



# Numerical implementation of complex orthogonalization, parallel transport on Stiefel bundles, and analyticity

Daniele Avitabile<sup>a,\*</sup>, Thomas J. Bridges<sup>b</sup>

<sup>a</sup> Department of Engineering Mathematics, University of Bristol, Bristol BS8 1TR, England, United Kingdom

<sup>b</sup> Department of Mathematics, University of Surrey, Guildford, Surrey GU2 7XH, England, United Kingdom

## ARTICLE INFO

### Article history:

Received 28 January 2009

Received in revised form

16 February 2010

Accepted 24 February 2010

Available online 6 March 2010

Communicated by J. Bronski

### Keywords:

Orthonormalization

Stiefel bundles

Constraints

Geometric integration

Poincaré index

Newton's method

## ABSTRACT

Numerical integration of complex linear systems of ODEs depending analytically on an eigenvalue parameter are considered. Complex orthogonalization, which is required to stabilize the numerical integration, results in non-analytic systems. It is shown that properties of eigenvalues are still efficiently recoverable by extracting information from a non-analytic characteristic function. The orthonormal systems are constructed using the geometry of Stiefel bundles. Different forms of continuous orthogonalization in the literature are shown to correspond to different choices of connection one-form on the Stiefel bundle. For the numerical integration, Gauss–Legendre Runge–Kutta algorithms are the principal choice for preserving orthogonality, and performance results are shown for a range of GLRK methods. The theory and methods are tested by application to example boundary value problems including the Orr–Sommerfeld equation in hydrodynamic stability.

© 2010 Elsevier B.V. All rights reserved.

## 1. Introduction

Linear systems of ordinary differential equations (ODEs) depending analytically on a parameter  $\lambda \in \mathbb{C}$  arise in a number of applications. They are typically of the form

$$\mathbf{y}_x = \mathbf{A}(x, \lambda)\mathbf{y}, \quad \mathbf{y} \in \mathbb{C}^n, \quad a \leq x \leq b \quad (1)$$

where  $\mathbf{A}(x, \lambda)$  is a given  $n \times n$  complex matrix depending smoothly on  $x$  and analytically on  $\lambda$ , for all  $\lambda \in \Lambda$  where  $\Lambda \subset \mathbb{C}$  is an open simply-connected subset. This system is accompanied by homogeneous boundary conditions

$$\mathbf{B}(\lambda)\mathbf{y}(a) = \mathbf{0}_{n-k}, \quad \mathbf{C}(\lambda)\mathbf{y}(b) = \mathbf{0}_k, \quad (2)$$

where  $\mathbf{B}(\lambda)$  is a  $(n-k) \times n$  matrix of rank  $(n-k)$  and  $\mathbf{C}(\lambda)$  is a  $k \times n$  matrix of rank  $k$ , with  $1 < k < n$ . Both  $\mathbf{B}(\lambda)$  and  $\mathbf{C}(\lambda)$  are analytic for all  $\lambda \in \Lambda$ . Values of  $\lambda$  for which this system has non-trivial solutions are eigenvalues. Analytic boundary-value problems (BVPs) of this type arise in hydrodynamic stability (e.g. [1,2]), in the linear stability of solitary waves and fronts (e.g. [3–5]), and in a variety of other applications (e.g. [6,7]).

Theoretically, the construction of solutions of this problem is straightforward. Let  $\mathbf{Y}_a(\lambda)$  be any  $n \times k$  matrix, depending

analytically on  $\lambda$ , whose columns satisfy the boundary condition at  $x = a$ . Let  $\mathbf{Y}(x, \lambda)$  be the  $n \times k$  matrix whose columns are the solutions of the ODE (1) satisfying  $\mathbf{Y}(a, \lambda) = \mathbf{Y}_a(\lambda)$ . A value  $\lambda \in \Lambda$  is an eigenvalue if  $\mathbf{Y}(b, \lambda)$  satisfies the boundary condition at  $x = b$ ; that is if the analytic function  $\Delta(\lambda)$  vanishes, where

$$\Delta(\lambda) = \det[\mathbf{C}(\lambda)\mathbf{Y}(b, \lambda)]. \quad (3)$$

In the context of boundary-value problems this function is typically called the “characteristic function”, and in the context of the stability of solitary waves this function is called the Evans function. Since  $\Delta(\lambda)$  is analytic, properties of its zeros can be studied using Cauchy's theorem, and there is a wide range of other results from analytic function theory that can be appealed to.

For most problems in hydrodynamic stability, and many problems in the stability of solitary waves, little progress can be made analytically and numerical methods are needed to solve the BVP. There are three distinct problems that arise in the numerical construction of solutions. The first is how to construct analytic starting vectors (columns of  $\mathbf{Y}_a(\lambda)$ ). That such a basis can always be chosen *analytically* follows from the Gohberg–Rodman theorem.<sup>1</sup> However, the numerical construction of such an analytic basis can

\* Corresponding author.

E-mail address: [d.avitabile@bristol.ac.uk](mailto:d.avitabile@bristol.ac.uk) (D. Avitabile).

<sup>1</sup> The Gohberg–Rodman theorem states there exists a uniformly analytic basis for the kernel of a matrix which depends analytically on a parameter; see Theorem S6.1 on page 388 of [8].

be non-trivial. A constructive theorem of Kato (see Section II.4.2 of [9] and Section 2 of [10]) can be used to compute the basis, and such a strategy for computing the columns of  $Y_a(\lambda)$  is proposed in [10].

In this paper we will restrict attention to the second and third numerical difficulties associated with the BVP: the integration from  $x = a$  to  $x = b$  and the numerical study of the zeros of the analytic function  $\Delta(\lambda)$ . Since analyticity of starting vectors is not considered, we can simplify matters by assuming henceforth that the boundary conditions do not depend on  $\lambda$

$$\mathbf{B}\mathbf{y}(a) = \mathbf{0}_{n-k}, \quad \mathbf{C}\mathbf{y}(b) = \mathbf{0}_k, \quad (4)$$

where  $\mathbf{B}$  is a constant  $(n - k) \times n$  matrix of rank  $(n - k)$  and  $\mathbf{C}$  is a constant  $k \times n$  matrix of rank  $k$ .

The initial-value problem to be integrated numerically is

$$\mathbf{Y}_x = \mathbf{A}(x, \lambda)\mathbf{Y}, \quad \mathbf{Y} \in \mathbb{C}^{n \times k}, \quad \text{with } \mathbf{Y}(a, \lambda) \in \text{Ker}(\mathbf{B}). \quad (5)$$

In the problems of interest in hydrodynamic stability and stability of waves, numerical integration of this problem is notoriously unstable: it is difficult to maintain linear independence of the columns of  $\mathbf{Y}(x, \lambda)$  numerically. This problem is overcome by using the freedom in the choice of basis for the columns of  $\mathbf{Y}(x, \lambda)$ . Let  $\mathbf{Y}_a$  be any fixed  $n \times k$  matrix of rank  $k$  in the kernel of  $\mathbf{B}$ . Then

$$\text{Ker}(\mathbf{B}) = \{\tilde{\mathbf{Y}}_a \in \mathbb{C}^{n \times k} : \tilde{\mathbf{Y}}_a = \mathbf{Y}_a \mathbf{m}, \mathbf{m} \in \mathbb{C}^{k \times k} \text{ with } \det(\mathbf{m}) \neq 0\}.$$

We are free to choose any basis for  $\text{Ker}(\mathbf{B})$ , and since the ODE is linear, we can choose any basis for the solution space as a function of  $x$  and  $\lambda$  as well. Let

$$\mathbf{Y}(x, \lambda) = \mathbf{V}(x, \lambda)\mathbf{m}(x, \lambda).$$

Instead of integrating (5) we integrate

$$\begin{aligned} \mathbf{V}_x &= \mathbf{A}(x, \lambda)\mathbf{V} - \mathbf{V}\mathbf{g}, \quad \text{with } \mathbf{V}(a, \lambda) = \mathbf{Y}_a \\ \mathbf{m}_x &= \mathbf{g}\mathbf{m}, \quad \text{with } \mathbf{m}(a, \lambda) = \mathbf{m}_a, \end{aligned} \quad (6)$$

where  $\mathbf{m}_a$  is any invertible  $k \times k$  complex matrix. We are free to choose the  $k \times k$  matrix  $\mathbf{g}(x, \lambda)$ . The numerical strategy is to choose  $\mathbf{g}(x, \lambda)$  so that the integration is stable and robust. The choice of  $\mathbf{g}$  does not affect the zeros of the complex function since

$$\begin{aligned} \Delta(\lambda) &= \det[\mathbf{C}\mathbf{Y}(b, \lambda)] = \det[\mathbf{C}\mathbf{V}(b, \lambda)\mathbf{m}(b, \lambda)] \\ &= \det[\mathbf{m}(b, \lambda)] \det[\mathbf{C}\mathbf{V}(b, \lambda)], \end{aligned}$$

and  $\det[\mathbf{m}(b, \lambda)] \neq 0$  because  $\det(\mathbf{m}_a) \neq 0$  by construction, and the determinant of  $\mathbf{m}(x, \lambda)$  cannot vanish, due to the Abel–Liouville Theorem for ODEs.

In *continuous orthogonalization* the function  $\mathbf{g}$  is chosen so that the columns of  $\mathbf{V}$  are orthonormal. The condition for orthonormality of the columns of  $\mathbf{V}$  is  $\mathbf{V}^H \mathbf{V} = \mathbf{I}_k$  where the superscript  $H$  denotes the complex conjugate transpose. The manifold defined by this condition, the *Stiefel manifold*, is a compact submanifold of  $\mathbb{C}^{n \times k}$ . Hence, if orthogonalization can be preserved numerically, the trajectories of the differential equation for  $\mathbf{V}(x, \cdot)$  are bounded. However, there is no unique choice for  $\mathbf{g}$  that results in orthogonalization. Indeed, we will show that the choice

$$\mathbf{g} = \mathbf{V}^H \mathbf{A} \mathbf{V} - \mathbf{S}(x, \mathbf{V}),$$

where  $\mathbf{S}(x, \mathbf{V})$  is any skew-Hermitian matrix depending smoothly on its arguments, generates a path on the Stiefel manifold. Hence there is a continuum-infinite number of choices of  $\mathbf{g}$  that give an orthonormal path. There are at least three choices in the literature: choose  $\mathbf{S}$  associated with Gram–Schmidt orthogonalization [11–14], choose  $\mathbf{S} = 0$  [15,16,5], and choose  $\mathbf{S}$  using orthogonal projection from  $\mathbb{C}^{n \times k}$  onto the tangent space of the Stiefel manifold [2,17]. However, the full range of this non-uniqueness has yet to be exploited. Other choices of  $\mathbf{g}$  may optimize other aspects of the numerics. One can even work directly on the Grassmannian. This involves choosing an atlas for the Grassmannian and chart

switching and leads to matrix Riccati equations. An efficient algorithm for working directly on the Grassmannian has recently been proposed [18].

In this paper the differential geometry of *Stiefel bundles* will be used as a basis for constructing the numerical algorithm. The bundle structure of Stiefel manifolds is developed in Section 2. It is a generalization of ideas in [2,19,17,20] where the differential geometry of Stiefel manifolds was used.

There is however another problem with *complex orthogonalization*: it does not preserve analyticity. The function  $\Delta(\lambda)$  will not in general be analytic when  $\mathbf{V}$  is orthonormal. For example, taking  $k = 1$  and the vector-valued analytic function  $\mathbf{u} : \Lambda \rightarrow \mathbb{C}^n$ , the QR-decomposition is just

$$\mathbf{u}(\lambda) = \frac{\mathbf{u}(\lambda)}{\|\mathbf{u}(\lambda)\|} \|\mathbf{u}(\lambda)\| := \mathbf{Q}(\lambda, \bar{\lambda})\mathbf{R}(\lambda, \bar{\lambda}),$$

and the norm is not analytic since  $\|\mathbf{u}(\lambda)\|^2 = \overline{\mathbf{u}(\lambda)}^T \mathbf{u}(\lambda)$  and so it depends on both  $\lambda$  and  $\bar{\lambda}$ . This loss of analyticity does not affect the numerical integration of the ODE from  $x = a$  to  $x = b$ , but it introduces a non-analytic component into  $\Delta(\lambda)$ . Suppose a QR-decomposition of  $\mathbf{Y}(x, \lambda)$  is used. Then

$$\mathbf{Y}(x, \lambda) = \mathbf{Q}(x, \lambda, \bar{\lambda})\mathbf{R}(x, \lambda, \bar{\lambda}),$$

and so

$$\Delta(\lambda) = \det[\mathbf{C}\mathbf{Q}(b, \lambda, \bar{\lambda})\mathbf{R}(b, \lambda, \bar{\lambda})] = q(\lambda, \bar{\lambda})r(\lambda, \bar{\lambda}),$$

where  $r(\lambda, \bar{\lambda}) = \det[\mathbf{R}(b, \lambda, \bar{\lambda})]$ . The analytic function  $\Delta(\lambda)$  is a product of two non-analytic functions.

This problem of non-analyticity was first studied by Humpherys and Zumbrun [5]. Their approach is to solve for *both*  $\mathbf{Q}(x, \lambda, \bar{\lambda})$  and the determinant of  $\mathbf{R}(x, \lambda, \bar{\lambda})$  and then the analytic function  $\Delta(\lambda)$  can be recovered in full. The approach proposed in this paper is complementary: the idea is to see how much information can be extracted from  $q(\lambda, \bar{\lambda})$  without computing  $\mathbf{R}(x, \lambda, \bar{\lambda})$ .

By construction  $\det[\mathbf{R}(x, \lambda, \bar{\lambda})] \neq 0$  and so  $r(\lambda, \bar{\lambda}) \neq 0$ . Therefore  $\Delta(\lambda) = 0$  if and only if  $q(\lambda, \bar{\lambda}) = 0$ . In principle it is sufficient for finding eigenvalues to compute only  $q$  and discard  $r$ . However,  $q$  is not analytic. On the other hand, it has an interesting property which makes it almost analytic: we prove that the index of  $q(\lambda, \bar{\lambda})$ , considered as a vectorfield on the plane, equals the number of zeros of  $q$  in that region, including multiplicity. Therefore there is an analogue of Cauchy’s theorem for  $q(\lambda, \bar{\lambda})$ .

The use of Newton’s method to find zeros of  $q(\lambda, \bar{\lambda})$  has to be refined. The function  $q$  has to be considered as a function of  $\lambda$  and  $\bar{\lambda}$ . Even here there are surprises. Near zeros of  $q(\lambda, \bar{\lambda})$  it behaves like an analytic function, and so Newton iteration assuming analyticity still converges when the initial condition is sufficiently close to a zero. We present performance results for Newton’s method with and without the hypothesis of analyticity.

Although the induced system on the Stiefel manifold preserves orthogonality exactly, numerical methods have to be chosen carefully to respect this property. In this paper we test a range of Runge–Kutta algorithms, principally the Gauss–Legendre Runge–Kutta (GLRK) algorithms. It is often stated in the literature that GLRK methods preserve the Stiefel manifold (indeed any quadratic invariant) to machine accuracy. However, this requires iterating to convergence at each step, a result first observed in [13]. Our results show that preservation of the Stiefel manifold is intimately connected with the tolerance of this iteration. We present performance results for 4th, 6th and 8th order implicit GLRK methods, and compare them with explicit RK as well as explicit RK with stabilization. There are other geometric integration methods that preserve the Stiefel manifold and are competitive. The principal example is the class of Munthe–Kaas RK methods, and performance results for these methods are reported in [20–22]. We do not compare speed, as the main interest in this paper is how well the Stiefel manifold is preserved, and the impact on analyticity when complex orthogonalization is used.

## 2. Differential geometry of Stiefel bundles

For each  $\lambda \in \Lambda$ , solutions of (5) are paths  $\mathbf{Y}(x, \lambda)$  in  $\mathbb{C}^{n \times k}$ , which is a linear vector space. However, solution paths  $\mathbf{Y}(x, \lambda)$  are required to have rank  $k$  for all  $x$ , and this brings in a manifold structure.

Take the standard Hermitian metric on  $\mathbb{C}^{n \times k}$ ,

$$\langle \mathbf{U}, \mathbf{V} \rangle := \text{Tr}(\mathbf{U}^H \mathbf{V}), \quad \text{with norm } \|\mathbf{U}\| := \sqrt{\text{Tr}(\mathbf{U}^H \mathbf{U})}.$$

Then the Stiefel manifold

$$V_k(\mathbb{C}^n) = \{\mathbf{Q} \in \mathbb{C}^{n \times k} : \mathbf{Q}^H \mathbf{Q} = \mathbf{I}_k\} \quad (7)$$

can be viewed as the image of a projection [23]

$$\Pi(\mathbf{Y}) = \arg \min_{\mathbf{Q} \in V_k(\mathbb{C}^n)} \|\mathbf{Y} - \mathbf{Q}\|^2, \quad \mathbf{Y} \in \mathbb{C}^{n \times k}.$$

In the case  $k = 1$  this projection reduces to

$$\Pi(\mathbf{Y}) = \frac{\mathbf{Y}}{\|\mathbf{Y}\|}, \quad \mathbf{Y} \in \mathbb{C}^n \setminus \{0\}.$$

The image of  $\Pi : \mathbb{C}^{n \times k} \rightarrow V_k(\mathbb{C}^n)$  is the point  $\mathbf{Q} \in V_k(\mathbb{C}^n)$  which minimizes the distance from  $\mathbf{Y}$  to  $V_k(\mathbb{C}^n)$ . The differential

$$d\Pi_{\mathbf{Y}} : T_{\mathbf{Y}}\mathbb{C}^{n \times k} \rightarrow T_{\mathbf{Q}}V_k(\mathbb{C}^n),$$

provides an easy way to find the tangent space of  $V_k(\mathbb{C}^n)$ . The tangent space at each point is a subspace of  $\mathbb{C}^{n \times k}$  of the form (see pages 639–640 of [23])

$$T_{\mathbf{Q}}V_k(\mathbb{C}^n) = \{\mathbf{F} \in \mathbb{C}^{n \times k} : \mathbf{F} = \mathbf{Q}^\perp \mathbf{B} + \mathbf{Q}\mathbf{C}\},$$

where  $\mathbf{C}$  is skew-Hermitian and  $\mathbf{B}$  is an  $(n - k) \times k$  matrix. The dimension of the Stiefel manifold is then the dimension of skew-Hermitian matrices plus the dimension of  $n \times (n - k)$  matrices

$$\dim V_k(\mathbb{C}^n) = k(n - k) + \frac{1}{2}k(k - 1) \quad (\text{complex dimension}).$$

By a dimension count, the normal space to  $T_{\mathbf{Q}}V_k(\mathbb{C}^n)$  in  $\mathbb{C}^{n \times k}$  has dimension  $\frac{1}{2}k(k + 1)$ . Using the Hermitian inner product on  $\mathbb{C}^{n \times k}$ , the normal space is

$$N_{\mathbf{Q}}V_k(\mathbb{C}^n) = \{\mathbf{W} \in \mathbb{C}^{n \times k} : \mathbf{W} = \mathbf{Q}\mathbf{D}\},$$

with  $\mathbf{D}$  a  $k \times k$  Hermitian-symmetric matrix.

When  $\mathbf{Q} \in V_k(\mathbb{C}^n)$  then  $\mathbf{g}$  in (6) is chosen so that the right-hand side of (6) is in the tangent space,

$$\mathbf{Q}_x = \mathbf{A}(x, \lambda)\mathbf{Q} - \mathbf{Q}\mathbf{g} \in T_{\mathbf{Q}}V_k(\mathbb{C}^n). \quad (8)$$

Using the above characterization of the tangent space gives

$$\mathbf{Q}_x = (\mathbf{I} - \mathbf{Q}\mathbf{Q}^H)\mathbf{A}\mathbf{Q} + \mathbf{Q}\text{skew}_{\mathbb{C}}(\mathbf{Q}^H\mathbf{A}\mathbf{Q}), \quad (9)$$

where  $\text{skew}_{\mathbb{C}}(\mathbf{F}) := \frac{1}{2}(\mathbf{F} - \mathbf{F}^H)$ . Hence

$$\mathbf{g} = \text{sym}_{\mathbb{C}}(\mathbf{Q}^H\mathbf{A}\mathbf{Q}), \quad \text{sym}_{\mathbb{C}}(\mathbf{F}) := \frac{1}{2}(\mathbf{F} + \mathbf{F}^H).$$

### 2.1. Stiefel manifold as a principal fiber bundle

Stiefel manifolds can also be characterized as fiber bundles. The total space is  $V_k(\mathbb{C}^n)$ , the base manifold is the Grassmannian  $G_k(\mathbb{C}^n)$  and the fiber is the unitary group  $U(k)$ ,

$$\begin{array}{ccc} U(k) & \rightarrow & V_k(\mathbb{C}^n) \\ & & \downarrow \\ & & G_k(\mathbb{C}^n). \end{array}$$

Indeed, the Stiefel bundle is a principal fiber bundle since  $U(k)$  acts freely on  $V_k(\mathbb{C}^n)$  and the fiber is a group (cf. Section 5.2 of [24]). The unitary group is defined by

$$U(k) = \{\mathbf{m} \in \mathbb{C}^{k \times k} : \mathbf{m}^H \mathbf{m} = \mathbf{I}_k\}.$$

It is a Lie group of real dimension  $k^2$ . The Lie algebra of  $U(k)$ , denoted  $u(k)$ , consists of complex  $k \times k$  skew-Hermitian matrices with the Lie bracket given by the commutator. The unitary group is non-abelian for  $k > 1$ . The determinant of a unitary matrix lies on the unit circle

$$\det : U(k) \rightarrow U(1) \cong S^1.$$

The action of  $U(k)$  on  $V_k(\mathbb{C}^n)$  is by a right action

$$\Phi_{\mathbf{m}}(\mathbf{Q}) = \mathbf{Q}\mathbf{m}, \quad \text{for any } \mathbf{Q} \in V_k(\mathbb{C}^n), \text{ and } \mathbf{m} \in U(k).$$

Some properties of the fiber bundle structure of  $V_k(\mathbb{C}^n)$  have been used in other contexts (e.g. in numerical linear algebra [19], in the computation of Lyapunov exponents [17,20]). Here further properties of the geometry of the bundle formulation are used, in particular our main new observation is the role of the ‘‘universal connection’’ on the Stiefel bundle in the construction of induced differential equations for numerical integration.

To lighten the notation, let

$$P := V_k(\mathbb{C}^n), \quad M := G_k(\mathbb{C}^n), \quad G := U(k) \quad \text{and} \quad \mathfrak{g} := u(k).$$

The principal fiber bundle is denoted by  $(P, \pi, M)$  where  $\pi : P \rightarrow M$  is the natural projection which is a smooth submersion.

The tangent space of  $P$  has a natural subspace at each point called the *vertical subspace*

$$V_p P = \{v \in T_p P : d_p \pi(v) = 0\},$$

and  $\cup_{p \in P} V_p P$  is a distribution on  $P$ . There is no unique complement. However, the existence of a complement, called the *horizontal subspace*  $H_p P$  at each  $p \in P$ , gives a splitting of the tangent space

$$T_p P = H_p P \oplus V_p P, \quad \text{for each } p \in P. \quad (10)$$

When this distribution of horizontal subspaces is also equivariant, it is called a *connection* on the principal fiber bundle (cf. Section 6.1 of [24], or [25,26]). A precise definition of the horizontal subspace in terms of a connection one-form will be given below.

This splitting in (10) is useful as it gives a precise way to decompose the right-hand side of (8). Indeed, the different choices of orthogonalization in the literature correspond to different choices of horizontal distribution.

For the case of the Stiefel bundle, the vertical subspace is

$$V_{\mathbf{Q}}V_k(\mathbb{C}^n) = \{\mathbf{V} \in T_{\mathbf{Q}}V_k(\mathbb{C}^n) : \mathbf{V} = \mathbf{Q}\mathbf{n}, \mathbf{n} \in u(k)\}.$$

### 2.2. Connection one-forms on the Stiefel bundle

A connection one-form on a principal fiber bundle is a  $\mathfrak{g}$ -valued one-form  $\Omega_p(v)$ . For each  $p \in P$ , it is a mapping from  $T_p P$  to  $\mathfrak{g}$  satisfying [24–26]

- $\Omega_p(\xi_p) = \xi$  for all  $p \in P, \xi \in \mathfrak{g}$ ,
- $\Omega_{\Phi_g(p)}(T_p \Phi_g v_p) = \text{Ad}_g \Omega(v_p)$  for all  $p \in P, g \in G$ .

In this definition,

$$\xi_p := \left. \frac{d}{ds} \right|_{s=0} \Phi_{g(s)}(p),$$

$$T_p \Phi_g := \left. \frac{d}{ds} \right|_{s=0} \Phi_g(p(s)),$$

$$\text{Ad}_g \Omega_p(v_p) := g^{-1} \Omega(v_p) g.$$

The connection one-form is useful because of the following property

$$H_p P = \{v_p \in T_p P : \Omega_p(v_p) = 0\}.$$

The horizontal distribution in the splitting (10) is determined by the kernel of  $\Omega_p$ . There is a one-to-one correspondence between the choice of horizontal distribution and a connection one-form [24].

On the Stiefel bundle there is a particularly important connection one-form: the *universal connection* [27]. Reverting back to Stiefel manifold notation, the universal connection on the Stiefel manifold is

$$\Omega_Q(\mathbf{F}) = \mathbf{Q}^H \mathbf{F}, \quad \text{for } \mathbf{Q} \in V_k(\mathbb{C}^n), \mathbf{F} \in T_{\mathbf{Q}}V_k(\mathbb{C}^n).$$

Let's verify that  $\Omega_Q$  is indeed a connection one form. Any element in  $T_{\mathbf{Q}}V_k(\mathbb{C}^n)$  is of the form

$$\mathbf{F} = \mathbf{Q}^\perp \mathbf{B} + \mathbf{Q} \mathbf{C}, \tag{11}$$

where  $\mathbf{C}$  is skew-Hermitian and so  $\mathbf{C} \in u(k)$ . Now

$$\mathbf{Q}^H \mathbf{F} = \mathbf{Q}^H \mathbf{Q}^\perp \mathbf{B} + \mathbf{Q}^H \mathbf{Q} \mathbf{C} = \mathbf{C} \in u(k),$$

verifying that the image of  $\Omega_Q$  is in  $u(k)$ . Check the action on tangent vectors to the group orbit,

$$\left. \frac{d}{ds} \right|_{s=0} \Phi_{g(s)}(\mathbf{Q}) = \mathbf{Q} \xi, \quad \text{for some } \xi \in u(k),$$

and so

$$\Omega_Q(\xi_p) = \Omega_Q(\mathbf{Q} \xi) = \mathbf{Q}^H \mathbf{Q} \xi = \xi,$$

verifying the second property of a connection one-form.

Since  $U(k)$  is a matrix group  $T_p \Phi_g(p) = v_p g$  for a tangent vector  $v_p \in T_p P$ . Hence for any  $\mathbf{m} \in U(k)$

$$\begin{aligned} \Omega_{\Phi_m(\mathbf{Q})}(T_{\mathbf{Q}} \Phi_m \mathbf{F}) &= (\mathbf{Q} \mathbf{m})^H \mathbf{F} \mathbf{m} \\ &= \mathbf{m}^H \mathbf{Q}^H \mathbf{F} \mathbf{m} \\ &= \mathbf{m}^{-1} \Omega_{\mathbf{Q}}(\mathbf{F}) \mathbf{m} \\ &= \text{Ad}_g \Omega_{\mathbf{Q}}(\mathbf{F}). \end{aligned}$$

This computation verifies that  $\Omega_Q$  is a connection one-form on the Stiefel bundle. The kernel of  $\Omega_Q$  is a horizontal subspace of the tangent space of the Stiefel bundle at each point.

Apply this result to (8). Choose  $\mathbf{g}$  so that the right-hand side is horizontal

$$\begin{aligned} 0 &= \Omega_Q(\mathbf{A} \mathbf{Q} - \mathbf{Q} \mathbf{g}) \\ &= \mathbf{Q}^H (\mathbf{A} \mathbf{Q} - \mathbf{Q} \mathbf{g}) \\ &= \mathbf{Q}^H \mathbf{A} \mathbf{Q} - \mathbf{Q}^H \mathbf{Q} \mathbf{g} \\ &= \mathbf{Q}^H \mathbf{A} \mathbf{Q} - \mathbf{g}. \end{aligned}$$

Hence the right hand side of (8) is horizontal if  $\mathbf{g} = \mathbf{Q}^H \mathbf{A} \mathbf{Q}$ . Substituting this expression into the right-hand side of (8) gives

$$\mathbf{Q}_x = (\mathbf{I} - \mathbf{Q} \mathbf{Q}^H) \mathbf{A}(x, \lambda) \mathbf{Q}, \tag{12}$$

which is precisely the Drury–Davey algorithm [15,16]. There is an interesting consequence of this characterization: the flow (12) defines a *parallel transport* along the bundle. The trajectory travels from  $x = a$  to  $x = b$  with the minimum of rotation of the fibers. Parallel transport is the generalization to arbitrary bundles of the concept of a geodesic [26]. This system is in some sense optimal – *with respect to the universal connection*. This property of the Eq. (12) might explain the excellent behaviour of this algorithm found by Humpherys and Zumbun [5]. However, this system is not optimal in every sense. Differences between the orthonormalization schemes show up when the impact on analyticity is considered, and in the numerical implementation.

Parallel transport is defined as follows. Consider any curve  $\mathbf{c}(x)$  in  $G_k(\mathbb{C}^n)$  passing through some point  $\mathbf{c}(0) = \mathbf{c}_0 \in G_k(\mathbb{C}^n)$ . For any point  $\mathbf{Q} \in V_k(\mathbb{C}^n)$  in the fiber over  $\mathbf{c}_0$ , there is a unique curve  $\mathbf{c}^*(x)$  in  $V_k(\mathbb{C}^n)$ , call the *horizontal lift* of  $\mathbf{c}$ , which passes through  $\mathbf{Q}$ , projects to  $\mathbf{c}(x)$ , and is everywhere horizontal:  $\frac{d}{dx} \mathbf{c}^*(x) \in \text{Kernel}(\Omega_{\mathbf{Q}})$ . If  $\mathbf{c}(x_1) = \mathbf{c}_1$  and  $\mathbf{c}(x_2) = \mathbf{c}_2$ , each point  $\mathbf{Q}_1$  in the fiber over  $\mathbf{c}_1$  is connected by a unique horizontal lift of  $\mathbf{c}(x)$  to a point  $\mathbf{Q}_2$  in the fiber over  $\mathbf{c}_2$ . The map from  $\mathbf{Q}_1$  to  $\mathbf{Q}_2$  is called *parallel displacement* along  $\mathbf{c}(x)$ . Thus parallel displacement determines a map from  $\pi^{-1}(\mathbf{c}_1)$  to  $\pi^{-1}(\mathbf{c}_2)$ .

Horizontal lifting and parallel displacement depend on the choice of connection in the principal fiber bundle.

### 3. What is continuous orthogonalization?

Let  $\mathbf{Q}(x) \in \mathbb{C}^{n \times k}$  be a continuous path on the interval  $a \leq x \leq b$ . This path is a continuous orthogonal path if in addition

$$\mathbf{Q}(x)^H \mathbf{Q}(x) = \mathbf{I}_k, \quad \text{for all } x \in [a, b].$$

Now suppose  $\mathbf{Q}(x)$  is a  $C^1$  path. Then the path preserves orthogonality if

$$\frac{d}{dx} \mathbf{Q}(x)^H \mathbf{Q}(x) + \mathbf{Q}(x)^H \frac{d}{dx} \mathbf{Q}(x) = 0 \quad \forall x \in [a, b], \quad \text{and}$$

$$\mathbf{Q}(a)^H \mathbf{Q}(a) = \mathbf{I}_k.$$

We argued earlier that  $\mathbf{Q}_x$  should lie in the tangent space of  $V_k(\mathbb{C}^n)$ . However, this tangent space has many subspaces, and so orthogonalization can take many forms. Let  $\mathbf{S}(x, \mathbf{V})$  be any skew-symmetric  $k \times k$  matrix depending smoothly on its entries. Given a differential equation  $\mathbf{Q}_x = \mathbf{A} \mathbf{Q}$  with  $\mathbf{Q}(a) \in V_k(\mathbb{C}^n)$ , then the path  $\mathbf{Q}(x)$  satisfying the ODE is an orthonormal path if

$$\mathbf{Q}_x = (\mathbf{I}_n - \mathbf{Q} \mathbf{Q}^H) \mathbf{A}(x, \lambda) + \mathbf{Q} \mathbf{S}(x, \mathbf{Q}), \quad \text{with } \mathbf{Q}(a) \in V_k(\mathbb{C}^n),$$

for any choice of  $\mathbf{S}$ . Special cases include (a) Gram–Schmidt orthogonalization [11–14], (b) parallel transport ( $\mathbf{S} = 0$ ) [15,16,5], and orthogonal projection from  $\mathbb{C}^{n \times k}$  onto the tangent space of the Stiefel manifold ( $\mathbf{S} = \text{skew}_{\mathbb{C}}(\mathbf{Q}^H \mathbf{A} \mathbf{Q})$ ) [2,17].

Different choices of  $\mathbf{S}$  provide different horizontal paths through the Stiefel bundle. We conjecture there is a one-to-one correspondence between a choice of  $\mathbf{S}$  is continuous orthogonalization and a choice of connection one-form on the Stiefel bundle. A geometric view of continuous orthogonalization is then a choice of connection one-form for construction of the ODE on the tangent space of the Stiefel manifold.

There is a curiosity that arises when parallel transport is used, and it is related to the concept of *holonomy* in differential geometry. Take the vectorfield on the Stiefel bundle (9) as a starting point, and compare it with the parallel transport equation

$$\mathbf{W}_x = (\mathbf{I} - \mathbf{W} \mathbf{W}^H) \mathbf{A} \mathbf{W}.$$

Now let

$$\mathbf{W}(x) = \mathbf{Q}(x) \mathbf{m}(x),$$

where  $\mathbf{m}(x)$  is a path in  $U(k)$ . The matrix  $\mathbf{m}(x)$  will determine the precise amount of movement up or down the fiber to assure that  $\mathbf{W}(x)$  is parallel. Substituting this expression into (9) and using the equation for  $\mathbf{W}$  gives

$$\mathbf{m}_x = -[\text{skew}_{\mathbb{C}}(\mathbf{Q}^H \mathbf{A} \mathbf{Q})] \mathbf{m}.$$

The trace of a skew-Hermitian matrix is purely imaginary. Hence the determinant of  $\mathbf{m}$  gives a path on the unit circle. This is the so-called *geometric phase* of the path. It is an analogue of Berry's phase in quantum mechanics. A view of Berry's phase from the point of view of Stiefel bundles is given in [28]. Hence, given any path of the differential equation defined by parallel transport there is an attendant phase shift. These phases appear to have many interesting properties but will not be studied here. The recent Ph.D. thesis of Way [29] explores these phases in the context of numerical integration of boundary-value problems.

### 4. Analyticity and orthogonalization

Orthogonalization will not in general preserve analyticity. Analyticity is useful because Cauchy's Theorem can be used to determine the number of zeros of an analytic function in a region, it simplifies the use of Newton's method, and in general the many properties of analytic functions can be appealed to.

Suppose  $\mathbf{Q}$  is the induced path on  $V_k(\mathbb{C}^n)$  associated with the ODE (1). Then  $\mathbf{Y}(x, \lambda) = \mathbf{Q}(x, \lambda)\mathbf{R}(x, \lambda)$  and substitution into (9) shows that

$$\mathbf{R}_x = \text{sym}_{\mathbb{C}}(\mathbf{Q}^H \mathbf{A} \mathbf{Q})\mathbf{R}, \quad \mathbf{R}(a, \lambda) = \mathbf{I}_k \tag{13}$$

where  $\text{sym}_{\mathbb{C}}(\mathbf{F}) = \frac{1}{2}(\mathbf{F} + \mathbf{F}^H)$ . The important property here is that the coefficient matrix of the ODE (13) is Hermitian. Now look at the effect of this decomposition on  $\Delta(\lambda)$

$$\begin{aligned} \Delta(\lambda) &= \det[\mathbf{C}\mathbf{Y}(b, \lambda)] \\ &= \det[\mathbf{C}\mathbf{Q}(b, \lambda, \bar{\lambda})\mathbf{R}(b, \lambda, \bar{\lambda})] \\ &= \det[\mathbf{C}\mathbf{Q}(b, \lambda, \bar{\lambda})] \det[\mathbf{R}(b, \lambda, \bar{\lambda})] \\ &:= q(\lambda, \bar{\lambda})r(\lambda, \bar{\lambda}). \end{aligned}$$

But

$$\begin{aligned} \frac{d}{dx} \det[\mathbf{R}(x, \cdot)] &= \text{Trace} \left( \mathbf{R}(x, \cdot)^{-1} \frac{d}{dx} \mathbf{R}(x, \cdot) \right) \det[\mathbf{R}(x, \cdot)] \\ &= \text{Trace}(\text{sym}_{\mathbb{C}}(\mathbf{Q}^H \mathbf{A} \mathbf{Q})) \det[\mathbf{R}(x, \cdot)], \end{aligned}$$

and the trace of a hermitian matrix is real. Hence there exists a real-valued function  $\tau(\lambda, \bar{\lambda})$  such that  $r(\lambda, \bar{\lambda}) = \exp[\tau(\lambda, \bar{\lambda})]$ , and so

$$\Delta(\lambda) = q(\lambda, \bar{\lambda}) \exp[\tau(\lambda, \bar{\lambda})], \quad \tau \in \mathbb{R}. \tag{14}$$

Although the right-hand side is analytic, it is the product of two non-analytic functions, one of which is real and never zero. Hence  $\Delta(\lambda) = 0$  if and only if  $q(\lambda, \bar{\lambda}) = 0$ .

Contrast this result with the effect of using the parallel transport equation (12). Let  $\mathbf{Y}(x, \lambda) = \mathbf{W}(x, \lambda, \bar{\lambda})\mathbf{R}(x, \lambda, \bar{\lambda})$  and substitute into (12). The matrix  $\mathbf{R}(x, \lambda, \bar{\lambda})$  satisfies

$$\tilde{\mathbf{R}}_x = (\mathbf{W}^H \mathbf{A} \mathbf{W})\tilde{\mathbf{R}}, \quad \tilde{\mathbf{R}}(a, \lambda) = \mathbf{I}_k. \tag{15}$$

In this case the coefficient matrix  $(\mathbf{W}^H \mathbf{A} \mathbf{W})$  is not hermitian and so the  $q - \tilde{r}$  decomposition of  $\Delta(\lambda)$  is:

$$\begin{aligned} \Delta(\lambda) &= \det[\mathbf{C}^H \mathbf{Y}(b, \lambda)] = \det[\mathbf{C}^H \mathbf{Q}(b, \lambda, \bar{\lambda})] \det[\tilde{\mathbf{R}}(b, \lambda, \bar{\lambda})] \\ &:= q(\lambda, \bar{\lambda})\tilde{r}(\lambda, \bar{\lambda}), \end{aligned}$$

and in this case the term  $\tilde{r}(\lambda, \bar{\lambda})$  is complex valued.

#### 4.1. The index of the non-analytic function $q(\lambda, \bar{\lambda})$

When a function is analytic, Cauchy's Theorem gives a precise count of the zeros in a simply-connected region of the complex plane. When a function is not analytic the analogue is "the index" (sometimes called Poincaré index), a special case of the topological degree [30]. However, the index does not give a precise count of zeros in general, but only a sufficient condition for the existence of at least one zero. However, we will show that the function  $q(\lambda, \bar{\lambda})$  retains a key property of the analytic function  $\Delta(\lambda)$ : the index equals the number of zeros in the chosen set.

The index of a vectorfield  $\mathbf{u}(x, y) := (u(x, y), v(x, y))$  relative to a set in the  $(x, y)$ -plane is defined as follows. Let  $\mathcal{C}$  be a simple closed curve in the  $(x, y)$  plane with simply-connected interior  $\mathcal{D}$  and let  $(x(t), y(t)), 0 \leq t \leq 2\pi$  be a parametrization of  $\mathcal{C}$ . Then

$$\text{Ind}(\mathbf{u}; \mathcal{D}) = \frac{1}{2\pi} \int_{\mathcal{C}} \frac{u dv - v du}{u^2 + v^2} = \frac{1}{2\pi} \int_0^{2\pi} \frac{u \dot{v} - v \dot{u}}{u^2 + v^2} dt.$$

When  $u + iv$  is an analytic function, we can use Cauchy's Theorem and the index equals the number of zeros in the region  $\mathcal{D}$ .

Let  $q = u + iv$  and  $\lambda = x + iy$ , and separate  $\Delta$  into real and imaginary parts

$$\Delta_r := \text{Re}(\Delta) = e^\tau u \quad \text{and} \quad \Delta_i := \text{Im}(\Delta) = e^\tau v.$$

**Lemma.** Let  $\mathcal{C}$  be a Jordan curve in the  $\lambda$ -plane with interior  $\mathcal{D}$ , and suppose  $\Delta$  is analytic in  $\mathcal{D}$ . The Cauchy index of  $\Delta$  is equal to the Poincaré index of the vectorfield  $\mathbf{u} = (u, v)$

$$\text{Ind}(\Delta, \mathcal{D}) = \text{Ind}(\mathbf{q}, \mathcal{D}).$$

**Proof.** Cauchy's theory for the analytic function  $\Delta$  states that

$$\text{Ind}(\Delta, \mathcal{D}) = \frac{1}{2\pi i} \oint \frac{\Delta'(\lambda)}{\Delta(\lambda)} d\lambda.$$

Parameterizing the  $\lambda$  curve by  $t \in [0, 2\pi)$  and substituting the expressions for  $\Delta_r$  and  $\Delta_i$ , we find

$$\begin{aligned} \text{Ind}(\Delta, \mathcal{D}) &= \frac{1}{2\pi} \int_0^{2\pi} \frac{\Delta_r \dot{\Delta}_i - \Delta_i \dot{\Delta}_r}{\Delta_r^2 + \Delta_i^2} dt \\ &= \frac{1}{2\pi} \int_0^{2\pi} \frac{ue^\tau(\dot{t}ve^\tau + \dot{v}e^\tau) - ve^\tau(\dot{t}ue^\tau + \dot{u}e^\tau)}{e^{2\tau}(u^2 + v^2)} dt \\ &= \frac{1}{2\pi} \int_0^{2\pi} \frac{u\dot{v} - v\dot{u}}{u^2 + v^2} dt = \text{Ind}(\mathbf{q}, \mathcal{D}). \quad \square \end{aligned}$$

Hence the Poincaré index of the non-analytic function  $\mathbf{q}$  is equal – including multiplicity – to the number of zeros of  $\Delta(\lambda)$  in the region.

Computing the index is equivalent to computing the topological degree of  $q$ . Several algorithms have been proposed in the literature for computing the topological degree (e.g. [31]). In Section 8 a simple algorithm based on the trapezoidal rule and numerical differentiation is proposed.

### 5. Example boundary value problems

Three examples will be used for illustration of the theory. First a two-dimensional example. Consider

$$\phi_{xx} + \lambda\phi = 0, \quad 0 < x < \pi, \lambda \in \mathbb{C}, \tag{16}$$

where  $\phi(x, \lambda)$  is a complex-valued function, with boundary conditions

$$\phi(0, \lambda) = 0 \quad \text{and} \quad \phi_x(\pi, \lambda) = 0.$$

The exact eigenvalues are

$$\lambda_k = \left(k + \frac{1}{2}\right)^2, \quad k = 0, 1, 2, \dots,$$

with eigenfunctions

$$\phi_n(k, \lambda_k) = a_k \sin(k + 1/2)x,$$

where  $a_k$  is an arbitrary constant.

The second is a four-dimensional example proposed by Ng and Reid [32],

$$\phi_{xxxx} - \lambda^4\phi = 0, \quad 0 < x < 1, \lambda \in \mathbb{C}. \tag{17}$$

Here  $\phi(x, \lambda)$  is a complex-valued function, with boundary conditions at  $x = 0$

$$\phi(0, \lambda) = \phi_x(0, \lambda) = 0,$$

and boundary conditions at  $x = 1$ ,

$$\phi(1, \lambda) = \phi_x(1, \lambda) = 0.$$

The exact solution is given in [32]. The eigenvalues are all real and satisfy

$$\cos \lambda = \frac{1}{\cosh \lambda}.$$

The lowest eigenvalue is  $\lambda_1 \approx 4.73$ .

The third example is the Orr–Sommerfeld equation from hydrodynamic stability theory. It is a notoriously stiff system when the Reynolds number  $R$  is large, so it provides a good test case for complex orthogonalization. Indeed, it is the equation that inspired much of the early work on complex orthogonalization [15,16,2].

The Orr–Sommerfeld equation can be expressed in the form

$$\mathbf{y}_x = \mathbf{A}(x, \lambda)\mathbf{y}, \quad \mathbf{y} \in \mathbb{C}^4, \quad a \leq x \leq b \tag{18}$$

with homogeneous boundary conditions

$$\begin{aligned} \mathbf{B}\mathbf{y}(a, \lambda) = \mathbf{0}_2 \quad \text{and} \quad \mathbf{C}\mathbf{y}(b, \lambda) = \mathbf{0}_2, \quad \text{with} \\ \mathbf{B} = \mathbf{C} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \end{aligned} \tag{19}$$

The matrix  $\mathbf{A}$  depends smoothly on  $x$  and  $\lambda$ ,

$$\mathbf{A}(x, \lambda) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \alpha^2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -i\alpha R U''(x) & 0 & \gamma(x, \lambda) & 0 \end{bmatrix}, \tag{20}$$

$$\gamma(x, \lambda) = \alpha^2 + i\alpha R U(x) + \lambda R.$$

Eq. (18) arises for example when investigating the linear stability of the streamwise velocity profile  $U(x)$  in a plane channel,  $x$  being the direction normal to the channel walls. The explicit form for  $U(x)$  for plane channel flow is

$$U(x) = \frac{4U_0}{(a-b)^2} (-ab + (a+b)x - x^2), \quad a \leq x \leq b,$$

where  $U_0$  is a constant, representing the maximum velocity at the channel centre. In the numerics,  $U_0$  is normalized to unity, and  $a, b$  are taken to be  $-1, +1$  respectively.

The velocity field is perturbed by harmonic waves with wavenumber  $\alpha \in \mathbb{R}$  and wavespeed  $c \in \mathbb{C}$ . In the matrix  $\mathbf{A}$ ,  $R$  is the Reynolds number, whereas  $\lambda = -i\alpha c$ . If, for a given value of the complex parameter  $\lambda$ , there is a solution to the boundary value problem (18)–(19), then  $\lambda$  is said to be an eigenvalue and the corresponding solution an eigenmode. If  $\text{Re } \lambda > 0$  for some  $\lambda$  then  $U(x)$  is linearly unstable.

### 6. From geometry to numerical algorithms

Denote by GLRK4 (GLRK6, GLRK8) the two-stage (respectively three and four stage) GLRK methods of order four (respectively six and eight). They are symplectic, fully implicit RK collocation schemes, based on shifted Legendre polynomials and they preserve quadratic invariants [33,34].

Let  $\mathbf{H}(\mathbf{Q}, x) \in \text{Skew}_n\mathbb{C}$ , then the differential equation  $\mathbf{Q}_x = \mathbf{H}(\mathbf{Q}, x)\mathbf{Q}$  possesses the invariant  $\mathbf{I}(\mathbf{Q}) = \mathbf{Q}^H\mathbf{Q}$  and a GLRK scheme preserves  $\mathbf{I}$  to machine accuracy. Any ODE on the Stiefel manifold can be represented in this form. Consider

$$\mathbf{Q}_x = (\mathbf{I}_n - \mathbf{Q}\mathbf{Q}^H)\mathbf{A}\mathbf{Q} + \mathbf{Q}\mathbf{S},$$

where  $\mathbf{S}$  is any  $(x, \mathbf{Q})$  dependent skew-Hermitian matrix. Then subtraction of the trivial term

$$-\mathbf{A}^H(\mathbf{I}_n - \mathbf{Q}\mathbf{Q}^H)\mathbf{Q},$$

converts the equation into skew-symmetric form

$$\mathbf{Q}_x = \mathbf{H}(\mathbf{Q}, x)\mathbf{Q}, \tag{21}$$

with

$$\mathbf{H}(\mathbf{Q}, x) = 2\text{skew}_{\mathbb{C}}[(\mathbf{I}_n - \mathbf{Q}\mathbf{Q}^H)\mathbf{A}] + \mathbf{Q}\mathbf{S}\mathbf{Q}^H. \tag{22}$$

**Table 1**

Values of  $|q(\lambda_0, \bar{\lambda}_0)|$  for different numerical methods. The eigenvalue corresponds to the parameters  $\alpha = 1.020547$ ,  $R = 5772.2218$ ,  $c_0 = 0.2640003$  and  $\lambda_0 = -i\alpha c_0$ . For all methods  $h = 10^{-5}$ . Underlined digits represent discrepancies with the exact value  $\|\mathbf{Q}(1, \lambda_0, \bar{\lambda}_0)\|_F = \sqrt{2}$ .

Method	$\varepsilon$	$ q(\lambda_0, \bar{\lambda}_0) $	$\ \mathbf{Q}(1, \lambda_0, \bar{\lambda}_0)\ _F$
ERK4	–	$9.948160578828703 \times 10^{-08}$	<u>1.41421356272021</u>
PERK4	–	$7.712189768227568 \times 10^{-09}$	1.41421356237310
GLRK4	$10^{-16}$	$2.477341583320470 \times 10^{-09}$	<u>1.41421356237309</u>
GLRK6	$10^{-16}$	$2.192808264608588 \times 10^{-10}$	<u>1.41421356237309</u>
GLRK8	$10^{-16}$	$2.186958413939397 \times 10^{-10}$	<u>1.41421356237309</u>

The GLRK methods are fully implicit, hence at the  $i$ -th step  $\mathbf{Q}_i \mapsto \mathbf{Q}_{i+1}$ , a nonlinear system of algebraic equations of the form

$$\mathbf{Z}_{i,r} = h \sum_{j=1}^s a_{rj} \mathbf{F}(\mathbf{Q}_i + \mathbf{Z}_{i,j}, x_i + c_j h) \quad r = 1, \dots, s \tag{23}$$

has to be solved, and then the solution will be updated

$$\mathbf{Q}_{i+1} = \mathbf{Q}_i + h \sum_{j=1}^s b_j \mathbf{F}(\mathbf{Q}_i + \mathbf{Z}_{i,j}, x_i + c_j h). \tag{24}$$

We solved (23) via a fixed point iteration with initial guess  $\mathbf{Z}_{i0} = \mathbf{0}$  for all  $r$  and  $i$ . This fixed-point algorithm is known to converge for  $h$  sufficiently small [13]. This algorithm is not the most efficient, though it is the easiest to implement. An interesting alternative is proposed by McLachlan [35].

There are starting algorithms which provide optimal initial guesses and sensibly reduce the number of iterations,  $r$ , needed for the convergence to the fixed point [36]. These algorithms are crucial for large systems of equations; in our case,  $r$  is between 2 and 4, and  $\mathbf{Q} \in \mathbb{C}^{4 \times 2}$ , and in view of the relatively modest size of the system (23) we opted for the simpler implementation.

Also connected with the fixed point iteration is the stopping criterion. The iteration is stopped when the condition  $\|\mathbf{Z}_{[k+1]} - \mathbf{Z}_{[k]}\|_F \leq \varepsilon$  is met. The choice of the parameter  $\varepsilon$  plays an important role in how well the Stiefel manifold is preserved.

Since  $\mathbf{Q}_i$  and  $\sum_j b_j \mathbf{F}_j$  can have very different orders of magnitude, the rounding errors coming from their sum can be more significant than the ones coming from the solution of (23) itself. To overcome this problem, compensated summations have been implemented both in the calculation of  $\sum_j b_j \mathbf{F}_j$  and in the addition of this term and  $\mathbf{Q}_i$ . For details on the compensated summation see [37,33].

#### 6.1. Explicit RK methods and stabilization

Explicit RK methods can also be used if a stabilization term is added. Let

$$\mathbf{P}(\mathbf{Q}) := \mathbf{Q}^H \mathbf{Q} - \mathbf{I}_k.$$

Then modification of (21) to

$$\mathbf{Q}_x = \mathbf{H}(\mathbf{Q}, x)\mathbf{Q} - \gamma \mathbf{P}(\mathbf{Q}) = \widehat{\mathbf{M}}(\mathbf{Q}, x)\mathbf{Q} \tag{25}$$

with  $\gamma \in \mathbb{R}^+$ , still defines a flow on  $V_k(\mathbb{C}^n)$ , since on the Stiefel manifold  $\mathbf{P}(\mathbf{Q})$  vanishes identically, but

$$\mathbf{P}_x = -2\gamma \text{sym}_{\mathbb{C}}[(\mathbf{I}_k + \mathbf{P})\mathbf{P}] = -2\gamma \mathbf{P} - 2\gamma \mathbf{P}^2,$$

and so the manifold  $\mathbf{P} = \mathbf{0}$  is attracting. Here we will use the split step method of [38]; see also [17,5] for variations on this algorithm.

For the calculation a standard explicit RK scheme of fourth order is used in the first step and the Euler step is repeated two times for the second step, as suggested in [17]. This coupled scheme is denoted by PERK4. When the explicit RK method is used *without stabilization* it is denoted by ERK4.

**Table 2**  
Behavior of GLRK4 method for different values of the tolerance  $\varepsilon$ . Upper table: the value of  $|q(\lambda_0, \bar{\lambda}_0)|$  remains essentially unchanged. Lower table: the lower the tolerance, the bigger the number of average iterations needed to solve (23), the more the norm of  $\mathbf{Q}$  approaches the exact value  $\sqrt{2}$ . Parameters are those of Table 1.

Method	$\varepsilon$	Iterations	$ q(\lambda_0, \bar{\lambda}_0) $
PERK4	–	–	$7.712188274879859140943714270361545 \times 10^{-08}$
GLRK4	$10^{-16}$	6.2	$2.476828587317406227675273663784758 \times 10^{-09}$
GLRK4	$10^{-19}$	8.0	$2.476943132282032274442788242396517 \times 10^{-09}$
GLRK4	$10^{-25}$	10.1	$2.476942894862731906931829792273646 \times 10^{-09}$
GLRK4	$10^{-30}$	12.1	$2.476942894862759650910360742558537 \times 10^{-09}$
Method	$\varepsilon$	Iterations	$\ \mathbf{Q}(1, \lambda_0, \bar{\lambda}_0)\ _F$
PERK4	–	–	1.41421356237309504880168872406560
GLRK4	$10^{-16}$	6.2	1.41421356237309559882931008932402
GLRK4	$10^{-19}$	8.0	1.41421356237309504802583329658602
GLRK4	$10^{-25}$	10.1	1.41421356237309504880168885669407
GLRK4	$10^{-30}$	12.1	1.41421356237309504880168872423216

**7. Performance of GLRK and ERK methods**

First we test the performance of the complex orthonormalization. The codes were written in both MATLAB and FORTRAN 90 and tested on a Linux (Fedora) workstation and Unix (Solaris) Sun workstation.

As a benchmark, known eigenvalues of the Orr–Sommerfeld equation are used. Drazin and Reid [1] report that for  $\alpha = 1.020547$  and  $R = 5772.2218$  the critical eigenvalue is  $c_0 = 0.264003$ . We integrated the Orr–Sommerfeld equation on  $V_2(\mathbb{C}^4)$  in  $-1 \leq x \leq 1$  using these parameters and  $U(x) = 1 - x^2$ . For the range of integrators we computed  $|q(\lambda_0, \bar{\lambda}_0)|$  and the Frobenius norm of  $\mathbf{Q}(1, \lambda_0, \bar{\lambda}_0)$  which are expected to be null and  $\sqrt{2}$  respectively. Table 1 reports the results.

Focussing on the modulus of  $q$ , it can be observed that it gets closer to zero as the order of the scheme increases. Among the fourth order codes, however, GLRK4 with  $\varepsilon = 10^{-16}$  gives the best result. If the preservation of  $V_2(\mathbb{C}^4)$  is examined, the best result is obtained by PERK4, whereas ERK4 is much less accurate, as expected. The GLRKx codes, exhibit a very small drift off the Stiefel manifold. This drift behaviour is negligible, since after  $2 \times 10^5$  iterations the drift is of the order of  $10^{-14}$ , but its total dependence on the tolerance  $\varepsilon$  used in the fixed point iterations of (23) is of interest. It is indeed reasonable to expect that the smaller  $\varepsilon$  is, the more accurate the solutions  $\mathbf{Z}_{i,r}$  are, and the closer  $\mathbf{Q}_{i+1}$  will be to the Stiefel manifold. In fact, using double precision arithmetic, one is actually forced to choose  $\varepsilon$  around  $10^{-16}$ , and hence we recalculated using quadruple precision. This shifts substantially the zero precision, so that lower values for the tolerance are allowed. Table 2 reports the results for PERK4 and for GLRK4 for different values of  $\varepsilon$ . In quadruple precision, the Stiefel manifold is preserved with accuracy driven exactly by  $\varepsilon$ , and we pay for this improvement with an increase of the average number of iterations needed to solve (23). For the absolute value of  $q$  minor changes are observable. Unless otherwise stated, in the remainder of the article we will report calculations in quadruple precision and  $\varepsilon = 10^{-30}$ , even if we will present results with less digits.

**8. Contour integration and computation of the index**

In this section algorithms for the numerical integration of index  $\text{Ind}(\mathbf{q}, \mathcal{D})$  for a simply connected region  $\mathcal{D}$  are presented. The contour integral is approximated with the trapezoidal rule,

$$\begin{aligned} \text{Ind}(\mathbf{q}, \mathcal{D}) &= \frac{1}{2\pi} \int_0^{2\pi} \frac{u\dot{v} - v\dot{u}}{u^2 + v^2} dt \\ &\approx \frac{1}{n} \sum_{\nu=0}^{n-1} \frac{u(t_\nu)\dot{v}(t_\nu) - v(t_\nu)\dot{u}(t_\nu)}{u(t_\nu)^2 + v(t_\nu)^2}, \\ t_\nu &= \nu 2\pi/n, \nu = 0, \dots, n-1. \end{aligned} \tag{26}$$

The trapezoidal rule guarantees fast convergence for periodic functions [39]. There are however other sources of error:

- Computation of  $u, v$ : the fact that  $u(t_\nu)$  and  $v(t_\nu)$  are obtained as the real and imaginary parts of  $q(\lambda(t_\nu), \bar{\lambda}(t_\nu))$  which is found from numerical integration of an ODE from  $x = a$  to  $x = b$ .
- Computation of  $\dot{u}, \dot{v}$ : the derivatives along the contour need to be approximated.

We employ two approaches to numerical differentiation. The first is a centered difference approximation

$$\dot{u}(t_\nu) = \frac{u(t_\nu + \delta) - u(t_\nu - \delta)}{2\delta} + \mathcal{O}(\delta^2). \tag{27}$$

Another possibility is to employ complex differentiation [40,41]

$$\dot{u}(t_\nu) = \frac{\text{Im}(u(t_\nu + i\delta))}{\delta} + \mathcal{O}(\delta^2). \tag{28}$$

The latter expression is obtained by extending  $u$  to be a complex function and assuming  $u$  to be real on the real axis. Expanding  $u$  in a Taylor series leads to

$$u(t_\nu + i\delta) = u(t_\nu) + i\delta\dot{u}(t_\nu) - \frac{\delta^2}{2}\ddot{u}(t_\nu) - i\frac{\delta^3}{6}\dddot{u}(t_\nu) + \dots$$

Now taking the imaginary part and dividing by  $\delta$  shows that complex differentiation has the same truncation error as centered differences. Complex differentiation (28) is much better behaved for very small  $\delta$  (as, unlike finite differences (27), it does not suffer from subtractive cancellation errors) but it requires complexification which is cumbersome for complex-valued functions.

To test the index computation, first consider the 2D example (16). Recast (16) into a system with coefficient matrix

$$\mathbf{A}(\lambda) = \begin{bmatrix} 0 & 1 \\ -\lambda & 0 \end{bmatrix}.$$

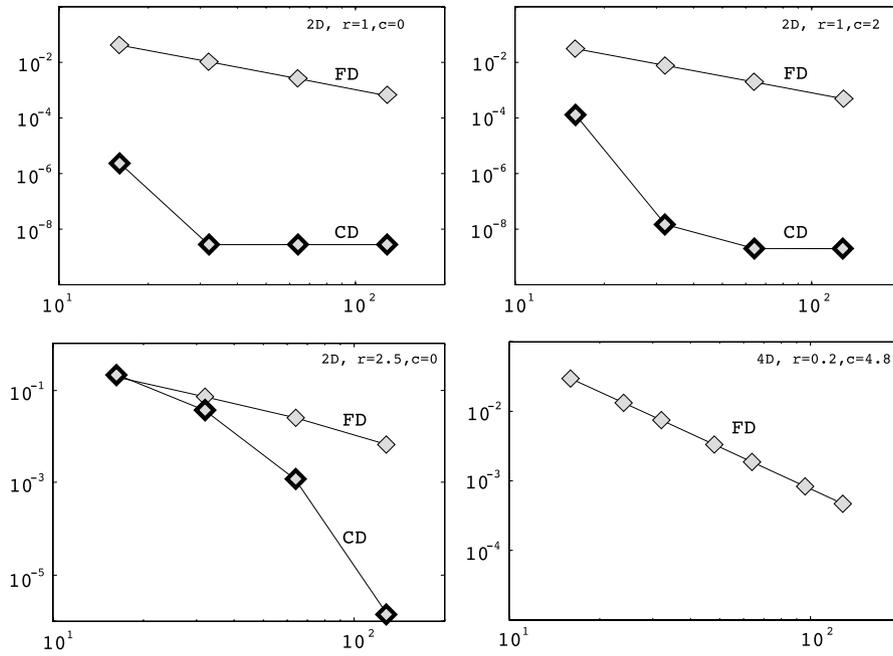
Then continuous orthogonalization with stabilization leads to the system

$$\begin{aligned} \mathbf{Q}_x &= \text{skew}_c[\mathbf{A}(\lambda) - 2\mathbf{Q}\mathbf{Q}^H \text{sym}_c(\mathbf{A})]\mathbf{Q} - \gamma\mathbf{Q}P(\mathbf{Q}), \\ \mathbf{Q} &\in V_1(\mathbb{C}^2) \end{aligned} \tag{29}$$

on the interval  $0 \leq x \leq \pi$ . This system is integrated using PERK4. At  $x = \pi$ ,  $q(\lambda, \bar{\lambda}) = Q_2(\pi, \lambda, \bar{\lambda})$ , the second component of  $\mathbf{Q}$ .

Integrating (29) makes available  $u(t_\nu), v(t_\nu)$ , in that  $q(\lambda(t_\nu), \bar{\lambda}(t_\nu)) = Q_2(\pi, \lambda(t_\nu), \bar{\lambda}(t_\nu)) = u(t_\nu) + iv(t_\nu)$  and this allows us also to compute  $\dot{u}(t_\nu), \dot{v}(t_\nu)$ , via finite differences (27).

On the other hand, when employing complex differentiation, integrating (29) is not sufficient. Indeed, in order to perform complex differentiation (28), we need to be able to access  $\text{Im}(u(t_\nu + i\delta))$



**Fig. 1.** Convergence of the numerical computations of  $\text{Ind}(\mathbf{A}, \mathcal{D})$  for increasing number of sampling points on  $\partial\mathcal{D}$ . Label FD refers to derivatives computed with finite-differences, while CD to derivatives computed with complex differentiation. Top-left: two-dimensional example with  $\mathcal{D} = \mathcal{D}_1$ . Top-right: two-dimensional example with  $\mathcal{D} = \mathcal{D}_2$ . Bottom-left: two-dimensional example with  $\mathcal{D} = \mathcal{D}_3$ . Bottom-right: four-dimensional example with  $\mathcal{D} = \mathcal{D}_4$ . Two-dimensional ODE integrations are computed with PERK4 and  $h = 0.006$ . Four-dimensional computation is with PERK4 and  $h = 10^{-4}$ . For complex differentiation we chose  $\delta = 10^{-4}$ .

and  $\text{Im}(v(t_v + i\delta))$  separately, whereas setting  $\lambda = \lambda(t_v + i\delta)$  and integrating (29) will only provide us with  $Q_2$

$$Q_2(t_v + i\delta) = \text{Re}(u(t_v + i\delta)) + \text{Re}(v(t_v + i\delta)) + i[\text{Im}(u(t_v + i\delta)) + \text{Im}(v(t_v + i\delta))].$$

In the expression for  $Q_2$ , the imaginary parts of  $u(t_v + i\delta)$  and  $v(t_v + i\delta)$  are bundled together, therefore we can not use this approach for complex differentiation. Hence the strategy is to first rewrite (29) in real coordinates, that is, we pose  $\mathbf{Q} = \mathbf{U} + i\mathbf{V}$ , denote by  $\mathbf{F}(\mathbf{U}, \mathbf{V})$  the right-hand side of Eq. (29), and then evolve

$$\mathbf{U}_x = \text{Re}(\mathbf{F}(\mathbf{U}, \mathbf{V}))$$

$$\mathbf{V}_x = \text{Im}(\mathbf{F}(\mathbf{U}, \mathbf{V})).$$

This system is a four-dimensional real system and in order to compute  $u(t_v + i\delta)$  and  $v(t_v + i\delta)$  it needs to be complexified, leading to an eight-dimensional system.

Contour integrals were computed using both finite-difference and complex differentiation, in the three domains:

- $\mathcal{D}_1$ , the disk of radius 1 centered in the origin, for which  $\text{Ind}(\mathbf{A}, \mathcal{D}_1) = 1$ .
- $\mathcal{D}_2$ , the disk of radius 1 centered in  $(2, 0)$ , for which  $\text{Ind}(\mathbf{A}, \mathcal{D}_2) = 1$ .
- $\mathcal{D}_3$ , the disk of radius 2 centered in the origin, for which  $\text{Ind}(\mathbf{A}, \mathcal{D}_3) = 2$ .

In Fig. 1 we collect convergence results for the contour integral computed via (26) as the number of gridpoints  $n$  increases. The integral converges as expected and the complex differentiation always performs better than finite-differences, even with a relatively large value of  $\delta$ . On the other hand the double complexification leads to large systems in the case of 4 dimensions and higher and so the centered finite difference is preferred in these cases.

Now apply the contour integration to the four-dimensional example (17). The eigenvalues are all real, the lowest eigenvalue is  $\lambda_1 \approx 4.73$ , and the next eigenvalue is far enough away that a unit circle contour around  $\lambda_1$  will contain only one eigenvalue. A convergence study for the approximation of the contour integral for  $\mathcal{D}_4$ , the disk of radius 0.2 centered around  $\lambda_1$  is shown in Fig. 1: the finite-difference approximation for the derivatives in the contour is satisfactory, giving algebraic convergence.

## 9. Newton's method and analyticity

Since  $q$  is non-analytic, the simplest approach to using Newton's method is to separate the real and imaginary parts and treat it as a map from  $\mathbb{R}^2$  to  $\mathbb{R}^2$ . Define  $\varphi: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,

$$\zeta := (\text{Re } \lambda, \text{Im } \lambda)^T$$

$$\varphi(\zeta) := (\text{Re } q(\lambda, \bar{\lambda}), \text{Im } q(\lambda, \bar{\lambda}))^T.$$

The Newton sequence for  $\mathbf{q}(\lambda, \bar{\lambda}) = 0$  is then

$$\zeta_{[k+1]} = \zeta_{[k]} - \mathbf{J}^{-1}(\zeta_{[k]})\varphi(\zeta_{[k]}) \quad k = 0, 1, \dots, c, \quad (30)$$

where  $\mathbf{J}$  is the Jacobian of  $\varphi$ .

To test this algorithm, it is applied to the Orr–Sommerfeld equation. Let  $\zeta^* = (\text{Re } \lambda_0, \text{Im } \lambda_0)^T$  be a known root of  $\mathbf{q} = 0$  and perturb it. Namely we take  $\zeta_{[0]} \in \partial\mathcal{B}_\rho(\zeta^*)$  where  $\mathcal{B}_\rho(\zeta^*)$  is the open ball centred in  $\zeta^*$  with radius  $\rho$ . In our codes we are able to set  $\rho$  and to obtain a random direction for the perturbation.

Both  $\zeta_{[k]}$  and the value of the Jacobian need to be computed for the Newton iteration. For a given  $\zeta_{[k]}$  (namely for a fixed  $\lambda$ ) we integrate the Orr–Sommerfeld equation restricted to the Stiefel manifold and from  $\mathbf{Q}(1, \lambda, \bar{\lambda})$  compute  $\varphi(\zeta_{[k]})$ .

The elements of the Jacobian are approximated with a standard second order finite difference expression

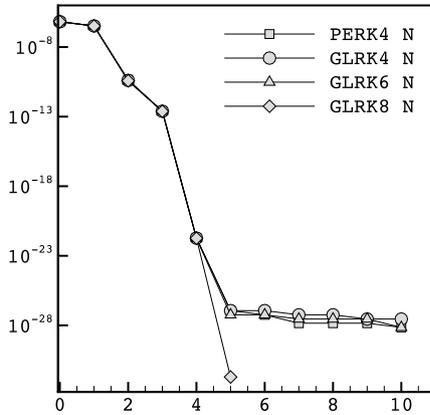
$$\frac{\partial \varphi_i}{\partial \zeta_j}(\zeta) \approx \frac{\varphi_i(\zeta + \delta \zeta_j \mathbf{e}_j) - \varphi_i(\zeta - \delta \zeta_j \mathbf{e}_j)}{2\delta \zeta_j} \quad i, j = 1, 2, \quad (31)$$

where the parameters  $\delta \zeta_j$  can be chosen conveniently small.

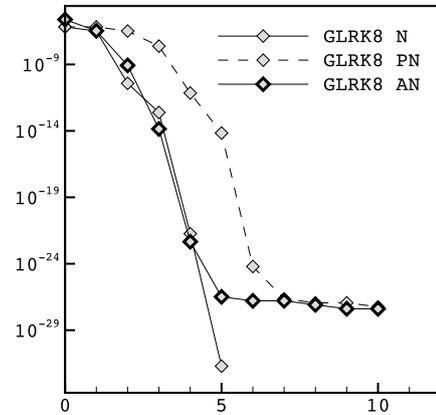
In order to perform a single Newton step five function evaluations are required (five integrations of the Orr–Sommerfeld equation). Fig. 2 shows the convergence for the Newton iteration  $\|\varphi(\zeta_{[k]})\| = |q(\lambda_{[k]}, \bar{\lambda}_{[k]})|$  as a function of  $k$ . The algorithm GLRK8 N gives the best result in terms of convergence and of course is the most accurate in the integration of the Orr–Sommerfeld equation. As reported in Table 3, the best approximation for the eigenvalue with  $R = 5772.2218$ ,  $\alpha = 1.020547$  is  $c \approx 0.2640002080337$ , for which  $|q(\lambda, \bar{\lambda})| \approx 10^{-32}$ . Considering that the computations

**Table 3**  
 Roots of  $q(\lambda, \bar{\lambda})$  obtained with different Newton iterations and different integrators. Newton (N) and Pseudo-Newton (PN). Parameters  $\alpha = 1.020547$ ,  $R = 5772.2218$ . Drazin and Reid [1] reports  $c = 0.2640003$ .

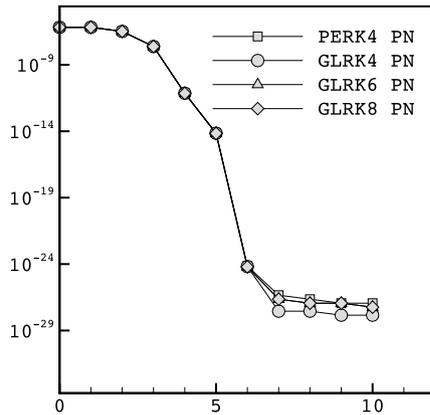
Method	$c$	$ q(\lambda, \bar{\lambda}) $
PERK4 N	$0.2640027091196 + i0.2147621517184 \times 10^{-05}$	$0.714 \times 10^{-28}$
GLRK4 N	$0.2640010449625 + i0.7207359743415 \times 10^{-06}$	$0.286 \times 10^{-27}$
GLRK6 N	$0.2640002081762 - i0.2646932244137 \times 10^{-10}$	$0.714 \times 10^{-28}$
GLRK8 N	$0.2640002080337 - i0.4270687919662 \times 10^{-09}$	$0.194 \times 10^{-31}$
PERK4 PN	$0.2640027091196 + i0.2147621517184 \times 10^{-05}$	$0.114 \times 10^{-26}$
GLRK4 PN	$0.2640010449625 + i0.7207359743415 \times 10^{-06}$	$0.143 \times 10^{-27}$
GLRK6 PN	$0.2640002081762 - i0.2646932244137 \times 10^{-10}$	$0.571 \times 10^{-27}$
GLRK8 PN	$0.2640002080337 - i0.4270687919662 \times 10^{-09}$	$0.571 \times 10^{-27}$



**Fig. 2.** Absolute value of  $q(\lambda, \bar{\lambda})$  versus Newton iterations for different integrators. Parameters are the same as Table 1.



**Fig. 4.** Convergence of  $q(\lambda, \bar{\lambda})$  for GLRK8 with Newton (N), Pseudo-Newton (PN) and Newton with analyticity conjecture (AN). Parameters are the same as Table 1.



**Fig. 3.** Absolute value of  $q(\lambda, \bar{\lambda})$  versus Pseudo-Newton iterations for different integrators. Parameters as in Table 1.

have been performed in extended precision, this result represents a refinement of the value  $c_0 = 0.2640003$  reported in [1]. Both PERK4 N and GLRK4 N show a slower convergence, and to a less accurate value of  $c$ . For instance the projection explicit method gives  $c \approx 0.264003$ , less accurate in the real as well as in the imaginary part.

**9.1. Pseudo-Newton iteration**

Since the codes PERK4 N and GLRKx N spend most of their time evaluating the Jacobian of (30), it is reasonable to implement their Pseudo-Newton (PN) versions. This is obtained by computing the Jacobian every  $m$  iterations. More precisely the iterations are

$$\zeta_{[k+1]} = \zeta_{[k]} - J(\zeta_{[k']})^{-1} \varphi(\zeta_{[k]}) \quad k = 0, 1, \dots \quad (32)$$

where

$$k' = \begin{cases} 0 & \text{if } k = 0 \\ m \lfloor (k-1)/m \rfloor + 1 & \text{if } k \neq 0. \end{cases}$$

In our algorithms we set  $m = 2$ , so that the number of function evaluations (here integrations of the Orr–Sommerfeld equation) is considerably reduced. This, of course, penalises the convergence of the method, as one can observe in Fig. 3. A direct comparison between Newton and Pseudo-Newton iterations for GLRK8 is also in Fig. 4 and it is evident that the curve relative to Newton iterations is always well below the Pseudo-Newton one. The third curve represents an attempt to treat  $q$  as a pseudo-analytic function. This latter observation is a curiosity witnessed in the numerics: that  $q$  near a zero of  $q$  behaves asymptotically like an analytic function. The Pseudo-Newton scheme needs many more iterations to converge, but approximately the same CPU time as the Newton one.

**10. Concluding remarks**

Questions associated with complex orthogonalization and analyticity have been studied. The main new results are: (a) there is a continuum-infinite number of ways to implement continuous orthonormalization; (b) there is a link with the differential geometry of Stiefel bundles, and the simplest implementation of continuous orthonormalization corresponds to parallel transport; (c) the index of the non-analytic Evans function (characteristic function) is the same including multiplicity as the complex analytic form; (d) the index can be efficiently computed with the trapezoidal rule and numerical differentiation; (e) the efficiency of GLRK methods for orthogonalization is intimately connected with how accurately the nonlinear equations are solved at each step; (f) GLRK8 gives optimal results, both in term of convergence of  $|q(\lambda, \bar{\lambda})|$  and of accuracy of the eigenvalue – indeed computed

results for Orr–Sommerfeld are more accurate than previous results in the literature; and (g) Newton's method applied to the non-analytic characteristic function is almost as efficient as Newton's method applied to the analytic form.

## Acknowledgements

The authors are grateful for helpful discussions or helpful comments from Gianne Derks (Surrey), Simon Malham (Heriot-Watt), Sebastian Reich (Potsdam), Björn Sandstede (Brown), and Kevin Zumbrun (Indiana). Daniele Avitabile's research was funded by EPSRC, EP/E032249/1 *Applied Mathematics: Making it Real* grant.

## References

- [1] P.G. Drazin, W.H. Reid, *Hydrodynamic Stability*, in: Cambridge Mathematical Library, Cambridge University Press, Cambridge, 2004.
- [2] T.J. Bridges, The Orr–Sommerfeld equation on a manifold, *Proc. R. Soc. Lond. A* 455 (1988) 3019–3040. 1999.
- [3] J. Alexander, R. Gardner, C.K.R.T. Jones, A topological invariant arising in the stability analysis of travelling waves, *J. Reine. Angew. Math.* 410 (1990) 167–212.
- [4] B. Sandstede, Stability of travelling waves, in: *Handbook of Dynamical Systems*, 2, 2002, pp. 983–1055.
- [5] J. Humpherys, K. Zumbrun, An efficient shooting algorithm for Evans function calculations in large systems, *Physica D* 220 (2006) 116–126.
- [6] U.M. Ascher, R.D. Russell, R.M.M. Mattheij, *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*, Prentice-Hall Inc., Englewood Cliffs, NJ, 1988.
- [7] F. Austin, T.J. Bridges, A bundle view of boundary-value problems: generalizing the Gardner–Jones bundle, *J. Differential Equations* 189 (2003) 412–439.
- [8] I. Gohberg, L. Rodman, Analytic matrix functions with prescribed local data, *J. Anal. Math.* 18 (40) (1981) 90–128.
- [9] T. Kato, *Perturbation Theory for Linear Operators*, Springer-Verlag, New York, 1985.
- [10] J. Humpherys, B. Sandstede, K. Zumbrun, Efficient computation of analytic bases in Evans function analysis of large systems, *Numer. Math.* 103 (4) (2006) 631–642.
- [11] I. Goldhirsch, P.-L. Sulem, S.A. Orszag, Stability and Lyapunov stability of dynamical systems: a differential approach and a numerical method, *Physica D* 27 (3) (1987) 311–337.
- [12] L. Dieci, E.S. Van Vleck, Computation of a few Lyapunov exponents for continuous and discrete dynamical systems, *Appl. Numer. Math.* 17 (3) (1995) 275–291.
- [13] L. Dieci, R.D. Russell, E.S. van Vleck, Unitary integrators and application to continuous orthogonalization techniques, *SIAM J. Numer. Anal.* 31 (1) (1994) 261–281.
- [14] L. Dieci, R.D. Russell, E.S. Van Vleck, On the computation of Lyapunov exponents for continuous dynamical systems, *SIAM J. Numer. Anal.* 34 (1) (1997) 402–423.
- [15] L.O. Drury, Numerical solution of Orr–Sommerfeld type equations, *J. Comput. Phys.* 37 (1) (1980) 133–139.
- [16] A. Davey, An automatic orthonormalization method for solving stiff boundary value problems, *J. Comput. Phys.* 51 (2) (1983) 343–356.
- [17] T.J. Bridges, S. Reich, Computing Lyapunov exponents on a Stiefel manifold, *Physica D* 156 (2001) 219–238.
- [18] V. Ledoux, S.J.A. Malham, V. Thummler, Grassmannian spectral shooting, *Math. Comput.* (2009) (in press).
- [19] A. Edelman, T.A. Arias, S.T. Smith, The geometry of algorithms with orthogonality constraints, *SIAM J. Matrix Anal. Appl.* 20 (2) (1998) 303–353.
- [20] S. Krogstad, A low complexity Lie group method on the Stiefel manifold, *BIT* 43 (1) (2003) 107–122.
- [21] E. Celledoni, B. Owren, A class of intrinsic schemes for orthogonal integration, *SIAM J. Numer. Anal.* 40 (6) (2002) 2069–2084.
- [22] E. Celledoni, B. Owren, On the implementation of Lie group methods on the Stiefel manifold, *Numer. Algorithms* 32 (2–4) (2003) 163–183.
- [23] J.H. Manton, Optimization algorithms exploiting unitary constraints, *IEEE Trans. Signal Process* 50 (3) (2002) 635–650.
- [24] C.J. Isham, *Modern Differential Geometry for Physicists*, in: *Lecture Notes in Physics*, vol. 61, World Scientific, London, 2003.
- [25] S. Kobayashi, K. Nomizu, *Foundations of Differential Geometry*, vol. 1, John Wiley and Sons, New York, 1963.
- [26] M. Spivak, *A Comprehensive Introduction to Differential Geometry*, vol. II, Publish or Perish Inc., Houston TX, 1979.
- [27] M.S. Narasimhan, S. Ramanan, Existence of universal connections, *Amer. J. Math.* 83 (1961) 563–572.
- [28] A. Bohm, L.J. Boya, A. Mostafazadeh, G. Rudolph, Classification theorem for principal fibre bundles, Berry's phase, and exact cyclic evolution, *J. Geom. Phys.* 12 (1993) 13–28.
- [29] R. Way, Dynamics in the Hopf bundle, the geometric phase and implications for dynamical systems, Ph.D. Thesis, Department of Mathematics, University of Surrey, England, UK, 2009.
- [30] M.P. do Carmo, *Differential Forms and Applications*, in: *Universitext*, Springer-Verlag, Berlin, 1994.
- [31] K.H. Ko, T. Sakkalis, N.M. Patrikalakis, A reliable algorithm for computing the topological degree of a mapping in  $\mathbb{R}^2$ , *Appl. Math. Comput.* 196 (2008) 666–678.
- [32] B.S. Ng, W.H. Reid, An initial value method for eigenvalue problems using compound matrices, *J. Comput. Phys.* 30 (1979) 125–136.
- [33] E. Hairer, C. Lubich, G. Wanner, *Geometric Numerical Integration*, in: *Springer Series in Computational Mathematics*, vol. 31, Springer-Verlag, Berlin, 2002.
- [34] B. Leimkuhler, S. Reich, *Simulating Hamiltonian Dynamics*, in: *Cambridge Monographs on Applied and Computational Mathematics*, vol. 14, Cambridge University Press, Cambridge, 2004.
- [35] R.I. McLachlan, A new implementation of symplectic Runge–Kutta methods, *SIAM J. Sci. Comput.* 29 (2007) 1637–1649.
- [36] M.P. Laburta, Construction of starting algorithms for the RK–Gauss methods, *J. Comput. Appl. Math.* 90 (2) (1998) 239–261.
- [37] N.J. Higham, The accuracy of floating point summation, *SIAM J. Sci. Comput.* 14 (4) (1993) 783–799.
- [38] U.M. Ascher, H. Chin, S. Reich, Stabilization of DAEs and invariant manifolds, *Numer. Math.* 67 (2) (1994) 131–149.
- [39] J.A.C. Weideman, Numerical integration of periodic functions: a few examples, *Amer. Math. Monthly* 109 (2002) 21–36.
- [40] W. Squire, G. Trapp, Using complex variables to estimate derivatives of real functions, *SIAM Rev.* 40 (1998) 110–112.
- [41] L.F. Shampine, Accurate numerical derivatives in MATLAB, *ACM Trans. Math. Software* 33 (article 26) (2007) 17.