JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 26, No. 2, March–April 2003

# Timescale Analysis for Nonlinear Dynamical Systems

K. D. Mease,* S. Bharadwaj,† and S. Iravanchy‡
*University of California, Irvine, California 92697*

**Insight into the behavior and simplified control of a nonlinear dynamical system can be gained by analyzing the timescale structure. Near an equilibrium point, the eigenvalues and eigenvectors for the linearized system provide the necessary information. Nonlinear systems often operate on multiple timescales away from equilibrium, but there has been no general systematic approach to determine these timescales and the associated geometric structure of the state space. A timescale analysis method based on Lyapunov exponents and vectors is synthesized, and its theoretical basis is established. As an initial demonstration the method is applied to an example system, for which the timescale structure is known by other means, and is shown to yield the correct results.**

## Introduction

If a system operates on two or more disparate timescales, there is an opportunity for decomposing a mathematical model of the system on the basis of the timescale separation and thus simplifying the analysis and design of the system. Flight guidance problems provide particular motivation. Guidance problems are typically nonlinear, require reaching a target in finite time, and often do not have near-equilibrium or near-periodic solutions. Guidance problems are frequently formulated as optimal control problems because high performance is required. Especially for autonomous onboard guidance, simplification of the guidance problem is highly desirable. There have been a number of successes in simplifying guidance problems on the basis of timescale separation in the nonlinear translational dynamics using the singular perturbation approach.[1] The steps in the singular perturbation approach are 1) to express the equations of motion in singularly perturbed form and 2) to use the singular perturbation method to construct a solution to the guidance problem as the composite of solutions to reduced-order problems, one for each timescale.[1,2] Whereas the second step is systematic and has a theoretical basis,[2] the first step has been ad hoc, with a few exceptions mentioned in the next paragraph, and requires a priori insight into the timescale structure for success.

Several techniques have been proposed for casting the equations of motion in an appropriate singularly perturbed form. For a given set of state variables, several authors[3–5] have nondimensionalized the state variables, by scaling altitude by a characteristic altitude, time by a characteristic time, etc., thereby identifying a small parameter and the singularly perturbed structure of the dynamics. Rather than scaling variables, some researchers[6,7] have inserted small parameters in the equations of motion to achieve a singularly perturbed form, thus embedding the original system in a family of systems parameterized by the small parameters. For either of these approaches to be successful, the state variables must have been selected appropriately, implying either good fortune or most likely some a priori knowledge of the timescale structure. Other investigators[6,8,9] have derived full or partial state transformations via the solution of partial differential equations to eliminate slow/fast coupling in the equations of motion. The derived transformation is from a given set of state variables to a new set of state variables, and a priori knowledge of which subset of the given state variables is primarily responsible for the fast motion is required. The feature that a pure fast variable should be constant during the slow behavior has also guided the construction of appropriate fast variables. Both the scaling and decoupling approaches can also be found in the general singular perturbations literature.[2]

Some fundamental questions arise regarding the timescales of a nonlinear dynamic system. For a linear time-invariant system or a nonlinear system operating in the neighborhood of an equilibrium point, the timescales are characterized by eigenvalues. For a periodic linear system, or a nonlinear system operating in the neighborhood of a periodic orbit, the timescales are characterized by Floquet exponents. In the analysis and design of nonlinear systems operating in transient regions of the state space, that is, away from equilibrium or periodic solutions, it would be beneficial to determine the timescale structure of the dynamics. How should timescales be defined? How can they be computed? If we think of the "dynamic system" as the physical system we are modeling, are the timescales an inherent property of the system that would be present in any mathematical model of the system, independent of the coordinates one uses, or are the timescales coordinate dependent?

Answers to these questions are developed in this paper, drawing heavily from the foundations laid by mathematicians, in the area now called dynamical systems, starting with the work of Lyapunov.[10] Some other key contributors are Perron, Hadamard, Oseledec,[11] and Barreira and Pesin.[12] Our work also draws from the more applied mathematics of Lorenz,[13] Greene and Kim,[14] Goldhirsch et al.,[15] and Lam.[16] The contribution of the present paper is to extract and adapt results from dynamical systems theory to the timescale analysis of multiple timescale guidance-type problems. "Guidance-type" refers to the mathematical features stated in the first paragraph; our results would be applicable in other areas with problems with the same features. There are three steps in reaching our goal: 1) resolving theoretical and practical issues in determining the timescale structure in the linearized dynamics about trajectories, 2) resolving the theoretical and practical issues in determining the timescale structure of the nonlinear system from that of the linearized dynamics, and 3) establishing effective computation methods. This paper primarily addresses the first step and includes results from several conference papers[17–19] and a dissertation.[20]

Lyapunov[10] introduced characteristic exponents, now called Lyapunov exponents, as a general means of determining the timescales and stability of nonlinear systems. Under certain hypotheses, Oseledec[11] showed that the Lyapunov exponents define a geometric structure associated with the linearized dynamics along a trajectory of the nonlinear system. Barreira and Pesin[12] developed a theory for translating the linear structure to nonlinear structure, generalizing the stable and unstable manifold theorems for equilibria to trajectories, building on the earlier work of Perron and Hadamard. There is an extensive literature on Lyapunov exponents in mathematics (in particular, dynamical systems), physics, and numerous application areas, much of which concerns chaotic

Received 18 April 2001; revision received 2 August 2002; accepted for publication 9 October 2002. 

*Professor, Department of Mechanical and Aerospace Engineering. Associate Fellow AIAA.

†Graduate Research Assistant, Department of Mechanical and Aerospace Engineering; currently Research Staff, Electronic and Photonic Systems Group, General Electric India Technology Center, Bangalore 560 066, India.

‡Graduate Research Assistant, Department of Mechanical and Aerospace Engineering.

attractors. With an eye toward model decomposition in the spirit of boundary-layer-type singular perturbations, we are interested in not only the exponents but also the Lyapunov vectors. Together the Lyapunov exponents and vectors characterize quantitatively how a sphere of initial states propagates into an ellipsoid of final states, according to the linearized dynamics along a trajectory of a dynamic system. Computing the Lyapunov exponents and vectors for trajectories spanning the region of interest in the state space provides of means of determining the timescale structure for the region; the Lyapunov vectors in particular indicate appropriate coordinates for decomposing the system. The Lyapunov vectors have received far less attention than the Lyapunov exponents. Lorenz[13] obtained the Lyapunov exponents and vectors via a singular value decomposition, which is the approach we use in this paper. Greene and Kim[14] and Goldhirsch et al.[15] derived evolution equations for the Lyapunov exponents and vectors and investigated their asymptotic properties. These investigators[13–15] used the Lyapunov exponents and vectors to determine the dimension and structure of a chaotic attractor. Abarbanel et al.[21] use finite time Lyapunov exponents to characterize local system behavior on a chaotic attractor. Vastano and Moser[22] used Lyapunov vectors and kinematic eigenvalues, quantities related to the Lyapunov exponents, to understand fluid flow mechanisms. Wiesel has contributed to the theory of finite-time Lyapunov exponents[23] and applied them to trajectory tracking law design.[24]

This paper has some tutorial content for the purpose of making the mathematics we are drawing from understandable to more readers. Whereas linear algebra and vector spaces provide the means of geometrically characterizing linear systems, differential geometry provides the means of geometrically characterizing nonlinear systems. The use of differential geometry in this paper is thus unavoidable, but we have attempted to write the paper in a way that does not require prior knowledge of differential geometry. The geometric structure of a two-timescale system is introduced in the next section using a simple example in which the geometry can be visualized as a distorted version of the familiar geometry of a linear system. This concrete example of the timescale-induced geometric structure of a nonlinear system clarifies the objective of the subsequent developments.

## Nonlinear Two-Timescale System Geometry: Motivating Example

Consider two different state variable representations of the same dynamic system. In terms of the state variables $w_1$ and $w_2$, the dynamics are linear, time invariant (LTI):

$$\dot{\boldsymbol{w}} = A\,\boldsymbol{w} = \begin{bmatrix} -1 & 0 \\ 0 & -10 \end{bmatrix} \boldsymbol{w} \tag{1}$$

In terms of the alternative state variables $x_1$ and $x_2$, the dynamics are nonlinear, time invariant (NTI):

$$\begin{aligned} \dot{x}_1 &= -x_1 - a(x_1 + 19x_2)(x_1 + x_2) - 18a^2(x_1 + x_2)^3 \\ \dot{x}_2 &= -10x_2 - 2a(4x_1 - 5x_2)(x_1 + x_2) + 18a^2(x_1 + x_2)^3 \end{aligned} \tag{2}$$

The nonlinear coordinate transformation that relates these two representations is $w_1 = x_1 - a(x_1 + x_2)^2$ and $w_2 = x_2 + a(x_1 + x_2)^2$. The LTI system has a two-timescale structure because the eigenvalues of $A$ are $-1$ and $-10$, and the structure is uniform because the system is LTI. The nonlinear system will have uniform two-timescale structure, if the state transformation does not significantly modify the timescales. In the $(w_1, w_2)$ space, the two-timescale behavior is evident. The general solution has the form $\boldsymbol{w}(t) = w_2(0)e^{-10t}\boldsymbol{v}_1 + w_1(0)e^{-t}\boldsymbol{v}_2$, where $\boldsymbol{w} = (w_1, w_2)^T$ is the state vector and $\boldsymbol{v}_1 = (0, 1)^T$ and $\boldsymbol{v}_2 = (1, 0)^T$ are the eigenvectors of $A$. At each point in the space, the vertical direction, $\boldsymbol{v}_1$, is the direction of fast motion evolving at the rate $e^{-10t}$, and the horizontal direction, $\boldsymbol{v}_2$, is the direction of slow motion evolving at the rate $e^{-t}$.

From Fenichel's geometric perspective,[25] the trajectories of a two-timescale system are organized in the state space by a slow manifold and a family of fast manifolds. In the $(w_1, w_2)$ space (Fig. 1a), the slow manifold is the $w_1$ axis and the vertical line through each point on the slow manifold is a fast manifold. Trajectories from initial points off the slow manifold proceed quickly in the vertical direction, that is, along the fast manifold through the initial point, toward the slow manifold at the fast rate $e^{-10t}$ and subsequently proceed along the slow manifold at the slower rate $e^{-t}$. This description is only approximate for finite timescale separation because motion in the slow direction does occur during the approach to the slow manifold, and thus, the trajectory does not exactly follow the fast manifold on which it begins. At each point in the state space, the timescales for the LTI system are given by the eigenvalues of the matrix $A$, and the fast and slow directions of the motion are given by the eigenvectors of the matrix $A$.

The trajectories of the NTI system are shown in the $(x_1, x_2)$ space in Fig. 1b for $a = 0.01$. Note that the two-timescale structure is still present, but the slow and fast manifolds are now curved rather than straight. In a sufficiently small neighborhood of the equilibrium point, the timescale structure can be determined from the eigenvalues and eigenvectors for the linearized dynamics. In guidance problems, however, our interest is typically in a compact transient region $\mathcal{X}$, the operating region, such as shown in Fig. 1b. For our example NTI system, there is two-timescale structure in this region, and thus the potential exists for simplified analysis and design.

Our ultimate goal is to determine state variables that reflect the timescale structure and allow the dynamics to be decomposed on this basis. In our contrived example, the goal is to be given the $\boldsymbol{x}$ representation and determine the $\boldsymbol{w}$ representation. The $(w_1, w_2)$ coordinate curves are superimposed in Fig. 1b to indicate the information we desire; viewed in the $\boldsymbol{x}$ frame, $(w_1, w_2)$ are curvilinear coordinates. The steps we achieve in this paper are to determine the characteristic numbers (the Lyapunov exponents) that give the timescales and to determine the corresponding vectors (the Lyapunov vectors) that give at each point in the region $\mathcal{X}$ the directions of the desired curvilinear coordinates. The Lyapunov vectors are shown in Fig. 1b at three points in the region $\mathcal{X}$. The bold vectors are the Lyapunov vectors, one arrowhead for slow and two for fast, and give the correct directions. The shorter, lighter vectors are the local eigenvectors; they do not uniformly give the desired directions. Note that the equivalent LTI system was employed in this motivating example as a means of more simply introducing the two-timescale structure. Our approach does not require that the dynamics in the desired state variables be linear, as appropriate because in general a nonlinear dynamic system cannot be transformed into a linear dynamic system.

In the remainder of the paper we will synthesize and justify a method of timescale analysis based on finite time Lyapunov exponents and vectors. After establishing our assumptions and terminology, we define and discuss the timescale information that can be obtained at each point in the state space by integrating the linearized dynamics along the trajectory of the nonlinear system through that point. In general the timescale information is metric and coordinate dependent; we treat both the Euclidean metric and the general Riemannian metric, as well as the effect of coordinate transformations. The timescale information characterizes the average exponential behavior in the linearized dynamics over a time interval. We consider the properties of the timescale information in the infinite time limit, in particular the convergence and metric independence. We next consider the convergence rate with regard to computational feasibility. Convergence requires the presence of two or more sufficiently disparate timescales. This is a fortunate result that means that if multiple timescale behavior exists, then computing the corresponding timescale information is feasible. Finally a procedure is outlined and applied to the example of this section.

## Terminology, Assumptions, and Evolution Equations

Our objective is to characterize the timescale structure of a system whose state $\boldsymbol{x}$ evolves in the space $\mathbb{R}^m$ according to the NTI dynamics

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) \tag{3}$$

where $\dot{\boldsymbol{x}}$ is the time derivative of the state and $\boldsymbol{f}$ is a smooth function of the state. Equation (3) is said to define a vector field on the state space. Although our work is motivated by flight guidance problems, for which one would expect a control-dependent vector

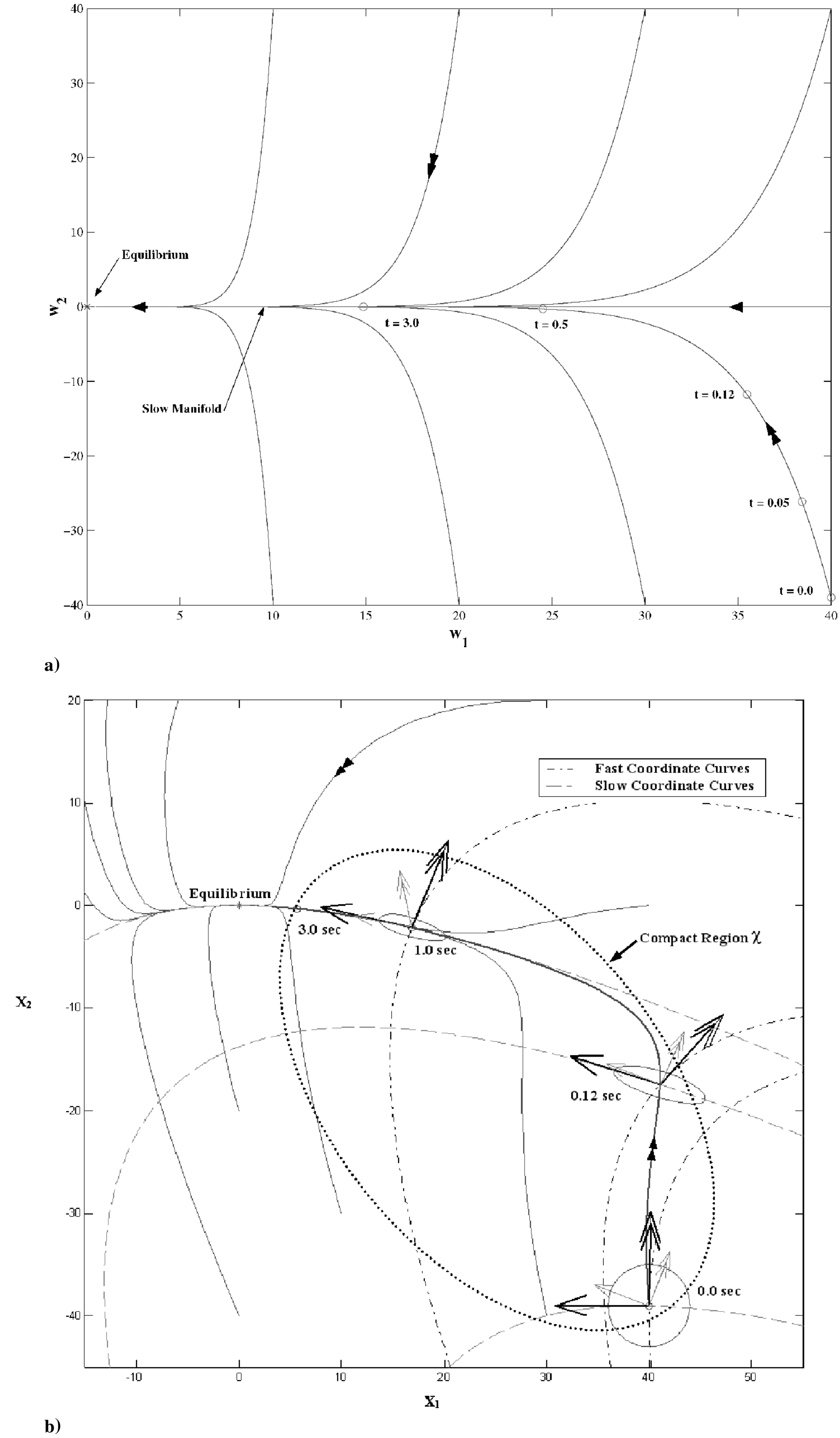


**Fig. 1 State portraits with two-timescale structure for a) an LTI system and b) a related NTI system.**

field, $\dot{x} = f(x, u)$ with $u$ the vector of control variables, we do not explicitly consider control dependence. Our analysis of $\dot{x} = f(x)$ is relevant for a controlled system in any of the following cases: 1) $u$ is determined by a feedback law and $f(x)$ represents the closed-loop dynamics, 2) $f(x)$ is the open-loop dynamics and we do not plan to alter the timescale structure by control activity, or 3) $\dot{x} = f(x)$ is the Hamiltonian system for an optimally controlled system, with $x$ being the combined state and costate vector.

Unlike an LTI system, the timescale structure of a nonlinear system can be different in different regions of the state space. Thus, we focus on a compact connected set $\mathcal{X}$, a subset of the state space $\mathbb{R}^m$, that is an operating region whose timescale structure we wish to determine. The solution, or orbit, evolving from the initial point $x \in \mathcal{X}$, is denoted by $x(t) = \phi(t, x)$, where $\phi$ satisfies $\partial\phi(t, x)/\partial t = f[\phi(t, x)]$, for each value of $t$ under consideration, and $\phi(0, x) = x$. [We use $\phi(t, x)$ rather than $x(t)$ because we want

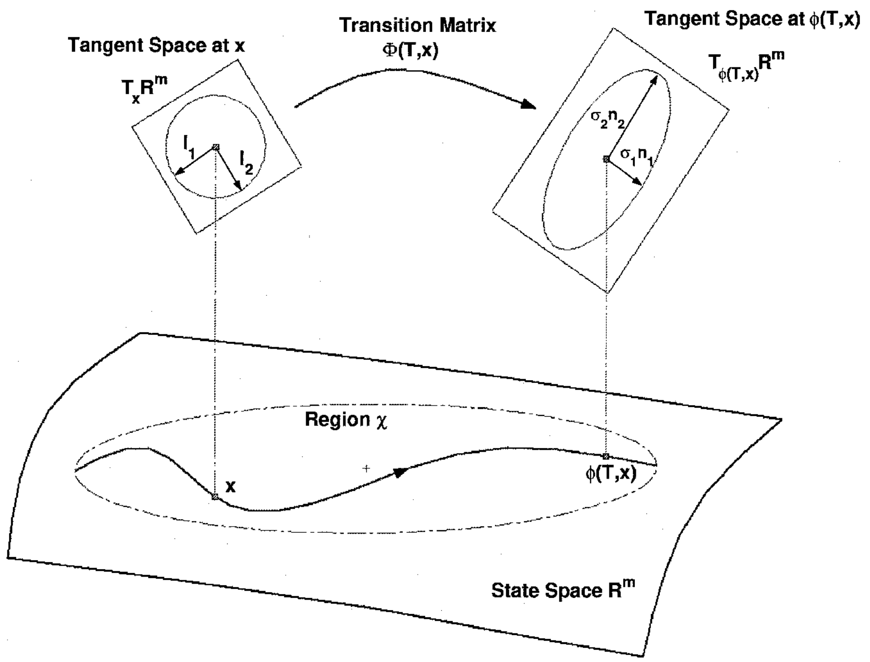


**Fig. 2 Evolution of a circle to an ellipse.**

to use $x$ for the initial state.] We assume that the orbit exists at least as long as it remains within $\mathcal{X}$. For each $x \in \mathcal{X}$, there is a maximal time interval $\mathcal{T}(x) = [\underline{t}(x), \bar{t}(x)]$ that includes zero, such that $\phi(t, x) \in \mathcal{X}$ for all $t \in \mathcal{T}(x)$. The minimum time to reach transversely the boundary of $\mathcal{X}$ from $x$ is $\bar{t}(x)$ for forward time propagation and $\underline{t}(x)$ for backward time propagation. If $x$ is on the boundary of $\mathcal{X}$, then one of these times can be zero. If the orbit never leaves $\mathcal{X}$ in forward time, then $\bar{t}(x)$ is taken to be infinite and similarly for backward time. If $\bar{t}(x) = \infty$ for all $x \in \mathcal{X}$, then $\mathcal{X}$ is a positively invariant set for the dynamics under consideration; if $\underline{t}(x) = -\infty$ for all $x \in \mathcal{X}$, then $\mathcal{X}$ is a negatively invariant set; if $\mathcal{T}(x) = (-\infty, \infty)$ for all $x \in \mathcal{X}$, then $\mathcal{X}$ is invariant.

The associated linear dynamic system is

$$\frac{\partial v}{\partial t} = F[\phi(t, x)]v \tag{4}$$

where $F = \partial f/\partial x$ and $v(t, x)$, an $m$-dimensional vector, can be interpreted as a small perturbation in $x$ that evolves approximately according to the linear dynamics, or as a tangent vector of any size that evolves exactly according to the linear dynamics. In this paper, we shall develop a methodology for computing and interpreting the timescale structure for the linear dynamic system associated with orbit segments covering $\mathcal{X}$. The next step, which is beyond the scope of this paper, is to transfer the linear timescale structure to the nonlinear system, for example, using the linear results to determine slow and fast manifolds in $\mathcal{X}$. The rigorous theory supporting this step is given by Barreira and Pesin.[12]

The solution to Eq. (4) can be expressed as $v(t, x) = \Phi(t, x)v(0, x)$, where $\Phi$ is the transition matrix. $\Phi$ satisfies Eq. (4), that is, $\partial\Phi/\partial t = F[\phi(t, x)]\Phi$; also, $\Phi(0, x) = I$, where $I$ is the $m \times m$ identity matrix. The dependence of $v$ on $x$ is due to the dependence of $F$ on $\phi(t, x)$, that is, $x$ indicates the initial point of the orbit with respect to which the dynamics in Eq. (4) are linearized. $\Phi$ is related to the nonlinear solution operator $\phi$ by $\Phi(t, x) = \partial\phi(t, x)/\partial x$. Interpreting $v$ as a tangent vector, the system composed of Eqs. (3) and (4) describes the evolution of a point with coordinates $(x, v)$ in a $2m$-dimensional space $T\mathcal{X}$ referred to as the tangent bundle. Spivak[26] is a reference for the differential geometric concepts we use here. For each point $x \in \mathcal{X}$, there is an associated tangent space $T_x\mathcal{X}$, the space of all possible vectors tangent to orbits passing through $x$. Note that the vector $\dot{x} = f(x)$ is just one of these possible tangent vectors. The tangent bundle is the union of the tangent spaces, $T\mathcal{X} = \bigcup_{x\in\mathcal{X}} T_x\mathcal{X}$. Focusing on a state space orbit $\phi(t, x)$ passing through $x$ at $t = 0$, for $v(0, x) \in T_x\mathcal{X}$, we have $v(t, x) \in T_{\phi(t,x)}\mathcal{X}$, where $v(t, x) = \Phi(t, x)v(0, x)$; in words, the vector $v(0, x)$ in the tangent space at $x$ evolves to the vector $v(t, x)$ in the tangent space at $\phi(t, x)$. See Fig. 2 for the case $m = 2$, where $T$ is the propagation time and the vectors $l_1$, $l_2$, $n_1$, and $n_2$ will be defined later. If we have $k$ linearly independent vector fields $f_1(x), \ldots, f_k(x)$ defined on $\mathcal{X}$ that vary smoothly with $x$, we can define at each $x$ a $k$-dimensional subspace $\Lambda(x) = \text{span}\{f_1(x), \ldots, f_k(x)\}$. If $k = m$, then $\Lambda(x) = T_x\mathcal{X}$, and for each $x$, the vector fields provide a basis for $T_x\mathcal{X}$. If $k < m$, then $\Lambda(x)$ is a subspace of $T_x\mathcal{X}$, and $\Lambda = \cup_{x\in\mathcal{X}}\Lambda(x)$ is called a subbundle of the tangent bundle $T\mathcal{X}$. If a collection of $r \le m$ subspaces of $T_x\mathcal{X}$ can be ordered such that $\Lambda_1(x) \subset \Lambda_2(x) \subset \cdots \subset \Lambda_r(x) = T_x\mathcal{X}$, then this collection defines a filtration of $T_x\mathcal{X}$. Note that the subspaces of the filtration satisfy for $i = 1, \ldots, r-1$ the conditions $\dim\Lambda_i(x) < \dim\Lambda_{i+1}(x)$, and $\Lambda_i(x)$ is a subspace of $\Lambda_{i+1}(x)$.

The simplest case to conceptualize is for an equilibrium point, that is, $x$ is a point $x_{\text{eq}}$ such that $f(x_{\text{eq}}) = 0$. Then the "orbit" is just the equilibrium point, that is, $\phi(t, x_{\text{eq}}) = x_{\text{eq}}$ for all $t$ and $\mathcal{T}(x_{\text{eq}}) = (-\infty, \infty)$. If we choose our state-space set $\mathcal{X}$ to be this single point, then for any tangent vector $v(0, x_{\text{eq}}) \in T_{x_{\text{eq}}}\mathcal{X}$ we only need to consider this tangent space because $v(t, x) \in T_{x_{\text{eq}}}\mathcal{X}$ for all $t$. When $e_1, \ldots, e_m$ denote the generalized (real) eigenvectors of $F(x_{\text{eq}})$, a filtration of $T_{x_{\text{eq}}}\mathcal{X}$ can be defined by $\Lambda_1 = \text{span}\{e_1\}$, $\Lambda_2 = \text{span}\{e_1, e_2\}, \ldots, \Lambda_m = \text{span}\{e_1, e_2, \ldots, e_m\}$. Alternatively one could apply Gram–Schmidt orthogonalization to the eigenvectors, commencing with $e_1$ and working up in index and use the resulting orthogonal basis to define the same filtration.

We use $G$ to define the norm by which the length of a vector is measured. $G$, in the most general case we consider, is an $m \times m$ matrix-valued smooth bounded function on $\mathcal{X}$ that is symmetric and positive definite for all $x \in \mathcal{X}$. In other words, $G$ defines a Riemannian metric on $\mathcal{X}$. The associated inner product for vectors $u, v \in T_x\mathcal{X}$ is $\langle u, v\rangle_{G(x)} = u^T G(x)v$. The length of $v \in T_x\mathcal{X}$ is $\|v\|_{G(x)} = \langle v, v\rangle_{G(x)}^{1/2}$. The simplest case is when $G = I$, the constant identity matrix, which corresponds to the Euclidean metric.

## Singular Value Decomposition of Transition Matrix

The transition matrix $\Phi(t, x)$, for any $x \in \mathcal{X}$ and $t \in \mathcal{T}(x)$, has a singular value decomposition[27] (SVD)

$$\Phi = N\Sigma L^T \tag{5}$$

where $N$ and $L$ are $m \times m$ orthogonal matrices with respect to the Euclidean metric and $\Sigma$ is a diagonal matrix with positive diagonal elements called the singular values $\sigma_i$, $i = 1, \ldots, m$. Although the results of this paper can be adapted to the case of nondistinct singular values, to simplify the presentation we assume that the singular values are distinct. We assume the ordering $0 < \sigma_1 < \sigma_2 < \cdots < \sigma_m$.

Because the SVD of $\Phi(t, \boldsymbol{x})$ plays a central role, we discuss its construction and geometric interpretation, suppressing the arguments of $\Phi$ to simplify the notation. Let $\gamma_i$, $i = 1, \ldots, m$ denote the eigenvalues of $\Phi^T\Phi$. Because $\Phi$ is square and nonsingular, $\Phi^T\Phi$ is a symmetric positive definite matrix, and all its eigenvalues are positive. The eigenvalues can have multiplicity greater than one, but consistent with our assumption of distinct singular values, we assume distinct eigenvalues. Define the eigenspace associated with the eigenvalue $\gamma_i$ as the nonzero linear subspace $U_i \subset \mathbb{R}^m$ such that $\Phi^T\Phi\boldsymbol{v} = \gamma_i\boldsymbol{v}$ for all vectors $\boldsymbol{v} \in U_i$ (Ref. 28). There are $m$ uniquely defined one-dimensional eigenspaces under the assumption of distinct eigenvalues. Each eigenspace can be written as the span of a unit length vector, and we use the term eigenvector to refer to this vector. The eigenvectors are, thus, uniquely defined up to multiplication by $\pm 1$. Because of the symmetry of $\Phi^T\Phi$, the eigenvectors are mutually orthogonal. The eigenvalues and corresponding eigenvectors can always be ordered such that $\gamma_1 < \gamma_2 < \cdots < \gamma_m$, and this ordering is assumed in the remainder of the paper.

Next define $N = \Phi L \Sigma^{-1}$, where $\Sigma$ is the diagonal matrix with diagonal elements $\sigma_i = +\sqrt{\gamma_i}$, $i = 1, \ldots, m$, and $L$ is the matrix whose columns $\boldsymbol{l}_i$, $i = 1, \ldots, m$, comprise a set of eigenvectors for $\Phi^T\Phi$. Because the vectors $\boldsymbol{l}_i$, $i = 1, \ldots, m$, constitute an orthonormal basis for $\mathbb{R}^m$, the matrix $L$ is orthogonal, that is, $L^T L = I$. It can be shown that $N$ is also orthogonal ($N^T N = I$) and that the column vectors of $N$, denoted by $\boldsymbol{n}_i$, $i = 1, \ldots, m$, are the eigenvectors of $\Phi\Phi^T$, that is, $N$ satisfies $\Phi\Phi^T N = N\Sigma^2$. Rearranging $N = \Phi L \Sigma^{-1}$, we obtain the SVD of the transition matrix $\Phi = N\Sigma L^T$.

The transition matrix, because it is square and invertible, also has the polar decomposition[15,27]

$$\Phi = QS \tag{6}$$

where $S$ is the symmetric positive definite $m \times m$ matrix that satisfies $S^2 = \Phi^T\Phi$ and $Q = \Phi S^{-1}$. This follows directly from the SVD. We have $\Phi = N(L^T L)\Sigma L^T = (NL^T)(L\Sigma L^T) = QS$. It can be shown that the matrix $Q = NL^T$ is a rotation matrix.

## Timescale Information

### Finite Time Lyapunov Exponents and Vectors

A sphere of initial conditions for $\boldsymbol{v} \in T_x\mathcal{X}$, propagated along an orbit of the nonlinear system according to the linearized dynamics, evolves into an ellipsoid.[11] By analyzing this behavior, one can characterize the timescales of the nonlinear system and the corresponding geometric structure.[13–15,23,29] A vector $\boldsymbol{v} \in T_x\mathcal{X}$, propagated for $T$ units of time along the orbit $\phi(t, \boldsymbol{x})$, evolves to the vector $\boldsymbol{v}_T = \Phi(T, \boldsymbol{x})\boldsymbol{v}$. The Euclidean length of the initial vector is $\|\boldsymbol{v}\| = (\boldsymbol{v}^T\boldsymbol{v})^{1/2}$ and of the final vector is $\|\boldsymbol{v}_T\| = \|\Phi(T, \boldsymbol{x})\boldsymbol{v}\| = (\boldsymbol{v}^T\Phi^T\Phi\boldsymbol{v})^{1/2}$. Given the positive definiteness of $\Phi^T\Phi$, it is clear that a unit sphere of initial conditions, expressed as the set $\mathcal{S}(\boldsymbol{x}) = \{\boldsymbol{v} \in T_x\mathcal{X} : \|\boldsymbol{v}\| = 1\}$, propagates to an ellipsoid of final conditions $\mathcal{S}[\phi(T, \boldsymbol{x})] = \{\boldsymbol{v}_T \in T_{\phi(T,x)}\mathcal{X} : \boldsymbol{v}_T = \Phi(T, \boldsymbol{x})\boldsymbol{v}$, for some $\boldsymbol{v} \in \mathcal{S}(\boldsymbol{x})\}$. The same statement can be made for a non-Euclidean metric, only in this case the "sphere" is defined by the property $\|\boldsymbol{v}\|_G = [\boldsymbol{v}^T G(\boldsymbol{x})\boldsymbol{v}]^{1/2} = 1$ and the "ellipsoid" is adapted similarly.

The ratio of the lengths of an initial nonzero vector and its corresponding final vector, $\sigma = \|\boldsymbol{v}_T\|_{G[\phi(T,x)]}/\|\boldsymbol{v}\|_{G(x)}$, is a multiplier that characterizes the net expansion (growth), if $\sigma > 1$, or contraction, if $\sigma < 1$, of the vector over the time interval $(0, T)$. The square of this multiplier is

$$\sigma^2(T, \boldsymbol{x}, \boldsymbol{v}) = \frac{\|\boldsymbol{v}_T\|^2_{G[\phi(T,x)]}}{\|\boldsymbol{v}\|^2_{G(x)}} = \frac{\|\Phi(T, \boldsymbol{x})\boldsymbol{v}\|^2_{G[\phi(T,x)]}}{\|\boldsymbol{v}\|^2_{G(x)}}$$

$$= \frac{\boldsymbol{v}^T\Phi^T(T, \boldsymbol{x})G[\phi(T, \boldsymbol{x})]\Phi(T, \boldsymbol{x})\boldsymbol{v}}{\boldsymbol{v}^T G(\boldsymbol{x})\boldsymbol{v}} \tag{7}$$

[the Rayleigh quotient in linear algebra (see Ref. 27)].

In the case of the Euclidean metric, $G \equiv I$, the information in the SVD of the transition matrix $\Phi$ allows us to characterize the multipliers for all of the vectors in the tangent space $T_x\mathcal{X}$. Postmultiplying both sides of $\Phi(T, \boldsymbol{x}) = N\Sigma L^T$ by $L$ and using the orthogonality property $L^T L = I$, we have

$$\Phi L = N\Sigma \tag{8}$$

This equation allows us to interpret Fig. 2 and more generally to characterize the propagation of the (hyper)sphere into the (hyper)ellipsoid in a space of any finite dimension. Equation (8) indicates that the column vectors of $L$ evolve into the directions given by the column vectors of $N$. The column vectors of $L$ should be viewed as vectors in $T_x\mathcal{X}$, whereas the column vectors of $N$ should be viewed as vectors in $T_{\phi(T,x)}\mathcal{X}$. The diagonal elements, denoted by $\sigma_i$, of $\Sigma$ determine the amount of expansion or contraction. This is seen clearly if we write Eq. (8) in the form $\Phi\boldsymbol{l}_i = \sigma_i\boldsymbol{n}_i$, $i = 1, \ldots, m$. The vectors $\boldsymbol{n}_i$, $i = 1, \ldots, m$, are the directions of the principal axes of the ellipsoid to which the sphere of initial conditions has evolved. The vectors $\boldsymbol{l}_i$, $i = 1, \ldots, m$, are the vectors in $T_x\mathcal{X}$ that evolve into the principal axes of the ellipsoid. The $\sigma_i$, $i = 1, \ldots, m$, are the lengths of the principal semi-axes in the case when the initial sphere has unit radius. When an arbitrary initial vector is represented using the orthonormal basis provided by the column vectors of $L$ by $\boldsymbol{v} = c_1\boldsymbol{l}_1 + \cdots + c_m\boldsymbol{l}_m$ and $\Phi(T, \boldsymbol{x})\boldsymbol{l}_i = \sigma_i\boldsymbol{n}_i$, or equivalently $\Phi(T, \boldsymbol{x})L = N\Sigma$, is used the corresponding final vector is

$$\boldsymbol{v}_T = \sum_{i=1}^{m} c_i\sigma_i\boldsymbol{n}_i = N\Sigma\boldsymbol{c}$$

where $\boldsymbol{c} = (c_1, \ldots, c_m)$ is the coordinate vector for $\boldsymbol{v}$ in the basis composed of the column vectors of $L$.

We mention, but do not consider in this paper, the alternative factorization $\Phi(T, \boldsymbol{x}) = QR$, where $Q$ is an orthogonal matrix [distinct from that in Eq. (6)] and $R$ is upper triangular,[29,30] which offers another useful geometric interpretation involving the propagation of parallelepipeds and the growth/decay rates for the enclosed volumes.

Given sets $S_1$ and $S_2$ with $S_1 \subset S_2$, the notation $S_2 \backslash S_1$ refers to the set containing all of the elements in $S_2$ except for those that are also elements of $S_1$; in this sense, it is $S_2$ minus $S_1$. For example, if $S_2$ contains all of the points in a plane and $S_1$ contains the points of one line in this plane, then $S_2 \backslash S_1$ is the plane with the line of points extracted. Note that $S_2$ is still two-dimensional in this example.

*Proposition 1:* Euclidean Metric Case, $G = I$. When the SVD of the transition matrix $\Phi(T, \boldsymbol{x}) = N\Sigma L^T$ has $m$ distinct singular values, the singular values define a filtration of nested subspaces of $T_x\mathcal{X}$

$$\{0\} = L_0 \subset L_1 \subset L_2 \subset \cdots \subset L_m = T_x\mathcal{X} \tag{9}$$

by the conditions $\sigma(\boldsymbol{v}) = \sigma_1$ for all $\boldsymbol{v} \in L_1 \backslash L_0$ and for $i = 2, \ldots, m$, $\sigma_{i-1} < \sigma(\boldsymbol{v}) \le \sigma_i$ for all $\boldsymbol{v} \in L_i \backslash L_{i-1}$. The subspaces can be represented in terms of the column vectors of $L$, the matrix factor in the SVD, as

$$L_1 = \mathrm{span}\{\boldsymbol{l}_1\},\ L_2 = \mathrm{span}\{\boldsymbol{l}_1, \boldsymbol{l}_2\}, \ldots, L_m = \mathrm{span}\{\boldsymbol{l}_1, \ldots, \boldsymbol{l}_m\} \tag{10}$$

[The idea that the multipliers define a filtration is due to Oseledec[11] (also see Refs. 12 and 31)].

*Proof:* When an arbitrary initial vector is represented by $\boldsymbol{v} = c_1\boldsymbol{l}_1 + \cdots + c_m\boldsymbol{l}_m = L\boldsymbol{c}$, the squared multiplier is

$$\sigma^2 = \frac{\boldsymbol{c}^T\Sigma^2\boldsymbol{c}}{\boldsymbol{c}^T\boldsymbol{c}} = \sum_{i=1}^{m} \hat{c}_i^2\sigma_i^2 \tag{11}$$

where $\hat{c}_i^2 = c_i^2/(c_1^2 + \cdots + c_m^2)$. Because

$$\sum_{i=1}^{m} \hat{c}_i^2 = 1, \qquad \hat{c}_i^2 \ge 0, \qquad i = 1, \ldots, m$$

$\sigma^2$ is a convex combination of the squared singular values of $\Phi(T, \boldsymbol{x})$. With no additional constraints on $\hat{c}_i^2$, $i = 1, \ldots, m$, we have $\sigma_1^2 \le \sigma^2(\boldsymbol{v}) \le \sigma_m^2$, or equivalently $\sigma_1 \le \sigma(\boldsymbol{v}) \le \sigma_m$, for any nonzero vector $\boldsymbol{v}$. For the moment, define the $L_i$ subspaces, $i = 1, \ldots, m$, as in Eq. (10). If $\boldsymbol{v} \in L_1$, then $\hat{c}_2^2 = \cdots = \hat{c}_m^2 = 0$ and $\hat{c}_1^2 = 1$. Thus, $\hat{\boldsymbol{c}} = (1, 0, \ldots, 0)$ is the only possibility; the initial vector $\boldsymbol{v} = c_1\boldsymbol{l}_1$ is mapped to the final vector $\boldsymbol{v}_T = c_1\sigma_1\boldsymbol{n}_1$ and the value of $\sigma$ is $\sigma_1$. If $\boldsymbol{v} \in L_2 \backslash L_1$, then $\hat{c}_3^2 = \cdots = \hat{c}_m^2 = 0$, $\hat{c}_2 \ne 0$, and $\sigma^2 = \hat{c}_1^2\sigma_1^2 + \hat{c}_2^2\sigma_2^2$, with $\hat{c}_1^2 + \hat{c}_2^2 = 1$, is a convex combination. Thus, $\sigma^2$ is restricted to the interval $\sigma_1^2 < \sigma^2 \le \sigma_2^2$. Note that $(\hat{c}_1, \hat{c}_2) = (1, 0)$ is not allowed because it corresponds to $v = \hat{c}_1^2 l_1 + \hat{c}_2^2 l_2$ being in $L_1$. When continued in this manner, the proposition is proved. In particular the $L_i$ subspaces can be defined by the multipliers rather than by the column vectors of $L$. The column vectors of $L$ are just one basis that can be used to represent the subspaces in the filtration. This basis is distinguished by being orthonormal. □

Consider the propagation of tangent vectors in the reverse direction from $x$ to $\phi(-T, x)$. The transition matrix for this mapping is $\Phi(-T, x)$, and its SVD is $\Phi(-T, x) = N^- \Sigma^- (L^-)^T$. The factors $N^-$, $\Sigma^-$, and $L^-$ are different from those for the forward matrix $\Phi(T, x)$. In the backward time case, we assume the singular values are distinct and ordered as $\sigma_1^- > \cdots > \sigma_m^-$. The singular values define a filtration

$$T_x\mathcal{X} = L_1^- \supset L_2^- \supset \ldots \supset L_m^- \supset L_{m+1}^- = \{0\} \tag{12}$$

in that $\sigma^-(v) = \sigma_m^-$ for all $v \in L_m^- \backslash L_{m+1}^-$ and for $i = 1, \ldots, m-1$, $\sigma_i^- \geq \sigma(v) > \sigma_{i+1}^-$ for all $v \in L_i^- \backslash L_{i+1}^-$. The multiplier $\sigma^-(v)$ here is for backward time propagation, meaning integrating Eqs. (3) and (4) from an initial condition $(x, v)$ at $t = 0$ to $t = -T$. The subspaces can be represented in terms of the column vectors of $L^-$ as

$$L_1^- = \text{span}\left\{l_1^-, \ldots, l_m^-\right\}, \ldots,$$

$$L_{m-1}^- = \text{span}\left\{l_{m-1}^-, l_m^-\right\}, \; L_m^- = \text{span}\left\{l_m^-\right\} \tag{13}$$

We can also determine $L^-$ by forward integration by starting at the point $\phi(-T, x)$, computing $\Phi[T, \phi(-T, x)]$, and then performing the SVD $\Phi = N\Sigma L^T$. It follows that $L^- = N$, that is, $L^-$ for backward propagation is equal to $N$ for forward propagation over the same orbit segment.

Because our timescale information concerns how the lengths of vectors change, it is in general dependent on how the length of a vector is measured, that is, it is metric dependent. To achieve consistency in the timescale information under a state variable transformation, the metric must also be transformed such that the length of a given vector is the same for either state variable representation. We will describe later the circumstances under which the timescale information is metric independent. Consider a smooth, invertible state variable transformation $x = h(z)$ and let $H(z) = \partial h/\partial z$. It follows that the pair $(x, v)$, the coordinate representation of a point in the tangent bundle, transforms to $(z, u)$ where $v = H(z)u$. The inverse state transformation is denoted by $z = h^{-1}(x)$. In the following we label variables with a superscript $x$ or $z$ to denote the coordinate system to which they pertain.

*Proposition 2(a):* Effect of State/Metric Transformation. If the metric

$$G(z) = H^T(z)H(z) \tag{14}$$

is used in the $(z, u)$ formulation, the timescale information, the triplet $(\Sigma, L^x, N^x)$, obtained in the $(x, v)$ formulation using the Euclidean metric transforms to $(\Sigma, L^z, N^z)$, where $L^x = H(z)L^z$ and $N^x = H(z_T)N^z$, with $z_T = h^{-1}[\phi(T, x)]$, and the singular values are invariant under the transformation. $L^z$ and $N^z$ are $G$ orthogonal, meaning that they satisfy the relations $(L^z)^T G(z) L^z = I$ and $(N^z)^T G(z_T) N^z = I$.

*Proof:* The column vectors of $L^x$ and $N^x$ are vectors in $T_x\mathcal{X}$ and $T_{\phi(T,x)}\mathcal{X}$, respectively. The appropriate transformations[32] are, thus, $L^x = H(z)L^z$ and $N^x = H(z_T)N^z$. The $G$ orthogonality of the transformed matrices is verified by

$$(L^z)^T G(z) L^z = (H^{-1}L^x)^T H^T H H^{-1} L^x = (L^x)^T L^x = I \tag{15}$$

and similarly for $N^z$. The squared multiplier in terms of the transformed vector $u$ and the metric $G$ is

$$\sigma^2 = \frac{\|u_T\|_G^2}{\|u\|_G^2} = \frac{\|\Phi^z(T, x)u\|_G^2}{\|u\|_G^2} = \frac{u^T(\Phi^z)^T G(z_T)\Phi^z u}{u^T G(z) u} \tag{16}$$

The transition matrix $\Phi^x$ transforms to $\Phi^z = H^{-1}(z_T)\Phi^x H(z)$. Using the SVD for $\Phi^x$ yields $\Phi^z = N^z \Sigma (L^z)^T G(z)$. Using this result and $u = L^z c$, where $c$ is the coordinate vector for $u$ in the $L^z$ basis, we obtain

$$\sigma^2 = \frac{u^T(\Phi^z)^T G(z_T)\Phi^z u}{u^T G(z) u} = \frac{c^T \Sigma^2 c}{c^T c} = \sum_{i=1}^{m} \hat{c}_i^2 \sigma_i^2 \tag{17}$$

where $\hat{c}_i^2 = c_i^2/(c_1^2 + \cdots + c_m^2)$ and $c_i$, $i = 1, \ldots, m$ are the components of $c$. The rest of the proof proceeds just as the one for the Euclidean case. □

As in the $x$ representation, we have $\Phi^z L^z = N^z \Sigma$. Because a filtration can be defined by the characteristic multipliers alone as in proposition 1, it follows that if the multipliers are the same for two representations of the same dynamic system, then the corresponding filtrations are the same. The difference introduced by the transformation from $(x, v)$ to $(z, u)$ is that now the filtration is represented using the column vectors of $L^z$ rather than those of $L^x$ and the column vectors of $L^z$ are orthonormal with respect to the non-Euclidean metric $G$.

*Proposition 2b:* For a dynamic system $\dot{z} = f^z(z)$ and a non-Euclidean metric $G(z)$, the timescale information $(\Sigma, L^z, N^z)$ can be obtained by 1) factoring $G(z) = H^T(z)H(z)$, 2) performing an SVD of the matrix $H(z_T)\Phi^z H^{-1}(z) = N\Sigma L^T$, and 3) computing $L^z = H^{-1}(z)L$ and $N^z = H^{-1}(z_T)N$.

This proposition follows easily from proposition 2a, and so we do not provide the proof. Note that in proposition 2a we start with a coordinate transformation and specify the metric required for consistency in the timescale information. In proposition 2b, we start with a non-Euclidean metric and provide a means of computing the timescale information. If the metric $G$ has arisen from a coordinate transformation, then the matrix $H$ comes from this transformation; otherwise $H$ can be any square root of $G$.

*Definition 1:* In the Euclidean case, the numbers

$$\mu_i(T, x) = (1/T)\ln \sigma_i(T, x), \qquad i = 1, \ldots, m \tag{18}$$

associated with the singular values $\sigma_i(T, x)$, $i = 1, \ldots, m$, of a transition matrix $\Phi(T, x)$ are called the finite-time Lyapunov exponents.[21,23] They are the exponential rates associated with the principal axes of the ellipsoid and define boundaries within the continuous range of exponential rates at which the lengths of tangent vectors change (refer to Proposition 1). The column vectors $l_i(T, x)$, $i = 1, \ldots, m$, of the orthogonal matrix $L$ of the SVD for $\Phi(T, x)$ are called the finite-time Lyapunov vectors. In the non-Euclidean case, we use the same terminology for the timescale information described in propositions 2a and 2b.

*Definition 2:* The spectrum[11] of $\Phi(T, x)$ is $Sp(T, x) = \{\mu_i(T, x), i = 1, \ldots, m\}$, that is, the set of Lyapunov exponents.

Thus far we have characterized the timescale information at one point in the region $\mathcal{X}$. The timescale information is dependent on $T$ and $x$. In principle, we could compute the timescale information at every point in $\mathcal{X}$ and thereby determine the timescale structure for the entire region $\mathcal{X}$. If the timescale properties are approximately uniform on $\mathcal{X}$, then the Lyapunov exponents will not change much with $x$ and $T$; on the other hand, the column vectors of $L$ and $N$ will in general rotate as $x$ varies. Greene and Kim[14] and Wiesel[33] have derived differential equations for propagating $L$ and $N$ along an orbit; these could be used to propagate $L$ determined at one $x$ to other points on the orbit through $x$ and similarly for $N$. As noted, the timescale structure determined for $\mathcal{X}$ is dependent on the propagation time $T$. There are cases, however, where the timescale structure approaches a limit as $T$ is increased; these are discussed later.

### Kinematic Eigenvalues

In the preceding subsection, we characterized the average dynamic behavior over a time interval by analyzing the transition matrix. In this subsection, we introduce a differential equation that allows us to identify the functions of time whose averages are the finite time Lyapunov exponents and to identify a distinguished basis of solutions with regard to diagonalizing the linearized dynamics, a basis that is related to the Lyapunov vectors.

Consider the simple case of a scalar nonlinear equation $\dot{x} = f(x)$ and the associated linear dynamics $\dot{v} = F[\phi(t, x)]v$ for an orbit

$\phi(t, x)$. Let $F[\phi(t, x)]$ be denoted by $\rho(t)$, suppressing the $x$ dependence to simplify the appearance. We refer to $\rho(t)$ as the normalized instantaneous rate, in that it is equal to $\dot{v}/v$. The transition "matrix" (a scalar in this case) is

$$\Phi(t) = \exp\left[\int_0^t \rho(\tau)\,d\tau\right]$$

For the time interval $T$, we define the characteristic exponent by

$$\bar{\rho} = \frac{1}{T}\int_0^T \rho(\tau)\,d\tau$$

the average value of $\rho$ over the interval. If $\rho$ is constant, then $\rho$ is the traditional eigenvalue and $\bar{\rho} = \rho$. In the case of a nonconstant $\rho(t)$, we still have $v(T) = e^{\bar{\rho}T} v(0)$ so that $\bar{\rho}$ is the exponential rate that characterizes the average behavior of $v$ over the interval $T$, although the behavior is not in general exponential, that is, $v(t) = e^{\bar{\rho}t} v(0)$ does not hold in general for all $t \in (0, T)$.

Now consider the general $m$-dimensional case. To take the viewpoint of the preceding paragraph, the notion of the normalized instantaneous rate must be generalized. We write $\boldsymbol{v}(t) = \|\boldsymbol{v}(t)\|_G \boldsymbol{e}(t)$, where $\boldsymbol{e}(t) = \boldsymbol{v}(t)/\|\boldsymbol{v}(t)\|_G$ is the unit vector in the direction of $\boldsymbol{v}(t)$. Because $\boldsymbol{v}(t)$ satisfies the equation $\dot{\boldsymbol{v}} = F[\phi(t, \boldsymbol{x})]\boldsymbol{v}$, it follows that the differential equation satisfied by $\boldsymbol{e}(t)$ is, replacing $F[\phi(t, \boldsymbol{x})]$ by $F(t)$ to simplify the notation,

$$\dot{\boldsymbol{e}} = [F(t) - \rho(t)I]\boldsymbol{e} \tag{19}$$

where

$$\rho(t) = \frac{1}{\|\boldsymbol{v}(t)\|}\frac{d}{dt}\|\boldsymbol{v}(t)\| = \frac{1}{2}\boldsymbol{e}^T[F^T(t)G + \dot{G} + GF(t)]\boldsymbol{e} \tag{20}$$

is the normalized instantaneous rate of change of the magnitude of $\boldsymbol{v}$. The quadratic form for $\rho(t)$ is derived by differentiating $\|\boldsymbol{v}(t)\|_G = [\boldsymbol{v}^T(t)G\boldsymbol{v}(t)]^{1/2}$. For the Euclidean metric we have $\rho(t) = 0.5\boldsymbol{e}^T[F^T(t) + F(t)]\boldsymbol{e}$. When $\rho(t)I$ is subtracted from $F(t)$ in Eq. (19), $\boldsymbol{e}(t)$ tracks the direction of $\boldsymbol{v}$ while maintaining unit length. The function $\rho(t)$ depends on the vector $\boldsymbol{v}(t)$, or equivalently the vector $\boldsymbol{e}(t)$, under consideration, although our notation does not indicate this dependence.

Integrating Eq. (19) along a trajectory of the nonlinear system and simultaneously computing $\rho$ using the quadratic form in Eq. (20) allows one to monitor the evolution of the direction, and the local exponential rate of change of the magnitude, of any initial vector $\boldsymbol{v} \in T_x\mathcal{X}$. For any set of linearly independent initial vectors, $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$, which are subsequently normalized to unit length initial vectors, Eqs. (19) and (20) generate a time-varying basis $\boldsymbol{e}_1(t), \ldots, \boldsymbol{e}_m(t)$ and corresponding instantaneous rates $\rho_1(t), \ldots, \rho_m(t)$. Let $D(t)$ denote the diagonal matrix with $\rho_1(t), \ldots, \rho_m(t)$ as diagonal elements, and let $E(t)$ denote the matrix with columns $\boldsymbol{e}_1(t), \ldots, \boldsymbol{e}_m(t)$.

From the results of the preceding subsection, we know that for the timescale information $(\Sigma, L, N)$, the initial condition $E(0) = L$ will evolve to $E(T) = N$. Observe that Eq. (20) implies $\partial/\partial t\|\boldsymbol{v}(t)\|_G = \rho(t)\|\boldsymbol{v}(t)\|_G$; it makes sense to define the average exponent, as we did in the scalar case,

$$\bar{\rho} = \frac{1}{T}\int_0^T \rho(t)\,dt \tag{21}$$

where $T$ is the time interval over which the average is computed. For a given basis set represented by $E(t)$, we use $\bar{D}$ to denote the corresponding diagonal matrix of average exponents. This allows us to write $\Phi(T, \boldsymbol{x})E(0) = E(T)\exp(\bar{D}\cdot T)$. Comparison with Eq. (8), $\Phi(T, \boldsymbol{x})L = N\Sigma$, proves the following proposition showing the equality of the average instantaneous rates and the finite time Lyapunov exponents when $E(0) = L$.

*Proposition 3:* The initialization $E(0) = L$ generates a normalized solution basis $E(t)$ such that $E(T) = N$ and the corresponding $\exp(\bar{D}\cdot T) = \Sigma$. We have

$$\bar{\rho}_i = \frac{1}{T}\int_0^T \rho_i(t)\,dt = \mu_i = \frac{1}{T}\ln\sigma_i \tag{22}$$

for $i = 1, \ldots, m$.

*Definition 3:* The instantaneous rates $\rho_1(t), \ldots, \rho_m(t)$ when Eq. (19) is initialized by $E(0) = L$ are the kinematic eigenvalues[30] for the linear system $\dot{\boldsymbol{v}} = F(t)\boldsymbol{v}$.

We capture timescale information about the characteristic rates at which the magnitudes of vectors change in the kinematic eigenvalues $\rho_1(t), \ldots, \rho_m(t)$ and their averages $\bar{\rho}_1, \ldots, \bar{\rho}_m$, the latter of which are the finite time Lyapunov exponents. Note that we do not have a means of directly computing the kinematic eigenvalues; we must first calculate $L$ at an initial point $\boldsymbol{x}$ from the transition matrix and use it to initialize Eq. (19). Then integrating Eq. (19), we can compute the kinematic eigenvalues and $E$ at subsequent points $\phi(t, \boldsymbol{x})$. At the initial point $\boldsymbol{x}$, the column vectors of $L$ define the associated spectral filtration in $T_x\mathcal{X}$. At subsequent points $\phi(t, \boldsymbol{x})$ along the trajectory, the associated spectral filtrations can be represented using $E(t)$; however, $E(t)$ does not remain orthogonal, but it is again at $t = T$ where $E(T) = N$. Alternatively, one can derive a differential equation for propagating $L$ along the trajectory such that the orthogonality is preserved.[14,20,23]

There is also timescale information regarding the rates at which the directions of the column vectors of either $E$ or $L$ change along the trajectory. The method investigated in this paper does not extract this timescale information. Note that when $G = I$ it would be equivalent to write Eq. (20) as $\rho(t) = \boldsymbol{e}^T F\boldsymbol{e}$; however, writing it as we have clarifies that the skew-symmetric part of $F$, namely, $\frac{1}{2}(F - F^T)$, does not contribute to $\rho(t)$. For example, the symmetric part of the constant matrix

$$F = \begin{bmatrix} 0 & \omega \\ -\omega & 0 \end{bmatrix} \tag{23}$$

is the zero matrix and both kinematic eigenvalues for $\dot{\boldsymbol{v}} = F\boldsymbol{v}$ will be zero, despite the fact that vectors evolving according to $\dot{\boldsymbol{v}} = F\boldsymbol{v}$ will rotate with frequency $\omega$. Wiesel[34] has investigated a means of adding an imaginary part to the Lyapunov exponents to capture this additional timescale information in analogy with the LTI case. Our focus on the magnitude rates, rather than the rotational rates, is consistent with our interest in multiple timescale systems of the boundary-layer type.

## Timescale Information in the Infinite Time Limit: Convergence and Metric Independence

In general the timescale information depends on the choice of metric and coordinates. Given a particular choice, we can ensure that the Lyapunov exponents are invariant under a coordinate transformation and that the Lyapunov vectors transform to mutually orthogonal vectors by transforming the metric appropriately as in proposition 2a. In this manner, consistency in the timescale information can be achieved; however, the initial choice of metric and coordinates is arbitrary. Thus the timescales can be manipulated as noted by Greene and Kim,[32] and we must conclude that timescales are in general not inherent features of a nonlinear dynamic system.

The typical starting point for timescale analysis is a differential equation model, expressed in a particular set of coordinates. Unless there is a compelling reason to employ a non-Euclidean metric, one would use the Euclidean metric. The timescale information characterizes the system behavior in terms of exponential rates of expansion or contraction of tangent vectors along orbits. The Euclidean metric is constant along orbits, so that any exponential behavior is due to the dynamics. A striking illustration of manipulating timescales is that there exists[35] a coordinate transformation $\boldsymbol{z} = \boldsymbol{h}(\boldsymbol{x})$ such that $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x})$ becomes $\dot{\boldsymbol{z}} = \boldsymbol{c}$, where $\boldsymbol{c}$ is a constant vector, locally in the neighborhood of any point $\boldsymbol{x}$ that is not an equilibrium point. Any timescale behavior that existed in the $\boldsymbol{x}$ representation has been absorbed into the coordinate transformation. On the other hand, there exists a class of metrics and coordinates for which the Lyapunov exponents and the induced filtrations are invariant for sufficiently long averaging time; this class is broad enough to make the timescale information useful. Simply stated, the variations across this class only cause subexponential variations in the timescale behavior and such variations diminish with increasing averaging time. In this section we provide the details for understanding this feature and also address the issue of convergence for both Lyapunov exponents and vectors. This requires consideration of the limit of infinite averaging time. To consider the limiting case of the averaging time going to infinity, we assume, just for this section, that $\mathcal{X}$ is invariant with respect to the flow, that is, for any $x \in \mathcal{X}$, $\phi(t, x) \in \mathcal{X}$ for all $t \in (-\infty, \infty)$.

*Definition 4:* For any $(x, v) \in T\mathcal{X}$, the infinite time upper and lower forward Lyapunov exponents are defined by

$$\bar{\mu}(x, v) = \overline{\lim}_{T \to \infty} \frac{1}{T} \ln\left(\frac{\|\Phi(T, x)v\|}{\|v\|}\right) \quad (24)$$

$$\underline{\mu}(x, v) = \underline{\lim}_{T \to \infty} \frac{1}{T} \ln\left(\frac{\|\Phi(T, x)v\|}{\|v\|}\right) \quad (25)$$

where $\|\cdot\|$ is the metric induced norm and $\overline{\lim}$ and $\underline{\lim}$ are the limit supremum and infimum, respectively. The infinite time upper and lower backward Lyapunov exponents $\bar{\mu}^-(x, v)$ and $\underline{\mu}^-(x, v)$ are defined the same way except with $\Phi(-T, x)$ replacing $\Phi(T, x)$.

Under the general assumption of this paper that the vector field $f(x)$ is smooth and the assumption of this section that $\mathcal{X}$ is invariant in addition to being compact, it follows that $F$ is bounded from above and below on any trajectory $\phi(t, x)$ with $x \in \mathcal{X}$ for all $t$. Consequently, the finite time exponent for any nonzero vector $v \in T_x\mathcal{X}$ will be bounded for all time, and the limits defined in Eqs. (24) and (25) will exist and be finite.[15,36] A simple illustration of the need for the limit supremum is $\overline{\lim}_{T \to \infty}(1/T) \ln e^{T \sin T} = \overline{\lim}_{T \to \infty} \sin T = 1$; whereas the limit does not exist, the $\overline{\lim}$ does. The $\underline{\lim}$ for this example is $-1$.

The relationship between an infinite time upper Lyapunov exponent $\bar{\mu}_i(x)$ and the corresponding finite time Lyapunov exponent $\mu_i(T, x)$ is $\bar{\mu}_i(x) = \overline{\lim}_{T \to \infty} \mu_i(T, x)$. In the infinite time limit, the exponents depend only on $x$. Whereas in the finite time setting we could only say in proposition 1 that, for all $v \in L_i \backslash L_{i-1}$, $\mu(T, x, v)$ is in the interval $\mu_{i-1} < \mu(T, x, v) \le \mu_i$; in the infinite time setting we can say that, for all $v \in L_i \backslash L_{i-1}$, $\bar{\mu}(x, v) = \bar{\mu}_i(x)$. This is because in the infinite time limit the component of $v$ corresponding to the largest exponent will dominate as illustrated later in this section.

Let $L(T, x)$ denote the matrix whose columns are the Lyapunov vectors at $x$ obtained with an averaging time $T$, ordered as we have been assuming. Define $L(x) = \lim_{T \to \infty} L(T, x)$. We can also consider the limiting behavior of $L^-$ in the SVD of $\Phi(-T, x) = N^- \Sigma^- (L^-)^T$ for backward time propagation. The column vectors of $L^-$ are interpreted as vectors in $T_x\mathcal{X}$. Define $L^-(x) = \lim_{T \to \infty} L^-(-T, x)$. Assume that the limits for $L$ and $L^-$ exist for now; the convergence of the Lyapunov vectors will be addressed later. At each $x \in \mathcal{X}$, the set $\{l_i(x), i = 1, \ldots, m\}$ is an orthonormal basis for $T_x\mathcal{X}$, as is the set $\{l_i^-(x), i = 1, \ldots, m\}$; this latter set is ordered such that $\mu_1^- > \cdots > \mu_m^-$. We can define filtrations of $T_x\mathcal{X}$ using the two bases, namely, $L_1 \subset L_2 \subset \cdots \subset L_m$ and $L_1^- \supset L_2^- \supset \cdots \supset L_m^-$, where $L_i^- = \text{span}\{l_i^-, l_{i+1}^-, \ldots, l_m^-\}$.

*Definition 5:* When the forward (respectively, backward) infinite time exponents have the property $\bar{\mu}_i(x) = \underline{\mu}_i(x)$ [respectively, $\bar{\mu}_i^-(x) = \underline{\mu}_i^-(x)$], for $i = 1, \ldots, m$, the system is said to be forward regular (respectively, backward regular) at $x$. The system is Lyapunov regular at $x$ if 1) it is forward and backward regular at $x$, 2) $\bar{\mu}_i(x) = -\bar{\mu}_i^-(x)$, $i = 1, \ldots, m$, 3) the forward and backward filtrations have the same dimension, 4) there exists a decomposition $T_x\mathcal{X} = \mathcal{E}_1(x) \oplus \cdots \oplus \mathcal{E}_m(x)$ into invariant subbundles such that $L_i(x) = \mathcal{E}_1(x) \oplus \cdots \oplus \mathcal{E}_i(x)$ and $L_i^-(x) = \mathcal{E}_i(x) \oplus \cdots \oplus \mathcal{E}_m(x)$, $i = 1, \ldots, m$, and 5) for any $v \in \mathcal{E}_i(x) \backslash \{0\}$, the $\lim_{t \to \pm\infty}(1/t) \ln \|\Phi(t, x)v\| = \bar{\mu}_i(x)$, where $\oplus$ denotes direct sum. We note that Barreira and Pesin (Ref. 12, p. 26) define forward regular using $\bar{\mu}_i = -\bar{\mu}_i^*$ where $\bar{\mu}_i^*$, $i = 1, \ldots, m$, are the Lyapunov exponents for the associated adjoint system. They say the Lyapunov exponents are exact when $\bar{\mu}_i(x) = \underline{\mu}_i(x)$, for $i = 1, \ldots, m$. However, when the linear system matrix $F$ is bounded, as assumed here, forward regularity and exactness are equivalent properties (Ref. 12, p. 26). We also note that, under our assumption that we always have $m$ distinct exponents, condition 3 is automatically satisfied.

Given the forward and backward filtrations for a Lyapunov regular system, we can construct the invariant subbundles by $\mathcal{E}_i(x) = L_i(x) \bigcap L_i^-(x)$, $i = 1, \ldots, m$. These subbundles are the generalizations of the eigenspaces for LTI systems. By invariant, we mean that given a solution $[\phi(t, x), \Phi(t, x)v]$ with $v \in \mathcal{E}_i(x)$, we have $\Phi(t, x)v \in \mathcal{E}_i[\phi(t, x)]$ for all $t$. All vectors in $\mathcal{E}_i(x)$ are characterized by the Lyapunov exponent $\bar{\mu}_i(x)$. In general the Lyapunov vectors are metric dependent because they are required to be mutually orthogonal with respect to a specific metric. On the other hand, because the filtrations are defined by the Lyapunov exponents, it follows that if the exponents are invariant for a given class of metrics, the filtrations and their intersections will also be invariant for that class.

For a general nonlinear system the infinite time Lyapunov exponents at an equilibrium point $x_{eq}$ are[10] the real parts of the eigenvalues of the Jacobian matrix $F(x_{eq})$, the system is Lyapunov regular on $\mathcal{X} = \{x_{eq}\}$, and the invariant subbundles (subspaces in this case) are the (real) eigenspaces. For the case of $m$ distinct exponents, the limiting values of the finite time Lyapunov vectors $l_i$, $i = 1, \ldots, m$, are related to the eigenvectors of $F(x_{eq})$ in the following manner. Let $v_i$, $i = 1, \ldots, m$, be the normalized eigenvectors, assumed for simplicity of discussion to be real and ordered such that the corresponding eigenvalues go from smallest to largest. Then $l_1 = v_1$, $l_2$ is orthogonal to $l_1$ and lies in the subspace spanned by $v_1$ and $v_2$, etc. In other words, the $l_i$ vectors are obtained by applying Gram–Schmidt orthogonalization to the eigenvectors, starting with $v_1$ and moving successively upward in index. Similarly the $l_i^-$ vectors (and the $n_i$ vectors) are obtained by applying Gram–Schmidt orthogonalization to the eigenvectors, starting with $v_m$ and moving successively downward in index. Because the metric defines orthogonality, different metrics will produce different sets of $l_i$ and $l_i^-$ vectors except for the common starting vectors $l_1 = v_1$ and $l_m^- = v_m$. However, if the Lyapunov exponents are metric invariant, as they are for $\mathcal{X} = \{x_{eq}\}$, then the filtration is metric invariant, even though the orthonormal basis used to represent the filtration is metric dependent. If we intersect the forward and backward filtrations, we will get the eigenspaces, independent of the metric used. Thus, the Lyapunov exponents and vectors are directly related to the eigenvalues and eigenvectors of the Jacobian matrix in the special case of analyzing the timescale structure associated with an equilibrium point; of course this should be the case because the eigenvalues and eigenvectors correctly characterize the time-scale structure at an equilibrium point.

For the case in which $\mathcal{X}$ is a periodic orbit, the Lyapunov exponents are[10] the real parts of the Floquet exponents. For the case of distinct real exponents, at each point on the periodic orbit, the Lyapunov vectors are the orthogonal counterparts of the eigenvectors of the monodromy matrix for that point. The Lyapunov vectors rotate along the periodic orbit in such a way that they coincide with their initial directions one period later.

*Definition 6:* For an $m$-dimensional dynamic system on a compact set $\mathcal{X}$ with finite time Lyapunov exponents $\mu_1 < \mu_2 < \cdots < \mu_m$, neighboring exponents $\mu_j(T, x)$ and $\mu_{j+1}(T, x)$ are uniformly distinct at $x$, if there exists a $T$ independent spectral gap $\Delta\mu_i(x)$ and an averaging time $T_1(x) < \bar{t}(x)$ such that $\mu_{j+1}(T, x) - \mu_j(T, x) > \Delta\mu_j(x)$ for all $T \in [T_1(x), \bar{t}(x)]$. Note in this definition that $\bar{t}(x)$ can be finite. $T_1(x)$ is introduced to eliminate any initial transient period that is not representative of the subsequent behavior. In the case of finite $\bar{t}(x)$, now $T_1(x)$ should be a small fraction of $\bar{t}(x)$. Figure 3 provides an example. The definition is adapted from Goldhirsch et al.[15]

Goldhirsch et al.[15] developed evolution equations for $\mu_i(T, x)$ and the vectors $\{l_i(T, x), i = 1, \ldots, m\}$ for the Euclidean metric case that clarified the convergence behavior of this timescale information. Their results are extended for a general metric by the following lemma; the proof of which follows theirs.

*Lemma:* Over intervals of the averaging time $T$ during which the finite time Lyapunov exponents are distinct, the finite time Lyapunov exponents and vectors at $x \in \mathcal{X}$ evolve with the averaging time $T$ according to the differential equations

$$\frac{\partial}{\partial T} \exp[2\mu_i(T, x)T] = \exp[2\mu_i(T, x)T] l_i^T Q^T \left[F^T(T) G_T + \dot{G}_T + G_T F(T)\right] Q l_i \quad (26)$$

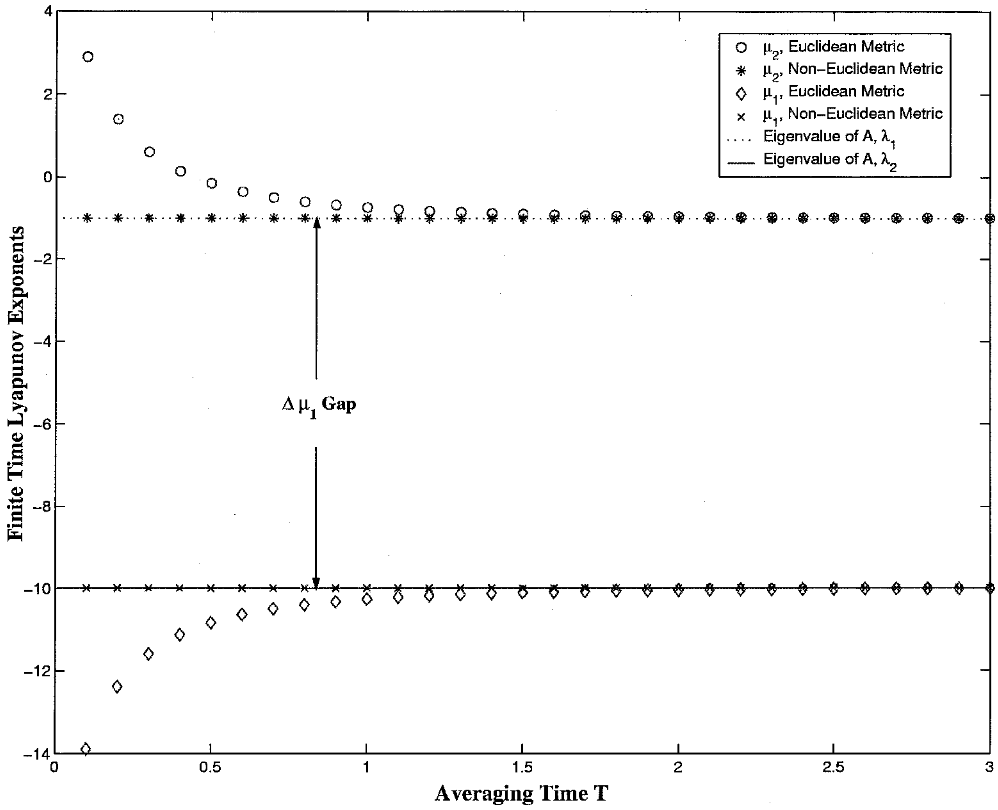


**Fig. 3 Finite time Lyapunov exponents as functions of the averaging time *T* for Euclidean and non-Euclidean metrics.**

$$\frac{\partial}{\partial T}\boldsymbol{l}_i = \sum_{k=1}^{i-1} \times \frac{\boldsymbol{l}_k^T Q^T\left[F^T(T)G_T + \dot{G}_T + G_T F(T)\right]Q\boldsymbol{l}_i}{\exp\{[\mu_i(T,\boldsymbol{x}) - \mu_k(T,\boldsymbol{x})]T\} - \exp\{[\mu_k(T,\boldsymbol{x}) - \mu_i(T,\boldsymbol{x})]T\}}\boldsymbol{l}_k + c_{ii}\boldsymbol{l}_i + \sum_{k=i+1}^{m} \times \frac{\boldsymbol{l}_k^T Q^T[F^T(T)G_T + \dot{G}_T + G_T F(T)]Q\boldsymbol{l}_i}{\exp\{[\mu_i(T,\boldsymbol{x}) - \mu_k(T,\boldsymbol{x})]T\} - \exp\{[\mu_k(T,\boldsymbol{x}) - \mu_i(T,\boldsymbol{x})]T\}}\boldsymbol{l}_k \tag{27}$$

where $F = F[\phi(T,\boldsymbol{x})]$ and $Q$ is the rotation matrix in the generalized polar decomposition $\Phi = QG_0^{-1}S$ defined by $Q = NL^TG_0$ and $S = G_0L\Sigma L^TG_0$ with both $Q$ and $S$ depending on $\boldsymbol{x}$ and $T$. The coefficient $c_{ii}$, not specified here, is determined such that $\boldsymbol{l}_i$ evolves continuously with unit length.

*Proof:* Express $\boldsymbol{l}_i(T+\Delta T,\boldsymbol{x})$ to first-order in $\Delta T$ as

$$\boldsymbol{l}_i(T+\Delta T,\boldsymbol{x}) = \boldsymbol{l}_i(T,\boldsymbol{x}) + \Delta T\,\Sigma_{j=1}^{m} c_{ij}\boldsymbol{l}_j(T,\boldsymbol{x}) \tag{28}$$

Substitute this and the first-order approximations

$$\Phi(T+\Delta T,\boldsymbol{x}) = \Phi(T,\boldsymbol{x}) + \Delta T F(T)\Phi(T,\boldsymbol{x})$$

$$G[\phi(T+\Delta T,\boldsymbol{x})] = G[\phi(T,\boldsymbol{x})] + \Delta T\dot{G}[\phi(T,\boldsymbol{x})]$$

$$\exp[2(T+\Delta T)\mu_i(T+\Delta T,\boldsymbol{x})] = \exp[2T\mu_i(T,\boldsymbol{x})] + \Delta T\frac{\mathrm{d}}{\mathrm{d}T}\exp[2T\mu_i(T,\boldsymbol{x})]$$

into the equation

$$[\Phi(T+\Delta T,\boldsymbol{x})]^T G[\phi(T+\Delta T,\boldsymbol{x})]\Phi(T+\Delta T,\boldsymbol{x})\boldsymbol{l}_i(T+\Delta T,\boldsymbol{x}) = \exp[2(T+\Delta T)\mu_i(T+\Delta T,\boldsymbol{x})]G(\boldsymbol{x})\boldsymbol{l}_i(T+\Delta T,\boldsymbol{x})$$

Premultiply both sides of the resulting equation by $\boldsymbol{l}_k^T(T,\boldsymbol{x})$. The zeroth-order terms in $\Delta T$ cancel out. The first-order terms lead to the first equation in the lemma when $k = i$ and define the coefficients $c_{ik}$ when $k \neq i$. Substituting these coefficients into Eq. (28) and taking the limit as $\Delta T \to 0$ of $[\boldsymbol{l}(T+\Delta T,\boldsymbol{x}) - l(T,\boldsymbol{x})]/\Delta T$ leads to the second equation in the lemma. □

*Theorem:* If the neighboring Lyapunov exponents $\mu_j(T,\boldsymbol{x})$ and $\mu_{j+1}(T,\boldsymbol{x})$ are uniformly distinct and the metric $G$ is continuously differentiable on $\mathcal{X}$, then as $T$ increases the subspace $L_j(T,\boldsymbol{x}) = \text{span}\{\boldsymbol{l}_1(T,x),\ldots,\boldsymbol{l}_j(T,x)\}$ converges to a fixed subspace $L_j(\boldsymbol{x})$ at least at the rate $\exp[-\Delta\mu_j T]$ where $\Delta\mu_j$ is the spectral gap.

*Proof:* In Eq. (27), the rate of change of $\boldsymbol{l}_i$ is expressed as a linear combination of the Lyapunov vectors at the corresponding time. The continuity of $F^T(\boldsymbol{x})G(\boldsymbol{x}) + \dot{G}(\boldsymbol{x}) + G(\boldsymbol{x})F(\boldsymbol{x})$ on the compact set $\mathcal{X}$ allows us to establish finite upper and lower bounds for the quadratic forms that appear in the coefficients. Thus, one can show that the coefficients of the finite time Lyapunov vectors $\{\boldsymbol{l}_{j+1},\ldots,\boldsymbol{l}_m\}$ for the rates of change of $\boldsymbol{l}_1,\ldots,\boldsymbol{l}_j$ go to zero at least at the rate $\exp(-\Delta\mu_j T)$. Hence, the changes in the vectors $\boldsymbol{l}_1,\ldots,\boldsymbol{l}_j$ become confined to the subspace $L_j(\boldsymbol{x})$ that they span, and the subspace itself converges. □

By applying a similar approach for backward propagation, we could show that a uniform gap between $\mu_j^-(-T,\boldsymbol{x})$ and $\mu_{j+1}^-(-T,\boldsymbol{x})$ leads to the convergence of $L_{j+1}^-$. Although only the Euclidean metric was considered, the paper by Goldhirsch et al.[15] contains the basic ideas in the preceding theorem and its proof, as well as the result that an individual Lyapunov vector $\boldsymbol{l}_i$ will converge if $\mu_i$ is uniformly distinct with respect to both $\mu_{i+1}$ and $\mu_{i-1}$ at the rate proportional to $\max(\{\exp[\mu_i(T,\boldsymbol{x}) - \mu_{i+1}(T,\boldsymbol{x})]T\}; \exp\{[\mu_{i-1}(T,\boldsymbol{x}) - \mu_i(T,\boldsymbol{x})]T\})$. For $i = 1$ and $m$, there is only one neighbor $\mu$; one of the two exponential terms is undefined, and the remaining one is the convergence rate. A case that can arise is that there is one persistent gap in the exponents that leads to the convergence of a subspace, but there are no other gaps, and thus, the Lyapunov vectors in the subspace will not converge. If the finite time Lyapunov exponents $\mu_i(T,\boldsymbol{x})$ and $\mu_{i+1}(T,\boldsymbol{x})$ are not uniformly distinct, then $\underline{\mu}_i(\boldsymbol{x})$, $\bar{\mu}_i(\boldsymbol{x})$, $\underline{\mu}_{i+1}(\boldsymbol{x})$ and $\bar{\mu}_{i+1}(\boldsymbol{x})$ will exist, but the intervals $[\underline{\mu}_i(\boldsymbol{x}), \bar{\mu}_i(\boldsymbol{x})]$ and $[\underline{\mu}_{i+1}(\boldsymbol{x}), \bar{\mu}_{i+1}(\boldsymbol{x})]$ will overlap. The Lyapunov vectors may or may not converge; if they do converge, the convergence will not be exponential.

Consider a two-dimensional nonlinear system $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x})$ with the region of interest $\mathcal{X}$ a compact invariant set. We want to investigate the relationship between the timescale information for the Euclidean metric and the timescale information for a non-Euclidean metric $G(\boldsymbol{x})$. For an initial point $\boldsymbol{x} \in \mathcal{X}$ and a time interval $T$, let $\boldsymbol{l}_1, \boldsymbol{l}_2, \boldsymbol{n}_1, \boldsymbol{n}_2$, $\mu_1$, and $\mu_2$, all functions of $(T, \boldsymbol{x})$, denote the timescale information obtained from an SVD assuming the Euclidean metric. An arbitrary initial vector in $T_x\mathcal{X}$, expressed by $\boldsymbol{v} = c_1\boldsymbol{l}_1 + c_2\boldsymbol{l}_2$, propagates to $\boldsymbol{v}(T) = c_1 e^{\mu_1 T}\boldsymbol{n}_1 + c_2 e^{\mu_2 T}\boldsymbol{n}_2$. The multiplier for the net change in length of this vector, according to the metric $G(\boldsymbol{x})$, from $t = 0$ to $t = T$ is

$$\sigma(T) = \frac{\|\boldsymbol{v}(T)\|_{G_T}}{\|\boldsymbol{v}(0)\|_{G_0}} = \left\{\frac{c_1^2\exp(2\mu_1 T)\boldsymbol{n}_1^T G_T\boldsymbol{n}_1 + 2c_1c_2\exp[(\mu_1+\mu_2)T]\boldsymbol{n}_1^T G_T\boldsymbol{n}_2 + c_2^2\exp(2\mu_2 T)\boldsymbol{n}_2^T G_T\boldsymbol{n}_2}{c_1^2\boldsymbol{l}_1^T G_0\boldsymbol{l}_1 + 2c_1c_2\boldsymbol{l}_1^T G_0\boldsymbol{l}_2 + c_2^2\boldsymbol{l}_2^T G_0\boldsymbol{l}_2}\right\}^{\frac{1}{2}}$$

$$= \exp(\mu_2 T)\left(\frac{c_1^2\exp[2(\mu_1-\mu_2)T]\boldsymbol{n}_1^T G_T\boldsymbol{n}_1 + 2c_1c_2\exp[(\mu_1-\mu_2)T]\boldsymbol{n}_1^T G_T\boldsymbol{n}_2 + c_2^2\boldsymbol{n}_2^T G_T\boldsymbol{n}_2}{c_1^2\boldsymbol{l}_1^T G_0\boldsymbol{l}_1 + 2c_1c_2\boldsymbol{l}_1^T G_0\boldsymbol{l}_2 + c_2^2\boldsymbol{l}_2^T G_0\boldsymbol{l}_2}\right)^{\frac{1}{2}} \tag{29}$$

where $G_0$ and $G_T$ are the values of $G$ at the initial and final points, respectively. Now assume that the finite time Lyapunov exponents $\mu_1$ and $\mu_2$ are uniformly distinct, that is, there exist a gap $\Delta\mu_i(\boldsymbol{x})$ and an averaging time $T_1(\boldsymbol{x})$ such that $\mu_2(T,\boldsymbol{x}) - \mu_1(T,\boldsymbol{x}) > \Delta\mu_1(\boldsymbol{x})$ for all $T > T_1(\boldsymbol{x})$. Computing the upper Lyapunov exponent, we have for $c_2 = 0$, that is, for $\boldsymbol{v} = c_1\boldsymbol{l}_1$ that

$$\bar{\mu}(\boldsymbol{x}) = \overline{\lim}_{T\to\infty}\mu_1(T,\boldsymbol{x}) + \overline{\lim}_{T\to\infty}\frac{1}{T}\ln\left[\frac{\boldsymbol{n}_1^T(T,\boldsymbol{x})G_T\boldsymbol{n}_1(T,\boldsymbol{x})}{\boldsymbol{l}_1^T G_0\boldsymbol{l}_1}\right]^{\frac{1}{2}}$$

$$= \overline{\lim}_{T\to\infty}\mu_1(T,\boldsymbol{x}) \tag{30}$$

whereas for $c_2 \neq 0$, that is, for any vector not in the subspace given by span$\{\boldsymbol{l}_1\}$

$$\bar{\mu}(\boldsymbol{x}) = \overline{\lim}_{T\to\infty}\mu_2(T,\boldsymbol{x})$$

$$+ \overline{\lim}_{T\to\infty}\frac{1}{T}\ln\left(\frac{c_2^2\boldsymbol{n}_2^T(T,\boldsymbol{x})G_T\boldsymbol{n}_2(T,\boldsymbol{x})}{c_1^2\boldsymbol{l}_1^T G_0\boldsymbol{l}_1 + 2c_1c_2\boldsymbol{l}_1^T G_0\boldsymbol{l}_2 + c_2^2\boldsymbol{l}_2^T G_0\boldsymbol{l}_2}\right)^{\frac{1}{2}}$$

$$= \overline{\lim}_{T\to\infty}\mu_2(T,\boldsymbol{x}) \tag{31}$$

The second terms on the right-hand sides of Eqs. (30) and (31) vanish because the expression within the brackets has finite upper and lower bounds on $\mathcal{X}$, given the continuity of $G(\boldsymbol{x})$ and the compactness of $\mathcal{X}$. Thus, the difference between the finite time exponents for the metric $G$ and for the Euclidean metric diminishes to zero with increasing $T$. Analogous results showing the metric independence of $\underline{\mu}_1(\boldsymbol{x})$ and $\underline{\mu}_2(\boldsymbol{x})$ can be obtained. Proving the invariance of the Lyapunov exponents with respect to smooth coordinate transformations can be done in a similar way.[37]

An observation particularly relevant for practical computation is that the metric effect for any $G$ with subexponential growth or decay along the orbits in $\mathcal{X}$ will diminish with increasing $T$. In other words, if the metric-induced changes to the length of tangent vectors, relative to their Euclidean lengths, are subexponential, for example, constant or proportional to $T^n$ for any finite $n$, then the difference in the exponents for the Euclidean metric and $G$ diminishes as $T$ increases. On the other hand, if the metric effect is exponential, then it will alter the exponents. The coordinate dependence of the exponents is very similar. Relative to the exponents for a particular set of coordinates, the difference in the exponents for a new set of coordinates without changing the metric will diminish with $T$, if the coordinate transformation does not alter the linear behavior along orbits at the exponential level.

The upper Lyapunov exponents given by Eqs. (30) and (31) exist that is, the finite time exponents $\mu_1$ and $\mu_2$ will converge in the $\overline{\lim}$ sense, for the reasons given earlier. Denote the exponent given by Eq. (30) by $\bar{\mu}_1(\boldsymbol{x})$ and that given by Eq. (31) by $\bar{\mu}_2(\boldsymbol{x})$; all vectors in $T_x\mathcal{X}$ have one of these two exponents. Also $\underline{\mu}_1(\boldsymbol{x})$ and $\underline{\mu}_2(\boldsymbol{x})$ exist. If the timescale behavior is uniform along the orbit, which would be indicated by constant kinematic eigenvalues, then the upper and lower exponents will be equal. In the absence of uniformity, in general $\bar{\mu}_1(\boldsymbol{x}) \neq \underline{\mu}_1(\boldsymbol{x})$ and $\bar{\mu}_2(\boldsymbol{x}) \neq \underline{\mu}_2(\boldsymbol{x})$, that is, the system will not be forward regular at $\boldsymbol{x}$. However, there will exist a finite time $T_2$ beyond which the uniformly distinct finite time exponents $\mu_1(T,\boldsymbol{x})$ and $\mu_2(T,\boldsymbol{x})$ lie in disjoint intervals, namely, for $T > T_2$, $\mu_1(T,\boldsymbol{x}) \in [\underline{\mu}_1(\boldsymbol{x}), \bar{\mu}_1(\boldsymbol{x})]$, and $\mu_2(T,\boldsymbol{x}) \in [\underline{\mu}_2(\boldsymbol{x}), \bar{\mu}_2(\boldsymbol{x})]$. For the case of nonuniform timescale behavior, the theory guaranteeing the convergence of the Lyapunov exponents in the limit (not just limsup) requires[12,31] ergodic or recurrent behavior, a periodic orbit being a special case of recurrent behavior.

The results for a compact and invariant set $\mathcal{X}$ that we have just obtained for the general two-dimensional case can be extended to $m$ dimensions.

## Timescale Analysis Procedure

With the theory and understanding of the timescale information established, we can now describe how one would compute the timescale structure of a nonlinear dynamic system. For developing theory requiring limits, it is important for $\mathcal{X}$ to be invariant to ensure the existence of orbits on an infinite time interval. When the interest is in numerically computing timescale information, the critical property of $\mathcal{X}$ is that there is useful timescale information that converges to sufficient accuracy over the available time intervals. In general, it will not be feasible, but nor is it necessary, to resolve the complete timescale structure. What is important is to resolve the timescales that are disparate relative to the region $\mathcal{X}$. If there are two or more sufficiently disparate timescales, then the associated timescale structure can be resolved.

We say the behavior is uniform on $\mathcal{X}$, if the spectrum $Sp(T,\boldsymbol{x})$ is approximately uniform in $\boldsymbol{x}$ and $T$ for $T > T_1$, where $T_1$ is as defined in definition 6. We look for gaps that split the spectrum into smaller subsets. For example, if there are $m_{fc}$ large negative exponents, $m_s$ near-zero exponents, and $m_{fe}$ large positive exponents, with $m_{fc} + m_s + m_{fe} = m$, we would have a splitting of the form $Sp(T,\boldsymbol{x}) = Sp^{fc}(T,\boldsymbol{x}) \cup Sp^s(T,\boldsymbol{x}) \cup Sp^{fe}(T,\boldsymbol{x})$ where $Sp^{fc}$, $Sp^s$, and $Sp^{fe}$ are the subsets corresponding to the fast-contracting, slow, and fast-expanding behaviors, respectively. For the timescale structure induced by the splitting to be computable the gaps in the spectrum must be sufficiently large relative to the characteristic time interval for $\mathcal{X}$ (to be defined). If these conditions are satisfied, then we can resolve the tangent space structure $T_x\mathcal{X} = \mathcal{E}^{fc}(\boldsymbol{x}) \oplus \mathcal{E}^s(\boldsymbol{x}) \oplus \mathcal{E}^{fe}(\boldsymbol{x})$, referred to as a splitting of the tangent space. A special case of this splitting is $T_x\mathcal{X} = \mathcal{E}^{fc}(\boldsymbol{x}) \oplus \mathcal{E}^s(\boldsymbol{x})$, as in our motivating example, where there is only one gap and there is no fast expanding behavior.

The steps in computing and interpreting timescale information are the following:

1) Choose the compact region $\mathcal{X}$ of the state space to be analyzed and define a Riemannian metric on $\mathcal{X}$. In the absence of a compelling reason to use a non-Euclidean metric, use the Euclidean metric.

2) Compute, for a range of averaging times $T$, the timescale information along trajectories that collectively sample the region $\mathcal{X}$. Compute the corresponding kinematic eigenvalues to determine the

degree of uniformity in the timescale structure. If the timescale structure is not uniform, one or more subsets of $\mathcal{X}$ may have uniform structure, and it may be of interest to analyze each separately.

3) Identify any persistent, sufficiently large spectral gaps. Determine a characteristic time interval $\bar{t}_c$ for $\mathcal{X}$, an average maximum $T$ for well-chosen starting points. For example, if $\mathcal{X}$ is a transient region for which orbits enter on one side and exit on the other side, then $\bar{t}_c$ would be the characteristic time to traverse $\mathcal{X}$. Determine if there are any consistent gaps in the spectrum indicating timescale separation in the dynamics on $\mathcal{X}$. If there is a spectral gap $\Delta\mu_j$ and $\bar{t}_c$ is at least several times as long as $(\Delta\mu_j)^{-1}$, such as shown in Fig. 3, then the subspaces $L_j(T, \boldsymbol{x})$ and $L^-_{j+1}(-T, \boldsymbol{x})$ will converge. If $\mathcal{X}$ is invariant, then a persistent gap of any size will ensure convergence. If there are additional sufficiently large gaps, then additional timescale structure can be resolved.

4) Resolve the timescale structure. For subspaces $L_j(\boldsymbol{x})$ and $L^-_{j+1}(\boldsymbol{x})$ that converge, $L_j(\boldsymbol{x})$ will have minimum and maximum Lyapunov exponents for averaging times in the interval $[T_1, \bar{t}(\boldsymbol{x})]$. To consistently deal with forward time, we will talk about $N_{j+1}(\boldsymbol{x})$, which is the same subspace as $L^-_{j+1}(\boldsymbol{x})$. $N_{j+1}(\boldsymbol{x})$ will have minimum and maximum Lyapunov exponents for the same interval of averaging times. These minimum and maximum exponents bound the average exponential rates at which vectors in the two subspaces grow or decay. The difference between the maximum exponent for $L_j(\boldsymbol{x})$ and the minimum exponent for $N_{j+1}(\boldsymbol{x})$, that is, $\Delta\mu_j(\boldsymbol{x})$, is a lower bound on the gap between the exponents for the two subspaces. For a subspace that converges, the associated exponents will most generally continue to vary as $T$ is increased. If any neighboring exponents associated with the subspace remain distinct and a sufficiently large, persistent gap exists, then a splitting of this subspace can be resolved.

For the purpose of model decomposition, it is the geometric structure induced by relatively large gaps in the Lyapunov exponents that is of interest. Fortunately, it is exactly this structure that is resolvable. The exponents themselves need not converge for this structure to be computed. Although we have assumed that the finite time Lyapunov exponents are distinct to simplify the presentation, there could be values of $T$ for which there are fewer than $m$ distinct exponents. The exponents still define a filtration at such values of $T$; however, the number of elements of the filtration will be reduced to the number of distinct exponents. A persistent, sufficiently large gap in the exponents will still lead to subspace convergence; however, in the general case the evolution equations (26) and (27) cannot be used to prove this due to the singularities associated with repeated exponents.

The next step for model decomposition would be to use the subspace structure to determine appropriate coordinates. This involves integrating vector fields or more generally distributions. This step and the theory supporting it are beyond the scope of this paper.

## Numerical Example

We return to the motivating example described at the beginning of the paper. Recall that we have two coordinate representations of a dynamic system. In the $(w_1, w_2)$ representation, the dynamics are linear, and the timescale properties are easily deduced from the eigenvalues and eigenvectors of the system matrix $A$ in Eq. (1). In the $(x_1, x_2)$ representation the dynamics are nonlinear. Provided that the state transformation and/or the metric we use do not significantly alter the timescales, which is the case here, the correct result is that the computed Lyapunov exponents and vectors should converge to the eigenvalues and transformed, orthogonalized eigenvectors of $A$. Because the matrix for transforming tangent vectors is state dependent, the transformed, orthogonalized eigenvectors are state dependent. We consider an orbit segment in the transient region $\mathcal{X}$ shown in Fig. 1b (the orbit with the times indicated along it) to determine if the convergence rate is fast enough to allow the correct timescale information to be computed. There is a strategy for this example that we want to avoid using. The nonlinear system has a globally attracting equilibrium point at the origin, and the timescale information is the same at the equilibrium point as it is in $\mathcal{X}$ due to the underlying linearity. With a long enough averaging time, the orbit will spend most of the time in the vicinity of the origin and the exponents will be dictated by the timescales near the origin, that is, by the eigenvalues of the Jacobian matrix at the origin, which has the same eigenvalues as the $A$ matrix of the linear system. We would like to determine the correct timescale information from the orbit segment in $\mathcal{X}$ because $\mathcal{X}$ is the region for which we are interested in knowing the timescale structure. A general method cannot rely on extracting timescale information from the vicinity of an equilibrium point because in general there will not be an equilibrium point that is attracting and has the correct timescale information for the region $\mathcal{X}$.

Assuming the Euclidean metric in the $\boldsymbol{w}$ system, $G^w = I$, an appropriate metric in the $\boldsymbol{x}$ system, if the length of a tangent vector is to be invariant under the transformation, is

$$G^x(\boldsymbol{x}) = H^T(\boldsymbol{x})H(\boldsymbol{x}) \tag{32}$$

where

$$H(\boldsymbol{x}) = \frac{\partial \boldsymbol{h}(\boldsymbol{x})}{\partial \boldsymbol{x}} = \begin{bmatrix} 1 - 2a(x_1 + x_2) & -2a(x_1 + x_2) \\ 2a(x_1 + x_2) & 1 + 2a(x_1 + x_2) \end{bmatrix}$$

The linearized dynamics for the $\boldsymbol{x}$ representation are $\dot{\boldsymbol{v}} = F(x_1(t), x_2(t))\boldsymbol{v}$, where

$$F = \begin{bmatrix} -1 - 2a(x_1 + 10x_2) - \eta(x_1, x_2) & -2a(10x_1 + 19x_2) - \eta(x_1, x_2) \\ 2a(-8x_1 + x_2) + \eta(x_1, x_2) & -10 + 2a(x_1 + 10x_2) + \eta(x_1, x_2) \end{bmatrix} \tag{33}$$

and $\eta(x_1, x_2) = 54a^2(x_1 + x_2)^2$.

For this simple low-order system, we employ a straightforward SVD-based method to determine the Lyapunov exponents and vectors. (Dieci et al.[30] provide a survey of methods of computing Lyapunov exponents and links to the vast literature on these methods.) The computations are done using MATLAB®. The nonlinear and linearized systems, Eqs. (3) and (4), are integrated simultaneously using Gear's method (ODE15s). For the Euclidean metric, the finite time Lyapunov exponents and vectors for the interval $T$ are computed by applying the SVD function to $\Phi(T, \boldsymbol{x})$. For a general metric $G$, we compute the timescale information as stated in proposition 2b.

We determine the timescale structure for the orbit of the dynamic system [Eq. (2)] with $a = 0.01$, associated with the initial point $(x_1, x_2) = (40, -39)$, as shown in Fig. 1b. The finite time Lyapunov exponents at $\boldsymbol{x} = (40, -39)$ for averaging times up to three, at which time the orbit leaves $\mathcal{X}$, are shown in Fig. 3. There is a minimum gap $\Delta\mu_1$ of 9; thus, convergence of the Lyapunov vectors should occur for $T > 0.3$, depending on the convergence tolerance. Because it takes three time units to traverse the region, it is feasible to average long enough for convergence. Because the system dimension $m = 2$, convergence of the subspace $L_1$ at our initial point implies convergence of $\boldsymbol{l}_1$ at that point. Our numerical results confirm that $\boldsymbol{l}_1$ at $\boldsymbol{x} = (40, -39)$ converges to a fixed direction. For the Euclidean metric, convergence requires about $T = 0.3$, which is $3\Delta\mu_1^{-1}$, whereas for the metric $G$, the convergence is much faster. The limiting $\boldsymbol{l}_1$ vector is the same for both metrics. From our knowledge of the equivalent linear system in this contrived example, we independently determine $L^x = H^{-1}(\boldsymbol{x})L^w$, where $L^w$ is constructed from orthogonalizing the eigenvectors of the LTI system, and find agreement to the level of precision we are using. Let $\boldsymbol{v}^w_1$ and $\boldsymbol{v}^w_2$ denote the eigenvectors of $A$ corresponding to the eigenvalues $\lambda_1 = -10$ and $\lambda_2 = -1$, respectively. We find that $\boldsymbol{l}^x_1$, the first column of $L^x$, points in the same direction as the transformed eigenvector of $A$ corresponding to $\lambda_1$, that is, $\boldsymbol{l}^x_1 = H^{-1}(\boldsymbol{x})\boldsymbol{v}^w_1$. Similarly, examining $N^x$ for $T = 1$, as obtained from the SVD, we find that $\boldsymbol{n}^x_2$ is the transformed eigenvector of $A$ corresponding to $\lambda_2$, that is, $\boldsymbol{n}^x_2 = H^{-1}[\phi(1, \boldsymbol{x})]\boldsymbol{v}^w_2$.

The Lyapunov vectors $\boldsymbol{l}^x_1$ and $\boldsymbol{n}^x_2$ [or equivalently $(\boldsymbol{l}^-_2)^x$] denoted by the longer arrows [one arrowhead for slow $(\boldsymbol{n}^x_2)$, two for fast $(\boldsymbol{l}^x_1)$] are shown in Fig. 1b at several points along the orbit under consideration. These vectors correctly represent the slow and fast directions, in that they are tangent to the slow and fast coordinate curves (denoted by dashed and dash–dotted lines, respectively). In other words, the fast/slow splitting $T_x\mathcal{X} = \mathcal{E}^{fc}(\boldsymbol{x}) \oplus \mathcal{E}^s(\boldsymbol{x})$ is given

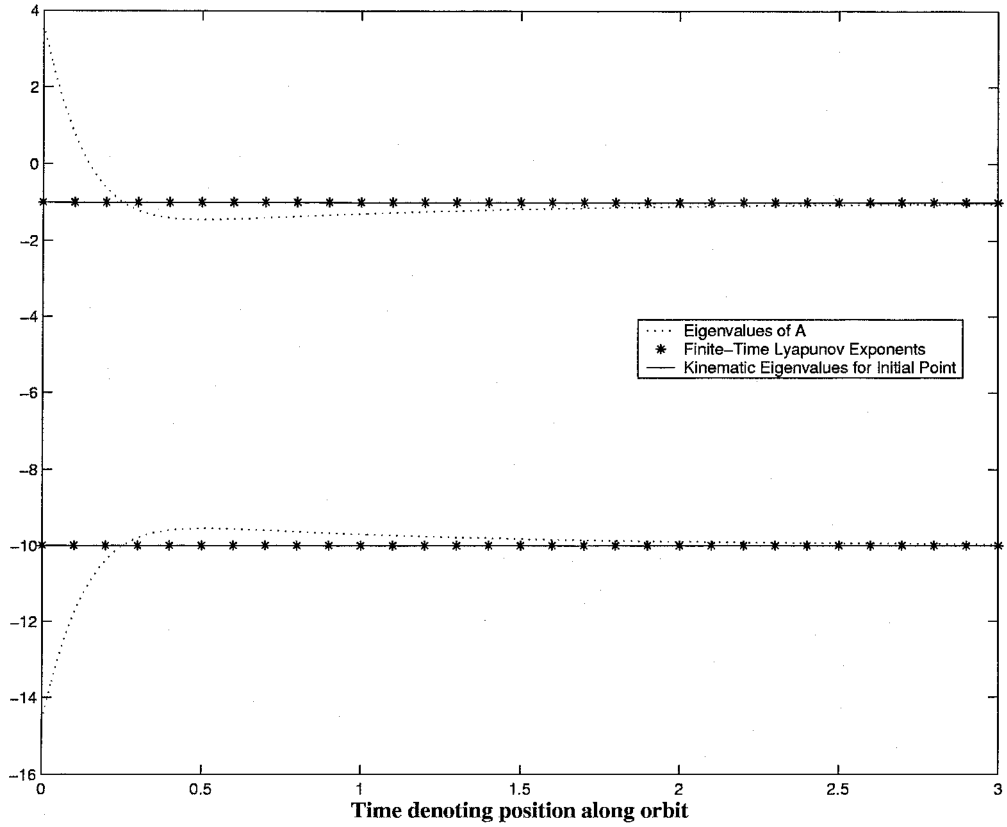


**Fig. 4 Lyapunov exponents, kinematic eigenvalues, and eigenvalues of $F$ along the orbit shown in Fig. 1b.**

by $\mathcal{E}^{fc}(\boldsymbol{x}) = \text{span}\{\boldsymbol{l}_1^x(x)\}$ and $\mathcal{E}^s(\boldsymbol{x}) = \text{span}\{\boldsymbol{n}_2^x(x)\}$. The eigenvectors of the Jacobian matrix $F$, denoted by the shorter, lighter arrows, do not in general correctly indicate the slow and fast directions. The calculation of $\boldsymbol{n}_2^x$ at $t = 0$ requires either forward or backward integration along the orbit segment leading to the $t = 0$ point. This segment lies partly outside of $\mathcal{X}$. The region $\mathcal{X}$ could be enlarged to include this segment. Another option is to use evolution equations for $N$, developed by Greene and Kim,[32] to propagate $N$ computed at a later time point, such as $t = 1$, back to the $t = 0$ point.

The Lyapunov exponents converge to $\mu_1 = -10$ and $\mu_2 = -1$ and are, thus, consistent with the eigenvalues of the equivalent linear system. Figure 3 shows that, if the exponents are computed with the non-Euclidean metric $G$ induced by the transformation, the convergence is very rapid; if the exponents are computed with the Euclidean metric, it takes about one time unit for the metric effect to become negligible. Thus, we are able to extract the correct information from the region $\mathcal{X}$ for either the Euclidean or non-Euclidean metric. When the Euclidean metric was used, the Lyapunov exponents were computed at many points along the orbit using an averaging time of $T = 1$ in each case. They are plotted in Fig. 4. Comparison of the exponents with the eigenvalues of the Jacobian matrix $F$ shows that the Jacobian eigenvalues provide good approximations once the orbit reaches the slow manifold ($t \approx 1$) but are inaccurate off the slow manifold; the Lyapunov exponents on the other hand are uniformly accurate. Figure 4 also shows the kinematic eigenvalues (see definition 3) for the non-Euclidean metric, computed by initializing Eqs. (19) and (20) with the value of $L$ for the point $(x_1, x_2) = (40, -39)$, marked $t = 0$ in Fig. 1b. The kinematic eigenvalues are constant along the trajectory and, thus, indicate that the timescale behavior is uniform. The kinematic eigenvalues for the Euclidean metric (not shown) are different than $-1$ and $-10$ initially, due to local exponential-level effects of the coordinate transformation; this is why the convergence of the exponents is slower, as already mentioned, for the Euclidean metric.

For this simple example it is clear that the vector fields $\boldsymbol{n}_2^x(\boldsymbol{x})$ and $\boldsymbol{l}_1^x(\boldsymbol{x})$ could be integrated to obtain the slow and fast curvilinear coordinate curves, respectively. As stated earlier, the general treatment of this step is beyond the scope of this paper.

## Conclusions

A step has been taken toward developing a general method of determining timescale structure in finite-dimensional nonlinear dynamic systems. The timescale information consists of finite time Lyapunov exponents and vectors. The supporting theory for the interpretation and computation of the timescale information has been synthesized. Away from equilibrium points, periodic orbits, and certain other invariant sets, the timescale structure of a dynamic system is in general not an inherent system feature, that is, invariant under arbitrary coordinate and/or metric transformations. However, regarding the usual situation of analyzing a differential equation model in a particular set of coordinates using the Euclidean metric, the timescale structure is metric and coordinate independent over the class of metric and coordinate changes that do not alter the average exponential behavior in the region of interest, and our method can identify this structure and provide the information necessary to determine the appropriate coordinates. Thus, our method is more general than scaling and other methods that assume the initial model is already in appropriate coordinates and that the task is only to identify which coordinates are slow and which are fast. In contrast to the eigenvalues and eigenvectors of the Jacobian matrix for the linearized dynamics, which do not in general correctly identify the timescale structure, the Lyapunov exponents and vectors do, in principle, correctly identify the timescale structure. On the other hand, the computation of the Lyapunov exponents and vectors is more demanding. As a first step in investigating the computational feasibility and accuracy of the timescale information, the method was verified for a low-order problem in which the correct information is known by independent means.

## Acknowledgments

This research has been supported by the National Science Foundation under Grant CMS-9610304. Discussions with S.-H. Lam at Princeton University started the first author on this research. Discussions with L.-S. Young of the Courant Institute were instrumental in understanding various aspects of dynamical systems theory relevant to this paper.

## References


[1]Naidu, D. S., and Calise, A. J., "Singular Perturbations and Time Scales in Guidance and Control of Aerospace Sytems: A Survey," *Journal of Guidance, Control, and Dynamics*, Vol. 24, No. 6, 2001, pp. 1057–1078.

[2]Kokotovic, P. V., Khalil, H. K., and O'Reilly, J., *Singular Perturbation Methods in Control: Analysis and Design*, Academic Press, New York, 1986, pp. 1–28.

[3]Ashley, H., "Multiple Scaling in Flight Vehicle Dynamic Analysis—A Preliminary Look," AIAA Paper 67-560, Aug. 1967.

[4]Calise, A. J., "Optimal Thrust Control with Proportional Navigation Guidance," *Journal of Guidance, Control, and Dynamics*, Vol. 3, No. 4, 1980, pp. 312–318.

[5]Calise, A. J., Markopoulos, N., and Corban, J. E., "Nondimensional Forms for Singular Perturbation Analyses of Aircraft Energy Climbs," *Journal of Guidance, Control and Dynamics*, Vol. 17, No. 3, 1994, pp. 584–590.

[6]Kelley, H. J., "Aircraft Maneuver Optimization by Reduced-Order Approximations," *Control and Dynamic Systems*, edited by C. T. Leondes, Academic Press, New York, 1973, pp. 131–178.

[7]Shinar, J., "On Applications of Singular Perturbation Techniques in Nonlinear Optimal Control," *Automatica*, Vol. 19, No. 2, 1983, pp. 203–211.

[8]Ardema, M. D., and Rajan, N., "Separation of Time Scales in Aircraft Trajectory Optimization," *Journal of Guidance, Control, and Dynamics*, Vol. 8, No. 2, 1985, pp. 275–278.

[9]Kelley, H. J., Cliff, E. M., and Weston, A. R., "Energy State Revisited," *Optimal Control Applications and Methods*, Vol. 7, No. 2, 1986, pp. 195–200.

[10]Lyapunov, A. M., "The General Problem of Stability of Motion," *International Journal of Control*, Vol. 55, No. 3, 1992, pp. 531–773.

[11]Oseledec, V. I., "A Multiplicative Ergodic Theorem. Lyapunov Characteristic Numbers for Dynamical Systems," *Transactions of the Moscow Mathematical Society*, Vol. 19, 1968, pp. 197–231.

[12]Barreira, L., and Pesin, Y. B., *Lyapunov Exponents and Smooth Ergodic Theory*, Univ. Lecture Series, Vol. 23, American Mathematical Society, Providence, RI, 2002, p. 26.

[13]Lorenz, E. N., "The Local Structure of a Chaotic Attractor in Four-Dimensions,"*Physica D*, Vol. 13D, Nos. 1–2, 1984, pp. 90–104.

[14]Greene, J. M., and Kim, J. S., "The Calculation of Lyapunov Spectra," *Physica D*, Vol. 24, Nos. 1–3, 1987, pp. 213–225.

[15]Goldhirsch, I., Sulem, P. L., and Orszag, S. A., "Stability and Lyapunov Stability of Dynamical Systems: A Differential Approach and A Numerical Method," *Physica D*, Vol. 27, No. 3, 1987, pp. 311–337.

[16]Lam, S. H., "Singular Perturbation for Stiff Equations Using Numerical Computations," *Lectures in Applied Mathematics*, Vol. 24, Pt. 2, 1986, pp. 513–522.

[17]Bharadwaj, S., and Mease, K. D., "A New Invariance Property of Lyapunov Characteristic Directions," *Proceedings of the American Control Conference*, Vol. 6, American Automatic Control Council, Evanston, IL, 1999, pp. 3800–3804.

[18]Bharadwaj, S., and Mease, K. D., "Geometric Structure of Multiple Time-Scale Nonlinear Systems," *Proceedings of the 14th World Congress*, Vol. 5, International Federation of Automatic Control, Laxenburg, Austria, 1999, pp. 527–532.

[19]Mease, K. D., Iravanchy, S., Bharadwaj, S., and Fedi, M., "Time Scale Analysis for Nonlinear Systems," AIAA Paper 2000-4591, Aug. 2000.

[20]Bharadwaj, S., "Geometric Structure of Multiple Time-Scale Nonlinear Dynamical System," Ph.D Dissertation, Mechanical and Aerospace Engineering Dept., Univ. of California, Irvine, CA, Aug. 1999.

[21]Abarbanel, H. D. I., Brown, R., and Kennel, M. B., "Variation of Lyapunov Exponents on a Strange Attractor," *Journal of Nonlinear Science*, Vol. 1, 1991, pp. 175–199.

[22]Vastano, J. A., and Moser, R. D., "Short-Time Lyapunov Exponent Analysis and the Transition to Chaos in Taylor-Couette Flow," *Journal of Fluid Mechanics*, Vol. 233, 1991, pp. 83–118.

[23]Wiesel, W. E., "Modal Feedback Control on Chaotic Trajectories," *Physical Review E*, Vol. 49, No. 3, 1994, pp. 1990–1996.

[24]Wiesel, W. E., "Optimal Pole Placement in Time-Dependent Linear Systems," *Journal of Guidance, Control, and Dynamics*, Vol. 18, No. 5, 1994, pp. 995–999.

[25]Fenichel, N., "Geometric Singular Perturbation Theory for Ordinary Differential Equations," *Journal of Differential Equations*, Vol. 31, No. 1, 1979, pp. 53–98.

[26]Spivak, M., *Differential Geometry*, Vol. 1, Publish or Perish, Berkeley, CA, 1979, pp. 86–230.

[27]Strang, G., *Linear Algebra and Its Applications*, Harcourt, Brace, Jovanovich, San Diego, CA, 1988, pp. 442–451.

[28]Dickinson, B., "Matrices and Linear Algebra," *The Control Handbook*, edited by W. S. Levine, CRC Press, Boca Raton, 1996, pp. 33–50.

[29]Geist, K., Parlitz, U., and Lauterborn, W., "Comparison of Different Methods for Computing Lyapunov Exponents," *Progress of Theoretical Physics*, Vol. 83, No. 5, 1990, pp. 875–893.

[30]Dieci, L., Russell, R. D., and Van Vleck, E. S., "On the Computation of Lyapunov Exponents for Continuous Dynamical Systems," *SIAM Journal on Numerical Analysis*, Vol. 34, No. 1, 1997, pp. 402–423.

[31]Young, L.-S., "Ergodic Theory of Differentiable Dynamical Systems," *Real and Complex Dynamics*, edited by A. Branner and P. G. Hjorth, NATO ASI Ser., Kluwer Academic, Norwell, MA, 1995, pp. 201–226.

[32]Greene, J. M., and Kim, J. S., "Introduction of the Metric Tensor into Linearized Evolution Equations,"*Physica D*, Vol. 36, Nos. 1–2, 1989, pp. 83–91.

[33]Wiesel, W. E., "Continuous Time Algorithm for Lyapunov Exponents I," *Physical Review E*, Vol. 47, No. 3, 1994, pp. 3686–3691.

[34]Wiesel, W. E., "Extended Lyapunov Exponents," *Physical Review A: General Physics*, Vol. 46, No. 12, 1992, pp. 7480–7491.

[35]Arnol'd, V. I., *Ordinary Differential Equations*, Springer-Verlag, New York, 1992, p. 104.

[36]Hahn, W., *Stability of Motion*, Springer-Verlag, New York, 1967, p. 33.

[37]Eichhorn, R., Linz, S. J., and Hanggi, P., "Transformation Invariance of Lyapunov Exponents," *Chaos, Solitons, and Fractals*, Vol. 12, 2001, pp. 1377–1383.