PERGAMON

Chaos, Solitons and Fractals 12 (2001) 1305–1322

CHAOS
SOLITONS & FRACTALS

www.elsevier.nl/locate/chaos

Convergence analysis of Davidchack and Lai's algorithm for finding periodic orbits

Aaron Klebanoff *, Erik Bollt

Mathematics Department, United States Naval Academy, 572C Holloway Rd, Annapolis, MD 21402-5002, USA

Accepted 3 May 2000

Communicated by L.E. Reichl

Abstract

We rigorously study a recent algorithm due to Davidchack and Lai (DL) [Davidchack RL, Lai Y-C. Phys Rev E 1999;60(5):6172–5] for efficiently locating complete sets of hyperbolic periodic orbits for chaotic maps. We give theorems concerning sufficient conditions on convergence and also describing variable sized basins of attraction of initial seeds, thus pointing out a particularly attractive feature of the DL-algorithm. We also point out the true role of involutary matrices which is different from that implied by Schmelcher and Diakonos [Schmelcher P, Diakonos FK. Phys Rev E 1998;57(3):2739–46] and propagated by Davidchack and Lai.

1. Introduction

Periodic orbits have been rightly called by Cvitanovic [8–10] the "skeleton" of chaos. It was Poincaré [11] who first noted the central role played by periodic orbits in chaotic dynamics. A well-regarded definition of chaos [12] requires the existence of an infinite number of periodic orbits, but periodic orbits are useful for more than just defining chaos. The "periodic orbit theory" provides that periodic orbits embedded within a chaotic attractor are useful in statistical characterizations of the attractor to describe such quantities as average position, escape rates from hyperbolic repellers [13], and other averages of invariant measure by cycle expansions [8–10]. Such summation formulae are closely related to the trace of transfer operator formulae [14]. For Axiom A diffeomorphisms, for example, Bowen [31] proved that the asymptotic growth rate of the number of periodic orbits is determined by topological entropy, and he went on to prove several important results concerning the limit distribution of periodic orbits for such systems [30,32–34]. Similarly, Grebogi et al. [15] showed that the invariant measure of hyperbolic systems can be estimated in terms of an appropriately weighted distribution of periodic orbits, and the formulae are conjectured to hold for non-hyperbolic systems [16,17]. There have been similarly motivated rigorous studies which prove that periodic orbits determine the behavior of a system, including that of Katok [35] concerning the asymptotic distribution and strength of hyperbolicity at periodic orbits in determining Lyapunov exponents and entropy, and likewise discussions can be found in de la Llave [36] as well as in Ruelle [41], of the fact that SRB measures can be characterized by Lyapunov exponents at periodic orbits in a weak limit with increasing period. Once we have invariant measure, other quantities such as fractal dimension, and Lyapunov

* Corresponding author. Permanent address: Mathematics Department CM 144, Rose-Hulman Institute of Technology, 5500 Wabash Avenue, Terre Haute, IN 47803, USA.

E-mail addresses: aaron.klebanoff@rose-hulman.edu (A. Klebanoff), bollt@usna.edu (E. Bollt).

0960-0779/01/$ - see front matter
PII: S0960-0779(00)00099-0

exponents can also be expressed in terms of limits involving periodic orbits. Similarly, a periodic orbit measure (POM) property has been proven [18] for certain one-dimensional maps ("eventually onto maps") that invariant measure is the weak limit of unweighted averages over subsequences of the periodic orbits of increasing order, and the property is conjectured to hold for more general dynamical systems.

Topological descriptions of the chaotic attractor are also accessible given the periodic orbits. For example, a large set of periodic orbits is highly constraining to the symbolic dynamics, and it has been shown that such a set can be used to infer a generating partition [19,20]. It has also been shown [21–25] that the periodic orbits can be used to extract topological invariants via knot theoretic techniques, by associating an appropriate template; this approach has also been used to extract a generating partition from periodic orbits. Such analysis based on periodic orbits is physical, having been successfully demonstrated for a laser system [21–25]. Symbolic dynamics analysis based on template analysis of periodic orbits has also been experimentally demonstrated for an NMR laser [26] and similar techniques have been used for Belousov–Zhabotinskii reaction data [4]. In the pruning front theory introduced by Cvitanovic [8–10], and made rigorous by de Carvalho [37], some of the main ideas of kneading theory [42] are generalized to certain homeomorphisms of the plane. That is, a two-parameter parameterized curve $P(a,b)$ through a symbolic plane representing all possible symbolic sequences of a full Smale horseshoe map [38] completely determines the set of allowed orbits found in the non-wandering set of the Henon maps $H_{a,b}$, and in particular, the complete set of allowed periodic orbits is determined. Furthermore, there is an analog of the kneading theory concept of universal family, by which considering a candidate (periodic) orbit ordered relative to the pruning fronts, one can understand topological bifurcation of such dynamical systems, from the simple, toward the complicated, as the parameters (a,b) are varied.

Centrally important to accurately estimating invariant properties of a dynamical system in terms of periodic orbits is a good algorithm to collect complete collections of such orbits, since even a few missing periodic orbits are expected to skew results. We will focus our discussion on discrete dynamical systems $x_{k+1} = f(x_k)$, $f : \mathbb{R}^n \to \mathbb{R}^n$, which includes flows when considered by Poincaré surface of section. Finding a period-p point, $x = f^p(x)$ is equivalent to finding a root of $g(x) = f^p(x) - x$. While low ordered periodic points are not difficult to find, say by Newton's method, higher ordered periodic points are numerically sensitive, which is obvious considering two facts: (a) the derivative Df^p becomes large with large p; (b) the number of periodic orbits is expected to grow asymptotically as the exponent of topological entropy. Techniques exist to extract periodic orbits from chaotic time-series [4,26,27] often based on close recurrences. Such methods are useful for time-series via the time-delay embedding technique, and reasonable when we want one or several periodic orbits, but are not reliable when it comes to collecting complete collections of periodic orbits.

In this paper, we will give complete convergence analysis of a recent and efficient algorithm to find such complete sets of periodic orbits, but first we give some history. In 1989, Biham and Wenzel [28,29] introduced an algorithm which in principle finds all periodic orbits of an arbitrary high period, but the technique is highly specialized and usually only applied to the Hénon map. In 1997, Schmelcher and Diakonos [2] introduced a generally applicable algorithm which apparently finds complete sets of periodic orbits of arbitrary period. Schmelcher and Diakonos showed that a (matrix) C could be found so that

$$x_{k+1} = x_k + hCg(x_k) \tag{1}$$

would converge to any saddle point given a small enough number h. The iteration scheme (1) can be understood as Euler's method for solving the ordinary differential equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = Cg(x). \tag{2}$$

Fixed points for the differential equation are stable if Cg has eigenvalues of negative real part, so if h is chosen small enough, C can be chosen appropriately. Different choices of C stabilize points that were saddles for the original map so that one has a better chance of finding periodic orbits for the original map. Schmelcher and Diakonos report finding complete sets of periodic orbits for various maps [2,3].

To improve efficiency, Davidchack and Lai [1] used an *almost* implicit Euler method for the differential equation (2). The implicit Euler routine for this ODE is

$$x_{k+1} = x_k + hCg(x_{k+1}).$$

We estimate $g(x_{k+1}) \approx g(x_k) + J(x_k)(x_{k+1} - x_k)$ to yield

$$x_{k+1} = x_k + hCg(x_k) + hCJ(x_k)(x_{k+1} - x_k)$$

and define $\beta := 1/h$ to get

$$x_{k+1} - x_k = (\beta I - CJ_k)^{-1} Cg_k.$$

Davidchack and Lai used

$$x_{k+1} - x_k = (\beta \|g_k\| I - CJ_k)^{-1} Cg_k, \tag{3}$$

adding the $\|g_k\|$ term which makes $\mathrm{DH}(\bar{x}) = 0$ at any fixed point $\bar{x}$. This method retains the qualitative nature of the differential equation (2) in that we can predict which C will be best for certain types of fixed points based on the eigenvalues of Cg.

The one-dimensional case sheds some light on the behavior of the iteration scheme (3). If $\beta = 0$, then the algorithm is Newton's method. This modified Newton's method still has the benefit of being super-attracting for all $\beta \geqslant 0$ and $C = \pm 1$, i.e., if $g(\bar{x}) = 0$, then $H(\bar{x}) = \bar{x}$ and $H'(\bar{x}) = 0$. Unlike Newton's method, we can focus on roots of different stability characteristics by means of the "switching parameter" C as is demonstrated in Fig. 1(a). By increasing the "tuning" parameter β, we enlarge the basin of attraction for the roots focussed on by C, and decrease the measure of the set in which orbits of H diverge while increasing computation time (see Fig. 1(b)).

An appealing feature of this algorithm is its apparent ability to efficiently find complete sets of periodic orbits. There are two factors that affect its efficiency. First is the quadratic convergence in the neighborhood of a root which is due to the fact that if x is near a root of g, then $\|g\| \approx 0$ and the algorithm is approximately Newton's method. The second factor is the choice of the seeds. The majority of the computation

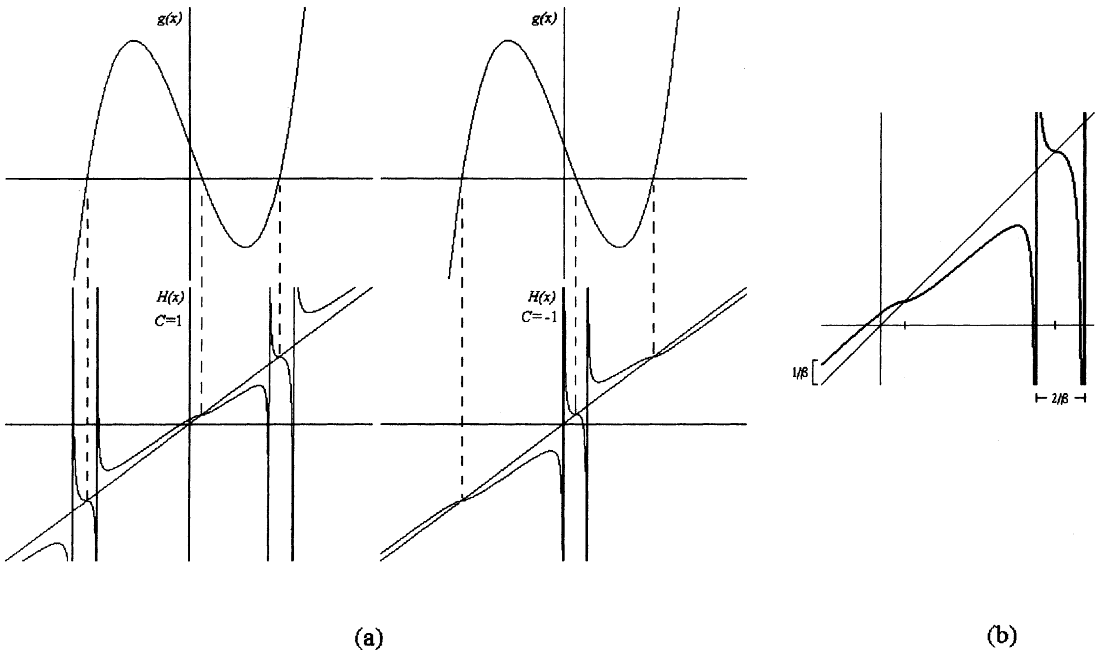

Fig. 1. (a) Roots $\bar{x}$ satisfying $g'(\bar{x}) < 0$ have a large basin of attraction when $C = 1$. Switching to $C = -1$ enlarges the basin of attraction for roots satisfying $g'(\bar{x}) > 0$. (b) For the case $C = 1$ and for β sufficiently large, the diameter of the basin of attraction for roots satisfying $g'(\bar{x}) > 0$ is $2/\beta$. Furthermore, close to roots satisfying $g'(\bar{x}) < 0$, $H(x)$ is bounded within $1/\beta$ of x which implies linear convergence rates far from the root.

time is spent finding periodic orbits that were already found or failing to converge at all. Therefore, choosing seeds wisely can reduce computation time many times over but choosing seeds unwisely can yield an incomplete set of periodic orbits and/or increase computation time.

The purpose of this paper is to give a rigorous convergence analysis of the Davidchack-Lai (DL)-algorithm, which we consider to be an attractive and efficient general new algorithm to find periodic orbits. Perhaps the two most important issues in choosing a good numerical procedure are good seeding and stability. It is well known that Newton's method can present very complicated basins of attraction, and hence studying global convergence properties of even single variable (complex) functions is difficult, with major work in this area by Smale [39,40], and can itself even present chaotic dynamics [43]. The layout of this paper is as follows. In Section 2, we review the DL-algorithm. Then in Section 3, we give our main convergence theorem which shows sufficient conditions for the DL-algorithm to converge, and then also show the main importance of this algorithm is its abilities to control the size of basins of attraction. We give several examples in Section 4, which show both strengths and failings of this algorithm, as well as various kinds of general difficulties which can arise in searching for periodic orbits. Finally, in Appendix A, we discuss the role played by the switching matrices, which is not in complete agreement with [2].

2. DL-algorithm

Assume that the map

$$f(x) : \mathbb{R}^n \to \mathbb{R}^n \tag{4}$$

has a chaotic orbit and that subsequently, it has a dense set of periodic orbits. Define

$$g = g(x) = f^p(x) - x, \tag{5}$$

and for typographical clarity, define

$$s = s(x) = \|g(x)\| \quad \text{and} \quad J = J(x) = Dg = \frac{\partial g}{\partial x}. \tag{6}$$

We will assume that orbits of all periods exist so that all periodic points of period-$p-1$ are used as seeds. An algorithm [1] which finds roots of $g(x)$ is as follows:

1. Find all fixed and period-2 points (in general, this can be done by using a fine-mesh grid to seed Newton's method).
2. List the $2^n n!$ $n \times n$ orthogonal "switching" matrices C with only ± 1 non-zero entries (for example, see Table 1).
3. Using each point of every distinct periodic orbit of period-$p-1$ as a seed x_0, choose a matrix C from Table 1 and a number $\beta = \beta_1 \geqslant 0$ and generate the sequence $\{x_k\}$ using the *iteration scheme*

$$x_{k+1} = H(x_k), \tag{7}$$

Table 1
All possible values for C in one and two dimensions

1D	2D	
$C_1 = 1$	$C_1 = I$	$C_5 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
$C_2 = -1$	$C_2 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$	$C_6 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$
	$C_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	$C_7 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$
	$C_4 = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$	$C_8 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

where

$$H(x) = x + [\beta s I - CJ]^{-1} Cg \tag{8}$$

and I is the usual identity matrix.
4. If the sequence $\{x_k\}$ converges to a root of g, i.e., a periodic point of f, then iterate f to find a complete period of the new-found orbit.
5. Repeat the last two steps for each matrix in Table 1.
6. Repeat the last three steps for increasing values of β until we can find no more.

Remark 1. Since $[\beta s I - CJ]^{-1} Cg = O(\beta^{-1})$ for β large, we expect the number of iterations required for convergence to depend linearly on β. Experience suggests that very few convergent sequences are lost by the following rule: assume the sequence will not converge if it fails to satisfy $\|x_{k+1} - x_k\| < \varepsilon$ for some small number ε with $k \leqslant \max\{100, 4\beta\}$. Even if we throw out a sequence that might have converged to a periodic orbit not yet found, it is possible for another seed to yield the same periodic orbit.

Remark 2. We always found a critical β^* such that for all $\beta \geqslant \beta^*$ no more periodic orbits were found as shown in [1]. Since we expect the number of periodic orbits to grow exponentially with the period, we also expect β^* to grow exponentially with the period. So, we can experimentally determine β^* for relatively low periods, and then determine a rule for choosing β^* for high periods.

Remark 3. If there are no period-$p - 1$ points, Davidchack and Lai [1] proposed the following scheme. Use an incomplete set of period-p points (found, say, by random seeding) to find period-$p + 1$ points. Then use the period-$p + 1$ points as seeds to find the rest of the period-p points which in turn are used to find the rest of the period-$p + 1$ points.

Remark 4. We found examples (see Section 4) where the DL-algorithm fails to find period-3 orbits using the period-2 orbits as seeds. Although the period-3 orbits are relatively easy to find by other means, this gives us reason to question the algorithm's ability to find complete sets of periodic orbits as claimed. At issue is whether the basins for convergence to *all* the period-p points contain the period-$p - 1$ points.

Remark 5. If the fixed points for the vector field $g : \mathbb{R}^2 \to \mathbb{R}^2$ has no eigenvalues with zero real part, then there is always a C taken from $\{C_1, \ldots, C_6\}$ so that the corresponding fixed point for the vector field Cg has all negative real parts. So, for efficiency we can omit C_7 and C_8 from our list in step 2 of the algorithm.

3. Convergence theorems

In this section, we begin by showing sufficient conditions under which the DL-algorithm is guaranteed to converge to a fixed point. We then show that in one dimension, this sufficient basin grows with increasing β. First some preliminary results are required.

Lemma 1 ([5], p. 253). *If $J(x)$ exists for all x in a convex region $X \subseteq \mathbb{R}^n$, and if a constant γ exists with*

$$\|J(x) - J(y)\| \leqslant \gamma \|x - y\| \quad \text{for all } x,\ y \in X,$$

then for all x, $y \in X$, the estimate

$$\|g(x) - g(y) - J(y)(x - y)\| \leqslant \frac{\gamma}{2} \|x - y\|^2$$

holds.

Definition 1. We call an $n \times n$ orthogonal matrix C with only ± 1 non-zero entries a *switching* matrix.

Lemma 2. *There are $2^n n!$ $n \times n$ switching matrices. Furthermore, $\|C\| = 1$ and $\|Cv\| = \|v\|$ for any $v \in \mathbb{R}^n$.*

Proof. Standard counting arguments give the first part. Since C is an orthogonal matrix, we have $C^{\mathrm{T}} = C^{-1}$. So,

$$\|C\|^2 = \max_{\|v\|=1} \|Cv\|^2 = \max_{\|v\|=1} v^{\mathrm{T}} C^{\mathrm{T}} C v = \max_{\|v\|=1} v^{\mathrm{T}} v = 1.$$

Furthermore, since $\|C\| = \|C^{\mathrm{T}}\| = 1$, for any $v \in \mathbb{R}^n$, we have

$$\|v\| = \|C^{\mathrm{T}} C v\| \leqslant \|C^{\mathrm{T}}\| \cdot \|Cv\| \leqslant \|Cv\| \leqslant \|v\|,$$

implying $\|Cv\| = \|v\|$. □

Definition 2. We say that C stabilizes J if all eigenvalues of CJ have negative real part.

The next result is similar to the proof for convergence of Newton's method given in [5]. This theorem shows that the algorithm converges for essentially any interesting map g. Note that the set inside which orbits converge is typically small. However, the theorem allows us to be sure that regardless of our choice of the switching matrix C or parameter β, for "nice" maps we will always be able to find fixed points with this algorithm. We later show that by tightening restrictions on the map and by carefully choosing C, we can enlarge the region in which convergence is guaranteed by increasing β.

Theorem 3. *Let X be a convex set and $g : X \to \mathbb{R}^n$ be continuous and differentiable for all $x \in X$. For $x_0 \in X$ let positive constants r, α, b, γ, h be given with the following properties*:

$$S_r(x_0) = \{x : \|x - x_0\| < r\} \subseteq X,$$
$$h = \alpha b \gamma < 1,$$
$$r = \alpha/(1 - h).$$

For typographical clarity, we further define

$$g_k = g(x_k),$$
$$s_k = \|g(x_k)\|,$$
$$J_k = J(x_k).$$

For a given switching matrix C and number $\beta \geqslant 0$, let $g(x)$ have the properties:
(i) $\|J(x) - J(y)\| \leqslant \gamma \|x - y\|$ *for all* $x,\ y \in X$,
(ii) $[\beta s I - CJ]^{-1}$ *exists and satisfies* $\|[\beta s I - CJ]^{-1}\| \leqslant b$ *for all* $x \in X$,
(iii) $\|[\beta s_0 I - CJ_0]^{-1} C s_0\| \leqslant \alpha$,
(iv) *If* $\beta > 0, s \leqslant \frac{\alpha\gamma}{2\beta}$.
Then:
(1) *Beginning at x_0, each point*

$$x_{k+1} = x_k + [\beta s_k I - CJ_k]^{-1} C g_k, \quad k = 0, 1, \ldots$$

is well defined and satisfies $x_k \in S_r(x_0)$ for all $k \geqslant 0$.
(2) $\lim_{k\to\infty} x_k = \bar{x}$ *exists and satisfies* $\bar{x} \in S_r(x_0)$ *and* $g(\bar{x}) = 0$.
(3) *The convergence rate is at least quadratic.*

Proof. (1) Since $[\beta s I - CJ]^{-1}$ exists for all $x_k \in X$, x_{k+1} is well defined for each $k \geqslant 0$. To show that $x_k \in S_r(x_0)$, we proceed by mathematical induction. By definition of r, $k = 0$ follows immediately. By assumption (iii) we have $x_k \in S_r(x_0)$ for $k = 1$. Now, if $x_j \in S_r(x_0)$ for $j = 0, 1, 2, \ldots, k$, then from assumption (ii) and the property of the switching matrix C

$$\begin{aligned}\|x_{k+1} - x_k\| &= \|[\beta s_k I - CJ_k]^{-1} C g_k\| \leqslant b\|C g_k\| \\ &= b\|C g_k - C g_{k-1} - CJ_{k-1}(x_k - x_{k-1}) + \beta s_{k-1} I (x_k - x_{k-1})\| \leqslant b\|C g_k - C g_{k-1}\end{aligned}$$

$$-CJ_{k-1}(x_k - x_{k-1})\| + \beta s_{k-1}\|x_k - x_{k-1}\|$$
$$= b\|g_k - g_{k-1} - J_{k-1}(x_k - x_{k-1})\| + \beta s_{k-1}\|x_k - x_{k-1}\|$$

since the definition of x_k implies

$$Cg_k - (\beta s_k I - CJ_k)(x_{k+1} - x_k) = 0. \tag{9}$$

Due to Lemma 1 and assumption (iv)

$$\|x_{k+1} - x_k\| \leqslant \frac{b\gamma}{2}\|x_k - x_{k-1}\|^2 + b\beta s_{k-1}\|x_k - x_{k-1}\| \leqslant \frac{h}{2}\left[\|x_k - x_{k-1}\|^2/\alpha + \|x_k - x_{k-1}\|\right].$$

Note that for $j = 0$, by assumption (iii), we have

$$\|x_1 - x_0\| \leqslant \alpha,$$

and then for $j = 1, 2, 3, \ldots, k$ we have

$$\begin{aligned}
\|x_2 - x_1\| &\leqslant \alpha h, \\
\|x_3 - x_2\| &\leqslant \frac{\alpha}{2}(h^3 + h^2) \leqslant \alpha h^2, \\
\|x_4 - x_3\| &\leqslant \frac{\alpha}{2}(h^5 + h^3) \leqslant \alpha h^3, \\
&\vdots \\
\|x_{k+1} - x_k\| &\leqslant \alpha h^k,
\end{aligned} \tag{10}$$

so by our induction hypothesis

$$\|x_{k+2} - x_{k+1}\| \leqslant \frac{h}{2}\left[\|x_{k+1} - x_k\|^2/\alpha + \|x_{k+1} - x_k\|\right] \leqslant \frac{h}{2}(\alpha h^{2k} + \alpha h^k) \leqslant \alpha h^{k+1}. \tag{11}$$

By the triangle inequality and the fact that $h < 1$, we also have

$$\|x_{k+1} - x_0\| \leqslant \|x_{k+1} - x_k\| + \|x_k - x_{k-1}\| + \cdots + \|x_1 - x_0\| \leqslant \alpha(h^k + h^{k-1} + \cdots + 1) < \alpha\frac{1}{1-h} = r.$$

So, we have shown that $x_{k+1} \in S_r(x_0)$.

(2) Let $m \geqslant k$ be natural numbers. Eq. (10) and the triangle inequality imply

$$\begin{aligned}
\|x_m - x_k\| &\leqslant \|x_m - x_{m-1}\| + \|x_m - x_{m-1}\| + \cdots + \|x_{k+1} - x_k\| \leqslant \alpha(h^{m-1} + h^{m-2} + \cdots + h^k) \\
&\leqslant \alpha h^k(1 + h + \cdots + h^{m-k-1}) < \frac{\alpha h^k}{1-h} = rh^k.
\end{aligned}$$

Set $\epsilon := rh^{\hat{k}}$, and choose m and k so that $m \geqslant k \geqslant \hat{k}$. Since $0 < h < 1$, $\{x_k\}$ is a Cauchy sequence in $S_r(x_0)$, so that the closure of $S_r(x_0)$ has a limit point $\bar{x} \in \overline{S_r(x_0)}$. Finally, to show that $g(\bar{x}) = 0$, note that because of assumption (i) and the fact that $x_k \in S_r(x_0)$ for all $k \geqslant 0$

$$\|J_k - J_0\| = \|J(x_k) - J(x_0)\| \leqslant \gamma\|x_k - x_0\| < \gamma r.$$

The triangle inequality implies

$$\|J_k\| \leqslant \|J_k - J_0\| + \|J_0\| < \gamma r + \|J_0\| =: K.$$

Eq. (9) and assumption (iv) further imply that

$$s_k = \|Cg_k\| = \|(\beta s_k I - CJ_k)(x_{k+1} - x_k)\| < \left(\frac{\alpha\gamma}{2} + K\right)\|x_{k+1} - x_k\|. \tag{12}$$

So

$$\lim_{k\to\infty} s_k = \lim_{k\to\infty}\|g_k\| = 0$$

and the continuity of g in $\overline{S_r(x_0)}$ implies that

$$\lim_{k\to\infty} \|g_k\| = \|g(\bar{x})\| = 0.$$

(3) Define $\Delta_k = \|x_{k+1} - x_k\|$. Note that inequalities (11) and (12) imply

$$\Delta_k \leqslant bs_k < \frac{b\gamma}{2}\Delta_{k-1}^2 + b\beta\left(\frac{\alpha\gamma}{2} + K\right)\Delta_{k-1}^2 \leqslant \left[\frac{h}{2\alpha} + \beta\left(\frac{h}{2} + bK\right)\right]\Delta_{k-1}^2.$$

So, the convergence is at least quadratic and the proof is complete. □

Remark 6. Fig. 2 shows that the DL algorithm converges linearly until it gets sufficiently close to the root and accelerates to quadratic convergence in agreement with (3) in Theorem 3.

Remark 7. Theorem 3 gives sufficient conditions for convergence. The sufficient region can be quite small and is dependent on β and C.

An important feature of the DL-algorithm is that the size of the basin of attraction can be enlarged by choosing a larger value of β. The following theorem shows that the length of the interval containing $\bar{x}$ where $H(x)$ is increasing with a slope less than 1 is an increasing function of β.

Theorem 4. *Suppose* $g(\bar{x}) = 0$, $g'(\bar{x}) \neq 0$, *and* $\beta \geqslant 0$. *Define*

$$H(x) := x + \frac{Cg(x)}{\beta|g(x)| - Cg'(x)}$$

and also define $N_{\beta,C}(\bar{x})$ *to be the open interval containing* $\bar{x}$ *and satisfying* $|H'(x)| < 1$ *for all* $x \in N_{\beta,C}(\bar{x})$. *Then, if*

$$-2 < \frac{g(x)g''(x) - [g'(x)]^2}{[\beta|g(x)| - Cg'(x)]^2} < 0 \tag{13}$$

is well defined, then we can choose a number $C^* = \pm 1$ *such that if* $0 \leqslant \beta_2 \leqslant \beta_1$, *then* $N_{\beta_2,C^*}(\bar{x}) \subseteq N_{\beta_1,C^*}(\bar{x})$.

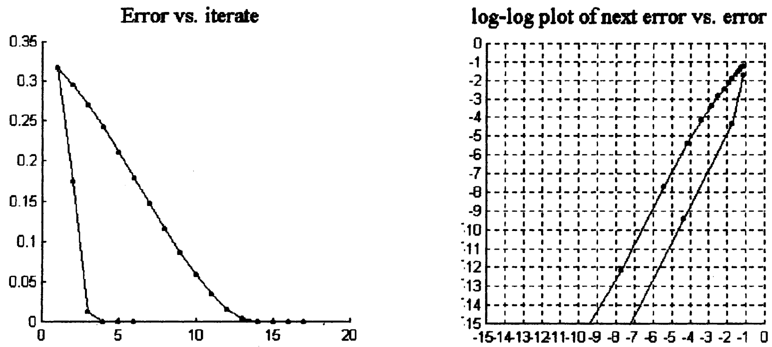

Fig. 2. Convergence data for Newton's method and the DL-method. The function used is $g(x_1, x_2) = (x_2(x_1 - 1/3), x_1(x_2 - 2))$ seeded with $x_0 = (0.23, 1.7)$ which is inside the basin of attraction for $\bar{x} = (0.\bar{3}, 2)$. The DL-algorithm uses $C = C_6 = -I$ and $\beta = 25$. The plot on the left shows the distance from x_k to $\bar{x}$, Δ_k, as a function of k. The top curve is for the DL-method, and the bottom is for Newton's method. The plot on the right is a graph of $\log \Delta_{k+1}$ versus $\log \Delta_k$. The slope of the curve as the error goes to zero (as the logarithms go to $-\infty$) is the order of convergence. Both Newton's method and the DL-method are parallel with slope 2 near the bottom, while the DL-method has slope 1 near the top in agreement with expectations.

Proof. We begin by differentiating H

$$H'(x) = 1 + C\frac{g'(x)\left(\beta|g(x)| - Cg'(x)\right) - g(x)\left(\beta|g(x)|' - Cg''(x)\right)}{[\beta|g(x)| - Cg'(x)]^2}$$
$$= 1 + C\frac{C(g(x)g''(x) - [g'(x)]^2) + \beta\left(g'(x)|g(x)| - g(x)|g(x)|'\right)}{[\beta|g(x)| - Cg'(x)]^2}.$$

We now consider the difference $\Delta := g'(x)|g(x)| - g(x)|g(x)|'$. On any interval where $g(x) \geqslant 0$, $\Delta = 0$. Note that in fact, on any interval where $g(x) \leqslant 0$, $\Delta = g'(x)(-g(x)) - g(x)(-g'(x)) = 0$. So

$$H'(x) = 1 + \frac{g(x)g''(x) - [g'(x)]^2}{[\beta|g(x)| - Cg'(x)]^2}$$

and the bound in (13) implies that $H'(x)$ is well defined and satisfies $|H'(x)| < 1$ on $N_{\beta,C}(\bar{x})$. Now, let $\beta_2 \leqslant \beta_1$.

Case 1: $g'(\bar{x}) < 0$. Choose $C = 1$. Suppose $x \in N_{\beta_2,1}(\bar{x})$. Then

$$-2 < \frac{g(x)g''(x) - [g'(x)]^2}{[\beta_2|g(x)| - Cg'(x)]^2} \leqslant \frac{g(x)g''(x) - [g'(x)]^2}{[\beta_1|g(x)| - Cg'(x)]^2} < 0,$$

since $0 < \beta_2|g(x)| - Cg'(x) \leqslant \beta_1|g(x)| - Cg'(x)$. Therefore, $x \in N_{\beta_1,1}(\bar{x})$.

Case 2: $g'(\bar{x}) > 0$. Choose $C = -1$. Similar to case 1. □

Theorem 4 does not generalize in higher dimensions. This is due in part to the fact that the inverse of $\beta s - CJ$ is a quotient of two functions of β while in the one-dimensional problem, β arises only in the denominator.

The contraction mapping theorem guarantees any point x satisfying $\|\partial H/\partial x\| < 1$ will converge to a unique fixed point contained in the set.

Remark 8. Unfortunately, the following is not true. Define the neighborhood of a fixed point of g by $N_{\beta,C}(\bar{x}) = \{x : \|\partial H/\partial x\| \leqslant 1\}$. There exists a C so that $0 \leqslant \beta_2 \leqslant \beta_1$ implies $N_{\beta_2,C^*}(\bar{x}) \subseteq N_{\beta_1,C^*}(\bar{x})$.

We can show it is not true by example. We compute the derivative of $H(x) = x - [\beta sI - CJ]^{-1}Cg$ for the simple function g given in Section 4.1. By writing the derivative symbolically

$$\frac{\partial H}{\partial x}(x;\beta,C) = [\beta sI - CJ]^{-1}\left[\left(I \otimes \frac{\beta s}{g^{\mathrm{T}}g}g^{\mathrm{T}}J - C\frac{\partial J}{\partial x}\right)\left([\beta sI - CJ]^{-1} \otimes I\right)(Cg \otimes I) + \beta sI\right],$$

where "$\otimes$" is the Kronecker product, it is easier to graph the curves $\partial H/\partial x = 1$ in the x_1x_2-plane for various values of β. Since there are only eight Cs to check, we must produce only 16 graphs to verify that the result does not hold.

Remark 9. It is natural to compare the DL-algorithm with Newton's method since $\beta = 0$ yields Newton's method. On the other hand, the DL-step-size decreases with increasing β once β is sufficiently large, say when $\beta \geqslant \beta^* > 0$. Since an iteration scheme that takes smaller steps is less likely to over-shoot the targeted root, we also compared the DL-algorithm with a quasi-Newton method in which before each Newton step, $J_k^{-1}g_k$, is taken, we check that $\|g_{k+1}\| < \|g_k\|$. If not, we try steps of $2^{-j}J_k^{-1}g_k$ for $j = 1, 2, \ldots$ until $\|g_{k+1}\| < \|g_k\|$. Typically, we find larger convergence basins with quasi-Newton than with Newton, but we never find that quasi-Newton has large enough basins to permit the seeding technique of DL. We state this more precisely in the next remark.

Remark 10. Experience suggests the following: although we have been unable to verify it in general: Let $\tilde{S}_{\beta,C}(\bar{x}) = \{x_k : x_k \to \bar{x}$ using the DL iteration scheme$\}$. Then, let $S_{\beta,C}(\bar{x})$ be the largest simply connected subset of $\tilde{S}_{\beta,C}(\bar{x})$ containing $\bar{x}$ ($\tilde{S}_{\beta,C}(\bar{x})$ can be disconnected because of the large initial steps). Then, there exists a switching matrix C^* such that $\beta \geqslant 0$ implies that $S_{0,C^*}(\bar{x}) \subseteq S_{\beta,C^*}(\bar{x})$. In other words, the DL-algorithm always provides a larger basin of convergence than Newton's method. The same appears to be true for a quasi-Newton scheme.

4. Examples

We show through examples that the DL-algorithm for finding periodic orbits is quite a general root finding technique that finds periodic orbits – stable or not. In fact, while Schmelcher and Diakonos originally used their switching matrices to determine the type of instability of an unstable periodic orbit (UPO) found, the switching matrix does not generally tell us the stability type of the orbits it will find with the DL-algorithm. Both SD [2] and DL [1] report that the majority of periodic orbits are found with $C_1 = I$. For the DL-algorithm, there is nothing inherently special about $C_1 = I$ except that it is the first switching matrix used. In fact, if we used C_2 first, we should expect that most periodic orbits would be found by it. For a particular map, there are switching matrices that will most efficiently find the periodic orbits, but there is no way in general to determine which those are. In 1 and 2 dimensions, one can easily show on a case-by-case basis that all fixed points of the vector field Cg can be stabilized by the Cs listed in Table 1. We do not know if this is true in general, but we have found no examples to the contrary.

4.1. A simple map

The following example was designed to show application of the preceding theorems, and so it was constructed to simplify checking each condition of Theorem 3. Let

$$g(x) = \begin{cases} g_1(x_1, x_2) & = & x_2(x_1 - 1/3), \\ g_2(x_1, x_2) & = & x_1(x_2 - 2). \end{cases} \tag{14}$$

Our main objective is to find periodic orbits of maps. So, we think of g as the function whose roots are fixed points of $f(x) = x + g(x)$. In this case, f has two fixed points: $\hat{x} = (0,0)$, a saddle with corresponding eigenvalues $1 + 1/3\sqrt{6}$ and $1 - 1/3\sqrt{6}$ and $\bar{x} = (1/3, 2)$, a source with corresponding eigenvalues 3 and 4/3. Equivalently, g has roots $\hat{x}$ and $\bar{x}$ with corresponding eigenvalues $1/3\sqrt{6}, -1/3\sqrt{6}$ and $2, 1/3$, respectively. Since g_1 and g_2 are quadratic in their variables, we expect the boundary separating the basins of attraction for the two roots to lie close to the curve $\det(J) = 0$, where J is the Jacobian matrix of g. In this case, $\det(J) = 0$ is $x_2 = 2 - 6x_1$. Fig. 3(b) shows how $\det(J) = 0$ splits the plane as well as the basins of attraction for the two roots of g under Newton's method. The quasi-Newton method produces identical convergence basins for this example. In more complicated maps, the quasi-Newton typically has larger convergence basins than for Newton's method. The DL-method stabilizes the root for the vector field. As we see in Figs. 3(a) and 4, neither root is stable for g, but the origin $\hat{x}$ is stable for $C_5 g$ and the non-trivial root $\bar{x}$ is stabilized for $C_6 g$. It is quite noticeable how much larger an appropriate choice of C can make the basin of attraction for a particular root.

4.2. Julia set for quadratic map

The chaotic set need not be attracting. The Julia set for $Q(z) = z^2 - 0.13 + 0.76i$, the boundary of the black region shown in Fig. 5(a), has a dense chaotic repeller. Since Q is a one-dimensional polynomial, there are exactly 2^p period-p points making this a prime candidate for testing the reliability of the DL-algorithm. The method succeeds despite having only sources and a single sink for the zeros of $g(z) = Q^p(z) - z$, $p = 1, 2, \ldots$ As always, we can view the "skeleton" of the chaotic set by plotting its periodic orbits as shown in Fig. 5(b).

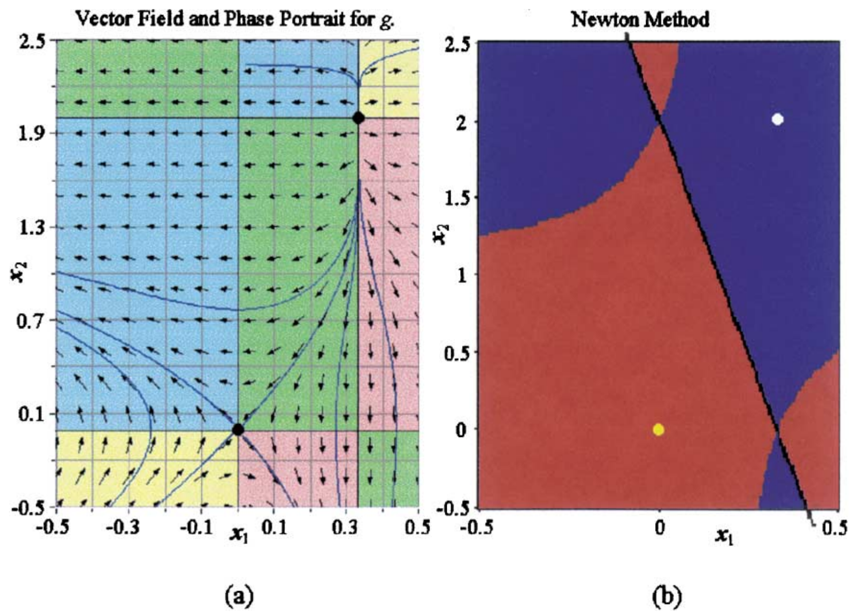

Fig. 3. (a) Regions are bounded by nullclines of g in Section 4.1 and color coded by the sign of g_1 and g_2 to make it easier to see the directions of the vector field flow. The origin is a saddle and the non-trivial root is a source. (b) Seeds (points in the plane) are colored blue if Newton's method converges to the non-trivial root (white dot) within 40 iterations and red for the origin (yellow dot). We also superimposed the line $\det(J) = 0$. The quasi-Newton method produces the same basins for this simple map.

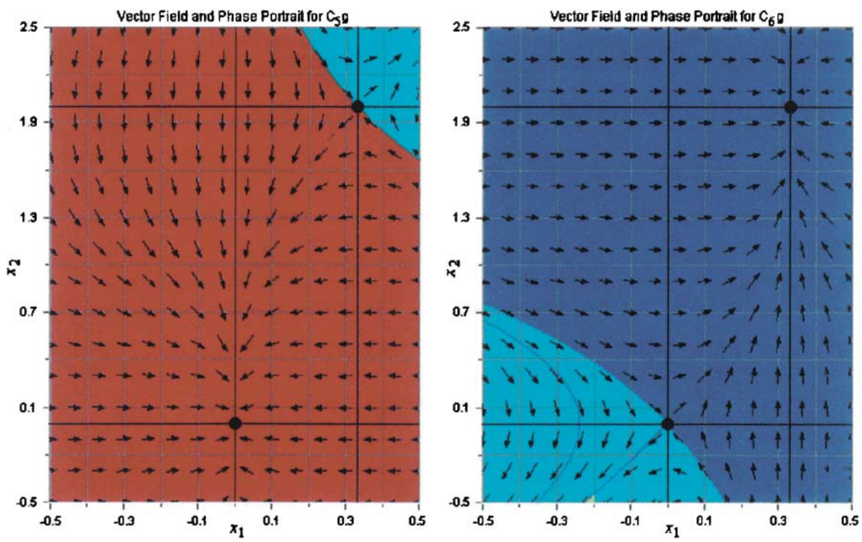

Fig. 4. The origin is stable in the vector field C_5g, and the non-trivial root is stable in the vector field for C_6g. But, since switching matrices only swap the order and switching signs in g, the nullclines of any vector field Cg are always the same.

4.3. The standard map

We studied the DL-algorithm for various values of parameter K in the standard map

$$x_k = \left(x_{k-1} - \frac{K}{2\pi}\sin(2\pi y_{k-1})\right) \bmod 1,$$
$$y_k = (x_k + y_{k-1}) \bmod 1.$$

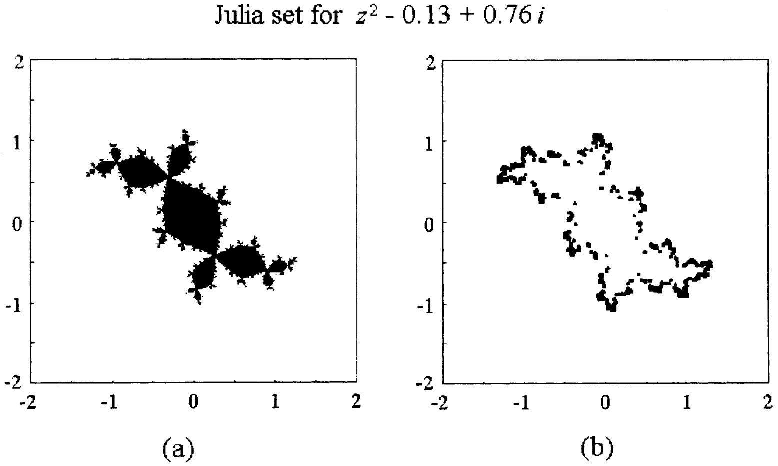

Fig. 5. (a) The Julia set for $z^2 - 0.13 + 0.76i$ is the boundary of the black region. All points in the black region but not on the boundary are attracted to a stable period-3 orbit. All points outside the black region diverge to infinity. Almost all points on the boundary lie on a chaotic orbit – so this is where we find all periodic orbits except for the attracting period-3 orbit. (b) The 986 prime period-10 points of the Julia set for $z^2 - 0.13 + 0.76i$.

A typical case in which chaotic sets coexist with KAM-like invariant circles is $K = 1.25$ [7]. One typically expects that for large enough values of β the DL-algorithm will succeed as long as we wait long enough. In this case, however, we find that while Newton's method fails to find as many UPOs as the DL-algorithm does, that we do best for very small β values and find ourselves gradually finding fewer UPOs as we increase β in stark contrast to the behavior of the algorithm reported in [1]. Since large β gives small steps for the DL-method, orbits become trapped in invariant regions containing their seed, potentially blocking them from unfound roots. Furthermore, seeding only with period-2 points will not yield any period-3 orbits – regardless of C or β so we generate seeds by following a chaotic orbit. Although the quasi-Newton method has much larger convergence regions than Newton's, it too fails to find the four period-3 orbits when seeded with only the period-2 points.

4.4. The Tinkerbell map

Nusse and Yorke [6] reported finding 64 period-10 UPOs using a quasi-Newton method in the Tinkerbell map

$$f(x) = \begin{cases} f_1(x_1, x_2) = x_1^2 - x_2^2 + 0.9x_1 - 0.6013x_2, \\ f_2(x_1, x_2) = 2x_1x_2 + 2x_1 + 0.5x_2. \end{cases}$$

The quasi-Newton method fails to find the majority of period-p points when seeded with the period-$p-1$ points. In contrast, the DL-algorithm only fails to find the period-3 orbits from the period-2 seeds, but starting from period-3, it appears to find complete sets of UPOs of all higher periods. So, millions of randomly chosen seeds are required for the quasi-Newton routine to find the period-10 orbits, while the DL-algorithm uses only the $9 \times 56 = 504$ period-9 points as seeds and exploits the larger basin of attraction that the right switching matrix provides. With six switching matrices, we require only 3024 applications of the DL-iterator to find 101 period-10 UPOs making the DL-algorithm a much faster routine.

4.5. A problem with the seeding algorithm

Here is a simple example to illustrate why we often have to modify the approach of seeding with previous orbits. Consider the map

$$f(x) = 10x \bmod 1, \quad 0 \leqslant x < 1.$$

After finding the nine fixed points and 90 period-2 points, the DL-algorithm proceeds as follows. We use each of the 90 period-2 points as seeds, once with $C = 1$ and again with $C = -1$ and start with an arbitrary but small value of $\beta = \beta_1 \geqslant 0$. We then repeat this process for an increasing sequence of β-values and record the number of period-3 UPOs found as a function of β, $N_3(\beta)$. The graph of $N_3(\beta)$ versus β will typically level off after which we assume that we have found all of the periodic points. In this case, however, the method will find at most $90 \times 2 = 180$ UPOs or only $3 \times 180 = 540$ out of the 990 period-3 points. The cause of this problem is overcome if $p > 4$, but one can similarly construct higher dimensional baker-like maps in which the number of UPOs must grow too fast early on to be seeded by the UPOs of the previous period. Such considerations make it likely that the transformations which Kaloshin [44] has recently shown produce superexponential growth of the number of periodic orbits can produce difficulties for the DL-seeding scheme.

5. Conclusion

In this paper, we have given a rigorous analysis on convergence properties of the DL-algorithm which finds periodic orbits. We have shown that we generally expect the root finder to converge, the only matter being to which root. We have shown that a sufficient basin of attraction has a size which is adjustable with a control parameter β which we consider to be one of the algorithms most attractive features. We have also shown the asymptotic rate of convergence to be quadratic. Furthermore, we have given many examples which show the general utility of this algorithm. We suggest that this algorithm should also become the general root finder in numerical analysis settings in which: (1) there are many roots to be found and thus the flexibility of different seeds *and* different switching matrices improves success in finding all of them, (2) speed is an issue. In our numerical experiments, we have reported that experience suggests that the algorithm's most attractive feature is even better than we can rigorously prove. It seems that the basin of attraction of a given fixed point usually increases significantly with appropriately chosen β control parameter, although our experience with the standard map suggests that such statements cannot be general. In future work, we plan to investigate an appropriate probabilistic measure which says roughly that the chance of choosing a convergent seed increases monotonically with β.

Acknowledgements

This work was supported by the National Science Foundation under grant DMS-9704639.

Appendix A. Are involutary matrices important?

When Schmelcher and Diakonos [2,3] described their method for finding UPOs, they seemed to imply that involutary matrices C, i.e., matrices satisfying $C^2 = I$, played a role in their technique. They inferred that switching matrices are involutary (not all of them are). Furthermore, they used a result that applied to certain involutary matrices and implied that this would also follow for the switching matrices. We complete and correct their arguments below.

The proof of a theorem similar to the following was sketched by Schmelcher and Diakonos [2]. We complete and generalize it enough to show that it does not easily follow from this theorem that switching matrices and involutary matrices are related as suggested by Schmelcher and Diakonos.

Theorem 5. *Let J be an arbitrary $n \times n$ invertible matrix with n distinct eigenvalues: $\lambda_1, \lambda_2, \ldots, \lambda_n$ corresponding to eigenvectors $v_1, v_2, \ldots, v_n$. Then, there exists a matrix C satisfying:*
(i) $C^2 = aI$ (*for some positive constant a*);
(ii) $\|C\| = 1$;
such that every eigenvalue of CJ has negative real part.

Proof. Let $A = CJ$. Since J has n distinct eigenvalues, it is diagonalizable by a similarity transformation P

$$P^{-1}JP = \Lambda := \mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}.$$

So

$$P^{-1}AP = P^{-1}CJP = P^{-1}CP\Lambda. \tag{A.1}$$

Construct the diagonal matrix $C_D = \mathrm{diag}\{c_1, c_2, \ldots, c_n\}$ as follows:

$$c_i = \begin{cases} 1 & \text{if } \mathrm{Re}(\lambda_i) < 0, \\ -1 & \text{if } \mathrm{Re}(\lambda_i) > 0 \end{cases}$$

and find $\tilde{C}$ satisfying $C_D = P^{-1}\tilde{C}P$. Define $C = \tilde{C}/\|\tilde{C}\|$ so that $\|C\| = 1$ and $C^2 = \tilde{C}^2/\|\tilde{C}\|^2 = PC_D^2P^{-1}/\|\tilde{C}\|^2 = PIP^{-1}/\|\tilde{C}\|^2 = aI$ where $a = 1/\|\tilde{C}\|^2$. Finally, since

$$P^{-1}CP = \frac{1}{\|\tilde{C}\|}P^{-1}PC_DP^{-1}P = \frac{C_D}{\|PC_DP^{-1}\|},$$

Eq. (A.1) implies that $A = CJ$ is similar to a diagonal matrix with all negative entries on the diagonal. □

Remark 11. The matrix $\tilde{C}$ in Theorem 5 is involutary but does not in general have norm 1. Since all switching matrices have norm 1, the question we now seek to answer is whether we can use the method of proof in Theorem 5 to construct a switching matrix given (generic) arbitrary J. The corollary that follows shows that eigenvectors of CJ and J are the same if C is constructed as in Theorem 5.

Remark 12. In the construction of C_D in Theorem 5, generalize

$$c_i = \begin{cases} |c_i| & \text{if Re } (\lambda_i) < 0, \\ -|c_i| & \text{if Re } (\lambda_i) > 0, \end{cases}$$

so that $\{\mathrm{Re}(c_i\lambda_i),\ i = 1, \ldots, n\}$ are distinct. The result in Theorem 5 still holds.

Corollary 6. *If the eigenpairs of J are* (λ_i, v_i), *then the eigenpairs of CJ are* $(\pm a\lambda_i, v_i)$ *with* $a = 1/\|\tilde{C}\|$ *where we choose* "+" *if* $\mathrm{Re}(\lambda_i) < 0$ *and* "−" *if* $\mathrm{Re}(\lambda_i) > 0$.

Proof. Since the eigenvalues and eigenvectors for J collectively satisfy $JP = P\Lambda$,

$$(CJ)P = CP\Lambda = \frac{1}{\|\tilde{C}\|}PC_DP^{-1}P\Lambda = P(aC_D\Lambda). \qquad \square$$

Theorem 7. *Assume J is an* $n \times n$ *matrix with n distinct eigenvalues, none of which have zero real part.*
1. *If all eigenvalues of J have negative real part, then* $C = I$ *stabilizes J.*
2. *If all eigenvalues of J have positive real part, then* $C = -I$ *stabilizes J.*
3. *Let P be a matrix of eigenvectors for J, and let* P_C *be a matrix of eigenvectors for a switching matrix C. If J has eigenvalues with both positive and negative real part, then we can only construct a switching matrix from Theorem 5 if each column of P is parallel to a column of* P_C.

Proof.

(1) J is already stabilized so the result is immediate.

(2) Let P and $\Lambda = \mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ be a matrix of eigenvectors and corresponding eigenvalues, respectively. Then $JP = P\Lambda$, so $(-I)JP = -1 \cdot IP\Lambda = P(-\Lambda)$. So, $C = -I$ stabilizes J.

(3) Let P_C and Λ_C be a matrix of eigenvectors and corresponding eigenvalues for a switching matrix C (other than $\pm I$) so that $CP_C = P_C\Lambda_C$. Note that since J has n distinct eigenvalues, that it also has n distinct eigenvectors. Subsequently, any matrix (S, D) pair satisfying $CS = SD$ where D is diagonal must be matrices of eigenvectors and eigenvalues. So, the following must be true in order for $CP = PC_D$:

(i) $C_D = a\Lambda_C$ for some constant a;

(ii) each column of P is parallel to a column of P_C. □

Example 1. With probability zero will the columns of P and P_C be parallel. So, if we randomly construct J with the constraints that it will have distinct eigenvalues with non-zero real part, Theorem 5 will fail to produce a switching matrix. For example, let

$$J = \begin{bmatrix} -1 & -2 \\ -3 & -2 \end{bmatrix}.$$

Then

$$P = \begin{bmatrix} -1 & 2 \\ 1 & 3 \end{bmatrix} \quad \text{and} \quad \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & -4 \end{bmatrix}.$$

We construct

$$C_D = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

and then find

$$C = \frac{5}{\sqrt{26}+1} \begin{bmatrix} -1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -3/5 & 2/5 \\ 1/5 & 1/5 \end{bmatrix} = \frac{1}{\sqrt{26}+1} \begin{bmatrix} -1 & 4 \\ 6 & 1 \end{bmatrix} \approx \begin{bmatrix} -0.16396 & 0.65584 \\ 0.98376 & 0.16396 \end{bmatrix},$$

which is not a switching matrix.

Example 2. Here is an example of a rare case when Theorem 5 permits a switching matrix that is neither I nor $-I$. Let

$$J = \begin{bmatrix} -3/2 & -25/4 \\ -1 & -3/2 \end{bmatrix}.$$

Then

$$P = \begin{bmatrix} -2 & 5 \\ 2 & 5 \end{bmatrix} \quad \text{and} \quad \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & -4 \end{bmatrix}.$$

By Theorem 5,

$$C_D = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

so that

$$C = \begin{bmatrix} -2 & 5 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1/4 & 1/4 \\ 1/10 & 1/10 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Note that

$$P_C = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad \Lambda_C = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

so that the two required conditions are met in step 3 of the proof.

Remark 13. Schmelcher and Diakonos seemed to imply that Theorem 5 guaranteed that there is always a switching matrix that can stabilize J. Subsequently, Davidchack and Lai continued to use the switching matrices without showing concern for this issue, but Theorem 7 shows this is not always possible.

Remark 14. The theorems of this section do not guarantee that there is always a switching matrix for an arbitrary diagonalizable J. However, we have yet to find an example of a matrix J that cannot be stabilized. Consider, for example, the 5×5 matrix

$$J = \begin{bmatrix} 1 & -3 & 0 & -2 & 5 \\ 3 & 2 & 5 & -7 & 1 \\ 2 & 8 & 1 & 0 & -4 \\ 4 & 9 & -2 & 3 & 0 \\ 4 & 4 & 7 & -4 & 3 \end{bmatrix},$$

whose five eigenvalues are

$$\begin{aligned} \lambda_1 &= -0.4387597197, \\ \lambda_{2,3} &= 0.08849259349 \pm 5.728217720i, \\ \lambda_{4,5} &= 5.130887266 \pm 5.061384383i. \end{aligned}$$

Theorem 5 will not yield a switching matrix. However, we can easily find a switching matrix by noting two important facts:

(1) Switching matrices permute and switch signs of rows of the matrix they multiply on the left.

(2) Eigenvalues lie inside Gershgorin circles whose centers are given by the diagonal entries.

If the diagonal entries of J are large enough in magnitude and negative, then we should get negative eigenvalues. The largest magnitude entry is the 9 in row 4, so at first one might try moving row 4 to row 2 and switching signs. However, by doing this we cannot make use of the 8 in row 3 which will guarantee us a negative eigenvalue once it is negated. So, we first move row 4 to 1 and 3 to 2 and negate rows placing -4 and -8 on the diagonal. We next move row 5 to row 3 and negate to place a -7 on the diagonal. Finally, we can move row 2 to 4 without any sign change, and row 1 to 5 with a sign change. The switching matrix

$$C = \begin{bmatrix} 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

does this and stabilizes J since

$$CJ = \begin{bmatrix} -4 & -9 & 2 & -3 & 0 \\ -2 & -8 & -1 & 0 & 4 \\ -4 & -4 & -7 & 4 & -3 \\ 3 & 2 & 5 & -7 & 1 \\ -1 & 3 & 0 & 2 & -5 \end{bmatrix}$$

has eigenvalues

$$\lambda_{1,2} = -11.40252568 \pm 2.407836061i,$$
$$\lambda_3 = -4.883901295,$$
$$\lambda_4 = -2.925593730,$$
$$\lambda_5 = -0.3854536062.$$

References

[1] Davidchack RL, Lai Y-C. Efficient algorithm for detecting unstable periodic orbits in chaotic systems. Phys Rev E 1999;60(5):6172–5.
[2] Schmelcher P, Diakonos FK. General approach to the localization of unstable periodic orbits in chaotic dynamical systems. Phys Rev E 1998;57(3):2739–46.
[3] Schmelcher P, Diakonos FK. Detecting unstable periodic orbits of chaotic dynamical systems. Phys Rev Lett 1997;78(25): 4733–6.
[4] Lathrop DP, Kostelich EJ. Characterization of an experimental strange attractor by periodic orbits. Phys Rev A 1989;40(7): 4028–31.
[5] Stoer J, Bulirsch R. Introduction to numerical analysis. New York: Springer; 1980.
[6] Nusse HE, Yorke JA. Dynamics: numerical explorations. vol. 101, 2nd ed. Applied mathematical sciences series. New York: Springer, 1997.
[7] Meiss J. Simplectic, variational principles and transport. Rev Mod Phys 1992;64(3):795–848.
[8] Cvitanovic P. Invariant measurements of strange sets in terms of cycles. Phys Rev Lett 1988;61:2729–32.
[9] Cvitanovic P. Periodic orbits as the skeleton of classical and quantum chaos. Physica D 1991;51:138–51.
[10] Cvitanovic P. Dynamical averaging in terms of periodic orbits. Physica D 1995;83:109–23.
[11] Poincaré H. Les mé thodes nouvelles de la méchanique céleste. Paris: Guthier-Villars; 1892.
[12] Devaney RL. An introduction to chaotic dynamical systems. 2nd ed. Redwood City, CA: Addison-Wesley, 1989.
[13] Artuso R, Aurell E, Cvitanovic P. Recycling of strangesets: I cycle expansions. Nonlinearity 1990;3:325–59.
[14] Bauer O, Mainieri R. The convergence of chaotic integrals. Chaos, Solitons & Fractals 1997;7(3):361–7.
[15] Grebogi C, Ott E, Yorke JA. Unstable periodic orbits and the dimension of multifractal attractors. Phys Rev A 1988;37: 1711.
[16] Lai YC, Nagai Y, Grebogi C. Characterization of the natural measure by unstable periodic orbits in chaotic attractors. Phys Rev Lett 1997;79:649.
[17] Lai Y-C. Characterization of the natural measure by unstable periodic orbits in nonhyperbolic chaotic systems. Phys Rev E 1997;56(6):6531–9.
[18] Gora P, Byers W, Boyarsky A. Measures on periodic orbits for continuous transformations on the interval. Stoch Anal Applic 1991;9(3):263–70.
[19] Badii R, Brun E, Finardi M, Flepp L, Holzner R, Parisi J, Reyl C, Simonet J. Progress in the analysis of experimental chaos through periodic orbits. Rev Mod Phys 1994;66:1389.
[20] Davidchack R, Lai Y-C, Bollt E, Dhamala M. Estimating generating partitions of chaotic systems by unstable periodic orbits. Phys Rev E 2000;61(2):1353–6.
[21] Lefranc M, Glorieux P, Papoff F, Molesti F, Arimondo E. Combining topological analysis and symbolic dynamics to describe a strange attractor and its crises. Phys Rev Lett 1994;73:1364–7.
[22] Boulant G, Bielawski S, Derozier D, Lefranc M. Experimental observation of a chaotic attractor with a reverse horseshoe topological structure. Phys Rev E 1997;55(4):3801–4.
[23] Boulant G, Lefranc M, Bielawski S, Derozier D. Horseshoe templates with global torsion in a driver laser. Phys Rev E 1997;55(4):5082–91.
[24] Plumecoq J, Lefranc M. From template analysis to generating partitions I: Periodic orbits, knots and symbolicencodings, preprint, 1999.
[25] Plumecoq J, Lefranc M. From template analysis to generating partitions II: Characterizations of the symbolic encodings, preprint, 1999.
[26] Tufillaro NB, Holzner R, Flepp L, Brun E, Finardi M, Badii R. Template analysis for a chaotic NMR laser. Phys Rev E 1991;44:4786.
[27] Auerback D, Cvitanovic P, Eckmann J-P, Guaratne G, Procaccia I. Exploring chaotic motion through periodic orbits. Phys Rev Lett 1987;58(23):2387–8.
[28] Biham O, Wenzel W. Characterization of unstable periodic orbits in chaotic attractors and repellers. Phys Rev Lett 1989; 63:819.
[29] Biham O, Wenzel W. Unstable periodic orbits andthe symbolic dynamics of the complex Hénon map. Phys Rev A 1990;42(9):4639.
[30] Bowen R. Equilibrium states and the ergodic theory of Anosov diffeomorphisms. In: Lecture Notes in Mathematics, vol. 470. New York: Springer, 1975.

[31] Bowen R. Topological entropy and Axiom A. Proc Symp Pure Math, A.M.S., Providence R.I. 1970;14:23–41.
[32] Bowen R. Periodic points and measures for Axiom A diffeomorphisms. Trans Amer Math Soc 1971;154:377–97.
[33] Bowen R. Some systems with unique equilibrium state. Math Systems Theory 1975;8:193–203.
[34] Bowen R. Periodic orbits for hyperbolic flows. Amer J Math 1972;94:1–30.
[35] Katok A. Lyapunov exponents, entropy and periodic orbits for diffeomorphisms. Inst Hautes Études Sci Publ Math 1980;51: 137–73.
[36] de la Llave R. Smooth conjugacy and S-R-B measures for uniformly and non-uniformly hyperbolic systems. Comm Math Phys 1992;150(2):289–320.
[37] de Carvalho A. Pruning fronts and the formation of horseshoes. Ergodic Theory Dynam Syst 1999;19(4):851–94.
[38] Smale S. Differentiable dynamical systems. Bull AMS 1967;73:747–817.
[39] Smale S. On the efficiency of algorithms of analysis. Bull Amer Math Soc (NS) 1985;13(2):87–121.
[40] Smale S. The fundamental theorem of algebra and complexity theory. Bull Amer Math Soc (NS) 1981;4(1):1–36.
[41] Ruelle D. A measure associate with Axiom A attractors. Am J Math 1976;98:619–54.
[42] Milnor J, Thurston W. On iterated maps of the interval (I and II), Princeton: Princeton University Press, 1977.
[43] Strang G. A chaotic search for i. Coll. Math. J. 1991;221:3–12.
[44] Kaloshin V. Generic diffeomorphisms with superexponential growth of number of periodic orbits, Stony Brook IMS Preprint, #1999/2.