Rigorous Shadowing of Numerical Solutions of Ordinary Differential Equations by Containment

Wayne B. Hayes (wayne@cs.toronto.edu)

Kenneth R. Jackson (krj@cs.toronto.edu)

Department of Computer Science, University of Toronto,

10 King's College Rd., Toronto, ON, M5S 3G4, CANADA[1]

draft of December 3, 2001

Abstract

An exact trajectory of a dynamical system lying close to a numerical trajectory is called a *shadow*. We present a general-purpose method for proving the existence of finite-time shadows of numerical ODE integrations of arbitrary dimension in which some measure of hyperbolicity is present and there is either 0 or 1 expanding modes, or 0 or 1 contracting modes. Much of the rigor is provided automatically by interval arithmetic and validated ODE integration software that is freely available. The method is a generalization of a previously published *containment* process that was applicable only to two-dimensional maps. We extend it to handle maps of arbitrary dimension with the above restrictions, and finally to ODEs. The method involves building n-cubes around each point of the discrete numerical trajectory through which the shadow is guaranteed to pass at appropriate times. The proof consists of two steps: first, the rigorous computational verification of a simple geometric property we call the *inductive containment property*; and second, a simple geometric argument showing that this property implies the existence of a shadow. The computational step is almost entirely automated and easily adaptable to any ODE problem. The method allows for the rescaling of time, which is a necessary ingredient for successfully shadowing ODEs. Finally, the method is local, in the sense that it builds the shadow inductively, requiring information only from the most recent integration step, rather than more global information typical of several other methods. The method produces shadows of comparable length and distance to all currently published results.

[1] *The authors thank NSERC for financial support for this research.*

1 Introduction

Consider the *initial value problem* (IVP) for an autonomous *ordinary differential equation* (ODE)

$$\mathbf{y}'(t) = \mathbf{f}(\mathbf{y}(t)), \tag{1}$$

$$\mathbf{y}(t_0) = \mathbf{y}_0, \tag{2}$$

where the ODE (1) is called the *defining equation*, (2) is called the *initial condition*, $\mathbf{y}$ is an n-dimensional vector and f an n-dimensional vector-valued function. Standard forward error analysis (*e.g.,* Dahlquist and Björck (1974) or Kahaner, Moler, and Nash (1989)) tells us that, for a large class of ODEs, it is impossible in fixed-precision arithmetic to produce a numerical solution to an IVP which remains uniformly close to the exact solution for a long time. As a result, forward error bounds are impractical in such cases. However, one can often guarantee that a numerical solution remains uniformly close not to the solution starting at the initial condition specified, but instead to the exact solution to the ODE (1) starting at a nearby initial condition. In other words, if one allows the initial condition to have a non-zero error, just as one is satisfied with a non-zero error at all other times (Murdock 1995), then it may be possible to guarantee that the numerical solution remains uniformly close to *some* exact solution for a long time. Such an exact solution is called a *shadow* of the numerical solution.

Backward error analysis is a general term applied to methods of error analysis that relate a numerical solution to the exact solution of a "nearby" problem (Corless 1994). Defect-based and other backward error analyses allow a time-dependent perturbation to the defining equation while leaving the initial condition untouched. Shadowing does the opposite: the defining equation is left unchanged, while the initial condition is perturbed. For many physical systems which are modelled using ODEs, the governing equations are extremely well-known, and virtually all error is introduced by imprecise knowledge of initial conditions and/or by numerical error in the computation of the solution. In these contexts, shadowing may a be more appropriate method of error analysis than defect-based methods. On the other hand, rigorous shadowing as presented in this paper and elsewhere is extremely expensive. Whereas non-rigorous defect controlled methods are of roughly equal expense compared to more traditional integration methods, rigorous shadowing requires validated ODE integration, which at present tends to be several orders of magnitude more expensive in both time and memory than non-validated methods, even for low-dimensional problems.

Thus, the goal of shadowing should not be to validate every numerical solution computed, but instead to study under what conditions we can expect a numerical solution to have a shadow.

Procedures for finding shadows usually involve some sort of fixed-point method. These include non-rigorous numerical methods akin to Newton's method (Grebogi, Hammel, Yorke, and Sauer 1990; Quinlan and Tremaine 1992; Hayes 1995), and methods that employ a theorem to prove the existence of a shadow, usually relying on Brouwer's Fixed-Point Theorem or the Newton-Kantorovich Theorem (Sauer and Yorke 1991; Chow and Palmer 1991, 1992; Chow and Van Vleck 1994).

An important advance has been the realization that ODEs differ fundamentally from maps in that they have errors in *time* as well as space. By employing a *rescaling of time*, shadow lengths for ODEs can be increased by several orders of magnitude (Coomes, Koçak, and Palmer 1994b, 1995a, 1995b; Van Vleck 1995), even allowing the proof of existence of periodic trajectories near periodic pseudo-trajectories (Coomes, Koçak, and Palmer 1994a; Coomes, Koçak, and Palmer 1997). Shadowing has also been used to demonstrate that conservative integrations that approximately satisfy a first integral can have shadows that exactly satisfy it (Coomes 1997), and that more explicit control of the numerical error in the stable versus unstable subspaces can lead to better shadowing results (Van Vleck 2000). An interesting application has been to prove that a chaotic trajectory exists near an apparently chaotic pseudo-trajectory (Stoffer and Palmer 1999). Hayes (2001) provides a more detailed survey of ODE shadowing results.

This paper extends the work of Grebogi, Hammel, Yorke, and Sauer (1990), who introduced an elegant geometrical method called *containment* for proving the existence of shadows. Their proof is valid for iterated maps in two dimensions, and is also practical for two-dimensional ODE problems which do not require a rescaling of time. We extend their results to maps of arbitrary dimension in which some measure of hyperbolicity is present and there is either 0 or 1 expanding modes, or 0 or 1 contracting modes, and introduce a new method complementary to containment that facilitates a rescaling of time. In contrast to the above methods that use a fixed-point result, containment, including our new rescaling of time, uses an entirely geometrical argument. We rigorously verify the conditions of our theorems using validated ODE integration (Nedialkov 1999; Nedialkov, Jackson, and Corliss 1999), and demonstrate that containment is capable of proving the existence of shadows of IVPs for ODEs that are of comparable quality to any currently in the literature. We also demonstrate how containment can reproduce the proof of chaos given by Stoffer and Palmer (1999).

The outline of the paper is as follows. Section 2 presents the ideas for the proofs of containment in an informal, geometrical setting. We present the actual proofs in section 3. Formally, these proofs break into two steps. First, we must prove that the numerical trajectory satisfies a certain property called the *Inductive Containment Property* (ICP for short). The ICP can be proven computationally to hold using a validated ODE integrator; we defer discussion of how this is done until section 4. Second, we must show that a numerical trajectory that satisfies the ICP has a shadow. We prove this for maps in n dimensions for the cases that there is either one expanding direction while all the others contract (§3.1), one contracting direction while all the others expand (§3.2), or all directions either expand or contract (§3.3). Generalizing our scheme to apply effectively to ODE integrations requires an extension allowing a rescaling of time. One such extension is presented in section 5. Section 6 presents experimental results, followed in Section 7 by our conclusions.

2 Informal description of containment

Although containment was the first method introduced for proving the existence of finite-time shadows of numerical orbits, it has not, to our knowledge, been pursued beyond its initial conception. In this paper we demonstrate that, at least in the restricted cases discussed, containment is about as strong as any method currently in the literature.

2.1 Definitions

In this paper, an *orbit* is a discrete sequence of points, a *solution* is a continuous curve, and a trajectory more generally refers to either an orbit or a solution depending upon the context. The prefix *pseudo-* will be used to denote an approximate orbit, solution, or trajectory, although sometimes it will be omitted if the meaning is clear from the context.

Shadowing of numerical orbits was first applied to iterated maps.

Definition: An *orbit* of an *iterated map* consists of a sequence of points $\mathbf{x}_i$ generated by the recurrence $\mathbf{x}_{i+1} = \varphi(\mathbf{x}_i)$ for some map φ.

Definition: a *homeomorphism* is a map which is continuous, one-to-one, and onto.

For our purposes, we restrict φ to being a homeomorphism.

Definition: A *pseudo-orbit*, or *noisy* orbit, for φ satisfies $\mathbf{y}_{i+1} = \varphi(\mathbf{y}_i) + \delta_i$, where δ_i is the noise introduced at step i. If $\|\delta_i\| < \delta$ for all i, then it is called a δ-*pseudo-orbit* for φ.

Definition: The exact orbit $\{\mathbf{x}_i\}_{i=0}^N$ is an ε-shadow of the pseudo-orbit $\{\mathbf{y}_i\}_{i=0}^N$ if $\|\mathbf{y}_i - \mathbf{x}_i\| < \varepsilon$ for $i = 0, \ldots, N$.

Numerical solutions to ODEs can often be viewed as iterated maps by defining $\mathbf{x}_{i+1} = \varphi_{h_i}(\mathbf{x}_i)$, where φ_{h_i} is the *time-h_i* solution operator for the IVP (1,2). The time-h_i solution operator is a homeomorphism as long as $\mathbf{f}$ in (1) is bounded and Lipschitz continuous over the domain of interest (Ascher, Mattheij, and Russell 1988). For small h_i, a *one-step numerical method* approximates φ_{h_i} by $\tilde{\varphi}_{h_i}$ and then computes a sequence of discrete points $\mathbf{y}_{i+1} = \tilde{\varphi}_{h_i}(\mathbf{y}_i)$ representing approximations to $\mathbf{y}(t_{i+1})$, where $t_{i+1} = t_i + h_i$. We will term such a discrete sequence of points a *pseudo-trajectory.* If the pseudo-trajectory satisfies a *local error tolerance* of δ such that $\|\mathbf{y}_{i+1} - \varphi_{h_i}(\mathbf{y}_i)\| \leq \delta$, then we call it a *$\delta$-pseudo-trajectory.* If h_i is constant, we can drop it as a subscript and treat the pseudo-trajectory as a pseudo-orbit of the iterated map $\varphi \equiv \varphi_h$.

2.1.1 Hyperbolicity and pseudo-hyperbolicity

One of the most important concepts in shadowing is that of *hyperbolicity.* Essentially, a system of ODEs is hyperbolic if the variational equation along a solution $\mathbf{y}(t)$ displays *exponential dichotomy* (Palmer 1988). This means that a perturbation δ to the solution $\mathbf{y}(t)$ at time $t = t_0$, $\mathbf{z}(t_0) = \mathbf{y}(t_0) + \delta$ produces a new solution $\mathbf{z}(t)$ with one of two properties: if δ lies in the *stable subspace* of $\mathbf{y}(t)$, then $\mathbf{z}(t)$ converges exponentially to $\mathbf{y}(t)$ as t increases; if δ lies outside the stable subspace (*i.e.,* if it has any component in the *unstable subspace*), then $\mathbf{z}(t)$ diverges exponentially away from $\mathbf{y}(t)$ as t increases. More details can be found in Hayes (2001) or Palmer (1988). If a system is hyperbolic, then the angle between the stable and unstable subspaces is always bounded away from 0 (Grebogi, Hammel, Yorke, and Sauer 1990).

This paper deals not with hyperbolic systems, but with systems whose pseudo-trajectories are shadowable for finite but nontrivial lengths of time even though they are not hyperbolic. For this to occur, a system must display pseudo-hyperbolicity. We say that a system is *pseudo-hyperbolic* if a small perturbation to a trajectory $\mathbf{y}(t)$ produces a new solution $\mathbf{z}(t)$ which falls into one of two classes: those which *tend to* diverge exponentially away from $\mathbf{y}(t)$ as t increases, and those that *tend to* converge exponentially towards $\mathbf{y}(t)$ as t increases. In addition, $\mathbf{z}(t)$ should behave in this manner over non-trivial periods of time. In short, a pseudo-hyperbolic system should "mimic" the behaviour of a hyperbolic system over finite but nontrivial periods of time. This notion could be quantified by, for example, attempting to find the stable and unstable subspaces using *refinement* (Grebogi, Hammel, Yorke, and Sauer 1990; Quinlan and Tremaine 1992; Hayes 1995; Hayes 2001), and then performing least-squares fits to exponential curves of the growth and decay of these subspaces.

2.2 Containment in two dimensions

The first studies of shadows of pseudo-hyperbolic systems appear to be Beyn (1987) and Hammel, Yorke, and Grebogi (1987). Hammel, Yorke, and Grebogi (1988) and Grebogi, Hammel, Yorke, and Sauer (1990) (hereafter GHYS) provide the first proof of the existence of a shadow for a non-hyperbolic system over a non-trivial length of time. Their method consists of two parts. First, they *refine* a noisy trajectory using an iterative method that produces a nearby trajectory with less noise. When refinement converges to the point that the noise is of order the machine precision, they invoke *containment*, which can prove the existence of a nearby exact trajectory. Their method, which we now describe, is directly applicable only to two-dimensional maps.

Let $\{\mathbf{y}_i\}_{i=0}^{N} \subset \mathbf{R}^2$ be a two-dimensional δ-pseudo-orbit of φ. As i increases, orbits separated from each other by a small distance along the expanding direction diverge on average away from each other, while orbits separated by a small distance along the contracting direction approach each other on average. The containment process consists of building a parallelogram M_i around each point $\mathbf{y}_i$ of the pseudo-orbit such that two sides $C_i^{\pm 1}$ are approximately normal to, and separated from each other along, the contracting direction while the other two sides $E_i^{\pm 1}$ are approximately normal to, and separated from each other along, the expanding direction.[2] The diameter of M_i will bound the distance from the pseudo-orbit to the shadow. In order to prove the existence of a shadow, the image of M_i under φ must intersect M_{i+1} such that $\varphi(M_i)$ makes a "plus sign" with M_{i+1} (Figure 1).

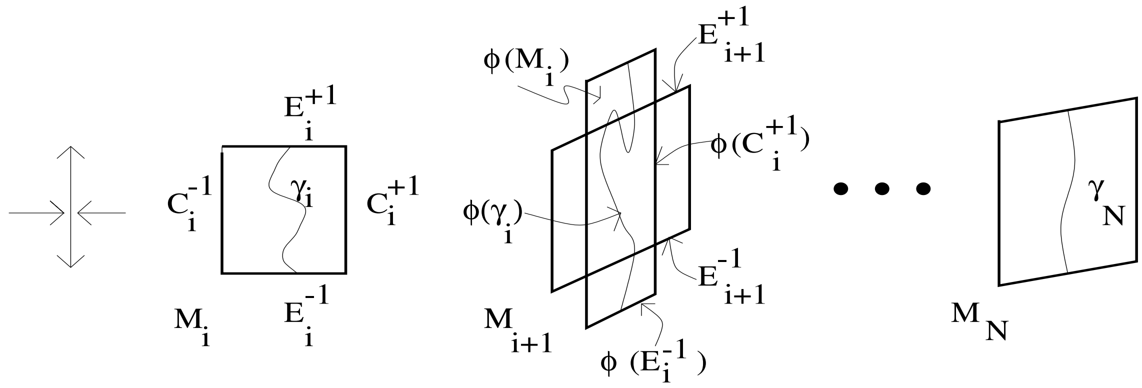


Figure 1: Containment in two dimensions, reproduced from GHYS. The horizontal direction is contracting, and the vertical direction is expanding.

[2]Note that this naming convention is exactly opposite to that of GHYS, because in two dimensions they emphasized the direction to which the sides of M_i were *parallel*. In higher dimensions, the faces of an n-cube are not parallel to a unique direction, and it is the direction along which a face is separated from the centre of the n-cube that matters. We change the naming convention now to avoid confusion later.

To ensure this property holds, GHYS require a bound on the second derivative of φ, and the amounts of expansion and contraction need to be resolvable to within the machine precision. The proof of the existence of an exact orbit then relies on the following argument. For any $i \in \{0, 1, \ldots, N-1\}$ let γ_i be a continuous curve in M_i connecting the expanding sides E_i^{-1} and E_i^{+1}. Its image $\varphi(\gamma_i)$ is then stretched such that there is a section of $\varphi(\gamma_i)$ lying wholly within M_{i+1}, and in particular $\varphi(\gamma_i)$ leaves M_{i+1} through the expanding sides $E_{i+1}^{\pm 1}$ at both ends. Let γ_{i+1} be a continuous sub-section of $\varphi(\gamma_i)$ lying wholly within M_{i+1} connecting the expanding sides $E_{i+1}^{\pm 1}$. Repeating this process along the orbit produces γ_N lying wholly within the final parallelogram M_N. Then any point $\mathbf{x}_N \in \gamma_N$ traced backwards via φ^{-1} yields a point $\mathbf{x}_i \in \gamma_i \subset M_i, \quad i = N-1, \ldots, 1, 0$. Note that $\{\mathbf{x}_i\}_{i=0}^N$ is an exact orbit. Moreover, since $\mathbf{x}_i, \mathbf{y}_i \in M_i$, we infer that $\|\mathbf{x}_i - \mathbf{y}_i\| \leq \varepsilon$, where ε bounds the diameter of $M_i, \ i = 0, \ldots, N$. Thus, $\{\mathbf{x}_i\}_{i=0}^N$ is an ε-shadow of $\{\mathbf{y}_i\}_{i=0}^N$. We make the intuitive argument described here rigorous in section 3.

With this picture in mind, there is a nice geometric interpretation of the requirement that the angle between the stable and unstable directions be bounded away from 0: if the angle gets too small, then the parallelogram essentially loses a dimension, and $\varphi(M_i)$ can not make a "plus sign" with M_{i+1}. Practically speaking, this occurs when the angle becomes comparable with the noise amplitude of the pseudo-orbit. Hence, the more accurate the orbit, the longer it can be shadowed (Grebogi, Hammel, Yorke, and Sauer 1990; Quinlan and Tremaine 1992).

2.3 Containment in 3 dimensions

The process described by GHYS is not directly applicable to systems with more than 2 dimensions, and GHYS provided no indication of how it could be extended beyond 2 dimensions. We describe how the method can be extended to 3 dimensions, in which there are precisely two interesting cases:

(i) 1 expanding direction, and 2 contracting (Figure 2). Assume that the z direction is expanding, while the x and y directions are contracting. (We assume, for simplicity of exposition and for ease of drawing, that these three directions are roughly orthogonal, although in practice they need only be resolvable from each other.) Then, analogously to the 2-dimensional argument, assume we can draw a *cube* M_i of diameter no larger than ε around each noisy point $\mathbf{y}_i$, and, for $i = 0, 1, \ldots, N-1$, assume we can verify that $\varphi(M_i)$ maps over M_{i+1} so that φ stretches M_i into a long, thin tube, a segment of which lies wholly in M_{i+1}. Then, precisely as in the 2-dimensional case, we can

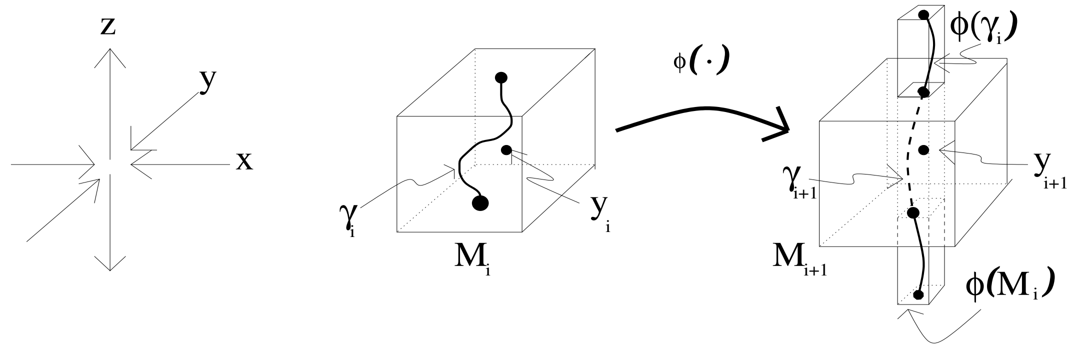


Figure 2: Containment in 3D, case (i): 1 expanding direction and 2 contracting.

prove an ε-shadow of $\{\mathbf{y}_i\}_{i=0}^N$ exists as follows. We introduce a curve γ_i that runs approximately along the expanding (vertical) direction from any point on the top of M_i to its bottom. If $\varphi(M_i)$ maps over M_{i+1} as in Figure 2, then we are guaranteed that a contiguous section of $\varphi(\gamma_i)$ lies inside M_{i+1}, connecting its top and bottom along the expanding direction. This segment of $\varphi(\gamma_i)$ becomes γ_{i+1}. Any point $\mathbf{x}_N \in \gamma_N \subset M_N$ can be traced backwards via φ^{-1} to a point $\mathbf{x}_i \in \gamma_i \subset M_i$ for $i = 0, 1, \ldots, N-1$. As in the 2-dimensional case, $\{\mathbf{x}_i\}_{i=0}^N$ is an ε-shadow of $\{\mathbf{y}_i\}_{i=0}^N$.

(ii) 2 expanding and 1 contracting direction (Figure 3). Assume now that the z (vertical) direction is contracting, while the x and y directions are expanding. Assume that we can again draw a cube M_i of diameter no larger than ε around each noisy point $\mathbf{y}_i$ such that $\varphi(M_i)$ maps over M_{i+1} so that φ flattens M_i into a thin slice, cutting M_{i+1} into 3 pieces, the middle piece of which contains a contiguous section of $\varphi(M_i)$. We can prove

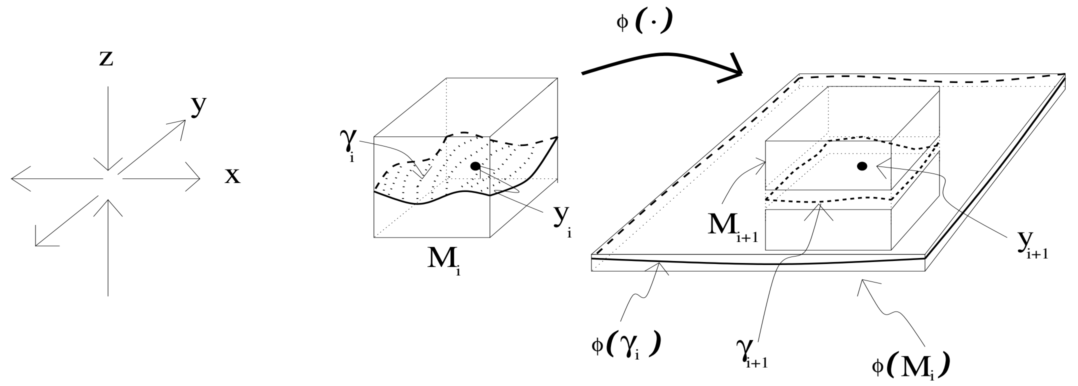


Figure 3: Containment in 3D, case (ii): 2 expanding directions and 1 contracting.

an ε-shadow of $\{\mathbf{y}_i\}_{i=0}^N$ exists as follows. Let γ_i be a *surface* whose boundary connects all of the expanding sides of M_i so that, under the mapping, $\varphi(\gamma_i)$ is stretched in all the directions it has extent (both horizontal directions), and is "compressed" along the direction it has measure zero (vertical). Then, we are guaranteed that there is a contiguous segment of $\varphi(\gamma_i)$ lying wholly in M_{i+1} and connecting all of its expanding sides. We call this surface γ_{i+1}. Then any point $\mathbf{x}_N \in \gamma_N \subset M_N$ can be traced back via φ^{-1} to a point $\mathbf{x}_i \in \gamma_i \subset M_i$ for $i = 0, 1, \ldots, N-1$. As in the 2-dimensional case, $\{\mathbf{x}_i\}_{i=0}^N$ is an ε-shadow of $\{\mathbf{y}_i\}_{i=0}^N$.

We make the intuitive arguments described above rigorous in section 3.

It seems intuitively clear that we can replace "cube" with "n-cube", "surface" with "manifold", and the above argument still applies in arbitrarily high dimension. The crucial points appear to be that γ has dimension equal to the number of expanding directions, and that its border must "wrap around" all the expanding sides of M_i.[3]

3 Containment theorems and proofs

3.1 Containment in n dimensions with one expanding direction

Here we provide a proof of what we call the $(n, 1)$-Inductive Containment Theorem: the n-dimensional case in which precisely one direction is expanding while all the others contract. Previous proofs of containment required explicit *a priori* bounds on spatial derivatives, whereas our proof requires no such bounds.[4]

Let M_i be a parallelepiped in $\mathbf{R}^n$ with faces F_i^j, for $i = 0, \ldots, N$ and $j = \pm 1, \ldots, \pm n$, with opposite signs in the superscript representing opposite faces of a parallelepiped. Without loss of generality, we assume that the first direction is the "expanding" one. We will denote the union of a set of faces by listing all of them in the superscript; for example, $F_i^{\pm 2, \ldots, \pm n}$ represents the set of all the faces of M_i except F_i^{-1} and F_i^{+1}. Let $\partial_E M_i \equiv F_i^{-1} \cup F_i^{+1} \equiv F_i^{\pm 1}$, and $\partial_C M_i \equiv \bigcup_{j=2}^n F_i^{-j} \cup F_i^{+j} \equiv F_i^{\pm 2, \ldots, \pm n}$. Let $\varphi : \mathbf{R}^n \to \mathbf{R}^n$ be a homeomorphism. Let *int* X represent the interior of X. Then M_i and M_{i+1} satisfy the $(n, 1)$-*Inductive Containment Property* (called ICP for short throughout this subsection) if

(1) $\varphi(F_i^{\pm 1}) \cap M_{i+1} = \emptyset$, and $\varphi(F_i^{-1})$ and $\varphi(F_i^{+1})$ are on opposite sides of the infinite slab between the two hyperplanes containing F_{i+1}^{-1} and F_{i+1}^{+1}.

[3] Discussion of how the phrase "wrap around" generalizes to higher dimensions is beyond the scope of this discussion, although it can be defined precisely by means of homotopy theory (see for example Munkres 1975).

[4] Of course, our validated ODE integration (Nedialkov 1999) must compute bounds on derivatives in order to compute enclosures, but these bounds are not *a priori*; they are computed on-the-fly, and if a bounds check fails, we can always try a smaller timestep to compensate.

(2) $\exists Q_{i+1}$, a parallelepiped in $\mathbf{R}^n$ with each face G^j_{i+1} parallel to the corresponding face F^j_{i+1} of M_{i+1} for $j = \pm 1, \ldots, \pm n$ such that

(i) $\varphi(M_i) \subset int\ Q_{i+1}$,

(ii) $F^{\pm 2,\ldots,\pm n}_{i+1} \cap Q_{i+1} = \emptyset$, and $\forall j \in \{2, \ldots, n\}$, F^{-j}_{i+1} and F^{+j}_{i+1} are on opposite sides of the infinite slab between the two hyperplanes containing G^{-j}_{i+1} and G^{+j}_{i+1}.

Let $\gamma_0 \subset M_0$ be a simple curve joining F^{-1}_0 to F^{+1}_0, but otherwise remaining in the interior of M_0. That is,

$$\gamma_0 \cap F^{-1}_0 \neq \emptyset \ \wedge \ \gamma_0 \cap F^{+1}_0 \neq \emptyset \ \wedge \ int\ \gamma_0 \subset int\ M_0.$$

Theorem 1 ($(n,1)$-Inductive Containment Theorem). *If M_i and M_{i+1} satisfy the $(n,1)$-ICP $\forall i = 0, \ldots, N-1$, then $\forall i = 0, \ldots, N$*

$$\exists \textit{ simple curve } \gamma_i \subseteq \varphi^i(\gamma_0) \textit{ s.t. } \gamma_i \cap F^{-1}_i \neq \emptyset \ \wedge \ \gamma_i \cap F^{+1}_i \neq \emptyset \ \wedge \ int\ \gamma_i \subset int\ M_i. \quad (3)$$

That is, γ_i touches the boundary of M_i in precisely two places, connecting F^{-1}_i to F^{+1}_i, but otherwise remains entirely inside M_i.

Proof. By induction on i. The proof of the base case $i = 0$ is immediate, by the definition of γ_0. For the inductive case, assume $\exists$ a simple curve $\gamma_i \subseteq \varphi^i(\gamma_0)$ such that $\gamma_i \cap F^{-1}_i \neq \emptyset \ \wedge \ \gamma_i \cap F^{+1}_i \neq \emptyset \ \wedge \ int\ \gamma_i \subset int\ M_i$. From ICP(1) and $\varphi(F^{\pm 1}_i) \subset Q_{i+1}$ and the fact that Q_{i+1} is convex, we know that Q_{i+1} intersects both F^{-1}_{i+1} and F^{+1}_{i+1}; and from ICP(2ii), Q_{i+1} does not intersect $F^{\pm 2,\ldots,\pm n}_{i+1}$. Thus, since Q_{i+1} is convex, $Q_{i+1} - M_{i+1}$ is disconnected by the slab defined in ICP(1) into two disjoint components[5] say Q^{-1}_{i+1} and Q^{+1}_{i+1}, each containing one of $\varphi(F^{\pm 1}_i)$, by ICP(1). Without loss of generality, assume $\varphi(F^j_i) \subset Q^j_{i+1}$, $j = \pm 1$. Now, consider one of the components, say Q^{-1}_{i+1}. It contains one of the two endpoints of $\varphi(\gamma_i)$, since one endpoint is in $\varphi(F^{-1}_i) \subset Q^{-1}_{i+1}$, while the other endpoint of $\varphi(\gamma_i)$ is in $\varphi(F^{+1}_i) \subset Q^{+1}_{i+1}$. Since γ_i is a simple curve and φ is a homeomorphism, $\varphi(\gamma_i)$ is a simple curve. Now, $Q^{-1}_{i+1} \cap Q^{+1}_{i+1} = \emptyset$, and $\varphi(\gamma_i)$ connects the two. Thus, $\varphi(\gamma_i)$ must cross the boundary of Q^{-1}_{i+1}. This boundary consists of exactly two mutually exclusive patches, one of which is a subset of ∂Q_{i+1}, the other a subset of F^{-1}_{i+1}. Since $\varphi(\gamma_i) \subset \varphi(M_i) \subset int\ Q_{i+1}$, we infer that $\varphi(\gamma_i) \cap \partial Q_{i+1} = \emptyset$, and so $\varphi(\gamma_i)$ leaves Q^{-1}_{i+1} through F^{-1}_{i+1}. A similar argument

[5]This is because F^{-1}_{i+1} and F^{+1}_{i+1} are each patches of an $n-1$ dimensional hyperplane residing in n dimensions, and so they each disconnect any convex set they intersect, as long as that convex set does not intersect their boundaries ∂F^{-1}_{i+1} and ∂F^{+1}_{i+1}, respectively.

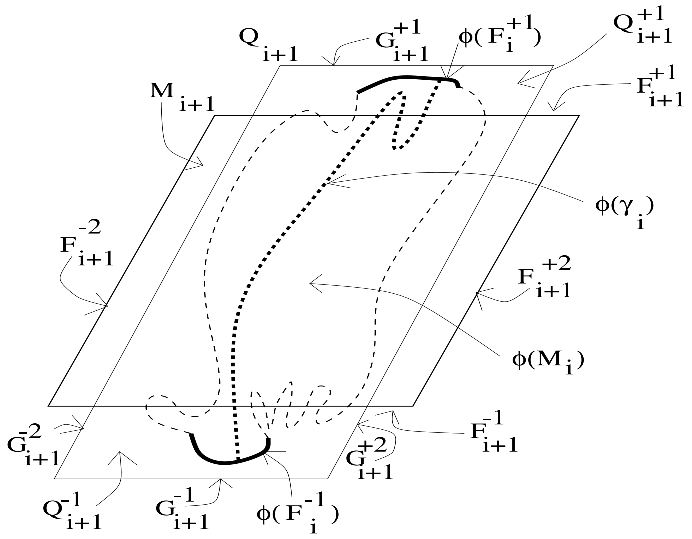


Figure 4: The image $\varphi(M_i)$ and M_{i+1} for 2 dimensions. The dark curves at the bottom and top are $\varphi(F_i^{\pm 1})$. The dashed curves at the left and right are $\varphi(F_i^{\pm 2})$.

shows that $\varphi(\gamma_i)$ leaves Q_{i+1}^{+1} through F_{i+1}^{+1}. Thus, $\varphi(\gamma_i) \cap F_{i+1}^{j} \neq \emptyset$, $j = \pm 1$. It remains to show that there exists a segment γ_{i+1} of $\varphi(\gamma_i)$ which is a simple curve and maintains the property defined in (3).

Since $\varphi(\gamma_i)$ is a simple curve, there exists a parameterization $\gamma(t)$ for $t \in [0,1]$ such that $\gamma([0,1]) = \varphi(\gamma_i)$ and $\gamma(t)$ is a homeomorphism (Munkres 1975). Let $s^j = \varphi(\gamma_i) \cap F_{i+1}^{j}$, $j = \pm 1$. Now, s^{-1} and s^{+1} are disjoint sets since $F_{i+1}^{-1} \cap F_{i+1}^{+1} = \emptyset$, and they are compact because (1) F_{i+1}^{j} for $j = \pm 1$ are compact, (2) γ_i is compact, φ is a homeomorphism and so $\varphi(\gamma_i)$ is compact, (3) and the intersection of two compact sets in $\mathbf{R}^n$ is compact. Finally, $\gamma^{-1}(s^{\pm 1})$ are compact because γ is a homeomorphism. To prove that there exists a simple curve $\gamma_{i+1} \subset \varphi(\gamma_i)$ such that $\gamma_{i+1} \cap F_{i+1}^{-1} \neq \emptyset$, $\gamma_{i+1} \cap F_{i+1}^{+1} \neq \emptyset$, and $int\ \gamma_{i+1} \subset int\ M_{i+1}$, we need to show that there exist two points in $[0,1]$, one each from $\gamma^{-1}(s^{-1})$ and $\gamma^{-1}(s^{+1})$, such that no points from either set are between them. This will prove that there exists a simple curve, which is a section of $\varphi(\gamma_i)$, that connects F_{i+1}^{-1} to F_{i+1}^{+1} without otherwise intersecting ∂M_{i+1}. To this end, let $G = \gamma^{-1}(s^{-1})$ and $R = \gamma^{-1}(s^{+1})$ and note that G and R are compact, disjoint, nonempty subsets of $[0,1]$. The following lemma completes the proof. $\square$

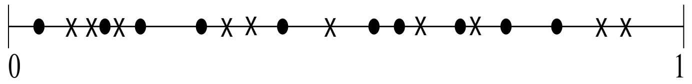


Figure 5: Schematic representation of the sets $\gamma^{-1}(s^{-1})$ (dots) and $\gamma^{-1}(s^{+1})$ ($\times$'s).

Lemma 2. *Let G and R be (possibly infinite) disjoint, compact, nonempty subsets of* $[0, 1]$. *Then* $\exists g \in G,\ r \in R$ *such that* $(g, r) \cap (G \cup R) = \emptyset$, *where we have assumed without loss of generality than* $g < r$.

Proof. Consider the function $f(x, y) = |x - y|$ over the subset $G \times R$ of the plane. Since f is continuous and $G \times R$ is compact, f attains its minimum at some point $(g, r) \in G \times R$. That is, $|g - r| \leq |g' - r'|$ for any other $g' \in G,\ r' \in R$. Thus, there is no element of either set G or R between g and r, and so the open interval (g, r) is disjoint from $G \cup R$. $\square$

Theorem 3 (Shadowing Containment Theorem). *Let φ be a homeomorphism. Let $\{M_i\}_{i=0}^N$ be a sequence of parallelepipeds enclosing a pseudo-trajectory $\{\mathbf{y}_i\}_{i=0}^N$. Let ε be the maximum diameter of M_i for $i = 0, 1, \ldots, N$. Let $\gamma_i \subset M_i, \gamma_i \neq \emptyset,\ i = 0, \ldots, N$, and let $\gamma_{i+1} \subseteq \varphi(\gamma_i),\ i = 0, \ldots, N - 1$. Then $\exists$ an ε-shadow $\{\mathbf{x}_i\}_{i=0}^N$ of $\{\mathbf{y}_i\}_{i=0}^N$. That is, there is an exact trajectory $\{\mathbf{x}_i\}_{i=0}^N$ of φ such that $\|\mathbf{x}_i - \mathbf{y}_i\| < \varepsilon,\ i = 0, \ldots, N$.*

Proof. Since φ is a homeomorphism, φ^{-1} is a well-defined function. Pick any point $\mathbf{x}_N \in \gamma_N$, and recursively define $\mathbf{x}_i = \varphi^{-1}(\mathbf{x}_{i+1}),\ i = N - 1, N - 2, \ldots, 0$. Since $\gamma_{i+1} \subseteq \varphi(\gamma_i)$, $\varphi^{-1}(\gamma_{i+1}) \subseteq \gamma_i$, and so by induction $\mathbf{x}_i \in \gamma_i$ for $i = N, N - 1, \ldots, 0$. Since $\mathbf{y}_i \in M_i$ and $\mathbf{x}_i \in \gamma_i \subset M_i$, $\|\mathbf{y}_i - \mathbf{x}_i\| \leq \mathrm{diam}(M_i) \leq \varepsilon,\ i = 0, \ldots, N$. $\square$

Thus, applying Theorem 3 to an orbit satisfying the $(n, 1)$-Inductive Containment Property implies the existence of a shadow.

Remark: Note that Theorem 3 is independent of the number of dimensions n, and independent of the number of expanding and contracting directions, because the only parts of the Inductive Containment Theorem that are used are the conclusions that $\gamma_{i+1} \subseteq \varphi(\gamma_i)$ for $i = 0, 1, \ldots, N - 1$ and $\emptyset \neq \gamma_i \subset M_i$ for $i = 0, \ldots, N$. As will be seen, the $(n, n - 1)$ Inductive Containment Theorem also asserts this property. The 0-expanding and 0-contracting directions are handled separately. We conjecture that the general (n, k)-Inductive Containment Theorem (work in progress) will also assert this property, so that the above Shadowing Containment Theorem is applicable to the general (n, k) case, in which k directions are expanding and $n - k$ are contracting.

3.2 Containment in n dimensions with one contracting direction

Remark: This case could immediately be proved by applying the one-expanding-direction theorem in backward time, since a system with one contracting direction in forward time is equivalent to a system with one expanding direction in backward time. However, we provide a different proof because, although it is not applicable to the general case, we believe a proof of the following form is more likely to be generalizable.

Remark: In the one expanding direction case, γ is one dimensional; in the one contracting direction case, γ is $(n-1)$-dimensional.

Let M_i be a parallelepiped in $\mathbf{R}^n$ with faces F_i^j, for $i = 0, \ldots, N$ and $j = \pm 1, \ldots, \pm n$, with opposite signs of j representing opposite faces of a parallelepiped. Without loss of generality, let the nth direction be the "contracting" one. We will denote the union of a set of faces by listing all of them in the superscript; for example, $F_i^{\pm 1, \ldots, \pm(n-1)}$ represents the set of all faces except F_i^{-n} and F_i^{+n}.

Let $\varphi : \mathbf{R}^n \to \mathbf{R}^n$ be a homeomorphism. Let *int* X represent the interior of X. Then M_i and M_{i+1} satisfy the $(n, n-1)$-*Inductive Containment Property* (called ICP for short throughout this section) if

(1) $\varphi(F_i^{\pm 1, \ldots, \pm(n-1)}) \cap M_{i+1} = \emptyset$ and $\forall j \in \{1, \ldots, n-1\}$ $\varphi(F_i^{-j})$ and $\varphi(F_i^{+j})$ are on opposite sides of the infinite slab between the two hyperplanes containing F_{i+1}^{-j} and F_{i+1}^{+j}.

(2) $\exists Q_{i+1}$, a parallelepiped in $\mathbf{R}^n$ with each face G_{i+1}^j parallel to the corresponding face F_{i+1}^j of M_{i+1} for $j = \pm 1, \ldots, \pm n$ such that

   (i) $\varphi(M_i) \subset \mathit{int}\ Q_{i+1}$,

   (ii) $F_{i+1}^{\pm n} \cap Q_{i+1} = \emptyset$ and F_{i+1}^{-n} and F_{i+1}^{+n} are on opposite sides of the infinite slab between the two hyperplanes containing G_{i+1}^{-n} and G_{i+1}^{+n}.

Lemma 4. *In an n-dimensional cube, the border of each face is contained in the union of all the other faces except the one opposite itself. That is,* $\partial F_i^j \subset \bigcup_{\substack{k \in \{\pm 1, \ldots, \pm n\} \\ k \neq \pm j}} F_i^k$.

Proof. Without loss of generality, consider the n-dimensional unit cube $C = [0,1]^n$. The border ∂C of C is $\partial C = \{\mathbf{p} = (p_1, \ldots, p_n)^T \in C \mid p_k \in \{0,1\} \text{ for some } k\}$. A face F of C is defined by assigning 0 or 1 to precisely one of the co-ordinates p_k of $\mathbf{p}$ (with the choice between 0 and 1 being the choice between opposite faces), and freeing all the other co-ordinates $p_j, j \neq k$ to roam in $[0,1]$. Without loss of generality, let $p_k = 0$. That is, $F = \{(p_1, \ldots, p_n)^T \in C \mid p_k = 0\}$. Note that F is a hypercube of dimension $n-1$. Thus, analogously to ∂C, we can define $\partial F = \{\mathbf{q} = (q_1, \ldots, q_n)^T \in F \mid q_j \in \{0,1\} \text{ for some } j \neq k\}$. However, if a co-ordinate q_j for some $j \neq k$ is 0 or 1, then in addition to being in F, $\mathbf{q}$ is in some other face F' of C. Namely, F' is either the face which has all points $q_j = 0$ or the one that has all points $q_j = 1$. Furthermore, F' can be any face of C except F or the one opposite F by choosing j and $q_j \in \{0, 1\}$ appropriately. $\square$

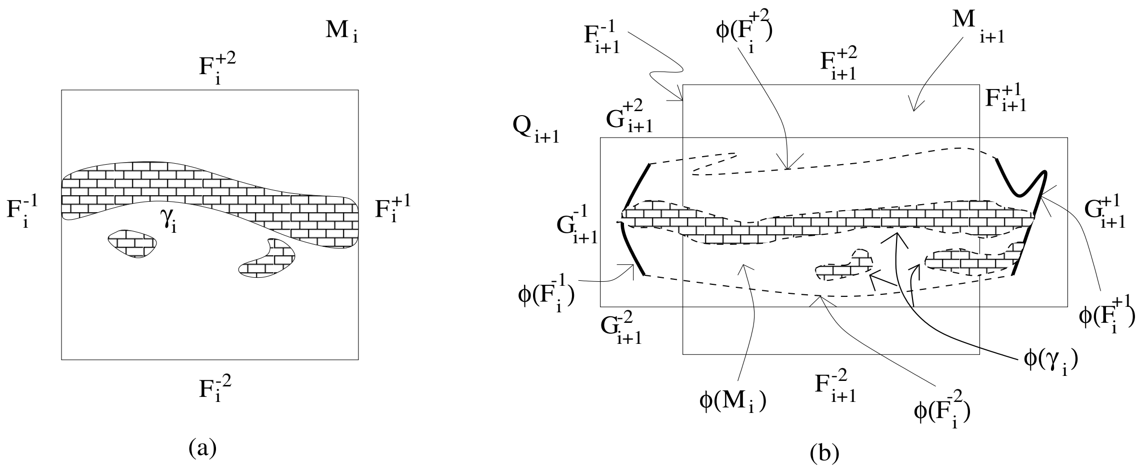


Figure 6: Schematic representation of the one-contracting-dimension proof in two dimensions. (a) γ_i, M_i and the faces F_i^j of M_i. (b) The dark curves at the left and right represent $\varphi(F_i^{\pm 1})$. The dashed curves at the top and bottom are $\varphi(F_i^{\pm 2})$. The proof of Theorem 6 shows that no path in M_{i+1} connects F_{i+1}^{-2} to F_{i+1}^{+2} without intersecting $\gamma_{i+1} = \varphi(\gamma_i) \cap M_{i+1}$.

Lemma 5. *Both $\varphi(F_i^{-n})$ and $\varphi(F_i^{+n})$ path disconnect F_{i+1}^{-n} from F_{i+1}^{+n} in M_{i+1}.*

Proof. By Lemma 4, the border of F_i^{-n} is contained in $F_i^{\pm 1, \ldots, \pm(n-1)}$. Since φ is a homeomorphism, the border of $\varphi(F_i^{-n})$ is contained in $\varphi(F_i^{\pm 1, \ldots, \pm(n-1)})$, which by ICP(1) is disjoint from M_{i+1}. However, by ICP(2i), $\varphi(F_i^{-n})$ is contained in the slab between G_{i+1}^{-n} and G_{i+1}^{+n}. Since $\varphi(F_i^{-n})$ is homeomorphic to a face of M_{i+1} and it lies in the slab between G_{i+1}^{-n} and G_{i+1}^{+n} of Q_{i+1} and its border lies outside M_{i+1}, it path disconnects M_{i+1} somewhere inside the slab between G_{i+1}^{-n} and G_{i+1}^{+n}. Furthermore, by ICP(2ii) the slab between G_{i+1}^{-n} and G_{i+1}^{+n} is contained in the slab between F_{i+1}^{-n} and F_{i+1}^{+n}, so F_{i+1}^{-n} and F_{i+1}^{+n} must be path disconnected from each other in M_{i+1} by $\varphi(F_i^{-n})$. The same argument applies to $\varphi(F_i^{+n})$. $\square$

Let $\gamma_0 \subset M_0$ be a set constructed to path disconnect F_0^{-n} from F_0^{+n} in M_0. That is, any path in M_0 that connects F_0^{-n} to F_0^{+n} must intersect γ_0.

Theorem 6 **($(n, n-1)$ Inductive Containment Theorem).** *If M_i and M_{i+1} satisfy the $(n, n-1)$-ICP for all $i = 0, \ldots, N-1$, then*

$$\forall i = 0, \ldots, N \;\; \exists \gamma_i \subseteq \varphi^i(\gamma_0) \text{ s.t. } \gamma_i \subset M_i \text{ and } \gamma_i \text{ path disconnects } F_i^{-n} \text{ from } F_i^{+n} \text{ in } M_i.$$

Proof. By induction on i. The proof of the base case $i = 0$ is immediate, by the definition of γ_0. For the inductive case, assume there exists a set $\gamma_i \subseteq \varphi^i(\gamma_0)$ such that $\gamma_i \subset M_i$ and γ_i path disconnects F_i^{-n} from F_i^{+n} in M_i. We claim that $\gamma_{i+1} = \varphi(\gamma_i) \cap M_{i+1}$ path disconnects F_{i+1}^{-n} from F_{i+1}^{+n} in M_{i+1}. Assume to the contrary that there exists a path β in M_{i+1} from F_{i+1}^{-n} to F_{i+1}^{+n} that does not intersect γ_{i+1}. By Lemma 5, $\varphi(F_i^{-n})$ and $\varphi(F_i^{+n})$ each path disconnect F_{i+1}^{-n} from F_{i+1}^{+n} in M_{i+1}. Thus, β intersects both $\varphi(F_i^{-n})$ and $\varphi(F_i^{+n})$, and connects the two, but by assumption does not intersect γ_{i+1}. Since $\beta \subset M_{i+1}$, ICP(1) implies $\beta \cap \varphi(F_i^{\pm 1,\dots,\pm(n-1)}) = \emptyset$. Thus, β intersects $\varphi(\partial M_i)$ only in $\varphi(F_i^{\pm n})$. Without loss of generality, assume that β intersects each only once. That is, it enters $\varphi(M_i)$ through $\varphi(F_i^{-n})$ and leaves through $\varphi(F_i^{+n})$. Thus, $\beta' = \beta \cap \varphi(M_i)$ connects $\varphi(F_i^{-n})$ to $\varphi(F_i^{+n})$ without leaving $\varphi(M_i)$, and, by assumption, without intersecting γ_{i+1}. Since $\gamma_{i+1} = \varphi(\gamma_i) \cap M_{i+1}$ and $\beta' \subseteq \beta \subset M_{i+1}$, β' does not intersect $\varphi(\gamma_i)$ either. However, φ is a homeomorphism, and applying φ^{-1} to $\varphi(F_i^{\pm n})$, $\varphi(M_i)$, and β', we see that $\varphi^{-1}(\beta')$ is a path from F_i^{-n} to F_i^{+n} that remains in M_i and does not intersect γ_i. This contradicts our induction hypothesis. Thus, any path in M_{i+1} that connects F_{i+1}^{-n} to F_{i+1}^{+n} must intersect γ_{i+1}. To complete the proof, note that γ_{i+1} satisfies $\gamma_{i+1} \subseteq \varphi^{i+1}(\gamma_0)$, $\gamma_{i+1} \subset M_{i+1}$, and γ_{i+1} path disconnects F_{i+1}^{-1} from F_{i+1}^{+1} in M_{i+1}. ☐

Theorem 6 allows us to apply Theorem 3 to prove in this case that a shadow exists for any noisy trajectory $\{\mathbf{y}_i\}_{i=0}^N$ such that $\mathbf{y}_i \in M_i, \quad i = 0, \dots, N$.

3.3 Containment with zero contracting or zero expanding directions

For completeness, we mention the trivial cases in which all directions are contracting, or all directions are expanding. We call these the $(n, 0)$ and (n, n) cases, respectively. The former case is entirely trivial, because the problem is stable: if $\varphi(M_i) \subset M_{i+1}$ for all i, then clearly any exact solution starting in M_0 will be in M_i for all $i > 0$. Similarly, if all directions are expanding, then we apply the same argument in the reverse direction: if $\varphi^{-1}(M_{i+1}) \subset M_i$ for all i, then any exact solution *finishing* in M_N, traced backwards, lies in M_i for $i = N-1, N-2, \dots, 0$.

3.4 Discussion

The four cases $(n, 0)$, $(n, 1)$, $(n, n-1)$, and (n, n) cover *all* cases for $n = 1, 2, 3$. That is, the theorems in this paper can prove the existence of shadows for any n-dimensional system, $n \leq 3$, in which some measure of pseudo-hyperbolicity is present. Furthermore, although the proofs, for simplicity, only deal with a single function φ, the induction argument could just as easily use a *different* function φ_i at each step. In particular, φ_i could be the ODE time-h_i solution operator φ_{h_i}. Thus, modulo a rescaling of time (which we discuss below), the above proofs can be used to find shadows of noisy trajectories of ODE systems, as well as maps, with up to three dependent variables. They can also be used in the case of n dependent variables, with the restriction that solutions have either one expanding and $n-1$ contracting directions, or one contracting and $n-1$ expanding directions.

3.5 Proving the existence of chaotic orbits

Following the analysis of Stoffer and Palmer (1999), we describe how to use containment to prove the existence of chaotic orbits. We quote directly from their introduction.

> The idea is to construct two periodic pseudo-orbits which happen to be close to each other at some point. We call this the branching point. Then it is possible to construct an infinite number of pseudo-orbits as follows. You follow one or the other of the periodic pseudo-orbits. When you reach the branching point you either stay on your periodic orbit for at least one more loop or else you switch to the other periodic orbit. Each time you arrive at the branching point you can again choose to stay or to switch, *ad infinitum.* Assume that for each such pseudo-orbit there is a unique *orbit* of the system which is close to the pseudo-orbit. Then the dynamical system indeed behaves chaotically, at least in a certain neighbourhood of the two periodic pseudo-orbits. (Stoffer and Palmer 1999)

To use this approach together with containment to prove the existence of a chaotic orbit, assume the first orbit has a sequence of N parallelepipeds M_i satisfying (n,k)-ICP with $M_N = M_0$. Then the (n,k)-Inductive Containment Theorem can be invoked *ad infinitum* around this periodic pseudo-orbit, and proves the existence of infinitely long exact orbits that remain in the vicinity of this pseudo-orbit[6]. Similarly, assume the second orbit has a sequence of P parallelepipeds Q_i satisfying (n,k)-ICP with $Q_P = Q_0$. Assume further that $M_i = Q_j$ for some i,j. Then $M_i = Q_j$ is the branching point, and the (n,k)-Inductive Containment Theorem can be invoked *ad infinitum* around *both* of these pseudo-orbits, and each time we pass $M_i = Q_j$ we can choose which pseudo-orbit to follow. The (n,k)-Inductive Containment Theorem proves that a shadow follows us as we go.

[6]Slightly more is required to prove the existence of periodic orbits, or to prove uniqueness.

4 Two ways to verify the Inductive Containment Property

We present two methods of verifying that the general Inductive Containment Property holds for a given pseudo-trajectory derived from the numerical solution of an ODE. (Two more are presented in Hayes (2001).) We note in passing that both of these schemes (as well as the other two discussed in Hayes (2001)) could easily be adapted to the simpler problem of maps. Each has strengths and weaknesses, which we will discuss. Each one requires the use of interval arithmetic, or a validated ODE integrator if φ derives from an ODE. The validated ODE integrator that we use is called VNODE (Nedialkov 1999; Nedialkov, Jackson, and Corliss 1999). VNODE works with n-dimensional parallelepipeds, and satisfies the following property: given an n-dimensional parallelepiped A and a timestep h, VNODE will return an n-dimensional parallelepiped B such that $\varphi_h(A) \subset B$, where φ_h is the time-h solution operator. For the purposes of this description, we will denote the output B as $\bar{\varphi}_h(A)$. Thus,

$$\varphi_h(A) \subset \bar{\varphi}_h(A).$$

We will usually omit the timestep parameter h; we will talk only of φ, keeping in mind that in the induction, φ can be different for each step.

4.1 Direct integration of all $2n$ faces

The most direct method of building a pair of parallelepipeds M_i, M_{i+1} satisfying the Inductive Containment Property is to integrate each face of M_i individually, building M_{i+1} explicitly to satisfy the Inductive Containment Property. See Figure 7. A separate validated ODE integration is performed on each of the $2n$ faces of M_i, as shown in Figure 7(a). A face F is simply represented by an n-dimensional parallelepiped which has zero width in one particular dimension. The image of F under φ lies inside the box $\bar{\varphi}(F)$. In Figure 7(b), the vertical direction is depicted as expanding, while the horizontal direction is contracting. In the right half of Figure 7(b), the rectangles with thin solid boundaries depict $\bar{\varphi}(F)$ for each face F of M_i. The thin dotted box is the boundary of Q_{i+1}, which is a parallelepiped convex hull enclosing the images of all the faces, thus ensuring ICP(2i). M_{i+1} (heavy dashed line) is built as follows: for the expanding directions $j = \pm 1, \ldots, \pm k$, F_{i+1}^j is aligned strictly inside the inner boundary of $\bar{\varphi}(F_i^j)$, thus enforcing ICP(1). For the contracting directions $j = \pm(k+1), \ldots, \pm n$, F_{i+1}^j is aligned strictly outside the boundary of Q_{i+1}, thus enforcing ICP(2ii). Note that it is probably sufficient if the contracting faces of M_{i+1} are aligned strictly outside the outer boundary of $\bar{\varphi}(F_i^j), j = \pm(k+1), \ldots, \pm n$.

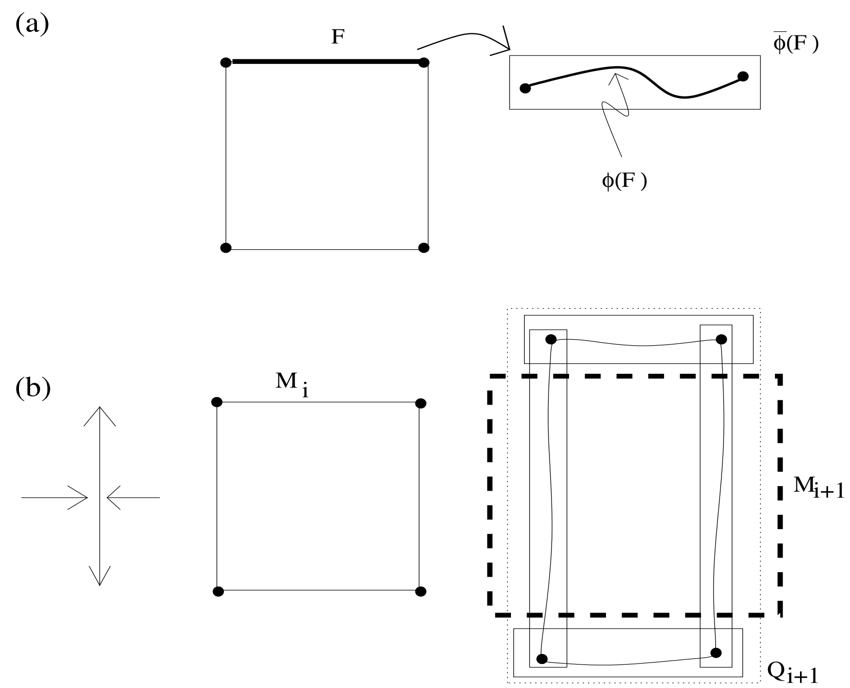


Figure 7: Schematic diagram of the direct $2n$ face-integration approach to proving the Inductive Containment Property. (a) An enclosure of the image of each individual face is computed. (b) From these enclosures, M_{i+1} is built. Q_{i+1} is a parallelepiped enclosing $\varphi(M_i)$ by enclosing the enclosures of all its faces. For simplicity, M_i, Q_{i+1} and M_{i+1} are all drawn axis-aligned, although in reality they may be arbitrarily oriented (as long as they are identically oriented).

This choice would give a slightly smaller M_{i+1}, and thus tighter containment, although it would complicate the proof of the Inductive Containment Theorem slightly.

As with any shadowing method, the system becomes hard to shadow when it lacks pseudo-hyperbolicity. This means either that the nominally contracting directions fail to contract enough for our method to detect contraction, or the nominally expanding directions fail to expand enough for our method to resolve the two faces opposite each other in that direction. In the former case, the width of M_i grows in the contracting directions as i increases, resulting in a growing shadowing distance. In the latter case, we cannot resolve the two opposite faces because the bounding boxes of their images overlap, as depicted in Figure 8.

The direct $2n$ face integration method was the first method of proving the Inductive Containment Property that we implemented in our code. However, we found that this method has several drawbacks. The most catastrophic problem is that the current implementation of VNODE is not designed to handle parallelepipeds that have some dimensions initially of zero thickness, and thus it is not always capable of providing tight enclosures of

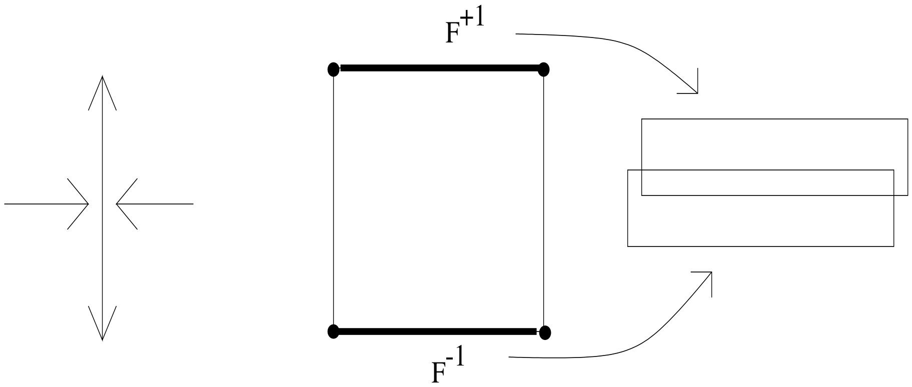


Figure 8: How the $2n$ direct face integration method can fail when not enough expansion occurs. Here, the vertical direction is the nominally expanding direction.

the images of faces. This problem is particularly bad if the initially zero width dimension lies along an expanding direction. Then, numerical errors very quickly compound to give an enclosure which is useless for containment purposes (see Figure 9). The amount by which

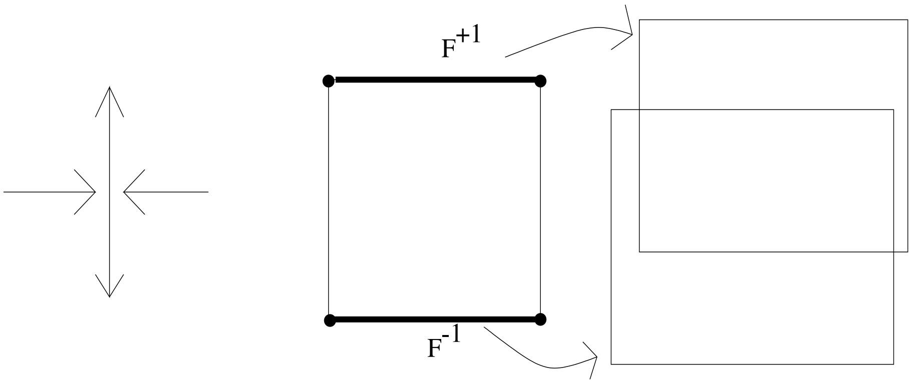


Figure 9: How the $2n$ direct face integration method *actually* fails, because the current implementation of VNODE cannot produce tight enclosures of the images of faces. Again, the vertical direction is the nominally expanding direction.

a validated ODE enclosure over-estimates the error is called the "excess". It is well-known that most current implementations of validated ODE integration have a large excess. It may be possible to decrease the excess for this particular application (Nedialkov, personal communication), although it may not be worth the effort given that we introduce below a more efficient and accurate method for proving ICP.

Finally, we note that the $2n$ direct face integration method is expensive: it requires $2n$ validated ODE integrations per shadow step; we will show below that there exist ways of verifying ICP with fewer validated integrations per shadow step. On the other hand, each of the $2n$ validated integrations is independent of all the others, so they could be performed in parallel.

4.2 Forward-backward iterative method

The above method can be used on arbitrary n-dimensional systems, but requires $2n$ validated ODE integrations for each containment step. Since validated ODE integrations are very expensive, we would like to find a way to ensure that the Inductive Containment Property holds using fewer integrations, especially for large n. The following method demonstrates that it is possible to verify ICP using an iterative method that we have found empirically to require about 3–4 validated integrations per step on average, independent of n. This method rigorously verifies ICP in the cases for which we have proven the Inductive Containment Theorem, and is the method we actually used to produce our numerical results.

We first look at the simple two-dimensional case in which one of the directions is expanding, while the other is contracting. To begin, assume that the only information provided by our validated ODE integration is an outer bound $\bar{\varphi}(M_i)$ on $\varphi(M_i)$. Then, it is *not* possible to verify the $(2,1)$-ICP with only one validated integration, because this information can only prove contraction, not expansion, as shown in Figure 10. In both Figures 10(a) and

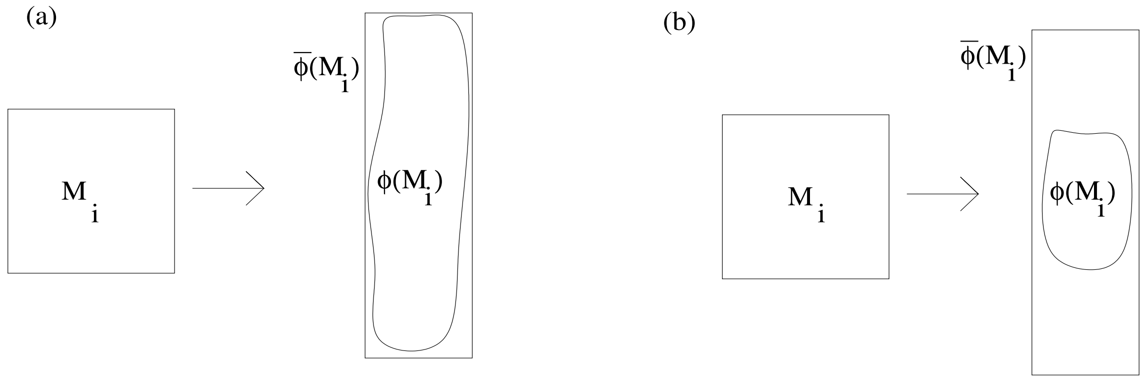


Figure 10: Enclosure methods can prove contraction, but not expansion.

10(b), $\bar{\varphi}(M_i)$ is a valid enclosure of $\varphi(M_i)$. In both Figures, $\bar{\varphi}(M_i)$ can be used to prove that $\varphi(M_i)$ has contracted in the horizontal direction. However, enclosure methods cannot directly prove expansion, as Figure 10(b) illustrates: although $\bar{\varphi}(M_i)$ is a valid enclosure of $\varphi(M_i)$, it is not a very good one, because the actual image $\varphi(M_i)$ of M_i has not expanded in any direction. To solve this problem, we perform two validated integrations. Refer to Figure 11(a). The first integration (solid rectangles) is a forward integration that provides $\bar{\varphi}(M_i)$,

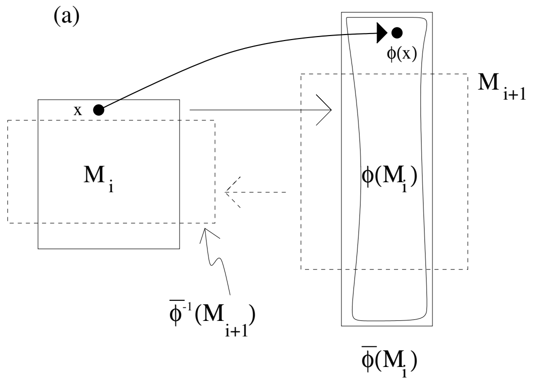


Figure 11: (a) The two validated integrations required to prove (2, 1)-ICP. (b) A potential problem, which is solved by doing a (cheap) point integration of one point on each expanding face, to verify there are points of $\varphi(M_i)$ on both side of M_{i+1}.

which in turn gives us a bound on the size of $\varphi(M_i)$ in the contracting directions (depicted as the horizontal direction in the Figure). Now, assume we can find an M_{i+1} which satisfies ICP not with $\varphi(M_i)$, but with $\bar{\varphi}(M_i)$. (If we cannot find such an M_{i+1}, then our method fails and we cannot prove the existence of a shadow beyond step i.) A validated integration *backwards* (dashed rectangles) is then performed on M_{i+1}, giving $\bar{\varphi}^{-1}(M_{i+1})$. If $\bar{\varphi}^{-1}(M_{i+1})$ proves that *contraction* has occured in the nominally expanding directions when moving back from M_{i+1} to M_i, then we argue that expansion in forward time has occurred, as follows. Choose any $\mathbf{x} \in M_i - \bar{\varphi}^{-1}(M_{i+1})$. Since $\mathbf{x} \notin \bar{\varphi}^{-1}(M_{i+1}) \supset \varphi^{-1}(M_{i+1})$, this implies $\varphi(\mathbf{x}) \in \varphi(M_i) - M_{i+1}$. Since $F_i^{\pm 1} \subset M_i - \bar{\varphi}^{-1}(M_{i+1})$, this tells us that $\varphi(F_i^{\pm 1}) \cap M_{i+1} = \emptyset$. This is insufficient to prove ICP(1), as illustrated in Figure 11(b): perhaps $\bar{\varphi}(M_i)$ is a loose enclosure of $\varphi(M_i)$, and all of $\varphi(M_i)$ is actually on one side of M_{i+1}. To verify that this is not the case, we pick one point on each of F_i^{+1} and F_i^{-1} and perform a validated point integration (which can be done cheaply) of each to verify that they land on opposite sides of M_{i+1}.[7] Since there is exactly one expanding direction, M_{i+1} cuts $\bar{\varphi}(M_i)$ into two disjoint sets, and a simple continuity argument shows that the two faces in their entirety land on opposite sides of M_{i+1}, thus verifying ICP(1). A similar argument in reverse time shows that the chosen M_{i+1} also verifies ICP(2ii).

The argument of the previous paragraph clearly applies just as well in n dimensions when there is one expanding direction and and $n-1$ contracting directions, for the same reasons that the two-dimensional proof of containment is easily transformed into Theorem 1

[7] We have found empirically that this problem must be very rare, because it has not happened even once during our experiments. We suspect that it may be possible to prove ICP without this extra point integration, but we have not devoted much thought to this matter.

(Hayes 2001). To prove that it also works when there is one contracting direction and $n-1$ expanding directions, note that there is a precise symmetry between the two cases (one expanding *vs.* one contracting): if we simultaneously reverse the order of $\{M_i\}_{i=0}^N$ giving $L_i = M_{N-i}$ and let $\psi = \varphi^{-1}$, then the above argument applies to the sequence $\{L_i\}_{i=0}^N$ using ψ as the homeomorphism. Thus, by symmetry, this method is also rigorous in the case that there is one contracting direction and $n-1$ expanding ones.

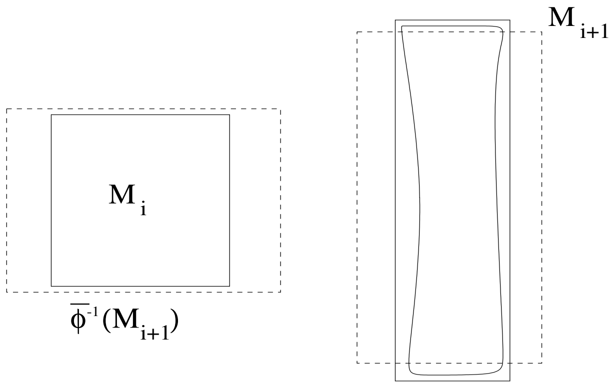


Figure 12: Shortcomings of the two-integration method: sometimes it can not prove expansion even if the M_{i+1} is valid.

Figure 12 illustrates that it is possible to choose an M_{i+1} that satisfies ICP, but for which we cannot *verify* ICP holds. This occurs when M_{i+1} is chosen to be "almost as large" as $\bar{\varphi}(M_i)$ in the expanding directions; then, the excess when computing $\bar{\varphi}^{-1}(M_{i+1})$ swamps the contraction that occurs when integrating the expanding direction backwards in time. We solve this problem by iteratively shrinking M_{i+1} in the nominally expanding directions until $\bar{\varphi}^{-1}(M_{i+1})$ fits inside M_i in those directions. If we shrink M_{i+1} to size zero in the expanding direction without being able to integrate it backwards to fit inside M_i, then the method fails, and we cannot prove the existence of a shadow beyond step i. We have found empirically that, when the algorithm is succeeding, no more than 2 to 3 backward integrations are usually required, independent of n. The number of backwards integrations is occasionally significantly larger, when the system encounters areas of non-hyperbolicity.

If the system were hyperbolic, then the nominally expanding directions would always expand, and the nominally contracting directions would always contract. However, in systems that are only pseudo-hyperbolic, the nominally expanding directions may expand most of the time, but not always; and vice versa for the contracting directions. One of the reasons our shadowing method can fail is if a nominally expanding direction contracts too much or for too long a time (Figure 13). Then, the expanding dimensions of M_i can

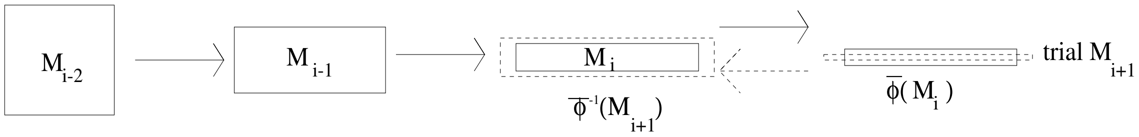


Figure 13: Example of the nominally expanding direction contracting too much for our integrator to prove contraction in the backwards direction.

become so small that no backwards integration from M_{i+1} can fit inside M_i in the nominally expanding directions.

We note that this method is *not* parallelizable across dimensions in the fashion that the previous method is. These and other methods are discussed in more detail in Hayes (2001).

4.3 Implementation issues and discussion

In the original paper that described containment, Grebogi, Hammel, Yorke, and Sauer (1990) appear to have used boxes M_i of fixed size, and found that smaller boxes seemed to work better. In contrast, our method dynamically grows and shrinks the M_i as i progresses, in an effort to maintain the Inductive Containment Property. In fact, we find it advantageous to choose the expanding dimension of M_i to be fairly large, to allow us to "absorb" possible future non-expansion, in an effort to avoid the situation depicted in Figure 13. Similarly, we choose the contracting dimensions to be relatively small, to avoid the opposite effect (allowing us to "absorb" non-contraction without the nominal contracting dimensions becoming too large). Practically, we find that our "boxes" can be extremely long and thin: typically, they are of length 10^{-3} to 10^{-6} in the expanding dimensions, and as small as 10^{-12} to 10^{-14} in the contracting dimensions.

Referring once again to Figure 13, we note that when containment fails, the "expanding" dimension of M_i has often shrunk to almost the same size as the contracting dimension, and both can be quite small (say, 10^{-12}), whereas when containment is "working", the expanding dimension of M_i can be several orders of magnitude larger than the contracting dimension. It is interesting to note that this implies that the hardest parts of an orbit to shadow are the places where our bounds on the distance between the noisy and shadow orbits are *smallest*, *i.e.,* where we can prove that they are unusually close together. This appears counter-intuitive, but may be related to the one-dimensional result of Chow and Palmer (1991), in which they proved that shadows must maintain a minimum distance from the noisy orbit.

5 Rescaling time

5.1 Informal description

Containment as presented thus far has put no restrictions on φ other than that it is a homeomorphism. As has also been mentioned, all of our theorems and proofs have been based on a single application of φ, and there is no explicit connection between the φ used at one step, and the one used on the next. Thus, everything said thus far is also applicable if we allow φ to change between steps. In particular, at each step we could use the time-h_i solution operator φ_{h_i} with h_i being the length of the ODE integration timestep taken at step i. The resulting method for shadowing numerical ODE integrations has been dubbed the *Map Method* by Coomes, Koçak, and Palmer (1994b, 1995a, 1995b). However, ODE integrations suffer from errors in time. For systems in which the $\mathbf{y}'$ direction lacks even pseudo-hyperbolicity, errors in time (which manifest themselves in phase space as errors in the $\mathbf{y}'$ direction) can lead to short shadowing times that can be dramatically increased if time is *rescaled*. In this section, we describe how containment can be augmented to allow for the rescaling of time.

Our idea for rescaling time in containment was inspired in part by the rescaling of time developed by Coomes, Koçak, and Palmer (1994b, 1995a) (although our proofs are very different from theirs), and partly by the idea of the *Poincaré section*, also known as a *Poincaré map* or *return map*. There are several variations on this idea, but the one that concerns us is the following. Assume that the solution to an ODE is "almost periodic", in the sense that the solution passes through some given plane $\mathcal{H}$ approximately every T time units, where $\mathcal{H}$ is approximately perpendicular to the trajectory at the point it crosses the plane. The Poincaré map generates the sequence of points where the trajectory intersects $\mathcal{H}$. To accomplish the general rescaling of time, we modify this idea to remove the almost-periodic requirement of the orbit, and simply place a plane $\mathcal{H}_i$ in the vicinity of the solution at time t_i, placed so that $\mathcal{H}_i$ is approximately perpendicular to $\mathbf{y}'(t_i)$.

To facilitate containment, we must extend the idea of the Poincaré section to encompass a small ensemble of solutions. To that effect, we wish to take a set $M_{i-1} \subset \mathcal{H}_{i-1}$, where the diameter of M_{i-1} is small, and place an $(n-1)$-dimensional hyperplane $\mathcal{H}_i$ approximately normal to the flow in the vicinity of $\varphi_{h_{i-1}}(M_{i-1})$. Then we define the Poincaré section of the set $\varphi_{h_{i-1}}(M_{i-1})$ pointwise as follows. Let Δh_{i-1} bound the time interval over which the ensemble $\varphi_{h_{i-1}}(M_{i-1})$ crosses $\mathcal{H}_i$:

$$\forall \mathbf{x} \in M_{i-1} \exists h \in [h_{i-1} - \Delta h_{i-1}/2, h_{i-1} + \Delta h_{i-1}/2] \text{ s.t. } \varphi_h(\mathbf{x}) \in \mathcal{H}_i,$$

where we assume that for each $\mathbf{x}$, the h chosen is unique. That is, we take the point-by-point Poincaré section of the points in M_{i-1} with respect to the plane $\mathcal{H}_i$. We call this a *splash* operation, because we imagine that the points in M_{i-1}, evolving via φ_h for $h \in [h_{i-1} - \Delta h_{i-1}/2, h_{i-1} + \Delta h_{i-1}/2]$, "splash" through $\mathcal{H}_i$ approximately simultaneously, and we assume that each trajectory intersects $\mathcal{H}_i$ precisely once during that interval. See Figure 14.

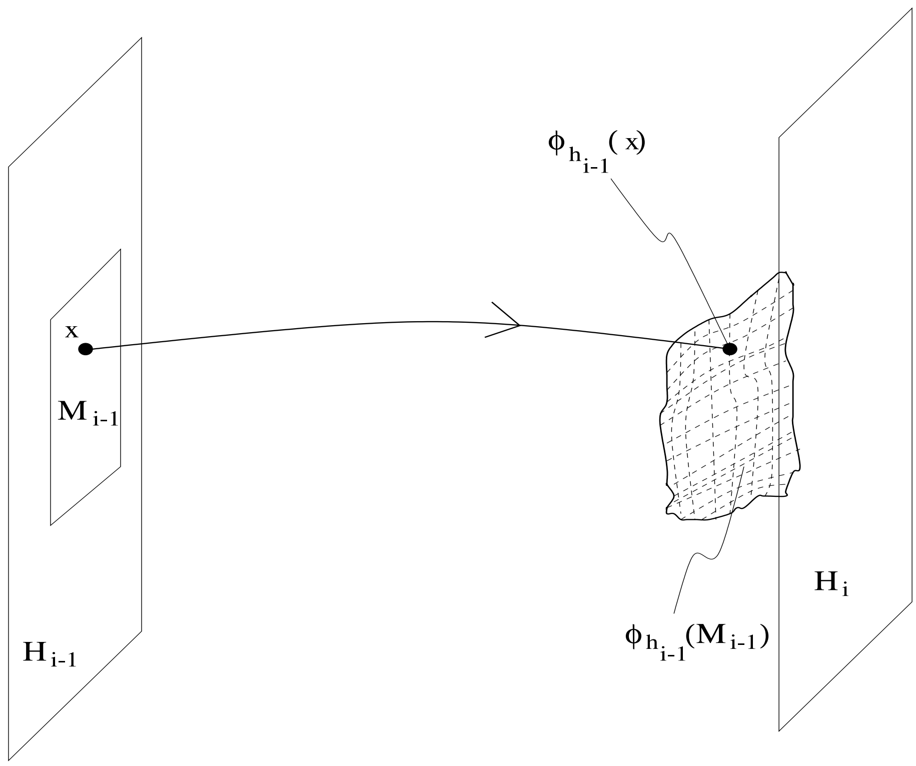


Figure 14: The "splash" operation depicted for a two-dimensional ensemble evolving in a three-dimensional configuration space. M_{i-1} is embedded in the plane $\mathcal{H}_{i-1}$, and evolves through one timestep to $\varphi_{h_{i-1}}(M_{i-1})$. As depicted, the ensemble is about to splash through $\mathcal{H}_i$.

Our intent is to build $(n-1)$-dimensional parallelepipeds M_i inside $\mathcal{H}_i$, and then show that the point-by-point Poincaré section at $\mathcal{H}_i$—the splash operation—is a homeomorphism. We can then directly apply the previously proved containment theorems to the $(n-1)$-dimensional M_i's which are each contained in the $(n-1)$-dimensional hyperplane $\mathcal{H}_i$, for an ODE system of n equations.

We note that since rescaling time via the splash operation effectively deletes one dimension from the problem, and our map containment theorems are rigorous in three dimensions, this means that the methods presented in this paper are capable of rigorously shadowing ODE solutions of up to four dimensions, as long as a rescaling of time is applied.

5.2 Theorem: splash is a homeomorphism

Refer to Figure 15. Let Q_i be an n-dimensional parallelepiped. Let $F_i^{\pm 1}$ be the two opposing faces of Q_i that are approximately normal to $\mathbf{y}'$ inside Q_i, and let $\mathbf{v}_i$ be the unit normal vector to these two faces, with $\mathbf{v}_i$ pointing from F_i^{-1} to F_i^{+1}. That is, $\mathbf{v}_i$ is approximately parallel to $\mathbf{y}'$ inside Q_i. Let D be the distance between F_i^{-1} and F_i^{+1} along $\mathbf{v}_i$. Let the infinite hyper-planes containing F_i^{-1} and F_i^{+1} be H_i^{-1} and H_i^{+1}, respectively, and let Z_i be the closed infinite slab between them. Let B_i be a parallelepiped with faces parallel to Q_i satisfying $Q_i \subset B_i \subset Z_i$, with two of the faces of B_i contained in $H_i^{\pm 1}$. Let $\{\mathbf{f}(\mathbf{x}) \cdot \mathbf{v}_i \mid \mathbf{x} \in B_i\} \subset [v_0, v_1]$, and assume $0 < v_0 \leq v_1$.

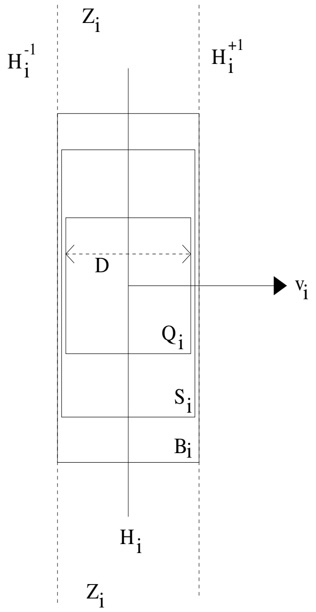


Figure 15: The objects used in Lemmas 7–10. Note that the left and right sides of Q_i, S_i, B_i, and Z_i are all in the planes H_i^{-1}, H_i^{+1}, respectively; they have been drawn distinct for illustrative purposes only.

Lemma 7. *If a trajectory remains in B_i while it is in Z_i, then it remains in Z_i for at least time $\underline{\varepsilon}_i^t \equiv D/v_1$ and at most $\bar{\varepsilon}_i^t \equiv D/v_0$.*

Proof. Let $\mathbf{y}(t)$ be a trajectory that remains in B_i while it is in Z_i. Let $z(t) = \mathbf{y}(t) \cdot \mathbf{v}_i$. Since $0 < v_0 \leq z'(t) \leq v_1$, and the width of B_i in the $\mathbf{v}_i$ direction is D, the maximum time to cross B_i is D/v_0, while the minimum time to cross is D/v_1. □

Let $\bar{\mathbf{f}}(B_i)$ be an enclosure of $\{\mathbf{f}(\mathbf{x}) \mid \mathbf{x} \in B_i\}$. Let S_i be a parallelepiped enclosure of $\{Z_i \cap (Q_i + h\bar{\mathbf{f}}(B_i)) \mid h \in [-\bar{\varepsilon}_i^t, \bar{\varepsilon}_i^t]\}$ and assume $S_i \subseteq B_i$.

Remark: S_i is intended to enclose how far a trajectory can drift from Q_i along the direction approximately perpendicular to $\mathbf{y}'$ as it travels across Z_i. This is required because a point in Q_i may not remain in Q_i when it is "splashed" onto $\mathcal{H}_i$. The following lemma formalizes this statement.

Lemma 8. *Any trajectory intersecting Q_i remains in S_i while in Z_i, and thus remains in B_i as well.*

Proof. Since $S_i \subseteq B_i$, $\bar{\mathbf{f}}(B_i)$ bounds $\mathbf{y}' \equiv \mathbf{f}$ inside S_i. Since a trajectory remaining in B_i as it crosses Z_i does so in time $\leq \bar{\varepsilon}_i^t$, and since $S_i \subset B_i$, $\{h\bar{\mathbf{f}}(B_i) \mid h \in [-\bar{\varepsilon}_i^t, \bar{\varepsilon}_i^t]\}$ encloses the maximum possible distance from Q_i that a trajectory can travel in time $|\bar{\varepsilon}_i^t|$ while it remains in B_i. Thus, since $Q_i \subset S_i \subseteq B_i$, $\{Q_i + h\bar{\mathbf{f}}(B_i) \mid h \in [-\bar{\varepsilon}_i^t, \bar{\varepsilon}_i^t]\}$ encloses the position of any trajectory $\mathbf{y}(t)$ that is within time $\bar{\varepsilon}_i^t$ of intersecting Q_i, unless $\mathbf{y}(t)$ leaves Z_i during that time. Intersecting with Z_i completes the proof. $\square$

Let $\mathcal{H}_i$ be any plane perpendicular to $\mathbf{v}_i$ which intersects the interior of Q_i. That is, $\mathcal{H}_i$ lies strictly between H_i^{-1} and H_i^{+1}.

Lemma 9. *Every trajectory intersecting Q_i intersects $\mathcal{H}_i$ at precisely one point while it crosses Z_i.*

Proof. Let $\mathbf{y}(t)$ be a trajectory that intersects Q_i. By Lemma 8, $\mathbf{y}(t)$ remains in $S_i \subseteq B_i$ while it crosses Z_i. Let $z(t) = \mathbf{y}(t) \cdot \mathbf{v}_i$. Let the z co-ordinates of $H_i^{-1}, \mathcal{H}_i, H_i^{+1}$ be z_{-1}, z_0, z_{+1}, respectively. While the trajectory remains in $S_i \subseteq B_i$, $z'(t) \geq v_0 > 0$, and, since $z(t)$ is continuous, it increases monotonically while $\mathbf{y}(t)$ remains in S_i, taking on each value between z_{-1} and z_{+1} precisely once, by the Intermediate Value Theorem. In particular, it takes on the value z_0 precisely once, and thus crosses $\mathcal{H}_i$ precisely once. $\square$

Assume Q_i is an enclosure of $\varphi_{h_{i-1}}(M_{i-1})$. For a point $\mathbf{x} \in M_{i-1}$, let $\varphi_{i-1}(\mathbf{x})$ be the unique point in $\mathcal{H}_i$ defined by Lemma 9. Let $\bar{M}_i = S_i \cap \mathcal{H}_i$. Clearly, $\bar{M}_i$ is an enclosure of $\varphi_{i-1}(M_{i-1})$. To show that φ_{i-1} applied to M_{i-1} is a homeomorphism, we need to show it is continuous and one-to-one. We first prove it is one-to-one.

Let $\varepsilon^t > 0$ be given. Recall $\bar{\varepsilon}_i^t$ as defined in Lemma 7.

Assumption 1: Assume $\bar{\varepsilon}_i^t < \varepsilon^t$ and $\nexists$ distinct $\mathbf{x}, \mathbf{y} \in M_{i-1}$ such that $\mathbf{y} = \varphi_t(\mathbf{x})$ for $|t| < \varepsilon^t$.

Each of the *Assumptions* introduced in this section is assumed to hold throughout the remainder of section, once it is introduced.

Lemma 10. *Each point in $\varphi_{i-1}(M_{i-1})$ comes from only one point in M_{i-1}.*

Proof. Assume to the contrary that there exist distinct $\mathbf{x}, \mathbf{y} \in M_{i-1}$ such that $\varphi_{i-1}(\mathbf{x}) = \varphi_{i-1}(\mathbf{y}) = \mathbf{z} \in \bar{M}_i$. Since $\varphi_{h_{i-1}}(\mathbf{x}), \varphi_{h_{i-1}}(\mathbf{y})$ both splash to $\mathbf{z}$, they are on the same trajectory, and since they are both in Q_i, the time-shift between them is $\leq \bar{\varepsilon}_i^t$. Thus, $\exists t_1, t_2$ such that $\varphi_{t_1}(\mathbf{x}) = \mathbf{z} = \varphi_{t_2}(\mathbf{y})$ with $|t_1 - t_2| \leq \bar{\varepsilon}_i^t$. Then $\mathbf{y} = \varphi_{t_1 - t_2}(\mathbf{x})$, contradicting Assumption 1. $\square$

Theorem 11. *φ_{i-1} applied to M_{i-1} is one-to-one.*

Proof. Lemma 9 proves that $\varphi_{i-1}(M_{i-1})$ is many-to-one, and Lemma 10 proves it is one-to-many. Thus, it is one-to-one. $\square$

We now prove that $\varphi_{i-1}(\mathbf{x})$ is continuous for all $\mathbf{x} \in M_{i-1}$.

Assumption 2: $\varphi_t(\mathbf{x})$ exists and is continuous in both t and $\mathbf{x}$ for all $\mathbf{x} \in M_{i-1}$ and all t such that $\varphi_t(\mathbf{x}) \in B_i$. Note that this is true as long as $\mathbf{f}$ is Lipschitz continuous (Stuart and Humphries 1996, Theorem 2.1.12).

We will need the use of the following theorem:

Theorem 12. *If $\mathbf{y}$ and $\mathbf{z}$ each satisfy the differential equation $\mathbf{y}'(t) = \mathbf{f}(\mathbf{y}(t))$ on the interval $[t_0, t_1]$ and $\mathbf{f}$ is Lipschitz continuous with constant L, then for all $t \in [t_0, t_1]$,*

$$\|\mathbf{y}(t) - \mathbf{z}(t)\| \leq \|\mathbf{y}(t_0) - \mathbf{z}(t_0)\| e^{L(t-t_0)}.$$

Proof. See Theorem 112J of Butcher (1987). $\square$

For a point $\mathbf{x} \in M_{i-1}$, let $h_{i-1}(\mathbf{x})$ be that unique timestep defined by $\varphi_{h_{i-1}(\mathbf{x})}(\mathbf{x}) \in \mathcal{H}_i$. That is, $\varphi_{i-1}(\mathbf{x})$ specifies *where* $\mathbf{x}$ goes, and $h_{i-1}(\mathbf{x})$ specifies how long it takes to get there.

Lemma 13. *If $\mathbf{f}$ is Lipschitz continuous, then for all $\mathbf{x}_0 \in M_{i-1}$, $h_{i-1}(\mathbf{x})$ is continuous at $\mathbf{x} = \mathbf{x}_0$.*

Proof. For simplicity, we will drop the subscript from $h_{i-1}(\mathbf{x})$ during this proof.

Let L be the Lipschitz constant for $\mathbf{f}$. Then by Theorem 12, for any $\mathbf{x}_0, \mathbf{x}$,

$$\|\varphi_{h(\mathbf{x}_0)}(\mathbf{x}) - \varphi_{h(\mathbf{x}_0)}(\mathbf{x}_0)\| \leq \|\mathbf{x} - \mathbf{x}_0\| e^{Lh(\mathbf{x}_0)} \equiv \delta_3(\mathbf{x}, \mathbf{x}_0).$$

Since we are only interested in the behaviour of $h(\mathbf{x})$ in a neighborhood of $\mathbf{x}_0$, choose $\mathbf{x} \in M_{i-1}$ close enough to $\mathbf{x}_0$ so that $\varphi_{h(\mathbf{x}_0)}(\mathbf{x}) \in B_i$. Now, since $\varphi_{h(\mathbf{x}_0)}(\mathbf{x}_0) \in \mathcal{H}_i$, the distance from $\varphi_{h(\mathbf{x}_0)}(\mathbf{x})$ to $\mathcal{H}_i$ is also bounded above by $\delta_3(\mathbf{x}, \mathbf{x}_0)$. Since $\varphi_{h(\mathbf{x}_0)}(\mathbf{x}) \in B_i$, the maximum time to intersect $\mathcal{H}_i$ is $\delta_3(\mathbf{x}, \mathbf{x}_0)/v_0$. Thus, $h(\mathbf{x}) \in [h(\mathbf{x}_0) - \delta_3(\mathbf{x}, \mathbf{x}_0)/v_0, h(\mathbf{x}_0) + \delta_3(\mathbf{x}, \mathbf{x}_0)/v_0]$. The continuity of $h(\mathbf{x})$ at $\mathbf{x}_0$ follows by letting $\mathbf{x} \to \mathbf{x}_0$. $\square$

Lemma 14. *$\varphi_{i-1}(\mathbf{x})$ is continuous for all $\mathbf{x} \in M_{i-1}$.*

Proof. By definition, $\varphi_{i-1}(\mathbf{x}) = \varphi_{h_{i-1}(\mathbf{x})}(\mathbf{x})$, and by construction, $\varphi_{h_{i-1}(\mathbf{x})}(\mathbf{x}) \in S_i \subseteq B_i$. Since the composition of two continuous functions is continuous and Lemma 13 asserts that $h_{i-1}(\mathbf{x})$ is continuous, Assumption 2 directly implies that $\varphi_{i-1}(\mathbf{x})$ is continuous. □

Thus, $\varphi_{i-1}(\mathbf{x}) \equiv \varphi_{h_{i-1}(\mathbf{x})}(\mathbf{x})$ is the unique splash point of $\mathbf{x}$ in $\mathcal{H}_i$.

Finally, the second part of Assumption 1 cannot be taken for granted. The following lemma is applied at step i to give us the second part of Assumption 1 at step $i+1$.

Let W_i be an infinite slab with width $E > D$ in the $\mathbf{v}_i$ direction, parallel to Z_i such that $Z_i \subset W_i$. Let C_i be a parallelepiped with sides parallel to Q_i, also with a width of E in the $\mathbf{v}_i$ direction, satisfying $M_i \subset C_i \subset W_i$, where M_i is built inside $\mathcal{H}_i$ to satisfy ICP with M_{i-1} under φ_{i-1}. Let $E_{+1} > 0$ be the distance from $\mathcal{H}_i$ in the $\mathbf{v}_i$ direction to the face of W_i, and let $E_{-1} > 0$ be the distance to the opposite face of W_i. Note that $E_{-1} + E_{+1} = E$. Let $\{\mathbf{f}(\mathbf{x}) \cdot \mathbf{v}_i \mid \mathbf{x} \in C_i\} \subset [u_0, u_1]$, and assume $0 < u_0 \leq u_1$. Let $\bar{\mathbf{f}}(C_i)$ be an enclosure of $\{\mathbf{f}(\mathbf{x}) \mid \mathbf{x} \in C_i\}$. Let T_i be a parallelepiped enclosure of $\{W_i \cap (M_i + h\bar{\mathbf{f}}(C_i)) \mid h \in [-\varepsilon^t, \varepsilon^t]\}$, and assume $T_i \subseteq C_i$.

Assumption 3: Assume $E/u_1 > \varepsilon^t$. That is, the minimum crossing time of C_i is greater than ε^t.

Lemma 15. *$\nexists$ distinct $\mathbf{x}, \mathbf{y} \in M_i$ such that $\mathbf{y} = \varphi_t(\mathbf{x})$ for $|t| < \varepsilon^t$.*

Proof. Substituting M_i for Q_i, W_i for Z_i, T_i for S_i, and C_i for B_i in Lemmas 7–9, we see that

1) If a trajectory remains in C_i while it is in W_i, then it remains in W_i for at least time E/u_1 and at most E/u_0. By a similar argument, the minimum and maximum times between such a trajectory entering C_i and intersecting $\mathcal{H}_i$ are E_{-1}/u_1 and E_{-1}/u_0, respectively; and the corresponding times between such a trajectory intersecting $\mathcal{H}_i$ and exiting C_i are E_{+1}/u_1 and E_{+1}/u_0.

2) Any trajectory intersecting M_i remains in T_i while it is in W_i, and thus remains in C_i.

3) Every trajectory intersecting M_i intersects $\mathcal{H}_i$ at precisely one point while it remains in W_i, where $\mathcal{H}_i \subset W_i$ and $\mathcal{H}_i$ is parallel to the planes enclosing W_i.

Thus, by point (3), to intersect $\mathcal{H}_i$ more than once inside M_i, a trajectory must, at least, first traverse the distance from $\mathcal{H}_i$ to ∂C_i, exit and then re-enter C_i, and traverse the distance from ∂C_i back to $\mathcal{H}_i$. By point (1), it takes time at least $E_{-1}/u_1 + E_{+1}/u_1 = E/u_1$ to do so. By Assumption 3, $E/u_1 > \varepsilon^t$. Thus, no trajectory can intersect M_i, exit T_i, and then re-enter T_i to again intersect M_i in time less than ε^t. $\square$

Remark: The base case of the induction is produced by substituting M_0 for M_i in Lemma 15, after building suitable W_0, C_0, and T_0.

5.3 Algorithmic details

Algorithmic verification of the requirements for the above theorems and lemmas is fairly straightforward: Q_i is simply the enclosure of $\varphi_{h_{i-1}}(M_{i-1})$ given to us by VNODE; the size of B_i is computed heuristically in an effort to ensure that $S_i \subseteq B_i$, and if our first guess is incorrect we simply increase its size until $S_i \subseteq B_i$, or fail if increasing the size of B_i results in $0 \in \{\mathbf{f}(\mathbf{x}) \cdot \mathbf{v}_i \mid \mathbf{x} \in B_i\}$; ε^t, which is an upper bound on the time error introduced at each step by the rescaling of time, must currently be pre-chosen by trial and error, although we believe that good, simple heuristics for choosing it probably exist. The sole complication is to maintain the property that Q_i has a pair of faces approximately normal to $\mathbf{y}'$ inside Q_i. Note that VNODE maintains a rotation matrix A_i which represents the orientation of the parallelepiped Q_i. Let the columns of A_i be $\mathbf{a}_i^j, j = 1, \ldots, n$. We simply assign $\mathbf{a}_i^1$ to be parallel to our best estimate of $\mathbf{y}'(t_i)$. VNODE then ensures that $\mathbf{a}_{i+1}^1$ evolves via the variational equation to be approximately parallel to $\mathbf{y}'(t_{i+1})$. To account for the slow buildup of error that would allow $\mathbf{a}_i^1$ to drift away from $\mathbf{y}'(t_i)$, we reset $\mathbf{a}_i^1$ to be parallel to the computed $\mathbf{y}'(t_i)$ at each timestep. This corresponds to rotating Q_i about its centre by a small angle θ, computed by solving

$$\cos(\theta) = \frac{\mathbf{a}_i^1 \cdot \mathbf{y}'(t_i)}{\|\mathbf{a}_i^1\| \, \|\mathbf{y}'(t_i)\|},$$

where $\mathbf{a}_i^1$ is the vector computed via evolution of the ODE from the previous timestep, and $\mathbf{y}'(t_i)$ is the value of $\mathbf{y}'$ computed directly from the right hand side of the ODE at the current timestep. The largest distance a point in Q_i will move as a result of this rotation is $r\theta$, where r is the distance of the furthest corner in Q_i from its centre. Thus, after rotating Q_i by θ, we increase its size by $r\theta$ in all directions, thus ensuring that it still encloses $\varphi_{h_{i-1}}(M_{i-1})$.

A simple variable stepsize algorithm was used: whenever containment of a particular step succeeds, we increase the stepsize by a small factor; whenever it fails, we decrease the stepsize by a factor of 2. We do not explicitly fail due to small stepsize, because too small a stepsize results in failures in other parts of the method, for example as depicted in Figure 13.

6 Results and discussion

In this section, we present results of our containment method for ODEs, compare our results to those of others, discuss some of the interesting implementation details of our method, and comment on observations of the behaviour of our method, including how it sometimes fails.

6.1 Quantitative comparisons with other methods

6.1.1 The Lorenz system of equations

The Lorenz equations (Lorenz 1963),

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \sigma(y-x) \\ \rho x - y - xz \\ xy - \beta z \end{pmatrix}, \tag{4}$$

define a dissipative dynamical system (*i.e.*, energy is not conserved) which was originally constructed to be a very simplified weather model. It can be shown (Coomes, Koçak, and Palmer 1995a) that under the Lorenz equations, the set

$$U = \{(x,y,z) : \rho x^2 + \sigma y^2 + \sigma(z-2\rho)^2 \leq \sigma\rho^2\beta^2/(\beta-1)\}$$

is *forward invariant*: any solution that is in U at time t_0 remains in U for all time $t \geq t_0$. All the methods discussed in this section solve the Lorenz equations with the classical parameter values $\sigma = 10, \rho = 28, \beta = 8/3$ (Lorenz 1963). It is easy to show that for these parameter values, the cube $[0,15]^3$ lies in U, and so for our experiments we chose initial conditions randomly inside this cube. A set of initial conditions in this cube will invariably produce a solution whose three-dimensional shape has been dubbed the "Lorenz butterfly" (Figure 16). Schematically, the Lorenz butterfly consists of two two-dimensional disks in three-space with a "bridge" between them. The two disks together are termed a "chaotic attractor", because solutions tend to remain in the disks, but jump chaotically from one to the other and back again. Solutions lack pseudo-hyperbolicity in the direction of the flow (Van Vleck 1995; Coomes, Koçak, and Palmer 1994b, 1995a), and so a rescaling of

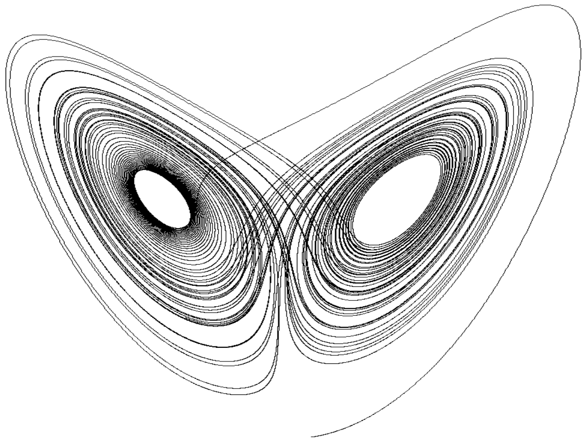


Figure 16: The "Lorenz Butterfly".

time is required to shadow them effectively. As should be clear from Figure 16 and the above description, in addition to the $\mathbf{y}'$ direction, at any given point a solution has one contracting direction, which is perpendicular to the disk currently housing the solution, and one expanding direction, directed radially from the centre of the disk. Provided a rescaling of time is employed, solutions to the Lorenz equations display remarkable pseudo-hyperbolicity for extremely long periods of time. Thus, this system is a prime first candidate for testing shadowing methods.

We will compare our results to the only other published results on shadowing the Lorenz equations using a rescaling of time: Van Vleck (1995), whose results could be made rigorous but currently are not; and Coomes, Koçak, and Palmer, (1994b, 1995a), whose results are completely rigorous.

First, with *no* rescaling of time (the "map method"), Van Vleck gives two examples of shadows with a local error[8] of about 10^{-5} lasting 1.04 and 1.38 time units; Coomes *et al.* have six examples with local error of about 10^{-13} lasting 9.7, 9.8, 9.9, 9.9, 86, and 126 time units. For this paper, we have simulated hundreds of shadows with various local errors. We have found that with local errors of about 10^{-5}, containment finds shadows that last between 1 and 30 time units, with a median and mean of about 20. With local errors of 10^{-13}, we find shadows lasting between 10 and 1000 time units, again with a mean and median about halfway through that range. Thus, it appears that, without a rescaling

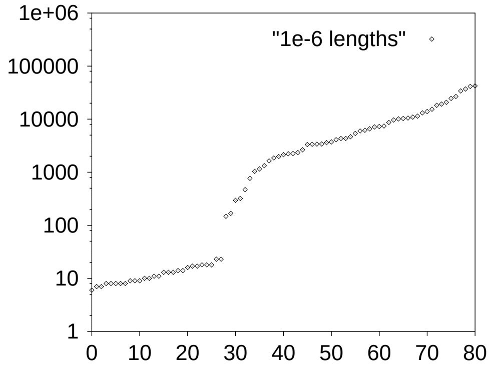
[8]All other methods used constant timesteps, and so the local errors are implicitly *per-unit-step*. The local errors used in the current paper were normalized to have comparable size per-unit-step, even though variable stepsize methods were used both for the validated ODE integration (Nedialkov 1999), and for choosing the size of shadow steps.

Author	local error	global error	Map Method	Rescaling Time
VV	10^{-6}	10^{-5}	1–2	$10^2 \sim 10^4$
Hayes	10^{-6}	10^{-5}	$10 \sim 50$	$10^3 \sim 10^5$
CKP	10^{-13}	10^{-9}	$10 \sim 100$	$\geq 10^5$
Hayes	10^{-13}	10^{-9}	$10 \sim 1000$	$\geq 7.7\times 10^5$

Table 1: Comparison of shadow lengths for the Lorenz system. VV=Van Vleck (1995); CKP = Coomes, Koçak, and Palmer (1994b, 1995a).

of time, the containment method is capable of finding shadows that are about an order of magnitude longer than other existing methods.

With a rescaling of time, Van Vleck gives many examples of shadows (with a local error of about 10^{-6}) ranging from 10^2 to 10^4 time units. Coomes *et al.* (with a local error of 10^{-13}) give six examples of shadows lasting at least 10^5 time units; they do not attempt finding longer shadows, so in fact their method may be capable of finding shadows longer than 10^5. The corresponding numbers for containment are 10^2 to 10^5 for local errors of 10^{-6}, and 10^2 to almost 10^6 for local errors of 10^{-13}. The results are summarized in Table 1.[9] It is clear that containment is at least as powerful as the other existing methods. It



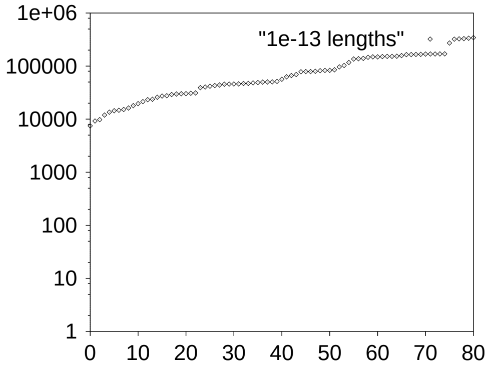


Figure 17: Distribution of shadow lengths computed by containment with a rescaling of time. Each Figure shows a sorted list of shadow lengths for 80 simulations of the Lorenz equations. The horizontal axis is simply a label for each shadow; the vertical axis is its length. The magnitude of the noise (*i.e.,* the local error) in the noisy orbits is about 10^{-6} in the left graph, and 10^{-13} in the right.

is worth noting that our results for local errors of 10^{-13} were produced using only a 17th order Taylor series, whereas Coomes *et al.* used a Taylor series of 31st order.

Figure 17 shows two sets of results of shadow lengths, including the rescaling of time.

[9] Our attempts to find the longest possible shadows for the latter case have been repeatedly confounded by having either workstation or disk crashes while our simulations were running. The longest shadow we've observed is thus 7.7×10^5, even though, had our machines not crashed, the shadows may have been longer.

The first is for eighty solutions with local error of approximately 10^{-6}, and the second for eighty solutions with local error of approximately 10^{-13}. The sharp increase in shadow lengths occuring just left of centre in the first Figure is probably due to the fact that, other than choosing $\mathbf{v}_0$ (cf. Figure 15 on page 26) to be parallel to $\mathbf{y}'(t_0)$, the directions of the faces of M_0 are currently chosen at random. As a result, we sometimes choose nominally expanding and contracting directions that are not sufficiently close to the actual expanding and contracting directions. Thus, many shadows fail early due to this problem. However, if our nominally chosen directions are (by luck) close enough to the actual ones, then we get over this hump to find much longer shadows. There is probably a more clever way to choose the initial M_0, but we have not yet studied this problem closely. This problem becomes less pronounced as the local error decreases, and is virtually absent in the right figure, which has local error $\delta = 10^{-13}$.

In addition, our shadowing distances (*i.e.,* the maximum distance between the shadow and the numerical trajectory) are comparable to the methods of the above authors: for orbits with noise 10^{-6} and 10^{-13}, our method and those of Van Vleck and Coomes, Koçak and Palmer find shadowing distances of approximately 10^{-5} and 10^{-9}, respectively. For containment, these sizes are based on ε^t and the maximum size of M_i over all i, which are at least in part user-controlled. For Van Vleck and Coomes *et al.* the shadowing distances are computed analytically based upon global bounds of various computed quantities.

6.1.2 Other systems of equations

We have reproduced the shadowing experiments of several other authors, usually getting comparable results, as illustrated in Table 2. We discussed results for the Lorenz system in the previous section. In this section, we provide results for three other problems.

Forced damped pendulum We first compare our results for the forced damped pendulum problem,

$$y'' + ay' + \sin y = b \cos t,$$

to those of Grebogi, Hammel, Yorke, and Sauer (1990), Sauer and Yorke (1991), and Chow and Van Vleck (1994). These authors use the values $a = 0.2, b = 2.4$ and $a = 1, b = 2.4$, with initial conditions $(y, y') = (0, 0)$, and mention that they get similar results with other pairs of values of a, b and initial conditions. We used the above two pairs of values for a, b and various random initial conditions in the unit square $[0, 1]^2$. We convert the second

System	Auth.	δ	ε	ε_t	L	Comment
Lorenz						
	VV	10^{-6}	10^{-5}		10^4	NR
	Hayes	10^{-6}	10^{-5}	2.5×10^{-5}	10^3–10^5	
	CKP	10^{-13}	10^{-9}		$\geq 10^5$	
	Hayes	10^{-13}	10^{-9}	2.5×10^{-9}	$\geq 7.7 \times 10^5$	
Forced Damped Pendulum						
	SY	10^{-18}	10^{-9}		3×10^4	high machine precision
	Hayes	10^{-15}	10^{-6}	10^{-3}	10^3–3×10^4	
	CVV	10^{-6}	10^{-3}		10^4	NR
	Hayes	10^{-6}	10^{-5}	10^{-3}	10^3	
	CVV	10^{-11}	10^{-8}		10^3	NR
	Hayes	10^{-11}	10^{-8}	10^{-3}	10^3	
Forced van der Pol						periodic attractor
	VV	10^{-5}	10^{-4}		10^4	NR
	Hayes	10^{-5}	10^{-6}	3×10^{-5}	$\geq 10^5$	
Logistic Equation						
	CVV	10^{-7}	5×10^{-6}		9.22	$y_0 = 0.01$, fixed L, NR
	Hayes	10^{-7}	10^{-6}		9.22	
	CVV	10^{-7}	5×10^{-6}		18.46	$y_0 = 10^{-4}$, fixed L, NR
	Hayes	10^{-7}	10^{-6}		18.46	

Table 2: Comparison of shadow lengths for four systems. For our results, the lengths shown are typical results after attempting many trials with the given local and global errors; the results of others are taken from their respective publications. Legend: δ = local error; ε = global space error; ε_t = global time error (if none is listed for our method then we did not rescale time); L = shadow length; CKP = Coomes, Koçak, and Palmer (1994b, 1995a); SY = Sauer and Yorke (1991); CVV = Chow and Van Vleck (1994a); VV = Van Vleck (1995); NR=not rigorous.

order equation to two first-order equations by assigning $y_1 = y, y_2 = y'$, giving

$$y_1' = y_2,$$

$$y_2' = b\cos t - \sin y_1 - a y_2.$$

Grebogi, Hammel, Yorke, and Sauer (1990) and Sauer and Yorke (1991) use extended precision arithmetic with a machine epsilon of 10^{-29} to generate a trajectory with local truncation error rigorously bounded by 10^{-18} per step, which allows them to find a shadow of length 3×10^4 and rigorous maximum distance 10^{-9} from their noisy trajectory. In comparison, we use standard IEEE754 floating-point numbers and arithmetic, and obtain a local truncation error of about 10^{-15} at best, so our shadow distances are significantly less stringent at 10^{-6}, and tend to be shorter, although in a few instances we successfully found shadows of length $\sim 3 \times 10^4$. Given that Sauer and Yorke used higher precision, we are not surprised that our shadows tend to be shorter and not as close as theirs. Comparing our results to Chow and Van Vleck (1994), we see our method is capable of rigorously proving the existence of a shadow which is closer, but lasts for a shorter time, than their method does; on the other hand, our result is rigorous, whereas theirs is not, because they do not rigorously bound numerical errors before applying their theorem.

The primary problem with shadowing this system appears to be that it is non-autonomous. We currently handle a non-autonomous system by converting it to an autonomous system with one component of our solution, y_0, representing time: $y_0(0) = t_0,\ \ y_0'(t) = 1$. This has several drawbacks: (1) the new component is decidedly non-hyperbolic; (2) assuming we can solve the linear system $y' = 1$ exactly, the interval representing y_0 then accumulates roundoff error and as time progresses, the error in y_0 grows; (3) this is exacerbated by the minimum absolute error in y_0 increasing as $\varepsilon_{mach} t$, where ε_{mach} is the machine precision; (4) finally, the error in the computation of $\cos(y_0)$ adds to the error. These drawbacks, however, do not seem to adequately explain our poor shadowing results for this system. Perhaps the difficulties would vanish if a native procedure for validated integration of non-autonomous systems were used, or if we used higher precision, as did Sauer and Yorke (1991).

Forced van der Pol The forced van der Pol equation,

$$x'' + \alpha(x^2 - 1)x' + x = \beta\cos(\omega t),$$

is studied by Van Vleck (1995). He defines the parameters implicitly with $\alpha = k = \sigma = 2/5$, where $k = \beta/(2\alpha)$ and $\sigma = (1-\omega^2)/\alpha$, and uses the initial conditions $(x, x') = (0, 0)$. We try this initial condition, as well as others chosen randomly in the unit square $[0, 1]^2$, and we convert the second order equation to two first-order equations by assigning $y_1 = x, y_2 = x'$, giving

$$y_1' = y_2,$$

$$y_2' = \beta \cos(\omega t) - (y_1^2 - 1)\alpha y_2 - y_1.$$

This equation has a hyperbolic periodic attractor which all solutions approach asymptotically, and so this system is easy to shadow. With a local truncation error of 10^{-6}, Van Vleck found numerical shadows of length 10^4 and distance 10^{-4}, while we went significantly further, finding rigorous shadows lasting 10^5 and longer with a distance of 10^{-6}. Since solutions asymptotically approach a periodic solution that is hyperbolic, we conjecture that containment could be maintained indefinitely.

Logistic equation Finally, the logistic equation,

$$y' = y(1-y), \quad y(0) = \zeta,\ 0 < \zeta \ll 1,$$

was studied by Chow and Van Vleck (1994). In this problem, there is an unstable fixed point at $y = 0$ and a stable fixed point at $y = 1$. Chow and Van Vleck attempt shadowing two solutions, both starting at $y(0) = \zeta$ and integrating until $y(T) \approx 1 - \zeta$. If $\zeta = 10^{-2}$, then $T \approx 9.22$, and if $\zeta = 10^{-4}$, then $T \approx 18.46$. In both cases, we use a local truncation error of $\delta = 10^{-7}$. We find that we easily match their results, noting again that ours are rigorous, while theirs are not. In fact, we find that we can prove the existence of these shadows for $\varepsilon \approx 10\delta$ for δ down to about 10^{-14}.

6.2 Qualitative comparisons with other methods

First and foremost, our method has only been proven to work in a limited number of special cases. Generalizing the (n, k)-Inductive Containment Property to arbitrary (n, k) is straightforward (Hayes 2001). Proving that it implies the existence of a shadow is more difficult, and is in progress. See Hayes (2001) for more discussion.

Although containment is rigorous, it appears to be less robust than non-rigorous methods. For example, in two examples out of three, the non-rigorous results of Chow and Van Vleck (1994) produced shadows that were about an order of magnitude longer than we could produce using containment. In addition, Hayes (1995) presented convincing evidence that the gravitational n-body problem is shadowable, and yet containment could prove the existence of shadows lasting only 1% as long as those found nonrigorously in Hayes (1995). Even worse, the VNODE package (Nedialkov 1999) is capable of providing a validated enclosure of an IVP for the n-body problem which is about ten times as long as the containment-produced shadow! Clearly, if an enclosure of an IVP exists, then a shadow exists for the associated point solution for at least as long. Thus, at least for some problems, our implementation of containment is incapable of finding shadows even though they exist. This does not necessarily imply that the theorems proved in Section 2.2 are deficient; it probably means that our implementation for verifying that the Inductive Containment Property holds can be improved, for example by reducing the excess of the validated numerical integrator.

Our method requires some *a priori* guesses; for example, the maximum and minimum sizes of the M_i, and the maximum time rescaling ε^t need to be chosen before the algorithm runs. We typically had to choose these numbers by trial and error for each problem; if a certain ε^t did not work, for example, we often found that increasing it or decreasing the maximum size of M_i would allow us to find longer or closer shadows, respectively. Van Vleck's (1995) method also requires some *a priori* guesswork to make a rescaling of time work. Although Coomes, Koçak and Palmer do not discuss their choice of parameters, it is likely that they require significant guesswork to find parameters that satisfy their theorems as well. Finally, *all* shadowing methods currently in the literature appear to require guesswork to discover the number of expanding and contracting dimensions, and to choose a local error δ which is stringent enough to satisfy their respective theorems.

It is also not trivial to see how containment could be parallelized, since each M_i depends on M_{i-1}. Possibly an iterative method that guesses all the $\{M_i\}_{i=0}^{N}$ and then iteratively refines them in parallel could be constructed; this may also be related to two-point boundary value problems (Ascher, Mattheij, and Russell 1988).

On the other hand, containment appears to have several advantages over other methods.

- First and foremost, the method of proof is simple and easy to understand.

- We use an (almost) off-the-shelf validated integrator (Nedialkov 1999) to verify that ICP holds; this integrator is almost as easy-to-use as any standard integrator, and thus getting the code "up and running" on a new problem usually takes only a few minutes. Another advantage of this simplicity is that it requires the user to have no deeper understanding of the system than knowing the defining equations.[10]

- Although the success of containment may depend, of course, upon global properties of the system, the method itself is local. By that we mean that it requires information only from the previous step to extend the length of the shadow. Several other methods require computing, storing, and updating global information such as the extent of non-hyperbolicity (cf. Chow and Palmer's p parameter (1991, 1992)).

7 Conclusions

We have extended the simple and elegant *containment* method of producing shadows from two-dimensional maps to maps of arbitrary dimension in which some measure of hyperbolicity is present and there is either 0 or 1 expanding modes, or 0 or 1 contracting modes, and added a rescaling of time to allow containment to work better for ODEs. We have demonstrated that this new method produces shadows of ODE integrations that are of comparable quality and length to any currently in the literature, and noted how it can be used to prove the existence of chaos.

References

Ascher, U. M., R. M. M. Mattheij, and R. D. Russell (1988). *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*. Prentice-Hall Series in Computational Mathematics. Prentice-Hall.

Beyn, W.-J. (1987). On Invariant Closed Curves for One-Step Methods. *Numer. Math. 51*, 103–122.

Butcher, J. C. (1987). *The numerical analysis of ordinary differential equations*. Wiley.

Chow, S.-N. and K. J. Palmer (1991). On the numerical computation of orbits of dynamical systems: the one-dimensional case. *Dynamics and Differential Equations 3*, 361–380.

Chow, S.-N. and K. J. Palmer (1992). On the numerical computation of orbits of dynamical systems: the higher dimensional case. *Journal of Complexity 8*, 398–423.

Chow, S.-N. and E. S. Van Vleck (1994, July). A Shadowing Lemma Approach to Global Error Analysis for Initial Value ODEs. *SIAM Journal on Scientific Computing 15*(4),

[10] Some may consider this a disadvantage.

959–976.

Coomes, B. A. (1997, January). Shadowing orbits of ordinary differential equations on invariant submanifolds. *Transactions of the American Mathematical Society 349*(1), 203–216.

Coomes, B. A., H. Koçak, and K. J. Palmer (1994a). Periodic shadowing. In P. Kloeden and K. Palmer (Eds.), *Chaotic Numerics*, Volume 172 of *Contemporary Mathematics*, pp. 115–130. Amer. Math. Soc., Providence, RI.

Coomes, B. A., H. Koçak, and K. J. Palmer (1994b). Shadowing orbits of ordinary differential equations. *Journal of Computational and Applied Mathematics 52*, 35–43.

Coomes, B. A., H. Koçak, and K. J. Palmer (1995a). Rigorous computational shadowing of orbits of ordinary differential equations. *Numerische Mathematik 69*, 401–421.

Coomes, B. A., H. Koçak, and K. J. Palmer (1995b). A shadowing theorem for ordinary differential equations. *Z angew Math Phys (ZAMP) 46*, 85–106.

Coomes, B. A., H. Koçak, and K. J. Palmer (1997). Long periodic shadowing. *Numerical Algorithms 14*, 55–78.

Corless, R. M. (1994). Error Backward. *Contemporary Mathematics 172*, 31–62.

Dahlquist, G. and Å. Björck (1974). *Numerical Methods*. Prentice-Hall series in Automatic Computation. Prentice-Hall.

Grebogi, C., S. M. Hammel, J. A. Yorke, and T. Sauer (1990, 24 September). Shadowing of Physical Trajectories in Chaotic Dynamics: Containment and Refinement. *Physical Review Letters 65*(13), 1527–1530.

Hammel, S. M., J. A. Yorke, and C. Grebogi (1987). Do Numerical Orbits of Chaotic Dynamical Processes Represent True Orbits? *Journal of Complexity 3*, 136–145.

Hammel, S. M., J. A. Yorke, and C. Grebogi (1988). Numerical Orbits of Chaotic Dynamical Processes Represent True Orbits. *Bull. Am. Math. Soc. 19*, 465–470.

Hayes, W. (1995, January). Efficient Shadowing of High Dimensional Chaotic Systems with the Large Astrophysical N-body Problem as an Example. Master's thesis, Dept. of Computer Science, University of Toronto. Available on the web at `http://www.cs.toronto.edu/NA/reports.html#hayes95` and via anonymous ftp from `ftp.cs.utoronto.ca/pub/reports/na/hayes-95-msc`.

Hayes, W. B. (2001). *Rigorous shadowing of numerical solutions of ordinary differential equations by containment*. Ph. D. thesis, Department of Computer Science, University of Toronto.

Kahaner, D., C. Moler, and S. Nash (1989). *Numerical Methods and Software*. Prentice-Hall series in Computational Mathematics. Prentice-Hall.

Lorenz, E. N. (1963, March). Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences 20*(2), 130–141. Also reprinted in *Chaos* (1984), by Hao Bai-Lin, World Scientific Publishing Co., Singapore.

Munkres, J. R. (1975). *Topology: a first course*. Prentice-Hall, Englewood Cliffs, New Jersey.

Murdock, J. (1995). Shadowing multiple elbow orbits: an application of dynamical systems to perturbation theory. *J. Diff. Eq. 119*(1), 224–247.

Nedialkov, N. S. (1999). *Computing rigorous bounds on the solution of an initial value problem for an ordinary differential equation*. Ph. D. thesis, Department of Computer Science, University of Toronto.

Nedialkov, N. S., K. R. Jackson, and G. F. Corliss (1999). Validated solutions of initial value problems for ordinary differential equations. *Applied Mathematics and Computation 105*(1), 21–68.

Palmer, K. J. (1988). Exponential Dichotomies, the Shadowing Lemma and Transversal Homoclinic Points. In U. Kirchgraber and H. O. Walther (Eds.), *Dynamics Reported*, Volume 1. Wiley and Teubner.

Quinlan, G. D. and S. Tremaine (1992). On the reliability of gravitational N-body integrations. *Monthly Notices of the Royal Astronomical Society 259*, 505–518.

Sauer, T. and J. A. Yorke (1991). Rigorous Verification of Trajectories for the Computer Simulation of Dynamical Systems. *Nonlinearity 4*, 961–979.

Stoffer, D. and K. J. Palmer (1999). Rigorous verification of chaotic behaviour of maps using validated shadowing. *Nonlinearity 12*, 1683–1698.

Stuart, A. M. and A. R. Humphries (1996). *Dynamical Systems and Numerical Analysis*. Cambridge University Press.

Van Vleck, E. S. (1995). Numerical Shadowing Near Hyperbolic Trajectories. *SIAM Journal on Scientific Computing 16*(5), 1172–1189.

Van Vleck, E. S. (2000). Numerical shadowing using componentwise bounds and a sharper fixed point result. *SIAM Journal on Scientific Computing 22*(3), 787–801.