LOW-RANK DYNAMICS FOR COMPUTING EXTREMAL POINTS OF REAL PSEUDOSPECTRA∗

NICOLA GUGLIELMI† AND CHRISTIAN LUBICH‡

Abstract. We consider the real ε-pseudospectrum of a real square matrix, which is the set of eigenvalues of all real matrices that are ε-close to the given matrix, where closeness is measured in either the 2-norm or the Frobenius norm. We characterize extremal points and compare the situation with that for the complex ε-pseudospectrum. We present differential equations for rank-1 and rank-2 matrices for the computation of the real pseudospectral abscissa and radius. Discretizations of the differential equations yield algorithms that are fast and well suited for sparse large matrices. Based on these low-rank differential equations, we further obtain an algorithm for drawing boundary sections of the real pseudospectrum with respect to both the 2-norm and the Frobenius norm.

Key words. real pseudospectrum, real pseudospectral abscissa, real stability radius

AMS subject classifications. 15A18, 65K05

DOI. 10.1137/120862399

1. Introduction. In this paper a novel approach to computing extremal points of real pseudospectra is presented and analyzed. We are interested in computing effects of perturbations on the spectrum of a given real matrix A. The framework consists of

(i) fixing a class of perturbations (here we consider the class of real and the class of complex perturbations);

(ii) fixing a norm to measure the size of the possible perturbations (here we consider the Frobenius and the spectral norm).

We measure the largest possible spectral abscissa or radius of the perturbed spectrum or, in other words, the pseudospectral abscissa or radius. In the literature these quantities are widely studied since they allow one to analyze stability properties and robustness of stable linear dynamical systems; see, e.g., [HP05]. The most common tool is the unstructured complex pseudospectrum, which means that independently of the structure of the matrix, unstructured perturbations of a given norm are considered. When one restricts the analysis to structured perturbations, challenging mathematical difficulties arise, because the characterization in terms of singular values of the standard unstructured pseudospectrum is lost when we put constraints on the perturbations.

In this paper we focus our attention on real perturbations of real matrices, but in an appendix we consider also the case of complex perturbations for comparison. In particular we are able to exploit a low-rank property of critical perturbations, which allows us to devise efficient methods for computing the extremal points. In the literature very few papers discuss real pseudospectra (for example, [BRQ98] and [Ru06]) and their computation appears to be difficult.
For a square matrix $A$, let $\Lambda(A)$ denote its spectrum. Now let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, and let the norm $\| \cdot \|$ be the 2-norm $\| \cdot \|_2$ or the Frobenius norm $\| \cdot \|_F$ on $\mathbb{K}^{n \times n}$. For real $\varepsilon > 0$, the $\varepsilon$-pseudospectrum of a matrix $A \in \mathbb{K}^{n \times n}$ is given by

$$\Lambda_\varepsilon^{\mathbb{K},\| \cdot \|}(A) = \{ \lambda \in \mathbb{C} : \lambda \in \Lambda(A + E) \text{ for some } E \in \mathbb{K}^{n \times n} \text{ with } \| E \| \leq \varepsilon \}.$$

It holds true that

$$\Lambda_\varepsilon^{\mathbb{C},\| \cdot \|_2}(A) = \Lambda_\varepsilon^{\mathbb{C},\| \cdot \|_F}(A), \quad \text{but in general } \Lambda_\varepsilon^{\mathbb{R},\| \cdot \|_2}(A) \neq \Lambda_\varepsilon^{\mathbb{R},\| \cdot \|_F}(A).$$

(The equality of the Frobenius norm and 2-norm pseudospectra in the complex case is due to the fact that the complex perturbation matrix $E$ can be chosen to be of rank 1. For rank-1 matrices, the two norms coincide.)

In the following we will sometimes write $\Lambda_\varepsilon^{\mathbb{R},2}(A)$ and $\Lambda_\varepsilon^{\mathbb{R},F}(A)$ for brevity and omit the superscripts when they are irrelevant or clear from the context.

We will characterize critical perturbations associated with points in the pseudospectrum that have locally maximum real part. By a rightmost eigenvalue of $A$ or point in $\Lambda_\varepsilon(A)$ we mean one with largest real part. By a locally rightmost point in $\Lambda_\varepsilon(A)$ we mean a point $z$ which is the rightmost point in $\Lambda_\varepsilon(A) \cap \mathcal{N}$, where $\mathcal{N}$ is some neighborhood of $z$. The $\varepsilon$-pseudospectral abscissa of $A$ is the largest of the real parts of the elements of the pseudospectrum, i.e.,

$$\alpha_\varepsilon^{\mathbb{K},\| \cdot \|}(A) = \max\{ \Re z : z \in \Lambda_\varepsilon^{\mathbb{K},\| \cdot \|}(A) \}.$$

Similarly, the $\varepsilon$-pseudospectral radius of $A$ is

$$\rho_\varepsilon^{\mathbb{K},\| \cdot \|}(A) = \max\{ |\lambda| : \lambda \in \Lambda_\varepsilon^{\mathbb{K},\| \cdot \|}(A) \}.$$

These quantities are closely related to the stability radius, which is the perturbation size $\varepsilon$ such that $\alpha_\varepsilon^{\mathbb{K},\| \cdot \|}(A) = 0$ (for a continuous-time dynamical system) or $\rho_\varepsilon^{\mathbb{K},\| \cdot \|}(A) = 0$ (for a discrete-time dynamical system); see [SVT96] for a first algorithm to compute the 2-norm real stability radius.

We will make frequent use of the following standard perturbation result for eigenvalues; see, e.g., [Kat95, section II.1.1]. Here and in the following, we denote $\dot{\cdot} = d/dt$.

**Lemma 1.1.** Consider the differentiable $n \times n$ matrix valued function $C(t)$ for $t$ in a neighborhood of 0. Let $\lambda(t)$ be an eigenvalue of $C(t)$ converging to a simple eigenvalue $\lambda_0$ of $C_0 = C(0)$ as $t \to 0$. Let $x_0$ and $y_0$ be left and right eigenvectors, respectively, of $C_0$ corresponding to $\lambda_0$, that is, $(C_0 - \lambda_0 I)y_0 = 0$ and $x_0^*(C_0 - \lambda_0 I) = 0$. Then $x_0^*y_0 \neq 0$, and $\lambda(t)$ is differentiable near $t = 0$ with

$$\dot{\lambda}(0) = \frac{x_0^*\dot{C}(0)y_0}{x_0^*y_0}.$$

The paper is organized as follows. In section 2 we study extremal points of the Frobenius norm real pseudospectrum and derive differential equations for rank-2 matrices that lead to the pseudospectral abscissa. The same program is carried out in section 3 for the 2-norm real pseudospectrum. Section 4 presents algorithms for computing the real pseudospectral abscissa with respect to both the Frobenius norm and the 2-norm. They are obtained by using an adaptive time stepping for the differential equations of sections 2 and 3. This yields a sequence of real perturbations $E_n$ of rank 2 such that the real parts of the rightmost eigenvalues of $A + \varepsilon E_n$ grow monotonically.
with $n$ and tend to a locally rightmost point in the real $\varepsilon$-pseudospectrum of $A$. Numerical examples with dense and sparse matrices are presented in section 5. Section 6 gives the analogous differential equations for the pseudospectral radius. Section 7 presents a method for drawing sections of real pseudospectra, which is based on an adaptation of the algorithm to compute the pseudospectral radius. To the best of our knowledge, so far there have not been any algorithms presented for drawing the contour of real Frobenius pseudospectra, and those known for 2-norm real pseudospectra (see [KKK10]) can only be applied to matrices of small size. In an appendix, some analogous results for the case of complex pseudospectra are given for comparison.

2. Frobenius norm real pseudospectra.

2.1. Low-rank property of extremizers and steepest ascent dynamics.

2.1.1. Statement of results for Frobenius norm extremizers. We begin by stating the main results of this subsection.

Theorem 2.1. For $A \in \mathbb{R}^{n \times n}$, let $\lambda_*$ be a locally rightmost point in the Frobenius norm real pseudospectrum $\Lambda^2_F(A)$. Let $E_*$ be a real matrix of unit Frobenius norm such that $\lambda_*$ is an eigenvalue of $A + \varepsilon E_*$. If $\lambda_*$ is a simple eigenvalue, then $E_*$ has the following low-rank property:

(a) If $\lambda_*$ is real, then $E_*$ has rank 1.
(b) If $\text{Im} \lambda_* \neq 0$, then $E_*$ has rank 2.

The result will be proven in the course of this subsection using a steepest ascent differential equation. We will also obtain the following characterization of extremizers.

Theorem 2.2. Let $E_*$ be a real matrix of unit Frobenius norm. Let $\lambda_*$ be a simple eigenvalue of $A + \varepsilon E_*$, with left and right eigenvectors $x$ and $y$, respectively, both of unit norm and with $x^* y > 0$. Then the following two statements are equivalent:

1. Every differentiable path $(E(t), \lambda(t))$ (for small $t \geq 0$) such that $\|E(t)\|_F \leq 1$ and $\lambda(t)$ is an eigenvalue of $A + \varepsilon E(t)$, with $E(0) = E_*$ and $\lambda(0) = \lambda_*$, has $\Re \lambda(0) \leq 0$.
2. $E_*$ is a positive multiple of $\Re(x^* y)$.

There is a rotated version of Theorem 2.2, which is useful for characterizing boundary points of $\Lambda^2_F(A)$ with an outer normal $e^{i\theta}$ and which is proved in the same way by noting that $\Re(e^{-i\theta} \lambda) = (x e^{i\theta})^* E y / (x^* y)$ and by replacing $x$ with $x e^{i\theta}$.

Corollary 2.3. Fix an angle $\theta \in \mathbb{R}$. In the situation of Theorem 2.2, the following statements are equivalent:

1. Every differentiable path $(E(t), \lambda(t))$ (for small $t \geq 0$) such that $\|E(t)\|_F \leq 1$ and $\lambda(t)$ is an eigenvalue of $A + \varepsilon E(t)$, with $E(0) = E_*$ and $\lambda(0) = \lambda_*$, has $\Re(e^{-i\theta} \lambda(0)) \leq 0$.
2. $E_*$ is a positive multiple of $\Re(e^{i\theta} x^* y)$.

2.1.2. Steepest ascent direction. Our aim is to construct a family of real matrices $A + \varepsilon E(t)$, where $\|E(t)\|_F = 1$ such that $\lim_{t \to \infty} E(t) = E_\infty$ and an eigenvalue of $A + \varepsilon E_\infty$ is a locally rightmost point of $\Lambda_\varepsilon(A)$. The perturbation matrices $E(t)$ shall be chosen such that the derivative $E(t)$ gives the maximum possible increase in $\Re \lambda(t)$ for the rightmost eigenvalue $\lambda(t)$ of $A + \varepsilon E(t)$. To comply with the constraint $\|E\|_F^2 = 1$, we must have $\frac{d}{dt} \|E(t)\|_F^2 = 0$, i.e.,

$$\langle E, \dot{E} \rangle = 0,$$

where $\langle X, Y \rangle = \sum_{i,j} x_{ij} y_{ij} = \text{trace}(X^T Y)$ denotes the inner product associated with the Frobenius norm. In view of Lemma 1.1 for $C(t) = A + \varepsilon E(t)$, we are led to the following optimization problem.
Lemma 2.4. Let $E \in \mathbb{R}^{n \times n}$ be a real matrix of unit Frobenius norm, and let $x, y \in \mathbb{C}^n$ be given nonzero complex vectors. The solution of the optimization problem

\begin{equation}
Z_* = \arg \max_{\|Z\|_F = 1, (E, Z) = 0} \text{Re} (x^* Z y)
\end{equation}

is given by

\begin{equation}
\mu Z_* = \text{Re} (x y^*) - \langle E, \text{Re} (x y^*) \rangle E,
\end{equation}

where $\mu$ is the Frobenius norm of the matrix on the right-hand side in (2.2).

Proof. The result follows on noting

\[ \text{Re} (x^* Z y) = \langle Z, \text{Re} (x y^*) \rangle \]

and the fact that the inner product with a given vector (which here is a matrix) is maximized over a subspace by orthogonally projecting the vector onto that subspace. The expression in (2.2) is the orthogonal projection of $\text{Re} (x y^*)$ to the orthogonal complement of the span of $E$.

A similar result is given in [BK04] in the analysis of condition numbers of complex eigenvalues under real perturbations.

2.1.3. Steepest ascent differential equation. Lemmas 1.1 and 2.4 suggest considering the following differential equation on the manifold of real $n \times n$ matrices of unit Frobenius norm:

\begin{equation}
\dot{E} = \text{Re} (x y^*) - \langle E, \text{Re} (x y^*) \rangle E,
\end{equation}

where $x(t)$ and $y(t)$ are left and right eigenvectors, respectively, to a simple eigenvalue $\lambda(t)$ of $A + \epsilon E(t)$, both of unit norm and with $x(t)^* y(t) > 0$. We are mainly interested in the rightmost eigenvalue of $A + \epsilon E(t)$.

We immediately have the following monotonicity result.

Theorem 2.5. Let $E(t)$ of unit Frobenius norm satisfy the differential equation (2.3). If $\lambda(t)$ is a simple eigenvalue of $A + \epsilon E(t)$, then

\begin{equation}
\text{Re} \dot{\lambda}(t) \geq 0.
\end{equation}

Proof. Though the result follows directly from Lemmas 1.1 and 2.4, we give here an explicit computation of $\text{Re}(x(t)^* \dot{E}(t)y(t))$, which is claimed to be nonnegative. Observe that (omitting again the omnipresent argument $t$)

\[
\text{Re}(x^* \text{Re}(x y^*) y) = \|\text{Re}(x y^*)\|_F^2,
\]

\[
\text{Re}(x^* Ey) = \langle E, \text{Re}(x y^*) \rangle.
\]

By the Cauchy–Schwarz inequality,

\[
|\langle E, \text{Re}(x y^*) \rangle| \leq \|E\|_F \|\text{Re}(x y^*)\|_F = \|\text{Re}(x y^*)\|_F.
\]

As a consequence, using (2.3),

\begin{equation}
\text{Re}(x^* \dot{E} y) = \|\text{Re}(x y^*)\|_F^2 - \langle E, \text{Re}(x y^*) \rangle^2 \geq 0,
\end{equation}

implying (2.4) by Lemma 1.1.
2.1.4. Stationary points. The stationary points of (2.3) are characterized as follows.

**Theorem 2.6.** The following statements are equivalent along solutions of (2.3):
1. Re ˙λ = 0.
2. ˙E = 0.
3. E is a real multiple of Re(xy∗).

**Proof.** This follows directly from (2.5) and Lemma 1.1.

2.1.5. Proof of Theorem 2.2. We first show that the negation of statement 1 implies the negation of 2. If there is some path E(t) through E∗ such that Re ˙λ(0) > 0, then the maximization property of Lemma 2.4, together with Lemma 1.1, shows that also the solution path of (2.3) passing through E∗ is such a path. Hence E∗ is not a stationary point of (2.3), and Theorem 2.6 then yields that E∗ is not a real multiple of Re(xy∗).

Conversely, if E∗ is not a real multiple of Re(xy∗), then E∗ is not a stationary point of (2.3), and Theorems 2.6 and 2.5 yield that Re ˙λ(0) > 0 along the solution path of (2.3). Further, if E∗ = −µRe(xy∗) with µ > 0, then along the path E(t) = e−tE∗ we have that Re(x∗E(0)y) = µ||Re(xy∗)||2 > 0, and hence, by Lemma 1.1, Re ˙λ(0) > 0 in contradiction to statement 1.

2.1.6. Proof of Theorem 2.1. By Theorem 2.2, rank(E∗) = rank(Re(xy∗)) ≤ 2. Let x and y be left and right eigenvectors, respectively, to the eigenvalue λ∗ of A + εE∗, both of unit norm and normalized such that x∗y > 0.

(a) If λ∗ is real, then x and y can be chosen real, and hence Re(xy∗) = xyT is of rank 1.

(b) We separate the real and imaginary parts in x = xR + ixI and y = yR + iyI. If Re(xy∗) = xRy∗R + xIy∗I is of rank 1, then xR and xI are linearly dependent or yR and yI are linearly dependent. Let us first assume the latter. In this case there is a real θ such that y = cos(θ)w + i sin(θ)w for some nonzero real vector w. Rotating both x and y by e−iθ does not change the required property x∗y > 0, so we can assume without loss of generality that y is a real eigenvector of the real matrix A + εE∗, which implies that the corresponding eigenvalue λ∗ is real. The argument is analogous when xR and xI are linearly dependent.

2.2. Rank-2 dynamics. Theorem 2.1 motivates a search for a differential equation on the manifold of rank-2 matrices,

\[ \mathcal{M}_2 = \{ E \in \mathbb{R}^{n \times n} : \text{rank}(E) = 2 \}, \]

which leads to the real pseudospectral abscissa when it is the real part of a point in \( \Lambda_{x,t}^{\mathbb{R},F} (A) \) that has nonvanishing imaginary part. We proceed similarly to subsection 2.1.3, but in addition require the derivative ˙E to lie in the tangent space \( T_E \mathcal{M}_2 \).

When the pseudospectral abscissa corresponds to a pseudospectral point on the real axis, the real pseudospectral abscissa equals the complex pseudospectral abscissa, and we can use the rank-1 dynamics of [GL11] (or also the rank-1 iterative method proposed in [GO11]).

2.2.1. Rank-2 matrices and their tangent matrices. In this subsection we follow [KL07]. Every real rank-2 matrix E of dimension \( n \times n \) can be written in the form

\[ E = USV^T, \]
where \( U \in \mathbb{R}^{n \times 2} \) and \( V \in \mathbb{R}^{n \times 2} \) have orthonormal columns, i.e.,
\[
(2.7) \quad U^T U = I_2, \quad V^T V = I_2
\]
(with the identity matrix \( I_2 \) of dimension 2), and \( S \in \mathbb{R}^{2 \times 2} \) is nonsingular. The singular value decomposition yields \( S \) diagonal, but here we will not assume a special form of \( S \). The representation \( (2.6) \) is not unique: replacing \( U \) by \( \tilde{U} = UP \) and \( V \) by \( \tilde{V} = VQ \) with orthogonal matrices \( P, Q \in \mathbb{R}^{2 \times 2} \), correspondingly, replacing \( S \) by \( \tilde{S} = P^T S Q \), yields the same matrix \( E = USV^T = \tilde{USV}^T \).

As a substitute for the nonuniqueness in the decomposition \( (2.6) \), we will use a unique decomposition in the tangent space. Let \( \mathcal{V}_{n,2} \) denote the Stiefel manifold of real \( n \times 2 \) matrices with orthonormal columns. The tangent space at \( U \in \mathcal{V}_{n,2} \) is
\[
(2.8) \quad \mathcal{T}_U \mathcal{V}_{n,2} = \{ \dot{U} \in \mathbb{R}^{n \times 2} : \dot{U}^T U + U^T \dot{U} = 0 \} = \{ \dot{U} \in \mathbb{R}^{n \times 2} : U^T \dot{U} \text{ is skew-symmetric} \}.
\]
As is shown in [KL07], every tangent matrix \( \dot{E} \in \mathcal{T}_E \mathcal{M}_2 \) is of the form
\[
(2.9) \quad U^T \dot{U} = 0, \quad V^T \dot{V} = 0.
\]
We note the following lemma from [KL07].

**Lemma 2.7.** The orthogonal projection onto the tangent space \( \mathcal{T}_E \mathcal{M}_2 \) at \( E = USV^T \in \mathcal{M}_2 \) is given by
\[
(2.10) \quad P_E(Z) = Z - (I - UU^T)Z(I - VV^T)
\]
for \( Z \in \mathbb{R}^{n \times n} \).

### 2.2.2. A differential equation for rank-2 matrices.

In the differential equation \( (2.3) \) we replace the right-hand side by its orthogonal projection onto \( \mathcal{T}_E \mathcal{M}_2 \), so that solutions starting with rank 2 will retain rank 2 for all times:
\[
(2.11) \quad \dot{E} = P_E \left( \text{Re} (x^y) - \langle E, \text{Re} (x^y) \rangle E \right),
\]
where again \( x(t) \) and \( y(t) \) are right and left eigenvectors, respectively, to a simple eigenvalue \( \lambda(t) \) of \( A + \varepsilon E(t) \), both of unit norm and with \( x(t)^* y(t) > 0 \).

Since \( E \in \mathcal{T}_E \mathcal{M}_2 \), we have \( P_E(E) = E \) and \( \langle E, Z \rangle = \langle E, P_E(Z) \rangle \), and hence the differential equation can be rewritten as
\[
(2.12) \quad \dot{E} = P_E(\text{Re} (x^y)) - \langle E, P_E(\text{Re} (x^y)) \rangle E,
\]
which differs from \( (2.3) \) only in that \( \text{Re} (x^y) \) is replaced by its orthogonal projection to \( \mathcal{T}_E \mathcal{M}_2 \). This also shows that \( \langle E, \dot{E} \rangle = 0 \), so that the unit norm of \( E \) is conserved along solutions of \( (2.12) \).

To obtain the differential equation in a form that uses the factors in \( E = USV^T \) rather than the full \( n \times n \) matrix \( E \), we use the following result.
LEMMA 2.8 ([KL07, Prop. 2.1]). For \( E = USV^T \in M_2 \) with nonsingular \( S \in \mathbb{R}^{2 \times 2} \) and with \( U \in \mathbb{R}^{n \times 2} \) and \( V \in \mathbb{R}^{n \times 2} \) having orthonormal columns, the equation \( \dot{E} = P_E(Z) \) is equivalent to \( \dot{E} = USV^T + U\dot{S}V^T + US\dot{V}^T \), where
\[
\dot{S} = U^T Z V, \\
\dot{U} = (I - UU^T) Z V S^{-1}, \\
\dot{V} = (I - VV^T) Z^T U S^{-T}.
\]

With \( Z = \text{Re}(xy^*) - \langle E, \text{Re}(xy^*) \rangle E \), this yields that the differential equation (2.12) for \( E = USV^T \) is equivalent to the following system of differential equations for \( S, U, V \), where we abbreviate \( g = UTx \in \mathbb{C}^2 \), \( h = VTy \in \mathbb{C}^2 \):
\[
\dot{S} = \text{Re}(gh^*) - \text{Re}(g^*Sh)S, \\
\dot{U} = \text{Re}((x - Ug)h^*)S^{-1}, \\
\dot{V} = \text{Re}((y - Vh)g^*)S^{-T}.
\]

2.2.3. Monotonicity of eigenvalues. We have the following analogue of Theorem 2.5.

THEOREM 2.9. Let \( E(t) \) satisfy the differential equation (2.12). If \( \lambda(t) \) is a simple eigenvalue of \( A + \varepsilon E(t) \), then
\[
\text{Re} \lambda(t) \geq 0.
\]

Proof. The proof closely follows that of Theorem 2.5. We note that
\[
\text{Re}(x^* P_E(\text{Re}(xy^*))y) = \langle \text{Re}(xy^*), P_E(\text{Re}(xy^*)) \rangle = \|P_E(\text{Re}(xy^*))\|^2_F
\]
and
\[
\langle E, P_E(\text{Re}(xy^*)) \rangle = \langle E, \text{Re}(xy^*) \rangle = \text{Re} \langle x^*Ey \rangle.
\]
We then have
\[
\text{Re}(x^* \dot{E}y) = \|P_E(\text{Re}(xy^*))\|^2_F - \langle E, P_E(\text{Re}(xy^*)) \rangle^2 \geq 0,
\]
which implies (2.15) by Lemma 1.1.

2.2.4. Stationary points. The stationary points of (2.12) can be characterized as in Theorem 2.6, with \( P_E(\text{Re}(xy^*)) \) in place of \( \text{Re}(xy^*) \). We then have the following preliminary result, which will be needed for proving the more favorable characterization in Theorem 2.11.

LEMMA 2.10. The following statements are equivalent along solutions of (2.12) provided that \( \lambda \) is not an eigenvalue of \( A \) and \( \lambda \notin \mathbb{R} \):
1. \( \text{Re} \dot{\lambda} = 0. \)
2. \( \dot{E} = 0. \)
3. \( E \) is a real multiple of \( P_E(\text{Re}(xy^*)) \).

Proof. Assume first that \( P_E(\text{Re}(xy^*)) \neq 0. \) If \( \text{Re} \dot{\lambda} = 0. \), then (2.16) shows that \( E \) is a real multiple of \( P_E(\text{Re}(xy^*)) \). This implies \( \dot{E} = 0 \) by (2.12) and \( \|E\|_F = 1 \). Finally, by Lemma 1.1, \( \dot{E} = 0 \) implies \( \dot{\lambda} = 0. \).

On the other hand, if \( P_E(\text{Re}(xy^*)) = 0. \), then (2.10) shows that \( \text{Re}(xy^*) = (I - UU^T)\text{Re}(xy^*)(I - VV^T) \). With \( X = (\text{Re}x, \text{Im}x), Y = (\text{Re}y, \text{Im}y) \in \mathbb{R}^{n \times 2} \) we have
\[
\text{Re}(xy^*) = XY^T,
\]
and hence the above formula translates into \( XY^T = (I - UU^T)XY^T(I - VV^T) \). Multiplying this from the right by \( V \) and from the left by \( U^T \) gives \( XY^TV = 0 \) and \( U^TXY^T = 0 \), which implies

\[
Y^TV = 0, \quad U^TX = 0.
\]

This yields that for \( E = USV^T \) we have \( Ey = 0 \) and \( x^*E = 0 \), which shows that \( Ay = \lambda y \) and \( x^*A = \lambda x^* \). This contradicts the assumption. \( \square \)

The following theorem, together with Theorem 2.6, shows that, apart from exceptional cases where the eigenvalue \( \lambda \) of \( A + \varepsilon E \) is also an eigenvalue of \( A \), the differential equations (2.3) and (2.12) have the same stationary points of rank 2.

**Theorem 2.11.** The following statements are equivalent along solutions of (2.12)

1. \( \Re \lambda = 0 \).
2. \( E = 0 \).
3. \( E \) is a real multiple of \( \Re(xy^*) \).

**Proof.** (a) If \( E \) is a stationary point of (2.3), then \( E \) is a multiple of \( \Re(xy^*) \) by Theorem 2.6. Since \( E \in T_{(E, M_2)} \), we have \( P_E(E) = E \) and hence \( P_E(\Re(xy^*)) = \Re(xy^*) \), so that \( E \) is indeed a multiple of \( P_E(\Re(xy^*)) \) and, therefore, by Lemma 2.10, a stationary point of (2.12).

(b) If \( E \) is stationary for (2.12), then Lemma 2.10 states that \( E = \mu P_E(\Re(xy^*)) \) for some real \( \mu \), and so we have

\[
\Re(x^*Ey) = (\Re(xy^*), E) = (\Re(xy^*), \mu P_E(\Re(xy^*))) = (P_E(\Re(xy^*)), \mu P_E(\Re(xy^*))) = \frac{1}{\mu} \|E\|^2_F \neq 0.
\]

Now let the real rank-2 matrix \( E \) be written as \( E = USV^T \) with invertible \( S \in \mathbb{R}^{2 \times 2} \) and \( U, V \in \mathbb{R}^{n \times 2} \) with orthonormal columns. We form the matrices \( X = (x_R, x_I) \in \mathbb{R}^{n \times 2} \) and \( Y = (y_R, y_I) \in \mathbb{R}^{n \times 2} \) from the real and imaginary parts of the left and right eigenvectors \( x \) and \( y \), respectively, and introduce the real \( 2 \times 2 \) matrices \( G = U^TX \), \( H = V^TY \). At a stationary point the first equation of (2.14) shows, since \( \Re(g^*Sh) = \Re(x^*Ey) \neq 0 \), that \( \Re(gh^*) = GH^T \) is a nonzero multiple of the invertible matrix \( S \), and hence both \( G \) and \( H \) are invertible. The second equation of (2.14) then yields that \( 0 = \Re((x - Uy)h^*) = (X - UG)H^T \) and hence \( X = UG \), that is, \( x = Uy \). Similarly, by the last equation in (2.14), \( y = Vh \). Then,

\[
(I - UU^T)\Re(xy^*) (I - VV^T) = (I - UU^T)U\Re(gh^*) V^T(I - VV^T) = 0,
\]

and hence, by Lemma 2.7,

\[
\frac{1}{\mu} E = P_E(\Re(xy^*)) = \Re(xy^*),
\]

so that \( E \) is a stationary point of (2.3) by Theorem 2.6. \( \square \)

**2.2.5. Rotated variant.** Theorems 2.9 and 2.11 hold with \( e^{-i\theta} \dot{\lambda} \) in place of \( \dot{\lambda} \) when \( xy^* \) is replaced by \( e^{i\theta}xy^* \) in the differential equation (2.12) and in item 3 of Theorem 2.11. The proof remains the same.

This allows us to draw the pseudospectrum close to extremal points, as long as the boundary remains convex and the points on the arc centered in the rightmost point of the pseudospectrum are indeed locally rightmost points of the rotated matrix \( e^{-i\theta}A \) (here the considered differential equations are not applied to a real matrix but to a complex multiple of it).
3. 2-norm real pseudospectra.

3.1. Extremizers.

3.1.1. Statement of results for 2-norm extremizers. We begin by stating the results of this subsection. We have the following analogue of Theorem 2.1.

**Theorem 3.1.** For $A \in \mathbb{R}^{n \times n}$, let $\lambda_*$ be a locally rightmost point in the 2-norm real pseudospectrum $\Lambda^2_2(A)$. Let $E$ be a real matrix of unit 2-norm such that $\lambda_*$ is an eigenvalue of $A + \varepsilon E$. If $\lambda_*$ is a simple eigenvalue, then there is a matrix $E_*$ such that $A + \varepsilon E_*$ has the eigenvalue $\lambda_*$ with the same left and right eigenvectors as $A + \varepsilon E$, and $E_*$ has the following low-rank property:

(a) If $\lambda$ is real, then $E_*$ has rank 1.

(b) If $\text{Im} \lambda \neq 0$, then $E_*$ has rank 2. Moreover, both nonzero singular values of $E_*$ are equal to 1.

A characterization of the rank-2 extremizers is given in the following 2-norm analogue of Theorem 2.2.

**Theorem 3.2.** Assume that $E_* = U V^T$, where $U, V \in \mathbb{R}^{n \times 2}$ have orthonormal columns. Let $\lambda_*$ be a simple, nonreal eigenvalue of $A + \varepsilon E_*$, with $\lambda(0) = \lambda_*$. Let $x = x_R + i x_I$ and $y = y_R + i y_I$, respectively, both of unit norm and with $x^* y > 0$. Then the following two statements are equivalent:

1. Every differentiable path $(E(t), \lambda(t))$ (for small $t \geq 0$) such that $\|E(t)\|_2 \leq 1$ and $\lambda(t)$ is an eigenvalue of $A + \varepsilon E(t)$, with $E(0) = E_*$ and $\lambda(0) = \lambda_*$, has $\text{Re} \lambda(0) \leq 0$.

2. $X$ and $U$ have the same range, $Y$ and $V$ have the same range, and the real $2 \times 2$ matrix $U^T X Y^T V$ is symmetric and positive definite.

There is the following rotated version, which is proved in the same way.

**Corollary 3.3.** Fix an angle $\theta \in \mathbb{R}$. In the situation of Theorem 3.2, the following statements are equivalent:

1. Every differentiable path $(E(t), \lambda(t))$ (for small $t \geq 0$) such that $\|E(t)\|_2 \leq 1$ and $\lambda(t)$ is an eigenvalue of $A + \varepsilon E(t)$, with $E(0) = E_*$ and $\lambda(0) = \lambda_*$, has $\text{Re} (e^{-i \theta} \lambda(0)) \leq 0$.

2. $X$ and $U$ have the same range, $Y$ and $V$ have the same range, and the real $2 \times 2$ matrix $U^T X e^{i \theta} Y^T V$ is symmetric and positive definite, where $J = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$.

3.1.2. Proof of Theorem 3.1. The real case (a) is known [GL11]. It remains to prove (b), and so we assume $\lambda_*$ to be nonreal in the following. We denote by $x = x_R + i x_I$ and $y = y_R + i y_I$ the left and right eigenvectors, respectively, both of unit norm and with $x^* y > 0$. We set

$$X = (x_R, x_I) \in \mathbb{R}^{n \times 2}, \quad Y = (y_R, y_I) \in \mathbb{R}^{n \times 2}.$$ 

Let $\hat{U}, \hat{V} \in \mathbb{R}^{n \times n}$ be orthogonal matrices whose first two columns $U, V \in \mathbb{R}^{n \times 2}$ span the range of $X, Y$, respectively:

$$X = \hat{U} \begin{pmatrix} G \\ 0 \end{pmatrix} = U G, \quad Y = \hat{V} \begin{pmatrix} H \\ 0 \end{pmatrix} = V H \quad \text{with} \ G, H \in \mathbb{R}^{2 \times 2}.$$ 

Let $E(t)$, for small $t \geq 0$, be a continuously differentiable path on the polygonal set of matrices of unit 2-norm, with $E(0) = E$, which we write as

$$E(t) = \hat{U} \hat{S}(t) \hat{V}^T, \quad \|\hat{S}(t)\|_2 = 1.$$
Let $\lambda(t)$ be the rightmost eigenvalue of $A + \varepsilon E(t)$, so that $\lambda(0) = \lambda_*$. Since $\lambda_*$ is locally rightmost, we have (with $\dot{\ } = d/dt$)

$$0 \geq \text{Re} \dot{\lambda}(0) = \frac{\varepsilon}{x^*y} \text{Re}(x^* \dot{E}(0)y)$$

and we observe, denoting by $S(t) \in \mathbb{R}^{2 \times 2}$ the left upper $2 \times 2$ block of $\dot{S}(t)$,

$$0 \geq \text{Re}(x^* \dot{E}(0)y) = x_R^T \dot{E}(0)y_R + x_I^T \dot{E}(0)y_I = \text{trace}(X^T \dot{E}(0)Y) = \text{trace}(G^T \dot{S}(0)H).$$

This inequality holds for every path $\dot{S}(t)$ of unit norm, which is possible only if $\|S(0)\|_2 = 1$ since otherwise we could choose $\dot{S}(0)$ arbitrarily. Below we will show that in fact both singular values of $S(0)$ are equal to 1. Therefore, $\dot{S}(0)$ must have the form

$$\dot{S}(0) = \begin{pmatrix} S(0) & 0 \\ 0 & B \end{pmatrix},$$

with a matrix $B$ with $\|B\|_2 \leq 1$. This yields that $E_* = US(0)V^T$ has rank 2 and $A + \varepsilon E_*$ has the same eigenvalue $\lambda_*$ as $A + \varepsilon E$ with the same left and right eigenvectors $x$ and $y$.

We prove now that $S(0)$ has rank 2. If $S(0)$ were of rank 1, it could be written as

$$S(0) = P \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} Q^T$$

with orthogonal $2 \times 2$ matrices $P, Q$. Choosing

$$S(t) = P \begin{pmatrix} 1 & 0 \\ 0 & \pm t \end{pmatrix} Q^T, P \begin{pmatrix} \sqrt{1 - t^2} & \pm t \\ 0 & 0 \end{pmatrix} Q^T, P \begin{pmatrix} \sqrt{1 - t^2} & 0 \\ \pm t & 0 \end{pmatrix} Q^T,$$

we find that the inequality $\text{trace}(G^T \dot{S}(0)H) \leq 0$ implies

$$0 = \text{trace}(G^T P e_i e_j^T Q^T H) = e_j^T Q^T H G^T P e_i \text{ for } (i, j) \in \{(2, 2), (1, 2), (2, 1)\},$$

so that

$$Q^T H G^T P = \begin{pmatrix} * & 0 \\ 0 & 0 \end{pmatrix},$$

and therefore $G$ or $H$ must be of rank 1. This yields that $x_R$ and $x_I$ are linearly dependent or $y_R$ and $y_I$ are linearly dependent. Let us first assume the latter. Then there is a real $\theta$ such that $y = \cos(\theta)w + i \sin(\theta)w$ for some real vector $w$. Rotating both $x$ and $y$ by $e^{-i\theta}$ does not change the required property $x^*y > 0$, so we can assume without loss of generality that $y$ is a real eigenvector of the real matrix $A + \varepsilon E_*$, which implies that the corresponding eigenvalue $\lambda_*$ is real. The argument is analogous when $x_R$ and $x_I$ are linearly dependent.

Therefore, $\text{Im} \lambda_* \neq 0$ implies that the rank of $E_*$ is 2. It remains to show that both singular values of $E_*$ or, equivalently, of $S(0)$, are equal to 1.

We show that a second singular value less than 1 leads to a contradiction. We write the singular value decomposition of $S(0)$ as

$$S(0) = P \begin{pmatrix} 1 & 0 \\ 0 & \sigma_2 \end{pmatrix} Q^T.$$
with orthogonal $2 \times 2$ matrices $P, Q$. Assume $\sigma_2 < 1$. We choose
\[
S(t) = P \begin{pmatrix} 1 & 0 \\ 0 & \sigma_2 \pm t \end{pmatrix} Q^T, \quad P \begin{pmatrix} a(t) & \pm t \\ 0 & b(t) \end{pmatrix} Q^T, \quad P \begin{pmatrix} a(t) & 0 \\ \pm t & b(t) \end{pmatrix} Q^T,
\]
such that in the latter two cases the matrices $S(0)$ have the singular values 1 and $\sigma_2$ and $\dot{a}(0) = \dot{b}(0) = 0$. As before, the inequality $\text{trace}(G^T \dot{S}(0) H) \leq 0$ then implies that $Q^T H G^T P$, and hence $G$ or $H$ is of rank 1. As we have seen above, this is not possible when $\text{Im} \lambda \neq 0$.

### 3.1.3. An auxiliary result.

**Lemma 3.4.**

(a) Let $B$ be a real square matrix. Then we have $\text{trace}(BZ) = 0$ for every skew-symmetric matrix $Z$ if and only if $B$ is symmetric.

(b) Let $B$ be a symmetric matrix. Then we have $\text{trace}(BM) \geq 0$ for every symmetric positive semidefinite matrix $M$ if and only if $B$ is positive semidefinite.

This is probably well known, but for the convenience of the reader we include the simple proof.

**Proof.** (a) We note that for a skew-symmetric matrix $Z$ we have
\[
\text{trace}(BZ) = \sum_{i,j} b_{ij}z_{ij} = \sum_{i<j} (b_{ij} - b_{ji})z_{ij}.
\]
(b) We diagonalize $M = QAQ^T$ with an orthogonal matrix $Q = (q_1, \ldots, q_n)$ and $\Lambda = \text{diag}(\lambda_i)$ with $\lambda_i \geq 0$ and note that
\[
\text{trace}(BM) = \text{trace}(Q^T BQ Q^T M Q) = \sum_i \lambda_i q_i^T B q_i,
\]
which yields the result. \(\square\)

### 3.1.4. Proof of Theorem 3.2.

1 implies 2: We denote by $\mathcal{R}(W)$ the range of a matrix $W$. We know from the proof of Theorem 3.1 that $\mathcal{R}(X) = \mathcal{R}(U)$, $\mathcal{R}(Y) = \mathcal{R}(V)$, and we write $X = UG$, $Y = VH$ with invertible $2 \times 2$ matrices $G, H$. We note that $B^T := GH^T = U^T X Y^T V$ is the matrix which we want to show is symmetric and positive definite. We consider the path of matrices of unit 2-norm,
\[
E(t) = U e^{tZ} V^T \quad \text{with a skew-symmetric matrix } Z,
\]
and the corresponding path of eigenvalues $\lambda(t)$ of $A + \varepsilon E(t)$ with $\lambda(0) = \lambda$. By 1, we have
\[
0 \geq \text{Re}(x^* \dot{E}(0)y) = \text{trace}(X^T \dot{E}(0)Y) = \text{trace}(G^T Z H) = \text{trace}(H G^T Z) = \text{trace}(BZ),
\]
which holds for every skew-symmetric matrix $Z$. By Lemma 3.4, $B$ is symmetric. We next consider the path
\[
E(t) = U e^{-tM} V^T \quad \text{with a symmetric positive semidefinite matrix } M
\]
consisting of matrices of 2-norm $\leq 1$. We then have by 1 that
\[
0 \geq \text{Re}(x^* \dot{E}(0)y) = -\text{trace}(BM)
\]
for every symmetric positive semidefinite matrix $M$. By Lemma 3.4, $B$ is positive semidefinite. Since $B$ is invertible, it is actually positive definite. This proves 2.
implies 1: Let \( E(t) \) be a continuously differentiable path of matrices of unit 2-norm, such that \( E(0) = UV^T \). We extend \( U, V \in \mathbb{R}^{n \times 2} \) to orthogonal matrices \( \hat{U}, \hat{V} \in \mathbb{R}^{n \times n} \) and write \( \hat{U} = (U, U^\perp) \), \( \hat{V} = (V, V^\perp) \). We write \( E(t) \) as

\[
E(t) = \hat{U}(t)\hat{S}(t)\hat{V}(t)^T,
\]

where we may choose \( \hat{U}(t) = (U(t), U^\perp) \) with \( R(U(t)) = R(U) \) and \( U^T(t)\hat{U}(t) = 0 \), and \( \hat{V}(t) = (V(t), V^\perp) \) with \( R(V(t)) = R(V) \) and \( V^T(t)\hat{V}(t) = 0 \). By 2, this implies \( X^T\hat{U}(0) = 0, Y^T\hat{V}(0) = 0 \), and therefore, with the symmetric and positive definite matrix \( B = X^TU^TY \) and with \( \hat{S}(t) \) the upper left \( 2 \times 2 \) block of \( \hat{S}(t) \),

\[
\Re(x^*\hat{E}(0)y) = \text{trace}(X^T\hat{E}(0)Y) = \text{trace}(X^TU^TY) = \text{trace}(BS(0)^T) = \frac{1}{2}\text{trace}(B(S(0) + \dot{S}(0)^T)).
\]

We claim that \( \dot{S}(0) + \dot{S}(0)^T \) is negative semidefinite. This holds because \( S(t) = e^{t\hat{S}(0)} + O(t^2) \), and a positive eigenvalue of \( \dot{S}(0) + \dot{S}(0)^T \) would imply that the 2-norm of \( e^{t\hat{S}(0)} \) grows linearly with \( t \), which is in contradiction to the fact that \( S(t) \) has norm \( \leq 1 \). With Lemma 3.4, we now conclude that

\[
\Re(x^*\hat{E}(0)y) \leq 0,
\]

which yields 1.

3.2. Rank-2 dynamics.

3.2.1. Rank-2 matrices with unit singular values and their tangent matrices. Theorem 3.1 motivates us to search for a differential equation on the manifold

\[
\mathcal{M} = \{ E \in \mathbb{R}^{n \times n} : E \text{ has rank 2 and both nonzero singular values of } E \text{ are equal to } 1 \}
\]

that leads to the real pseudospectal abscissa when it is the real part of a point in \( \Lambda(A) \) that has imaginary part greater than 0.

When the pseudospectral abscissa is given by a pseudospectral point on the real axis, the real pseudospectral abscissa equals the complex pseudospectral abscissa, and we can use the rank-1 dynamics of [GL11].

We represent matrices in \( \mathcal{M} \) in a nonunique way as

\[
E = UQV^T,
\]

where \( U, V \in \mathbb{R}^{n \times 2} \) have orthonormal columns and \( Q \in \mathbb{R}^{2 \times 2} \) is an orthogonal matrix. For a given choice of \( U, V, Q \), tangent matrices \( \dot{E} \in T_E\mathcal{M} \) can then be uniquely written as (cf. [KL07])

\[
(3.1) \quad \dot{E} = \dot{U}QV^T + U\dot{Q}V^T + UQ\dot{V}^T \quad \text{with} \quad U^T\dot{U} = 0, V^T\dot{V} = 0,
\]

\( Q^T\dot{Q} \) is skew-symmetric.

Introducing the auxiliary \( 2 \times 2 \) matrix \( \dot{Q} \) allows us to impose the very helpful orthogonality relations \( U^T\dot{U} = 0, V^T\dot{V} = 0 \).
3.2.2. A differential equation for rank-2 matrices with unit singular values. Our objective is to find differential equations for the factors \( U(t), V(t), Q(t) \) of \( E(t) = U(t)Q(t)V(t)^T \) along which the rightmost eigenvalue \( \lambda(t) \) of \( A + \varepsilon E(t) \) is monotonically increasing with \( t \) and which have attractive stationary points at which the extremality condition of Theorem 3.2 is satisfied. Assuming that \( \lambda(t) \) is simple for all \( t \) under consideration, we denote the left and right eigenvectors associated with \( \lambda(t) \) by \( x(t) \) and \( y(t) \), respectively, both of unit norm and with \( x^*(t)y(t) > 0 \). Let \( X(t) \) and \( Y(t) \) denote the real \( n \times 2 \) matrices whose columns contain the real and imaginary parts of \( x(t) \) and \( y(t) \), respectively. We omit the argument \( t \) in the following when its presence is clear from the context.

Motivated by the orthogonality relations in (3.1) and the objective that \( R(X) = R(U) \) when \( \dot{U} = 0 \) and \( R(Y) = R(V) \) when \( \dot{V} = 0 \), we make the ansatz, with \( R, S \in \mathbb{R}^{2 \times 2} \) and real \( \tau \),

\[
\begin{align*}
\dot{U} &= (I - UU^T)XR, \\
\dot{V} &= (I - VV^T)YS, \\
\dot{Q} &= Q \begin{pmatrix} 0 & \tau \\ -\tau & 0 \end{pmatrix}.
\end{align*}
\]

(3.2)

We now maximize \( \text{Re} \dot{\lambda} = \varepsilon \text{Re}(x^*\dot{E}y)/x^*y \) under the normalizing constraint that \( \dot{E} \) has unit Frobenius norm. With the abbreviations

\[
G = U^T X, \quad H = V^T Y, \quad M = X^T(I - UU^T)X, \quad N = Y^T(I - VV^T)Y
\]

we have, inserting (3.1) and (3.2),

\[
\text{Re}(x^*\dot{E}y) = \text{trace}(X^T\dot{E}Y) \\
= \text{trace}(MRHQ + G^T\dot{Q}H + G^TQS^TN) \\
= \text{trace}(QHM + HG^T\dot{Q} + S^TNG^TQ) \\
= \langle MH^TQ^T, R \rangle + \langle GH^T, \dot{Q} \rangle + \langle NG^TQ, S \rangle
\]

with the Frobenius inner product \( \langle A, B \rangle = \text{trace}(A^TB) \). We further note that

\[
\langle GH^T, \dot{Q} \rangle = \langle Q^TGH^T, \begin{pmatrix} 0 & \tau \\ -\tau & 0 \end{pmatrix} \rangle = \tau(Q^TGH^T - (Q^TGH^T)^T)_{1,2},
\]

where \( (\cdot)_{1,2} \) denotes the (1, 2)-entry of a matrix. The normalizing constraint becomes, on inserting (3.1) and (3.2),

\[
1 = \|\dot{E}\|_F^2 = \text{trace}(R^TMR) + 2\tau^2 + \text{trace}(S^TNS).
\]

We thus have an optimization problem for \( z = \text{vec}(R, S, \tau) \) of the type

\[
\max_c c^Tz \quad \text{s.t.} \quad z^TKz = 1
\]

with a symmetric positive semidefinite matrix \( K \), which has the optimality condition

\[
c - \mu Kz = 0
\]
with a positive Lagrange multiplier $\mu$. (Observe that, if $\mu$ is negative, then $c^T z = \mu z^T K z$ is also negative.)

In our situation, the optimality condition reads as

$$
\begin{pmatrix}
\mu MN \\
\mu NS \\
\frac{1}{2} \tau
\end{pmatrix}
= 
\begin{pmatrix}
MH^T Q^T \\
NG^T Q \\
2 \text{skew}(Q^T GH^T)_{1,2}
\end{pmatrix},
$$

where $\text{skew}(B) = \frac{1}{2} (B - B^T)$. Since we are interested only in the ascent direction, we ignore the scalar factor $\mu$, and so we are led to choose in (3.2)

$$
\begin{pmatrix}
R \\
S \\
\tau
\end{pmatrix}
= 
\begin{pmatrix}
H^T Q^T \\
G^T Q \\
\text{skew}(Q^T GH^T)_{1,2}
\end{pmatrix}.
$$

This gives us the following system of differential equations for $U, V, Q$:

$$
\begin{align*}
\dot{U} &= (I - UU^T)XY^TVQ^T, \\
\dot{V} &= (I - VV^T)YX^TUQ, \\
\dot{Q} &= Q \text{skew}(Q^T U^T XY^TV).
\end{align*}
$$

We recall that here $X$ and $Y$ are $n \times 2$ matrices containing the real and imaginary parts of the left and right eigenvectors $x$ and $y$, of unit norm and with $x^y > 0$, corresponding to an eigenvalue $\lambda$ of $A + \varepsilon E$ with $E = UQV^T$.

3.2.3. Monotonicity of eigenvalues. Our derivation of these differential equations immediately gives us the following monotonicity result.

**Theorem 3.5.** Let $E(t) = U(t)Q(t)V(t)^T$ with $U(t), V(t), Q(t)$ satisfying the differential equations (3.3) and with initial values such that $U(0), V(0) \in \mathbb{R}^{n \times 2}$ have orthonormal columns and $Q(0) \in \mathbb{R}^{2 \times 2}$ is an orthogonal matrix. If $\lambda(t)$ is a simple eigenvalue of $A + \varepsilon E(t)$, then

$$
\text{Re } \dot{\lambda}(t) \geq 0.
$$

3.2.4. Stationary points. As the following result shows in comparison with Theorem 3.2, stationary points of (3.3) are extremizers if $Q^T U^T XY^TV$ is positive definite. Note that $UQ$ here plays the role of $U$ in Theorem 3.2.

**Theorem 3.6.** If $\text{Im } \lambda \neq 0$, then $(U, V, Q)$ is a stationary point of (3.3) if and only if $X$ and $U$ have the same range, $Y$ and $V$ have the same range, and the real $2 \times 2$ matrix $Q^T U^T XY^TV$ is symmetric.

**Proof.** From (3.3) it is immediate that $(U, V, Q)$ is a stationary point if and only if $R(X) \subset R(U)$, $R(Y) \subset R(V)$, and $Q^T U^T XY^TV$ is symmetric. Moreover, we know from the proof of Theorem 3.1 that $X$ and $Y$ are of rank 2 when $\text{Im } \lambda \neq 0$. □

4. Algorithms for computing the real pseudospectral abscissa. In order to compute the pseudospectral abscissa we use the explicit Euler method to discretize the differential equations (2.14) and (3.3) for the Frobenius and the 2-norm case, respectively. At each discretization step, according to monotonicity of the exact eigenvalue (see Theorems 2.5 and 2.9), we aim to increase the real part of the rightmost eigenvalue of $A + \varepsilon E$. The adaptive step size control is based on this growth criterion for the real part, which is satisfied for sufficiently small step sizes because
of the monotonicty property of the differential equation. Since the objective is to find the rightmost point in the $\varepsilon$-pseudospectrum and not to follow accurately the trajectory of the differential equation, the accuracy of the discrete approximation is not controlled. In this way we obtain a sequence $(\lambda_n, E_n)$ where $\text{Re}\lambda_n$ grows monotonically and can stop only when $E_n$ arrives at a stationary point as characterized by Theorems 2.11 and 2.2.

Let $\gamma > 1$ be a given scaling factor for the step size control, and let $\text{Orth}(B)$ denote the orthogonal matrix obtained by orthogonalizing the columns of a given (rectangular) matrix $B$ via QR decomposition.

4.1. The Frobenius norm case. Given $E_n = U_nS_nV_n^T \approx E(t_n)$ of rank 2 and unit Frobenius norm, where $U_n, V_n \in \mathbb{R}^{n \times 2}$ with $U_n^T U_n = V_n^T V_n = I_2$ and $S_n \in \mathbb{R}^{2 \times 2}$ invertible with $\|S_n\|_F = 1$, and given $x_n$ and $y_n$ left and right eigenvectors of $A + \varepsilon E_n$ associated with its rightmost eigenvalue $\lambda_n$, with $x_n^* y_n > 0$, we determine the following approximations at time $t_{n+1} = t_n + \delta_n$, as stated in Algorithm 1, by applying a step of the Euler method with step size $\delta_n$ to (2.14).

Algorithm 1: Euler step; Frobenius norm case.

<table>
<thead>
<tr>
<th>Data: $E_n, U_n, V_n, S_n, x_n, y_n, \lambda_n,$ and $\rho_n$ (step size predicted the previous step)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Result: $E_{n+1}, U_{n+1}, V_{n+1}, S_{n+1}, x_{n+1}, y_{n+1}, \lambda_{n+1},$ and $\rho_{n+1}$</td>
</tr>
<tr>
<td>begin</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>9</td>
</tr>
</tbody>
</table>

*Initial conditions.* Let $x_0$ and $y_0$ be the left and right eigenvectors associated with the rightmost eigenvalue of $A$. We use the following initial condition which in most
cases increases the spectral abscissa of \( A + \varepsilon E_0 \) with respect to that of \( A \):

\[
E_0 = \text{Re} \left( x_0 y_0^* \right) / \| \text{Re} \left( x_0 y_0^* \right) \|_F, \quad U_0 S_0 V_0^T = E_0,
\]

where \( U_0, V_0 \) and \( S_0 \) (diagonal) are determined by the singular value decomposition of \( E_0 \). This choice is suggested by the formula for the steepest ascent direction (2.2). Since the algorithm may converge to a stationary point that is only locally optimal, we suggest starting from a few different (possibly randomly chosen) initial values.

Remarks. The step size control (see points 7 and 8 in Algorithm 1) ensures that the sequence \( \text{Re}(\lambda_n) \) is monotone and that the step size is increased after a successful step which has not required any step size rejection.

The cost of the algorithm is essentially that of the computation of rightmost eigenvalues and associated eigenvectors of rank-2 perturbations of the matrix \( A \), which can be computed efficiently for large sparse matrices \( A \), combining the approach considered in [MeR96] with the Sherman–Morrison formula.

We have observed in several examples, similarly to what happens in the complex case [GL11], that close to stationary points the step size oscillates due to the stability restriction of the Euler method. In fact the step size should be chosen as large as possible due to the fact that the right-hand sides in (2.14) vanish. On the other hand, the Euler method becomes unstable for a too large step size so that the step oscillates close to the stability boundary of the method. Remedies similar to those proposed in [GL11] can also be used in the present context.

### 4.2. The 2-norm case.

The algorithm, still based on Euler discretization, is similar to the previous one and is derived by integrating (3.3). Recall that for a complex vector \( v_n \) we indicate by \( v_{nR} \) its real part and by \( v_{nI} \) its imaginary part. Given \( E_n = U_n Q_n V_n^T \approx E(t_n) \) of rank 2 and unit 2-norm, where \( U_n, V_n \in \mathbb{R}^{n \times 2} \) with \( U_n^T U_n = V_n^T V_n = I_2 \) and \( Q_n \in \mathbb{R}^{2 \times 2} \) orthogonal, we determine the following approximation at the time instant \( t_{n+1} = t_n + \delta_n \), as stated in Algorithm 2, by applying an Euler step to (3.3).

**Algorithm 2:** Euler step; spectral norm case.

Input and output data analogous to Algorithm 1 with \( Q_n, Q_{n+1} \) replacing \( S_n, S_{n+1} \)

```
begin
1  Set \( \delta = \rho_n \)
2  Let \( X_n = (x_{nR}, x_{nI}) \) and \( Y_n = (y_{nR}, y_{nI}) \)
3  Compute
   \[
   P_n = \text{skew}(Q_n^T U_n X_n Y_n^T V_n)
   \]
   \[
   \hat{U}_{n+1} = U_n + \delta \left( (I - U_n U_n^T) X_n Y_n^T V_n Q_n^T \right)
   \]
   \[
   \hat{V}_{n+1} = V_n + \delta \left( (I - V_n V_n^T) Y_n X_n^T U_n Q_n \right)
   \]
   \[
   \hat{Q}_{n+1} = Q_n e^{\delta P_n}.
   \]
4  Orthogonalize the computed matrices
   \[
   U_{n+1} = \text{Orth}(\hat{U}_{n+1}), \quad V_{n+1} = \text{Orth}(\hat{V}_{n+1}), \quad Q_{n+1} = \text{Orth}(\hat{Q}_{n+1}).
   \]
5  Set \( E_{n+1} = U_{n+1} Q_{n+1} V_{n+1}^T \)
6  6 to 9 are analogous to the Frobenius case.
```
**Initial conditions.** In this case we make use of the following initial condition, which in most cases allows us to increase the spectral abscissa of $A + \varepsilon E_0$ with respect to the spectral abscissa of $A$:

$$
U_0 = \text{Orth}(X_0), \quad X_0 = (x_{0R}, x_{0I}),
V_0 = \text{Orth}(Y_0), \quad Y_0 = (y_{0R}, y_{0I}),
Q_0 = I_2,
$$

which is suggested by the analysis in section 3.2.

**5. Examples.** We illustrate the behavior of the presented algorithms in three examples: a small dense matrix and two large sparse matrices. We can exploit the sparsity to efficiently compute the rightmost eigenvalue using the MATLAB routine `eigs`, which is an interface for ARPACK, a well-known code implementing the implicitly restarted Arnoldi method [LS96, LSY98]. Here we focus our attention on real pseudospectra and compute approximate solutions to the presented systems of differential equations by a constant step size Euler discretization.

**Example 1.** We consider here the $6 \times 6$ matrix from [GL11],

$$
A = \begin{pmatrix}
0.1019 & -0.8350 & 0.2966 & -0.0756 & -2.2079 & 1.2682 \\
1.1813 & -1.4224 & -0.8664 & 0.8003 & -1.3413 & 1.3547 \\
-1.2457 & -0.1737 & -1.1910 & -0.3194 & -0.2909 & 0.8230 \\
-0.7830 & -0.5115 & -0.0109 & 0.8860 & 0.4878 & 0.3246 \\
-0.5740 & 0.0268 & 1.1950 & -0.1729 & 0.9966 & -0.8003 \\
-0.3815 & -0.4476 & -0.9740 & 1.4030 & 1.0361 & 0.7399
\end{pmatrix}.
$$

The following three cases are of interest:

(i) The case of $\mathbb{K} = \mathbb{R}$ and $\| \cdot \| = \| \cdot \|_2$.

In Figure 5.1 we show the trajectories of $\lambda(t)$ associated with the computed solutions of (3.3). The boundaries of the real 2-norm pseudospectra have been computed by the MATLAB package Seigtool by Karow, Kokkoupolou and Kressner [KKK10] who make use of the algorithm proposed by Qiu, Tits, and Yang [QTY95] to compute pseudospectral sets.

![Solution trajectories of system (3.3) for the matrix A of (5.1) with values of $\varepsilon = 10^{-0.5}$ (dark grey), $\varepsilon = 10^{-0.25}$ (light gray) and $\varepsilon = 1$ (gray in the print version; blue in the online version) inside real 2-norm pseudospectra. Initial values have been chosen according to (4.1). Right: zoom close to the rightmost points.](image-url)
Fig. 5.2. Solution trajectories of system (2.14) for the matrix $A$ of (5.1) with $\varepsilon = 1$, starting from randomly chosen initial values. The complex pseudospectrum is also drawn (light gray in the print version; yellow in the online version) to show the reduction of the pseudospectral abscissa. Right picture: zoom close to the rightmost point.

(ii) The case $K = \mathbb{R}$ and $\| \cdot \| = \| \cdot \|_F$.

In Figure 5.2 we show the trajectories of $\lambda(t)$ associated with the computed solutions of (2.14) corresponding to four different random initial values. Since no algorithms are available to draw the real Frobenius norm pseudospectrum, we plot the classical complex pseudospectrum to visualize the difference between the pseudospectral abscissa in the two cases. We have checked numerically that the computed stationary point is indeed a locally rightmost point.

(iii) The case $K = \mathbb{C}$.

This case is fully discussed and illustrated in [GL11]. In order to compare with subsequent figures we show in Figure 5.3 (right) the complex pseudospectrum for several values of $\varepsilon$.

Fig. 5.3. Boundary of 2-norm real (left) and complex (right) pseudospectra with values $\varepsilon = 10^{-0.5}, 10^{-0.25}, \varepsilon = 10^0$ of matrix (5.1). The colorbar indicates $\log_{10}(\varepsilon)$.

Example 2: A moderately large sparse matrix. We consider the matrix tols340.mtx from the matrix market library [BBD+95]. It arises in aeroelasticity and is a highly nonnormal matrix of dimension 340, with 2196 nonzero entries.

Let us consider the case of $K = \mathbb{R}$ and $\| \cdot \| = \| \cdot \|_2$. 
DIFFERENTIAL EQUATIONS FOR REAL AND COMPLEX PSEUDOSPECTRA

Fig. 5.4. On the left we draw the spectrum (in black) of the matrix \( \text{mtx340} \) and the boundary of the real 2-norm pseudospectrum corresponding to \( \varepsilon = 1 \) and \( \varepsilon = 10^{-0.25} \). We also show the corresponding solution trajectories of system (3.3); initial values have been chosen according to (4.1). Middle: zoom close to the starting points. Right: zoom close to the rightmost points.

In Figure 5.4 we see the trajectories of rightmost eigenvalues starting with a randomly chosen initial perturbation \( E_0 \); it is interesting to observe the asymptotic behavior which is similar to that of the complex pseudospectrum; that is, the trajectory converges almost vertically to the stationary value; moreover, also the initial phase of the integration of the ODEs is remarkable since the trajectories turn out to leave the starting points very slowly.

### Table 5.1

*Computed complex and real pseudospectral abscissas for the matrix tols340.mtx in the Frobenius and spectral norms.*

<table>
<thead>
<tr>
<th>( \varepsilon )</th>
<th>( | \cdot |_2, F ), ( K = \mathbb{C} )</th>
<th>( | \cdot |_2, K = \mathbb{R} )</th>
<th>( | \cdot |_F, K = \mathbb{R} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 10^{-0.5} )</td>
<td>0.24512 \cdot 10^2</td>
<td>0.24329 \cdot 10^2</td>
<td>0.17018 \cdot 10^2</td>
</tr>
<tr>
<td>( 10^{-0.25} )</td>
<td>0.47737 \cdot 10^2</td>
<td>0.44540 \cdot 10^2</td>
<td>0.30910 \cdot 10^2</td>
</tr>
<tr>
<td>( 10^0 )</td>
<td>1.10379 \cdot 10^2</td>
<td>1.10092 \cdot 10^2</td>
<td>0.58684 \cdot 10^2</td>
</tr>
<tr>
<td>( 10^{+0.25} )</td>
<td>2.86315 \cdot 10^2</td>
<td>2.85322 \cdot 10^2</td>
<td>1.63090 \cdot 10^2</td>
</tr>
</tbody>
</table>

**Example 3: A large sparse matrix.** We consider the sparse matrix \( A \) of dimension 2961 from the Eigtool demo that is denoted as Elliptic-PDE matrix. As the name indicates, it arises from the discretization of an elliptic boundary value problem; see [Nist] for a detailed description. We choose \( \| \cdot \| = \| \cdot \|_2, K = \mathbb{R} \) and \( \varepsilon = 10^{-1.5} \) and \( \varepsilon = 10^{-1} \).

Starting from a random rank-2 normalized matrix \( E_0 \), we run a constant step size Euler discretization with \( h = 0.2 \) and obtain an estimated accuracy of about 8 digits after 155 (for \( \varepsilon = 10^{-1.5} \)) and 134 (for \( \varepsilon = 10^{-1} \)) steps respectively.

Given the high dimension of the matrix, it is far from possible to compute the real 2-norm pseudospectrum with any available software.

Figure 5.5 shows the standard complex pseudospectrum (for the two considered values of \( \varepsilon \)) and the rightmost eigenvalues associated with the matrices \( A + \varepsilon E(t_n) \) obtained by the numerical method applied on a uniform grid \( \{ t_n \}_{n=0}^N \). Interestingly, we note that for \( \varepsilon = 10^{-1.5} \) the sequence seems to converge to the rightmost point of the complex pseudospectrum, which may lead one to wonder whether the real and
The complex pseudospectra are drawn in order to visualize the position of computed locally rightmost points with respect to the standard complex pseudospectral boundary. Right: zoom close to the rightmost point.

This does not occur in the case $\varepsilon = 10^{-1}$, as it appears clearly on the right.

It is worthwhile to remark that both computations are obtained in less than a minute on a portable PC with i7 CPU with clock frequency 1.73 GHz and a RAM of 8 GB.

6. Extensions to compute the pseudospectral radius.

The case $K = \mathbb{C}$ is discussed in [GL11]. Consider now the case $K = \mathbb{R}$ and $\| \cdot \| = \| \cdot \|_F$.

At $E \in \mathcal{M}_2$ (the manifold of real rank-2 matrices considered in section 2), let $x, y$ be left and right eigenvectors with $x^* y > 0$ to the eigenvalue $\lambda$ of largest absolute value of $A + \varepsilon E$. We modify the differential equation (2.12) to

\[
\dot{E} = P_E(\text{Re}(\lambda x^* y)) - \langle E, P_E(\text{Re}(\lambda x y^* )) \rangle E;
\]

that is, we replace $x$ by $\lambda x$ in (2.3) and correspondingly in the differential equations (2.14) for $S, U,$ and $V$. We then obtain the following analogue of Theorem 2.9 for the modulus instead of the real part.

**Theorem 6.1.** Let $E(t)$ satisfy the differential equation (6.1). If $\lambda(t)$ is a simple eigenvalue of $A + \varepsilon E(t)$, then $|\lambda(t)|$ increases monotonically with $t$.

**Proof.** We have

\[
\frac{d}{dt} |\lambda|^2 = 2 \text{Re}(\lambda \dot{\lambda}) = 2 \text{Re} \left( \lambda \varepsilon \frac{x^* \dot{E} y}{x^* y} \right) = 2 \varepsilon \text{Re} \left( \frac{(\lambda x)^* \dot{E} y}{x^* y} \right).\]

As in the proof of Theorem 2.9, with $\lambda x$ in place of $x$, we then have

\[
\text{Re}((\lambda x)^* \dot{E} y) = \| P_E(\text{Re}((\lambda x)y^*)) \|^2_F - \langle E, P_E(\text{Re}((\lambda x)y^*)) \rangle^2 \geq 0,
\]

and the result follows. \(\square\)
Finally, consider the case $\mathbb{K} = \mathbb{R}$ and $\| \cdot \| = \| \cdot \|_2$.

Let $x, y$ be left and right eigenvectors with $x^* y > 0$ to the eigenvalue $\lambda$ of largest absolute value of $A + \varepsilon E$. We modify the differential equation (3.3) to

$$\begin{align*}
\dot{U} &= (I - UU^T) X \Lambda^T V Q^T, \\
\dot{V} &= (I - VV^T) Y (X \Lambda)^T U Q, \\
\dot{Q} &= Q \text{skew}(Q^T U^T X \Lambda^T Y V^T),
\end{align*}$$

(6.3)

where $\Lambda = \text{Re}(\lambda) I + \text{Im}(\lambda) J$ with

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Similarly to the previous case we have the following monotonicity result.

**Theorem 6.2.** Let $E(t) = U(t)Q(t)V(t)^T$ with $U(t), V(t), Q(t)$ satisfying the differential equations (6.3) and with initial values such that $U(0), V(0) \in \mathbb{R}^{n \times 2}$ have orthonormal columns and $Q(0) \in \mathbb{R}^{2 \times 2}$ is an orthogonal matrix. If $\lambda(t)$ is a simple eigenvalue of $A + \varepsilon E(t)$, then $|\lambda(t)|$ increases monotonically with $t$.

We remark that the stationary points of (6.1) and (6.3) are characterized as in the rotated version of the previous sections with $e^{i \theta} = \lambda / |\lambda|$.

Note that—analogously to the complex pseudospectrum case—the algorithm for the $\varepsilon$-pseudospectral radius is obtained (see [GL11]) by replacing $x$ by $\lambda x$ in the ODEs, where $\lambda$ is the eigenvalue of largest modulus instead of the rightmost one. This means that the algorithms explained in section 4 for the $\varepsilon$-pseudospectral abscissa can be adapted to compute the $\varepsilon$-pseudospectral radius by simply replacing $x_n$ by $\lambda_n x_n$, where $\lambda_n$ is the eigenvalue of maximal modulus of $A + \varepsilon E_n$, and $x_n, y_n$ are the associated eigenvectors. We illustrate the results of this section in the following example.

**Example 1.** Here we reconsider the $6 \times 6$ matrix (5.1). In Figure 6.1 we show the trajectories of (6.3), starting from random initial values, to the points of maximum modulus of $\Lambda_{\mathbb{R}, \| \cdot \|_2}$ for several values of $\varepsilon$.

In Figure 6.2 we show two trajectories of (6.1), starting from random initial values, to the point of maximum modulus of $\Lambda_{\mathbb{R}, \| \cdot \|_2}$ for $\varepsilon = 1$ together with the boundary of the complex pseudospectrum $\Lambda_{\mathbb{C}, \| \cdot \|_F}$.

![Fig. 6.1. Trajectories—starting from randomly chosen initial values—to the points with maximum modulus and boundary of 2-norm real pseudospectra with values $\varepsilon = 10^{-0.5}, 10^{-0.25}, \varepsilon = 10^0$ of matrix (5.1). The colorbar indicates $\log_{10}(\varepsilon)$. The circle in black is centered at the origin and has radius $\rho_1(A)$ so that it is tangential to $\Lambda_{\mathbb{C}, \| \cdot \|_F}$. Right: zoom close to the extremal points.](image)
Fig. 6.2. Solution trajectories of system (6.1) for the matrix $A$ of (5.1) with $\varepsilon = 1$. The circle with radius $p_{\varepsilon}(A)$ is drawn in black and the complex pseudospectrum is drawn (in yellow in the online version) to show the reduction of the pseudospectral radius. Right: zoom close to the extremal point.

7. Drawing the boundary of the real pseudospectrum. Consider the real Frobenius case (the 2-norm case is analogous), which is particularly interesting since there are no algorithms available for drawing the boundary of the $\varepsilon$-pseudospectrum, even for a small matrix.

The idea is the following: if we consider a point $z$ which is located outside the $\varepsilon$-pseudospectrum of a matrix $A$, by using a simple variant of the method we have presented to compute the $\varepsilon$-pseudospectral radius, we can calculate the point on the boundary of the $\varepsilon$-pseudospectrum having minimal distance from $z$. By a simple shifting argument, which we discuss here, this is reduced to computing the distance to the origin of the $\varepsilon$-pseudospectrum of a matrix $B$, which is assumed not to contain the origin.

Assume that $\Lambda_{\varepsilon,\parallel \cdot \parallel_F}(B)$ does not contain the origin; we obtain a complementary algorithm of the one considered in section 6 to compute the pseudospectral radius, which computes the minimal (instead of the maximal) distance of $\Lambda_{\varepsilon,\parallel \cdot \parallel_F}(B)$ to the origin. In order to do this, we replace $\lambda$ by $-\mu$ in the system of ODEs (6.1), that is,

\begin{equation}
\dot{\mu} = P_E(\text{Re}(-\mu xy^*)) - \langle E, P_E(\text{Re}(-\mu xy^*))) \rangle E,
\end{equation}

where $x(t)$ and $y(t)$ are left and right eigenvectors, both of unit norm and with $x(t)^*y(t) > 0$ for all $t$, to a continuous path of simple eigenvalues $\mu(t)$ of $B + \varepsilon E(t)$ having smallest modulus. For the numerical discretization of (7.1) we make use of an algorithm which is very similar to the first one described in section 4, where the step size control is devised here to guarantee the property of monotonic decrease of the sequence $|\mu_n|$.

In this way we follow a steepest descent direction (this is due to the minus sign) of the eigenvalue $\mu$ of smallest modulus of the usual matrix valued function $A + \varepsilon E(t)$. We are able to prove that the trajectory $\mu(t)$ converges to a local minimum of $|\mu|$ for $\mu \in \sigma_{\varepsilon}(B)$. In order to obtain this result, we make use of arguments similar to those used for the $\varepsilon$-pseudospectral radius computation, which we omit here for the sake of brevity.
Exploiting the shifting property (for arbitrary \( z \in \mathbb{C} \))

\[
\Lambda^R_z\|F\,(A - zI) = \Lambda^R_z\|F\,(A) - z = \{ w \in \mathbb{C} : w = \lambda - z \text{ and } \lambda \in \Lambda^R_z\|F\,(A) \},
\]

we can integrate (7.1) with \( B = A - zI \) and compute the point of minimal modulus of the \( \varepsilon \)-pseudospectrum of \( A - zI \), where \( z \in \mathbb{C} \) is a chosen point external to \( \Lambda^R_z\|F\,(A) \).

The computed point, which we denote as \( \mu_{\text{min}}(z) \), is then shifted by \( z \) to give the point on the boundary of the \( \varepsilon \)-pseudospectrum with minimal distance to \( z \).

This procedure is the basis for the development of an algorithm for tracking the boundary of the \( \varepsilon \)-pseudospectrum.

For example, if we are interested drawing the rightmost section of \( \Lambda^R_z\|F\,(A) \), that is, the curve which delimits the right boundary of the pseudospectrum, we first compute the pseudospectral abscissa \( \alpha^R_z\|F\,(A) \). Afterward we can consider the segment \( z(s) = \alpha^R_z\|F\,(A) + \delta + is \), where \( \delta > 0 \) is a small constant and \( s \in [\underline{s}, \overline{s}] \) is a suitable interval. In order to compute the right boundary of the \( \varepsilon \)-pseudospectrum we repeatedly integrate (7.1) with \( B(s) = A - z(s)I \) and shift the computed optimal values \( \mu_{\text{min}}(z(s)) \) by \( z(s) \). An illustrative example is shown in Figure 7.1, where we have discretized the segment \([\underline{s}, \overline{s}]\) with respect to \( s \) and used as initial guess \( E_0 \), for the problem with parameter \( s_{k+1} = s_k + \Delta s \), the stationary matrix \( E \) of the previous problem with \( s = s_k \).

Similarly, we can draw the boundary of the \( \varepsilon \)-pseudospectrum close to the point with largest modulus. Let \( z = \rho^R\|F\,(A)e^{i\theta} \) be a point on the \( \varepsilon \)-pseudospectrum with maximal modulus; we can consider the arc \( z(s) = (\rho^R\|F\,(A) + \delta)e^{i\phi(s)} \), where \( \delta > 0 \) is a small constant and \( s \in [\underline{s}, \overline{s}] \) is a suitable interval such that \( \phi(\underline{s}) = \theta - \eta/2, \phi(\overline{s}) = \theta + \eta/2 \), where \( \eta > 0 \).

![Fig. 7.1](image)

**Fig. 7.1.** Rightmost section (blue in online version) of \( \Lambda^R_z\|F\,(A) \) for the matrix (5.1) compared to the unstructured pseudospectrum \( \Lambda^C_z\|F\,(A) \) (yellow in online version). Right: zoom.

We consider again Example (5.1). In Figure 7.1 we plot the computed rightmost section of the \( \varepsilon \)-pseudospectrum for \( \varepsilon = 1 \). It is interesting to note in the right figure the needle (nonsmooth) structure of \( \Lambda^R_z\|F\,(A) \) on the real axis, which looks like the one computed (by Seigtool) for the 2-norm case (see Figure 5.3). It is actually the same needle, because on the real axis the real \( \varepsilon \)-pseudospectrum coincides with the complex \( \varepsilon \)-pseudospectrum, as the extremal perturbation is a real rank-1 matrix.

In Figure 7.2 we plot the computed section of the \( \varepsilon \)-pseudospectrum for \( \varepsilon = 1 \) close to the point of maximal modulus.
Any other section of the $\varepsilon$-pseudospectrum can be plotted in a similar way.

**Appendix A. Complex pseudospectra.** The “usual” complex pseudospectra with respect to the 2-norm and their extremal points have been considered more thoroughly in the literature; see, e.g., [Tr99, TE05, GL11]. Here we give results analogous to those of section 2 for the Frobenius-norm complex pseudospectra. In the complex case, the perturbations can be chosen of rank 1 for each point in the $\varepsilon$-pseudospectrum. Since the Frobenius norm and the 2-norm of rank-1 matrices coincide, the extremizers are the same for both the 2-norm and the Frobenius norm, and hence the results are valid for both norms.

**A.1. Rank-1 property of extremizers and steepest ascent dynamics.**

**A.1.1. Extremizers.**

**Theorem A.1.** Let $A \in \mathbb{C}^{n \times n}$ and fix an angle $\theta \in \mathbb{R}$. Let $E_*$ be a complex matrix of unit Frobenius norm, and let $\lambda_*$ be a simple eigenvalue of $A + \varepsilon E_*$, with left and right eigenvectors $x$ and $y$, respectively, both of unit norm and with $x^* y > 0$. Then the following two statements are equivalent:

1. Every differentiable path $(E(t), \lambda(t))$ (for small $t \geq 0$) such that $\|E(t)\|_F \leq 1$ and $\lambda(t)$ is an eigenvalue of $A + \varepsilon E(t)$, with $E(0) = E_*$ and $\lambda(0) = \lambda_*$, has $\text{Re}(e^{-i\theta} \lambda(0)) \leq 0$.

2. $E_*$ is a positive multiple of $e^{i\theta} xy^*$.

In particular, Frobenius norm extremizers necessarily have rank 1. They are also extremizers in the 2-norm.

The proof follows closely that of Theorem 2.2, with straightforward modifications. We do not give all details but formulate the analogues of the results in section 2 that yield Theorem A.1 for $\theta = 0$ (and hence also for arbitrary angle $\theta$).

**A.1.2. Steepest ascent direction.**

**Lemma A.2.** Let $E \in \mathbb{C}^{n \times n}$ be a nonzero complex matrix of unit Frobenius norm, and let $x, y \in \mathbb{C}^n$ be given nonzero complex vectors. The solution of the optimization problem in $\mathbb{C}^{n \times n}$,

$$Z_* = \arg \max_{\|Z\|_F = 1, \text{Re}(E,Z) = 0} \text{Re}(x^* Z y)$$

(A.1)
is given by

\[(A.2) \quad \mu Z_\ast = xy^* - \text{Re} \langle E, xy^* \rangle E,\]

where \(\mu\) is the Frobenius norm of the matrix on the right-hand side.

**A.1.3. Steepest ascent differential equation.** Lemmas 1.1 and A.2 suggest considering the differential equation on the manifold of complex \(n \times n\) matrices of unit Frobenius norm

\[(A.3) \quad \dot{E} = xy^* - \text{Re} \langle E, xy^* \rangle E,
\]

where \(x(t)\) and \(y(t)\) are left and right eigenvectors, respectively, to a simple eigenvalue \(\lambda(t)\) of \(A + \varepsilon E(t)\), both of unit norm and with \(x(t)^*y(t) > 0\). We have again monotonic growth of the real part of \(\lambda(t)\).

**Theorem A.3.** Let \(E(t) \in \mathbb{C}^{n \times n}\) of unit Frobenius norm satisfy the differential equation \((A.3)\). If \(\lambda(t)\) is a simple eigenvalue of \(A + \varepsilon E(t)\), then

\[(A.4) \quad \text{Re} \dot{\lambda}(t) \geq 0.\]

We note that here, in place of \((2.5)\),

\[(A.5) \quad \text{Re}(x^* \dot{E}y) = \|xy^*\|^2_F - (\text{Re} \langle E, xy^* \rangle)^2 \geq 0.\]

**A.1.4. Stationary points.** The stationary points of \((A.3)\) are characterized as follows.

**Theorem A.4.** The following statements are equivalent along solutions of \((A.3)\):

1. \(\text{Re} \dot{\lambda} = 0\).
2. \(\dot{E} = 0\).
3. \(E\) is a real multiple of \(xy^*\).

**A.2. Rank-1 dynamics.**

**A.2.1. Rank-1 matrices and their tangent matrices.** We denote by \(M_1\) the manifold of (complex) rank-1 matrices of dimension \(n \times n\) and write \(E \in M_1\) in a nonunique way as

\[E = \sigma uv^*,\]

where \(\sigma \in \mathbb{C}\) and \(u, v \in \mathbb{C}^n\) have unit norm. Tangent matrices \(\dot{E} \in T_E M_1\) are of the form \(\dot{E} = \sigma \dot{u}v^* + \sigma uv^* \dot{v}^*\), where \(\dot{\sigma} \in \mathbb{C}\), \(\dot{u}, \dot{v} \in \mathbb{C}^n\) are uniquely determined by \(\sigma, u, v\) if we impose the orthogonality conditions \(u^*\dot{u} = 0\), \(v^*\dot{v} = 0\).

**Lemma A.5.** The orthogonal projection onto the tangent space \(T_E M_1\) at \(E = \sigma uv^* \in M_1\) is given by

\[(A.6) \quad P_E(Z) = Z - (I - uu^*)Z(I - vv^*)\]

for \(Z \in \mathbb{C}^{n \times n}\).

**A.2.2. A differential equation for rank-1 matrices.** In the differential equation \((A.3)\) we replace the right-hand side by its orthogonal projection to \(T_E M_1\):

\[(A.7) \quad \dot{E} = P_E \left( xy^* - \text{Re} \langle E, xy^* \rangle E \right),\]
where again \( x(t) \) and \( y(t) \) are right and left eigenvectors, respectively, to a simple eigenvalue \( \lambda(t) \) of \( A + \varepsilon E(t) \), both of unit norm and with \( x(t)^* y(t) > 0 \).

**Lemma A.6.** For \( E = \sigma uv^* \in \mathcal{M}_1 \) with nonzero \( \sigma \in \mathbb{C} \) and with \( u \in \mathbb{C}^n \) and \( v \in \mathbb{C}^n \) of unit norm, the equation \( \dot{E} = P_E(Z) \) is equivalent to \( \dot{E} = \dot{\sigma} uv^* + \sigma \dot{u} v^* + \sigma \dot{v}^* \), where

\[
\dot{\sigma} = u^* Z v, \\
\dot{u} = (I - uu^*) Z v \sigma^{-1}, \\
\dot{v} = (I - vv^*) Z^* u \sigma^{-1}.
\]

(A.8)

With \( Z = xy^* - \text{Re}(E, xy^*) E \), this yields that the differential equation (A.7) for \( E = \sigma uv^* \) is equivalent to the following system of differential equations for \( \sigma, u, v \), where we abbreviate \( \alpha = u^* x \in \mathbb{C} \), \( \beta = v^* y \in \mathbb{C}^2 \):

\[
\dot{\sigma} = \alpha \beta - \text{Re}(\overline{\alpha} \beta \sigma) \sigma, \\
\dot{u} = (x - \alpha u) \beta \sigma^{-1}, \\
\dot{v} = (y - \beta v) \alpha \sigma^{-1}.
\]

(A.9)

If \( \sigma \) is of unit modulus, then the first differential equation can be equivalently written as

\[ \dot{\sigma} = i \text{Im}(\alpha \beta \sigma) \sigma, \]

which shows that \( \sigma \) remains of unit modulus. We note that the differential equations (A.9) are equivalent to the differential equations (4.3) and (4.4) of [GL11], which were found by an educated guess, but stated without derivation.

**A.2.3. Monotonicity of eigenvalues.**

**Theorem A.7.** Let \( E(t) \in \mathcal{M}_1 \) of unit Frobenius norm satisfy the differential equation (A.7). If \( \lambda(t) \) is a simple eigenvalue of \( A + \varepsilon E(t) \), then

\[ \text{Re} \dot{\lambda}(t) \geq 0. \]

(A.10)

This is Theorem 4.2 of [GL11]. We note that a calculation yields

\[ \text{Re}(x^* \dot{E} y) = (\text{Im}(\overline{\alpha} \beta \sigma))^2 + |\alpha|^2 \cdot \|y - \beta v\|^2 + |\beta|^2 \cdot \|x - \alpha u\|^2 \geq 0. \]

(A.11)

**A.2.4. Stationary points.** The next result is a direct consequence of (A.11), since \( \alpha = \beta = 0 \) implies that \( \lambda \) is an eigenvalue of \( A \) with left and right eigenvectors \( x \) and \( y \).

**Theorem A.8.** The following statements are equivalent along solutions of (A.7) provided that \( \lambda \) is not an eigenvalue of \( A \):

1. \( \text{Re} \dot{\lambda} = 0 \).
2. \( \dot{E} = 0 \).
3. \( E \) is a real multiple of \( xy^* \).

**Acknowledgments.** The authors thank the referees for their careful reading of the paper and their useful and constructive remarks.
REFERENCES


