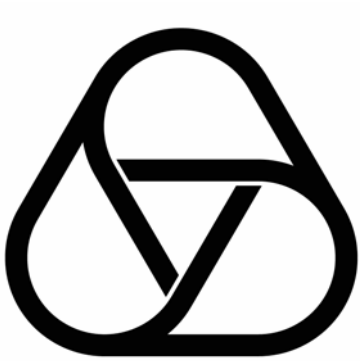
Oberwolfach Preprints



OWP 2013 - 01

NICOLA GUGLIELMI, DANIEL KRESSNER, AND CHRISTIAN LUBICH

Low Rank Differential Equations for Hamiltonian Matrix Nearness Problems

Mathematisches Forschungsinstitut Oberwolfach gGmbH
Oberwolfach Preprints (OWP) ISSN 1864-7596

Oberwolfach Preprints (OWP)

Starting in 2007, the MFO publishes a preprint series which mainly contains research results related to a longer stay in Oberwolfach. In particular, this concerns the Research in Pairs-Programme (RiP) and the Oberwolfach-Leibniz-Fellows (OWLF), but this can also include an Oberwolfach Lecture, for example.

A preprint can have a size from 1 - 200 pages, and the MFO will publish it on its website as well as by hard copy. Every RiP group or Oberwolfach-Leibniz-Fellow may receive on request 30 free hard copies (DIN A4, black and white copy) by surface mail.

Of course, the full copy right is left to the authors. The MFO only needs the right to publish it on its website *www.mfo.de* as a documentation of the research work done at the MFO, which you are accepting by sending us your file.

In case of interest, please send a **pdf file** of your preprint by email to *rip@mfo.de* or *owlf@mfo.de*, respectively. The file should be sent to the MFO within 12 months after your stay as RiP or OWLF at the MFO.

There are no requirements for the format of the preprint, except that the introduction should contain a short appreciation and that the paper size (respectively format) should be DIN A4, "letter" or "article".

On the front page of the hard copies, which contains the logo of the MFO, title and authors, we shall add a running number (20XX - XX).

We cordially invite the researchers within the RiP or OWLF programme to make use of this offer and would like to thank you in advance for your cooperation.

Imprint:

Mathematisches Forschungsinstitut Oberwolfach gGmbH (MFO)
Schwarzwaldstrasse 9-11
77709 Oberwolfach-Walke
Germany

Tel +49 7834 979 50
Fax +49 7834 979 55
Email admin@mfo.de
URL www.mfo.de

The Oberwolfach Preprints (OWP, ISSN 1864-7596) are published by the MFO.

14 January 2013

Low rank differential equations for Hamiltonian matrix nearness problems

Nicola Guglielmi · Daniel Kressner · Christian Lubich

Abstract For a Hamiltonian matrix with purely imaginary eigenvalues, we aim to determine the nearest Hamiltonian matrix such that some or all eigenvalues leave the imaginary axis. Conversely, for a Hamiltonian matrix with all eigenvalues lying off the imaginary axis, we look for a nearest Hamiltonian matrix that has a pair of imaginary eigenvalues. The Hamiltonian matrices can be allowed to be complex or restricted to be real. Such Hamiltonian matrix nearness problems are motivated by applications such as the analysis of passive control systems. They are closely related to the problem of determining extremal points of Hamiltonian pseudospectra. We obtain a characterization of optimal perturbations, which turn out to be of low rank and are attractive stationary points of low-rank differential equations that we derive. This permits us to give fast algorithms - which show quadratic convergence - for solving the considered Hamiltonian matrix nearness problems.

Keywords Hamiltonian pseudospectrum, passivity radius, algebraic Riccati equations, low-rank dynamics, differential equations on Stiefel manifolds

Mathematics Subject Classification (2000) 15A18, 65K05, 93B36, 93B40, 49N35, 65F15, 93B52, 93C05

Nicola Guglielmi
Dipartimento di Matematica Pura ed Applicata, Università degli Studi di L' Aquila, Via Vetoio - Loc. Coppito, I-67010 L' Aquila, Italy
E-mail: guglielm@univaq.it

Daniel Kressner
EPFL-SB-MATHICSE-ANCHP, Station 8, CH-1015 Lausanne, Switzerland
E-mail: daniel.kressner@epfl.ch

Christian Lubich
Mathematisches Institut, Universität Tübingen, Auf der Morgenstelle 10, D–72076 Tübingen, Germany
E-mail: lubich@na.uni-tuebingen.de

1 Introduction

In this paper we propose and study algorithms for the following optimization problems:

(A) *Given a Hamiltonian matrix with no eigenvalues on the imaginary axis, find a nearest Hamiltonian matrix having some purely imaginary eigenvalue.*

(B) *Given a Hamiltonian matrix with all eigenvalues on the imaginary axis, find a nearest Hamiltonian matrix such that arbitrarily close to that matrix there exist Hamiltonian matrices with eigenvalues off the imaginary axis.*

The notion of "nearest" depends on the choice of norm, for which we consider the matrix 2-norm. The Hamiltonian matrices can be allowed to be complex or restricted to be real.

Such Hamiltonian matrix nearness problems arise in several important applications: in the solution of algebraic Riccati equations whose Hamiltonian matrix has eigenvalues on the imaginary axis (see, e.g., [And98, Hes09, PaVL81]), in passivation of linear time invariant control systems (see, e.g., [ABKMM11]), and in the stability of gyroscopic systems (see, e.g., [HKLP00]).

Both problems (A) and (B) are closely related to the problem of finding extremal (locally leftmost or rightmost) points of the structured pseudospectrum (for this notion see, e.g., [BK04, WA05, Ru06, HP05, TE05, KKK10, Ka11]) of a given matrix $A \in \mathcal{M}$ with respect to perturbed matrices on a matrix manifold $\mathcal{M} \subset \mathbb{C}^{n\times n}$ that are ε-close to A in a norm $\|\cdot\|$:

$$\Lambda_\varepsilon(A, \mathcal{M}, \|\cdot\|) = \{\lambda \in \mathbb{C} : \lambda \text{ is an eigenvalue of some } B \in \mathcal{M} \text{ with } \|B - A\| \le \varepsilon\}.$$

We will always consider the matrix 2-norm $\|\cdot\| = \|\cdot\|_2$ in this paper. In the Hamiltonian case considered here we take $\mathcal{M}$ as the real-linear space of *Hamiltonian matrices*, i.e., those satisfying (with even dimension $n = 2d$)

$$JA \text{ is complex hermitian (or real symmetric), where } J = \begin{pmatrix} 0 & I_d \\ -I_d & 0 \end{pmatrix}.$$

We recall that the eigenvalues of a Hamiltonian matrix lie symmetric to the imaginary axis.

The paper is organized as follows. In Section 2 we characterize extremal *complex* Hamiltonian perturbations, which correspond to a locally leftmost or rightmost point in the ε-pseudospectrum. In Section 3 we deal with extremal *real* Hamiltonian perturbations. In both cases, in analogy to [GL11, GL11b, GL12] (see also [GO11]) we derive low-rank matrix differential equations that have the extremizers as attractive stationary points. The ranks are 2 in the complex case and 4 in the real case.

Our approach to solving the Hamiltonian matrix nearness problems (A) and (B) consists of a two-level procedure, where we determine an extremizer

for a fixed perturbation size ε by following the differential equation into a stationary point, and then optimize over ε.

From an algorithmic point of view we are interested in robust variants of (A) and (B). We therefore introduce a thin vertical strip symmetric with respect to the imaginary axis and compute distances with respect to this strip instead of the imaginary axis itself, in agreement with [ABKMM11].

In Section 5 we present a fast algorithm for solving problem (A) and in Section 6 for problem (B). The first algorithm computes a Hamiltonian perturbation of minimal norm such that some pair of eigenvalues of a Hamiltonian matrix with purely imaginary spectrum is moved outside a vertical strip lying symmetric with respect to the imaginary axis, of a preassigned arbitrarily small size. The second algorithm computes a Hamiltonian perturbation of minimal norm such that a pair of eigenvalues of a Hamiltonian matrix A with spectrum bounded away from the imaginary axis is taken inside a given small strip that is symmetric with respect to the imaginary axis. In the examples we shall also deal with the following third problem, which is related to problem (B).

(C) *Given a Hamiltonian matrix with all eigenvalues on the imaginary axis, find a nearest Hamiltonian matrix such that arbitrarily close to that matrix there exist Hamiltonian matrices with all eigenvalues off the imaginary axis.*

This is associated to passivity measures. In Section 7 we present in some detail two interesting applications: passivation of linear control systems and gyroscopic stability.

There are various extensions to the present work, which we do not report here: While we consider only the matrix 2-norm in this paper, a related but different theory and corresponding algorithms can be given also for the Frobenius norm. Analogous matrix nearness problems to this paper can be studied also for symplectic matrices, where the unit circle assumes the role of the imaginary axis for the Hamiltonian case. Another interesting extension is to Hamiltonian matrices with additional (block) structure, as they arise in control systems.

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 $\dot{} = d/dt$.

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 λ_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 λ_0, that is, $(C_0 - \lambda_0 I)y = 0$ and $x^*(C_0 - \lambda_0 I) = 0$. Then, $x_0^* y_0 \neq 0$ and $\lambda(t)$ is differentiable near $t = 0$ with*

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

2 Rank-2 dynamics to extremal points in complex Hamiltonian pseudospectra

In this section we study Hamiltonian perturbations to the given Hamiltonian matrix A that correspond to extremal (locally leftmost or rightmost) points in the Hamiltonian ε-pseudospectrum. It is shown that to each extremal point the corresponding extremizer can be chosen of rank at most 2. We then proceed to derive and study a differential equation on the manifold of Hamiltonian rank-2 matrices that has the rank-2 extremizers as attractive stationary points.

2.1 Extremizers

Theorem 2.1 *For a Hamiltonian matrix $A \in \mathbb{C}^{n\times n}$ and $\varepsilon > 0$, let $\lambda_* \notin i\mathbb{R}$ be a locally rightmost point in the Hamiltonian 2-norm ε-pseudospectrum, implying that there exists a Hamiltonian matrix $E \in \mathbb{C}^{n\times n}$ of unit 2-norm such that λ_* is an eigenvalue of $A+\varepsilon E$. If λ_* is a simple eigenvalue of $A+\varepsilon E$ and not an eigenvalue of A, then there exists a unique matrix E_* of rank 2 such that λ_* is an eigenvalue of $A+\varepsilon E_*$ with the same left and right eigenvectors as for $A+\varepsilon E$. Moreover, the nonzero eigenvalues of JE_* are $+1$ and -1.*

The proof of Theorem 2.1 shows that there always exists an extremizer of rank 2, but there are also extremizers of any rank ≥ 2.

A characterization of the rank-2 extremizers is given in the following result.

Theorem 2.2 *Assume that $E_* = J^T V \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} V^*$, where $V \in \mathbb{C}^{n\times 2}$ has orthonormal columns. Let $\lambda_* \notin i\mathbb{R}$ be a simple eigenvalue of $A+\varepsilon E_*$, with left and right eigenvectors x and y, respectively, both of unit norm and with $x^*y > 0$. Let $Y = (Jx, y) \in \mathbb{C}^{n\times 2}$, and $T \in \mathbb{C}^{2\times 2}$ be defined by*

$$T = RPR^* \quad \text{with} \quad R = V^*Y \in \mathbb{C}^{2\times 2} \quad \text{and} \quad P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then the following two statements are equivalent:

1. *Every differentiable path $(E(t), \lambda(t))$ (for small $t \geq 0$) such that $E(t)$ is Hamiltonian with $\|E(t)\|_2 \leq 1$ and $\lambda(t)$ is an eigenvalue of $A+\varepsilon E(t)$, with $E(0) = E_*$ and $\lambda(0) = \lambda_*$, has $\operatorname{Re}\dot{\lambda}(0) \leq 0$.*
2. *Y and V have the same range, and $T = \begin{pmatrix} +\tau_1 & 0 \\ 0 & -\tau_2 \end{pmatrix}$ with $\tau_1, \tau_2 > 0$.*

Proof (of Theorem 2.1) (a) Let $\widehat{V} \in \mathbb{C}^{n\times n}$ be a unitary matrix obtained from a QR factorization of $Y = (Jx, y)$:

$$Y = \widehat{V} \begin{pmatrix} R \\ 0 \end{pmatrix} \quad \text{with } R \in \mathbb{C}^{2\times 2}.$$

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

$$E(t) = J^T \widehat{V} \widehat{S}(t) \widehat{V}^*$$

with a hermitian matrix

$$\widehat{S}(t) \in \mathbb{C}^{n\times n} = \begin{pmatrix} S_{11}(t) & S_{21}^*(t) \\ S_{21}(t) & S_{22}(t) \end{pmatrix}, \quad S_{11}(t) \in \mathbb{C}^{2\times 2},$$

We can choose the first 2 columns of the unitary matrix $\widehat{V}$ such that

$$S_{11}(0) = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix}.$$

Clearly, σ_1, σ_2 are real and have absolute value at most 1. In the following, we will in fact show that $|\sigma_1| = |\sigma_2| = 1$ and that σ_1, σ_2 have opposite sign.

(b) Let $\lambda(t)$ be the rightmost eigenvalue of $A + \varepsilon E(t)$, so that $\lambda(0) = \lambda_*$. Since λ_* is locally rightmost, we have

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

Denoting by $\langle A, B\rangle = \operatorname{trace}(A^* B)$ the standard matrix inner product and by $A_{\text{herm}} = \frac{1}{2}(A + A^*)$ the hermitian part of a matrix, we observe

$$\begin{aligned} 0 \geq \operatorname{Re}(x^* \dot{E}(0) y) &= \langle (Jxy^*)_{\text{herm}}, J\dot{E}(0)\rangle = \langle YPY^*, \widehat{V}\dot{\widehat{S}}(0)\widehat{V}^*\rangle \\ &= \langle \widehat{V}^* YPY^* \widehat{V}, \dot{\widehat{S}}(0)\rangle = \langle RPR^*, \dot{S}_{11}(0)\rangle = \langle T, \dot{S}_{11}(0)\rangle. \end{aligned} \tag{2.1}$$

This inequality holds for every path $\widehat{S}(t)$ of unit norm, which in turn implies that it holds for every path $S_{11}(t)$ of norm at most one.

Suppose that one of the diagonal entries of $S_{11}(0)$, say σ_2, has absolute value less than one. Then, locally, the following three different paths are admissible:

$$S_{11}^{(1)}(t) = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \pm t \end{pmatrix}, \; S_{11}^{(2)}(t) = \begin{pmatrix} a(t) & \pm t \\ \pm t & b(t) \end{pmatrix}, \; S_{11}^{(3)}(t) = \begin{pmatrix} a(t) & \mp it \\ \pm it & b(t) \end{pmatrix},$$

where in the latter two cases $a(t), b(t)$ are chosen such that the matrices have the eigenvalues σ_1 and σ_2 and $\dot{a}(0) = \dot{b}(0) = 0$. We then obtain from (2.1) that

$$\begin{aligned} \langle T, \dot{S}_{11}^{(1)}(0)\rangle = \pm t_{22} &\implies t_{22} = 0, \\ \langle T, \dot{S}_{11}^{(2)}(0)\rangle = \pm(t_{21} + \overline{t_{21}}) &\implies \operatorname{Re}(t_{21}) = 0, \\ \langle T, \dot{S}_{11}^{(3)}(0)\rangle = \pm(t_{21} - \overline{t_{21}}) &\implies \operatorname{Im}(t_{21}) = 0. \end{aligned}$$

In particular, this means that $T = RPR^*$ is singular and hence R itself is singular. By the definition of R, this would imply that the eigenvectors Jx and y are linearly dependent, which can only happen for an eigenvalue λ_* on the imaginary axis, in contradiction to our assumption. Therefore, $|\sigma_1| = |\sigma_2| = 1$.

(c) Now suppose that σ_1 and σ_2 have the same sign, so that $S_{11}(0) = \pm I_2$. We then choose

$$S_{11}(t) = \pm e^{-tM}$$

with a hermitian positive definite matrix M, so that $\dot{S}(0) = \mp M$. The inequality $0 \geq \langle T, \dot{S}(0)\rangle = \mp \langle T, M\rangle$ would then hold for every positive definite hermitian M, which would imply that T is positive semi-definite (negative semi-definite). This is not possible, as P and therefore also $T = RPR^*$ is an indefinite nonsingular matrix. In view of this contradiction, $S_{11}(0)$ must have two eigenvalues of different sign, $+1$ and -1.

(d) Finally, $|\sigma_1| = |\sigma_2| = 1$ implies that $S_{21}(0) = 0$, since otherwise one of the first two columns of $\hat{S}(0)$ has norm larger than one, in contradiction to $\|\hat{S}(0)\|_2 = 1$. This shows that $A + \varepsilon E_*$, with

$$E_* = J^T \hat{V} \begin{pmatrix} S_{11}(0) & 0 \\ 0 & 0 \end{pmatrix} \hat{V}^*,$$

satisfies $(A+\varepsilon E_*)y = (A+\varepsilon E)y = \lambda_* y$ and $x^*(A+\varepsilon E_*) = x^*(A+\varepsilon E) = \lambda_* x^*$. The rank-2 matrix E_* is unique because the block structure of $\widehat{S}(0)$ and the orthogonality of the columns of $\widehat{V}$ show that for every other rank-2 reduction of $E(t)$ we have $Ey = 0$ or $EJx = 0$, which would imply that λ is an eigenvalue of A, contrary to the assumption. □

Proof (of Theorem 2.2)

1. implies 2.: We know from the proof of Theorem 2.1 that Y and V have the same range, and we write $Y = VR$ with an invertible 2×2 matrix R. We note that $T = RPR^*$ is the matrix of which we want to show that it is diagonal with entries of different sign. Let us denote

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

We consider the path of matrices of unit 2-norm

$$JE(t) = Ve^{tZ} De^{-tZ} V^* \quad \text{with a skew-hermitian matrix } Z \in \mathbb{C}^{2\times 2}$$

and the corresponding path of eigenvalues $\lambda(t)$ of $A + \varepsilon E(t)$ with $\lambda(0) = \lambda_*$. By *1.*, we have

$$0 \geq \mathrm{Re}(x^* \dot{E}(0) y) = \langle YPY^*, \dot{E}(0)\rangle = \langle T, ZD - DZ\rangle\,,$$

which holds for every skew-hermitian matrix Z. The matrix $ZD - DZ$ is hermitian with zero diagonal. It thus follows that T is diagonal. Since $T = RPR^*$ is hermitian, indefinite and nonsingular, its diagonal entries must be real and have opposite sign. Using the paths

$$E(t) = Ve^{-tM} De^{-tM} V^* \quad \text{with a diagonal positive semi-definite matrix } M$$

we then conclude from

$$0 \geq \mathrm{Re}(x^* \dot{E}(0) y) = -\langle T, MD + DM\rangle$$

that the $(1,1)$ entry of T is positive and the $(2,2)$ entry is negative. This proves *2.*

2. implies 1.: Let $E(t)$ be a continuously differentiable path of Hamiltonian matrices of 2-norm at most one, with $E(0) = J^T V D V^*$. We extend $V \in \mathbb{C}^{n\times 2}$ to a unitary matrix $\widehat{V} \in \mathbb{C}^{n\times n}$ and write $\widehat{V} = (V, V^\perp)$. We write $E(t)$ as

$$JE(t) = \widehat{V}(t)\widehat{S}(t)\widehat{V}(t)^*$$

where we may choose $\widehat{V}(t) = (V(t), V^\perp)$ with range $\mathcal{R}(V(t)) = \mathcal{R}(V)$ and $V^*(t)\dot{V}(t) = 0$. By *2.*, this implies $Y^*\dot{V}(0) = 0$, and therefore, with the matrix $T = RPR^* = V^*YPY^*V$ satisfying $T = \begin{pmatrix} \tau_1 & 0 \\ 0 & -\tau_2 \end{pmatrix}$ with $\tau_1, \tau_2 > 0$, and with $S(t)$ denoting the upper left 2×2 block of $\widehat{S}(t)$,

$$\begin{aligned}
\mathrm{Re}(x^*\dot{E}(0)y) &= \langle YPY^*, \dot{E}(0)\rangle \\
&= \langle YPY^*, \dot{\widehat{V}}(0)\widehat{S}(0)\widehat{V}(0)^* + \widehat{V}(0)\dot{\widehat{S}}(0)\widehat{V}(0)^* + \widehat{V}(0)\widehat{S}(0)\dot{\widehat{V}}(0)^*\rangle \\
&= \langle \begin{pmatrix} R \\ 0 \end{pmatrix} P(R^*, 0), \dot{\widehat{S}}(0)\rangle = \langle RPR^*, \dot{S}(0)\rangle = \langle T, \dot{S}(0)\rangle \\
&= \tau_1 \dot{s}_{11}(0) - \tau_2 \dot{s}_{22}(0) \le 0,
\end{aligned}$$

since $s_{11}(0) = 1$ and $s_{22}(0) = -1$ imply $\dot{s}_{11}(0) \le 0$ and $\dot{s}_{22}(0) \ge 0$. This yields *1.* □

2.2 Rank-2 dynamics for the complex case

Theorem 2.1 motivates to search for a differential equation on the manifold

$$\mathcal{M} = \{E \in \mathbb{C}^{n\times n} : \quad E \text{ is Hamiltonian of rank 2 and the nonzero eigenvalues of } JE \text{ are equal to } +1 \text{ and } -1\}$$

such that it leads to a rightmost point in the Hamiltonian 2-norm pseudospectrum. We represent matrices in $\mathcal{M}$ in a non-unique way as

$$JE = VQV^*,$$

where $V \in \mathbb{C}^{n\times 2}$ has orthonormal columns and $Q \in \mathbb{C}^{2\times 2}$ is a hermitian unitary matrix with eigenvalues $+1$ and -1. Such a Q can be written as a complex Householder matrix

$$Q = I_2 - 2uu^*, \qquad u \in \mathbb{C}^2, \quad u^*u = 1.$$

For a given choice of V and Q, tangent matrices $\dot{E} \in T_E\mathcal{M}$ can then be uniquely written as (cf. [KL07])

$$J\dot{E} = \dot{V}QV^* + V\dot{Q}V^* + VQ\dot{V}^* \quad \text{with} \quad V^*\dot{V} = 0,\ u^*\dot{u} = 0. \tag{2.2}$$

We have given up a unique representation of E for the benefit of the very useful orthogonality relations for $\dot{V}$ and $\dot{u}$.

Our objective is to find differential equations for $u(t)$ and $V(t)$ along which the rightmost eigenvalue $\lambda(t)$ of $A+\varepsilon E(t)$ increases monotonically with t and which have attractive stationary points where the extremality condition of Theorem 2.2 is satisfied. Assuming that $\lambda(t)$ is simple for all t under consideration, we denote the left and right eigenvectors by $x(t)$ and $y(t)$, respectively, both of unit norm and with $x(t)^*y(t) > 0$. We let $Y(t) = (Jx(t), y(t))$. We omit the argument t in the following when its presence is clear from the context.

Motivated by the orthogonality relations in (2.2) and the objective that Y and V have the same range when $\dot{V} = 0$, we make the ansatz, with as yet unknown $f \in \mathbb{C}^2$ and $G \in \mathbb{C}^{2\times 2}$,

$$\begin{aligned} \dot{u} &= (I_2 - uu^*)f \\ \dot{V} &= (I_n - VV^*)YG. \end{aligned} \tag{2.3}$$

We now maximize $\operatorname{Re}\dot{\lambda} = \varepsilon \operatorname{Re}(x^*\dot{E}y)/x^*y$ under the normalizing constraint that $\dot{E}$ has unit Frobenius norm. With $R = V^*Y$ and $T = RPR^*$ - with P given by $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ - we have, inserting (2.2) and (2.3),

$$\begin{aligned} \operatorname{Re}(x^*\dot{E}y) &= \langle (Jxy^*)_{\mathrm{herm}}, JE\rangle = \langle YPY^*, J\dot{E}\rangle \\ &= \langle YPY^*, \dot{V}QV^* + V\dot{Q}V^* + VQ\dot{V}^*\rangle \\ &= \langle YPR^*Q, \dot{V}\rangle + \langle T, \dot{Q}\rangle + \langle QRPY^*, \dot{V}^*\rangle \\ &= \langle T, \dot{Q}\rangle + 2\operatorname{Re}\langle YPR^*Q, \dot{V}\rangle. \end{aligned}$$

We further note that, with the hermitian positive semi-definite 2×2 matrices

$$L = I_2 - uu^*, \quad M = Y^*(I_n - VV^*)Y,$$

we have

$$\begin{aligned} &\langle T, \dot{Q}\rangle = -2\langle T, \dot{u}u^* + u\dot{u}^*\rangle = -4\operatorname{Re}\langle Tu, \dot{u}\rangle = -4\operatorname{Re}\langle LTu, f\rangle \\ &2\operatorname{Re}\langle YPR^*Q, \dot{V}\rangle = 2\operatorname{Re}\langle MPR^*Q, G\rangle. \end{aligned}$$

The normalizing constraint becomes, on inserting (2.2),

$$1 = \|J\dot{E}\|_F^2 = \|\dot{V}QV^*\|_F^2 + \|V\dot{Q}V\|_F^2 + \|VQ\dot{V}^*\|_F^2 = \|\dot{Q}\|_F^2 + 2\|\dot{V}\|_F^2.$$

Because of (2.3) we obtain

$$\begin{aligned} &\|\dot{Q}\|_F^2 = 4\|u\dot{u}^* + \dot{u}u^*\|_F^2 = 8\|\dot{u}\|_2^2 = 8\, f^*Lf \\ &\|\dot{V}\|_F^2 = \operatorname{trace}(G^*MG). \end{aligned}$$

Summing up, we want to find $f \in \mathbb{C}^2$ and $G \in \mathbb{C}^{2\times 2}$ that maximize

$$-4\operatorname{Re}\langle LTu, f\rangle + 2\operatorname{Re}\langle MPR^*Q, G\rangle$$

under the quadratic constraint

$$1 = 8\, f^*Lf + 2\operatorname{trace}(G^*MG).$$

We thus have an optimization problem for $z = \mathrm{vec}(f, G)$ of the type

$$\mathrm{Re}(c^* z) \to \max, \qquad z^* K z = 1$$

with a symmetric positive semi-definite matrix K. Up to a positive factor, this has the same solution as the quadratic minimization problem with a linear constraint

$$\tfrac{1}{2} z^* K z \to \min, \quad \mathrm{Re}(c^* z) = 1,$$

which has the optimality condition

$$Kz = \mu c$$

with a positive Lagrange multiplier μ. In our situation, this equation reads

$$\begin{pmatrix} 8Lf \\ 2MG \end{pmatrix} = \mu \begin{pmatrix} -4LTu \\ 2MPR^*Q \end{pmatrix}.$$

Ignoring the scalar factor μ, we are thus led to choose in (2.3)

$$\begin{pmatrix} f \\ G \end{pmatrix} = \begin{pmatrix} -\frac{1}{2}Tu \\ PR^*Q \end{pmatrix}.$$

This gives us the following system of differential equations for u and V:

$$\begin{aligned} \dot{u} &= -\tfrac{1}{2}\,(I_2 - uu^*)Tu \\ \dot{V} &= (I_n - VV^*)YPY^*VQ. \end{aligned} \tag{2.4}$$

We recall that here $Y = (Jx, y)$ contains the left and right eigenvectors x and y, of unit norm and with $x^*y > 0$, corresponding to an eigenvalue λ of $A + \varepsilon E$ with $JE = VQV^*$ and $Q = I_2 - 2uu^*$ with $u \in \mathbb{C}^2$ of unit norm. Moreover, $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $T = V^*YPY^*V$. Our derivation of these differential equations immediately gives the following monotonicity result.

Theorem 2.3 *Let $E(t) = J^T V(t)(I - 2u(t)u(t)^*)V(t)^*$ with $u(t), V(t)$ satisfying the differential equations* (2.4) *and with initial values such that $u(0) \in \mathbb{C}^2$ has unit Euclidean norm and $V(0) \in \mathbb{C}^{n\times 2}$ has orthonormal columns. If $\lambda(t)$ is a simple eigenvalue of $A + \varepsilon E(t)$, then*

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

As the following result shows in comparison with Theorem 2.2, stationary points of (2.4) are extremizers. Let $u^\perp$ be a vector of unit norm orthogonal to u, and let $\widetilde{T} = (u^\perp, u)^* T (u^\perp, u)$.

Theorem 2.4 *If $\lambda \notin i\mathbb{R}$, then (u, V) is a stationary point of* (2.4) *if and only if Y and V have the same range and $\widetilde{T}$ is diagonal.*

We note that the transformation $\widetilde{V} = V(u^\perp, u)$ yields $JE = \widetilde{V} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \widetilde{V}^*$, and therefore $\widetilde{T}$ corresponds to T of Theorem 2.2. $\widetilde{T}$ has the same eigenvalues as the hermitian matrix $T = RPR^*$, which are of different sign.

Proof From (2.4) it is immediate that (V, u) is a stationary point if and only if the range of Y is contained in the range of V and $(u^\perp)^* T u = 0$. Since Y has rank 2 for an eigenvalue $\lambda \notin i\mathbb{R}$, this yields the result. □

2.3 An illustrative example

Consider the Hamiltonian matrix

$$\begin{pmatrix} \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} & \frac{1}{8} & \frac{1}{9} \\ \frac{1}{5} & \frac{1}{6} & \frac{1}{7} & \frac{1}{8} & \frac{1}{9} & \frac{1}{10} \\ \frac{1}{6} & \frac{1}{7} & \frac{1}{8} & \frac{1}{9} & \frac{1}{10} & \frac{1}{11} \\ -1 & -\frac{1}{2} & -\frac{1}{3} & -\frac{1}{4} & -\frac{1}{5} & -\frac{1}{6} \\ -\frac{1}{2} & -\frac{1}{3} & -\frac{1}{4} & -\frac{1}{5} & -\frac{1}{6} & -\frac{1}{7} \\ -\frac{1}{3} & -\frac{1}{4} & -\frac{1}{5} & -\frac{1}{6} & -\frac{1}{7} & -\frac{1}{8} \end{pmatrix} \tag{2.5}$$

Its eigenvalues are given by

$$\pm 0.4060\mathbf{i}, \qquad \pm 0.002078\mathbf{i}, \qquad \pm 0.0000027 45\mathbf{i}.$$

In Figure 2.3 we show the boundary of the Hamiltonian ε-pseudospectrum, computed by Structured Eigtool [KKK10], and the path of rightmost eigenvalues of $A + \varepsilon E(t)$ for $\varepsilon = 0.1$, where $E(t) = J^T V(t)(I - 2u(t)u(t)^*)V(t)^*$ is computed by solving (2.4).

We note in the right picture of Figure 2.3 that the trajectory approaches the rightmost point of the Hamiltonian ε-pseudospectrum tangentially, as it happens for the unstructured ε-pseudospectrum case. This can be explained by analogous arguments.

3 Rank-4 dynamics to extremal points in real Hamiltonian pseudospectra

3.1 Extremizers

We have the following real analog of Theorem 2.1.

Theorem 3.1 *For a Hamiltonian matrix $A \in \mathbb{R}^{n\times n}$ and $\varepsilon > 0$, let $\lambda_* \notin \mathbb{R} \cup i\mathbb{R}$ be a locally rightmost point in the Hamiltonian 2-norm ε-pseudospectrum, implying that there exists a real Hamiltonian matrix $E \in \mathbb{R}^{n\times n}$ of unit 2-norm such that λ_* is an eigenvalue of $A+\varepsilon E$. If λ_* is a simple eigenvalue of $A+\varepsilon E$*

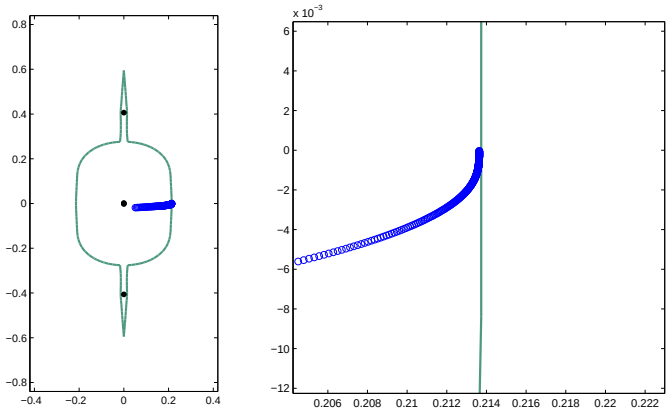


Fig. 2.1 Hamiltonian pseudospectrum of the matrix A in Example (2.5) for $\varepsilon = 0.1$ and trajectory of the rightmost eigenvalue $\lambda(t)$ (in blue) associated to the solution of (2.4). Right picture: zoom close to the rightmost point.

and not an eigenvalue of A, then there exists a unique matrix E_* of rank 4 such that λ_* is an eigenvalue of $A + \varepsilon E_*$ with the same left and right eigenvectors as for $A+\varepsilon E$. Moreover, the nonzero eigenvalues of JE_* are $+1$ and -1, each of multiplicity 2.

For $\lambda_* \in \mathbb{R}$ we have the rank-2 matrix E_* of Theorem 2.1, which is real for real λ_*. The following characterization of the rank-4 real extremizers is given in the following simultaneous analog of Theorem 2.2 and of [GL12, Theorem 3.2] on extremizers for unstructured real perturbations.

Theorem 3.2 *Assume that $E_* = J^T V \operatorname{diag}(1,1,-1,-1) V^T$, where $V \in \mathbb{R}^{n\times 4}$ has orthonormal columns. Let $\lambda_* \notin \mathbb{R} \cup i\mathbb{R}$ be a simple eigenvalue of $A+\varepsilon E_*$, with left and right eigenvectors x and y, respectively, both of unit norm and with $x^*y > 0$. Let $Y = (Jx_R, y_R, Jx_I, y_I) \in \mathbb{R}^{n\times 4}$ and $T \in \mathbb{R}^{4\times 4}$ be defined by*

$$T = RPR^T \quad \text{with} \quad R = V^T Y \in \mathbb{R}^{4\times 4} \quad \text{and} \quad P = \begin{pmatrix} P_2 & 0 \\ 0 & P_2 \end{pmatrix} \text{ with } P_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then the following two statements are equivalent:

1. *Every differentiable path $(E(t), \lambda(t))$ (for small $t \geq 0$) such that $E(t)$ is real Hamiltonian with $\|E(t)\|_2 \leq 1$ and $\lambda(t)$ is an eigenvalue of $A+\varepsilon E(t)$, with $E(0) = E_*$ and $\lambda(0) = \lambda_*$, has $\operatorname{Re} \dot\lambda(0) \leq 0$.*
2. *Y and V have the same range, and $T = \begin{pmatrix} +T_1 & 0 \\ 0 & -T_2 \end{pmatrix}$ with symmetric positive definite blocks $T_1, T_2 \in \mathbb{R}^{2\times 2}$.*

Proof (of Theorem 3.1) Let $\widehat{V} \in \mathbb{R}^{n\times n}$ be an orthogonal matrix obtained from a QR factorization of $Y = (Jx_R, y_R, Jx_I, y_I)$:

$$Y = \widehat{V} \begin{pmatrix} R \\ 0 \end{pmatrix} \quad \text{with } R \in \mathbb{R}^{4\times 4}.$$

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

$$E(t) = J^T \widehat{V} \widehat{S}(t) \widehat{V}^T$$

with a symmetric matrix

$$\widehat{S}(t) \in \mathbb{R}^{n\times n} = \begin{pmatrix} S_{11}(t) & S_{21}^T(t) \\ S_{21}(t) & S_{22}(t) \end{pmatrix}, \quad S_{11}(t) \in \mathbb{R}^{4\times 4},$$

We can choose the first 4 columns of the orthogonal matrix $\widehat{V}$ such that $S_{11}(0)$ is diagonal. The diagonal entries σ_j are real and have absolute value at most 1. Using the same arguments as in the proof of Theorem 2.1 it is shown that $|\sigma_j| = 1$ for $j = 1, 2, 3, 4$ and that two of the σ_j are positive and two are negative. This implies that $S_{21}(0) = 0$, which shows that $A + \varepsilon E_*$, with the rank-4 matrix

$$E_* = J^T \hat{V} \begin{pmatrix} S_{11}(0) & 0 \\ 0 & 0 \end{pmatrix} \hat{V}^T,$$

satisfies $(A+\varepsilon E_*)y = (A+\varepsilon E)y = \lambda_* y$ and $x^*(A+\varepsilon E_*) = x^*(A+\varepsilon E) = \lambda_* x^*$. By the same argument as in the proof of Theorem 2.1, E_* is the unique rank-4 matrix with this property. □

Proof (of Theorem 3.2)

1. implies 2.: We know from the proof of Theorem 3.1 that Y and V have the same range, and we write $Y = VR$ with an invertible 4×4 real matrix R. We denote

$$D = \mathrm{diag}(1, 1, -1, -1).$$

We consider the path of matrices of unit 2-norm

$$JE(t) = Ve^{tZ} De^{-tZ} V^T \quad \text{with a skew-symmetric matrix } Z \in \mathbb{R}^{4\times 4}$$

and the corresponding path of eigenvalues $\lambda(t)$ of $A + \varepsilon E(t)$ with $\lambda(0) = \lambda_*$. By *1.*, we have

$$0 \geq \mathrm{Re}(x^* \dot{E}(0) y) = \langle YPY^T, \dot{E}(0) \rangle = \langle T, ZD - DZ \rangle,$$

which holds for every skew-symmetric matrix Z. The matrix $ZD - DZ$ is symmetric and its upper and lower diagonal 2×2 blocks are zero. It thus follows that T is block-diagonal with 2×2 blocks. Moreover, $T = RPR^T$ is symmetric and nonsingular. Using the paths

$$E(t) = Ve^{-tM} De^{-tM} V^T \quad \text{with positive semi-definite matrices } M$$

that are block diagonal with 2×2 blocks, we then conclude from

$$0 \geq \mathrm{Re}(x^* \dot{E}(0) y) = -\langle T, MD + DM \rangle$$

that the upper diagonal block of T is positive definite and the lower block is negative definite; compare [GL12, Section 3] for an analogous situation in the characterization of extremizers among general real ε-perturbations. This proves *2.*

2. implies 1.: This is proven by a direct generalization of the corresponding argument in the proof of Theorem 2.2. □

3.2 Rank-4 dynamics for the real case

In view of the preceding theorem, we search for a differential equation for matrices of the form

$$JE = VQV^T,$$

where $V \in \mathbb{R}^{n\times 4}$ has orthonormal columns and $Q \in \mathbb{R}^{4\times 4}$ is a symmetric orthogonal matrix with double eigenvalues $+1$ and -1, to be determined in the form of a double Householder transformation

$$Q = I_4 - 2UU^T, \qquad U \in \mathbb{R}^{4\times 2},\ \ U^TU = I_2.$$

For a given choice of V and Q, the tangent matrices $\dot{E} \in T_E\mathcal{M}$ can then be uniquely written as (cf. [KL07])

$$J\dot{E} = \dot{V}QV^T + V\dot{Q}V^T + VQ\dot{V}^T \quad \text{with} \quad V^T\dot{V} = 0,\ U^T\dot{U} = 0. \tag{3.1}$$

In the same way as in Section 2.2, we derive the following differential equations for V and U:

$$\begin{aligned} \dot{U} &= -\tfrac{1}{2}\,(I_4 - UU^T)TU \\ \dot{V} &= (I_n - VV^T)YPY^TVQ. \end{aligned} \tag{3.2}$$

Let us recall that $Y = (Jx_R, y_R, Jx_I, y_I) \in \mathbb{R}^{n\times 4}$ contains 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 λ of $A + \varepsilon E$ with $JE = VQV^T$ and $Q = I - 2UU^T$ with U having two orthonormal columns. Further, $P \in \mathbb{R}^{4\times 4}$ and $T \in \mathbb{R}^{4\times 4}$ are defined as in Theorem 3.2. We obtain again the desired monotonicity result.

Theorem 3.3 *Let $E(t) = J^TV(t)(I - 2U(t)U(t)^T)V(t)^T$ with $U(t), V(t)$ satisfying the differential equations* (3.2) *and with initial values such that $U(0) \in \mathbb{R}^{4\times 2}$ and $V(0) \in \mathbb{R}^{n\times 4}$ have orthonormal columns. If $\lambda(t)$ is a simple eigenvalue of $A + \varepsilon E(t)$, then* $\operatorname{Re}\dot{\lambda}(t) \geq 0$.

Let $\widetilde{U} \in \mathbb{R}^{4\times 4}$ be an orthogonal matrix such that its last two columns are those of U, and let $\widetilde{T} = \widetilde{U}^TT\widetilde{U}$. The following characterization of the stationary points is then obtained by direct inspection of (2.4).

Theorem 3.4 *If $\lambda \notin \mathbb{R} \cup i\mathbb{R}$, then (U, V) is a stationary point of* (3.2) *if and only if Y and V have the same range and $\widetilde{T}$ is block-diagonal with 2×2 blocks.*

Note that the transformation $\widetilde{V} = V\widetilde{U}$ yields $JE = \widetilde{V}\operatorname{diag}(1, 1, -1, -1)\widetilde{V}^T$, and therefore $\widetilde{T}$ corresponds to T of Theorem 2.2. Since $T = RPR^T$, we know that $\widetilde{T}$ always has two positive and two negative eigenvalues.

4 An algorithm to compute locally rightmost points in the Hamiltonian pseudospectrum and their extremizers

We discretize the differential equation (2.4) or (3.2) by the explicit Euler method with an adaptively chosen stepsize. At each discretization step, we require – according to the underlying property of the exact solutions – that the real part of the rightmost eigenvalue of $A + \varepsilon E$ is increased. In this way the method determines a sequence (λ_n, E_n) such that $\operatorname{Re}\lambda_n > \operatorname{Re}\lambda_{n-1}$, until E_n approaches a stationary point. We consider here the case where the matrix A and the perturbations are complex.

Given $E_n = J^{\mathrm{T}} V_n Q_n V_n^* \approx E(t_n)$ of rank 2 and unit 2-norm, where $V_n \in \mathbb{C}^{n\times 2}$ with $V_n^* V_n = I_2$, $u_n \in \mathbb{C}^2$ such that $\|u_n\| = 1$ and $Q_n = I_2 - 2u_n u_n^* \in \mathbb{C}^{2\times 2}$ orthogonal, Algorithm 1 determines an approximation at the time instant $t_{n+1} = t_n + h_n$. In this algorithm, $\gamma > 1$ denotes a given scaling factor for the stepsize control and $\operatorname{Orth}(B)$ denotes the unitary matrix obtained by orthogonalizing the columns of a given (rectangular) matrix B.

Algorithm 1: adaptive Euler step

Data: $E_n, V_n, Q_n, x_n, y_n, \lambda_n$ and ρ_n (step size predicted by previous step)

Result: $E_{n+1}, V_{n+1}, Q_{n+1}, x_{n+1}, y_{n+1}, \lambda_{n+1}$ and ρ_{n+1}

begin

1 Set $h = \rho_n$, $Y_n = (Jx_n, y_n)$ and $T_n = V_n^* Y_n P Y_n^* V_n \in \mathbb{C}^{2\times 2}$

2 Compute

$$\widehat{u}_{n+1} = u_n - \frac{1}{2} h Q_n T_n u_n, \quad u_{n+1} = \frac{\widehat{u}_{n+1}}{\|\widehat{u}_{n+1}\|}$$

$$\widehat{V}_{n+1} = V_n + h\,(I_n - V_n V_n^*)\, Y_n P Y_n^* V_n Q_n, \qquad V_{n+1} = \operatorname{Orth}\big(\widehat{V}_{n+1}\big)$$

$$Q_{n+1} = I_2 - 2u_{n+1}u_{n+1}^*.$$

3 Set $E_{n+1} = J^{\mathrm{T}} V_{n+1} Q_{n+1} V_{n+1}^*$

4 Compute the rightmost eigentriple $\widehat{\lambda}, \widehat{x}$ and $\widehat{y}$ of $A + \varepsilon E_{n+1}$

5 **if** $\operatorname{Re}(\widehat{\lambda}) \le \operatorname{Re}(\lambda_n)$ **then**
  reject the step, reduce the step size to $h := h/\gamma$ and repeat from 3
else
  accept the step: set $h_{n+1} = h$, $\lambda_{n+1} = \widehat{\lambda}$, $x_{n+1} = \widehat{x}$ and $y_{n+1} = \widehat{y}$

6 **if** $h_{n+1} = \rho_n$ **then**
  increase the stepsize to $\rho_{n+1} := \gamma\rho_n$
else
  set $\rho_{n+1} = \rho_n$

7 Proceed to next step.

Initial conditions. We make use of the following initial condition:

$$\widetilde{E}_0 = J^{\mathrm{T}}(Jx_0y_0^*)_{\mathrm{herm}}, \qquad V_0S_0V_0^* \;=\; J\widetilde{E}_0,$$

where S_0 is diagonal. Then we choose, according to the sign of the (diagonal) entries of S_0,

$$Q_0 = \begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix}.$$

For conciseness we do not formulate the algorithm for the real Hamiltonian case, which has a similar structure to the one we have described in this section.

We remark that the most expensive step of the algorithm is the computation of a leading eigentriple, which – for a large sparse matrix – can be done efficiently by using the Arnoldi method with implicit restarting, as implemented in ARPACK [LeSY98].

5 An algorithm for Problem (A)

In this section, we use Algorithm 1 for addressing a generalization of Problem (A) from the introduction:

Given a Hamiltonian matrix A with no eigenvalues on the imaginary axis and a given constant $\delta > 0$, find a nearest (in the 2-norm) Hamiltonian matrix B having some eigenvalue δ-close to the imaginary axis, that is, such that

$$\min |\mathrm{Re}(\lambda)| \le \delta, \quad \text{where } \lambda \text{ is an eigenvalue of } B.$$

The algorithm has the goal to compute

$$\varepsilon^* = \inf\{\varepsilon \ge 0 : \Lambda_\varepsilon(A, \mathcal{M}, \|\cdot\|) \cap S_\delta \neq \emptyset\}, \tag{5.1}$$

where $\mathcal{M}$ is the space of Hamiltonian matrices and $S_\delta = \{z \in \mathbb{C} :\ |\mathrm{Re}(z)| \le \delta\}$.

5.1 The left pseudospectral abscissa

In order to compute the value of ε^* defined in (5.1), we make use of the following definition. The left ε-*pseudospectral abscissa* of a Hamiltonian matrix A is the real part of the right-most point of the Hamiltonian ε-pseudospectrum in the left complex plane $\mathbb{C}^- = \{z : \mathrm{Re}(z) \le 0\}$, i.e.,

$$\alpha_\varepsilon^\ell(A) = \max\{\mathrm{Re}\, z \ : z \in \Lambda_\varepsilon(A, \mathcal{M}, \|\cdot\|) \cap \mathbb{C}^-\}. \tag{5.2}$$

Starting from $\varepsilon > 0$ such that $\alpha_\varepsilon^\ell(A) < 0$, we want to compute a root ε^* of the equation

$$\alpha_\varepsilon^\ell(A) = -\delta.$$

We make the following generic assumption for all ε near ε^*.

Assumption 5.1 *Let $\lambda(\varepsilon)$ with $\operatorname{Re}\lambda(\varepsilon) \neq 0$ be a locally rightmost point in the Hamiltonian ε-pseudospectrum of A. Then $\lambda(\varepsilon)$ is a* simple *eigenvalue of the corresponding perturbed matrix $A + \varepsilon E(\varepsilon)$ (with an extremizer $E(\varepsilon)$, of unit norm and Hamiltonian).*

Under Assumption 5.1, the corresponding locally rightmost point $\lambda(\varepsilon)$ of the pseudospectrum is a smooth function of ε and the same holds for suitably normalized eigenvectors $x(\varepsilon)$ and $y(\varepsilon)$. In order to derive an equation for ε to approximate ε^*, we can compute the derivative of the left ε-pseudospectral abscissa,

$$\alpha_\varepsilon^\ell(A) = \operatorname{Re}\lambda(\varepsilon),$$

with respect to ε.

Theorem 5.1 *Let $\lambda(\varepsilon)$, $\varepsilon \in [\underline{\varepsilon}, \overline{\varepsilon})$, be a branch of locally rightmost points of the complex or real Hamiltonian ε-pseudospectrum such that $\operatorname{Re}\lambda(\varepsilon) \neq 0$ for all ε, and Assumption 5.1 holds. Let $x(\varepsilon)$ and $y(\varepsilon)$ be left and right eigenvectors of $A+\varepsilon E(\varepsilon)$ to the eigenvalue $\lambda(\varepsilon)$, with Hamiltonian $E(\varepsilon)$ of unit norm. We then have*

$$\frac{d\operatorname{Re}\lambda(\varepsilon)}{d\varepsilon} = \operatorname{Re}\frac{x(\varepsilon)^* E(\varepsilon) y(\varepsilon)}{x(\varepsilon)^* y(\varepsilon)} \geq 0. \tag{5.3}$$

Proof By Lemma 1.1 we obtain, indicating by $'$ differentiation with respect to ε,

$$\lambda'(\varepsilon) = \frac{x(\varepsilon)^* \Big(E(\varepsilon) + \varepsilon E'(\varepsilon)\Big) y(\varepsilon)}{x^*(\varepsilon) y(\varepsilon)}.$$

In order to prove the theorem we have to show that

$$\operatorname{Re}\left(x(\varepsilon)^* E'(\varepsilon) y(\varepsilon)\right) = 0. \tag{5.4}$$

The maximality property of the real part of the eigenvalue $\lambda(\varepsilon)$ of $A+\varepsilon E(\varepsilon)$ yields that

$$\operatorname{Re}\left(x(\varepsilon)^* E'(\varepsilon) y(\varepsilon)\right) \leq 0.$$

Now suppose that for some $\varepsilon_0 \in (\underline{\varepsilon}, \overline{\varepsilon})$, this inequality would actually be a strict inequality. Consider $\widetilde{E}(\varepsilon)$ of unit norm such that $\widetilde{E}(\varepsilon_0) = E(\varepsilon_0)$ and $\widetilde{E}'(\varepsilon_0) = -E'(\varepsilon_0)$. Then, for all ε sufficiently close to ε_0, we would have that the corresponding eigenvalue $\widetilde{\lambda}(\varepsilon)$ of $A + \varepsilon\widetilde{E}(\varepsilon)$ satisfies $\operatorname{Re}\widetilde{\lambda}(\varepsilon) > \operatorname{Re}\lambda(\varepsilon)$. This, however, contradicts the optimality of $A + \varepsilon E(\varepsilon)$ and hence (5.4) must hold.

The non-negativity of $\dfrac{d\operatorname{Re}\lambda(\varepsilon)}{d\varepsilon}$ is due to the monotonicity of pseudospectral sets. □

5.2 The algorithm

Let the distance $\delta > 0$ be given, with δ small, and let ε^M denote the limit of ε^* (see (5.1)) as $\delta \to 0^+$. Close to the imaginary axis we make use of the second order expansion

$$\varepsilon \approx \varepsilon^M - \gamma \alpha_\varepsilon^\ell(A)^2 \tag{5.5}$$

with $\gamma > 0$. Given ε_k, we use Theorem 5.1 and estimate γ and ε^M by γ_k and ε_k^M using the formulæ (the first of which is obtained by differentiating (5.5) with respect to ε),

$$\gamma_k = \frac{x^*(\varepsilon_k) y(\varepsilon_k)}{2|\alpha_{\varepsilon_k}^\ell(A)| \operatorname{Re}\left(x^*(\varepsilon_k)\left(E(\varepsilon_k)\right) y(\varepsilon_k)\right)}$$

$$\varepsilon_k^M = \varepsilon_k + \gamma_k (\alpha_{\varepsilon_k}^\ell)^2$$

and then compute

$$\varepsilon_{k+1} = \varepsilon_k^M - \gamma_k \delta^2.$$

This results in Algorithm 2, where tol is a tolerance controlling the desired accuracy of the computed optimal ε. (Note that tol cannot be too small due to the fact that (5.5) is not exact.)

Algorithm 2: Basic algorithm for solving problem (A)

Data: δ, tol and ε_0 (such $\alpha_{\varepsilon_0}^\ell(A) < 0$)
Result: ε_f
begin
1 Set Reject = False and $k = 0$
2 **while** $|\alpha_{\varepsilon_k}^\ell(A) - \delta| \geq$ tol **do**
3 **if** Reject = False **then**
Store ε_k and $\alpha_{\varepsilon_k}^\ell$ into the memory
4 Compute γ_k and ε_k^M
5 Set $\varepsilon_{k+1} = \varepsilon_k^M - \gamma_k \delta^2$
6 Set $k = k + 1$
else
Set $\varepsilon_k = \dfrac{\varepsilon_k + \varepsilon_{k-1}}{2}$
7 Compute $\alpha_{\varepsilon_k}^\ell$ by integrating (2.4) (or (3.2)) with initial datum $E(\varepsilon_{k-1})$ (for $k \geq 1$).
8 **if** $\alpha_{\varepsilon_k}^\ell(A) \approx 0$ *(with tolerance tol)* **then**
Set Reject = True
else
Set Reject = False
9 Set $\varepsilon_f = \varepsilon_k$

In order to further refine the value obtained from Algorithm 2, we can apply, starting from the final value ε_f, the Newton iteration:

$$\varepsilon_{n+1} = \varepsilon_n - \left(\alpha^{\ell}_{\varepsilon_n}(A) + \delta\right) \frac{x(\varepsilon_n)^* y(\varepsilon_n)}{\mathrm{Re}\left(x(\varepsilon_n)^* E(\varepsilon_n) y(\varepsilon_n)\right)}, \qquad n = 0, 1, \ldots$$

which yields quadratic convergence.

Since values $\varepsilon > \varepsilon^M$ lead to rejected steps, it is necessary to choose a starting value $\varepsilon_0 < \varepsilon^M$ and possibly underestimate the initial predicted values for ε_k. A good choice for ε_0 is the unstructured distance from the imaginary axis, that is the norm of the smallest unstructured matrix F such that $A + F$ has a purely imaginary eigenvalue. Such a distance can be effectively computed by combining the method proposed in [GO11, GL11] with a Newton iteration.

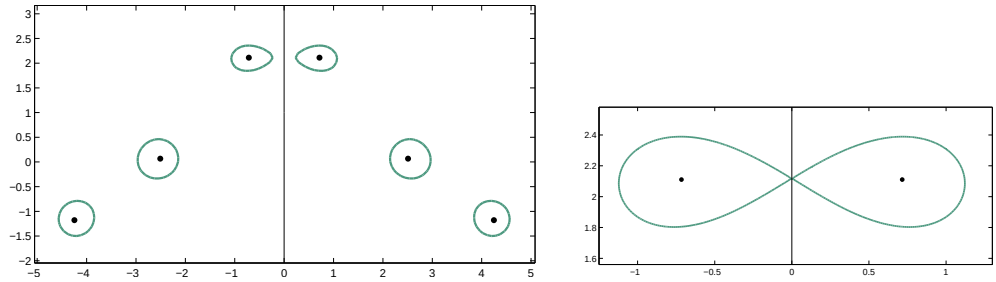


Fig. 5.1 Hamiltonian 2-norm pseudospectra of the matrix A in Example 5.1 for $\varepsilon = 10^{-0.9}$ (left picture) and $\varepsilon = \varepsilon_4$ in Table 5.1 (right picture).

Example 5.1 Consider the Hamiltonian matrix

$$A = \begin{pmatrix} 4 & 1 & 3-\mathbf{i} & -2 & -2+2\mathbf{i} & 1-\mathbf{i} \\ \mathbf{i} & -1-\mathbf{i} & 1+2\mathbf{i} & -2-2\mathbf{i} & 0 & 1+\mathbf{i} \\ -2+\mathbf{i} & 1-\mathbf{i} & -1+2\mathbf{i} & 1+\mathbf{i} & 1-\mathbf{i} & 0 \\ -4 & -3 & -1+\mathbf{i} & -4 & \mathbf{i} & 2+\mathbf{i} \\ -3 & -2 & -1-2\mathbf{i} & -1 & 1-\mathbf{i} & -1-\mathbf{i} \\ -1-\mathbf{i} & -1+2\mathbf{i} & 0 & -3-\mathbf{i} & -1+2\mathbf{i} & 1+2\mathbf{i} \end{pmatrix}$$

We set $\delta = 10^{-4}$ and $\varepsilon_0 = 0.1$. The following Tables 5.1 and 5.2 report the results for the 2-norm. Two steps have been rejected by detecting the presence of purely imaginary eigenvalues. The observed convergence is fast.

A refinement of the value of the approximation of ε^* can be obtained by applying a Newton iteration.

Example 5.2 We consider a random 30×30 Hamiltonian matrix with $\delta = 10^{-3}$ and $\mathrm{tol} = 10^{-6}$. Table 5.3 illustrates the observed behavior (there appears only one rejection).

Also in this case we apply a Newton iteration starting from ε_3.

k	ε_k	$\alpha^{\ell}_{\varepsilon_k}$
0	0.1	−0.37063002644925
1	0.**1**4178200003380	−0.08483455322963
2	0.**1442230**2327426	−0.00482210859774
3	0.**1442309578**447	−0.00010118673406
4	0.**14423095786597**	−0.00009999646472

Table 5.1 Computed values of ε and $\alpha^{\ell}_{\varepsilon}$ for Example 5.1.

k	ε_k	$\alpha^{\ell}_{\varepsilon_k}$
3	0.**1442309578**447	−0.00010118673406
4	0.**14423095786**600	−0.00009999253623
5	0.**14423095786549**	−0.00009999990084

Table 5.2 Computed values of ε and $\alpha^{\ell}_{\varepsilon}$ for Example 5.1 by a Newton iteration starting from $\varepsilon = \varepsilon_3$ in Table 5.1.

k	ε_k	$\alpha^{\ell}_{\varepsilon_k}$
0	0.2	−1.03039154166432
1	0.**220**59314066343	−0.05527116512772
2	0.**22065271**799409	−0.00101261933010
3	0.**22065271898506**	−0.00099987217016

Table 5.3 Computed values of ε and $\alpha^{\ell}_{\varepsilon}$ for Example 5.2.

k	ε_k	$\alpha^{\ell}_{\varepsilon_k}$
3	0.**22065271**898506	−0.00099987217016
4	0.**2206527192**0303	−0.00099999710718
5	0.**22065271920521**	−0.00099999996525

Table 5.4 Computed values of ε and $\alpha^{\ell}_{\varepsilon}$ for Example 5.2 by Newton refinement.

6 Outer algorithms for Problems (B) and (C)

In this section, we use Algorithm 1 for addressing a generalization of Problem (B) from the introduction: Given a Hamiltonian matrix A with all eigenvalues on the imaginary axis and a given constant $\delta > 0$, find a nearest Hamiltonian matrix B having some eigenvalue δ-far from the imaginary axis, that is such that

$$\max |\mathrm{Re}(\lambda)| = \delta, \quad \text{where } \lambda \text{ is an eigenvalue of } B.$$

The infimum value $\varepsilon = \|B - A\|$ such that the above holds for some $\delta > 0$ is denoted by ε^m.

6.1 The algorithm for problem (B)

The algorithm for solving Algorithm (B) proceeds similarly as the one presented in the previous section for problem (A). In particular, the Newton refinement strategy works exactly in the same way.

Let the distance $\delta > 0$ be given and assume δ to be small. (Otherwise, we make directly use of the Newton iteration).

We have to solve the equation

$$\alpha_\varepsilon(A) = \delta,$$

where $\alpha_\varepsilon(A)$ denotes the pseudospectral abscissa, that is, the real part of the rightmost point in the Hamiltonian ε-pseudospectrum.

Close to ε^m we make use of the second order expansion, with $\gamma > 0$,

$$\varepsilon \approx \varepsilon^m + \gamma \alpha_\varepsilon(A)^2. \tag{6.1}$$

Given ε_k, we estimate γ_k and ε_k^m by the following formulæ:

$$\gamma_k = \frac{x^*(\varepsilon_k) y(\varepsilon_k)}{2\alpha_{\varepsilon_k}(A)\,\mathrm{Re}\,(x^*(\varepsilon_k)\,(E(\varepsilon_k))\,y(\varepsilon_k))}$$

$$\varepsilon_k^m = \varepsilon_k - \gamma_k(\alpha_{\varepsilon_k})^2$$

and then compute

$$\varepsilon_{k+1} = \varepsilon_k^m + \gamma_k \delta^2.$$

This results in Algorithm 3, where ε^u denotes the corresponding distance for the unstructured pseudospectrum.

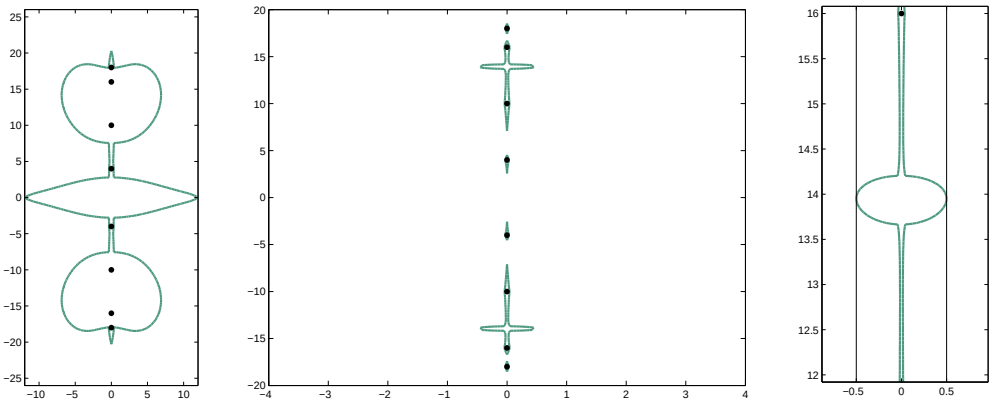


Fig. 6.1 2-norm Hamiltonian pseudospectra of the matrix A in Example 6.1 for $\delta = 0.5$ and $\varepsilon_0 = 0.5$ (left picture) and $\varepsilon = \varepsilon_5$ in Table 6.1 (middle picture). Zoom close to the rightmost section (right picture).

Algorithm 3: Basic algorithm for solving problem (B)

Data: δ, tol and ε_0 (such that $\alpha_{\varepsilon_0}(A) = 0$) and $r \in [1/2, 1)$
Result: ε_f
begin
1 Set Reject = False and $k = 0$
2 **while** $|\alpha_{\varepsilon_k}(A) - \delta| \geq$ tol **do**
3 **if** Reject = False **then**
Store ε_k and α_{ε_k} into the memory
4 Compute γ_k and ε_k^m
5 Set $\varepsilon_{k+1} = \varepsilon_k^m + \gamma_k \delta^2$
6 **if** $\varepsilon_{k+1} \leq \varepsilon^u$ **then**
Set $\varepsilon_{k+1} = r\varepsilon_k$
else
Proceed
7 Set $k = k + 1$
else
Set $\varepsilon_k = \dfrac{\varepsilon_k + \varepsilon_{k-1}}{2}$
8 Compute α_{ε_k} by integrating (2.4) (or (3.2)) with initial datum $E(\varepsilon_{k-1})$ (for $k \geq 1$).
9 **if** $\alpha_{\varepsilon_k}(A) \approx 0$ *(with tolerance tol)* **then**
Set Reject = True
else
Set Reject = False
10 Set $\varepsilon_f = \varepsilon_k$

Example 6.1 Consider the Hamiltonian matrix from [ABKMM11]:

$$A = \begin{pmatrix} -73 & -86 & 54 & -99 & 93 & -58 & 80 & 77 \\ 1 & -4 & 59 & 54 & -58 & -61 & 4 & 1 \\ -24 & -31 & -4 & -86 & 80 & 4 & 27 & 26 \\ -26 & -24 & 1 & -73 & 77 & 1 & 26 & 24 \\ -24 & -26 & -1 & -77 & 73 & -1 & 24 & 26 \\ -26 & -27 & -4 & -80 & 86 & 4 & 31 & 24 \\ -1 & -4 & 61 & 58 & -54 & -59 & 4 & -1 \\ -77 & -80 & 58 & -93 & 99 & -54 & 86 & 73 \end{pmatrix}$$

Setting tol $= 10^{-6}$, we have obtained for $\delta = 0.5$ and $\delta = 0.01$ the results reported in Table 6.1 and 6.2 (where we have set $r = 0.5$ and have counted two rejections).

In Figure 6.1 we show the initial and final pseudospectra by the proposed Algorithm when $\delta = 0.5$. In the right picture we show a section of the final pseudospectrum, which lies - as expected - in the strip $\{z : \ |\mathrm{Re}(z)| \leq \delta\}$.

Remark 6.1 It is important, at the initial step $\varepsilon = \varepsilon_0$, to find an initial perturbation E such that $A + \varepsilon E$ does not have all eigenvalues on the imaginary axis, otherwise we cannot produce a trajectory to a locally righmost point of the Hamiltonian pseudospectrum by integrating (2.4). In subsequent steps, the optimal matrix E from the previous step usually provides a good choice for the initial perturbation.

k	ε_k	α_{ε_k}
0	0.5	6.87824338551888
1	0.25	3.91845140594614
2	0.125	1.41478802483750
3	0.**10**722950104024	0.52551651210602
4	0.**10695**570418704	0.50001145712079
5	0.**10695558424707**	0.50000313432414

Table 6.1 Computed values of ε and α_ε for Example 6.1 with $\delta = 0.5$.

k	ε_k	α_{ε_k}
0	0.5	6.87824338551888
1	0.25	3.91845140594614
2	0.125	1.41478802483750
3	0.**104**69321189086	0.18533034160397
4	0.**10433**09049728	0.00345257718485
5	0.**10433196518822**	0.00010001932585

Table 6.2 Computed values of ε and α_ε for Example 6.1 with $\delta = 0.0001$.

6.2 A comparison for the real Hamiltonian case

In [ABKMM11, page 31] the authors report the results for Example 6.1 using real Hamiltonian perturbations. After setting $\delta = 10^{-1}$ they obtain by their algorithm a perturbation with norm $0.16768\ldots$ (and full rank).

Applying our algorithm we compute an optimal value

$$\varepsilon = 0.104437573679967 \quad \text{for which} \quad \alpha_\varepsilon(A) = 0.099999720172499$$

and obtain the optimal perturbation

$$\varepsilon E = 10^{-2} \times \begin{pmatrix} 2.4533 & 2.4889 & 0.1535 & -0.0234 & -1.2265 & -1.4012 & 1.3594 & -1.1337 \\ 3.1866 & 3.0415 & 0.3055 & 0.0969 & -1.4012 & -1.5915 & 1.0038 & -1.7817 \\ -0.2804 & -7.1985 & 2.9955 & 2.4656 & 1.3594 & 1.0038 & 0.8808 & 3.6684 \\ 6.9138 & -0.2359 & 3.3212 & 2.4674 & -1.1337 & -1.7817 & 3.6684 & 1.9671 \\ -2.0387 & -3.6151 & 1.8683 & 1.1127 & -2.4533 & -3.1866 & 0.2804 & -6.9138 \\ -3.6151 & -0.8938 & -0.9839 & -1.3348 & -2.4889 & -3.0415 & 7.1985 & 0.2359 \\ 1.8683 & -0.9839 & 1.7394 & 1.4531 & -0.1535 & -0.3055 & -2.9955 & -3.3212 \\ 1.1127 & -1.3348 & 1.4531 & 1.2232 & 0.0234 & -0.0969 & -2.4656 & -2.4674 \end{pmatrix},$$

where only the leading 5 digits are shown.

The eigenvalues of $A + \varepsilon E$ are the following (to a 6 digit accuracy):

$$\pm 0.1000 - 13.9365\mathbf{i}, \qquad \pm 0.1000 + 13.9365\mathbf{i}, \qquad \pm 4.3068\mathbf{i}, \qquad \pm 17.6914\mathbf{i},$$

and hence 4 eigenvalues lie on the boundary of the strip S_δ.

In some runs, it turns out that we follow a branch of locally (but not globally) extremal points of the real Hamiltonian pseudospectra and determine a final value

$$\varepsilon = 0.14182248093698 \quad \text{for which} \quad \operatorname{Re}\lambda(\varepsilon) = 0.10000024415943.$$

Here the eigenvalues of the corresponding perturbed matrix $A+\varepsilon E$ are the following (to 6 digits accuracy):

$$\pm 0.1000, \qquad \pm 6.6450\mathbf{i}, \qquad \pm 16.2781\mathbf{i}, \qquad \pm 18.4005\mathbf{i},$$

so that only 2 (real) eigenvalues lie on the boundary of the strip.

6.3 An algorithm for problem (C)

A natural way to deal with problem (C) from the introduction is a repeated use of a method for solving (B). In fact, after the computation of the first Hamiltonian perturbation, which moves a pair of imaginary eigenvalues on the boundary of the considered thin strip, one can apply a Hamiltonian Schur decomposition to the perturbed matrix and obtain an Hamiltonian block structure on which it is possible to apply again the method on a Hamiltonian matrix of smaller dimension, as done in [ABKMM11].

Let us consider again Example 6.1. Computing the Hamiltonian Schur form of the matrix $A + \varepsilon E$, by an orthogonal symplectic matrix Q, we obtain the following Hamiltonian matrix,

$$Q(A+\varepsilon E)Q^{\mathrm{T}} = \begin{pmatrix} F_{11} & F_{12} & G_{11} & G_{12} \\ \mathrm{O} & F_{22} & G_{21} & G_{22} \\ \mathrm{O} & \mathrm{O} & -F_{11}^{\mathrm{T}} & \mathrm{O} \\ \mathrm{O} & H_{22} & -F_{12}^{\mathrm{T}} & -F_{22}^{\mathrm{T}} \end{pmatrix} \tag{6.2}$$

where the block

$$\hat{A} = \begin{pmatrix} F_{22} & G_{22} \\ H_{22} & -F_{22}^{\mathrm{T}} \end{pmatrix}, \quad \text{with} \quad F_{22} = \begin{pmatrix} -1.3952 & -6.0587 \\ -5.6920 & -7.6024 \end{pmatrix},$$

and

$$G_{22} = \begin{pmatrix} 0.3318 & -4.2732 \\ -4.2732 & -33.0414 \end{pmatrix}, \quad H_{22} = \begin{pmatrix} -3.4288 & 3.8850 \\ 3.8850 & 12.8889 \end{pmatrix},$$

has the same purely imaginary eigenvalues as $A + \varepsilon E$, namely $\pm 4.3068\mathbf{i}$, and $\pm 17.6914\mathbf{i}$.

Applying Algorithm 3 to the matrix $\hat{A}$ we are able to compute the distance $\hat{\varepsilon} = 2.777017377181622$ and the following optimal perturbation

$$\hat{\varepsilon}\hat{E} = \begin{pmatrix} -0.4573 & -2.4489 & -1.0526 & 0.6305 \\ -2.4489 & -0.1079 & 0.6305 & -1.1426 \\ -1.0526 & 0.6305 & 0.4573 & 2.4489 \\ 0.6305 & -1.1426 & 2.4489 & 0.1079 \end{pmatrix}.$$

Combining the steps in this and the previous subsection, this implies that the 2-norm of an optimal perturbation that takes all eigenvalues of A outside the strip $|\mathrm{Re}(z)| \leq 0.1$, is at most 2.8815.

7 Applications

As mentioned in the introduction, Hamiltonian matrices play an important role in applications from control theory and gyroscopic systems. We will now illustrate how the algorithms developed in this paper can be used in such applications.

7.1 Linear control systems

Consider a control system

$$\begin{cases} \dot{x} = A\,x + B\,u, \\ y = C\,x + D\,u \end{cases} \tag{7.1}$$

with zero initial condition $x(0) = 0$ and matrices A, B, C and D either complex or real. Here x is the state, u is the input and y is the output. In the very common case where A is an Hurwitz matrix and D is a square nonsingular matrix, passivity with respect to the supply rate $s(u(t), y(t)) = y^*(t)u(t) + u^*(t)y(t)$, is equivalent to the following dissipation inequality (for $t_1 > t_0$):

$$\frac{d}{dt}\theta\left(x(t)\right) \leq s\left(u(t), y(t)\right), \tag{7.2}$$

where θ, the so-called storage function, is a suitable nonnegative function which generalizes the concept of Lyapunov function to an open system. This means that for a passive system the change in internal storage $\theta\left(x(t_1)\right) - \theta\left(x(t_0)\right)$ can never exceed what is supplied to the system.

A well-known result (see [Ant05]) states that passivity is equivalent to the property that the associated Hamiltonian matrix, with $R = D^*D - I$ and $S = DD^* - I$,

$$M = \begin{pmatrix} F & G \\ H & -F^* \end{pmatrix} \tag{7.3}$$
$$\text{with} \quad F = A - BR^{-1}D^*C,\ G = -BR^{-1}B^*,\ H = C^*S^{-1}C,$$

has no purely imaginary eigenvalues.

7.1.1 Distance to bounded-realness

Let us now suppose that (7.1) is not passive. In order to make it passive, it is possible to perturb it in a way that the corresponding perturbed Hamiltonian matrix has all eigenvalues bounded away from the imaginary axis. The minimal norm of such a perturbation is called *distance to bounded-realness*. This quantity is computed by solving Problem (C) and consequently by repeatedly solving Problem (B). On the other hand, the robustness of a passive system can be computed by solving problem (A), that is, computing the distance of the underlying Hamiltonian matrix from the set of Hamiltonian matrices having at least a pair of imaginary eigenvalues.

As mentioned in [ABKMM11], passivation is very important for those systems arising from passive physical models which lose the passivity property after being finitely approximated or discretized, so that it is important to recover it in the finite dimensional model. The same problem occurs after applying most model reduction procedures, so that the reduced model may lose the passivity property of the non-reduced problem.

Recently in [MeXu08] the authors have developed a matrix perturbation theory for Hamiltonian matrices and in [ABKMM11] a method is proposed to compute the distance to bounded-realness. Computing minimal norm perturbations is a difficult nonconvex optimization problem. In [ABKMM11] suboptimal perturbations are constructed, allowing to get upper bounds for the minimal norm perturbations.

Example 7.1 We consider the following linear control system from [BBK89]:

$$A = \begin{pmatrix} -0.08 & 0.83 & 0 & 0 \\ -0.83 & -0.08 & 0 & 0 \\ 0 & 0 & -0.7 & 9 \\ 0 & 0 & -9 & -0.7 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 1 & -1 \\ 0 & 0 \end{pmatrix}, \quad C^{\mathrm{T}} = \begin{pmatrix} 0.4 & 0.6 \\ 0 & 0 \\ 0.4 & 1 \\ 0 & 0 \end{pmatrix}$$

and $D = \operatorname{diag}(0.3, -0.15)$. The resulting Hamiltonian matrix M of (7.3) reads (we represent it by using 5 digits)

$$\begin{pmatrix} -0.0402 & 0.8300 & -0.0216 & 0 & 2.1219 & 0 & 0.0759 & 0 \\ -0.8300 & -0.0800 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.2239 & 0 & -0.4147 & 9.0000 & 0.0759 & 0 & 2.1219 & 0 \\ 0 & 0 & -9.0000 & -0.7000 & 0 & 0 & 0 & 0 \\ -0.5441 & 0 & -0.7896 & 0 & 0.0402 & 0.8300 & -0.2239 & 0 \\ 0 & 0 & 0 & 0 & -0.8300 & 0.0800 & 0 & 0 \\ -0.7896 & 0 & -1.1988 & 0 & 0.0216 & 0 & 0.4147 & 9.0000 \\ 0 & 0 & 0 & 0 & 0 & 0 & -9.0000 & 0.7000 \end{pmatrix}$$

which has all purely imaginary eigenvalues:

$$\pm 9.6211\mathbf{i}, \qquad \pm 8.4523\mathbf{i}, \qquad \pm 1.5256\mathbf{i}, \qquad \pm 0.4499\mathbf{i}$$

If we apply – as a first step of passivation – Algorithm 3 with $\delta = 0.3$ we obtain the perturbation $\Delta =$

$$\frac{1}{10}\begin{pmatrix} 0.0009 & -0.0003 & 0.0209 & -0.1642 & -0.0197 & 0.0002 & -0.0478 & -0.2058 \\ 0.0004 & 0.0000 & 0.0153 & 0.0030 & 0.0002 & -0.0001 & 0.0192 & -0.0065 \\ -0.0459 & -0.0013 & -2.6639 & -0.0986 & -0.0478 & 0.0192 & 1.0948 & -0.2580 \\ 0.0009 & -0.0039 & -0.1114 & -2.6521 & -0.2058 & -0.0065 & -0.2580 & 1.1025 \\ 0.0008 & -0.0000 & 0.0324 & -0.0112 & -0.0009 & -0.0004 & 0.0459 & -0.0009 \\ -0.0000 & 0.0000 & 0.0001 & 0.0023 & 0.0003 & -0.0000 & 0.0013 & 0.0039 \\ 0.0324 & 0.0001 & 1.0931 & -0.2579 & -0.0209 & -0.0153 & 2.6639 & 0.1114 \\ -0.0112 & 0.0023 & -0.2579 & 1.0835 & 0.1642 & -0.0030 & 0.0986 & 2.6521 \end{pmatrix}$$

of spectral norm $\varepsilon = 0.2894$. The perturbed Hamiltonian matrix $M + \Delta$ has the eigenvalues

$$\pm 0.3000 \pm 9.0370\mathbf{i}, \qquad \pm 1.5266\mathbf{i}, \qquad \pm 0.4497\mathbf{i},$$

so that 4 eigenvalues have been pushed away from a strip of width $\delta = 0.3$.

The second (and final) step of Algorithm 3 is obtained in an analogous way to Example 6.1. Computing the Hamiltonian Schur form of the matrix $M + \Delta$, by an orthogonal symplectic matrix Q we obtain an Hamiltonian matrix in the form (6.2) with the block

$$\hat{M} = \begin{pmatrix} F_{22} & G_{22} \\ H_{22} & -F_{22}^{\mathrm{T}} \end{pmatrix}, \quad \text{where} \quad F_{22} = \begin{pmatrix} 0.0 & 0.3201 \\ -1.1328 & 0.0993 \end{pmatrix},$$

and

$$G_{22} = \begin{pmatrix} 1.0826 & -0.0619 \\ -0.0619 & 0.2931 \end{pmatrix}, \quad H_{22} = \begin{pmatrix} -1.8178 & 0.0 \\ 0.0 & 0.5137 \end{pmatrix}.$$

Algorithm 3 applied to $\hat{M}$ provides the optimal perturbation

$$\hat{\Delta} = \begin{pmatrix} 0.0025 & -0.3267 & 0.1380 & -0.0040 \\ -0.3267 & -0.0059 & -0.0040 & 0.1379 \\ 0.1380 & -0.0040 & -0.0025 & 0.3267 \\ -0.0040 & 0.1379 & 0.3267 & 0.0059 \end{pmatrix}$$

of norm $\hat{\varepsilon} = 0.3546$. The imaginary eigenvalues of $\hat{M}$ (the same of $M + \Delta$) are moved by the perturbation above to $\pm 0.3000 \pm 0.9801\mathbf{i}$.

Finally, the 2-norm of an *optimal* Hamiltonian perturbation moving the eigenvalues of M outside the strip $|\mathrm{Re}(z)| \leq 0.3$ is *at most* $\varepsilon + \hat{\varepsilon} = 0.6440$.

We remark that it is in general not possible to express the final perturbation as a matrix of the form (7.3). Dealing with this additional structure is beyond the scope of this paper. We intend to treat this problem in a future work, building on ideas and techniques of the present paper.

7.1.2 Passivity radius

Similarly, for a passive system, it is interesting to compute the so-called passivity radius (see e.g. [OvVD05]), that is, its distance to the closest non passive system. A surrogate for an exact computation of the smallest structured perturbation

$$\begin{pmatrix} A_\Delta & B_\Delta \\ C_\Delta & D_\Delta \end{pmatrix}$$

which destroys passivity can be obtained by solving problem (A) for the associated Hamiltonian matrix (7.3), which gives an indication of the distance to passivity.

Example 7.2 We consider the following passive linear control system:

$$A = \begin{pmatrix} -8 & -4 & -1.5 \\ 4 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \quad C^{\mathrm{T}} = \begin{pmatrix} 1 \\ 1 \\ 0.75 \end{pmatrix}, \quad D = (-0.75)\,.$$

Indeed the eigenvalues of the associated Hamiltonian matrix (7.3),

$$M = \begin{pmatrix} -11.4286 & -7.4286 & -4.0714 & 9.1429 & 0 & 0 \\ 4.0000 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.0000 & 0 & 0 & 0 & 0 \\ -2.2857 & -2.2857 & -1.7143 & 11.4286 & -4.0000 & 0 \\ -2.2857 & -2.2857 & -1.7143 & 7.4286 & 0 & -1.0000 \\ -1.7143 & -1.7143 & -1.2857 & 4.0714 & 0 & 0 \end{pmatrix}$$

are

$$\pm 6.5856, \qquad \pm 2.5784, \qquad \pm 0.5173$$

The smallest Hamiltonian perturbation which moves a pair of eigenvalues of M δ-close to the imaginary axis (with $\delta = 10^{-4}$) is

$$\Delta = \frac{1}{100} \begin{pmatrix} -0.0500 & -1.1629 & 3.5652 & 0.5637 & 0.8216 & 1.0869 \\ -0.0797 & -2.6869 & 7.0393 & 0.8216 & 1.0293 & 0.9200 \\ -0.1232 & -6.1570 & 14.1464 & 1.0869 & 0.9200 & -0.5256 \\ -0.0042 & -0.0633 & 0.2424 & 0.0500 & 0.0797 & 0.1232 \\ -0.0633 & 3.4474 & -3.5051 & 1.1629 & 2.6869 & 6.1570 \\ 0.2424 & -3.5051 & -2.3758 & -3.5652 & -7.0393 & -14.1464 \end{pmatrix}$$

where $\|\Delta\| = 0.1826$ and the eigenvalues of $M + \Delta$ are

$$\pm 6.4616, \qquad \pm 2.9138, \qquad \pm 0.0001.$$

Again, Δ cannot be expressed in terms of perturbations of A, B, C and D and a direct analysis of such structured optimal perturbations will be pursued in future. In [OvVD05] the authors propose a method for computing the passivity radius. The method depends on computing the smallest structured indefinite perturbation to a Hermitian matrix that makes it singular, which is obtained by solving a unimodal optimization problem. We believe that a method based on similar techniques to those exploited in this paper might be a promising alternative.

7.2 Gyroscopic systems

We consider a gyroscopic system of ordinary differential equations taking the form

$$M\ddot{x}(t) + G\dot{x}(t) + Kx(t) = 0, \tag{7.4}$$

where M is Hermitian positive definite, K is Hermitian, and G is skew-Hermitian. The stability of (7.4) is determined by the eigenvalues of the quadratic eigenvalue problem

$$\big(\lambda^2 M + \lambda G + K\big)x = 0.$$

Provided that M is sufficiently well conditioned, we can compute a Cholesky decompositon $M = R^*R$ and consider instead the eigenvalue problem

$$\big(\lambda^2 I + \lambda\hat{G} + \hat{K}\big)x = 0, \tag{7.5}$$

with $\hat{G} = (R^*)^{-1}GR^{-1}$ and $\hat{K} = (R^*)^{-1}GR^{-1}$. The quadratic eigenvalue problem (7.5) admits the Hamiltonian linearization [MeWa00]

$$A = \begin{pmatrix} -\frac{1}{2}\hat{G} & -\frac{1}{4}\hat{G}^2 - \hat{K} \\ I & -\frac{1}{2}\hat{G} \end{pmatrix}.$$

Clearly, a gyroscopic system (7.4) can only be stable if all eigenvalues of A are purely imaginary. If the matrices M, G, K are subject to uncertainties or parameter dependence, it is of interest to provide a bound on the norm of the perturbations under which the system remains stable. A possibly rather conservative bound can be obtained from computing the minimal Hamiltonian perturbation that moves some eigenvalues of A off the imaginary axis:

$$\gamma(A) = \inf\big\{\|\Delta A\|_2 :\ \Delta A \text{ is Hamiltonian}, \Lambda(A + \Delta A) \not\subset \mathrm{i}\mathbb{R}\big\}. \tag{7.6}$$

This is problem (B), which can be solved by Algorithm 3.

Example 7.3 We consider Example 5 in [HKLP00] stemming from a Galerkin discretization of a partial differential equation that models deflation of an elastic pipe conveying fluid.

Using as parameters a mass equal to 4 ($M = 4I$) and a load parameter $l = 0.75$ we obtain a stable system, whose Hamiltonian matrix has eigenvalues

$$\pm 0.2947\mathbf{i}, \quad \pm 1.8343\mathbf{i}, \quad \pm 4.3391\mathbf{i}, \quad \pm 7.8429\mathbf{i}, \quad \pm 12.3440\mathbf{i}, \quad \pm 17.8826\mathbf{i}.$$

Applying Algorithm 3 we find that there exists a Hamiltonian perturbation εE of norm $\varepsilon = 0.0959 > \gamma(A)$ which moves a pair of eigenvalues out of the strip of radius $\delta = 0.1$. The corresponding eigenvalues of $A + \varepsilon E$ are

$$\pm 0.1000, \quad \pm 1.8395\mathbf{i}, \quad \pm 4.3395\mathbf{i}, \quad \pm 7.7982\mathbf{i}, \quad \pm 12.2359\mathbf{i}, \quad \pm 17.9382\mathbf{i},$$

which means that the two eigenvalues of A of smallest modulus coalesced and were moved off the strip.

The obtained distance ε identifies a sort of stability radius for the gyroscopic system in terms of the associated Hamiltonian matrix. The distance to instability in terms of perturbations to the original matrices M, G, K is an interesting topic to investigate, which is beyond the scope of this paper.

Acknowledgements

This work has been partially developed at the Mathematisches Forschungsinstitut Oberwolfach. We are grateful for the very kind hospitality.

We also wish to thank the Italian M.I.U.R. and the I.N.d.A.M.-G.N.C.S. and the German DFG Priority Research Programme 1324 for support.

References

[ABKMM11] R. Alam, S. Bora, M. Karow, V. Mehrmann, and J. Moro. Perturbation theory for Hamiltonian matrices and the distance to bounded-realness. *SIAM J. Matrix Anal. Appl.*, 32(2):484–514, 2011.

[And98] B. Anderson. Riccati equations, network theory and Brune synthesis: old solutions for contemporary problems, *Dynamical systems, control, coding, computer vision*, series Progr. Systems Control Theory, 251–25, 1998.

[Ant05] A. Antoulas. *Approximation of large-scale dynamical systems.* volume 6 of *Advances in Design and Control* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005.

[BBK89] S. Boyd, V. Balakrishnan, and P. Kabamba. A bisection method for computing the H_∞ norm of a transfer matrix and related problems. *Math. Control Signals Systems*, 2(3):207–219, 1989.

[BK04] R. Byers and D. Kressner. On the condition of a complex eigenvalue under real perturbations. *BIT*, 44(2):209–214, 2004.

[GL11] N. Guglielmi and C. Lubich. Differential equations for roaming pseudospectra: paths to extremal points and boundary tracking. *SIAM J. Numer. Anal.*, 49(3):1194–1209, 2011.

[GL11b] N. Guglielmi and C. Lubich. Erratum/addendum: Differential equations for roaming pseudospectra: paths to extremal points and boundary tracking. *SIAM J. Numer. Anal.*, 50(2):977–981, 2012.

[GL12] N. Guglielmi and C. Lubich. Low-rank dynamics for computing extremal points of real and complex pseudospectra. *SIAM J. Matrix Anal. Appl.*, in press, 2013.

[GO11] N. Guglielmi and M. Overton. 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.

[HLW06] E. Hairer, Ch. Lubich, and G. Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2006. Structure-preserving algorithms for ordinary differential equations.

[Hes09] J. Hespanha. *Linear systems theory.* Princeton University Press, Princeton, NJ, 2009.

[HP05] D. Hinrichsen and A. J. Pritchard. *Mathematical Systems Theory. I*, volume 48 of *Texts in Applied Mathematics*. Springer-Verlag, Berlin, 2005. Modelling, state space analysis, stability and robustness.

[HKLP00] R.O. Hryniv, W. Kliem, P. Lancaster and C. Pommer. A precise bound for gyroscopic stabilization. *ZAMM Z. Angew. Math. Mech.*, 80(8):507–516, 2000.

[KKK10] M. Karow, E. Kokiopoulou, and D. Kressner. On the computation of structured singular values and pseudospectra. *Systems Control Lett.*, 59(2):122–129, 2010.

[Ka11] M. Karow. Structured Pseudospectra for Small Perturbations. *SIAM J. Matrix Anal. Appl.*, 32(1):1383–1398, 2011.

[Kat95] T. Kato. Perturbation theory for linear operators. *Classics in Mathematics*. Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.

[KL07] O. Koch and Ch. Lubich. Dynamical low-rank approximation. *SIAM J. Matrix Anal. Appl.*, 29(2):434–454, 2007.

[LeSY98] R. B. Lehoucq, D. C. Sorensen, and C. Yang. *ARPACK users' guide*, volume 6 of *Software, Environments, and Tools*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1998. Solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods.

[MeWa00] V. Mehrmann, and D. Watkins. Structure-preserving methods for computing eigenpairs of large sparse skew-Hamiltonian/Hamiltonian pencils. *SIAM J. Sci. Comput.*, 22:1905–1925, 2000.

[MeXu08] V. Mehrmann, and H. Xu. Perturbation of purely imaginary eigenvalues of Hamiltonian matrices under structured perturbations. *Electron. J. Linear Algebra*, 17:234–257, 2008.

[OvVD05] M. L. Overton and P. Van Dooren. On computing the complex passivity radius. *Proc. 4th IEEE Conference on Decision and Control (ECC-CDC05, Seville, Spain, December 2005)*, 5805:7960–7964, 2005.

[PaVL81] C. Paige and C. Van Loan. A Schur decomposition for Hamiltonian matrices-*Linear Algebra Appl.*, 41:11–32, 1981.

[Ru06] S. M. Rump. Eigenvalues, pseudospectrum and structured perturbations *Linear Algebra Appl.*, 413(2-3):567–593, 2006.

[Tr99] L.N. Trefethen. Computation of Pseudospectra. *Acta Numerica*, 8:247–295, 1999.

[TE05] L. N. Trefethen and M. Embree. *Spectra and Pseudospectra*. Princeton University Press, Princeton, NJ, 2005. The behavior of nonnormal matrices and operators.

[WA05] T. Wagenknecht and J. Agarwal. Structured pseudospectra in structural engineering. *International Journal for Numerical Methods in Engineering*, 64(13):1735–1751, 2005.