Physica 27D (1987) 311–337
North-Holland, Amsterdam

STABILITY AND LYAPUNOV STABILITY OF DYNAMICAL SYSTEMS: A DIFFERENTIAL APPROACH AND A NUMERICAL METHOD

Isaac GOLDHIRSCH
Department of Fluid Mechanics and Heat Transfer, Faculty of Engineering, Tel-Aviv University, Ramat-Aviv, Tel-Aviv, 69978, Israel

Pierre-Louis SULEM
School of Mathematical Sciences, Tel-Aviv University, Ramat-Aviv, Tel-Aviv, 69978, Israel and CNRS, Observatoire de Nice, France

and

Steven A. ORSZAG
Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544, USA

Received 14 October 1985
Revised manuscript received 6 October 1986

A set of differential equations for the eigenvalues and eigenvectors of the stability matrix of a dynamical system, as well as for the Lyapunov exponents and the corresponding eigenvectors, is derived. The rate of convergence of the Lyapunov eigenvectors is shown to be exponential. The eigenvectors of the stability matrix can be grouped into sets, each spanning a subspace which converges at an exponential rate. It is demonstrated that, generically, the real parts of the eigenvalues of the stability matrix equal the corresponding Lyapunov exponents. This statement has been tested numerically. The values of the Lyapunov exponents, μ_i, are shown to be related to the corresponding finite time values of the Lyapunov exponents (e.g. those deduced from a finite time numerical simulation), $\mu_i(t)$, by: $\mu_i(t)=\mu_i+(b_i+\xi_i(t))/t$. The b_i's are constants and $\xi_i(t)$ are "noise" terms of zero mean. This observation leads to a method of extrapolation, which has been used to predict Lyapunov exponents from a finite amount of data. It is shown that the use of the standard (numerical) methods to compute Lyapunov exponents introduces an error a_i/t in the value of $\mu_i(t)$, where the a_i's are constants. Thus the standard method has a rate of convergence which is the same as that of the exact $\mu_i(t)$'s. Finally, we have shown how one can compute the eigenvectors associated with each of the eigenvalues of the stability matrix as well as the Lyapunov eigenvectors.

1. Introduction

One of the most meaningful ways to characterize a nonlinear dynamical system is that suggested by Lyapunov [1–5]. It is basically a generalization of linear stability analysis (which applies to perturbations of steady state solutions) to unsteady and even chaotic situations.

Consider a dynamical system of dimension n, whose variables are $(x_i;\ 1\le i\le n)$. Let $\boldsymbol{x}$ be the vector $(x_1, x_2, \dots, x_n)$. The dynamical system is defined by the equations: $\dot{x}_i = F_i(\boldsymbol{x})$ or

$$\dot{\boldsymbol{x}} = \boldsymbol{F}(\boldsymbol{x}). \tag{1.1}$$

For a given solution of (1.1), $\boldsymbol{x}(t)$, define the (Hessenberg) matrix $\mathbf{G}$ by its elements,

$$G_{ij}(\boldsymbol{x}(t)) = \left.\frac{\partial F_i(\boldsymbol{x})}{\partial x_j}\right|_{\boldsymbol{x}=\boldsymbol{x}(t)} \tag{1.2}$$

The stability equation can now be defined as

$$\dot{\boldsymbol{y}} = \mathbf{G}\boldsymbol{y}, \tag{1.3}$$

where y is a time-dependent n-dimensional vector. The vector y can be regarded as a small perturbation δx. The solution of (1.3) can be written as

$$y(t_2) = \mathbf{M}(t_2, t_1)\, y(t_1) \tag{1.4}$$

and the matrix $\mathbf{M}$ can be written as a time-ordered product,

$$\mathbf{M}(t_2, t_1) = T_+ \, \mathrm{e}^{\int_{t_1}^{t_2} \mathbf{G}(x(t'))\, \mathrm{d}t'}. \tag{1.5}$$

It has been proven by Oseledec [1] that for almost every initial point $x(t_1)$ there exists an orthonormal set of vectors $\{ f_i(t_1)\}$ $(1 \le i \le n)$ such that

$$\mu_i = \lim_{t_2 \to \infty} \frac{1}{t_2 - t_1} \log \| \mathbf{M}(t_2, t_1)\, f_i(t_1) \|, \tag{1.6}$$

which means that the limit on the right side of eq. (1.6) exists and equals a number μ_i for $1 \le i \le n$. It has been argued [2] that for ergodic systems the set $\{\mu_i\}$ does not depend on the initial point (up to set of measure zero in phase space) and so the numbers $\{\mu_i\}$ are global properties of the attractor of the dynamical system. The immediate significance of the numbers $\{\mu_i\}$, known as the Lyapunov exponents, is that nearby trajectories, separated initially by an infinitesimal distance in the direction of f_i will separate essentially at a rate of $\mathrm{e}^{\mu_i t}$. {Obviously a separation at a rate of $\mathrm{e}^{\mu_i t + \beta_i t^{\alpha}}$ where $\alpha < 1$ corresponds by eq. (1.6) to a Lyapunov exponent μ_i. Other types of separation are possible as well. For short times, the rate of separation need not be exponential.} A positive value of a μ_i indicates an exponential loss of correlation between two initial trajectories "starting of" at nearby points. This kind of separation of trajectories is one of the most important ingredients of chaotic dynamics. Indeed, one of the widely accepted definitions of a chaotic state is the existence of at least one positive Lyapunov exponent. When all Lyapunov exponents are negative, the attractor is a stable fixed point. A stable limit cycle has one zero exponent (corresponding to a perturbation tangent to the cycle) and all other exponents are negative. A two-torus has two zero exponents for a similar reason. Every attractor of a smooth dynamical system [such as given by eq. (1.1) for smooth $f(x)$] has at least one zero Lyapunov exponent corresponding to a perturbation tangent to the trajectory.

The Lyapunov exponents are believed to be related to other global properties of attractors. An interesting suggestion was made by Kaplan and Yorke [6], who defined a dimension D_{KY} as follows:

$$D_{\mathrm{KY}} = j + \frac{\mu_1 + \mu_2 + \cdots + \mu_j}{|\mu_{j+1}|}, \tag{1.7}$$

where j is the smallest integer for which $\mu_1 + \mu_2 + \cdots + \mu_j > 0$ and $\mu_1 + \mu_2 + \cdots + \mu_{j+1} < 0$. It turns out that for many systems that have been investigated, the Kaplan–Yorke dimension is close to the correlation dimension [7] and to the capacity dimension [8] (in the cases where it can be measured). The sum of positive Lyapunov numbers is the Kolmogoroff entropy of a dynamical system [9]. All in all, the Lyapunov exponents have proven very useful for the characterization of dynamical systems.

In contrast to the exponents $\{\mu_i\}$, the eigenvectors $\{ f_i\}$ are known to be position-dependent in the attractor. They are only "local" properties. While is is intuitively clear that the f_i's carry important local geometrical and dynamical information, we know of little progress in deducing such information [10].

It is interesting to note that Lyapunov stability analysis is not a generalization of linear stability analysis to time-dependent systems. Indeed, linear stability analysis would involve the computation of the eigenvalues of the matrix $\mathbf{M}(t_2, t_1)$, which solves the linearized equation of motion, eq. (1.3). These eigenvalues can be written as $\mathrm{e}^{\lambda_i(t)t}$, with time-dependent "exponents" $\lambda_i(t)$. The latter, unlike the Lyapunov exponents, are, in general, complex quantities and there is no parallel of the Oseledec theorem for them. Define the quantities

$e^{\mu_i(t)t}$ as the eigenvalues of $\mathbf{M}^\dagger(t,0)\mathbf{M}(t,0)$, which is a symmetric matrix; thus the $\mu_i(t)$'s are real numbers. We shall call them the time-dependent Lyapunov exponents. Let $f_i(t)$ be a (normalized) eigenvector of $\mathbf{M}^\dagger\mathbf{M}$ corresponding to $\mu_i(t)$. Then

$$\|\mathbf{M}(t,0)\, f_i(t)\| = e^{\mu_i(t)t}. \tag{1.8}$$

We shall show below that the time-dependent Lyapunov exponents converge to the respective Lyapunov exponents and that $f_i(t) \to f_i$ (see section 3). Thus Lyapunov stability analysis involves the investigation of the eigenvalues of $\mathbf{M}^\dagger\mathbf{M}$ rather than those of $\mathbf{M}$. In spite of this difference, we shall argue below (see section 4) that (generically) to each μ_i there exists a corresponding λ_i such that $\mu_i = \operatorname{Re}\lambda_i$. Counterexamples to this statement {we are grateful to the referees for helping us elucidate this point} can easily be constructed (see section 4). In what follows we shall call the quantities $\lambda_i(t)$ the (time-dependent) stability exponents. Let $\{e_i(t)\}$ be a right eigenvector of $\mathbf{M}(t,0)$ corresponding to $\lambda_i(t)$. If $\mathbf{M}$ has a degenerate spectrum, the number of $\{e_i\}$'s can be less than the dimension (n) of $\mathbf{M}$. We shall consider below only the case when $\operatorname{Re}\lambda_i$ are nondegenerate (i.e. to each i there corresponds a different rate of separation). Since $\mathbf{M}$ is real by construction, any complex eigenvalue of $\mathbf{M}$, $a+ib$, has a corresponding complex eigenvalue $a-ib$, in contrast to our assumption of nondegeneracy of the real parts of the eigenvalues of $\mathbf{M}$. Thus the λ_i's are all real under the above assumption. The advantage in considering the stability exponents is that an initial ($t=0$) separation of trajectories in the direction of $(e_i(t))$ is stretched in the same direction by $e^{\lambda_i t}$ after a time t. Moreover, as is shown below (section 2), it we arrange the λ_i's in decreasing order $\lambda_1 > \lambda_2 > \cdots > \lambda_n$ then $e_n(t)$ becomes asymptotically (in time) a fixed direction and any subspace spanned by $\{e_n(t), e_{n-1}(t), \ldots, e_{n-k}(t)\}$ for $0 \le k \le n$ becomes asymptotically a fixed subspace. In contrast, the direction of $\mathbf{M}(t,0)\, f_i$ is not fixed since f_i is not an eigenvector of $\mathbf{M}$. Actually, $\mathbf{M}(t,0)$ rotates the direction of the set of vectors f_i. To see this notice that:

$$\left(\mathbf{M}(t,0)\, f_j\right)^\dagger \cdot \mathbf{M}(t,0)\, f_i = f_j^+ \mathbf{M}^\dagger(t,0)\mathbf{M}(t,0)\, f_i. \tag{1.9}$$

Since f_i is (for large enough times) an approximate eigenvector of $\mathbf{M}^\dagger\mathbf{M}$,

$$f_j^+ \mathbf{M}^\dagger(t,0) \cdot \mathbf{M}(t,0)\, f_i \approx \delta_{i,j}\, e^{2\mu_i t}, \tag{1.10}$$

which shows that the mutually orthogonal vectors f_i are transformed into other mutually orthogonal vectors under the operation of $\mathbf{M}$. The eigenvectors of $\mathbf{M}$ may be of more direct geometrical meaning than those of $\mathbf{M}^\dagger\mathbf{M}$. This point, however, will be pursued in a different publication.

The computation of Lyapunov exponents for model systems [3–5, 11–17] and experimental systems is an important endeavor. Bennetin et al. [11] suggested a method to calculate the largest Lyapunov exponent. Later, they [12] and Shimada et al. [13] suggested algorithms to compute all Lyapunov exponents of a dynamical system. Their method, which is called below the standard method, is very successful for low-dimensional dynamical systems, when the equations of motion are known and soluble at least numerically. The slow rate of convergence of the Lyapunov exponents to their asymptotic values [5] makes the straightforward use of the standard method inapplicable to large systems. The necessity of solving the equations of motion (as well as the linearized equations of motion) make the method impractical for experimental systems. Other methods [14, 17] were suggested to compute Lyapunov exponents in such cases. In the present paper we show that for large enough times (which are much shorter than the convergence times of the $\mu_i(t)$) the time-dependent Lyapunov exponents satisfy

$$\mu_i(t) = \mu_i + \frac{b_i + \xi_i(t)}{t}, \tag{1.11}$$

where b_i are constants and $\xi_i(t)$ are "noise" terms of zero average and relatively small amplitudes. We use this property to predict μ_i from the knowledge of $\mu_i(t)$ for a relatively short time. This method can be employed in conjunction with different methods for obtaining Lyapunov exponents. We believe it should be useful for the analysis of experimental data and of numerical data for large systems.

While our approach is mostly nonrigorous, it is by far simpler than the exact mathematical approaches to the problem at hand. Some of our results, such as the existence of Lyapunov eigenvectors [1] are known to the mathematical community. The derivation here seems to us more transparent. Other results such as rates of convergence of eigenvectors and eigenvalues (which are important for physical applications) are new, to the best of our knowledge; so is the suggested connection between stability and Lyapunov stability (section 4).

The structure of this paper is as follows. In section 2 we present a derivation of differential equations for stability eigenvectors and exponents, and some conclusions based on the resulting equations. Section 3 provides a derivation of differential equations for the Lyapunov exponents and their corresponding eigenvectors; an analysis of the resulting equations is presented as well. Section 4 deals with the connection between Lyapunov and stability exponents, and presents a conjecture that these sets of exponents are generically equal. Section 5 presents a differential formulation of the standard (Bennetin et al. [11–12] and Shimada et al. [13]) method for obtaining Lyapunov exponents. It also shows how the method can be improved to yield more accurate results and how it can be used to find stability exponents and eigenvectors. An analysis of the resulting differential formulation reveals several properties of the variables associated with this method. Section 6 presents numerical results with emphasis on an extrapolation method, which enables us to obtain Lyapunov exponents from restricted amounts of data. Section 7 offers a summary and conclusions.

2. Equations of motion for the stability eigenvalues and eigenvectors

In the present section we derive differential equations of motion for the eigenvalues and eigenvectors of the stability matrix $\mathbf{M}(t,0)$, cf. eqs. (1.4)–(1.5). As mentioned in the introduction, we assume the spectrum of $\mathbf{M}$ to be nondegenerate, in the sense that $\lambda_1 > \lambda_2 > \cdots > \lambda_n$. Subsection 2.1 presents the derivation of the equations and subsection 2.2 is devoted to an analysis of the resulting equations.

2.1. *The differential equations*

Denote the right eigenvector of $\mathbf{M}(t,0)$ corresponding to $\lambda_i(t)$ by $|e_i(t)\rangle$ and the corresponding left eigenvector by $\langle e_i(t)|$. It follows from eqs. (1.3)–(1.5) that

$$\dot{\mathbf{M}}(t,0) = \mathbf{G}(x(t))\mathbf{M}(t,0). \tag{2.1}$$

We now use perturbation theory to express $|e_i(t+\tau)\rangle$ and $\lambda_i(t+\tau)$ in terms of $|e_i(t)\rangle$ and $\lambda_i(t)$ to the first order in τ. To this order

$$\mathbf{M}(t+\tau,0) = \mathbf{M}(t,0) + \tau\mathbf{G}(x(t))\mathbf{M}(t,0). \tag{2.2}$$

By definition of the eigenvectors and eigenvalues

$$\mathbf{M}(t+\tau,0)|e_i(t+\tau)\rangle = \mathrm{e}^{\lambda_i(t+\tau)(t+\tau)}|e_i(t+\tau)\rangle, \tag{2.3}$$

a similar relation holding for $\tau = 0$. We also assume the normalization

$$\langle e_i(t) | e_j(t) \rangle = \delta_{ij}, \quad 1 \le i, j \le n. \tag{2.4}$$

Since nondegeneracy is assumed one can expand $|e_i(t+\tau)\rangle$ in terms of the complete set $\{|e_j(t)\rangle\}$. To first order in τ

$$|e_i(t+\tau) = |e_i(t)\rangle + \tau \sum_{j=1}^{n} c_{ij} |e_i(t)\rangle, \tag{2.5}$$

where c_{ij} are time-dependent expansion coefficients. To first order in τ we have

$$\lambda_i(t+\tau) = \lambda_i(t) + \tau \frac{\mathrm{d}\lambda_i}{\mathrm{d}t}, \quad 1 \le i \le n. \tag{2.6}$$

Substituting eqs. (2.2), (2.5) and (2.6) in eq. (2.3) we obtain, to the first order in τ,

$$\begin{aligned} &\mathbf{M}(t,0)|e_i(t)\rangle + \tau \mathbf{G}(x(t))\mathbf{M}(t,0)|e_i(t)\rangle + \tau \sum_{j=1}^{n} c_{ij}\mathbf{M}(t,0)|e_j(t)\rangle \\ &\quad = \mathrm{e}^{\lambda_i(t)}|e_i(t)\rangle + \tau \mathrm{e}^{\lambda_i(t)t}\left[\frac{\mathrm{d}(\lambda_i(t)t)}{\mathrm{d}t}|e_i\rangle + \sum_{j=1}^{n} c_{ij}|e_j(t)\rangle\right]. \end{aligned} \tag{2.7}$$

Upon multiplying eq. (2.7) by $\langle e_j|$ from the left, we obtain (using $\mathbf{M}(t,0)|e_k(t)\rangle = \mathrm{e}^{\lambda_k(t)t}|e_k(t)\rangle$)

$$c_{ij} = \frac{\langle e_j(t)|\mathbf{G}(x(t))|e_i(t)\rangle}{1 - \mathrm{e}^{(\lambda_j(t)-\lambda_i(t))\tau}}, \quad \text{for } j \ne i. \tag{2.8}$$

Multiplying eq. (2.7) by $\langle e_i(t)|$ from the left we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}(\lambda_i(t)t) = \langle e_i(t)|\mathbf{G}(x(t))|e_i(t)\rangle. \tag{2.9}$$

Notice that (as is usual in perturbation theory) c_{ii} has not been fixed. Since the value of c_{ii} only determines the normalization and phase of $|e_j\rangle$, we may assume without loss of generality that $c_{ii} = 0$. Noting from eq. (2.5) that when $\tau \to 0$

$$\frac{\mathrm{d}}{\mathrm{d}t}|e_i(t)\rangle = \sum_{j=1}^{n} c_{ij}|e_j(t)\rangle \tag{2.10}$$

and, using (2.8) and $c_{ii} = 0$, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}|e_i(t)\rangle = \sum_{j \ne i} \frac{|e_j(t)\rangle\langle e_j(t)|\mathbf{G}(x(t))|e_i(t)\rangle}{1 - \mathrm{e}^{(\lambda_j(t)-\lambda_i(t))t}}. \tag{2.11}$$

A similar procedure for the left eigenvectors leads to

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle e_i(t)| = \sum_{j\neq i} \frac{\langle e_i(t)|\mathbf{G}(x(t))|e_j(t)\rangle\langle e_j(t)|}{\mathrm{e}^{(\lambda_i(t)-\lambda_j(t))t}-1}. \tag{2.12}$$

Eqs. (2.9), (2.11) and (2.12) are the sought differential equations for the eigenvalues and eigenvectors of the stability matrix $\mathbf{M}(t,0)$.

2.2. *Analysis of the equations for the eigenvalues and eigenvectors*

In the present subsection we use eqs. (2.9), (2.11) and (2.12) of the previous subsection to elucidate some properties of the eigenvalues and eigenvectors of $\mathbf{M}(t,0)$. We assume that beyond a time t_0 there exists a real number Δ such that

$$|\lambda_i(t)-\lambda_j(t)| > \Delta, \quad \text{for } t>t_0; \quad i\neq j. \tag{2.13}$$

When this property holds we call the spectrum strongly nondegenerate. Consider first eq. (2.12). Since the spectrum is assumed to be strongly nondegenerate it is obvious that

$$(\lambda_i(t)-\lambda_k(t))t \xrightarrow[t\to\infty]{} \infty, \quad \text{when } i<k, \tag{2.14a}$$

$$(\lambda_i(t)-\lambda_k(t))t \xrightarrow[t\to\infty]{} -\infty, \quad \text{when } i>k. \tag{2.14b}$$

Since, by eq. (2.11),

$$\frac{\mathrm{d}}{\mathrm{d}t}|e_i\rangle = \sum_{j\neq i} S_{ij}(t)\cdot|e_j\rangle, \tag{2.15}$$

where $S_{ij}(t)$ can be inferred from eq. (2.11), it follows that the matrix $\mathbf{e}$, whose columns are defined to be the vectors $|e_i\rangle$, has a time-independent determinant (the time derivative of a column is a linear combination of the other columns). Thus the norm of a single $|e_i\rangle$ can become unbounded in time if and only if the determinant of the matrix $\hat{\mathbf{e}}$, whose columns are the normalized right eigenvectors, becomes very small. This case is argued in section 4 to be nongeneric and thus we assume the $|e_i\rangle$'s to have bounded norms for all $t>t_0$. In this case one also has (by section 4) $\langle\hat{e}_i(t)|\hat{e}_j(t)\rangle$ of order unity, where $|\hat{e}_i(t)\rangle$ and $\langle\hat{e}_j(t)|$ are the normalized $|e_i(t)\rangle$ and $\langle e_j(t)|$, respectively. It follows, using eq. (2.4) that both $\langle e_i(t)|$ and $|e_i(t)\rangle$ have bounded norms. Assuming further that the attractor is bounded, it follows that the matrix $\mathbf{G}(x)$ is bounded. Using now eq. (2.14), it follows from eqs. (2.11)–(2.12) that

$$\frac{\mathrm{d}\langle e_1(t)|}{\mathrm{d}t} = \mathcal{O}(\mathrm{e}^{(\lambda_2-\lambda_1)t}),$$

$$\frac{\mathrm{d}\langle e_2(t)|}{\mathrm{d}t} = -\langle e_2(t)|\mathbf{G}(x(t))|e_1(t)\rangle\langle e_1(t)| + \mathcal{O}(\max(\mathrm{e}^{(\lambda_2-\lambda_1)t},\mathrm{e}^{(\lambda_3-\lambda_2)t})),$$

$$\frac{\mathrm{d}\langle e_3(t)|}{\mathrm{d}t} = -\langle e_3(t)|\mathbf{G}(x(t))|e_1(t)\rangle\langle e_1(t)| - \langle e_3|\mathbf{G}(x(t))|e_2(t)\rangle\langle e_2(t)| + \mathcal{O}(\max(\mathrm{e}^{(\lambda_3-\lambda_2)t},\mathrm{e}^{(\lambda_4-\lambda_3)t})),$$

and so on.

It follows that $\langle e_1(t)|$ converges to a fixed vector $\langle e_1|$, the subspace spanned by $\langle e_1(t)|$ and $\langle e_2(t)|$ converges to a fixed subspace, so does the subspace spanned by $\langle e_1(t)|$, $\langle e_2(t)|$ and $\langle e_3(t)|$, etc. The convergence rate of these respective subspaces is exponential. A similar conclusion can be drawn using eqs. (2.11) and (2.13). Here $|e_n(t)\rangle$ converges exponentially to a fixed vector, $|e_{n-1}(t)\rangle$ and $|e_n(t)\rangle$ will asymptotically span a fixed subspace and so on.

Upon performing a Gram–Schmidt orthogonalization procedure on $\langle e_1(t)|, \ldots, \langle e_n(t)|$ (starting from $\langle e_1|$), one obtains an orthonormal set of vectors $\langle g_1(t)|, \langle g_2(t)|, \ldots, \langle g_n(t)|$ such that $\langle g_1(t)| \propto \langle e_1(t)|$, $\langle g_1(t)|$ and $\langle g_2(t)|$ span the same subspace spanned by $\langle e_1|$ and $\langle e_2|$ and so on. The vectors $\{\langle g_i|; 1 \le i \le n\}$ are asymptotically constant. Performing the same procedure on $|e_1(t)\rangle, \ldots, |e_n(t)\rangle$, but this time starting from $|e_n(t)\rangle$, then orthonormalizing $|e_{n-1}(t)\rangle$ with respect to $|e_n(t)\rangle$, etc., we obtain an orthonormal set $|f_1(t)\rangle, \ldots, |f_n(t)\rangle$. The vector $|f_i(t)\rangle$ is the transpose of $\langle g_i(t)|$. To see this, note that by construction

$$\langle g_i(t)| = \sum_{j=1}^{i} r_{ij}(t) \langle e_j(t)|, \tag{2.16a}$$

$$|f_k(t)\rangle = \sum_{j=k}^{n} s_{kj}(t) |e_j(t)\rangle, \tag{2.16b}$$

where $r_{ij}(t)$ and $s_{kj}(t)$ are coefficients. Hence

$$\langle g_i(t)|f_k(t)\rangle = 0, \quad \text{for } i < k. \tag{2.17}$$

Thus, $\langle g_1|$ is orthogonal to $\langle |f_k\rangle; 2 \le k \le n\}$ and by the orthogonality of the $|f_k\rangle$'s, $\langle g_1|$ is the adjoint of $|f_1\rangle$. The proof is carried further by induction. By eq. (2.16b), provided $s_{kk}(t)$ does not vanish, $\mathbf{M}(t,0)|f_k(t)\rangle$ is dominated by $s_{kk}(t)\,e^{\lambda_k t}|e_k(t)\rangle$. Thus

$$\|\mathbf{M}(t,0)|f_k(t)\rangle\| \approx e^{\lambda_k t} \tag{2.18}$$

for asymptotic times.

A comparison of eq. (2.18) and eq. (1.8) suggests that $\{|f_k(t)\rangle\}$ are the eigenvectors of $\mathbf{M}^\dagger\mathbf{M}$. Further connections between the eigenvectors of $\mathbf{M}$ and those of $\mathbf{M}^\dagger\mathbf{M}$ are analysed in the forthcoming sections.

Finally, we turn to eq. (2.4), which can be rewritten as

$$\lambda_i(t) = \frac{1}{t} \int_0^t dt' \langle e_i(t')|\mathbf{G}(x(t'))|e_i(t')\rangle, \tag{2.19}$$

i.e. $\lambda_i(t)$ is the time average of the diagonal element $\langle e_i|\mathbf{G}|e_i\rangle$. Assuming the average of this element on the attractor exists and is independent of the trajectory ("ergodicity") and denoting it by λ_i we have

$$\lambda_i(t) = \lambda_i + \frac{c_i(t)}{t}, \tag{2.20}$$

where c_i is a time-dependent function. Denoting by b_i the average of $c_i(t)$ (which we assume to exist) we have

$$\lambda_i(t) = \lambda_i + \frac{b_i + \xi_i(t)}{t}, \tag{2.21}$$

where $\xi_i(t)$ has zero average by definition. We have numerical evidence (see section 6) that $\xi_i(t)$ averages out to zero on relatively short time scales. Thus, on a coarse time scale,

$$\lambda_i(t) \approx \lambda_i + \frac{b_i}{t}, \tag{2.22}$$

which shows that the rate of convergence of $\lambda_i(t)$ to λ_i is relatively slow.

3. Equations of motion for the Lyapunov eigenvalues and eigenvectors

In the previous section we derived differential equations for the eigenvalues and eigenvectors of the stability matrix $\mathbf{M}(t,0)$. Although we presented plausibility arguments for the connections between these quantities and the Lyapunov exponents and eigenvectors, these connections are not of general nature. In fact (see section 4) it is easy to construct examples in which the Lyapunov exponents do not equal the stability exponents. Nevertheless we find that it is still interesting to investigate the stability of a system as a problem in its own right.

In the present section we turn to the problem of Lyapunov stability. In subsection 3.1 we derive differential equations for the eigenvalues and eigenvectors of $\mathbf{M}^\dagger\mathbf{M}$ and in subsection 3.2 we analyze the results. Since the method used here and that used in the previous section are identical we shall present the derivations very briefly.

3.1. *The differential equations*

Since $\mathbf{M}^\dagger(t,0)\mathbf{M}(t,0)$ is a symmetric matrix, it can be represented as a square [18, 19] of a symmetric matrix $\mathbf{S}(t,0)$,

$$\mathbf{M}^\dagger\mathbf{M} = \mathbf{S}^2, \tag{3.1}$$

where the time dependence of the matrices was suppressed (as done below) for notational simplicity. Define a matrix $\mathbf{R}$,

$$\mathbf{R} \equiv \mathbf{M}\mathbf{S}^{-1}. \tag{3.2}$$

It is easy to see that $\mathbf{R}$ is a real unitary matrix (a rotation matrix). Also $\mathbf{M} = \mathbf{R}\mathbf{S}$. Thus we can regard the action of $\mathbf{M}$ on a (initial) vector as that of stretching (by $\mathbf{S}$; assume for example that the initial vector is an eigenvector of $\mathbf{S}$) followed by a rotation. It follows from eq. (2.1) that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{M}^\dagger\mathbf{M} = \mathbf{M}^\dagger(\mathbf{G}^\dagger + \mathbf{G})\mathbf{M}. \tag{3.3}$$

Using eqs. (3.1) and (3.2),

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{S}^2 = \mathbf{S}\mathbf{R}^\dagger(\mathbf{G}^\dagger + \mathbf{G})\mathbf{R}\mathbf{S}. \tag{3.4}$$

Define now the eigenvalues of $\mathbf{S}$ to be $e^{\mu_i(t)t}$ ($\mu_i(t)$ are obviously real) and the corresponding eigenvectors by $|\mu_i\rangle$. Obviously: $\langle\mu_i| = |\mu_i\rangle^\dagger$. Applying the method of the previous section leads to the following differential equations for $|\mu_i\rangle$ and μ_i:

$$\frac{\mathrm{d}}{\mathrm{d}t}|\mu_i\rangle = \sum_{k\neq i} \frac{1}{\mathrm{e}^{(\mu_i-\mu_k)t} - \mathrm{e}^{(\mu_k-\mu_i)t}} \langle\mu_k|\mathbf{R}^\dagger(\mathbf{G}^\dagger+\mathbf{G})\mathbf{R}|\mu_i\rangle \tag{3.5}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{2\mu_i t} = \mathrm{e}^{2\mu_i t}\langle\mu_i|\mathbf{R}^\dagger(\mathbf{G}^\dagger+\mathbf{G})\mathbf{R}|\mu_i\rangle, \tag{3.6}$$

where the time dependence of $\mu_i(t)$ has been omitted. It follows from (3.6) that

$$\mu_i(t) = \frac{1}{t}\int_0^t \langle\mu_i|\mathbf{R}^\dagger\cdot\frac{\mathbf{G}+\mathbf{G}^\dagger}{2}\cdot\mathbf{R}|\mu_i\rangle. \tag{3.7}$$

Eqs. (3.5)–(3.6) are the sought differential equations for the eigenvalues and eigenvectors of $\mathbf{M}^\dagger\mathbf{M}$. Eq. (3.7) shows that the Lyapunov exponents, like the stability exponents in the previous section are time averages of some diagonal matrix element. By its definition (see also next subsection), the ith Lyapunov exponent is

$$\mu_i = \lim_{t\to\infty}\mu_i(t). \tag{3.8}$$

3.2. *Analysis of the differential formulation*

Since $\mathbf{S}$ is a real symmetric matrix, we can assume without loss of generality that the $|\mu_i\rangle$'s are an orthonormal set. It is easy to check that eq. (3.5) leaves the norm of these eigenvectors unchanged in time. Since $\mathbf{R}$ is a unitary matrix, it follows that $\mathbf{R}|\mu_i\rangle$ has a norm of unity. Assuming that the attractor is bounded, it follows that the matrix element $\langle\mu_k|\mathbf{R}^\dagger(\mathbf{G}^\dagger+\mathbf{G})\cdot\mathbf{R}|\mu_i\rangle$, for $1\leq i,k\leq n$, is bounded for all times. Assuming further a nondegenerate Lyapunov spectrum, it follows from eq. (3.5) that the eigenvectors $|\mu_i\rangle$ converge to fixed vectors, $|f_i\rangle$, at an exponential rate,

$$\frac{\mathrm{d}}{\mathrm{d}t}|\mu_i\rangle \propto \max\{\mathrm{e}^{(\mu_{i+1}-\mu_i)t};\ \mathrm{e}^{(\mu_i-\mu_{i-1})t}\}, \tag{3.9}$$

where an ordering of the μ_i's has been assumed: $\mu_1>\mu_2>\mu_3>\cdots>\mu_n$.

Arguments, similar to those presented in subsection 2.1 lead to the following conjecture for the form of $\mu_i(t)$ [based on eq. (3.7)]:

$$\mu_i(t) = \mu_i + \frac{b_i+\xi_i(t)}{t}, \tag{3.10}$$

where $\mu_i \equiv \lim_{t\to\infty}\mu_i(t)$ are the Lyapunov exponents, b_i are constants and $\xi_i(t)$ are "noise" terms whose time average is zero. This conjecture has been tested numerically (see section 6).

So far we have shown that the eigenvectors of $\mathbf{M}^\dagger\mathbf{M}$ converge to fixed vectors $|\mu_i(\infty)\rangle$. In order to show that the latter are indeed the Lyapunov vectors $|f_i\rangle$ it remains to show that $\|\mathbf{M}(t,0)|\mu_i(\infty)\rangle\| \propto \mathrm{e}^{\mu_i t}$. This is done below. Eq. (3.5) can be rewritten as

$$|\dot{\mu}_i(t)\rangle = \sum_{k\neq i} s_{ik}|\mu_k(t)\rangle, \tag{3.11}$$

where s_{ik} can be identified from eq. (3.5). For large enough times

$$s_{ik} = u_{ik}(t)\,\mathrm{e}^{-|\mu_i - \mu_k|t}, \tag{3.12}$$

where $u_{ik}(t)$ is bounded. Integrating eq. (3.11) from a time t to infinity yields

$$|\mu_i(\infty)\rangle = |\mu_i(t)\rangle + \sum_{k\neq i} \int_t^{\infty} s_{ik}(t')|\mu_k(t')\rangle \mathrm{d}t'. \tag{3.13}$$

Iterating this integral equation,

$$|\mu_i(\infty)\rangle = |\mu_i(t)\rangle + \sum_{k\neq i} \int_t^{\infty} s_{ik}(t')\,\mathrm{d}t'|\mu_k(t)\rangle + \sum_{k\neq i}\sum_{l\neq k} \int_t^{\infty} \mathrm{d}t' \int_{t'}^{\infty} \mathrm{d}t'' s_{ik}(t') s_{kl}(t'')|\mu_1(t)\rangle + \cdots . \tag{3.14}$$

The coefficients of the various vectors on the right side of eq. (3.14) can be bounded (and estimated) using eq. (3.12). The result is

$$\begin{aligned} |\mu_i(\infty)\rangle = {} & |\mu_i(t)\rangle + \sum_{k\neq i} g_{ik} \frac{1}{|\mu_{ik}|} \mathrm{e}^{-|\mu_{ik}|t} |\mu_k(t)\rangle \\ & + \sum_{i\neq k}\sum_{l\neq k} g_{ikl} \frac{1}{|\mu_{kl}|(|\mu_{ik}| + |\mu_{kl}|)} \mathrm{e}^{-(|\mu_{ik}|+|\mu_{kl}|)t} |\mu_l(t)\rangle + \cdots, \end{aligned} \tag{3.15}$$

where $\mu_{ik} \equiv \mu_i - \mu_k$ and obviously the g's are order one. Applying $\mathbf{S}(t,0)$ to eq. (3.15) yields

$$\begin{aligned} \mathbf{S}(t,0)|\mu_i(\infty)\rangle = \mathrm{e}^{\mu_i(t)t} \Big\{ & |\mu_i(t)\rangle + \sum_{k\neq i} \frac{g_{ik}}{|\mu_{ik}|} \mathrm{e}^{(-|\mu_{ik}|+\mu_{ki})t} |\mu_k(t)\rangle \\ & + \sum_{k\neq i}\sum_{l\neq k} g_{ikl} \frac{1}{|\mu_{kl}|(|\mu_{ik}| + |\mu_{kl}|)} \mathrm{e}^{-(|\mu_{ik}|+|\mu_{kl}|)t+\mu_{li}t} |\mu_l(t)\rangle + \cdots \Big\}. \end{aligned} \tag{3.16}$$

A term like $|\mu_{ik}| + \mu_{ik}$ is positive unless $\mu_{ik} < 0$, i.e. $i > k$ (since then $\mu_k > \mu_i$). Similarly, $|\mu_{ik}| + |\mu_{kl}| + \mu_{li}$ is positive unless $i > k > l$. Consequently, the only terms in the curly bracket in eq. (3.16) that are not exponentially small are

$$|\mu_i(t)\rangle + \sum_{k<i} \frac{g_{ik}}{|\mu_{ik}|} |\mu_k(t)\rangle + \sum_{l<k<i} \frac{g_{ikl}}{|\mu_{kl}|(|\mu_{ik}| + |\mu_{kl}|)} |\mu_1\rangle + \cdots . \tag{3.17}$$

It is easy to see that this vector has a norm which is larger than one (all the vectors in the summation are orthogonal to $|\mu_i(t)\rangle$) and bounded. Hence

$$\| \mathbf{S}(t,0)|\mu_i(\infty)\rangle \| = \mathrm{e}^{\mu_i(t)t} N, \tag{3.18}$$

where N is order 1. Since $\mathbf{R}$ is unitary it follows that for $\mathbf{M}$ $(= \mathbf{RS})$

$$\| \mathbf{M}(t,0)|\mu_i(\infty)\rangle \| = \mathrm{e}^{\mu_i(t)t} N, \tag{3.19}$$

which shows that $|\mu_i(\infty)\rangle$ are indeed the Lyapunov eigenvectors.

4. Relations between Lyapunov and stability eigenvalues and eigenvectors

As mentioned in section 2, there is no theorem (to the best of our knowledge) regarding the long time properties of the stability eigenvalues and eigenvectors. In the present section we wish to present an analysis of some properties of stability eigenvalues and eigenvectors and relate them to properties of the Lyapunov eigenvalues and eigenvectors. We believe that the stability matrix $\mathbf{M}(t,0)$, being a generalization to unsteady states of the steady state stability matrix, should be of importance (and may be useful for nonlinear stability analysis around unsteady states) and its investigation worthwhile. In what follows we assume the Lyapunov spectrum to be nondegenerate. Consider the time-dependent matrix $\mathbf{M} = \mathbf{RS}$. Let $|e_i(t)\rangle$ be a right (normalized) eigenvector of $\mathbf{M}$, as in the previous sections. $|e_i(t)\rangle$ can be represented as a linear combination of $\{|\mu_i(t)\rangle\}$, the eigenvectors of $\mathbf{S}$,

$$|e_i(t)\rangle = \sum_{j=1}^{n} c_{ij}|\mu_j(t)\rangle, \tag{4.1}$$

where c_{ij} are time-dependent coefficients. Applying $\mathbf{M} = \mathbf{RS}$ to both sides of eq. (4.1) we obtain

$$e^{\lambda_i t}\sum_{j=1}^{n} c_{ij}|\mu_j(t)\rangle = \mathbf{R}\sum_{j=1}^{n} c_{ij}\,e^{\mu_j t}|\mu_j(t)\rangle, \tag{4.2}$$

where $\{\lambda_i\}$ and $\{\mu_i\}$ are the time-dependent stability and Lyapunov exponents, respectively. It follows from eq. (4.2) that

$$\sum_{j=1}^{n} \langle\mu_k|\mathbf{R}|\mu_j\rangle\, e^{\mu_j t}c_{ij} = e^{\lambda_i t}c_{ik}. \tag{4.3}$$

Denoting: $R_{kj} = \langle\mu_k|\mathbf{R}|\mu_j\rangle$, defining a diagonal matrix $\mathbf{D}$ whose entries are $e^{\mu_j t}$ and a vector $\boldsymbol{c}^{(i)}$ whose elements are c_{ij}, we obtain

$$\mathbf{RD}\boldsymbol{c}^{(i)} = e^{\lambda_i t}\boldsymbol{c}^{(i)}. \tag{4.4}$$

Hence

$$\mathbf{D}\boldsymbol{c}^{(i)} = e^{\lambda_i t}\mathbf{R}^{\dagger}\boldsymbol{c}^{(i)}, \tag{4.5}$$

$$c_k^{(i)} = \frac{e^{\lambda_i t}}{e^{\mu_k t}}(\mathbf{R}^{\dagger}\boldsymbol{c}^{(i)})k. \tag{4.6}$$

If $\lambda_i = \mu_i$ and nondegeneracy of the $\{\mu_i\}$ is assumed (and since $\|\mathbf{R}^{\dagger}\boldsymbol{c}^{(i)}\| = 1$ by the unitarity of $\mathbf{R}$), we find that $c_k^{(i)}$ is exponentially small for $i > k$.

When $i < k$ we can rewrite (4.6),

$$(\mathbf{R}^{\dagger}\boldsymbol{c}^{(i)})_k = \frac{e^{\mu_k t}}{e^{\lambda_i t}}c_k^{(i)}, \tag{4.7}$$

i.e. $(\mathbf{R}^{\dagger}\boldsymbol{c}^{(i)})_k$ is exponentially small for $k > i$. Thus to exponential accuracy

$$\mathbf{R}^{\dagger}\boldsymbol{c}^{(i)} \approx \begin{pmatrix} a_1^{(i)} \\ \vdots \\ a_i^{(i)} \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \tag{4.8}$$

where $\{a_k^{(i)};\ 1 \le k \le i\}$ are the elements of $\mathbf{R}^{\dagger}\boldsymbol{c}^{(i)}$ which need not be exponentially small by the above argument. Hence

$$\boldsymbol{c}^{(i)} \approx \mathbf{R}\cdot\boldsymbol{a}^{(i)}, \tag{4.9}$$

$\boldsymbol{a}^{(i)}$ being the vector on the right side of eq. (4.8). However $c_{(k)}^{(i)}$ is vanishingly small for $k < i$. Thus:

$$0 \approx \sum_{j=1}^{i} R_{kj}a_j^{(i)}, \quad \text{for } k < i. \tag{4.10}$$

All in all we obtain

$$c_i^{(i)}\delta_{ki} \approx \sum_{j=1}^{i} R_{kj}a_j^{(i)}, \quad \text{for } 1 \le k \le i. \tag{4.11}$$

Denote by $\tilde{\mathbf{R}}_{(i)}$ the matrix consisting of the elements of R_{kj} for $1 \le k, j \le i$, $\tilde{\boldsymbol{a}}_{(i)}$ the "reduced" vector consisting of the first i elements of $\boldsymbol{a}$. Then by (4.11)

$$\tilde{\boldsymbol{a}}_{(i)} = c_i^{(i)} \tilde{\mathbf{R}}_{(i)}^{-1} \cdot \hat{\mathbf{e}}_{(i)}, \tag{4.12}$$

provided $\tilde{\mathbf{R}}_{(i)}^{-1}$ exists; moreover we demand that $\det \tilde{\mathbf{R}}_{(i)}$ not be exponentially small for the above construction to be valid. $\hat{\mathbf{e}}_i$ is a unit vector whose ith element is 1 and all other elements are zero. The value of $c_i^{(i)}$ is determined by the normalization condition $\|\boldsymbol{a}_{(i)}\| = 1$. What happens when $\det \tilde{\mathbf{R}}_{(i)}$ is exponentially small? To see this consider the next paragraph.

Assume that there is a stability exponent λ_i which satisfies

$$\mu_{i+1} + \Delta < \lambda_i < \mu_i - \Delta \tag{4.13}$$

for some (small) fixed value of Δ and for some large enough value of t, for which $e^{-\Delta t}$ can be considered exponentially negligible. Returning to eq. (4.6), we find that $c_k^{(i)} \approx 0$ for $k \le i$ and $(\mathbf{R}^\dagger \boldsymbol{c}^{(i)})_k \approx 0$ for $k > i$. Consequently,

$$\tilde{\mathbf{R}}_{(i)} \cdot \tilde{\boldsymbol{a}}_{(i)} \approx 0, \tag{4.14}$$

where $\tilde{\mathbf{R}}_{(i)}$ is defined as before and $\tilde{\boldsymbol{a}}_{(i)}$ is, as before, the projection of $\mathbf{R}^\dagger \boldsymbol{c}^{(i)}$ on the first i directions. Thus $\det \tilde{\mathbf{R}}_{(i)} \approx 0$, i.e. this determinant must be exponentially small. Since we believe that in the generic case, there is no reason for $\det \tilde{\mathbf{R}}_i \approx 0$, we conjecture that for large enough times t, $\det \tilde{\mathbf{R}}_{(i)}$ is generically $\mathcal{O}(1)$ for most of the time. In particular, $\lambda_1 = \mu_1$ if R_{11} is not exponentially small for exponentially large times.

Notice that when $\det \tilde{\mathbf{R}}_i \approx 0$, $\boldsymbol{c}^{(i)}$ corresponding to the right eigenvector has exponentially vanishing elements $c_k^{(i)} \approx 0$ for $k \le i$. A similar analysis shows that $d_k^{(i)}$ corresponding to the left ith eigenvector satisfies $d_k^{(i)} \approx 0$ for $k > i$. Thus the left and right eigenvectors are "almost orthogonal" (and the set of right eigenvectors is "almost linearly dependent"):

$$\langle \hat{e}_i(t) | \hat{e}_i(t) \rangle \approx e^{-\gamma_i t}, \tag{4.15}$$

where the hat means normalized. If we define $|e_i(t)\rangle = |\hat{e}_i(t)\rangle$ and demand $\langle e_i(t)|e_i(t)\rangle = 1$ we obtain that the norm of $\langle e_i(t)|$ is exponentially large. This case contradicts the assumptions made in section 2 leading to equality of Lyapunov and stability exponents (these assumptions are equivalent to assuming $\det \tilde{\mathbf{R}}_i \approx \mathcal{O}(1)$ for all i).

It is instructive to consider a two-dimensional example to clarify the conjecture just made. Let us write in this case (using the representation $\mathbf{M}$ in the $|\boldsymbol{\mu}_i\rangle$ space, which is asymptotically fixed)

$$\mathbf{M}(t,0) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} e^{\mu_1 t} & 0 \\ 0 & e^{\mu_2 t} \end{pmatrix}, \tag{4.16}$$

where θ is a time-dependent angle. The characteristic equation for $\mathbf{M}$ is

$$\Lambda^2 - \Lambda \cos\theta (e^{\mu_1 t} + e^{\mu_2 t}) + e^{(\mu_1 + \mu_2)t} = 0. \tag{4.17}$$

It is obvious that when $\cos\theta = 0$ ($R_{11} = 0$) the stability spectrum is degenerate ($|\Lambda_+| = |\Lambda_-|$). When $\cos\theta \approx e^{-\alpha t}$, α being an $\mathcal{O}(1)$ number, the spectrum is not degenerate, yet the stability exponents differ from the Lyapunov exponents. Let us analyze this possibility further. It follows from eq. (4.17) that

$$\Lambda_\pm = \frac{b \pm \sqrt{b^2 - 4c}}{2}, \tag{4.18}$$

where $b \equiv \cos\theta (e^{\mu_1 t} + e^{\mu_2 t})$ and $c = e^{(\mu_1 + \mu_2)t}$. Hence

$$\frac{\Lambda_-}{\Lambda_+} = \frac{1 - \sqrt{1 - 4c/b^2}}{1 + \sqrt{1 - 4c/b^2}}. \tag{4.19}$$

If we demand Λ_- and Λ_+ to be "exponentially separated" for large times (i.e. Λ_-/Λ_+ is exponentially small), it is necessary to have $4c/b^2 \propto e^{-\alpha t}$ where α is an $\mathcal{O}(1)$ number. Let's assume, for simplicity, that $4c/b^2 = 4A\,e^{-\alpha t}$, where A is a fixed number. Recalling the definitions of c and b

we find

$$\frac{e^{(\mu_2-\mu_1)t}}{\cos^2\theta(1+e^{(\mu_2-\mu_1)t})^2}=A\,e^{-\alpha t}, \tag{4.20}$$

thus

$$\cos\theta \approx \frac{1}{\sqrt{A}}e^{((\mu_2-\mu_1+\alpha)/2)t} \tag{4.21}$$

and

$$\Lambda_+ \approx \frac{1}{\sqrt{A}}e^{((\mu_1+\mu_2+\alpha)/2)t}, \tag{4.22}$$

$$\Lambda_- \approx \sqrt{A}\,e^{((\mu_1+\mu_2-\alpha)/2)t}. \tag{4.23}$$

Notice that, since $|\cos\theta|<1$, we must have

$$\alpha \le \mu_1-\mu_2. \tag{4.24}$$

Thus it seems that one can have a nondegenerate stability spectrum which is different from the corresponding Lyapunov spectrum. Let us calculate now the corresponding $\mathbf{G}$ [cf. eq. (2.1)],

$$\mathbf{G}=\dot{\mathbf{M}}\mathbf{M}^{-1}, \tag{4.25}$$

or

$$\mathbf{G}=\dot{\mathbf{R}}\mathbf{R}^{-1}+\mathbf{R}\dot{\mathbf{S}}\mathbf{S}^{-1}\mathbf{R}^{-1}. \tag{4.26}$$

Assuming the μ_i's to be fixed in the above example, we obtain for asymptotic times (up to exponentially small corrections)

$$\mathbf{G}\approx\begin{pmatrix}\mu_2 & 0\\ 0 & \mu_1\end{pmatrix}, \tag{4.27}$$

which shows that $\mathbf{G}$ tends to a constant matrix. In an ergodic (and smooth) system this means that $\mathbf{G}$ *is* a fixed matrix on the attractor and the stability eigenvalues are μ_1 and μ_2. Thus the assumption that the stability spectrum is nondegenerate for large times and different from the Lyapunov spectrum has led to a contradiction. On the basis of this example, we conjecture that when the stability spectrum is nondegenerate, the dynamical system is smooth and ergodic and the attractor is bounded – the Lyapunov exponents equal the stability exponents. We also believe this to be the generic case. Indeed, in all systems we have investigated numerically (see section 6), this was the case. The geometrical meaning of this conjecture is that, generically, stretching and shrinking of relative separations of trajectories, starting off at a given point, happen in fixed directions in phase space (in the sense described in section 2).

Another consequence of the type of analysis presented in this section is that the Lyapunov exponents do not depend (generically) on the (initial) point in phase space with respect to which they are computed. Consider a trajectory $x(t)$ which is a solution of a dynamical system. The stability matrix $\mathbf{M}$ satisfies

$$\mathbf{M}(t_2,t_1)\cdot\mathbf{M}(t_1,t_0)=\mathbf{M}(t_2,t_0), \tag{4.28}$$

where the right side of eq. (4.28) corresponds to the stability with respect to reference point $x(t_0)$ and $\mathbf{M}(t_2,t_1)$ is related to stability with respect to $x(t_1)$. In the limit $t_2\to\infty$, the eigenvalues of $\mathbf{M}(t_2,t_0)$ and those of $\mathbf{M}(t_2,t_1)$ should (generically) be the same (to exponential accuracy) since $\mathbf{M}(t_1,t_0)$ is a fixed matrix. As shown before, if $\mathbf{M}(t_1,t_0)$ has some special properties (e.g. vanishing of some determinants of partial matrices) this may not be the case. Similarly, one can show that the Lyapunov exponents deduced from $\mathbf{M}^\dagger(t_2,t_1)\mathbf{M}(t_2,t_1)$ and those deduced from $\mathbf{M}^\dagger(t_2,t_0)\mathbf{M}(t_2,t_0)$ as $t_2\to\infty$ are (generically) the same. If a given trajectory is dense on the attractor and if the Lyapunov spectrum is not discontinuous on it, we can expect the Lyapunov spectrum to be a global property of the attractor.

5. Differential formulation of the standard method

In this section we present a differential formulation of the method suggested by Bennetin et al. [11, 12] and Shimada et al. [13]. Using this formulation we can, among other things, reproduce analytically some of the results which were proven by the above authors with the aid of geometrical methods. We find that the "finite time" Lyapunov exponents found the standard way have an error of order $1/t$. Expressions for the eigenvectors are found as well. In subsection 5.1 we review the standard method. Subsection 5.2 offers a derivation of the differential version. Subsection 5.3 presents an analysis of the results found in subsection 5.2. In subsection 5.4 we find the Lyapunov and stability exponents and the associated eigenvectors.

5.1. *Review of the standard method*

If one wishes to compute the first k Lyapunov exponents, using the methods of refs. 11–13, one first chooses k orthogonal vectors as initial conditions for eq. (1.3). The standard choice is $\hat{e}_1 = (1,0,0,\dots)$; $\hat{e}_2 = (0,1,0,0,\dots)$ etc. Eq. (1.3) is then solved up to time T for each of the initial conditions yielding vectors $y_1^{(1)}, y_2^{(1)}, \dots, y_k^{(1)}$. These vectors are orthonormalized, using a Gram–Schmidt procedure to yield

$$\begin{aligned}\hat{w}_1 &= \frac{y_1}{\| y_1 \|},\\ \hat{w}_2 &= \frac{y_2 - (y_2 \cdot w_1^+) w_1}{\| y_2 - (y_2 \cdot w_1^+) w_1 \|},\end{aligned} \tag{5.1}$$

and so on. The norms in the denominators are kept; denote them by $N_1^{(1)}$, $N_2^{(1)}$, $N_3^{(1)}$, etc. The procedure is repeated using $\hat{w}_i$ as initial conditions for eq. (3.1) and a time T of integration. The resulting vectors $y_1^{(2)}, y_2^{(2)}, \dots, y_k^{(2)}$ are orthonormalized again yielding norms $N_1^{(2)}, N_2^{(2)}, \dots, N_k^{(2)}$ in the denominators and a new set of orthonormal conditions. The process is repeated for r iterations. One defines then

$$\mu_{j(r)} = \frac{\sum_{m=1}^{r} \log N_j(m)}{rT}. \tag{5.2}$$

If $\mu_{j(r)}$ is close enough to $\mu_{j(r-1)}$ (according to a predetermined criterion) one identifies this quantity with the Lyapunov exponent μ_j. It has been shown that, indeed, when $r \to \infty$, $\mu_{j(r)} \to \mu_j$.

As mentioned before, this method converges relatively slowly and is impractical for systems having many degrees of freedom.

5.2. *The differential version*

The basic difference between the procedure in this subsection and that in the previous one is that here we take the time interval T to be infinitesimal and derive differential equations. Let $\hat{e}_i(0)$, $1 \le i \le n$ be an orthonormal set of n vectors, say $\hat{e}_{j,i} = \delta_{i,j}$ (the jth component of the ith vector). Let $\mathbf{M}$ be as defined in section (1). Define

$$e_i(t) = \mathbf{M}(t,0)\hat{e}_i(0) \tag{5.3}$$

and let $\hat{e}_i(t)$ be an orthonormalization of the set $\{e_i(t);\ 1 \le i \le n\}$ done by starting a Gram–Schmidt procedure with $i=1$ and proceeding through increasing indices to $i=n$. The vectors $\hat{e}_i(t)$ for t being an integer multiple of T obviously coincide with the $\hat{e}_i$'s in the previous subsection. By construction, there exists a triangular matrix $\mathbf{d}$ (i.e. $d_{ij} = 0$ for $i < j$) such that

$$e_i(t) = \sum_{j=1}^{i} d_{ij} \hat{e}_j(t). \tag{5.4}$$

For an infinitesimally small time interval τ (to first order in τ)

$$e_1(t+\tau) = (\mathbf{1} + \tau \mathbf{G}(x(t))) e_1(t). \tag{5.5}$$

Hence

$$\hat{e}_1(t+\tau) = \frac{e_1(t) + \tau \mathbf{G}(x(t)) e_1(t)}{\| e_1(t) + \tau \mathbf{G}(x(t)) e_1(t) \|}, \tag{5.6}$$

or

$$\begin{aligned}\hat{e}_1(t+\tau) = \hat{e}_1(t) + \tau\big(&\mathbf{G}(x(t)) \cdot \hat{e}_1(t)\\ &- \hat{e}_1(\hat{e}_1(t) \cdot \mathbf{G}(x(t)) \cdot \hat{e}_1(t))\big).\end{aligned}$$

Hence

$$\frac{\mathrm{d}\hat{e}_1}{\mathrm{d}t} = \mathbf{G}\hat{e}_1 - G_{11}\hat{e}_1, \tag{5.7}$$

where $G_{ij} \equiv (\hat{e}_i(t) \cdot \mathbf{G}(x(t))\hat{e}_j(t))$ and the time dependence of the vectors has been suppressed.

To find an equation for $\hat{e}_2$ we define

$$y_2(t+\tau) = (\mathbf{1} + \tau\mathbf{G}(x(t))\hat{e}_2(t)) \tag{5.8}$$

and

$$u_2(t+\tau) = y_2(t+\tau) - (y_2(t+\tau) \cdot \hat{e}_1(t+\tau))\hat{e}_1(t+\tau) \tag{5.9}$$

and

$$\hat{e}_2(t+\tau) = \frac{u_2(t+\tau)}{\|u_2(t+\tau)\|}. \tag{5.10}$$

An equivalent procedure would be the direct use of eq. (5.3). We have preferred to exploit the equivalence of eq. (5.3) and the $\hat{e}_i$'s as defined in 5.1. The result to first order in τ is

$$\hat{e}_2(t+\tau) = \hat{e}_2(t) + \tau(\mathbf{G}\hat{e}_2 - G_{22}\hat{e}_2 - (G_{12} + G_{21})\hat{e}_1),$$

or

$$\frac{\mathrm{d}\hat{e}_2}{\mathrm{d}t} = \mathbf{G}\hat{e}_2 - G_{22}\hat{e}_2 - (G_{12} + G_{21})\hat{e}_1. \tag{5.11}$$

Similarly it can be shown by induction that

$$\frac{\mathrm{d}\hat{e}_i}{\mathrm{d}t} = \mathbf{G}\hat{e}_i - G_{ii}\hat{e}_i - \sum_{j=1}^{i-1} (G_{ij} + G_{ji})\hat{e}_j. \tag{5.12}$$

Since $\mathbf{G}$ is assumed to be known at all times, eqs. (5.12) are a closed set of equations for the orthogonal set of vectors $\hat{e}_i(t)$.

It is convenient to find a set of equations for the variables d_{ij} as defined in eq. (5.4). Since the vectors $\hat{e}_i(t)$ are orthonormal it follows from eq. (3.4) that

$$d_{ij} = e_i(t) \cdot \hat{e}_j(t). \tag{5.13}$$

(We omit conjugate signs in scalar products since all vectors are real here.) Hence

$$\dot{d}_{ij} = \dot{e}_i \cdot \hat{e}_j + e_i \cdot \dot{\hat{e}}_j.$$

Since $\dot{e}_i = \mathbf{G} \cdot e_i$ by construction and $\dot{\hat{e}}_i$ is given by eq. (5.12), we obtain

$$\dot{d}_{ij} = \hat{e}_j\mathbf{G} \cdot e_i + e_i \cdot \left(\mathbf{G}\hat{e}_j - G_{jj}\hat{e}_j - \sum_{l=1}^{j-1} (G_{j,1} + G_{l,j})\hat{e}_l\right). \tag{5.14}$$

Using eq. (5.4) to express the e_i's in terms of $\hat{e}_j$'s and using the orthonormality of the $\hat{e}_i$'s we obtain

$$\dot{d}_{ij} = \sum_{k=1}^{n} d_{ik}\left(G_{jk} + G_{kj} - \sum_{l=1}^{j-1} (G_{jl} + G_{lj})\delta_{k,l}\right) - G_{jj}d_{ij}, \tag{5.15}$$

or

$$\dot{d}_{ij} = \sum_{k=1}^{n} (G_{jk} + G_{kj})d_{ik} - \sum_{k=1}^{j-1} (G_{jk} + G_{kj})d_{ik} - G_{jj}d_{ij}, \tag{5.16}$$

or

$$\dot{d}_{ij} = \sum_{k=j}^{n} d_{ik}(G_{jk} + G_{kj}) - G_{jj}d_{ij}.$$

Since $d_{ik} = 0$ for $i < k$ it follows that

$$\dot{d}_{ii} = G_{ii}\, d_{ii}, \tag{5.17}$$

and, in general,

$$\dot{d}_{ij} = \sum_{k=j}^{i} d_{ik}(G_{jk} + G_{kj}) - G_{jj}d_{ij}. \tag{5.18}$$

For $(n=3)$ these equations can be written in matrix form,

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} d_{11} & 0 & 0 \\ d_{21} & d_{22} & 0 \\ d_{31} & d_{32} & d_{33} \end{pmatrix} = \begin{pmatrix} d_{11} & 0 & 0 \\ d_{21} & d_{22} & 0 \\ d_{31} & d_{32} & d_{33} \end{pmatrix} \times \begin{pmatrix} G_{11} & 0 & 0 \\ G_{12}+G_{21} & G_{22} & 0 \\ G_{31}+G_{13} & G_{32}+G_{23} & G_{33} \end{pmatrix}, \tag{5.19}$$

with an obvious generalization for any n,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{d} = \mathbf{d}\cdot\boldsymbol{\pi}, \tag{5.20}$$

where $\boldsymbol{\pi}$ denotes the matrix on the right side of eq. (5.19) or its generalization to arbitrary n.

It follows from eq. (5.3) that

$$M_{ji} = \hat{\boldsymbol{e}}_j(0)\cdot\boldsymbol{e}_i(t), \tag{5.21}$$

or, using eq. (5.13),

$$M_{ji} = \sum_{k=1}^{n} d_{ik}(t)\hat{e}_{jk}(t), \tag{5.22}$$

where $\hat{e}_{jk}(t)$ is the jth component of $\hat{\boldsymbol{e}}_k(t)$.

Defining a matrix $\hat{\mathbf{e}}$ whose columns are the vectors $\hat{\boldsymbol{e}}_r$, we obtain from (5.22)

$$\mathbf{M} = \hat{\mathbf{e}}\cdot\mathbf{d}^{\dagger}, \tag{5.23}$$

where $\mathbf{d}^{\dagger}$ is the transpose of $\mathbf{d}$.

Eqs. (5.12) and (5.18) or (5.20) constitute the sought differential version of the discrete method.

5.3. *Analysis of the differential version*

It follows from eq. (5.4) that the functions $d_{ii}(t)$ are merely the norms of $N_i^{(r)}$ defined in subsection 5.1 (in the differential limit). Thus for long times

$$|d_{ii}(t)| \approx \mathrm{e}^{\mu_i t}. \tag{5.24}$$

Hence

$$\mu_i = \lim_{t\to\infty}\frac{1}{t}\log|d_{ii}(t)|. \tag{5.25}$$

It thus follows from eq. (5.16) that

$$\mu_i = \lim_{t\to\infty}\frac{1}{t}\int_0^t \mathrm{d}t'\,G_{ii}(t'),$$

or

$$\mu_i = \lim_{t\to\infty}\frac{1}{t}\int_0^t \mathrm{d}t'\,\hat{\boldsymbol{e}}_i^{\dagger}(t)\cdot\mathbf{G}(t)\hat{\boldsymbol{e}}_i(t). \tag{5.26}$$

Since $\mathbf{G}(t)$ is known all one has to do to obtain the Lyapunov exponents is to compute the $\hat{e}_i(t)$'s from eq. (5.12) and substitute in eq. (5.26) (numerically the procedure is slightly different). Note that $(1/t)\int_0^t \mathrm{d}t'\,G_{ii}(t')$ is not $\mu_i(t)$, unlike the case of eq. (3.7).

This is so because the matrix $\mathbf{d}$ is not diagonal, it is lower triangular. Let us analyze the terms d_{ij}. Let

$$\boldsymbol{h}_i(t) = \mathbf{R}^{\dagger}\boldsymbol{e}_i(t), \tag{5.27}$$

where $\mathbf{R}^{\dagger}$ is defined in eq. (3.2). Let $\hat{\boldsymbol{h}}_i$ be the set of vectors obtained from $\boldsymbol{h}_i$ by Gram–Schmidt orthonormalization. Obviously,

$$\hat{\boldsymbol{h}}_i(t) = \mathbf{R}^{\dagger}(t,0)\hat{e}_i(t). \tag{5.28}$$

It follows from eq. (5.13) that

$$d_{ij} = \boldsymbol{h}_i(t)\cdot\hat{\boldsymbol{h}}_j(t). \tag{5.29}$$

Using eq. (3.2) and eq. (5.27),

$$\boldsymbol{h}_i(t) = \mathbf{S}(t,0)\hat{e}_i(0). \tag{5.30}$$

By expanding $\hat{e}_i(0)$ in eq. (5.30) in the vectors $\{|\mu_j\rangle\}$, it is easy to see that

$$\boldsymbol{h}_i(t) = \sum_{j=1}^{N} r_{ij}\,\mathrm{e}^{\mu_j t}|\mu_j\rangle, \tag{5.31}$$

where r_{ij} are time-dependent coefficients.

It is easy to check that

$$\hat{h}_1(t) = \alpha_{11}|\mu_1\rangle + \alpha_{12}\,\mathrm{e}^{\mu_{21}t}|\mu_2\rangle + \cdots ,$$
$$\hat{h}_2 = \alpha_{21}\,\mathrm{e}^{\mu_{21}}|\mu_1\rangle + \alpha_{22}|\mu_2\rangle + \cdots , \tag{5.32}$$

where α_{ij} are $\mathcal{O}(1)$ coefficients. In general, it is easy to prove (by induction) that

$$\hat{h}_i(t) = \sum_{j=1}^{n} \alpha_{ij}\,\mathrm{e}^{-|\mu_{ij}|t}|\mu_j\rangle. \tag{5.33}$$

Hence, using (5.29),

$$d_{ij} \propto \mathrm{e}^{\mu_j t}. \tag{5.34}$$

Since the vectors $|\mu_i\rangle$ converge at exponential rates to fixed vectors, it is easy to verify that the coefficients α_{ij} in eq. (5.32) converge to fixed quantities (which depend on the initial conditions) and thus, to leading order,

$$d_{ij} = A_{ij}\,\mathrm{e}^{\mu_j t}, \tag{5.35}$$

where A_{ij}'s are constants. It is convenient to define the following quantities:

$$\bar{d}_{ij} = \begin{cases} \dfrac{d_{ij}}{d_{jj}}, & i \geq j, \\ 0, & i < j. \end{cases} \tag{5.36}$$

By eq. (5.35), the quantities $\bar{d}_{ij}$ should converge at exponential rates to constants. Indeed, using eq. (5.18) we find

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{d}_{ij} = \sum_{k=j+1}^{i} \frac{d_{kk}}{d_{jj}}\bar{d}_{ik}\left(G_{jk} + G_{kj}\right), \quad \text{for } i > j. \tag{5.37}$$

Since $d_{kk}/d_{jj} \approx \mathrm{e}^{(\mu_k - \mu_j)t}$ and $k > j$, we have established the exponential rate of convergence of the $\bar{d}_{ij}$.

5.4. *Lyapunov and stability eigenvalues and eigenvectors*

Define a diagonal matrix $\mathbf{D}$ by

$$D_{ii} = d_{ii}. \tag{5.38}$$

Hence

$$\mathbf{d}^{\dagger} = \mathbf{D} \cdot \bar{\mathbf{d}}^{\dagger}, \tag{5.39}$$

where $\bar{\mathbf{d}}$ is a matrix whose entries are $\bar{d}_{ij}$. As proven before $\bar{\mathbf{d}}$ tends to a constant matrix at asymptotic times. The eigenvalue equation

$$\mathbf{M}(t,0)|e_i(t)\rangle = \mathrm{e}^{\lambda_i(t)t}|e_i(t)\rangle \tag{5.40}$$

can now be rewritten as (see eq. (5.23))

$$\mathbf{D}\bar{\mathbf{d}}^{\dagger}|e_i\rangle = \mathrm{e}^{\lambda_i(t)t}\hat{e}^{-1}|e_i\rangle. \tag{5.41}$$

Define the vectors $\bar{d}_i$ as

$$\bar{d}_1 = \left(\bar{d}_{11}, \bar{d}_{21}, \bar{d}_{31}, \ldots\right),$$
$$\bar{d}_2 = \left(0, \bar{d}_{22}, \bar{d}_{32}, \ldots\right), \tag{5.42}$$

etc., i.e. these are the rows of $\bar{\mathbf{d}}$. It follows from eq. (5.41) that

$$d_{jj}\left(\bar{d}_j \cdot |e_i(t)\rangle\right) = \mathrm{e}^{\lambda_i(t)}\left(\hat{e}^{-1}|e_i\rangle\right)_j, \tag{5.43}$$

where subscript j means the jth component. For $j < i$ divide both sides of eq. (3.40) by d_{jj}. Hence, for asymptotic t (and assuming $\lambda_i = \mu_i$)

$$\bar{d}_j \cdot |e_i\rangle = 0, \quad \text{for } j < i. \tag{5.44}$$

Thus $|e_n\rangle$ is orthogonal to $\bar{d}_1, \bar{d}_2, \ldots, \bar{d}_{n-1}$, $|e_{n-1}\rangle$ is orthogonal to $\bar{d}_1, \bar{d}_2, \ldots, \bar{d}_{n-2}$, and so on. Let $D_1, D_2, \ldots, D_n$ be an orthonormal set of vectors obtained from the $\bar{d}_i$'s by the Gram–Schmidt procedure, starting with $\bar{d}_1$.

Consequently, $|e_n\rangle$ is orthogonal to $D_1, D_2, \ldots, D_{n-1}$, $|e_{n-1}\rangle$ is a linear combination of D_n and D_{n-1}, and so on. The orthonormalization of $|e_1\rangle, \ldots, |e_n\rangle$, starting with $|e_n\rangle$ and proceeding by decreasing n will give the vectors $\boldsymbol{D}_1, \boldsymbol{D}_2, \ldots, \boldsymbol{D}_n$. This very orthonormalization was shown in section 2 to yield the invariant vectors $|f_i\rangle$. We conclude that

$$|f_i\rangle = \boldsymbol{D}_i \tag{5.45}$$

for asymptotic times. Since we have proven that the quantities $\bar{d}_{ij}$ converge to time-independent constants, we indeed have a consistency check on the result of section 2, as well as a closed formula, which enables one to compute the vectors $|f_i\rangle$ using the differential version of the standard method. Eq. (5.45) also enables us to compute the eigenvalues and eigenvectors of $\mathbf{M}$ for finite large times. Denote

$$d_{ii} = \eta_i \, \mathrm{e}^{\lambda_i(t)t}. \tag{5.46}$$

Then eq. (5.43) can be rewritten as

$$\eta_i \frac{d_{jj}}{d_{ii}} (\bar{\boldsymbol{d}}_j | e_i\rangle) = (\hat{e}^{-1}|e_i\rangle)_j \quad \text{for } i<j, \tag{5.47a}$$

$$\eta_i (\bar{\boldsymbol{d}}_j | e_i\rangle) = (\hat{e}^{-1}|e_i\rangle)_i, \quad \text{for } i=j, \tag{5.47b}$$

$$\eta_i (\bar{\boldsymbol{d}}_j | e_i\rangle) = \frac{d_{ii}}{d_{jj}} (\hat{e}^{-1}|e_i\rangle)_j, \quad \text{for } i>j. \tag{5.47c}$$

Noting that $d_{ii}/d_{jj} \approx \mathrm{e}^{(\lambda_j(t)-\lambda_i(t))t}$, we see that all coefficients in eqs. (5.47) are bounded and that it is a straightforward matter (solving a determinant) to obtain η_i and then $|e_i\rangle$. In the limit of large time (basically when $\mathrm{e}^{(\lambda_i-\lambda_j)t}$ is large enough for all $i<j$) we can present analytical expressions for η_i and $|e_i\rangle$. In this limit eqs. (5.47) reduce to

$$(\hat{e}^{-1}|e_i\rangle)_j = 0, \quad \text{for } i<j, \tag{5.48a}$$

$$\eta_i (\bar{\boldsymbol{d}}_i | e_i\rangle) = (\hat{e}^{-1}|e_i\rangle)_i, \quad \text{for } i=j, \tag{5.48b}$$

$$\bar{\boldsymbol{d}}_j | e_i\rangle = 0, \quad \text{for } i>j. \tag{5.48c}$$

Eq. (5.48c) can be rewritten as

$$\boldsymbol{D}_j | e_i\rangle = 0, \quad \text{for } i>j. \tag{5.48d}$$

Hence, it follows from (5.48d) that

$$|e_n\rangle = \boldsymbol{D}_n \tag{5.49}$$

and from (5.48b)

$$\eta_n = \frac{\left(\hat{e}^{-1} \cdot \boldsymbol{D}_n\right)_n}{\left(\bar{\boldsymbol{d}}_n \cdot \boldsymbol{D}_n\right)}. \tag{5.50}$$

For $i<n$ it follows from (5.48d) that $|e_i\rangle$ can be expressed as

$$|e_i\rangle = \sum_{k=i}^{n} s_{ik} \boldsymbol{D}_k. \tag{5.51}$$

We can fix $s_{i,n} = 1$ without loss of generality, leading to

$$|e_i\rangle = \sum_{k=i}^{n-1} s_{ik} \boldsymbol{D}_k + \boldsymbol{D}_n. \tag{5.52}$$

Noting that since the matrix $\hat{\mathbf{e}}$ is orthogonal, $\hat{\mathbf{e}}^{-1} = \mathbf{e}^{\mathrm{T}}$. Thus, it follows from (5.48a) that

$$\hat{e}_j | e_i\rangle = 0, \quad \text{for } i<j. \tag{5.53}$$

As we recall, $\hat{e}_j$ is the jth column of $\hat{\mathbf{e}}$. Multiplying eq. (5.52) by $\hat{e}_j$ and using (5.53) we obtain

$$\sum_{k=i}^{n-1} s_{ik}\left(\hat{e}_j \boldsymbol{D}_k\right) + \left(e_j \boldsymbol{D}_n\right) = 0, \quad \text{for } j>i. \tag{5.54}$$

The quantities $(\hat{e}_j \boldsymbol{D}_k)$ are known. Denote

$$\left(\hat{e}_j \boldsymbol{D}_k\right) = H_{j-1,k} \tag{5.55}$$

and

$$\left(e_j \boldsymbol{D}_n\right) = -\mathbf{r}_j.$$

The matrices $\mathbf{s}$ and $\mathbf{H}$ are defined for indices $i \le \alpha \le n-1$. Rewriting (5.54) in matrix form,

$$\mathbf{Hs} = \mathbf{r}, \tag{5.56}$$

we find

$$\mathbf{s} = \mathbf{H}^{-1}\mathbf{r} \tag{5.57}$$

(provided $\mathbf{H}^{-1}$ exists), which, together with (5.52) solves for the eigenvectors. Substituting $|e_i\rangle$ in (5.48b) we can now obtain η_i. Notice that, by construction, η_i is a bounded quantity. Thus for large enough times

$$\lambda_i(t) = \frac{\log d_{ii}}{t} + \frac{\log \eta_i}{t}, \tag{5.58}$$

hence the asymptotic difference between the numbers obtained by the discrete methods $(\log d_{ii})/t$, and the real eigenvalues is $\mathcal{O}(1/t)$. Recalling that $(\lambda_i(t) - \lambda_i(\infty)) \approx \mathcal{O}(1/t)$ for large times, we see that the difference between the numbers obtained in the framework of the discrete method and the true λ_i's is not the reason for the (relatively) slow convergence of the discrete algorithms. It is merely due to the slow convergence of the λ_i's themselves.

Next we turn to a calculation for the Lyapunov exponents. We recall that the diagonal elements of the $\mathbf{d}(t)$ matrix $d_{ii}(t)$ serve to define the Lyapunov exponents $(d_{ii} \approx \mathrm{e}^{\mu_i t})$ in the standard method. Therefore it is convenient to define the eigenvalue of $\mathbf{M}^\dagger\mathbf{M}$ as $d_{ii}^2 q_i$, where q_i is a correction term. It follows from eq. (5.23) that

$$\mathbf{M}^\dagger\mathbf{M} = \mathbf{d}\mathbf{d}^\dagger \tag{5.59}$$

and from eq. (5.39) that

$$\mathbf{M}^\dagger\mathbf{M} = \bar{\mathbf{d}}\mathbf{D}^2\bar{\mathbf{d}}^\dagger. \tag{5.60}$$

Thus the eigenvector $|\mu_i\rangle$ satisfies

$$\bar{\mathbf{d}}\mathbf{D}^2\bar{\mathbf{d}}^\dagger|\mu_i\rangle = q_i d_{ii}^2|\mu_i\rangle. \tag{5.61}$$

Hence

$$\mathbf{D}^2\bar{\mathbf{d}}^\dagger|\mu_i\rangle = q_i d_{ii}^2\bar{\mathbf{d}}^{-1}|\mu_i\rangle. \tag{5.62}$$

The kth element in eq. (5.62) satisfies

$$d_{kk}^2\left(\bar{\mathbf{d}}^\dagger|\mu_i\rangle\right)_k = q_i d_{ii}^2\left(\bar{\mathbf{d}}^{-1}|\mu_i\rangle\right)_k. \tag{5.63}$$

Thus, up to exponentially small errors,

$$\left(\bar{\mathbf{d}}^{-1}|\mu_i\rangle\right)_k = 0, \quad \text{for } i < k, \tag{5.64a}$$

$$\left(\bar{\mathbf{d}}^\dagger|\mu_i\rangle\right)_k = 0, \quad \text{for } i > k, \tag{5.64b}$$

$$\left(\bar{\mathbf{d}}^\dagger|\mu_i\rangle\right)_i = q_i\left(\bar{\mathbf{d}}^{-1}|\mu_i\rangle\right)_i. \tag{5.64c}$$

Eq. (5.64c) determines q_i. It follows from (5.64a) that

$$|\mu_i\rangle = \bar{\mathbf{d}}\cdot\begin{pmatrix} a_1^{(i)} \\ a_2^{(i)} \\ \vdots \\ a_i^{(i)} \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{5.65}$$

where the a's are undetermined numbers. Using now (5.64b) we obtain from (5.65)

$$(\bar{\mathbf{d}}^\dagger\bar{\mathbf{d}}\boldsymbol{a}^{(i)})_k = 0, \quad \text{for } k < i. \tag{5.66}$$

The value of $(\bar{\mathbf{d}}^\dagger\bar{\mathbf{d}}\boldsymbol{a}^{(i)})_i$ can be chosen at will (but positive) since it only determines the norm of $\boldsymbol{a}^{(i)}$. We choose it to be 1. Define now a square matrix of dimension i, $\boldsymbol{\Delta}$, by

$$\Delta_{kj} = (\bar{\mathbf{d}}^\dagger\bar{\mathbf{d}})_{k,j}, \quad 1 \le k, j \le i, \tag{5.67}$$

and similarly define an i-dimensional vector $\tilde{\boldsymbol{a}}_i$ which consists of the first i elements of $\boldsymbol{a}^{(i)}$. We obtain

$$\left(\boldsymbol{\Delta}\cdot\tilde{\boldsymbol{a}}_{(i)}\right)_k = \delta_{ki}, \tag{5.68}$$

hence

$$a_m^{(i)} = \Delta_{mi}^{-1}. \tag{5.69}$$

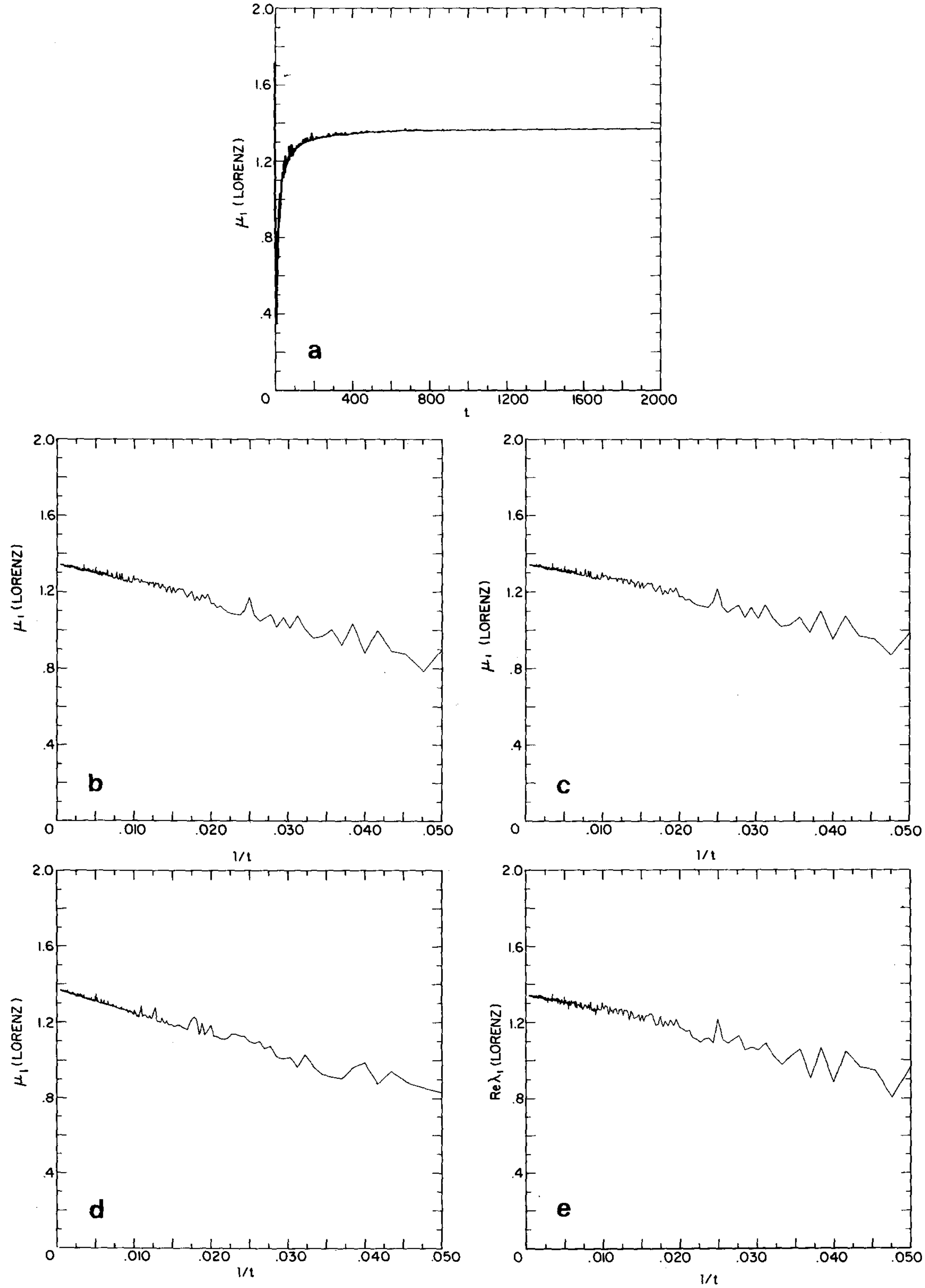

Fig. 1. Lyapunov exponents and stability exponents for the Lorenz model ($r = 40$, $\sigma = 16$, $b = 4$). a) $\mu_1(t)$ vs. t (accuracy per time unit 10^{-4}, standard method); b) $\mu_1(t)$ vs. $1/t$ (accuracy per time unit 10^{-4}, differential standard method); c) $\mu_1(t)$ vs. $1/t$ (corrected differential standard method); d) $\mu_1(t)$ vs. $1/t$ (discrete standard method); e) $\lambda_1(t)$ vs. $1/t$; f) $\mu_3(t)$ vs. t; g) $\mu_3(t)$ vs. $1/t$. See text for details.

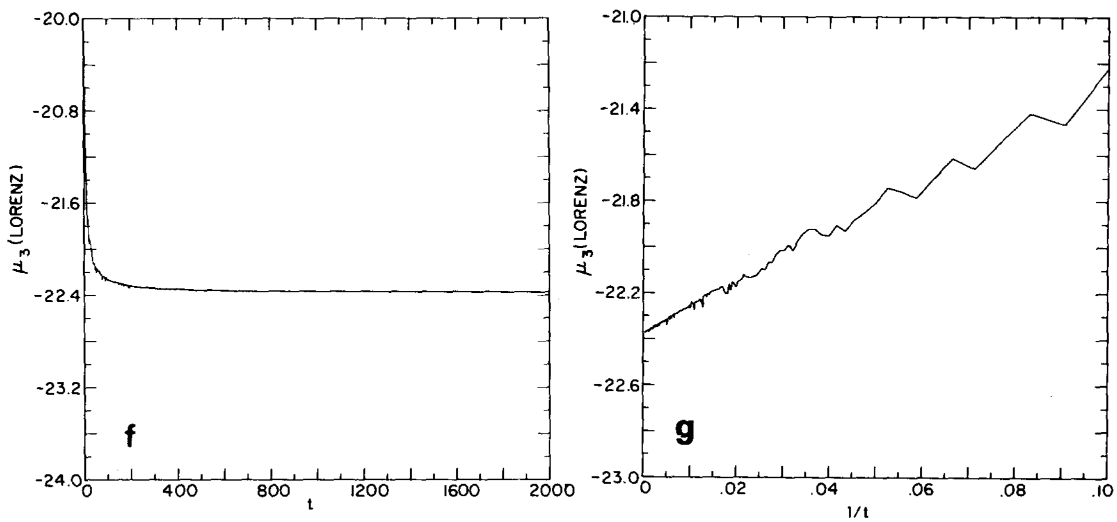

Fig. 1 (Continued).

In particular, it follows from eqs. (5.64c) and (5.65) that $q_i = 1/a_i^{(i)}$. Hence

$$q_i = \frac{1}{\Delta_{ii}^{-1}}. \tag{5.70}$$

Thus

$$\mu_i(t) = \frac{\ln d_{ii}(t)}{t} - \frac{\ln \Delta_{ii}^{-1}}{2t}. \tag{5.71}$$

Note that the matrix Δ is composed of "constant" elements, up to exponentially small corrections. Hence the difference between $\mu_i(t)$ and the value of the exponent deduced from the standard method is c_i/t where c_i is a constant.

6. Numerical results

In this section we present numerical results for Lyapunov and stability exponents, we show how the method of extrapolation works and we state other numerical results, whose details are not presented. The models that have been investigated are: Lorenz [20] (fig. 1 and table 1), Rabinovich–Fabrikant [21] (fig. 2 and table II), the 14-mode Currey [22] model (fig. 3 and table III), Mackay–Glass [23] (fig. 4 and table IV) and Knobloch–Weiss [24] (fig. 5 and table V).

We have found that in *all* of the above cases the following holds:

1) The Lyapunov exponents are nondegenerate.
2) The stability exponents equal the Lyapunov exponents.
3) The Lyapunov eigenvectors converge at exponential rates (as predicted in section 2).
4) The stability eigenvectors form subspaces converging at exponential rates (see subsection 2.2).
5) Except for a short "transient" time all Lyapunov exponents have the form: $\mu_i(t) = \mu_i + (b_i + \xi_i(t))/t$, where ξ_i is an oscillating function of zero average and relatively small amplitude.
6) The smaller the Lyapunov exponent, the faster it and the corresponding eigenvector (or stability subspace) converge.
7) The Kaplan–Yorke dimension converges much faster than the individual Lyapunov

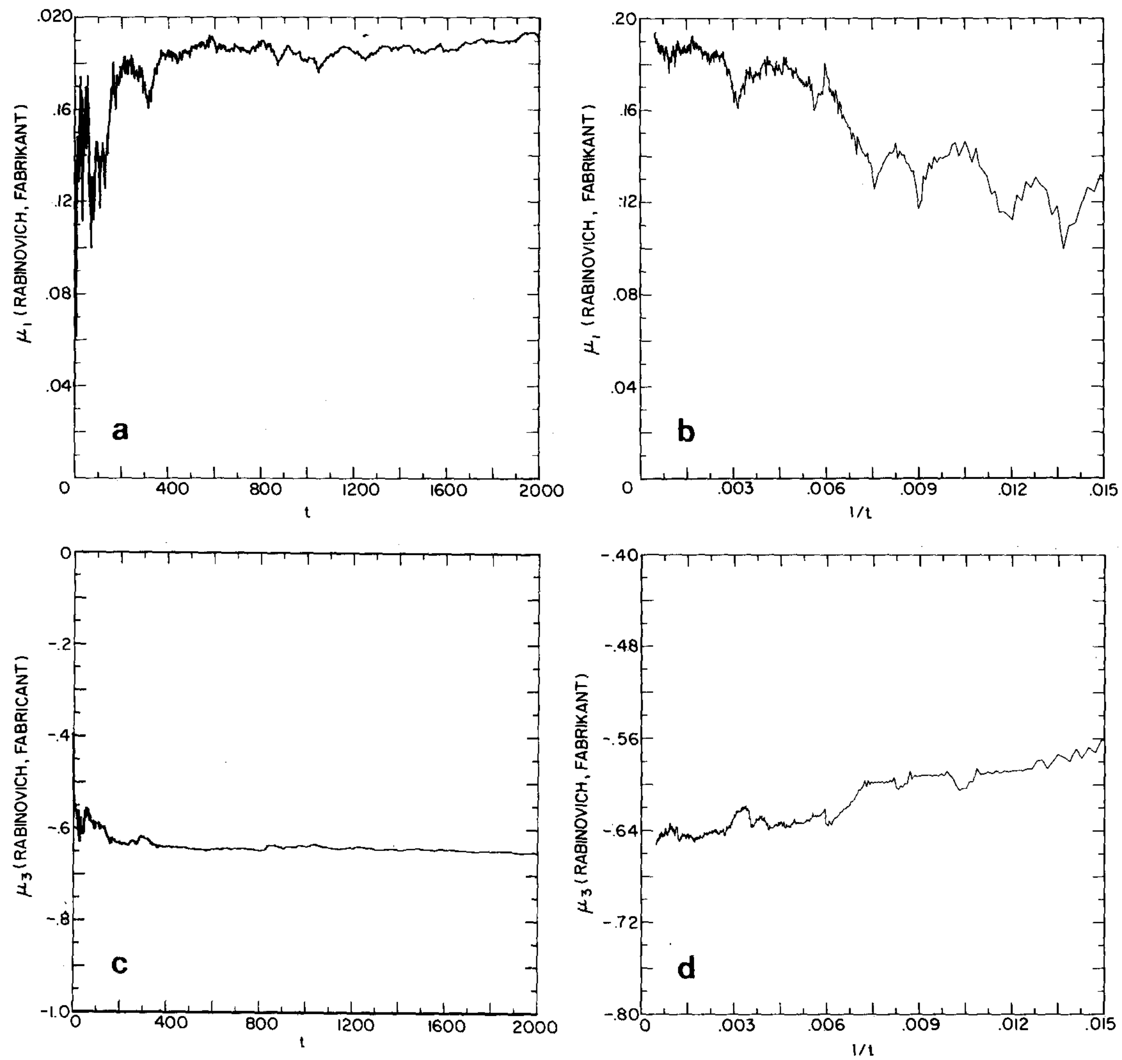

Fig. 2. Lyapunov exponents for the Rabinovich–Fabrikant equations ($\gamma = 0.87$, $\alpha = 1.1$). a) μ_1 vs. t; b) μ_1 vs. $1/t$; c) μ_3 vs. t; d) μ_3 vs. $1/t$.

Table I
Extrapolation results for μ_1: the Lorenz model

t_1	t_2	μ_1	b	σ
20	40	1.36	−10.84	0.03
25	100	1.37	−11.29	0.03
50	100	1.34	−9.18	0.02
100	200	1.390	−14.14	0.007
100	500	1.374	−12.10	0.005
100	1000	1.375	−12.19	0.004
100	1500	1.374	−12.03	0.003
100	2000	1.374	−12.09	0.003

Table II
Extrapolation results for μ_1: the Rabinowitz–Fabrikant model

t_1	t_2	μ_1	b	σ
100	300	0.21	−3.40	0.007
150	300	0.19	−3.43	0.005
150	500	0.19	−3.54	0.005
150	2000	0.19	−3.58	0.003

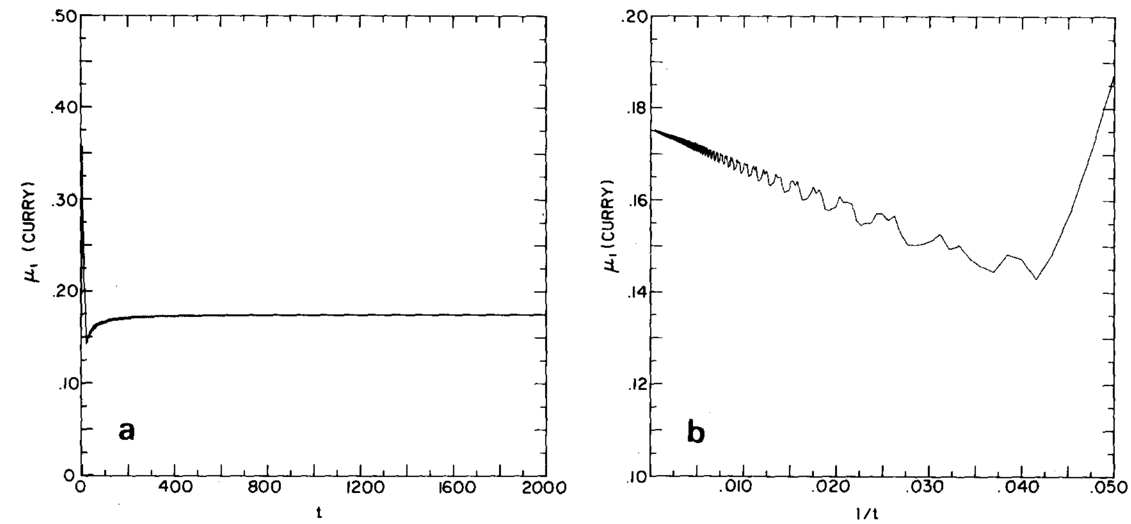

Fig. 3. Lyapunov exponents for the 14-mode Curry equations. a) μ_1 vs. t; b) μ_1 vs. $1/t$.

exponents. By the time $\mu_i(t)$ attains the asymptotic form (as in statement (5)) the Kaplan–Yorke dimension has basically converged (typically up to 1%).

8) The numerical value of the computed Lyapunov exponents depends slightly on the accuracy of the integration scheme (we have used fourth order Runge–Kutta with several time step sizes).
9) Using the form of $\mu_i(t)$ given in statement (5) one can find the values of Lyapunov exponent from short time series. Smaller Lyapunov exponents tend to this form faster.
10) The "noise" $\xi_i(t)$ is neither white nor gaussian.

Next we present a detailed account of some of our numerical results.

Fig. 1 displays results for the Lorenz models. The parameters are $b = 4$, $\sigma = 16$ and $r = 40$. Fig. 1a shows $\mu_1(t)$ as a function of time. The results were obtained using the differential version of the standard method and relative accuracy per time unit of 10^{-4}. Fig. 1b shows $\mu_1(t)$ as a function of $1/t$. The average linear behaviour is prominent. Notice that the asymptotic value of $\mu_1(t)$ in this case is slightly different from that in fig. 1a since here only a relative accuracy of 10^{-2} per time unit was used. The "average" straight line "starts" at about a time of 20, when $\mu_1(t) = 0.9$ is still far from its asymptotic value. Indeed, table I shows extrapolated values of $\mu_1(t)$, using the following formula:

$$L = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \left(\mu_1(t') - \mu_1 - \frac{b_1}{t'} \right)^2 \mathrm{d}t'. \tag{6.1}$$

The minimization of L yields values for μ_1, b_1 and the mean square deviation of the fit, σ. Notice that a better than 1% accuracy is obtained when $t_1 = 20$ and $t_2 = 40$. Fig. 1c presents $\mu_1(t)$ as computed using the results of subsection 5.4, namely, a correction to the standard method yielding $\mu_i(t)$ up to exponentially small errors has been used. Identical results are obtained by a direct use of the differential equations for Lyapunov exponents and eigenvectors. As predicted the slope of the graph

Table III
Extrapolation results for μ_1: the Curry model

t_1	t_2	μ_1	b	σ
25	100	0.163	−0.309	0.030
25	200	0.166	−0.293	0.025
25	250	0.167	−0.286	0.023
25	1000	0.172	−0.248	0.01

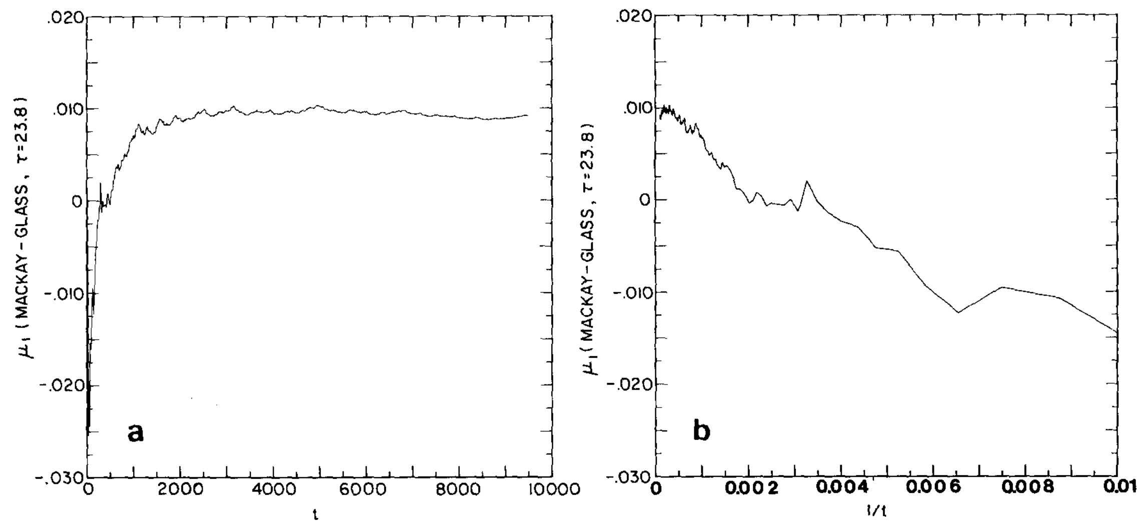

Fig. 4. Lyapunov exponents for the Mackay–Glass equations ($\tau = 23.8$). a) μ_1 vs. t; b) μ_1 vs. $1/t$.

of $\mu_1(t)$ versus $1/t$ in this case is different than that obtained by a naive use of the standard method. The asymptotic value is the same and, obviously, the same extrapolation method that was described above works here. Fig. 1d shows a similar behaviour obtained from the discrete version of the standard method. In fig. 1e we present the results of a computation of $\operatorname{Re} \lambda_1(t)$ plotted as a function of $1/t$. For some short times, λ_1 may have an imaginary part (i.e. the real part of the spectrum is degenerate). For long enough times $\lambda_1(t)$ is real and very close in value to $\mu_1(t)$. As can be observed from this plot, $\lambda_1(t)$ has the same asymptotic behaviour as $\mu_1(t)$ has. The data for $\lambda_1(t)$ was computed using the method described in section 5 as well as by the differential method described in section 2. In fig. 1f we show $\mu_3(t)$ as a function of time. The results presented here have been obtained by the standard method. Fig. 1f displays the fact that the $\mu_3(t)$ converges to its asymptotic value at an average linear rate in $1/t$.

Fig. 2a shows $\mu_1(t)$ for the Rabinovich–Fabrikant equations (with $\gamma = 0.87$ and $\alpha = 1.1$). A plot of the same quantity as a function of $1/t$ once again shows an average linear behaviour. The "noise" $\xi_1(t)$ in this case is more pronounced than in the Lorenz case. Table II shows the results of the extrapolation method described before, as applied to this case. A 5% accuracy for μ_1 is obtained from data for times ranging from 100 to 300 units (100 is still in the transient regime). A better than 0.5% accuracy is already obtained using $t_1 = 150$ and $t_2 = 300$. In this range the actual value of $\mu_1(t)$ is by about 35% less than the asymptotic value. Fig. 2c shows $\mu_3(t)$ for the same model. Note the faster approach of $\mu_3(t)$ to its asymptotic value compared to that of $\mu_1(t)$. Fig. 2d shows $\mu_3(t)$ as a function of $1/t$ and the same average linear behaviour is observed. It is also interesting to note that $\mu_1(t)$ on the average increases towards its asymptotic value whereas $\mu_3(t)$ decreases towards its.

Fig. 3 presents similar results for the 14-mode Curry model. Table III shows how well the ex-

Table IV
Extrapolation results for μ_1: the Mackay–Glass equation

t_1	t_2	μ_1	b	σ
95	1900	0.0090	−2.77	0.0002
190	1900	0.0098	−3.42	0.0001
1900	8000	0.0096	0.625	0.0003

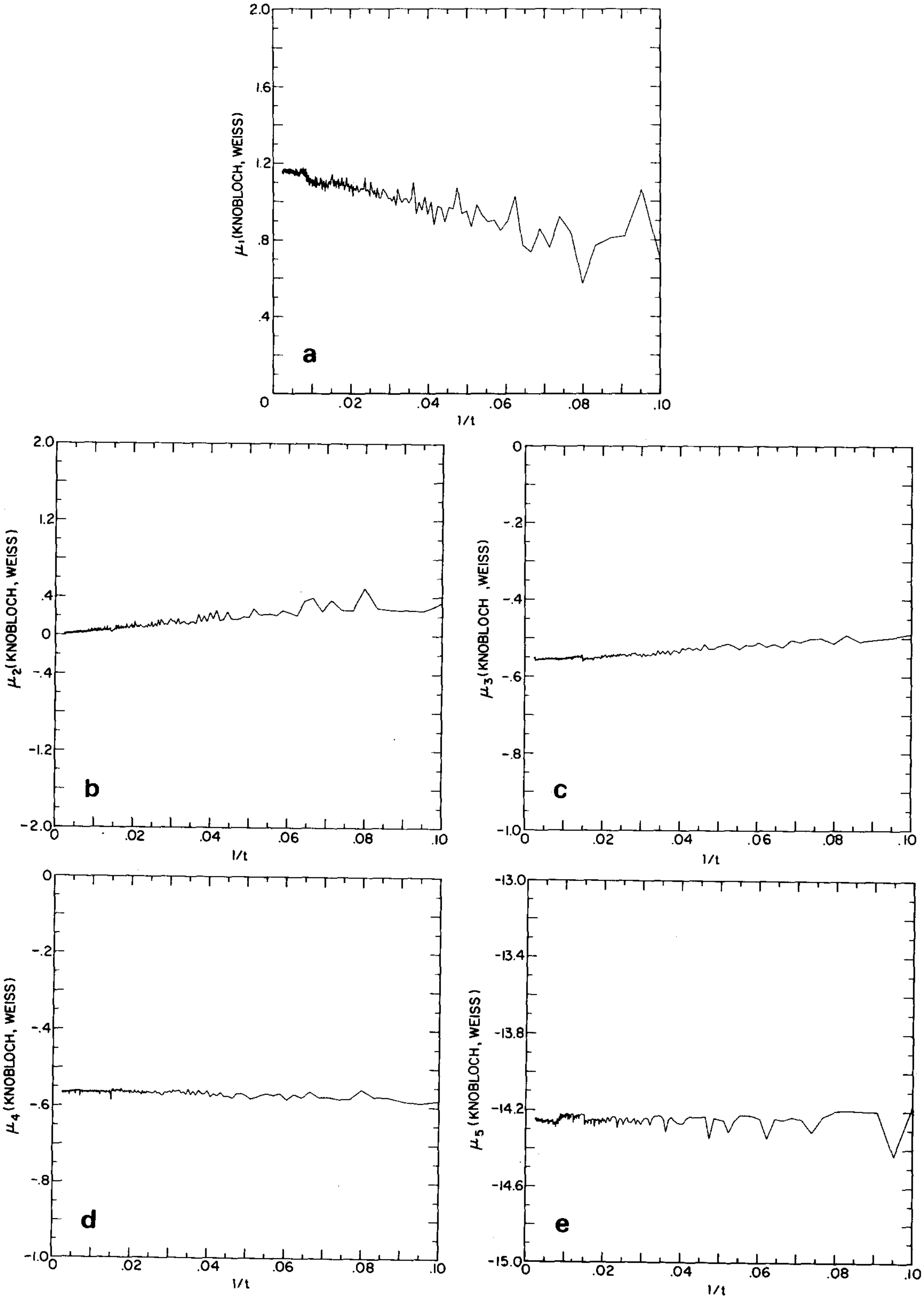

Fig. 5. Lyapunov exponents for the Knobloch–Weiss equations ($q = 5$, $\Lambda = 2$, $\xi = 0.4$, $r = 50$ and $\sigma = 10$). a) μ_1 vs. $1/t$; b) μ_2 vs. $1/t$; c) μ_3 vs. $1/t$; d) μ_4 vs. $1/t$; e) μ_5 vs. $1/t$.

Table V
Extrapolation results for μ_1: the Knobloch–Weiss model

t_1	t_2	μ_1	b
10	50	1.15	−4.18
20	100	1.15	−4.02
20	400	1.17	−4.85
100	200	1.19	−7.20
100	400	1.17	−4.23
20	200	1.17	−4.85

trapolation method works. Notice the dependence of the fitted slope on t_2. It is due to the arbitrary choice of an initial point (t_1). The average slope is to be found using several values of t_1.

As a further example, in the case of a system having an infinite number of degrees of freedom, we present results for the Mackay–Glass equations ($t = 23.8$) in fig. 4. Here μ_1 is relatively small and the noise $\xi(t)$ has some very low frequency components. Nevertheless, a relatively short time simulation is sufficient for the extrapolation to yield a good result (see table IV). Notice the "local" decrease in $\mu_1(t)$ for $2000 < t < 8000$. A fit in this region still yields a good value for μ_1.

So far we have demonstrated our method on the leading Lyapunov exponent and, in some cases, another Lyapunov exponent. Fig. 5 shows the behaviour of $\{\mu_i(t), 1 \le i \le 5\}$ for the Knobloch–Weiss model, as a function of $1/t$. Note the average linear behaviour of all of these plots, the fact that $\mu_1(t)$ increases towards its asymptotic value and the other $\mu_i(t)$ decrease towards their respective asymptotic values. Finally, note that the smaller or less dominant the Lyapunov exponent, the faster it converges to its asymptotic value. The parameters used here are: $q = 5$, $\Lambda = 2$, $\xi = 0.4$, $r = 50$ and $\sigma = 10$. Table V shows the results of the extrapolation method for μ_1. Similar results have been obtained for all other μ_i.

7. Summary and conclusions

We have constructed differential equations for the eigenvalues and eigenvectors of the stability matrix $\mathbf{M}$ and the corresponding matrix $\mathbf{M}^\dagger\mathbf{M}$. We have also constructed a differential version of the standard methods. The resulting equations allow one to obtain some insights into the nature of the stability problem in unsteady or even chaotic situations. Some of the results, such as the exponential rate of convergence of the eigenvectors of $\mathbf{M}^\dagger\mathbf{M}$, are rigorous and are derived in a simple and straightforward way. Others, such as the relations between stability and Lyapunov exponents, are essentially conjectures based upon plausible arguments and numerical results. One of the important ingredients of our formalism is, in our view, its simplicity; compare for example to the various proofs of the Oseledec theorem [1].

A question that had occupied us is: why do the eigenvectors of $\mathbf{M}^\dagger\mathbf{M}$ converge at exponential rates (and are, in a sense, local properties of the attractor) whereas the corresponding eigenvalues converge at a very slow rate? We believe that our formula for the asymptotic form of $\mu_i(t)$ answers, at least partially, this question. The information on the value of μ_i is already contained in $\mu_i(t)$ by the time the eigenvectors have converged. However, it is masked by a term b_i/t and a noisy term. This may be taken to mean that the exponents themselves are "local properties" inasmuch as the eigenvectors are. An argument showing that the Lyapunov exponents are independent of position on a given trajectory, showing that they are truly global quantities has also been given (see section 4).

We have shown how the asymptotic form of the $\mu_i(t)$'s can be used in order to infer the Lyapunov exponents μ_i from limited amounts of data. This observation should be useful for the analysis of experimental data or numerical data for systems having many degrees of freedom. A disappointing fact concerning the Lyapunov exponents is that one cannot improve, in principle, on the standard method (except in computational stability and efficiency) since the error introduced by this method is $\mathcal{O}(1/t)$, which is similar to the convergence rate of the Lyapunov exponents themselves.

A point which has not been taken up in this work is the nature of the noise $\xi_i(t)$ and its relation to the actual dynamics of a system [25]. This will be done in a future publication. Another set of quantities which should be investigated is the Lyapunov (and stability) eigenvectors (for an interesting application, see ref. 10). We believe they contain much information on the geometry and dynamics on the attractor (and they have the advantage of being smooth fields), in spite of the fact that they are representation dependent (in physical problems one is normally interested in a representation in terms of physical variables).

The theory of dynamical systems has shown that many intuitive arguments can be wrong and that "physical reasoning" can be often erroneous in systems having highly nontrivial topological and dynamical properties. One should thus try to develop a new physical intuition and phenomenology which is suitable for the kind of phenomena displayed by dynamical systems, in order to tackle physical problems which are not amenable to exact analysis. We hope that the present work is a small step in this direction.

Acknowledgements

This work was partially supported by the Israel-U.S. Binational Science Foundation under grant number 84-00073, the Air Force Office of Scientific Research under Contract 83-0089, and the Office of Naval Research under Contracts N00014-82-C-0451 and N00014-85-K-0201. One of us (I.G.) gratefully acknowledges support from the Alice and Beno Gitter Foundation.

References

[1] V.I. Oseledec, A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems, Trans. Moscow Math. Soc. 19 (1968) 197. The original work of Lyapunov can be found in: Ann. Math. Study 17 (1977) (Princeton).
For simplified proofs see M.S. Raghunathatan, Israel J. Math. 32 (1979) 356; and H. Kesten, in: Random Matrices and Their Applications, J.E. Cohen, H. Kesten and C.M. Newman, eds., in AMS Series in Contemporary Mathematics.

[2] For a down-to-earth review see: J.D. Farmer, E. Ott and J.A. Yorke, The dimension of chaotic attractors, Physica 7D (1983) 153.

[3] D.A. Russell, J.E. Hanson and E. Ott, Phys. Rev. Lett. 45 (1980) 1175.

[4] J. Froyland, Lyapunov exponents for multidimensional orbits, Phys. Lett. A 97 (1983) 8.

[5] J.D. Farmer, Chaotic attractors of an infinite-dimensional dynamical system, Physica 4D (1982) 366.

[6a] J. Kaplan and J. Yorke, Functional Differential Equations and Approximation of Fixed Points, H.O. Peitgen and H.O. Walther, eds. (Springer, Berlin, 1979), p. 228.

[6b] P. Frederickson, J.L. Kaplan, E.D. Yorke and J.A. Yorke, J. Diff. Eq. 49 (1983) 185.

[6c] For a different suggestion, see: H. Mori, Prog. Theor. Phys. 63 (1980) 1044.

[7] P. Grassberger and I. Procaccia, Characterization of strange attractors, Phys. Rev. Lett. 50 (1983) 396.

[8] B.B. Mandelbrot, Fractals - Form, Chance and Dimension (Freeman, San Francisco, CA, 1977).

[9] P. Grassberger and I. Procaccia, Estimation of the Kolmogorov entropy from a chaotic signal, Phys. Rev. A 28 (1983) 2591; and references therein.

[10] B.J. Bayley, Ph.D. Thesis (1986) and to be published.

[11] G. Bennetin, L. Galgani and J.M. Strelcyn, Phys. Rev. A 14 (1976) 2338.

[12] G. Bennetin, L. Galgani, A. Giorgilli and J.M. Strelcyn, Meccanica 15 (1980) 9, 21.

[13] I. Shimada and T. Nagashima, Prog. Theor. Phys. 61 (1979) 1605.

[14] M. Sano and Y. Sawada, Phys. Rev. Lett. 55 (1985) 1089.

[15] A. Brandstater, J. Swift, H.L. Swinney, A. Wolf, J.D. Farmer, E. Jen and P.J. Crutchfield, Phys. Rev. Lett. 51 (1983) 1442.

[16] A. Wolf and J. Swift, in: Statistical Physics and Chaos in Fusion Plasmas, W. Horton and L. Reichl, eds. (Wiley, New York, 1983).

[17] A. Wolf, J. Swift, H.L. Swinney and J.A. Vastano, Determining Lyapunov exponents from a time series, Physica 16D (1985) 285.

[18] We are grateful to B.J. Bayly for this suggestion.

[19] Since this paper was submitted, we have learned of a similar (unpublished) method by Bernard Legras. We are grateful to him for sending his results to us.

[20] E.N. Lorenz, Deterministic nonperiodic flow, J. Atmos. Sci. 20 (1983) 130.

[21] M.I. Rabinovich and A.L. Fabrikant, Stochastic self-modulation of waves in nonequilibrium media, Soviet Physics JETP 50 (1979) 311.

[22] J.H. Curry, Comm. Math. Phys. 60 (1978) 193.

[23] M.C. Mackey and L. Glass, Oscillation and chaos in physiological control systems, Science 197 (1977) 287.

[24] E. Knobloch and N.O. Weiss, Bifurcations in a model of magnetoconvection, Physica 9D (1983) 379.
See also E. Knobloch, N.O. Weiss and L.N. Da Costa, Oscillatory and steady convection in a magnetic field, JFM 113 (1981) 153.

[25] See however, I. Procaccia, Physica Scripta T9 (1985) 40.