Carleman embedding and Lyapunov exponents

R. F. S. Andrade

Additional information on J. Math. Phys.

Carleman embedding and Lyapunov exponents

R. F. S. Andrade

Instituto de Física da Universidade Federal da Bahia, 40000-Salvador, Bahia, Brazil

We investigate the solutions of those autonomous systems with quadratic nonlinearities in a N-dimensional vector space together with the solutions of their first variational equation systems by means of the Carleman embedding. An iterative procedure based on this result is developed to evaluate the Lyapunov exponents of the considered systems. We test the method by giving some results for the Lyapunov exponents of the Lorenz model.

PACS numbers: 02.30.Hq

I. INTRODUCTION

The concepts of Carleman embedding[1] and Lyapunov exponents are related to the study of nonlinear differential equation systems and have been intensively investigated recently.[2–6] Loosely speaking the former is a method for solving the proposed system after embedding it into an infinite set of linear equations, whereas the latter reports about the asymptotic behavior of two trajectories (with respect to each other) which are close together at the initial time.

In this paper we bring these concepts together in that we use the method of Carleman to evaluate Lyapunov exponents. We consider here those autonomous systems in a N-dimensional vector space with the form

$$\frac{dy_i}{dt} = F_i(\{y_i\})$$

$$= \sum_j a_{ij} y_j + \sum_{jl} b_{ijl} y_j y_l, \quad i,j,l = 1,2,...,N,$$

$$(1)$$

where the a_{ij} and the b_{ijl} are constant.

System (1) describes many different models in hydrodynamic and chemical reactions, etc., which display nontrivial dynamics, with the presence, e.g., of strange attractors and sequences of period-doubling bifurcations. The results we get here can be easily extended to include systems with higher polynomial nonlinearities.

The central point of our method lies on the fact that the infinite matrix M (to be defined later) resulting from the Carleman embedding of the system (1) gives rise not only to the solution of (1) itself, but also to the solution of the system of first-variational equations of (1),

$$\frac{d\delta y_i}{dt} = \sum_j a_{ij}\,\delta y_j$$

$$+ \sum_{jl} b_{ijl}\,(y_j(t)\delta y_l + y_l(t)\delta y_j), \quad (2)$$

whose knowledge is required for the evaluation of the Lyapunov exponents. The solution of (2) indicates how the difference δy between two trajectories evolves with time, and we have explicitly written $y_j(t)$ on the rhs of (2) to call attention to the fact that it is *not* a simple autonomous linear equation system at all.

Generally one can find approximate solutions for (1) and (2) with the help of numerical techniques. The method developed here gives a clear insight into the relation between the solutions of (1) and (2), and has simplifying advantages for the numerical computing that is required if we want to give the explicit values of the Lyapunov exponents.

This article is organized as follows: In Sec. II we introduce the concept of the Carleman embedding and use it to write down the solutions of (1) and (2). The equivalence between the Taylor series solution of (2) and the one given by the Carleman embedding is explicitly proven. In the Sec. III we define the Lyapunov exponents and discuss their meaning. Finally, a discussion of the numerical procedure, some results for the Lorenz model, and concluding remarks are presented in the Sec. IV.

II. THE CARLEMAN EMBEDDING

A. The solution of the basic system

Let us write the system (1) with the help of vector notation;

$$\frac{dx}{dt} = Ax + Bx^{[2]} \quad (3)$$

where x is the column vector of N components y_i and A is the matrix of components a_{ij}. The nonlinear terms are described by the $N \times N^2$ matrix B [$B_{ik} = b_{ijl}$, where $k = (j,l)$ and is to be labeled according to the lexicographical order], and by the vector $x^{[2]} = x \otimes x$, where $\otimes$ denotes the Kronecker product. It is convenient to introduce the Lth Kronecker power by $x^{[L]} = x^{[L-1]} \otimes x, x^{[1]} = x$, as well as the matrices

$$A_L = A \otimes I^{[L-1]} + I \otimes A_{L-1}, \qquad A_1 = A,$$
$$L = 2,3,\cdots,$$
$$B_L = B \otimes I^{[L-1]} + I \otimes B_{L-1}, \qquad B_1 = B,$$
$$(4)$$

with I denoting the $N \times N$ identity matrix. They will appear throughout the work with the method of Carleman.

The Carleman embedding of the system (3) amounts to first considering the components of $x^{[2]}$ as independent variables, and then to writing down their equations of motion. It can be easily shown that

$$\frac{dx^{[2]}}{dt} = A_2 x^{[2]} + B_2 x^{[3]}. \quad (5)$$

Then we consider successively the $x^{[3]}$, $x^{[4]}, \cdots$, as independent, which leads to the infinite linear system

$$\frac{d}{dt}\begin{pmatrix} x^{[1]} \\ x^{[2]} \\ x^{[3]} \\ \vdots \end{pmatrix} = \begin{pmatrix} A_1 & B_1 & 0 & 0 & \cdots \\ 0 & A_2 & B_2 & 0 & \cdots \\ 0 & 0 & A_3 & B_3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}\begin{pmatrix} x^{[1]} \\ x^{[2]} \\ x^{[3]} \\ \vdots \end{pmatrix}. \qquad (6)$$

We write X and M for the infinite-dimensional vector and matrix appearing in (6), and let their block components be denoted by capital indices as $X_L = x^{[L]}$ and

$$M_{L,K} = A_L \delta_{L,K} + B_L \delta_{L,K+1}.$$

The component X_1 of the solution of (6) should also furnish the solution of the system (3). Nevertheless one must proceed carefully, for the system (6) may have a broader class of solutions than that of (3); e.g., one can find C^∞ solutions for (6) whose component X_1 does not satisfy (3).[7] However, the analytical solutions of (3) and of X_1 in (6) are equal, which can be shown by comparing the terms of their Taylor series.

The formal analytical solution for X in (6) is $X(t) = e^{MT}X(0)$. If we are only interested in the component X_1, we may write

$$X_1(t) = \sum_L (e^{MT})_{1,L}\, X_L(0)$$

$$= \sum_L \left[I + Mt + \frac{M^2 t^2}{2!} + \cdots \right]_{1,L} X_L(0). \qquad (7)$$

We introduce now a generalized power operation between the $N \times N$ matrix A and the $N \times N^2$ matrix B which enables the writing of both the Carleman solution and the general term of the Taylor series solution of (1) and (2) in a very compact way. So, if $n, m \in Z$ we define the matrix $\langle A^n | B^m \rangle$ by

$$\langle A^n | B^m \rangle = 0 \text{ if } n \text{ or } m < 0, \qquad (8a)$$

$$\langle A^0 | B^0 \rangle = I \qquad (8b)$$

$$\langle A^n | B^m \rangle = \langle A^n | B^{m-1} \rangle B_m + \langle A^{n-1} | B^m \rangle A_{m+1}, \qquad (8c)$$

where the A_n and B_m are the matrices defined in (4).

The definitions (8c) implies that the elements of $\langle A^n | B^m \rangle$ given in terms of those of A and B contain always an n-fold product of elements of A and an m-fold product of those of B.

This justifies calling the definition (8) a kind of generalized power. In particular we have $\langle A^n | B^0 \rangle = A^n$ and $\langle A^0 | B^m \rangle = B_1 B_2 B_3 \cdots B_m$.

Now it is straightforward to verify that

$$(M^n)_{1,L} = \langle A^{n+1-L} | B^{L-1} \rangle, \qquad (9)$$

and after inserting (9) into (7) we get

$$X_1(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \sum_{L=1}^{n+1} \langle A^{n+1-L} | B^{L-1} \rangle X_L(0). \qquad (10)$$

The expression (10), which was derived only with the help of the Carleman method, constitutes also the Taylor series solution of (1), which can be directly verified by using the definition (8) to evalutate the nth time derivative of x in (1). In the next subsection we illustrate the use of the matrices defined in (8) to verify a similar result for the case of the first variational system (2).

B. The solution of the variational equation system

Now we will show the solution of (2) may be given with the help of the Carleman embedding matrix as

$$\delta x(t; x_0, \delta x_0) = \sum_{L=1}^{\infty} (e^{Mt})_{1,L} \frac{\partial X_L(0)}{\partial x(0)} \delta x_0. \qquad (11)$$

In order to prove Eq. (11) it is convenient to write (2) in terms of Kronecker operations to get

$$\frac{d}{dt} \delta x = A \delta x + B(x \otimes \delta x + \delta x \otimes x). \qquad (12)$$

Further, we define the following $N^L \times N$ matrices Y_L, $L \geqslant 1$:

$$Y_L = X \otimes Y_{L-1} + I \otimes X^{[L-1]}, \quad L = 2, 3, \cdots,$$

$$Y_1 = I. \qquad (13)$$

The Y_L are very important for our proof, since it can be shown that they satisfy the relation

$$(Y_L)_{ij} = \partial(X_L)_i / \partial y_j. \qquad (14)$$

We can also use the Y_L to bring (12) into a more compact form

$$\frac{d}{dt} \delta x = (A + BY_2) \delta x. \qquad (15)$$

The solution of (15) written with the aid of the Taylor series is

$$\delta x(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \frac{d^n}{dt^n} \delta x \bigg|_{t=0}.$$

Using the bracket operation defined in (8) and Y_L, we will show that

$$\frac{d^n}{dt^n} \delta x \bigg|_{t=0} = \sum_{L=1}^{n+1} \langle A^{n+1-L} | B^{L-1} \rangle Y_L(0) \delta x_0. \qquad (16)$$

We proceed by induction over n. For $n = 1$, (16) is evident, for it reduces to (15). If we assume (16) to be valid for $n-1$, we get

$$\frac{d^n}{dt^n} \delta x = \sum_{L=1}^{n} \langle A^{n-L} | B^{L-1} \rangle \frac{d}{dt}(Y_L \delta x). \qquad (17)$$

In order to go further we need identity (18), which can be easily derived with some algebraic manipulations with the help of Kronecker operations and of Eq. (15),

$$\frac{d}{dt}(Y_L \delta x) = (A_L Y_L + B_L Y_{L+1}) \delta x. \qquad (18)$$

After inserting (18) into (17) we arrive at

$$\frac{d^n}{dt^n} \delta x$$
$$= \sum_{L=1}^{n} \langle A^{n-L} | B^{L-1} \rangle [A_L Y_L + B_L Y_{L+1}] \delta x, \qquad (19)$$

and using the definition of the bracket operation it follows that

$$\frac{d^n}{dt^n} \delta x = \sum_{L=1}^{n} [\langle A^{n+1-L} | B^{L-1} \rangle$$
$$- \langle A^{n+1-L} | B^{L-2} \rangle B_{L-1}] Y_L \delta x$$
$$+ \sum_{L=1}^{n} \langle A^{n-L} | B^{L-1} \rangle B_L Y_{L+1} \delta x. \qquad (20)$$

We now observe that all the terms in the second and third series cancel each other with exception of

$$-\langle A^n|B^{-1}\rangle B_0 Y_1 = 0$$

and

$$\langle A^0|B^{n-1}\rangle B_n Y_{n+1} = \langle A^0|B^n\rangle Y_{n+1}.$$

Then Eq. (16) follows immediately.

The solution of (12) given by the Taylor series is thus

$$\delta x(t;x_0,\delta x_0)$$

$$= \sum_{n=0}^{\infty} \sum_{L=1}^{n+1} \frac{t^n}{n!} \langle A^{n+1-L}|B^{L-1}\rangle Y_L(0)\delta x_0, \qquad (21)$$

which reduces to (11) if we use (9) and (14).

The Carleman method shows very clearly the relation between the solutions of (1) and (2). This fact has not only a formal relevance, but also for practical purposes it enables the development of a numerical procedure furnishing the solutions of both equations which are required for the evaluation of the Lyapunov exponents.

III. LYAPUNOV EXPONENTS

Let us now introduce the notation which is most used when one is concerned with the problem of defining Lyapunov exponents. Let us write the solution of (3) as

$$x(t) = T^t x_0, \quad x_0 = x(t=0), \qquad (22)$$

where the map T^t, which describes the evolution of any point in the phase space, is the flow induced by the vector field $F(x)$. Let $DT^t x_0$ be the matrix of partial derivatives

$$(DT^t x_0)_{ij} = \frac{\partial(T^t x_0)_i}{\partial(x_0)_j}. \qquad (23)$$

The map $DT_t x_0$ may be so interpreted: two trajectories starting at x and $x + \delta x$, where δx is a small vector, will differ by $DT^t x_0 \delta x$ at a time t. Hence DT^t entails the same information as the solution of (2), and if knowledge of the explicit form of T^t is not possible, DT^t is formed by collecting the N column eigenvectors of (2). The rhs of (2) is time dependent through $x(t)$ and by the above quoted method one can only give an approximation for DT^t which is valid only for short time intervals. On the other hand, the Carleman method enables the evaluation of DT^t by direct differentiation of T^t [identified with the time-evolution operator in (7)] as has become clear from the last section.

The concept of Lyapunov exponents of the dynamical system (1) may be introduced in different ways.[2,3,8,9] A very illustrative one is to say that the point x_0 of the phase space has Lyapunov exponents $\lambda_1 < \lambda_2 \cdots < \lambda_N$ if there exist subspaces $E_1 \subset E_2 \subset \cdots \subset E_N$ of our vector space with the following properties:

$$E_j = \left\{ \delta x \mid \lim_{t\to\infty} \frac{1}{t} \ln \left[\frac{\|DT^t x_0 \delta x\|}{\|\delta x\|} \right] \leqslant \lambda_j \right\}, \qquad (24a)$$

$$\dim E_j = j, \qquad (24b)$$

$$\lim_{t\to\infty} \frac{1}{t} \ln \left[\frac{\|DT^t x_0 \delta x\|}{\|\delta x\|} \right] = \lambda_j$$

if $\delta x \in E_j$ but $\delta x \notin E_{j-1}$. \qquad (24c)

Definition (24) becomes easy to visualize if we consider a linear autonomous system with real eigenvalues. In such a case the Lyapunov exponents λ_j coincide with the eigenvalues of the system. The subspace E_j is the set of all vectors which are written as linear combinations of the eigenvectors corresponding to the j smallest eigenvalues of the system.

The largest eigenvalue λ_j of E_j is projected out by (24c). Two trajectories differing by $\delta x \in E_j$ will converge or diverge according to the negative or positive sign of λ_j. Since an arbitrary vector δx will almost always belong to E_N, the λ_N is the exponent which really decides about the diverging character of the trajectories. Moreover, it is easier to evaluate than any other λ_j.

The question concerning the existence of such exponents has been answered by Oseledec.[10] The convergence of the limits in (24) is assured for almost all x_0, provided a T^t invariant measure μ exists. The concrete evaluation of Lyapunov exponents has also been investigated, e.g., in references.[2,3] Our method presented in Sec. IV goes along the line of some of the ideas of the quoted works. We make use of the Carleman method, which besides simplifying the evaluation of the DT^t, offers the possibility of easily increasing the accuracy of the numerical work.

IV. APPROXIMATE SCHEME AND DISCUSSION

The major difficulty we are faced with while working with the Carleman embedding lies in the exact determination of the time evolution operator e^{Mt} which, as has become clear from the previous discussion, is equivalent to summing up the Taylor series for the trajectory $x(t)$. However, we can use the Carleman method as a starting point for approximative schemes which are useful in connection with computing facilities. In what follows we will take the Lorenz model[11]

$$\dot{x} = \sigma(y - x),$$
$$\dot{y} = -y + rx - xz, \qquad (25)$$
$$\dot{z} = -bz + xy$$

as an example and evaluate its largest Lyapunov exponent λ_3. This model belongs to the class of systems described by (1), and has been chosen as testing object because some results on its Lyapunov exponents have been published, which will be used for comparison with ours.[3] In (25), σ, r, and b are constant parameters, while x, y, and z denote the variables y_i, $i = 1,2,3$.

We consider the Carleman solution (10) for the Lorenz model. Our approximation scheme starts by cutting off M to get M_s, which contains only the first $s + 1$ diagonal and the s upper-diagonal blocks. Then the approximate time-evolution operator $e^{M_s t}$ (determined, e.g., by numerical methods), applied to the cut off vector formed by the first $s + 1$ block components X_L, leads to an approximate trajectory which differs from the exact one by terms of the order t^{s+1}. This fact induces us to proceed as usual in the numerical methods for differential equations: If we want to integrate Eq. (1) until a time T, we divide it in subintervals at length τ such that the error introduced in the approximate solution is sufficiently small, and integrate until T step by step. The iteration procedure for the trajectory with the help of the matrix M_s is based

TABLE I. Values for the largest Lyapunov exponent of the Lorenz model with the standard values of the parameters $\sigma = 10$, $b = 8/3$, $r = 28$. s characterizes the cutoff matrix M_s and gives also the accuracy of integration, across a step τ, of the order of τ^s. We have taken $\tau = 0.01$, and the values of λ are given after a number n of 1.4×10^5 iterations. We list the results for five arbitrary trajectories, characterized by their starting points.

Start Point \ s	4	5	6
(1.0,1.0,1.0)	0.8981	0.9009	0.9023
(-0.1,27.0,0.1)	0.9084	0.8969	0.9059
(3.0,-2.0,5.0)	0.9024	0.9045	0.9017
(-9.0,11.0,12.0)	0.9011	0.9093	0.9032
(-0.1,27.0,-0.1)	0.9153	0.9089	0.9046

TABLE II. Results for the Lyapunov exponent of the Lorenz model with the Shimada values for the parameters $\sigma = 16$, $b = 4$, $r = 40$. Here we have also $\tau = 0.01$ and $n = 1.4 \times 10^5$. The starting point of the trajectory used in the evaluation of λ is (1.0, 1.0, 1.0); s has the same meaning as in the Table I.

s	4	5	6
λ	1.3704	1.3620	1.3707

on the following scheme:

(i) $X_1(n\tau) = \sum_{L=1}^{s+1} (e^{M_s \tau})_{1,L} \, X_L(n\tau - \tau)$,

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (26)

(ii) $X_L(n\tau) = X_1(n\tau)^{[L]}$.

For the Lyapunov exponents λ_3, which we will hereafter call $\lambda (x_0, \delta x)$, we have

$\lambda (x_0, \delta x) = \lim_{n \to \infty} \frac{1}{n\tau} \ln \left[\frac{\| DT^{n\tau} x_0 \, \delta x_0 \|}{\| \delta x \|} \right]$

$= \lim_{n \to \infty} \frac{1}{n\tau}$ $\qquad\qquad\qquad\qquad$ (27)

$\times \ln \left[\frac{\| DT^\tau(x_{n-1}) DT^\tau(x_{n-2}) \cdots DT^\tau(x_0) \delta x \|}{\| \delta x \|} \right]$,

where we have used the properties

$DT^{(\tau + \eta)} x = DT^\tau(T^\eta x) DT^\tau x$

and

$x_n = x(n\tau)$. $\qquad\qquad\qquad\qquad\qquad\qquad$ (28)

Equation (28) follows immediately from (11) in the exact summation and is valid also within our approximation scheme if τ is small and s large. Writing $DT^\tau(x_{n-1}) \delta x_{n-1} = \delta x_n \| DT^\tau(x_{n-1}) \delta x_{n-1} \|$, $\delta x_0 = \delta x$, we get from (27),

$\lambda (x_0, \delta x) = \lim_{n \to \infty} \frac{1}{n\tau} \ln \left[\prod_{k=0}^{n-1} \frac{\| DT^\tau(x_k) \delta x_k \|}{\| \delta x_k \|} \right]$

$= \lim_{n \to \infty} \frac{1}{n\tau} \sum_{k=0}^{n-1} \ln \frac{\| DT^\tau(x_k) \delta x_k \|}{\| \delta x_k \|}$. $\qquad$ (29)

$\lambda (x_0, \delta x)$ is then evaluated within our scheme with the help of the expressions (26) and (29). In each step we evaluate the $X_1(n\tau)$ and the $DT^\tau(x_{n-1}) \delta x_{n-1}$, besides the $X_L(n\tau)$, δx_n, and $DT^\tau(x_n)$ which are necessary to go further with the iterative procedure.

We present some of our results for the Lorenz model obtained with the help of the above described method in Tables I and II. Those in Table II may be compared with the calculations of Shimada,[3] who has obtained the same value for λ as ours. The results in the Table I, where values of λ for different start points and increasing size of the cut off matrix M_s are presented, show a rather good convergence and support the suggestion that, for the Lorenz model, λ should not depend upon the starting point.

We see two major practical advantages in our approximation scheme. In evaluating the Lyapunov exponents we just have to work with one time-evolution operator, which takes into account both the trajectory and its variation. This makes the separate integration of the Eqs. (1) and (2) unnecessary, requires just one basic time interval τ within the procedure, ensures equal accuracy for the evaluated time development of x and δx, and reduces the sources of possible rounding off errors.

The second interesting point refers to the possibility of easily increasing the accuracy of the procedure without any further difficult than increasing the size of the cut off matrix M_s. This remains in contrast with the normally used Runge–Kutta methods: despite their major advantage of requiring only the knowledge of the first time derivative of the variables, they become increasingly complicated if we want to get higher accuracy, and new formulae must be derived for each new case.

Concluding, we point out that the Carleman embedding has provided a useful method for evaluating Lyapunov exponents of a large class of systems. It is based on a rigorous analytic result and on an iterative numerical procedure to avoid the problem of the impossibility of determining the exact time-evolution operator. Whereas the solution of (1) by means of the Carleman embedding has already been analyzed, an investigation of the interdependency between the first variational equation and the Carleman method was needed.

Moreover, the evaluation of Lyapunov exponents enables us to test the viability of iterating the Carleman method itself, for Lyapunov exponents are quantities which may be used for comparision between our results and others. On the other hand, comparing individual trajectories of systems with sensitive dependence on the initial point, as that of Lorenz, which were obtained by two different numerical methods, makes no sense at all.

[1] T. Carleman, Acta Mathematica **59**, 63 (1932).
[2] G. Benettin, L. Galgani, and J. M. Strelcyn, Phys. Rev. A **14**, 2338 (1976).
[3] I. Shimada and T. Nagashima, Prog. Theor. Phys. **61**, 1605 (1979).
[4] W. H. Steeb and F. Wilhelm, J. Math. Anal. Appl. **77**, 601 (1980).
[5] R. F. S. Andrade and A. Rauh, Phys. Lett. A **82**, 276 (1980).
[6] L. Brenig and V. Fairen, J. Math. Phys. **22**, 649 (1981).

[7]J. Schröter, "Some Comments on the Carleman Linearization Procedure," Preprint-AG Theoretische Physik, Universität Paderborn, Paderborn, West Germany.

[8]V. V. Nemytskii and V. V. Stepanov, *Qualitative Theory of Differential Equations* (Princeton U. P., Princeton, NJ, 1960).

[9]O. E. Lanford, "Strange Attractors and Turbulence," in *Hydrodynamic Instabilities and Transition to Turbulence*, edited by H. L. Swinney and J. P. Gollub (Springer, Berlin, 1981).

[10]V. I. Oseledec, Trans. Moscow Math. Soc. **22**, 649 (1968).

[11]E. Lorenz, J. Atmos. Sci. **20**, 130 (1963).