The Lyapunov Characteristic Exponents and their computation

Charalampos Skokos[12]

[1] Astronomie et Systèmes Dynamiques, IMCCE, Observatoire de Paris, 77 avenue Denfert–Rochereau, F-75014, Paris, France

[2] Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Strasse 38, D-01187, Dresden, Germany
`hskokos@pks.mpg.de`

> *For want of a nail the shoe was lost.*
> *For want of a shoe the horse was lost.*
> *For want of a horse the rider was lost.*
> *For want of a rider the battle was lost.*
> *For want of a battle the kingdom was lost.*
> *And all for the want of a horseshoe nail.*
> ***For Want of a Nail*** (proverbial rhyme)

Summary. We present a survey of the theory of the Lyapunov Characteristic Exponents (LCEs) for dynamical systems, as well as of the numerical techniques developed for the computation of the maximal, of few and of all of them. After some historical notes on the first attempts for the numerical evaluation of LCEs, we discuss in detail the multiplicative ergodic theorem of Oseledec [102], which provides the theoretical basis for the computation of the LCEs. Then, we analyze the algorithm for the computation of the maximal LCE, whose value has been extensively used as an indicator of chaos, and the algorithm of the so–called 'standard method', developed by Benettin et al. [14], for the computation of many LCEs. We also consider different discrete and continuous methods for computing the LCEs based on the QR or the singular value decomposition techniques. Although, we are mainly interested in finite–dimensional conservative systems, i. e. autonomous Hamiltonian systems and symplectic maps, we also briefly refer to the evaluation of LCEs of dissipative systems and time series. The relation of two chaos detection techniques, namely the fast Lyapunov indicator (FLI) and the generalized alignment index (GALI), to the computation of the LCEs is also discussed.

Key words: Lyapunov exponents; Multiplicative ergodic theorem; Numerical techniques; Dynamical systems; Chaos; Variational equations; Tangent map; Chaos detection methods

1 Introduction

One of the basic information in understanding the behavior of a dynamical system is the knowledge of the spectrum of its *Lyapunov Characteristic Expo-*

nents (LCEs). The LCEs are asymptotic measures characterizing the average rate of growth (or shrinking) of small perturbations to the solutions of a dynamical system. Their concept was introduced by Lyapunov when studying the stability of nonstationary solutions of ordinary differential equations [96], and has been widely employed in studying dynamical systems since then. The value of the maximal LCE (mLCE) is an indicator of the chaotic or regular nature of orbits, while the whole spectrum of LCEs is related to entropy (Kolmogorov–Sinai entropy) and dimension–like (Lyapunov dimension) quantities that characterize the underlying dynamics.

By *dynamical system* we refer to a physical and/or mathematical system consisting of a) a set of l real state variables $x_1, x_2 \ldots, x_l$, whose current values define the state of the system, and b) a well–defined rule from which the evolution of the state with respect to an independent real variable (which is usually referred as the time t) can be derived. We refer to the number l of state variables as the *dimension* of the system, and denote a state using the vector $\mathbf{x} = (x_1, x_2 \ldots, x_l)$, or the matrix $\mathbf{x} = [\, x_1 \; x_2 \; \ldots \; x_l \,]^{\mathrm{T}}$ notation, where $(^{\mathrm{T}})$ denotes the transpose matrix. A particular state $\mathbf{x}$ corresponds to a point in an l–dimensional space $\mathcal{S}$, the so–called *phase space* of the system, while a set of states $\mathbf{x}(t)$ parameterized by t is referred as an *orbit* of the dynamical system.

Dynamical systems come in essentially two types:

1. *Continuous dynamical systems* described by differential equations of the form

$$\dot{\mathbf{x}} = \frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, t),$$

with dot denoting derivative with respect to a continuous time t and $\mathbf{f}$ being a set of l functions $f_1, f_2 \ldots, f_l$ known as the *vector field*.

2. *Discrete dynamical systems* or *maps*, described by difference equations of the form

$$\mathbf{x}_{n+1} = \mathbf{f}(\mathbf{x}_n),$$

with $\mathbf{f}$ being a set of l functions $f_1, f_2 \ldots, f_l$ and $\mathbf{x}_n$ denoting the vector $\mathbf{x}$ at a discrete time $t = n$ (integer).

Let us now define the term *chaos*. In the literature there are many definitions. A brief and concise presentation of them can be found for example in [90]. We adopt here one of the most famous definitions of chaos due to Devaney [35, p. 50], which is based on the topological approach of the problem.

Definition 1. *Let V be a set and $\mathbf{f} : V \to V$ a map on this set. We say that $\mathbf{f}$ is `chaotic` on V if*

1. $\mathbf{f}$ *has sensitive dependence on initial conditions.*
2. $\mathbf{f}$ *is topologically transitive.*
3. *periodic points are dense in V.*

Let us explain in more detail the hypothesis of this definition.

Definition 2. f $: V \to V$ *has* `sensitive dependence on initial conditions` *if there exists* $\delta > 0$ *such that, for any* $\mathbf{x} \in V$ *and any neighborhood* Δ *of* $\mathbf{x}$, *there exist* $\mathbf{y} \in \Delta$ *and* $n \geq 0$, *such that* $|\mathbf{f}^n(\mathbf{x}) - \mathbf{f}^n(\mathbf{y})| > \delta$, *where* $\mathbf{f}^n$ *denotes* n *successive applications of* $\mathbf{f}$.

Practically this definition implies that there exist points arbitrarily close to $\mathbf{x}$ which eventually separate from $\mathbf{x}$ by at least δ under iterations of $\mathbf{f}$. We point out that not all points near $\mathbf{x}$ need eventually move away from $\mathbf{x}$ under iteration, but there must be at least one such point in every neighborhood of $\mathbf{x}$.

Definition 3. f $: V \to V$ *is said to be* `topologically transitive` *if for any pair of open sets* U, $W \subset V$ *there exists* $n > 0$ *such that* $\mathbf{f}^n(U) \cap W \neq \emptyset$.

This definitions implies the existence of points which eventually move under iteration from one arbitrarily small neighborhood to any other. Consequently, the dynamical system cannot be decomposed into two disjoint invariant open sets.

From Definition 1 we see that a chaotic system possesses three ingredients: a) unpredictability because of the sensitive dependence on initial conditions, b) indecomposability, because it cannot be decomposed into noninteracting subsystems due to topological transitivity, and c) an element of regularity because it has periodic points which are dense.

Usually, in physics and applied sciences, people focus on the first hypothesis of Definition 1 and use the notion of chaos in relation to the sensitive dependence on initial conditions. The most commonly employed method for distinguishing between regular and chaotic motion, which quantifies the sensitive dependence on initial conditions, is the evaluation of the mLCE χ_1. If $\chi_1 > 0$ the orbit is chaotic. This method was initially developed at the late 70's based on theoretical results obtained at the end of the 60's.

The concept of the LCEs has been widely presented in the literature from a practical point of view, i. e. the description of particular numerical algorithms for their computation [54, 44, 62, 92, 36]. Of course, there also exist theoretical studies on the LCEs, which are mainly focused on the problem of their existence, starting with the pioneer work of Oseledec [102]. In that paper the Multiplicative Ergodic Theorem (MET), which provided the theoretical basis for the numerical computation of the LCEs, was stated and proved. The MET was the subject of several theoretical studies afterwards [108, 114, 76, 141]. A combination of important theoretical and numerical results on LCEs can be found in the seminal papers of Benettin et al. [13, 14], written almost 30 years ago, where an explicit method for the computation of all LCEs was developed.

In the present report we focus our attention both on the theoretical framework of the LCEs, as well as on the numerical techniques developed for their computation. Our goal is to provide a survey of the basic results on these issues obtained over the last 40 years, after the work of Oseledec [102]. To

this end, we present in detail the mathematical theory of the LCEs and discuss its significance without going through tedious mathematical proofs. In our approach, we prefer to present the definitions of various quantities and to state the basic theorems that guarantee the existence of the LCE, citing at the same time the papers where all the related mathematical proofs can be found. We also describe in detail the various numerical techniques developed for the evaluation of the maximal, of few or even of all LCEs, and explain their practical implementation. We do not restrict our presentation to the so–called *standard method* developed by Benettin et al. [14], as it is usually done in the literature (see e. g. [54, 44, 92]), but we include in our study modern techniques for the computation of the LCEs like the discrete and continuous methods based on the singular value decomposition (SVD) and the QR decomposition procedures.

In our analysis we deal with finite–dimensional dynamical systems and in particular with autonomous Hamiltonian systems and symplectic maps defined on a compact manifold, meaning that we exclude cases with escapes in which the motion can go to infinity. We do not consider the rather exceptional cases of completely chaotic systems and of integrable ones, i. e. systems that can be solved explicitly to give their variables as single–valued functions of time, but we consider the most general case of 'systems with divided phase space' [30, p. 19] for which *regular*[1] (quasiperiodic) and *chaotic orbits* co–exist. In such systems one sees both regular and chaotic domains. But the regular domains contain a dense set of unstable periodic orbits, which are followed by small chaotic regions. On the other hand, the chaotic domains contain stable periodic orbits that are followed by small islands of stability. Thus, the regular and chaotic domains are intricately mixed. However, there are regions where order is predominant, and other regions where chaos is predominant.

Although in our report the theory of LCEs and the numerical techniques for their evaluation are presented mainly for *conservative systems*, i.e. system that preserve the phase space volume, these techniques are not valid only for such models. For completeness sake, we also briefly discuss at the end of the report the computation of LCEs for *dissipative systems*, for which the phase space volume decreases on average, and for time series.

We tried to make the paper self–consistent by including definitions of the used terminology and brief overviews of all the necessary mathematical notions. In addition, whenever it was considered necessary, some illustrative examples have been added to the text in order to clarify the practical implementation of the presented material. Our aim has been to make this review of use for both the novice and the more experienced practitioner interested in LCEs. To this end, the reader who is interested in reading up on detailed technicalities is provided with numerous signposts to the relevant literature.

Throughout the text bold lowercase letters denote vectors, while matrices are represented, in general, by capital bold letters. We also note that the

---

[1] Regular orbits are often called *ordered orbits* (see e. g. [30, p. 18]).

most frequently used abbreviations in the text are: LCE(s), Lyapunov Characteristic Exponent(s); p–LCE, Lyapunov Characteristic Exponent of order p; mLCE, maximal Lyapunov Characteristic Exponent; p–mLCE, maximal Lyapunov Characteristic Exponent of order p; MET, multiplicative ergodic theorem; SVD, Singular Value Decomposition; PSS, Poincaré surface of section; FLI, fast Lyapunov indicator; GALI, generalized alignment index.

The paper is organized as follows:

In Section 2 we present the basic concepts of Hamiltonian systems and symplectic maps, emphasizing on the evolution of orbits, as well as of deviation vectors about them. In particular, we define the so–called variational equations for Hamiltonian systems and the tangent map for symplectic maps, which govern the time evolution of deviation vectors. We also provide some simple examples of dynamical systems and derive the corresponding set of variational equations and the corresponding tangent map.

Section 3 contains some historical notes on the first attempts for the application of the theoretical results of Oseledec [102] for the actual computation of the LCEs. We recall how the notion of exponential divergence of nearby orbits was eventually quantified by the computation of the mLCE, and we refer to the papers where the mLCE or the spectrum of LCEs were computed for the first time.

The basic theoretical results on the LCEs are presented in Section 4 following mainly the milestone papers of Oseledec [102] and Benettin et al. [13, 14]. In Section 4.1 the basic definitions and theoretical results of LCEs of various orders are presented. The practical consequences of these results on the computation of the LCEs of order 1 and of order $p > 1$ are discussed in Sections 4.2 and 4.3 respectively. Then, in Section 4.4 the MET of Oseledec [102] is stated in its various forms, while its consequences on the spectrum of LCEs for conservative dynamical systems are discussed in Section 4.5.

Section 5 is devoted to the computation of the mLCE χ_1, which is the oldest chaos indicator used in the literature. In Section 5.1 the method for the computation of the mLCE is discussed in great detail and the theoretical basis of its evaluation is explained. The corresponding algorithm is presented in Section 5.2, while the behavior of χ_1 for regular and chaotic orbits is analyzed in Section 5.3.

In Section 6 the various methods for the computation of part or of the whole spectrum of LCEs are presented. In particular, in Section 6.1 the standard method developed in [119, 14], is presented in great detail, while the corresponding algorithm is given in Section 6.2. In Section 6.3 the connection of the standard method with the discrete QR decomposition technique is discussed and the corresponding QR algorithm is given, while Section 6.4 is devoted to the presentation of other techniques for computing few or all LCEs, which are based on the SVD and QR decomposition algorithms.

In Section 7 we briefly refer to various chaos detection techniques based on the analysis of deviation vectors, as well as to a second category of chaos indicators based on the analysis of the time series constructed by the coordi-

nates of the orbit under consideration. The relation of two chaos indicators, namely the fast Lyapunov indicator (FLI) and the generalized alignment index (GALI), to the computation of the LCEs is also discussed.

Although the main topic of our presentation is the theory and the computation of the LCEs for conservative dynamical systems, in the last section of our report some complementary issues related to other types of dynamical systems are concisely presented. In particular, Section 8.1 is devoted to the computation of the LCEs for dissipative systems, while in Section 8.2 some basic features on the numerical computation of the LCEs from a time series are presented.

Finally, in the appendix A we present some basic elements of the exterior algebra theory in connection to the evaluation of wedge products, which are needed for the computation of the volume elements appearing in the definitions of the various LCEs.

2 Autonomous Hamiltonian systems and symplectic maps

In our study we consider two main types of conservative dynamical systems:

1. Continuous systems corresponding to an *autonomous Hamiltonian system* of N degrees (ND) of freedom having a Hamiltonian function

$$H(q_1, q_2, \ldots, q_N, p_1, p_2, \ldots, p_N) = h = \text{constant}, \qquad (1)$$

where q_i and p_i, $i = 1, 2, \ldots, N$ are the generalized coordinates and conjugate momenta respectively. An orbit in the $l = 2N$–dimensional phase space $\mathcal{S}$ of this system is defined by a vector

$$\mathbf{x}(t) = (q_1(t), q_2(t), \ldots, q_N(t), p_1(t), p_2(t), \ldots, p_N(t)),$$

with $x_i = q_i$, $x_{i+N} = p_i$, $i = 1, 2, \ldots, N$. The time evolution of this orbit is governed by the Hamilton equations of motion, which in matrix form are given by

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) = \left[\frac{\partial H}{\partial \mathbf{p}} \quad -\frac{\partial H}{\partial \mathbf{q}} \right]^{\mathrm{T}} = \mathbf{J}_{2N} \cdot \mathbf{DH}, \qquad (2)$$

with $\mathbf{q} = (q_1(t), q_2(t), \ldots, q_N(t))$, $\mathbf{p} = (p_1(t), p_2(t), \ldots, p_N(t))$, and

$$\mathbf{DH} = \left[\frac{\partial H}{\partial q_1} \; \frac{\partial H}{\partial q_2} \; \cdots \; \frac{\partial H}{\partial q_N} \; \frac{\partial H}{\partial p_1} \; \frac{\partial H}{\partial p_2} \; \cdots \; \frac{\partial H}{\partial p_N} \right]^{\mathrm{T}}.$$

Matrix $\mathbf{J}_{2N}$ has the following block form

$$\mathbf{J}_{2N} = \begin{bmatrix} \mathbf{0}_N & \mathbf{I}_N \\ -\mathbf{I}_N & \mathbf{0}_N \end{bmatrix},$$

with $\mathbf{I}_N$ being the $N \times N$ identity matrix and $\mathbf{0}_N$ being the $N \times N$ matrix with all its elements equal to zero. The solution of (2) is formally written with respect to the induced flow $\Phi^t : \mathcal{S} \to \mathcal{S}$ as

$$\mathbf{x}(t) = \Phi^t \left(\mathbf{x}(0) \right). \tag{3}$$

2. Symplectic maps of $l = 2N$ dimensions having the form

$$\mathbf{x}_{n+1} = \mathbf{f}(\mathbf{x}_n). \tag{4}$$

A *symplectic map* is an area–preserving map whose *Jacobian matrix*

$$\mathbf{M} = \mathbf{Df}(\mathbf{x}) = \frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_{2N}} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_{2N}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_{2N}}{\partial x_1} & \frac{\partial f_{2N}}{\partial x_2} & \cdots & \frac{\partial f_{2N}}{\partial x_{2N}} \end{bmatrix},$$

satisfies

$$\mathbf{M}^{\mathrm{T}} \cdot \mathbf{J}_{2N} \cdot \mathbf{M} = \mathbf{J}_{2N}. \tag{5}$$

The state of the system at the discrete time $t = n$ is given by

$$\mathbf{x}_n = \Phi^n \left(\mathbf{x}_0 \right) = \left(\mathbf{f} \right)^n \left(\mathbf{x}_0 \right), \tag{6}$$

where $\left(\mathbf{f} \right)^n \left(\mathbf{x}_0 \right) = \mathbf{f}(\mathbf{f}(\cdots \mathbf{f}(\mathbf{x}_0) \cdots))$, n times.

2.1 Variational equations and tangent map

Let us now turn our attention to the (continuous or discrete) time evolution of deviation vectors $\mathbf{w}$ from a given reference orbit of a dynamical system. These vectors evolve on the *tangent space* $\mathcal{T}_{\mathbf{x}}\mathcal{S}$ of $\mathcal{S}$. We denote by $d_{\mathbf{x}}\Phi^t$ the linear mapping which maps the tangent space of $\mathcal{S}$ at point $\mathbf{x}$ onto the tangent space at point $\Phi^t(\mathbf{x})$, and so we have $d_{\mathbf{x}}\Phi^t : \mathcal{T}_{\mathbf{x}}\mathcal{S} \to \mathcal{T}_{\Phi^t(\mathbf{x})}\mathcal{S}$ with

$$\mathbf{w}(t) = d_{\mathbf{x}}\Phi^t \, \mathbf{w}(0), \tag{7}$$

where $\mathbf{w}(0)$, $\mathbf{w}(t)$ are deviation vectors with respect to the given orbit at times $t = 0$ and $t > 0$ respectively.

In the case of the Hamiltonian system (1) an initial deviation vector $\mathbf{w}(0) = (\delta x_1(0), \delta x_2(0), \dots, \delta x_{2N}(0))$ from the solution $\mathbf{x}(t)$ (3) evolves on the tangent space $\mathcal{T}_{\mathbf{x}}\mathcal{S}$ according to the so–called *variational equations*

$$\dot{\mathbf{w}} = \mathbf{Df}(\mathbf{x}(t)) \cdot \mathbf{w} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}(t)) \cdot \mathbf{w} = \left[\mathbf{J}_{2N} \cdot \mathbf{D}^2\mathbf{H}(\mathbf{x}(t)) \right] \cdot \mathbf{w} =: \mathbf{A}(t) \cdot \mathbf{w}, \tag{8}$$

with $\mathbf{D}^2\mathbf{H}(\mathbf{x}(t))$ being the Hessian matrix of Hamiltonian (1) calculated on the reference orbit $\mathbf{x}(t)$ (3), i. e.

$$\mathbf{D}^2\mathbf{H}(\mathbf{x}(t))_{i,j} = \left.\frac{\partial^2 H}{\partial x_i \partial x_j}\right|_{\varPhi^t(\mathbf{x}(0))} \quad , \quad i,j = 1,2,\ldots,2N.$$

We underline that equations (8) represent a set of *linear differential equations* with respect to $\mathbf{w}$, having time dependent coefficients since, matrix $\mathbf{A}(t)$ depends on the particular reference orbit, which is a function of time t. The solution of (8) can be written as

$$\mathbf{w}(t) = \mathbf{Y}(t) \cdot \mathbf{w}(0), \tag{9}$$

where $\mathbf{Y}(t)$ is the so–called *fundamental matrix* of solutions of (8), satisfying the equation

$$\dot{\mathbf{Y}}(t) = \mathbf{Df}(\mathbf{x}(t)) \cdot \mathbf{Y}(t) = \mathbf{A}(t) \cdot \mathbf{Y}(t) \ , \ \text{ with } \ \mathbf{Y}(0) = \mathbf{I}_{2N}. \tag{10}$$

In the case of the symplectic map (4) the evolution of a deviation vector $\mathbf{w}_n$, with respect to a reference orbit $\mathbf{x}_n$, is given by the corresponding *tangent map*

$$\mathbf{w}_{n+1} = \mathbf{Df}(\mathbf{x}_n) \cdot \mathbf{w}_n = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}_n) \cdot \mathbf{w}_n =: \mathbf{M}_n \cdot \mathbf{w}_n. \tag{11}$$

Thus, the evolution of the initial deviation vector $\mathbf{w}_0$ is given by

$$\mathbf{w}_n = \mathbf{M}_{n-1} \cdot \mathbf{M}_{n-2} \cdot \ldots \cdot \mathbf{M}_0 \cdot \mathbf{w}_0 =: \mathbf{Y}_n \cdot \mathbf{w}_0, \tag{12}$$

with $\mathbf{Y}_n$ satisfying the relation

$$\mathbf{Y}_{n+1} = \mathbf{M}_n \cdot \mathbf{Y}_n = \mathbf{Df}(\mathbf{x}_n) \cdot \mathbf{Y}_n, \ \text{ with } \ \mathbf{Y}_0 = \mathbf{I}_{2N}. \tag{13}$$

2.2 Simple examples of dynamical systems

As representative examples of dynamical systems we consider a) the well–known 2D Hénon–Heiles system [72], having the Hamiltonian function

$$H_2 = \frac{1}{2}(p_x^2 + p_y^2) + \frac{1}{2}(x^2 + y^2) + x^2 y - \frac{1}{3}y^3, \tag{14}$$

with equations of motion

$$\dot{\mathbf{x}} = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{p}_x \\ \dot{p}_y \end{bmatrix} = \mathbf{J}_4 \cdot \mathbf{DH}_2 = \mathbf{J}_4 \cdot \begin{bmatrix} x + 2xy \\ y + x^2 - y^2 \\ p_x \\ p_y \end{bmatrix} \Rightarrow \begin{cases} \dot{x} &= p_x \\ \dot{y} &= p_y \\ \dot{p}_x &= -x - 2xy \\ \dot{p}_y &= -y - x^2 + y^2 \end{cases}, \tag{15}$$

and b) the 4–dimensional (4d) symplectic map

$$\begin{aligned} x_{1,n+1} &= x_{1,n} + x_{3,n} \\ x_{2,n+1} &= x_{2,n} + x_{4,n} \\ x_{3,n+1} &= x_{3,n} - \nu \sin(x_{1,n+1}) - \mu[1 - \cos(x_{1,n+1} + x_{2,n+1})] \\ x_{4,n+1} &= x_{4,n} - \kappa \sin(x_{2,n+1}) - \mu[1 - \cos(x_{1,n+1} + x_{2,n+1})] \end{aligned} \quad (\text{mod } 2\pi), \tag{16}$$

with parameters ν, κ and μ. All variables are given (mod 2π), so $x_{i,n} \in [\pi, \pi)$, for $i = 1, 2, 3, 4$. This map is a variant of Froeschlé's 4d symplectic map [52] and its behavior has been studied in [31, 123]. It is easily seen that its Jacobian matrix satisfies equation (5).

2.3 Numerical integration of variational equations

When dealing with Hamiltonian systems the variational equations (8) have to be integrated simultaneously with the Hamilton equations of motion (2). Let us clarify the issue by looking to a specific example. The variational equations of the 2D Hamiltonian (14) are

$$
\dot{\mathbf{w}} = \begin{bmatrix} \dot{\delta x} \\ \dot{\delta y} \\ \dot{\delta p_x} \\ \dot{\delta p_y} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1-2y & -2x & 0 & 0 \\ -2x & -1+2y & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \delta x \\ \delta y \\ \delta p_x \\ \delta p_y \end{bmatrix} \Rightarrow
$$
$$
\begin{cases} \dot{\delta x} = \delta p_x \\ \dot{\delta y} = \delta p_y \\ \dot{\delta p_x} = (-1-2y)\delta x + (-2x)\delta y \\ \dot{\delta p_y} = (-2x)\delta x + (-1+2y)\delta y \end{cases} . \tag{17}
$$

This system of differential equations is linear with respect to δx, δy, δp_x, δp_y, but it cannot be integrated independently of system (15) since the x and y variables appear explicitly in it. Thus, if we want to follow the time evolution of an initial deviation vector $\mathbf{w}(0)$ with respect to a reference orbit with initial condition $\mathbf{x}(0)$, we are obliged to integrate simultaneously the whole set of differential equations (15) and (17).

A numerical scheme for integrating the variational equations (8), which exploits their linearity and is particularly useful when we need to evolve more than one deviation vectors is the following. Solving the Hamilton equations of motion (2) by any numerical integration scheme we obtain the time evolution of the reference orbit (3). In practice this means that we know the values $\mathbf{x}(t_i)$ for $t_i = i\,\Delta t$, $i = 0, 1, 2, \ldots$, where Δt is the integration time step. Inserting this numerically known solution to the variational equations (8) we end up with a linear system of differential equations with constant coefficients for every time interval $[t_i, t_i + \Delta t)$, which can be solved explicitly.

For example, in the particular case of Hamiltonian (14), the system of variational equations (17) becomes

$$
\begin{aligned} \dot{\delta x} &= \delta p_x \\ \dot{\delta y} &= \delta p_y \\ \dot{\delta p_x} &= [-1-2y(t_i)]\,\delta x + [-2x(t_i)]\,\delta y \\ \dot{\delta p_y} &= [-2x(t_i)]\,\delta x + [-1+2y(t_i)]\,\delta y \end{aligned} \quad , \text{ for } t \in [t_i, t_i + \Delta t), \tag{18}
$$

which is a linear system of differential equations with constant coefficients and thus, easily solved. In particular, equations (18) can by considered as the Hamilton equations of motion corresponding to the Hamiltonian function

$$
H_V(\delta x, \delta y, \delta p_x, \delta p_y) =
$$

$$
\tfrac{1}{2}\left(\delta p_x^2 + \delta p_y^2\right) + \tfrac{1}{2}\left\{[1+2y(t_i)]\,\delta x^2 + [1-2y(t_i)]\,\delta y^2 + 2\,[2x(t_i)]\,\delta x \delta y\right\} . \tag{19}
$$

The Hamiltonian formalism (19) of the variational equations (18) is a specific example of a more general result. In the case of the usual Hamiltonian function

$$H(\mathbf{q}, \mathbf{p}) = \frac{1}{2} \sum_{i=1}^{N} p_i^2 + V(\mathbf{q}), \qquad (20)$$

with $V(\mathbf{q})$ being the potential function, the variational equations (8) for the time interval $[t_i, t_i + \Delta t)$ take the form (see e. g. [12])

$$\dot{\mathbf{w}} = \begin{bmatrix} \dot{\boldsymbol{\delta}\mathbf{q}} \\ \dot{\boldsymbol{\delta}\mathbf{p}} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_N & \mathbf{I}_N \\ -\mathbf{D}^2\mathbf{V}(\mathbf{q}(t_i)) & \mathbf{0}_N \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{\delta}\mathbf{q} \\ \boldsymbol{\delta}\mathbf{p} \end{bmatrix}$$

with $\boldsymbol{\delta}\mathbf{q} = (\delta q_1(t), \delta q_2(t), \dots, \delta q_N(t))$, $\boldsymbol{\delta}\mathbf{p} = (\delta p_1(t), \delta p_2(t) \dots, \delta p_N(t))$, and

$$\mathbf{D}^2\mathbf{V}(\mathbf{q}(t_i))_{jk} = \frac{\partial^2 V(\mathbf{q})}{\partial q_j \partial q_k}\bigg|_{\mathbf{q}(t_i)} \quad, \quad j,k = 1, 2, \dots, N.$$

Thus, the tangent dynamics of (20) is represented by the Hamiltonian function (see e. g. [105])

$$H_V(\boldsymbol{\delta}\mathbf{q}, \boldsymbol{\delta}\mathbf{p}) = \frac{1}{2} \sum_{j=1}^{N} \delta p_i^2 + \frac{1}{2} \sum_{j,k}^{N} \mathbf{D}^2\mathbf{V}(\mathbf{q}(t_i))_{jk} \delta q_j \delta q_k.$$

2.4 Tangent dynamics of symplectic maps

In the case of symplectic maps, the dynamics on the tangent space, which is described by the tangent map (11), cannot be considered separately from the phase space dynamics determined by the map (4) itself. This is because the tangent map depends explicitly on the reference orbit $\mathbf{x}_n$.

For example, the tangent map of the 4d map (16) is

$$\begin{aligned} \delta x_{1,n+1} &= \delta x_{1,n} + \delta x_{3,n} \\ \delta x_{2,n+1} &= \delta x_{2,n} + \delta x_{4,n} \\ \delta x_{3,n+1} &= a_n \delta x_{1,n} + b_n \delta x_{2,n} + (1 + a_n)\delta x_{3,n} + b_n \delta x_{4,n} \\ \delta x_{4,n+1} &= b_n \delta x_{1,n} + c_n \delta x_{2,n} + b_n \delta x_{3,n} + (1 + c_n)\delta x_{4,n} \end{aligned} \qquad (21)$$

with

$$\begin{aligned} a_n &= -\nu \cos(x_{1,n+1}) - \mu \sin(x_{1,n+1} + x_{2,n+1}) \\ b_n &= -\mu \sin(x_{1,n+1} + x_{2,n+1}) \\ c_n &= -\kappa \cos(x_{2,n+1}) - \mu \sin(x_{1,n+1} + x_{2,n+1}) \end{aligned},$$

which explicitly depend on $x_{1,n}$, $x_{2,n}$, $x_{3,n}$, $x_{4,n}$. Thus, the evolution of a deviation vector requires the simultaneous iteration of both the map (16) and the tangent map (21).

3 Historical introduction: The early days of LCEs

Prior to the discussion of the theory of the LCEs and the presentation of the various algorithms for their computation, it would be interesting to go back in time and see how the notion of LCEs, as well as the nowadays taken for granted techniques for evaluating them, were formed.

The LCEs are asymptotic measures characterizing the average rate of growth (or shrinking) of small perturbations to the orbits of a dynamical system, and their concept was introduced by Lyapunov [96]. Since then they have been extensively used for studying dynamical systems. As it has already been mentioned, one of the basic features of chaos is the sensitive dependence on initial conditions and the LCEs provide quantitative measures of response sensitivity of a dynamical system to small changes in initial conditions. For a chaotic orbit at least one LCE is positive, implying exponential divergence of nearby orbits, while in the case of regular orbits all LCEs are zero. Therefore, the presence of positive LCEs is a signature of chaotic behavior. Usually the computation of only the mLCE χ_1 is sufficient for determining the nature of an orbit, because $\chi_1 > 0$ guarantees that the orbit is chaotic.

Characterization of the chaoticity of an orbit in terms of the divergence of nearby orbits was introduced by Hénon and Heiles [72] and further used by several authors (e. g. [48, 51, 52, 131, 22, 21]). In these studies two initial points were chosen very close to each other, having phase space distance of about $10^{-7} - 10^{-6}$, and were evolved in time. If the two initial points were located in a region of regular motion their distance increased approximately linearly with time, while if they were belonging to a chaotic region the distance exhibited an exponential increase in time (Figure 1).

Although the theory of LCEs was applied to characterize chaotic motion by Oseledec [102], quite some time passed until the connection between LCEs and exponential divergence was made clear [10, 106]. It is worth mentioning that Casartelli et al. [21] defined a quantity, which they called 'stochastic parameter', in order to quantify the exponential divergence of nearby orbits, which was realized afterwards in [10] to be an estimator of the mLCE for $t \to \infty$.

So, the mLCE χ_1 was estimated for the first time in [10], as the limit for $t \to \infty$ of an appropriate quantity $X_1(t)$, which was obtained from the evolution of the phase space distance of two initially close orbits. In this paper some nowadays well–established properties of $X_1(t)$ were discussed, like for example the fact that $X_1(t)$ tends to zero in the case of regular orbits following a power law $\propto t^{-1}$, while it tends to nonzero values in the case of chaotic orbits (Figure 2). The same algorithm was immediately applied for the computation of the mLCE of a dissipative system, namely the Lorenz system [99].

The next improvement of the computational algorithm for the evaluation of the mLCE was introduced in [34], where the variational equations were used for the time evolution of deviation vectors instead of the previous approach of the simultaneous integration of two initially close orbits. This more direct
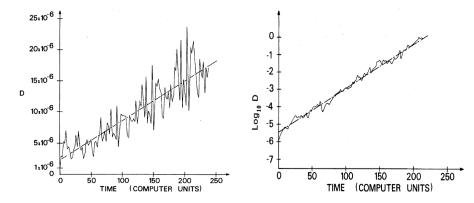
Fig. 1. Typical behavior of the time evolution of the distance D between two initially close orbits in the case of regular and chaotic orbits. The particular results are obtained for a 2D Hamiltonian system describing a Toda lattice of two particles with unequal masses (see [22] for more details). The initial Euclidian distance of the two orbits in the 4–dimensional phase space is $D_0 = 10^{-6}$. D exhibits a linear (on the average) growth when the two orbits are initially located in a region of regular motion (left panel), while it grows exponentially in the case of chaotic orbits (right panel). The big difference in the values of D between the two cases is evident since the two panels have the same horizontal (time) axis but different vertical ones. In particular, the vertical axis is linear in the left panel and logarithmic in the right panel (after [22]).

approach constituted a significant improvement for the computation of the mLCE since it allowed the use of larger integration steps, diminishing the real computational time and also eliminated the problem of choosing a suitable initial distance between the nearby orbits.

In [11] a theorem was formulated, which led directly to the development of a numerical technique for the computation of some or even of all LCEs, based on the time evolution of more than one deviation vectors, which are kept linearly independent through a Gram–Schmidt orthonormalization procedure (see also [9]). This method was explained in more detail in [119], where it was applied to the study of the Lorenz system and was also presented in [12], where it was applied to the study of an ND Hamiltonian system with N varying from 2 to 10.

The theoretical framework, as well as the numerical method for the computation of the maximal, some or even all LCEs were given in the seminal papers of Benettin et al. [13, 14]. In [14] the complete set of LCEs was calculated for several different Hamiltonian systems, including four and six dimensional maps. In Figure 3 we show the results of [14] concerning the 3D Hamiltonian system of [34]. The importance of the papers of Benettin et al. [13, 14] is reflected by the fact that almost all methods for the computation of the LCEs are more or less based on them. Immediately the ideas presented in
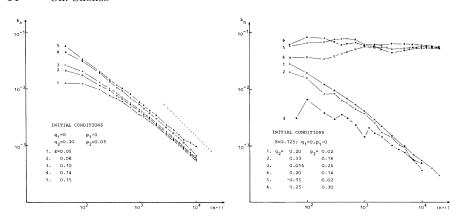
Fig. 2. Evolution of $X_1(t)$ (denoted as k_n) with respect to time t (denoted by $n \times \tau$) in log–log scale for several orbits of the Hénon–Heiles system (14). In the left panel $X_1(t)$ is computed for 5 different regular orbits at different energies H_2 (denoted as E) and it tends to zero following a power law $\propto t^{-1}$. A dashed straight line corresponding to a function proportional to t^{-1} is also plotted. In the right panel the evolution of $X_1(t)$ is plotted for three regular orbits (curves 1–3) and three chaotic ones (curves 4–6) for $H_2 = 0.125$. Note that the values of the initial conditions given in the two panels correspond to $q_1 = x$, $q_2 = y$, $p_1 = p_x$, $p_2 = p_y$ in (14) (after [10]).

[13, 14] were used for the computation of the LCEs for a variety of dynamical systems like infinite–dimensional systems described by delay differential equations [46], dissipative systems [44], conservative systems related to Celestial Mechanics problems [53, 55], as well as for the determination of the LCEs from a time series [144, 118].

4 Lyapunov Characteristic Exponents: Theoretical treatment

In this section we define the LCEs of various orders presenting also the basic theorems which guarantee their existence and provide the theoretical background for their numerical evaluation. In our presentation we basically follow the fundamental papers of Oseledec [102] and of Benettin et al. [13] where all the theoretical results of the current section are explicitly proved.

We consider a continuous or discrete dynamical system defined on a differentiable manifold $\mathcal{S}$. Let $\Phi^t(\mathbf{x})$ denote the state at time t of the system which at time $t = 0$ was at $\mathbf{x}$ (see equations (3) and (6) for the continuous and discrete case respectively). For the action of Φ^t over two successive time intervals t and s we have the following composition law

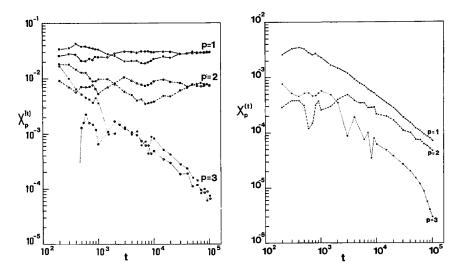$$\Phi^{t+s} = \Phi^t \circ \Phi^s.$$

Fig. 3. Time evolution of appropriate quantities denoted by $X_p^{(t)}$, $p = 1, 2, 3$, having respectively as limits for $t \to \infty$ the first three LCEs χ_1, χ_2, χ_3, for two chaotic orbits (left panel) and one regular orbit (right panel) of the 3D Hamiltonian system initially studied in [34] (see [14] for more details). In both panels $X_3^{(t)}$ tends to zero implying that $\chi_3 = 0$. This is due to the fact that Hamiltonian systems have at least one vanishing LCE, namely the one corresponding to the direction along the flow (this property is explained in Section 4.5). On the other hand, χ_1 and χ_2 seem to get nonzero values (with $\chi_1 > \chi_2$) for chaotic orbits, while they appear to vanish for regular orbits (after [14]).

The tangent space at $\mathbf{x}$ is mapped onto the tangent space at $\varPhi^t(\mathbf{x})$ by the differential $d_{\mathbf{x}}\varPhi^t$ according to equation (7). The action of $\varPhi^t(\mathbf{x})$ is given by equation (9) for continuous systems and by equation (12) for discrete ones. Thus, the action of $d_{\mathbf{x}}\varPhi^t$ on a particular initial deviation vector $\mathbf{w}$ of the tangent space is given by the multiplication of matrix $\mathbf{Y}(t)$ for continuous systems or $\mathbf{Y}_n$ for discrete systems with vector $\mathbf{w}$. From equations (9) and (12) we see that the action of $d_{\mathbf{x}}\varPhi^t$ over two successive time intervals t and s satisfies the composition law

$$d_{\mathbf{x}}\varPhi^{t+s} = d_{\varPhi^s(\mathbf{x})}\varPhi^t \circ d_{\mathbf{x}}\varPhi^s. \tag{22}$$

This equation can be written in the form

$$\mathbf{R}(t + s, \mathbf{x}) = \mathbf{R}(t, \varPhi^s(\mathbf{x})) \cdot \mathbf{R}(s, \mathbf{x}), \tag{23}$$

where $\mathbf{R}(t, \mathbf{x})$ is the matrix corresponding to $d_{\mathbf{x}}\varPhi^t$. We note that since $\mathbf{Y}(0) = \mathbf{Y}_0 = \mathbf{I}_{2N}$ we get $d_{\mathbf{x}}\varPhi^0\mathbf{w} = \mathbf{w}$ and $\mathbf{R}(0, \mathbf{x}) = \mathbf{I}_{2N}$. A function $\mathbf{R}(t, \mathbf{x})$ satisfying relation (23) is called a *multiplicative cocycle* with respect to the dynamical system $\varPhi^t$.

Let $\mathcal{S}$ be a measure space with a normalized measure μ such that

$$\mu(\mathcal{S}) = 1 \ , \ \ \mu\left(\Phi^t \mathcal{A}\right) = \mu(\mathcal{A}) \tag{24}$$

for $\mathcal{A} \subset \mathcal{S}$. Suppose also that a smooth Riemannian metric $\|\ \|$ is defined on $\mathcal{S}$. We consider the multiplicative cocycle $\mathbf{R}(t, \mathbf{x})$ corresponding to $d_{\mathbf{x}}\Phi^t$ and we are interested in its asymptotic behavior for $t \to \pm\infty$. Since, as mentioned by Oseledec [102], the case $t \to +\infty$ is analogous to the case $t \to -\infty$, we restrict our treatment to the case $t \to +\infty$, where time is increasing. In order to clarify what we are practically interested in let us consider a nonzero vector $\mathbf{w}$ of the tangent space $\mathcal{T}_{\mathbf{x}}\mathcal{S}$ at $\mathbf{x}$. Then the quantity

$$\lambda_t(\mathbf{x}) = \frac{\|d_{\mathbf{x}}\Phi^t \mathbf{w}\|}{\|\mathbf{w}\|}$$

is called the *coefficient of expansion in the direction of* $\mathbf{w}$. If

$$\limsup_{t \to \infty} \frac{1}{t} \ln \lambda_t(\mathbf{x}) > 0$$

we say that exponential diverge occurs in the direction of $\mathbf{w}$. Of course the basic question we have to answer is whether the *characteristic exponent* (also called *characteristic exponent of order 1*)

$$\lim_{t \to \infty} \frac{1}{t} \ln \lambda_t(\mathbf{x})$$

exists.

We will answer this question in a more general framework without restricting ourselves to multiplicative cocycles. So, the results presented in the following Section 4.1, are valid for a general class of matrix functions, a subclass of which contains the multiplicative cocycles which are of more practical interest to us, since they describe the time evolution of deviation vectors for the dynamical systems we study.

4.1 Definitions and basic theorems

Let $\mathbf{A}_t$ be an $n \times n$ matrix function defined either on the whole real axis or on the set of integers, such that $\mathbf{A}_0 = \mathbf{I}_n$, for each time t the value of function $\mathbf{A}_t$ is a nonsingular matrix and $\|\mathbf{A}_t\|$ the usual 2–norm of $\mathbf{A}_t$[2]. In particular, we consider only matrices $\mathbf{A}_t$ satisfying

---

[2] The 2–norm $\|\mathbf{A}\|$ of an $n \times n$ matrix $\mathbf{A}$ is induced by the 2–norm of vectors, i. e. the usual Euclidean norm $\|\mathbf{x}\| = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}$, by

$$\|\mathbf{A}\| = \max_{\mathbf{x} \neq 0} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|}$$

and is equal to the largest eigenvalue of matrix $\sqrt{\mathbf{A}^{\mathrm{T}}\mathbf{A}}$.

$$\max\left\{\|\mathbf{A}_t\|, \|\mathbf{A}_t^{-1}\|\right\} \leq e^{ct} \tag{25}$$

with $c > 0$ a suitable constant.

Definition 4. *Considering a matrix function $\boldsymbol{A}_t$ as above and a nonzero vector $\mathbf{w}$ of the Euclidian space $\mathbb{R}^n$ the quantity*

$$\chi(\boldsymbol{A}_t, \mathbf{w}) = \limsup_{t \to \infty} \frac{1}{t} \ln \|\boldsymbol{A}_t \mathbf{w}\| \tag{26}$$

is called the `1-dimensional Lyapunov Characteristic Exponent` *or the* `Lyapunov Characteristic Exponent of order 1 (1-LCE)` *of $\boldsymbol{A}_t$ with respect to vector $\mathbf{w}$.*

For simplicity we will usually refer to 1–LCEs as LCEs.

We note that the value of the norm $\|\mathbf{w}\|$ does not influence the value of $\chi(\mathbf{A}_t, \mathbf{w})$. For example, considering a vector $\beta\mathbf{w}$, with $\beta \in \mathbb{R}$ a nonzero constant, instead of $\mathbf{w}$ in Definition 4, we get the extra term $\ln|\beta|/t$ (with $|\ |$ denoting the absolute value) in equation (26) whose limiting value for $t \to \infty$ is zero and thus does not change the value of $\chi(\mathbf{A}_t, \mathbf{w})$. More importantly, the value of the LCE is independent of the norm appearing in equation (26). This can be easily seen as follows: Let us consider a second norm $\|\ \|'$ satisfying the inequality

$$\beta_1\|\mathbf{w}\| \leq \|\mathbf{w}\|' \leq \beta_2\|\mathbf{w}\|$$

for some positive real numbers β_1, β_2, and for all vectors $\mathbf{w}$. Such norms are called *equivalent* (see e.g. [73, §5.4.7]). Then, by the above–mentioned argument it is easily seen that the use of norm $\|\ \|'$ in (26) leaves unchanged the value of $\chi(\mathbf{A}_t, \mathbf{w})$. Since all norms of finite dimensional vector spaces are equivalent, we conclude that the LCEs do not depend on the chosen norm.

Let $\mathbf{w}_i$, $i = 1, 2, \ldots, p$ be a set of linearly independent vectors in $\mathbb{R}^n$, E^p be the subspace generated by all $\mathbf{w}_i$ and $\mathrm{vol}_p(\mathbf{A}_t, E^p)$ be the volume of the p–parallelogram having as edges the p vectors $\mathbf{A}_t\mathbf{w}_i$. This volume is computed as the norm of the wedge product of these vectors (see Appendix A for the definition of the wedge product and the actual evaluation of the volume)

$$\mathrm{vol}_p(\mathbf{A}_t, E^p) = \|\mathbf{A}_t\mathbf{w}_1 \wedge \mathbf{A}_t\mathbf{w}_2 \wedge \cdots \wedge \mathbf{A}_t\mathbf{w}_p\|.$$

Let also $\mathrm{vol}_p(\mathbf{A}_0, E^p)$ be the volume of the initial p–parallelogram defined by all $\mathbf{w}_i$, since $\mathbf{A}_0$ is the identity matrix. Then the quantity

$$\lambda_t(E^p) = \frac{\mathrm{vol}_p(\mathbf{A}_t, E^p)}{\mathrm{vol}_p(\mathbf{A}_0, E^p)}$$

is called the *coefficient of expansion in the direction of E^p* and it depends only on E^p and not on the choice of the linearly independent set of vectors. Obviously for an 1–dimensional subspace E^1 the coefficient of expansion is $\|\mathbf{A}_t\mathbf{w}_1\|/\|\mathbf{w}_1\|$. If the limit

$$\lim_{t\to\infty} \frac{1}{t} \ln \lambda_t(E^p)$$

exits it is called the *characteristic exponent of order p in the direction of E^p*.

Definition 5. *Considering the linearly independent set $\mathbf{w}_i$, $i = 1, 2, \ldots, p$ and the corresponding subspace E^p of $\mathbb{R}^n$ as above, the* `p-dimensional Lyapunov` `Characteristic Exponent` *or the* `Lyapunov Characteristic Exponent of` `order p (p-LCE)` *of $\mathbf{A}_t$ with respect to subspace E^p is defined as*

$$\chi(\boldsymbol{A}_t, E^p) = \limsup_{t\to\infty} \frac{1}{t} \ln \mathrm{vol}_p(\boldsymbol{A}_t, E^p). \tag{27}$$

Similarly to the case of the 1–LCE, the value of the initial volume $\mathrm{vol}_p(\mathbf{A}_0, E^p)$, as well as the used norm, do not influence the value of $\chi(\mathbf{A}_t, E^p)$.

From (25) and the Hadamard inequality (see e. g. [102]), according to which the Euclidean volume of a p–parallelogram does not exceed the product of the lengths of its sides, we conclude that the LCEs of equations (26) and (27) are finite.

From the definition of the LCE it follows that

$$\chi(\mathbf{A}_t, c_1\mathbf{w}_1 + c_2\mathbf{w}_2) \leq \max\left\{\chi(\mathbf{A}_t, \mathbf{w}_1), \chi(\mathbf{A}_t, \mathbf{w}_2)\right\}$$

for any two vectors $\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{R}^n$ and $c_1, c_2 \in \mathbb{R}$ with $c_1, c_2 \neq 0$, while the Hadamard inequality implies that if $\mathbf{w}_i$, $i = 1, 2, \ldots, n$ is a basis of $\mathbb{R}^n$ then

$$\sum_{i=1}^{n} \chi(\mathbf{A}_t, \mathbf{w}_i) \geq \limsup_{t\to\infty} \frac{1}{t} \ln |\det \mathbf{A}_t| \tag{28}$$

where $\det \mathbf{A}_t$ is the determinant of matrix $\mathbf{A}_t$.

It can be shown that for any $r \in \mathbb{R}$ the set of vectors $\{\mathbf{w} \in \mathbb{R}^n : \chi(\mathbf{A}_t, \mathbf{w}) \leq r\}$ is a vector subspace of $\mathbb{R}^n$ and that the function $\chi(\mathbf{A}_t, \mathbf{w})$ with $\mathbf{w} \in \mathbb{R}^n$, $\mathbf{w} \neq 0$ takes at most n different values, say

$$\nu_1 > \nu_2 > \cdots > \nu_s \text{ with } 1 \leq s \leq n. \tag{29}$$

For the subspaces
$$L_i = \{\mathbf{w} \in \mathbb{R}^n : \chi(\mathbf{A}_t, \mathbf{w}) \leq \nu_i\} \tag{30}$$

we have
$$\mathbb{R}^n = L_1 \supset L_2 \supset \cdots \supset L_s \supset L_{s+1} \stackrel{\text{def}}{=} \{0\}, \tag{31}$$

with $L_{i+1} \neq L_i$ and $\chi(\mathbf{A}_t, \mathbf{w}) = \nu_i$ if and only if $\mathbf{w} \in L_i \setminus L_{i+1}$ for $i = 1, 2, \ldots, s$. So in descending order each LCE 'lives' in a space of dimensionality less than that of the preceding exponent. Such a structure of linear spaces with decreasing dimension, each containing the following one, is called a *filtration*.

Definition 6. *A basis $\mathbf{w}_i$, $i = 1, 2, \ldots, n$ of $\mathbb{R}^n$ is called normal if $\sum_{i=1}^{n} \chi(\boldsymbol{A}_t, \mathbf{w}_i)$ attains a minimum at this basis. In other words, the basis $\mathbf{w}_i$, is a* `normal basis` *if*

$$\sum_{i=1}^{n} \chi(\boldsymbol{A}_t, \mathbf{w}_i) \leq \sum_{i=1}^{n} \chi(\boldsymbol{A}_t, \mathbf{g}_i)$$

where $\mathbf{g}_i$, $i = 1, 2, \ldots, n$ is any other basis of $\mathbb{R}^n$.

A normal basis $\mathbf{w}_i$, $i = 1, 2, \ldots, n$ is not unique but the numbers $\chi(\mathbf{A}_t, \mathbf{w}_i)$ depend only on $\mathbf{A}_t$ and not on the particular normal basis and are called the LCEs of function $\mathbf{A}_t$. By a possible permutation of the vectors of a given normal basis we can always assume that $\chi(\mathbf{A}_t, \mathbf{w}_1) \geq \chi(\mathbf{A}_t, \mathbf{w}_2) \geq \cdots \geq \chi(\mathbf{A}_t, \mathbf{w}_n)$.

Definition 7. *Let $\mathbf{w}_i$, $i = 1, 2, \ldots, n$ be a normal basis of $\mathbb{R}^n$ and $\chi_1 \geq \chi_2 \geq \cdots \geq \chi_n$, with $\chi_i \equiv \chi(\boldsymbol{A}_t, \mathbf{w}_i)$, $i = 1, 2, \ldots, n$, the LCEs of these vectors. Assume that value ν_i, $i = 1, 2, \ldots, s$ appears exactly $k_i = k_i(\nu_i) > 0$ times among these numbers. Then k_i is called the* `multiplicity` *of value ν_i and the collection (ν_i, k_i) $i = 1, 2, \ldots, s$ is called the* `spectrum of LCEs`*.*

In order to clarify the used notation we stress that χ_i, $i = 1, 2, \ldots, n$ are the n (possibly nondistinct) LCEs, satisfying $\chi_1 \geq \chi_2 \geq \cdots \geq \chi_n$, while ν_i, $i = 1, 2, \ldots, s$ represent the s $(1 \leq s \leq n)$, different values the LCEs have, with $\nu_1 > \nu_2 > \cdots > \nu_s$.

Definition 8. *The matrix function $\boldsymbol{A}_t$ is called* `regular` *as $t \to \infty$ if for each normal basis $\mathbf{w}_i$, $i = 1, 2, \ldots, n$ it holds that*

$$\sum_{i=1}^{n} \chi(\boldsymbol{A}_t, \mathbf{w}_i) = \liminf_{t \to \infty} \frac{1}{t} \ln |\det \boldsymbol{A}_t|,$$

which, due to (28) leads to

$$\liminf_{t \to \infty} \frac{1}{t} \ln |\det \boldsymbol{A}_t| = \limsup_{t \to \infty} \frac{1}{t} \ln |\det \boldsymbol{A}_t|$$

guaranteeing that the limit

$$\lim_{t \to \infty} \frac{1}{t} \ln |\det \boldsymbol{A}_t|$$

exists, is finite and equal to

$$\lim_{t \to \infty} \frac{1}{t} \ln |\det \boldsymbol{A}_t| = \sum_{i=1}^{n} \chi(\boldsymbol{A}_t, \mathbf{w}_i) = \sum_{i=1}^{s} k_i \nu_i.$$

We can now state a very important theorem for the LCEs:

Theorem 1. *If the matrix function $\boldsymbol{A}_t$ is regular then the LCEs of all orders are given by equations (26) and (27) where the $\limsup\limits_{t\to\infty}$ is substituted by $\lim\limits_{t\to\infty}$*

$$\chi(\boldsymbol{A}_t, \mathbf{w}) = \lim_{t\to\infty} \frac{1}{t} \ln \|\boldsymbol{A}_t \mathbf{w}\| \tag{32}$$

$$\chi(\boldsymbol{A}_t, E^p) = \lim_{t\to\infty} \frac{1}{t} \ln \mathrm{vol}_p(\boldsymbol{A}_t, E^p). \tag{33}$$

In particular, for any p–dimensional subspace $E^p \subseteq \mathbb{R}^n$ we have

$$\chi(\boldsymbol{A}_t, E^p) = \sum_{j=1}^{p} \chi_{i_j}. \tag{34}$$

with a suitable sequence $1 \le i_1 \le i_2 \le \cdots \le i_p \le n$.

The part of the theorem concerning equations (32) and (33) was proved by Oseledec in [102], while equation (34), although was not explicitly proved in [102], can be considered as a rather easily proven byproduct of the results presented there. Actually, the validity of equation (34) was shown in [13].

4.2 Computing LCEs of order 1

Let us now discuss how we can use Theorem 1 for the numerical computation of LCEs, starting with the computation of LCEs of order 1.

As we have already mentioned in (29), the LCE takes at most n different values ν_i , $i = 1, 2, \ldots, s$, $1 \le s \le n$. If we could know a priori the sequence (31) of subspaces L_i $i = 1, 2, \ldots, s$ of $\mathbb{R}^n$ we would, in principle, be able to compute the values ν_i of all LCEs. This could be done by taking an initial vector $\mathbf{w}_i \in L_i \setminus L_{i+1}$ and compute

$$\nu_i = \lim_{t\to\infty} \frac{1}{t} \ln \|\mathbf{A}_t \mathbf{w}_i\| \ , \quad i = 1, 2, \ldots, s. \tag{35}$$

Now apart from $L_1 = \mathbb{R}^n$ all the remaining subspaces L_i, $i = 2, 3, \ldots, s$ have positive codimension $\mathrm{codim}(L_i)$ $(= \dim \mathbb{R}^n - \dim L_i > 0)$ and thus, vanishing Lebesgue measure. Then a random choice of $\mathbf{w} \in \mathbb{R}^n$ would lead to the computation of χ_1 from (35), because, in principle $\mathbf{w}$ will belong to L_1 and not to the subspaces L_i $i = 2, \ldots, s$. Let us consider a simple example in order to clarify this statement.

Suppose that L_1 is the usual 3–dimensional space $\mathbb{R}^3$, $L_2 \subset L_1$ is a particular 2–dimensional plane of $\mathbb{R}^3$, e. g. the plane $z = 0$, $L_3 \subset L_2$ is a particular 1–dimensional line e. g. the x axis (Figure 4(a)) and the corresponding LCEs are $\chi_1 > \chi_2 > \chi_3$ with multiplicities $k_1 = k_2 = k_3 = 1$. For this case we have $\dim L_1 = 3$, $\dim L_2 = 2$, $\dim L_3 = 1$ and $\mathrm{codim}(L_1) = 0$, $\mathrm{codim}(L_2) = 1$, $\mathrm{codim}(L_3) = 2$. Concerning the measures μ of these subspaces of $\mathbb{R}^3$, it is
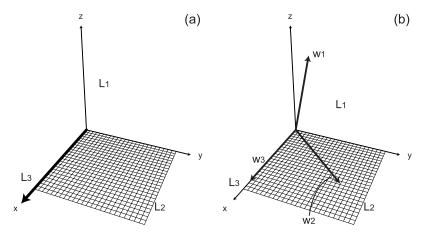
Fig. 4. (a) A schematic representation of the sequence of subspaces (31) where L_1 identifies with $\mathbb{R}^3$, $L_2 \subset L_1$ is represented by the xy plane and the x axis is considered as the final subspace $L_3 \subset L_2$. (b) A random choice of a vector in $L_1 \equiv \mathbb{R}^3$ will result with probability one to a vector belonging to L_1 and not to L_2, like vector $\mathbf{w}_1$. Vectors $\mathbf{w}_2$, $\mathbf{w}_3$ belonging respectively to $L_2 \setminus L_3$ and to L_3 are not random since their coordinates should satisfy certain conditions. In particular, the z coordinate of $\mathbf{w}_2$ should be zero, while both the z and y coordinate of $\mathbf{w}_3$ should vanish. The use of $\mathbf{w}_1$, $\mathbf{w}_2$, $\mathbf{w}_3$ in (35) leads to the computation of χ_1, χ_2 and χ_3 respectively.

obvious that $\mu(L_2) = \mu(L_3) = 0$, since the measure of a surface or of a line in the 3–dimensional space $\mathbb{R}^3$ is zero.

If we randomly choose a vector $\mathbf{w} \in \mathbb{R}^3$ it will belong to L_1 and not to L_2, i. e. having its z coordinate different from zero and thus, equation (35) would lead to the computation of the mLCE χ_1. Vector $\mathbf{w}_1$ in Figure 4(b) represents such a random choice. In order to compute χ_2 from (35) we should choose vector $\mathbf{w}$ not randomly but in a specific way. In particular, it should belong to L_2 but not to L_3, so its z coordinate should be equal to zero. Thus this vector should have the form $\mathbf{w} = (w_1, w_2, 0)$ with $w_1, w_2 \in \mathbb{R}$, $w_2 \neq 0$, like vector $\mathbf{w}_2$ in Figure 4(b). Our choice will become even more specific if we would like to compute χ_3 because in this case $\mathbf{w}$ should be of the form $\mathbf{w} = (w_1, 0, 0) \neq \mathbf{0}$ with $w_1 \in \mathbb{R}$. Vector $\mathbf{w}_3$ of Figure 4(b) is a choice of this kind.

From this example it becomes evident that a random choice of vector $\mathbf{w}$ in (35) will lead to the computation of the largest LCE χ_1 with probability one. One more comment concerning the numerical implementation of equation (35) should be added here. Even if in some special examples one could happen to know a priori the subspaces L_i $i = 1, 2, \ldots, s$, so that one could choose $\mathbf{w} \in L_i \setminus L_{i+1}$ with $i \neq 1$ then the computational errors would eventually lead to the numerical computation of χ_1. Such an example was presented in [14].

4.3 Computing LCEs of order $p > 1$

Let us now turn our attention to the computation of p–LCEs with $p > 1$. Equation (34) of Theorem 1 actually tells us that the p–LCE $\chi(\mathbf{A}_t, E^p)$ can take at most $\binom{n}{p}$ distinct values, i. e. as many as all the possible sums of p 1–LCEs out of n are. Now, as the choice of a random vector $\mathbf{w} \in \mathbb{R}^n$, or in other words, of a random 1–dimensional subspace of $\mathbb{R}^n$ produced by $\mathbf{w}$, leads to the computation of the maximal 1–LCE, the random choice of a p–dimensional subspace E^p of $\mathbb{R}^n$, or equivalently the random choice of p linearly independent vectors $\mathbf{w}_i$ $i = 1, 2, \ldots, p$, leads to the computation of the *maximal p–LCE (p–mLCE)* which is equal to the sum of the p largest 1–LCEs

$$\chi(\mathbf{A}_t, E^p) = \sum_{i=1}^{p} \chi_i. \tag{36}$$

This relation was formulated explicitly in [11, 9] and proved in [13] but was implicitly contained in [102]. The practical importance of equation (36) was also clearly explained in [119]. Benettin et al. [13] gave a more rigorous form to the notion of the random choice of E^p, which is essential for the derivation of (36), by introducing a condition that subspace E^p should satisfy. They named this condition *Condition R* (at random). According to Condition R a p–dimensional space $E^p \subset \mathbb{R}^n$ is chosen at random if for all $j = 2, 3, \ldots, s$ we have

$$\dim(E^p \cap L_j) = \max\left\{ 0, p - \sum_{i=1}^{j-1} k_i \right\} \tag{37}$$

where L_j belongs to the sequence of subspaces (31) and k_i is the multiplicity of the LCE ν_i (Definition 7).

In order to clarify these issues let us consider again the example presented in Figure 4, where we have three distinct values for the 1–LCEs $\chi_1 > \chi_2 > \chi_3$ with multiplicities $k_1 = k_2 = k_3 = 1$. In this case the 2–LCE can take one of the three possible values $\chi_1 + \chi_2$, $\chi_2 + \chi_3$, $\chi_1 + \chi_3$, while the 3–LCE takes only one possible value, namely $\chi_1 + \chi_2 + \chi_3$.

The computation of the 2–LCE requires the choice of two linearly independent vectors $\mathbf{w}_1$, $\mathbf{w}_2$ and the application of equation (33). The two vectors $\mathbf{w}_1$, $\mathbf{w}_2$ define a 2–dimensional plane E^2 in $\mathbb{R}^3$ and $\chi(\mathbf{A}_t, E^2)$ practically measures the time rate of the coefficient of expansion of the surface of the parallelogram having as edges the vectors $\mathbf{A}_t \mathbf{w}_1$, $\mathbf{A}_t \mathbf{w}_2$.

By choosing the two vectors $\mathbf{w}_1$, $\mathbf{w}_2$ randomly we define a random plane E^2 in $\mathbb{R}^3$ which intersects the subspace L_2 (plane xy) along a line, i. e. $\dim(E^2 \cap L_2) = 1$ and the subspace L_3 (x axis) at a point, i. e. $\dim(E^2 \cap L_3) = 0$ (Figure 5(a)). This random choice of plane E^2 satisfies Condition R (37) and thus, equation (33) leads to the computation of the 2–mLCE, namely $\chi_1 + \chi_2$. This result can be also understood in the following way. Plane E^2 in Figure 5(a) can be considered to be spanned by two vectors $\mathbf{w}_1$, $\mathbf{w}_2$ such that $\mathbf{w}_1 \in L_1$
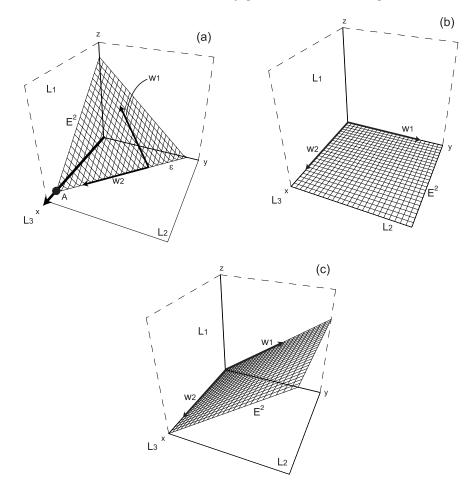
Fig. 5. Possible choices of the 2–dimensional space E^2 for the computation of the 2–LCE in the example of Figure 4, where $\mathbb{R}^3$ is considered as the tangent space of a hypothetical dynamical system. In each panel the chosen 'plane' E^2 is drawn, as well as one of its possible basis constituted of vectors $\mathbf{w}_1$, $\mathbf{w}_2$. (a) a random choice of E^2 leads to a plane intersecting L_2 along line ϵ ($\dim(E^2 \cap L_2) = 1$) and L_1 at point A ($\dim(E^2 \cap L_3) = 0$). In this case equation (33) gives $\chi(\mathbf{A}_t, E^2) = \chi_1 + \chi_2$. More carefully made choices of E^2 (which are obviously not made at random) results to configurations leading to the computation of $\chi_2 + \chi_3$ (b) and $\chi_1 + \chi_3$ (c) from equation (33). In these cases E^2 does not satisfy Condition R (37) since $\dim(E^2 \cap L_2) = 2$, $\dim(E^2 \cap L_3) = 1$ in (b) and $\dim(E^2 \cap L_2) = 1$, $\dim(E^2 \cap L_3) = 1$ in (c).

but not in its subspace L_2 and $\mathbf{w}_2 \in L_2$ but not in its subspace L_3. Then the expansion of $\mathbf{w}_1 \in L_1 \setminus L_2$ is determined by the LCE χ_1 and the expansion of $\mathbf{w}_2 \in L_2 \setminus L_3$ by the LCE χ_2. These 1–dimensional expansion rates result to an expansion rate equal to $\chi_1 + \chi_2$ for the surface defined by the two vectors.

Other more carefully designed choices of the E^2 subspace lead to the computation of the other possible values of the 2–LCE. If for example $\mathbf{w}_1 \in L_2 \setminus L_3$ and $\mathbf{w}_2 \in L_3$ (Figure 5(b)) we have $E^2 = L_2$ with $\dim(E^2 \cap L_2) = 2$ and $\dim(E^2 \cap L_3) = 1$. In this case the expansion of $\mathbf{w}_1$ is determined by the LCE χ_2 and of $\mathbf{w}_2$ by χ_3, and so the computed 2–LCE is $\chi_2 + \chi_3$. Finally, a choice of E^2 of the form presented in Figure 5(c) leads to the computation of $\chi_1 + \chi_3$. In this case the plane E^2 is defined by $\mathbf{w}_1 \in L_1 \setminus L_2$ and $\mathbf{w}_2 \in L_3$ and intersects subspaces L_2 and L_3 along the line corresponding to L_3, i. e. $\dim(E^2 \cap L_2) = 1$ and $\dim(E^2 \cap L_3) = 1$. It can be easily checked that for the last two choices of E^2 (Figures 5(b) and (c)), for which the computed 2–LCE does not take its maximal possible value, Condition R (37) is not satisfied, as one should have expected from the fact that these choices correspond to carefully designed configurations and not to a random process.

Similarly to the case of the computation of the 1–LCEs we note that, even if in some exceptional case one could know a priori the subspaces L_i $i = 1, 2, \ldots, s$, so that one could choose $\mathbf{w}_i$ $i = 1, 2, \ldots, p$ to span a particular subspace E^p in order to compute a specific value of the p–LCE, smaller than $\sum_{i=1}^{p} \chi_i$ (like in Figures 5(b) and (c)), the inevitable computational errors would eventually lead to the numerical computation of the maximal possible value of the p–LCE.

Summarizing we point out that the practical implementation of Theorem 1 guarantees that a random choice of p initial vectors $\mathbf{w}_i$ $i = 1, 2, \ldots, p$ with $1 \leq p \leq n$ generates a space E^p which satisfies Condition R (37) and leads to the actual computation of the corresponding p–mLCE, namely $\chi_1 + \chi_2 + \ldots + \chi_p$. This statement, which was originally presented in [11, 9], led to the standard algorithm for the computation of all LCEs presented in [14]. This algorithm is analyzed in Section 6.1.

4.4 The Multiplicative Ergodic Theorem

After presenting results concerning the existence and the computation of the LCEs of all orders for a general matrix function $\mathbf{A}_t$, let us restrict our study to the case of multiplicative cocycles $\mathbf{R}(t, \mathbf{x})$, which are matrix functions satisfying equation (23). The multiplicative cocycles arise naturally in discrete and continuous dynamical systems as was explained in the beginning of Section 4.

In particular, we consider the multiplicative cocycle $d_{\mathbf{x}} \Phi^t$ which maps the tangent space at $\mathbf{x} \in \mathcal{S}$ to the tangent space at $\Phi^t(\mathbf{x}) \in \mathcal{S}$ for a dynamical system defined on the differentiable manifold $\mathcal{S}$. We recall that $\mathcal{S}$ is a measure space with a normalized measure μ and that Φ^t is a diffeomorphism on $\mathcal{S}$, i. e. Φ^t is a measurable bijection of $\mathcal{S}$ which preserves the measure μ (24) and whose inverse is also measurable. We remark that in measure theory we

disregard sets of measure 0. In this sense Φ^t is called measurable if it becomes measurable upon disregarding from $\mathcal{S}$ a set of measure 0. Quite often we will us the expression *'for almost all* $\mathbf{x}$ *with respect to measure* μ*'* for the validity of a statement, implying that the statement is true for all points $\mathbf{x}$ with the possible exception of a set of points with measure 0.

A basic property of the multiplicative cocycles is their regularity, since Theorem 1 guarantees the existence of characteristic exponents and the finiteness of the LCEs of all orders for regular multiplicative cocycles. Thus, it is important to determine specific conditions that multiplicative cocycles should fulfill in order to be regular. Such conditions were first provided by Oseledec [102] who also formulated and proved the so–called *Multiplicative Ergodic Theorem* (MET), which is often referred as *Oseledec's theorem.*

The MET gives information about the dynamical structure of a multiplicative cocycle $\mathbf{R}(t, \mathbf{x})$ and its asymptotic behavior for $t \to \infty$. The application of the MET for the particular multiplicative cocycle $d_{\mathbf{x}}\Phi^t$ provides the theoretical framework for the computation of the LCEs for dynamical systems. The MET is one of the milestones in the study of ergodic properties of dynamical systems and it can be considered as a sort of a spectral theorem for random matrix products [113]. As a testimony to the importance of this theorem one can find several alternative proofs for it in the literature. The original proof of Oseledec [102] applies both to continuous and discrete systems. In view to the application to algebraic groups, Raghunathan [108] devised a simple proof of the MET, which nevertheless could not guarantee the finiteness of all LCEs. Although Raghunathan's results apply only to maps, an extension to flows, following the ideas of Oseledec, was given by Ruelle [114]. Benettin et al. [13] proved a somewhat different version of the theorem being mainly interested to its application on Hamiltonian flows and symplectic maps. Alternative proofs can also be found in [76, 141].

In [102] Oseledec proved that a multiplicative cocycle $\mathbf{R}(t, \mathbf{x})$ is regular and thus, the MET is applicable to it, if it satisfies the condition

$$\sup_{|t| \leq 1} \ln^+ \|\mathbf{R}^{\pm}(t, \mathbf{x})\| \in L^1(\mathcal{S}, \mu), \ ^3 \tag{38}$$

where $\ln^+ a = \max\{0, \ln a\}$. From (38) we obtain the estimate

$$\|\mathbf{R}(t, \mathbf{x})\| \leq e^{J(\mathbf{x})|t|} \tag{39}$$

for $t \to \pm\infty$ for almost all $\mathbf{x}$ with respect to μ, where $J(\mathbf{x})$ is a measurable function. From (39) it follows that $\mathbf{R}(t, \mathbf{x})$, considered as a function of t for

---

[3] We recall that a measurable function $f : \mathcal{S} \to \mathbb{R}$ (or $\mathbb{C}$) of the measure space $(\mathcal{S}, \mu)$ belongs to the space $L^1(\mathcal{S}, \mu)$ if its absolute value has a finite Lebesgue integral, i. e.

$$\int |f| d\mu < \infty \ .$$

fixed $\mathbf{x}$, satisfies equation (25). Benettin et al. [13] considered a slightly different version of the MET with respect to the one presented in [102]. Their version was adapted to the framework of a continuous or discrete dynamical system with Φ^t being a diffeomorphism of class C^1, i. e. both Φ^t and its inverse are continuously differentiable. They formulated the MET for the particular multiplicative cocycle $d_{\mathbf{x}}\Phi^t$, which they proved to be regular. Since our presentation is mainly focused on autonomous Hamiltonian systems and symplectic maps we will also state the MET for the specific cocycle $d_{\mathbf{x}}\Phi^t$. The version of the MET we present is mainly based on [102, 114, 13] and combines different formulations of the theorem given by various authors over the years.

Theorem 2 (Multiplicative Ergodic Theorem – MET). *Consider a dynamical system as follows: Let its phase space $\mathcal{S}$ be an n–dimensional compact manifold with a normalized measure μ, $\mu(\mathcal{S}) = 1$ and a smooth Riemannian metric $\| \, \|$. Consider also a measure–preserving diffeomorphism Φ^t of class C^1 satisfying*

$$\Phi^{t+s} = \Phi^t \circ \Phi^s \ ,$$

with t denoting time and having real (continuous system) or integer (discrete system) values. Then for almost all $\mathbf{x} \in \mathcal{S}$, with respect to measure μ we have:

1. *The family of multiplicative cocycles $d_{\mathbf{x}}\Phi^t : \mathcal{T}_{\mathbf{x}}\mathcal{S} \to \mathcal{T}_{\Phi^t(\mathbf{x})}\mathcal{S}$, where $\mathcal{T}_{\mathbf{x}}\mathcal{S}$ denotes the tangent space of $\mathcal{S}$ at point $\mathbf{x}$, is regular.*
2. *The LCEs of all orders exist and are independent of the choice of the Riemannian metric of $\mathcal{S}$.*
   In particular, for any $\mathbf{w} \in \mathcal{T}_{\mathbf{x}}\mathcal{S}$ the finite limit

$$\chi(\mathbf{x}, \mathbf{w}) = \lim_{t \to \infty} \frac{1}{t} \ln \|d_{\mathbf{x}}\Phi^t \mathbf{w}\| \tag{40}$$

exists and defines the LCE of order 1 (1–LCE). There exists at least one normal basis $\mathbf{v}_i$, $i = 1, 2, \ldots, n$ of $\mathcal{T}_{\mathbf{x}}\mathcal{S}$ for which the corresponding (possibly nondistinct) 1–LCEs $\chi_i(\mathbf{x}) = \chi(\mathbf{x}, \mathbf{v}_i)$ are ordered as

$$\chi_1(\mathbf{x}) \geq \chi_2(\mathbf{x}) \geq \cdots \geq \chi_n(\mathbf{x}). \tag{41}$$

Assume that the value $\nu_i(\mathbf{x})$, $i = 1, 2, \ldots, s$ with $s = s(\mathbf{x})$, $1 \leq s \leq n$ appears exactly $k_i(\mathbf{x}) = k_i(\mathbf{x}, \nu_i) > 0$ times among these numbers. Then the spectrum of LCEs $(\nu_i(\mathbf{x}), k_i(\mathbf{x}))$, $i = 1, 2, \ldots, s$ is a measurable function of $\mathbf{x}$, and as $\mathbf{w} \neq 0$ varies in $\mathcal{T}_{\mathbf{x}}\mathcal{S}$, $\chi(\mathbf{x}, \mathbf{w})$ takes one of these s different values

$$\nu_1(\mathbf{x}) > \nu_2(\mathbf{x}) > \cdots > \nu_s(\mathbf{x}). \tag{42}$$

It also holds

$$\sum_{i=1}^{s} k_i(\mathbf{x})\nu_i(\mathbf{x}) = \lim_{t \to \infty} \frac{1}{t} \ln |\det d_{\mathbf{x}}\Phi^t|. \tag{43}$$

For any p-dimensional ($1 \leq p \leq n$) subspace $E^p \subseteq \mathcal{T}_{\mathbf{x}}\mathcal{S}$, generated by a linearly independent set $\mathbf{w}_i$, $i = 1, 2, \ldots, p$ the finite limit

$$\chi(\mathbf{x}, E^p) = \lim_{t \to \infty} \frac{1}{t} \ln \mathrm{vol}_p(d_{\mathbf{x}} \Phi^t, E^p), \tag{44}$$

where $\mathrm{vol}_p(d_{\mathbf{x}} \Phi^t, E^p)$ *is the volume of the p–parallelogram having as edges the vectors* $d_{\mathbf{x}} \Phi^t \mathbf{w}_i$*, exists and defines the LCE of order p (p–LCE). The value of* $\chi(\mathbf{x}, E^p)$ *is equal to the sum of p 1–LCEs* $\chi_i(\mathbf{x})$*,* $i = 1, 2, \ldots, n$*.*

3. The set of vectors

$$L_i(\mathbf{x}) = \{\mathbf{w} \in \mathcal{T}_{\mathbf{x}}\mathcal{S} : \chi(\mathbf{x}, \mathbf{w}) \le \nu_i(\mathbf{x})\} \ , \ \ 1 \le i \le s$$

is a linear subspace of $\mathcal{T}_{\mathbf{x}}\mathcal{S}$ *satisfying*

$$\mathcal{T}_{\mathbf{x}}\mathcal{S} = L_1(\mathbf{x}) \supset L_2(\mathbf{x}) \supset \cdots \supset L_s(\mathbf{x}) \supset L_{s+1}(\mathbf{x}) \stackrel{def}{=} \{0\}. \tag{45}$$

If $\mathbf{w} \in L_i(\mathbf{x}) \setminus L_{i+1}(\mathbf{x})$ *then* $\chi(\mathbf{x}, \mathbf{w}) = \nu_i(\mathbf{x})$ *for* $i = 1, 2, \ldots, s$*. The multiplicity* $k_i(\mathbf{x})$ *of values* $\nu_i(\mathbf{x})$ *is given by* $k_i(\mathbf{x}) = \dim L_i(\mathbf{x}) - \dim L_{i+1}(\mathbf{x})$*.*

4. The symmetric positive–defined matrix

$$\Lambda_{\mathbf{x}} = \lim_{t \to \infty} \left(\mathbf{Y}^{\mathrm{T}}(t) \cdot \mathbf{Y}(t) \right)^{1/2t}$$

exists. $\mathbf{Y}(t)$ *is the matrix corresponding to* $d_{\mathbf{x}} \Phi^t$ *and is defined by equations (10) and (13) for continuous and discrete dynamical systems respectively. The logarithms of the eigenvalues of* $\Lambda_{\mathbf{x}}$ *are the s distinct 1–LCEs (42) of the dynamical system. The corresponding eigenvectors are orthogonal (since* $\Lambda_{\mathbf{x}}$ *is symmetric), and for the corresponding eigenspaces* $V_1(\mathbf{x}), V_2(\mathbf{x}), \ldots, V_s(\mathbf{x})$ *we have*

$$k_i(\mathbf{x}) = \dim V_i(\mathbf{x}) \ , \ \ L_i(\mathbf{x}) = \bigoplus_{r=i}^{s} V_r(\mathbf{x}) \ \text{ for } \ i = 1, 2, \ldots, s.$$

Thus, $\mathcal{T}_{\mathbf{x}}\mathcal{S}$ *is decomposed as*

$$\mathcal{T}_{\mathbf{x}}\mathcal{S} = V_1(\mathbf{x}) \oplus V_2(\mathbf{x}) \oplus \cdots \oplus V_s(\mathbf{x}),$$

and for every nonzero vector $\mathbf{w} \in V_i(\mathbf{x})$*,* $i = 1, 2, \ldots, s$*, we get*

$$\chi(\mathbf{x}, \mathbf{w}) = \nu_i(\mathbf{x}).$$

A short remark is necessary here. The regularity of $d_{\mathbf{x}} \Phi^t$, which guarantees the validity of equations (40) and (44) and the finiteness of the LCEs of all orders, should not be confused with the regular nature of orbits of the dynamical system. Regular orbits have all their LCEs equal to zero (see also the discussion in Section 5.3).

Benettin et al. [11, 13] have formulated also the following theorem which provides the theoretical background for the numerical algorithm they presented in [14] for the computation of all LCEs.

Theorem 3. *Under the assumptions of the MET, the p–LCE of any p-dimensional subspace $E^p \subseteq \mathcal{T}_\mathbf{x}\mathcal{S}$ satisfying Condition R (37), is equal to the sum of the p largest 1–LCEs (41):*

$$\chi(\mathbf{x}, E^p) = \lim_{t \to \infty} \frac{1}{t} \ln \mathrm{vol}_p(d_\mathbf{x}\Phi^t, E^p) = \sum_{i=1}^{p} \chi_i(\mathbf{x}). \tag{46}$$

4.5 Properties of the spectrum of LCEs

Let us now turn our attention to the structure of the spectrum of LCEs for ND autonomous Hamiltonian systems and for $2N$d symplectic maps, which are the main dynamical systems we are interested in. Such systems preserve the phase-space volume, and thus, the r. h. s. of (43) vanishes. So for the sum of all the 1–LCEs we have

$$\sum_{i=1}^{2N} \chi_i(\mathbf{x}) = 0. \tag{47}$$

The symplectic nature of these systems gives indeed more. It has been proved in [13] that the spectrum of LCEs consists of pairs of values having opposite signs

$$\chi_i(\mathbf{x}) = -\chi_{2N-i+1}(\mathbf{x}) \ , \ \ i = 1, 2, \ldots, N. \tag{48}$$

Thus, the spectrum of LCEs becomes

$$\chi_1(\mathbf{x}) \geq \chi_2(\mathbf{x}) \geq \cdots \geq \chi_N(\mathbf{x}) \geq -\chi_N(\mathbf{x}) \geq \cdots \geq -\chi_2(\mathbf{x}) \geq -\chi_1(\mathbf{x}).$$

For autonomous Hamiltonian flows we can say something more. Let us first recall that for a general differentiable flow on a compact manifold without stationary points at least one LCE must vanish [13, 70]. This follows from the fact that, in the direction along the flow a deviation vector grows only linearly in time. So, in the case of a Hamiltonian flow, due to the symmetry of the spectrum of LCEs (48), at least two LCEs vanish, i. e.

$$\chi_N(\mathbf{x}) = \chi_{N+1}(\mathbf{x}) = 0,$$

while the presence of any additional independent integral of motion leads to the vanishing of another pair of LCEs.

Let us now study the particular case of a periodic orbit of period T, such that $\Phi^T(\mathbf{x}) = \mathbf{x}$, following [9, 12]. In this case $d_\mathbf{x}\Phi^T$ is a linear operator on the tangent space $\mathcal{T}_\mathbf{x}\mathcal{S}$ so that for any deviation vector $\mathbf{w}(0) \in \mathcal{T}_\mathbf{x}\mathcal{S}$ we have

$$\mathbf{w}(T) = \mathbf{Y} \cdot \mathbf{w}(0), \tag{49}$$

where $\mathbf{Y}$ is the constant matrix corresponding to $d_\mathbf{x}\Phi^T$. Suppose that $\mathbf{Y}$ has $2N$ (possibly complex) eigenvalues λ_i, $i = 1, 2, \ldots, 2N$, whose magnitudes can be ordered as

$$|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_{2N}|.$$

Let $\hat{\mathbf{w}}_i$, $i = 1, 2, \ldots, 2N$, denote the corresponding unitary eigenvectors. Then for $\mathbf{w}(0) = \hat{\mathbf{w}}_i$ equation (49) implies

$$\mathbf{w}(kT) = \lambda_i^k \hat{\mathbf{w}}_i \ , \ \ k = 1, 2, \ldots \tag{50}$$

and so we conclude from (40) that

$$\chi(\mathbf{x}, \hat{\mathbf{w}}_i) = \frac{1}{T} \ln |\lambda_i| = \chi_i(\mathbf{x}), \ \ i = 1, 2, \ldots, 2N.$$

Furthermore for a deviation vector

$$\mathbf{w}(0) = c_1 \hat{\mathbf{w}}_1 + c_2 \hat{\mathbf{w}}_2 + \ldots + c_{2N} \hat{\mathbf{w}}_{2N}$$

with $c_i \in \mathbb{R}$, $i = 1, 2, \ldots, 2N$, it follows from (50) that the first nonvanishing coefficient c_i eventually dominates the evolution of $\mathbf{w}(t)$ and we get $\chi(\mathbf{x}, \mathbf{w}) = \chi_i$. In this case we can define a filtration similar to the one presented in (45) by defining $L_1 = [\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_2, \ldots, \hat{\mathbf{w}}_{2N}] = \mathcal{T}_{\mathbf{x}}\mathcal{S}$, $L_2 = [\hat{\mathbf{w}}_2, \ldots, \hat{\mathbf{w}}_{2N}]$, $\ldots$, $L_{2N} = [\hat{\mathbf{w}}_{2N}]$, $L_{2N+1} = [0]$, where $[\ ]$ denotes the linear space spanned by vectors $\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_2, \ldots, \hat{\mathbf{w}}_{2N}$ and so on. It becomes evident that a random choice of an initial deviation vector $\mathbf{w}(0) \in \mathcal{T}_{\mathbf{x}}\mathcal{S}$ will lead to the computation of the mLCE $\chi_1(\mathbf{x})$ since, in general, $\mathbf{w}(0) \in L_1 \setminus L_2$.

So, in the case of an unstable periodic orbit where $|\lambda_1| > 1$ we get $\chi_1(\mathbf{x}) > 0$, which implies that nearby orbits diverge exponentially from the periodic one. These orbits are not called chaotic, although their mLCE is larger than zero, but simply 'unstable'. In fact, unstable periodic orbits exist also in integrable systems. Since the measure of periodic orbits in a general dynamical system has zero measure, periodic orbits (stable and unstable) are rather exceptional.

In the general case of a nonperiodic orbit we are no more allowed to use concepts as eigenvectors and eigenvalues because the linear operator $d_{\mathbf{x}}\Phi^t$ maps $\mathcal{T}_{\mathbf{x}}\mathcal{S}$ into $\mathcal{T}_{\Phi^t(\mathbf{x})}\mathcal{S} \neq \mathcal{T}_{\mathbf{x}}\mathcal{S}$, while eigenvectors are intrinsically defined only for linear operators of a linear space into itself. Nevertheless, in the case of nonperiodic orbits the MET proves the existence of the LCEs and of filtration (45). In a way, the MET provides an extention of the linear stability analysis of periodic orbits to the case of nonperiodic ones, although one should always keep in mind that the LCEs are related to the real and positive eigenvalues of the symmetric, positive–defined matrix $\mathbf{Y}^{\mathrm{T}}(t) \cdot \mathbf{Y}(t)$ [63, 98]. On the other hand, linear stability analysis involves the computation of the eigenvalues of the nonsymmetric matrix $\mathbf{Y}(t)$, which solves the linearized equations of motion (10) for Hamiltonian flows or (13) for maps. These eigenvalues are real or come in pairs of complex conjugate pairs and, in general, they are not directly related to the LCEs which are real numbers.

An important property of the LCEs is that they are constant in a connected chaotic domain. This is due to the fact that every nonperiodic orbit in the

same connected chaotic domain covers densely this domain, thus, two different orbits of the same domain are in a sense dynamically equivalent. The unstable periodic orbits in this chaotic domain have in general LCEs that are different from the constant LCEs of the nonperiodic orbits. This is due to the fact that the periodic orbits do not visit the whole domain, thus, they cannot characterize its dynamical behavior. In fact, different periodic orbits have different LCEs.

5 The maximal LCE

From this point on, in order to simplify our notation, we will not explicitly write the dependence of the LCEs on the specific point $\mathbf{x} \in \mathcal{S}$. So, in practice, considering that we are referring to a specific point $\mathbf{x} \in \mathcal{S}$, we denote by χ_i the LCEs of order 1 and by $\chi_i^{(p)}$ the LCEs of order p.

For the practical determination of the chaotic nature of orbits a numerical computation of the mLCE χ_1 can be employed. If the studied orbit is regular $\chi_1 = 0$, while if it is chaotic $\chi_1 > 0$, implying exponential divergence of nearby orbits. The computation of the mLCE has been used extensively as a chaos indicator after the introduction of numerical algorithms for the determination of its value at late 70's [10, 99, 8, 34, 14].

Apart from using the mLCE as a criterion for the chaoticity or the regularity of an orbit its value also attains a 'physical' meaning and defines a specific time scale for the considered dynamical system. In particular, the inverse of the mLCE, which is called *Lyapunov time*

$$t_L = \frac{1}{\chi_1} \tag{51}$$

gives an estimate of the time needed for a dynamical system to become chaotic and in practice measures the time needed for nearby orbits of the system to diverge by e (see e. g [30, p. 508]).

5.1 Computation of the mLCE

The mLCE can be computed by the numerical implementation of equation (40). In Section 4.2 we showed that a random choice of the initial deviation vector $\mathbf{w}(0) \in \mathcal{T}_\mathbf{x}\mathcal{S}$ leads to the numerical computation of the mLCE. We recall that the deviation vector $\mathbf{w}(t)$ at time $t > 0$ is determined by the action of the operator $d_\mathbf{x}\Phi^t$ on the initial deviation vector $\mathbf{w}(0)$ according to equation (7)

$$\mathbf{w}(t) = d_\mathbf{x}\Phi^t \, \mathbf{w}(0). \tag{52}$$

This equation represents the solution of the variational equations (8) or the evolution of a deviation vector under the action of the tangent map (11), and takes the form (9) and (12) respectively. We emphasize that, both the

variational equations and the equations of the tangent map are linear with respect to the tangent vector $\mathbf{w}$, i. e.

$$d_{\mathbf{x}}\Phi^t\left(a\,\mathbf{w}\right) = a\,d_{\mathbf{x}}\Phi^t\mathbf{w}, \;\; \text{for any } a \in \mathbb{R}. \tag{53}$$

In order to evaluate the mLCE of an orbit with initial condition $\mathbf{x}(0)$, one has to follow simultaneously the time evolution of the orbit itself and of a deviation vector $\mathbf{w}$ from this orbit with initial condition $\mathbf{w}(0)$. In the case of a Hamiltonian flow (continuous time) we solve simultaneously the Hamilton equations of motion (2) for the time evolution of the orbit and the variational equations (8) for the time evolution of the deviation vector. In the case of a symplectic map (discrete time) we iterate the map (4) for the evolution of the orbit simultaneously with the tangent map (11), which determines the evolution of the tangent vector. The mLCE is then computed as the limit for $t \to \infty$ of the quantity

$$X_1(t) = \frac{1}{t}\ln\frac{\|d_{\mathbf{x}(0)}\Phi^t\,\mathbf{w}(0)\|}{\|\mathbf{w}(0)\|} = \frac{1}{t}\ln\frac{\|\mathbf{w}(t)\|}{\|\mathbf{w}(0)\|}, \tag{54}$$

often called *finite time mLCE*. So, we have

$$\chi_1 = \lim_{t\to\infty} X_1(t). \tag{55}$$

The direct numerical implementation of equations (54) and (55) for the evaluation of χ_1 meets a severe difficulty. If, for example, the orbit under study is chaotic, the norm $\|\mathbf{w}(t)\|$ increases exponentially with increasing time t, leading to numerical overflow, i. e. $\|\mathbf{w}(t)\|$ attains very fast extremely large values that cannot be represented in the computer. This difficulty can be overcome by a procedure which takes advantage of the linearity of $d_{\mathbf{x}}\Phi^t$ (53) and of the composition law (22). Fixing a small time interval τ we express time t with respect to τ as $t = k\tau$, $k = 1, 2, \ldots$. Then for the quantity $X_1(t)$ we have

$$\begin{aligned}
X_1(k\tau) &= \frac{1}{k\tau}\ln\frac{\|\mathbf{w}(k\tau)\|}{\|\mathbf{w}(0)\|}\\
&= \frac{1}{k\tau}\ln\left(\frac{\|\mathbf{w}(k\tau)\|}{\|\mathbf{w}((k-1)\tau)\|}\frac{\|\mathbf{w}((k-1)\tau)\|}{\|\mathbf{w}((k-2)\tau)\|}\cdots\frac{\|\mathbf{w}(2\tau)\|}{\|\mathbf{w}(\tau)\|}\frac{\|\mathbf{w}(\tau)\|}{\|\mathbf{w}(0)\|}\right)\\
&= \frac{1}{k\tau}\sum_{i=1}^{k}\ln\frac{\|\mathbf{w}(i\tau)\|}{\|\mathbf{w}((i-1)\tau)\|} \Rightarrow
\end{aligned}$$

$$X_1(k\tau) = \frac{1}{k\tau}\sum_{i=1}^{k}\ln\frac{\|d_{\mathbf{x}(0)}\Phi^{i\tau}\,\mathbf{w}(0)\|}{\|d_{\mathbf{x}(0)}\Phi^{(i-1)\tau}\,\mathbf{w}(0)\|}. \tag{56}$$

Denoting by D_0 the norm of the initial deviation vector $\mathbf{w}(0)$

$$D_0 = \|\mathbf{w}(0)\|,$$

we get for the evolved deviation vector at time $t = k\tau$

$$
\begin{aligned}
d_{\mathbf{x}(0)}\Phi^{i\tau}\,\mathbf{w}(0) &= d_{\mathbf{x}(0)}\Phi^{\tau+(i-1)\tau}\,\mathbf{w}(0) \overset{(22)}{=} d_{\Phi^{(i-1)\tau}(\mathbf{x}(0))}\Phi^{\tau}\Big(d_{\mathbf{x}(0)}\Phi^{(i-1)\tau}\mathbf{w}(0)\Big) \\
&\overset{(53)}{=} \frac{\|d_{\mathbf{x}(0)}\Phi^{(i-1)\tau}\,\mathbf{w}(0)\|}{D_0}\; d_{\Phi^{(i-1)\tau}(\mathbf{x}(0))}\Phi^{\tau}\left(\frac{d_{\mathbf{x}(0)}\Phi^{(i-1)\tau}\mathbf{w}(0)}{\|d_{\mathbf{x}(0)}\Phi^{(i-1)\tau}\,\mathbf{w}(0)\|}D_0\right) \Rightarrow
\end{aligned}
$$

$$
\frac{d_{\mathbf{x}(0)}\Phi^{i\tau}\,\mathbf{w}(0)}{\|d_{\mathbf{x}(0)}\Phi^{(i-1)\tau}\,\mathbf{w}(0)\|} = \frac{d_{\Phi^{(i-1)\tau}(\mathbf{x}(0))}\Phi^{\tau}\left(\frac{d_{\mathbf{x}(0)}\Phi^{(i-1)\tau}\mathbf{w}(0)}{\|d_{\mathbf{x}(0)}\Phi^{(i-1)\tau}\,\mathbf{w}(0)\|}D_0\right)}{D_0}. \tag{57}
$$

Let us now denote by

$$
\hat{\mathbf{w}}((i-1)\tau) = \frac{d_{\mathbf{x}(0)}\Phi^{(i-1)\tau}\mathbf{w}(0)}{\|d_{\mathbf{x}(0)}\Phi^{(i-1)\tau}\,\mathbf{w}(0)\|}D_0,
$$

the deviation vector at point $\Phi^{(i-1)\tau}(\mathbf{x}(0))$ having the same direction with $\mathbf{w}((i-1)\tau)$ and norm D_0, and by D_i its norm after its evolution for τ time units

$$
D_i = \|d_{\Phi^{(i-1)\tau}(\mathbf{x}(0))}\Phi^{\tau}\hat{\mathbf{w}}((i-1)\tau)\|.
$$

Using this notation we derive from equation (57)

$$
\ln\frac{\|d_{\mathbf{x}(0)}\Phi^{i\tau}\,\mathbf{w}(0)\|}{\|d_{\mathbf{x}(0)}\Phi^{(i-1)\tau}\,\mathbf{w}(0)\|} = \ln\frac{D_i}{D_0} = \ln\alpha_i, \tag{58}
$$

with α_i being the local coefficient of expansion of the deviation vector for a time interval of length τ when the corresponding orbit evolves from position $\Phi^{(i-1)\tau}(\mathbf{x}(0))$ to position $\Phi^{i\tau}(\mathbf{x}(0))$ ($\ln\alpha_i/\tau$ is also called *stretching number* [135][30, p. 257]).

From equations (55), (56) and (58) we conclude that the mLCE χ_1 can be computed as

$$
\chi_1 = \lim_{k\to\infty} X_1(k\tau) = \lim_{k\to\infty}\frac{1}{k\tau}\sum_{i=1}^{k}\ln\frac{D_i}{D_0} = \lim_{k\to\infty}\frac{1}{k\tau}\sum_{i=1}^{k}\ln\alpha_i. \tag{59}
$$

Since the initial norm D_0 can have any arbitrary value, one usually set it to $D_0 = 1$. Equation (59) implies that practically χ_1 is the limit value, for $t \to \infty$, of the mean of the stretching numbers along the studied orbit [14, 57, 135].

5.2 The numerical algorithm

In practice, for the evaluation of the mLCE we follow the evolution of a unitary initial deviation vector $\hat{\mathbf{w}}(0) = \mathbf{w}(0)$, $\|\mathbf{w}(0)\| = D_0 = 1$ and every $t = \tau$ time units we replace the evolved vector $\mathbf{w}(k\tau)$, $k = 1, 2, \ldots$, by vector $\hat{\mathbf{w}}(k\tau)$

having the same direction but norm equal to 1 ($\|\hat{\mathbf{w}}(k\tau)\| = 1$). Before each new renormalization the corresponding α_k is computed and χ_1 is estimated from equation (59).

More precisely at $t = \tau$ we have $\alpha_1 = \|\mathbf{w}(\tau)\|$. Then we define a unitary vector $\hat{\mathbf{w}}(\tau)$ by renormalizing $\mathbf{w}(\tau)$ and using it as an initial deviation vector we evolve it along the orbit from $\mathbf{x}(\tau)$ to $\mathbf{x}(2\tau)$ according to equation (52), having $\mathbf{w}(2\tau) = d_{\mathbf{x}(\tau)}\Phi^\tau\,\hat{\mathbf{w}}(\tau)$. Then we define $\alpha_2 = \|\mathbf{w}(2\tau)\|$ and we estimate χ_1 (see Figure 6). We iteratively apply the above described procedure until
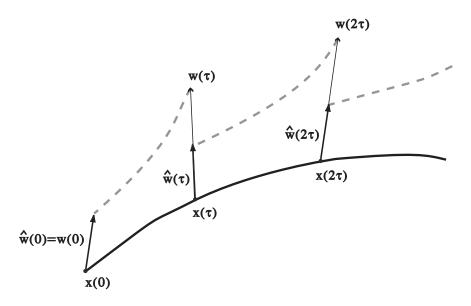

Fig. 6. Numerical scheme for the computation of the mLCE χ_1. The unitary deviation vector $\hat{\mathbf{w}}((i-1)\tau)$, $i = 1, 2, \ldots$, is evolved according to the variational equations (8) (continuous time) or the equations of the tangent map (11) (discrete time) for $t = \tau$ time units. The evolved vector $\mathbf{w}(i\tau)$ is replaced by a unitary vector $\hat{\mathbf{w}}(i\tau)$ having the same direction with $\mathbf{w}(i\tau)$. For each successive time interval $[(i-1)\tau, i\tau]$ the quantity $\alpha_i = \|\mathbf{w}(i\tau)\|$ is computed and χ_1 is estimated from equation (59).

a good approximation of χ_1 is achieved. The algorithm for the evaluation of the mLCE χ_1 is described in pseudo–code in Table 1.

Instead of utilizing the variational equations or the tangent map for the evolution of a deviation vector in the above described algorithm, one could integrate equations (2) or iterate equations (4) for two orbits starting nearby and estimate $\mathbf{w}(t)$ by difference. Indeed, this approach, influenced by the rough idea of divergence of nearby orbits introduced in [72], was initially adopted for the computation of the mLCE [10, 99, 8]. This technique was abandoned after a while as it was realized that the use of explicit equations for the evolution

Table 1. The algorithm for the computation of the mLCE χ_1 as the limit for $t \to \infty$ of $X_1(t)$ according to equation (59). The program computes the evolution of $X_1(t)$ as a function of time t up to a given upper value of time $t = T_M$ or until $X_1(t)$ attains a very small value, smaller than a low threshold value X_{1m}.

Input: 1. Hamilton equations of motion (2) and variational equations (8), or
           equations of the map (4) and of the tangent map (11).
     2. Initial condition for the orbit $\mathbf{x}(0)$.
     3. Initial *unitary* deviation vector $\mathbf{w}(0)$.
     4. Renormalization time τ.
     5. Maximal time: T_M and minimum allowed value of $X_1(t)$: X_{1m}.

Step 1 **Set** the stopping flag, $SF \leftarrow 0$, and the counter, $k \leftarrow 1$.
Step 2 **While** $(SF = 0)$ **Do**
       Evolve the orbit and the deviation vector from time $t = (k-1)\tau$
       to $t = k\tau$, i. e. **Compute** $\mathbf{x}(k\tau)$ and $\mathbf{w}(k\tau)$.
Step 3   **Compute** current value of $\alpha_k = \|\mathbf{w}(k\tau)\|$.
       Compute and **Store** current value of $X_1(k\tau) = \frac{1}{k\tau} \sum_{i=1}^{k} \ln \alpha_i$.
Step 4   Renormalize deviation vector by **Setting** $\mathbf{w}(k\tau) \leftarrow \mathbf{w}(k\tau)/\alpha_k$.
Step 5   **Set** the counter $k \leftarrow k+1$.
Step 6   **If** $[(k\tau > T_M)$ or $(X_1((k-1)\tau) < X_{1m})]$ **Then**
          Set $SF \leftarrow 1$.
       End If
    End While
Step 7 **Report** the time evolution of $X_1(t)$.

---

of deviation vectors was more reliable and efficient [34, 119, 14], although in some cases it is used also nowadays (see e. g. [145]).

5.3 Behavior of $X_1(t)$ for regular and chaotic orbits

Let us now discuss in more detail the behavior of the computational scheme for the evaluation of the mLCE for the cases of regular and chaotic orbits.

    The LCE of regular orbits vanish [10, 23] due to the linear increase with time of the norm of deviation vectors. We illustrate this behavior in the case of an ND Hamiltonian system, but a similar analysis can be easily carried out for symplectic maps. In such systems regular orbits lie on N–dimensional tori. If such tori are found around a stable periodic orbit, they can be accurately described by N formal integrals of motion in involution, so that the system would appear locally integrable. This means that we could perform a local transformation to action–angle variables, considering as actions $J_1, J_2, \ldots, J_N$ the values of the N formal integrals, so that Hamilton's equations of motion, locally attain the form

$$\dot{J}_i = 0, \ \dot{\theta}_i = \omega_i(J_1, J_2, \ldots, J_N), \ i = 1, 2, \ldots, N. \tag{60}$$

These equations can be easily integrated to give

$$J_i(t) = J_{i0}, \ \theta_i(t) = \theta_{i0} + \omega_i(J_{10}, J_{20}, \ldots, J_{N0})\, t, \ i = 1, 2, \ldots, N,$$

where J_{i0}, θ_{i0}, $i = 1, 2, \ldots, N$ are the initial conditions of the studied orbit.

By denoting as ξ_i, η_i, $i = 1, 2, \ldots, N$ small deviations of J_i and θ_i respectively, the variational equations (8) of system (60), describing the evolution of a deviation vector are

$$\dot{\xi}_i = 0, \ \ \dot{\eta}_i = \sum_{j=1}^{N} \omega_{ij} \cdot \xi_j, \ \ i = 1, 2, \ldots, N,$$

where

$$\omega_{ij} = \left. \frac{\partial \omega_i}{\partial J_j} \right|_{\mathbf{J}_0}, \ \ i, j = 1, 2, \ldots, N,$$

and $\mathbf{J}_0 = (J_{10}, J_{20}, \ldots, J_{N0}) = \text{constant}$, represents the N–dimensional vector of the initial actions. The solution of these equations is:

$$\begin{aligned} \xi_i(t) &= \xi_i(0) \\ \eta_i(t) &= \eta_i(0) + \left[\sum_{j=1}^{N} \omega_{ij} \xi_j(0) \right] t, \end{aligned} \quad i = 1, 2, \ldots, N. \tag{61}$$

From equations (61) we see that an initial deviation vector $\mathbf{w}(0)$ with coordinates $\xi_i(0)$, $i = 1, 2, \ldots, N$ in the action variables and $\eta_i(0)$, $i = 1, 2, \ldots, N$ in the angles, i. e. $\mathbf{w}(0) = (\xi_1(0), \xi_2(0), \ldots, \xi_N(0), \eta_1(0), \eta_2(0), \ldots, \eta_N(0))$, evolves in time in such a way that its action coordinates remain constant, while its angle coordinates increase linearly in time. This behavior implies an almost linear increase of the norm of the deviation vector. To see this, let us assume that vector $\mathbf{w}(0)$ has initially unit magnitude, i. e.

$$\sum_{i=1}^{N} \xi_i^2(0) + \sum_{i=1}^{N} \eta_i^2(0) = 1$$

whence the time evolution of its norm is given by

$$\|\mathbf{w}(t)\| = \left\{ 1 + \left[\sum_{i=1}^{N} \left(\sum_{j=1}^{N} \omega_{ij} \xi_j(0) \right)^2 \right] t^2 + \left[2 \sum_{i=1}^{N} \left(\eta_i(0) \sum_{j=1}^{N} \omega_{ij} \xi_j(0) \right) \right] t \right\}^{1/2}.$$

This implies that the norm for long times grows linearly with t

$$\|\mathbf{w}(t)\| \propto t. \tag{62}$$

So, from equation (54) we see that for long times $X_1(t)$ is of the order $\mathcal{O}(\ln t / t)$, which means that $X_1(t)$ tends asymptotically to zero, as $t \to \infty$ like t^{-1}. This asymptotic behavior is evident in numerical computations of the mLCE of regular orbits, as we can see for example in the left panel of Figure 2.

The asymptotic behavior of $X_1(t)$ for regular orbits, described above, represents a particular case of a more general estimation presented in [63]. In particular, Goldhirsch et al. [63] showed that, in general, after some initial

transient time the value of the mLCE χ_1 is related to its finite time estimation by

$$X_1(t) = \chi_1 + \frac{b + z(t)}{t}, \tag{63}$$

where b is a constant and $z(t)$ is a 'noise' term of zero mean. According to their analysis, this approximate formula is valid both for regular and chaotic orbits. It is easily seen that from (63) we retrieve again the asymptotic behavior $X_1(t) \propto t^{-1}$ for the case of regular orbits ($\chi_1 = 0$).
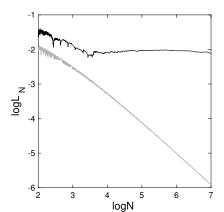
In the case of chaotic orbits the variation of $X_1(t)$ is usually irregular for relatively small t and only for large t the value of $X_1(t)$ stabilizes and tends to a constant positive value which is the mLCE χ_1. If for example the value of χ_1 is very small then initially, for small and intermediate values of t, the term proportional to t^{-1} dominates the r. h. s. of equation (63) and $X_1(t) \propto t^{-1}$. As t grows the significance of term $(b + z(t))/t$ diminishes and eventually the value of χ_1 becomes dominant and $X_1(t)$ stabilizes. It becomes evident that for smaller values of χ_1 the larger is the time required for $X_1(t)$ to reach its limiting value, and consequently $X_1(t)$ behaves as in the case of regular orbits, i. e. $X_1(t) \propto t^{-1}$ for larger time intervals. This behavior is clearly seen in Figure 7 where the evolution of $X_1(t)$ of chaotic orbits with small mLCE is shown. In particular, the values of the mLCE are $\chi_1 \approx 8 \cdot 10^{-3}$ (left panel) and $\chi_1 \approx 1.6 \cdot 10^{-7}$ (right panel). In both panels the evolution of $X_1(t)$ of regular orbits (following the power law $\propto t^{-1}$) is also plotted in order to facilitate the comparison between the two cases.

6 Computation of the spectrum of LCEs

While the knowledge of the mLCE χ_1 can be used for determining the regular ($\chi_1 = 0$) or chaotic ($\chi_1 > 0$) nature of orbits, the knowledge of part, or of the whole spectrum of LCEs, provides additional information on the underlying dynamics and on the statistical properties of the system, and can be used for measuring the fractal dimension of strange attractors in dissipative systems.

In Section 4.5 it was stated that, for Hamiltonian systems the existence of an integral of motion results to a pair of zero values in the spectrum of LCEs. As an example of such case we refer to the Hamiltonian system studied in [12]. This system has one more integral of motion apart from the Hamiltonian function and so 4 LCEs were always found to be equal to zero. Thus, the determination of the number of LCEs that vanish can be used as an indicator of the number of the independent integrals of motion that a dynamical system has.

It has been also stated in Section 4.5 that the spectrum of the LCEs of orbits in a connected chaotic region is independent of their initial conditions. So, we have a strong indication that two chaotic orbits belong to connected chaotic regions if they exhibit the same spectrum. As an example of this situation we refer to the case studied in [3] of two chaotic orbits of a 16D
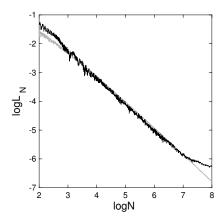
Fig. 7. Evolution of $X_1(t)$ (denoted as L_N) with respect to the discrete time t (denoted as N) in log–log scale for regular (grey curves) and chaotic (black curves) orbits of the 4d map (16) (left panel) and of a 4d map composed of two coupled 2d standard maps (right panel) (see [122] for more details). For regular orbits $X_1(t)$ tends to zero following a power law decay, $X_1(t) \propto t^{-1}$. For chaotic orbits $X_1(t)$ exhibits for some initial time interval the same power law decay before stabilizing to the positive value of the mLCE χ_1. The length of this time interval is larger for smaller values of χ_1. The chaotic orbits have $\chi_1 \approx 8 \cdot 10^{-3}$ (left panel) and $\chi_1 \approx 1.6 \cdot 10^{-7}$ (right panel) (after [122]).

Hamiltonian system having similar spectra of LCEs but very different initial conditions.

Vice versa, the existence of different LCEs spectra of chaotic orbits provides strong evidence that these orbits belong to different chaotic regions of the phase space that do not communicate. In [14] two chaotic orbits, previously studied in [34], were found to have significantly different spectra of LCEs and they were considered to belong to different chaotic regions which were called the 'big' (corresponding to the largest χ_1) and the 'small' chaotic sea. It is worth mentioning that the numerical results of [14] suggested the possible existence of an additional integral of motion for the 'small' chaotic sea, since χ_2 seemed to vanish. This assumption was in accordance to the results of [34] where such an integral was formally constructed.

The spectrum of LCEs is also related to two important quantities namely, the metric entropy, also called *Kolmogorov–Sinai (KS) entropy h*, and the *information dimension* D_1, which are trying to quantify the statistical properties of dynamical systems. For the explicit definition of these quantities, as well as detailed discussion of their relation to the LCEs the reader is referred for example to [9, 46, 54, 44] [92, p. 304–305] for the KS entropy and to [79, 46, 47, 66, 44] for the information dimension.

In particular, Pesin [106] showed that under suitable smoothness conditions the relation between the KS entropy h and the LCEs is given by

$$h = \int_{\mathcal{M}} \left[\sum_{\chi_i(\mathbf{x})>0} \chi_i(\mathbf{x}) \right] d\mu,$$

where the sum is extended over all positive LCEs and the integral is defined over a specified region $\mathcal{M}$ of the phase space $\mathcal{S}$.

Kaplan and Yorke [79] introduced a quantity, which they called the *Lyapunov dimension*

$$D_L = j + \frac{\sum_{i=1}^{j} \chi_i}{|\chi_{j+1}|} \tag{64}$$

where j is the largest integer for which $\chi_1 + \chi_2 + \ldots + \chi_j \geq 0$. The *Kaplan–Yorke conjecture* states that the information dimension D_1 is equal to the Lyapunov dimension D_L, i. e.

$$D_1 = D_L, \tag{65}$$

for a typical system, and thus, it can be used for the determination of the fractal dimension of strange attractors. The meaning of the word 'typical' is that it is not hard to construct examples where equation (65) is violated (see e. g [47]). But the claim is that these examples are pathological in that the slightest arbitrary change of the system restores the applicability of (65) and that such violation has 'zero probability' of occurring in practice. The validity of the Kaplan–Yorke conjecture has been proved in some cases [146, 87] although a general proof has not been achieved yet. We note that in the case of a $2N$D conservative system D_L is equal to the dimension of the whole space, i. e. $D_L = 2N$, because $j = 2N$ in (64) since $\sum_{i=1}^{2N} \chi_i = 0$ according to equation (47).

So, it becomes evident that developing an efficient algorithm for the numerical evaluation of few or of all LCEs is of great importance for the study of dynamical systems. In this section we present the different methods developed over the years for the computation of the spectrum of LCEs, focusing on the method suggested by Benettin et al. [14], the so–called *standard method*.

6.1 The standard method for computing LCEs

The basis for the computation of few or even of all LCEs is Theorem 3, which states that the computation of a p–LCE from equation (44), considering

a random choice of p $(1 < p \leq 2N)$ linearly independent initial deviation vectors, leads to the evaluation of the p–mLCE $\chi_1^{(p)}$, which is equal to the sum of the p largest 1–LCEs (46).

In order to evaluate the p–mLCE of an orbit with initial condition $\mathbf{x}(0)$, one has to follow simultaneously the time evolution of the orbit itself and of p linearly independent deviation vectors with initial conditions $\mathbf{w}_1(0), \mathbf{w}_2(0), \ldots, \mathbf{w}_p(0)$ (using the variational equations (8) or the equations of the tangent map (11)). Then, the p–mLCE is computed as the limit for $t \to \infty$ of the quantity

$$
\begin{aligned}
X^{(p)}(t) &= \frac{1}{t} \ln \frac{\mathrm{vol}_p \left(d_{\mathbf{x}(0)} \Phi^t\, \mathbf{w}_1(0), d_{\mathbf{x}(0)} \Phi^t\, \mathbf{w}_2(0), \cdots, d_{\mathbf{x}(0)} \Phi^t\, \mathbf{w}_p(0) \right)}{\mathrm{vol}_p \left(\mathbf{w}_1(0), \mathbf{w}_2(0), \ldots, \mathbf{w}_p(0) \right)} \\
&= \frac{1}{t} \ln \frac{\| \mathbf{w}_1(t) \wedge \mathbf{w}_2(t) \wedge \cdots \wedge \mathbf{w}_p(t) \|}{\| \mathbf{w}_1(0) \wedge \mathbf{w}_2(0) \wedge \cdots \wedge \mathbf{w}_p(0) \|} = \frac{1}{t} \ln \frac{\| \bigwedge_{i=1}^{p} \mathbf{w}_i(t) \|}{\| \bigwedge_{i=1}^{p} \mathbf{w}_i(0) \|}, \quad (66)
\end{aligned}
$$

which is also called the *finite time p–mLCE*. So we have

$$
\chi_1^{(p)} = \chi_1 + \chi_2 + \cdots + \chi_p = \lim_{t \to \infty} X^{(p)}(t). \tag{67}
$$

We recall that the quantity $\mathrm{vol}_p \left(\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_p \right)$ appearing in the above definition is the volume of the p–parallelogram having as edges the vectors $\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_p$ (see equations (106) and (105) in Appendix A).

The direct numerical implementation of equations (66) and (67) faces one additional difficulty apart from the fast growth of the norm of deviation vectors discussed in Section 5.1. This difficulty is due to the fact that when at least two vectors are involved (e. g. for the computation of $\chi_1^{(2)}$), the angles between their directions become too small for numerical computations.

This difficulty can be overcome on the basis of the following simple remark: an invertible linear map, as $d_{\mathbf{x}(0)} \Phi^t$, maps a linear p–dimensional subspace onto a linear subspace of the same dimension, and the coefficient of expansion of any p–dimensional volume under the action of any such linear map (like for example $\| \bigwedge_{i=1}^{p} \mathbf{w}_i(t) \| / \| \bigwedge_{i=1}^{p} \mathbf{w}_i(0) \|$ in our case) does not depend on the initial volume [14]. Since the numerical value of $\| \bigwedge_{i=1}^{p} \mathbf{w}_i(0) \|$ does not depend on the choice of the orthonormal basis of the space (see Appendix A for more details), in order to show the validity of this remark we will consider an appropriate basis which will facilitate our calculations.

In particular, let us consider an orthonormal basis $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \ldots, \hat{\mathbf{e}}_p\}$ of the p–dimensional space $E^p \subseteq \mathcal{T}_{\mathbf{x}(0)} \mathcal{S}$ spanned by $\{\mathbf{w}_1(0), \mathbf{w}_2(0), \ldots, \mathbf{w}_p(0)\}$. This basis can be extended to an orthonormal basis of the whole $2N$–dimensional space $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \ldots, \hat{\mathbf{e}}_p, \hat{\mathbf{e}}_{p+1}, \ldots, \hat{\mathbf{e}}_{2N}\}$ and $E^p \subseteq \mathcal{T}_{\mathbf{x}(0)} \mathcal{S}$ can be written as the direct sum of E^p and of the $(2N - p)$–dimensional subspace E' spanned by $\{\hat{\mathbf{e}}_{p+1}, \ldots, \hat{\mathbf{e}}_{2N}\}$

$$
\mathcal{T}_{\mathbf{x}(0)} \mathcal{S} = E^p \bigoplus E'.
$$

Consider also the $2N \times p$ matrix $\mathbf{W}(0)$ having as columns the coordinates of vectors $\mathbf{w}_i(0)$, $i = 1, 2, \ldots, p$ with respect to the complete orthonormal basis

$\hat{\mathbf{e}}_j$, $j = 1, 2, \ldots, 2N$, in analogy to equation (102). Since $\mathbf{w}_i(0) \in E^p$ this matrix has the form

$$\mathbf{W}(0) = \begin{bmatrix} \widetilde{\mathbf{W}}(0) \\ \mathbf{0}_{(2N-p) \times p} \end{bmatrix}$$

where $\widetilde{\mathbf{W}}(0)$ is a square $p \times p$ matrix and $\mathbf{0}_{(2N-p) \times p}$ is the $(2N - p) \times p$ matrix with all its elements equal to zero. Then, according to equations (105) and (106) the volume of the initial p–parallelogram is

$$\left\| \bigwedge_{i=1}^{p} \mathbf{w}_i(0) \right\| = \left| \det \widetilde{\mathbf{W}}(0) \right|, \tag{68}$$

since $\det \widetilde{\mathbf{W}}^{\mathrm{T}}(0) = \det \widetilde{\mathbf{W}}(0)$ for the square matrix $\widetilde{\mathbf{W}}(0)$.

Each deviation vector is evolved according to equation (7) and it can be computed through equation (9) or (12), with $\mathbf{Y}(t)$ being the $2N \times 2N$ matrix representing the action of $d_{\mathbf{x}(0)}\Phi^t$. By doing a similar choice for the basis of the $\mathcal{T}_{\Phi^t(\mathbf{x}(0))}\mathcal{S}$ space, equation (102) gives for the evolved vectors

$$\begin{bmatrix} \mathbf{w}_1(t) \ \mathbf{w}_2(t) \ \cdots \ \mathbf{w}_p(t) \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{e}}_1 \ \hat{\mathbf{e}}_2 \ \cdots \ \hat{\mathbf{e}}_p \end{bmatrix} \cdot \mathbf{Y}(t) \cdot \mathbf{W}(0) = \begin{bmatrix} \hat{\mathbf{e}}_1 \ \hat{\mathbf{e}}_2 \ \cdots \ \hat{\mathbf{e}}_p \end{bmatrix} \cdot \mathbf{W}(t).$$

Writing $\mathbf{Y}(t)$ as

$$\mathbf{Y}(t) = \begin{bmatrix} \mathbf{Y}_1(t) \ \mathbf{Y}_2(t) \end{bmatrix}$$

where $\mathbf{Y}_1(t)$ is the $2N \times p$ matrix formed from the first p columns of $\mathbf{Y}(t)$ and $\mathbf{Y}_2(t)$ is the $2N \times (2N - p)$ matrix formed from the last $2N - p$ columns of $\mathbf{Y}(t)$, $\mathbf{W}(t)$ assumes the form

$$\mathbf{W}(t) = \mathbf{Y}_1(t) \cdot \widetilde{\mathbf{W}}(0).$$

Then from equation (105) we get

$$\left\| \bigwedge_{i=1}^{p} \mathbf{w}_i(t) \right\| = \sqrt{\det \left(\widetilde{\mathbf{W}}^{\mathrm{T}}(0) \cdot \mathbf{Y}_1^{\mathrm{T}}(t) \cdot \mathbf{Y}_1(t) \cdot \widetilde{\mathbf{W}}(0) \right)}$$

$$= \sqrt{\det \widetilde{\mathbf{W}}^{\mathrm{T}}(0) \det \left(\mathbf{Y}_1^{\mathrm{T}}(t) \cdot \mathbf{Y}_1(t) \right) \det \widetilde{\mathbf{W}}(0)}$$

$$= |\det \widetilde{\mathbf{W}}(0)| \sqrt{\det \left(\mathbf{Y}_1^{\mathrm{T}}(t) \cdot \mathbf{Y}_1(t) \right)}. \tag{69}$$

Thus, from equations (68) and (69) we conclude that the coefficient of expansion

$$\frac{\left\| \bigwedge_{i=1}^{p} \mathbf{w}_i(t) \right\|}{\left\| \bigwedge_{i=1}^{p} \mathbf{w}_i(0) \right\|} = \sqrt{\det \left(\mathbf{Y}_1^{\mathrm{T}}(t) \cdot \mathbf{Y}_1(t) \right)}$$

does not depend on the initial volume but it is an intrinsic quantity of the subspaces defined by the properties of $d_{\mathbf{x}(0)}\Phi^t$. Note that in the particular

case of $p = 2N$ the coefficient of expansion is equal to $|\det \mathbf{Y}(t)|$ in accordance to equation (43). An alternative way of expressing this property is that, for two sets of linearly independent vectors $\{\mathbf{w}_1(0), \mathbf{w}_2(0), \ldots, \mathbf{w}_p(0)\}$ and $\{\mathbf{f}_1(0), \mathbf{f}_2(0), \ldots, \mathbf{f}_p(0)\}$ spanning the same p–dimensional subspace of $\mathcal{T}_{\mathbf{x}(0)}\mathcal{S}$, the relation

$$\frac{\|\bigwedge_{i=1}^{p} \mathbf{w}_i(t)\|}{\|\bigwedge_{i=1}^{p} \mathbf{w}_i(0)\|} = \frac{\|\bigwedge_{i=1}^{p} \mathbf{f}_i(t)\|}{\|\bigwedge_{i=1}^{p} \mathbf{f}_i(0)\|} \tag{70}$$

holds [119].

Let us now describe the method for the actual computation of the p–mLCE. Similarly to the computation of the mLCE we fix a small time interval τ and define quantity $X^{(p)}(t)$ (66) as

$$X^{(p)}(k\tau) = \frac{1}{k\tau} \sum_{i=1}^{k} \ln \frac{\|\bigwedge_{j=1}^{p} d_{\mathbf{x}(0)}\Phi^{i\tau}\,\mathbf{w}_j(0)\|}{\|\bigwedge_{j=1}^{p} d_{\mathbf{x}(0)}\Phi^{(i-1)\tau}\,\mathbf{w}_j(0)\|} = \frac{1}{k\tau} \sum_{i=1}^{k} \ln \gamma_i^{(p)} \tag{71}$$

where $\gamma_i^{(p)}$, $i = 1, 2, \ldots$, is the coefficient of expansion of a p–dimensional volume from $t = (i-1)\tau$ to $t = i\tau$. According to equation (70) $\gamma_i^{(p)}$ can be computed as the coefficient of expansion of the p–parallelogram defined by any p vectors spanning the same p–dimensional space. A suitable choice for this set is to consider an orthonormal set of vectors $\{\hat{\mathbf{w}}_1((i-1)\tau), \hat{\mathbf{w}}_2((i-1)\tau), \ldots, \hat{\mathbf{w}}_p((i-1)\tau)\}$ giving to equation (71) the simplified form

$$X^{(p)}(k\tau) = \frac{1}{k\tau} \sum_{i=1}^{k} \ln \gamma_i^{(p)} = \frac{1}{k\tau} \sum_{i=1}^{k} \ln \left\| \bigwedge_{j=1}^{p} d_{\mathbf{x}((i-1)\tau)}\Phi^{\tau}\,\hat{\mathbf{w}}_j((i-1)\tau) \right\|. \tag{72}$$

Thus, from equations (67) and (72) we get

$$\chi_1^{(p)} = \chi_1 + \chi_2 + \cdots + \chi_p = \lim_{k\to\infty} \frac{1}{k\tau} \sum_{i=1}^{k} \ln \gamma_i^{(p)} \tag{73}$$

for the computation of the p–mLCE. This equation is valid for $1 \leq p \leq 2N$ since in the extreme case of $p = 1$ it is simply reduced to equation (59) with $\alpha_i \equiv \gamma_i^{(1)}$. In order to estimate the values of χ_i, $i = 1, 2, \ldots, p$, which is our actual goal, we compute from (73) all the $\chi_1^{(p)}$ quantities and evaluate the LCEs from

$$\chi_i = \chi_1^{(i)} - \chi_1^{(i-1)}, \quad i = 2, 3, \ldots, p \tag{74}$$

with $\chi_1^{(1)} \equiv \chi_1$ [119].

Benettin et al. [14] noted that the p largest 1–LCEs can be evaluated at once by computing the evolution of just p deviation vectors for a particular choice of the orthonormalization procedure, namely performing the Gram–Schmidt orthonormalization method.

Let us discuss the Gram–Schmidt orthonormalization method in some detail. Let $\mathbf{w}_j(i\tau)$, $j = 1, 2, \ldots, p$ be the evolved deviation vectors $\hat{\mathbf{w}}_j((i-1)\tau)$ from time $t = (i-1)\tau$ to $t = i\tau$. From this set of linearly independent vectors we construct a new set of orthonormal vectors $\hat{\mathbf{w}}_j(i\tau)$ from equations

$$\mathbf{u}_1(i\tau) = \mathbf{w}_1(i\tau), \ \ \gamma_{1i} = \|\mathbf{u}_1(i\tau)\|, \ \ \hat{\mathbf{w}}_1(i\tau) = \frac{\mathbf{u}_1(i\tau)}{\gamma_{1i}}, \tag{75}$$

$$\mathbf{u}_2(i\tau) = \mathbf{w}_2(i\tau) - \langle\mathbf{w}_2(i\tau), \hat{\mathbf{w}}_1(i\tau)\rangle\hat{\mathbf{w}}_1(i\tau), \ \ \gamma_{2i} = \|\mathbf{u}_2(i\tau)\|, \ \ \hat{\mathbf{w}}_2(i\tau) = \frac{\mathbf{u}_2(i\tau)}{\gamma_{2i}},$$

$$\mathbf{u}_3(i\tau) = \mathbf{w}_3(i\tau) - \langle\mathbf{w}_3(i\tau), \hat{\mathbf{w}}_1(i\tau)\rangle\hat{\mathbf{w}}_1(i\tau) - \langle\mathbf{w}_3(i\tau), \hat{\mathbf{w}}_2(i\tau)\rangle\hat{\mathbf{w}}_2(i\tau),$$

$$\gamma_{3i} = \|\mathbf{u}_3(i\tau)\|, \ \ \hat{\mathbf{w}}_3(i\tau) = \frac{\mathbf{u}_3(i\tau)}{\gamma_{3i}},$$

$$\vdots$$

which are repeated up to the computation of $\hat{\mathbf{w}}_p(i\tau)$. We remark that $\langle\mathbf{w}, \mathbf{u}\rangle$ denotes the usual inner product of vectors $\mathbf{w}$, $\mathbf{u}$. The general form of the above equations, which is the core of the Gram–Schmidt orthonormalization method, is

$$\mathbf{u}_k(i\tau) = \mathbf{w}_k(i\tau) - \sum_{j=1}^{k-1}\langle\mathbf{w}_k(i\tau), \hat{\mathbf{w}}_j(i\tau)\rangle\hat{\mathbf{w}}_j(i\tau),$$

$$\gamma_{ki} = \|\mathbf{u}_k(i\tau)\|, \ \ \hat{\mathbf{w}}_k(i\tau) = \frac{\mathbf{u}_k(i\tau)}{\gamma_{ki}}, \tag{76}$$

for $1 \leq k \leq p$.

As we will show in Section 6.3 the volume of the p–parallelogram having as edges the vectors $d_{\mathbf{x}((i-1)\tau)}\Phi^\tau \hat{\mathbf{w}}_j((i-1)\tau) = \mathbf{w}_j(i\tau)$, $j = 1, 2, \ldots, p$ is equal to the volume of the p–parallelogram having as edges the vectors $\mathbf{u}_j(i\tau)$, i. e.

$$\left\|\bigwedge_{j=1}^{p} d_{\mathbf{x}((i-1)\tau)}\Phi^\tau \hat{\mathbf{w}}_j((i-1)\tau)\right\| = \left\|\bigwedge_{j=1}^{p} \mathbf{u}_j(i\tau)\right\|. \tag{77}$$

Since vectors $\mathbf{u}_j(i\tau)$ are normal to each other, the volume of their p–parallelogram is equal to the product of their norms. This leads to

$$\gamma_i^{(p)} = \left\|\bigwedge_{j=1}^{p} \mathbf{u}_j(i\tau)\right\| = \prod_{j=1}^{p} \gamma_{ji}. \tag{78}$$

Then, equation (73) takes the form

$$\chi_1^{(p)} = \chi_1 + \chi_2 + \cdots + \chi_p = \lim_{k\to\infty} \frac{1}{k\tau}\sum_{i=1}^{k}\ln\left(\prod_{j=1}^{p}\gamma_{ji}\right).$$

Using now equation (74) we are able to evaluate the 1–LCE χ_p as

$$\chi_p = \chi_1^{(p)} - \chi_1^{(p-1)} = \lim_{k \to \infty} \frac{1}{k\tau} \sum_{i=1}^{k} \ln \frac{\prod_{j=1}^{p} \gamma_{ji}}{\prod_{j=1}^{p-1} \gamma_{ji}} = \lim_{k \to \infty} \frac{1}{k\tau} \sum_{i=1}^{k} \ln \gamma_{pi}.$$

In conclusion we see that the value of the 1–LCE χ_p with $1 < p \leq 2N$ can be computed as the limiting value, for $t \to \infty$, of the quantity

$$X_p(k\tau) = \frac{1}{k\tau} \sum_{i=1}^{k} \ln \gamma_{pi},$$

i. e.

$$\chi_p = \lim_{k \to \infty} X_p(k\tau) = \lim_{k \to \infty} \frac{1}{k\tau} \sum_{i=1}^{k} \ln \gamma_{pi}, \tag{79}$$

where γ_{ji}, $j = 1, 2, \ldots, p$, $i = 1, 2, \ldots$ are quantities evaluated during the successive orthonormalization procedures (equations (75) and (76)). Note that for $p = 1$ equation (79) is actually equation (59) with $\alpha_i \equiv \gamma_{1i}$.

6.2 The numerical algorithm for the standard method

In practice, in order to compute the p largest 1–LCEs with $1 < p \leq 2N$ we follow the evolution of p initially orthonormal deviation vectors $\hat{\mathbf{w}}_j(0) = \mathbf{w}_j(0)$ and every $t = \tau$ time units we replace the evolved vectors $\mathbf{w}_j(k\tau)$ $j = 1, 2, \ldots, p$, $k = 1, 2, \ldots$ by a new set of orthonormal vectors produced by the Gram-Schmidt orthonormalization method (76). During the orthonormalization procedure the quantities γ_{jk} are computed and $\chi_1, \chi_2, \ldots, \chi_p$ are estimated from equation (79). This algorithm is described in pseudo–code in Table 2 and can be used for the computation of few or even all 1–LCEs. A Fortran code of this algorithm can be found in [144], while [117] contains a similar code developed for the computer algebra platform "Mathematica" (Wolfram Research Inc.).

Let us illustrate the implementation of this algorithm in the particular case of the computation of the 2 largest LCEs χ_1 and χ_2. As shown in Figure 8 we start our computation with two orthonormal deviation vectors $\mathbf{w}_1(0)$ and $\mathbf{w}_2(0)$ which are evolved to $\mathbf{w}_1(\tau)$, $\mathbf{w}_2(\tau)$ at $t = \tau$. Then according to the the Gram-Schmidt orthonormalization method (75) we define vectors $\mathbf{u}_1(\tau)$ and $\mathbf{u}_2(\tau)$. In particular, $\mathbf{u}_1(\tau)$ coincides with $\mathbf{w}_1(\tau)$ while, $\mathbf{u}_2(\tau)$ is the component of vector $\mathbf{w}_2(\tau)$ in the direction perpendicular to vector $\mathbf{u}_1(\tau)$. The norms of these two vectors define the quantities $\gamma_{11} = \|\mathbf{u}_1(\tau)\|$, $\gamma_{21} = \|\mathbf{u}_2(\tau)\|$ needed for the estimation of χ_1, χ_2 from equation (79). Then vectors $\hat{\mathbf{w}}_1(\tau)$ and $\hat{\mathbf{w}}_2(\tau)$ are defined as unitary vectors in the directions of $\mathbf{u}_1(\tau)$ and $\mathbf{u}_2(\tau)$ respectively. Since the unitary vectors $\hat{\mathbf{w}}_1(\tau)$, $\hat{\mathbf{w}}_2(\tau)$ are normal by construction they constitute the initial set of orthonormal vectors for the next iteration of the algorithm. From Figure 8 we easily see that the parallelograms

Table 2. The standard method. The algorithm for the computation of the p largest LCEs $\chi_1, \chi_2, \ldots, \chi_p$ as limits for $t \to \infty$ of quantities $X_1(t), X_2(t), \ldots, X_p(t)$ (71), according to equation (79). The program computes the evolution of $X_1(t), X_2(t), \ldots, X_p(t)$ with respect to time t up to a given upper value of time $t = T_M$ or until any of the quantities $X_1(t), X_2(t), \ldots, X_p(t)$ attain a very small value, smaller than a low threshold value X_m.

Input: 1. Hamilton equations of motion (2) and variational equations (8), or
   equations of the map (4) and of the tangent map (11).
   2. Number of desired LCEs p.
   3. Initial condition for the orbit $\mathbf{x}(0)$.
   4. Initial *orthonormal* deviation vectors $\mathbf{w}_1(0), \mathbf{w}_2(0), \ldots, \mathbf{w}_p(0)$.
   5. Renormalization time τ.
   6. Maximal time: T_M and minimum allowed value of $X_1(t)$,
   $X_2(t), \ldots, X_p(t)$: X_m.

Step 1 **Set** the stopping flag, $SF \leftarrow 0$, and the counter, $k \leftarrow 1$.
Step 2 **While** $(SF = 0)$ **Do**
   Evolve the orbit and the deviation vectors from time $t = (k-1)\tau$
   to $t = k\tau$, i. e. **Compute** $\mathbf{x}(k\tau)$ and $\mathbf{w}_1(k\tau), \mathbf{w}_2(k\tau), \ldots, \mathbf{w}_p(k\tau)$.
Step 3   Perform the **Gram-Schmidt orthonormalization** procedure
   according to equation (76):
   Do for $j = 1$ to p
      Compute current vectors $\mathbf{u}_j(k\tau)$ and values of γ_{jk}.
      Compute and **Store** current values of $X_j(k\tau) = \frac{1}{k\tau} \sum_{i=1}^{k} \ln \gamma_{ji}$.
      Set $\mathbf{w}_j(k\tau) \leftarrow \mathbf{u}_j(k\tau)/\gamma_{jk}$.
   End Do
Step 4   **Set** the counter $k \leftarrow k + 1$.
Step 5   **If** $[(k\tau > T_M)$ or (Any of $X_j((k-1)\tau) < X_m, j = 1, 2, \ldots, p)]$ **Then**
      Set $SF \leftarrow 1$.
   End If
   End While
Step 6 **Report** the time evolution of $X_1(t), X_2(t), \ldots, X_p(t)$.

defined by vectors $\mathbf{w}_1(\tau), \mathbf{w}_2(\tau)$ and by vectors $\mathbf{u}_1(\tau)$ and $\mathbf{u}_2(\tau)$ have the same area. This equality corresponds to the particular case $p = 2$, $i = 1$ of equation (77). Evidently, since vectors $\mathbf{u}_1(\tau)$, $\mathbf{u}_2(\tau)$ are perpendicular to each other, we have $\text{vol}_2(\mathbf{u}_1(\tau), \mathbf{u}_2(\tau)) = \gamma_{11}\gamma_{21}$ in accordance to equation (78).

6.3 Connection between the standard method and the QR decomposition

Let us rewrite equations (75) of the Gram-Schmidt orthonormalization procedure, by solving them with respect to $\mathbf{w}_j(i\tau)$, $j = 1, 2, \ldots, p$, with $1 < p \leq 2N$

$$\mathbf{w}_1(i\tau) = \gamma_{1i}\hat{\mathbf{w}}_1(i\tau) \tag{80}$$
$$\mathbf{w}_2(i\tau) = \langle \hat{\mathbf{w}}_1(i\tau), \mathbf{w}_2(i\tau) \rangle \hat{\mathbf{w}}_1(i\tau) + \gamma_{2i}\hat{\mathbf{w}}_2(i\tau)$$
$$\mathbf{w}_3(i\tau) = \langle \hat{\mathbf{w}}_1(i\tau), \mathbf{w}_3(i\tau) \rangle \hat{\mathbf{w}}_1(i\tau) + \langle \hat{\mathbf{w}}_2(i\tau), \mathbf{w}_3(i\tau) \rangle \hat{\mathbf{w}}_2(i\tau) + \gamma_{3i}\hat{\mathbf{w}}_3(i\tau)$$

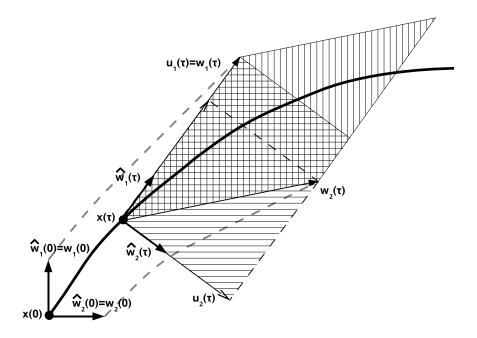**Fig. 8.** Numerical scheme for the computation of the 2 largest LCEs χ_1, χ_2 according to the standard method. The orthonormal deviation vectors $\mathbf{w}_1(0)$, $\mathbf{w}_2(0)$ are evolved according to the variational equations (8) (continuous time) or the equations of the tangent map (11) (discrete time) for $t = \tau$ time units. The evolved vectors $\mathbf{w}_1(\tau)$, $\mathbf{w}_2(\tau)$, are replaced by a set of orthonormal vectors $\hat{\mathbf{w}}_1(\tau)$, $\hat{\mathbf{w}}_2(\tau)$, which span the same 2–dimensional vector space, according to the Gram-Schmidt orthonormalization method (76). Then these vectors are again evolved and the same procedure is iteratively applied. For each successive time interval $[(i-1)\tau, i\tau]$, $i = 1, 2, \ldots$, the quantities $\gamma_{1i} = \|\mathbf{u}_1(i\tau)\|$, $\gamma_{2i} = \|\mathbf{u}_2(i\tau)\|$ are computed and χ_1, χ_2 are estimated from equation (79).

$$\vdots$$

and get the general form

$$\mathbf{w}_k(i\tau) = \sum_{j=1}^{k-1} \langle \hat{\mathbf{w}}_j(i\tau), \mathbf{w}_k(i\tau) \rangle \hat{\mathbf{w}}_j(i\tau) + \gamma_{ki} \hat{\mathbf{w}}_k(i\tau), \quad k = 1, 2, \ldots, p.$$

This set of equations can be rewritten in matrix form as follows:

$$\begin{bmatrix} \mathbf{w}_1(i\tau) \ \mathbf{w}_2(i\tau) \ \cdots \ \mathbf{w}_p(i\tau) \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{w}}_1(i\tau) \ \hat{\mathbf{w}}_2(i\tau) \ \cdots \ \hat{\mathbf{w}}_p(i\tau) \end{bmatrix} \cdot$$

$$\cdot \begin{bmatrix} \gamma_{1i} & \langle \hat{\mathbf{w}}_1(i\tau), \mathbf{w}_2(i\tau) \rangle & \langle \hat{\mathbf{w}}_1(i\tau), \mathbf{w}_3(i\tau) \rangle & \cdots & \langle \hat{\mathbf{w}}_1(i\tau), \mathbf{w}_p(i\tau) \rangle \\ 0 & \gamma_{2i} & \langle \hat{\mathbf{w}}_2(i\tau), \mathbf{w}_3(i\tau) \rangle & \cdots & \langle \hat{\mathbf{w}}_2(i\tau), \mathbf{w}_p(i\tau) \rangle \\ 0 & 0 & \gamma_{3i} & \cdots & \langle \hat{\mathbf{w}}_3(i\tau), \mathbf{w}_p(i\tau) \rangle \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & & \gamma_{pi} \end{bmatrix} .$$

So the $2N \times p$ matrix $\mathbf{W}(i\tau) = \begin{bmatrix} \mathbf{w}_1(i\tau) \ \mathbf{w}_2(i\tau) \ \cdots \ \mathbf{w}_p(i\tau) \end{bmatrix}$, having as columns the linearly independent deviation vectors $\mathbf{w}_j(i\tau)$, $j = 1, 2, \ldots, p$ is written as a product of the $2N \times p$ matrix $\mathbf{Q} = \begin{bmatrix} \hat{\mathbf{w}}_1(i\tau) \ \hat{\mathbf{w}}_2(i\tau) \ \cdots \ \hat{\mathbf{w}}_p(i\tau) \end{bmatrix}$, having as columns the coordinates of the orthonormal vectors $\hat{\mathbf{w}}_j(i\tau)$, $j = 1, 2, \ldots, p$ and satisfying $\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I}_p$, and of an upper triangular $p \times p$ matrix $\mathbf{R}(i\tau)$ with positive diagonal elements

$$\mathbf{R}_{jj}(i\tau) = \gamma_{ji}, \ \ j = 1, 2, \ldots, p, \ \ i = 1, 2, \ldots.$$

From equations (80) we easily see that $\langle \hat{\mathbf{w}}_j(i\tau), \mathbf{w}_j(i\tau) \rangle = \gamma_{ji}$ and so matrix $\mathbf{R}(i\tau)$ can be also expressed as

$$\mathbf{R}(i\tau) = \begin{bmatrix} \langle \hat{\mathbf{w}}_1(i\tau), \mathbf{w}_1(i\tau) \rangle & \langle \hat{\mathbf{w}}_1(i\tau), \mathbf{w}_2(i\tau) \rangle & \cdots & \langle \hat{\mathbf{w}}_1(i\tau), \mathbf{w}_p(i\tau) \rangle \\ 0 & \langle \hat{\mathbf{w}}_2(i\tau), \mathbf{w}_2(i\tau) \rangle & \cdots & \langle \hat{\mathbf{w}}_2(i\tau), \mathbf{w}_p(i\tau) \rangle \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & \langle \hat{\mathbf{w}}_p(i\tau), \mathbf{w}_p(i\tau) \rangle \end{bmatrix} .$$

The above procedure is the so–called QR decomposition of a matrix. In practice, we proved by actually constructing the $\mathbf{Q}$ and $\mathbf{R}$ matrices via the Gram-Schmidt orthonormalization method, the following theorem:

Theorem 4. *Let $\boldsymbol{A}$ be an $n \times m$ $(n \geq m)$ matrix with linearly independent columns. Then $\boldsymbol{A}$ can be uniquely factorized as*

$$\boldsymbol{A} = \boldsymbol{Q} \cdot \boldsymbol{R},$$

where $\boldsymbol{Q}$ is an $n \times m$ matrix with orthogonal columns, satisfying $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q} = \boldsymbol{I}_m$ and $\boldsymbol{R}$ is an $m \times m$ invertible upper triangular matrix with positive diagonal entries.

Although we presented the QR decomposition through the Gram-Schmidt orthonormalization procedure this decomposition can also be achieved by others, computationally more efficient techniques like for example the Householder transformation [62][107, §2.10].

Observing that the quantities γ_{ji}, $j = 1, 2 \ldots, p$, $i = 1, 2 \ldots$, needed for the evaluation of the LCEs through equation (79) are the diagonal elements of $\mathbf{R}(i\tau)$ we can implement a variant of the standard method for the computation on the LCEs, which is based on the QR decomposition procedure [44, 62, 36, 40]. Similarly to the procedure followed in Section 6.2, in order to compute the p $(1 < p \leq 2N)$ largest LCEs we follow the evolution of p initially orthonormal

deviation vectors $\hat{\mathbf{w}}_j(0) = \mathbf{w}_j(0)$, $j = 1, 2 \ldots, p$, which can be considered as columns of a $2N \times p$ matrix $\mathbf{Q}(0)$. Every $t = \tau$ time units the matrix $\mathbf{W}(i\tau)$, $i = 1, 2, \ldots$, having as columns the deviation vectors

$$d_{\mathbf{x}((i-1)\tau)}\Phi^\tau \, \hat{\mathbf{w}}_j((i-1)\tau) = \mathbf{w}_j(i\tau), \;\; j = 1, 2, \ldots, p,$$

i. e. the columns of $\mathbf{Q}((i-1)\tau)$ evolved in time interval $[(i-1)\tau, i\tau]$ by the action of $d_{\mathbf{x}((i-1)\tau)}\Phi^\tau$, undergoes the QR decomposition procedure

$$\mathbf{W}(i\tau) = \mathbf{Q}(i\tau) \cdot \mathbf{R}(i\tau) \tag{81}$$

and the new $\mathbf{Q}(i\tau)$ is again evolved for the next time interval $[i\tau, (i+1)\tau]$, and so on and so forth. Then the LCEs are estimated from the values of the diagonal elements of matrix $\mathbf{R}(i\tau)$ as

$$\chi_p = \lim_{k \to \infty} \frac{1}{k\tau} \sum_{i=1}^{k} \ln \mathbf{R}_{pp}(i\tau). \tag{82}$$

The corresponding algorithm is presented in pseudo-code in Table 3. From the above–presented analysis it becomes evident that the standard method developed by Shimada and Nagashima [119] and Benettin et al. [14] for the computation of the LCEs, is practically a QR decomposition procedure performed by the Gram–Schmidt orthonormalization method, although the authors of these papers formally do not refer to the QR decomposition. We note that both the standard method and the QR decomposition technique presented here can be used for the computation of part $(p < 2N)$ or of the whole $(p = 2N)$ spectrum of LCEs.

   As a final remark on the QR decomposition technique let us show the validity of equation (77) by considering the QR decomposition of matrix $\mathbf{W}(i\tau)$ (81). According to equations (105) and (106) we have

$$\left\| \bigwedge_{j=1}^{p} \mathbf{w}_j(i\tau) \right\| = \sqrt{\det \left(\mathbf{W}^{\mathrm{T}}(i\tau) \cdot \mathbf{W}(i\tau) \right)}$$

$$= \sqrt{\det \left(\mathbf{R}^{\mathrm{T}}(i\tau) \cdot \mathbf{Q}^{\mathrm{T}}(i\tau) \cdot \mathbf{Q}(i\tau) \cdot \mathbf{R}(i\tau) \right)}$$

$$= \sqrt{\det \mathbf{R}^{\mathrm{T}}(i\tau) \det \mathbf{R}(i\tau)} = |\det \mathbf{R}(i\tau)|$$

$$= \prod_{j=1}^{p} \gamma_{ji} = \prod_{j=1}^{p} \|\mathbf{u}_j(i\tau)\| = \left\| \bigwedge_{j=1}^{p} \mathbf{u}_j(i\tau) \right\|,$$

where the identities $\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I}_p$ and $\det \mathbf{R}(i\tau) = \prod_{j=1}^{p} \gamma_{ji}$ have been used.

6.4 Other methods for computing LCEs

Over the years several methods have been proposed and applied for computing the numerical values of the LCEs. The standard method we discussed so

Table 3. Discrete QR decomposition. The algorithm for the computation of the p largest LCEs $\chi_1, \chi_2, \ldots, \chi_p$ according to the QR decomposition method. The program computes the evolution of $X_1(t), X_2(t), \ldots, X_p(t)$ with respect to time t up to a given upper value of time $t = T_M$ or until any of the these quantities becomes smaller than a low threshold value X_m.

Input: 1. Hamilton equations of motion (2) and variational equations (8), or
equations of the map (4) and of the tangent map (11).
2. Number of desired LCEs p.
3. Initial condition for the orbit $\mathbf{x}(0)$.
4. Initial matrix $\mathbf{Q}(0)$ having as columns *orthonormal* deviation vectors
$\mathbf{w}_1(0)$, $\mathbf{w}_2(0)$, ..., $\mathbf{w}_p(0)$.
5. Time interval τ between successive QR decompositions.
6. Maximal time: T_M and minimum allowed value of $X_1(t)$,
$X_2(t)$, ..., $X_p(t)$: X_m.

Step 1	**Set** the stopping flag, $SF \leftarrow 0$, and the counter, $k \leftarrow 1$.
Step 2	**While** $(SF = 0)$ **Do**
	Evolve the orbit and the matrix $\mathbf{Q}((k-1)\tau)$ from time $t = (k-1)\tau$ to $t = k\tau$, i. e. **Compute** $\mathbf{x}(k\tau)$ and $\mathbf{W}(i\tau)$.
Step 3	Perform the **QR decomposition** of $\mathbf{W}(i\tau)$ according to (81): **Compute** $\mathbf{Q}(k\tau)$ and $\mathbf{R}(k\tau)$.
	Compute and **Store** current values of $X_j(k\tau) = \frac{1}{k\tau} \sum_{i=1}^{k} \ln \mathbf{R}_{jj}(i\tau)$, $j = 1, 2 \ldots, p$.
Step 4	**Set** the counter $k \leftarrow k + 1$.
Step 5	**If** $[(k\tau > T_M)$ or (Any of $X_j((k-1)\tau) < X_m$, $j = 1, 2, \ldots, p)]$ **Then** **Set** $SF \leftarrow 1$.
	End If
	End While
Step 6	**Report** the time evolution of $X_1(t), X_2(t), \ldots, X_p(t)$.

far, is the first and probably the simplest method to address this problem. As we showed in Section 6.3 the standard method, which requires successive applications of the Gram-Schmidt orthonormalization procedure, is practically equivalent to the QR decomposition technique.

The reorthonormalization of deviation vectors plays an indispensable role for computing the LCEs and the corresponding methods can be distinguished in discrete and continuous methods. The *discrete methods* iteratively approximate the LCEs in a finite number of (discrete) time steps and therefore apply to both continuous and discrete dynamical systems [62, 36, 40]. The standard method and its QR decomposition version, are discrete methods. A method is called *continuous* when all relevant quantities are obtained as solutions of certain ordinary differential equations, which maintain orthonormality of deviation vectors continuously. Therefore such methods can only be formulated for continuous dynamical systems and not for maps. The use of continuous orthonormalization for the numerical computation of LCEs was first proposed by Goldhirsch et al. [63] and afterwards developed by several authors [67, 62, 36, 40, 26, 110, 109, 94, 38].

Discrete and continuous methods are based on appropriate decomposition of matrices performed usually by the QR decomposition or by the SVD procedure. The discrete QR decomposition, which has already been presented in Section 6.3 is the most frequently used method and has proved to be quite efficient and reliable. The continuous QR decomposition, as well as methods based on the SVD procedure are discussed in some detail at the end of the current section.

Variants of these techniques have been also proposed by several authors. Let us briefly refer to some of them. Rangarajan et al. [110] introduced a method for the computation of part or of the whole spectrum of LCEs for continuous dynamical systems, which does not require rescaling and renormalization of vectors. The key feature of their approach is the use of explicit group theoretical representations of orthogonal matrices, which leads to a set of coupled ordinary differential equations for the LCEs along with the various angles parameterizing the orthogonal matrices involved in the process. Ramasubramanian and Sriram [109] showed that the method is competitive with the standard method and the continuous QR decomposition.

Carbonell et al. [20] proposed a method for the evaluation of the whole spectrum of LCEs by approximating the differential equations describing the evolution of an orbit of a continuous dynamical system and their associated variational equations by two piecewise linear sets of ordinary differential equations. Then an SVD or a QR decomposition–based method is applied to these two new sets of equations, allowing us to obtain approximations of the LCEs of the original system. An advantage of this method is that it does not require the simultaneous integration of the two sets of piecewise linear equations.

Lu et al. [94] proposed a new continuous method for the computation of few or of all LCEs, which is related to the QR decomposition technique. According to their method one follows the evolution of orthogonal vectors, similarly to the QR method, but does not require them to be necessarily orthonormal. By relaxing the length requirement Lu et al. [94] established a set of recursive differential equations for the evolution of these vectors. Using symplectic Runge–Kutta integration schemes for the evolution of these vectors they succeeded in preserving automatically the orthogonality between any two successive vectors. Normalization of vectors occurs whenever the magnitude of any vector exceeds given lower or upper bounds.

Chen et al. [24] proposed a simple discrete QR algorithm for the computation of the whole spectrum of LCEs of a continuous dynamical system. Their method is based on a suitable approximation of the solution of variational equations by assuming that the Jacobian matrix remains constant over small integration time steps. Thus, the scheme requires the numerical solution of the $2N$ equations of motion but not the solution of the $(2N)^2$ variational equations since their solution is approximated by an explicit expression involving the computed orbit. This approach led to a computationally fast evaluation of the LCEs for various multidimensional dynamical systems studied in [24].

It is worth mentioning here a completely different approach, with respect to the above–mentioned techniques, which was developed at the early 80's. In particular, Frøyland proposed in [60] an algorithm for the computation of LCEs, which he claimed to be quite efficient in the case of low–dimensional systems, and applied it to the Lorenz system [61]. The basic idea behind this algorithm is the implementation of appropriate differential equations describing the time evolution of volume elements around the orbits of the dynamical system, instead of defining these volumes through deviation vectors whose evolution is governed by the usual variational equations (8).

Apart from the actual numerical computation of the values of the LCEs, methods for the theoretical estimation of those values have been also developed. For example, Li and Chen [90] provided a theorem for the estimation of lower and upper bounds for the values of all LCEs in the case of discrete maps. These results were also generalized for the case of continues dynamical systems [91]. The validity of these estimates was demonstrated by a comparison between the estimated bounds and the numerically computed spectrum of LCEs of some specific dynamical systems [90, 91].

Finally, let us refer to a powerful analytical method which allows one to verify the existence of positive LCEs for a dynamical system, the so–called *cone technique*. The method was suggested by Wojtkowski [142] and has been extensively applied for the study of chaotic billiards [142, 143, 43, 97] and geodesic flows [41, 42, 19]. A concise description of the techniques can also be found in [7] and [25, §3.13]. Considering the space $\mathbb{R}^n$ a *cone* $\mathcal{C}_\gamma$, with $\gamma > 0$, centered around $\mathbb{R}^{n-k}$ is

$$\mathcal{C}_\gamma = \left\{ (\mathbf{u}, \mathbf{v}) \in \mathbb{R}^k \times \mathbb{R}^{n-k} : \|\mathbf{u}\| < \gamma \|\mathbf{v}\| \right\} \cup (\mathbf{0}, \mathbf{0}). \qquad (83)$$

Note that $\{\mathbf{0}\} \times \mathbb{R}^{n-k} \subset \mathcal{C}_\gamma$ for every γ. In the particular case of $n = 3$, $k = 2$, $\mathcal{C}_\gamma$ corresponds to the usual 3–dimensional cone, while in the case of the plane ($n = 2$) a cone $\mathcal{C}_\gamma$ around an axis L is the set of vectors of $\mathbb{R}^2$ that make angle $\phi < \arctan \gamma$ with the line L. In the case of Hamiltonian systems (and symplectic maps) a cone can get the simple form $\delta \mathbf{q} \cdot \delta \mathbf{p} > 0$. Finding an invariant family of cones (83) in $\mathcal{T}_\mathbf{x} \mathcal{S}$, which are mapped strictly into themselves by $d_\mathbf{x} \Phi^t$, guarantees that the values of the $n - k$ largest LCEs are positive [142, 143]. We emphasize that the cone technique is not used for the explicit numerical computation of the LCEs, but for the analytical proof of the existence of positive LCEs, providing at the same time some bounds for their actual values.

Continuous QR decomposition methods

The QR decomposition methods allow the computation of all or of the p ($1 < p < 2N$) largest LCEs. Let us discuss in more detail the developed procedure for both cases following mainly [62, 36, 94].

Computing the complete spectrum of LCEs

The basic idea of the method is to avoid directly solving the differential equation (10), by requiring $\mathbf{Y}(t) = \mathbf{Q}(t)\mathbf{R}(t)$ where $\mathbf{Q}(t)$ is orthogonal and $\mathbf{R}(t)$ is upper triangular with positive diagonal elements, according to Theorem 4. With this decomposition, one can write equation (10) into the form

$$\mathbf{Q}^{\mathrm{T}}\dot{\mathbf{Q}} + \dot{\mathbf{R}}\mathbf{R}^{-1} = \mathbf{Q}^{\mathrm{T}}\mathbf{A}\mathbf{Q},$$

where, for convenience, we dropped out the explicit dependence of the matrices on time t, i. e. $\mathbf{Q}(t) \equiv \mathbf{Q}$. Since $\mathbf{Q}^{\mathrm{T}}\dot{\mathbf{Q}}$ is skew and $\dot{\mathbf{R}}\mathbf{R}^{-1}$ is upper triangular, one reads off the differential equations

$$\dot{\mathbf{Q}} = \mathbf{Q}\mathbf{S}, \tag{84}$$

where $\mathbf{S}$ is the skew symmetric matrix

$$\mathbf{S} = \mathbf{Q}^{\mathrm{T}}\dot{\mathbf{Q}}$$

with elements

$$\mathbf{S}_{ij} = \begin{cases} (\mathbf{Q}^{\mathrm{T}}\mathbf{A}\mathbf{Q})_{ij} & i > j \\ 0 & i = j \\ -(\mathbf{Q}^{\mathrm{T}}\mathbf{A}\mathbf{Q})_{ji} & i < j \end{cases}, \quad i,j = 1, 2, \ldots, 2N, \tag{85}$$

and

$$\frac{\dot{\mathbf{R}}_{pp}}{\mathbf{R}_{pp}} = (\mathbf{Q}^{\mathrm{T}}\mathbf{A}\mathbf{Q})_{pp}, \quad p, = 1, 2, \ldots, 2N \tag{86}$$

where $\mathbf{R}_{pp}$ are the diagonal elements of $\mathbf{R}$. As we have already seen in equation (82) the LCEs are related to the elements $\mathbf{R}_{pp}$, through

$$\chi_p = \lim_{t\to\infty} \frac{1}{t} \ln \mathbf{R}_{pp}(t).$$

Thus, in order to compute the spectrum of LCEs only equations (84) and (86) have to be solved simultaneously with the equations of motion (2). In practice, the knowledge of matrix $\mathbf{R}$ is not necessary for the actual computation of the LCEs. Noticing that

$$\frac{d}{dt}(\ln \mathbf{R}_{pp}) = \frac{\dot{\mathbf{R}}_{pp}}{\mathbf{R}_{pp}} = (\mathbf{Q}^{\mathrm{T}}\mathbf{A}\mathbf{Q})_{pp} = \mathbf{q}_p \cdot \mathbf{A}\mathbf{q}_p, \tag{87}$$

where $\mathbf{q}_p$ is the pth column vector of $\mathbf{Q}$, we can compute the LCEs using

$$\chi_p = \lim_{t\to\infty} \frac{1}{t} \int_0^t \mathbf{q}_p \cdot \mathbf{A}\mathbf{q}_p \, dt.$$

In practice, the LCEs can be estimated through a recursive formula. Let

$$X_p(k\tau) = \frac{1}{k\tau} \int_0^{k\tau} \mathbf{q}_p \cdot \mathbf{A}\mathbf{q}_p dt.$$

Then we have

$$X_p\left((k+1)\tau\right) = \frac{1}{(k+1)\tau} \int_0^{(k+1)\tau} \mathbf{q}_p \cdot \mathbf{A}\mathbf{q}_p dt$$

$$= \frac{1}{(k+1)\tau} \int_0^{k\tau} \mathbf{q}_p \cdot \mathbf{A}\mathbf{q}_p dt + \frac{1}{(k+1)\tau} \int_{k\tau}^{(k+1)\tau} \mathbf{q}_p \cdot \mathbf{A}\mathbf{q}_p dt \ .$$

Replacing the first integral with $k\tau X_p(k\tau)$ we get

$$X_p\left((k+1)\tau\right) = \frac{k}{k+1} X_p(k\tau) + \frac{1}{(k+1)\tau} \int_{k\tau}^{(k+1)\tau} \mathbf{q}_p \cdot \mathbf{A}\mathbf{q}_p dt, \qquad (88)$$

and

$$\chi_p = \lim_{k\to\infty} X_p(k\tau). \qquad (89)$$

The basic difference between the discrete QR decomposition method presented in Section 6.3, and the continuous QR method presented here, is that in the first method the orthonormalization is performed numerically at discrete time steps, while the latter method seeks to maintain the orthogonality via solving differential equations that encode the orthogonality continously.

Computation of the $p > 1$ largest LCEs

If we want to compute the p largest LCEs, with $1 < p < 2N$, we change equation (10) to

$$\dot{\mathbf{Y}}(t) = \mathbf{A}(t)\,\mathbf{Y}(t) \ , \ \text{ with } \ \mathbf{Y}(0)^{\mathrm{T}}\mathbf{Y}(0) = \mathbf{I}_p. \qquad (90)$$

where $\mathbf{Y}(t)$ is in practice, the $2N \times p$ matrix having as columns the p deviation vectors $\mathbf{w}_1(t), \mathbf{w}_2(t), \dots, \mathbf{w}_p(t)$. Applying Theorem 4 we get $\mathbf{Y}(t) = \mathbf{Q}(t)\mathbf{R}(t)$ where $\mathbf{Q}(t)$ is orthogonal so that the identity $\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I}$ holds but not the $\mathbf{Q}\mathbf{Q}^{\mathrm{T}} = \mathbf{I}$. Then from equation (90) we get

$$\dot{\mathbf{R}} = \left(\mathbf{Q}^{\mathrm{T}}\mathbf{A}\mathbf{Q} - \mathbf{S}\right)\mathbf{R}$$

where $\mathbf{S} = \mathbf{Q}^{\mathrm{T}}\dot{\mathbf{Q}}$ is a $p \times p$ matrix whose elements are given by equation (85) for $i, j = 1, 2, \dots, p$. Since $\mathbf{R}$ is invertible, from the relations

$$\dot{\mathbf{R}}\mathbf{R}^{-1} = \mathbf{Q}^{\mathrm{T}}\mathbf{A}\mathbf{Q} - \mathbf{S}$$

and

$$\dot{\mathbf{Q}} = \mathbf{A}\mathbf{Q} - \mathbf{Q}\dot{\mathbf{R}}\mathbf{R}^{-1}, \qquad (91)$$

we obtain

$$\dot{\mathbf{Q}} = \left(\mathbf{A} - \mathbf{Q}\mathbf{Q}^{\mathrm{T}}\mathbf{A} + \mathbf{Q}\mathbf{S}\mathbf{Q}^{\mathrm{T}}\right)\mathbf{Q},$$

or

$$\dot{\mathbf{Q}} = \mathbf{H}(\mathbf{Q}, t)\mathbf{Q}, \tag{92}$$

with

$$\mathbf{H}(\mathbf{Q}, t) = \mathbf{A} - \mathbf{Q}\mathbf{Q}^{\mathrm{T}}\mathbf{A} + \mathbf{Q}\mathbf{S}\mathbf{Q}^{\mathrm{T}}.$$

Notice that the matrix $\mathbf{H}(\mathbf{Q}, t)$ in not necessarily skew–symmetric, and the term $\mathbf{Q}\mathbf{Q}^{\mathrm{T}}$ is responsible for lack of skew–symmetry in $\mathbf{H}$. Of course for $p = 2N$ equation (92) reduces to equation $\dot{\mathbf{Q}} = \mathbf{Q}\mathbf{S}$ (84). The evolution of the diagonal elements of $\mathbf{R}$ are again governed by equation (86), but for $p < 2N$, and so the p largest LCEs can be computed again from equations (87)–(89).

The main difference with respect to the case of the computation of the whole spectrum is the numerical difficulties arising in solving equation (92), since $\mathbf{H}$ is not skew–symmetric as was matrix $\mathbf{S}$ in equation (84). Due to this difference usual numerical integration techniques fail to preserve the orthogonality of matrix $\mathbf{Q}$.

A central observation of [36] is that the matrix $\mathbf{H}$ has a weak skew–symmetry property. The matrix $\mathbf{H}$ is called weak skew–symmetric if

$$\mathbf{Q}^{\mathrm{T}}\left(\mathbf{H}(\mathbf{Q}, t) + \mathbf{H}^{\mathrm{T}}(\mathbf{Q}, t)\right)\mathbf{Q} = \mathbf{0}, \text{ whenever } \mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I}_p.$$

A matrix $\mathbf{H}$ is said to be strongly skew–symmetric if it is skew–symmetric, i. e. $\mathbf{H}^{\mathrm{T}} = -\mathbf{H}$. Christiansen and Rugh [26] proposed a method according to which, the numerically unstable equations (91) for the continuous orthonormalization could be stabilized by the addition of an appropriate dissipation term. This idea was also used in [18], where it was shown that it is possible to reformulate equation (92) so that $\mathbf{H}$ becomes strongly skew–symmetric and thus, achieve a numerically stable algorithm for the computation of few LCEs.

Discrete and continuous methods based on the SVD procedure

An alternative way of evaluating the LCEs is obtained by applying the SVD procedure on the fundamental $2N \times 2N$ matrix $\mathbf{Y}(t)$, which defines the evolution of deviation vectors through equations (9) and (12) for continuous and discrete systems respectively. According to the SVD algorithm a $2N \times p$ matrix ($p \leq 2N$) $\mathbf{B}$ can be written as the product of a $2N \times p$ column–orthogonal matrix $\mathbf{U}$, a $p \times p$ diagonal matrix $\mathbf{F}$ with positive or zero elements σ_i, $i = 1, \ldots, p$ (the so–called *singular values*), and the transpose of a $p \times p$ orthogonal matrix $\mathbf{V}$:

$$\mathbf{B} = \mathbf{U} \cdot \mathbf{F} \cdot \mathbf{V}^{\mathrm{T}}.$$

We note that matrices $\mathbf{U}$ and $\mathbf{V}$ are orthogonal so that:

$$\mathbf{U}^{\mathrm{T}} \cdot \mathbf{U} = \mathbf{V}^{\mathrm{T}} \cdot \mathbf{V} = \mathbf{I}_p. \tag{93}$$

For a more detailed description of the SVD method, as well as an algorithm for its implementation the reader is referred to [107, Section 2.6] and references therein. The SVD is unique up to permutations of corresponding columns, rows and diagonal elements of matrices $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{F}$. Advanced numerical techniques for the computation of the singular values of a product of many matrices can be found for example in [130, 101].

So, for the purposes of our study let

$$\mathbf{Y} = \mathbf{U} \cdot \mathbf{F} \cdot \mathbf{V}^{\mathrm{T}}, \tag{94}$$

where we dropped out as before, the explicit dependence of the matrices on time t. In those cases where all singular values are different, a unique decomposition can be achieved by the additional request of a strictly monotonically decreasing singular value spectrum, i. e. $\sigma_1(t) > \sigma_2(t) > \cdots > \sigma_{2N}(t)$. Multiplying equation (94) with the transpose

$$\mathbf{Y}^{\mathrm{T}} = \mathbf{V} \cdot \mathbf{F}^{\mathrm{T}} \cdot \mathbf{U}^{\mathrm{T}},$$

from the left we get

$$\mathbf{Y}^{\mathrm{T}} \cdot \mathbf{Y} = \mathbf{V} \cdot \mathbf{F}^{\mathrm{T}} \cdot \mathbf{U}^{\mathrm{T}} \cdot \mathbf{U} \cdot \mathbf{F} \cdot \mathbf{V}^{\mathrm{T}} = \mathbf{V} \cdot \mathrm{diag}(\sigma_i^2(t)) \cdot \mathbf{V}^{\mathrm{T}}, \tag{95}$$

where equation (93) has been used. From equation (95) we see that the eigenvalues of the diagonal matrix $\mathrm{diag}(\sigma_i^2(t))$, i. e. the squares of the singular values of $\mathbf{Y}(t)$, are equal to the eigenvalues of the symmetric matrix $\mathbf{Y}^{\mathrm{T}}\mathbf{Y}$. Then from point 4 of the MET we conclude that the LCEs are related to the singular values of $\mathbf{Y}(t)$ through [62, 130]

$$\chi_p = \lim_{t \to \infty} \frac{1}{t} \ln \sigma_i(t), \ \ p = 1, 2, \ldots, 2N,$$

which implies that the LCEs can be evaluated as the limits for $t \to \infty$ of the time rate of the logarithms of the singular values.

Theoretical aspects of the SVD technique, as well as a detailed study of its ability to approximate the spectrum of LCEs can be found in [101, 37, 38]. Continuous [67, 62, 39] and discrete [130] versions of the SVD algorithm have been applied for the computation of few or of all LCEs, although this approach is not widely used. A basic problem of these methods is that they fail to compute the spectrum of LCEs if it is degenerate, i. e. when two or more LCEs are equal or very close to each other, due to the appearance of ill–conditioned matrices.

7 Chaos detection techniques

A simple, qualitative way of studying the dynamics of a Hamiltonian system is by plotting the successive intersections of its orbits with a Poincaré surface

of section (PSS) (e. g. [72] [92, p. 17–20]). Similarly, in the case of symplectic maps one simply plots the phase space of the system. This qualitative method has been extensively applied to 2d maps and to 2D Hamiltonians, since in these systems the PSS is a 2–dimensional plane. In such systems one can visually distinguish between regular and chaotic orbits since the points of a regular orbit lie on a torus and form a smooth closed curve, while the points of a chaotic orbit appear randomly scattered. In 3D Hamiltonian systems (or 4d symplectic maps) however, the PSS (or the phase space) is 4–dimensional and the behavior of the orbits cannot be easily visualized. Things become even more difficult and deceiving for multidimensional systems. One way to overcome this problem is to project the PSS (or the phase space) to spaces with lower dimensions (see e.g. [139, 140, 105]) although these projections are often very complicated and difficult to interpret. Thus, we need fast and accurate numerical tools to give us information about the regular or chaotic character of orbits, mainly when the dynamical system has many degrees of freedom.

The most commonly employed method for distinguishing between regular and chaotic behavior is the evaluation of the mLCE χ_1, because if $\chi_1 > 0$ the orbit is chaotic. The main problem of using the value of χ_1 as an indicator of chaoticity is that, in practice, the numerical computation may take a huge amount of time, in particular for orbits which stick to regular ones for a long time before showing their chaotic behavior. Since χ_1 is defined as the limit for $t \to \infty$ of the quantity $X_1(t)$ (54), the time needed for $X_1(t)$ to converge to its limiting value is not known a priori and may become extremely long. Nevertheless, we should keep in mind that the mLCE gives us more information than just characterizing an orbit as regular or chaotic, since it also quantifies the notion of chaoticity by providing a characteristic time scale for the studied dynamical system, namely the Lyapunov time (51).

In order to address the problem of the fast and reliable determination of the regular or chaotic nature of orbits, several methods have been developed over the years with varying degrees of success. These methods can be divided in two major categories: Some are based on the study of the evolution of deviation vectors from a given orbit, like the computation of χ_1, while others rely on the analysis of the particular orbit itself.

Among other chaoticity detectors, belonging to the same category with the evaluation of the mLCE, are the fast Lyapunov indicator (FLI) [58, 59, 56, 89, 49, 69] and its variants [4, 5], the smaller alignment index (SALI) [122, 124, 125] and its generalization, the so–called generalized alignment index (GALI) [126, 127], the mean exponential growth of nearby orbits (MEGNO) [28, 29], the relative Lyapunov indicator (RLI) [115, 116], the average power law exponent (APLE) [95], as well as methods based on the study of spectra of quantities related to the deviation vectors like the stretching numbers [57, 93, 135, 138], the helicity angles (the angles of deviation vectors with a fixed direction) [32], the twist angles (the differences of two succes-

sive helicity angles) [33], or the study of the differences between such spectra [88, 136].

In the category of methods based on the analysis of a time series constructed by the coordinates of the orbit under study, one may list the frequency map analysis of Laskar [83, 86, 84, 85], the '0–1' test [64, 65], the method of the low frequency spectral analysis [137, 81], the 'patterns method' [120, 121], the recurrence plots technique [147, 148] and the information entropy index [100]. One could also refer to several ideas presented by various authors that could be used in order to distinguish between chaoticity and regularity, like the differences appearing for regular and chaotic orbits in the time evolutions of their correlation dimension [50], in the time averages of kinetic energies related to the virial theorem [74] and in the statistical properties of the series of time intervals between successive intersections of orbits with a PSS [80].

A systematic and detailed comparative study of the efficiency and reliability of the various chaos detection techniques has not been done yet, although comparisons between some of the existing methods have been performed sporadically in studies of particular dynamical systems [122, 125, 132, 133, 82, 95, 6].

Let us now focus our attention on the behavior of the FLI and of the GALI and on their connection to the LCEs. The FLI was introduced as an indicator of chaos in [58, 59] and after some minor modifications in its definition, it was used for the distinction between resonant and not resonant regular motion [56, 49]. The FLI is defined as

$$\text{FLI}(t) = \sup_t \ \ln \|\mathbf{w}(t)\|,$$

where $\mathbf{w}(t)$ is a deviation vector from the studied orbit at point $\mathbf{x}(t)$, which initially had unit norm, i. e. $\|\mathbf{w}(0)\| = 1$. In practice, $\text{FLI}(t)$ registers the maximum length that an initially unitary deviation vector attains from the beginning of its evolution up to the current time t. Using the notation appearing in equation (59), the FLI can be computed as

$$\text{FLI}(k\tau) = \sup_k \ \sum_{i=1}^{k} \ln \frac{D_i}{D_0} = \sup_k \ \sum_{i=1}^{k} \ln \alpha_i,$$

with the initial norm D_0 of the deviation vector being $D_0 = 1$.

According to equation (62) the norm of $\mathbf{w}(t)$ increases linearly in time in the case of regular orbits. On the other hand, in the case of chaotic orbits the norm of any deviation vector exhibits an exponential increase in time, with an exponent which approximates χ_1 for $t \to \infty$. Thus, the norm of a deviation vector reaches rapidly completely different values for regular and chaotic orbits, which actually differ by many orders of magnitude. This behavior allows FLI to discriminate between regular orbits, for which FLI has relatively small values, and chaotic orbits, for which FLI gets very large values.

The main difference of FLI with respect to the evaluation of the mLCE by equation (59) is that FLI registers the *current* value of the norm of the

deviation vector and does not try to compute the limit value, for $t \to \infty$, of the mean of stretching numbers as χ_1 does. By dropping the time average requirement of the stretching numbers, FLI succeeds in determining the nature of orbits faster than the computation of the mLCE.

The generalized alignment index of order p (GALI_p) is determined through the evolution of $2 \leq p \leq 2N$ initially linearly independent deviation vectors $\mathbf{w}_i(0)$, $i = 1, 2, \ldots, p$ and so it is more related to the computation of many LCEs than to the computation of the mLCE. The evolved deviation vectors $\mathbf{w}_i(t)$ are normalized from time to time in order to avoid overflow problems, but their directions are left intact. Then, according to [126] GALI_p is defined to be the volume of the p–parallelogram having as edges the p unitary deviation vectors $\hat{\mathbf{w}}_i(t)$, $i = 1, 2, \ldots, p$

$$\mathrm{GALI}_p(t) = \|\hat{\mathbf{w}}_1(t) \wedge \hat{\mathbf{w}}_2(t) \wedge \cdots \wedge \hat{\mathbf{w}}_p(t)\|. \qquad (96)$$

In [126] the value of GALI_p is computed according to equation (105), while in [2, 127] a more efficient numerical technique based on the SVD algorithm is applied. From the definition of GALI_p it becomes evident that if at least two of the deviation vectors become linearly dependent, the wedge product in (96) becomes zero and the GALI_p vanishes.

In the case of a chaotic orbit all deviation vectors tend to become linearly dependent, aligning in the direction which corresponds to the mLCE and GALI_p tends to zero exponentially following the law [126]:

$$\mathrm{GALI}_p(t) \sim e^{-[(\chi_1 - \chi_2) + (\chi_1 - \chi_3) + \cdots + (\chi_1 - \chi_p)]t},$$

where $\chi_1, \chi_2, \ldots, \chi_p$ are the p largest LCEs. On the other hand, in the case of regular motion all deviation vectors tend to fall on the N–dimensional tangent space of the torus on which the motion lies. Thus, if we start with $p \leq N$ general deviation vectors they will remain linearly independent on the N–dimensional tangent space of the torus, since there is no particular reason for them to become linearly dependent. As a consequence GALI_p remains practically constant for $p \leq N$. On the other hand, GALI_p tends to zero for $p > N$, since some deviation vectors will eventually become linearly dependent, following a particular power law which depends on the dimensionality N of the torus and the number p of deviation vectors. So, the generic behavior of GALI_p for regular orbits lying on N–dimensional tori is given by [126]

$$\mathrm{GALI}_p(t) \sim \begin{cases} \text{constant if } 2 \leq p \leq N \\ \frac{1}{t^{2(p-N)}} \quad \text{if } N < p \leq 2N \end{cases}. \qquad (97)$$

The different behavior of GALI_p for regular orbits, where it remains different from zero or tends to zero following a power law, and for chaotic orbits, where it tends exponentially to zero, makes GALI_p an ideal indicator of chaoticity independent of the dimensions of the system [126, 127, 15]. GALI_p is a generalization of the SALI method [122, 124, 125] which is related to the

evolution of only two deviation vectors. Actually $GALI_2 \propto SALI$. However, $GALI_p$ provides significantly more detailed information on the local dynamics, and allows for a faster and clearer distinction between order and chaos. It was shown recently [27, 127] that $GALI_p$ can also be used for the determination of the dimensionality of the torus on which regular motion occurs.

As we discussed in Section 6.1 the alignment of all deviation vectors to the direction corresponding to the mLCE is a basic problem for the computation of many LCEs, which is overcome by successive orthonormalizations of the set of deviation vectors. The GALIs on the other hand, exploit exactly this 'problem' in order to determine rapidly and with certainty the regular or chaotic nature of orbits.

It was shown in Section 4.1 that the values of all LCEs (and therefore the value of the mLCE) do not depend on the particular used norm. On the other hand, the quantitative results of all chaos detection techniques based on quantities related to the dynamics of the tangent space on a finite time, depend on the used norm, or on the coordinates of the studied system. For example, the actual values of the finite time mLCE $X_1(t)$ (54) will be different for different norms, or for different coordinates, although its limiting value for $t \to \infty$, i. e. the mLCE χ_1, will be always the same. Other chaos detection methods, like the FLI and the GALI, which depend on the current values of some norm–related quantities and not on their limiting values for $t \to \infty$, will attain different values for different norms and/or coordinate systems. Although the values of these indices will be different, one could expect that their qualitative behavior would be independent of the chosen norm and the used coordinates, since these indices depend on the geometrical properties of the deviation vectors. For example, the GALI quantifies the linear dependence or independence of deviation vectors, a property which obviously does not depend on the particular used norm or coordinates. Indeed, some arguments explaining the independence of the behavior of the GALI method on the chosen coordinates can be found in [126]. Nevertheless, a systematic study focused on the influence of the used norm on the qualitative behavior of the various chaos indicators has not been performed yet, although it would be of great interest.

8 LCEs of dissipative systems and time series

The presentation of the LCEs in this report was mainly done in connection to conservative dynamical systems, i. e. autonomous Hamiltonian flows and symplectic maps. The restriction to conservative systems is not necessary since the theory of LCEs, as well as the techniques for their evaluation are valid for general dynamical systems like for example dissipative ones. In addition, within what is called time series analysis (see e.g. [78]) it is of great interest to measure LCEs in order to understand the underlying dynamics that produces any time series of experimental data. For the completeness of our presentation

we devote the last section of our report to a concise survey of results concerning the LCEs of dissipative systems and time series.

8.1 Dissipative systems

In contrast to Hamiltonian systems and symplectic maps for which the conservation of the phase space volume is a fundamental constraint of the motion, a dissipative system is characterized by a decrease of the phase space volume with increasing time. This leads to the contraction of motion on a surface of lower dimensionality than the original phase space, which is called *attractor*. Thus any dissipative dynamical system will have at least one negative LCE, the sum of all its LCEs (which actually measures the contraction rate of the phase space volume through equation (43)) is negative and after some initial transient time the motion occurs on an attractor.

Any continuous n–dimensional dissipative dynamical system without a stationary point (which is often called a *fixed point*) has at least one LCE equal to zero [70] as we have already discussed in Section 4.5. For regular motion the attractor of dissipative flows represents a fixed point having all its LCEs negative, or a quasiperiodic orbit lying on a p–dimensional torus ($p < n$) having p zero LCEs while the rest $n - p$ exponents are negative. For dissipative flows in three or more dimensions there can also exist attractors having a very complicated geometrical structure which are called 'strange'.

Strange attractors have one or more positive LCEs implying that the motion on them is chaotic. The exponential expansion indicated by a positive LCE is incompatible with motion on a bounded attractor unless some sort of folding process merges separated orbits. Each positive exponent corresponds to a direction in which the system experiences the repeated stretching and folding that decorrelates nearby orbits on the attractor. A simple geometrical construction of a hypothetical strange attractor where orbits are bounded despite the fact that nearby orbits diverge exponentially can be found in [92, Sect. 1.5].

The numerical methods for the evaluation of the mLCE, of the p ($1 < p < n$) largest LCEs and of the whole spectrum of them, presented in Sections 5 and 6, can be applied also to dissipative systems. Actually, many of these techniques were initially used in studies of dissipative models [99, 119, 61, 62]. For a detailed description of the dynamical features of dissipative systems, as well as of the behavior of LCEs for such systems the reader is referred, for example, to [103, 44] [92, Sect. 1.5, Chapt. 7 and 8] and references therein.

8.2 Computing LCEs from a time series

A basic task in real physical experiments is the understanding of the dynamical properties of the studied system by the analysis of some observed time series of data. The knowledge of the LCEs of the system is one important step towards the fulfillment of this goal. Usually, we have no knowledge of the nonlinear

equations that govern the time evolution of the system which produces the experimental data. This lack of information makes the computation of the spectrum of LCEs of the system a hard and challenging task.

The methods developed for the determination of the LCEs from a scalar time series have as starting point the technique of *phase space reconstruction* with *delay coordinates* [104, 134, 112] [78, Chapt. 3 and 9]. This technique is used for recreating a d–dimensional phase space to capture the behavior of the dynamical system which produces the observed scalar time series.

Assume that we have N_D measurements of a dynamical quantity x taken at times $t_n = t_0 + n\tau$, i. e. $x(n) \equiv x(t_0 + n\tau)$, $n = 0, 1, 2, \ldots, N_D - 1$. Then we produce $N_d = N_D - (d-1)T$ d–dimensional vectors $\mathbf{x}(t_n)$ from the x's as

$$\mathbf{x}(t_n) = \big[\, x(n)\ x(n+T)\ \ldots\ x(n + (d-1)T) \,\big]^{\mathrm{T}},$$

where T is the (integer) delay time. With this procedure we construct N_d points in a d–dimensional phase space, which can be treated as successive points of a hypothetical orbit. We assume that the evolution of $\mathbf{x}(t_n)$ to $\mathbf{x}(t_{n+1})$ is given by some map and we seek to evaluate the LCEs of this orbit.

The first algorithm to compute LCEs for a time series was introduced by Wolf et al. [144]. According to their method (which is also referred as the *direct method*), in order to compute the mLCE we first locate the nearest neighbor (in the Euclidean sense) $\mathbf{x}(t_k)$, to the initial point $\mathbf{x}(t_0)$ and define the corresponding deviation vector $\mathbf{w}(t_0) = \mathbf{x}(t_0) - \mathbf{x}(t_k)$ and its length $L(t_0) = \|\mathbf{w}(t_0)\|$. The points $\mathbf{x}(t_0)$ and $\mathbf{x}(t_k)$ are considered as initial conditions of two nearby orbits and are followed in time. Then the mLCE is evaluated by the method discussed in Section 5.2, which approximates deviation vectors by differences of nearby orbits. So, at some later time t_{m_1} (which is fixed a priori or determined by some predefined threshold violation of the vector's length) the evolved deviation vector $\mathbf{w}'(t_{m_1}) = \mathbf{x}(t_{m_1}) - \mathbf{x}(t_{k+m_1})$ is normalized and its length $L'(t_{m_1}) = \|\mathbf{w}'(t_{m_1})\|$ is registered. The 'normalization' of the evolved deviation vector is done by looking for a new data point, say $\mathbf{x}(t_l)$, whose distance $L(t_{m_1}) = \|\mathbf{x}(t_{m_1}) - \mathbf{x}(t_l)\|$ from the studied orbit is small and the corresponding deviation vector $\mathbf{w}(t_{m_1}) = \mathbf{x}(t_{m_1}) - \mathbf{x}(l)$ has the same direction with $\mathbf{w}'(t_{m_1})$. Of course with finite amount of data, one cannot hope to find a replacement point $\mathbf{x}(l)$ which falls exactly on the direction of $\mathbf{w}'(t_{m_1})$ but chooses a point that comes as close as possible. Assuming that such point is found the procedure is repeated and an estimation $X_1(t_{m_n})$ of the mLCE χ_1 is obtained by an equation analogous to equation (56):

$$X_1(t_{m_n}) = \frac{1}{t_{m_n} - t_0} \sum_{i=1}^{n} \ln \frac{L'_1(t_{m_i})}{L(t_{m_{i-1}})},$$

with $m_0 = 0$. A Fortran code of this algorithm with fixed time steps between replacements of deviation vectors is given in [144].

Generalizing this technique by evolving simultaneously $p > 1$ deviation vectors, i. e. following the evolution of the orbit under study, as well as of p

nearby orbits, we can, in principle, evaluate the p–mLCE $\chi_1^{(p)}$ of the system, which is equal to the sum of the p largest 1–LCEs (see equation (67)). Then the values of χ_i $i = 1, 2, \ldots, p$ can be computed from equation (74). This procedure corresponds to a variant of the standard method for computing the LCEs, presented in [119] and discussed in Section 6.1, in that deviation vectors are defined as differences of neighboring orbits. The implementation of this approach requires the repeated replacement of the deviation vectors, i. e. the replacement of the p points close to the evolved orbit under consideration, when the lengths of the vectors exceed some threshold value. This replacement should be done in a way that the volume of the corresponding p–parallelogram is small, and in particular smaller than the replaced volume, and the new p vectors point more or less to the same direction like the old ones. This procedure is explained in detail in [144] for the particular case of the computation of $\chi_1^{(2)} = \chi_1 + \chi_2$, where a triplet of points is involved.

It is clear that in order to achieve a good replacement of the evolved p vectors, which will lead to a reliable estimation of the LCEs, the numerical data have to satisfy many conditions. Usually this is not feasible due to the limited number of data points. So the direct method of [144] does not yield very precise results for the LCEs. Another limitation of the method, which was pointed out in Wolf et al. [144], is that it should not be used for finding negative LCEs which correspond to shrinking directions, due to a cut off in small distances implied mainly by the level of noise of the experimental data. An additional disadvantage of the direct method is that many parameters which influence the estimated values of the LCEs like the embedding dimension d, the delay time T, the tolerances in direction angles during vector replacements and the evolution times between replacements, have to be tuned properly in order to obtain reliable results.

A different approach for the computation of the whole spectrum of LCEs is based on the numerical determination of matrix $\mathbf{Y}_n$, $n = 1, 2, \ldots$, of equation (12), which defines the evolution of deviation vectors in the reconstructed phase space. This method was introduced in [118] and was studied in more detail in [44, 45] (see also [78, Chapt. 11]). According to this approach, often called the *tangent space method*, matrix $\mathbf{Y}_n$ is evaluated for each point of the studied orbit through local linear fits of the data. In particular, for every point $\mathbf{x}(t_n)$ of the orbit we find all its neighboring points, i. e. points whose distance from $\mathbf{x}(t_n)$ is less than a predefined small value ϵ. Each of these point define a deviation vector. Then we find the next iteration of all these points and see how these vectors evolve. Keeping only the evolved vectors having length less than ϵ we evaluate matrix $\mathbf{Y}_n$ through a least–square–error algorithm. By repeating this procedure for the whole length of the studied orbit we are able to evaluate at each point of the orbit matrix $\mathbf{Y}_n$ which defines the evolution of deviation vectors over one time step. Then by applying the QR decomposition version of the standard method, which was presented in Section 6.3, we estimate the values of the LCEs. The corresponding algorithm

is included in the TISEAN software package of nonlinear time series analysis methods developed by Hegger et al. [71]. It is also worth mentioning that Brown et al. [17] improved the tangent space method by using higher order polynomials for the local fit.

If, on the other hand, we are interested only in the evaluation of the mLCE of a time series we can apply the algorithm proposed by Rosenstein et al. [111] and Kantz [77]. The method is based on the statistical study of the evolution of distances of neighboring orbits. This approach is in the same spirit of Wolf et al. [144] although being simpler since it compares distances and not directions. A basic difference with the direct method is that for each point of the reference orbit not one, but several neighboring orbits are evaluated leading to improved estimates of the mLCE with smaller statistical fluctuations even in the case of small data sets. This algorithm is also included in the TISEAN package [71], while its Fortran and C codes can be found in [78, Appendix B].

Acknowledgments

Appendix

A Exterior algebra and wedge product: Some basic notions

We present here some basic results of the exterior algebra theory along with an introduction to the theory of wedge products following [1] and textbooks such as [128, 68, 129]. We also provide some simple illustrative examples of these results.

Let us consider an M–dimensional vector space V over the field of real numbers $\mathbb{R}$. The *exterior algebra* of V is denoted by $\Lambda(V)$ and its multiplication, known as the *wedge product* or the *exterior product*, is written as $\wedge$. The wedge product is associative:

$$(\mathbf{u} \wedge \mathbf{v}) \wedge \mathbf{w} = \mathbf{u} \wedge (\mathbf{v} \wedge \mathbf{w})$$

for $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and bilinear

$$(c_1\mathbf{u} + c_2\mathbf{v}) \wedge \mathbf{w} = c_1(\mathbf{u} \wedge \mathbf{w}) + c_2(\mathbf{v} \wedge \mathbf{w}),$$
$$\mathbf{w} \wedge (c_1\mathbf{u} + c_2\mathbf{v}) = c_1(\mathbf{w} \wedge \mathbf{u}) + c_2(\mathbf{w} \wedge \mathbf{v}),$$

for $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $c_1, c_2 \in \mathbb{R}$. The wedge product is also alternating on V

$$\mathbf{u} \wedge \mathbf{u} = \mathbf{0}$$

for all vectors $\mathbf{u} \in V$. Thus we have that

$$\mathbf{u} \wedge \mathbf{v} = -\mathbf{v} \wedge \mathbf{u}$$

for all vectors $\mathbf{u}, \mathbf{v} \in V$ and

$$\mathbf{u}_1 \wedge \mathbf{u}_2 \wedge \cdots \wedge \mathbf{u}_k = \mathbf{0} \tag{98}$$

whenever $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k \in V$ are linearly dependent. Elements of the form $\mathbf{u}_1 \wedge \mathbf{u}_2 \wedge \cdots \wedge \mathbf{u}_k$ with $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k \in V$ are called k–*vectors*. The subspace of $\Lambda(V)$ generated by all k–vectors is called the k–*th exterior power of* V and denoted by $\Lambda^k(V)$.

Let $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \ldots, \hat{\mathbf{e}}_M\}$ be an orthonormal basis of V, i. e. $\hat{\mathbf{e}}_i$, $i = 1, 2, \ldots, M$ are linearly independent vectors of unit magnitude and

$$\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j = \delta_{ij}$$

where '$\cdot$' denotes the inner product in V and

$$\delta_{ij} = \begin{cases} 1 \text{ for } i = j \\ 0 \text{ for } i \neq j \end{cases}.$$

It can be easily seen that the set

$$\{\hat{\mathbf{e}}_{i_1} \wedge \hat{\mathbf{e}}_{i_2} \wedge \cdots \wedge \hat{\mathbf{e}}_{i_k} \mid 1 \leq i_1 < i_2 < \cdots < i_k \leq M\} \tag{99}$$

is a basis of $\Lambda^k(V)$ since any wedge product of the form $\mathbf{u}_1 \wedge \mathbf{u}_2 \wedge \cdots \wedge \mathbf{u}_k$ can be written as a linear combination of the k–vectors of equation (99). This is true because every vector $\mathbf{u}_i$, $i = 1, 2, \ldots, k$ can be written as a linear combination of the basis vectors $\hat{\mathbf{e}}_i$, $i = 1, 2, \ldots, M$ and using the bilinearity of the wedge product this can be expanded to a linear combination of wedge products of those basis vectors. Any wedge product in which the same basis vector appears more than once is zero, while any wedge product in which the basis vectors do not appear in the proper order can be reordered, changing the sign whenever two basis vectors change places. The dimension d_k of $\Lambda^k(V)$ is equal to the binomial coefficient

$$d_k = \dim \Lambda^k(V) = \binom{M}{k} = \frac{M!}{k!(M - k)!}.$$

Ordering the elements of basis (99) of $\Lambda^k(V)$ according to the standard *lexicographical order*

$$\boldsymbol{\omega}_i = \hat{\mathbf{e}}_{i_1} \wedge \hat{\mathbf{e}}_{i_2} \wedge \cdots \wedge \hat{\mathbf{e}}_{i_k}, \ \ 1 \le i_1 < i_2 < \cdots < i_k \le M, \ \ i = 1, 2, \cdots, d_k, \ \ (100)$$

any k–vector $\bar{\mathbf{u}} \in \Lambda^k(V)$ can be represented as

$$\bar{\mathbf{u}} = \sum_{i=1}^{d_k} \bar{u}_i \boldsymbol{\omega}_i \ , \ \ \bar{u}_i \in \mathbb{R}. \tag{101}$$

A k–vector which can be written as the wedge product of k linear independent vectors of V is called *decomposable*. Of course, if the k vectors are linearly dependent we get the zero k–vector (98). Note that not all k–vectors are decomposable. For example the 2–vector $\bar{\mathbf{u}} = \mathbf{e}_1 \wedge \mathbf{e}_2 + \mathbf{e}_3 \wedge \mathbf{e}_4 \in \Lambda^2(\mathbb{R}^4)$ is not decomposable as it cannot be written as $\mathbf{u}_1 \wedge \mathbf{u}_2$ with $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{R}^4$.

Let us consider a decomposable k–vector $\bar{\mathbf{u}} = \mathbf{u}_1 \wedge \mathbf{u}_2 \wedge \cdots \wedge \mathbf{u}_k$. Then the coefficients $\bar{u}_i$ in (101) are the minors of matrix $\mathbf{U}$ having as columns the coordinates of vectors $\mathbf{u}_i$, $i = 1, 2, \ldots, k$ with respect to the orthonormal basis $\hat{\mathbf{e}}_i$, $i = 1, 2, \ldots, M$. In matrix form we have

$$\begin{bmatrix} \mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_k \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{e}}_1 \ \hat{\mathbf{e}}_2 \ \cdots \ \hat{\mathbf{e}}_M \end{bmatrix} \cdot \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1k} \\ u_{21} & u_{22} & \cdots & u_{2k} \\ \vdots & \vdots & & \vdots \\ u_{M1} & u_{M2} & \cdots & u_{Mk} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{e}}_1 \ \hat{\mathbf{e}}_2 \ \cdots \ \hat{\mathbf{e}}_M \end{bmatrix} \cdot \mathbf{U}$$

$$(102)$$

where u_{ij}, $i = 1, 2, \ldots, M$, $j = 1, 2, \ldots, k$ are real numbers. Then, the wedge product $\mathbf{u}_1 \wedge \mathbf{u}_2 \wedge \cdots \wedge \mathbf{u}_k$ is written as

$$\bar{\mathbf{u}} = \mathbf{u_1} \wedge \mathbf{u}_2 \wedge \cdots \wedge \mathbf{u}_k = \sum_{i=1}^{d_k} \bar{u}_i \boldsymbol{\omega}_i =$$

$$\sum_{1 \le i_1 < i_2 < \cdots < i_k \le M} \begin{vmatrix} u_{i_1 1} & u_{i_1 2} & \cdots & u_{i_1 k} \\ u_{i_2 1} & u_{i_2 2} & \cdots & u_{i_2 k} \\ \vdots & \vdots & & \vdots \\ u_{i_k 1} & u_{i_k 2} & \cdots & u_{i_k k} \end{vmatrix} \hat{\mathbf{e}}_{i_1} \wedge \hat{\mathbf{e}}_{i_2} \wedge \cdots \wedge \hat{\mathbf{e}}_{i_k}, \tag{103}$$

where the sum is performed over all possible combinations of k indices out of the M total indices and $|\ |$ denotes the determinant. So, the coefficient of a particular k–vector $\hat{\mathbf{e}}_{i_1} \wedge \hat{\mathbf{e}}_{i_2} \wedge \cdots \wedge \hat{\mathbf{e}}_{i_k}$ is the determinant of the $k \times k$ submatrix of the $M \times k$ matrix of coefficients appearing in equation (102) formed by its i_1, i_2, ..., i_k rows.

The inner product on V induces an *inner product* on each vector space $\Lambda^k(V)$ as follows: Considering two decomposable k–vectors

$$\bar{\mathbf{u}} = \mathbf{u_1} \wedge \mathbf{u}_2 \wedge \cdots \wedge \mathbf{u}_k \ \text{ and } \ \bar{\mathbf{v}} = \mathbf{v_1} \wedge \mathbf{v}_2 \wedge \cdots \wedge \mathbf{v}_k,$$

with $\mathbf{u}_i, \mathbf{v}_j \in V$, $i, j = 1, 2, \ldots, k$, the inner product of $\bar{\mathbf{u}}, \bar{\mathbf{v}} \in \Lambda^k(V)$ is defined by

$$\langle \bar{\mathbf{u}}, \bar{\mathbf{v}} \rangle_k \stackrel{\text{def}}{=} \begin{vmatrix} \mathbf{u}_1 \cdot \mathbf{v}_1 & \mathbf{u}_1 \cdot \mathbf{v}_2 & \cdots & \mathbf{u}_1 \cdot \mathbf{v}_k \\ \mathbf{u}_2 \cdot \mathbf{v}_1 & \mathbf{u}_2 \cdot \mathbf{v}_2 & \cdots & \mathbf{u}_2 \cdot \mathbf{v}_k \\ \vdots & \vdots & & \vdots \\ \mathbf{u}_k \cdot \mathbf{v}_1 & \mathbf{u}_k \cdot \mathbf{v}_2 & \cdots & \mathbf{u}_k \cdot \mathbf{v}_k \end{vmatrix} = \left| \mathbf{U}^\mathrm{T} \cdot \mathbf{V} \right| \tag{104}$$

where $\mathbf{U}$, $\mathbf{V}$ are matrices having as columns the coefficients of vectors $\mathbf{u}_i$, $\mathbf{v}_i$, $i = 1, 2, \ldots, k$ with respect to the orthonormal $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \ldots, \hat{\mathbf{e}}_M\}$ (see equation (102)). Since every element of $\Lambda^k(V)$ is a sum of decomposable element, this definition extends by bilinearity to any k–vector. Obviously for the basis (100) of $\Lambda^k(V)$ we have

$$\langle \boldsymbol{\omega}_i, \boldsymbol{\omega}_j \rangle_k = \delta_{ij} \ , \ i, j = 1, 2, \ldots, d_k,$$

implying that the basis is orthonormal. Inner product (104) defines a *norm* $\| \ \|$ for k–vectors by

$$\|\bar{\mathbf{u}}\| = \sqrt{\langle \bar{\mathbf{u}}, \bar{\mathbf{u}} \rangle_k} = \sqrt{\left| \mathbf{U}^\mathrm{T} \cdot \mathbf{U} \right|}.$$

Thus, the norm of a decomposable k–vector (103) is given by

$$\|\bar{\mathbf{u}}\| = \|\mathbf{u_1} \wedge \mathbf{u_2} \wedge \cdots \wedge \mathbf{u}_k\| = \sqrt{\left| \mathbf{U}^\mathrm{T} \cdot \mathbf{U} \right|} = \left(\sum_{i=1}^{d_k} \bar{u}_i^2 \right)^{1/2} =$$

$$\left\{ \sum_{1 \le i_1 < i_2 < \cdots < i_k \le M} \begin{vmatrix} u_{i_1 1} & u_{i_1 2} & \cdots & u_{i_1 k} \\ u_{i_2 1} & u_{i_2 2} & \cdots & u_{i_2 k} \\ \vdots & \vdots & & \vdots \\ u_{i_k 1} & u_{i_k 2} & \cdots & u_{i_k k} \end{vmatrix}^2 \right\}^{1/2}, \tag{105}$$

and it measures the volume $\mathrm{vol}(P_k)$ of the k–parallelogram P_k having as edges the k vectors $\mathbf{u_1}, \mathbf{u_2}, \cdots, \mathbf{u}_k$, since this volume is defined as (see e. g. [75, p. 472])

$$\mathrm{vol}(P_k) = \sqrt{\left| \mathbf{U}^\mathrm{T} \cdot \mathbf{U} \right|} \ . \tag{106}$$

The use of a different orthonormal basis does not change the numerical value of $\mathrm{vol}(P_k)$. This can be easily seen as follows: Let $\hat{\mathbf{f}}_i$, $i = 1, 2, \cdots, M$ be a different orthonormal basis of V related to basis $\hat{\mathbf{e}}_i$ through

$$\begin{bmatrix} \hat{\mathbf{e}}_1 & \hat{\mathbf{e}}_2 & \cdots & \hat{\mathbf{e}}_M \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{f}}_1 & \hat{\mathbf{f}}_2 & \cdots & \hat{\mathbf{f}}_M \end{bmatrix} \cdot \mathbf{A}$$

where $\mathbf{A}$ is an *orthogonal* $M \times M$ matrix, i. e. $\mathbf{A}^{-1} = \mathbf{A}^\mathrm{T}$. From equation (102) we get

$$\begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{f}}_1 & \hat{\mathbf{f}}_2 & \cdots & \hat{\mathbf{f}}_M \end{bmatrix} \cdot \mathbf{A} \cdot \mathbf{U},$$

whence the volume $\mathrm{vol}'(P_k)$ with respect to the new basis $\hat{\mathbf{f}}_i$, $i = 1, 2, \cdots, M$ is given by

$$\mathrm{vol}'(P_k) = \sqrt{\left|(\mathbf{A} \cdot \mathbf{U})^{\mathrm{T}} \cdot \mathbf{A} \cdot \mathbf{U}\right|} = \sqrt{\left|\mathbf{U}^{\mathrm{T}} \cdot \mathbf{A}^{-1} \cdot \mathbf{A} \cdot \mathbf{U}\right|} = \sqrt{\left|\mathbf{U}^{\mathrm{T}} \cdot \mathbf{U}\right|} = \mathrm{vol}(P_k),$$

where the orthogonality of $\mathbf{A}$ was used. This result is not surprising since an orthogonal matrix corresponds to a rotation that leaves unchanged the norms of vectors and the angles between them.

Finally we note that the equality

$$\left|\mathbf{U}^{\mathrm{T}}\mathbf{U}\right| = \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq M} \begin{vmatrix} u_{i_1 1} & u_{i_1 2} & \cdots & u_{i_1 k} \\ u_{i_2 1} & u_{i_2 2} & \cdots & u_{i_2 k} \\ \vdots & \vdots & & \vdots \\ u_{i_k 1} & u_{i_k 2} & \cdots & u_{i_k k} \end{vmatrix}^2$$

appearing in equation (105) is the so–called *Lagrange identity* (e. g. [68, p. 108], [16, p. 103]).

A.1 An illustrative example

In order to illustrate the content of the previous section we consider here a specific example. Let V be the vector space of $M = 4$–dimensional real vectors, i. e. $V = \mathbb{R}^4$ and

$$\hat{\mathbf{e}}_1 = (1,0,0,0) \ , \ \hat{\mathbf{e}}_2 = (0,1,0,0) \ , \ \hat{\mathbf{e}}_3 = (0,0,1,0) \ , \ \hat{\mathbf{e}}_4 = (0,0,0,1) \ , \quad (107)$$

the usual orthonormal basis of $\mathbb{R}^4$. Then the lexicographically ordered orthonormal basis (100) of the $d_2 = 6$–dimensional vector space $\Lambda^2(\mathbb{R}^4)$ is

$$\begin{aligned} \boldsymbol{\omega}_1 &= \hat{\mathbf{e}}_1 \wedge \hat{\mathbf{e}}_2 \ , \ \ \boldsymbol{\omega}_2 = \hat{\mathbf{e}}_1 \wedge \hat{\mathbf{e}}_3 \ , \ \ \boldsymbol{\omega}_3 = \hat{\mathbf{e}}_1 \wedge \hat{\mathbf{e}}_4 \ , \\ \boldsymbol{\omega}_4 &= \hat{\mathbf{e}}_2 \wedge \hat{\mathbf{e}}_3 \ , \ \ \boldsymbol{\omega}_5 = \hat{\mathbf{e}}_2 \wedge \hat{\mathbf{e}}_4 \ , \ \ \boldsymbol{\omega}_6 = \hat{\mathbf{e}}_3 \wedge \hat{\mathbf{e}}_4 \ . \end{aligned} \quad (108)$$

The $\Lambda^3(\mathbb{R}^3)$ vector space has dimension $d_3 = 4$ and the set

$$\begin{aligned} \mathbf{y}_1 &= \hat{\mathbf{e}}_1 \wedge \hat{\mathbf{e}}_2 \wedge \hat{\mathbf{e}}_3 \ , \ \ \mathbf{y}_2 = \hat{\mathbf{e}}_1 \wedge \hat{\mathbf{e}}_2 \wedge \hat{\mathbf{e}}_4 \ , \\ \mathbf{y}_3 &= \hat{\mathbf{e}}_1 \wedge \hat{\mathbf{e}}_3 \wedge \hat{\mathbf{e}}_4 \ , \ \ \mathbf{y}_4 = \hat{\mathbf{e}}_2 \wedge \hat{\mathbf{e}}_3 \wedge \hat{\mathbf{e}}_4 \ , \end{aligned}$$

as an orthonormal basis, while the $d_4 = 1$–dimensional vector space $\Lambda^4(\mathbb{R}^4)$ is spanned by vector

$$\mathbf{x}_1 = \hat{\mathbf{e}}_1 \wedge \hat{\mathbf{e}}_2 \wedge \hat{\mathbf{e}}_3 \wedge \hat{\mathbf{e}}_4.$$

Let us now consider 4 linearly independent vectors $\mathbf{u_1}$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{u}_4$ of $\mathbb{R}^4$ and the matrix

$$\mathbf{U} = [u_{ij}] = [\,\mathbf{u_1} \ \mathbf{u_2} \ \mathbf{u}_3 \ \mathbf{u}_4\,] = \begin{bmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ u_{21} & u_{22} & u_{23} & u_{24} \\ u_{31} & u_{32} & u_{33} & u_{34} \\ u_{41} & u_{42} & u_{43} & u_{44} \end{bmatrix} \ , \ \ i,j = 1,2,3,4,$$

having as columns the coordinates of these vectors with respect to the basis (107) of $\mathbb{R}^4$.

Considering basis (108) of $\Lambda^2(\mathbb{R}^4)$ the 2–vector $\mathbf{u}_1 \wedge \mathbf{u}_2$ is given by

$$\mathbf{u_1} \wedge \mathbf{u}_2 = \begin{vmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{vmatrix} \boldsymbol{\omega}_1 + \begin{vmatrix} u_{11} & u_{12} \\ u_{31} & u_{32} \end{vmatrix} \boldsymbol{\omega}_2 + \begin{vmatrix} u_{11} & u_{12} \\ u_{41} & u_{42} \end{vmatrix} \boldsymbol{\omega}_3 +$$
$$\begin{vmatrix} u_{21} & u_{22} \\ u_{31} & u_{32} \end{vmatrix} \boldsymbol{\omega}_4 + \begin{vmatrix} u_{21} & u_{22} \\ u_{41} & u_{42} \end{vmatrix} \boldsymbol{\omega}_5 + \begin{vmatrix} u_{31} & u_{32} \\ u_{41} & u_{42} \end{vmatrix} \boldsymbol{\omega}_6$$

according to equation (103). For the norm of this vector we get from equations (104) and (105):

$$\|\mathbf{u_1} \wedge \mathbf{u}_2\|^2 = \begin{vmatrix} \|\mathbf{u}_1\|^2 & \mathbf{u}_1 \cdot \mathbf{u}_2 \\ \mathbf{u}_2 \cdot \mathbf{u}_1 & \|\mathbf{u}_2\|^2 \end{vmatrix} = \begin{vmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{vmatrix}^2 + \begin{vmatrix} u_{11} & u_{12} \\ u_{31} & u_{32} \end{vmatrix}^2 +$$
$$\begin{vmatrix} u_{11} & u_{12} \\ u_{41} & u_{42} \end{vmatrix}^2 + \begin{vmatrix} u_{21} & u_{22} \\ u_{31} & u_{32} \end{vmatrix}^2 + \begin{vmatrix} u_{21} & u_{22} \\ u_{41} & u_{42} \end{vmatrix}^2 + \begin{vmatrix} u_{31} & u_{32} \\ u_{41} & u_{42} \end{vmatrix}^2 ,$$

where $\| \ \|$ is used also for denoting the usual Euclidian norm of a vector.

In a similar way we conclude that the norm of the 3–vector produced by $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$

$$\mathbf{u}_1 \wedge \mathbf{u}_2 \wedge \mathbf{u}_3 = \begin{vmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{vmatrix} \mathbf{y}_1 + \begin{vmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{41} & u_{42} & u_{43} \end{vmatrix} \mathbf{y}_2 +$$
$$\begin{vmatrix} u_{11} & u_{12} & u_{13} \\ u_{31} & u_{32} & u_{33} \\ u_{41} & u_{42} & u_{43} \end{vmatrix} \mathbf{y}_3 + \begin{vmatrix} u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \\ u_{41} & u_{42} & u_{43} \end{vmatrix} \mathbf{y}_4$$

is

$$\|\mathbf{u}_1 \wedge \mathbf{u}_2 \wedge \mathbf{u}_3\|^2 = \begin{vmatrix} \|\mathbf{u}_1\|^2 & \mathbf{u}_1 \cdot \mathbf{u}_2 & \mathbf{u}_1 \cdot \mathbf{u}_3 \\ \mathbf{u}_2 \cdot \mathbf{u}_1 & \|\mathbf{u}_2\|^2 & \mathbf{u}_2 \cdot \mathbf{u}_3 \\ \mathbf{u}_3 \cdot \mathbf{u}_1 & \mathbf{u}_3 \cdot \mathbf{u}_2 & \|\mathbf{u}_3\|^2 \end{vmatrix} =$$
$$\begin{vmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{vmatrix}^2 + \begin{vmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{41} & u_{42} & u_{43} \end{vmatrix}^2 + \begin{vmatrix} u_{11} & u_{12} & u_{13} \\ u_{31} & u_{32} & u_{33} \\ u_{41} & u_{42} & u_{43} \end{vmatrix}^2 + \begin{vmatrix} u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \\ u_{41} & u_{42} & u_{43} \end{vmatrix}^2 ,$$

while the norm of the 4–vector produced by $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{u}_4$

$$\mathbf{u_1} \wedge \mathbf{u}_2 \wedge \mathbf{u}_3 \wedge \mathbf{u}_4 = |\mathbf{U}| \, \mathbf{x}_1$$

is given by

$$\|\mathbf{u}_1 \wedge \mathbf{u}_2 \wedge \mathbf{u}_3 \wedge \mathbf{u}_4\|^2 = \begin{vmatrix} \|\mathbf{u}_1\|^2 & \mathbf{u}_1 \cdot \mathbf{u}_2 & \mathbf{u}_1 \cdot \mathbf{u}_3 & \mathbf{u}_1 \cdot \mathbf{u}_4 \\ \mathbf{u}_2 \cdot \mathbf{u}_1 & \|\mathbf{u}_2\|^2 & \mathbf{u}_2 \cdot \mathbf{u}_3 & \mathbf{u}_2 \cdot \mathbf{u}_4 \\ \mathbf{u}_3 \cdot \mathbf{u}_1 & \mathbf{u}_3 \cdot \mathbf{u}_2 & \|\mathbf{u}_3\|^2 & \mathbf{u}_3 \cdot \mathbf{u}_4 \\ \mathbf{u}_4 \cdot \mathbf{u}_1 & \mathbf{u}_4 \cdot \mathbf{u}_2 & \mathbf{u}_4 \cdot \mathbf{u}_3 & \|\mathbf{u}_4\|^2 \end{vmatrix} = |\mathbf{U}|^2 .$$

References

1. Allen L, Bridges TJ (2002) Numerical exterior algebra and the compound matrix method. Numerische Mathematik 92:197–232
2. Antonopoulos C, Bountis T (2006) Detecting Order and Chaos by the Linear Dependence Index (LDI) Method. ROMAI Journal 2:1–13
3. Antonopoulos C, Bountis T, Skokos Ch (2006) Chaotic dynamics of N–degree of freedom Hamiltonian systems. Int. J. Bif. Chaos 16:1777–1793
4. Bario R (2005) Sensitivity tools vs. Poincaré sections. Chaos Solit. Fract. 25:711–726
5. Bario R (2006) Painting chaos: a gallery of sensitivity plots of classical problems. Int. J. Bif. Chaos 16:2777–2798
6. Bario R, Borczyk W, Breiter S (2009) Spurious structures in chaos indicators maps. Chaos Solit. Fract. (in press)
7. Barreira L, Pesin Y (2006) Smooth ergodic theory and nonuniformly hyperbolic dynamics. In: Hasselblatt B, Katok A (eds.) Handbook of dynamical systems. Vol. 1B, 57–263, Elsevier
8. Benettin G, Strelcyn J–M (1978) Numerical experiments of the free motion of a point mass moving in a plane convex region: Stochastic transition and entropy. Phys. Rev. A 17:773–785
9. Benettin G, Galgani L (1979) Lyapunov characteristic exponents and stochasticity. In: Laval G, Grésillon D (eds.) Intrinsic Stochasticity in Plasmas, 93–114, Edit. Phys. Orsay
10. Benettin G, Galgani L, Strelcyn J–M (1976) Kolmogorov entropy and numerical experiments Phys. Rev. A 14:2338–2344
11. Benettin G, Galgani L, Giorgilli A, Strelcyn J–M (1978) Tous les nombres caractéristiques sont effectivement calculables. C. R. Acad. Sc. Paris Sér. A 286:431–433
12. Benettin G, Froeschlé C, Scheidecker JP (1979) Kolmogorov entropy of a dynamical system with an increasing number of degrees of freedom. Phys. Rev. A 19:2454–2460
13. Benettin G, Galgani L, Giorgilli A, Strelcyn J–M (1980) Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; A method for computing all of them. Part 1: theory. Meccanica March:9–20
14. Benettin G, Galgani L, Giorgilli A, Strelcyn J–M (1980) Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; A method for computing all of them. Part 2: Numerical application. Meccanica March:21–30
15. Bountis T, Manos T, Christodoulidi H (2008) Application of the GALI method to localization dynamics in nonlinear systems. J. Comp. Appl. Math. (in press), nlin.CD/0806.3563
16. Bourbaki N (1958) Éléments de mathématique, Livre II: Algèbre, Chapitre 3. Hermann, Paris
17. Brown R, Bryant P, Abarbanel HDI (1991) Computing the Lyapunov spectrum of a dynamical system from an observed time series. Phys. Rev. A 43:2787–2806
18. Bridges TJ, Reich S (2001) Computing Lyapunov exponents on a Stiefel manifold. Physica D 156:219–238
19. Burns K, Donnay V (1997) Embedded surfaces with ergodic geodesic flow. Int. J. Bif. Chaos 7:1509–1527

20. Carbonell F, Jimenez JC, Biscay R (2002) A numerical method for the computation of the Lyapunov exponents of nonlinear ordinary differential equations. Appl. Math. Comput. 131:21–37
21. Casartelli M, Diana E, Galgani L, Scotti A (1976) Numerical computations on a stochastic parameter related to the Kolmogorov entropy. Phys. Rev. A 13:1921–1925
22. Casati G, Ford J (1975) Stochastic transition in the unequal–mass Toda lattice. Phys. Rev. A 12:1702–1709
23. Casati G, Chirikov BV, Ford J (1980) Marginal local instability of quasi–periodic motion. Phys. Let. A 77:91–94
24. Chen Z–M, Djidjeli K, Price WG (2006) Computing Lyapunov exponents based on the solution expression of the variational system. Appl. Math. Comput. 174:982–996
25. Chernov N, Markarian R (2006) Chaotic billiards. Mathematical Surveys and Monographs, Vol. 127. American Mathematical Society
26. Christiansen F, Rugh HH (1997) Computing Lyapunov spectra with continuous Gram–Schmidt orthonormalization. Nonlinearity 10:1063–1072
27. Christodoulidi H, Bountis T (2006) Low-dimensional quasiperiodic motion in Hamiltonian systems. ROMAI Journal 2:37–44
28. Cincotta PM, Simó C (2000) Simple tools to study global dynamics in non–axisymmetric galactic potentials–I. Astron. Astroph. Supp. Ser. 147:205–228
29. Cincotta PM, Giordano CM, Simó C (2003) Phase space structure of multi–dimensional systems by means of the mean exponential growth factor of nearby orbits. Physica D 182:151–178
30. Contopoulos G (2002) Order and Chaos in Dynamical Astronomy. Springer, Berlin Heidelberg New York
31. Contopoulos G, Giorgilli A (1988) Bifurcations and complex instability in a 4–dimensional symplectic mapping. Meccanica 23:19–28
32. Contopoulos G, Voglis N (1996) Spectra of stretching numbers and helicity angles in dynamical systems. Cel. Mech. Dyn. Astron. 64:1–20
33. Contopoulos G, Voglis N (1997) A fast method for distinguishing between ordered and chaotic orbits. Astron. Astroph. 317:73–81
34. Contopoulos G, Galgani L, Giorgilli A (1978) On the number of isolating integrals in Hamiltonian systems. Phys. Rev. A 18:1183–1189
35. Devaney RL (1989) An Introduction to Chaotic Dynamical Systems. 2nd Ed. Addison–Wesley Publishing Company, New York
36. Dieci L, Van Vleck ES (1995) Computation of a few Lyapunov exponents for continuous and discrete dynamical systems. Appl. Num. Math. 17:275–291
37. Dieci L, Van Vleck ES (2002) Lyapunov spectral intervals: theory and computation. SIAM J. Numer. Anal. 40:516–542
38. Dieci L, Elia C (2006) The singular value decomposition to approximate spectra of dynamical systems. Theoretical aspects. J. Diff. Eq. 230:502–531
39. Dieci L, Lopez L (2006) Smooth singular value decomposition on the symplectic group and Lyapunov exponents approximation. Calcolo 43:1–15
40. Dieci L, Russell RD, Van Vleck ES (1997) On the computation of Lyapunov exponents for continuous dynamical systems. SIAM J. Numer. Anal. 34:402–423
41. Donnay V (1988) Geodesic flow on the two-sphere, Part I: Positive measure entropy. Erg. Theory Dyn. Syst. 8:531–553

42. Donnay V (1988) Geodesic flow on the two-sphere, Part II: Ergodicity. Lect. Notes Math. 1342:112–153

43. Donnay V, Liverani C (1991) Potentials on the two-torus for which the Hamiltonian flow is ergodic. Commun. Math. Phys. 135:267–302

44. Eckmann J–P, Ruelle D (1985) Ergodic theory of chaos and strange attractors. Rev. Mod. Phys. 57:617–656

45. Eckmann J–P, Oliffson Kamphorst S, Ruelle D, Ciliberto S (1986) Liapunov exponents from time series. Phys. Rev. A 34:4971–4979

46. Farmer JD (1982) Chaotic attractors of an infinite–dimensional dynamical system. Physica D 4:366–393

47. Farmer JD, Ott E, Yorke JA (1983) The dimension of chaotic attractors. Physica D 7:153–180

48. Ford J, Lunsford GH (1970) Stochastic behavior of resonant nearly linear oscillator systems in the limit of zero nonlinear coupling. Phys. Rev. A 1:59–70

49. Fouchard M, Lega E, Froeschlé Ch, Froeschlé C (2002) On the relationship between the fast Lyapunov indicator and periodic orbits for continuous flows. Cel. Mech. Dyn. Astron. 83:205–222

50. Freistetter F (2000) Fractal dimensions as chaos indicators. Cel. Mech. Dyn. Astron. 78:211–225

51. Froeschlé C (1970) Numerical study of dynamical systems with three degrees of freedom II. Numerical displays of four–dimensional sections. Astron. Astroph. 5:177–183

52. Froeschlé C (1972) Numerical study of a four–dimensional mapping. Astron. Astroph. 16:172–189

53. Froeschlé C (1984) The Lyapunov characteristic exponents - Applications to Celestial Mechanics. Cel. Mech. Dyn. Astron. 34:95–115

54. Froeschlé C (1984) The Lyapunov characteristic exponents and applications. J. de Méc. Théor. et Appl. Numéro spécial 101–132

55. Froeschlé C (1985) The Lyapunov characteristic exponents and applications to the dimension of the invariant manifolds and chaotic attractors. In: Szebehely VG (ed.) Stability of the Solar System and its minor natural and artificial bodies, 265–282, D. Reidel Publishing Company

56. Froeschlé C, Lega E (2000) On the structure of symplectic mappings. The fast Lyapunov indicator: a very sensitive tool. Cel. Mech. Dyn. Astron. 78:167–195

57. Froeschlé C, Froeschlé Ch, Lohinger E (1993) Generalized Lyapunov characteristic indicators and corresponding Kolmogorov like entropy of the standard mapping. Cel. Mech. Dyn. Astron. 56:307–314

58. Froeschlé C, Lega E, Gonczi R (1997) Fast Lyapunov indicators. Application to asteroidal motion. Cel. Mech. Dyn. Astron. 67:41–62

59. Froeschlé C, Gonczi R, 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. 45:881–886

60. Frøyland J (1983) Lyapunov exponents for multidimensional orbits. Phys. Let. A 97:8-10

61. Frøyland J, Alfsen KH (1984) Lyapunov–exponent spectra for the Lorenz model. Phys. Rev. A 29:2928–2931

62. Geist K, Parlitz U, Lauterborn W (1990) Comparison of different methods for computing Lyapunov exponents. Prog. Theor. Phys. 83:875–893

63. Goldhirsch I, Sulem P–L, Orszag SA (1987) Stability and Lyapunov stability of dynamical systems: a differential approach and a numerical method. Physica D 27:311–337
64. Gottwald GA, Melbourne I (2004) A new test for chaos in deterministic systems. Proc. Roy. Soc. London A 460:603–611
65. Gottwald GA, Melbourne I (2005) Testing for chaos in deterministic systems with noise. Physica D 212:100–110
66. Grassberger P, Procaccia I (1983) Measuring the strangeness of strange attractors. Physica D 9:189–208
67. Greene JM, Kim J–S (1987) The calculation of Lyapunov spectra. Physica D 24:213–225
68. Greub W (1978) Multilinear Algebra. 2nd Ed. Springer, Berlin Heidelberg New York
69. Guzzo M, Lega E, Froeschlé C (2002) On the numerical detection of the effective stability of chaotic motions in quasi–integrable systems. Physica D 163:1–25
70. Haken H (1983) At least one Lyapunov exponent vanishes if the trajectory of an attractor does not contain a fixed point. Phys. Let. A 94:71–72
71. Hegger R, Kantz H, Schreiber T (1999) Practical implementation of nonlinear time series methods: The TISEAN package. Chaos 9:413–435
72. Hénon M, Heiles C (1964) The applicability of the third integral of motion: some numerical experiments. Astron. J. 69:73–79
73. Horn RA, Johnson CR (1985) Matrix Analysis. Cambridge University Press
74. Howard JE (2005) Discrete virial theorem. Cel. Mech. Dyn. Astron. 92:219–241
75. Hubbard JH, Hubbard BB (1999) Vector Calculus, Linear Algebra and Differential Forms: A Unified Approach. Prentice Hall, New Jersey
76. Johnson BA, Palmer KJ, Sell GR (1987) Ergodic properties of linear dynamical systems. SIAM J. Math. Anal. 18:1–33
77. Kantz H (1994) A robust method to estimate the maximal Lyapunov exponent of a time series. Phys. Let. A 185:77-87
78. Kantz H, Schreiber T (1997) Nonlinear time series analysis. Cambridge University Press
79. Kaplan JL, Yorke JA (1979) Chaotic behavior of multidimensional difference equations. In: Peitgen H–O, Walter H–O (eds.) Functional Differential Equations and Approximations of Fixed Points, Lect. Notes Math. 730:204–227
80. Karanis GI, Vozikis ChL (2008) Fast detection of chaotic behavior in galactic potentials. Astron. Nachr. 329:403–412
81. Kotoulas T, Voyatzis G (2004) Comparative study of the 2:3 and 3:4 resonant motion with Neptune: An application of symplectic mappings and low frequency analysis. Cel. Mech. Dyn. Astron. 88:343–163
82. Kovács B (2007) About the efficiency of Fast Lyapunov Indicator surfaces and Small Alignment Indicator surfaces. PADEU, 19:221–236
83. Laskar J (1990) The chaotic motion of the Solar System: a numerical estimate of the size of the chaotic zones. Icarus 88:266–291
84. Laskar J (1993) Frequency analysis of multi–dimensional systems. Global dynamics and diffusion. Physica D 67:257–281
85. Laskar J (1999) Introduction to frequency map analysis. In: Simó C (ed.) Hamiltonian systems with three or more degrees of freedom, 134–150, Plenum Press

86. Laskar J, Froeschlé C, Celletti A (1992) The measure of chaos by the numerical analysis of the fundamental frequencies. Application to the standard map. Physica D 56:253–269
87. Ledrappier F, Young L–S (1988) Dimension formula for random transformations. Commun. Math. Phys. 117:529–548
88. Lega E, Froeschlé C (1998) Comparison of convergence towards invariant distributions for rotation angles, twist angles and local Lyapunov characteristic numbers. Planet. Space Sci. 46:1525–1534
89. Lega E, Froeschlé C (2001) On the relationship between fast Lyapunov indicator and periodic orbits for symplectic mappings. Cel. Mech. Dyn. Astron. 81:129–147
90. Li C, Chen G (2004) Estimating the Lyapunov exponents of discrete systems. Chaos 14:343–346
91. Li C, Xia X (2004) On the bound of the Lyapunov exponents of continuous systems. Chaos 14:557–561
92. Lichtenberg AJ, Lieberman MA (1992) Regular and Chaotic Dynamics. Second Edition. Springer, Berlin Heidelberg New York
93. Lohinger E, Froeschlé, Dvorak R (1993) Generalized Lyapunov exponents indicators in Hamiltonian dynamics: an application to a double star system. Cel. Mech. Dyn. Astron. 56:315–322
94. Lu J, Yang G, Oh H, Luo ACJ (2005) Computing Lyapunov exponents of continuous dynamical systems: method of Lyapunov vectors. Chaos Sol. Fract. 23:1879–1892
95. Lukes–Gerakopoulos G, Voglis N, Efthymiopoulos C (2008) The production of Tsallis entropy in the limit of weak chaos and a new indicator of chaoticity. Physica A 387:1907–1925
96. Lyapunov AM (1992) The General Problem of the Stability of Motion. Taylor and Francis, London (English translation from the French: Liapounoff A (1907) Problème général de la stabilité du mouvement. Annal. Fac. Sci. Toulouse 9:203–474. The French text was reprinted in Annals Math. Studies Vol. 17 Princeton Univ. Press (1947). The original was published in Russian by the Mathematical Society of Kharkov in 1892)
97. Markarian R (1994) Non-uniformly hyperbolic billiards. Annal. Fac. Sci. Toulouse 3:223–257
98. Mathiesen J, Cvitanović P (2008) Lyapunov exponents. In: Cvitanović P, Artuso R, Mainieri R, Tanner G, Vattay G Chaos: Classical and Quantum. Niels Bohr Institute, Copenhagen, `http://chaosbook.org/version12/`
99. Nagashima T, Shimada I (1977) On the C–system–like property of the Lorenz system. Prog. Theor. Phys. 58:1318–1320
100. Núñez JA, Cincotta PM, Wachlin FC (1996) Information entropy. An indicator of chaos. Cel. Mech. Dyn. Astron. 64:43–53
101. Oliveira S, Stewart DE (2000) Exponential splittings of products of matrices and accurately computing singular values of long products. Lin. Algebra Appl. 309:175–190
102. Oseledec VI (1968) A multiplicative ergodic theorem. Ljapunov characteristic numbers for dynamical systems. Trans. Moscow Math. Soc. 19:197–231
103. Ott E (1981) Strange attractors and chaotic motions of dynamical systems. Rev. Mod. Phys. 53:655–671
104. Packard NH, Crutchfiled JP, Farmer JD, Shaw RS (1980) Geometry from time series. Phys. Rev. Let. 45:712–716

105. Paleari S, Penati S (2008) Numerical methods and results in the FPU problem. Lect. Notes Phys. 728:239–282
106. Pesin YaB (1977) Characteristic Lyapunov exponents and smooth ergodic theory. Russian Math. Surveys 32:55–114
107. Press WH, Teukolsky SA, Vetterling WT, Flannery BP (2007) Numerical Recipes in Fortran. The Art of Scientific Computing. Cambridge University Press
108. Raghunathan MS (1979) A proof of Oseledec's multiplicative ergodic theorem. Isr. J. Math. 32:356–362
109. Ramasubramanian K, Sriram MS (2000) A comparative study of computation of Lyapunov spectra with different algorithms. Physica D 139:72–86
110. Rangarajan G, Habib S, Ryne RD (1998) Lyapunov exponents without rescaling and reorthogonalization. Phys. Rev. let. 80:3747–3750
111. Rosenstein MT, Collins JJ, De Luca CJ (1993) A practical method for calculating largest Lyapunov exponents from small data sets. Physica D 65:117–134
112. Roux J–C, Simoyi RH, Swinney HL (1983) Observation of a strange attractor. Physica D 8:257–266
113. Ruelle D (1979) Analycity properties of the characteristic exponents of random matrix products. Adv. Math. 32:68–80
114. Ruelle D (1979) Ergodic theory of differentiable dynamical systems. IHES Publ. Math. 50:275–306
115. Sándor Zs, Érdi B, Efthymiopoulos C (2000) The phase space structure around L_4 in the restricted three–body problem. Cel. Mech. Dyn. Astron. 78:113–123
116. Sándor Zs, Érdi B, Széll A, Funk B (2004) The relative Lyapunov indicator: an efficient method of chaos detection. Cel. Mech. Dyn. Astron. 90:127–138
117. Sandri M (1996) Numerical Calculation of Lyapunov Exponents. Mathematica J. 6:78–84
118. Sano M, Sawada Y (1985) Measurement of the Lyapunov spectrum from a chaotic time series. Phys. Rev. Let. 55:1082–1085
119. Shimada I, Nagashima T (1979) A numerical approach to ergodic problem of dissipative dynamical systems. Prog. Theor. Phys. 61:1605–1615
120. Sideris IV (2005) Characterization of chaos: A new, fast and effective measure. In: Gottesman ST, Buchler J–R (eds.) Nonlinear Dynamics in Astronomy and Astrophysics, Annals of the New York Academy of Science, 1045:79, The New York Academy of Sciences
121. Sideris IV (2006) Measure of orbital stickiness and chaos strength. Phys. Rev. E 066217
122. Skokos Ch (2001) Alignment indices: a new, simple method for determining the ordered or chaotic nature of orbits. J. Phys. A 34:10029–10043
123. Skokos Ch, Contopoulos G, Polymilis C (1997) Structures in the phase space of a four dimensional symplectic map. Cel. Mech. Dyn. Astron. 65:223–251
124. Skokos Ch, Antonopoulos Ch, Bountis TC, Vrahatis MN (2003) How does the smaller alignment index (SALI) distinguish order from chaos? Prog. Theor. Phys. Supp. 150:439–443
125. Skokos Ch, Antonopoulos Ch, Bountis TC, Vrahatis MN (2004) Detecting order and chaos in Hamiltonian systems by the SALI method. J. Phys. A 37:6269–6284
126. Skokos Ch, Bountis TC, Antonopoulos Ch (2007) Geometrical properties of local dynamics in Hamiltonian systems: The generalized alignment index (GALI) method. Physica D 231:30–54

127. Skokos Ch, Bountis TC, Antonopoulos Ch (2008) Detecting chaos, determining the dimensions of tori and predicting slow diffusion in Fermi–Pasta–Ulam lattices by the generalized alignment index method. Eur. Phys. J. Sp. Top. 165:5–14

128. Spivak M (1965) Calculus on Manifolds. Addison–Wesley, New York

129. Spivak M (1999) Comprehensive Introduction to Differential Geometry. Vol. 1. Publish or Perish Inc., Houston

130. Stewart DE (1997) A new algorithm for the SVD of a long product of matrices and the stability of products. Electr. Trans. Numer. Anal. 5:29–47

131. Stoddard SD, Ford J (1973) Numerical experiments on the stochastic behavior of a Lennard–Jones gas system. Phys. Rev. A 8:1504–1512

132. Süli Á (2006) Speed and efficiency of chaos detection methods. In: Süli Á, Freistetter F, Pál A (eds.) Proceedings of the 4th Austrian Hungarian workshop on Celestial Mechanics, 18:179–189, Publications of the Astronomy Department of the Eötvös University

133. Süli Á (2006) Motion indicators in the 2D standard map. PADEU 17:47–62

134. Takens F (1981) Detecting strange attractors in turbulence. Lect. Notes Math. 898:366–381

135. Voglis N, Contopoulos G (1994) Invariant spectra of orbits in dynamical systems. j. Phys. A 27:4899–4909

136. Voglis N, Contopoulos G, Efthymiopoulos C (1998) Method for distinguishing between ordered and chaotic orbits in four–dimensional maps. Phys. Rev. E 57:372–377

137. Voyatzis G, Ichtiaroglou S (1992) On the spectral analysis of trajectories in near–integrable Hamiltonian systems. J. Phys. A 25:5931–5943

138. Vozikis ChL, Varvoglis H, Tsiganis K (2000) The power spectrum of geodesic divergences as an early detector of chaotic motion. Astron. Astroph. 359:386–396

139. Vrahatis MN, Bountis TC, Kollmann M (1996) Periodic orbits and invariant surfaces of 4D nonlinear mappings. Int. J. Bif. Chaos 6:1425–1437

140. Vrahatis MN, Isliker H, Bountis TC (1997) Structure and breakdown of invariant tori in a 4–D mapping model of accelerator dynamics. Int. J. Bif. Chaos 7:2707–2722

141. Walters P (1993) A dynamical proof of the multiplicative ergodic theorem. Thans. Amer. Math. Soc. 335:245–257

142. Wojtkowski M (1985) Invariant families of cones and Lyapunov exponents. Erg. Theory Dyn. Syst. 5:145–161

143. Wojtkowski M (1986) Principles for the design of billiards with nonvanishing Lyapunov exponents. Commun. Math.Phys. 105:391–414

144. Wolf A, Swift JB, Swinney HL, Vastano JA (1985) Determining Lyapunov exponents from a time series. Physica D 16:285–317

145. Wu X, Huang T–Y, Zhang H (2006) Lyapunov indices with two nearby trajectories in a curved spacetime. Phys. Rev. E 74:083001

146. Young L–S (1982) Dimension, entropy and Lyapunov exponents. Erg. Theory Dyn. Syst. 2:109–124

147. Zou Y, Pazó D, Romano MC, Thiel M, Kurths J (2007) Distinguishing quasiperiodic dynamics from chaos in short–time series. Phys. Rev. E 76:016210

148. Zou Y, Thiel M, Romano MC, Kurths J (2007) Characterization of stickiness by means of recurrence. Chaos 17:043101