REAL LINEAR MATRIX ANALYSIS

MARKO HUHTANEN and OLAVI NEVANLINNA

Institute of Mathematics, Helsinki University of Technology, Box 1100, FIN-02015, Finland
E-mail: Marko.Huhtanen@hut.fi, Olavi.Nevanlinna@tkk.fi

1. Introduction and background. This survey paper deals with real linear matrix analysis, i.e., we are concerned with operators $\mathcal{A} : \mathbb{C}^n \to \mathbb{C}^p$ defined by

$$(1.1) \qquad x \mapsto \mathcal{A}x = Ax + A_\# \tau x$$

for a pair of matrices $A, A_\# \in \mathbb{C}^{p \times n}$. Here τ is the conjugation operator $\tau x = \overline{x}$ allowing us to interchangeably use the shorthand $A + A_\# \tau$ for $\mathcal{A}$. If $A_\# = 0$, then we have the ordinary matrix-vector product and the operator is $\mathbb{C}$-linear, while with $A = 0$ we are dealing with a conlinear, i.e., antilinear operator. If both A and $A_\#$ are nonzero, then the operator is $\mathbb{R}$-linear.

Our interest in developing real linear matrix analysis originates from the late 1990's with the motivation coming from three different sources. The first one is computational. In numerical linear algebra it is common knowledge that to iteratively solve a complex linear system, rewriting the problem in an equivalent real form is not advisable. The reason is the particular block structure of the coefficient matrix with its unfavorable spectral properties that is prone to lead to slow convergence of iterations as advocated by Freund [3]. Because of this, we got interested in doing the converse, i.e., in finding a way to rewrite any real matrix problem in an appropriate complex form and then to apply complex linear algebraic techniques.

The second source lies in operator theory. In [13, 14] Putinar and Shapiro consider a multiplication operator followed by a conjugation and an orthogonal projection. The resulting operator is $\mathbb{R}$-linear on a function space over $\mathbb{C}$. The motivation for their study was a paper by Friedrichs on planar elasticity [4]. However, typically the $\mathbb{R}$-linear structure is avoided by converting the problem to a $\mathbb{C}$-linear one, or by regarding the function space as being over $\mathbb{R}$. As a consequence, the related $\mathbb{R}$-linear eigenvalue problem is not considered.

The third source is in applied mathematics. In [11] a method to solve the inverse problem of reconstructing an unknown electrical conductivity from boundary measurements is proposed. After discretizations, a very large parameter dependent $\mathbb{R}$-linear system of equations in $\mathbb{C}^n$ is obtained. In practice the solving must be done iteratively. Since there are no such methods for systems of this type in $\mathbb{C}^n$, the problem is rewritten in $\mathbb{R}^{2n}$. To our mind this appeared somewhat unnatural.

This paper consists of eight sections as follows. In section 2 basic definitions and notation are set. In section 3 two families of $\mathbb{R}$-linear operators are given to illustrate various aspects of real linear matrix analysis. In section 4 we present factorizations of $\mathbb{R}$-linear operators. Two matrix representations of $\mathbb{R}$-linear operators are suggested in section 5. Section 6 deals with the spectrum and related spectral sets for $\mathbb{R}$-linear operators. In section 7 we consider various resolvent related operators as well as functions of $\mathbb{R}$-linear operators. Finally, in section 8 we consider the problem of approximating $\mathbb{R}$-linear operators.

Most of the results in this paper are given without proof. More details can be found in [2, 10, 8, 7]. See also [1] for an application to numerical integration of large systems of Hamiltonian ODEs.

2. Definitions and notation. In real linear matrix analysis we study operators $\mathcal{A}$: $\mathbb{C}^n \to \mathbb{C}^p$ satisfying

$$(2.1) \qquad \mathcal{A}(\alpha x + \beta y) = \alpha \mathcal{A}x + \beta \mathcal{A}y$$

for $\alpha, \beta \in \mathbb{R}$ and $x, y \in \mathbb{C}^n$. A special $\mathbb{R}$-linear operator is the conjugation operator

$$(2.2) \qquad x \mapsto \tau x = \bar{x}$$

which allows every $\mathbb{R}$-linear operator $\mathcal{A}$ to be represented uniquely in the form

$$(2.3) \qquad \mathcal{A} = A + A_{\#}\tau,$$

where A and $A_{\#}$ are $\mathbb{C}$-linear operators. Here $A = \frac{1}{2i}(i\mathcal{A} + \mathcal{A}i)$ is the $\mathbb{C}$-linear part of $\mathcal{A}$ and $A_{\#} = \frac{1}{2i}(i\mathcal{A} - \mathcal{A}i)\tau$ is its conlinear part. The conjugation operator satisfies $\tau^2 = I$ and $\tau A = \bar{A}\tau$, where conjugating A means conjugating the entries of any of its matrix representations. Then

$$\mathcal{A}\mathcal{B} = AB + A_{\#}\overline{B_{\#}} + \left(AB_{\#} + A_{\#}\overline{B} \right)\tau.$$

for two $\mathbb{R}$-linear operators $\mathcal{A}$ and $\mathcal{B}$ of conforming size. The sum operation is defined in an obvious way.

In what follows, we assume A and $A_{\#}$ are complex matrices of type p-by-n and denote by $\mathcal{M}_{p,n}$ the respective set of real linear operators. This gives rise to a complex normed space once the multiplication by a scalar $\alpha \in \mathbb{C}$ is defined from the left as

$$(2.4) \qquad (\alpha\mathcal{A})(x) = \alpha(\mathcal{A}x)$$

and the operator norm is set as

$$(2.5) \qquad ||\mathcal{A}|| = \max_{||x||=1} ||\mathcal{A}x||.$$

Occasionally we also use the Frobenius norm

$$\|\mathcal{A}\|_F = \sqrt{\|A\|_F^2 + \|A_\#\|_F^2}.$$

For $n = p$ we write $\mathcal{M}_n$ for short. With respect to the operator norm $\mathcal{M}_n$ is a complex Banach space but merely a real Banach algebra since only real scalars commute with all the elements.

Let $p(\lambda) = \sum_{j=0}^k \alpha_j \lambda^j$ be a polynomial. By the polynomial p in $\mathcal{A} \in \mathcal{M}_n$ we mean the $\mathbb{R}$-linear operator

$$(2.6) \qquad p(\mathcal{A}) = \sum_{j=0}^k \alpha_j \mathcal{A}^j.$$

The degree of $\mathcal{A}$, denoted by $\deg(\mathcal{A})$, is the degree of the monic polynomial of the least degree annihilating $\mathcal{A}$.

At times we employ the real inner product $(x, y)_\mathbb{R} = \operatorname{Re}(x, y)$ on $\mathbb{C}^n$, where $(x, y) = y^* x$ denotes the standard inner product of vectors $x, y \in \mathbb{C}^n$. Then, for the concept of adjoint, recall that the adjoint A^* of a matrix $A \in \mathbb{C}^{p \times n}$ is defined by

$$(Ax, y) = (x, A^* y) \quad \text{for every } x \in \mathbb{C}^n \text{ and } y \in \mathbb{C}^p.$$

Analogously, requiring $(Ax, y)_\mathbb{R} = (x, A^* y)_\mathbb{R}$ gives us the same matrix A^*.

DEFINITION 2.1. The *adjoint* of $\mathcal{A} \in \mathcal{M}_{p,n}$ is the unique $\mathcal{A}^* \in \mathcal{M}_{n,p}$ satisfying

$$(\mathcal{A}x, y)_\mathbb{R} = (x, \mathcal{A}^* y)_\mathbb{R} \quad \text{for every } x \in \mathbb{C}^n \text{ and } y \in \mathbb{C}^p.$$

Uniqueness follows readily. It is also straightforward that $\tau^* = \tau$ and

$$\mathcal{A}^* = (A + A_\# \tau)^* = A^* + A_\#{}^T \tau.$$

With respect to the standard complex inner product we have

$$(\mathcal{A}x, y) = (x, \mathcal{A}^* y) - 2i \operatorname{Im}(x, A_\#{}^T \overline{y}).$$

DEFINITION 2.2. An $\mathbb{R}$-linear operator $\mathcal{A}$ is *self-adjoint* if $\mathcal{A}^* = \mathcal{A}$.

Clearly, $\mathcal{A}$ is self-adjoint if and only if its $\mathbb{C}$-linear part is a Hermitian and conlinear part a symmetric matrix.

DEFINITION 2.3. An $\mathbb{R}$-linear operator $A + A_\# \tau \in \mathcal{M}_n$ is *diagonal* if A and $A_\#$ are diagonal matrices.

In view of this, there are many equivalent characterizations of normality for matrices. Two of them can be regarded as the most fundamental giving a motivation for the following analogues in the $\mathbb{R}$-linear case.

DEFINITION 2.4. An $\mathbb{R}$-linear operator $\mathcal{A} \in \mathcal{M}_n$ is *normal* if

$$\mathcal{A} = U \mathcal{D} U^*$$

for a unitary $U \in \mathbb{C}^{n \times n}$ and a diagonal $\mathcal{D} \in \mathcal{M}_n$.

DEFINITION 2.5. An $\mathbb{R}$-linear operator $\mathcal{A} \in \mathcal{M}_n$ is *star-commuting* if

$$\mathcal{A}\mathcal{A}^* = \mathcal{A}^* \mathcal{A}.$$

We say that an $\mathbb{R}$-linear operator $\mathcal{P} \in \mathcal{M}_n$ is a *projection* if $\mathcal{P}^2 = \mathcal{P}$.

EXAMPLE 1. If P is a real square matrix satisfying $P^2 = P$, then

$$\mathcal{P} = \frac{1}{2}(1 + \tau)P$$

is an $\mathbb{R}$-linear projection. If P is additionally symmetric, then $\mathcal{P}$ is a self-adjoint projection.

DEFINITION 2.6. An $\mathbb{R}$-linear operator $\mathcal{W} \in \mathcal{M}_{p,n}$ is an *isometry* if

$$||\mathcal{W}x|| = ||x|| \quad \text{for all} \ \ x \in \mathbb{C}^n.$$

If $p = n$, then $\mathcal{W}$ is said to be *unitary*.

An isometry preserves the lengths but not the angles, i.e., $(x, y) \neq (\mathcal{W}x, \mathcal{W}y)$ in general. Moreover, $\mathcal{W}$ is an isometry if and only if $\mathcal{W}^*\mathcal{W} = I$. In particular, unitary $\mathbb{R}$-linear operators are invertible. By the inverse of $\mathcal{A} \in \mathcal{M}_n$ we obviously mean $\mathcal{A}^{-1} \in \mathcal{M}_n$ satisfying $\mathcal{A}\mathcal{A}^{-1} = \mathcal{A}^{-1}\mathcal{A} = I$.

EXAMPLE 2. If $Q \in \mathbb{C}^{n \times k}$ satisfies $\mathrm{Re}(Q^*Q) = I$, then $\mathcal{W} = (I - QQ^*) - QQ^T\tau \in \mathcal{M}_n$ is unitary. For $k = 2$ such a $\mathcal{W}$ is called an $\mathbb{R}$-linear Householder transformation.

A convenient formula for the inverse is as follows.

PROPOSITION 2.7. *For* $\mathcal{A} \in \mathcal{M}_n$ *assume* A *and* $A - A_\#\overline{A}^{-1}\overline{A_\#}$ *are invertible. Then* $\mathcal{A}^{-1} = (I - A^{-1}A_\#\tau)(A - A_\#\overline{A}^{-1}\overline{A_\#})^{-1}$.

For $\mathcal{A} \in \mathcal{M}_{p,n}$ we denote by

$$\mathrm{Ker}(\mathcal{A}) = \{x \in \mathbb{C}^n : \ \mathcal{A}x = 0\}$$

the kernel of $\mathcal{A}$ and by

$$\mathrm{Ran}(\mathcal{A}) = \{y \in \mathbb{C}^p : \ \text{there exists } x \in \mathbb{C}^n \text{ such that } y = \mathcal{A}x\}$$

its range. These are linear manifolds over $\mathbb{R}$ but generally not subspaces of $\mathbb{C}^n$ (which we always regard as being over $\mathbb{C}$).

Basic linear algebraic manipulations yield

$$\mathrm{Ker}(\mathcal{A}) \oplus \mathrm{Ran}(\mathcal{A}^*) = \mathbb{C}^n,$$

and that $\mathcal{A}$ is invertible if and only if $\mathrm{Ker}(\mathcal{A}) = \{0\}$. This allows us to define the spectrum of $\mathcal{A} \in \mathcal{M}_n$ as

$$\sigma(\mathcal{A}) = \{\lambda \in \mathbb{C} : \ \lambda I - \mathcal{A} \text{ is not invertible}\}.$$

Then $\lambda \in \sigma(\mathcal{A})$ is called an eigenvalue and a nonzero $x \in \mathrm{Ker}(\lambda I - \mathcal{A})$ a corresponding eigenvector of $\mathcal{A}$.

For $\lambda \notin \sigma(\mathcal{A})$ the resolvent operator of $\mathcal{A}$ is defined by

$$(2.7) \qquad \qquad \lambda \mapsto \mathcal{R}(\lambda, \mathcal{A}) = (\lambda I - \mathcal{A})^{-1}.$$

Another object of interest is

$$(2.8) \qquad \qquad \lambda \mapsto \mathcal{C}(\lambda, \mathcal{A}) = \sum_{j=0}^{\infty} \frac{1}{\lambda^{j+1}} \mathcal{A}^j$$

for $|\lambda| > ||\mathcal{A}||$ and otherwise extended by analytic continuation.

EXAMPLE 3. The spectrum can be empty. In fact, for $\mathcal{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tau$ we have

$$\mathcal{R}(\lambda, \mathcal{A}) = \frac{1}{|\lambda|^2 + 1} \left(\overline{\lambda} I + \mathcal{A} \right)$$

which is real analytic and bounded in the whole complex plane. On the other hand,

$$\mathcal{C}(\lambda, \mathcal{A}) = \frac{1}{\lambda^2 + 1} (\lambda I + \mathcal{A}).$$

EXAMPLE 4. For the conjugation operator $\tau \in \mathcal{M}_n$ we have

$$\mathcal{R}(\lambda, \tau) = \frac{1}{|\lambda|^2 - 1} \left(\overline{\lambda} I + \tau \right)$$

so that the spectrum of τ is the unit circle. Further,

$$\mathcal{C}(\lambda, \tau) = \frac{1}{\lambda^2 - 1} \left(\lambda I + \tau \right),$$

with singularities at 1 and -1.

3. Two families of $\mathbb{R}$-linear operators. To get an idea of how $\mathbb{R}$-linear operators differ from the $\mathbb{C}$-linear ones, we list properties of two families of $\mathbb{R}$-linear operators of importance. We introduce circlets to generalize scalars and operets to generalize rank-1 matrices.

3.1. *Circlets.* The operator $\lambda - \tau$ in Example 4 can be regarded as an $\mathbb{R}$-linear generalization of the scalar λ. For more generality, we set the following definition.

DEFINITION 3.1. Assume $\alpha, \beta \in \mathbb{C}$. Then

$$\alpha + \beta \tau \in \mathcal{M}_n$$

is called a *circlet*.

Circlets act like scalars in the sense that their action is unambiguous without specifying the dimension n.

Sometimes it is useful to interpret a diagonal $\mathcal{A} \in \mathcal{M}_n$ as being the direct sum of circlets on $\mathbb{C}$.

For the operator norm of a circlet we have

(3.1) $$\|\alpha + \beta \tau\| = |\alpha| + |\beta|.$$

The resolvent operator equals

(3.2) $$\mathcal{R}(\lambda, \alpha + \beta \tau) = \frac{1}{|\lambda - \alpha|^2 - |\beta|^2} \left(\overline{\lambda} - \overline{\alpha} + \beta \tau \right),$$

so that the spectrum of $\alpha + \beta \tau$ is the circle of radius $|\beta|$ centered at α. Also

(3.3) $$\mathcal{C}(\lambda, \alpha + \beta \tau) = \frac{1}{\lambda^2 - (\alpha + \overline{\alpha})\lambda + |\alpha|^2 - |\beta|^2} \left(\lambda - \overline{\alpha} + \beta \tau \right)$$

having singularities at $\lambda = \operatorname{Re}\alpha \pm \sqrt{|\beta|^2 - (\operatorname{Im}\alpha)^2}$.

A circlet $\alpha + \beta \tau$ is self-adjoint if and only if $\alpha \in \mathbb{R}$.

Any circlet $\alpha + \beta \tau$ is normal whereas having either $\alpha \in \mathbb{R}$ or $\beta = 0$ is a necessary and sufficient condition for $\alpha + \beta \tau$ to be star-commuting.

A circlet $\alpha + \beta\tau$ is a projector if and only if $|\alpha| = |\beta|$ and $\alpha = \frac{1}{2}(1 + it)$ for $t \in \mathbb{R}$. A projector that is additionally self-adjoint takes the form $\frac{1}{2}(1 + e^{i\theta}\tau)$ for $\theta \in \mathbb{R}$.

The only unitary circlets are $e^{i\theta}$ and $e^{i\theta}\tau$ for $\theta \in \mathbb{R}$.

Circlets are always annihilated by a second degree polynomial. For $\alpha + \beta\tau$ to be nilpotent, a necessary and sufficient condition is that $\alpha + \beta\tau = t(i + e^{i\theta}\tau)$ for $t, \theta \in \mathbb{R}$.

PROPOSITION 3.2. *There exists a polynomial $p(\lambda) = \lambda^2 + \gamma\lambda$ such that $p(\alpha + \beta\tau) = 0$ if and only if $\alpha + \beta\tau$ is not invertible or it is a scalar.*

Proof. Assume $p(\alpha + \beta\tau) = 0$ and factor p as $p(\lambda) = (\lambda + \gamma)\lambda$. If $\alpha + \beta\tau$ is invertible, then it follows that $\alpha + \beta\tau = \gamma$.

For the converse, if $\alpha + \beta\tau$ is scalar, then the claim is true. If $\alpha + \beta\tau$ is not invertible, then $q(\alpha + \beta\tau) = 0$ for a second degree polynomial. The constant term must be zero since otherwise $\alpha + \beta\tau$ would be invertible. ∎

Assume $\gamma + \delta\tau$ is invertible and consider

$$(3.4) \qquad\qquad (\gamma + \delta\tau)(\alpha + \beta\tau)(\gamma + \delta\tau)^{-1},$$

i.e., the circlet $\alpha + \beta\tau$ has been "similarity transformed" by applying $\gamma + \delta\tau$.

PROPOSITION 3.3. *For $\alpha + \beta\tau$ there exists an invertible $\gamma + \delta\tau$ such that (3.4) equals either* (i) $\xi + \eta\tau$, (ii) $\xi + i\eta$, *or* (iii) $\xi + \eta(i + \tau)$, *with $\xi = \operatorname{Re}\alpha$ and $\eta \in \mathbb{R}$.*

3.2. *Operets.* To generalize rank-1 matrices, we make the following definition.

DEFINITION 3.4. Assume $u, v \in \mathbb{C}^p$ and $w \in \mathbb{C}^n$. Then

$$(u + v\tau)w^* \in \mathcal{M}_{p,n}$$

is called an *operet*.

For the operator norm of an operet we have

$$(3.5) \qquad\qquad \|(u + v\tau)w^*\| = \sqrt{\|u\|^2 + 2\,|v^*u| + \|v\|^2}\,\|w\|.$$

The resolvent operator of an operet $\mathcal{O} = (u + v\tau)w^* \in \mathcal{M}_n$ equals

$$\mathcal{R}(\lambda, \mathcal{O}) = \frac{1}{\lambda}\left(I + \frac{\overline{\lambda}u - \overline{w^*u}u + \overline{w^*v}v + (\lambda v - w^*uv + w^*vu)\tau}{|\lambda - w^*u|^2 - |w^*v|^2}w^*\right)$$

so that the spectrum of $\mathcal{O}$ consists of the circle of radius $|w^*v|$ centered at w^*u together with the origin, assuming $n > 1$. (For $n = 1$ an operet is simply a circlet.) Also

$$\mathcal{C}(\lambda, \mathcal{O}) = \frac{1}{\lambda}\left(I + \frac{\lambda u - \overline{w^*u}u + \overline{w^*v}v + (\lambda v - w^*uv + w^*vu)\tau}{\lambda^2 - (w^*u + u^*w)\lambda + |w^*u|^2 - |w^*v|^2}w^*\right),$$

where the dividing scalar operates from the left.

An operet $\mathcal{O}$ is normal when it can be written as

$$\mathcal{O} = u(\alpha + \beta\tau)u^* \quad \text{with a circlet} \ \ \alpha + \beta\tau \in \mathcal{M}_n.$$

If $\alpha \in \mathbb{R}$, then we have a self-adjoint operet. Moreover, a star-commuting $\alpha + \beta\tau$ gives rise to a star-commuting operet.

The sum of j operets can be given as $(U + V\tau)W^*$ with matrices $U, V \in \mathbb{C}^{p \times j}$ and $W \in \mathbb{C}^{n \times j}$.

Analogously to rank-1 matrices, operets arise as follows. For an $\mathbb{R}$-linear operator $A + A_\#\tau \in \mathcal{M}_{p,n}$, let u_j and v_j denote the columns of A and $A_\#$. If $\{e_i\}_{j=1}^n$ is the standard basis of $\mathbb{C}^n$, then

$$(3.6) \qquad \mathcal{A} = \sum_{j=1}^{n} (u_j + v_j\tau)e_j^*$$

represents $\mathcal{A}$ as the sum of operets. Using the standard basis is too restrictive for tasks such as representing $\mathcal{A}$ with the minimum number of operets.

DEFINITION 3.5. Let $\mathcal{A} \in \mathcal{M}_{p,n}$. The *right-rank* of $\mathcal{A}$ is right-rank$(\mathcal{A}) = \dim(\text{Ran}(\mathcal{A}^*)^\perp)^\perp$. The *left-rank* of $\mathcal{A}$ is left-rank$(\mathcal{A}) = \dim(\text{Ran}(\mathcal{A})^\perp)^\perp$.

For the minimum number of operets these give the following connection.

PROPOSITION 3.6. *The smallest number of operets needed to represent $\mathcal{A} \in \mathcal{M}_{p,n}$ as their sum equals* right-rank$(\mathcal{A})$.

Aside from the minimum number of terms, a central problem is to find a representation that is in some sense optimal; see section 8.

4. Decompositions of $\mathbb{R}$-linear operators.

An $\mathbb{R}$-linear operator $\mathcal{L} \in \mathcal{M}_n$ is said to be lower (resp. upper) triangular if L and $L_\#$ are lower (resp. upper) triangular matrices. Under fairly mild assumptions on $\mathcal{A} \in \mathcal{M}_n$, there exists an $\mathcal{LU}$-factorization of $\mathcal{A}$, i.e., $\mathcal{A} = \mathcal{LU}$ with $\mathcal{L}$ and $\mathcal{U}$ lower and upper triangular, respectively [2]. This decomposition can be used in solving $\mathbb{R}$-linear systems of equations.

An approach based on the Gram-Schmidt orthogonalization process does not seem to yield a $\mathcal{QR}$-factorization. In this respect $\mathbb{R}$-linear Householder transformations (see Example 2) are more fundamental and by applying them we have the anticipated decomposition.

THEOREM 4.1 ([2]). *For $\mathcal{A} \in \mathcal{M}_{p,n}$ there exists a unitary $\mathcal{Q} \in \mathcal{M}_p$ and an upper triangular $\mathcal{R} \in \mathcal{M}_{p,n}$ such that $\mathcal{A} = \mathcal{QR}$.*

An analogy of the Schur decomposition is as follows.

THEOREM 4.2 ([2]). *For $\mathcal{A} \in \mathcal{M}_n$ there exists a unitary $\mathcal{U} \in \mathcal{M}_n$ and an upper triangular $\mathcal{T} \in \mathcal{M}_n$ such that $\mathcal{A} = \mathcal{UTU}^*$.*

By employing Proposition 3.3 we can refine the diagonal entries in this factorization.

COROLLARY 4.3. *For $\mathcal{A} \in \mathcal{M}_n$ there exists an invertible $\mathcal{S} \in \mathcal{M}_n$ and an upper triangular $\mathcal{T} \in \mathcal{M}_n$ such that*

$$\mathcal{A} = \mathcal{STS}^{-1}$$

and each diagonal entry of $\mathcal{T}$ is of type (i) $\xi + \eta\tau$, (ii) $\xi + i\eta$, *or* (iii) $\xi + \eta(i + \tau)$, *with* $\xi, \eta \in \mathbb{R}$.

In view of Definition 2.4, star-commutativity is in essence a less stringent generalization of normality in the following sense.

THEOREM 4.4. *Let $\mathcal{A} \in \mathcal{M}_n$. Then $\mathcal{A} = \mathcal{U}\mathcal{D}\mathcal{U}^*$ with a unitary $\mathcal{U} \in \mathcal{M}_n$ and a diagonal $\mathcal{D} \in \mathcal{M}_n$ having diagonal entries of type* (i) $\xi + \eta\tau$ *or* (ii) $\xi + i\eta$, *with $\xi, \eta \in \mathbb{R}$ if and only if $\mathcal{A}$ is star-commuting.*

Proof. The other direction being clear, assume $\mathcal{A}$ is star-commuting. Let $\mathcal{A}$ be decomposed as in Theorem 4.2. Hence $\mathcal{A}$ is star-commuting if and only if $\mathcal{T}$ is star-commuting. For the linear part this implies

$$TT^* + T_\# T_\#{}^* = T^*T + T_\#{}^T \overline{T_\#}$$

which forces $\mathcal{T}$ to be diagonal, i.e., $\mathcal{T}$ is the direct sum of star-commuting circlets. It remains to perform a unitary similarity to deal with each of these circlets separately.

Having either $\alpha \in \mathbb{R}$ or $\beta = 0$ is a necessary and sufficient condition for a circlet $\alpha + \beta\tau$ to be star-commuting. If $\beta = re^{i\theta} \neq 0$, then $\alpha + \beta\tau = e^{\theta/2}(\alpha + r\tau)e^{-\theta/2}$. ∎

EXAMPLE 5. Take $\mathcal{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\tau$ of Example 3 which is star-commuting. Then with the unitary

$$\mathcal{U} = \frac{1}{2}\left(\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix} + \begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix}\tau \right)$$

we have $\mathcal{U}^*\mathcal{A}\mathcal{U} = \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix}$.

Ruling out the diagonal entries of type (ii) gives rise to the set of self-adjoint $\mathbb{R}$-linear operators.

COROLLARY 4.5. *Let $\mathcal{A} \in \mathcal{M}_n$. Then $\mathcal{A} = \mathcal{U}\mathcal{D}\mathcal{U}^*$ with a unitary $\mathcal{U} \in \mathcal{M}_n$ and a diagonal $\mathcal{D} \in \mathcal{M}_n$ having diagonal entries of type* (i) $\xi + \eta\tau$ *with $\xi, \eta \in \mathbb{R}$ if and only if $\mathcal{A}$ is self-adjoint.*

An interesting class of $\mathbb{R}$-linear operators consists of those $\mathcal{A} \in \mathcal{M}_n$ for which $A^*A_\#$ is skew-symmetric. Equivalently, the conlinear part of $\mathcal{A}^*\mathcal{A}$ vanishes.

THEOREM 4.6 ([8]). *Let $\mathcal{A} \in \mathcal{M}_n$. Then the conlinear part of $\mathcal{A}^*\mathcal{A}$ vanishes if and only if $\mathcal{A} = \mathcal{W}\mathcal{D}\mathcal{V}^*$ for a unitary $\mathcal{W} \in \mathcal{M}_n$, a diagonal $D \in \mathbb{C}^{n \times n}$ and unitary $V \in \mathbb{C}^{n \times n}$.*

The following conditions are also of interest.

THEOREM 4.7 ([8]). *Let $\mathcal{A} \in \mathcal{M}_n$. Then $\mathcal{A} = U\mathcal{D}V^*$ for unitary $U, V \in \mathbb{C}^{n \times n}$ and a diagonal $\mathcal{D} \in \mathcal{M}_n$ if and only if $AA_\#{}^T$ and $A^*A_\#$ are symmetric.*

If $U = V$, then $\mathcal{A}$ is normal. Equivalent conditions for this to hold follow by applying [5, Corollary 5.3].

THEOREM 4.8 ([8]). *Let $\mathcal{A} \in \mathcal{M}_n$. Then $\mathcal{A}$ is normal if and only if A is normal and $A_\#$, $(A + A^*)A_\#$, and $(A - A^*)A_\#$ are symmetric.*

There remains the open problem of characterizing those $\mathbb{R}$-linear operators $\mathcal{A} \in \mathcal{M}_n$ that can be decomposed as

$$\mathcal{A} = \mathcal{W}\mathcal{D}\mathcal{V}^*$$

with a unitary $\mathcal{W} \in \mathcal{M}_n$, a diagonal $\mathcal{D} \in \mathcal{M}_n$ and a unitary $V \in \mathbb{C}^{n \times n}$.

5. Matrix representations of $\mathbb{R}$-linear operators. Next we consider representing an $\mathbb{R}$-linear operator $\mathcal{A} \in \mathcal{M}_{p,n}$ with a $2p$-by-$2n$ matrix.

5.1. *Representing $\mathcal{A}$ in $\mathbb{R}^{2p \times 2n}$.* Define $\mathcal{F} = F + F_{\#}\tau \in \mathcal{M}_{2n,n}$ by setting

$$F = \frac{1}{2}\begin{bmatrix} I \\ -iI \end{bmatrix} \quad \text{and} \quad F_{\#} = \overline{F} = \frac{1}{2}\begin{bmatrix} I \\ iI \end{bmatrix}$$

so that $\mathcal{F} = (1+\tau)F$ and $\mathcal{F}x = \begin{bmatrix} \operatorname{Re} x \\ \operatorname{Im} x \end{bmatrix}$. Therefore $\operatorname{Ran}(\mathcal{R}) = \mathbb{R}^{2n}$. To reverse the action of $\mathcal{F}$ with a $\mathbb{C}$-linear operator, take $2F$ to have $2F^*\mathcal{F} = I$.

PROPOSITION 5.1. *We have $\mathcal{F}^*\mathcal{F} = I$ and $\mathcal{F}\mathcal{F}^* = \frac{1}{2} + \frac{1}{2}\tau$. Moreover,*

$$(\mathcal{F}x, \mathcal{F}y) = (x,y)_{\mathbb{R}}$$

for every $x, y \in \mathbb{C}^n$.

Observe that $\mathcal{F}$ is an isometry.

Define $\phi : \mathcal{M}_{p,n} \to \mathbb{R}^{2p \times 2n}$ by

(5.1) $$\phi(\mathcal{A}) = \begin{bmatrix} \operatorname{Re}(A + A_{\#}) & \operatorname{Im}(-A + A_{\#}) \\ \operatorname{Im}(A + A_{\#}) & \operatorname{Re}(A - A_{\#}) \end{bmatrix}.$$

We call $\phi(\mathcal{A})$ the real matrix representation of $\mathcal{A}$ due to the fact that $\mathcal{F}\mathcal{A} = \phi(\mathcal{A})\mathcal{F}$. (Here $\mathcal{F}$ is assumed to be of appropriate type without any reference to the dimension of the underlying space.) This identity leads to the standard way of dealing with $\mathbb{R}$-linear problems, i.e., by representing the operator with a real matrix while the vector spaces are regarded as being over $\mathbb{R}$. It follows readily that ϕ is bijective, isometric and satisfies $\phi(\mathcal{A} + \mathcal{B}) = \phi(\mathcal{A}) + \phi(\mathcal{B})$. Moreover, for $\mathcal{A}$ and $\mathcal{B}$ of conforming size we have $\phi(\mathcal{A}\mathcal{B}) = \phi(\mathcal{A})\phi(\mathcal{B})$.

From the above manipulations follow the identities

(5.2) $$\mathcal{A} = \mathcal{F}^*\phi(\mathcal{A})\mathcal{F} = 2F^*\phi(\mathcal{A})\mathcal{F} = 2\mathcal{F}^*\phi(\mathcal{A})F.$$

5.2. *Representing $\mathcal{A}$ in $\mathbb{C}^{2p \times 2n}$.* Define $\mathcal{G} = G + G_{\#}\tau \in \mathcal{M}_{2n,n}$ by setting

$$G = \frac{1}{\sqrt{2}}\begin{bmatrix} I \\ 0 \end{bmatrix} \quad \text{and} \quad G_{\#} = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ I \end{bmatrix}$$

to have an isometry. To reverse the action of $\mathcal{G}$ with a $\mathbb{C}$-linear operator, take $2G$ to have $2G^*\mathcal{G} = I$.

PROPOSITION 5.2. *We have $\mathcal{G}^*\mathcal{G} = I$ and $\mathcal{G}\mathcal{G}^* = \frac{1}{2}(I + [\begin{smallmatrix} 0 & I \\ I & 0 \end{smallmatrix}]\tau)$. Moreover,*

$$(\mathcal{G}x, \mathcal{G}y) = (x,y)_{\mathbb{R}}$$

for every $x, y \in \mathbb{C}^n$.

Define $\psi : \mathcal{M}_{p,n} \to \mathbb{C}^{2p \times 2n}$ by

(5.3) $$\psi(\mathcal{A}) = \begin{bmatrix} A & A_{\#} \\ \overline{A_{\#}} & \overline{A} \end{bmatrix}.$$

We call $\psi(\mathcal{A})$ the complex matrix representation of $\mathcal{A}$, because $\mathcal{G}\mathcal{A} = \psi(\mathcal{A})\mathcal{G}$. Now ψ is injective, isometric and satisfies $\psi(\mathcal{A} + \mathcal{B}) = \psi(\mathcal{A}) + \psi(\mathcal{B})$. Moreover, for $\mathcal{A}$ and $\mathcal{B}$ of conforming size $\psi(\mathcal{A}\mathcal{B}) = \psi(\mathcal{A})\psi(\mathcal{B})$ holds.

We also have the identities

(5.4) $\mathcal{A} = \mathcal{G}^*\psi(\mathcal{A})\mathcal{G} = 2G^*\psi(\mathcal{A})\mathcal{G} = 2\mathcal{G}^*\psi(\mathcal{A})G.$

Like with the singular value decomposition of a matrix, we say that a (not necessarily square) matrix D is diagonal if only its d_{jj}-entries are possibly nonzero.

THEOREM 5.3. *For $\mathcal{A} \in \mathcal{M}_{p,n}$ there exists an isometry $\mathcal{W}^* \in \mathcal{M}_{2p,p}$, a diagonal matrix $D \in \mathbb{C}^{2p \times 2n}$ with nonnegative entries and an isometry $V^* \in \mathbb{C}^{2n \times n}$ such that*

$$\mathcal{A} = \mathcal{W}DV^*.$$

Proof. From (5.4) we have $\mathcal{A} = 2\mathcal{G}^*\psi(\mathcal{A})G$. Hence, with the singular value decomposition $\psi(\mathcal{A}) = U\Sigma W^*$ of $\psi(\mathcal{A})$, set $\mathcal{W} = \mathcal{G}^*U$, $D = \sqrt{2}\Sigma$, $V = \sqrt{2}G^*W$. ∎

With the unitary matrix $E = \frac{1}{\sqrt{2}}\begin{bmatrix} I & iI \\ I & -iI \end{bmatrix}$ the real and complex matrix representations of $\mathcal{A}$ are related through

(5.5) $\phi(\mathcal{A}) = E^*\psi(\mathcal{A})E,$

i.e., the matrix representations $\phi(\mathcal{A})$ and $\psi(\mathcal{A})$ of $\mathcal{A}$ are unitarily similar.

6. Spectrum and other spectral sets for $\mathbb{R}$-linear operators.
Assume $\mathcal{A} \in \mathcal{M}_n$ and consider the $\mathbb{R}$-linear eigenvalue problem

(6.1) $\mathcal{A}x = Ax + A_\#\overline{x} = \lambda x$

for $\lambda \in \mathbb{C}$ and a nonzero vector $x \in \mathbb{C}^n$.

By separating $\lambda = \alpha + i\beta$ into its real and imaginary parts, associate with (6.1) the bivariate polynomial

(6.2) $p_\mathcal{A}(\alpha, \beta) = \det(\phi(\lambda I - \mathcal{A})) = \det(\psi(\lambda I - \mathcal{A}))$

called the characteristic bivariate polynomial of $\mathcal{A}$. Then λ is an eigenvalue of $\mathcal{A}$ if and only if $p_\mathcal{A}(\alpha, \beta) = 0$.

PROPOSITION 6.1. *The characteristic bivariate polynomial of $\mathcal{A} \in \mathcal{M}_n$ is of degree $2n$ such that*

$$p_\mathcal{A}(\alpha, \beta) = (\alpha^2 + \beta^2)^n + \text{lower order terms.}$$

The spectrum of $\mathcal{A}$ is thus a real algebraic plane curve in $\mathbb{R}^2 \simeq \mathbb{C}$ of degree $2n$ at most.

The spectrum function $\mathcal{A} \mapsto \sigma(\mathcal{A})$ is upper semi-continuous.

The spectrum can be empty; see Example 3. For conditions to have non-void spectrum, split $A_\#$ as $A_\# = S + T$, where $S = \frac{1}{2}(A_\# + A_\#^T)$ is the symmetric part of $A_\#$. In the following theorem we use this notation.

THEOREM 6.2 ([10]). *Split $\mathcal{A} \in \mathcal{M}_n$ as $\mathcal{A} = \mathcal{B} + T\tau$. If $||T^{-1}|| \, ||\mathcal{B}|| < 1$, then $\sigma(\mathcal{A}) = \emptyset$.*

If $n = 2$ and $A_\# = S$, then $\sigma(\mathcal{A}) \neq \emptyset$; see [9]. It is an open problem to prove or disprove this for $n > 2$.

If $\mu \in \mathbb{C}$, then $\sigma(\mu\mathcal{A}) = \sigma(\mathcal{A}\mu) = \mu\sigma(\mathcal{A})$ and $\sigma(\mu I + \mathcal{A}) = \mu + \sigma(\mathcal{A})$.

PROPOSITION 6.3. *Assume $\mathcal{A} \in \mathcal{M}_n$. Then $\sigma(\mathcal{A}^*) = \overline{\sigma(\mathcal{A})}$.*

If $\mathcal{S} \in \mathcal{M}_n$ is invertible, then in general we have only

$$\sigma(\mathcal{S}\mathcal{A}\mathcal{S}^{-1}) \cap \mathbb{R} = \sigma(\mathcal{A}) \cap \mathbb{R}$$

unless $\mathcal{S}$ is $\mathbb{C}$-linear.

THEOREM 6.4. *Assume* $\mathcal{A} \in \mathcal{M}_n$. *If* $S \in \mathbb{C}^{n \times n}$ *is invertible, then*

$$\sigma(S\mathcal{A}S^{-1}) = \sigma(\mathcal{A}).$$

For an upper triangular $\mathcal{T} \in \mathcal{M}_n$, let the circlets $\alpha_j + \beta_j \tau$ denote the diagonal entries of $\mathcal{T}$ for $1 \le j \le n$. This notation is used in the following theorem.

THEOREM 6.5 ([2]). *Assume* $\mathcal{A} \in \mathcal{M}_n$ *such that* $\mathcal{A} = S\mathcal{T}S^{-1}$ *with an invertible* $S \in \mathbb{C}^{n \times n}$ *and an upper triangular* $\mathcal{T} \in \mathcal{M}_n$. *Then* $\sigma(\mathcal{A})$ *is the union of the circles*

$$\{\lambda \in \mathbb{C} : \; |\alpha_j - \lambda| = |\beta_j|\}, \;\; j = 1, \ldots, n.$$

It is an open problem to give necessary and sufficient conditions on $\mathcal{A}$ to be upper triangularizable in the sense of this theorem.

COROLLARY 6.6. *If* $\mathcal{A} \in \mathcal{M}_n$ *is normal, then* $\sigma(\mathcal{A}) \ne \emptyset$.

By Example 5 the spectrum of a star-commuting $\mathbb{R}$-linear operator can be void.

EXAMPLE 6. In a conlinear similarity transformation the spectrum gets conjugated. In fact, let $\mathcal{S} = S\tau \in \mathcal{M}_n$ be invertible. Then $\mathcal{S}\mathcal{A}\mathcal{S}^{-1} = S\overline{\mathcal{A}}S^{-1}$, where $\overline{\mathcal{A}} = \overline{A} + \overline{A_{\#}}\tau$. But $\sigma(\overline{\mathcal{A}}) = \overline{\sigma(\mathcal{A})}$, so that $\sigma(\mathcal{S}\mathcal{A}\mathcal{S}^{-1}) = \overline{\sigma(\mathcal{A})}$.

If $\lambda \in \sigma(\mathcal{A})$, then $\mathrm{Ker}(\lambda I - \mathcal{A})$ is a linear manifold over $\mathbb{R}$. A nonzero subspace $V \subset \mathbb{C}^n$ is said to be invariant under $\mathcal{A}$ if $x \in V$ implies $\mathcal{A}x \in V$. Then $\sigma(\mathcal{A}_{|V}) \subset \sigma(\mathcal{A})$, where $\mathcal{A}_{|V}$ denotes the restriction of $\mathcal{A}$ to V.

THEOREM 6.7 ([10]). *Any* $\mathcal{A} \in \mathcal{M}_n$ *has an invariant subspace of dimension*

$$\min \{\deg(A)(\mathrm{rank}(A_{\#}) + 1), \deg(A_{\#}\tau)(\mathrm{rank}(A) + 1)\}$$

at most.

We also have the following bound in terms of the left-rank of $\mathcal{A}$.

PROPOSITION 6.8. *Let* $\mathcal{A} \in \mathcal{M}_n$. *Then* $\mathcal{A}$ *has an invariant subspace of dimension* left-rank$(\mathcal{A})$.

Since the spectrum of $\mathcal{A}$ is determined by the matrix representations of $\lambda I - \mathcal{A}$ through (6.2), the spectra of $\phi(\mathcal{A}) \in \mathbb{R}^{2n \times 2n}$ and $\psi(\mathcal{A}) \in \mathbb{C}^{2n \times 2n}$ are of interest. The similarity (5.5) yields the following proposition.

PROPOSITION 6.9. *Let* $\mathcal{A} \in \mathcal{M}_n$. *Then* $\sigma(\phi(\mathcal{A})) = \sigma(\psi(\mathcal{A}))$ *consists of* $2n$ *points, counting multiplicities, that are symmetrically located relative to the real axis.*

If $\mathcal{S} \in \mathcal{M}_n$ is invertible, then

$$\sigma(\psi(\mathcal{S}\mathcal{A}\mathcal{S}^{-1})) = \sigma(\psi(\mathcal{A})).$$

Consider $\psi(\mathcal{A})$. Then $\begin{bmatrix} u \\ v \end{bmatrix} \in \mathbb{C}^{2n}$ is an eigenvector corresponding to $\lambda \in \sigma(\psi(\mathcal{A}))$ if and only if

(6.3) $$(\mathcal{A} - \lambda I)\alpha u = -(\mathcal{A} - \overline{\lambda} I)\tau \alpha v \quad \text{for any} \ \ \alpha \in \mathbb{C}.$$

Hence we have $(\mathcal{A}^k - \lambda^k I)\alpha u = -(\mathcal{A}^k - \overline{\lambda}^k I)\tau \alpha v$ for $k \in \mathbb{N}$ and for any $\alpha \in \mathbb{C}$.

In the Schur decomposition $\mathcal{A} = \mathcal{U}\mathcal{T}\mathcal{U}^*$ of $\mathcal{A}$ (see Theorem 4.2), let the circlets $\alpha_j + \beta_j \tau$ denote the diagonal entries of $\mathcal{T}$ for $1 \leq j \leq n$. These can be used to determine the spectrum of $\psi(\mathcal{A})$ as follows.

PROPOSITION 6.10. *Let $\mathcal{A} = \mathcal{U}\mathcal{T}\mathcal{U}^*$ be the Schur decomposition of $\mathcal{A} \in \mathcal{M}_n$ with $\alpha_j + \beta_j \tau$ being the diagonal entries of $\mathcal{T}$ for $1 \leq j \leq n$. Then*

$$\prod_{j=1}^{n}(\lambda^2 - (\alpha_j + \overline{\alpha}_j)\lambda + |\alpha_j|^2 - |\beta_j|^2)$$

is the characteristic polynomial of $\psi(\mathcal{A})$.

Proof. Since $\psi(\mathcal{U}\mathcal{T}\mathcal{U}^*) = \psi(\mathcal{U})\psi(\mathcal{T})\psi(\mathcal{U}^*)$, the characteristic polynomials of $\psi(\mathcal{A})$ and $\psi(\mathcal{T})$ are equal. ∎

The spectra of $\mathcal{A}$ and $\psi(\mathcal{A})$ satisfy

$$\sigma(\mathcal{A}) \cap \mathbb{R} = \sigma(\psi(\mathcal{A})) \cap \mathbb{R},$$

so that if the matrix representations of $\mathcal{A}$ have real eigenvalues, then $\sigma(\mathcal{A}) \neq \emptyset$. This is always the case if $\mathcal{A}$ is self-adjoint, since then so is $\psi(\mathcal{A})$.

The connection between intersecting complex eigenvalues is more subtle. In fact, suppose $\lambda \in \sigma(\mathcal{A}) \cap \sigma(\psi(\mathcal{A}))$ so that

(6.4) $$\mathcal{A}x = \lambda x \quad \text{and} \quad \begin{bmatrix} A & A_{\#} \\ \overline{A_{\#}} & \overline{A} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \lambda \begin{bmatrix} u \\ v \end{bmatrix}$$

with nonzero vectors $x \in \mathbb{C}^n$ and $\begin{bmatrix} u \\ v \end{bmatrix} \in \mathbb{C}^{2n}$.

PROPOSITION 6.11. *Let $\mathcal{A} \in \mathcal{M}_n$ and assume (6.4) holds. Then either $u \neq \alpha x$ for any $\alpha \in \mathbb{C}$, $\overline{\lambda} \in \sigma(\mathcal{A})$, or $\mathcal{A}(\alpha x) = \lambda \alpha x$ for any $\alpha \in \mathbb{C}$.*

Proof. Suppose $u = \alpha x$ for some $\alpha \in \mathbb{C}$. If $\alpha = 0$, then (6.4) yields $\mathcal{A}\overline{v} = \overline{\lambda}\overline{v}$, so that $\overline{\lambda} \in \sigma(\mathcal{A})$. If $\alpha \neq 0$, then $\alpha u \neq 0$ and $\overline{\lambda} \in \sigma(\mathcal{A})$ by (6.3) assuming $v \neq 0$. If $v = 0$, then by (6.4) we may assume $x = u$ and that $A_{\#}\overline{x} = 0$. But then $\mathcal{A}(\alpha x) = \lambda \alpha x$ for any $\alpha \in \mathbb{C}$. ∎

Set $c_l(A, A_{\#}) = \{\lambda \in \mathbb{C} : |a_{l,l} - \lambda| = |a_{\#l,l}|\}$ for $\mathcal{A} \in \mathcal{M}_n$ and define

(6.5) $$\mathbf{A}_l(A, A_{\#}) = \Big\{\lambda \in \mathbb{C} : \text{dist}\big(\lambda, c_l(A, A_{\#})\big) \leq \sum_{j=1, j \neq l}^{n} (|a_{l,j}| + |a_{\#l,j}|)\Big\},$$

for $1 \leq l \leq n$. These "Geršgorin's annuli" are of use as follows.

THEOREM 6.12 ([10]). *Let $\mathcal{A} \in \mathcal{M}_n$. Then $\sigma(\mathcal{A}) \subset \cup_{j=1}^{n} \mathbf{A}_j(A, A_{\#})$.*

This can be applied to $\mathcal{A}^*$ and used with Proposition 6.3. Observe that for a diagonal $\mathbb{R}$-linear operator this is an equality.

Define the numerical range of $\mathcal{A} \in \mathcal{M}_n$ by

$$W(\mathcal{A}) = \{\lambda \in \mathbb{C}: \ \lambda = z^* \mathcal{A} z, \text{ with } ||z|| = 1\}.$$

We obtain a compact set such that $\sigma(\mathcal{A}) \subset W(\mathcal{A}) \subset W(\psi(\mathcal{A}))$. Observe though that $W(\mathcal{A})$ differs from the numerical range of $\psi(\mathcal{A})$ in general.

$W(\mathcal{A}+\mathcal{B}) \subset W(\mathcal{A})+W(\mathcal{B})$ holds for any real linear operators $\mathcal{A}$ and $\mathcal{B}$. Also $W(\mu\mathcal{A}) = \mu W(\mathcal{A})$ for $\mu \in \mathbb{C}$.

For the adjoint of $\mathcal{A}$ we have $W(\mathcal{A}^*) = \overline{W(\mathcal{A})}$.

We have $W(\mathcal{A}) = \{\lambda\}$ for $\lambda \in \mathbb{C}$ if and only if $A = \lambda I$ and $A_\#{}^T = -A_\#$. Moreover, $W(\mathcal{A}) \subset \mathbb{R}$ if and only if $A^* = A$ and $S = \frac{1}{2}(A_\# + A_\#{}^T) = 0$.

PROPOSITION 6.13. *Assume* $\mathcal{A} \in \mathcal{M}_n$. *Let* $T = \frac{1}{2}(A_\# - A_\#{}^T)$ *and split* $\mathcal{A}$ *as* $\mathcal{A} = \mathcal{B}+T\tau$. *Then* $W(\mathcal{A}) = W(\mathcal{B})$.

This leads to some peculiarities. For an illustration, define the numerical radius of $\mathcal{A}$ by $\nu(\mathcal{A}) = \max\{|\lambda|: \ \lambda \in W(\mathcal{A})\}$. Then $\nu(\mathcal{A}) = 0$ if and only if $A = 0$ and $A_\#{}^T = -A_\#$. We do have $\nu(\mathcal{A} + \mathcal{B}) \leq \nu(\mathcal{A}) + \nu(\mathcal{B})$ for $\mathcal{A}, \mathcal{B} \in \mathcal{M}_n$. Also, if $\mathcal{A}$ is normal, then $||\mathcal{A}|| = \nu(\mathcal{A})$.

PROPOSITION 6.14. *Let* $\mathcal{A} \in \mathcal{M}_n$. *Then* $\nu(\mathcal{A}) \leq ||\mathcal{A}||$. *Moreover, if* $A_\# = A_\#{}^T$, *then* $\frac{1}{2}||\mathcal{A}|| \leq \nu(\mathcal{A})$.

Proof. The first part is clear. For the second part,

$$\max_{||x||=1} ||\mathcal{A}x|| \leq \max_{||x||=1} \left|\left|\left(\frac{A+A^*}{2} + A_\#\tau\right)x\right|\right| + \max_{||x||=1} \left|\left|\left(\frac{A-A^*}{2}\right)x\right|\right|$$

which equals $\nu(\frac{A+A^*}{2} + A_\#\tau) + \frac{1}{2}\nu(A-A^*)$ since $A_\#{}^T = A_\#$. This is bounded by $\frac{1}{2}(\nu(A+A_\#\tau) + \nu(A^* + A_\#\tau) + \nu(A + A_\#\tau) + \nu(-A^* - A_\#\tau))$ so that the claim follows by the first part of the proposition. ∎

For another peculiarity, the numerical range need not be convex as the circlet τ on $\mathbb{C}$ illustrates. In view of this, define the polynomial numerical hull (see [12, p. 41]) of $\mathcal{A} \in \mathcal{M}_n$ of degree k by

$$V_k(\mathcal{A}) = \{\lambda \in \mathbb{C}: \ |p(\lambda)| \leq ||p(\mathcal{A})|| \text{ for all polynomials } p \text{ of degree } k \text{ at most}\}.$$

By mimicking the proof of [12, Theorem 2.10.5] we have the following analogy.

PROPOSITION 6.15. *Let* $\mathcal{A} \in \mathcal{M}_n$. *Then the convex hull of* $W(\mathcal{A})$ *equals* $V_1(\mathcal{A})$.

Let $\mathcal{A} \in \mathcal{M}_n$. Since for any $n \in \mathbb{N}$ we have $||\mathcal{A}^n|| = ||\psi(\mathcal{A})^n||$, the limit

$$(6.6) \qquad r(\mathcal{A}) = \lim_{n\to\infty} ||\mathcal{A}^n||^{1/n}$$

exists and equals the spectral radius of $\psi(\mathcal{A})$. In contrast, consider $\mathcal{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\tau$ of Example 3. For this we have $r(\mathcal{A}) = 1$ even though $\sigma(\mathcal{A}) = \emptyset$.

In terms of the resolvent operator of $\mathcal{A}$, if $\sum_{j=0}^{\infty} \left(\frac{1}{\lambda}\mathcal{A}\right)^j$ converges for $\lambda \in \mathbb{C}$, then

$$(6.7) \qquad \mathcal{R}(\lambda, \mathcal{A}) = \sum_{j=0}^{\infty} \left(\frac{1}{\lambda}\mathcal{A}\right)^j \frac{1}{\lambda}.$$

Therefore $\sigma(\mathcal{A}) \subset \gamma(\mathcal{A})$, where

$$\gamma(\mathcal{A}) = \left\{ \lambda \in \mathbb{C} : \ \lambda \neq 0, \ r\left(\frac{1}{\lambda}\mathcal{A}\right) \geq 1 \right\} \cup \{0\}.$$

This is a star-shaped and symmetric set with respect to the origin. Moreover, $\gamma(\mathcal{A})$ is preserved under an $\mathbb{R}$-linear similarity transformation even though the spectrum of $\mathcal{A}$ is not. This is because we have $r(\mathcal{S}\mathcal{A}\mathcal{S}^{-1}) = r(\mathcal{A})$ for any invertible $\mathcal{S} \in \mathcal{M}_n$.

EXAMPLE 7. Complex multiples of $\mathbb{R}$-linear operators behave very differently in taking powers. For an illustration, the circlet $t(i + \tau)$ is nilpotent for any $t \in \mathbb{R}$. However, $-i(i + \tau) = 1 - i\tau$ is not even power bounded.

7. The resolvent operator and functions of $\mathbb{R}$-linear operators. The resolvent operator $\mathcal{R}(\lambda, \mathcal{A})$ of $\mathcal{A} \in \mathcal{M}_n$ satisfies

$$(7.1) \qquad\qquad \lambda\mathcal{R}(\lambda, \mathcal{A}) - \mathcal{A}\mathcal{R}(\lambda, \mathcal{A}) = I$$

for λ outside $\sigma(\mathcal{A})$ such that for $|\lambda| > \|\mathcal{A}\|$ the series expansion (6.7) gives $\mathcal{R}(\lambda, \mathcal{A})$. The resolvent operator is real analytic but, in general, not analytic anywhere.

PROPOSITION 7.1. *Assume* $A + A_\# \tau \in \mathcal{M}_n$. *If for some* $\lambda_0 \in \mathbb{C}$ *the limit*

$$(7.2) \qquad\qquad \lim_{\lambda \to \lambda_0} \frac{1}{\lambda - \lambda_0} \left(\mathcal{R}(\lambda, \mathcal{A}) - \mathcal{R}(\lambda_0, \mathcal{A})\right)$$

exists, then $A_\# = 0$.

Proof. We have

$$\frac{1}{\lambda - \lambda_0} \left(\mathcal{R}(\lambda, \mathcal{A}) - \mathcal{R}(\lambda_0, \mathcal{A})\right) = \frac{-1}{\lambda - \lambda_0}\mathcal{R}(\lambda, \mathcal{A})(\lambda - \lambda_0)\mathcal{R}(\lambda_0, \mathcal{A}).$$

Using this for $\lambda - \lambda_0 \in \mathbb{R}$ gives $-\mathcal{R}(\lambda_0, \mathcal{A})^2$ for the limit (7.2). Combining this with taking the limit for $\lambda - \lambda_0 \in i\mathbb{R}$ forces the conlinear part $\mathcal{R}(\lambda_0, \mathcal{A})$ to equal zero. This implies that $A_\# = 0$. ∎

For an analytic function, consider $\mathcal{C}(\lambda, \mathcal{A})$ defined in (2.8). With this we have a solution to the Sylvester equation

$$(7.3) \qquad\qquad \lambda\mathcal{C}(\lambda, \mathcal{A}) - \mathcal{C}(\lambda, \mathcal{A})\mathcal{A} = I$$

in contrast to the equation (7.1) satisfied by the resolvent operator. We say that $\mathcal{C}(\lambda, \mathcal{A})$ is the *cosolvent* of $\mathcal{A}$.

DEFINITION 7.2. Let $\mathcal{A} \in \mathcal{M}_n$. The poles of the cosolvent of $\mathcal{A}$ are called the *proper values* of $\mathcal{A}$ and denoted by $\sigma_\mathcal{C}(\mathcal{A})$.

The set of proper values of $\mathcal{A}$ is nonempty and satisfies $\sigma_\mathcal{C}(\mathcal{A}) \subset \sigma(\psi(\mathcal{A}))$ being hence finite. To see this, we relate the resolvent and cosolvent by defining $H : \mathbb{C}^2 \to \mathbb{C}^{2n \times 2n}$ as

$$(7.4) \qquad\qquad (\lambda, \mu) \mapsto H(\lambda, \mu) = \begin{bmatrix} \lambda I - A & -A_\# \\ -A_\# & \mu I - \overline{A} \end{bmatrix}^{-1}.$$

Then by (5.4) we have the equality $\mathcal{R}(\lambda, \mathcal{A}) = 2G^*H(\lambda, \overline{\lambda})G$. For $\mathcal{C}(\lambda, \mathcal{A})$ we obtain $\mathcal{C}(\lambda, \mathcal{A}) = 2G^*H(\lambda, \lambda)G$ for λ outside $\sigma(\psi(\mathcal{A}))$. Consequently, $\sigma_\mathcal{C}(\mathcal{A}) \subset \sigma(\psi(\mathcal{A}))$.

If $S \in \mathbb{C}^{n \times n}$ is invertible, then $\sigma_\mathcal{C}(S\mathcal{A}S^{-1}) = \sigma_\mathcal{C}(\mathcal{A})$.

In (2.6) we defined polynomials in $\mathcal{A}$. A way to extend this to more general analytic functions is as follows. Assume $f(z) = \sum_{j=0}^{\infty} \alpha_j z^j$ converges for $|z| < R$ and $\mathcal{A} \in \mathcal{M}_n$ satisfies $r(\mathcal{A}) < R$. Then

$$f(\mathcal{A}) = \sum_{j=0}^{\infty} \alpha_j \mathcal{A}^j.$$

EXAMPLE 8. Functions of $\mathbb{R}$-linear operators must be handled with some care. To illustrate this aspect, let $f(z) = \frac{1}{\lambda - z}$ so that $R = |\lambda|$. Then for $r(\mathcal{A}) < R$ we have $f(\mathcal{A}) = \mathcal{C}(\lambda, \mathcal{A})$. On the other hand, take $g(z) = \lambda - z$. Then $g(\mathcal{A})^{-1} = \mathcal{R}(\lambda, \mathcal{A})$.

We still need more generality to deal with the case $r(\mathcal{A}) \geq R$. To this end we introduce a holomorphic functional calculus.

Recall that the scalar multiplication is defined from the left for the elements of $\mathcal{M}_{p,n}$. To emphasize this in holomorphic functional calculus, we write path integrals with $d\zeta$ followed by the integrand.

THEOREM 7.3. *Let $\mathcal{A} \in \mathcal{M}_n$ and assume f is analytic in a domain Ω containing $\sigma_{\mathcal{C}}(\mathcal{A})$. If Γ is a piecewise smooth contour surrounding $\sigma_{\mathcal{C}}(\mathcal{A})$ in Ω, then*

$$f(\mathcal{A}) = \frac{1}{2\pi i} \int_{\Gamma} d\zeta \, f(\zeta) \mathcal{C}(\zeta, \mathcal{A})$$

is well defined.

Proof. By (5.4) we have $\mathcal{A} = 2G^* \psi(\mathcal{A}) \mathcal{G}$. We first show that from this it follows that for any polynomial p

(7.5) $p(\mathcal{A}) = 2G^* p(\psi(\mathcal{A})) \mathcal{G}.$

To see this, observe that $\alpha \psi(\mathcal{A}) - \psi(\alpha \mathcal{A}) = (\alpha - \overline{\alpha}) \begin{bmatrix} 0 & 0 \\ A_\# & A \end{bmatrix}$ for any $\alpha \in \mathbb{C}$. Therefore $p(\psi(\mathcal{A})) = \psi(p(\mathcal{A})) + \begin{bmatrix} 0 & 0 \\ M & N \end{bmatrix}$ for $M, N \in \mathbb{C}^{n \times n}$. This in turn implies that $G^* p(\psi(\mathcal{A})) = G^* \psi(p(\mathcal{A}))$ and (7.5) follows.

Let $\widetilde{\Gamma}$ be a smooth contour surrounding $\sigma(\psi(\mathcal{A}))$ so that

(7.6) $p(\psi(\mathcal{A})) = \frac{1}{2\pi i} \int_{\widetilde{\Gamma}} d\zeta \, p(\zeta) \left(\zeta I - \psi(\mathcal{A}) \right)^{-1}.$

Using $\mathcal{C}(\zeta, \mathcal{A}) = 2G^* H(\zeta, \zeta) \mathcal{G}$ and $H(\zeta, \zeta) = \left(\zeta I - \psi(\mathcal{A}) \right)^{-1}$ with (7.5) and (7.6) yields

$$p(\mathcal{A}) = \frac{1}{2\pi i} \int_{\widetilde{\Gamma}} d\zeta \, p(\zeta) \mathcal{C}(\zeta, \mathcal{A}).$$

Since $\sigma_{\mathcal{C}}(\mathcal{A}) \subset \sigma(\psi(\mathcal{A}))$, the path $\widetilde{\Gamma}$ can be deformed to Γ. We can conclude that this functional calculus is consistent when applied to polynomials.

The general case follows by approximating f with polynomials. ∎

If we have $A_\# = 0$, then this yields the classical holomorphic functional calculus for matrices.

By deforming the path of integration, compute

(7.7) $\mathcal{K}_j = \frac{1}{2\pi i} \int_{\Gamma_j} d\zeta \, \mathcal{C}(\zeta, \mathcal{A}),$

where Γ_j surrounds a single proper value of $\mathcal{A}$. Taking the sum over all the proper values of $\mathcal{A}$, we have $\sum \mathcal{K}_j = I$ although $\mathcal{K}_j$ is not a projector and does not commute with $\mathcal{A}$ in general.

Observe that the holomorphic functional calculus proposed does not give rise to a homomorphism.

EXAMPLE 9. To see that we have $(fg)(\mathcal{A}) \neq f(\mathcal{A})g(\mathcal{A})$ in general, consider $\mathcal{A} = \tau$ on $\mathbb{C}$. With $f(z) = z$ and $g(z) = iz$ we obtain $(fg)(\mathcal{A}) = i$ while $f(\mathcal{A})g(\mathcal{A}) = -i$.

For another illustration of this aspect, consider the following two versions of the Laplace transformation of the exponential function.

PROPOSITION 7.4. *Let $\mathcal{A} \in \mathcal{M}_n$. Then*

$$\int_0^\infty e^{-\lambda t} e^{t\mathcal{A}} dt = \mathcal{C}(\lambda, \mathcal{A}) \quad and \quad \int_0^\infty e^{-t(\lambda I - \mathcal{A})} dt = \mathcal{R}(\lambda, \mathcal{A}).$$

Proof. For the first part, since $e^{t\mathcal{A}} = \sum_{j=0}^\infty \frac{t^j}{j!} \mathcal{A}^j$, multiply this by $e^{-\lambda t}$ from the left and integrate termwise.

For the second part, differentiate termwise $e^{t(\mathcal{A}-\lambda I)} = \sum_{j=0}^\infty \frac{t^j}{j!}(\mathcal{A}-\lambda I)^j$ to have $\frac{d}{dt} e^{t(\mathcal{A}-\lambda I)} = (\mathcal{A}-\lambda I)e^{t(\mathcal{A}-\lambda I)}$. The claim follows after integrating. ∎

In this connection, define the logarithmic norm

$$\mu(\mathcal{A}) = \max\{\mathrm{Re}\,\lambda : \lambda \in W(\mathcal{A})\}$$

of $\mathcal{A} \in \mathcal{M}_n$. We have $\mu(\mathcal{A}) = \mu(\psi(\mathcal{A}))$ and therefore $||e^{t\mathcal{A}}|| \leq e^{t\mu(\mathcal{A})}$.

To end this section, if f is analytic and $A_\# = 0$, then $\sigma(f(\mathcal{A})) = f(\sigma(\mathcal{A}))$ by the classical spectral mapping theorem for matrices. At the other extreme, assume $A = 0$ and consider $\mathcal{B}$ which is a polynomial in $\mathcal{A}$. Then

(7.8) $$\mathcal{B} = s(A_\# \overline{A_\#}) + A_\# q(\overline{A_\#} A_\#)\tau$$

for polynomials s and q. With this notation we have the following.

PROPOSITION 7.5 ([10]). *Assume $\lambda \in \sigma(A_\# \tau)$. Then $s(|\lambda|^2) + \lambda q(|\lambda|^2) \in \sigma(\mathcal{B})$, where $\mathcal{B}$ is defined in (7.8).*

In general the spectral mapping theorem breaks down for $\mathbb{R}$-linear operators.

EXAMPLE 10. Consider $\mathcal{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tau$ for which $\sigma(\mathcal{A}) = \emptyset$ by Example 3. We have $\sigma(\mathcal{A}^2) = \{-1\}$. We can conclude that even Proposition 7.5 is not exhaustive.

A polynomial in a circlet remains a circlet. Since the spectrum of a circlet is either a point or a circle, this imposes a restriction on allowable functions in a spectral mapping theorem. Recall that circles are mapped to circles by linear fractional transformations

(7.9) $$f(z) = \frac{\alpha - \beta z}{\gamma - \delta z} \quad \text{with} \quad \alpha, \beta, \gamma, \delta \in \mathbb{C}.$$

We obtain a spectral mapping theorem with this for $\mathcal{B} = (\alpha I - \beta \mathcal{A})(\gamma I - \delta \mathcal{A})^{-1}$.

THEOREM 7.6. *Let $\mathcal{A} \in \mathcal{M}_n$ and assume $\mathcal{B} = (\alpha I - \beta \mathcal{A})(\gamma I - \delta \mathcal{A})^{-1}$ is defined. Then*

$$\sigma(\mathcal{B}) = f(\sigma(\mathcal{A})),$$

where f is given by (7.9).

Proof. Assume an $\mathbb{R}$-linear operator $\mathcal{M} \in \mathcal{M}_n$ is invertible. Then for $x \neq 0$

$$\mathcal{M}x = \lambda x \;\Leftrightarrow\; \mathcal{M}^{-1}(\lambda x) = \frac{1}{\lambda}(\lambda x),$$

so that $\sigma(\mathcal{M}^{-1}) = \frac{1}{\sigma(\mathcal{M})}$. Using this and assuming $\delta \neq 0$ we obtain

$$\mathcal{B} = (\alpha I - \beta \mathcal{A})(\gamma I - \delta \mathcal{A})^{-1} = \frac{\beta}{\delta} I + \left(\alpha - \frac{\gamma\beta}{\delta} \right)(\gamma I - \delta \mathcal{A})^{-1}$$

so that $\sigma(\mathcal{B}) = \frac{\beta}{\delta} + \frac{\alpha - \gamma\beta/\delta}{\gamma - \delta\sigma(\mathcal{A})}$ proving the claim. For $\delta = 0$ the claim is true. ∎

8. Approximating $\mathbb{R}$-linear operators. Denote by $\mathbb{P}_j \subset \mathbb{C}^{n \times n}$ the set of orthogonal projectors of rank j at most. Finding the singular value decomposition of a matrix $A \in \mathbb{C}^{p \times n}$ can be regarded as equivalent to finding

$$(8.1) \qquad \min_{P \in \mathbb{P}_j} |||A(I - P)|||$$

for a unitarily invariant norm $|||\cdot|||$. Adapting this to a real linear operator $\mathcal{A} \in \mathcal{M}_{p,n}$, consider the problem of finding

$$(8.2) \qquad \min_{P \in \mathbb{P}_j} |||\mathcal{A}(I - P)|||.$$

Somewhat surprisingly, the approximations depend on the unitarily invariant norm used. Here we concentrate on the operator and the Frobenius norm.

Observe that if $P = WW^* \in \mathbb{P}_j$ with $W \in \mathbb{C}^{n \times j}$ having orthonormal columns, then $\mathcal{A}P = (U + V\tau)W^*$, where $U = AW$ and $V = A_{\#}\overline{W}$ are of type p-by-j. In other words, the problem (8.2) is equivalent to approximating $\mathcal{A}$ with the sum of j operets.

Consider first (8.2) in the Frobenius norm. Since the Frobenius norm of a matrix is preserved under taking the complex conjugate, we obtain for $P = WW^*$

$$(8.3) \qquad ||\mathcal{A}(I - P)||_F^2 = ||A - AWW^*||_F^2 + ||\overline{A_{\#}} - \overline{A_{\#}}WW^*||_F^2.$$

Therefore, to find $W \in \mathbb{C}^{n \times j}$ with orthonormal columns solving this problem, we can employ the singular value decomposition of the matrix $\left[\frac{A}{A_{\#}} \right] \in \mathbb{C}^{2p \times n}$.

DEFINITION 8.1. *The approximation numbers of $A + A_{\#}\tau \in \mathcal{M}_{p,n}$ with respect to the Frobenius norm are the singular values of the matrix $\left[\frac{A}{A_{\#}} \right] \in \mathbb{C}^{2p \times n}$.*

Equivalently, the approximation numbers with respect to the Frobenius equal the square roots of the eigenvalues of the $\mathbb{C}$-linear part of $\mathcal{A}^*\mathcal{A}$. We denote them by $\sigma_{F,j}(\mathcal{A})$.

We have $||\mathcal{A}||_F = (\sum_{j=1}^k \sigma_{F,j}(\mathcal{A})^2)^{1/2}$, where $k = \text{right-rank}(\mathcal{A})$.

Take the singular value decomposition $\left[\frac{A}{A_{\#}} \right] = U\Sigma W^*$ of $\left[\frac{A}{A_{\#}} \right]$. Then represent $\mathcal{A}$ as the sum of operets as

$$(8.4) \qquad \mathcal{A} = \sum_{l=1}^n (u_l + v_l\tau)w_l^*,$$

where $W = [w_1 \, w_2 \, \cdots \, w_n]$ so that $[u_1 \, u_2 \, \cdots \, u_n] = AW = [Aw_1 \, Aw_2 \, \cdots \, Aw_n]$ while $[v_1 \, v_2 \, \cdots \, v_n] = A_{\#}\overline{W} = [A_{\#}\overline{w_1} \, A_{\#}\overline{w_2} \, \cdots \, A_{\#}\overline{w_n}]$. From this we recover the approximation

numbers through $\sigma_{F,j}(\mathcal{A}) = (||u_j||^2 + ||v_j||^2)^{1/2}$. Moreover, the terms in this expansion are zero for $l >$ right-rank$(\mathcal{A})$.

If we restrict the operet $(u_j + v_j\tau)w_j^*$ to the subspace spanned by w_j, then its average norm equals $\sigma_{F,j}(\mathcal{A})$. Thus, with respect to this averaging, the representation (8.4) yields an optimally decreasing expansion of $\mathcal{A}$.

The approximation numbers with respect to the Frobenius norm are not preserved under taking the adjoint.

Consider (8.2) now in the operator norm. Then we call the values of (8.2) the approximation numbers of $\mathcal{A}$ with respect to the operator norm and denote them by $\sigma_j(\mathcal{A})$ in accordance with the classical notation.

PROPOSITION 8.2 ([8]). *Assume* $\mathcal{A} \in \mathcal{M}_{p,n}$ *with* $k =$ right-rank$(\mathcal{A})$. *If* $\mathcal{W} \in \mathcal{M}_p$ *and* $W \in \mathbb{C}^{n \times n}$ *are unitary, then*

$$\sigma_j(\mathcal{A}) = \sigma_j(\mathcal{W}^*\mathcal{A}W) \ \ and \ \ \sigma_{F,j}(\mathcal{A}) = \sigma_{F,j}(\mathcal{W}^*\mathcal{A}W)$$

for $j = 1, 2, \ldots, k$.

Finding $\sigma_j(\mathcal{A})$ appears to be a difficult problem. We do have the following bound.

THEOREM 8.3 ([8]). *Assume* $\mathcal{A} \in \mathcal{M}_{p,n}$ *with* $k =$ right-rank$(\mathcal{A})$. *Then*

$$\sigma_{F,j}(\mathcal{A}) \leq \sigma_j(\mathcal{A}) \leq \sqrt{2}\,\sigma_{F,j}(\mathcal{A}) \quad for \ j = 1, \ldots, k.$$

In some instances finding $\sigma_j(\mathcal{A})$ is easy.

COROLLARY 8.4. *Assume* $A^*A_{\#}$ *is skew-symmetric. Then*

$$\sigma_{F,j}(\mathcal{A}) = \sigma_j(\mathcal{A}) \quad for \ j = 1, \ldots, k.$$

Let $p = n$. The problem of diagonalizing a bilinear form under orthogonal substitutions resulted in the earliest versions of the singular value decomposition [6, Chapter 3]. For an analogy, assume $W_1, W_2 \in \mathbb{C}^{n \times n}$ are unitary and consider

$$(8.5) \qquad\qquad\qquad \mathcal{A} = W_1\mathcal{D}W_2^*$$

with a diagonal $\mathcal{D} \in \mathcal{M}_n$ having the diagonal entries $\alpha_j + \beta_j\tau$ for $1 \leq j \leq n$; see Theorem 4.7. Then $(|\alpha_j|^2 + |\beta_j|^2)^{1/2}$ are the approximation numbers of $\mathcal{A}$ with respect to the Frobenius norm, after arranging these numbers in non-increasing order. In this case solving (8.2) in the operator norm is also straightforward, i.e., to find V, choose those columns of W_2 that correspond to k largest values of $|\alpha_j| + |\beta_j|$.

References

[1]   T. Eirola, *Krylov integrators for Hamiltonian systems*, talk given at Workshop on Exponential Integrators (Innsbruck, 2004), http://techmath.uibk.ac.at/numbau/alex/events/conference2004.html.

[2]   T. Eirola, M. Huhtanen and J. von Pfaler, *Solution methods for $\mathbb{R}$-linear problems in $\mathbb{C}^n$*, SIAM J. Matrix Anal. Appl. 25 (2004), 804–828.

[3]   R. Freund, *On conjugate-gradient type methods for linear systems with complex-symmetric coefficient matrices*, SIAM J. Sci. Stat. Comput. 13 (1992), 425–448.

[4]   K. Friedrichs, *On certain inequalities for analytic functions and for functions of two variables*, Trans. Amer. Math. Soc. 1 (1937), 321–364.

[5]   Y. P. Hong and R. A. Horn, *On simultaneous reduction of families of matrices to triangular or diagonal form by unitary congruences*, Lin. and Multilin. Alg. 17 (1985), 271–288.

[6]   R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge Univ. Press, Cambridge, 1987.

[7]   M. Huhtanen, *Real linear Kronecker product operations*, Linear Algebra Appl. 418 (2006), 347–361.

[8]   M. Huhtanen and O. Nevanlinna, *Approximating real linear operators*, Studia Math.179 (2007), 7–25.

[9]   M. Huhtanen and O. Nevanlinna, *The real linear resolvent and cosolvent operators*, J. Operator Theory, to appear.

[10]   M. Huhtanen and J. von Pfaler, *The real linear eigenvalue problem in $\mathbb{C}^n$*, Linear Algebra Appl. 394 (2005), 169–199.

[11]   J. L. Mueller and S. Siltanen, *Direct reconstructions of conductivities from boundary measurements*, SIAM J. Sci. Comput. 24 (2003), 1232–1266.

[12]   O. Nevanlinna, *Convergence of Iterations for Linear Equations*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, 1993.

[13]   M. Putinar and H. S. Shapiro, *The Friedrichs operator of a planar domain*, in: Complex Analysis, Operators, and Related Topics, Oper. Theory Adv. Appl. 113, Birkhäuser, Basel, 2000, 303–330.

[14]   M. Putinar and H. S. Shapiro, *The Friedrichs operator of a planar domain. II*, in: Recent Advances in Operator Theory and Related Topics (Szeged, 1999), Oper. Theory Adv. Appl. 127, Birkhäuser, Basel, 2001, 519–551.