z\| \leq \nu e^{-\sigma}} \|Df(z)\| \leq \frac{2\pi m}{\nu\sigma} \sup_{\|z\| \leq \nu} \|f(z)\|, \tag{2.5}$$

as well as

$$\sup_{\|z\| \leq \nu e^{-\sigma}} \|\partial_i \partial_j f(z)\| \leq \frac{4\pi^2}{\nu^2\sigma^2} \sup_{\|z\| \leq \nu} \|f(z)\| \quad \text{and} \quad \sup_{\|z\| \leq \nu e^{-\sigma}} \|D^2 f(z)\| \leq \frac{4\pi^2 m^2}{\nu^2\sigma^2} \sup_{\|z\| \leq \nu} \|f(z)\|. \tag{2.6}$$

Finally we consider two analytic functions which play an important role in the discussion to follow. The first is the geometric series in several complex variables. This is the function $f \colon \mathbb{D}_0(1) \to \mathbb{C}$ given by

$$f(z) = f(z_1, \ldots, z_n) = \sum_{|\beta|=0}^{\infty} z^{\beta} = \sum_{\beta_1=0}^{\infty} \cdots \sum_{\beta_n=0}^{\infty} z_1^{\beta_1} \cdot \ldots \cdot z_n^{\beta_n},$$

which is analytic on the open unit polydisk (but neither bounded nor continuous on the boundary). We note that if $0 < s < 1$ then

$$\sup_{\|z\| \leq s} |f(z)| \leq \frac{1}{(1 - s)^n},$$

as is seen by summing each geometric series in turn.

Next we consider the function $g\colon \mathbb{D}_0(1) \to \mathbb{C}$ given by

$$g(z) = \nabla f(z) \cdot z,$$

where f is the geometric series in n complex variables. We have that

$$g(z) = z_1 \partial_1 f(z) + \ldots + z_n \partial_n f(z) = \sum_{|\beta|=1}^{\infty} (\beta_1 + \ldots + \beta_n) z_1^{\beta_1} \cdot \ldots \cdot z_n^{\beta_n},$$

and have that if $0 < s < 1$ then

$$\sup_{\|z\| \le s} |g(z)| \le \sum_{|\beta|=1}^{\infty} |\beta| s^{|\beta|} = \frac{ns}{(1-s)^{n+1}},$$

as

$$\partial_i f(z) = \frac{1}{1-z_1} \cdot \ldots \cdot \frac{1}{(1-z_i)^2} \cdot \ldots \cdot \frac{1}{1-z_n}.$$

These remarks allow us to sum two infinite series which arise in the sequel when we bound certain second order remainder terms and their derivatives. We define the quantities

$$C_n(s) \equiv \sum_{|\beta|=2}^{\infty} s^{|\beta|-2}, \qquad \text{and} \qquad K_n(s) \equiv \sum_{|\beta|=2}^{\infty} |\beta| s^{|\beta|-2}. \tag{2.7}$$

We note that β is an n-dimensional multi-index, so that $C_n(s)$ and $K_n(s)$ are dimension dependent constants. Indeed we have that

$$\begin{aligned} C_n(s) &= \frac{1}{s^2} \left(\sum_{|\beta|=0}^{\infty} s^{|\beta|} - \sum_{|\beta|=1} s - \sum_{|\beta|=0} 1 \right) \\ &= \frac{1}{s^2} \left(\frac{1}{(1-s)^n} - s \begin{pmatrix} n \\ n-1 \end{pmatrix} - 1 \right) \\ &= \frac{1 - (1-s)^n - ns(1-s)^n}{s^2 (1-s)^n} \\ &= \frac{n^2 - (1+sn) \sum_{k=2}^{n} (-1)^k \begin{pmatrix} n \\ k \end{pmatrix} s^{k-2}}{(1-s)^n}. \end{aligned} \tag{2.8}$$

Similarly

$$\begin{aligned} K_n(s) &= \frac{1}{s^2} \left(\sum_{|\beta|=1}^{\infty} |\beta| s^{|\beta|} - \sum_{|\beta|=1} s \right) \\ &= \frac{1}{s^2} \left(\frac{ns}{(1-s)^{n+1}} - ns \right) \\ &= \frac{n \left(1 - (1-s)^{n+1}\right)}{s(1-s)^{n+1}} \\ &= \frac{n \left(n - 1 + ns - (1-s) \sum_{k=2}^{n} (-1)^k \begin{pmatrix} n \\ k \end{pmatrix} s^{k-1} \right)}{(1-s)^{n+1}}. \end{aligned} \tag{2.9}$$

Note that the last two lines in each computation make it clear that $C_n(s), K_n(s)$ approach non-zero, dimension dependent constants as $s \to 0$, while Equations (2.8) and (2.9) give explicit formulas which are especially convenient to work with in numerical computations.

2.3. Some Banach Spaces and Banach Algebras. Define the function space

$$\mathcal{W}_\nu^n = \{f \colon \mathbb{D}_0(\nu) \subset \mathbb{C}^n \to \mathbb{C} \ : \ \|f\|_\nu < \infty\},$$

where $\|f\|_\nu$ is the norm defined in Equation (2.4). $\mathcal{W}$ is sometimes called the positive Wiener algebra on $\mathbb{D}_0(\nu)$. Note that if $f, g \in \mathcal{W}_\nu^n$ then $f \cdot g$ is analytic and bounded on $\mathbb{D}_0(\nu)$, and extends to a continuous function on $\overline{\mathbb{D}_0(\nu)}$. In fact $\mathcal{W}_\nu^n$ is a Banach Algebra as suggested by the name. To see this let

$$f(z) = \sum_{|\beta|=0}^{\infty} a_\beta z^\beta, \qquad \text{and} \qquad g(z) = \sum_{|\beta|=0}^{\infty} b_\beta z^\beta,$$

and assume that $\|f\|_\nu, \|g\|_\nu < \infty$. Then

$$\begin{aligned}
\|f \cdot g\|_\nu &= \sum_{|\beta|=0}^{\infty} \left| \sum_{\beta_1+\beta_2=\beta} a_{\beta_1} b_{\beta_2} \right| \nu^{|\beta|} \\
&\leq \sum_{|\beta|=0}^{\infty} \sum_{\beta_1+\beta_2=\beta} |a_{\beta_1}||b_{\beta_2}| \nu^{|\beta|} \\
&= \left(\sum_{|\beta_1|=0}^{\infty} |a_{\beta_1}| \nu^{|\beta_1|} \right) \left(\sum_{|\beta_2|=0}^{\infty} |b_{\beta_2}| \nu^{|\beta_2|} \right) \\
&= \|f\|_\nu \|g\|_\nu < \infty.
\end{aligned}$$

Since $f \in \mathcal{W}_\nu^n$ is continuous on the boundary of $\mathbb{D}_0(\nu)$ and $\|f\|_\nu$ bounds the sup norm of f, we have the bounds on the decay rates of the power series coefficients of f

$$|a_\beta| \leq \frac{\|f\|_\nu}{\nu^{|\beta|}},$$

by applying the Cauchy estimate of Equation 2.3.

Now we define the main function spaces needed in the remainder of the the present work. We consider

$$\mathcal{W}_\nu^{k,n} = \left\{ f \colon \mathbb{D}_0^k(\nu) \to \mathbb{C}^n \mid f(\theta) = \sum_{|\alpha|=0}^{\infty} a_\alpha \theta^\alpha \ \text{ and } \ \|f\|_\nu = \sum_{|\alpha|=0}^{\infty} \|a_\alpha\| \nu^{|\alpha|} < \infty \right\},$$

$$\mathcal{X}_\nu^{k,n} = \left\{ f \colon \mathbb{D}_0^k(\nu) \to \mathbb{C}^n \mid f(\theta) = \sum_{|\alpha|=1}^{\infty} a_\alpha \theta^\alpha \ \text{ and } \ \|f\|_{\nu,1} = \sum_{|\alpha|=1}^{\infty} \|a_\alpha\| \nu^{|\alpha|} < \infty \right\},$$

and

$$\mathcal{H}_\nu^{k,n} = \left\{ f \colon \mathbb{D}_0^k(\nu) \to \mathbb{C}^n \mid f(\theta) = \sum_{|\alpha|=2}^{\infty} a_\alpha \theta^\alpha \ \text{ and } \ \|f\|_{\nu,2} = \sum_{|\alpha|=2}^{\infty} \|a_\alpha\| \nu^{|\alpha|} < \infty \right\}.$$

Note that in all cases $a_\alpha \in \mathbb{C}^n$.

Remarks 2.3.

- $\mathcal{W}_\nu^n = \mathcal{W}_\nu^{n,1}$ is just another notation for the Wiener algebra defined above.
- $\mathcal{X}_\nu^{k,1}$ and $\mathcal{H}_\nu^{k,1}$ are sub-Banach algebras of $\mathcal{W}_\nu^{k,1}$.
- $\mathcal{X}_\nu^{k,n}$ and $\mathcal{H}_\nu^{k,n}$ are sub-Banach spaces of $\mathcal{W}_\nu^{k,n}$ as the $\|\cdot\|_{\nu,1}$ and $\|\cdot\|_{\nu,2}$ norms are just the norms induced by the $\|\cdot\|_\nu$ norm. Nevertheless we sometimes find it suggestive to employ the subscripts as a reminder of which terms are present in the quantities being estimated. On the other hand we will suppress the subscripts on the norms when convenient.

The following Lemma collects some useful facts about the spaces and algebras defined above, and facilitates the nonlinear analysis of Section 2.4. The proof requires only the elementary Banach Algebra properties and the Mean Value Theorem, and is omitted for the sake of brevity. It is important to note that the functions f, g are assumed to be zero only to first order, but that the operations result in functions which are zero to at least second order.

Lemma 2.4. *Let $f, g \in \mathcal{X}_\nu^{k,n}$ and $\beta \in \mathbb{N}^n$ be an n-dimensional multi-index with $|\beta| \geq 2$. Then,*

(I) The function $f^\beta \colon \mathbb{D}_0^k(\nu) \to \mathbb{C}$ defined by

$$f^\beta(\theta) = f_1^{\beta_1}(\theta) \cdot \ldots \cdot f_n^{\beta_n}(\theta),$$

has $f^\beta \in \mathcal{H}_\nu^{k,1}$ and

$$\|f^\beta\|_{\nu,2} \leq \|f\|_{\nu,1}^{|\beta|}. \tag{2.10}$$

(II) For $\delta > 0$ suppose that $\|f\|_{\nu,1}, \|g\|_{\nu,1} \leq \delta$. Then $f^\beta - g^\beta \in \mathcal{H}_\nu^{k,1}$ and

$$\|f^\beta - g^\beta\|_{\nu,2} \leq |\beta|\delta^{|\beta|-1}\|f - g\|_{\nu,1}.$$

If f and g are equal to first order (i.e. if $Df(0) = Dg(0)$) then $f - g \in \mathcal{H}_\nu^{k,n}$ and the above becomes

$$\|f^\beta - g^\beta\|_{\nu,2} \leq |\beta|\delta^{|\beta|-1}\|f - g\|_{\nu,2}.$$

2.4. A Little Nonlinear Analysis: Operator Equation On $\mathcal{H}_\nu^{k,n}$. The goal of this section is to prove an abstract existence and uniqueness result for a certain nonlinear operator equation on $\mathcal{H}_\nu^{k,n}$ which arises when we study linear approximations of stable/unstable manifolds. The point is that exactly the same operator equation comes up in the case of both diffeomorphisms and differential equations, modulo the linear terms. Then it is reasonable to study the operator equation in the abstract, for an arbitrary boundedly invertible linear operator.

Before turning to this equation we we first establish some basic estimates for the composition of a function $R \in \mathcal{H}_\rho^{n,n}$ with functions $f, g \in \mathcal{X}_\nu^{k,n}$ under appropriate hypotheses. These estimates are essential in the sequel when we consider the truncation error associated with the linear approximation of stable manifolds. In particular it may help the reader to think of R as the second order Taylor remainder of the map/vector field expanded at the fixed point/equilibria and f, g as candidates for the parameterization of the stable manifold.

Lemma 2.5 (**Composition Lemma**). *Fix $\rho > 0$, suppose that $R \in \mathcal{H}_\rho^{n,n}$, and let $\hat{M} \geq 0$ be such that*

$$\sup_{\|z\| \leq \rho} \|R(z)\| \leq \hat{M}.$$

Let $0 < \hat{\delta} < \rho$ and let $C_n(\hat{\delta}/\rho), K_n(\hat{\delta}/\rho)$ be the dimension dependent sums defined in Equation (2.7).

(1): Let $f \in \mathcal{X}_\nu^{k,n}$ with $\|f\|_{\nu,1} \leq \hat{\delta}$ (so that the image of f is contained in the interior of the domain of R). Then $R \circ f \in \mathcal{H}_\nu^{k,n}$ and

$$\|R[f]\|_{\nu,2} \leq \frac{\hat{M} C_n(\hat{\delta}/\rho)}{\rho^2} \hat{\delta}^2. \tag{2.11}$$

(2): Let $f, g \in \mathcal{X}_\nu^{k,n}$, $\|f\|_{\nu,1}, \|g\|_{\nu,1} \leq \hat{\delta}$, and $Df(0) = Dg(0)$. Then $R[f] - R[g] \in \mathcal{H}_\nu^{k,n}$ and

$$\|R[f] - R[g]\|_{\nu,2} \leq \frac{\hat{M} K_n(\hat{\delta}/\rho)}{\rho^2} \hat{\delta} \|f - g\|_{\nu,2}. \tag{2.12}$$

Remark 2.6. Note that $R \in \mathcal{H}_\rho^{n,n}$ implies $\|R\|_{\rho,2} < \infty$, and we could take $\hat{M} = \|R\|_{\rho,2}$. On the other hand the bound on $\hat{M}$ could be obtained in some other way, for example using the Lagrange form of the Taylor remainder as in the applications below.

Proof. We have that

$$R(z) = \sum_{|\beta|=2}^{\infty} R_\beta \, z^\beta,$$

converges absolutely for all $\|z\| \leq \rho$ as $R \in \mathcal{H}_\rho^{n,n}$. Moreover the Cauchy estimates of equation (2.3) give that

$$\|R_\beta\| \leq \frac{\hat{M}}{\rho^{|\beta|}},$$

as R is continuous on the boundary.

If $f \in \mathcal{X}_\nu^{k,n}$ and $\|f\|_{\nu,1} \leq \hat{\delta}$ then the composition $R[f]$ is analytic on $\mathbb{D}_0^k(\nu)$ as the image of f is contained in the domain of R (i.e. $\hat{\delta} < \rho$). Moreover $R[f](0) = 0$ and $DR[f](0) = 0$ as

$$R[f](\theta) = \sum_{|\beta|=2}^{\infty} R_\beta f(\theta)^\beta,$$

and $f^\beta \in \mathcal{H}_\nu^{k,n}$ for all $|\beta| \geq 2$ by part (I) of Lemma 2.4.

To see that $R[f] \in \mathcal{H}_\nu^{k,n}$ we consider

$$\begin{aligned}
\|R[f(\theta)]\|_{\nu,2} &= \left\| \sum_{|\beta|=2}^{\infty} R_\beta [f(\theta)]^\beta \right\|_{\nu,2} \\
&\leq \sum_{|\beta|=2}^{\infty} \|R_\beta\| \|[f(\theta)]^\beta\|_{\nu,2} \\
&\leq \sum_{|\beta|=2}^{\infty} \|R_\beta\| \|f\|_{\nu,1}^{|\beta|} \\
&\leq \sum_{|\beta|=2}^{\infty} \frac{\hat{M}}{\rho^{|\beta|}} \hat{\delta}^{|\beta|} \\
&\leq \frac{\hat{M}}{\rho^2} \hat{\delta}^2 \sum_{|\beta|=2}^{\infty} \left(\frac{\hat{\delta}}{\rho} \right)^{|\beta|-2} \\
&\leq \frac{\hat{M} C_n(\hat{\delta}/\rho)}{\rho^2} \hat{\delta}^2,
\end{aligned}$$

where we use (I) of Lemma 2.4 to pass from the second to the third line. This proves part (1).

Now take $f, g \in \mathcal{X}_\nu^{k,n}$, with $Df(0) = Dg(0)$, and $\|f\|_{\nu,1}, \|g\|_{\nu,1} \leq \hat{\delta}$. Again it is clear that $R[f] - R[g]$ is analytic on $\mathbb{D}_0^k(\nu)$ as $\hat{\delta} < \rho$. By writing

$$R[f] - R[g] = \sum_{|\beta|=2}^{\infty} R_\beta \left(f(\theta)^\beta - g(\theta)^\beta \right),$$

we see that $(R[f] - R[g])(0) = 0$ and $D(R[f] - R[g])(0) = 0$ as (III) of Lemma 2.4 gives that $f^\beta - g^\beta \in \mathcal{H}_\nu^{k,n}$ for all $|\beta| \geq 2$. That $R[f] - R[g] \in \mathcal{H}_\nu^{k,n}$ now follows from part (1) of the present lemma and the fact that $\mathcal{H}_\nu^{k,n}$ is a Banach space.

Finally to obtain the Lipschitz bound on R we consider that $f - g \in \mathcal{H}_\nu^{k,n}$ (as the linear terms cancel) and have

$$\begin{aligned}
\|R[f] - R[g]\|_{\nu,2} &= \left\| \sum_{|\beta|=2}^{\infty} R_\beta \left([f(\theta)]^\beta - [g(\theta)]^\beta \right) \right\|_{\nu,2} \\
&\leq \sum_{|\beta|=2}^{\infty} \|R_\beta\| \left\| (f^\beta - g^\beta) \right\|_{\nu,2} \\
&\leq \sum_{|\beta|=2}^{\infty} \frac{\hat{M}}{\rho^{|\beta|}} |\beta| \hat{\delta}^{|\beta|-1} \|f - g\|_{\nu,2} \\
&\leq \frac{\hat{M}}{\rho^2} \hat{\delta} \left(\sum_{|\beta|=2}^{\infty} \frac{|\beta|}{\rho^{|\beta|-2}} \hat{\delta}^{|\beta|-2} \right) \|f - g\|_{\nu,2} \\
&\leq \frac{\hat{M} K_n(\hat{\delta}/\rho)}{\rho^2} \; \hat{\delta} \|f - g\|_{\nu,2}.
\end{aligned}$$

Here we have passed from the second to the third line using (III) of Lemma 2.4. Similarly, we are able to use the $\|\cdot\|_{\nu,2}$ norm throughout by (III) of Lemma 2.4 and the assumption that $Df(0) = Dg(0)$. This completes the proof of part (2). $\square$

Remark 2.7. The statement and proof of the theorem above goes through without change if we let $0 < s < 1$ be a positive number with $\hat{\delta}/\rho < s$ and replace $C_n(\hat{\delta}/\rho)$, $K_n(\hat{\delta}/\rho)$ with $C_n(s)$ and $K_n(s)$.

Remark 2.8 (Matrix norm inherited from $\mathcal{X}_\nu^{k,n}$). Note that an $n \times k$ complex matrix A can be viewed as an element of $\mathcal{X}_\nu^{n,k}$ in the usual way, i.e. letting A act on $\theta \in \mathbb{C}^k$ by matrix multiplication. In this setting we think of A as a power series with only linear terms. Let a_{e_j} denote the columns of A, so that $A = [a_{e_1}, \ldots, a_{e_k}]$. Then

$$\begin{aligned}
\|A\|_{\nu,1} &= \sum_{|\alpha|=1} \|a_\alpha\| \nu^{|\alpha|} \\
&= \sum_{j=1}^{k} \|a_{e_j}\| \nu \\
&= \sum_{j=1}^{k} \max_{1 \le i \le n} |a_{ij}| \nu \\
&= \|A\| \nu,
\end{aligned}$$

where $\|A\|$ is the matrix norm defined in Section 2.2. (This justifies the choice of norm).

We now simplify our notation and let $\|\cdot\|_\nu = \|\cdot\|_{\nu,2}$ denote the norm on $\mathcal{H}_\nu^{k,n}$ as these are our primary spaces throughout the remainder of the discussion.

Theorem 2.9. *Suppose that $R \in \mathcal{H}_\rho^{n,n}$ and that*

$$\sup_{\|z\| \le \rho} \|R(z)\| \le \hat{M}.$$

Let $A \in Mat_{n\times k}(\mathbb{C})$ be an $n \times k$ complex matrix and let $a > 0$ be a constant with $\|A\| \le a$. Finally suppose that $\mathfrak{L}$ is a boundedly invertible linear operator defined on $\mathcal{H}_\nu^{k,n}$ and that $\widetilde{C}$ is a positive constant so that $\mathfrak{L}^{-1} \colon \mathcal{H}_\nu^{k,n} \to \mathcal{H}_\nu^{k,n}$ has

$$\|\mathfrak{L}^{-1}\|_{\mathcal{H}_\nu^{k,n}} \le \widetilde{C}.$$

We assume that ν and δ are positive constants having $a\nu + \delta < \rho$ and that s is a positive constant with

$$\frac{a\nu + \delta}{\rho} < s < 1. \tag{2.13}$$

Now let $C_n(s)$ and $K_n(s)$ be constants defined in Equation (2.7), and assume that the inequalities

$$\frac{\widetilde{C}\hat{M}C_n(s)}{\rho^2}\left(a\nu + \delta\right)^2 - \delta \le 0, \tag{2.14}$$

and

$$\frac{\widetilde{C}\hat{M}K_n(s)}{\rho^2}\,(a\nu+\delta)<1, \tag{2.15}$$

are satisfied.

Then the functional equation

$$\mathfrak{L}[h](\theta)=R[A\theta+h(\theta)], \tag{2.16}$$

has a unique solution $h\in\mathcal{H}_\nu^{k,n}$ *satisfying the bound*

$$\|h\|_\nu\leq\delta.$$

Remark 2.10. Note that we make no assumption about the range of the linear operator $\mathfrak{L}$. What is important is the the inverse of $\mathfrak{L}$ is defined on and maps into $\mathcal{H}_\nu^{k,n}$.

Proof. Let

$$U_\delta=\{h\in\mathcal{H}_\nu^{k,n}\mid\|h\|_\nu\leq\delta\},$$

and note that U_δ is closed in $\mathcal{H}_\nu^{k,n}$. Define the nonlinear operator $\Phi\colon\mathcal{H}_\nu^{k,n}\to\mathcal{H}_\nu^{k,n}$ by

$$\Phi[h](\theta)=\mathfrak{L}^{-1}R[A\theta+h(\theta)]\quad\text{ for any }\quad h\in\mathcal{H}_\nu^{k,n}, \tag{2.17}$$

and observe that h is a solution of Equation (2.16) if and only if h is a fixed point of Equation (2.17).

To see that Φ is well defined take any $h\in\mathcal{H}_\nu^{k,n}$ and note that $A\theta+h(\theta)\in\mathcal{X}_\nu^{k,n}$. Then $R[A\theta+h(\theta)]\in\mathcal{H}_\nu^{k,n}$ by Lemma 2.5 (with $f(\theta)=A\theta+h(\theta)$ and $a\nu+\delta=\hat{\delta}$). Moreover (by taking again $\hat{\delta}=a\nu+\delta$) the estimate of Equation (2.11) in Lemma 2.5 gives

$$\|R[A+h]\|_\nu\leq\frac{\hat{M}C_n(s)}{\rho^2}(a\nu+\delta)^2<\infty.$$

Then $\Phi[h]\in\mathcal{H}_\nu^{k,n}$ by the assumption that $\mathfrak{L}\colon\mathcal{H}_\nu^{k,n}\to\mathcal{H}_\nu^{k,n}$ is invertible, and combining the previous inequality with Equation (2.14) gives

$$\|\Phi[h]\|_\nu\leq\|\mathfrak{L}^{-1}\|\|R[A+h]\|_\nu\leq\frac{\widetilde{C}\hat{M}C_n(s)}{\rho^2}(a\nu+\delta)^2\leq\delta.$$

This shows that $\Phi\colon U_\delta\to U_\delta$.

Now let $h_1,h_2\in U_\delta$. The Estimate given by Equation (2.12) of Lemma 2.5 (with $f=A+h_1$, $g=A+h_2$, and $a\nu+\delta=\hat{\delta}$) combined with Equation 2.14 gives

$$\begin{aligned}\|\Phi[h_1]-\Phi[h_2]\|_\nu&=\|\mathfrak{L}^{-1}\|\|R[h_1]-R[h_2]\|_\nu &(2.18)\\ &\leq\widetilde{C}\frac{\hat{M}K_n(s)}{\rho^2}(a\nu+\delta)\|h_1-h_2\|_\nu. &(2.19)\\ &\equiv\kappa\|h_1-h_2\|_\nu. &(2.20)\end{aligned}$$

Here $\kappa < 1$ follows from the assumption of Equation (2.15), and we conclude the existence of a unique fixed point of Φ on U_δ by the Contraction Mapping Theorem. ∎

The previous theorem requires a bound on the size of the function R on some disk where R is analytic. The following standard estimate provides such a bound when R is the second order Taylor remainder of an analytic function f expanded at z_0. The bound requires only knowledge of the size of the second partial derivatives of f in a neighborhood of z_0. We include the proof for the sake of completeness.

LEMMA 2.11. *Suppose that* $f\colon \mathbb{D}^n_{z_0}(\rho) \to \mathbb{C}^n$ *is analytic on* $\mathbb{D}^n_{z_0}(\rho)$ *and that* $M > 0$ *provides a bound of the form*

$$\max_{1\le j\le n}\max_{|\beta|=2}\sup_{z\in\mathbb{D}^n_{z_0}(\rho)}\left\|\frac{\partial^{|\beta|}}{\partial z^\beta}f_j(z)\right\| \le M.$$

Suppose in addition that $N_f \le n^2$ *is the maximum over* $1 \le j \le n$ *of the number of second partial derivatives of* f_j *which are not identically zero on* $\mathbb{D}^n_{z_0}(\rho)$, *i.e.*

$$N_f = \max_{1\le j\le n} \mathit{card}\left(\left\{\beta\in\mathbb{N}^n \,|\, |\beta| = 2 \text{ and } \partial^\beta f_j \not\equiv 0\right\}\right). \tag{2.21}$$

For any $\eta \in \mathbb{D}^n_0(\rho)$ *denote the second order Taylor expansion of* f *at* z_0 *by*

$$f(z_0+\eta) = f(z_0) + Df(z_0)\eta + R_{z_0}(\eta) \quad \eta\in\mathbb{D}^n_0(\rho).$$

Then the second order remainder function $R_{z_0}\colon \mathbb{D}^n_0(\rho) \to \mathbb{C}^n$ *is analytic in the variable* η, *has* $R(0) = 0$, $DR(0) = 0$, *and satisfies the bound*

$$\sup_{\|\eta\|\le\rho}\|R_{z_0}(\eta)\| \le N_f M\rho^2. \tag{2.22}$$

Proof. The Lagrange form of the Taylor remainder written out at second order this gives

$$[R_{z_0}]_j(\eta) = \sum_{|\beta|=2}\frac{2}{\beta!}\eta^\beta\int_0^1(1-t)\frac{\partial^{|\beta|}}{\partial z^\beta}f_j(z_0+t\eta)\,dt.$$

Then clearly R is zero to second order at $\eta = 0$. An application of Morera's theorem shows that $[R_{z_0}]_j$ is analytic for each $1 \le j \le n$, due to the fact that all partial derivatives of f are analytic.

Let $\sigma_j(\beta)$ be the function which is zero if $\partial^\beta f_j \equiv 0$ and one otherwise. A simple counting argument shows that $\sum_{|\beta|=2} 2\sigma_j(\beta)/\beta! \le N_f$. Then the bound claimed in Equation (2.22) follows from

$$\begin{aligned}\sup_{\|\eta\|<\rho}\|R_{z_0}(\eta)\| &= \max_{1\le j\le n}\sup_{\|\eta\|<\rho}\|[R_{z_0}]_j(\eta)\|\\ &\le \max_{1\le j\le n}\sum_{|\beta|=2}\frac{2\sigma_j(\beta)}{\beta!}\rho^{|\beta|}\int_0^1|1-t|\left|\frac{\partial^{|\beta|}}{\partial z^\beta}f_j(z_0+t\eta)\right|dt\\ &\le N_f M\rho^2.\end{aligned}$$

∎

2.5. A Little More Nonlinear Analysis: Newton-Like Operators on $\mathbb{R}^n$. In the present section we suppose that $\mathbb{R}^n$ is endowed with any norm whatsoever. Let $\|\cdot\|\colon \mathbb{R}^n \to \mathbb{R}^+$ denote this norm. For $r > 0$ and $x_0 \in \mathbb{R}^n$ let

$$B_{x_0}(r) := \{x \in \mathbb{R}^n \, : \, \|x - x_0\| < r\},$$

denote the ball of radius r about x_0 induced by the norm. If $A\colon \mathbb{R}^n \to \mathbb{R}^n$ is a linear operator then we let

$$\|A\| := \sup_{\eta \in \mathbb{R}^n, \|\eta\| = 1} \|A\eta\|,$$

be the operator norm induced by the given norm on $\mathbb{R}^n$.

In the applications to computer assisted proof for higher dimensional connecting orbits discussed in the sequel we exploit the following Newton-Kantorovich theorem. For the sake of clarity we state the theorem only for the case of maps on $\mathbb{R}^n$, as this is the only case required in the sequel. However the proof goes through more generally on a Banach space modulo the usual modifications.

Theorem 2.12 (Solution of a Newton-Like Equation with Approximate Inverse). *Suppose that $x_0 \in \mathbb{R}^n$, $r > 0$ and that $F\colon \overline{B_{x_0}(r)} \to \mathbb{R}^n$ is a $C^{1,\kappa}$ mapping. More explicitly we assume that F is differentiable on $\overline{B_{x_0}(r)}$ and that there is a $\kappa > 0$ so that for all $x, y \in \overline{B_{x_0}(r)}$ we have the Lipschitz bound*

$$\|DF(x) - DF(y)\| \leq \kappa \|x - y\|.$$

In addition, assume that $A \in Mat_{n\times n}(\mathbb{R})$ is a real $n \times n$ matrix with

$$\|I - ADF(x_0)\| \leq \delta < 1, \qquad \text{(Approximate Inverse)}, \tag{2.23}$$

and that

$$\|AF(x_0)\| \leq \epsilon. \qquad \text{(Approximate Solution)} \tag{2.24}$$

If

$$\|A\|\kappa r^2 - r(1-\delta) + \epsilon \leq 0, \tag{2.25}$$

and

$$\delta + \|A\| \, \kappa \, r < 1, \tag{2.26}$$

then there exists a unique $x^* \in \overline{B_{x_0}(r)}$ so that

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

Remarks 2.13.

- In applications the hypotheses of the theorem are checked via rigorous interval arithmetic. Since we will apply the theorem in a setting where n is quite large, the fact that the theorem does not require an exact inverse for the matrix $DF(x_0)$ is very useful. (Compare to the standard Newton-Kantorovich Theorem, for example Theorem 2.7 in [54]).

- We also note that it is rarely necessary to check explicitly that A is invertible. For example if the condition given by Equation (2.23) holds, then it follows by the Neumann theorem that $ADF(x_0)$, and hence A is invertible.
- The proof of the Theorem is standard and follows by showing that the Newton-like operator $T\colon \overline{B_{x_0}(r)} \to \mathbb{R}^n$ defined by

$$T(x) = x - AF(x),$$

  is a contraction mapping. We omit the details.

The following lemma allows us to conclude that the differential of F is nondegenerate at the zero of F given by Theorem 2.12. We include the elementary proof for the sake of completeness.

Lemma 2.14 (Nondegeneracy of a Solution of the Newton-Like Equation). *Suppose that F, x_0, A, and r satisfy the hypotheses of Theorem 2.12, and that x^* is the unique solution of $F = 0$ given by the theorem. Then $DF(x^*)$ is invertible and*

$$\left\| [DF(x^*)]^{-1} \right\| \leq \frac{\|A\|}{1 - \|A\|\kappa r - \delta}.$$

Proof. We have that

$$\begin{aligned} ADF(x^*) &= ADF(x^*) - ADF(x_0) + ADF(x_0) + I - I \\ &= I - [A(DF(x_0) - DF(x^*)) + (I - ADF(x_0))], \end{aligned} \tag{2.27}$$

which is invertible by the Neumann Theorem as

$$\|A(DF(x_0) - DF(x^*)) + (I - ADF(x_0))\| \leq \|A\|\kappa r + \delta < 1,$$

by hypothesis. Indeed we have that

$$\left\| [I - [A(DF(x_0) - DF(x^*)) + (I - ADF(x_0))]]^{-1} \right\| \leq \frac{1}{1 - \|A\|\kappa r - \delta}.$$

Since the product of two square matrices is invertible if and only if the matrices are invertible singly, we have that $DF(x)^*$ is invertible and that $[ADF(x^*)]^{-1} = DF(x^*)^{-1}A^{-1}$. Combining the previous observation with Equation (2.27) gives

$$[DF(x^*)]^{-1} = [I - [A(DF(x_0) - DF(x^*)) + (I - ADF(x_0))]]^{-1} A,$$

and

$$\|[DF(x^*)]^{-1}\| \leq \frac{\|A\|}{1 - \|A\|\kappa r - \delta},$$

as desired. ∎

3. Rigorous Bounds on the Linear Approximation of $W^s(p)$ for Diffeomorphisms. Throughout this section we make the following assumptions.

M1: Let $p \in \mathbb{C}^n$ and suppose that $f(p) = p$. Let $f\colon \mathbb{D}_p^n(\rho) \to \mathbb{C}^n$ be analytic and suppose that f has bounded and continuous second partial derivatives on $\overline{\mathbb{D}_p^n(\rho)}$.

M2: Let $Df(p)$ be diagonalizable over $\mathbb{C}$ and hyperbolic in the sense of diffeomorphisms.

Let $\lambda_1, \ldots, \lambda_n$ denote the eigenvalues of $Df(p)$. Since p is a hyperbolic fixed point there is a $0 \leq k \leq n$ so that the eigenvalues $\lambda_1, \ldots, \lambda_k$ are stable in the sense of diffeomorphisms. Assuming that $k \geq 1$ we have that

$$|\lambda_j| < 1, \qquad \text{for} \quad 1 \leq j \leq k.$$

Let Λ and Σ denote respectively the $k \times k$ diagonal matrix of stable eigenvalues and the $n \times n$ diagonal matrix of all the eigenvalues. Since $Df(p)$ is diagonalizable we choose eigenvectors $\xi_1, \ldots, \xi_n \in \mathbb{C}^n$, and note that these vectors are unique up to scaling. We let $\xi_1, \ldots, \xi_k$ be the eigenvectors associated with the stable eigenvalues (we call these the stable eigenvectors). Let

$$A = [\xi_1 | \ldots | \xi_k], \qquad \text{and} \qquad Q = [\xi_1 | \ldots | \xi_n],$$

denote respectively the $n \times k$ matrix whose columns are the stable eigenvectors and the $n \times n$ matrix whose columns are the complete collection of eigenvectors.

Definition 3.1. *Let $\alpha \in \mathbb{N}^k$ be a multi-index with $|\alpha| \geq 2$. We say that the collection of stable eigenvalues $\{\lambda_1, \ldots, \lambda_k\}$ has a resonance at α (in the sense of diffeomorphisms) if*

$$\lambda_1^{\alpha_1} \cdot \ldots \cdot \lambda_k^{\alpha_k} - \lambda_j = 0,$$

for some $1 \leq j \leq k$. If there are no resonances for any $\alpha \in \mathbb{N}^k$ then we say that the eigenvalues $\{\lambda_1, \ldots, \lambda_k\}$ are non-resonant.

The next lemma shows that in order to establish the non-resonance of a collection of stable eigenvalues we only have to check a finite number of conditions.

Lemma 3.2. *Let μ_*, μ^* be positive constants with*

$$0 < \mu_* \leq \min_{1 \leq j \leq k} |\lambda_j| \leq \max_{1 \leq j \leq k} |\lambda_j| \leq \mu^* < 1,$$

and let N_0 a positive integer with

$$\mu_* - (\mu^*)^{N_0} > 0.$$

If for each k-dimensional multi-index α with $2 \leq |\alpha| \leq N_0$ and every $1 \leq j \leq k$ we have

$$\lambda_1^{\alpha_1} \cdot \ldots \cdot \lambda_k^{\alpha_k} - \lambda_j \neq 0,$$

then the eigenvalues $\{\lambda_1, \ldots, \lambda_k\}$ are non-resonant.

Proof. Since the eigenvalues are assumed to be non-resonant for $|\alpha| \leq N_0$ we only have to check the remaining cases. So assume that $|\alpha| > N_0$ and consider

$$\begin{aligned} |\lambda_1^{\alpha_1} \cdot \ldots \cdot \lambda_k^{\alpha_k} - \lambda_j| &\geq ||\lambda_1^{\alpha_1} \cdot \ldots \cdot \lambda_k^{\alpha_k}| - |\lambda_j|| \\ &\geq \mu_* - (\mu^*)^{N_0} > 0, \end{aligned}$$

as

$$|\lambda_1^{\alpha_1} \cdot \ldots \cdot \lambda_k^{\alpha_k}| = |\lambda_1|^{\alpha_1} \cdot \ldots \cdot |\lambda_k|^{\alpha_k} \leq (\mu^*)^{|\alpha|} \leq (\mu^*)^{N_0} < \mu_* \leq |\lambda_j|.$$

□

Finally, in addition to **M1-M2** we make the following assumption.

M3: Assume that the set of stable eigenvalues are non-resonant in the sense of diffeomorphisms.

We also define the constant N_f as in Equation (2.21).

Definition 3.1 (Validation Values for the Linear Approximation of $W^s(p)$: Maps Case). We call the collection of positive constants $\mu_*, \mu^*, N_0, \widetilde{C}, a$, and M *a collection of validation values for the diffeomorphism f at the fixed point p* if μ_*, μ^*, and N_0 are as in Lemma 3.2 and

$$\max_{1\leq j\leq n} \max_{|\beta|=2} \sup_{z\in\mathbb{D}_p(\rho)} \left|\frac{\partial^{|\beta|}}{\partial z^\beta} f_j(z)\right| \leq M < \infty, \tag{3.1}$$

$$\|A\| \leq a, \tag{3.2}$$

and

$$\|Q\|\|Q^{-1}\| \max\left(\frac{1}{\mu_* - (\mu^*)^{N_0}}, \max_{\substack{1\leq j\leq k \\ 2\leq|\alpha|\leq N_0}} \left\{\frac{1}{|\lambda_1^{\alpha_1}\cdot\ldots\cdot\lambda_k^{\alpha_k} - \lambda_j|}\right\}\right) \leq \widetilde{C}. \tag{3.3}$$

Theorem 3.2. *Under assumptions **M1-M3** suppose that $\mu_*, \mu^*, N_0, \widetilde{C}, a$, and M are a collection of validation values in the sense of Definition 3.1 and let ν, δ and s be positive constants with*

$$\frac{a\nu+\delta}{\rho} < 1, \tag{3.4}$$

$$\widetilde{C}N_f M C_n(s)(a\nu+\delta)^2 - \delta \leq 0, \tag{3.5}$$

and

$$\widetilde{C}N_f M K_n(s)(a\nu+\delta) < 1. \tag{3.6}$$

Then there is a unique analytic function $h\colon \mathbb{D}_0^k(\nu) \to \mathbb{C}^n$ having that

$$P(\theta) = p + A\theta + h(\theta),$$

is an exact solution to Equation (2.1) and hence $P[\mathbb{D}_0^k(\nu)] = W^s_{loc}(p)$. Moreover

$$\sup_{\|\theta\|\leq\nu} |h(\theta)| \leq \|h\|_\nu \leq \delta.$$

3.1. Proof of Theorem 3.2. Suppose that $P(\theta) = p + A\theta + h(\theta)$ is an exact solution of Equation (2.1) so that

$$f\left[p + A\theta + h(\theta)\right] = p + A\Lambda\theta + h(\Lambda\theta).$$

Since f is analytic we Taylor expand the left hand side to second order and obtain

$$\begin{aligned} f\left[p + A\theta + h(\theta)\right] &= f(p) + Df(p)[A\theta + h(\theta)] + R_p[A\theta + h(\theta)] \\ &= p + Df(p)A\theta + Df(p)h(\theta) + R_p[A\theta + h(\theta)]. \end{aligned}$$

This gives

$$h(\Lambda\theta) - Df(p)h(\theta) = Df(p)A\theta - A\Lambda\theta + R_p[A\theta + h(\theta)].$$

However note that

$$\begin{aligned} Df(p)A\theta - A\Lambda\theta &= (Df(p)\theta_1\xi_1 + \ldots Df(p)\theta_k\xi_k) - (\lambda_1\theta_1\xi_1 + \ldots + \lambda_k\theta_k\xi_k) \\ &= 0, \end{aligned} \tag{3.7}$$

as for each $1 \leq j \leq k$ we have that $\theta_j\xi_j$ is an eigenvector for $Df(p)$ associated associated with the eigenvalue λ_j. Then the truncation error function h solves the equation

$$\mathfrak{L}[h](\theta) = R_p[A\theta + h(\theta)], \tag{3.8}$$

where $\mathfrak{L}$ is the linear operator defined by

$$\mathfrak{L}[h](\theta) = h(\Lambda\theta) - Df(p)h(\theta). \tag{3.9}$$

Since this has the form given by Equation (2.16) the proof is complete as soon as we show that $\mathfrak{L}$ and R_p satisfy the hypotheses of Theorem 2.9. Using that $N_f M \rho^2$ satisfies the condition given by Equation (3.1) in the definition of validation values we have that R_p is bound by $\hat{M} = N_f M \rho^2$ again by Lemma 2.11. Then R_p satisfies the hypotheses of Theorem 2.9 (note that there is a cancelation of the ρ^2 terms). What remains is to show that $\mathfrak{L}$ also satisfies the hypotheses of Theorem 2.9, and this is established by the following Lemma, which completes the proof.

LEMMA 3.3. *Assume **M1-M3** and let μ_*, μ^*, N_0, and $\widetilde{C}$ be as in Definition 3.1. Then the expression given by Equation (3.9) defines a boundedly invertible linear operator on $\mathcal{H}_\nu^{k,n}$. Moreover we have that*

$$\|\mathfrak{L}^{-1}\|_{\mathcal{H}_\nu^{k,n}} \leq \widetilde{C},$$

where $\widetilde{C}$ is as in Equation (3.3) of Definition 3.1.

Proof. Let $\eta \in \mathcal{H}_\nu^{k,n}$ and consider the equation

$$\mathfrak{L}[q](\theta) = q(\Lambda\theta) - Df(p)q(\theta) = \eta(\theta) \qquad \text{for} \qquad \theta \in \mathbb{D}_0^k(\nu),$$

with q unknown. Making the substitution $q = Qw$ gives that

$$Q\left[w(\Lambda\theta) - \Sigma w(\theta)\right] = \eta(\theta),$$

or

$$w(\Lambda\theta) - \Sigma w(\theta) = Q^{-1}\eta(\theta),$$

with w unknown. Since $\eta \in \mathcal{H}_\nu^{k,n}$ we have

$$\eta(\theta) = \sum_{|\alpha|=2}^{\infty} \eta_\alpha \theta^\alpha.$$

We are seeking a solution $w \in \mathcal{H}_\nu^{k,n}$ so let

$$w(\theta) = \sum_{|\alpha|=2}^{\infty} w_\alpha \theta^\alpha.$$

Then

$$\sum_{|\alpha|=2}^{\infty} \left(\lambda_1^{\alpha_1} \cdot \ldots \cdot \lambda_k^{\alpha_k} - \Sigma\right) w_\alpha \theta^\alpha = \sum_{|\alpha|=2}^{\infty} Q^{-1} \eta_\alpha \theta^\alpha.$$

Matching like powers gives and solving for the components of the w_α coefficients gives

$$(w_\alpha)_j = \frac{1}{\lambda_1^{\alpha_1} \cdot \ldots \cdot \lambda_k^{\alpha_k} - \lambda_j} \left(Q^{-1}\eta_\alpha\right)_j. \tag{3.10}$$

Note that these coefficients are well defined for all multi-indices $|\alpha| \geq 2$ by **M3**. Taking Equation (3.10) as the *definition* of the components of $w(\theta)$ we see that

$$\begin{aligned} \|w\|_\nu &= \sum_{|\alpha|=2}^{\infty} |w_\alpha| \nu^{|\alpha|} \\ &\leq \max \left(\frac{1}{\mu_* - (\mu^*)^{N_0}}, \max_{\substack{1 \leq j \leq k \\ 2 \leq |\alpha| \leq N_0}} \left\{ \frac{1}{|\lambda_1^{\alpha_1} \cdot \ldots \cdot \lambda_k^{\alpha_k} - \lambda_j|} \right\} \right) \|Q^{-1}\| \|\eta\|_\nu. \end{aligned}$$

Defining $v(\theta) = Qw(\theta)$ and working backwards we see that $\mathfrak{L}[v](\theta) = \eta(\theta)$ and moreover

$$\|q\|_\nu \leq \widetilde{C} \|\eta\|_\nu.$$

Defining $\mathfrak{L}^{-1}[\eta](\theta) = q(\theta)$ and taking the supremum over all η with norm one we have

$$\|\mathfrak{L}^{-1}\|_{\mathcal{H}_\nu^{k,n}} \leq \widetilde{C},$$

as desired. ∎

4. Rigorous Bounds on the Linear Approximation of $W^s(p)$ for Differential Equations. In this section we make the following assumptions.

- **F1**: Let $p \in \mathbb{C}^n$ have $f(p) = 0$. Let $f \colon \mathbb{D}_p^n(\rho) \to \mathbb{C}^n$ be analytic with bounded and continuous second partial derivatives on $\overline{\mathbb{D}_p^n(\rho)}$.
- **F2**: Let $Df(p)$ be diagonalizable over $\mathbb{C}$ and hyperbolic in the sense of differential equations.

Let $\lambda_1, \ldots, \lambda_n$ denote the eigenvalues of $Df(p)$. Since p is a hyperbolic equilibria there is a $0 \leq k \leq n$ so that the eigenvalues $\lambda_1, \ldots, \lambda_k$ have

$$\mathrm{real}(\lambda_k) \leq \ldots \leq \mathrm{real}(\lambda_1) < 0.$$

Again we take Λ and Σ to be the $k \times k$ diagonal matrix of stable eigenvalues and the $n \times n$ diagonal matrix of all the eigenvalues, choose eigenvectors $\xi_1, \ldots, \xi_n \in \mathbb{C}^n$, let $\xi_1, \ldots, \xi_k$ be the eigenvectors associated with the stable eigenvalues, and define

$$A = [\xi_1 | \ldots | \xi_k], \quad \text{and} \quad Q = [\xi_1 | \ldots | \xi_n],$$

to be respectively the $n \times k$ matrix whose columns are the stable eigenvectors and the $n \times n$ matrix whose columns are the complete collection of eigenvectors.

DEFINITION 4.1. *Let $\alpha \in \mathbb{N}^k$ be a multi-index with $|\alpha| \geq 2$. We say that the collection of stable eigenvalues $\{\lambda_1, \ldots, \lambda_k\}$ has a resonance (in the sense of differential equations) at α if*

$$\alpha_1\lambda_1 + \ldots + \alpha_k\lambda_k - \lambda_j = 0,$$

for some $1 \leq j \leq k$. If there are no resonances for any $\alpha \in \mathbb{N}^k$ then we say that the eigenvalues $\{\lambda_1, \ldots, \lambda_k\}$ are non-resonant.

Again this is actually a finite set of conditions.

LEMMA 4.2. *Let μ_*, μ^* be positive constants with*

$$0 < \mu_* \leq \min_{1 \leq j \leq k} |\mathrm{real}(\lambda_j)| \leq \max_{1 \leq j \leq k} |\mathrm{real}(\lambda_j)| \leq \mu^*,$$

and let N_0 be a positive integer so that

$$N_0\mu_* - \mu^* > 0.$$

If for each k-dimensional multi-index α with $2 \leq |\alpha| \leq N_0$ and every $1 \leq j \leq k$ we have

$$\alpha_1\lambda_1 + \ldots \alpha_k\lambda_k - \lambda_j \neq 0,$$

then the eigenvalues $\{\lambda_1, \ldots, \lambda_k\}$ are non-resonant.

Proof. By hypothesis we have only to rule out resonances of order greater than N_0. Then let $\alpha \in \mathbb{N}^k$ have $|\alpha| > N_0$ and define $\lambda_j = a_j + ib_j$ for all $1 \leq j \leq k$. Recalling that $a_j < 0$ for each $1 \leq j \leq k$ we have

$$\begin{aligned}
|\alpha_1 a_1 + \ldots + \alpha_k a_k| &= \alpha_1|a_1| + \ldots + \alpha_k|a_k| \\
&\geq \alpha_1\mu_* + \ldots + \alpha_k\mu_* \\
&\geq N_0\mu_* \\
&\geq \mu^* \\
&\geq |a_j| \\
&\geq 0.
\end{aligned}$$

Then

$$\begin{aligned}
|\alpha_1\lambda_1 + \ldots \alpha_k\lambda_k - \lambda_j| &\geq |\alpha_1 a_1 + \ldots \alpha_k a_k - a_j| \\
&\geq ||\alpha_1 a_1 + \ldots \alpha_k a_k| - |a_j|| \\
&\geq ||\alpha|\mu_* - \mu^*| \\
&> N_0\mu_* - \mu^* \\
&> 0,
\end{aligned}$$

giving that $\lambda_1, \ldots, \lambda_k$ are non-resonant for all $|\alpha| > N_0$. Since we assumed that there were no resonances for $2 \leq |\alpha| \leq N_0$ we have that the eigenvalues are non-resonant. ∎

F3: Assume that the stable collection of eigenvalues are non-resonant in the sense of differential equations.

Again we take N_f to be defined as in Equation (2.21).

DEFINITION 4.1 (Validation Values for the Linear Approximation of $W^s(p)$: Flows Case). We call the collection of positive constant $\mu_*, \mu^*, N_0, \widetilde{C}, a,$ and M *a set of validation values for f at p* if μ_*, μ^*, and N_0 are as in Lemma 4.2 and

$$\max_{1\le j\le n}\max_{|\beta|=2}\sup_{z\in\mathbb{D}_p(\rho)}\left|\frac{\partial^{|\beta|}}{\partial z^\beta}f_j(z)\right| \le M < \infty, \tag{4.1}$$

$$\|A\| \le a, \tag{4.2}$$

and

$$\|Q\|\|Q^{-1}\| \max\left(\frac{1}{N_0\mu_* - \mu^*}, \max_{\substack{1\le j\le k\\ 2\le|\alpha|\le N_0}}\left\{\frac{1}{|\alpha_1\lambda_1+\ldots+\alpha_k\lambda_k-\lambda_j|}\right\}\right) \le \widetilde{C}. \tag{4.3}$$

THEOREM 4.2. *Under assumptions **F1-F3** suppose that $\mu_*, \mu^*, N_0, \widetilde{C}, a,$ and M are a collection of validation values in the sense of Definition 4.1 and let ν, δ and s be positive constants with*

$$\frac{a\nu+\delta}{\rho} < s < 1, \tag{4.4}$$

$$\widetilde{C}N_f M C_n(s)(a\nu+\delta)^2 - \delta \le 0, \tag{4.5}$$

and

$$\widetilde{C}N_f M K_n(s)(a\nu+\delta) < 1. \tag{4.6}$$

Then there is a unique analytic function $h\colon \mathbb{D}_0^k(\nu)\to\mathbb{C}^n$ having that

$$P(\theta) = p + A\theta + h(\theta),$$

is an exact solution to the invariance Equation (2.2) and hence $P[\mathbb{D}_0^k(\nu)] = W^s_{loc}(p)$. Moreover

$$\sup_{\|\theta\|\le\nu}|h(\theta)| \le \|h\|_\nu \le \delta.$$

4.1. Proof of Theorem 4.2. The proof is similar to before. Let $P(\theta) = p + A\theta + h(\theta)$ be an exact solution of the invariance Equation (2.2). Then

$$f\left[p + A\theta + h(\theta)\right] = D[p + A\theta + h(\theta)]\Lambda\theta.$$

We Taylor expand the left hand side to second order and obtain

$$\begin{aligned} f\left[p + A\theta + h(\theta)\right] &= f(p) + Df(p)[A\theta + h(\theta)] + R_p[A\theta + h(\theta)] \\ &= Df(p)A\theta + Df(p)h(\theta) + R_p[A\theta + h(\theta)]. \end{aligned}$$

This gives

$$Dh(\theta)\Lambda\theta - Df(p)h(\theta) = Df(p)A\theta - A\Lambda\theta + R_p[A\theta + h(\theta)].$$

We exploit again the cancelations shown in Equation (3.7) and obtain

$$\mathfrak{L}[h](\theta) = R_p[A\theta + h(\theta)], \tag{4.7}$$

where

$$\mathfrak{L}[h](\theta) = Dh(\theta)\Lambda\theta - Df(p)h(\theta). \tag{4.8}$$

This has the form given by Equation (2.16) and R_p satisfies the hypotheses of Theorem 2.9 with $\hat{M} = N_f M\rho^2$ through Lemma 2.11. The proof is completed by the following Lemma.

LEMMA 4.3. *Assume **F1-F3** and let μ_*, μ^*, N_0, and $\widetilde{C}$ be as in Definition 4.1. Then the expression given by Equation (4.8) defines a boundedly invertible linear operator on $\mathcal{H}_\nu^{k,n}$ with*

$$\|\mathfrak{L}^{-1}\|_{\mathcal{H}_\nu^{k,n}} \leq \widetilde{C},$$

and $\widetilde{C}$ is as in Equation (4.3) of Definition 4.1.

Proof. We take $\eta \in \mathcal{H}_\nu^{k,n}$ and consider

$$\mathfrak{L}[q](\theta) = Dq(\theta)\Lambda\theta - Df(p)q(\theta) = \eta(\theta) \qquad \text{for} \qquad \theta \in \mathbb{D}_0^k(\nu).$$

Let $q = Qw$ so that

$$Dw(\theta)\Lambda\theta - \Sigma w(\theta) = Q^{-1}\eta(\theta),$$

Formally then

$$\eta(\theta) = \sum_{|\alpha|=2}^{\infty} \eta_\alpha \theta^\alpha.$$

and

$$w(\theta) = \sum_{|\alpha|=2}^{\infty} w_\alpha \theta^\alpha,$$

so that

$$\sum_{|\alpha|=2}^{\infty} \left(\alpha_1\lambda_1 + \ldots + \alpha_k\lambda_k - \Sigma\right) w_\alpha \theta^\alpha = \sum_{|\alpha|=2}^{\infty} Q^{-1}\eta_\alpha\theta^\alpha,$$

or

$$\left(w_\alpha\right)_j = \frac{1}{\alpha_1\lambda_1 + \ldots + \alpha_k\lambda_k - \lambda_j}\left(Q^{-1}\eta_\alpha\right)_j. \tag{4.9}$$

These coefficients are well defined by **F3**. Then defining $w(\theta)$ by Equation (4.9) and $q(\theta) = Qw(\theta)$ gives

$$\|q\|_\nu \leq \widetilde{C}\|\eta\|_\nu.$$

Since $\mathfrak{L}^{-1}[\eta](\theta) = q(\theta)$ taking the supremum over all η with norm one gives the desired bound

$$\|\mathfrak{L}^{-1}\|_{\mathcal{H}_\nu^{k,n}} \leq \widetilde{C}.$$

∎

5. Example Computations. We now present some example computations which illustrate the effectiveness of the machinery developed above. As a first example we compute rigorous bounds for the two dimensional stable manifold of the Lorenz system at the origin. We present a step by step breakdown of the computer assisted proof as the necessary steps are completely representative of what is done in general. Then we present the results of a number of validated computations for the delayed Hénon mapping, in dimensions through one hundred.

5.1. Lorenz. Let $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ be the Lorenz Vector Field given by

$$f(x,y,z) = \left(\begin{array}{c} \sigma(y-x) \\ x(\rho - z) - y \\ xy - \beta z \end{array} \right), \tag{5.1}$$

with σ, β, ρ parameters. For all values of the parameters the system has an equilibria at the origin. In this section we consider the validated computer assisted error bounds for the linear approximation of the stable manifold given by Theorem 4.2.

First consider the classical parameter values $\sigma = 10$, $\beta = 8/3$, and $\rho = 28$. Of course the eigenvalue and eigenvectors can be computed by hand for this system. However we use the *IntLab* program "verrifyeig" as this illustrates the procedure we must use in higher dimensions. For these parameters the eigenvalues associated with the differential of the equilibrium at the origin have

$$\lambda_1^s \in B(-22.82772345116345, 7.12 \times 10^{-15}),$$

$$\lambda_2^s \in B(-2.66666666666666, 4.45 \times 10^{-15}),$$

$$\lambda^u \in B(11.82772345116346, 7.12 \times 10^{-15}),$$

and we choose associated eigenvectors enclosed by

$$\xi_1^s \in B\left(\left[\begin{array}{c} -0.61481678521647 \\ 0.78866996938902 \\ -0.00000000000000 \end{array} \right], 2.23 \times 10^{-16} \right),$$

$$\xi_2^s \in B\left(\left[\begin{array}{c} 0.0 \\ 0.0 \\ 1.0 \end{array} \right], 1.21 \times 10^{-323} \right),$$

and

$$\xi^u \in B\left(\left[\begin{array}{c} -0.41650417819290 \\ -0.90913380178489 \\ 0.00000000000000 \end{array} \right], 1.67 \times 10^{-16} \right).$$

We let $Q = [\xi_1^s|\xi_2^s|\xi^u]$, $A = [\xi_1^s|\xi_2^s]$, and use the *IntLab* function "inv($\cdot$)" to compute an inclosure of Q^{-1}. Set $\varepsilon_{\text{tol}} = 2.25 \times 10^{-16}$ to be the desired tolerance of the final result. We present a step-by-step break down of the steps required in order to choose validation values and check the hypotheses of Theorem 4.2 for this problem.

Validated Computation Steps:

(1) First choose a radius R for the polydisk on which to bound the second partial derivatives of the vector field. Since f is quadratic in this case the choice of R is completely arbitrary. We take $R = 1$. (In the theorem we had the notation $R = \rho$. But in this section we reserve ρ for the Lorenz parameter).

(2) We note that

$$Df(x, y, z) = \begin{pmatrix} -\sigma & \sigma & 0 \\ \rho - z & -1 & -x \\ y & x & -\beta \end{pmatrix},$$

and the only non-zero second partials are

$$\frac{\partial^2}{\partial x \partial z} f_2 = \frac{\partial^2}{\partial z \partial x} f_2 = -1$$

(so this counts as two), and

$$\frac{\partial^2}{\partial y \partial x} f_3 = \frac{\partial^2}{\partial x \partial y} f_3 = 1,$$

(two again). Then $N_f = 2$ as this is the componentwise max, and we define

$$\max_{1 \leq j \leq 3} \max_{|\beta|=2} \sup_{z \in \mathbb{D}_p(\rho)} |\partial^\beta f_j(z)| \leq 1 = M.$$

(3) Using interval arithmetic we check that

$$\|Q\| \leq 1.698, \qquad \|A\| \leq a = 1, \qquad \text{and} \qquad \|Q^{-1}\| \leq 1.59.$$

(4) We take $\mu_* = 2.667$ and $\mu^* = 22.828$. Note that

$$8\mu_* - \mu^* < -1.494, \qquad \text{and} \qquad 9\mu_* - \mu^* > 1.172.$$

Then it is sufficient to take $N_0 = 9$. We check that

$$\frac{1}{N_0 \mu_* - \mu^*} \leq 0.854,$$

and

$$\max_{1 \leq j \leq 2} \max_{2 \leq |\alpha| \leq 9} \left(\frac{1}{|\alpha_1 \lambda_1^s + \alpha_2 \lambda_2^s - \lambda_j^s|} \right) \leq 0.853.$$

(5) Then $\widetilde{C} = \|Q\| \|Q^{-1}\| 0.854 \leq 2.291$.

(6) We choose $s = 10^{-6}$ and have $n = 3$ (the dimension of the phase space). Then we compute $C_3(s) = 6.02$ and $K_3(s) = 12.04$.

(7) We seek a final truncation error on the order of machine epsilon. Then set $\sqrt{\varepsilon_{\text{tol}}} = 1.582 \times 10^{-8}$.

(8) Using interval arithmetic we compute

$$\widetilde{C} N_f M C_3(s) \leq 54.972 \equiv B.$$

We check that

$$2aB\nu \leq 1.739 \times 10^{-6} < 1,$$

and see that the discriminant has

$$(2aB\nu - 1)^2 - 4B^2a^2\nu^2 \geq 0.9999965 > 0.$$

Denote by $[\delta_*, \delta^*]$ the set of all δ satisfying Equation (4.5). Using interval arithmetic we check that

$$[\delta_*, \delta^*] \subset \left[1.375 \times 10^{-14}, 0.01819\right].$$

Then we choose

$$\delta = 1.4 \times 10^{-14}.$$

(9) We check that

$$\frac{a\nu + \delta}{\rho} \leq 1.59 \times 10^{-8} < s = 10^{-6}.$$

(10) Finally we check that

$$\widetilde{C} M N_f K_3(s)(a\nu + \delta) \leq 1.74 \times 10^{-6} < 1,$$

and have that all the hypotheses of Theorem 4.2 are verified. The computation has succeeded.

Then the parameterization given by the linear approximation has

$$P_1(\theta_1, \theta_2) = A \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix},$$

and we have that there is an analytic truncation error function

$$h \colon B(1.582 \times 10^{-8}) \subset \mathbb{R}^2 \to \mathbb{R}^3.$$

Then $P(\theta) = A\theta + h(\theta)$ and

$$\sup_{|\theta_1|, |\theta_2| \leq 1.582 \times 10^{-8}} |h(\theta_1, \theta_2)| \leq 1.4 \times 10^{-14}.$$

Since the size of our domain is roughly 10^{-8} and the size of our error is roughly 10^{-14} this says that we have roughly six significant figures in the approximation. Another way to think about it is that we asked for $\epsilon_{\text{tol}} = 2.5 \times 10^{-16}$ accuracy. Then we picked a disk of radius 10^{-8} and obtained a final accuracy of 10^{-14}. Then the "constant" in our expected quadratic bounds is on the order of 10^2.

Table 5.1 records the results of a number of other computations with various validation domain radii-ν. We see that we can take the validation domain as large as 10^{-3}, however we only obtain roughly one significant figure of accuracy.

5.2. The Delayed Hénon Map. The Delayed Hénon Map gives an example dynamical system in any dimension n. This is a map $f \colon \mathbb{R}^n \to \mathbb{R}^n$ given by

$$f(x_1, \ldots, x_n) = \begin{pmatrix} 1 - ax_1^2 + bx_n \\ x_1 \\ \vdots \\ x_{n-1} \end{pmatrix}, \tag{5.2}$$

Domain Size	s	Error δ
3.16×10^{-8}	10^{-3}	5.51×10^{-14}
1×10^{-7}	10^{-3}	5.51×10^{-13}
3.17×10^{-7}	10^{-3}	5.51×10^{-12}
3.16×10^{-6}	10^{-3}	5.51×10^{-10}
3.16×10^{-4}	10^{-3}	5.71×10^{-6}
1×10^{-3}	10^{-2}	9.5×10^{-4}

TABLE 5.1
Domain Size Versus Validated Error Bounds in Lorenz

Dim	Domain Size ν	s	Error δ	$\widetilde{C}$	$C_n(s)$	$K_n(s)$	(sec)
3	3.162×10^{-8}	10^{-2}	4.48×10^{-18}	59.14	6.11	9.28	0.31
6	3.162×10^{-8}	10^{-2}	1.869×10^{-16}	159	21.58	37.68	0.51
12	3.162×10^{-8}	10^{-3}	1.02×10^{-14}	792	81.79	156	1.66
25	7.07×10^{-9}	10^{-3}	2.95×10^{-14}	4571	328	634	13.1
50	7.07×10^{-9}	10^{-3}	1.15×10^{-12}	$20,296$	1298	2568	108
100	7.07×10^{-10}	10^{-5}	4.12×10^{-13}	$87,498$	5052	$10,001$	882

TABLE 5.2
Validated error bounds for the delayed Hénon map in various dimensions. Linear approximation of the co-dimension one stable manifold (so if the dimension of the phase space is n then the dimension of the stable manifold is $n-1$). We remark that the constant $\widetilde{C}$ is problem dependent, so that column 5 provides information relevant only to the delayed Hénon system. On the other hand the constants C_n and K_n are dimension dependent.

with parameters $a, b \in \mathbb{R}$. In [69] numerical evidence is provided for the conjecture that the map exhibits chaotic behavior for dimensions $n = 2, \ldots, 100$ when $a = 1.6$ and $b = 0.1$. We use these parameters throughout, and obtain validated bounds for the linear approximation of the stable manifold in various dimensions.

The map has fixed points at $p_\pm = (x_\pm, \ldots, x_\pm) \in \mathbb{R}^n$, where

$$x_\pm = \frac{b - 1 \pm \sqrt{(1-b)^2 + 4a}}{2a}.$$

The fixed point p_+ has one dimensional unstable manifold and co-dimension one stable manifold for every dimension. We compute the validated bounds for the linear approximation of the stable manifold in a number of dimensions. The results are recorded in Table 5.2.

REMARKS 5.1.

(I) (**Eigenvector Rescaling**) For the computations on the delayed Hénon map we rescale the stable eigenvectors as

$$A = 0.001\,[\xi_2| \ldots |\xi_n]$$

for the linear part of the parameterization, where ξ_j are the eigenvectors scaled to length one. This has the effect of allowing a larger parameter radius ν for a given error bound δ. Or, thinking of it another way, this rescales the parameterization to have a smaller derivative. The rescaling is useful when computing the connecting orbits in Section 7, as we are able to obtain validated bounds on the derivative in a larger neighborhood. (More precisely, in the terms of Theorem 7.2, this allows us to find a larger r_κ, making the condition given by Equation (7.5) easier to fulfill).

(II) (**Curse of Dimension**) Note that in order to maintain an accurate approximation as the dimension increases it is necessary to decrease the size of the validated domain for the linear approximation. Decreasing s helps a little as well. However the real problem is the growth of the constants $C_n(0)$ and $K_n(0)$. In addition, for the delayed Hénon map the stable eigenvalues get closer and closer to resonant as the dimension increases. This leads to poor scaling of the constant $\tilde{C}$ as dimension increases.

(III) (**Poor Cauchy Bounds**) One way to obtain bounds on the derivative of the truncation error is by applying the Cauchy Bounds of Lemma 2.2. Beginning with the decomposition

$$P(\theta) = p_0 + A\theta + h(\theta),$$

we utilize the fact that $h(\theta)$ is analytic with

$$\sup_{\|z\| \leq \nu} \|h(z)\| \leq \delta.$$

However we note that in higher dimensions this often results in very poor bounds on the derivative. For example if we consider the delayed Hénon system in 25 dimensions and take a "loss of domain parameter" $\sigma = 0.5$ then, after consulting Lemma 2.2, we have the bound

$$\sup_{\|z\| \leq e^{-0.5}\nu} \|D^2 h(z)\| \leq \frac{4\pi^2 k^2}{\sigma^2 \nu^2}\delta = \frac{4\pi^2 24^2}{(0.5)^2 (7.07 \times 10^{-9})^2} 2.95 \times 10^{-14} \approx 5.36 \times 10^7.$$

Here $k = 24$ is the dimension of the stable manifold and the rest of the constants for the Cauchy Bound are simply read off of Table 5.2. Since $e^{-1/2}\nu = 4.28 \times 10^{-10}$, we have a bound on the second derivative for all z in a polydisk of this radius. However the bound itself is quite poor.

(IV) (**Even More on Poor Cauchy Bounds**) The reason for generally poor Cauchy Bounds when applied to the linear approximation can be seen in yet another way. Note that the Cauchy Bounds give

$$\|D^2 h\| \approx \frac{C_1}{\nu^2}\delta,$$

but $\delta \approx C_2 \nu^2$ for the linear approximation (i.e. the truncation error associated with the approximation of P by it's tangent space has quadratic accuracy). After the cancelation of δ with ν^2 in the denominator we are left with a bound on the second derivative which is like $C_1 C_2$, the dimension dependent constants. In higher dimensions we should expect that this leads to very poor Cauchy bounds, more or less independent of the dynamical system under consideration. This suggests the need for improved derivative bounds on P, which is the topic addressed in Section 6.

6. Linear Approximation of the Partial Derivatives of the Parameterization and Computer Assisted Error Bounds. In this section we develop a method which provides much better bounds on the first derivative of the parameterization, but which in turn requires more work than the classical Cauchy bounds. We develop these improved bounds only for the case of maps (discrete time dynamical systems) as this is what we require in the sequel. However one could establish similar results for differential equations. The idea behind the improved bounds is to compute, using rigorous numerics, some information about the second order terms of the power series of the parameterization. These second order terms need not be stored, or manipulated. Instead they are used in order to study the functional equation for the first partial derivatives of the parameterization. Again, we only approximate the partial derivatives to first order, hence the need for information about the second order coefficients of the parameterization itself.

We note that higher order derivatives could be estimated by similar methods. However for the purposes of the present work it is enough that we obtain good bounds only for the first derivatives. Second derivatives are then bound applying Cauchy Bounds to the first derivatives, a strategy which provides enough control for the applications considered in Section 7.

6.1. Functional Equation and Estimates. Suppose that $P \in \mathcal{W}_\nu^{k,n}$ is a solution of Equation (2.1) and let $R_i \colon \mathbb{D}_0^k(\nu) \to \mathbb{C}^n$ be given by

$$\frac{\partial}{\partial \theta_i} P(\theta) = R_i(\theta) = \sum_{|\alpha|=0}^{\infty} r_\alpha \theta^\alpha,$$

where for the moment we consider r_α unknown (of course the coefficients r_α are determined by the coefficients of P). Taking the partial derivative of Equation (2.1) with respect to θ_i for $1 \leq i \leq k$ gives that R_i is a solution of the functional equation

$$Df[P(\theta)]R_i(\theta) - \lambda_i\, R_i[\Lambda\theta] = 0. \tag{6.1}$$

We note that this is a linear operator equation for R_i, which is the i-th partial derivative of P.

Let $P(\theta) = p_0 + A\theta + h(\theta)$ where $h \in \mathcal{H}_\nu^{k,n}$ with $\|h\|_\nu \leq \delta$. We denote the power series for P by

$$P(\theta) = \sum_{|\alpha|=0}^{\infty} p_\alpha \theta^\alpha.$$

Our goal is to find an $n \times k$ matrix B_i, a function $\tilde{h}_i \in \mathcal{H}_\nu^{k,n}$, and a number δ_i so that

$$R_i(\theta) = \xi_i + B_i\theta + \tilde{h}_i(\theta)$$

and $\|\tilde{h}_i\|_\nu \leq \delta_i < \infty$.

First we denote the columns of the unknown matrix by $B_i = [b_1| \ldots |b_k]$ and note that

$$b_j = \frac{\partial}{\partial \theta_j} R_i(0) = \frac{\partial^2}{\partial \theta_j \partial \theta_i} P(0) = \begin{cases} p_{e_i+e_j} & \text{if} \quad i \neq j \\ 2p_{2e_i} & \text{if} \quad i = j \end{cases} \tag{6.2}$$

as can be seen by differentiating the series expansion of P. Here $e_j = (0, \ldots, 1, \ldots, 0)$ is the vector with one in the j-th component and zeros elsewhere.

Remark 6.1. The expression for b_i given by Equation (6.2) is valuable as the power series coefficients for P can be computed efficiently using the power matching techniques for the formal series developed in [54]. For example when studying the delayed Hénon system we have that the second order coefficients for P are given by

$$p_{e_i+e_j} = Q[\Sigma - \lambda_i\lambda_j \mathrm{Id}]^{-1} Q^{-1} s_{e_i+e_j},$$

where

$$s_{e_i+e_j} = \begin{pmatrix} a\xi_i^1 \xi_j^1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Then

$$B_i = 2[p_{e_i+e_1}|\dots|p_{e_i+e_k}]. \tag{6.3}$$

Remark 6.2. While the explicit formula of the previous remark is useful for numerical computations, it will benefit us in a moment to have a more general (albeit implicit) expression for the matrix B_i. The computation is based on the observation that

$$B_i = D_\theta R_i(0).$$

We begin by expanding

$$Df[P(\theta)] = Df[p_0 + A\theta + h(\theta)] = Df[p_0] + D^2 f(p_0)[A\theta + h(\theta)] + \tilde{N}(A\theta + h(\theta)),$$

where $\tilde{N}$ is the second order Taylor remainder for Df. Inserting this expression into Equation (6.1) gives

$$\left[Df(p_0) + D^2 f(p_0)A\theta + D^2 f(p_0)h(\theta) + \tilde{N}(A\theta + h(\theta))\right] R_i(\theta) - \lambda_j R_i[\Lambda\theta] = 0.$$

Taking the total derivative of this expression with respect to θ and evaluating at $\theta = 0$ gives that the matrix B_i is the solution of the matrix equation

$$Df(p_0)B_i - \lambda_i B_i \Lambda = -D^2 f(p_0)A\xi_i. \tag{6.4}$$

This expression provides useful cancelations in the argument to follow.

We now return to the linear approximation of R_i and see that Equation (6.1) is

$$Df[P(\theta)]R_i(\theta) =$$

$$\left[Df(p_0) + D^2 f(p_0)A\theta + D^2 f(p_0)h(\theta) + \tilde{N}(A\theta + h(\theta))\right] \left(\xi_i + B_i\theta + \tilde{h}_i(\theta)\right),$$

on the left and

$$\lambda_i R_i[\Lambda\theta] = \lambda_i \xi_i + \lambda_i\, B_i \Lambda\theta + \lambda_i\, \tilde{h}_i(\Lambda\theta),$$

on the right. Expanding the products and exploiting the cancelations given by the eigenvalue/eigenvector relationship

$$Df(p_0)\xi_i - \lambda_i \xi_i = 0,$$

and the relationship

$$Df(p_0)B_i\theta - \lambda_i B_i \Lambda\theta + D^2 f(p_0)(A\theta, \xi_i) = 0,$$

given by Equation (6.4), we collect terms in involving $\tilde{h}_i$ on the left and have

$$Df(p_0)\tilde{h}_i(\theta) - \lambda_i\tilde{h}_i(\Lambda\theta) + \bar{N}[A\theta + h(\theta)]\tilde{h}_i(\theta) = G(\theta). \tag{6.5}$$

Here

$$\bar{N}[A\theta + h(\theta)] = D^2 f(p_0)[A\theta + h(\theta)] + \tilde{N}[A\theta + h(\theta)],$$

and

$$G(\theta) = -D^2 f(p_0)(h(\theta), \xi_i) - \bar{N}[A\theta + h(\theta)]B_i\theta - \tilde{N}[A\theta + h(\theta)]\xi_i. \tag{6.6}$$

Proceeding as in the proof of Lemma 3.3 we define the linear operator

$$\mathfrak{L}_i[q](\theta) \equiv Df(p_0)q(\theta) - \lambda_i q[\Lambda\theta].$$

Assuming for the moment that $\mathfrak{L}_i$ is invertible, we define the linear operator

$$C[\tilde{h}_i](\theta) = \mathfrak{L}_i^{-1}\left[\bar{N}(A\theta + h(\theta))\tilde{h}_i(\theta)\right],$$

and rewrite Equation (6.5) as

$$\mathfrak{L}_i\left[\tilde{h}_i(\theta) - C[\tilde{h}_i](\theta)\right] = G(\theta).$$

We use the Neumann Theorem in order to solve and obtain the formula

$$\tilde{h}_i(\theta) = [\mathrm{Id} - C]^{-1}\,\mathfrak{L}_i^{-1}[G](\theta),$$

(where of course we must assume that $\|C\|_{\mathcal{H}_\nu^{k,n}} < 1$). The next lemma shows that $\mathfrak{L}_i$ is invertible under the same non-resonance assumptions of Lemma 3.3

LEMMA 6.3. *If the stable eigenvalues $\lambda_1, \ldots, \lambda_k$ for the differential $Df(p_0)$ are non-resonant, then the linear operator defined by*

$$\mathfrak{L}_i[q](\theta) = Df(p_0)q(\theta) - \lambda_i q[\Lambda\theta],$$

is boundedly invertible on $\mathcal{H}_\nu^{k,n}$. Moreover we have the bound

$$\left\|\mathfrak{L}_i^{-1}\right\|_{\mathcal{H}_\nu^{k,n}} \le \tilde{C},$$

where $\tilde{C}$ is the constant from Lemma 3.3.

Proof. As in the proof of Lemma 3.3 we let $p \in \mathcal{H}_\nu^{k.n}$ be given by

$$p(\theta) = \sum_{|\alpha|=2}^{\infty} p_\alpha \theta^\alpha,$$

and consider the equation

$$\mathfrak{L}_i[q](\theta) = Df(p_0)q(\theta) - \lambda_i q[\Lambda\theta] = p(\theta).$$

Making the substitution $q(\theta) = Qw(\theta)$ as in the proof of Lemma 3.3 and matching like powers of θ gives that the power series coefficients of $q(\theta)$ are (at least formally) given by

$$q_\alpha = Q\left[\Sigma - \lambda_1^{\alpha_1}\cdot\ldots\cdot\lambda_i^{\alpha_i+1}\cdot\ldots\cdot\lambda_k^{\alpha_k}\mathrm{Id}\right]^{-1}Q^{-1}p_\alpha.$$

Since the stable eigenvalues are assumed to be non-resonant we have that these coefficients are well defined to all orders. The desired bound on the inverse of the linear operator is then obtained by observing that for each $\alpha \geq 2$ we have

$$\begin{aligned}
\left\| \left[\Sigma - \lambda_1^{\alpha_1} \cdot \ldots \cdot \lambda_i^{\alpha_i+1} \cdot \ldots \cdot \lambda_k^{\alpha_k} \mathrm{Id}\right]^{-1} \right\| &= \max_{1 \leq j \leq n} \frac{1}{|\lambda_j - \lambda_1^{\alpha_1} \cdot \ldots \cdot \lambda_i^{\alpha_i+1} \cdot \ldots \cdot \lambda_k^{\alpha_k}|} \\
&\leq \sup_{2 \leq |\alpha|} \max_{1 \leq j \leq n} \frac{1}{|\lambda_j - \lambda_1^{\alpha_1} \cdot \ldots \cdot \lambda_k^{\alpha_k}|} \\
&\leq \tilde{C}.
\end{aligned}$$

□

The remainder of the discussion is simplified somewhat once we define the following terms. Suppose that

$$Df(p_0 + z) = Df(p_0) + D^2 f(p_0) z + \tilde{N}(z),$$

where $\tilde{N} \colon \mathbb{D}_0^n(\rho) \to \mathbb{C}^n$ is the second order bounded analytic Taylor remainder associated with $Df(p_0 + z)$ on $\mathbb{D}_0^n(\rho)$. In order to simplify the statement of our results we also let $\bar{N}(z) = D^2 f(p_0) z + \tilde{N}(z)$. Suppose that $P \in \mathcal{W}_\nu^{k,n}$ is given by

$$P(\theta) = p_0 + A\theta + h(\theta),$$

and that

$$f[P(\theta)] = P[\Lambda\theta],$$

with $h \in \mathcal{H}_\nu^{k,n}$ having

$$\|h\|_\nu \leq \delta.$$

The following definition collects the constants necessary in order to obtain validated bounds on the partial derivatives of P.

Definition 6.4 (Validation Values for Partial Derivatives of the Truncation Error Function Associated with the Linear Approximation of a Stable Manifold). A collection of positive constants $\nu, \delta, \rho, \tilde{C}, C_{\tilde{N}}, C_G$, and $C_{\bar{N}}$ are called *validation values for the partial derivative problem* if $\tilde{C}$ is as in Definition 3.1,

$$\tilde{N}(z) \leq C_{\tilde{N}} \|z\|^2, \tag{6.7}$$

$$\bar{N}(z) \leq C_{\bar{N}} \|z\|, \tag{6.8}$$

for all $\|z\| < \rho$, and

$$\|D^2 f(p_0)\| \|\xi_i\| \delta + C_{\bar{N}} (\|A\|\nu + \delta) \|B_i\| \nu + C_{\tilde{N}} \|\xi_i\| (\|A\|\nu + \delta)^2 \leq C_G. \tag{6.9}$$

The proof of the following theorem is contained in the preceding discussion.

Theorem 6.5. *Suppose that $\delta, \nu, \rho, \tilde{C}, C_{\tilde{N}}, C_{\bar{N}}$, and C_G are validation values in the sense of Definition 6.4. Suppose that*

$$\|A\|\nu + \delta < \rho, \tag{6.10}$$

$$\tilde{C} C_{\tilde{N}}(\|A\|\nu + \delta) < 1, \tag{6.11}$$

and assume that $\delta_i > 0$ is a positive constant with

$$\frac{\tilde{C}}{1 - \tilde{C} C_{\tilde{N}}(\|A\|\nu + \delta)} C_G \leq \delta_i. \tag{6.12}$$

Then there exists a unique $\tilde{h}_i \in \mathcal{H}_\nu^{k,n}$ so that

$$\frac{\partial}{\partial \theta_i} P(\theta) = \xi_i + B_i \theta + \tilde{h}_i(\theta).$$

Moreover

$$\left\| \tilde{h}_i \right\|_\nu \leq \delta_i.$$

Remark 6.6. Observe that the quantity on the left hand side of Equation (6.9) is quadratic in ν, at least when we think of δ as being quadratic in ν. Then Theorem 6.5 provides bounds on the error associated with the linear approximation of the partial derivatives of P which are (as one expects) quadratic in the domain size ν.

6.2. Numerical Examples of Derivative Bounds. Using Theorem 6.5 we are able to obtain substantially improved bounds on the size of the first and second derivatives of the truncation error h associated with the linear approximation of the stable manifold. In order to illustrate the method we compute bounds for the approximations discussed in Section 5.2 and tabulated in Table 5.2.

The computations using Theorem 6.5 applied to the delayed Hénon map are reported in Table 6.2. We note that while the second derivative bounds do grow with dimension, the loss of order is nowhere near as aggressive as discussed in Section 5.2. Note for example that in dimension 25 we obtain bounds on the second derivative of roughly size 6, as opposed to size roughly 10^7 as obtained when using Cauchy Bounds alone.

Another advantage of this method is that the bounds on the first derivative are valid on the same polydisk as the bounds on the size of the truncation error itself. We only give up domain in order to bound the second derivatives, but even here we only have to deal with a factor of ν rather than ν^2 in the denominator of the estimates (i.e we bound $D^2 h$ by applying the first order Cauchy bounds to the $R_i = \partial_i h$).

Note that by dimension 50 we are loosing control on the bounds on the second derivative. Again we remark that if needed the second order bounds could be improved by developing an analogue of Theorem 6.5 for second order mixed partials of P. This is not a project we undertake here as in Section 7 we consider the delayed Hénon system in dimension only up to 20.

7. Existence of Connecting Orbits by Method of Projected Boundary Conditions. In order to formalize the discussion of connecting orbits we establish the following notation.

P1 Assume $f\colon \mathbb{R}^n \to \mathbb{R}^n$ 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 respectively.

P2 Assume $U_s \subset \mathbb{R}^{n_s}$ is an open neighborhood of the origin and that $P\colon U_s \to \mathbb{R}^n$ is a parameterization of $W^s_{\text{loc}}(p)$.

Dimension	Domain Size ν	Validated bound on $\|Dh\|_\nu$	$\|D^2h\|_{e^{-\sigma}\nu}$	(sec)
3	3.162×10^{-8}	1.44×10^{-14}	9.07×10^{-7}	0.09
6	3.162×10^{-8}	1.75×10^{-13}	3.2×10^{-5}	0.41
12	3.162×10^{-8}	6.25×10^{-13}	1.05×10^{-2}	1.73
25	7.07×10^{-9}	4.993×10^{-12}	5.86	10.4
50	7.07×10^{-9}	1.22×10^{-9}	$6,393$	93.3

TABLE 6.1

Validated error bounds for the partial derivatives of the parameterization for the delayed Hénon map in various dimensions. For all second derivative computations we take $\sigma = 0.8$ and apply the Cauchy Bounds of Lemma 2.2. Then the second derivative bounds hold on a polydisk of radius roughly 0.449ν in every case. The remaining parameters are set as in the computations reported in Table 5.2

P3 Assume $U_u \subset \mathbb{R}^{n_u}$ is an open neighborhood of the origin and that $Q \colon U_u \to \mathbb{R}^n$ is a parameterization of $W^u_{\text{loc}}(q)$.

We define the *connecting orbit operator* to be the function $F \colon \mathbb{R}^{n(k+1)} \to \mathbb{R}^{n(k+1)}$ given by

$$F(\theta, x, \phi) = \begin{bmatrix} f^{-1}(x_1) & - & Q(\theta) \\ f^{-1}(x_2) & - & x_1 \\ \vdots & & \\ f^{-1}(x_j) & - & x_{j-1} \\ f(x_j) & - & x_{j+1} \\ \vdots & & \\ f(x_{k-1}) & - & x_k \\ f(x_k) & - & P(\phi) \end{bmatrix} \tag{7.1}$$

where $x = (x_1, \ldots, x_k) \in \mathbb{R}^{nk}$, $\theta \in \mathbb{R}^{n_u}$, and $\phi \in \mathbb{R}^{n_s}$. Observe that if $(\theta_*, x_*, \phi_*) \in \mathbb{R}^{n(k+1)}$ is a zero of F then for example the orbit of $Q(\theta_*)$ is a heteroclinic connection for f from q to p.

REMARK 7.1 (Formulation of the Connecting Orbit Equation). There are many variations of the connecting orbit operator defined above. We chose to use f as well as f^{-1} in the definition of the operator F because the inverse map has a stabilizing effect on the numerics. In the applications to the delayed Hénon mapping considered in the present paper one can work out the explicit formula for f^{-1} and its derivative, so that the presence of the inverse presents no problem.

In other applications the inverse map may be inconvenient to work with and instead we may consider the equivalent system

$$G(\theta, x, \phi) = \begin{bmatrix} Q(\theta) & - & x_1 \\ f(x_1) & - & x_2 \\ & \vdots & \\ f(x_k) & - & P(\phi) \end{bmatrix}.$$

In fact the operator G can be studied via a Newton method even in the case that f is not invertible, as long as $Df(x_j)$ is invertible for $1 \leq j \leq k$, i.e. as long as the connecting orbit does not intersect the singularity set of f. Additional constraints are necessary if one wishes to study orbits which intersect the singularity set, or orbits which arise as non-transverse intersections of the stable and unstable manifolds.

The following theorem considers the map $F(\theta, x, \phi)$ given by Equation (7.1) in a neighborhood of an approximate solution $(\theta_0, x_0, \phi_0) \in \mathbb{R}^{n(k+1)}$. Let $\|\phi\|_{n_s}, \|\theta\|_{n_u}, \|x_i\|_n$ denote respectively the norms on $\mathbb{R}^{n_s}$, $\mathbb{R}^{n_u}$, and $\mathbb{R}^n$. For $x = (x_1, \ldots, x_k) \in \mathbb{R}^{nk}$ let $\|x\|_{nk} := \max_{1 \leq i \leq k} \|x_i\|_n$. Then for any $(\theta, x, \phi) \in \mathbb{R}^{n(k+1)}$ we employ the norm

$$\|(\theta, x, \phi)\| := \max\left(\|\theta\|_{n_u}, \|\phi\|_{n_s}, \|x\|_{nk}\right).$$

For $(\theta_0, x_0, \phi_0) \in \mathbb{R}^{n(k+1)}$ and $r > 0$ this norm induces the ball

$$B_{(\theta_0, x_0, \phi_0)}(r) := \{(\theta, x, \phi) \in \mathbb{R}^{n(k+1)} \, : \, \|(\theta_0, x_0, \phi_0) - (\theta, x, \phi)\| < r\}.$$

The following theorem is used to establish the existence of a zero of F near the approximate solution (θ_0, x_0, ϕ_0).

Theorem 7.2 (Existence of a Transverse Connecting Orbit). *Let $\theta_0 \in \mathbb{R}^{n_u}$, $(x_0^1, \ldots, x_0^k) = x_0 \in \mathbb{R}^{nk}$, $\phi_0 \in \mathbb{R}^{n_u}$, and A be an invertible $n(k+1) \times n(k+1)$ matrix. Suppose that $\hat{\varepsilon}$, $\hat{\delta}$, r_κ, and κ are positive constants with*

$$\|AF(\theta_0, x_0, \phi_0)\| \leq \hat{\varepsilon} \tag{7.2}$$

$$\|Id - ADF(\theta_0, x_0, \phi_0)\| \leq \hat{\delta} \tag{7.3}$$

$$\sup_{(\theta, x, \phi) \in B_{(\theta_0, x_0, \phi_0)}(r)} \|D^2 F(\theta, x, \phi)\| \leq \kappa. \tag{7.4}$$

Then for any $r > 0$ with

$$r \leq r_\kappa, \tag{7.5}$$

$$\|A\|\kappa r^2 - r\left(1 - \hat{\delta}\right) + \hat{\varepsilon} \leq 0, \tag{7.6}$$

and

$$\hat{\delta} + \|A\|\kappa r < 1, \tag{7.7}$$

there is a unique $(\hat{\theta}, \hat{x}, \hat{\phi}) \in B_{(\theta_0, x_0, \phi_0)}(r)$ so that $F(\hat{\theta}, \hat{x}, \hat{\phi}) = 0$. The orbit under f of any of the points $P(\hat{\theta})$, $\hat{x}_1, \ldots, \hat{x}_k$, or $Q(\hat{\phi})$ is a transverse heteroclinic orbit from p to q. (Each of these gives rise to the same connecting orbit. If $p = q$ the orbit is a transverse homoclinic).

For a more complete exposition of the theorem we refer to [54]. Here we make the following remarks.

Remarks 7.3.

(I) The maps P and Q are rarely known explicitly. Instead we usually have polynomials P_{N_u}, Q_{N_s} and analytic tails $h_u \colon \mathbb{D}_0^{n_u}(\nu_u) \to \mathbb{C}^n$ and $h_s \colon \mathbb{D}_0^{n_s}(\nu_s) \to \mathbb{C}^n$ so that

$$P(\theta) = P_{N_u}(\theta) + h_u(\theta) \quad \text{ for all } \quad \theta \in \mathbb{D}_0^{n_u}(\nu_u),$$

and

$$Q(\phi) = Q_{N_s}(\phi) + h_s(\phi) \quad \text{ for all } \quad \phi \in \mathbb{D}_0^{n_s}(\nu^s).$$

Moreover the functions h_u, h_s are not known explicitly, however there are $\delta_u, \delta_s > 0$ (usually very small) so that

$$\sup_{\|\theta\| \leq \nu_u} \|h_u(\theta)\| \leq \delta_u \quad \text{and} \quad \sup_{\|\phi\| \leq \nu_s} \|h_s(\phi)\| \leq \delta_s.$$

The polynomials P_{N_u} and Q_{N_s} and rigorous bounds δ_s, δ_u are computed using the methods of [54] when high order approximations are desired, and with the methods of the present work when the linear approximation is used.

(II) These unknown errors δ_u, δ_s must be propagated through the argument. In other words F must be written as

$$F(\theta, x, \phi) = F_{N_u, N_s}(\theta, x, \phi) + H(\theta, \phi),$$

where F_{N_u, N_s} is the maps given by Equation (7.1) with P and Q replaced by P_{N_u} and Q_{N_s} and

$$H(\theta, \phi) = \begin{bmatrix} h_u(\theta) \\ 0 \\ \vdots \\ 0 \\ h_s(\phi) \end{bmatrix}.$$

Then (θ_0, x_0, ϕ_0) is obtained by numerically solving the equation $F_{N_u, N_s}(\theta, x, \phi) \approx 0$, usually with a numerical Newton scheme.

(III) The constants $\hat{\varepsilon}$, $\hat{\delta}$ and $\hat{\kappa}$ depend on the known bounds in the truncation errors. For example $\hat{\varepsilon}$ should be a bound on

$$\|AF(\theta_0, x_0, \phi_0)\| \leq \|AF_{N_u, N_s}(\theta_0, x_0, \phi_0)\| + \|A\|\delta_u + \|A\|\delta_s \leq \hat{\varepsilon}.$$

(IV) Similar considerations allow us to obtain appropriate constraints for $\hat{\delta}$ and κ, however these require bounds on $\|Dh_u\|$, $\|Dh_s\|$, as well as on $\|D^2 h_u\|$ and $\|D^2 h_s\|$. When using the high order techniques of [54] we can bound both the first and second derivatives of the truncation errors using the Cauchy bounds of Lemma 2.2. When the stable or unstable manifolds are approximated using the low order approximation of the present work it is highly advantageous to use the methods of Section 6 in order to bound the derivative of the truncation error, and then only apply the Cauchy bounds of Lemma 2.2 to the partial derivative bounds in order to bound the second derivatives.

(V) It is shown in [54] that if $(\hat{\theta}, \hat{x}, \hat{\phi})$ is a zero of Equation (7.1) for which $DF(\hat{\theta}, \hat{x}, \hat{\phi})$ is invertible then the resulting connecting orbit is transverse (see Corollary 5.3 of that reference). In the present work since we are using Theorem 2.12 in order to prove the existence of a zero, we use Lemma 2.14 in order to get that the non-degeneracy of the zero and hence the transversality of the connecting orbit. Then the whole argument of [54] goes through with A only an approximate inverse of $DF(\theta_0, x_0, \phi_0)$.

(VI) Since A is only required to be an approximate inverse, the quantity $\|\mathrm{Id} - ADF(\theta_0, x_0, \phi_0)\|$ must be bound rigorously. This is done by considering

$$\|\mathrm{Id} - ADF(\theta_0, x_0, \phi_0)\| \leq \|\mathrm{Id} - ADF_{N_s, N_u}(\theta_0, x_0, \phi_0)\| + \|A\|\|DH(\theta_0, \phi_0)\|,$$

where $\|DH(\theta_0, \phi_0)\| \leq \max(\|Dh_u\|_{\nu_u}, \|Dh_s\|_{\nu_s})$, and the latter bounds are obtained using the ideas from (IV) of the present remark. We also note that the computation of $\|\mathrm{Id} - ADF_{N_u,N_s}(\theta_0, x_0, \phi_0)\|$ requires that the interval matrix $DF_{N_u,N_s}(\theta_0, x_0, \phi_0)$ is multiplied by the floating point matrix A giving an interval enclosure of $ADF_{N_u,N_s}(\theta_0, x_0, \phi_0)$. The subtraction $\mathrm{Id} - ADF_{N_u,N_s}(\theta_0, x_0, \phi_0)$ is then performed in interval arithmetic, as is the resulting matrix norm. When the dimension of the problem is high the matrix-matrix multiplication is then the most expensive portion of the proof. (Both of the matrices A and $DF_{N_u,N_s}(\theta_0, x_0, \phi_0)$ are $n(k+1) \times n(k+1)$). However this multiplication is still cheaper than the matrix inversion required in [54].

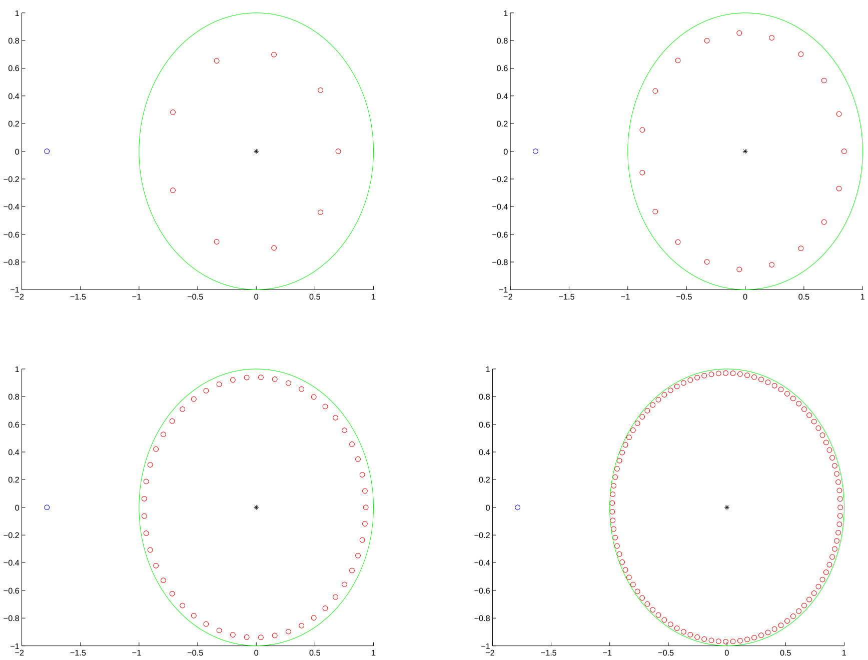

FIG. 8.1. ***Spectrum of** $Df(p_0)$ **for the delayed Hénon map**: The figure illustrates the configuration of the spectrum of the map in dimension* 10 *(top left),* 20 *(top right),* 50 *(bottom left), and* 100 *(bottom right). The dominant feature is the accumulation of the spectrum on the unit circle as the dimension of the system increases. This accounts for the slow approach of the connecting orbits along the stable manifold when the systems dimension is large.*

8. Computer Assisted Proof of Transverse Homoclinic Connecting Orbits/Topological Horseshoes in Higher Dimensions: a case study of the delayed Hénon Mapping. We have implemented the algorithm suggested by Theorem 7.2 in order to prove the existence of some connecting orbits for the delayed Hénon system. Presently we report the results of a number validated computations in dimensions three and 20. In both cases we use the methods of [54] in order to compute the one dimensional unstable manifold (using a polynomial approximation of order 25) with validated error bounds. We apply the methods of the present work in order to compute rigorous error bounds associated with the linear approximation of the co-dimension one stable manifold. There are two difficulties which arise as we change the dimension of the computation. The first difficulty, which is completely general,

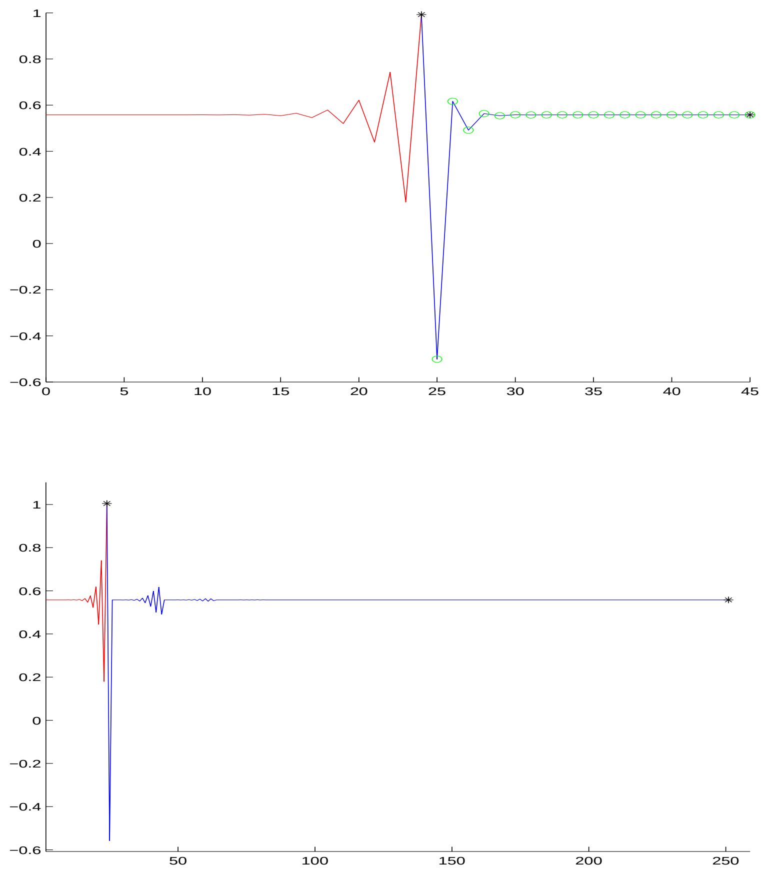

Fig. 8.2. ***Time Series Representation of the Connecting Orbits:*** *The figure illustrates, in dimension* 3 *and* 20*, the orbit segment which is used in the computer assisted proof of the existence of the transverse connecting orbit. In both proofs the blue segment illustrates the first component of* x_0*, or the "flight" of the orbit from the local unstable manifold to the local stable. The black dots in each frame illustrate the boundary conditions of the orbit. The red portion of the orbit is on the local stable manifold, and is only included in the figure in order to illustrate that the orbit is homoclinic. The "error bars" on the figures, i.e. the distance from the orbit shown to the actual homoclinic orbit, have width less then* 2.5×10^{-9} *in both cases.*

is the growth of the dimension dependent constants. This makes the conditions of Theorem 7.2 harder and harder to meet as the dimension increases.

However there is an even greater, system dependent difficulty which arises for the delayed Hénon map, and this is illustrated in Figure 8.1. The fact is that as dimension is increased for the delayed Hénon map the stable spectrum of $Df(p_0)$ accumulates on the unit circle. This means that the local dynamics get slower and slower, and that it takes a larger and larger number of iterates in order to get an orbit segment into a small enough neighborhood of the origin that our rigorous error bounds apply. This is illustrated in Figure 8.2, which compares the connecting orbit segments used in the computer assisted proofs in dimension 3 and 20. Figure 8.3 illustrates the "bursting" dynamics of connecting orbits in dimension 10 and 20.

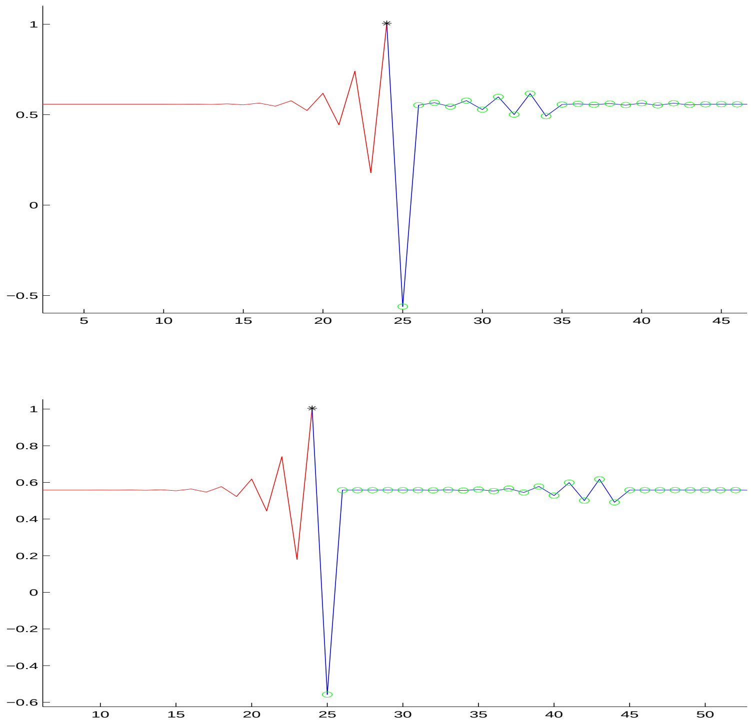

FIG. 8.3. ***Effect of Dimension on Homoclinic Dynamics**: The figure shows a closeup on the first component of the homoclinic orbit in dimension ten and twenty (then the bottom figure is a close up on the bottom frame of Figure* 8.2. *Observe that increasing the delay (which is actually the same thing as increasing the dimension) introduces "bursting" in the asymptotic behavior of the orbit. The calm periods between these bursts is longer and longer as dimension increases. (Compare also with the upper graph in Figure* 8.2 *for the dynamics in three dimensions).*

REMARKS 8.1.

(I) **A Few Comments About the Twenty Dimensional Proof**: The largest proof in the series is the proof of the transverse connecting orbit in 20 dimensions. This orbit lies in the transverse intersection of the one dimensional unstable and nineteen dimensional stable manifold. We remark that the "time of flight" for the numerical connection (i.e. the number of iterates required for the orbit to transition from the image of the unstable parameterization to the image of the stable parameterization) is 220. Since the dimension of the phase space is 20 the result is that the connecting orbit operator F is a map in 4560 dimensions. Interval arithmetic computations of the essential parameters in the proof come out to be

$$\hat{\epsilon} = 2.34 \times 10^{-9}, \qquad \hat{\delta} = 1.14 \times 10^{-10},$$

$$r_\kappa = 7.36 \times 10^{-9}, \qquad \kappa = 41.91,$$

and $\|A\| = 8,679$. We find that

$$r = 2.355 \times 10^{-9},$$

satisfies the hypotheses given by Equation (7.5), (7.6), and (7.5) of Theorem 7.2. Then there exists a true transverse connecting orbit within an r-neighborhood of the orbit shown in the lower frame of Figure 8.

(II) **Possible Future Improvements on the Numerical Implementation**: As noted in Remark 7.3 (VII), the computation of

$$\|\mathrm{Id} - ADF(\theta_0, x_0, \phi_0)\|,$$

is the most expensive portion of the numerical computation. In fact even the storage of 4560×4560 interval matrices is substantial drain on the resources of a standard laptop. However the current implementation does not take into account the fact that DF is a structured/banded matrix (see [54] for the explicit formula). In fact of the $4560^2 = 20,793,600$ entries of $DF(\theta_0, x_0, \phi_0)$ only 9699, or roughly five percent, of the entries are non-zero. The performance of the computation would benefit substantially from storage and multiplication algorithms which utilized this structure. However the computations discussed in this section are only meant to illustrate the fact that the linear theory developed in the present work enables validated computation of connecting orbits in higher dimensions. The implementation developed so far is offered as a proof of concept.

(IV) **Co-dimension one stable manifold:** the proof does not exploit the fact that the stable manifold is co-dimension one, and hence the stable manifold could be represented as the graph of a real valued function over the stable eigenspace. One reason for neglecting this fact is simply to illustrate the procedure when one manifold is "high dimensional" and the other is "low-dimensional". The fact that the low-dimensional manifold is in this case one dimensional is not essential. A second reason for employing the Parameterization Method rather than exploiting methods based on a "graph transformation" is that the graph transform contains terms requiring the composition of the unknown function with itself. The Parameterization Method on the other hand requires only the composition of the unknown function with the known dynamics f. The composition with f is sometimes easier to compute with.

9. Acknowledgments. The Author would like to thank an anonymous referee for carefully reading the first submitted version of this manuscript. The final work was greatly improved by the suggestions and comments of this referee. Thanks also to Rafael de la Llave for several illuminating discussions in the summer of 2012 at the DyToComp conference in Bedlewo Poland. These conversations were instrumental to the formulation of the present work. Thanks to J.P. Lessard, J.B. van den Berg, and Konstantin Mischaikow for reading earlier versions of the manuscript and making comments, corrections, and suggestions which informed and influenced the final product. The author was partially supported by a National Science Foundation Grant DSM 1318172.

REFERENCES

[1] L. V. Ahlfors. Complex Analysis. An Introduction to the Theory of Analytic Functions of One Complex Variable. Third Edition. International Series in Pure and Applied Mathematics. *McGraw-Hill Book Co.*, New York, 1978.

[2] Z. Arai, and K. Mischaikow. Rigorous computations of homoclinic tangencies. *SIAM J. Appl. Dyn. Syst.* 5 (2006), no. 2, 280-292.

[3] D. Ambrosi, G. Arioli, and H. Koch. A Homoclinic Solution for Excitation Waves on a Contractile Substratum. SIAM Journal on Applied Dynamical Systems, Vol. 11, No. 4, pp. 1533-1542, 2012.

[4] G. Arioli. Periodic Orbits, Symbolic Dynamics and Topological Entropy for the Restricted 3-Body Problem. *Comm. Math. Phys.* 231 (2002), no. 1, 1-24.

[5] G. Arioli, and P. Zgliczyński. Symbolic dynamics for the Hénon-Heiles Hamiltonian on the Critical Level. J. Differential Equations 171 (2001), no.1, 173-202.

[6] G. Arioli, and P. Zgliczyński. The Hénon-Helies Hamiltonian Near the Critical Energy Level-Some Rigorous Results. *Nonlinearity* 16 (2003), no. 5, 1833-1852.

[7] J.B. van den Berg, J.P. Lessard, K. Mischaikow, and J.D. Mireles James. Rigorous numerics for symmetric connecting orbits: even homoclinics of the Gray-Scott. SIAM Journal on Mathematical Analysis, Vol 43 (2011), Issue 4, pp. 1557–1594. www.math.rutgers.edu / ~jmireles /grayScottPage.html

[8] CAPD - *Computer assisted proofs in dynamics*, a software package for rigorous numerics, available online at `http://capd.ii.uj.edu.pl`.

[9] Martin Berz and Kyoko Makino, *Verified integration of ODEs and flows using differential algebraic methods on high-order Taylor models*, Reliab. Comput., 4 (1998), pp. 361–369.

[10] Martin Berz, *COSY INFINITY Version 8 reference manual.* Technical Report MSUCL-1088, National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, MI 48824, 1997.

[11] W. Beyn. The Numerical Computation of Connecting Orbits in Dynamical Systems. *IMA Journal of Numerical Analysis*, 9, $379-405$, 1990.

[12] W. Beyn, and J. Kleinkauf. The Numerical Computation of Homoclinic Orbits for Maps. *SIAM J. Numer. Anal.* 34 (1997), no.3, 1207-1236.

[13] W. Beyn, and J. Kleinkauf. Numerical Approximation of Homoclinic Chaos. Dynamical Numerical Analysis (Atlantl, GA, 1995). *Numer. Algorithms* 14 (1997), no. 1-3, 25-53.

[14] 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(2):283-328, (2003).

[15] 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(2):329-360, (2003).

[16] X. Cabré, E. Fontich, and R. de la Llave. The Parameterization Method for Invariant Manifolds. III. Overview and applications. *J. Differential Equations*, 218(2):444-515, (2005).

[17] 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), no. 6, 1311-1336.

[18] 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), no. 9, 2029-2058.

[19] M. Capiński, and C. Simó. Computer assisted proof for normally hyperbolic invariant manifolds. *Nonlinearity*, Vol. 25, No. 7, pg. 1997-2026 (2012).

[20] M. Capiński, and P. Zgliczyński. Cone conditions and covering relations for topologically normally hyperbolic invariant manifolds. *Discrete Contin. Dyn. Syst.*, Vol. 30, No. 3, pg. 641-670 (2011).

[21] M. Capinski. Covering Relations and the Existence of Topologically Normally Hyperbolic Invariant Sets. *Discrete Contin. Dyn. Syst.* 23 (2009), no. 3, 705-725.

[22] M. Capinski, and P. Roldan. 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.

[23] B. Coomes, H. Koçak, and K. Palmer. Homoclinic Shadowing. *J. Dynam. Differential Equations* 17 (2005), no.1, 175-215.

[24] S. Day, Y. Hiraoka, K. Mischaikow, and T. Ogawa. Rigorous numerics for global dynamics: a study of the Swift-Hohenberg equation. *SIAM J. Appl. Dyn. Syst.* 4 (2005), no. 1, 1-31.

[25] E.J. Doedel, and M.J. Friedman. Numerical computation of heteroclinic orbits. *Journal of Computational and Applied Mathematics* 26 (1989) 155–170.

[26] E.J. Doedel, and M.J. Friedman. Numerical computation and continuation of invariant manifolds connecting fixed points. *SIAM J. Numer. Anal.* Vol. 28 (1991), No. 3, pp. 789–808.

[27] H. Dullin, and J. Meiss. Quadratic Volume-Preserving Maps: Invariant Circles and Bifurcations. *SIAM J. Appl. Dyn. Syst.* 8 (2009), no. 1, 76-128.

[28] 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. *Electronic Research Announcements in Mathematical Sciences* Vol 16, pp. 9-22, (2009).

[29] 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), no. 5,

909-932.
[30] M. Gidea, and P. Zgliczyński. Covering Relations for Multidimensional Dynamical Systems. *J. Differential Equations* 202 (2004), no.1, 59-80.
[31] M. Gidea and P. Zgliczyński. Covering Relations for Multidimensional Dynamical Systems. II. *J. Differential Equations*, 202(1):59–80, 2004.
[32] R.H. Goodman, and J.K. Wróbel. High-Order Adaptive Method for Computing Two-dimensional Invariant Manifolds of Maps. (Submitted to Communications in Nonlinear Science and Numerical Simulation).
[33] R.H. Goodman, and J.K. Wróbel. High-Order Bisection Method for Computing Invariant Manifolds of Two-Dimensional Maps. International Journal of Bifurcation and Chaos, 21, pp. 2017-2042 (2011).
[34] A. Guillamon and G. Huguet. A Computational and Geometric Approach to Phase Resetting Curves and Surfaces. *SIAM Journal on Applied Dynamical Systems*, Vol 8, Issue 3, pp. 1005-1042 (2009).
[35] 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), no. 6, 1261-1300.
[36] 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), no. 2, 530-579.
[37] Y. Hiraoka. Construction of approximate solutions fo rigorous numerics of symmetric homoclinic orbits. *Workshop on "Pattern Formation Problems in Dissipative Systems" and "Mathematical Modeling and Analysis for Nonlinear Phenomena",* RIMS Kôkyûroku Bessatsu, B3, pp. 11-23. Res. Inst. Math. Sci. (RIMS) Kyoto, 2007.
[38] Y. Hiraoka. Rigorous numerics for symmetric homoclinic orbits in reversible dynamical systems. Kybernetika, Vol. 43, No. 6, pp. 797-806, 2007.
[39] W. M Hirsch, and C.C. Pugh. Stable manifolds and hyperbolic sets. Global Analysis (Berkeley 1960), Proc. Sympos. Pure Math Vol 14, pp 133–163, 1970.
[40] Jens Hoefkens and Martin Berz, *Controlling the wrapping effect in the solution of ODEs for asteroids*, Reliab. Comput., 9 (2003), pp. 21–41.
[41] Johnson T, Tucker W (2011) A note on the convergence of parametrised non-resonant invariant manifolds. Qual Theory Dyn Syst 10(1):107–121,
[42] V. R. Korostyshevskiy, and T. Wanner. A Hermite spectral method for the computation of homoclinic orbits and associated functionals. J. Comput. Appl. Math., 206(2):986-1006, 2007.
[43] 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), 763–791.
[44] R. de la Llave, A. González, A. Jorba, and J. Villanueva. KAM theory without action-angle variables, *Nonlinearity*, 18(2) pp. 855-895, (2005).
[45] R. de la Llave. A tutorial on KAM theory, Smooth ergodic theory and its applications (Seattle, WA, 1999), *Proc. Sympos. Pure Math.*, Vol 69, pp. 175-292, 2001.
[46] R. de la Llave, and J.D. Mireles James. Parameterization of Invariant Manifolds by Reducibility for Volume Preserving and Symplectic Maps, Discrete and Continuous Dynamical Systems, Vol. 32, No. 12, pp. 4321-4360, 2012
[47] J.P. Lessard, and C. Reinhardt. Rigorous numerics for nonlinear differential equations using Chebyshev series. SIAM Journal on Numerical Analysis, 52(1): 1-22, 2014.
[48] J.P. Lessard, and M. Gameiro. Efficient rigorous numerics for higher dimensional PDEs via one-dimensional estimates. SIAM Journal on Applied Dynamical Systems, 12(1): 204-245, 2013.
[49] J. B. van den Berg, J.P. Lessard, and K. Mischaikow. Global smooth solution curves using rigorous branch following. Mathematics of Computation, 79 (271), 1565-1584, 2010.
[50] J.P. Lessard, J.D. Mireles James, and Christian Reinhardt. Computer assisted proof of transverse saddle-to-saddle connecting orbits for first order vector fields, Journal of Dynamics and Differential Equations, Vol. 26, Issue 2, pp. 267-313, 2014.
[51] J.D. Mireles James. Polynomial Approximation of One Parameter Families of (Un)Stable Manifolds with Rigorous Computer Assisted Error Bounds. (Submitted).
[52] J.D. Mireles James. Quadratic Volume-Preserving Maps: (Un)stable Manifolds, Hyperbolic Dynamics, and Vortex Bubble Bifurcations. Journal of Nonlinear Science, Vol. 23, No. 4, pp. 585-615, 2013.
[53] J.D. Mireles James, and Hector Lomelí. Computation of Heteroclinic Arcs with Application to the Volume Preserving Hénon Family. *SIAM J. Applied Dynamical Systems*, Volume 9,

Issue 3, pp 919-953, 2010.

[54] J.D. Mireles James, and K. Mischaikow. Rigorous Computation of (Un)Stable Manifolds and Connecting Orbits for Analytic Maps. SIAM Journal on Applied Dynamical Systems, Vol. 12, No. 2, pp. 957-1006, 2013.

[55] J.D. Mireles James, and K. Mischaikow. `cap_ImCoLab`: *Computer Assisted Proof for Invariant Manifolds and Connecting Orbits Library in MatLab/IntLab* (2012) `http://www.math.rutgers.edu/~jmireles/tanglePaperPage.html`

[56] A. Neumaier, and T. Rage. Rigorous Chaos Verification in Discrete Dynamical Systems, *Physica D* 67 (1994), no. 4, pp 327-346.

[57] S. Newhouse, M. Berz, J. Grote, and K. Makino. On the Estimation of Topological Entropy on Surfaces. Geometric and probabilistic structures in dynamics, 243-270, *Contemp. Math.*, 469, *Amer. Math. Soc., Providence, RI,* 2008.

[58] J.M. Ortega. The Newton-Kantorovich Theorem. *Amer. Math. Monthly* 75 1968 658-660.

[59] K. J. Palmer. Exponential Dichotomies, the Shadowing Lemma and Transversal Homoclinic Points. Dynamics reporten, Vol. 1, 265-306, *Dynam. Report. Ser. Dynam. Systems Appl.*, 1, Wiley, Chichester, 1988.

[60] Y.B. Pesin. Characteristic Lyapunov Exponents and Smooth Ergodic Theory. Russian Mathematical Surveys, Vol. 32, Num. 4, pp. 55-114, 1977.

[61] W. Rudin. *Functional Analysis*, International Series in Pure and Applied Mathematics, Second Edition, McGraw-Hill Inc, New York, 1991

[62] D. Ruelle. Ergodic theory of differentiable dynamical systems. Publications Mathématiques de I'HÉS 50, pp. 27-58, 1979.

[63] S. Rump. Verification Methods: Rigorous Results Using Floating-Point Arithmetic. *Acta Numer.* 19 (2010), 287-449.

[64] S. Rump. *INTLAB - INTerval LABoratory*, Tibor Csendes , Developments in Reliable Computing, Kluwer Academic Publishers (1999), 77–104, `http://www.ti3.tuhh.de/rump/`

[65] D. Salamon. Morse Theory, the Conley index and Floer homologh. Bull. London Math. Soc. 22(2):113 - 140, 1990.

[66] M. Schwarz. Morse Homology. Progress in Math. 111, Birkhauser Verlag, Basel, 1993.

[67] L.P. Silnikov. A Case of the Existence of a Denumerable Set of Periodic Motions. (Russian) Dokl. Akad. Nauk SSSR 160, (1965), 558-561.

[68] S. Smale. Diffeomorphisms with Many Periodic Points. 1965 Differential and Combinatorial Topology (A Symposium in Honor of Marston Morse) pp. 63-80 *Princeton Univ. Press*, Princeton, N.J.

[69] J.C. Sprott. High-Dimensional Dynamics in the Delayed Hénon Map. *Electronic Journal of Theoretical Physics* 3 (2006), no. 12, 19-35.

[70] D. Stoffer, and K. Palmer. Rigorous Verification of Chaotic Behavior of Maps Using Validated Shadowing. *Nonlinearity* 12 (1999), no. 6, 1683-1698.

[71] D. Wilczak. Abundancs of Heteroclinic and Homoclinic Orbits for the Hyperchaotic Rössler System. *Discrete Contin. Dyn. Syst. Ser. B 11* (2009), no. 4, 1039-1055.

[72] A. Wittig. "Rigorous High-Precision Enclosures of Fixed Points and Their Invariant Manifolds," (Ph.D Dissertation) Michigan State University, 2011.

[73] E. Witten. Supersymmetry and Morse theory. J. Differential Geom., 17(4): 661-692 (1983), 1982.

[74] 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), no. 2-3, 107-126.

[75] P. Zgliczyński. Covering Relations, Cone Conditions and the Stable Manifold Theorem. *J. Differential Equations*, 246(5):1774–1819, 2009.