NUMERICAL ORBITS OF CHAOTIC PROCESSES REPRESENT TRUE ORBITS

STEPHEN M. HAMMEL, JAMES A. YORKE AND CELSO GREBOGI

1. Introduction. It is the nature of chaotic processes of the form $x_{n+1} = f(x_n)$, where $f\colon \mathbf{R}^d \to \mathbf{R}^d$, that different orbits starting close together will move apart rapidly. When following an orbit numerically, a common accuracy is about 14 digits. For chaotic systems such as the logistic map (or in two-dimensions, the Hénon map), distances between two nearby orbits on average grow geometrically on every iterate. For example, it is not unusual that the distance roughly doubles on every iterate. At that rate two true orbits starting 10^{-14} apart will be more than 1 unit apart after just 50 iterates: the error will be of the same order of magnitude as the variables themselves. The implication is that roundoff error on just the first step is sufficient to destroy totally the ability to predict just 50 iterates later.

While a numerical orbit will diverge rapidly from the true orbit with the same initial point, there often exists a different true orbit with a slightly different initial point which stays near the noisy orbit for a long time. We have developed rigorous numerical procedures to prove there exists a true orbit which stays near the noisy orbit of a given chaotic process for a long time.

We begin by defining the shadowing property. The term pseudo-orbit is used to describe a numerically generated noisy orbit.

DEFINITION. $\{p_n\}_{n=a}^b$ is a δ_f-pseudo-orbit for f if $|p_{n+1} - f(p_n)| < \delta_f$ for $a \le n \le b$, where f is a D-dimensional map and p is a D-dimensional vector representing the dynamical variables. We are particularly interested in the case where a and b are finite.

DEFINITION. A *true orbit* $\{x_n\}$ satisfies $x_{n+1} = f(x_n)$.

DEFINITION. The true orbit $\{x_n\}$ δ_x-*shadows* $\{p_n\}$ on $[a,b]$ if $|x_n - p_n| < \delta_x$.

For comparison of different dynamical systems, we use the convention that each coordinate has been rescaled to the interval $[-1, 1]$ or the square $[-1, 1] \times [-1, 1]$ before the quantities δ_f and δ_x are evaluated.

2. Two-dimensional shadowing. In this work, we investigate shadowing in two-dimensional diffeomorphisms. The first system to be examined is the Hénon map, which is not uniformly hyperbolic. Anosov and Bowen [1, 2] proved shadowing results for uniformly hyperbolic systems. The systems we study are not hyperbolic. See [3, 5, 6] for one-dimensional shadowing

---

results. Typical pseudo-orbits cannot be shadowed for all time. We shall use the following form of the Hénon map, with $\mathbf{x}_n = (u_n, v_n)$:

$$(1) \qquad \mathbf{x}_{n+1} = \mathbf{f}(\mathbf{x}_n) = \begin{pmatrix} u_{n+1} \\ v_{n+1} \end{pmatrix} = \begin{pmatrix} 1 - Au_n^2 + v_n \\ -Ju_n \end{pmatrix},$$

where J is the determinant of the Jacobian matrix of the map. We examine this map for $J = -0.3$, and for values of A and initial conditions (u_0, v_0) which yield a positive Lyapunov exponent along the orbit.

The pseudo-orbits we generate, and hence the theorems we state, are independent of the arithmetic round-off technique used by a particular computer. To insure this, we define a truncation operator $\mathbf{T}(\mathbf{x})$ which truncates each coordinate of the vector $\mathbf{x}$ to the B most significant bits. We study $B = 48$. The specific orbits we will analyze are of the form $\mathbf{p}_{n+1} = \mathbf{T}(\mathbf{f}(\mathbf{p}_n))$. In practice, this machine-independent pseudo-orbit is determined by having the computer find an approximation to $\mathbf{f}(\mathbf{p}_n)$ using 96-bit precision arithmetic. This approximation is machine-dependent, but we have an upper bound on the size of the error in the evaluation of $\mathbf{f}$. Truncation of this vector to B-bit precision results in a machine-independent vector provided that the error in the 96-bit evaluation of $\mathbf{f}(\mathbf{p}_n)$ does not affect the truncated 48-bit result. Therefore each such vector is checked before truncation to verify that the errors do not corrupt the value of $\mathbf{T}(\mathbf{f}(\mathbf{p}_n))$.

Our goal is simply to generate numerical orbits and then to calculate rigorously *how close* a true orbit is, and to obtain lower bounds for *how long* it stays close.

THEOREM 1. *Let $\mathbf{f}$ be defined by (1), with $A = 1.4$ and $\mathbf{p}_0 = (0,0)$. For $N = 10^7$, the δ_f-pseudo-orbit $\{\mathbf{p}_n\}_{n=0}^N$ with $\mathbf{p}_{n+1} = \mathbf{T}(\mathbf{f}(\mathbf{p}_n))$, is δ_x-shadowed by a true orbit within $\delta_x = 10^{-8}$.*

When the Hénon pseudo-orbit is rescaled to the region $[-1, 1] \times [-1, 1]$, we can show $\delta_f = 2^{-46} \simeq 10^{-14}$. The long shadowing time is striking when compared to the great rate at which orbits of this system diverge from each other. The values A and J are the most commonly investigated values. We have chosen the initial point $(0,0)$ to appear to be "generic": we have obtained comparable results for a variety of other initial conditions and parameter values.

We have a similar result for the Ikeda map. This map is a two-dimensional dissipative map describing the electromagnetic field within a laser ring cavity. It is convenient to formulate it as a map of the complex plane to itself: $\mathbf{g} \colon \mathbf{C} \to \mathbf{C}$. Thus, $\mathbf{z}_n \in \mathbf{C}$ represents both the amplitude and phase of the field at a fixed point in the ring cavity, and

$$(2) \qquad \mathbf{z}_{n+1} = \mathbf{g}(\mathbf{z}_n) = \gamma + R\mathbf{z}_n \exp\left[i\left(\kappa - \frac{\alpha}{1 + |\mathbf{z}_n|^2}\right)\right]$$

gives the field after one round trip of the signal within the cavity. Further details about the map and its properties, including relevant parameter values, can be found in [4].

THEOREM 2. *Let* $\mathbf{g}$ *be defined by* (2), *with* $\alpha = 5.5$, $\gamma = 0.85$, $\kappa = 0.4$, $R = 0.9$, *and* $\mathbf{p_0} = (0,0)$. *For* $N = 10^6$, *the pseudo-orbit* $\{\mathbf{p_n}\}_{n=0}^N$ *with* $\mathbf{p}_{n+1} = \mathbf{T}(\mathbf{g}(\mathbf{p_n}))$, *is* δ_x-*shadowed by a true orbit within* $\delta_x = 10^{-8}$.

In order to indicate the relative magnitudes of the quantities we expect to find, we present the following summary:

CONJECTURE. *For a typical dissipative map* $\mathbf{f} \colon \mathbf{R}^2 \to \mathbf{R}^2$ *with a positive Lyapunov exponent and a small noise amplitude* $\delta_f > 0$, *we expect to find* $\delta_x \leq \sqrt{\delta_f}$ *for an orbit length* $N \simeq 1/\sqrt{\delta_f}$.

OUTLINE OF PROOF OF THEOREMS. Our goal is to show that a true orbit $\{\mathbf{x_n}\}$ exists near the pseudo-orbit $\{\mathbf{p_n}\}$. We construct a sequence of small parallelograms A_n (near $\mathbf{p_n}$) within which this true orbit must reside. It will follow from a topological argument that there exists a true orbit $\{\mathbf{x_n}\}$ such that $\mathbf{x_n} \in A_n$ for all n. We require that for each $n \geq 1$, $\mathbf{f}(A_n)$ maps across A_{n+1} as shown in Figure 1, and in particular: (i) $A_{n+1} \cap \mathbf{f}(A_n) \neq \varnothing$, and (ii) each A_n has distinguished sides C_n^1 and C_n^2 such that $\mathbf{f}(A_n) \cap C_{n+1}^i \neq \varnothing$, while $\mathbf{f}(C_n^i) \cap A_{n+1} = \varnothing$, $i = 1, 2$. The construction of the parallelograms will depend on the stable and unstable unit vectors, denoted $\hat{\mathbf{c}}_n$ and $\hat{\mathbf{e}}_n$, of a numerical orbit. These unit vectors are computed using the fact that on the average, the map is contracting along stable directions $\{\hat{\mathbf{c}}_n\}$, and expanding along unstable directions $\{\hat{\mathbf{e}}_n\}$.
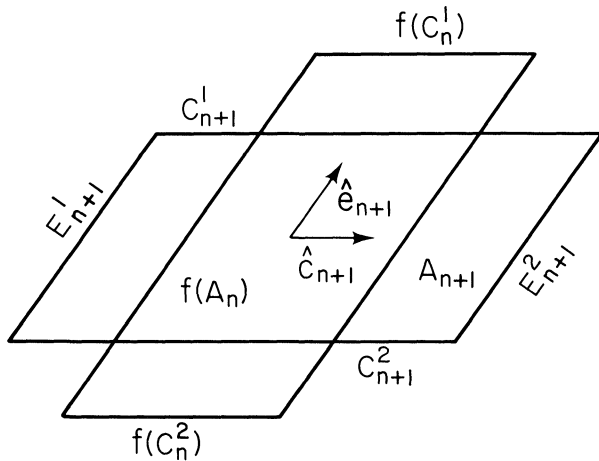

FIGURE 1. The numerical procedure verifies that $\mathbf{f}$ maps A_n across A_{n+1}.

It is then possible to define

(3)
$$J_0 = \bigcap_{j=0}^N \mathbf{f}^{-j}(A_j) \neq \varnothing.$$

There exists $\mathbf{x_0} \in J_0$ such that $\mathbf{x_n} = \mathbf{f}^n(\mathbf{x_0}) \in A_n$, $0 \leq n \leq N$.

Using 96-bit arithmetic, a less-noisy orbit $\{\tilde{\mathbf{p}}_n\}$ is constructed which remains uniformly near $\{\mathbf{p}_n\}$. The point $\tilde{\mathbf{p}}_n$ will be the center of A_n. Let $\mathbf{\Pi}_n$ represent the one-step error

$$(4) \qquad\qquad \mathbf{\Pi}_{n+1} = \mathbf{p}_{n+1} - \mathbf{f}(\mathbf{p}_n),$$

where by definition $|\mathbf{\Pi}_n| < \delta_f$. The refined orbit $\{\tilde{\mathbf{p}}_n\}$ is constructed by setting $\tilde{\mathbf{p}}_n = \mathbf{p}_n + \mathbf{\Phi}_n$, where $\mathbf{\Phi}_n$ is constructed so as to remain small and to satisfy

$$(5) \qquad\qquad \mathbf{\Phi}_{n+1} \simeq L_n \mathbf{\Phi}_n - \mathbf{\Pi}_{n+1},$$

where L_n is the derivative of $\mathbf{f}$ evaluated at $\mathbf{p}_n$.

The sequence $\{\mathbf{\Phi}_n\}$ is found as follows. Using the unit vectors $\{\hat{\mathbf{e}}_n\}$ and $\{\hat{\mathbf{c}}_n\}$, we can then represent $\mathbf{\Phi}_n$ and $\mathbf{\Pi}_n$ as: $\mathbf{\Phi}_n = \alpha_n \hat{\mathbf{e}}_n + \beta_n \hat{\mathbf{c}}_n$, and $\mathbf{\Pi}_n = \eta_n \hat{\mathbf{e}}_n + \varsigma_n \hat{\mathbf{c}}_n$. We use (5) to find approximate values for the coefficients $\alpha_n, \beta_n, \eta_n$, and ς_n:

$$(6) \qquad \mathbf{\Phi}_{n+1} = L_n(\alpha_n \hat{\mathbf{e}}_n + \beta_n \hat{\mathbf{c}}_n) - (\eta_{n+1}\hat{\mathbf{e}}_{n+1} + \varsigma_{n+1}\hat{\mathbf{c}}_{n+1}).$$

In (4), $\mathbf{f}(\mathbf{p}_n)$ can be well approximated by using $\tilde{\mathbf{f}}(\mathbf{p}_n)$, the numerical image of $\mathbf{p}_n$ using 96-bit arithmetic. Thus, $\{\eta_n\}$ and $\{\varsigma_n\}$ can be calculated for each n directly, and $\{\alpha_n\}$ and $\{\beta_n\}$ can be determined recursively using

$$(7) \qquad \alpha_{n+1} = |L_n\hat{\mathbf{e}}_n|\alpha_n - \eta_{n+1}, \qquad \beta_{n+1} = |L_n\hat{\mathbf{c}}_n|\beta_n - \varsigma_{n+1}.$$

This is made computationally stable by calculating the coefficients β_n for the stable direction $\hat{\mathbf{c}}_n$ by starting at the initial point $n = 0$, and the coefficients α_n in the unstable direction $\hat{\mathbf{e}}_n$ starting at the end point $n = N$:

$$(8) \qquad \begin{aligned} \alpha_n &= \frac{\alpha_{n+1} + \eta_{n+1}}{|L_n\hat{\mathbf{e}}_n|}, \qquad \alpha_N = 0, \\ \beta_{n+1} &= \beta_n|L_n\hat{\mathbf{c}}_n| - \varsigma_{n+1}, \qquad \beta_0 = 0. \end{aligned}$$

The refined pseudo-orbit $\{\tilde{\mathbf{p}}_n\}$ is less noisy than the original pseudo-orbit $\{\mathbf{p}_n\}$. There is no a priori guarantee that this procedure will work for arbitrary initial conditions, although it does work in the cases reported here.

Finally we can indicate how the numerical computations insure the proper behavior of the sequence of parallelograms A_n. The boundary of the parallelogram A_n consists of two edges C_n^i parallel to the unit vector $\hat{\mathbf{c}}_n$, and two edges E_n^i parallel to the unit vector $\hat{\mathbf{e}}_n$, as shown in Figure 1. Thus, the boundary of the parallelogram is defined by the 4 line segments: $\partial A_n = E_n^1 \cup E_n^2 \cup C_n^1 \cup C_n^2$. They are centered on the points $\tilde{\mathbf{p}}_n$. $\tilde{\mathbf{f}}(\mathbf{x})$ will denote the higher-precision image (using a 96-bit mantissa in the arithmetic) of $\mathbf{x}$ under the map $\mathbf{f}$, and $\tilde{\mathbf{f}}$ is assumed to be $\tilde{\delta}_f$-noisy, that is, $|\tilde{\mathbf{f}}(\mathbf{x}) - \mathbf{f}(\mathbf{x})| < \tilde{\delta}_f$. In our computations we are able to keep the A_n so small that $|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\tilde{\mathbf{p}}_n) - D\mathbf{f}(\tilde{\mathbf{p}}_n)(\mathbf{x} - \mathbf{p}_n)| < 2^{-92}$ for $\mathbf{x} \in A_n$.

Recall that $\mathbf{f}$ refers to the true image of A_n. For two sets A and B, let $d_{\inf}(A, B) = \inf\{d(x, y) \colon x \in A \text{ and } y \in B\}$. To insure that the qualitative intersection shown in Figure 1 occurs, we shall construct $\{A_n\}$ subject to the constraints: $d_{\inf}(\tilde{\mathbf{f}}(E_n^i), E_{n+1}^i) > \tilde{\delta}_f$ and $d_{\inf}(C_n^i, \tilde{\mathbf{f}}^{-1}(C_{n+1}^i)) > \tilde{\delta}_f$, for $i, j = 1, 2$. Similar to the calculation of coefficients $\{\alpha_n\}$ and $\{\beta_n\}$ for the

refinement process above, the sides $\{E_n\}$ are recursively tested under forward iteration, while the sides $\{C_n\}$ are tested under iteration by the inverse map. These conditions on the construction of A_n insure that the *true* image of A_n overlaps A_{n+1} correctly, as shown in Figure 1.

Since there is a true orbit in $\{A_n\}$, it is straightforward to compute the distance of the furthest point of A_n from $\mathbf{p}_n$, and thence to compute the maximum of these distances over all $\mathbf{p}_n$. This yields an upper bound on δ_x.

REFERENCES

1. D. V. Anosov, *Geodesic flows and closed Riemannian manifolds with negative curvature*, Proc. Steklov Inst. Math. **90** (1967).

2. R. Bowen, *ω-limit sets for Axiom A diffeomorphisms*, J. Differential Equations **18** (1975),

3. S. M. Hammel, J. A. Yorke, and C. Grebogi, *Do numerical orbits of chaotic dynamical processes represent true orbits?*, J. of Complexity **3** (1987), 136–145.

4. S. M. Hammel, C. K. R. T. Jones, and J. V. Moloney, *Global dynamical behaviour of the optical field in a ring cavity*, J. Opt. Soc. Amer. B **2** (1985), 552–564.

5. E. Coven, I. Kan and J. A. Yorke, *Pseudo-orbit shadowing in the family of tent maps*, Trans. Amer. Math. Soc. (to appear).

6. H. Nusse and J. A. Yorke, *Is every approximate trajectory of some process near an exact trajectory of a nearby process?*, preprint.

LABORATORY FOR PLASMA RESEARCH, UNIVERSITY OF MARYLAND, COLLEGE PARK, MARYLAND 20742 (address of S. M. Hammel and Celso Grebogi)

INSTITUTE FOR PHYSICAL SCIENCE AND TECHNOLOGY AND DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MARYLAND, COLLEGE PARK, MARYLAND 20742 (address of J. A. Yorke)