Pseudospectra for matrix pencils and stability of equilibria

Jos L.M. van Dorsselaer*

March 1996

Abstract. The concept of ε–pseudospectra for matrices, introduced by Trefethen and his co-workers, has been studied extensively since 1990. In this paper, ε–pseudospectra for matrix pencils, which are relevant in connection with generalized eigenvalue problems, are considered. Some properties as well as the practical computation of ε–pseudospectra for matrix pencils will be discussed. As an application, we demonstrate how this concept can be used for investigating the asymptotic stability of stationary solutions to time-dependent ordinary or partial differential equations; two cases, based on Burgers' equation, will be shown.

Key words: ε–pseudospectra, ε–pseudospectra for matrix pencils, (generalized) eigenvalue problems, equilibria of differential equations, stability of equilibria.

AMS subject classification: 65H17, 65L07, 15A18.

1 Introduction

In 1990, Trefethen and his co-workers introduced the concept op ε-pseudospectra (see, e.g., [12, 13, 18]). For a given $N \times N$ matrix A and $\varepsilon \geq 0$ the set $\Lambda_\varepsilon(A)$, the *ε-pseudospectrum* of A, is defined by

$$\Lambda_\varepsilon(A) = \{\lambda \in \mathbb{C} : \|(\lambda I - A)^{-1}\| \geq \varepsilon^{-1}\}, \tag{1.1}$$

with $\|(\lambda I - A)^{-1}\| = \infty$ whenever λ is an eigenvalue of A. Here $\|\cdot\|$ stands for the spectral norm, i.e. the matrix norm induced by the Euclidean vector norm, and I is the identity matrix. Definition (1.1) is equivalent to (see, e.g., [13])

$$\Lambda_\varepsilon(A) = \{\lambda \in \mathbb{C} : \exists\ E \text{ with } \|E\| \leq \varepsilon \text{ such that } \lambda \text{ is an eigenvalue of } A + E\} \tag{1.2}$$

*Mathematical Institute, University of Utrecht, P.O. Box 80.010, NL-3508 TA Utrecht, The Netherlands, E-mail: dorssela@math.ruu.nl. This research has been supported by the Netherlands Organization for Scientific Research (N.W.O.).

which explains the name pseudospectra.

When the matrix A is normal, its behaviour in numerical processes is determined by the location of the eigenvalues. For non-normal matrices, however, this does not need to be true; predictions based on eigenvalues may lead to wrong results. These phenomena occur e.g. in the convergence behaviour of iterative methods for linear equations (such as GMRES, compare [9, 12]), or in the stability behaviour of time-stepping methods for ordinary and partial differential equations [6, 13, 14]. In these cases knowledge of ε-pseudospectra may be useful for understanding the behaviour of the processes involving A.

The definitions (1.1), (1.2) are immediately connected to the eigenvalue problem $Av = \lambda v$. However, in many applications one has to deal with the generalized eigenvalue problem

$$Av = \lambda Bv \tag{1.3}$$

where B is a given $N \times N$ matrix, which may be singular (see, e.g., [5]).

In this paper we consider ε-pseudospectra for matrix pencils, so that there is a connection with eigenvalue problem (1.3); define for $\varepsilon \geq 0$ the set $\Lambda_\varepsilon(A,B)$, the *ε-pseudospectrum of the matrix pencil* (A,B) as

$$\Lambda_\varepsilon(A,B) = \{\lambda \in \mathbb{C} : \|(\lambda B - A)^{-1}\| \geq \varepsilon^{-1}\},$$

with $\|(\lambda B - A)^{-1}\| = \infty$ whenever λ is an eigenvalue of (1.3). This definition is equivalent to (see Section 2.1)

$$\Lambda_\varepsilon(A,B) = \{\lambda \in \mathbb{C} : \exists\ E \text{ with } \|E\| \leq \varepsilon \text{ such that } A + E - \lambda B \text{ is singular}\}. \tag{1.4}$$

Note that $\Lambda_\varepsilon(A,I) = \Lambda_\varepsilon(A)$, so our definition can be seen as a generalization of ε-pseudospectra to arbitrary matrix pencils (A,B). From (1.4) one sees that only perturbations of the matrix A are considered; one could also consider perturbations of both A and B as well. However, in this paper we will deal with a range of applications in which only perturbations of A will play a role. In [15] the set $\Lambda_\varepsilon(B^{-1}A)$ is considered for nonsingular matrices B; this set will be related to $\Lambda_\varepsilon(A,B)$ (in Section 2.1).

In this paper we demonstrate how ε-pseudospectra for matrix pencils can be used to investigate the asymptotic stability of stationary solutions to systems of ordinary differential equations, when inexact Jacobian matrices are used. To illustrate this approach, we consider two systems of ordinary differential equations; both systems have been obtained from semi-discretizing Burgers' equation, using a finite difference and a finite element method, respectively.

The paper is organized as follows. In Section 2.1 we present properties of $\Lambda_\varepsilon(A,B)$, and the computation of $\Lambda_\varepsilon(A,B)$ will be discussed in Section 2.2. In Section 3.1 we consider the application of ε-pseudospectra for investigating the asymptotic stability of equilibria. An illustration, based on Burgers' equation, will be given in Section 3.2.

2 Pseudospectra for matrix pencils

2.1 Some properties of pseudospectra

Unless stated otherwise, all matrices may have complex entries. The matrices A and B may be singular.

Note that the set $\Lambda_0(A,B) = \{\lambda \in \mathbb{C} : \lambda B - A \text{ is singular}\}$ is the *spectrum* (the set of eigenvalues) of the matrix pencil (A,B). In the remainder of this paper the spectrum of the matrix pencil (A,B) will be denoted by $\Lambda_0(A,B)$.

Most properties of $\Lambda_\varepsilon(A)$ can easily be generalized to $\Lambda_\varepsilon(A,B)$. An important observation is

$$\Lambda_\varepsilon(A,B) \subset \Lambda_{\varepsilon'}(A,B) \text{ for } 0 \le \varepsilon \le \varepsilon',$$

i.e. the sets $\Lambda_\varepsilon(A,B)$ are nested. Now some equivalent definitions of $\Lambda_\varepsilon(A,B)$ will be given; some of them can be useful in practice. Consider for given $N \times N$ matrices A and B and $\varepsilon > 0$ the following definitions ($\|u\|$ stands for the Euclidean norm of the vector u)

(2.1) $\Lambda_\varepsilon(A,B) = \{\lambda \in \mathbb{C} : \exists\ E \text{ with } \|E\| \le \varepsilon \text{ such that } A + E - \lambda B \text{ is singular}\}$,

(2.2) $\Lambda_\varepsilon(A,B) = \{\lambda \in \mathbb{C} : \exists\ E \text{ with } \|E\| = \varepsilon \text{ such that } A + E - \lambda B \text{ is singular}\}$,

(2.3) $\Lambda_\varepsilon(A,B) = \{\lambda \in \mathbb{C} : \exists\ U \in \mathbb{C}^{N,N} \text{ with } \|U\| = 1 \text{ such that } \|(\lambda B - A)U\| \le \varepsilon\}$,

(2.4) $\Lambda_\varepsilon(A,B) = \{\lambda \in \mathbb{C} : \exists\ u \in \mathbb{C}^{N} \text{ with } \|u\| = 1 \text{ such that } \|(\lambda B - A)u\| \le \varepsilon\}$,

(2.5) $\Lambda_\varepsilon(A,B) = \{\lambda \in \mathbb{C} : \lambda B - A \text{ is singular or } \|(\lambda B - A)^{-1}\| \ge \varepsilon^{-1}\}$,

(2.6) $\Lambda_\varepsilon(A,B) = \{\lambda \in \mathbb{C} : \text{ the smallest singular value of } \lambda B - A \text{ is } \le \varepsilon\}$.

Theorem 2.1. *For $N \ge 2$ the definitions* (2.1) – (2.6) *are equivalent.*

Proof. The equivalence (2.1) – (2.5) follows from a straightforward generalization of the proof for $B = I$ (see, e.g., [6, Theorem 4.2]). The equivalence of (2.5) and (2.6) is obvious. □

Before we discuss other properties, some examples will be presented. Consider the matrices

$$A_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \ (\alpha \in \mathbb{C}), \quad \text{and} \quad A_3 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Furthermore, $D[\gamma,\rho]$ stands for the disk in the complex plane with center γ and radius ρ. After some computations, using e.g. (2.6), one can see that $\Lambda_\varepsilon(A_1,I) = \Lambda_\varepsilon(A_1) = D[0,\varepsilon] \cup D[1,\varepsilon]$, $\Lambda_\varepsilon(A_2,I) = D\left[1, \sqrt{|\alpha|\varepsilon + \varepsilon^2}\right]$, $\Lambda_\varepsilon(A_3,I) = D[-i,\varepsilon] \cup D[i,\varepsilon]$, and $\Lambda_\varepsilon(A_1,A_3) = D\left[0, \sqrt{\varepsilon + \varepsilon^2}\right]$.

For singular matrices B the set $\Lambda_0(A,B)$ can be degenerated; for example $\Lambda_0(A,B) = \emptyset$ or $\Lambda_0(A,B) = \mathbb{C}$ is possible. In order to see how $\Lambda_\varepsilon(A,B)$ might behave for singular B, the sets $\Lambda_\varepsilon(I,A_1)$ and $\Lambda_\varepsilon(A_3,A_1)$ have been determined. For $\varepsilon < 1$ we have $\Lambda_\varepsilon(I,A_1) = D[1,\varepsilon]$, and $\Lambda_\varepsilon(I, A_1) = \mathbb{C}$ for $\varepsilon \geq 1$. Furthermore, $\Lambda_0(A_3, A_1) = \emptyset$ and $\Lambda_\varepsilon(A_3, A_1) = \{z \in \mathbb{C} : |z| \geq (1-\varepsilon^2)/\varepsilon\}$ for $\varepsilon > 0$, implying that $\Lambda_\varepsilon(A_3, A_1) = \mathbb{C}$ for $\varepsilon \geq 1$.

In these examples we observed that $\Lambda_\varepsilon(A, B)$ is bounded for nonsingular matrices B, while $\Lambda_\varepsilon(A, B) = \mathbb{C}$ is possible for singular B and ε sufficiently large. One might ask whether these phenomena occur in general. Another interesting question is whether $\Lambda_\varepsilon(A, B)$ can be empty for positive ε. The following theorem answers these questions. In order to formulate the result we introduce for the matrix pencil (A, B) with singular B the constant

$$\varepsilon^* = \min\{\, \|Au\| : u \in \mathbb{C}^N \text{ with } \|u\| = 1 \text{ and } Bu = 0\}.$$

Theorem 2.2. (a) *$\Lambda_\varepsilon(A, B)$ is bounded for nonsingular matrices B.*
(b) *For singular matrices B one has $\Lambda_\varepsilon(A, B) = \mathbb{C}$ for $\varepsilon \geq \varepsilon^*$.*
(c) *Let $B \neq O$ (the null matrix). Then $\Lambda_\varepsilon(A, B) \neq \emptyset$ for all $\varepsilon > 0$.*

Proof. For nonsingular matrices B we have $\|(zB - A)^{-1}\| = \|(zI - B^{-1}A)^{-1} B^{-1}\| \leq \|B^{-1}\| \cdot \|(zI - B^{-1}A)^{-1}\| \to 0$ as $|z| \to \infty$, and result (a) follows from (2.5).

Statement (b) follows immediately from (2.4).

Now (c) will be proved. It is clear that the result holds if $\Lambda_0(A, B) \neq \emptyset$. In the remainder of the proof we assume that $\Lambda_0(A, B) = \emptyset$. Let e_j be the jth unit vector in $\mathbb{C}^N$ and consider for $1 \leq j, k \leq N$ the functions $\varphi_{j,k}(z) = e_j{}^*(zB - A)^{-1} e_k$ (the entries of $(zB - A)^{-1}$). The functions $\varphi_{j,k}$ are analytic on $\mathbb{C}$, and according to Liouville's theorem (see, e.g., [1, p. 122]) $\varphi_{j,k}$ is unbounded or constant. If at least one of the functions $\varphi_{j,k}$ is unbounded, then $\|(zB - A)^{-1}\| \geq |\varphi_{j,k}(z)| \to \infty$ for some $|z| \to \infty$, and from (2.5) it follows that $\Lambda_\varepsilon(A, B) \neq \emptyset$ for all $\varepsilon > 0$. Otherwise, if all $\varphi_{j,k}$'s are constant, the entries of the matrix $(zB - A)^{-1}$ do not depend on z, which implies $B = O$.

□

Another important property of $\Lambda_\varepsilon(A)$ is that $\Lambda_\varepsilon(A)$ contains the disks with radius ε around the eigenvalues [18]; for normal matrices A the set $\Lambda_\varepsilon(A)$ is the union of these disks, but in general $\Lambda_\varepsilon(A)$ can be much larger (compare the examples presented above, or the pictures in [13, 14, 18], and in Section 3.2). For $B \neq O$ this inclusion can be generalized to

$$\Lambda_\varepsilon(A, B) \;\supset\; \bigcup\{D[\lambda, \varepsilon/\|B\|] : \lambda \in \Lambda_0(A, B)\}.$$

From our examples we conclude that $\Lambda_\varepsilon(A, B)$ can be much larger than the union of the disks with radius $\varepsilon/\|B\|$ around $\Lambda_0(A, B)$ — even when both A and B are normal.

For nonsingular matrices B the equality $\Lambda_0(A, B) = \Lambda_0(B^{-1}A, I)$ holds; one might ask whether there is a relation between $\Lambda_\varepsilon(A, B)$ and $\Lambda_\varepsilon(B^{-1}A)$ (the latter set is considered in [15]). From (2.1) one easily obtains the inclusions

$$\Lambda_{\varepsilon/\|B\|}(B^{-1}A) \;\subset\; \Lambda_\varepsilon(A, B) \;\subset\; \Lambda_{\varepsilon\|B^{-1}\|}(B^{-1}A), \tag{2.7}$$

which implies, e.g., $\Lambda_\varepsilon(A, B) = \Lambda_\varepsilon(B^{-1}A)$ for unitary matrices B.

In general it is not possible to determine $\Lambda_\varepsilon(A, B)$ analytically, and in practice one may compute $\Lambda_\varepsilon(A, B)$ numerically in some interesting region $V \subset \mathbb{C}$; usually V is a rectangle containing $\Lambda_0(A,B)$ (cf. Section 2.2). In the following we will discuss how the inclusion $\Lambda_\varepsilon(A,B) \subset V$ can be verified. In view of (2.5) this amounts to computing $\sup\{\|(\lambda B - A)^{-1}\| : \lambda \notin V\}$. In order to compute this quantity we introduce the constants

$$M_{\partial V} = \sup_{\lambda \in \partial V} \|(\lambda B - A)^{-1}\| \quad \text{and} \quad M_\infty = \lim_{|\lambda| \to \infty} \|(\lambda B - A)^{-1}\|$$

(∂V denotes the boundary of V). The quantity $M_\infty \in [0,\infty]$ always exists; for nonsingular matrices B one has $M_\infty = 0$, while $M_\infty \geq 1/\varepsilon^* > 0$ for singular matrices B. If the degree of the polynomial $p(\lambda) = \det(\lambda B - A)$ equals the rank of B, then M_∞ is finite; otherwise $M_\infty = \infty$ is possible (compare our examples).

Theorem 2.3. *Assume that $M_\infty < \infty$, $\Lambda_0(A,B) \subset V \backslash \partial V$ and $\mathbb{C} \backslash V$ is open and connected. Then*

$$\Lambda_\varepsilon(A,B) \subset V \quad \textit{for} \quad \varepsilon \leq 1/\max\left(M_{\partial V}, M_\infty\right).$$

Proof (suggested by L.N. Trefethen [19]). Define for $u, v \in \mathbb{C}^N$, $\|u\| = \|v\| = 1$, the function

$$\psi_{u,v}(z) = v^*(zB - A)^{-1}u.$$

This function is analytic outside $\Lambda_0(A,B)$, and from the maximum principle (see, e.g., [1, p. 134]) it follows that

$$\sup_{z \notin V} |\psi_{u,v}(z)| \;=\; \max\left(\sup_{z \in \partial V} |\psi_{u,v}(z)|,\ \lim_{|z| \to \infty} |\psi_{u,v}(z)|\right),$$

which implies

$$\sup_{z \notin V} \|(zB - A)^{-1}\| \;=\; \max\left(\sup_{z \in \partial V} \|(zB - A)^{-1}\|,\ \lim_{|z| \to \infty} \|(zB - A)^{-1}\|\right),$$

and the result follows from (2.5).

□

More results, examples, and discussion on ε-pseudospectra for matrices can be found in [18, 20].

Remark 2.4. In this paper we only deal with the spectral norm, although $\Lambda_\varepsilon(A,B)$ could be defined also for matrix norms $\|\cdot\|$ generated by an arbitrary vector norm on $\mathbb{C}^N$ (cf. [6]). The equivalence of (2.1) – (2.5), Theorem 2.2, and (2.7), remain valid in this case.

2.2 The computation of pseudospectra

Most techniques to compute $\Lambda_\varepsilon(A)$ numerically, can easily be adapted for $\Lambda_\varepsilon(A,B)$. Suppose one would like to know $\Lambda_\varepsilon(A,B)$ in some region $V \subset \mathbb{C}$. The set $\Lambda_\varepsilon(A,B)$ can be determined as follows (compare [18, 20]): compute for several $z \in V$ the quantity $\sigma_{\min}(z)$, the smallest singular value of $zB - A$, and draw some level curves of $\sigma_{\min}(z)$. In most applications V is a rectangle containing $\Lambda_0(A,B)$, and $\sigma_{\min}(z)$ is computed on a rectangular grid. When all entries of A and B are real, one can use the fact that $\Lambda_\varepsilon(A, B)$ is symmetric with respect to the real axis. The procedure described above has been used to produce the figures in Section 3.2.

For small N (the order of the matrices A and B) one can use e.g. MATLAB or LAPACK routines to compute $\sigma_{\min}(z)$. Unfortunately, this is not practical for large (sparse) matrices A and B, and the quantity $\sigma_{\min}(z)$ has to be computed in a different way. In [4] they use a variant of Davidson's method to compute the smallest eigenvalue of $(zI - A)^*(zI - A)$, and in [3] a continuation method is used to determine level curves of $\|(zI - A)^{-1}\|$; both techniques can be adapted for $\Lambda_\varepsilon(A, B)$.

Another approach is based on the following idea (cf. [17]): determine an $N \times m$ matrix W with $W^*W = I$, $m \ll N$, and approximate $\Lambda_\varepsilon(A, B)$ by $\Lambda_\varepsilon(W^*AW, W^*BW)$; the latter set can be computed easily. Different choices for W have been considered for $B = I$; a possibility described in [17] is to choose W as an orthonormal basis of some eigenvectors corresponding to the eigenvalues with maximal real part.

3 An application of pseudospectra

3.1 Asymptotic stability of equilibria and pseudospectra

In the study of time-dependent ordinary, or partial differential equations, one is often interested in the occurrence of equilibria and their stability behaviour, and how these equilibria and their stability properties depend on some parameter(s) in the model. Consider the systems of differential equations

$$BU'(t) = F(U(t), \mu) \quad \text{for} \quad t \geq 0, \tag{3.1}$$

with $B \in \mathbb{R}^{N,N}$, $\mu \in \mathbb{R}$, $U(t) \in \mathbb{R}^N$ for $t \geq 0$, and $F : \mathbb{R}^{N+1} \to \mathbb{R}^N$ is a smooth function. Here μ stands for a certain parameter. Although the solutions to (3.1) depend on μ, we will not express this dependence in order to keep the notation transparent. Also (systems of) time-dependent partial differential equations lead to (3.1), after discretizing the spatial derivative(s). Finite element discretizations often lead to (3.1) with $B \neq I$; problems in fluid dynamics often lead to problems of type (3.1) with singular B (cf., e.g., [5]).

The equilibria U_{eq} (sometimes called stationary solutions or steady states [5, 16]) to (3.1) satisfy $F(U_{\text{eq}}, \mu) = 0$. The equilibrium U_{eq} is called *asymptotically stable* [16] when $\lim_{t\to\infty} U(t) = U_{\text{eq}}$ for all $U(0)$ close to U_{eq}. In order to investigate whether U_{eq} is asymptotically stable, equation (3.1) is linearized around U_{eq} (note that $F(U_{\text{eq}}, \mu) = 0$):

$$BU'(t) = F'(U_{\text{eq}}, \mu) \cdot (U(t) - U_{\text{eq}}),$$

where $F'(y, \mu) = (\partial F_i(y, \mu)/\partial y_j) \in \mathbb{R}^{N,N}$ stands for the Jacobian matrix of the function F (with respect to y) evaluated at (y, μ). If all eigenvalues of the problem

$$F'(U_{\text{eq}}, \mu)v = \lambda Bv \tag{3.2}$$

have strictly negative real parts, then, when certain technical conditions are satisfied (cf., e.g., [16, p. 20]), the equilibrium U_{eq} to (3.1) is asymptotically stable. When at least one eigenvalue of (3.2) has a positive real part, the equilibrium U_{eq} is unstable [16, p. 20]. Hence investigating the asymptotic stability of equilibria amounts to compute the eigenvalues (with largest real part) of (3.2).

In practice one computes a solution (y, μ) satisfying $F(y, \mu) = 0$, by using (a variant of) Newton's method. It might be quite cumbersome to compute (and code) the Jacobian matrix $F'(y, \mu)$, so the question arises whether determination of $F'(y, \mu)$ can be avoided. A standard approach, when applying Newton's method, is to replace the entries of $F'(y, \mu)$ by difference quotients (cf. [16]). For small problems (i.e. N is small) this approach is feasible. When applying Newton's method for large problems, the resulting linear equations are often solved iteratively, e.g. by a Krylov subspace method [7]. Most Krylov methods do not need $F'(y, \mu)$ explicitly; only matrix-vector products $F'(y, \mu)v$ are required, and one can use the fact that

$$F'(y, \mu)\, v \;\approx\; \{F(y + hv, \mu) - F(y, \mu)\}/h \tag{3.3}$$

for small $h > 0$. In the actual execution of Newton's algorithm, the computation of $F'(y, \mu)v$ can be replaced by the right hand side of (3.3) [2, 7].

Suppose that some equilibria have been computed without using the exact Jacobian matrix $F'(U_{\mathrm{eq}}, \mu)$. One might ask whether it is possible to investigate the asymptotic stability of these equilibria *without* using $F'(U_{\mathrm{eq}}, \mu)$ explicitly. Let the matrix A be an approximation to $F'(U_{\mathrm{eq}}, \mu)$ (for example, the jth column of A might be the difference quotient in (3.3) with $v = e_j$, the jth unit vector) and consider the eigenvalue problem

$$Av = \lambda Bv \tag{3.4}$$

instead of (3.2). Now the question arises how we can get relevant information on the spectrum of (3.2) from (3.4). In order to answer this question, the concept of ε-pseudospectra is useful. The matrix $F'(U_{\mathrm{eq}}, \mu) = A + (F'(U_{\mathrm{eq}}, \mu) - A)$ can be seen as a perturbation of A. From (2.1) it follows that

$$\Lambda_0(F'(U_{\mathrm{eq}}, \mu), B) \subset \Lambda_\varepsilon(A, B) \;\text{ with }\; \varepsilon = \|F'(U_{\mathrm{eq}}, \mu) - A\|. \tag{3.5}$$

Hence when $\Lambda_\varepsilon(A, B)$ — with $\varepsilon = \|F'(U_{\mathrm{eq}}, \mu) - A\|$ — lies in the interior of the left half plane, we know that all eigenvalues to (3.2) have strictly negative real parts. In these situations we are able to conclude that the equilibrium U_{eq} is asymptotically stable, without computing any eigenvalue of (3.2)!

In order to make this practical, we need to estimate $\|F'(U_{\mathrm{eq}}, \mu) - A\|$ *without* computing the entries of $F'(U_{\mathrm{eq}}, \mu)$. Let Q be an $N \times N$ matrix with $Q^*Q = I$ and denote the jth column of Q by q_j. Assume that A is (implicitly) defined by

$$Aq_j = \{F(U_{\mathrm{eq}} + h_j q_j, \mu) - F(U_{\mathrm{eq}}, \mu)\}/h_j = F(U_{\mathrm{eq}} + h_j q_j, \mu)/h_j$$

for some small $h_j > 0$ (take e.g. $Q = I$). From $Q^*Q = I$ one obtains the estimate

$$\|F'(U_{\mathrm{eq}}, \mu) - A\| \le \sqrt{N} \cdot \max_{1 \le j \le N} \|F'(U_{\mathrm{eq}}, \mu)q_j - Aq_j\|; \tag{3.6}$$

the constant $\sqrt{N}$ in (3.6) cannot be replaced by a smaller constant. In order to estimate $\|F'(U_{\rm eq},\mu)q_j - Aq_j\|$, we note that Aq_j is an approximation to the directional derivative $F'(U_{\rm eq},\mu)q_j$; the error $F'(U_{\rm eq},\mu)q_j - Aq_j$ can be expressed in terms of second order partial derivatives of F. However, our goal was to avoid first order partial derivatives, so we certainly do not want to compute second order derivatives; these will also be approximated by finite difference quotients. From a Taylor expansion it follows that (assuming that F is sufficiently smooth)

$$F'(U_{\rm eq},\mu)\,q_j \;-\; Aq_j \;= \tag{3.7}$$
$$-\frac{h_j}{2}\left(\frac{F(U_{\rm eq}-h_jq_j,\mu)\;-\;2F(U_{\rm eq},\mu)\;+\;F(U_{\rm eq}+h_jq_j,\mu)}{{h_j}^2}\right)\;+\;\mathcal{O}({h_j}^2).$$

Hence one sees that

$$\tilde{\varepsilon} = \sqrt{N}\cdot\max_{1\le j\le N}\left\|-\frac{h_j}{2}\left(\frac{F(U_{\rm eq}-h_jq_j,\mu)\;-\;2F(U_{\rm eq},\mu)\;+\;F(U_{\rm eq}+h_jq_j,\mu)}{{h_j}^2}\right)\right\| \tag{3.8}$$

might be a reasonable estimate for an upper bound for $\|F'(U_{\rm eq},\mu)-A\|$. Note that it takes N additional function evaluations to compute $\tilde{\varepsilon}$.

In view of (3.7) and (3.8), one would like to choose h_j close to zero. However, rounding errors may effect the accuracy of the computed vector Aq_j whenever h_j is too small; rounding errors in $F(U_{\rm eq}+h_jq_j,\mu) - F(U_{\rm eq},\mu)$ are multiplicated by a factor $1/h_j$. The choice $h_j = \mathcal{O}(\sqrt{\xi})$, with ξ the machine precision, seems reasonable (see, e.g., [16, p. 38]).

In order to determine whether equilibria are asymptotically stable we may proceed as follows. Compute the eigenvalues of (3.4). Only when all these eigenvalues have strictly negative real parts, we have to investigate whether the equilibrium under consideration is really asymptotically stable. In order to validate this, one determines $\Lambda_\varepsilon(A,B)$ for $\varepsilon=\tilde{\varepsilon}$. If $\Lambda_{\tilde{\varepsilon}}(A,B)$ is in the interior of the left half plane, this almost implies that the equilibrium $U_{\rm eq}$ is asymptotically stable (note that $\tilde{\varepsilon}$ estimates an upper bound for $\|F'(U_{\rm eq},\mu)-A\|$, so one is not completely certain that $\tilde{\varepsilon}\ge\|F'(U_{\rm eq},\mu)-A\|$). If $\Lambda_\varepsilon(A,B)$ with $\varepsilon=\|F'(U_{\rm eq},\mu)-A\|$ is not entirely in the left half plane, then this does not need to imply that $U_{\rm eq}$ is unstable (compare (3.5)).

For large problems, it is not possible to compute all eigenvalues of (3.4). In such cases one attempts to compute only the eigenvalues with largest real parts; special algorithms for computing these eigenvalues have been developed (see [5, 10] and the references cited therein). The so-called Jacobi-Davidson QZ method [8] might also be suitable to determine some of the eigenvalues with largest real part. For large problems one may also use ε-pseudospectra to investigate the asymptotic stability of equilibria. Since we are only interested in $\Lambda_\varepsilon(A,B)$ for one single value ε (viz. $\varepsilon=\tilde{\varepsilon}$), the ideas described in [3] are useful for computing the boundary of $\Lambda_{\tilde{\varepsilon}}(A,B)$ (cf. Section 2.2).

Remark 3.1. The concept of ε-pseudospectra for matrix pencils is also useful for investigating the stability of time-stepping methods for linear initial value problems $B\,U'(t) = AU(t)$ with nonsingular matrix B. Formally one applies such time-stepping methods, e.g.

Runge-Kutta methods or linear multistep methods, to $U'(t) = B^{-1}AU(t)$. Roughly speaking, the numerical process for solving $U'(t) = B^{-1}AU(t)$ is stable if $\Lambda_\varepsilon(B^{-1}A)$ satisfies a certain condition with respect to the so-called stability region of the time-stepping method (see [13, 14]). By using the left inclusion in (2.7), this condition, dealing with $\Lambda_\varepsilon(B^{-1}A)$, can be replaced by a requirement for $\Lambda_\varepsilon(A, B)$; hence the theory in [13, 14] can be applied *without* determining $B^{-1}A$.

3.2 An illustration

As an illustration to the theory of Section 3.1 we consider Burgers' equation

$$(3.9)\qquad \begin{cases} \dfrac{\partial u}{\partial t}(x,t) \;+\; \dfrac{1}{2}\dfrac{\partial}{\partial x}\left[u(x,t)^2\right] \;=\; \mu\,\dfrac{\partial^2 u}{\partial x^2}(x,t), \quad 0 \le x \le 1, \quad t \ge 0, \\[2ex] u(0,t) = u(1,t) = 1, \end{cases}$$

where μ is a positive constant. Note that $u \equiv 1$ is a stationary solution to (3.9), and this solution is stable in the sense that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_0^1 |u(x,t)-1|^2\,\mathrm{d}x \;=\; -2\mu\int_0^1 \left[\tfrac{\partial u}{\partial x}(x,t)\right]^2\mathrm{d}x \;<\; 0$$

whenever u satisfies (3.9) and $u \not\equiv 1$.

We will consider two different schemes for discretizing the spatial derivatives in (3.1), viz. a finite difference and a finite element scheme. The scheme based on finite difference discretization leads to a problem of type (3.1) with $B = I$, while finite element discretization leads to (3.1) with $B \neq I$. For both methods an equidistant grid is used; for $N \in \mathbb{N}$ we define $\Delta x = 1/(N+1)$ (the mesh size) and $x_j = j\Delta x$ ($j = 1, 2, \ldots, N$).

In the finite difference scheme the diffusion term $\mu\frac{\partial^2 u}{\partial x^2}$ has been discretized by second order central differences, and for the non-linear advection term $\frac{1}{2}\frac{\partial}{\partial x}[u^2]$ we have used the upwind discretization $\frac{\partial}{\partial x}[u(x_j,t)^2] \approx \left(u(x_j,t)^2 - u(x_{j-1},t)^2\right)/\Delta x$. This leads to a system of type (3.1), with $B = I$ and $U_j(t) \approx u(x_j,t)$. Note that $U_{\mathrm{eq}} = (1,1,\ldots,1)^{\mathrm{T}}$ is an equilibrium for all $\mu > 0$.

In our numerical experiments (with MATLAB) we set $N = 64$, $\mu = \frac{1}{100}$ or $\mu = \frac{1}{1000}$, $U_{\mathrm{eq}} = (1,1,\ldots,1)^{\mathrm{T}}$, $h = h_j = 10^{-6}$ and $Q = I$. We found $\tilde{\varepsilon} = 3.7\cdot10^{-4}$ for $\mu = \frac{1}{100}$ as well as $\mu = \frac{1}{1000}$; note that $F(y,\mu)$ can be written as $F(y,\mu) = G(y) + \mu C y$ with a fixed matrix C, and therefore $\tilde{\varepsilon}$ does not depend on μ. Because all components of G are at most quadratic in the variables y_j, one can show that $F'(U_{\mathrm{eq}},\mu)q_j - Aq_j$ equals the right-hand side of (3.7) except for the $\mathcal{O}({h_j}^2)$ term. Hence $\tilde{\varepsilon}$ is an upper bound for $\|F'(U_{\mathrm{eq}},\mu) - A\|$. The quantity $\|F'(U_{\mathrm{eq}},\mu) - A\|$ has also been computed; note that $\|F'(U_{\mathrm{eq}},\mu) - A\|$ does not depend on μ, because μ appears only in the linear part of F. We found $\|F'(U_{\mathrm{eq}},\mu) - A\| = 6.5\cdot10^{-5} < \tilde{\varepsilon}$.

In Figure 1 the boundaries of the sets $\Lambda_\varepsilon(A)$ have been plotted for $\varepsilon = 10^{-1}$, $\tilde{\varepsilon}$ and $\varepsilon = \|F'(U_{\mathrm{eq}},\mu) - A\|$. For $\mu = \frac{1}{1000}$ and $\mu = \frac{1}{100}$, we conclude from our experiments that $\Lambda_{\tilde{\varepsilon}}(A)$ is in the interior of the left half plane, which implies that U_{eq} is asymptotically stable; this is nicely in agreement with the stability of the stationary solution $u \equiv 1$ to (3.9).

Finally, we note that the eigenvalues of $F'(U_{\text{eq}}, \mu)$ can be computed analytically (compare [11, p. 10]); for each $\mu > 0$ and $N \in \mathbb{N}$, these eigenvalues are real and negative. The eigenvalues of $F'(U_{\text{eq}}, \mu)$ are also displayed in Figure 1. This example illustrates nicely how ε-pseudospectra can be used to investigate the asymptotic stability of equilibria.

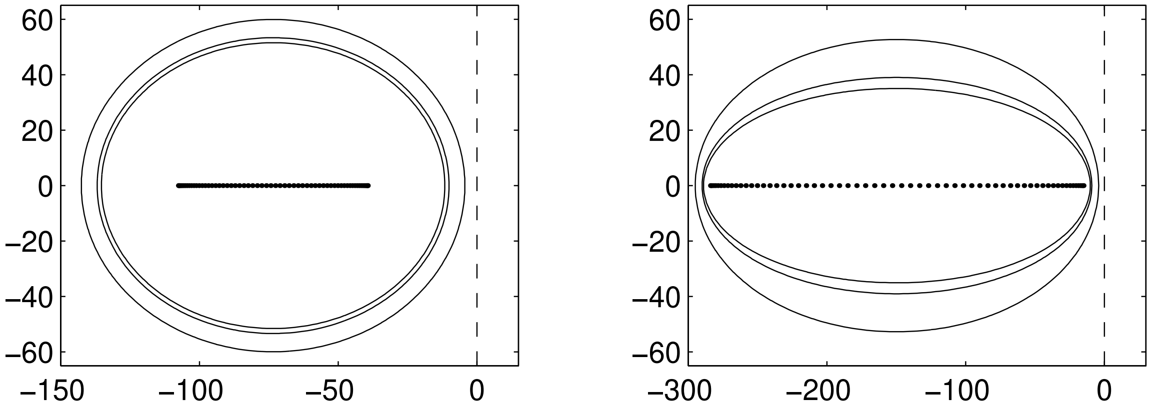

Figure 1. *The sets* $\Lambda_\varepsilon(A)$ *for* $\varepsilon = 10^{-1}$, $\tilde{\varepsilon}$, $\|F'(U_{\text{eq}}, \mu) - A\|$. *The left figure corresponds to* $\mu = \frac{1}{1000}$, *the right one to* $\mu = \frac{1}{100}$. *The dots are the eigenvalues of* $F'(U_{\text{eq}}, \mu)$, *and the imaginary axis is indicated by a dashed line.*

In order to illustrate the use of ε-pseudospectra for matrix pencils, we have also discretized the spatial derivatives in (3.9) with a finite element method. Piecewise linear elements with basis functions φ_j, given by $\varphi_j(x_i) = 0$, for $i \neq j$, and $\varphi_j(x_j) = 1$, are used. This leads to (3.1) with an $N \times N$ mass matrix $B = \frac{1}{6}\Delta x \cdot \text{tridiag}(1, 4, 1)$ and $U_j(t) \approx u(x_j, t)$.

In this case we also have $F(U_{\text{eq}}, \mu) = 0$ for all $\mu > 0$ and U_{eq} as above. Just as in the finite difference case we take $N = 64$, $\mu = \frac{1}{100}$ or $\mu = \frac{1}{1000}$, $U_{\text{eq}} = (1, 1, \ldots, 1)^{\mathrm{T}}$, $h = h_j = 10^{-6}$ and $Q = I$. Again the function $F(y, \mu)$ can be written as $F(y, \mu) = G(y) + \mu C y$, and the components of G are polynomials of degree at most two in y_j. This implies again that $\tilde{\varepsilon}$ is an upper bound for $\|F'(U_{\text{eq}}, \mu) - A\|$ (no $\mathcal{O}({h_j}^2)$ has been neglected). Both $\tilde{\varepsilon}$ and $\|F'(U_{\text{eq}}, \mu) - A\|$ do not depend on μ; we found $\tilde{\varepsilon} = 1.9 \cdot 10^{-6}$ and $\|F'(U_{\text{eq}}, \mu) - A\| = 3.3 \cdot 10^{-7}$ respectively.

In Figure 2 the boundaries of the sets $\Lambda_\varepsilon(A, B)$ have been plotted for $\varepsilon = 10^{-1}$, 10^{-3}, $\tilde{\varepsilon}$, and $\varepsilon = \|F'(U_{\text{eq}}, \mu) - A\|$. For $\mu = \frac{1}{1000}$ and $\mu = \frac{1}{100}$, we conclude from our experiments that $\Lambda_{\tilde{\varepsilon}}(A, B)$ is in the interior of the left half plane, which implies that the equilibrium U_{eq} is asymptotically stable, as in the finite difference case. Furthermore one can prove that for $\mu > 0$ and $N \in \mathbb{N}$ all eigenvalues $\lambda \in \Lambda_0(F'(U_{\text{eq}}, \mu), B)$ have strictly negative real parts; this follows from $\text{Re}\{(v^* F'(U_{\text{eq}}, \mu) v) \,/\, (v^* B v)\} < 0$ for all $v \in \mathbb{C}^N \backslash \{0\}$. This example provides a nice application of ε-pseudospectra for matrix pencils.

The set $\Lambda_0(F'(U_{\mathrm{eq}},\mu),B)$ has been computed numerically, and these eigenvalues are also displayed in Figure 2. For $\mu=\frac{1}{100}$, the computed eigenvalues are not symmetric with respect to the real axis; this behaviour is due to rounding errors.

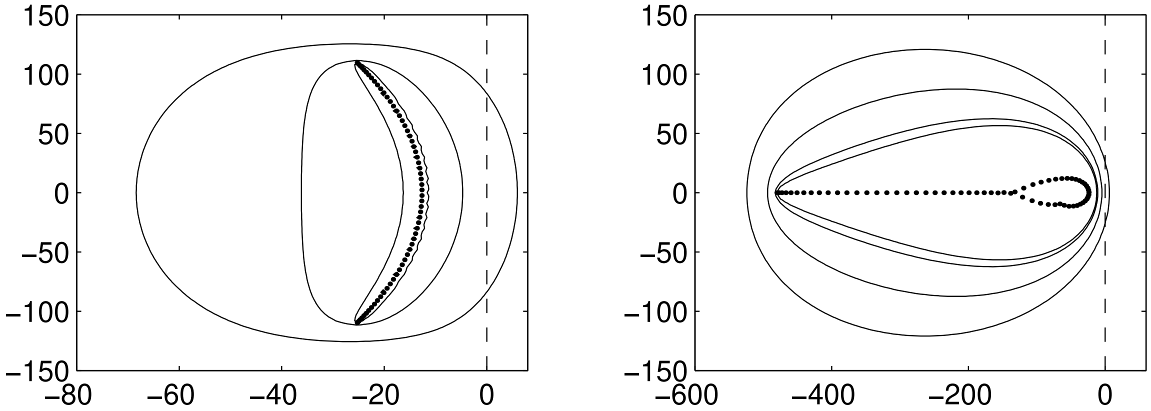

Figure 2. *The sets $\Lambda_\varepsilon(A,B)$ for $\varepsilon=10^{-1},\ 10^{-3},\ \hat{\varepsilon},\ \|F'(U_{\mathrm{eq}},\mu)-A\|$. The left figure corresponds to $\mu=\frac{1}{1000}$, the right one to $\mu=\frac{1}{100}$. The set $\Lambda_\varepsilon(A,B)$ with $\varepsilon=\|F'(U_{\mathrm{eq}},\mu)-A\|$ is not visible in the left picture. The dots are the eigenvalues of $F'(U_{\mathrm{eq}},\mu)v=\lambda Bv$, and the imaginary axis is indicated by a dashed line.*

Our experiments have also been carried out with $N=128$, instead of $N=64$. In all four cases we found that the corresponding sets $\Lambda_\varepsilon(A,B)$ are in the interior of the left half plane, implying that the equilibrium $U_{\mathrm{eq}}=(1,1,\ldots,1)^{\mathrm{T}}$ is asymptotically stable.

Acknowledgements. The author wishes to thank Gerard Sleijpen and Henk van der Vorst for valuable suggestions concerning the presentation of the paper.

References

[1] L.V. Ahlfors: *Complex Analysis*, 3rd Ed. Singapore: McGraw-Hill, 1979.

[2] P.N. Brown & Y. Saad: Hybrid Krylov methods for nonlinear systems of equations. *SIAM J. Sci. Stat. Comput.* **11**, 450-481 (1990).

[3] M. Brühl: A curve tracing algorithm for computing the pseudospectrum. Submitted for publication (1995).

[4] J.F. Carpraux, J. Erhel & M. Sadkane: Spectral portrait for non-Hermitian large sparse matrices. *Computing* **53**, 301-310 (1994).

[5] H.A. Dijkstra, M.J. Molemaker, A. van der Ploeg & E.F.F. Botta: An efficient code to compute non-parallel steady flows and their linear stability. *Computers and Fluids* **24**, 415-434 (1995).

[6] J.L.M. van Dorsselaer, J.F.B.M. Kraaijevanger & M.N. Spijker: Linear stability analysis in the numerical solution of initial value problems. *Acta Numerica 1993*, 199-237 (1993).

[7] D.R. Fokkema, G.L.G. Sleijpen & H.A. van der Vorst: Accelerated inexact Newton schemes for large systems of nonlinear equations. Preprint nr. 918, Dept. Math., Utrecht University, Utrecht, July 1995.

[8] D.R. Fokkema, G.L.G. Sleijpen & H.A. van der Vorst: Jacobi-Davidson style QR and QZ algorithms for the partial reduction of matrix pencils. Preprint nr. 941, Dept. Math., Utrecht University, Utrecht, January 1996.

[9] A. Greenbaum, V. Pták & Z. Strakoš: Any nonincreasing convergence curve is possible for GMRES. To appear in *SIAM J. Matrix Anal. Appl.*

[10] K. Meerbergen & D. Roose: The restarted Arnoldi method applied to iterative linear system solvers for computation of rightmost eigenvalues. To appear in *SIAM J. Matrix Anal. Appl.*

[11] A.R. Mitchell: *Computational Methods in Partial Differential Equations.* London: John Wiley & Sons, 1969.

[12] N.M. Nachtigal, S.C. Reddy & L.N. Trefethen: How fast are nonsymmetric matrix iterations? *SIAM J. Matrix Anal. Appl.* **13**, 778-795 (1992).

[13] S.C. Reddy & L.N. Trefethen: Lax-stability of fully discrete spectral methods via stability regions and pseudo-eigenvalues. *Comp. Meth. Appl. Mech. Eng.* **80**, 147-164 (1990).

[14] S.C. Reddy & L.N. Trefethen: Stability of the method of lines. *Numer. Math.* **62**, 235-267 (1992).

[15] A. Ruhe: The rational Krylov algorithm for large nonsymmetric eigenvalues – mapping the resolvent norms (pseudospectrum). Unpublished note (1994).

[16] R. Seydel: *Practical bifurcation and stability analysis*, 2nd Ed. New York: Springer, 1994.

[17] K.-C. Toh & L.N. Trefethen: Calculation of pseudospectra by the Arnoldi iteration. To appear in *SIAM J. Sci. Comput.* **17** (1996).

[18] L.N. Trefethen: Pseudospectra of matrices, in *Numerical Analysis 1991* (D.F. Griffiths and G.A. Watson, eds), 234 - 266. Harlow: Longman, 1992.

[19] L.N. Trefethen: Private communication (1995).

[20] L.N. Trefethen: *Spectra and Pseudospectra: The Behavior of Non-Normal Matrices and Operators*. Book in preparation.