Journal of Physics A:
Mathematical and
Theoretical

Computing Lyapunov exponents using weighted Birkhoff averages

To cite this article: E Sander and J D Meiss 2025 *J. Phys. A: Math. Theor.* **58** 355701

OPEN ACCESS
IOP Publishing
Journal of Physics A: Mathematical and Theoretical
J. Phys. A: Math. Theor. **58** (2025) 355701 (16pp)
https://doi.org/10.1088/1751-8121/adfac1

Computing Lyapunov exponents using weighted Birkhoff averages

E Sander[1,*] **and J D Meiss**[2]

[1] Department of Mathematical Sciences, George Mason University, Fairfax, VA 22030, United States of America
[2] Department of Applied Mathematics, University of Colorado, Boulder, CO 80309-0526, United States of America

E-mail: esander@gmu.edu and jdm@colorado.edu

Received 24 February 2025; revised 23 June 2025
Accepted for publication 12 August 2025
Published 27 August 2025

Abstract

The Lyapunov exponents of a dynamical system measure the average rate of exponential stretching along an orbit. Positive exponents are often taken as a defining characteristic of chaotic dynamics, with the size of the exponent indicating the strength of the chaos, or in the case of a negative exponent, a measure of the how far an orbit is from being chaotic. However, the standard orthogonalization-based method for computing Lyapunov exponents converges slowly—if at all. Many alternatively techniques have been developed to distinguish between regular and chaotic orbits, though most do not compute the exponents. We compute the Lyapunov spectrum in three ways: the standard method, the weighted Birkhoff average (WBA), and the 'mean exponential growth rate for nearby orbits' (MEGNO). The latter two improve convergence for nonchaotic orbits, but the WBA is fastest. However, for chaotic orbits the three methods converge at similar, slow rates. Since there is little computational cost for the WBA, we argue it should be used in any case. Though the original MEGNO method does not compute Lyapunov exponents, we show how to reformulate it as a weighted average that does.

Keywords: Lyapunov exponents, weighted Birkhoff averages, MEGNO, chaos

* Author to whom any correspondence should be addressed.

1. Introduction

Lyapunov exponents are a fundamental gauge of chaotic behavior in dynamical systems. They measure the asymptotic growth rate of the distance between a pair of (infinitesimally) close orbits, and the existence of a positive exponent, $\mu > 0$, is often taken as a primary indicator for 'sensitive dependence on initial conditions', one of the principal requirements for chaos. However, existing techniques for computing exponents typically converge—if at all—as $1/T$, and often no faster than $\ln(T)/T$, where T is the computed orbit length [1]. Indeed, it often is difficult to even determine if $\mu \neq 0$; for example, orbits that are very close to regular regions in a Hamiltonian system can be chaotic but have arbitrarily small maximal Lyapunov exponents [2].

Though there have been many attempts to accelerate this convergence, [3–6] such an acceleration seems to be difficult. Many authors have attempted to obviate this difficulty by developing other methods for distinguishing between regular and chaotic orbits. As reviewed by Skokos [7], Maffione *et al* [8], Skokos *et al* [9], such categorical tests for chaos include the FLI [10, 11] (fast Lyapunov indicator), MEGNO [1, 12] (mean exponential growth of nearby orbits), frequency analysis [13], SALI [14] (smaller alignment index), GALI [15, 16] (generalized alignment index), the 0-1 test for chaos [17], etc.

We too have developed a categorical test: we showed [18, 19] that the weighted Birkhoff average (WBA) [20] of a function on phase space can efficiently distinguish between regular and chaotic orbits by its convergence rates. We argued that—if this is the only goal—it would be more efficient to compute the WBA than the Lyapunov exponent or its categorical variants, since in this case only iterates of the map f and not those of its derivatives are needed. We compared the convergence of the WBA to that of conventional Lyapunov exponents [18] as well as to the 0-1 test [17]. More recently, Bazzani *et al* [21] compared the WBA to many of the other categorical tests.

However, these techniques are *categorical*; they do not try to compute accurate Lyapunov exponents. While classification can be valuable, the exponent itself is an important quantification of chaotic motion; for example, its inverse defines an effective prediction horizon, and invariant sets with larger exponents are quantifiably 'more' chaotic. In addition, it is important to distinguish the case of *hyper-chaos* where there is more than one positive exponent.

Can one develop new methods to accurately compute the Lyapunov spectrum? Our goal in this paper is to investigate such computations using the WBA. As we will see, for smooth maps and nonchaotic orbits, C^∞-smooth weighted averages can compute (non-positive) Lyapunov exponents with super-polynomial convergence: the error decreases faster than $1/T^k$ for all $k \in \mathbb{N}$. On the other hand, just as for functions on phase space [18], we will see that a weighted average usually does not improve the rate of convergence of the exponents when the orbit is chaotic.

Even though our result is largely negative in this regard, we believe it is valuable to quantify the accuracy of various methods. Moreover, since our technique requires essentially the same computational effort as the standard method, and does converge more rapidly in some cases, we propose that it should be used whenever a quantitative estimate for the Lyapunov exponents is desired. Moreover, we will show that the MEGNO method [12] can be reformulated as a weighted average, and as such gives an improved estimate for μ.

The context we consider is a discrete-time dynamical system, i.e. a differentiable map $f\colon M \to M$ on an n-dimensional phase space M. We let $x_0 \in M$ denote an initial point and $v_0 \in T_{x_0}M \cong \mathbb{R}^n$ a deviation vector. These evolve under the system

$$\begin{aligned} x_{t+1} &= f(x_t)\,, \\ v_{t+1} &= Df(x_t)\,v_t, \end{aligned} \tag{1}$$

for $t \in \mathbb{N}$. The Lyapunov exponent for $(x_0, v_0) \in TM$ is then the growth rate of the norm of v_t:

$$\begin{aligned} \mu_T(x_0, v_0) &= \frac{1}{T}\ln\left(\frac{\|v_T\|}{\|v_0\|}\right), \\ \mu(x_0, v_0) &= \limsup_{T\to\infty} \mu_T(x_0, v_0)\,. \end{aligned} \tag{2}$$

A system is 'regular' if the lim sup can be replaced by lim [22]. Convergence of μ_T as $T \to \infty$ almost everywhere with respect to an invariant measure was proven in Oseledec's multiplicative ergodic theorem under certain restrictions [23–26].

As is well known, the time-T exponent can also be written as the time average of the exponential stretching factors

$$\begin{aligned} \mu_T(x_0, v_0) &= \frac{1}{T}\sum_{t=0}^{T-1} s_t, \\ s_t &\equiv \ln\left(\frac{\|v_{t+1}\|}{\|v_t\|}\right), \end{aligned} \tag{3}$$

since the sum is telescoping. Note that (3) is the time average of the *stretching* function $S\colon TM \to \mathbb{R}$ on the tangent bundle, defined by

$$S(x, v) \equiv \ln\left(\frac{\|Df(x)\,v\|}{\|v\|}\right). \tag{4}$$

so that $s_t = S(x_t, v_t)$.

A similar process can be used to compute the spectrum of exponents. A standard technique is repeated application of Gram–Schmidt orthogonalization [22, 27–30]. Given an initial orthonormal basis $Q_0 = (q_0^{(1)}, q_0^{(2)}, \ldots q_0^{(n)})$, one iterates and then orthogonalizes:

$$\begin{aligned} p^{(j)} &= Df(x_t)\,q_t^{(j)} \\ z^{(j)} &= p^{(j)} - \textstyle\sum_{k=1}^{j-1} \frac{\langle p^{(j)}, z^{(k)}\rangle}{\|z^{(k)}\|^2} z^{(k)} \end{aligned} \quad , \quad j = 1, \ldots n.$$

Normalization of the orthogonal basis $Z = (z^{(1)}, \ldots, z^{(n)})$ then gives the scaling factors and a new orthonormal basis

$$r_{t+1}^{(j)} = \|z^{(j)}\|, \quad q_{t+1}^{(j)} = \frac{z^{(j)}}{r_{t+1}^{(j)}}.$$

Iterating this process along an orbit $\{x_t\}$ gives a sequence Q_t of orthogonal matrices and growth factors $r_t^{(j)}$. The spectrum of Lyapunov exponents then becomes

$$\begin{aligned} \mu_T^{(j)}(x_0) &= \frac{1}{T}\sum_{t=0}^{T-1} \ln\left(r_{t+1}^{(j)}\right), \\ \mu^{(j)}(x_0) &= \limsup_{T\to\infty} \mu_T^{(j)}(x_0)\,. \end{aligned} \tag{5}$$

We define the stretching for the j^{th} exponent by

$$R^{(j)}(x_t) \equiv \ln\left(r_t^{(j)}\right), \text{ and let } R = \left(R^{(1)}, \ldots, R^{(n)}\right). \tag{6}$$

The paper proceeds as follows. Section 2 recalls the WBA. In section 2.2 we show that the MEGNO method [1] can be reformulated as a weighted average method, but not with a weight that is C^∞ smooth at the endpoints. In section 3, we compare five weight functions for estimating the time average in (5). Several example maps that we think of as 'typical' are discussed in section 3.1. Finally, in section 3.2 we discuss some outliers; maps which have unexpected speed-up or slow-down of convergence. These examples include maps that have shear, that are noninvertible, and those with constant Jacobian. Our conclusions appear in section 4.

2. Weighted average methods

In this section, we review weighted average methods. Since (3) is a dynamical time average, its convergence is related to that implied by Birkhoff's ergodic theorem, which states that time averages equal space averages for $L^1(M,\mathbb{R})$ functions on an ergodic invariant set, see e.g. Billingsley [31]. Unfortunately, the convergence of such averages is typically slow, i.e. no faster than $1/T$ [32]. A technique for accelerating converge of time averages is the method of the WBA developed by Das *et al* [20]. In this work, an average like that in (3) for the stretching function (6) is replaced by

$$WB_T(R)(x_0) = \sum_{t=0}^{T-1} w_T(t) R(x_t). \tag{7}$$

Here $w_T : [0,T] \to \mathbb{R}$ is a normalized weight, which we write in the form

$$w_T(t) = \frac{1}{N_T} g\left(\tfrac{t}{T}\right), \quad N_T = \sum_{t=0}^{T-1} g\left(\tfrac{t}{T}\right), \tag{8}$$

for an unnormalized weight function $g : [0,1] \to \mathbb{R}^+$.

The acceleration of the convergence of WB_T for functions on phase space relies on the fact that $g(\tau)$ is a *bump function*: it vanishes at 0 and 1 and is smooth on the closed interval $[0,1]$. In particular it has been proven that when the orbit lies on an invariant torus on which the dynamics is conjugate to a rigid rotation with incommensurate frequency, and the map and conjugacy are sufficiently smooth, then such a bump function improves the convergence of the Birkhoff average of a sufficiently smooth function on phase space [20, 33, 34]. In the best cases, the convergence is super-polynomial (faster than $1/T^k$ for any $k \in \mathbb{Z}$) or even exponential.

Our goal in the current paper is to investigate when this improvement extends to computations of the Lyapunov spectrum (5).

2.1. Bump functions

In this paper we will test several weight functions g (8) to see how they influence the convergence of the time average (7) for the Lyapunov spectrum.

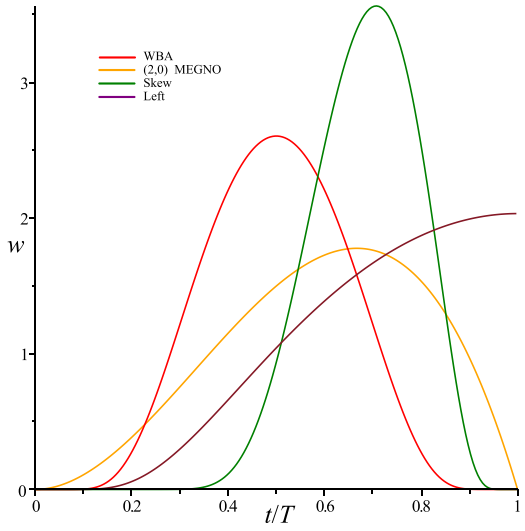

Figure 1. Normalized exponential bump (9), $(2,0)$ MEGNO (15), 'skew' (10) and 'left' (11) weight functions. The weights are normalized to have unit area, instead of using the sum as in (8).

The standard choice for g is the following C^∞ bump function [20], given by[3]

$$g_{\rm wba}(\tau) = \begin{cases} \mathrm{e}^{-(\tau(1-\tau))^{-1}} & \tau \in (0,1) \\ 0 & \tau = 0,1 \end{cases} . \tag{9}$$

We also consider the skewed bump function

$$g_{\rm skew}(\tau) = \begin{cases} \mathrm{e}^{-\left(\tau^2\left(1-\tau^2\right)\right)^{-1}} & \tau \in (0,1) \\ 0 & \tau = 0,1 \end{cases} . \tag{10}$$

This is still C^∞, but is no longer symmetric about $\tau = \frac{1}{2}$—it has a maximum at $\tau = \frac{1}{\sqrt{2}}$. As a third example, we will use the function

$$g_{\rm left}(\tau) = \begin{cases} \mathrm{e}^{-\left(\tau\left(1-\frac{1}{2}\tau\right)\right)^{-1}} & \tau \in (0,1) \\ 0 & \tau = 0,1 \end{cases} , \tag{11}$$

which is C^∞ and monotone increasing on $[0,1)$ but has a discontinuity at $\tau = 1$. This will test whether the computations of μ_T are more sensitive to initial transient behavior in the stretching of the vector $v_{\rm t}$. These functions are shown in figure 1.

2.2. MEGNO

Cincotta and Simo [12] presented a modification of the average (3), introduced to give a chaos indicator that they entitle the MEGNO. As reviewed by Cincotta and Giordano [1],

[3] Other possibilities with varying 'widths' were explored by Duignan and Meiss [33]. For time averages of functions on phase space, (9) was found to have the best convergence properties.

a generalized MEGNO, labeled by a pair integers (m,n), can be formulated. It is obtained by first computing a weighted average of the stretching factor (4) for the first t iterates:[4]

$$Y_{m,n}(t) = (m+1)\,t^n \sum_{j=0}^{t} j^m S(x_j, v_j)\,.$$

The (m,n) MEGNO is obtained as an additional time average of this quantity:

$$\bar{Y}_{(m,n)}(T) = \frac{1}{T^{m+n+1}} \sum_{t=0}^{T-1} Y_{m,n}(t)\,.$$

Note that we can reorder this double sum to obtain an expression that closely resembles the weighted average (7):

$$\bar{Y}_{(m,n)}(T) = \sum_{t=0}^{T-1} W_T^{meg}(t)\, S(x_t, v_t)\,,$$

where the MEGNO 'weight' is effectively

$$W_T^{\rm meg}(t) = \frac{m+1}{T^{m+n+1}} t^m \sum_{j=t}^{T-1} j^n\,. \tag{12}$$

Though $\bar{Y}_{(m,n)}$ now has a form similar to (7), the weight function is **not** normalized: MEGNO does not attempt to compute an accurate value for μ_T. We propose that a normalized weight function would be more appropriate, so we rescale (12) to define

$$w_T^{\rm meg}(t) = \frac{1}{N_T}\left(\frac{t}{T}\right)^m \frac{1}{T} \sum_{j=t}^{T-1} \left(\frac{j}{T}\right)^n, \tag{13}$$

where N_T is the normalization constant as in (8). Then the MEGNO-weighted average for the Lyapunov exponent is defined using this weight in (7).

According to Cincotta and Giordano [1], the most useful cases of MEGNO correspond to $(m,n) = (1,-1)$ and $(2,0)$. We will compare our results only with the second case since it was shown to converge more rapidly. When $n=0$ the sum in (13) is trivial: each term $(j/T)^0 = 1$ and so the sum is simply $T-t$. For $m=2$, the normalization factor becomes

$$N_T = \frac{1}{T^3} \sum_{t=0}^{T-1} t^2 (T-t) = \frac{T^2-1}{12T},$$

which then gives

$$w_T^{\rm meg}(t) = 12 \frac{t^2(T-t)}{T^2(T^2-1)}\,. \tag{14}$$

Note that this function vanishes at $t=0$ and T: it is a bump function like those in section 2.1. To compare more directly with these, we rescale time and set the interval to $[0,1]$ to obtain

$$g_{\rm meg}(\tau) = \tau^2(1-\tau)\,, \tag{15}$$

and then $w_T^{\rm meg}$ is given by the normalization (8). This weight is C^1 at the endpoint $\tau=0$, but only C^0 at $\tau=1$. This function is the orange curve in figure 1.

[4] To be consistent with the definition of the stretch (4), and the concept of a bump function, we shift the indices by one step from the notation of Cincotta and Giordano [1, equations (4.38)–(39)].

3. Numerical results

In this section, we compare the performance of the averages defined in section 2 by computing the Lyapunov spectrum for several example maps with regular and chaotic orbits. The first case, in section 3.1, exemplifies what we believe is the 'typical' behavior. We then describe in section 3.2 cases where the behavior is atypical due to special properties of the dynamics.

We first compute Lyapunov exponents using the Gram–Schmidt method and standard, unweighted average (5). Then we compute the exponents using the four weighting functions: exponential (9), skew (10), left (11), and $(2,0)$-MEGNO (15). In each case our goal is to understand how the averages converge as $T \to \infty$. The errors at time T could be computed if we knew the theoretical values of the exponents, $\mu^{(j)}$; however, in most cases these are not known.

Instead, we estimate the exponents by using the values $\mu^{(j)}_{T^*}$ at a fixed, large T^* to give an estimate of the 'true' answer. In order to avoid bias, rather than choosing a fixed 'truth', each method produces its own estimate. Nevertheless, we have verified that using the skew weighting as our 'absolute truth' (since as we will see below, it has the best convergence rates) does not change any of our convergence results. We will say that a Lyapunov exponent *converges as* T^{-k} if

$$|WB_T\left(R^{(j)}\right) - \mu^{(j)}_{T^*}| \sim T^{-k} \quad \text{for} \quad 1 \ll T \ll T^*,$$

i.e. if a log–log plot of the error has slope $-k$ over some interval.

3.1. Typical convergence

Recall that the WBA $WB_T(h)$ for a function $h \in C^\infty(M,\mathbb{R})$ on phase space converges slowly when an orbit is chaotic but for 'regular' orbits (those smoothly conjugate to incommensurate rotations) it converges at a rate determined by the smoothness of the weight function—for a C^∞ weight, such as (9), this can be super-polynomial [18–20, 34]. By contrast, the standard unweighted average nominally convergences at best [32] as T^{-1}.

Here we similarly observe that the Lyapunov spectrum also converges slowly whenever the orbit is chaotic, regardless of the method used. However for a regular orbit, the weighted averages do typically enhance the convergence of the Lyapunov spectrum.

As a first example, consider the three-dimensional 'discrete Lorenz map' [35]

$$\begin{aligned} x' &= y \\ y' &= -0.85x + \nu_2 y + yz \\ z' &= 0.95z - y^2. \end{aligned} \tag{16}$$

Here we fix two of the parameters ($\nu_1 = -0.85$ and $\nu_3 = 0.95$, in the notation of Gonchenko and Gonchenko [35]) and allow only the parameter ν_2 to vary. For (16), the determinant of the Jacobian is independent of the point: $\det(Df) = \nu_1\nu_3 = -0.8075$. This implies that the sum of the exponents should be

$$d^* = \ln\left(|\det(Df)|\right) \approx -0.213\,812\,223\,885\,3254.$$

However, in the computations below, we do not use this since we want to test convergence of the individual exponents (5).

Figure 2 shows the three Lyapunov exponents for this map as a function of ν_2, computed using the standard WBA weight (9) with the iterative Gram–Schmidt method (5). The orbit is arbitrarily chosen to start at $(0,-0.01,0.0001)$, and we discard the first 4000 iterates to

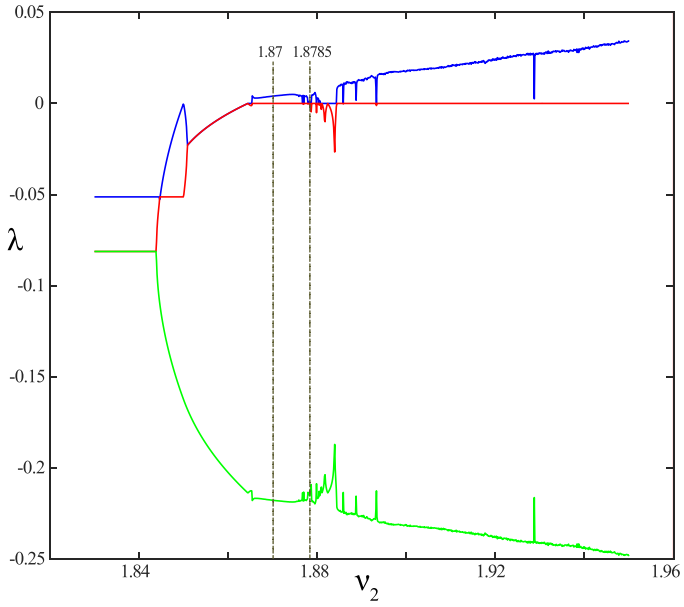

Figure 2. The three Lyapunov exponents for the discrete Lorenz map for 1000 values of $\nu_2 \in [1.83, 1.95]$. These were computed using WBA weight (9) with $T = 2(10)^4$. The dashed lines mark the values of ν_2 used for figures 3 and 4.

remove transients. The exponents are computed using (7) for the next $T = 2(10)^4$ iterates. In all cases, $|\mu^{(1)} + \mu^{(2)} + \mu^{(3)} - d^*| < 10^{-14}$, consistent with the constant Jacobian determinant. This excellent convergence follows from the fact that, up to floating point error, $\ln(r_t^{(1)}) + \ln(r_t^{(2)}) + \ln(r_t^{(3)}) = d^*$ for each t, so the sum of the averages is the average of the sum, with or without a weight function.

As shown in Gonchenko *et al* [35], the fixed point $(x, y, z) = (0, 0, 0)$ of the map (16) is stable up to $\nu_2 = 1.85$ where it undergoes a pitchfork bifurcation—at this point the largest exponent (blue in figure 2) hits zero. The newly created pair of fixed points lose stability at $\nu_2 \approx 1.8645$ in Neimark-Sacker bifurcations. The resulting pair of attracting circles or periodic orbits have basins of attraction limited by the stable and unstable manifolds of the origin, and there can be additional attractors. At $\nu_2 = 1.87$ there is a chaotic, Lorenz-like attractor, as shown in figure 3(a). This attractor has a single positive Lyapunov exponent and a tangential exponent of zero. Using the exponential weight (9) with $T^* = 2(10)^6$ gives the exponents

$$\mu \simeq (0.0039858, 0.0000000, -0.2177981)\,. \tag{17}$$

As ν_2 varies, the Lorenz-like attractor can collapse onto an attracting circle; for example this occurs at $\nu_2 = 1.8785$, see figure 4(a). For this attracting circle the maximal Lyapunov exponent is zero, and again using the exponential weight (9), we find (to higher accuracy)

$$\mu \simeq \begin{pmatrix} 0.000000000000000 \\ -0.000160991051261 \\ -0.213651232834031 \end{pmatrix}, \tag{18}$$

for the same T^*. These cases are marked with vertical dashed lines in figure 2.

The convergence of the three exponents using the five weight functions of section 2 is shown in figure 3 for $\nu_2 = 1.87$ and figure 4 for $\nu_2 = 1.8785$. These plots show the errors for

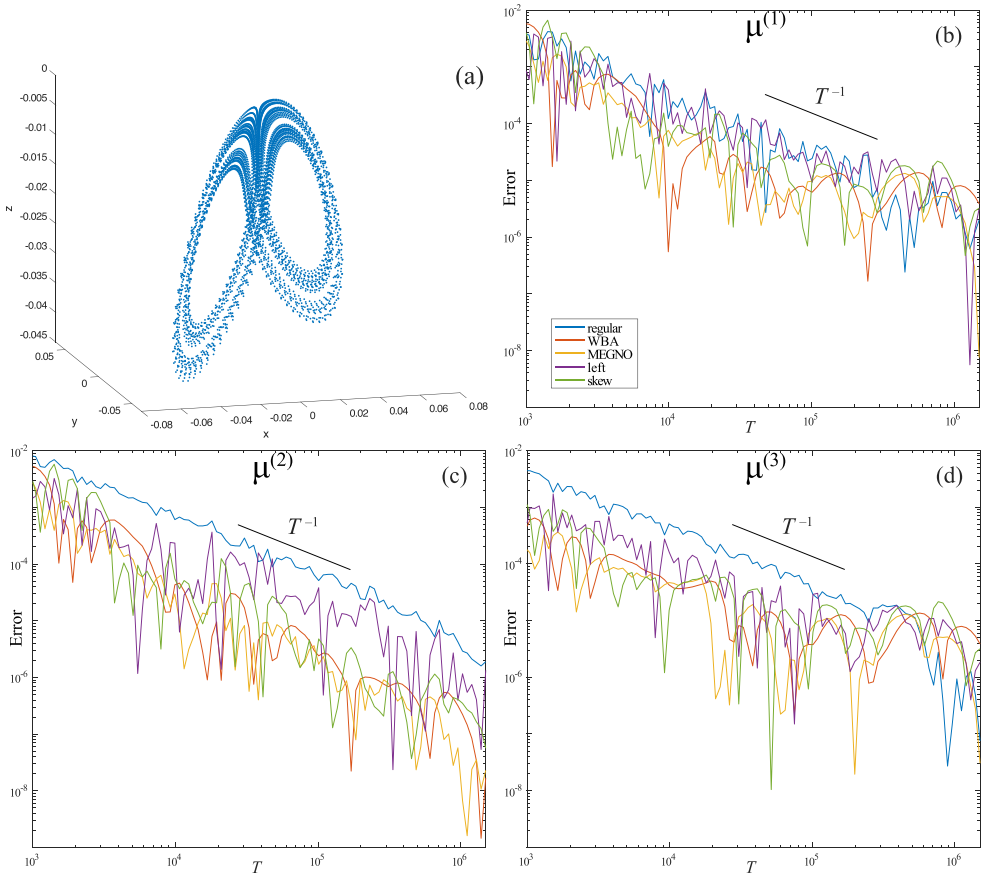

Figure 3. Panel (a): a chaotic attractor for the discrete Lorenz map (16) with $\nu_2 = 1.87$. Panels (b)–(d): corresponding errors for the three Lyapunov exponents (17) as a function of T. The curves correspond to the unweighted average (regular), exponential bump (WBA) (9), $(2,0)$-MEGNO (15), left (11), and skew (10) weights as labeled in (b). All methods appear to converge as T^{-1}, though quantitative errors for the WBA, MEGNO and skew weights are smaller.

100 logarithmically spaced values of $T \in [900, 1.5(10)^6]$. In each case the first 4000 iterates of the initial condition are discarded to remove transients. The error is estimated by comparing to $\mu^{(j)}_{T^*}$ with $T^* = 2(10)^6$.

For the chaotic attractor at $\nu_2 = 1.87$, figures 3(b)–(d) show that the values for each $\mu^{(j)}_T$ appear to converge as T^{-1} for all of the weights. Quantitatively, the unweighted and left weighted methods typically have the largest error, while—especially for the second and third exponents—the smoothly weighted methods do appear to have some reduction in error. Consistent with this, one might believe the results to 5 or 6 digits at $T = 10^6$.

By contrast, when the attractor is an invariant circle, figures 4(b)–(d) show that the C^∞ WBA and skew weights far outperform the other methods. The best convergence is for the skew weight; it reaches machine precision for all three exponents by $T = 3(10)^5$. The MEGNO

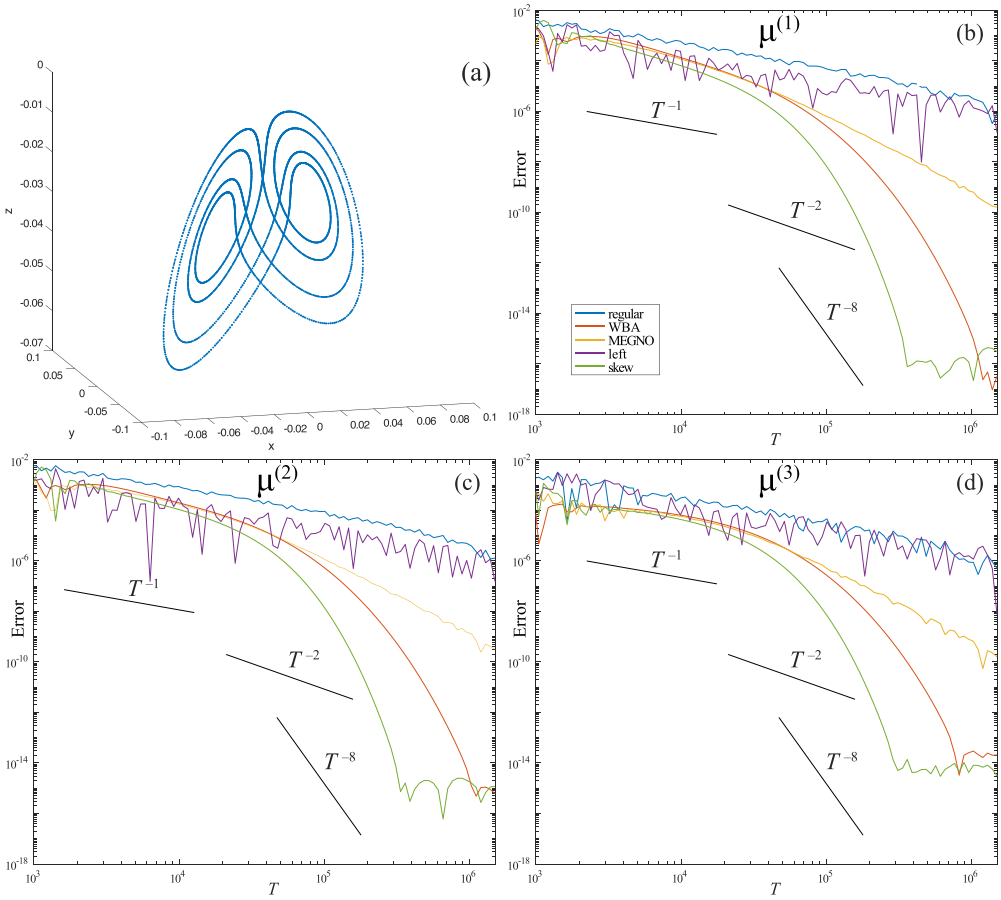

Figure 4. Panel (a): an invariant circle attractor for the discrete Lorenz map (16) with $\nu_2 = 1.8785$. Panels (b)–(d): corresponding errors for the three Lyapunov exponents (18) as a function of T. The curves for the different weights are labeled as in figure 3; however, note that the vertical scale is different here. The smoothly weighted methods converge much more quickly.

weight (15) also gives increased convergence at a rate nearing T^{-2}. The left and constant (regular) weights still converge as T^{-1}: even though the orbit is regular, there is no improvement since these weight functions are not continuous.

The behavior seen in figures 3 and 4 appears to be typical: we have seen similar performance for convergence of the Lyapunov spectrum in many other simulations for chaotic and nonchaotic orbits for various maps including the discrete Lorenz map for other parameter values, the classic Hénon map, the derived from Anosov map [36], and the 2D torus map studied in [20, section 3.7.3].

3.2. Outliers

We now describe several cases where the convergence does not follow the pattern seen in section 3.1.

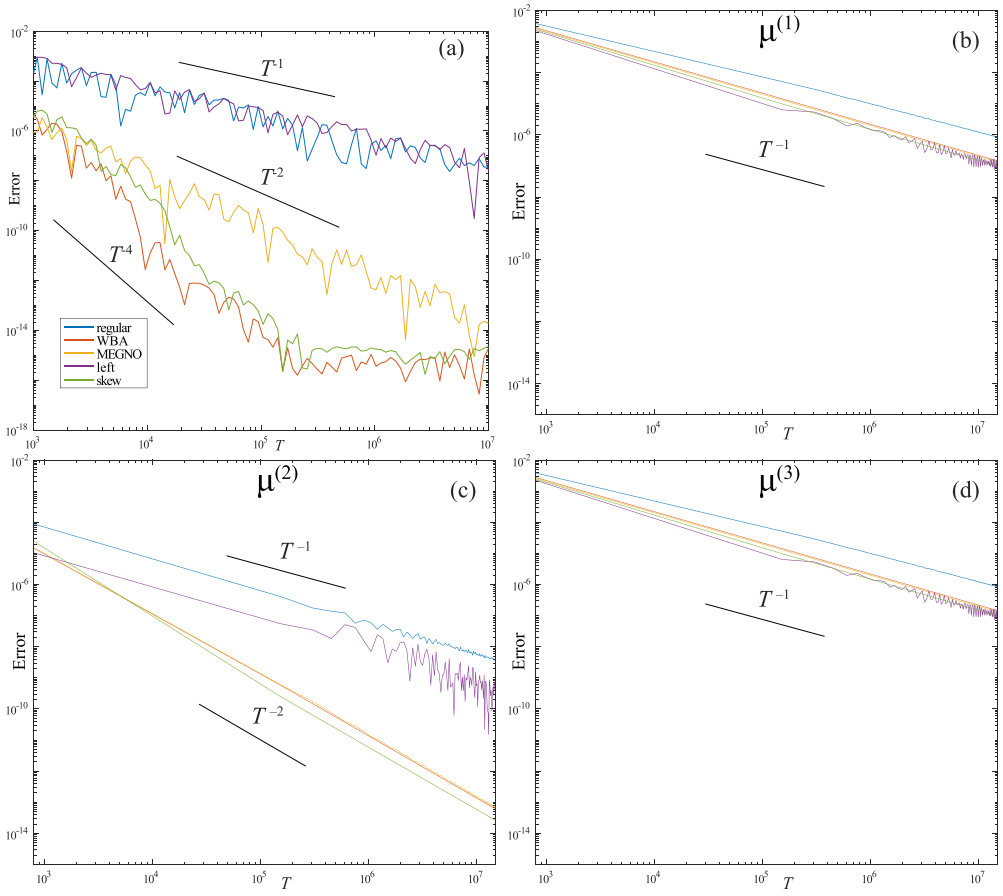

Figure 5. Convergence of weighted averages for an invariant torus of the map (19). Panel (a): time averages of $\cos(2\pi x)$. For the C^∞ weights this average converges to machine precision by $T \approx 3(10)^5$. Panel (b)–(d): convergence of the three Lyapunov exponents to zero. The errors appear to be smaller for the weighted averages, though–with the exception of $\mu^{(2)}$—the convergence rates are similar.

Dynamics with shear: integrable symplectic maps have families of invariant tori whose rotation vectors generically vary across tori: they have *shear*. This results in the linear growth of the length of vectors transverse to the tori, and this gives the slow convergence

$$\mu_T \sim \frac{\ln(T)}{T}$$

to zero for (3). It has long been known [37] that shear causes a similar slow convergence even when the map is not integrable, whenever the orbit lies on an invariant torus [9]. The weighted averages, as we describe below and see in figure 5, also exhibit this problem.

For example, consider the 3D volume-preserving map [19] for $(x, y, z) \in \mathbb{T}^2 \times \mathbb{R}$:

$$\begin{aligned} x' &= x + z' + \tfrac{1}{2}(\sqrt{5}-1) \mod 1 \\ y' &= y + 2(z')^2 + 0.4 \mod 1 \\ z' &= z - 0.02\,(\sin(2\pi x) + \sin(2\pi y) + \sin(2\pi\,(x-y))\,. \end{aligned} \tag{19}$$

The initial point $(x,y,z) = (0,0,-0.05)$ appears to lie on an invariant two-torus that is a graph over (x,y) on which the dynamics was shown [19] to have the incommensurate rotation vector $\omega \simeq (0.544\,519, 0.411\,571)$. All three of the Lyapunov exponents for this orbit are zero: the two tangential exponents vanish because the invariant set is a two-torus, and the transverse exponent is then zero because of volume-preservation. Nevertheless, since the rotation vector varies with z, the map has shear and the length of a vector transverse to the torus will grow linearly. This should result in slow convergence of the exponents.

The convergence of a Birkhoff average of the function $h = \cos(2\pi x)$ and of the largest Lyapunov exponent are compared in figure 5 for the five weight functions of section 2. The figure shows the averages for 100 values of $T \in [800, 1.5(10)^7]$. Since the map is volume preserving, no transient removal is needed. For the C^∞ weight functions the convergence rate of the Birkhoff average is excellent, as we would expect for a quasiperiodic orbit [19]. The errors for the Lyapunov exponents in panels (b)–(d) were computed by comparing to $\mu^{(j)} = 0$. Note that this convergence is very slow—even though the orbit is nonchaotic—though the weighted methods do exhibit some error reduction. As for the discrete Lorenz map, the sum of the three exponents is always near zero, to machine precision.

We observe similar behavior for other parameters and orbits of the map (19) as well as for nonchaotic orbits of the 2D Chirikov standard and the 4D Froeschlé maps.

Weak chaos: a dynamical system has *weak chaos* when the Lyapunov exponents are not positive, but it still has sensitive dependence on initial conditions. Such dynamics can lead to strange nonchaotic attractors (SNAs), where the orbit lies on a geometrically strange (fractal) set, but the Lyapunov exponent is not positive [38–44]. In this case we have previously observed that averages converge slowly for both Lyapunov exponents and functions on phase space, and adding a weight function does not improve convergence [33, 44].

Noninvertible maps: noninvertibility makes the computation of Lyapunov exponents more delicate. Denote the set of points where the Jacobian of the map is singular by

$$J_0 = \{x : \det(Df(x)) = 0\}.$$

If an invariant set intersects J_0, it may be nonsmooth and even self-intersecting [45–47]. If an ergodic component intersects J_0, even if it is smooth, some of the exponents will be undefined, since a zero determinant implies that $r_t^{(j)} = 0$ for some j, so that $\ln r_t^{(j)}$ is infinite. Moreover, the average (5) will also be undefined for initial conditions on the dense, countably infinite set of preimages of J_0. Of course it is still possible for the Lyapunov exponents to exist almost everywhere. However, from the numerical standpoint, if an orbit nears the dense set of singular points, then at least one $r_t^{(j)} \approx 0$, and this will lead to significant floating point errors.

This phenomenon is shown in figure 6 for a so-called Tinkerbell map [29]

$$\begin{aligned} x' &= 0.33x - 0.6y + x^2 - y^2 \\ y' &= 2x + 0.5y + 2xy\,. \end{aligned} \tag{20}$$

In this case J_0 is a circle centered at $(-0.2, -0.65)$ with radius $1/\sqrt{8}$ that intersects the attractor, see figure 6(a). The Lyapunov exponents for this attractor are nonpositive, $\mu \simeq (0, -0.14292)$, and as seen in figure 6(b), the convergence of $\mu^{(1)}$ for the C^∞ bump functions is excellent: since J_0 intersects the orbit transversally, $r_t^{(1)}$ is never near zero. However, $r_t^{(2)}$ does get arbitrarily close to zero infinitely often on the orbit. As seen in figure 6(c), this results in slow convergence of the second Lyapunov exponent for all methods. For this case,

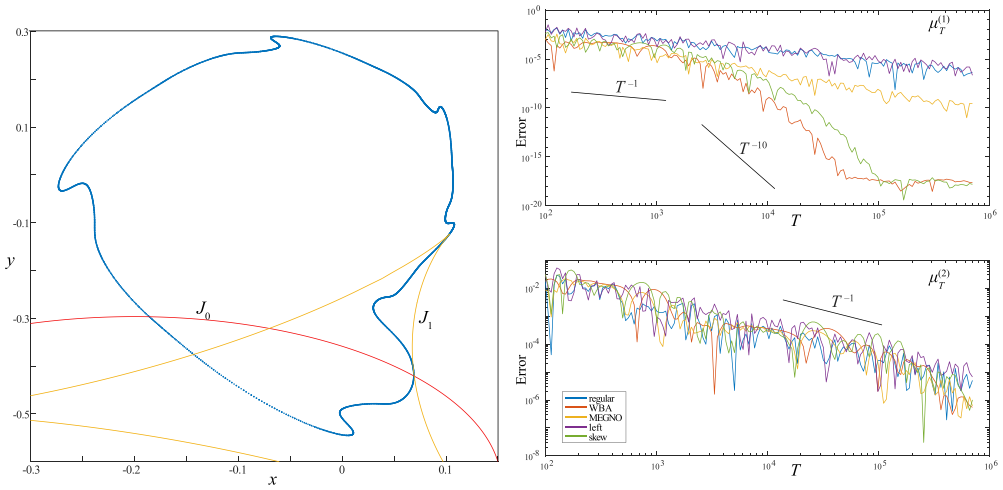

Figure 6. For the Tinkerbell map, the left panel depicts the attractor, the set of points with a singular Jacobian (J_0), and the first iterate of these points (J_1). The right panels show the errors for $\mu_T^{(1)}$ and $\mu_T^{(2)}$ as a function of T (note the difference in vertical scales). The second exponent converges slowly for all methods.

the attractor is smooth, and we observe that the convergence of $WB_T(h)$ using (9) for functions on phase space (such as $h = \cos(2\pi x)$ and the rotation number) is excellent (not shown). The implication is that for this case, the numerical difficulities are restricted to the smaller Lyapunov exponent.

Constant Jacobian: when the map f has a constant, hyperbolic Jacobian, the convergence rates of $WB_T(S)$, (7) are enhanced for smooth weight functions. Indeed, when $Df = A$ is constant, the method given in (5) is equivalent to a computing the singular values σ_j of A via normalized simultaneous power iteration. As long as $0 \leqslant \sigma_{j+1} < \sigma_j < \ldots < \sigma_1$ (the strict inequality is the generic case), we have $r_t^{(j)} \to \sigma_j$ with error $\mathcal{O}(|\sigma_{j+1}/\sigma_j|^t)$ [48]. As a result, the errors in the weighted average (5), occur only for the initial transients. The implication is that the left (11) and skew (10) weights perform better than the others, since they suppress the initial portion of the average. In contrast, when the orbit is chaotic, $WB_T(h)$ for functions h converges slowly for any weight function.

We have verified this improved convergence of $\mu_T^{(j)}$ and slow convergence of other averages for examples including Arnold's cat map and the skinny baker map [29], both of which are uniformly hyperbolic.

4. Conclusions

Lyapunov exponents are regarded not just as a categorical test but also as a quantitative estimate for chaos. For example, figure 2 shows more than a binary partition—it also gives detailed information on the onset of chaos. As is well-known, the computation of the Lyapunov exponent for orbits of a dynamical system can be formulated as a time average of the stretching function $S(x_t, v_t)$ (4) along a trajectory. It has been noted [20] that these averages can be computed using bump functions, similar to those used to compute Birkhoff averages of functions on

phase space. An advantage of this smoothing is that the results can be super-polynomially convergent when the trajectory is regular. We extended these ideas to compute the full Lyapunov spectrum (5) using weighted averages (7).

In section 3.1 we showed that a C^∞ weight function typically gives super-convergence of the Lyapunov spectrum on nonchaotic orbits, just as it does for functions on phase space. Thus a weighted averaging enhances the ability of a Lyapunov exponent computation to distinguish between regular an chaotic orbits, and in particular we showed that a C^∞ weight is best. As we saw in section 3.2, there are exceptions, including invariant sets that are tori with transverse shear; these are common in the Hamiltonian or symplectic case. Moreover, having nonpositive Lyapunov spectrum is not sufficient for super-convergence, as the example of weak-chaos shows. In addition, attractors of non-invertible maps, even if they are non-chaotic, can have slow convergence of some exponents due to singularities. Finally, convergence of exponents can be enhanced by a smooth weight for the simplest of chaotic systems: those that are uniformly hyperbolic with constant Jacobian.

As we showed in section 2, the MEGNO chaos detection method of Cincotta *et al* [1, 12] can be reformulated as a weighted time average which gives the maximal Lyapunov exponent. This reformulation shows that our scaled version of MEGNO is not just a detector of chaos, but gives a quantitative estimate of the exponent. The weight function in this case, however, is not C^∞ and this results in slower convergence of the average. Since C^∞ weight functions have essentially the same computational cost as less smooth functions, there seems to be no reason to use a less smooth weight.

The results in this paper still leave several open questions. Do these results extend to higher dimensional systems? Do similar results occur for the case of flows? Finally, is it possible to devise a technique for efficiently and accurately computing the Lyapunov spectrum for a typical, chaotic invariant set? As a final warning for the many computations of Lyapunov exponents: errors should be computed!

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments

E S was supported in part by the Simons Foundation under Award No. 636383. J D M was supported in part by the Simons Foundation under Award No. 601972. Thank you to the anonymous referees for their useful comments.

Conflict of interest

The authors have no conflicts to disclose.

ORCID iDs

E Sander 0000-0003-4478-3919
J D Meiss 0000-0002-0019-0356

References

[1] Cincotta P and Giordano C 2016 Theory and applications of the mean exponential growth factor of nearby orbits (MEGNO) method *Chaos Detection and Predictability Lecture Notes in Physics*) vol 915, ed C Skokos, G Gottwald and J Laskar (Springer) pp 93–126

[2] Bountis T and Skokos C 2012 *Complex Hamiltonian Dynamics* (*Springer Series in Synergetics)* (Springer) (https://doi.org/10.1007/978-3-642-27305-6)

[3] Christiansen F and Rugh H 1997 Computing Lyapunov spectra with continuous Gram-Schmidt orthonormalization *Nonlinearity* **10** 1063

[4] Geist K, Parlitz U and Lauterborn W 1990 Comparison of different methods for computing Lyapunov exponents *Prog. Phys.* **83** 875–93

[5] Rangarajan G, Habib S and Ryne R 1998 Lyapunov exponents without rescaling and reorthogonalization *Phys. Rev. Lett.* **80** 3747–50

[6] Sandor Z, Erdi B, Szell A and Funk B 2004 The relative Lyapunov indicator: an efficient method of chaos detection *Celestial Mech. Dynam. Astronom.* **90** 127–38

[7] Skokos C 2010 The Lyapunov characteristic exponents and their computation *Dynamics of Small Solar System Bodies and Exoplanets* ed J J Souchay and R Dvorak (Springer) pp 63–135

[8] Maffione N, Darriba L, Cincotta P and Giordano C 2011 A comparison of different indicators of chaos based on the deviation vectors: application to symplectic mappings *Celestial Mech. Dynam. Astronom.* **111** 285–307

[9] Skokos C, Gottwald G and Laskar J 2016 *Chaos Detection and Predictability* (*Lecture Notes in Physics*) vol 915 (Springer)

[10] Froeschle C, Gonczi R and Lega E 1997 The fast Lyapunov indicator: a simple tool to detect weak chaos. application to the structure of the main asteroidal belt *Planet. Space Sci.* **45** 881–6

[11] Froeschle C, Lega E and Gonczi R 1997 Fast Lyapunov indicators. application to asteroidal motion *Celest. Mech. and Dynam. Astron.* **67** 41–62

[12] Cincotta P and Simo C 2000 Simple tools to study global dynamics in non-axisymmetric galactic potentials *Astron. Astrophys.* **147** 205–28

[13] Laskar J 1993 Frequency analysis for multi-dimensional systems. Global dynamics and diffusion *Physica D* **67** 257–83

[14] Skokos C 2001 Alignment indices: a new, simple method for determining the ordered or chaotic nature of orbits *J. Phys. A* **34** 10029–43

[15] Skokos C, Bountis T C and Antonopoulos C 2007 Geometrical properties of local dynamics in Hamiltonian systems: the generalized alignment index (GALI) method *Physica D* **231** 30–54

[16] Moges H, Manos T and Skokos C 2020 On the behavior of the generalized alignment index (GALI) method for regular motion in multidimensional Hamiltonian systems *Nonlinear Phenom. Complex Syst.* **23** 153–64

[17] Gottwald G and Melbourne I 2009 On the implementation of the 0–1 test for chaos *SIAM J. Appl. Dyn. Sys.* **8** 129–45

[18] Sander E and Meiss J 2020 Birkhoff averages and rotational invariant circles for area-preserving maps *Physica D* **411** 132569

[19] Meiss J and Sander E 2021 Birkhoff averages and the breakdown of invariant tori in volume-preserving maps *Physica D* **428** 133048

[20] Das S, Saiki Y, Sander E and Yorke J 2017 Quantitative quasiperiodicity *Nonlinearity* **30** 4111

[21] Bazzani A, Giovannozzi M, Montanari C and Turchetti G 2023 Performance analysis of indicators of chaos for nonlinear dynamical systems *Phys. Rev. E* **107** 064209

[22] Dieci L and van Vleck E 2002 Lyapunov spectral intervals: theory and computation *SIAM J. Numer. Anal.* **40** 516–42

[23] Oseledec V 1968 A multiplicative ergodic theorem: Lyapunov characteristic numbers for dynamical systems *Trans. Moscow Math. Soc.* **19** 197–231 (available at: https://www.mathnet.ru/eng/mmo214)

[24] Ruelle D 1979 Ergodic theory of differentiable dynamical systems *Inst. Hautes Études Sci. Publ. Math.* **50** 27–58

[25] Raghunathan M 1979 A proof of Oseledec's multiplicative ergodic theorem *Isr. J. Math.* **32** 356–62

[26] Kuznetsov A 2012 *Hyperbolic Chaos: A Physicist's View* (Springer) (https://doi.org/10.1007/978-3-642-23666-2)

[27] Benettin G, Galgani L, Giorgilli A and Strelcyn J 1980 Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems: a method for computing all of them. part II *Meccanica* **15** 21–30
[28] Dieci L and van Vleck E 1995 Computation of a few Lyapunov exponents for continuous and discrete dynamical systems *Appl. Numer. Math.* **17** 275–91
[29] Alligood K, Sauer T and Yorke J 1997 *Chaos* (Springer) (available at: www.springer.com/us/book/9783642592812)
[30] Udwadia F and von Bremen H 2001 An efficient and stable approach for computation of Lyapunov characteristic exponents of continuous dynamical systems *Appl. Math. Comput.* **121** 219–59
[31] Billingsley P 1965 *Ergodic Theory and Information* (Wiley) (https://doi.org/10.1002/bimj.19680100113)
[32] Kachurovskii A 1996 The rate of convergence in ergodic theorems *Russ. Math. Surv.* **51** 653
[33] Duignan N and Meiss J 2023 Distinguishing between regular and chaotic orbits of flows by the weighted Birkhoff average *Physica D* **449** 133749
[34] Tong Z and Li Y 2024 Exponential convergence of the weighted Birkhoff average *J. Math. Pures Appl.* **188** 470–92
[35] Gonchenko S and Gonchenko A 2023 On discrete Lorenz-like attractors in three-dimensional maps with axial symmetry *Chaos* **33** 123104
[36] Coudene Y 2006 Pictures of hyperbolic dynamical systems *Notices Amer. Math. Soc.* **53** 8–13
[37] Casati G, Chirikov B and Ford J 1980 Marginal local instablitiy of quasi-periodic motion *Phys. Lett. A* **77** 91–94
[38] Ding M, Grebogi C and Ott E 1989 Evolution of attractors in quasiperiodically forced systems: from quasiperiodic to strange nonchaotic to chaotic *Phys. Rev.* **39A** 2593–8
[39] Glendinning P, Jäger T and Keller G 2006 How chaotic are strange non-chaotic attractors? *Nonlinearity* **19** 2005–22
[40] Yang Q, Li D, Chu Y, Li X and Grebogi C 2022 Existence and generation mechanisms of strange nonchaotic attractors in axially accelerating beam systems *Int. J. Bifur. Chaos Appl. Sci. Engrg.* **32** 2250212
[41] Li G, Li D, Wang C, Yue Y, Wen G and Grebogi C 2024 Quantifying strange property of attractors in quasiperiodically forced systems *Physica A* **633** 129417
[42] Durairaj P, Kanagaraj S, Zheng Z, Karthikeyan A and Rajagopal K 2024 Emergence of hidden strange nonchaotic attractors in a rational memristive map *Int. J. Bifur. Chaos Appl. Sci. Engrg.* **34** 2450017
[43] Zhao Y, Zhang Y and Du C 2024 Coexistence of attractors in a quasiperiodically forced Lozi map *Chaos Solitons Fractals* **187** 115381
[44] Meiss J D and Sander E 2025 Resonance and weak chaos in quasiperiodically-forced circle maps *Comm. Nonlinear Sci. Numer. Simulat.* **142** 108562
[45] Sander E 1999 Hyperbolic sets for noninvertible maps and relations *Discrete Contin. Dyn. Syst. A* **5** 339–57
[46] Sander E 2000 Homoclinic tangles for noninvertible maps *Nonlinear Analysis* **41** 259–76
[47] Josić K and Sander E 2004 The structure of synchronization sets for noninvertible systems *Chaos* **14** 249–62
[48] Sauer T 2018 *Numerical Analysis* 3rd edn (Pearson) (available at: www.pearson.com/en-us/pearsonplus/p/9780137982189)