Computing Lyapunov exponents on a Stiefel manifold

by **Thomas J. Bridges and Sebastian Reich**

*Department of Mathematics and Statistics, University of Surrey,
Guildford, Surrey GU2 7XH, England*

Abstract

The problem of numerical computation of a few Lyapunov exponents of finite-dimensional dynamical systems is considered from the viewpoint of the differential geometry of Stiefel manifolds. Whether one computes one, many or all Lyapunov exponents of a continuous dynamical system by time integration, discrete or continuous orthonormalization is essential for stable numerical integration. A differential geometric view of continuous orthogonalization suggests that one restrict the linearized vectorfield to a Stiefel manifold. However, the Stiefel manifold is not in general an attracting submanifold of the ambient Euclidean space: it is a constraint manifold with a weak numerical invariant. New numerical algorithms for this problem are then designed which use the fibre-bundle characterization of these manifolds. This framework leads to a new class of systems for continuous orthogonalization which have strong numerical invariance properties and the strong skew-symmetry property. Numerical integration of these new systems with geometric integrators leads to a new class of numerical methods for computing a few Lyapunov exponents which preserve orthonormality to machine accuracy. This idea is also taken a step further by making the Stiefel manifold an attracting invariant manifold in which case standard explicit Runge-Kutta algorithms can be used. This leads to an algorithm which requires only marginally more computation than a standard explicit integration without orthogonalization. This class of methods should be particularly effective for computing a few Lyapunov exponents for large-dimension dynamical systems. The new schemes are straightforward to implement. A test case is presented for illustration, and an example from dynamical systems is presented where a few Lyapunov exponents are computed for an array of coupled oscillators.

Table of Contents

— **April 2000 (revised February 2001, April 2001)** —

1. Introduction

Lyapunov exponents (LEs) are one of the most widely used qualitative measures of complexity in continuous dynamical systems, and numerical integration of the linearization about a trajectory, that is, linear systems of the form,

$$\mathbf{u}_t = \mathbf{A}(t)\,\mathbf{u}\,, \quad \mathbf{u} \in \mathbb{R}^n\,, \tag{1.1}$$

where $\mathbf{A}(t)$ is bounded for all $t \in \mathbb{R}$, is the basis of many of the methods for determining LEs in specific systems. An important property of the linear system (1.1) is that linear independence of vectors is preserved by the flow. On the other hand, linear systems obtained this way generally have widely separated growth and decay properties, and this provides a severe test for numerical methods. In general, using an off-the-shelf numerical integrator to integrate (1.1), starting with a set of linearly-independent vectors, will not in general maintain linear independence, because each of the vectors is attracted to the direction associated with the largest growth rate.

Let $\mathbf{\Phi}(t)$ be a fundamental matrix solution for (1.1) with $\mathbf{\Phi}(0)$ orthonormal, and let $\mathbf{e}_j \in \mathbb{R}^n$, $j = 1, \ldots, n$ be an orthonormal basis for $\mathbb{R}^n$. Then, the *characteristic numbers* are defined by

$$\lambda_j = \limsup_{t \to +\infty} \frac{1}{t} \log(\|\mathbf{\Phi}(t)\mathbf{e}_j\|)\,, \quad j = 1, \ldots, n\,.$$

When the sum of the characteristic numbers is minimized, the orthonormal basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is called normal and the λ_j, $j = 1, \ldots, n$ are called *Lyapunov exponents* (cf. DIECI, RUSSELL & VAN VLECK [1997], §1).

Lyapunov exponents are preserved under any bounded, smooth, transformation of the fundamental solution matrix (see Lemma 1.6 of DIECI, RUSSELL & VAN VLECK [1997]). Moreover, it follows from Oseledec's Theorem and Lemma 2.2 of DIECI & VAN VLECK [1995] that the k *largest* LEs can be obtained by integrating (1.1) starting with a k-dimensional randomly chosen orthonormal basis. We will assume throughout that the system is *regular* and the LEs are well defined (see DIECI, RUSSELL & VAN VLECK [1997] for a definition and examples of regularity). For regular systems, the *limsup* in the above expression for the LE can be replaced by *lim*.

The simplest and most well-known solution to the linear-dependence problem associated with the numerical integration of (1.1) is to use discrete orthogonalization. For example, given k independent orthonormal vectors as initial data for a linear system on $\mathbb{R}^n$, Gram-Schmidt orthogonalization can be imposed every few time steps to ensure linear independence over reasonably long time scales. For Lyapunov exponents, this technique appears to have been first used in the numerical algorithms proposed in BENETTIN, GALGANI, GIORGILLI & STRELCYN [1980] and is still widely used today (cf. GEIST, PARLITZ & LAUTERBORN [1990], DIECI, RUSSELL & VAN VLECK [1997], DIECI & VAN VLECK [1995] and references therein). For computation of the largest exponent, discrete orthogonalization reduces to numerically integrating (1.1), with a random initial vector of unit length in $\mathbb{R}^n$, using a standard numerical integrator such as an explicit Runge-Kutta algorithm, and then normalizing the length of the solution every few time steps.

There are several problems associated with the use of discrete orthogonalization to compute the k largest Lyapunov exponents. It is a discontinous process: there is no natural ordering of vectors at each orthogonalization step. Since numerical integration of (1.1) can be highly unstable, an explicit method will require very small time steps. For large systems the small time step coupled with orthogonalization at every step or every few steps can introduce unacceptable levels of round-off error, although there are no data on this in the literature. Finally, a curious problem with discrete orthogonalization is that it has difficulty with the computation of *negative* LEs (cf. DIECI, RUSSELL & VAN VLECK [1997], p. 414). However, discrete orthogonalization is still widely used, and for dynamical systems of small dimension is reasonably efficient (see the recent numerical experiments of RAMASUBRAMANIAN & SRIRAM [2000]).

Alternatively, *continuous orthogonalization* can be used, where the system (1.1) is transformed to maintain orthogonality continuously. Theoretically, this idea goes back to work of DILIBERTO [1950], and in a numerical setting, the idea appeared first in the literature on boundary-value problems (cf. DAVEY [1983], MEYER [1986], DIECI, RUSSELL & VAN VLECK [1994] and references therein).

For the computation of *all* the LEs, the numerical analysis for continuous orthogonalization was developed and formalized by DIECI, RUSSELL & VAN VLECK [1997]. In this case, a continuous QR decomposition leads to a differential equation on the orthogonal group, and a robust geometric scheme based on Gauss-Legendre

Runge-Kutta methods preserves orthogonalization to machine accuracy. Therefore, the theory and computation of *all* the LEs for regular systems – using either continuous or discrete orthogonalization – is now well developed, but it is computationally prohibitive for large-dimension dynamical systems.

The use of continuous orthogonalization for the numerical computation of a few LEs was first proposed by Goldhirsch, Sulem & Orszag [1987] (see also Greene & Kim [1987,1989]). The starting point is the induced linear system on $\mathbb{R}^{n\times k}$ associated with (1.1):

$$\mathbf{U}_t = \mathbf{A}(t)\,\mathbf{U}\,, \quad \mathbf{U}\in\mathbb{R}^{n\times k}\,. \tag{1.2}$$

Let $\mathbf{U}(t) = \mathbf{Q}(t)\mathbf{R}(t)$ where $\mathbf{Q}$ is an $n\times k$ matrix with orthonormal columns, and $\mathbf{R}$ is a $k\times k$ upper triangular matrix. Substitution of this QR-decomposition into (1.2) leads to the following uncoupled system for $\mathbf{Q}$

$$\mathbf{Q}_t = \mathbf{A}(t)\mathbf{Q} - \mathbf{Q}\mathbf{T}\,, \quad \text{where} \quad \mathbf{T}_{i,j} = \begin{cases} (\mathbf{Q}^T\mathbf{A}\mathbf{Q})_{i,j} + (\mathbf{Q}^T\mathbf{A}\mathbf{Q})_{j,i} & i<j \\ (\mathbf{Q}^T\mathbf{A}\mathbf{Q})_{i,j} & i=j \\ 0 & i>j \end{cases} \tag{1.3}$$

The Lyapunov exponents are determined from the diagonal elements of $\mathbf{T}$ (see Goldhirsch et al. [1987, §5], Geist et al. [1990, p. 881] for details of the extraction of the LEs). However, in Goldhirsch et al. [1987], a fourth-order explicit Runge-Kutta method was used, and no special consideration was given to the implications of the numerical integrator for preservation of orthogonality over long time intervals.

The first paper to consider the implications of choice of numerical integrator when computing a few LEs is Dieci & van Vleck [1995] (hereafter denoted by DV95). They first reformulate (1.3) as

$$\mathbf{Q}_t = \mathbf{H}(\mathbf{Q},t)\,\mathbf{Q}\,, \tag{1.4}$$

in order to identify its geometric structure. The matrix $\mathbf{H}(\mathbf{Q},t)$ is defined by

$$\mathbf{H}(\mathbf{Q},t) = [\mathbf{I}_n - \mathbf{Q}\mathbf{Q}^T]\mathbf{A}(t) + \mathbf{Q}\mathbf{S}\mathbf{Q}^T\,, \quad \text{where} \quad \mathbf{S}_{i,j} = \begin{cases} (\mathbf{Q}^T\mathbf{A}\mathbf{Q})_{i,j} & i>j \\ 0 & i=j \\ -(\mathbf{Q}^T\mathbf{A}\mathbf{Q})_{j,i} & i<j\,. \end{cases} \tag{1.5}$$

(see DV95, §2.2). A central observation of DV95 is that the matrix $\mathbf{H}$ has a *weak skew-symmetry* property. The matrix $\mathbf{H}$ is called weakly skew-symmetric if

$$\mathbf{Q}^T\{\,\mathbf{H}(\mathbf{Q},t) + \mathbf{H}(\mathbf{Q},t)^T\,\}\mathbf{Q} = \mathbf{0} \quad \text{whenever} \quad \mathbf{Q}^T\mathbf{Q} = \mathbf{I}_k\,. \tag{1.6}$$

A matrix $\mathbf{H}$ is said to be *strongly skew-symmetric* if it is skew-symmetric: $\mathbf{H}^T = -\mathbf{H}$. They established that known unitary integrators such as GL-RK methods would not generally preserve orthogonality of the computed solution of (1.5), and none of the existing standard methods would preserve orthogonality. Hybrid RK methods with solution-dependent weights can be devised but are technically complicated to implement, and DV95 do not recommend them for practical implementation, and instead suggest a "projected scheme". The projected scheme of DV95 reintroduces discrete orthogonalization as a post-processing step. The scheme consists of integrating (1.4) with a standard explicit RK method, coupled with discrete orthogonalization (modified Gram-Schmidt) at each or every few time steps. This approach requires more CPU time than discrete orthogonalization applied to (1.2), but the idea is that although the integration of (1.4) will in general be unstable, it is in principle *less unstable* than integrating (1.2). The Gram-Schmidt process is associated with a decomposition of the form $\mathbf{U} = \mathbf{Q}\mathbf{R}$ with $\mathbf{R}$ upper triangular. But another natural decomposition is $\mathbf{U} = \widetilde{\mathbf{Q}}\mathbf{H}$, where $\mathbf{H}$ is symmetric, sometimes called the *polar* decomposition. In fact recently, Higham [1997], proposed a modifcation of the above two-step scheme where the post-processing step consists of a polar decomposition – using a Schulz iteration – to stabilize the computing of a few LEs. This scheme is motivated by the fact that the orthonormal polar factor is the best orthonormal approximation of a matrix $\mathbf{U}\in\mathbb{R}^{n\times k}$ of rank k in the Frobenius norm.

Continuous orthogonalization can be approached by parametrizing the group $SO(n)$ by a set of $\frac{1}{2}n(n-1)$ angles, using a sequence of Householder reflections or Givens rotations, defined continuously using differential equations. This approach has recently been proposed independently by Rangarajan, Habib & Ryne [1998] and Dieci & van Vleck [1999]. This method has the advantage that orthogonality is built in, but it is technically much more complicated to implement than either discrete or continuous orthogonalization. Rangarajan et al. give examples in $\mathbb{R}^2$ and $\mathbb{R}^3$, and this has been extended to $\mathbb{R}^4$ by Ramasubramanian & Sriram [2000], who also show that it is competitive with discrete orthogonalization. Dieci & van Vleck [1999] give general expressions for parametrizing the orthogonal group. In general, the biggest obstacle to the use of parameterizations is the level of technicality involved, and the delicate bookkeeping necessary to switch parameterizations during the integration.

Another way to integrate systems of the form (1.5) when $\mathbf{H}$ is weakly skew-symmetric is the Munthe-Kaas Runge-Kutta (MK-RK) methods (cf. MUNTHE-KAAS [1998,1999] and MUNTHE-KAAS & ZANNA [1997]). These algorithms allow use of explicit Runge-Kutta methods in the Lie algebra and still preserve to machine accuracy the flow along the Lie group (in this case $SO(n)$), when coupled with a matrix exponentiation algorithm. As far as we are aware, these algorithms have not been tested on the Lyapunov exponent problem.

The subject of this paper is a differential-geometric formulation for the numerical computation of the k largest Lyapunov exponents of a system of the form (1.1) where $1 \leq k < n$ using continuous orthogonalization. One of the main results of this paper is that it is possible to reformulate (1.5) so that $\mathbf{H}$ is strongly skew-symmetric. Indeed, we show that it is natural to expect that $\mathbf{H}$ can be transformed into skew-symmetric form in general. Both of these issues are addressed by treating continuous orthogonalization as equivalent to the restriction of (1.2) to a compact submanifold, a Stiefel manifold, of $\mathbb{R}^{n\times k}$ and use the differential geometry of these manifolds to transform the system of ODEs.

Before addressing the geometry of this problem, it will be useful to recall why it is extremely desirable to reformulate (1.4) so that $\mathbf{H}$ is skew-symmetric. DIECI ET AL. [1994] show that a robust geometric numerical integration scheme, based on implicit Gauss-Legendre Runge-Kutta (GL-RK) methods, can be devised for systems of the form (1.4) *when* $\mathbf{H}$ *is skew-symmetric*, where orthogonality is preserved to machine accuracy during the integration. This preservation property is extremely important for integration over long time intervals.

A Stiefel manifold $V_k(\mathbb{R}^n)$ is defined as

$$V_k(\mathbb{R}^n) = \{\, \mathbf{Q} \in \mathbb{R}^{n\times k} \;:\; \mathbf{Q}^T\mathbf{Q} = \mathbf{I}_k \,\}\,.$$

Stiefel manifolds are a generalization of the unit sphere, and they reduce to S^{n-1} in the case $k = 1$, but in general they have very different geometric properties from spheres when $1 < k < n$. When $k = n$ the Stiefel manifold is $O(n)$, the orthogonal group. A Stiefel manifold is a regular submanifold of $\mathbb{R}^{n\times k}$ of dimension $nk - \frac{1}{2}k(k+1)$, and their geometry has been considerably studied because of their importance in the theory of fibre bundles (cf. STIEFEL [1935], STEENROD [1951], KAWAKUBO [1990]). Recently, they have been found to be useful in numerical linear algebra (cf. EDELMAN ET AL. [1998], ELDÉN & PARK [1999]).

A differential geometric definition of continuous orthonormalization is: restrict equation (1.2) to a Stiefel manifold. Constructively this means that the initial data should be a point on the manifold $V_k(\mathbb{R}^n)$ and the right-hand side of (1.2) should be in the tangent space of $V_k(\mathbb{R}^n)$ for each t. This restriction of (1.2) to a Stiefel manifold can be achieved in many different ways depending on the chosen projection operator. Independently of the projection, one always obtains a continuous factorization of the solutions $\mathbf{U}(t) = \mathbf{Q}(t)\mathbf{R}(t)$, where $\mathbf{Q}$ is in the Stiefel manifold and $\mathbf{R}$ is an invertible $k \times k$ matrix.

Different projections from $\mathbb{R}^{n\times k}$ onto the tangent space of $V_k(\mathbb{R}^n)$ will lead to different paths through the Stiefel manifold. Classical continuous orthogonalization – for example systems of the form (1.4) – will result in an $\mathbf{R}(t)$ which is upper triangular. However, the polar decomposition provides an interesting alternative to the upper-triangular decomposition (cf. HIGHAM [1997]).

If we take the standard Frobenius inner product on $\mathbb{R}^{n\times k}$, then the natural projection operator $\Pi_{\mathbf{Q}}$ from $\mathbb{R}^{n\times k}$ into the tangent space of $V_k(\mathbb{R}^n)$ leads to the following transformation of (1.2)

$$\mathbf{Q}_t = \Pi_{\mathbf{Q}}(\mathbf{A}(t)\mathbf{Q}) = \mathbf{A}(t)\mathbf{Q} + \mathbf{Q}\mathbf{g}\,, \quad \mathbf{Q}(t)\big|_{t=0} = \mathbf{Q}_0 \in V_k(\mathbb{R}^n)\,, \tag{1.7}$$

with $\mathbf{g}$ taking the form

$$\mathbf{g} = -\mathrm{sym}\,(\mathbf{Q}^T\mathbf{A}\mathbf{Q})\,, \tag{1.8}$$

where $\mathrm{sym}(\mathbf{B}) = \frac{1}{2}(\mathbf{B} + \mathbf{B}^T)$, for any square matrix $\mathbf{B}$ (cf. BRIDGES [1999]). The system (1.7) with $\mathbf{g}$ taking the form (1.8) is called a *Stiefel integrator*: it is the differential geometric interpretation of continuous orthogonalization based on the Frobenius metric and provides a *continuous polar decomposition*. This scheme can also be written in the form (1.4) by rearranging the right-hand side of the system in (1.7)

$$\mathbf{Q}_t = \mathbf{H}(\mathbf{Q},t)\mathbf{Q}\,, \quad \text{with} \quad \mathbf{H}(\mathbf{Q},\cdot) = (\mathbf{I}_n - \mathbf{Q}\mathbf{Q}^T)\mathbf{A} + \mathbf{Q}\,\mathrm{skew}(\mathbf{Q}^T\mathbf{A}\mathbf{Q})\mathbf{Q}^T\,,$$

where $\mathrm{skew}(\mathbf{B}) = \frac{1}{2}(\mathbf{B} - \mathbf{B}^T)$ for any square matrix $\mathbf{B}$. Note that the skew-symmetric matrix in the second term of $\mathbf{H}$ in this equation differs from that in (1.5).

Although the formulation (1.7)-(1.8) leads to the rigorous concept of restriction to and integration on a Stiefel manifold, it does not immediately lead to a robust numerical algorithm. In other words, straightforward numerical integration of (1.7) using a Runge-Kutta algorithm, or even a specialized numerical integration scheme such as a GL-RK scheme, does not produce a robust algorithm. Indeed the right-hand side of (1.7) is weakly skew-symmetric: when (1.7)-(1.8) is written in the form (1.4), $\mathbf{H}$ satisfies (1.6) but is *not* skew-symmetric. Moreover, treating (1.7) as a dynamical system on a constraint manifold, the manifold $V_k(\mathbb{R}^n)$ is a weak invariant.

The concepts of weak and strong numerical invariance are fundamental to the numerics of differential equations on constraint manifolds (cf. LEIMKUHLER & REICH [1994]). The constraint manifold for the system (1.7) is defined by $\mathbf{P}(\mathbf{Q}) = 0$, where

$$\mathbf{P}(\mathbf{Q}) = \mathbf{Q}^T\mathbf{Q} - \mathbf{I}_k\,. \tag{1.9}$$

The constraint $\mathbf{P}$ is called a *strong invariant* if $\frac{d}{dt}\mathbf{P}(\mathbf{Q}) = 0$ for all $\mathbf{Q}$, independent of the value of $\mathbf{P}(\mathbf{Q})$, and $\mathbf{P}$ is called a *weak invariant* if $\frac{d}{dt}\mathbf{P}(\mathbf{Q}) = 0$ for all $\mathbf{Q}$, only if $\mathbf{P}(\mathbf{Q})$=0. For any $\mathbf{Q}$ satisfying (1.7), $\mathbf{P}(\mathbf{Q})$ satisfies

$$\frac{d}{dt}\mathbf{P}(\mathbf{Q}) = \mathbf{g}\mathbf{P} + \mathbf{P}\mathbf{g}\,.$$

Since the right-hand side vanishes in general only if $\mathbf{P}$ is identically zero, the constraint $\mathbf{P}$ is a weak constraint. This weak invariance property is in general distinct from the weak skew-symmetry property.

One of the main results of this paper is that both the weak invariance property and the weak skew-symmetry property can be eliminated by using the fibre bundle characterization of Stiefel manifolds. Using the quotient group representation of Stiefel manifolds, $V_k(\mathbb{R}^n)$ can be represented as the base manifold of a principal fibre bundle with fibre $SO(n-k)$ and ambient bundle $SO(n)$ (cf. STEENROD [1951], KAWAKUBO [1991]). Results from differential geometry suggest that it should *always be possible to lift a vectorfield on $V_k(\mathbb{R}^n)$ to a vectorfield on $SO(n)$*. Indeed, once this is clear it is not difficult to modify the right-hand side of (1.7) to give it the skew-symmetry property. Write (1.4)-(1.5) or (1.7)-(1.8) in the general form

$$\mathbf{Q}_t = \mathbf{H}(\mathbf{Q},t)\mathbf{Q} \quad \text{with} \quad \mathbf{H}(\mathbf{Q},t) = (\mathbf{I}_n - \mathbf{Q}\mathbf{Q}^T)\mathbf{A}(t) + \Sigma(t)\,, \tag{1.10}$$

where $\Sigma(t)$ is *any* skew-symmetric matrix. Now, the right-hand side is mathematically equivalent on $V_k(\mathbb{R}^n)$ to

$$\mathbf{H}(\mathbf{Q},t) = [(\mathbf{I}_n - \mathbf{Q}\mathbf{Q}^T)\mathbf{A}(t) - \mathbf{A}(t)^T(\mathbf{I}_n - \mathbf{Q}\mathbf{Q}^T)] + \Sigma(t)\,,$$

since $\mathbf{A}(t)^T(\mathbf{I}_n - \mathbf{Q}\mathbf{Q}^T)\mathbf{Q} = 0$ when $\mathbf{Q} \in V_k(\mathbb{R}^n)$, and so,

$$\mathbf{Q}_t = \mathbf{H}(\mathbf{Q},t)\mathbf{Q} \quad \text{with} \quad \mathbf{H}(\mathbf{Q},t) = 2\,\mathrm{skew}[(\mathbf{I}_n - \mathbf{Q}\mathbf{Q}^T)\mathbf{A}(t)] + \Sigma(t)\,. \tag{1.11}$$

For $\mathbf{Q} \in V_k(\mathbb{R}^n)$, the system (1.11) is equivalent to either (1.5) or (1.7), depending on the chosen expression for $\Sigma(t)$, but (1.11) has eliminated both weaknesses of (1.5) or (1.7): $\mathbf{H}(\mathbf{Q},t)$ is now strongly skew-symmetric, and $V_k(\mathbb{R}^n)$ is a strong numerical constraint manifold. The details are given in §4. Note that $\mathbf{H}$ is *not* a projection onto the space of skew-symmetric matrices. It is the system (1.11), and related equations derived from this system, that we are proposing as the basis of the new algorithms for computing a few Lyapunov exponents.

CHRISTIANSEN & RUGH [1997] proposed that the unstable equations for continuous orthogonalization (1.3) could be stabilized by adding a dissipation term,

$$\mathbf{Q}_t = \mathbf{A}(t)\mathbf{Q} - \mathbf{Q}\mathbf{T} - \gamma\,\mathbf{Q}\mathbf{P}(\mathbf{Q})\,,$$

where γ is a positive real number. The equation for $\mathbf{P}(\mathbf{Q})$ is modified to

$$\frac{d}{dt}\mathbf{P}(\mathbf{Q}) = -\mathbf{P}\mathbf{T} - \mathbf{T}^T\mathbf{P} - 2\gamma\mathbf{P}(\mathbf{I}_k + \mathbf{P})\,.$$

Therefore if γ is chosen large enough, the manifold $\mathbf{P}(\mathbf{Q}) = \mathbf{0}$ is attracting. However this method is *conditionally attracting*: it is attracting only if γ satisfies certain solution dependent inequalities. Theoretical results of CHRISTIANSEN & RUGH show that γ must be greater than $-\lambda_n$ where λ_n is the lowest LE. However, numerical experiments of RAMASUBRAMANIAN & SRIRAM [2000] found that in practice γ must be significantly larger, and they found that the method failed completely on some problems.

Using the differential geometric framework, the dissipative term of CHRISTIANSEN & RUGH can be modifed to make the Stiefel manifold *unconditionally attracting*. If we now use the new form of the vectorfield (1.11) which has the strong skew-symmetry and the strong invariance property, and add an attracting term

$$\mathbf{Q}_t = \mathbf{H}(\mathbf{Q}, t)\mathbf{Q} - \gamma\, \mathbf{Q}\, \mathbf{P}(\mathbf{Q})\,,$$

where again γ is any positive real number, the equation for $\mathbf{P}(\mathbf{Q})$ reduces to

$$\frac{d}{dt}\mathbf{P}(\mathbf{Q}) = -2\gamma \mathbf{P}(\mathbf{I}_k + \mathbf{P})\,.$$

In this case the manifold $\mathbf{P}(\mathbf{Q})$ is unconditionally attracting.

In §5 we use theory of ASCHER, CHIN & REICH [1994] to prove that the modifed system is attracted uniformly in time to a perturbed Stiefel manifold. The advantage of the formulation is that any reasonably accurate off-the-shelf numerical integrator, such as a fourth-order explicit Runge-Kutta algorithm (RK4), can be used. This scheme can also be interpreted as a two-step method: using the split-step scheme of ASCHER, CHIN & REICH [1994] leads to a numerical calculation that is equivalent to the integration of (1.11) by some efficient explicit ODE solver followed by one step of a pseudo-Newton iteration, which we show is precisely a *Schulz iteration*, an established algorithm for optimal computation of the polar decomposition (cf. HIGHAM [1997]). The unconditionally attracting algorithm has the advantages that it is easy to implement, and requires marginally more computer time that the strong form of the equations. The algorithm therefore has excellent properties for computing a few LEs for orbits of large-dimension dynamical systems.

The paper is organized as follows. In §2, we treat the case of computing the largest Lyapunov exponent by restricting (1.1) to the unit sphere. This case provides a setting where the properties of weak invariant and weak skew-symmetry can be transformed in an elementary way. In §§3-4 the main algorithm for computing a few Lyapunov exponents on a Stiefel manifold is presented, and in §5 the details of the algorithm based on an attractive Stiefel manifold are presented. In §6, numerical results are presented for an example with an exact solution, and in §7 results of the computation of Lyapunov exponents for a large system of coupled oscillators are presented.

2. Computing Lyapunov exponents on the unit sphere

The standard way to compute the largest Lyapunov exponent is to integrate the linearized system, $\mathbf{u}_t = \mathbf{A}(t)\mathbf{u}$, with a standard integrator such as an explicit fourth-order Runge-Kutta method, and then scale $\mathbf{u}(t)$ to have unit length every few time steps. The continuous version of this algorithm is to restrict the linearized system to $V_1(\mathbb{R}^n)$ which is the unit sphere in $\mathbb{R}^n$.

The purpose of this section is twofold: the geometry of linear systems restricted to $V_1(\mathbb{R}^n)$ provides an elementary example where the difficulties of weak versus strong numerical invariance and weak versus strong skew-symmetry appear, and can be resolved in an elementary way, and secondly, this approach surprisingly leads to several new geometric algorithms for computing the largest Lyapunov exponent.

The linear system $\mathbf{u}_t = \mathbf{A}(t)\mathbf{u}$ for $\mathbf{u} \in \mathbb{R}^n$ can be restricted to a nonlinear ODE on S^{n-1} by modifying it to

$$\mathbf{q}_t = \mathbf{A}(t)\mathbf{q} + \mathbf{q}\,\mathrm{g}\,, \quad \text{with} \quad \mathbf{q}(0) = \mathbf{q}_0 \in \mathbb{R}^n \cap S^{n-1}\,, \tag{2.1a}$$

where

$$\mathrm{g} = -\mathbf{q}^T\mathbf{A}\mathbf{q}\,, \tag{2.1b}$$

or

$$\mathbf{q}_t = [\,\mathbf{I}_n - \mathbf{q}\mathbf{q}^T\,]\mathbf{A}(t)\mathbf{q}\,. \tag{2.2}$$

For each $\mathbf{q} \in V_1(\mathbb{R}^n)$, the right-hand side of (2.1a) or (2.2) is in the tangent space of S^{n-1}.

However, we can not expect a numerical integrator to keep the flow on the unit sphere in general: the unit sphere in this case is not necessarily a stable or attracting invariant submanifold of $\mathbb{R}^n$. We will rectify this below, after establishing the properties of the sphere for the system (2.1)-(2.2).

The unit sphere is a weak invariant for (2.1)-(2.2). The sphere can be considered as a constraint: $\mathrm{P}(\mathbf{q})=0$, with $\mathrm{P}(\mathbf{q})=\mathbf{q}^T\mathbf{q}-1$, and a straightforward calculation shows that $\mathrm{P}(\mathbf{q})$ satisfies

$$\frac{d}{dt}\mathrm{P}(\mathbf{q}) = 2\mathrm{g}\mathrm{P}(\mathbf{q})\,, \tag{2.3}$$

when $\mathbf{q}$ satisfies (2.1) or (2.2). The right-hand side of (2.3) does not vanish in general unless the constraint $\mathrm{P}(\mathbf{q})=0$ is satisfied exactly.

There is a trivial way to transform the constraint from weak to strong, by replacing g in (2.1b) by

$$\mathrm{g}_1 = -(\mathbf{q}^T\mathbf{q})^{-1}(\mathbf{q}^T\mathbf{A}\mathbf{q})\,, \tag{2.4}$$

and replacing (2.1a) by

$$\mathbf{q}_t = \mathbf{A}(t)\mathbf{q} + \mathbf{q}\,\mathrm{g}_1\,, \tag{2.5}$$

or

$$\mathbf{q}_t = [\,\mathbf{I}_n - \mathbf{q}(\mathbf{q}^T\mathbf{q})^{-1}\mathbf{q}^T\,]\mathbf{A}(t)\mathbf{q}\,. \tag{2.6}$$

Mathematically this system is equivalent to (2.1) when $\mathbf{q}\in S^{n-1}$, but numerically it is superior: for this system $\frac{d}{dt}\mathrm{P}(\mathbf{q})=0$ for *any* $\mathbf{q}$ satisfying (2.5), and so P is a strong numerical invariant. The modified system (2.5) is in fact Davey's algorithm for continuous orthonormalization, for the special case of a one-dimensional orthonormal subspace. (Compare (2.5) and (2.6) with equations (32)-(33) in DAVEY [1983] and Theorem 1 of MEYER [1986].)

However, even though the numerical integration of this modified equation will be substantially more reliable than (2.1)-(2.2), there are still potential difficulties. First, although the "inverse" in (2.6) is a scalar, its generalization to $k>1$ will require a pseudoinverse at each stage. Secondly, the sphere constraint is now a strong invariant, but use of a standard integrator – such as an RK4 method – may lead in general to a slow algebraic drift off the constraint surface; see LEIMKUHLER AND REICH [1994], and the numerical results in §6. This drift error can be remedied by using a GL-RK method, but when $k>1$ the form of the system (2.6) is such that orthonormality will not in general be preserved at intermediate steps in the required iteration.

Going back to (2.2), using the property that $(\mathbf{I}-\mathbf{q}\mathbf{q}^T)\mathbf{q}=0$ on the constraint manifold, we can add a term proportional to this to the right-hand side of (2.2):

$$\mathbf{q}_t = [\,(\mathbf{I}_n-\mathbf{q}\mathbf{q}^T)\,\mathbf{A}(t) - \mathbf{A}(t)^T(\mathbf{I}_n-\mathbf{q}\mathbf{q}^T)\,]\,\mathbf{q}\,, \tag{2.7}$$

or

$$\mathbf{q}_t = \mathbf{H}(\mathbf{q},t)\,\mathbf{q}\,, \tag{2.8}$$

with

$$\mathbf{H}(\mathbf{q},t) = 2\,\mathrm{skew}[(\mathbf{I}_n-\mathbf{q}\mathbf{q}^T)\mathbf{A}(t)]\,. \tag{2.9}$$

This apparently trivial modification to (2.1)-(2.2) results in P becoming a strong numerical invariant. More importantly, it leads to a form with *strong* skew-symmetry: the coefficient matrix $\mathbf{H}(\mathbf{q},t)$ is skew-symmetric. Note that the matrix $\mathbf{H}(\mathbf{q},t)$ is not strictly a projection of the matrix $(\mathbf{I}_n-\mathbf{q}\mathbf{q}^T)\mathbf{A}(t)$ in the right-hand side of (2.2) onto the space of skew-symmetric matrices: it is twice the projection.

There are in fact many ways that (2.1) can be written in the form (2.8) with skew-symmetric $\mathbf{H}(\mathbf{q},t)$. Let c_1 and c_2 be any two real numbers such that $c_1+c_2=1$. Then the system (2.8) with $\mathbf{H}$ any member of the following family of skew-symmetric matrices

$$\mathbf{H}(\mathbf{q},t) = 2\,\mathrm{skew}[\,c_2\mathbf{A} - \mathbf{q}\mathbf{q}^T(c_2\mathbf{A}+c_1\mathbf{A}^T)]\,, \tag{2.10}$$

is equivalent to (2.1) and has the strong invariant and strong skew-symmetry property ((2.9) is recovered by taking $c_1=0$).

The advantage of the formulation (2.8) with $\mathbf{H}(\mathbf{q},t)$ skew-symmetric, is that there are excellent numerical methods that are tailored to such systems: the DRV unitary iteration scheme coupled with a GL-RK integrator (DIECI, RUSSELL & VAN VLECK [1994]), and the MK-RK methods which tailor Runge-Kutta methods to Lie group and homogeneous spaces (cf. MUNTHE-KAAS [1998,1999], MUNTHE-KAAS & ZANNA [1997]).

The transformation between (2.1) and (2.8) is related to the principal fibre bundle structure of $SO(n)$. The Lie group $SO(n)$ can be characterized as a principal fibre bundle over S^{n-1} with the fibre being a copy of $SO(n-1)$ at each point (another way to say this is that the sphere can be characterized as the quotient of two Lie groups: $S^{n-1} = SO(n)/SO(n-1)$). Essentially, given a vectorfield on the base manifold S^{n-1} (such as (2.1)) there is well-defined horizontal lift of the vectorfield into the tangent space of the ambient bundle (see KOBAYASHI & NOMIZU [1969] for the differential geometry of this construction). We will not need this theory here, but it underlies the construction of (2.8)-(2.9) and will underlie the constructions of similar lifted vectorfields for Stiefel manifolds.

In principle, the transformation to (2.8) has made the sphere stable as an invariant manifold. We can however go one step further and make the sphere *attracting* by modifying (2.8) to

$$\mathbf{q}_t = \mathbf{H}(\mathbf{q},t)\mathbf{q} - \gamma\,\mathbf{q}\mathrm{P}(\mathbf{q})\,, \tag{2.11}$$

where γ is a positive constant. When $\mathbf{q} \in S^{n-1}$ the additional term is a null term. However,

$$\frac{d}{dt}\mathrm{P} = -2\gamma\,\mathrm{P} - 2\gamma\,\mathrm{P}^2\,, \tag{2.12}$$

when $\mathbf{q}$ satisfies (2.11) and therefore when $\mathrm{P} \neq 0$, there is exponential attraction onto the unit sphere. The advantage of this formulation over (2.8) is that stable numerical integration is possible without a special geometric integrator. Further details of this construction are given in §5 for the more general case of $k > 1$ and numerical results are presented in §6.

In summary, computing the largest Lyapunov exponent of (2.1) can be reduced to solving (2.8) numerically using a geometric integrator, starting with a random initial point on the unit sphere, or, integration of (2.11) with *any* reasonably accurate numerical integrator. In both cases the exponent is given by

$$\lambda_{max} = -\lim_{t\to\infty}\frac{1}{t}\int_0^t \mathrm{g}(s)\,ds\,, \tag{2.13}$$

when the limit exists, where g is defined in (2.1b).

3. Restricting linear systems to Stiefel manifolds

To compute a k-dimensional linear subspace of solutions of the linear system, $\mathbf{u}_t = \mathbf{A}(t)\mathbf{u}$, $\mathbf{u} \in \mathbb{R}^n$, the starting point is the induced ODE on $\mathbb{R}^{n\times k}$,

$$\mathbf{U}_t = \mathbf{A}(t)\,\mathbf{U}\,, \quad \mathbf{U}(t)\big|_{t=0} = \mathbf{U}_0 \in \mathbb{R}^{n\times k}\,. \tag{3.1}$$

Given any $n\times k$ matrix $\mathbf{U}(t)$ of rank k, the columns span a k-dimensional subspace, and any such subspace has an orthonormal basis. More generally, let $\mathbf{R}$ be *any* invertible $k\times k$ matrix, then the columns of $\mathbf{U} \in \mathbb{R}^{n\times k}$ and $\mathbf{Q} \in \mathbb{R}^{n\times k}$ span the same space when $\mathbf{U} = \mathbf{Q}\mathbf{R}$. Letting these matrices depend on t and substitution of this expression into (3.1) leads to the following equation for $\mathbf{Q}(t)$

$$\mathbf{Q}_t = \mathbf{A}(t)\,\mathbf{Q} + \mathbf{Q}\,\mathbf{g} \quad \text{with} \quad \mathbf{R}_t = -\mathbf{g}\,\mathbf{R}\,. \tag{3.2}$$

Given an expression for $\mathbf{g}$ this system of equations is closed.

Let $V_k(\mathbb{R}^n)$ be the Stiefel manifold of k-dimensional orthonormal subspaces of $\mathbb{R}^n$. If we require the columns of $\mathbf{Q}(t)$ to be an orthonormal basis for the space spanned by the columns of $\mathbf{U}(t)$, then the first equation of (3.2) describes the flow through a Stiefel manifold. This suggests that the natural way to determine $\mathbf{g}$ is to project the right-hand side of the equation for $\mathbf{Q}(t)$ onto the tangent space of a Stiefel manifold at the point $\mathbf{Q} \in V_k(\mathbb{R}^n)$ (cf. BRIDGES [1999], §2). A projection from $\mathbb{R}^{n\times k}$ onto the tangent space of $V_k(\mathbb{R}^n)$ at $\mathbf{Q} \in V_k(\mathbb{R}^n)$ is defined by

$$\Pi_{\mathbf{Q}}(\mathbf{Z}) = \mathbf{Z} - \mathbf{Q}\,\mathrm{sym}(\mathbf{Q}^T\mathbf{Z}) \quad \text{for any} \quad \mathbf{Z} \in \mathbb{R}^{n\times k}\,, \tag{3.3}$$

where $\mathrm{sym}(\mathbf{A})$ corresponds to projection of the square matrix $\mathbf{A}$ onto its symmetric part. Applying this projection to the equation for $\mathbf{Q}$ leads to the following nonlinear equation for the flow on a Stiefel manifold,

$$\mathbf{Q}_t = \mathbf{A}(t)\,\mathbf{Q} + \mathbf{Q}\,\mathbf{g}(t)\,, \quad \mathbf{Q}(t)\big|_{t=0} = \mathbf{Q}_0 \in \mathbb{R}^{n\times k} \cap V_k(\mathbb{R}^n)\,, \tag{3.4}$$

with

$$\mathbf{g}(t) = -\mathrm{sym}\,(\mathbf{Q}^T\mathbf{A}(t)\mathbf{Q})\,, \tag{3.5}$$

or

$$\mathbf{Q}_t = [\,\mathbf{I}_n - \mathbf{Q}\mathbf{Q}^T\,]\mathbf{A}(t)\,\mathbf{Q} + \mathbf{Q}\,\mathbf{S}\,, \quad \text{where} \quad \mathbf{S} = \mathrm{skew}(\mathbf{Q}^T\mathbf{A}(t)\mathbf{Q})\,. \tag{3.6}$$

Equations (3.4)-(3.5) generalize (2.1a)-(2.1b) and equation (3.6) generalizes (2.2), and the systems (3.4)-(3.6) have the same problem as (2.1)-(2.2): the constraint is a weak invariant. From a dynamical systems viewpoint, the Stiefel manifold is an invariant manifold of (3.4) or (3.6), but it is not in general a stable or attracting submanifold in the ambient space $\mathbb{R}^{n\times k}$.

Let $\mathbf{P}(\mathbf{Q}) = \mathbf{Q}^T\mathbf{Q} - \mathbf{I}_k$, then a straightforward calculation shows that

$$\frac{d}{dt}\mathbf{P}(\mathbf{Q}) = \mathbf{g}\mathbf{P} + \mathbf{P}\mathbf{g}\,, \tag{3.7}$$

when $\mathbf{Q}(t)$ is a solution of (3.4) or (3.5), and so in general $\frac{d}{dt}\mathbf{P} = 0$ only if $\mathbf{P}$ is identically zero.

The weak invariance can be transformed to strong invariance by introducing a pseudoinverse, as in (2.6), but this formulation will still have the weak skew-symmetry property. Here we show that a generalization of (2.8)-(2.9) is possible.

Using the fact that $(\mathbf{I}_n - \mathbf{Q}\mathbf{Q}^T)\mathbf{Q}$ vanishes on the Stiefel manifold, we can add a term proportional to this in (3.6),

$$\mathbf{Q}_t = [\,(\mathbf{I}_n - \mathbf{Q}\mathbf{Q}^T)\mathbf{A}(t) - \mathbf{A}(t)^T(\mathbf{I}_n - \mathbf{Q}\mathbf{Q}^T)\,]\,\mathbf{Q} + \mathbf{Q}\,\mathrm{skew}(\mathbf{Q}^T\mathbf{A}(t)\mathbf{Q})\,, \tag{3.8}$$

or, rearranging terms,

$$\mathbf{Q}_t = \mathbf{H}(\mathbf{Q},t)\,\mathbf{Q}\,, \quad \mathbf{Q}(t)\big|_{t=0} = \mathbf{Q}_0 \in \mathbb{R}^{n\times k}\cap V_k(\mathbb{R}^n)\,, \tag{3.9}$$

where

$$\mathbf{H}(\mathbf{Q},t) = \mathrm{skew}[\,\mathbf{A}(t) - 2\mathbf{Q}\mathbf{Q}^T\mathrm{sym}(\mathbf{A}(t))\,]\,. \tag{3.10}$$

It is immediate from the form of $\mathbf{H}(\mathbf{Q},t)$ that – for each fixed $t\in\mathbb{R}$ – it is a mapping from $V_k(\mathbb{R}^n)$ into the Lie algebra of $\mathbf{SO}(n)$, and the strong invariance of $\mathbf{P}$ is then an immediate consequence of the skew-symmetry of $\mathbf{H}(\mathbf{Q},t)$.

It is the system (3.9) with $\mathbf{H}(\mathbf{Q},t)$ in the form (3.10) that we propose as a new formulation for computing a few Lyapunov exponents, when integrated using a geometric integrator.

The theory behind the transformation between (3.2) and (3.9) is similar to that described in §2. The Lie group $SO(n)$ can also be characterized as a principal fibre bundle over $V_k(\mathbb{R}^n)$ with the fibre being a copy of $SO(n-k)$ at each point (another way to say this is that the Stiefel manifold can be characterized as the quotient of two Lie groups: $V_k(\mathbb{R}^n) = SO(n)/SO(n-k)$); cf. Steenrod [1951], Kawakubo [1991]. Essentially, given a vectorfield on the base manifold $V_k(\mathbb{R}^n)$ (such as (3.4) or (3.6)) there is well-defined horizontal lift of the vectorfield into the tangent space of the ambient bundle (see Kobayashi & Nomizu [1969]).

4. A few Lyapunov exponents on a Stiefel manifold

With the construction of §3, the computation of a few Lyapunov exponents reduces to solving a system on the Stiefel manifold of the form

$$\mathbf{Q}_t = \mathbf{H}(\mathbf{Q},t)\,\mathbf{Q}\,, \quad \mathbf{Q}(t)|_{t=0} = \mathbf{Q}_0 \in V_k(\mathbb{R}^n)\,, \tag{4.1}$$

where $\mathbf{Q}_0$ is a randomly chosen point on $V_k(\mathbb{R}^n)$, and $\mathbf{H}(\mathbf{Q},t)$ is skew-symmetric. Assuming the system is regular, the k largest Lyapunov exponents are deduced from the eigenvalues of $\mathbf{g}$ in (3.5). For computational purposes it is useful to transform $\mathbf{g}$ to upper triangular form. However, when $\mathbf{g}$ is in upper triangular form – in which case the Lyapunov exponents can be deduced from the diagonal entries – it is still possible to transform it to the skew-symmetric form (4.1). In fact, a system for $\mathbf{Q}$ with $\mathbf{g}$ in upper triangular form can be characterized as a different path on the Stiefel manifold.

Consider the canonical form for the $\mathbf{Q}$ equation restricted to a Stiefel manifold, equations (3.4)-(3.5). In canonical form, $\mathbf{g}(t)$ is the symmetric matrix (3.5). However, $\mathbf{g}(t)$ can also be chosen to be upper triangular as in GOLDHIRSCH ET AL. [1987] and DV95 (see equations (1.3) and (1.5) in the introduction). We will show that the transformation of (3.4) from $\mathbf{g}$ being symmetric to upper triangular is equivalent to choosing a different path in the Stiefel manifold.

Let $\mathbf{h}(t)$ be a one-parameter family of $k \times k$ orthogonal matrices satisfying $\mathbf{h}(0) = \mathbf{I}_k$. Then

$$\mathbf{W}(t) = \mathbf{Q}(t)\mathbf{h}(t)\,, \tag{4.2}$$

satisfies $\mathbf{W}(0) = \mathbf{Q}_0$ and $\mathbf{W}(t)$ is a path in $V_k(\mathbb{R}^n)$. Differentiating (4.2) and using (3.4) leads to

$$\mathbf{W}_t = \mathbf{A}(t)\mathbf{W} + \mathbf{W}\widetilde{\mathbf{g}}(t)\,, \quad \text{with} \quad \widetilde{\mathbf{g}} = \mathbf{h}^T\mathbf{g}\mathbf{h} + \mathbf{h}^T\mathbf{h}_t\,, \tag{4.3}$$

where $\mathbf{g}$ is the symmetric form (3.5). Given any symmetric matrix $\mathbf{g}$ it is possible to find an $\mathbf{h} \in SO(k)$ such that $\widetilde{\mathbf{g}}$ in (4.3) is upper triangular.

Now, take (4.3) with $\widetilde{\mathbf{g}}$ chosen in upper triangular form as in (1.3) as a starting point (i.e. choose $\widetilde{\mathbf{g}} = \mathbf{T}$ with $\mathbf{T}$ defined in (1.3)); then (4.3) can be recast as

$$\mathbf{W}_t = [\mathbf{I}_n - \mathbf{W}\mathbf{W}^T]\mathbf{A}(t)\mathbf{W} + \mathbf{W}\mathbf{S}\,,$$

with $\mathbf{S}$ the skew-symmetric matrix defined in (1.5). Following the construction in §3 it is now straightforward to transform this system into the form with strong skew-symmetry:

$$\mathbf{W}_t = \mathbf{H}(\mathbf{W},t)\,\mathbf{W} \quad \text{with} \quad \mathbf{H}(\mathbf{W},t) = 2\,\mathrm{skew}([\mathbf{I}_n - \mathbf{W}\mathbf{W}^T]\mathbf{A}(t)) + \mathbf{W}\mathbf{S}\mathbf{W}^T\,. \tag{4.4}$$

When this skew-symmetric system is integrated, instead of (4.1), then the k-largest Lyapunov exponents are deduced from the diagonal elements of the upper triangular matrix $\widetilde{\mathbf{g}}$,

$$\lambda_j = -\lim_{t\to\infty} \frac{1}{t}\int_0^t \widetilde{\mathbf{g}}_{jj}(s)\,ds\,, \quad j = 1,\ldots,k\,, \tag{4.5}$$

assuming the limits exist.

Both the systems (4.1) and (4.4) have the strong skew-symmetry property and therefore geometric integrators such as GL-RK with unitary iteration and MK-RK methods can be used and will maintain orthogonality – and hence preserve the Stiefel manifold – to machine accuracy. These methods are costlier than standard explicit RK methods, because iteration is required at each step in the GL-RK methods and exponentiation is required in the MK-RK methods. However, the elegance of preserving the Stiefel manifold to machine accuracy is well worth the cost for systems of modest size. Numerical experiments with this method are given in §6 and §7. In the next section we will show how systems of the form (4.1) or (4.4) can be modified in a straightforward way to make the Stiefel manifold attracting, and so standard explicit RK methods can be used.

5. Making the Stiefel manifold an attracting invariant manifold

Given a nonlinear ODE with a constraint manifold that has the strong invariant property, it is straightforward to make the constraint manifold attracting (cf. ASCHER, CHIN & REICH [1994]). In this section we modify the skew-symmetric form of the flow on the Stiefel manifold (i.e. forms (4.1) or (4.4)) – following the argument used to make the unit sphere attracting in §2 – to make the Stiefel manifold attracting.

Take as a starting point

$$\mathbf{Q}_t = \mathbf{H}(\mathbf{Q},t)\mathbf{Q}\,, \quad \text{with} \quad \mathbf{Q}(t)|_{t=0} = \mathbf{Q}_0 \in V_k(\mathbb{R}^n) \quad \text{and} \quad \mathbf{H}(\mathbf{Q},t)^T = -\mathbf{H}(\mathbf{Q},t)\,. \tag{5.1}$$

For this system the constraint of orthogonality, $\mathbf{P}(\mathbf{Q}) = \mathbf{0}$ with $\mathbf{P}(\mathbf{Q}) = \mathbf{Q}^T\mathbf{Q} - \mathbf{I}_k$, is a strong numerical invariant.

For any $\mathbf{F} \in \mathbb{R}^{n\times k}$, the modified equation

$$\mathbf{Q}_t = \mathbf{H}(\mathbf{Q},t)\mathbf{Q} + \mathbf{F}\mathbf{P}(\mathbf{Q})\,,$$

still has the property that the right-hand side is a mapping from $V_k(\mathbb{R}^n)$ into the tangent space of $V_k(\mathbb{R}^n)$ because $\mathbf{P}=0$ when $\mathbf{Q}\in V_k(\mathbb{R}^n)$. The simplest form for $\mathbf{F}$ which will force $V_k(\mathbb{R}^n)$ to be attracting is obtained by taking $\mathbf{F}=-\gamma\,\mathbf{Q}$, where γ is an arbitrary positive real number. The modified equation becomes

$$\mathbf{Q}_t=\mathbf{H}(\mathbf{Q},t)\mathbf{Q}-\gamma\,\mathbf{Q}\,\mathbf{P}(\mathbf{Q})\,. \tag{5.2}$$

The attractivity of $V_k(\mathbb{R}^n)$ follows from

$$\frac{d}{dt}\mathbf{P}=-2\gamma\,\mathbf{P}-2\gamma\,\mathbf{P}^2\,, \tag{5.3}$$

when $\mathbf{Q}\in V_k(\mathbb{R}^n)$ satisfies (5.2). With this property, an explicit RK method can be used to integrate (5.2) and any standard ODE software package can be applied.

To see the impact of computing the right-hand side of (5.2) when $\gamma\neq 0$, take $\mathbf{H}(\mathbf{Q},t)$ to be in the standard form (1.11) with $\Sigma(t)$ any skew symmetric matrix, then (5.2) becomes,

$$\mathbf{Q}_t=\left[\,(\mathbf{I}_n-\mathbf{Q}\mathbf{Q}^T)\mathbf{A}(t)\,\right]\;\mathbf{Q}+\Sigma(t)+(\mathbf{A}(t)^T-\gamma\,\mathbf{I}_n)\,\mathbf{Q}\,\mathbf{P}(\mathbf{Q})\,.$$

The first and second terms on the right-hand side correspond to the weak and unstable form, with the third term incorporating both the strong term and the attracting term. It is apparent that including the attracting term has insignificant effect on the operation count, and the strong term has the same order of operation count as the weak term. More importantly, however, is the fact that the scheme is unconditionally attracting, and elementary to implement.

Ascher, Chin & Reich [1994] suggest that another way to integrate systems of the form (5.2) is to use a splitting method. The $\mathbf{H}(\cdot,t)$ part is integrated with a standard efficient method of sufficient accuracy, such as RK4, and then the dissipative term can be integrated with an Euler-type step.

This algorithm is constructed as follows. Integrate (5.2) one step from t_n to t_{n+1} with $\gamma=0$ using RK4 for example – or any efficient high-order method – and call this solution $\widetilde{\mathbf{Q}}_{n+1}$. Then integrate

$$\mathbf{Q}_t=-\gamma\,\mathbf{Q}\,\mathbf{P}(\mathbf{Q})\,,$$

using a forward Euler step with the same stepsize Δt but with γ chosen so that $\gamma\Delta t=\frac{1}{2}$ and $\mathbf{Q}^n$ replaced by $\widetilde{\mathbf{Q}}_{n+1}$,

$$\mathbf{Q}_{n+1}=\widetilde{\mathbf{Q}}_{n+1}-\gamma\Delta t\,\widetilde{\mathbf{Q}}_{n+1}\mathbf{P}(\widetilde{\mathbf{Q}}_{n+1})\,.$$

Now choose $\gamma\Delta t=\frac{1}{2}$:

$$\mathbf{Q}_{n+1}=\widetilde{\mathbf{Q}}_{n+1}\left(\mathbf{I}_k-\tfrac{1}{2}\,\mathbf{P}(\widetilde{\mathbf{Q}}_{n+1})\right)\,. \tag{5.4}$$

Not only is this choice optimal, it leads to the interpretation of the second step of the split-scheme as one step of a *pseudo-Newton* iteration. This step is also equivalent to one step of the *Schulz iteration* in linear algebra, which is based on a Newton algorithm for computing the polar decomposition (cf. Higham [1997]).

To see the characterization as a pseudo-Newton step, apply Newton's method to the constraint: $\mathbf{Q}_{n+1}^T\mathbf{Q}_{n+1}=\mathbf{I}_k$ by letting $\mathbf{Q}^{n+1}=\widetilde{\mathbf{Q}}^{n+1}+\delta\mathbf{Q}^{n+1}$ with $\delta\mathbf{Q}^{n+1}$ in the *normal* space of $V_k(\mathbb{R}^n)$ at the point $\widetilde{\mathbf{Q}}^{n+1}$. Now, an element in the normal space of $V_k(\mathbb{R}^n)$ can be expressed form $\widetilde{\mathbf{Q}}^{n+1}\mathbf{\Lambda}$ where $\mathbf{\Lambda}$ is some *symmetric* matrix (cf. Bridges [1999], §2). Therefore substitution into the constraint and neglect of quadratic terms leads to the following form for one step of the Newton iteration:

$$\mathbf{\Lambda}\widetilde{\mathbf{Q}}_{n+1}^T\widetilde{\mathbf{Q}}_{n+1}+\widetilde{\mathbf{Q}}_{n+1}^T\widetilde{\mathbf{Q}}_{n+1}\mathbf{\Lambda}=\mathbf{I}_k-\widetilde{\mathbf{Q}}_{n+1}^T\widetilde{\mathbf{Q}}_{n+1}\,. \tag{5.5}$$

This expression can be simplified by approximating $\mathbf{\Lambda}$ with

$$\mathbf{\Lambda}=\tfrac{1}{2}\left(\mathbf{I}_k-\widetilde{\mathbf{Q}}_{n+1}^T\widetilde{\mathbf{Q}}_{n+1}\right)+\cdots=-\tfrac{1}{2}\,\mathbf{P}(\widetilde{\mathbf{Q}}_{n+1})+\cdots\,, \tag{5.6}$$

by using the following approximation: $\mathbf{\Lambda}\widetilde{\mathbf{Q}}_{n+1}^T\widetilde{\mathbf{Q}}_{n+1}+\widetilde{\mathbf{Q}}_{n+1}^T\widetilde{\mathbf{Q}}_{n+1}\mathbf{\Lambda}\approx 2\mathbf{\Lambda}$. Therefore one step of the pseudo-Newton approximation is

$$\begin{aligned}\mathbf{Q}_{n+1}&=\widetilde{\mathbf{Q}}_{n+1}+\delta\mathbf{Q}_{n+1}\\&=\widetilde{\mathbf{Q}}_{n+1}+\widetilde{\mathbf{Q}}_{n+1}\mathbf{\Lambda}\\&=\widetilde{\mathbf{Q}}_{n+1}\;(\mathbf{I}_k+\mathbf{\Lambda})\\&\approx\widetilde{\mathbf{Q}}_{n+1}\;\left(\mathbf{I}_k-\tfrac{1}{2}\,\mathbf{P}(\widetilde{\mathbf{Q}}_{n+1})\right)\,,\end{aligned}$$

using (5.6), which is to be compared with the forward Euler approximation (5.4).

A simple two-step algorithm is then

$$\begin{aligned} \widetilde{\mathbf{Q}}_{n+1} &= \phi^p_{\Delta t}(\mathbf{Q}_n) \\ \mathbf{Q}_{n+1} &= \widetilde{\mathbf{Q}}_{n+1}\left(\mathbf{I}_k - \tfrac{1}{2}\mathbf{P}(\widetilde{\mathbf{Q}}_{n+1})\right) , \end{aligned} \tag{5.7}$$

where $\phi^p_{\Delta t}(\mathbf{Q}_n)$ represents the discretization of (5.2) with $\gamma = 0$ by a method of $\mathcal{O}(\Delta t^p)$ accuracy, such as an explicit p-th order RK method. The second step in (5.7) is equivalent to the Schulz iteration proposed by Higham [1997] as a projection step. The advantage here is that the association with (5.2) enables us to show that the complete two-step algorithm is unconditionally attacting. In fact, one of the features of the split-step scheme (5.7) is that *the second Euler-type step does not reduce the order of accuracy of the first step*. In other words the overall accuracy of the two step scheme is $\mathcal{O}(\Delta t^p)$ (cf. Ascher, Chin & Reich [1994], Theorem 3.2).

6. A numerical test model for LEs

A model problem with known exact values for the LEs is constructed as follows. Let $\mathbf{D} \in \mathbb{R}^{n\times n}$ be a diagonal matrix with diagonal entries $D_{ii} = (3-i)/n\,, \quad i = 1,\ldots,n$. In the test cases presented below we will be taking $n = 20$.

Starting with the constant coefficient ODE

$$\mathbf{v}_t = \mathbf{D}\mathbf{v}, \qquad \mathbf{v} \in \mathbb{R}^n\,, \tag{6.1}$$

introduce the coordinate transformation $\mathbf{u} = \mathbf{W}(t)\mathbf{v}$, where $\mathbf{W}(t)$ is an orthogonal matrix defined by the ODE

$$\mathbf{W}_t = \mathbf{J}\mathbf{W}, \quad \mathbf{W}(0) = \mathbf{I}_n\,. \tag{6.2}$$

The matrix $\mathbf{J}$ is a constant skew-symmetric matrix. For the numerical experiments $\mathbf{J}$ is computed as the skew-symmetric part of a random $n \times n$ matrix generated in MATLAB using the command `rand(n)`. After transformation, the vector $\mathbf{u}(t)$ satisfies the time-dependent ODE

$$\mathbf{u}_t = \mathbf{A}(t)\mathbf{u}\,, \quad \mathbf{u} \in \mathbb{R}^n\,, \quad \text{with} \quad \mathbf{A}(t) = \mathbf{W}(t)\mathbf{D}\mathbf{W}(t)^T + \mathbf{J}\,. \tag{6.3}$$

The Lyapunov exponents for the system (6.3) are the diagonal elements of $\mathbf{D}$. We will compute the $k = 6$ largest exponents using continuous orthgonalization and a range of numerical schemes applied to (6.3).

In `Experiment (i)`, the strong system introduced in §4,

$$\mathbf{Q}_t = \mathbf{H}(\mathbf{Q},t)\mathbf{Q} \quad \text{with} \quad \mathbf{H}\ \text{skew-symmetric}, \tag{6.4}$$

is integrated using the second-order explicit midpoint method

$$\begin{aligned} \mathbf{Q}_{n+1/2} &= \mathbf{Q}_n + \tfrac{1}{2}\Delta t\mathbf{H}(\mathbf{Q}_n, t_n)\mathbf{Q}_n, \\ \mathbf{Q}_{n+1} &= \mathbf{Q}_n + \Delta t\mathbf{H}(\mathbf{Q}_{n+1/2}, t_{n+1/2})\mathbf{Q}_{n+1/2}\,. \end{aligned} \tag{6.5}$$

In `Experiment (ii)`, the unconditionally attracting form of the equations introduced in §5,

$$\mathbf{Q}_t = \mathbf{H}(\mathbf{Q},t)\mathbf{Q} - \gamma\mathbf{Q}(\mathbf{Q}^T\mathbf{Q} - \mathbf{I}_k)\,, \tag{6.6}$$

with $\mathbf{H}$ skew-symmetric is integrated using the explicit midpoint method followed by one step of the pseudo-Newton algorithm given in (5.7).

`Experiment (iii)` is the same as (ii) except that the pseudo-Newton iteration step is applied twice. In `Experiment (iv)`, the strong form (6.4) is integrated using the linearly implicit midpoint method

$$\begin{aligned} \mathbf{Q}_{n+1/2} &= \mathbf{Q}_n + \tfrac{1}{2}\Delta t\mathbf{H}(\mathbf{Q}_n, t_n)\mathbf{Q}_n, \\ \mathbf{Q}_{n+1} &= \mathbf{Q}_n + \Delta t\mathbf{H}(\mathbf{Q}_{n+1/2}, t_{n+1/2})\tfrac{1}{2}\left(\mathbf{Q}_{n+1} + \mathbf{Q}_n\right)\,. \end{aligned} \tag{6.7}$$

This method preserves the Stiefel manifold to round-off error.

In each case, the equations are integrated over a time-interval $[0, 1000]$ and a step-size $\Delta t = 0.1$ was used, except for `Experiment (i)` where $\Delta t = 0.05$ was required to satisfy the explicit stability condition. Since the Lyapunov exponents of (6.3) correspond to the diagonal entries of the constant matrix $\mathbf{D}$ and, hence, $\mathbf{g}_n =$ const. for n large enough, we have used the simplified numerical approximation $\lambda_i(t_n) = -(\mathbf{g}_n)_{ii}$, for $i = 1, \dots k$.

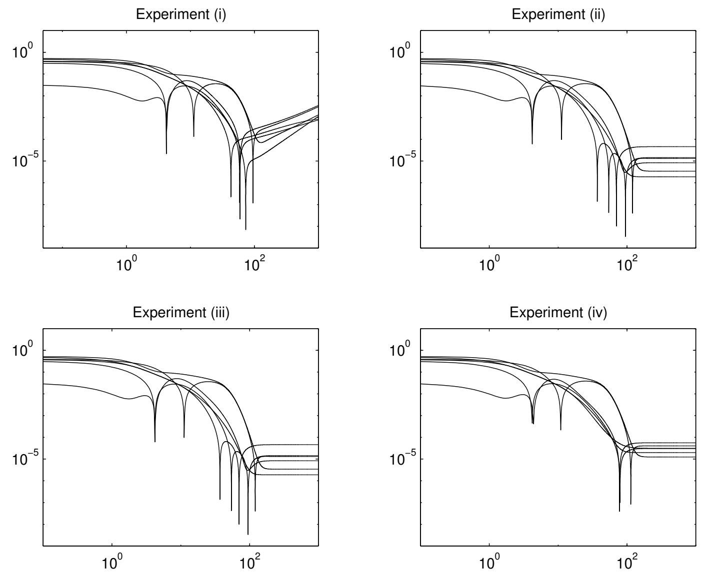

Figure 6.1. Errors in the computed Lyapunov exponents versus time

The numerical results for the four experiments can be found in Figure 6.1 and Figure 6.2. In `Experiment (i)` the increasing error at large times is clearly visible in Figure 6.1. On the other hand, all three of the other experiments yield a comparable accuarcy for the six largest Lyapunov exponents.

In Figure 6.2, the error is at the machine accuracy level for `Experiment (iii)` and `Experiment (iv)`. In Figure 6.2, the drift in the value of the orthogonality constraint for the modified implicit midpoint method is due to round-off errors in the required matrix inversion. It can also be seen from Figure 6.2 that one or two iterations with the pseudo-Newton iteration yields an excellent conservation of the orthogonality constraint.

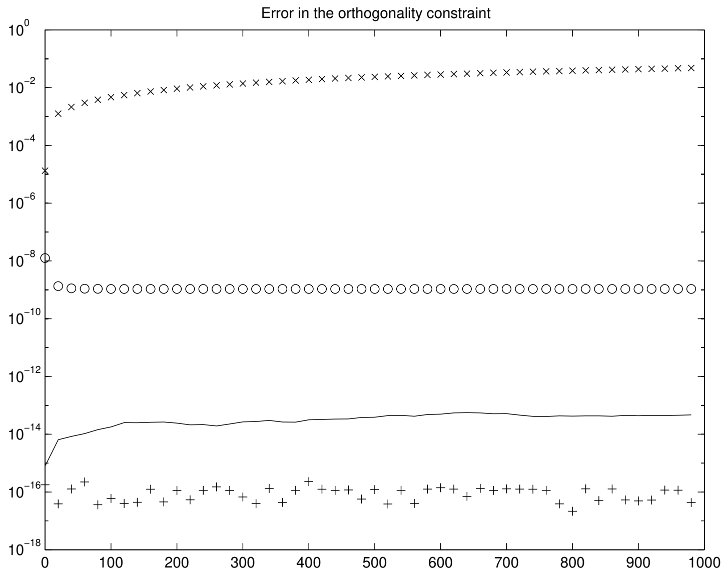

Figure 6.2. Error in the orthogonality constraint versus time. The symbols correspond to: (x) Experiment (i), (o) Experiment (ii), (+) Experiment (iii), (-) Experiment (iv).

We now perform a second test, which will illustrate the delicate nature of integration when the constraint is weak.

The weak formulation of continuous orthogonalization – for example, either (1.5) or (1.7) – has $V_k(\mathbb{R}^n)$ as a weak invariant manifold. In general, a weak invariant could be attracting or repelling, and this property could also depend on time. Therefore, in some cases, one could be fortunate and get attractivity of $V_k(\mathbb{R}^n)$ when integrating. For example, numerical experiments by GOLDHIRSCH ET AL. [1987] used an explicit numerical method to integrate the system (1.5) and did not report any instability.

As an experiment, we applied the explicit midpoint method to the weak form of the equations and computed the six largest Lyapunov exponents. The two plots on the left side of Figure 6.3 show the result of this experiment – it quickly becomes unstable. On the other hand, if only the three largest Lyapunov exponents – which are positive – are computed using exactly the same method, then the simulation was found to be stable! One might conclude from this that an explicit method, combined with the weak form of continuous orthogonalization, might work well if only positive exponents are computed. However, there is no reason to believe that this property will be repeated even for positive exponents of other systems. If an algorithm of any generality is required, integration of the weak form of the equations is unsatisfactory.

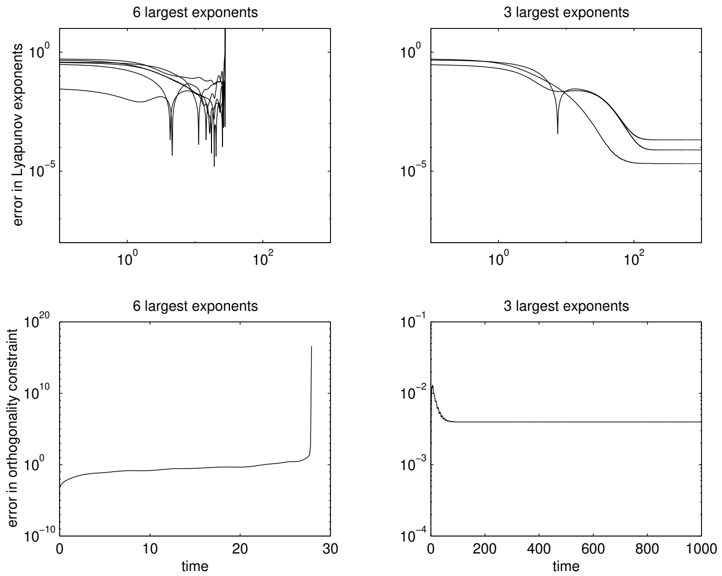

Figure 6.3. Numerical behavior of the integration of the weak form of the equations using an explicit method, when either 3 or 6 exponents are computed.

7. Computing LEs for a large system of coupled oscillators

In this section, we consider the computation of a few Lyapunov exponents using Stiefel integrators, of a large array of coupled oscillators of the form

$$\begin{aligned}
\frac{d^2y}{dt^2} &= -\alpha\,(y^2-1)\dot{y} - \omega^2 y\,,\\
\frac{d^2x_1}{dt^2} &= -d_1\,\dot{x}_1 - \beta\,\left[V'(x_1-x_N) - V'(x_2-x_1)\right] + \sigma\,y\,,\\
\frac{d^2x_i}{dt^2} &= -d_i\,\dot{x}_i - \beta\,\left[V'(x_i-x_{i-1}) - V'(x_{i+1}-x_i)\right]\,, \qquad i=2,\ldots,N\,,
\end{aligned} \tag{7.1}$$

where $N \geq 2$, the coefficients $d_i \geq 0$ are damping parameters, α, ω, β, σ are specified parameters, and $V(x)$ is a specified potential. The system is a model for a ring of weakly-damped oscillators with amplitude $x_i(t)$ and periodic boundary condition $x_{N+1} = x_1$. The ring is forced externally by the position component, $y(t)$, of the limit cycle of a van der Pol oscillator, and the potential $V(x)$ is taken to be a single-well Duffing potential $V(x) = (x^2/2) + (x^4/4)$.

This system is introduced in DRESSLER [1988, §II], and it has been used as a test case for computing Lyapunov exponents by GEIST ET AL. [1990, §5] and DV95, Example 5.3. We fix parameter values for α, β, σ and N following DV95: $\omega = 1.82$, $\alpha = 1$, $\beta = 1$, $\sigma = 4$, and $N = 5$. However we make the damping parameters smaller: $d_i = 0.0125$ for i odd and $d_i = 0.0075$ for i even. These values are chosen because they lead to more complex behaviour in the phase space, and more than two positive Lyapunov exponents. Three different numerical integrators are tested for computing Lyapunov exponents on a Stiefel manifold for this system.

The first numerical method uses the fourth-order GL-RK method for the numerical solution of the differential equation (7.1) and for the integration of the associated linear system. The associated linear system is transformed into a nonlinear skew-symmetric system on $V_k(\mathbb{R}^{12})$ following the formulation outlined in §4,

$$\mathbf{Q}_t = \mathbf{H}(\mathbf{Q},t)\mathbf{Q}, \quad \mathbf{Q} \in V_k(\mathbb{R}^{12}) \quad \text{with} \quad \mathbf{H}(\mathbf{Q},t) = -\mathbf{H}(\mathbf{Q},t)^T \quad \text{for any} \quad \mathbf{Q} \in V_k(\mathbb{R}^{12}),. \tag{7.2}$$

We compute only the four largest Lyapunov exponents ($k = 4$). The numerical solutions obtained from a GL-RK method using unitary iteration will preserve the Stiefel manifold up to the truncation error in the fixed point iterations used to implement the implicit Gauss-Legendre method. This approach for computing a few Lyapunov exponents is the most elegant, but it does have the drawback that the GL-RK schemes are implicit and, hence, are costly to implement.

The GL-RK integration is compared with methods based on a standard fourth-order explicit Runge-Kutta (RK4) method. In this case, the RK4 integrator will be used to integrate both (7.1) and (7.2) with the augmented stabilization term, i.e.

$$\mathbf{Q}_t = \mathbf{H}(\mathbf{Q},t)\mathbf{Q} - \gamma\mathbf{Q}\mathbf{P}(\mathbf{Q}). \tag{7.3}$$

We apply the two values $\gamma = 1$ and $\gamma = 10$.

All the simulations are run for a time-interval $[0, 4000]$ using a step-size of $\Delta t = 0.01$. A set of random initial conditions is used. The finite-time Lyapunov exponents are found by integrating

$$\frac{d}{dt}\Lambda_i = -g_{ii}(t), \qquad \lambda_i = \frac{\Lambda_i(t)}{t}$$

together with the equations (7.1) and (7.3). The numerically computed Lyapunov exponents can be found in Table 7.1 together with a floating point operation count. In the last row of Table 7.1 we include a run using discrete QR applied to (7.3) with $\gamma = 0$. The same RK4 routine is used to integrate, and then at the end of each time step the MATLAB `qr` routine is used to re-orthogonalize.

γ	Method	λ_1	λ_2	λ_3	λ_4	flops in MATLAB
$\gamma = 0$	GL4	0.1251	0.0937	0.0481	0.0231	7.0388e+10
$\gamma = 1$	RK4	0.1243	0.0872	0.0510	0.0226	1.6991e+10
$\gamma = 10$	RK4	0.1242	0.0882	0.0491	0.0218	1.6991e+10
$\gamma = 0$ (+ discrete QR)	RK4	0.1242	0.0883	0.0496	0.0209	1.6162e+10

Table 7.1. Computed LEs at $T = 4000$ and number of floating point operations as provide by the MATLAB `flops` routine.

The time evolution of the four Lyapunov exponents and the error of the orthogonality constraint can be found in Figure 7.1. We note that the magnitude of the Lyapunov exponents is still fluctuating at $T = 4000$ and therefore an even longer simulation would be required to find more accurate LEs for this system of equations.

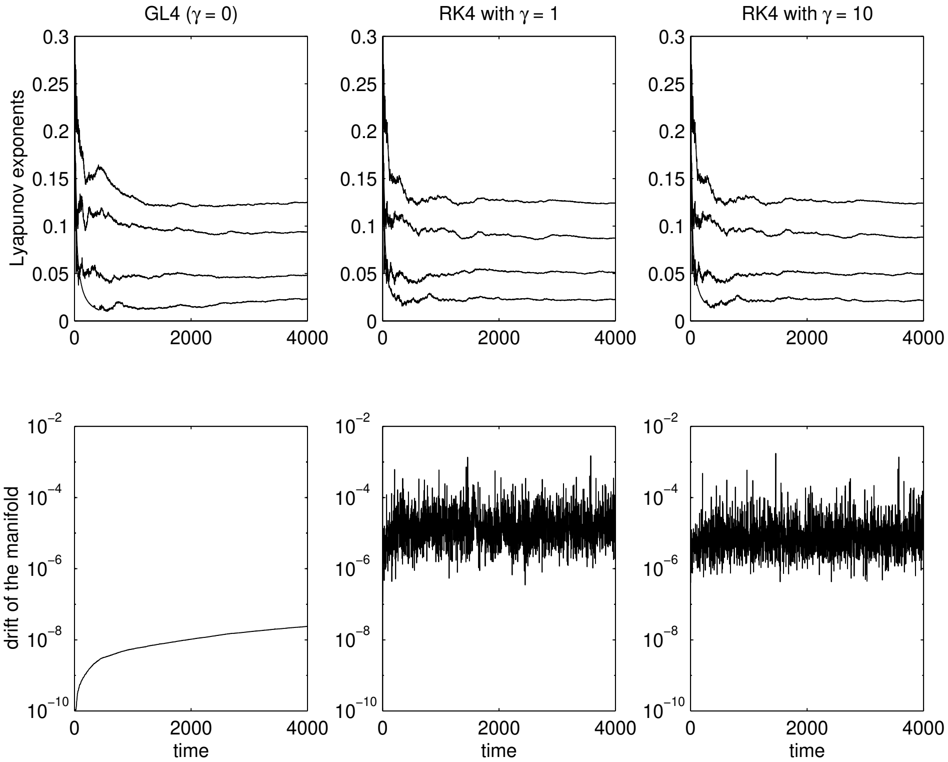

Figure 7.1. Four largest finite-time Lyapunov exponents and error in the orthogonality constraint versus time for GL4 ($\gamma = 0$) and RK4 with (i) $\gamma = 1$ and (ii) $\gamma = 10$.

An advantage of the unconditionally attracting algorithm (7.3) is that any off-the-shelf explicit integrator can be used, directly. For example, the equations (7.1) and (7.3) will be integrated numerically using the variable step-size MATLAB integration routine `ode45` with the standard parameter settings (`RelTol = 1.0e-3` and `AbsTol = 1.0e-6`). We use the formulation (7.3) with either $\gamma = 0$ or $\gamma = 1$ – see Table 7.2. No stabilization or post-processing is applied or required. The explicit algorithm is applied to (7.3) directly.

γ	`AbsTol` (`RelTol`)	λ_1	λ_2	λ_3	λ_4	flops in MATLAB
$\gamma = 0$	$1.0e - 06$ $(1.0e - 03)$	0.1270	0.8585	0.6982	0.7426	9.8204e+09
$\gamma = 1$	1.0e-06 $(1.0e - 03)$	0.1224	0.0882	0.0498	0.0226	1.244e+10
$\gamma = 1$	$1.0e - 08$ $(1.0e - 04)$	0.1266	0.0862	0.0508	0.0206	2.9122e+10

Table 7.2. Computed LEs at $T = 4000$ and number of floating point operations as provide by the MATLAB `flops` routine.

For comparison we also applied `ode45` to the formulation (7.3) with $\gamma = 1$ and smaller parameter values of `RelTol = 1.0e-4` and `AbsTol = 1.0e-8`.

The results of this experiment can be found in Figure 7.2 and Table 7.2. The inablity of a standard explicit variable step-size method to properly integrate the strong invariant but not attracting version – (7.3) with $\gamma = 0$ – can be clearly seen.

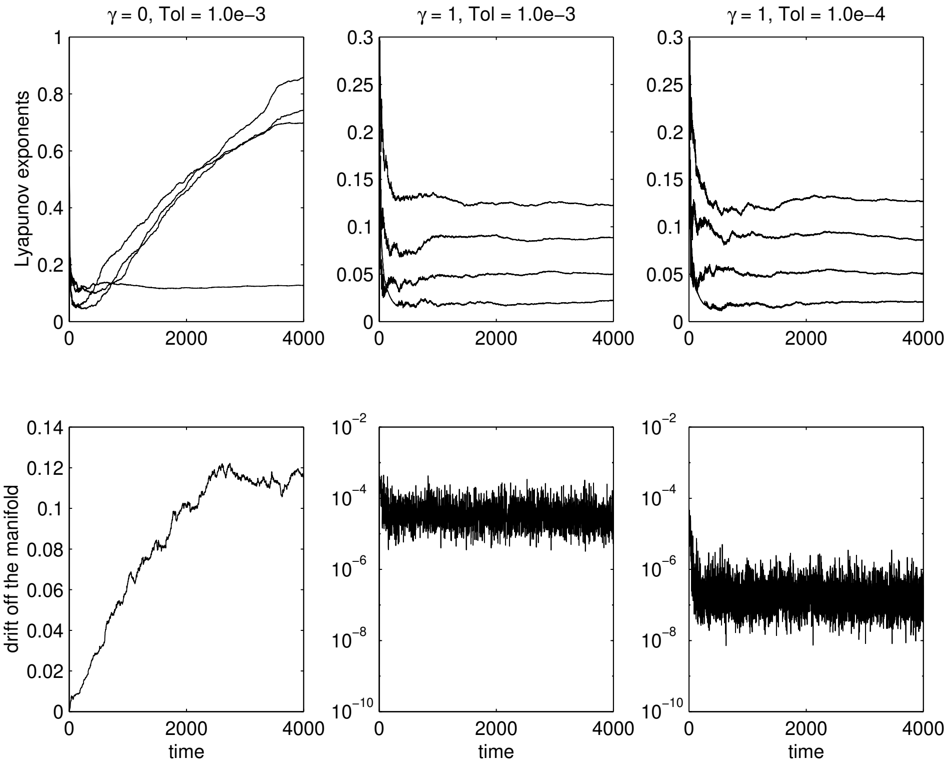

Figure 7.2. Four largest finite-time Lyapunov exponents and error in the orthogonality constraint versus time for the formulation (5.2) with (i) $\gamma = 1$ and (ii) $\gamma = 0$. The overall system of ODEs is integrated using the variable step-size MATLAB routine `ode45` with two different tolerances.

8. Concluding remarks

Several new algorithms have been presented which use the differential geometry of Stiefel manifolds in an essential way. Treating the Stiefel manifold as a submanifold of $\mathbb{R}^{n\times k}$, the equations associated with the linearization of a vectorfield about a basic state were transformed to have the "strong skew-symmetry" property. This property had heretofore been unknown for k independent orthonormal solutions of linear systems on $\mathbb{R}^n$ with $k < n$. Systems with this property are then in proper form for the use of geometric integrators such as GL-RK and MK-RK methods. Although GL-RK and MK-RK methods require more computer time – due to the implicit nature of GL-RK and the matrix exponentiation in MK-RK – they provide rigorously the property that the Stiefel manifold will be preserved to machine accuracy over very long time intervals.

The theory underlying the transformation from standard form for continuous orthonormalization to the strong skew-symmetric form is the principal fibre bundle characterization of Stiefel manifolds. These results suggest that the theory of principal fibre bundles may have something to offer in the design of numerical methods. The questions from the numerical point of view were quite clear. If a system of ODEs is required to satisfy a constraint, there are two geometric properties that are very desirable: that the constraint be a strong numerical invariant and that it have the strong skew-symmetry property. Our results indicate that not only is it straightforward to formulate the vector fields with the strong skew-symmetry property, this transformation is associated with a horizontal lift of a certain underlying principal fibre bundle. The whole idea of horizontal lifts may be useful in other areas of numerical analysis where certain geometric properties are known to be very desirable for numerics. The general question would start with a manifold $\mathcal{M}$ – in this case a sphere or a

Stiefel manifold – and then ask if there is an ambient bundle $\mathcal{A}$ – such as $SO(n)$ in this case – such that $\mathcal{A}$ is a principal fibre bundle over $\mathcal{M}$. If this is the case, then the main question would be: would a horizontal lift of the vectorfield on $\mathcal{M}$ to $\mathcal{A}$ lead to a more useful geometric structure?

The second family of new algorithms was based on making the Stiefel manifold an attracting invariant manifold. The resulting equations can be either solved by a standard numerical integrators or by the split-step scheme proposed in §5 with an $\mathcal{O}(\Delta t^p)$ accurate scheme for the first step, the overall accuracy is the same for the two steps. It would appear that the split-step scheme for the attracting Stiefel manifold has all the advantages of the scheme in §3-4. However, the split-scheme method will not preserve the Stiefel manifold to machine accuracy. It follows from Theorem 3.2 of Ascher, Chin & Reich [1994] that the split-step scheme for stabilizing strong invariants does not preserve the original manifold. On the other hand, it preserves a *perturbed invariant manifold* uniformly in time. The perturbed invariant manifold $\mathcal{M}_{\Delta t}$ is of $\mathcal{O}(\Delta t^{p+1})$ close to the original manifold, uniformly in time. The proof of Theorem 3.2 involves the use of Fenichel's invariant manifold theory (see Ascher, Chin & Reich [1994] for details). Application of this theory to the present case shows that the split-step scheme in §5 results in an attracting invariant manifold $V_k(\mathbb{R}^n, \Delta t)$, where

$$V_k(\mathbb{R}^n, \Delta t) = \{ \mathbf{Q} \in \mathbb{R}^{n\times k} \;\;:\;\; \mathbf{Q}^T\mathbf{Q} - \mathbf{I}_k = \mathcal{O}(\Delta t^{p+1}) \} \,.$$

when a scheme of $\mathcal{O}(\Delta t^p)$ is used in the first step.

References

[1] U.M. Ascher, H. Chin & S. Reich [1994] *Stabilization of DAEs and invariant manifolds*, Num. Math. **67**(2): 131–149.

[2] G. Benettin, L. Galgani, A. Giorgilli & J.M. Strelcyn [1980] *Lyapunov exponents for smooth dynamical systems and for Hamiltonian systems; a method for computing all of them. Part 1: theory and Part 2: numerical applcations*, Meccanica **15**: 9–20, 21–30.

[3] T.J. Bridges [1999] *The Orr-Sommerfeld equation on a manifold*, Proc. Roy. Soc. Lond. A **455**: 3019–3040.

[4] F. Christiansen & H.H. Rugh [1997] *Computing Lyapunov spectra with continuous Gram-Schmidt orthonormalization*, Nonlinearity **10**: 1063–1072.

[5] A. Davey [1983] *An automatic orthonormalization method for solving stiff boundary-value problems*, J. Comp. Phys. **51**: 343–356.

[6] L. Dieci, R.D. Russell & E.S. van Vleck [1994] *Unitary integrators and applications to continuous orthonormalization techniques*, SIAM J. Numer. Anal. **31**: 261–281.

[7] L. Dieci, R.D. Russell & E.S. van Vleck [1997] *On the computation of Lyapunov exponents for continuous dynamical systems*, SIAM J Numer. Anal. **34**: 402–423.

[8] L. Dieci & E.S. van Vleck [1995] *Computation of a few Lyapunov exponents for continuous and discrete dynamical systems*, Appl. Num. Math. **17**: 275–291.

[9] L. Dieci & E.S. van Vleck [1999] *Computation of orthonormal factors for fundamental solution matrices*, Num. Math. **83**: 599–620.

[10] S.P. Diliberto [1950] *On systems of ordinary differential equations*, in *Contributions to the Theory of Nonlinear Oscillations*, Ann. Math. Stud. **20**, Princeton University Press.

[11] U. Dressler [1988] *Symmetry property of Lyapunov spectra of a class of dissipative dynamical systems with viscous damping*, Phys. Rev. A **38**(4): 2103–2109.

[12] A. Edelman, T. Arias & S.T. Smith [1998] *The geometry of algorithms with orthogonality constraints*, SIAM J Matrix Anal. Appl. **20**: 303–353.

[13] L. Eldén & H. Park [1999] *A Procrustes Problem on the Stiefel Manifold*, Num. Math. **82**: 599–619.

[14] K. Geist, U. Parlitz & W. Lauterborn [1990] *Comparison of different methods for computing Lyapunov exponents*, Prog. Theor. Phys. **83**: 875–893.

[15] J.M. Greene & J.-S. Kim [1987] *The calculation of Lyapunov spectra*, PhysicaD **24**: 213–225.

[16] J.M. Greene & J.-S. Kim [1989] *Introduction of a metric tensor into linearized evolution equations*, PhysicaD **36**: 83–91.

[17] I. Goldhirsch, P.-L. Sulem & S.A. Orszag [1987] *Stability and Lyapunov stability of dynamical systems: a differential approach and a numerical method*, PhysicaD **27**: 311–337.

[18] D. Higham [1997] *Time-stepping and preserving orthonormality*, BIT **37**: 24–36.

[19] K. Kawakubo [1991] *The Theory of Transformation Groups*, Oxford University Press.

[20] S. Kobayashi & K. Nomizu [1969] *Foundations of Differential Geometry, Volume II*, John Wiley & Sons.

[21] B. Leimkuhler & S. Reich [1994] *Symplectic methods for constrained Hamiltonian systems*, Math. Comp. **63**, pp. 589–605.

[22] G. H. Meyer [1986] *Continuous orthonormalization for boundary-value problems*, J. Comp. Phys. **82**, pp. 248–262.

[23] H. Munthe-Kaas [1998] *Runge-Kutta methods on Lie groups*, BIT **38**, pp. 92–111.

[24] H. Munthe-Kaas [1999] *High-order Runge-Kutta methods on manifolds*, Appl. Num. Math. **29**, pp. 115–127.

[25] H. Munthe-Kaas & A. Zanna [1997] *Numerical integration of differential equations on homogeneous manifolds*, In F. Cucker & M. Shub, editors, *Foundations of Computational Mathematics*, published by Springer-Verlag, pp. 305–315.

[26] K. Ramasubramanian & M.S. Sriram [2000] *A comparative study of computation of Lyapunov spectra with different algorithms*, PhysicaD **139**: 72–86.

[27] G. Rangarajan, S. Habib & R.D. Ryne [1998] *Lyapunov exponents without rescaling and reorthogonalization* Phys. Rev. Lett. **80**(17): 3747–3750.

[28] N. Steenrod [1951] *The topology of fibre bundles*, Princeton University Press.

[29] E. Stiefel [1935] *Richtungsfelder und Fernparallelismus in $n-$dimensionalen Mannigfaltigkeiten*, Comm. Math. Helv. **8**, pp. 305–353.