# Distance Problems for Linear Dynamical Systems

Daniel Kressner and Matthias Voigt

[1] Daniel Kressner, ANCHP, EPF Lausanne, Station 8, CH-1015 Lausanne, Switzerland
e-mail:daniel.kressner@epfl.ch
[2] Matthias Voigt, Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, D-39106 Magdeburg, Germany
e-mail: voigtm@mpi-magdeburg.mpg.de

**Summary.** This chapter is concerned with distance problems for linear time-invariant differential and differential-algebraic equations. Such problems can be formulated as distance problems for matrices and pencils. In the first part, we discuss characterizations of the distance of a regular matrix pencil to the set of singular matrix pencils. The second part focuses on the distance of a stable matrix or pencil to the set of unstable matrices or pencils. We present a survey of numerical procedures to compute or estimate these distances by taking into account some of the historical developments as well as the state of the art.

## 1 Introduction

Consider a linear time-invariant differential-algebraic equation (DAE)

$$E\dot{x}(t) = Ax(t) + f(t), \quad x(0) = x_0, \tag{1}$$

with coefficient matrices $E, A \in \mathbb{R}^{n\times n}$, a sufficiently smooth inhomogeneity $f : [0,\infty) \to \mathbb{R}^n$, and an initial state vector $x_0 \in \mathbb{R}^n$. When $E$ is nonsingular, (1) is turned into a system of ordinary differential equations by simply multiplying with $E^{-1}$ on both sides. The more interesting case of singular $E$ arises, for example, when imposing algebraic constraints on the state vector $x(t) \in \mathbb{R}^n$. Equations of this type play an important role in a variety of applications, including electrical circuits [55, 54] and multi-body systems [48]. The analysis and numerical solution of more general DAEs (which also take into account time-variant coefficients, nonlinearities and time delays) is a central part of Volker Mehrmann's work, as witnessed by the monograph [43] and by the several other chapters of this book concerned with DAEs/descriptor systems. A recurring theme in Volker's work is the notion of *robustness*, as prominently expressed in the papers [2, 33].

This survey is concerned with distance measures that provide robust ways of assessing properties for (1). Before moving on with DAEs, let us illustrate the basic idea in the simpler setting of numerically deciding whether a given matrix $A \in \mathbb{R}^{n\times n}$ is singular. In the presence of roundoff error and other uncertainties, this problem is ill-posed: Square matrices are *generically* nonsingular and therefore the slightest perturbation very likely turns a possibly singular matrix into a nonsingular matrix. It is more sensible to ask whether $A$ is close to a singular matrix. In the absence of additional information on the uncertainty, this naturally leads to the problem of finding the smallest perturbation $\Delta A \in \mathbb{R}^{n\times n}$ such that $A + \Delta A$ is singular:

$$\delta(A) := \min\{\|\Delta A\| : \ A + \Delta A \text{ is singular}\}.$$

It is well known that this *distance to singularity* coincides with $\sigma_{\min}(A)$, the smallest singular value of $A$, when choosing the matrix 2-norm $\|\cdot\|_2$ or the Frobenius norm $\|\cdot\|_F$ for $\|\Delta A\|$. Note that $\sigma_{\min}(A)$ is the reciprocal of $\|A^{-1}\|_2$, the (relative) 2-norm condition number of matrix inversion. This witnesses a more general relation between the condition number of a problem and its distance to ill-posedness [16]. The quantity $\delta(A)$ is robust in the sense that it is only mildly affected by perturbations of $A$. To see this, consider a (slight) perturbation $A \mapsto \widetilde{A}$ that arises, e.g., due to roundoff error. By the triangular inequality it holds

$$\delta(A) - \|\widetilde{A} - A\| \leq \delta(\widetilde{A}) \leq \delta(A) + \|\widetilde{A} - A\|.$$

Such a robustness property holds more generally for all the distance measures discussed here.

In this survey, we will focus on distance measures for two of the most important properties of a DAE (1). The matrix pencil $sE - A$ associated with (1) is called

- *regular* if its characteristic polynomial $s \mapsto \det(sE - A)$ is not identically zero;
- *stable* if all finite eigenvalues of $sE - A$ are contained in $\mathbb{C}^-$, the open left complex halfplane .

Regularity guarantees the existence of a unique (classical) solution of the DAE for all consistent initial values [59]. On the other hand, the stability of $sE - A$ is equivalent to asymptotic stability of the homogeneous DAE (1) with $f(t) \equiv 0$ [17].

Numerically verifying the regularity and stability of a given pencil is a challenge; none of the straightforward approaches is guaranteed to remain robust under uncertainties in $E$ and $A$. This motivates considering the distance of a given regular or stable pencil to the set of singular or unstable pencils, respectively. Then one needs to devise algorithms for computing these distances or, at least, reliable and effective bounds. As we will see below, this is by no means trivial.

## 2 The Distance to Singularity of a Matrix Pencil

In this section we will discuss the distance of a given regular matrix pencil to the nearest singular matrix pencil, that is, a matrix pencil with vanishing characteristic polynomial. To our knowledge, such a notion of distance was first introduced and discussed by Byers, He, and Mehrmann [15]. Letting $\mathbb{R}[s]$ denote the ring of polynomials with coefficients in $\mathbb{R}$, the *distance to singularity* of a matrix pencil $sE - A$ is given by

$$\delta(E,A) := \min\left\{ \left\| \begin{bmatrix} \Delta E & \Delta A \end{bmatrix} \right\|_F : s(E+\Delta E) - (A+\Delta A) \in \mathbb{R}[s]^{n\times n} \text{ is singular} \right\}.$$

As explained in [15], explicit formulas for $\delta(E,A)$ can be obtained for special cases, for example when $n \leq 2$ or $E,A$ are scalar multiples of each other. Such explicit formulas are not known for the general case and, even worse, devising a numerical method for computing $\delta(E,A)$ or bounds thereof turned out to be an extremely difficult problem. Since the publication of [15], almost no progress has been made in this direction.

In view of the definition, one might attempt to check (nearby) singularity by inspecting the magnitudes of the coefficients of $\det(sE - A)$. This attempt is futile for most practical applications, because the polynomial coefficients exhibit a wildly different scaling as $n$ increases except for very particular situations, e.g., when all eigenvalues are (almost) on the unit circle. To obtain more meaningful measures, one therefore needs to consider other characterizations for the singularity of a pencil. In the following, we will discuss several such characterizations and their relation to $\delta(E,A)$.

### 2.1 Distance to Singularity and Structured Low-Rank Approximation

The Kronecker structure of $sE-A$ is intimately related to the ranks of certain block Toeplitz matrices constructed from $E,A$; see [9] for an overview. Specifically, let us consider

$$W_k = W_k(E,A) := \begin{bmatrix} A & & & \\ E & A & & \\ \ddots & \ddots & \ddots & \\ & & E & A \\ & & & E \end{bmatrix} \in \mathbb{R}^{(k+1)n \times kn}. \tag{2}$$

It can be shown that $sE-A$ is a regular pencil if and only if $W_k(E,A)$ has full column rank for all $k = 1, \ldots, n$. Setting

$$\gamma_k := \min \left\{ \left\| \begin{bmatrix} \Delta E & \Delta A \end{bmatrix} \right\|_{\mathrm{F}} : \operatorname{rank} W_k(E+\Delta E, A+\Delta A) < nk \right\},$$

it therefore holds that

$$\delta(E,A) = \min_{1 \le k \le n} \gamma_k = \gamma_n, \tag{3}$$

where the latter equality follows from the observation that the rank deficiency of $W_k$ implies the rank deficiency of $W_\ell$ for all $\ell \ge k$.

By (3), computing $\delta(E,A)$ is equivalent to finding a structured perturbation that makes $W_n$ rank deficient. Having wider applicability in signal processing, systems and control, such structured low-rank approximation problems have attracted quite some attention recently; see [45] for an overview.

To proceed from (3), one needs to replace the rank constraint by a simpler characterization. Clearly, the matrix $W_k(E+\Delta E, A+\Delta A)$ is rank deficient if and only if there exist vectors $x_1, \ldots, x_k \in \mathbb{R}^n$, not all equal to zero, such that

$$\begin{bmatrix} A+\Delta A & & & \\ E+\Delta E & A+\Delta A & & \\ & \ddots & \ddots & \\ & & E+\Delta E & A+\Delta A \\ & & & E+\Delta E \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_k \end{pmatrix} = 0. \tag{4}$$

Hence, we obtain the outer-inner optimization problem

$$\gamma_k = \min_{\substack{x \in \mathbb{R}^{kn} \\ x \neq 0}} f_k(x), \qquad f_k(x) = \min \left\{ \left\| \begin{bmatrix} \Delta E & \Delta A \end{bmatrix} \right\|_{\mathrm{F}} : W_k(E+\Delta E, A+\Delta A)x = 0 \right\}. \tag{5}$$

Its particular structure implies that (4) is equivalent to

$$\begin{bmatrix} \Delta E & \Delta A \end{bmatrix} Z_k = - \begin{bmatrix} E & A \end{bmatrix} Z_k \quad \text{with} \quad Z_k = \begin{bmatrix} 0 & x_1 & x_2 & \cdots & x_{k-1} & x_k \\ x_1 & x_2 & x_3 & \cdots & x_k & 0 \end{bmatrix}. \tag{6}$$

This shows that the inner optimization problem in (5) is a standard linear least-squares problem, admitting the explicit solution $f_k(x) = \left\| \begin{bmatrix} E & A \end{bmatrix} Z_k Z_k^\dagger \right\|_{\mathrm{F}}$, where $Z_k^\dagger$ denotes the Moore-Penrose pseudoinverse of $Z_k$. Thus,

$$\gamma_k = \min_{\substack{x \in \mathbb{R}^{kn} \\ x \neq 0}} \left\| \begin{bmatrix} E & A \end{bmatrix} Z_k Z_k^\dagger \right\|_{\mathrm{F}}$$

and, consequently,

$$\delta(E,A) = \min_{1\le k\le n} \min_{\substack{x\in\mathbb{R}^{kn}\\ x\neq 0}} \left\| \begin{bmatrix} E & A \end{bmatrix} Z_k Z_k^\dagger \right\|_{\mathrm{F}} = \min_{\substack{x\in\mathbb{R}^{n^2}\\ x\neq 0}} \left\| \begin{bmatrix} E & A \end{bmatrix} Z_n Z_n^\dagger \right\|_{\mathrm{F}}. \tag{7}$$

This so called *variable projection least-squares problem* can, in principle, be addressed by standard nonlinear optimization methods. Such an approach bears two major obstacles: (a) There may be many local minima. (b) The sheer number of variables ($n^2$ for $k = n$) restricts the scope of existing optimization methods to fairly small values of $n$.

Although the obstacles mentioned above have not been overcome yet, significant progress has been made in the availability of software for structured low-rank approximation problems. Specifically, the SLRA software package [46] covers the block Toeplitz structure of the matrix $W_k$ and can hence be used to compute $\gamma_k$. We have applied this software to all examples from [15]. For the matrix pencil

$$sE - A = s \begin{bmatrix} 0 & 1/\varepsilon \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 1/\varepsilon \\ 0 & 1 \end{bmatrix}$$

and $\varepsilon = 10^{-1}$, the value $\gamma_2 = 0.0992607\ldots$ returned by SLRA approximates the true value of $\delta(E,A)$ up to machine precision. This nice behavior is also observed for smaller values of $\varepsilon$ until around $10^{-7}$, below which SLRA signals an error.

For other examples from [15], SLRA seems to deliver local minima only or exhibits very slow convergence. For the $8\times 8$ matrix pencil arising from the model of a two-dimensional, three-link mobile manipulator [15, Example 14], SLRA returns:

| $\gamma_1$ | $\gamma_2$ | $\gamma_3$ | $\gamma_4$ | $\gamma_5$ | $\gamma_6$ | $\gamma_7$ | $\gamma_8$ |
|---|---|---|---|---|---|---|---|
| 0.0113 | 0.0169 | 0.0171 | 0.0277 | 0.0293 | 0.0998 | 0.0293 | 0.6171 |

These values do not reflect the fact that the exact value of $\gamma_k$ *decreases* as $k$ increases. This clearly indicates a need to further explore the potential of SLRA and related software for computing $\delta(E,A)$.

## 2.2 Distance to Singularity and Block Schur Form

For orthogonal matrices $Q,Z\in\mathbb{R}^{n\times n}$ we consider the equivalence transformation

$$Q^\mathsf{T}(sE-A)Z = s\begin{bmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{bmatrix} - \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where we partitioned the transformed pencil such that $sE_{11}-A_{11}\in\mathbb{R}[s]^{k-1\times k}$ and $sE_{22}-A_{22}\in\mathbb{R}[s]^{n-k+1\times n-k}$ for some $1\le k\le n$. The perturbation $s\Delta E-\Delta A$ defined by

$$Q^\mathsf{T}(s\Delta E-\Delta A)Z := s\begin{bmatrix} 0 & 0 \\ -E_{21} & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ -A_{21} & 0 \end{bmatrix}$$

yields

$$Q^\mathsf{T}(s(E+\Delta E)-(A+\Delta A))Z = s\begin{bmatrix} E_{11} & E_{12} \\ 0 & E_{22} \end{bmatrix} - \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}.$$

Since $sE_{11}-A_{11}$ and $sE_{22}-A_{22}$ are rectangular, and thus singular, the perturbed pencil $s(E+\Delta E)-(A+\Delta A)$ is singular as well. Partitioning $Q=\begin{bmatrix} Q_{k-1} & Q_{n-k+1}\end{bmatrix}\in\mathbb{R}^{n\times n}$ and $Z=\begin{bmatrix} Z_k & Z_{n-k}\end{bmatrix}\in\mathbb{R}^{n\times n}$ conformally, the de-regularizing perturbation satisfies

$$\left\| \begin{bmatrix} \Delta E & \Delta A \end{bmatrix} \right\|_{\mathrm{F}} = \left\| \begin{bmatrix} Q_{n-k+1}^{\mathsf{T}} E Z_k & Q_{n-k+1}^{\mathsf{T}} A Z_k \end{bmatrix} \right\|_{\mathrm{F}}.$$

Minimizing over all perturbations constructed as described above leads to the following minimax-like characterization [15, Sec. 4.2]:

$$\begin{aligned} \delta(E,A) = \min_{1\le k\le n} \Big\{ & \left\| \begin{bmatrix} Q_{n-k+1}^{\mathsf{T}} E Z_k & Q_{n-k+1}^{\mathsf{T}} A Z_k \end{bmatrix} \right\|_{\mathrm{F}} : \\ & Q_{n-k+1} \in \mathbb{R}^{n\times n-k+1},\, Q_{n-k+1}^{\mathsf{T}} Q_{n-k+1} = I_{n-k+1},\, \text{and } Z_k \in \mathbb{R}^{n\times k},\, Z_k^{\mathsf{T}} Z_k = I_k \Big\}. \end{aligned} \tag{8}$$

Computing $\delta(E,A)$ via this characterization amounts to solving $n$ optimization problems over Stiefel manifolds. Again, the possible presence of many local minima and the $O(n^2)$ degrees of freedom limit the usefulness of this characterization. A related idea was considered in [62] for computing the distance to the nearest uncontrollable system.

## 2.3 Lower and Upper Bounds for the Distance to Singularity

As discussed above, the characterizations (3) and (8) are of limited applicability for computing $\delta(E,A)$. It is therefore of interest to develop inexpensive lower and upper bounds, partially based on these characterizations.

**Lower bound from singular values of $W_k$.** A first lower bound proposed in [15] uses the matrix $W_k$ defined in (2). Let $s\Delta E - \Delta A$ be a minimum norm de-regularizing perturbation of $sE - A$. Then $W_k(E+\Delta E, A+\Delta A)$ is rank-deficient for some $k \le n$, and the inequalities

$$\sigma_{\min}(W_k(E,A)) \le \left\| W_k(\Delta E, \Delta A) \right\|_{\mathrm{F}} = \sqrt{k}\,\delta(E,A)$$

hold for this particular value $k$. Consequently,

$$\frac{\sigma_{\min}(W_n(E,A))}{\sqrt{n}} = \min_{1\le k\le n} \frac{\sigma_{\min}(W_k(E,A))}{\sqrt{k}} \le \delta(E,A). \tag{9}$$

To obtain a reliable lower bound, one needs to evaluate $\sigma_{\min}(W_k(E,A))$ for $k = n$, as its value can decrease by many orders of magnitude when increasing from $k$ to $k+1$, see [15, Subsec 5.1]. However, the computation of the smallest singular value of the $(k+1)n \times kn$-matrix $W_k(E,A)$ gets increasingly expensive as $k$ increases. For example, inverse iteration applied to $W_k^{\mathsf{T}} W_k$ requires $O(k\cdot n^3)$ operations and $O(k\cdot n^2)$ storage for the factorization of $W_k^{\mathsf{T}} W_k$, assuming its block tridiagonal structure is exploited.

**Lower bound from one-parameter optimization.** Another lower bound [15, Subsec 5.2] is obtained from the observation that $\alpha_0(E+\Delta E) - \beta_0(A+\Delta A)$ is singular for all scalars $\alpha_0, \beta_0 \in \mathbb{C}$ if the pencil $s(E+\Delta E) - (A+\Delta A)$ is singular. It follows that

$$\sigma_{\min}(\alpha_0 E - \beta_0 A) \le \left\| \alpha_0 \Delta E - \beta_0 \Delta A \right\|_{\mathrm{F}} \le \sqrt{|\alpha_0|^2 + |\beta_0|^2}\, \left\| \begin{bmatrix} \Delta E & \Delta A \end{bmatrix} \right\|_{\mathrm{F}}.$$

Defining

$$\mathbb{S}_{\mathbb{F}} := \Big\{ (\alpha,\beta) \in \mathbb{F}\times\mathbb{F} : |\alpha|^2 + |\beta|^2 = 1 \Big\},$$

we therefore obtain

$$\max_{(\alpha,\beta)\in\mathbb{S}_{\mathbb{R}}} \sigma_{\min}(\alpha E - \beta A) \le \max_{(\alpha,\beta)\in\mathbb{S}_{\mathbb{C}}} \sigma_{\min}(\alpha E - \beta A) \le \delta(E,A). \tag{10}$$

The first inequality is particularly suitable when $E, A$ are real and amounts to minimizing

$$g(t) := -\sigma_{\min}(\sin(t)E - \cos(t)A), \quad t \in [0,\pi].$$

Well-known properties of singular values imply that $g$ is piecewise smooth and Lipschitz-continuous. An efficient algorithm tailored to such a situation is described in [49]. The lower bounds (10) have been observed to be rather tight [15].

**Upper bound from common null space.** A simple upper bound is derived from the following observation. If $\operatorname{rank}\begin{bmatrix} E+\Delta E \\ A+\Delta A \end{bmatrix} < n$ or if $\operatorname{rank}\begin{bmatrix} E+\Delta E & A+\Delta A \end{bmatrix} < n$, then the pencil $s(E+\Delta E)-(A+\Delta A)$ is singular. Thus,

$$\delta(E,A) \le \min\left\{ \sigma_{\min}\left(\begin{bmatrix} E \\ A \end{bmatrix}\right), \sigma_{\min}\left(\begin{bmatrix} E & A \end{bmatrix}\right) \right\}, \tag{11}$$

which becomes an equality for $n = 1$ or $n = 2$.

**Upper bound from generalized Schur form.** The (real) generalized Schur form [23, Thm. 7.7.2] states that $sE - A \in \mathbb{R}[s]^{n\times n}$ can be reduced to quasi-triangular form by an orthogonal equivalence transformation:

$$Q^{\mathsf{T}}(sE-A)Z = \begin{bmatrix} sE_{11}-A_{11} & \cdots & sE_{1m}-A_{1m} \\ & \ddots & \vdots \\ & & sE_{mm}-A_{mm} \end{bmatrix}. \tag{12}$$

The diagonal blocks $sE_{ii} - A_{ii}$ are either $1\times 1$ (corresponding to a real eigenvalue, an infinite eigenvalue or a singular block) or $2\times 2$ (corresponding to a pair of complex conjugate eigenvalues).

Obviously, $s\widetilde{E} - \widetilde{A} := s(E+\Delta E) - (A+\Delta A)$ becomes singular when any of the diagonal blocks $s\widetilde{E}_{ii} - \widetilde{A}_{ii}$ in its generalized Schur form becomes singular. This directly gives the upper bound

$$\delta(E,A) \le \min_{1\le i\le m} \delta(E_{ii},A_{ii}) = \min_{1\le i\le m} \min\left\{ \sigma_{\min}\left(\begin{bmatrix} E_{ii} \\ A_{ii} \end{bmatrix}\right), \sigma_{\min}\left(\begin{bmatrix} E_{ii} & A_{ii} \end{bmatrix}\right) \right\},$$

where we used (11).

**Upper bound from singular values and vectors of $W_k$.** Let us come back to the least-squares problem (7), which implies

$$\delta(E,A) \le \left\| \begin{bmatrix} E & A \end{bmatrix} Z_k Z_k^{\dagger} \right\|_{\mathrm{F}} \le \left\| \begin{bmatrix} E & A \end{bmatrix} Z_k \right\|_{\mathrm{F}} \left\| Z_k^{\dagger} \right\|_{\mathrm{F}}$$

for any $2n\times(k+1)$ matrix $Z_k$ of the form (6), defined via vectors $x_1,\ldots,x_k \in \mathbb{R}^n$. We make a particular choice of $Z_k$ by considering the partitioning $x = \begin{pmatrix} x_1^{\mathsf{T}} & \cdots & x_k^{\mathsf{T}} \end{pmatrix}^{\mathsf{T}}$ of a (right) singular vector $x \in \mathbb{R}^{kn}$ belonging to $\sigma_{\min}(W_k)$. The structure of $W_k$ implies that $W_k x$ is the vectorization of $\begin{bmatrix} E & A \end{bmatrix} Z_k$, and therefore

$$\left\| \begin{bmatrix} E & A \end{bmatrix} Z_k \right\|_{\mathrm{F}} = \|W_k x\|_2 = \sigma_{\min}(W_k).$$

This gives the upper bounds

$$\delta(E,A) \le \left\| \begin{bmatrix} E & A \end{bmatrix} Z_k \right\|_{\mathrm{F}} \left\| Z_k^{\dagger} \right\|_{\mathrm{F}} \le \sigma_{\min}(W_k)/\sigma_{\min}(Z_k),$$

which are valid for every $k = 1,\ldots,n$. In [15], it is noted that the choice of $k$ is critical for this bound, since it might only be relatively tight for one $k$. As discussed above, the computation

of the smallest singular value and the corresponding singular vector gets expensive, if $k$ and/or $n$ are not small.

In [15], further upper bounds are presented which we do not summarize here. Moreover, several examples show that none of the bounds presented above is tight for all examples. A computationally attractive way to determine or estimate $\delta(E,A)$ thus remains an open problem.

The presentation above focused on real matrices with real perturbations. Up to minor modifications, all developments directly extend to complex matrices with complex perturbations.

## 2.4 Semi-Explicit DAEs

The consideration of general unstructured perturbations $\Delta E, \Delta A$ may become inappropriate when more information on the uncertainty in the coefficients $E,A$ if a DAE (1) is available. For example, for the special case of linear semi-explicit DAEs, the pencil $sE-A$ takes the form

$$sE-A = s\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

Since there is little reason to admit perturbations in the fixed matrix $E$, it makes sense to consider the modified distance to singularity given by

$$\delta_0(E,A) := \min\left\{ \|\Delta A\|_{\mathrm{F}} : sE-(A+\Delta A) \in \mathbb{C}[s]^{n\times n} \text{ is singular} \right\}. \tag{13}$$

At least in principle, it is straightforward to incorporate such linear constraints on the perturbation structure into the structured low-rank approximation framework of Section 2.1. However, it turns out that more can be said about (13).

By [61, Lem. 2.2.26], the matrix pencil $sE-A$ is regular if and only if the $(n-r)\times(n-r)$ rational matrix $G(s) := A_{21}(sI_r - A_{11})^{-1}A_{12} + A_{22}$ has full normal rank. Trivially, $sI_r - A_{11}$ has full normal rank for all $A_{11} \in \mathbb{R}^{r\times r}$ and, consequently, a de-regularizing perturbation is characterized by one of the properties

$$\operatorname{rank}\begin{bmatrix} A_{12}+\Delta A_{12} \\ A_{22}+\Delta A_{22} \end{bmatrix} < n-r \text{ or } \operatorname{rank}\begin{bmatrix} A_{21}+\Delta A_{21} & A_{22}+\Delta A_{22} \end{bmatrix} < n-r.$$

In other words we obtain

$$\delta_0(E,A) = \min\left\{ \sigma_{\min}\left(\begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix}\right), \sigma_{\min}\left(\begin{bmatrix} A_{21} & A_{22} \end{bmatrix}\right) \right\}.$$

This improves a result from [15], where such a statement has been shown only for the case that $r=1$ or $r=n-1$.

## 2.5 Low-Rank Perturbations

A rather different way of imposing structure is to constrain the perturbations to be of low rank. This leads to the concept of *rank-$(\kappa_E,\kappa_A)$ distance to singularity* proposed in [47]:

$$\begin{aligned}\delta_{\kappa_E,\kappa_A}(E,A) := \min\big\{ \big\| \begin{bmatrix} \Delta E & \Delta A \end{bmatrix} \big\|_{\mathrm{F}} : \Delta E, \Delta A \in \mathbb{C}^{n\times n}, \\ \operatorname{rank}\Delta E \le \kappa_E, \operatorname{rank}\Delta A \le \kappa_A, s(E+\Delta E)-(A+\Delta A) \text{ is singular} \big\}.\end{aligned}$$

Here, $\kappa_E, \kappa_A$ are supposed to be (much) smaller than $n$ and we consider the complex case because it is the more natural setting for the developments below.

Low-rank perturbations naturally arise in situations when only a few entries of $E$ and $A$ are subject to uncertainties. Another motivation to consider $\delta_{\kappa_E,\kappa_A}(E,A)$ arises from the suspicion that the low dimensionality of the manifold of rank-constrained matrices could help reduce the cost of nonlinear optimization methods for computing this distance.

Following [47], we now focus on $\delta_{0,1}(E,A)$, that is, $E$ is not perturbed at all and $A$ is subject to rank-1 perturbations:

$$sE - \left(A + \tau u v^{\mathsf{H}}\right) \tag{14}$$

with $u, v \in \mathbb{C}^n \setminus \{0\}$ and a nonzero scalar $\tau \in \mathbb{C}$ determining the perturbation level. It is assumed that $sE-A$ itself is not singular, and therefore $\Lambda(E,A)$, the set of (finite) eigenvalues of $sE-A$, does not coincide with $\mathbb{C}$. We define the rational function $Q : \mathbb{C} \setminus \Lambda(E,A) \to \mathbb{C}$ by

$$Q(s) := v^{\mathsf{H}} R(s) u, \qquad R(s) = (sE-A)^{-1}.$$

Because of

$$\det\left(sE - A - \tau u v^{\mathsf{H}}\right) = \left(1 - \tau Q(s)\right) \det\left(sE - A\right) = \tau\left(\tau^{-1} - Q(s)\right) \det\left(sE - A\right),$$

a scalar $s \in \mathbb{C} \setminus \Lambda(E,A)$ is an eigenvalue of (14) if and only if $\tau^{-1} = Q(s)$. Thus, the perturbed pencil (14) becomes singular if and only if $\tau^{-1} - Q(s)$ vanishes on $\mathbb{C} \setminus \Lambda(E,A)$. Since $Q(s)$ is a meromorphic function on $\mathbb{C}$ with at most $n$ poles, the latter condition is equivalent to

$$\tau^{-1} = Q(s_0), \quad 0 = \frac{\partial^j}{\partial s^j} Q(s)\bigg|_{s=s_0}, \quad j = 1, 2, \ldots, n, \tag{15}$$

for an arbitrary fixed $s_0 \in \mathbb{C} \setminus \Lambda(E,A)$. Note that $\frac{\partial^j}{\partial s^j} Q(s) = v^{\mathsf{H}} C_j(s) u$ with

$$C_j(s) := R(s)(ER(s))^j.$$

The smallest $\tau$ for which (15) holds gives the distance to singularity for *fixed* choices of $u, v$. Optimizing with respect to these vectors finally yields

$$\delta_{0,1}(E,A) = \min\Big\{ \left|v^{\mathsf{H}} R(s_0) u\right|^{-1} : u, v \in \mathbb{C}^n, \|u\|_2 = \|v\|_2 = 1, \\ v^{\mathsf{H}} C_j(s_0) u = 0,\ j = 1, \ldots, n \Big\} \tag{16}$$

for arbitrary fixed $s_0 \in \mathbb{C} \setminus \Lambda(E,A)$; see also [47, Thm. 7]. The constraints in (16) can be expressed as

$$0 = \operatorname{tr}\left(v^{\mathsf{H}} C_j(s_0) u\right) = \operatorname{tr}\left(u v^{\mathsf{H}} C_j(s_0)\right) = \left\langle C_j(s_0), v u^{\mathsf{H}} \right\rangle.$$

In other words, the matrix $v u^{\mathsf{H}}$ is orthogonal to all matrices $C_j(s_0)$ with respect to the matrix inner product $\langle \cdot, \cdot \rangle$. Equivalently, it holds that

$$v u^{\mathsf{H}} \in \mathscr{D} := \left(\operatorname{span}\left\{C_1(s_0), \ldots, C_n(s_0)\right\}\right)^{\perp} \subset \mathbb{C}^{n \times n}.$$

It turns out that the space $\mathscr{D}$ does not depend on the particular choice of $s_0$. Summarizing these developments, we arrive at

$$\delta_{0,1}(E,A) = \min\Big\{ \left|\operatorname{tr}(G R(s_0))\right|^{-1} : G^{\mathsf{H}} \in \mathscr{D}, \operatorname{rank} G = 1, \|G\|_{\mathrm{F}} = 1 \Big\}; \tag{17}$$

see [47, Thm. 13] for more details.

A difficulty in the nonconvex optimization problem (17) is that the two constraints $G^{\mathsf{H}} \in \mathscr{D}$ and $\operatorname{rank} G \leq 1$ need to be satisfied simultaneously. On the other hand, each of the two constraints individually constitutes a matrix manifold for which an orthogonal projection can be easily computed. This suggests the use of an alternating projection method on manifolds [44]. Starting from $G_0 = R(s_0)$ (or another suitable choice), the method proposed in [47] constructs the iteration

$$G_{2k+1} := \left(\Pi_{\mathscr{D}} G_{2k}^{\mathsf{H}}\right)^{\mathsf{H}}, \quad G_{2k+2} := \Psi(G_{2k+1}), \quad k = 0, 1, \ldots, \tag{18}$$

where $\Pi_{\mathscr{D}}$ is the orthogonal projection on the linear space $\mathscr{D}$ and $\Psi$ is the orthogonal projection on the manifold of matrices having rank at most 1 (which is simply a rank-1 truncation). It is shown in [47, Prop. 38] that *if* the sequence (18) converges to a matrix $G \neq 0$ then $G$ has rank one and the pencil $sE - \left(A + \operatorname{tr}(GR(s_0))^{-1} G\right)$ is singular. General results on alternating projection methods can be applied to study the local convergence behavior of (18).

The paper [47] also extensively covers the distance to singularity of a Hermitian matrix pencil under Hermitian rank-1 perturbations.

## 3 The Distance to Instability

In this section we turn to the problem of computing the distance to instability, sometimes also referred to as the stability radius, and closely related problems.

### 3.1 The Distance to Instability and Hamiltonian Matrices

The problem of computing the distance to instability in its simplest form is formulated for an ODE system

$$\dot{x}(t) = Ax(t) \tag{19}$$

with $A \in \mathbb{R}^{n\times n}$. The distance to instability is the smallest norm of a perturbation $\Delta$ such that at least one eigenvalue of $A + \Delta$ does not have negative real part. If $A$ itself is not stable, this distance is clearly 0. Assuming that $A$ is stable, this distance satisfies

$$\begin{aligned} r_{\mathbb{F}}^{\|\cdot\|}(A) &:= \min\left\{ \|\Delta\| : \Lambda(A+\Delta) \cap \overline{\mathbb{C}^+} \neq \emptyset \text{ for } \Delta \in \mathbb{F}^{n\times n} \right\} \\ &= \min\left\{ \|\Delta\| : \Lambda(A+\Delta) \cap \mathrm{i}\mathbb{R} \neq \emptyset \text{ for } \Delta \in \mathbb{F}^{n\times n} \right\}. \end{aligned}$$

The appropriate choice of the field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ for the entries of $\Delta$ depends on the application. As pointed out in [58, Chapter 50], one should work with complex perturbations, even for a real matrix $A$, when attempting to draw conclusions about the transient behavior of (19). On the other hand, if one is interested in stability robust to uncertainties in the entries of a real matrix $A$ then clearly $\mathbb{F} = \mathbb{R}$ is the preferred choice. Since $\mathbb{R} \subset \mathbb{C}$ it holds that

$$r_{\mathbb{C}}^{\|\cdot\|}(A) \leq r_{\mathbb{R}}^{\|\cdot\|}(A).$$

The norm for measuring the perturbations in the definition of $r_{\mathbb{F}}^{\|\cdot\|}(A)$ is usually chosen as $\|\cdot\| = \|\cdot\|_2$ or $\|\cdot\| = \|\cdot\|_{\mathrm{F}}$. The choice between these two norms has, at most, a limited impact on the distance to instability. As we will see below, there are always minimal perturbations that have rank 1 (for $\mathbb{F} = \mathbb{C}$) or rank 2 (for $\mathbb{F} = \mathbb{R}$). Thus,

$$r_{\mathbb{C}}^{\|\cdot\|_2}(A) = r_{\mathbb{C}}^{\|\cdot\|_F}(A), \qquad r_{\mathbb{R}}^{\|\cdot\|_2}(A) \le r_{\mathbb{R}}^{\|\cdot\|_F}(A) \le \sqrt{2}\, r_{\mathbb{R}}^{\|\cdot\|_2}(A).$$

In the following, we will therefore simply write $r_{\mathbb{C}}(A)$.

The papers [60, 35] were among the first to consider the distance to instability and established

$$r_{\mathbb{C}}(A) = \min_{\omega \in \mathbb{R}} \sigma_{\min}(\mathrm{i}\omega I_n - A). \tag{20}$$

This result is constructive and gives rise to a minimal perturbation of rank 1, which is obtained from the singular vectors belonging to $\sigma_{\min}(\mathrm{i}\omega I_n - A)$ for the optimal value of $\omega$. The characterization (20) is intimately related to the eigenvalues of the *Hamiltonian matrix*

$$\mathscr{H}(\alpha) := \begin{bmatrix} A & -\alpha I_n \\ \alpha I_n & -A^{\mathsf{T}} \end{bmatrix} \in \mathbb{R}^{2n \times 2n}. \tag{21}$$

In [14, 34], it is shown that $\alpha \ge r_{\mathbb{C}}(A)$ if and only if $H(\alpha)$ has at least one purely imaginary eigenvalue. Based on this result, Byers [14] proposed a bisection method that adapts the value of $\alpha$ by checking whether any of the eigenvalues of $\mathscr{H}(\alpha)$ are purely imaginary. This algorithm converges *globally* and is robust to roundoff error, provided that a structure-preserving algorithm for the Hamiltonian matrix $\mathscr{H}(\alpha)$ is used. As summarized in Chap. 1 by Bunse-Gerstner and Faßbender, the development of such algorithms is another central theme of Volker Mehrmann's work. We also refer to [4] for an overview and [1, 3] for the corresponding software.

## 3.2 The $\mathscr{H}_\infty$ Norm and Even Matrix Pencils

The results discussed in Section 3.1 have been extended into several different directions. One particularly important extension is concerned with the computation of the $\mathscr{H}_\infty$ norm for linear time-invariant control systems of the form

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t), \end{aligned} \tag{22}$$

where $sE - A \in \mathbb{R}[s]^{n \times n}$ is assumed to be regular, $B \in \mathbb{R}^{n \times m}$, and $C \in \mathbb{R}^{p \times n}$. Moreover, $u : [0,\infty) \to \mathbb{R}^m$ is an input control signal and $y : [0,\infty) \to \mathbb{R}^p$ is a measured output signal.

The transfer function of (22) is given by $G(s) = C(sE - A)^{-1}B + D$ and maps inputs to outputs in the frequency domain. We assume that $G(s)$ is analytic and bounded in $\mathbb{C}^+$, for which a sufficient condition is that $\Lambda(E,A) \subset \mathbb{C}^-$ and all infinite eigenvalues are geometrically simple. For such transfer functions, the $\mathscr{H}_\infty$ norm is defined as

$$\|G\|_{\mathscr{H}_\infty} := \sup_{s \in \mathbb{C}^+} \sigma_{\max}(G(s)) = \sup_{\omega \in \mathbb{R}} \sigma_{\max}(G(\mathrm{i}\omega)), \tag{23}$$

where $\sigma_{\max}$ denotes the largest singular value of a matrix.

Comparing (23) with (20), we obtain $\|G\|_{\mathscr{H}_\infty} = 1/r_{\mathbb{C}}(A)$ for the special case that $E = B = C = I_n$ and $D = 0$. To relate to the practically more relevant case of general $B, C$, we have to extend the notion of distance to instability and consider the structured complex stability radius [36, 37] given by

$$\begin{aligned} r_{\mathbb{C}}(A,B,C) :&= \min\left\{ \|\Delta\|_2 : \Lambda(A + B\Delta C) \cap \overline{\mathbb{C}^+} \neq \emptyset \text{ for } \Delta \in \mathbb{C}^{m \times p} \right\} \\ &= \min\left\{ \|\Delta\|_2 : \Lambda(A + B\Delta C) \cap \mathrm{i}\mathbb{R} \neq \emptyset \text{ for } \Delta \in \mathbb{C}^{m \times p} \right\}. \end{aligned} \tag{24}$$

This generalization accounts for perturbations of the system that have a feedback structure (that is, $y(t) = \Delta u(t)$) and thus assesses the robustness of stability with respect to external disturbances.

Provided that $A$ is stable, it holds [36] that

$$\|G\|_{\mathscr{H}_\infty} = \begin{cases} 1/r_{\mathbb{C}}(A,B,C) & \text{if } G(s) \not\equiv 0, \\ \infty & \text{if } G(s) \equiv 0, \end{cases} \tag{25}$$

with $G(s) = C(sI_n - A)^{-1}B$. In other words, a small $\mathscr{H}_\infty$ norm corresponds to a large robustness of stability of the system. The definition of the structured complex stability radius has been further generalized to cover $D \neq 0$ [38, Sec. 5.2] or $E \neq I_n$ [18, 17, 8]. However, in both cases the definition and interpretation of the structured complex stability radius becomes more cumbersome. For $D \neq 0$, the radius does not depend on the perturbations in an affine but in a linear fractional way. The case $E \neq I_n$ requires to carefully treat infinite eigenvalues and to account for perturbations that make the pencil $sE - A$ singular.

The most reliable methods for computing the $\mathscr{H}_\infty$ norm are based on the following extension of the connection between $r_{\mathbb{C}}(A)$ and the Hamiltonian matrix (21) discussed above; see [11, 10, 12] for $E = I_n$ and [5] for general $E$. Provided that $\alpha > \inf_{\omega\in\mathbb{R}} \sigma_{\max}(G(\mathrm{i}\omega))$, the inequality $\|G\|_{\mathscr{H}_\infty} \geq \alpha$ holds if and only if the *even matrix pencil*

$$s\mathscr{E} - \mathscr{A}(\alpha) := \left[\begin{array}{cc|cc} 0 & -sE^{\mathsf{T}} - A^{\mathsf{T}} & -C^{\mathsf{T}} & 0 \\ sE - A & 0 & 0 & -B \\ \hline -C & 0 & \alpha I_p & -D \\ 0 & -B^{\mathsf{T}} & -D^{\mathsf{T}} & \alpha I_m \end{array}\right] \in \mathbb{R}[s]^{2n+m+p\times 2n+m+p} \tag{26}$$

has purely imaginary eigenvalues. Such pencils are closely related to skew-Hamiltonian/Hamiltonian pencils, for which structure-preserving algorithms are discussed in Chap. 1 by Bunse-Gerstner and Faßbender; see [6, 7] for recently released software. By additionally exploiting eigenvalue and, optionally, eigenvector information, methods based on (26) can be implemented such that they converge globally quadratically or even faster [22].

The structured *real* stability radius $r_{\mathbb{R}}^{\|\cdot\|_2}(A,B,C)$ with respect to the matrix 2-norm is defined as in (24), but with the perturbation restricted to stay real: $\Delta \in \mathbb{R}^{m\times p}$. It turns out that the computation of $r_{\mathbb{R}}^{\|\cdot\|_2}(A,B,C)$ is more difficult compared to $r_{\mathbb{C}}(A,B,C)$. Provided that $G(s) = C(sI_n - A)^{-1}B$ is analytic and bounded in $\mathbb{C}^+$, the celebrated expression

$$r_{\mathbb{R}}^{\|\cdot\|_2}(A,B,C) = \left(\sup_{\omega\in\mathbb{R}} \inf_{\gamma\in(0,1]} \sigma_2\left(\begin{bmatrix} \mathrm{Re}(G(\mathrm{i}\omega)) & -\gamma\mathrm{Im}(G(\mathrm{i}\omega)) \\ \frac{1}{\gamma}\mathrm{Im}(G(\mathrm{i}\omega)) & \mathrm{Re}(G(\mathrm{i}\omega)) \end{bmatrix}\right)\right)^{-1} \tag{27}$$

holds [53], where $\sigma_2$ denotes the second largest singular value of a matrix. Again, this characterization is constructive in the sense that a minimal rank-2 perturbation can be constructed from singular vectors. The inner optimization problem in (27) is unimodal [53], implying that every local minimum is a global minimum and thus allowing for reliable and efficient numerical optimization. For the outer optimization problem, a numerical method similar to the stability radius and $\mathscr{H}_\infty$ norm computation is devised in [56].

The definition and computation of robust stability measures for linear delay DAEs is surveyed in Chap. 19 by Linh and Thuan.

### 3.3 The Distance to Instability and Pseudospectra

The approaches discussed above have been designed for relatively small problems and it is by no means clear that they can be well adapted to large-scale problems. This is mainly because of the lack of an efficient large-scale algorithm for deciding whether a Hamiltonian matrix or an even matrix pencil has purely imaginary eigenvalues; see [41] for a more detailed discussion.

Recent work by Guglielmi and Overton [31] has initiated the development of novel algorithms that are based on pseudospectra and appear to be more suitable for large-scale problems. Moreover, as we will see in Section 3.6 below, the framework offers much more flexibility for incorporating structure.

Given a matrix $A \in \mathbb{R}^{n\times n}$ and $\varepsilon > 0$, we consider the $\varepsilon$-pseudospectrum

$$\Lambda_\varepsilon^{\mathbb{F},\|\cdot\|}(A) := \big\{\lambda \in \mathbb{C} : \lambda \in \Lambda(A+\Delta) \text{ for some } \Delta \in \mathbb{F}^{n\times n},\ \|\Delta\| < \varepsilon\big\}$$

for $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and a matrix norm $\|\cdot\|$ (usually $\|\cdot\| = \|\cdot\|_2$ or $\|\cdot\| = \|\cdot\|_{\mathrm{F}}$). The pseudospectrum grows as $\varepsilon$ increases; see Figure 1 for an illustration.

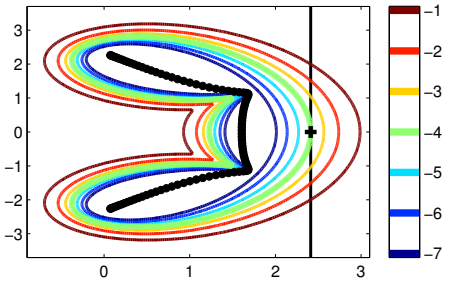


**Fig. 1.** Contour lines of the pseudospectrum $\Lambda_\varepsilon^{\mathbb{C},\|\cdot\|_2}(A)$ for the so called Grcar matrix and $\varepsilon = 10^{-1}, 10^{-2}, \ldots, 10^{-7}$. Black crosses denote eigenvalues and the black vertical line denotes the pseudospectral abscissa for $\varepsilon = 10^{-4}$.

**Basic idea.** By definition, the distance to instability for a stable matrix is the smallest value of $\varepsilon$ for which the $\varepsilon$-pseudospectrum touches the imaginary axis. In terms of the $\varepsilon$-pseudospectral abscissa

$$\alpha^{\mathbb{F},\|\cdot\|}(\varepsilon, A) := \sup\big\{\mathrm{Re}(\lambda) : \lambda \in \Lambda_\varepsilon^{\mathbb{F},\|\cdot\|}(A)\big\},$$

this means that $\varepsilon = r_{\mathbb{F}}^{\|\cdot\|}(A)$ satisfies

$$\alpha^{\mathbb{F},\|\cdot\|}(\varepsilon, A) = 0. \tag{28}$$

Hence, once we have an efficient method for evaluating the quantity $\alpha^{\mathbb{F},\|\cdot\|}(\varepsilon, A)$, root finding algorithms can be used to determine $r_{\mathbb{F}}^{\|\cdot\|}(A)$.

**Computation of the $\varepsilon$-pseudospectral abscissa.** In a large-scale setting, one benefits from the fact that every element of $\Lambda_\varepsilon^{\mathbb{F},\|\cdot\|}(A)$ can be realized by a low-rank perturbation. In particular, for complex pseudospectra we have

$$\begin{aligned}\Lambda_\varepsilon^{\mathbb{C}}(A) :&= \Lambda_\varepsilon^{\mathbb{C},\|\cdot\|_2}(A) = \Lambda_\varepsilon^{\mathbb{C},\|\cdot\|_{\mathrm{F}}}(A)\\ &= \big\{\lambda \in \mathbb{C} : \lambda \in \Lambda(A+\Delta) \text{ for some } \Delta \in \mathbb{C}^{n\times n},\ \mathrm{rank}\,\Delta = 1,\ \|\Delta\| < \varepsilon\big\};\end{aligned}$$

see [58] for this and many other properties of pseudospectra. The procedure proposed in [31] attempts to construct a sequence of perturbation matrices $\Delta_k := \varepsilon u_k v_k^{\mathsf{H}}$ with $\|u_k\|_2 = \|v_k\|_2 =$

1 for $k = 1, 2, \ldots$, such that one of the eigenvalues of $A + \Delta_k$ converges to the rightmost point of the $\varepsilon$-pseudospectrum. The following standard eigenvalue perturbation result is instrumental in determining these perturbations.

**Lemma 1 ([57]).** *Let $M_0, M_1 \in \mathbb{R}^{n\times n}$, $t \in \mathbb{R}$, and consider the matrix family $M(t) = M_0 + tM_1$. Let $\lambda(t)$ be an eigenvalue of $M(t)$ converging to a simple eigenvalue $\lambda_0$ of $M_0$ with corresponding right and left eigenvectors $x_0, y_0 \in \mathbb{C}^n$ for $t \to 0$. Then $y_0^{\mathsf{H}} x_0 \neq 0$ and $\lambda(t)$ is analytic near $t = 0$ with*

$$\left.\frac{\mathrm{d}\lambda(t)}{\mathrm{d}t}\right|_{t=0} = \frac{y_0^{\mathsf{H}} M_1 x_0}{y_0^{\mathsf{H}} x_0}.$$

For the initial perturbation $\Delta_1 = \varepsilon u_1 v_1^{\mathsf{H}}$ one computes the rightmost eigenvalue $\lambda_0$ of $A$ (assuming that it is simple) with corresponding right and left eigenvectors $x_0, y_0 \in \mathbb{C}^n$. The eigenvectors are normalized such that $y_0^{\mathsf{H}} x_0 > 0$, a property that is called RP-compatibility in [31]. Then for $A_1(t) := A + t\varepsilon u_1 v_1^{\mathsf{H}}$ with an eigenvalue $\lambda_1(t)$ converging to $\lambda_0$ for $t \to 0$ we obtain

$$\mathrm{Re}\left(\left.\frac{\mathrm{d}\lambda_1(t)}{\mathrm{d}t}\right|_{t=0}\right) = \varepsilon \frac{\mathrm{Re}\left(y_0^{\mathsf{H}} u_1 v_1^{\mathsf{H}} x_0\right)}{y_0^{\mathsf{H}} x_0} \leq \varepsilon \frac{\|y_0\|_2 \|x_0\|_2}{y_0^{\mathsf{H}} x_0}. \tag{29}$$

Equality in (29) holds for $u_1 = y_0/\|y_0\|_2$ and $v_1 = x_0/\|x_0\|_2$, i.e., this choice yields the maximal local growth of the real part of $\lambda_0$. For all subsequent perturbations, consider the matrix family

$$A_k(t) := A + \varepsilon u_{k-1} v_{k-1}^{\mathsf{H}} + t\varepsilon\left(u_k v_k^{\mathsf{H}} - u_{k-1} v_{k-1}^{\mathsf{H}}\right), \quad k = 2, 3, \ldots,$$

which constitutes rank-1 perturbations of norm $\varepsilon$ for $t = 0$ and $t = 1$. Assume that $A_k(t)$ has an eigenvalue $\lambda_k(t)$ converging to a (simple) rightmost eigenvalue $\lambda_{k-1}$ of $A + \Delta_{k-1} = A + \varepsilon u_{k-1} v_{k-1}^{\mathsf{H}}$ for $t \to 0$. Moreover, let $x_{k-1}, y_{k-1} \in \mathbb{C}^n$ with $y_{k-1}^{\mathsf{H}} x_{k-1} > 0$ be the corresponding right and left eigenvectors. Analogously to (29) we obtain

$$\begin{aligned}\mathrm{Re}\left(\left.\frac{\mathrm{d}\lambda_k(t)}{\mathrm{d}t}\right|_{t=0}\right) &= \varepsilon \frac{\mathrm{Re}\left(y_{k-1}^{\mathsf{H}}\left(u_k v_k^{\mathsf{H}} - u_{k-1} v_{k-1}^{\mathsf{H}}\right) x_{k-1}\right)}{y_{k-1}^{\mathsf{H}} x_{k-1}} \\ &\leq \varepsilon \frac{\|y_{k-1}\|_2 \|x_{k-1}\|_2 - \mathrm{Re}\left(y_{k-1}^{\mathsf{H}} u_{k-1} v_{k-1}^{\mathsf{H}} x_{k-1}\right)}{y_{k-1}^{\mathsf{H}} x_{k-1}}.\end{aligned} \tag{30}$$

Again, equality in (30) is achieved for the choice $u_k = y_{k-1}/\|y_{k-1}\|_2$ and $v_k = x_{k-1}/\|x_{k-1}\|_2$. As a consequence, the whole process of computing the pseudospectral abscissa consists of computing the rightmost eigenvalue of a matrix with corresponding right and left eigenvectors, constructing an optimal rank-1 perturbation by using these eigenvectors and repeating this procedure for the perturbed matrix until convergence. In [31] is is shown that this procedure is a fixed point iteration converging to a locally rightmost point of the $\varepsilon$-pseudospectrum (under a weak regularity assumption). A major weakness is that we cannot ensure to find a globally rightmost point and thus, at least in principle, only find a lower bound for the $\varepsilon$-pseudospectral abscissa. However, numerical examples reported in [31] indicate that this bound often attains the exact value.

**Connection to low-rank dynamics.** The above procedure is an iterative scheme on the manifold of rank-1 perturbations having norm $\varepsilon$. This naturally leads to the question whether

there exists a continuous path of such matrices converging to the desired optimal perturbation. Indeed, this question has a positive answer [27, 28]. Consider the differential equations

$$\begin{aligned}\dot{u}(t) &= \frac{\mathrm{i}}{2}\operatorname{Im}\left(u(t)^{\mathsf{H}}x(t)y(t)^{\mathsf{H}}v(t)\right) + \left(I_n - u(t)u(t)^{\mathsf{H}}\right)x(t)y(t)^{\mathsf{H}}v(t),\\ \dot{v}(t) &= \frac{\mathrm{i}}{2}\operatorname{Im}\left(v(t)^{\mathsf{H}}x(t)y(t)^{\mathsf{H}}u(t)\right) + \left(I_n - v(t)v(t)^{\mathsf{H}}\right)x(t)y(t)^{\mathsf{H}}u(t),\end{aligned} \tag{31}$$

where $x(t)$ and $y(t)$ are the right and left eigenvectors of unit norm with $y(t)^{\mathsf{H}}x(t) > 0$ corresponding to the rightmost eigenvalue $\lambda(t)$ of the matrix $A + \Delta(t) := A + \varepsilon u(t)v(t)^{\mathsf{H}}$. It has been shown in [27] that if $\lambda(t)$ is simple and smooth, it will tend to a locally rightmost point of the $\varepsilon$-pseudospectrum for $t \to \infty$. In fact, the iteration discussed above corresponds to the explicit Euler method applied to (31). Of course, other, potentially faster and adaptive methods can be used to discretize (31), a major advantage of the continuous formulation apart from its elegance.

**Computation of the distance to instability.** Now that we know how to efficiently determine $\alpha^{\mathbb{C}}(\varepsilon, A)$ for a fixed value of $\varepsilon$, we apply a root-finding algorithm to the nonlinear equation (28) in $\varepsilon$, in order to determine $\varepsilon_* = r_{\mathbb{C}}(A)$. In practice, Newton and Newton-bisection schemes turned out to be effective for this purpose [24, 8]. Let $\lambda(\varepsilon)$ be the rightmost point of $\Lambda_\varepsilon^{\mathbb{C}}(A)$ with the corresponding optimal perturbation matrix $\Delta(\varepsilon) := \varepsilon u(\varepsilon)v(\varepsilon)^{\mathsf{H}}$. Furthermore, let $x(\varepsilon)$ and $y(\varepsilon)$ be the corresponding right and left eigenvectors of $A + \Delta(\varepsilon)$ with $\|x(\varepsilon)\|_2 = \|x(\varepsilon)\|_2 = 1$ and $y(\varepsilon)^{\mathsf{H}}x(\varepsilon) > 0$. Suppose that the rightmost point of $\Lambda_{\widehat{\varepsilon}}^{\mathbb{C}}(A)$ is unique for a given $\widehat{\varepsilon}$. Then $\lambda(\cdot)$ is continuously differentiable at $\widehat{\varepsilon}$ and

$$\left.\frac{\mathrm{d}\lambda(\varepsilon)}{\mathrm{d}\varepsilon}\right|_{\varepsilon=\widehat{\varepsilon}} = \frac{1}{y(\widehat{\varepsilon})^{\mathsf{H}}x(\widehat{\varepsilon})}.$$

We thus have the ingredients for a Newton method to determine the root $\varepsilon_*$ of (28):

$$\varepsilon_{k+1} = \varepsilon_k - y(\varepsilon_k)^{\mathsf{H}}x(\varepsilon_k)\alpha^{\mathbb{C}}(\varepsilon_k, A).$$

**Acceleration.** The iterative procedures discussed above for computing $\alpha^{\mathbb{C}}(\varepsilon, A)$ and $r_{\mathbb{C}}(A)$ can be accelerated. In [42], subspace acceleration techniques for both quantities were proposed for which, under certain mild conditions, locally superlinear convergence was proven. In practice, these subspace methods exhibit a quite robust convergence behavior, their local convergence is observed to be even locally quadratic, and they can be much faster than the methods from [31, 24]. However, it is currently not clear how these methods can be extended to the more general situations discussed below. The vector extrapolation techniques from [51, 52], which sometimes achieve similar speedups, do not have this drawback.

**Real pseudospectra.** One of the beauties of the presented framework is that it seamlessly extends to real pseudospectra. In the real case, the corresponding perturbations have rank 2 and, consequently, the dynamical system (31) needs to be replaced by appropriate rank-2 dynamics. Such dynamics have been proposed and analyzed in [29], both for the matrix 2-norm and the Frobenius norm, yielding efficient algorithms for $\alpha^{\mathbb{R},\|\cdot\|_2}(\varepsilon, A)$ and $\alpha^{\mathbb{R},\|\cdot\|_{\mathrm{F}}}(\varepsilon, A)$. These results are used in [30] to design Newton-type algorithms for the computation of $r_{\mathbb{R}}^{\|\cdot\|_2}(A)$ and $r_{\mathbb{R}}^{\|\cdot\|_{\mathrm{F}}}(A)$.

### 3.4 The $\mathscr{H}_\infty$ Norm and Spectral Value Sets

The basic ideas from Section 3.3 can be generalized in a direct manner to structured complex stability radii $r_{\mathbb{C}}(A, B, C)$ with $B \in \mathbb{R}^{n\times m}$ and $C \in \mathbb{R}^{p\times n}$. Instead of $\Lambda_\varepsilon^{\mathbb{C}}(A)$, one has to consider structured complex pseudospectra of the form

$$\Lambda_\varepsilon^{\mathbb{C}}(A,B,C) = \left\{\lambda \in \mathbb{C} : \lambda \in \Lambda(A+B\Delta C) \text{ for some } \Delta \in \mathbb{C}^{m\times p},\, \|\Delta\| < \varepsilon\right\}.$$

By (25), this yields an approach for computing the $\mathscr{H}_\infty$ norm of a transfer function $G(s) = C(sI_n - A)^{-1}B$ with a possibly large $n$. However, in order to extend this to transfer functions $G(s) = C(sI_n - A)^{-1}B + D$ with nonzero $D$ one has to consider much more complicated pseudospectral structures (called spectral value sets in [38]), given by

$$\Lambda_\varepsilon^{\mathbb{C}}(A,B,C,D) = \left\{\lambda \in \mathbb{C} : \lambda \in \Lambda\left(A+B\Delta(I_p - D\Delta)^{-1}C\right) \right.$$
$$\left. \text{for some } \Delta \in \mathbb{C}^{m\times p},\, \|\Delta\| < \varepsilon\right\}.$$

As shown in [24], such spectral value sets can again be realized by rank-1 perturbations, allowing for an extension of the algorithms discussed above. A further extension has been made in [8] for transfer functions of the form $G(s) = C(sE-A)^{-1}B+D$. There, an embedding of the the original control system into system of larger dimension is used to eliminate $D$ which drastically simplifies the analysis of the algorithm. However, one has to consider structured pseudospectra of a matrix pencil instead of a matrix. Special care must be taken of possible perturbations of the infinite eigenvalues or perturbations that make the pencil singular. In [8] further improvements have been made with regard to the choice of the eigenvalues to follow during the computation of the pseudospectral abscissa. In particular, it is important to not only consider the location of the eigenvalues but also their sensitivity with respect to the perturbation structure. This is related to the concepts of controllability and observability of the underlying control system (22).

### 3.5 The Implicit Determinant Method

Another class of algorithms goes back to Freitag and Spence [19]. For the determination of the complex stability radius $r_{\mathbb{C}}(A)$ of a stable matrix $A \in \mathbb{R}^{n\times n}$ this approach again makes use of a spectral characterization involving the Hamiltonian matrix $\mathscr{H}(\alpha)$ defined in (21), however in a completely different way. Denote by $\alpha_*$ the smallest value of $\alpha$ such that $\mathscr{H}(\alpha)$ has a purely imaginary eigenvalue $\mathrm{i}\omega_*$, that is,

$$\left(\mathscr{H}(\alpha_*) - \mathrm{i}\omega_* I_{2n}\right) x_* = 0$$

for some eigenvector $x_* \in \mathbb{C}^{2n}$. This value of $\alpha_*$ coincides with $r_{\mathbb{C}}(A)$ and $\omega_*$ is an optimal frequency according to (20). Due to the spectral symmetry of $\mathscr{H}(\alpha_*)$, the eigenvalue $\mathrm{i}\omega_*$ generically forms a Jordan block of dimension two, which will be assumed throughout this section.

Motivated by related methods for determining bifurcation points in parameter-dependent nonlinear systems, the basic idea of [19] consists of setting up a (well-conditioned) system of two nonlinear equations for which $(\omega_*, \alpha_*)$ is a regular solution. For this purpose, a normalization vector $c \in \mathbb{C}^{2n}$ is chosen such that $c^{\mathsf{H}} x_* \neq 0$. Then the genericity assumption implies that the bordered matrix

$$\mathscr{M}(\omega,\alpha) := \begin{bmatrix} \mathscr{H}(\alpha) - \mathrm{i}\omega I_{2n} & \mathscr{J}c \\ c^{\mathsf{H}} & 0 \end{bmatrix} \quad \text{with} \quad \mathscr{J} = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$$

is nonsingular at $(\omega_*, \alpha_*)$ and hence it is also nonsingular for all pairs $(\omega, \alpha)$ in a neighborhood of $(\omega_*, \alpha_*)$. Thus, for all such pairs the linear system of equations

$$\begin{bmatrix} \mathscr{H}(\alpha) - \mathrm{i}\omega I_{2n} & \mathscr{J}c \\ c^{\mathsf{H}} & 0 \end{bmatrix} \begin{pmatrix} x(\omega,\alpha) \\ f(\omega,\alpha) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{32}$$

has a unique solution. By Cramer's rule it holds

$$f(\omega,\alpha) = \frac{\det(\mathscr{H}(\alpha) - \mathrm{i}\omega I_{2n})}{\det(\mathscr{M}(\omega,\alpha))}.$$

In particular, $f(\omega,\alpha) = 0$ is equivalent to $\det(\mathscr{H}(\alpha) - \mathrm{i}\omega I_{2n}) = 0$. From the algebraic properties of the eigenvalue $\mathrm{i}\omega_*$ it follows [19, Lem. 4(a)] that $\partial f(\omega_*,\alpha_*)/\partial\omega = 0$. In summary, $(\omega_*,\alpha_*)$ is a solution of the nonlinear system

$$0 = g(\omega,\alpha) := \begin{pmatrix} f(\omega,\alpha) \\ \partial f(\omega,\alpha)/\partial\omega \end{pmatrix}. \tag{33}$$

Moreover, it can be shown [19] that this solution is regular. Hence, Newton's method applied to $0 = g(\omega,\alpha)$ converges locally quadratically to $(\omega_*,\alpha_*)$.

To implement Newton's method, we need to evaluate $g(\omega,\alpha)$ and its Jacobian, which amounts to evaluating first- and second-order derivatives of $f(\omega,\alpha)$. By differentiating the relation (32), it turns out that all these derivatives can be computed by solving four linear systems with $\mathscr{M}(\omega,\alpha)$. Having these derivatives computed at the $k$th iterate $(\omega_k,\alpha_k)$, the next iterate of Newton's method is determined by first solving the $2\times 2$ linear system

$$\begin{bmatrix} \partial f(\omega_k,\alpha_k)/\partial\omega & \partial f(\omega_k,\alpha_k)/\partial\alpha \\ \partial^2 f(\omega_k,\alpha_k)/\partial\omega^2 & \partial^2 f(\omega_k,\alpha_k)/\partial\omega\partial\alpha \end{bmatrix} \begin{pmatrix} \Delta\omega_k \\ \Delta\alpha_k \end{pmatrix} = -g(\omega_k,\alpha_k),$$

and subsequently setting

$$\begin{pmatrix} \omega_{k+1} \\ \alpha_{k+1} \end{pmatrix} := \begin{pmatrix} \omega_k \\ \alpha_k \end{pmatrix} + \begin{pmatrix} \Delta\omega_k \\ \Delta\alpha_k \end{pmatrix}.$$

Numerical examples in [19] show the effectivity of this method. Since only a few linear systems have to be solved, it also has potential for large-scale systems. Similarly as for the pseudospectral approach, the method is only guaranteed to converge locally and it may converge to a solution of (33) that is different from $(\omega_*,\alpha_*)$. However, the situation is less bleak in practice; numerical results reveal that the method often converges to the correct solution for the initial values proposed in [19]. As discussed in [32], there is always the possibility to check whether $\mathscr{H}(\alpha_*)$ has purely imaginary eigenvalues, but this global optimality certificate may become too expensive for large-scale systems.

As discussed in [21], the described algorithm can be extended in a rather straightforward manner to $\mathscr{H}_\infty$ norm computations, even for the general case of descriptor systems. For this purpose, one only needs to replace the $(1,1)$-block in $\mathscr{M}(\omega,\alpha)$ by a Hamiltonian matrix $\widetilde{\mathscr{A}}(\alpha) - \mathrm{i}\omega\widetilde{\mathscr{E}}$ that can be derived from the even pencil $s\mathscr{E} - \mathscr{A}(\alpha)$ in (26); see also [61].

The extension to the real 2-norm stability radius proposed in [20] is more involved. As a basis, the $4\times 4$ Hamiltonian matrix constructed in [56] instead of the $2n\times 2n$ Hamiltonian matrix $H(\alpha)$ needs to be used. Moreover, the bordered matrix $\mathscr{M}(\omega,\alpha)$ is needs to be replaced by a matrix in three variables $\mathscr{M}(\omega,\alpha,\gamma)$ due to the fact that one has to optimize in (27) over two parameters $\omega$ and $\gamma$ instead of $\omega$ only.

### 3.6 Structured Distances and Variations

In this section, we briefly discuss existing work on structured distances to instability for structures beyond the real perturbations and fractional perturbations (related to linear control systems) considered above. Dealing with such structures is by no means simple, even

the (usually simpler) problem of the structured distance to singularity, also called structured singular value, often poses great difficulties; see [40] for summary. A notable exception are complex Hamiltonian perturbations, for which an expression based on a unimodal optimization problem, not unlike the one in (27), can be derived [39]. This is discussed in more detail in Chap. 8 by Bora and Karow, which also covers Volker's work on Hamiltonian perturbations.

The lack of simple characterizations for structured singular values complicates the development of efficient algorithms that guarantee global optimality, like Byers' bisection method, for structured distances to instability. In contrast, it is fairly straightforward to incorporate structure in the pseudospectra-based algorithms from Section 3.3. This is because finding a structured perturbation that is optimal in first order is a much simpler than finding a globally optimal perturbation. Considering Lemma 1, the former problem amounts to determining a structured perturbation $\Delta$ with $\|\Delta\| \leq \varepsilon$ such that the real part of $y_0^{\mathsf{H}} M_1 x_0$ becomes maximal. For the Frobenius norm, this optimization problem has an explicit solution in terms of the orthogonal projection of $x_0 y_0^{\mathsf{H}}$ onto the set of structured matrices, which becomes particularly simple in the usual situation when this set is a subspace or a manifold. To be able to address large-scale problems, one needs to restrict the obtained iteration to low-rank matrices, that is, one needs to develop a structured extension of the dynamical system (19). This part is significantly more challenging and has been addressed for (complex and real) Hamiltonian matrices [25], Toeplitz matrices [13], and symplectic matrices [26] so far.

Apart from structure preservation, the flexibility of pseudospectra-based algorithms is also also witnessed by several recent extensions, for example, to nonlinear eigenvalue problems [50] and to the regularization of solvent equations discussed in Chap. 3 by Lin and Schröder.

## 4 Summary and Conclusions

Distance measures are not only useful for quantifying the impact of uncertainty on properties of dynamical systems, they also lead to mathematically very interesting and challenging problems. They touch upon a diversity of current topics, including structured low-rank approximation, eigenvalue optimization, and low-rank dynamics. Despite the fact that distance measures are a classical topic in numerical linear algebra and control, significant progress has been made in incorporating structure and dealing with large-scale problems for robust stability calculations. On other topics, less progress has been made; in particular, the reliable estimation of the distance to the nearest singular matrix pencil remains a widely open problem.

## References

1. Benner, P., Byers, R., Barth, E.: Algorithm 800: Fortran 77 subroutines for computing the eigenvalues of Hamiltonian matrices I: the square reduced method. ACM Trans. Math. Softw. **26**(1), 49–77 (2000)
2. Benner, P., Byers, R., Mehrmann, V., Xu, H.: Robust numerical methods for robust control. Tech. Rep. 06-2004, Institut für Mathematik, TU Berlin (2004)
3. Benner, P., Kressner, D.: Algorithm 854: Fortran 77 subroutines for computing the eigenvalues of Hamiltonian matrices II. ACM Trans. Math. Softw. **32**(2), 352–373 (2006)

4. Benner, P., Kressner, D., Mehrmann, V.: Skew-Hamiltonian and Hamiltonian eigenvalue problems: Theory, algorithms and applications. In: Z. Drmač, M. Marušić, Z. Zutek (eds.) Proceedings of the Conference on Applied Mathematics and Scientific Computing, pp. 3–39. Springer-Verlag, Brijuni (2002)
5. Benner, P., Sima, V., Voigt, M.: $\mathscr{L}_\infty$-norm computation for continuous-time descriptor systems using structured matrix pencils. IEEE Trans. Automat. Control **57**(1), 233–238 (2012)
6. Benner, P., Sima, V., Voigt, M.: FORTRAN 77 subroutines for the solution of skew-Hamiltonian/Hamiltonian eigenproblems – Part I: Algorithms and applications. Preprint MPIMD/13-11, Max Planck Institute Magdeburg (2013). Available from `http://www.mpi-magdeburg.mpg.de/preprints/2013/11/`
7. Benner, P., Sima, V., Voigt, M.: FORTRAN 77 subroutines for the solution of skew-Hamiltonian/Hamiltonian eigenproblems – Part II: Implementation and numerical results. Preprint MPIMD/13-12, Max Planck Institute Magdeburg (2013). Available from `http://www.mpi-magdeburg.mpg.de/preprints/2013/12/`
8. Benner, P., Voigt, M.: A structured pseudospectral method for $\mathscr{H}_\infty$-norm computation of large-scale descriptor systems. Math. Control Signals Systems **26**(2), 303–338 (2014)
9. Boley, D.L.: The algebraic structure of pencils and block Toeplitz matrices. Linear Algebra Appl. **279**(1-3), 255–279 (1998)
10. Boyd, S., Balakrishnan, V.: A regularity result for the singular values of a transfer matrix and a quadratically convergent algorithm for computing its $L_\infty$-norm. Systems Control Lett. **15**(1), 1–7 (1990)
11. Boyd, S., Balakrishnan, V., Kabamba, P.: A bisection method for computing the $H_\infty$ norm of a transfer matrix and related problems. Math. Control Signals Systems **2**(3), 207–219 (1989)
12. Bruinsma, N.A., Steinbuch, M.: A fast algorithm to compute the $H_\infty$-norm of a transfer function matrix. Systems Control Lett. **14**(4), 287–293 (1990)
13. Buttà, P., Guglielmi, N., Noschese, S.: Computing the structured pseudospectrum of a Toeplitz matrix and its extreme points. SIAM J. Matrix Anal. Appl. **33**(4), 1300–1319 (2012)
14. Byers, R.: A bisection method for measuring the distance of a stable matrix to the unstable matrices. SIAM J. Sci. Stat. Comput. **9**, 875–881 (1988)
15. Byers, R., He, C., Mehrmann, V.: Where is the nearest non-regular pencil? Linear Algebra Appl. **285**, 81–105 (1998)
16. Demmel, J.W.: On condition numbers and the distance to the nearest ill-posed problem. Numer. Math. **51**(3), 251–289 (1987)
17. Du, N.H., Linh, V.H., Mehrmann, V.: Robust stability of differential-algebraic equations. In: A. Ilchmann, T. Reis (eds.) Surveys in Differential-Algebraic Equations I, Differential-Algebraic Equations Forum, chap. 2, pp. 63–95. Springer-Verlag, Berlin, Heidelberg (2013)
18. Du, N.H., Thuan, D.D., Liem, N.C.: Stability radius of implicit dynamic equations with constant coefficients on time scales. Systems Control Lett. **60**(8), 596–603 (2011)
19. Freitag, M.A., Spence, A.: A Newton-based method for the calculation of the distance to instability. Linear Algebra Appl. **435**(12), 3189–3205 (2011)
20. Freitag, M.A., Spence, A.: A new approach for calculating the real stability radius. BIT **54**(2), 381–400 (2014)
21. Freitag, M.A., Spence, A., Van Dooren, P.: Calculating the $H_\infty$-norm using the implicit determinant method. SIAM J. Matrix Anal. Appl. **35**(2), 619–634 (2014)

22. Genin, Y., Dooren, P.V., Vermaut, V.: Convergence of the calculation of $\mathscr{H}_\infty$ norms and related questions. In: Proc. 13th Symposium on Mathematical Theory of Networks and Systems, pp. 429–432. Padova (1998)
23. Golub, G.H., Van Loan, C.F.: Matrix Computations, 4th edn. The John Hopkins University Press, Baltimore (2013)
24. Guglielmi, N., Gürbüzbalaban, M., Overton, M.L.: Fast approximation of the $H_\infty$ norm via optimization of spectral value sets. SIAM J. Matrix Anal. Appl. **34**(2), 709–737 (2013)
25. Guglielmi, N., Kressner, D., Lubich, C.: Low-rank differential equations for Hamiltonian matrix nearness problems. Oberwolfach Preprint OWP 2013–01, Mathematisches Forschungsinstitut Oberwolfach (2013). To appear in Numer. Math.
26. Guglielmi, N., Kressner, D., Lubich, C.: Computing extremal points of symplectic pseudospectra and solving symplectic matrix nearness problems. Tech. rep. (2014). To appear in SIAM J. Matrix Anal. Appl.
27. Guglielmi, N., Lubich, C.: Differential equations for roaming pseudospectra: paths to extremal points and boundary tracking. SIAM J. Numer. Anal. **49**(3), 1194–1209 (2011)
28. Guglielmi, N., Lubich, C.: Erratum/addendum: Differential equations for roaming pseudospectra: paths to extremal points and boundary tracking. SIAM J. Numer. Anal. **50**(2), 977–981 (2012)
29. Guglielmi, N., Lubich, C.: Low-rank dynamics for computing extremal points of real pseudospectra. SIAM J. Matrix Anal. Appl. **34**(1), 40–66 (2013)
30. Guglielmi, N., Manetta, M.: Approximating real stability radii (2014). Available from `http://matematica.univaq.it/~guglielm/PAPERS/GuMaSR.pdf`
31. Guglielmi, N., Overton, M.L.: Fast algorithms for the approximation of the pseudospectral abscissa and pseudospectral radius of a matrix. SIAM J. Matrix Anal. Appl. **32**(4), 1166–1192 (2011)
32. He, C., Watson, G.A.: An algorithm for computing the distance to instability. SIAM J. Matrix Anal. Appl. **20**(1), 101–116 (1999)
33. Higham, N.J., Konstantinov, M., Mehrmann, V., Petkov, P.: The sensitivity of computational control problems. IEEE Control Systems Magazine **24**(1), 28–43 (2004)
34. Hinrichsen, D., Motscha, M.: Optimization problems in the robustness analysis of linear state space systems. In: Approximation and optimization (Havana, 1987), *Lecture Notes in Math.*, vol. 1354, pp. 54–78. Springer, Berlin (1988)
35. Hinrichsen, D., Pritchard, A.J.: Stability radii of linear systems. Systems Control Lett. **7**, 1–10 (1986)
36. Hinrichsen, D., Pritchard, A.J.: Stability radius for structured perturbations and the algebraic Riccati equation. Systems Control Lett. **8**, 105–113 (1986)
37. Hinrichsen, D., Pritchard, A.J.: Real and Complex Stability Radii: A Survey, *Progress in Systems and Control Theory*, vol. 6, pp. 119–162. Birkhäuser, Boston (1990)
38. Hinrichsen, D., Pritchard, A.J.: Mathematical Systems Theory I: Modeling, State Space Analysis, Stability and Robustness, *Texts Appl. Math.*, vol. 48. Springer-Verlag, New York (2005)
39. Karow, M.: $\mu$-values and spectral value sets for linear perturbation classes defined by a scalar product. SIAM J. Matrix Anal. Appl. **32**(3), 845–865 (2011)
40. Karow, M., Kokiopoulou, E., Kressner, D.: On the computation of structured singular values and pseudospectra. Systems Control Lett. **59**(2), 122–129 (2010)
41. Kressner, D.: Finding the distance to instability of a large sparse matrix. In: IEEE International Symposium on Computer-Aided Control Systems Design, Munich (2006)
42. Kressner, D., Vandereycken, B.: Subspace methods for computing the pseudospectral abscissa and the stability radius. SIAM J. Matrix Anal. Appl. **35**(1), 292–313 (2014)

43. Kunkel, P., Mehrmann, V.: Differential-Algebraic Equations. Analysis and Numerical Solution. EMS Publishing House, Zürich, Switzerland (2006)
44. Lewis, A.S., Malick, J.: Alternating projections on manifolds. Math. Oper. Res. **33**(1), 216–234 (2008)
45. Markovsky, I.: Low Rank Approximation: Algorithms, Implementation, Applications. Communications and Control Engineering. Springer (2012)
46. Markovsky, I., Usevich, K.: Software for weighted structured low-rank approximation. J. Comput. Appl. Math. **256**, 278–292 (2014)
47. Mehl, C., Mehrmann, V., Wojtylak, M.: On the distance to singularity via low rank perturbations. MATHEON-Preprint 1058, DFG-Forschungszentrum MATHEON (2014)
48. Mehrmann, V., Stykel, T.: Balanced truncation model reduction for large-scale systems in descriptor form. In: P. Benner, V. Mehrmann, D. Sorensen (eds.) Dimension Reduction of Large-Scale Systems, *Lect. Notes Comput. Sci. Eng.*, vol. 45, chap. 3, pp. 89–116. Springer-Verlag, Berlin, Heidelberg, New York (2005)
49. Mengi, E., Yildirim, E.A., Kiliç, M.: Numerical optimization of eigenvalues of Hermitian matrix functions. SIAM J. Matrix Anal. Appl. **35**(2), 699–724 (2014)
50. Michiels, W., Guglielmi, N.: An iterative method for computing the pseudospectral abscissa for a class of nonlinear eigenvalue problems. SIAM J. Sci. Comput. **34**(4), A2366–A2393 (2012)
51. Mitchell, T.: Robust and efficient methods for approximation and optimization of stability measures. Ph.D. thesis, NYU, USA (2014)
52. Mitchell, T., Overton, M.L.: Fast approximation of the $H_\infty$ norm via hybrid expansion-contraction using spectral value sets. Technical report (2014)
53. Qiu, L., Bernhardsson, B., Rantzer, A., Davison, E.J., Young, P.M., Doyle, J.C.: A formula for computation of the real stability radius. Automatica **31**(6), 879–890 (1995)
54. Reis, T.: Circuit synthesis of passive descriptor systems – a modified nodal approach. Int. J. Circ. Theor. Appl. **38**(1), 44–68 (2010)
55. Riaza, R.: Differential-Algebraic Systems. Analytical Aspects and Circuit Applications. World Scientific Publishing Co. Pte. Ltd., Singapore (2008)
56. Sreedhar, J., Dooren, P.V., Tits, A.L.: A fast algorithm to compute the real structured stability radius. pp. 219–230. Springer-Verlag (1996)
57. Stewart, G.W., Sun, J.G.: Matrix Perturbation Theory. Academic Press, New York (1990)
58. Trefethen, L.N., Embree, M.: Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators. Princeton University Press (2005)
59. Trenn, S.: Solution concepts for linear DAEs: A survey. In: A. Ilchmann, T. Reis (eds.) Surveys in Differential-Algebraic Equations I, Differential-Algebraic Equations Forum, pp. 137–172. Springer-Verlag, Berlin, Heidelberg (2013)
60. Van Loan, C.F.: How near is a stable matrix to an unstable matrix? Contemp. Math. **47**, 465–478 (1984)
61. Voigt, M.: On linear-quadratic optimal control and robustness of differential-algebraic systems. Dissertation, Fakultät für Mathematik, Otto-von-Guericke-Universität Magdeburg (2014). Submitted
62. Wicks, M., DeCarlo, R.A.: Computing the distance to an uncontrollable system. IEEE Trans. Automat. Control **36**(1), 39–49 (1991)