J. OPERATOR THEORY
58:2(2007), 229–250

THE REAL LINEAR RESOLVENT AND COSOLVENT OPERATORS

MARKO HUHTANEN and OLAVI NEVANLINNA

Communicated by Kenneth R. Davidson

ABSTRACT. For an $\mathbb{R}$-linear operator $\mathcal{A}$ the resolvent operator is defined outside the spectrum of $\mathcal{A}$ while the cosolvent operator is defined outside the proper values of $\mathcal{A}$. In this paper these two functions are studied. Series expansions are given. A new characterization for the eigenvalues of real matrices is obtained. The cosolvent operator is used to define and analyze analytic functions of $\mathcal{A}$. An application of this leads to a decomposition of $\mathbb{R}$-linear operators. Classes of structured $\mathbb{R}$-linear operators are considered.

KEYWORDS: *Eigenvalues, proper values, resolvent operator, cosolvent operator, minimal polynomial.*

MSC (2000): 47A10, 15A04.

INTRODUCTION

In real linear matrix analysis on $\mathbb{C}^n$ we are concerned with $\mathbb{R}$-linear operators defined as

$$x \longmapsto \mathcal{A}x = Ax + A_{\#}\overline{x} \tag{0.1}$$

for a pair of square matrices $A, A_{\#} \in \mathbb{C}^{n\times n}$. If $A_{\#} = 0$, then we have the ordinary matrix-vector product and the operator $\mathcal{A}$ is $\mathbb{C}$-linear, while with $A = 0$ we are dealing with a conlinear, i.e., antilinear operator. In this manner real linear matrix analysis extends classical matrix theory. For the background of this topic, its applications and computational tools it provides, see [3], [10], [8], [7], [2]. See also our recent survey paper [9].

In this paper we study functions associated with the eigenvalues and the proper values of $\mathcal{A}$. The set of eigenvalues of $\mathcal{A}$ consists of those $\lambda \in \mathbb{C}$ for which $\lambda I - \mathcal{A}$ is not invertible. The resolvent operator

$$\lambda \longmapsto \mathcal{R}(\lambda, \mathcal{A}) = (\lambda I - \mathcal{A})^{-1} \tag{0.2}$$

of $\mathcal{A}$, defined for λ outside the eigenvalues of $\mathcal{A}$, gives rise to a real analytic function. We derive various series expansions for the resolvent operator and make remarks on its growth properties. Classes of $\mathbb{R}$-linear operators are introduced for which the behavior of the resolvent operator can be regarded as understood, such as normal and circulant-like structures.

The minimal polynomial of $\mathcal{A}$ is defined to be the monic polynomial of the least degree annihilating $\mathcal{A}$. The set of proper values of $\mathcal{A}$ consists of the zeros of its minimal polynomial. Even though for matrices the eigenvalues and the proper values coincide, for $\mathbb{R}$-linear operators this does not hold in general. In fact, the set of eigenvalues of $\mathcal{A}$ is typically a continuum, and can even be empty, while the set of its proper values is always discrete and nonempty. The proper values of $\mathcal{A}$ determine the spectrum of the real matrix representation of $\mathcal{A}$. This interpreted conversely provides a new characterization for the eigenvalues of real matrices. The cosolvent operator of $\mathcal{A}$ is given by the analytic continuation of the series expansion

$$\lambda \longmapsto \mathcal{C}(\lambda, \mathcal{A}) = \sum_{k=0}^{\infty} \lambda^{-k-1} \mathcal{A}^k$$

which we show to be analytic exactly outside the proper values of $\mathcal{A}$. With the cosolvent operator functions of $\mathcal{A}$ can be defined and analyzed. Using this, a decomposition of $\mathcal{A}$ is introduced.

This paper is organized as follows. In Section 1, after a summary of basic facts from real linear matrix analysis, the $\mathbb{R}$-linear resolvent operator is considered starting from its various series expansions. In Section 2 classes of $\mathbb{R}$-linear operators are introduced in view of the relatively simple behavior of their resolvent and cosolvent operators. Section 3 deals first with polynomials in an $\mathbb{R}$-linear operator. With the minimal polynomial the proper values are characterized. A determinant-like scalar is obtained. To define more general functions, the cosolvent operator is applied. To operate with functions of $\mathbb{R}$-linear operators to vectors, in Section 4 some preliminary remarks on local aspects for $\mathbb{R}$-linear operators are made.

We remark that in this paper we are only concerned with $\mathbb{R}$-linear operators on $\mathbb{C}^n$ while $\mathbb{C}^n$ is regarded as representing a finite dimensional Hilbert space where the conjugation is defined. Infinite dimensional Hilbert spaces are not considered.

1. THE RESOLVENT OPERATOR

In $\mathbb{C}^n$, regarded as a vector space over $\mathbb{C}$, we use the standard inner product. We denote the set of $\mathbb{R}$-linear operators on $\mathbb{C}^n$ by $\mathcal{M}_n$. Defining the scalar multiplication from the left, $\mathcal{M}_n$ becomes a vector space over $\mathbb{C}$ of dimension $2n^2$ once the sum operation is defined in an obvious way. In view of the definition (0.1), we denote an $\mathbb{R}$-linear operator $\mathcal{A} \in \mathcal{M}_n$ interchangeably by $A + A_\# \tau$, where $\tau x = \overline{x}$ is the conjugation operator on $\mathbb{C}^n$.

The spectrum of $\mathcal{A} \in \mathcal{M}_n$, denoted by $\sigma(\mathcal{A})$, is the set of eigenvalues of $\mathcal{A}$, i.e., it consists of those points $\lambda \in \mathbb{C}$ for which $\lambda I - \mathcal{A}$ is not invertible. A nonzero vector $x \in \mathbb{C}^n$ with $\mathcal{A}x = \lambda x$ is called a respective eigenvector. The spectrum of $\mathcal{A}$ is an algebraic plane curve of degree $2n$ at most [3], [10]. For points λ outside $\sigma(\mathcal{A})$ the resolvent operator of $\mathcal{A}$ is defined by the formula (0.2).

In connection with the resolvent operator it might actually be natural to consider the entire family of linear fractional transformations of $\mathcal{A}$, i.e., those

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

with $\alpha, \beta \in \mathbb{C}$ and $\frac{\gamma}{\delta} \notin \sigma(\mathcal{A})$. This is because a spectral mapping theorem then holds [9]. Moreover, if $\lambda \in \sigma(\mathcal{A})$ and $\mathcal{A}x = \lambda x$, then $(\gamma - \delta\lambda)x$ is an eigenvector of $\mathcal{B}$ corresponding to the eigenvalue $\frac{\alpha-\beta\lambda}{\gamma-\delta\lambda}$. Hence, even though a complex multiple of x is no longer an eigenvector of $\mathcal{A}$ in general, it is still an eigenvector of a linear fractional transformed $\mathcal{A}$.

The operator norm of $\mathcal{A}$ is set as $\|\mathcal{A}\| = \max_{x\neq 0} \dfrac{\|\mathcal{A}x\|}{\|x\|}$, where $\|x\|$ denotes the 2-norm of a vector x. In addition to the straightforward bound $\|\mathcal{A}\| \leqslant \|A\| + \|A_\#\|$, let $\mathrm{S}\,(M) = \frac{1}{2}(M + M^T)$ denote the symmetric part of a matrix $M \in \mathbb{C}^{n\times n}$. Then there holds $\|\mathcal{A}\| \leqslant (\|A\|^2 + 2\|\mathrm{S}\,(A^* A_\#)\| + \|A_\#\|^2)^{1/2}$.

With these preliminaries, suppose $\mathcal{A}, \mathcal{B} \in \mathcal{M}_n$ and take $\lambda \in \mathbb{C}\backslash\mathbb{R}$. Because $(\lambda\mathcal{A})\mathcal{B} \neq \mathcal{A}(\lambda\mathcal{B})$ in general, $\mathcal{M}_n$ is not an algebra over $\mathbb{C}$. This defect has peculiar consequences such as the fact that Jacobson's lemma (see, e.g., [1]) fails to hold, i.e., if $\lambda I - \mathcal{A}\mathcal{B}$ is invertible, then $\lambda I - \mathcal{B}\mathcal{A}$ may not be invertible. The following, however, is proved in the classical way.

PROPOSITION 1.1. *Assume $\mathcal{A}, \mathcal{B} \in \mathcal{M}_n$. Then $\sigma(\mathcal{A}\mathcal{B}) \cap \mathbb{R} = \sigma(\mathcal{B}\mathcal{A}) \cap \mathbb{R}$ and if $\mathcal{B}$ is $\mathbb{C}$-linear, then $\sigma(\mathcal{A}\mathcal{B}) = \sigma(\mathcal{B}\mathcal{A})$.*

For the same reason other familiar identities must be handled with care in $\mathcal{M}_n$. In case of the resolvent identity, take μ and λ outside the spectrum of $\mathcal{A}$ to have

$$(\lambda I - \mathcal{A})^{-1} - (\mu I - \mathcal{A})^{-1} = (\lambda I - \mathcal{A})^{-1}(\mu - \lambda)(\mu I - \mathcal{A})^{-1}, \tag{1.1}$$

i.e., the scalar multiplication on the right must be performed between the resolvent operators. This yields

$$(\lambda I - \mathcal{A})^{-1}(\mu - \lambda)(\mu I - \mathcal{A})^{-1} = (\mu I - \mathcal{A})^{-1}(\mu - \lambda)(\lambda I - \mathcal{A})^{-1}$$

even though the resolvent operators do not commute.

Observe that the resolvent operator is not analytic.

PROPOSITION 1.2. *Assume $A + A_\#\tau \in \mathcal{M}_n$. If for some $\lambda_0 \in \mathbb{C}$ the following limit exists, then $A_\# = 0$:*

$$\lim_{\lambda\to\lambda_0} \frac{1}{\lambda - \lambda_0}(\mathcal{R}(\lambda, \mathcal{A}) - \mathcal{R}(\lambda_0, \mathcal{A})) \tag{1.2}$$

.

Proof. Denote $\mathcal{R}(\lambda_0, \mathcal{A}) = R + R_{\#}\tau$. Using the resolvent identity we have

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

Applying this for $\lambda - \lambda_0 \in \mathbb{R}$ we obtain $-\mathcal{R}(\lambda_0, \mathcal{A})^2$ for the limit (1.2). Combining this with taking the limit for $\lambda - \lambda_0 \in \mathrm{i}\mathbb{R}$ gives $R_{\#}\overline{R_{\#}} = 0$ and $R_{\#}\overline{R} = 0$. Then $\mathcal{R}(\lambda_0, \mathcal{A})^2 = R(R + R_{\#}\tau)$. Since this is invertible, also R must be invertible. Hence the conlinear part of $\mathcal{R}(\lambda_0, \mathcal{A})$ equals zero. This implies that $A_{\#} = 0$. ∎

For real analyticity of the resolvent operator, with the resolvent identity (1.1) we can compute partial derivatives. For the first order we have for $\lambda = s + \mathrm{i}t$

$$\frac{\partial}{\partial s}(\lambda I - \mathcal{A})^{-1}_{|\lambda=\lambda_0} = -(\lambda_0 I - \mathcal{A})^{-2} \text{ and } \frac{\partial}{\partial t}(\lambda I - \mathcal{A})^{-1}_{|\lambda=\lambda_0} = -(\lambda_0 I - \mathcal{A})^{-1}\mathrm{i}(\lambda_0 I - \mathcal{A})^{-1}.$$

Using these together with the product rule of differentiation, higher order partial derivatives follow readily. These yield us the Taylor series expansion

$$\mathcal{R}(s + \mathrm{i}t, \mathcal{A}) = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}(s - s_0)^j(t - t_0)^k\mathcal{A}_{j,k} \tag{1.3}$$

of the resolvent operator at $\lambda_0 = s_0 + \mathrm{i}t_0$ with the estimate $\|\mathcal{A}_{j,k}\| \leqslant \frac{(j+k)!}{j!k!}\|(\lambda_0 I - \mathcal{A})^{-1}\|^{j+k+1}$ so that $\max\{|s - s_0|, |t - t_0|\} < \frac{1}{2\|(\lambda_0 I - \mathcal{A})^{-1}\|}$ is a sufficient condition on convergence. As expected, the validity of the Taylor series expansion is linked with the growth properties of the resolvent operator.

Aside from the Taylor series (1.3), at infinity the resolvent operator can be expanded into the Neumann series as

$$\mathcal{R}(\lambda, \mathcal{A}) = \sum_{k=0}^{\infty}\left(\frac{1}{\lambda}\mathcal{A}\right)^k\frac{1}{\lambda} \tag{1.4}$$

being valid at least for $|\lambda| > \|\mathcal{A}\|$. Similarly, we have the second Neumann series

$$(\lambda I - \mathcal{A} - \mathcal{B})^{-1} = \sum_{k=0}^{\infty}((\lambda I - \mathcal{A})^{-1}\mathcal{B})^k(\lambda I - \mathcal{A})^{-1} \tag{1.5}$$

being valid at least for λ outside $\sigma(\mathcal{A})$ satisfying $\|(\lambda I - \mathcal{A})^{-1}\mathcal{B}\| < 1$. With this, fix λ outside the spectrum of $\mathcal{A}$ and assume $\mathcal{B} = -\mu \in \mathbb{C}$. Then

$$\lambda + \left\{\mu \in \mathbb{C} : \ \|(\lambda I - \mathcal{A})^{-1}\| < \frac{1}{|\mu|}\right\} \tag{1.6}$$

gives us a set in the complement of $\sigma(\mathcal{A})$. Hence for the points $\lambda + \mu$ in (1.6) we have yet another series expansion of the resolvent operator as the following, once $(\lambda I - \mathcal{A})^{-1}$ is known:

$$((\lambda + \mu)I - \mathcal{A})^{-1} = \sum_{k=0}^{\infty}((\lambda I - \mathcal{A})^{-1}\mu)^k(\lambda I - \mathcal{A})^{-1}.$$

For the growth of the norm of the resolvent operator we have an analogy of the classical estimate.

THEOREM 1.3. *Let $\mathcal{A} \in \mathcal{M}_n$. Then for any $\lambda \in \mathbb{C}$ we have*

$$\frac{1}{\operatorname{dist}(\lambda, \sigma(\mathcal{A}))} \leqslant \|(\lambda I - \mathcal{A})^{-1}\|. \tag{1.7}$$

Proof. We assume that $\sigma(\mathcal{A})$ is non-empty and that λ is not in the spectrum of $\mathcal{A}$ since otherwise the claim is true. Take $\mu \in \mathbb{C}$ such that $\mu + \lambda$ is the closest point in $\sigma(\mathcal{A})$ to λ. Then by (1.6) we must have $\|(\lambda I - \mathcal{A})^{-1}\| \geqslant \frac{1}{|\mu|}$. ∎

If the spectrum of $\mathcal{A}$ is empty, then it is understood that the left-hand side of the inequality (1.7) is zero.

DEFINITION 1.4. The adjoint of an $\mathbb{R}$-linear operator $\mathcal{A} = A + A_\# \tau$ is $\mathcal{A}^* = A^* + {A_\#}^T \tau$.

For every λ outside $\sigma(\mathcal{A})$ we have $\|(\lambda I - \mathcal{A})^{-1}\| = \|(\overline{\lambda} I - \mathcal{A}^*)^{-1}\|$ [3]. Therefore $\sigma(\mathcal{A}^*) = \overline{\sigma(\mathcal{A})}$.

If $\mathcal{A}^* = \mathcal{A}$, then $\mathcal{A}$ is said to be self-adjoint. If $\mathcal{A}^* \mathcal{A} = I$, then $\mathcal{A}$ is unitary.

PROPOSITION 1.5. *Let $A + A_\# \tau \in \mathcal{M}_n$ be self-adjoint with the real eigenvalues $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_k$. If $\lambda_{j+1} - \lambda_j > 2\|A_\#\|$, then $\sigma(\mathcal{A})$ is separated by a vertical line crossing the real axis at $\frac{\lambda_{j+1}+\lambda_j}{2}$.*

Proof. The field of values

$$F(\mathcal{A}) = \{x^* \mathcal{A} x \in \mathbb{C} : \text{ with } \|x\| = 1\} \tag{1.8}$$

of $\mathcal{A}$ contains $\sigma(\mathcal{A})$. With $\mathrm{i}\mathcal{A}$ we have the value $\mathrm{i}x^* A x + \mathrm{i}x^* A_\# \overline{x}$ such that $\mathrm{i}x^* A x$ is pure imaginary by the fact that A is Hermitian. Since $F(\mathrm{i}\mathcal{A}) = \mathrm{i}F(\mathcal{A})$, the imaginary part of any eigenvalue of $\mathcal{A}$ is bounded by $\|A_\#\|$.

With $r \in \mathbb{R}$ the operator norm of $\mathcal{R}(r, \mathcal{A})$ equals the reciprocal of the distance to the nearest real eigenvalue of $\mathcal{A}$ [10]. For $d = \frac{\lambda_{j+1}-\lambda_j}{2} > 0$ we thus have $\frac{1}{\operatorname{dist}(r, \sigma(\mathcal{A}))} \leqslant \frac{1}{d}$ by Theorem 1.3, where $r = \frac{\lambda_{j+1}+\lambda_j}{2}$. ∎

In the con-linear case all the conditions on the resolvent operator are necessarily circular symmetric with respect to the origin due to the following proposition.

PROPOSITION 1.6. *Let $\mathcal{A}$ be conlinear, i.e., $\mathcal{A} = A_\# \tau$ with $A_\# \in \mathbb{C}^{n \times n}$. Then $\lambda \longmapsto \|(\lambda I - \mathcal{A})^{-1}\|$ is circular symmetric with respect to the origin.*

Proof. Let $x \in \mathbb{C}^n$ be of unit length realizing $\displaystyle\min_{y \in \mathbb{C}^n, \|y\|=1} \|\lambda y - A_\# \overline{y}\|$ and suppose $\mu \in \mathbb{C}$ satisfies $|\mu| = |\lambda|$. Then for any $\theta \in \mathbb{R}$ we have $\|\mu e^{\mathrm{i}\theta} x - A_\# \overline{e^{\mathrm{i}\theta} x}\| = \|\mu e^{2\mathrm{i}\theta} x - A_\# \overline{x}\|$. Hence choosing θ such that $\mu e^{2\mathrm{i}\theta} = \lambda$ gives the same minimum. ∎

It is instructive to use the concanonical form of $A_{\#}$ to study the resolvent operator of $A_{\#}\tau$. For the concanonical form, recall that a $2k$-by-$2k$ quasi-Jordan block is

$$Q_{2k}(\mu) = \begin{bmatrix} 0 & I \\ J & 0 \end{bmatrix} = \begin{bmatrix} 0 & I \\ J_k(\mu) & 0 \end{bmatrix} \quad \text{with } \mu \in \mathbb{C}, \tag{1.9}$$

where $J_k(\mu)$ is a k-by-k Jordan block. A quasi-Jordan matrix is the direct sum of quasi-Jordan blocks. For any square matrix $A_{\#}$ we have $A_{\#} = S^{-1}(J \oplus Q)\overline{S}$ with a Jordan matrix J, a quasi-Jordan matrix Q and an invertible matrix S. This is the concanonical form of $A_{\#}$ of Hong and Horn [5]. Denote $D_{\#} = J \oplus Q$. Then

$$(\lambda I - A_{\#}\tau)^{-1} = S^{-1}\Big(\Big(\lambda I - \frac{1}{\overline{\lambda}} D_{\#}\overline{D_{\#}}\Big)^{-1} + (|\lambda|^2 D_{\#}{}^{-1} - \overline{D_{\#}})^{-1}\tau\Big)S$$

under sufficient assumptions on invertibility. Treating the parts of the resolvent operator separately, we have the bounds $\|R(\lambda)\| \leqslant \kappa(S)\big\|\big(\lambda I - \frac{1}{\overline{\lambda}}D_{\#}\overline{D_{\#}}\big)^{-1}\big\|$ as well as $\|R_{\#}(\lambda)\| \leqslant \kappa(S)\|(|\lambda|^2 D_{\#}{}^{-1} - \overline{D_{\#}})^{-1}\|$, where $\kappa(S) = \|S\|\|S^{-1}\|$.

In case the spectrum is empty, the lower bound (1.7) is vacuous. For a more versatile bound, independent of the size of the spectrum of $\mathcal{A}$, take $\lambda \in \mathbb{C}$ and a unit vector $x \in \mathbb{C}^n$ and set $c = \|\lambda x - \mathcal{A}x\|$. Then for any μ outside $\sigma(\mathcal{A})$ we have $1 = \|x\| = \|(\mu I - \mathcal{A})^{-1}(\mu I - \mathcal{A})x\| \leqslant \|(\mu I - \mathcal{A})^{-1}\|(|\lambda - \mu| + c)$, so that

$$\frac{1}{|\lambda - \mu| + c} \leqslant \|(\mu I - \mathcal{A})^{-1}\|. \tag{1.10}$$

In particular, if the spectrum is nonempty and λ is a nearest eigenvalue to μ, then we obtain (1.7) by picking x to be an eigenvector corresponding to λ. (And conversely, if a unit vector x is given, then to minimize c we should choose $\lambda = x^*\mathcal{A}x$.)

If $\mathcal{A}$ has an eigenvalue, then x and $\mathcal{A}x$ are parallel for any corresponding eigenvector x. Equivalently, we have $\big(I - \frac{xx^*}{x^*x}\big)\mathcal{A}x = 0$. This admits a generalization when we consider, for $x \in \mathbb{C}^n$ restricted to be of unit length, the local minima of the map

$$x \longmapsto \|((x^*\mathcal{A}x)I - \mathcal{A})x\|^2 = (\mathcal{A}x)^*\mathcal{A}x - |x^*\mathcal{A}x|^2 \tag{1.11}$$

on the unit sphere. We denote by $\delta(\mathcal{A})$ the set of points $\lambda = x^*\mathcal{A}x$, where $x \in \mathbb{C}^n$ is of unit length such that (1.11) attains a local minimum at x.

PROPOSITION 1.7. *Assume $\mathcal{A} \in \mathcal{M}_n$ and $\lambda = x^*\mathcal{A}x \in \delta(\mathcal{A})$. Then*

$$\|(\lambda I - \mathcal{A})x\|^{-1} = \|(\lambda I - \mathcal{A})^{-1}\|.$$

Proof. If $\lambda = x^*\mathcal{A}x \in \delta(\mathcal{A})$ is an eigenvalue of $\mathcal{A}$, then the claim is true. Hence suppose $x^*\mathcal{A}x$ is not an eigenvalue of $\mathcal{A}$. To show that the reciprocal of $\|(\lambda I - \mathcal{A})x\|$ equals $\|(\lambda I - \mathcal{A})^{-1}\|$, factor $\lambda I - \mathcal{A} = \mathcal{U}\mathcal{D}\mathcal{V}^*$ with unitary $\mathcal{U}, \mathcal{V} \in \mathcal{M}_n$ and $\mathcal{D} = D + D_{\#}\tau$ with diagonal matrices D and $D_{\#}$; see [3]. Hence we have

$\|(\lambda I - \mathcal{A})x\| = \|\mathcal{D}y\|$ with $y = \mathcal{V}^* x$. If

$$\|\mathcal{D}y\| > \|(\lambda I - \mathcal{A})^{-1}\| = \min_{z \in \mathbb{C}^n,\, \|z\|=1} \|\mathcal{D}z\|,$$

then choose y_ε of unit length arbitrarily close to y such that $\|\mathcal{D}y\| > \|\mathcal{D}y_\varepsilon\|$. This can be done since $\mathcal{D}$ is diagonal. Then with $x_\varepsilon = \mathcal{V}y_\varepsilon$ we have a vector of unit length arbitrarily close to x such that

$$\|\lambda x - \mathcal{A}x\| > \|\lambda x_\varepsilon - \mathcal{A}x_\varepsilon\| \geqslant \|(x_\varepsilon^* \mathcal{A} x_\varepsilon)x_\varepsilon - \mathcal{A}x_\varepsilon\|.$$

This is a contradiction and therefore $\|\mathcal{D}y\| = \frac{1}{\|\mathcal{D}^{-1}\|}$. ∎

We have $\delta(\mathcal{A}) \neq \emptyset$ by the fact that the map (1.11) is continuous and defined on a compact set. Moreover, there holds $\sigma(\mathcal{A}) \subset \delta(\mathcal{A}) \subset F(\mathcal{A})$, where $F(\mathcal{A})$ denotes the field of values of $\mathcal{A}$ defined in (1.8). Also $\delta(\lambda\mathcal{A}) = \lambda\delta(\mathcal{A})$ and $\delta(\lambda I + \mathcal{A}) = \lambda + \delta(\mathcal{A})$ for any $\lambda \in \mathbb{C}$.

EXAMPLE 1.8. Let $A_\# \in \mathbb{C}^{n\times n}$. Then $\delta(A_\#\tau)$ is circular symmetric with respect to the origin by the proof of Proposition 1.6. For $\mathcal{A} = \alpha I + A_\#\tau$ with $\alpha \in \mathbb{C}$ and ${A_\#}^T = -A_\#$ we have $\delta(\mathcal{A}) = \{\alpha\}$. To see that $\delta(\mathcal{A}) \neq \sigma(\mathcal{A})$ can hold, take $\mathcal{A} = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]\tau$ whose spectrum is empty.

An $\mathbb{R}$-linear operator $\mathcal{A}$ needs to be genuinely $\mathbb{R}$-linear for $\delta(\mathcal{A})$ to yield points outside the spectrum as follows.

THEOREM 1.9. *For $A \in \mathbb{C}^{n\times n}$ the map $\lambda \mapsto \|(\lambda I - A)^{-1}\|$ is not constant in an open set.*

Proof. Assume ζ is not in the spectrum of A and set $\widehat{A} = (\zeta I - A)^{-1}$. Then for $\triangle z = \triangle s + \mathrm{i}\triangle t$ small enough we have

$$M_{\triangle z} = ((\zeta + \triangle z)I - A)^{-1} = \sum_{k=0}^{\infty} (-1)^k \triangle z^k \widehat{A}^{k+1}.$$

Hence $M_{\triangle s} = \widehat{A}^* \widehat{A} + \triangle s(-\widehat{A}^* \widehat{A}^2 - \widehat{A}^{*2}\widehat{A} + \triangle s \cdots)$. If a unit vector x realizes the norm of $\widehat{A}$, we have

$$M_{\triangle s} x = \|\widehat{A}\|^2 x + \triangle s(-\widehat{A}^* \widehat{A}^2 - \widehat{A}^{*2}\widehat{A} + \triangle s \cdots)x. \tag{1.12}$$

If the norm of the resolvent is constant in a neighborhood of ζ, the last term in (1.12) must be identically zero. Hence $(\widehat{A}^* \widehat{A}^2 + \widehat{A}^{*2}\widehat{A})x = 0$. The same arguments with $M_{\mathrm{i}\triangle t}$ give $(\widehat{A}^* \widehat{A}^2 - \widehat{A}^{*2}\widehat{A})x = 0$. Summing these yield $\widehat{A}^* \widehat{A}^2 x = 0$, which is a contradiction since $\widehat{A}$ is invertible. ∎

Combining Proposition 1.7 with (1.12) we obtain the following corollary.

COROLLARY 1.10. *For $A \in \mathbb{C}^{n\times n}$ there holds $\delta(A) = \sigma(A)$.*

See Proposition 2.2 below for another family of $\mathbb{R}$-linear operators whose members $\mathcal{A}$ satisfy $\delta(\mathcal{A}) = \sigma(\mathcal{A})$.

With $\mathcal{A} = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right] \tau$ from Example 1.8 we can infer that the map

$$\lambda \longmapsto \log \|(\lambda I - \mathcal{A})^{-1}\|$$

need not be subharmonic (it has a local maximum at $\lambda = 0$ in this particular case). For subharmonicity in the $\mathbb{C}$-linear case, see [1].

2. NORMAL AND CIRCULANT-LIKE STRUCTURES

In what follows we give examples of $\mathbb{R}$-linear operators whose resolvent operator is readily computable. To this end, with $\alpha, \beta \in \mathbb{C}$, the $\mathbb{R}$-linear operator defined by

$$x \longmapsto \alpha x + \beta \overline{x} \tag{2.1}$$

is called a circlet and denoted by $\alpha + \beta\tau$. At times it is very useful to regard circlets as to extend scalars.

We have $(\lambda I - \alpha - \beta\tau)^{-1} = \frac{1}{|\lambda-\alpha|^2 - |\beta|^2}(\overline{\lambda} - \overline{\alpha} + \beta\tau)$ so that the spectrum of a circlet $\alpha + \beta\tau$ is the circle of radius $|\beta|$ centered at α. In particular, the set of circlets is a noncommutative ring. For the operator norm

$$\|(\lambda I - \alpha - \beta\tau)^{-1}\| = \frac{1}{|\lambda - \alpha| - |\beta|} \tag{2.2}$$

holds.

For more generality, an $\mathbb{R}$-linear operator $\mathcal{D} = D + D_{\#}\tau \in \mathcal{M}_n$ is said to be diagonal if both D and $D_{\#}$ are diagonal matrices. Equivalently, $\mathcal{D}$ is the direct sum of circlets on $\mathbb{C}$.

DEFINITION 2.1. $\mathcal{A} \in \mathcal{M}_n$ is normal if there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that $U^*\mathcal{A}U$ is diagonal.

If $\mathcal{A}$ is normal, then its spectrum consists of the circles corresponding to the spectra of the circlets on the diagonal of $U^*\mathcal{A}U$. In particular, then $\sigma(\mathcal{A}) \neq \emptyset$ so that by (2.2) we have the following result.

PROPOSITION 2.2. *If $\mathcal{A} \in \mathcal{M}_n$ is normal, then the equality holds in* (1.7).

Normal $\mathbb{R}$-linear operators can be identified as follows, where $H = \frac{1}{2}(A + A^*)$ and $K = \frac{1}{2\mathrm{i}}(A - A^*)$ denote the Hermitian and skew-Hermitian parts of $A \in \mathbb{C}^{n\times n}$.

THEOREM 2.3. *$\mathcal{A} \in \mathcal{M}_n$ is normal if and only if A is normal and $A_{\#}$, $HA_{\#}$ and $KA_{\#}$ are symmetric.*

Proof. Use Corollary 5.3 of [4] together with the fact that due to normality the matrices H and K commute to deduce that there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that $U^*\mathcal{A}U$ is diagonal. ∎

An algorithm to diagonalize a normal $\mathbb{R}$-linear operator $\mathcal{A}$ is obtain by applying plane rotations, as in Jacobi's method for normal matrices, to annihilate off-diagonal elements of A and $A_\#$.

Equality can hold in (1.7) for every λ even if $\mathcal{A}$ is not normal.

EXAMPLE 2.4. Let $\mathcal{A} : \mathbb{C}^3 \to \mathbb{C}^3$ be the direct sum of $\left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right] \tau$ on $\mathbb{C}^2$ and 0 on $\mathbb{C}$. Hence the resolvent operator $(\lambda I - \mathcal{A})^{-1}$ is the direct sum of

$$\frac{1}{|\lambda|^2+1}\left(\left[\begin{smallmatrix} \overline{\lambda} & 0 \\ 0 & \lambda \end{smallmatrix}\right] + \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right] \tau \right) \quad \text{on } \mathbb{C}^2 \quad \text{whose norm is } \frac{1}{\sqrt{|\lambda|^2+1}}, \tag{2.3}$$

and $\frac{1}{\lambda}$ on $\mathbb{C}$ whose norm is $\frac{1}{|\lambda|}$. Thus the norm of the resolvent of $\mathcal{A}$ equals the reciprocal of the distance to the spectrum even though $\mathcal{A}$ is not normal.

For a canonical family of normal $\mathbb{R}$-linear operators, in what follows we denote by $P \in \mathbb{C}^{n\times n}$ the "backward identity", i.e., the permutation matrix with ones on the diagonal joining the left lower corner with the right upper corner. By $\mathtt{C}$ we denote the set of circulant matrices.

DEFINITION 2.5. Let P be the backward identity. Then $A \in \mathbb{C}^{n\times n}$ is called a circulant-Hankel matrix if $A = PC$ for $C \in \mathtt{C}$.

In other words, a circulant-Hankel matrix has cyclically appearing antidiagonals. We denote such matrices by $P\mathtt{C}$.

EXAMPLE 2.6. If $A \in \mathtt{C}$ and $A_\# \in P\mathtt{C}$, then $\mathcal{A}$ is normal and can be diagonalized with $U = \frac{1}{\sqrt{n}} F_n$, where F_n is the Fourier matrix [10].

We do not know if the spectrum of a nonnormal $A + A_\#\tau \in \mathcal{M}_n$ can be empty in case $A_\#$ is symmetric while no restrictions are set on A. In the dimensions 1 and 2 we have the following suggestive result.

PROPOSITION 2.7. *Assume $\mathcal{A} \in \mathcal{M}_n$, with $n = 1$ or 2, has a symmetric antilinear part $A_\#$. Then $\sigma(\mathcal{A}) \neq \emptyset$.*

Proof. The case $n = 1$ is clear so let us consider $n = 2$. After a possible similarity transformation with a unitary $U \in \mathbb{C}^{2\times 2}$, we can assume the antilinear part of $\mathcal{A}$ to be diagonal, say $A_\# = \operatorname{diag}(d_1, d_2)$. Consider $(\lambda I - A)x - A_\#\overline{x} = 0$. We can assume $a_{12} \neq 0$ since otherwise the spectrum consists of circles and is thereby non-empty. Solving x_2 from the first equation gives $x_2 = \frac{1}{a_{12}}((\lambda - a_{11})x_1 - d_1\overline{x_1})$. Substituting this into the second equation yields

$$(-a_{21}a_{12} + (\lambda - a_{22})(\lambda - a_{11}) - d_1\overline{d_2})x_1 - (d_1(\lambda - a_{22}) + d_2(\overline{\lambda} - \overline{a_{11}}))\overline{x_1} = 0.$$

For some $\lambda \in \mathbb{C}$ this must have a non-zero solution since a scalar real linear equation $ax_1 - b\overline{x_1} = 0$ has a non-zero solution x_1 if and only if $|a| = |b|$. To see this, observe that $-a_{21}a_{12} + (\lambda - a_{22})(\lambda - a_{11}) - d_1\overline{d_2}$ is a polynomial, so it has a zero. Moreover, it is of degree two while $d_1(\lambda - a_{22}) + d_2(\overline{\lambda} - \overline{a_{11}})$ is of degree one so that it grows faster in absolute value as $|\lambda| \to \infty$. Therefore, being continuous, they must equal in modulus at some point $\lambda \in \mathbb{C}$. ∎

The proof is valid also in the case there exists an invertible $S \in \mathbb{C}^{2\times 2}$ such that $SA_{\#}\overline{S^{-1}}$ is diagonal.

The case $n = 2$ is of interest because then the eigenvalue problem can still be solved by hand as follows. With a unitary $U \in \mathbb{C}^{2\times 2}$ let $N = UAU^*$ be the Schur decomposition of A. Take $\mathcal{N} = U\mathcal{A}U^*$ and look at the eigenvalue problem

$$\begin{bmatrix} n_{11} & n_{12} \\ 0 & n_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} n_{\#11} & n_{\#12} \\ n_{\#21} & n_{\#22} \end{bmatrix} \begin{bmatrix} \overline{x_1} \\ \overline{x_2} \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

We assume $n_{\#21} \neq 0$ since otherwise the spectrum is known (Theorem 2.6 of [3]). Without loss of generality, we may assume $n_{11} = 0$ (after translating the problem by $n_{11}I$). Scale the eigenvector so that x_2 is of unit modulus, i.e., $x_2 = e^{i\theta_2}$ with $\theta_2 \in [0, 2\pi)$. Then solve x_1 from the second equation in terms of x_2 and λ insert it in the first equation. As a result, we end up having an equation of the form

$$|\lambda|^2 + b\lambda + c = 0 \quad \text{with } b = b(\theta_2), c = c(\theta_2) \in \mathbb{C}. \tag{2.4}$$

From this it is a simple task to find λ in terms of θ_2 (and all those $\theta_2 \in [0, 2\pi)$ for which there exists a solution λ give rise to the corresponding eigenvector x with $x_2 = e^{i\theta_2}$).

In view of these two dimensional manipulations, define a more general circulant-like structure as follows.

DEFINITION 2.8. $\mathtt{C} + P\mathtt{C} \subset \mathbb{C}^{n\times n}$ consists of the matrices representable as the sum of a circulant matrix and a circulant-Hankel matrix.

Observe that $\mathtt{C} + P\mathtt{C}$ is an algebra over $\mathbb{C}$.

PROPOSITION 2.9. *Let $A, A_{\#} \in \mathtt{C} + P\mathtt{C}$. Then $A + A_{\#}\tau$ is unitarily similar to an $\mathbb{R}$-linear operator with blocks of size 2-by-2 at most located at the respective positions.*

Proof. Take $U = \frac{1}{\sqrt{n}}F_n$, where F_n is the Fourier matrix. Then $U^*AU = D_1 + D_2 \oplus \widetilde{D}_1$, where D_1 is a diagonal matrix while D_2 is a scalar and $\widetilde{D}_1 \in \mathbb{C}^{(n-1)\times(n-1)}$ is antidiagonal, i.e., non-zero elements appear only on the diagonal joining the left lower corner with the right upper corner. Similarly we obtain $U^*A_{\#}\overline{U} = D_3 + D_4 \oplus \widetilde{D}_2$, where D_3 is a diagonal matrix while D_2 is a scalar and $\widetilde{D}_2 \in \mathbb{C}^{(n-1)\times(n-1)}$ is antidiagonal. Denote by e_j the standard basis vectors of $\mathbb{C}^n$. For n even, span$\{e_1, e_{\frac{n}{2}+1}\}$ and span$\{e_{j+1}, e_{n-j+1}\}$, for $j = 1, \ldots, \frac{n-2}{2}$, are mutually orthogonal invariant subspaces of $U^*\mathcal{A}U$. For n odd, span$\{e_1\}$ and span$\{e_{j+1}, e_{n-j+1}\}$, for $j = 1, \ldots, \frac{n-1}{2}$, are mutually orthogonal invariant subspaces of $U^*\mathcal{A}U$. The claim follows after applying the corresponding permutation similarity transformation. ∎

By the reasoning that led to the equation (2.4), a closed form solution to the spectrum can be given for $\mathbb{R}$-linear operators characterized by this proposition.

In this sense the eigenvalue problem can be regarded as understood for $\mathcal{A} \in \mathcal{M}_n$ with the parts from $\mathtt{C} + P\mathtt{C}$.

For another instance where (2.4) is of use, consider the following structure proposed in [8].

EXAMPLE 2.10. Assume the right-rank of $\mathcal{A}$ is at most j, i.e., $\mathcal{A}$ is representable as $\mathcal{A} = (U + V\tau)W^*$ with $U, V, W \in \mathbb{C}^{n\times j}$. Then for $\mathcal{A}^*$ this forces any of its eigenvectors to belong to the span of the columns of W. Hence, since this is an invariant subspace of $\mathcal{A}^*$, for $j = 2$ we can consider the restriction of $\mathcal{A}^*$ to it and find a closed form solution to the spectrum.

The same holds if the left-rank of $\mathcal{A}$ is at most two, i.e., we have $\mathcal{A} = W(U^* + V^T\tau)$ with $U, V, W \in \mathbb{C}^{n\times 2}$.

To end this section, observe that an $\mathbb{R}$-linear operator $\mathcal{A}$ can be split uniquely into the so-called real and imaginary parts as $\mathcal{A} = \mathcal{R} + \mathrm{i}\mathcal{I}$, where

$$\mathcal{R} = \frac{1}{2}(A + \overline{A_\#} + (A_\# + \overline{A})\tau) \quad \text{and} \quad \mathcal{I} = \frac{1}{2\mathrm{i}}(A - \overline{A_\#} + (A_\# - \overline{A})\tau),$$

i.e., the ranges of $\mathcal{R}$ and $\mathcal{I}$ belong to $\mathbb{R}^n$. For such parts (or $\mathbb{C}$-linearly similar to such parts) we have the following way of locating the spectrum.

PROPOSITION 2.11. *$\lambda = r\mathrm{e}^{\mathrm{i}\theta}$ is an eigenvalue of $M + \overline{M}\tau \in \mathcal{M}_n$ if and only if $r \in \sigma(\mathrm{e}^{-\mathrm{i}\theta}M + \mathrm{e}^{\mathrm{i}\theta}\overline{M})$.*

Proof. Assume $\mathcal{M}x = \lambda x$ with $\lambda = r\mathrm{e}^{\mathrm{i}\theta}$. Since the left-hand side belongs to $\mathbb{R}^n$, we must have $x = \mathrm{e}^{-\mathrm{i}\theta}y$ with $y \in \mathbb{R}^n$. Hence we obtain an equivalent standard eigenvalue problem $(\mathrm{e}^{-\mathrm{i}\theta}M + \mathrm{e}^{\mathrm{i}\theta}\overline{M})y = ry$. ∎

3. THE COSOLVENT OPERATOR AND FUNCTIONS OF $\mathbb{R}$-LINEAR OPERATORS

The eigenvalues of a matrix are determined by the zeros of its minimal polynomial. In this section we develop the respective structure in $\mathcal{M}_n$. In this case the role of the zeros of the minimal polynomial is taken by the so-called proper values. The resolvent operator is replaced with the cosolvent operator allowing us to define analytic functions of $\mathbb{R}$-linear operators.

For a tool needed in what follows, define $\psi : \mathcal{M}_n \to \mathbb{C}^{2n\times 2n}$ by

$$\psi(\mathcal{A}) = \begin{bmatrix} A & A_\# \\ \overline{A_\#} & \overline{A} \end{bmatrix}. \tag{3.1}$$

We call $\psi(\mathcal{A})$ the complex matrix representation of $\mathcal{A}$. The function ψ is injective and isometric and satisfies $\psi(\mathcal{A}+\mathcal{B}) = \psi(\mathcal{A}) + \psi(\mathcal{B})$ and $\psi(\mathcal{AB}) = \psi(\mathcal{A})\psi(\mathcal{B})$ for $\mathcal{A}, \mathcal{B} \in \mathcal{M}_n$. With the unitary matrix $E = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} I & \mathrm{i}I \\ I & -\mathrm{i}I \end{smallmatrix}\right]$ we obtain the so-called real matrix representation

$$\phi(\mathcal{A}) = E^*\psi(\mathcal{A})E = \begin{bmatrix} \mathrm{Re}(A + A_\#) & \mathrm{Im}(-A + A_\#) \\ \mathrm{Im}(A + A_\#) & \mathrm{Re}(A - A_\#) \end{bmatrix} \tag{3.2}$$

of $\mathcal{A}$. Then $\phi : \mathcal{M}_n \to \mathbb{R}^{2n\times 2n}$ is bijective and has the same properties as ψ. From this we can conclude that $\sigma(\psi(\mathcal{A})) = \sigma(\phi(\mathcal{A}))$ consists of at most $2n$ points and is symmetrically located with respect to the real axis.

Let now $q(z) = \sum_{j=0}^{k} c_j z^j$ be a polynomial with the coefficients $c_j = s_j + \mathrm{i}t_j \in \mathbb{C}$. Then the polynomial q in $\mathcal{A} \in \mathcal{M}_n$ is defined to be the $\mathbb{R}$-linear operator $q(\mathcal{A}) = \sum_{j=0}^{k} c_j \mathcal{A}^j$. The set of polynomials in $\mathcal{A}$, which we denote by $\mathbf{P}(\mathcal{A})$, is a subspace of $\mathcal{M}_n$ of dimension $2n$ at most. It is invariant under multiplications from the right by polynomials in $\mathcal{A}$ with real coefficients. In fact, if $q = q_1 q_2$ for two polynomials q_1 and q_2 such that q_2 has real coefficients, then $q(\mathcal{A}) = q_1(\mathcal{A})q_2(\mathcal{A})$.

If $\mathcal{A} = \mathcal{S}\mathcal{B}\mathcal{S}^{-1}$ for an invertible $\mathcal{S} \in \mathcal{M}_n$, then we have

$$q(\mathcal{A}) = \mathcal{S}\Big(\sum_{j=0}^{k}(s_j + \mathcal{S}^{-1}\mathrm{i}\mathcal{S}t_j)\mathcal{B}^j\Big)\mathcal{S}^{-1}. \tag{3.3}$$

In particular, $q(\mathcal{A}) = \mathcal{S}q(\mathcal{B})\mathcal{S}^{-1}$ in case q has real coefficients, or $\mathcal{S}$ is $\mathbb{C}$-linear. If $\mathcal{S}$ is conlinear, then $q(\mathcal{A}) = \mathcal{S}q^*(\mathcal{B})\mathcal{S}^{-1}$, where $q^*(z) = \overline{q(\overline{z})}$.

DEFINITION 3.1. The *minimal polynomial* of $\mathcal{A} \in \mathcal{M}_n$ is the monic polynomial q of the least degree annihilating $\mathcal{A}$. The zeros of q are called the proper values of $\mathcal{A}$.

The uniqueness of the minimal polynomial follows by the reasoning analogous to the $\mathbb{C}$-linear case. The degree of the minimal polynomial of $\mathcal{A} \in \mathcal{M}_n$ is bounded by $2n$ and if $A_\# = 0$, then by n [3]. In particular, the dimension of $\mathbf{P}(\mathcal{A})$ equals the degree of the minimal polynomial of $\mathcal{A}$. If $r \in \mathbb{R}$, then the degrees of the minimal polynomial of $\mathcal{A}$ and $rI + \mathcal{A}$ coincide.

Assume q is the minimal polynomial of $\mathcal{A} \in \mathcal{M}_n$. Then $\mathcal{A}$ is said to be algebraic of degree $\deg(q)$. The degree of $\mathcal{A}$ is denoted by $\deg(\mathcal{A})$. By (3.3), for an invertible $\mathbb{C}$-linear $\mathcal{S}$ we have $\deg(\mathcal{S}^{-1}\mathcal{A}\mathcal{S}) = \deg(\mathcal{A})$.

PROPOSITION 3.2. *Let $\mathcal{A} \in \mathcal{M}_n$. If the left-rank of $\mathcal{A}$ is j, then $\mathcal{A}$ is algebraic of degree $2j + 1$ at most.*

Proof. By Example 2.10, we may assume $\mathcal{A} = W(U^* + V^T\tau)$ with $U, V, W \in \mathbb{C}^{n\times j}$. Let $\mathcal{A}_W$ denote the restriction of $\mathcal{A}$ to the span of the columns of W which is an invariant subspace of $\mathcal{A}$. Then $p(\mathcal{A}_W) = 0$ for a polynomial of degree $2j$ at most. But then also $p(\mathcal{A})\mathcal{A} = 0$, since the range of $\mathcal{A}$ is in the span of the columns of W. Since the polynomial $q_2(z) = z$ has real coefficients, the polynomial pq_2 annihilates $\mathcal{A}$. ∎

PROPOSITION 3.3. *Let $A \in \mathbb{C}^{n\times n}$. Then $A + \overline{A}\tau \in \mathcal{M}_n$ is algebraic of degree $n+1$ at most.*

Proof. The restriction of $\mathcal{A}$ to $\mathbb{R}^n$ gives rise to an $\mathbb{R}$-linear operator with the minimal polynomial p of degree n at most. Since the range of $\mathcal{A}$ belongs to $\mathbb{R}^n$, we have $p(\mathcal{A})\mathcal{A} = 0$ and the polynomial $p(z)z$ thus annihilates $\mathcal{A}$. ∎

Observe that we can have a difference in taking the adjoint.

EXAMPLE 3.4. The degrees of the minimal polynomial of an $\mathbb{R}$-linear operator and its adjoint can differ. To see this, take a so-called operet $\mathcal{O} = (u + v\tau)w^*$ such that $d = w^*u \notin \mathbb{R}$ and $w^*v = 0$ with $v \neq 0$. Then $\mathcal{O}^2 = \mathcal{O}d$, i.e., we have $\mathcal{O}^{*2} - \overline{d}\mathcal{O}^* = 0$ so that $p(z) = z^2 - \overline{d}z$ is the minimal polynomial of $\mathcal{O}^*$. The minimal polynomial of $\mathcal{O}$ is $q(z) = z^3 - (d + \overline{d})z^2 + |d|^2 z$. Therefore also $\dim \mathbf{P}(\mathcal{O}) > \dim \mathbf{P}(\mathcal{O}^*)$. Observe though that p and q have two common zeros. This is not an accident; see Theorem 3.13.

Example 3.4 also illustrates that it is of interest to set the “right” polynomial in $\mathcal{A}$ by performing the scalar multiplication from the right.

DEFINITION 3.5. Let $q(z) = \sum_{j=0}^{k} c_j z^j$. Then the respective right polynomial in $\mathcal{A} \in \mathcal{M}_n$ is the $\mathbb{R}$-linear operator $q_r(\mathcal{A}) = \sum_{j=0}^{k} \mathcal{A}^j c_j$.

In general $q(\mathcal{A}) \neq q_r(\mathcal{A})$. By taking the adjoint of a polynomial in $\mathcal{A}$ we obviously obtain a right polynomial in $\mathcal{A}^*$. The set of right polynomials in $\mathcal{A}$ is a linear manifold over $\mathbb{R}$ of $\mathcal{M}_n$. Observe though that if the scalar multiplication in $\mathcal{M}_n$ had been defined from the right, then the set of right polynomials in $\mathcal{A}$ would be a subspace of $\mathcal{M}_n$ such that for $\mathcal{A}, \mathcal{B} \in \mathcal{M}_n$

$$(\mathcal{A}, \mathcal{B}) = \operatorname{tr}(B^*A + {B_\#}^T \overline{A_\#})$$

would give us an inner product on $\mathcal{M}_n$. Then computing the first linearly dependent element in the sequence $\{I, \mathcal{A}^*, \mathcal{A}^{2*}, \ldots\}$, for instance, by using the Gram-Schmidt process with respect to this inner product would yield us the minimal polynomial of $\mathcal{A}$, after taking the adjoint.

With the minimal polynomial we have a determinant-like scalar as follows.

PROPOSITION 3.6. *Let $q(z) = z^k + \sum_{j=0}^{k-1} c_j z^j$ be the minimal polynomial of $\mathcal{A} \in \mathcal{M}_n$. Then $\mathcal{A}$ is invertible if and only if $c_0 \neq 0$.*

Proof. Assume that $\mathcal{A}$ is invertible. If $c_0 = 0$, then $\left(\mathcal{A}^{k-1} + \sum_{j=1}^{k-1} c_j \mathcal{A}^{j-1}\right)\mathcal{A} = 0$. Hence applying $\mathcal{A}^{-1}$ from the right shows that q was not the minimal polynomial of $\mathcal{A}$, which is a contradiction. Therefore $c_0 \neq 0$.

Assume $\mathcal{A}$ is not invertible. If $c_0 \neq 0$, then $\mathcal{A}^k + \sum_{j=1}^{k-1} c_j \mathcal{A}^j = c_0 I$. Thus, $\frac{1}{c_0}\Big(\mathcal{A}^{k-1} + \sum_{j=1}^{k-1} c_j \mathcal{A}^{j-1}\Big)\mathcal{A} = I$, which contradicts the assumption that $\mathcal{A}$ is not invertible. Hence $c_0 = 0$. ∎

If $S \in \mathbb{C}^{n\times n}$ is invertible, then $\mathcal{A}$ and $S\mathcal{A}S^{-1}$ have the same minimal polynomial and hence the same c_0, i.e., c_0 is preserved in $\mathbb{C}$-linear similarity transformations.

Let $q(z) = z^k + \sum_{j=0}^{k-1} c_j z^j$ be the minimal polynomial of an invertible $\mathcal{A} \in \mathcal{M}_n$. Because then $c_0 \neq 0$, applying $\mathcal{A}^{-1}$ from the right to the identity $q(\mathcal{A}) = 0$ gives explicitly the inverse of $\mathcal{A}$ as a polynomial in $\mathcal{A}$ of degree $k-1$.

For more general functions of $\mathbb{R}$-linear operators, consider the Sylvester equation

$$\lambda \mathcal{C}(\lambda,\mathcal{A}) - \mathcal{C}(\lambda,\mathcal{A})\mathcal{A} = I. \tag{3.4}$$

whose solution $\mathcal{C}(\lambda,\mathcal{A})$ we call the cosolvent operator of $\mathcal{A}$. For $|\lambda| > \|\mathcal{A}\|$ we have the expansion $\mathcal{C}(\lambda,\mathcal{A}) = \sum_{k=0}^{\infty} \lambda^{-k-1}\mathcal{A}^k$ and otherwise extend this series by analytic continuation. Hence we obtain an analytic function outside its poles. The poles of the cosolvent operator of $\mathcal{A}$ are denoted by $\sigma_{\mathcal{C}}(\mathcal{A})$. In particular, if $\mathcal{A}$ is $\mathbb{C}$-linear, then its cosolvent operator coincides with its resolvent operator. In such a case $\sigma_{\mathcal{C}}(\mathcal{A})$ thus consist of at most n complex numbers.

Before inspecting properties of the cosolvent operator, we give two further examples. To this end, if $S \in \mathbb{C}^{n\times n}$ is invertible, then

$$\mathcal{C}(\lambda, S\mathcal{A}S^{-1}) = S\mathcal{C}(\lambda,\mathcal{A})S^{-1}. \tag{3.5}$$

EXAMPLE 3.7. For a circlet $\alpha + \beta\tau$ the cosolvent is

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

having singularities at $\lambda = \operatorname{Re}\alpha \pm \sqrt{|\beta|^2 - (\operatorname{Im}\alpha)^2}$. Consequently, for a normal $\mathcal{A} \in \mathcal{M}_n$ we can use this blockwise with (3.5) to find its cosolvent operator.

The above formula for the cosolvent operator of a circlet can also be used in the upper (equivalently lower) triangular case to locate $\sigma_{\mathcal{C}}(\mathcal{A})$. Recall that $A + A_{\#}\tau$ is said to be upper triangular if A and $A_{\#}$ are upper triangular matrices. If $S\mathcal{A}S^{-1}$ is upper triangular with an invertible $S \in \mathbb{C}^{n\times n}$, then the poles of the cosolvents on the diagonal of $S\mathcal{A}S^{-1}$ determine $\sigma_{\mathcal{C}}(\mathcal{A})$.

With these remarks, the conlinear case can be handled as follows.

EXAMPLE 3.8. Let $A_{\#} = S^{-1}(J \oplus Q)\overline{S}$ be the concanonical form of $A_{\#}$. Since for a Jordan matrix J we know $\sigma_{\mathcal{C}}(J\tau)$, to have $\sigma_{\mathcal{C}}(A_{\#}\tau)$ it suffices to find $\sigma_{\mathcal{C}}(Q_{2k}\tau)$

for a quasi-Jordan block $Q_{2k} = Q_{2k}(\mu) = \begin{bmatrix} 0 & I \\ J_k(\mu) & 0 \end{bmatrix} = \begin{bmatrix} 0 & I \\ J & 0 \end{bmatrix}$. The eigenvalues of Q_{2k} are $\mu^{1/2}$. Knowing this we obtain

$$\mathcal{C}(\lambda, Q_{2k}\tau) = \begin{bmatrix} \lambda(\lambda^2 I - \bar{J})^{-1} & 0 \\ 0 & \lambda(\lambda^2 I - J)^{-1} \end{bmatrix} + \begin{bmatrix} 0 & (\lambda^2 I - \bar{J})^{-1} \\ (\lambda^2 I - J)^{-1} J & 0 \end{bmatrix} \tau.$$

To study the cosolvent operator of $\mathcal{A}$ we define an auxiliary $\mathbb{C}$-linear operator $\mathbb{A} : \mathcal{M}_n \to \mathcal{M}_n$ as

$$\mathbb{A}(\mathcal{X}) = \mathcal{X}\mathcal{A}, \tag{3.6}$$

i.e., the "right multiplication by $\mathcal{A}$" on $\mathcal{M}_n$. In what follows we establish a connection between the spectrum of $\mathbb{A}$ and $\sigma_{\mathcal{C}}(\mathcal{A})$.

For a polynomial $p(z) = \sum_{j=0}^{k} c_j z^j$ with the coefficients $c_j = s_j + \mathrm{i}t_j \in \mathbb{C}$

$$p(\mathbb{A})(\mathcal{X}) = \sum_{j=0}^{k} c_j \mathcal{X}\mathcal{A}^j = \mathcal{X}\sum_{j=0}^{k} s_j \mathcal{A}^j + \mathrm{i}\mathcal{X}\sum_{j=0}^{k} t_j \mathcal{A}^j$$

for $\mathcal{X} \in \mathcal{M}_n$. Hence for a $\mathbb{C}$-linear $\mathcal{X}$ this equals $\mathcal{X}p(\mathcal{A})$. Therefore $p(\mathbb{A})(I) = p(\mathcal{A})$ in particular. Moreover, $\mathbf{P}(\mathcal{A})$ is an invariant subspace of $p(\mathbb{A})$ for any polynomial p such that

$$p(\mathbb{A})(q(\mathcal{A})) = p(\mathbb{A})(q(\mathbb{A})(I)) = pq(\mathcal{A}) \tag{3.7}$$

holds for any polynomial q. In case p has real coefficients, with $\mathcal{B} = p(\mathcal{A})$ we have the equality $\mathbb{B} = p(\mathbb{A})$.

The eigenvalue problem for the matrix representations (3.1) and (3.2) of $\mathcal{A}$ can be formulated in terms of the operator $\mathbb{A}$.

PROPOSITION 3.9. *Let $\mathcal{A} \in \mathcal{M}_n$. Then $\sigma(\mathbb{A}) = \sigma(\psi(\mathcal{A}))$.*

Proof. Assume $\lambda \in \sigma(\mathbb{A})$. Then $\mathbb{A}(\mathcal{X}) = \lambda\mathcal{X}$ if and only if $\psi(\mathcal{A})^* \begin{bmatrix} X^* \\ X_\#{}^* \end{bmatrix} = \overline{\lambda} \begin{bmatrix} X^* \\ X_\#{}^* \end{bmatrix}$. Hence any non-zero column of $\begin{bmatrix} X^* \\ X_\#{}^* \end{bmatrix}$ gives an eigenvector of $\psi(\mathcal{A})^*$ corresponding to $\overline{\lambda}$. Therefore $\lambda \in \sigma(\psi(\mathcal{A}))$.

Conversely, take an eigenvector of $\psi(\mathcal{A})^*$ corresponding to $\overline{\lambda}$ repeatedly n times to form the columns of $\begin{bmatrix} X^* \\ X_\#{}^* \end{bmatrix}$. Then $X + X_\#\tau$ gives an eigenvector of $\mathbb{A}$ corresponding to λ. ∎

Hence the spectrum of $\mathbb{A}$ is symmetrically located with respect to the real axis implying that as such it cannot yield $\sigma_{\mathcal{C}}(\mathcal{A})$. For a correct object, denote by $\mathbb{A}_I$ the restriction of $\mathbb{A}$ to the subspace $\mathbf{P}(\mathcal{A})$. By its invariance, there holds $\sigma(\mathbb{A}_I) \subset \sigma(\mathbb{A})$.

THEOREM 3.10. *Assume $\mathcal{A} \in \mathcal{M}_n$. Then $\mathcal{C}(\lambda, \mathcal{A}) = \mathcal{R}(\lambda, \mathbb{A}_I)(I)$ and*

$$\sigma_{\mathcal{C}}(\mathcal{A}) = \sigma(\mathbb{A}_I).$$

Proof. For $|\lambda| > \|\mathcal{A}\|$ we have $\mathcal{R}(\lambda, \mathbb{A}_I)(I) = \sum_{j=0}^{\infty} \lambda^{-j-1}\mathcal{A}^j = \mathcal{C}(\lambda, \mathcal{A})$ and hence both extend uniquely to have the same domain of definition. Applying (3.7) for any polynomial $q(z) = \sum_{j=0}^{k} c_j z^j$ we have $\mathcal{R}(\lambda, \mathbb{A}_I)(q(\mathcal{A})) = \sum_{j=0}^{k} c_j \mathcal{R}(\lambda, \mathbb{A}_I)(I)\mathcal{A}^j$ so that the spectrum of $\mathbb{A}_I$ consists of the poles of $\mathcal{R}(\lambda, \mathbb{A}_I)(I)$. ∎

Since the minimal polynomials of $\mathcal{A}$ and $\mathbb{A}_I$ coincide, we have the following corollary.

COROLLARY 3.11. *The set of proper values of $\mathcal{A}$ equals $\sigma_{\mathcal{C}}(\mathcal{A})$.*

We denote the proper values of $\mathcal{A}$ interchangeably by $\sigma_{\mathcal{C}}(\mathcal{A})$ and call $\sigma_{\mathcal{C}}(\mathcal{A})$ the proper values of $\mathcal{A}$. Consequently, there are at most $\deg(\mathcal{A})$ distinct proper values.

EXAMPLE 3.12. For an illustration of Corollary 3.11, let $A, A_\# \in \mathbb{C} + P\mathbb{C}$. Then by Proposition 2.9 we can consider 2-by-2 problems to locate $\sigma_{\mathcal{C}}(\mathcal{A})$. Hence we need to find the zero sets of polynomials of degree at most four, after finding the minimal polynomial for each 2-by-2 problem separately.

Under similarity transformations the proper values behave as follows.

THEOREM 3.13. *Let $\mathcal{A} \in \mathcal{M}_n$ and assume $\mathcal{S} \in \mathcal{M}_n$ is invertible. Then*

$$\sigma_{\mathcal{C}}(\mathcal{S}\mathcal{A}\mathcal{S}^{-1}) \subset \sigma(\psi(\mathcal{A})) \quad \textit{and} \quad \sigma_{\mathcal{C}}(\mathcal{S}\mathcal{A}^*\mathcal{S}^{-1}) \subset \sigma(\psi(\mathcal{A})).$$

Proof. Let first $\mathcal{S} = I$. For the first inclusion, the poles of $\mathcal{R}(\lambda, \mathbb{A}_I)(I)$ are among the poles of $\mathcal{R}(\lambda, \mathbb{A})$ which, by Proposition 3.9, are exactly the points of $\sigma(\psi(\mathcal{A}))$. Since $\sigma(\psi(\mathcal{A}))$ is located symmetrically with respect to the origin, we have $\sigma(\psi(\mathcal{A}^*)) = \sigma(\psi(\mathcal{A})^*) = \sigma(\psi(\mathcal{A}))$. Hence the second inclusion follows from $\sigma_{\mathcal{C}}(\mathcal{A}^*) \subset \sigma(\psi(\mathcal{A}^*))$.

The general case follows from the equality $\sigma(\psi(\mathcal{S}\mathcal{A}\mathcal{S}^{-1})) = \sigma(\psi(\mathcal{A}))$. ∎

This is quite striking because $\mathcal{A}$ and $\mathcal{A}^*$ can even have a different number of proper values; see Example 3.4. Observe also that the minimal polynomials (and their degrees) of $\mathcal{A}$ and $\mathcal{S}\mathcal{A}\mathcal{S}^{-1}$ differ in general.

From Theorem 3.13 it follows that the absolute values of the proper values of $\mathcal{A}$ are bounded by the norm of $\mathcal{A}$.

COROLLARY 3.14. *The set of algebraic elements of degree k at most is closed in $\mathcal{M}_n$ for any $k \in \mathbb{N}$.*

Proof. For any $\mathcal{A}$ we have $\|\mathcal{A}\| = \|\psi(\mathcal{A})\|$. Moreover, if q is the minimal polynomial of $\mathcal{A}$, then its roots are contained in $\sigma(\psi(\mathcal{A}))$. Using these two facts together with the steps of the proof of Theorem 3.2.6 in [1] gives the claim. ∎

PROPOSITION 3.15. *Let $\mathcal{A} \in \mathcal{M}_n$. Then $\sigma_{\mathcal{C}}(\mathcal{A}) \cap \mathbb{R} = \sigma(\mathcal{A}) \cap \mathbb{R} = \sigma(\psi(\mathcal{A})) \cap \mathbb{R}$.*

Proof. Since the second equality is know, we only need to prove the first one.

We have $\sigma_{\mathcal{C}}(\mathcal{A}) = \sigma(\mathbb{A}_I) \subset \sigma(\mathbb{A}) = \sigma(\psi(\mathcal{A}))$. On the other hand, since real numbers commute with the cosolvent operator, we also have $\sigma(\mathcal{A}) \cap \mathbb{R} \subset \sigma(\mathbb{A}_I) \cap \mathbb{R}$. But $\sigma(\mathcal{A}) \cap \mathbb{R} = \sigma(\psi(\mathcal{A})) \cap \mathbb{R}$. ∎

It thus follows that $\sigma_{\mathcal{C}}(rI + \mathcal{A}) = r + \sigma_{\mathcal{C}}(\mathcal{A})$ for $r \in \mathbb{R}$. This does not hold for complex translations r. To see this, take $\mathcal{A}$ with the empty spectrum and use the fact that $\sigma_{\mathcal{C}}(\mathcal{A})$ is never empty together with Proposition 3.15. To get an idea of the behavior of proper values in complex translations, look at circlets in Example 3.7.

Observe that the map

$$\mathcal{A} \longmapsto \sigma_{\mathcal{C}}(\mathcal{A}) \tag{3.8}$$

is not continuous in $\mathcal{M}_n$ in general. In fact, consider the $\mathbb{C}$-linear circlet $\alpha + \beta\tau$ with $\alpha = \mathrm{i}$ and $\beta = 0$ so that $\sigma_{\mathcal{C}}(\alpha + \beta\tau) = \{\mathrm{i}\}$. Take a perturbation $\varepsilon_1 + \varepsilon_2\tau$ with $\varepsilon_1, \varepsilon_2 \in \mathbb{R}$. Then $\mathrm{i} + \varepsilon_1 + \varepsilon_2\tau$ has the proper values $\varepsilon_1 \pm \sqrt{\varepsilon_2^2 - 1}$ given by the formula presented in Example 3.7.

In spite of this we have the following "extended" continuity of (3.8) by the fact that the map $\mathcal{A} \mapsto \sigma(\psi(\mathcal{A}))$ is continuous in $\mathcal{M}_n$.

THEOREM 3.16. *Assume $\mathcal{A} \in \mathcal{M}_n$. If $\lambda \in \sigma(\psi(\mathcal{A}))$, then either $\lambda \in \sigma_{\mathcal{C}}(\mathcal{A})$ or $\overline{\lambda} \in \sigma_{\mathcal{C}}(\mathcal{A})$.*

Proof. Let p be the minimal polynomial of $\mathbb{A}_I$ and assume that for $\lambda \in \sigma(\psi(\mathcal{A}))$ neither $\lambda \in \sigma_{\mathcal{C}}(\mathcal{A})$ nor $\overline{\lambda} \in \sigma_{\mathcal{C}}(\mathcal{A})$. Take the monic polynomial s of the smallest possible degree such that the product ps has real coefficients. Then neither λ nor $\overline{\lambda}$ is a zero of s. Moreover, $0 = p(\mathbb{A}_I)(s(\mathcal{A})) = ps(\mathcal{A})$ to which corresponds $\psi(ps(\mathcal{A})) = ps(\psi(\mathcal{A}))$ by the fact that the coefficients of ps are real. Hence ps annihilates also $\psi(\mathcal{A})$. This is a contradiction since neither λ nor $\overline{\lambda}$ was a zero of ps. ∎

We can thus conclude that knowing the proper values of $\mathcal{A}$ yields the spectrum of its matrix representations (3.1) and (3.2). This surprising fact could possibly be benefitted from in designing algorithms to compute eigenvalue approximations to real matrices.

We can also conclude that $\mathcal{A}$ is nilpotent if and only if $\sigma_{\mathcal{C}}(\mathcal{A}) = \{0\}$ by the more general consequence

$$\max_{\lambda \in \sigma_{\mathcal{C}}(\mathcal{A})} |\lambda| = \max_{\lambda \in \sigma_{\mathcal{C}}(\mathcal{A}^*)} |\lambda| = \lim_{j \to \infty} \|\mathcal{A}^j\|^{1/j} \tag{3.9}$$

of Theorem 3.16 corresponding to the classical Beurling-Gelfand formula. Hence conditions, for instance, on the power boundedness of $\mathcal{A}$ should be formulated in terms of the cosolvent operator of $\mathcal{A}$. To give an example, if $\mathcal{A}$ is star-commuting, i.e., it satisfies $\mathcal{A}^*\mathcal{A} = \mathcal{A}\mathcal{A}^*$, then the norm of $\mathcal{A}$ equals (3.9).

COROLLARY 3.17. *Assume $\lambda \in \sigma_{\mathcal{C}}(\mathcal{A}) \cap \sigma(\psi(\mathcal{A}))$. Then the order of the pole λ is the same for the cosolvent operator of $\mathcal{A}$ and the resolvent operator of $\psi(\mathcal{A})$.*

Proof. Let p be the minimal polynomial of $\mathbb{A}_I$ and take again the smallest degree monic polynomial s such that the product ps has real coefficients. Then $0 = p(\mathbb{A}_I)(s(\mathcal{A})) = ps(\mathcal{A})$ and therefore $ps(\psi(\mathcal{A})) = 0$, so that the degree of the minimal polynomial of $\psi(\mathcal{A})$ is bounded by the degree of ps. By this construction, the order of each zero λ_j of the polynomial ps is the maximum of the order of the zeros λ_j and $\overline{\lambda_j}$ of p.

Define $\mathcal{G} = G + G_\# \tau : \mathbb{C}^n \to \mathbb{C}^{2n}$ with $G = \frac{1}{\sqrt{2}}\begin{bmatrix} I \\ 0 \end{bmatrix}$ and $G_\# = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ I \end{bmatrix}$. Let $R(\lambda, \psi(\mathcal{A}))$ denote the resolvent operator of the matrix $\psi(\mathcal{A})$. Then we have $\mathcal{C}(\lambda, \mathcal{A}) = 2G^* R(\lambda, \psi(\mathcal{A}))\mathcal{G}$ for λ outside $\sigma(\psi(\mathcal{A}))$ so that the poles of $\mathcal{C}(\lambda, \mathcal{A})$ are not of higher order than the poles of $R(\lambda, \psi(\mathcal{A}))$. Hence ps is the minimal polynomial of $\psi(\mathcal{A})$. ∎

In view of Theorems 3.13 and 3.16, those invertible $\mathcal{S} \in \mathcal{M}_n$ that minimize $\deg(\mathcal{S}\mathcal{A}\mathcal{S}^{-1})$ are of interest. In fact, then for $\mathcal{B} = \mathcal{S}\mathcal{A}\mathcal{S}^{-1}$ we have $\mathcal{B}_I : \mathbf{P}(\mathcal{B}) \to \mathbf{P}(\mathcal{B})$ of the least complexity that still determines $\sigma(\psi(\mathcal{A}))$. For instance, if we obtain a $\mathbb{C}$-linear $\mathcal{B}$, then $\dim \mathbf{P}(\mathcal{B}) = \deg(\mathcal{B}) \leqslant n$.

Since at every point λ outside $\sigma_{\mathcal{C}}(\mathcal{A})$ the cosolvent operator belongs to $\mathbb{P}(\mathcal{A})$, it can be regarded as a simpler object than the resolvent operator of $\mathcal{A}$. Still, knowing the cosolvent operator yields the resolvent operator as follows.

PROPOSITION 3.18. *Let $\mathcal{A} \in \mathcal{M}_n$ and $\lambda, \mu, \zeta \in \mathbb{C}$. If $\mathcal{M} = \mathcal{C}(\zeta, \mathcal{A}) - \mathcal{C}(\mu, \mathcal{A})$ is invertible and $\lambda \notin \sigma(\mathcal{A})$, then*

$$(\lambda I - \mathcal{A})^{-1} = (\mathcal{M}\lambda - \zeta\mathcal{M} + (\mu - \zeta)\mathcal{C}(\mu, \mathcal{A}))^{-1}\mathcal{M}.$$

Proof. We have $\mathcal{M}(\lambda I - \mathcal{A}) = \mathcal{M}\lambda - \zeta\mathcal{M} + (\mu - \zeta)\mathcal{C}(\mu, \mathcal{A})$ by using the identity (3.4). Multiplying this with the inverses yields the claim. ∎

As a special case of the above equality we obtain "a cosolvent identity" $\lambda\mathcal{C}(\lambda, \mathcal{A}) - \mu\mathcal{C}(\mu, \mathcal{A}) = (\mathcal{C}(\lambda, \mathcal{A}) - \mathcal{C}(\mu, \mathcal{A}))\mathcal{A}$.

Aside from polynomials in $\mathcal{A}$ and its resolvent and cosolvent operator, let us now consider more elaborate functions of $\mathcal{A}$. The exponential of $\mathcal{A} \in \mathcal{M}_n$ is a natural next candidate being defined by the power series

$$\mathrm{e}^{\mathcal{A}} = \sum_{k=0}^{\infty} \frac{1}{k!}\mathcal{A}^k. \tag{3.10}$$

If $\mathcal{A}$ and $\mathcal{B}$ commute, then we have $\mathrm{e}^{\mathcal{A}+\mathcal{B}} = \mathrm{e}^{\mathcal{A}}\mathrm{e}^{\mathcal{B}}$.

EXAMPLE 3.19. For a circlet $\alpha + \beta\tau$ with $\alpha \in \mathbb{R}$ the parts commute and we obtain

$$\mathrm{e}^{\alpha+\beta\tau} = \mathrm{e}^{\alpha}\Big(\cos(\mathrm{i}|\beta|) - \frac{\mathrm{i}\beta}{|\beta|}\sin(\mathrm{i}|\beta|)\tau\Big). \tag{3.11}$$

In the self-adjoint case this can be used as follows.

PROPOSITION 3.20. *Let $\mathcal{A} \in \mathcal{M}_n$ be self-adjoint. Then* $\mathrm{e}^{\mathcal{A}} = \mathcal{U}\mathrm{e}^{\mathcal{D}}\mathcal{U}^*$ *with a unitary $\mathcal{U}$ and a diagonal $\mathcal{D} = D + D_\# \tau$ having $D \in \mathbb{R}^{n\times n}$.*

Proof. There exists a unitary $\mathcal{U}$ such that $\mathcal{U}^*\mathcal{A}\mathcal{U} = \mathcal{D}$ with a diagonal $\mathcal{D}$ having $D \in \mathbb{R}^{n\times n}$ [9]. (For an algorithm to find $\mathcal{U}$, see [10].) Since all the coefficients in the expansion of the exponential (3.10) are real, we have $\mathrm{e}^{\mathcal{A}} = \mathcal{U}\mathrm{e}^{\mathcal{D}}\mathcal{U}^*$. ∎

With convergent power series functions of $\mathcal{A}$ can be defined analogously. For more generality and by using the identity $\mathcal{C}(\lambda, \mathcal{A}) = \mathcal{R}(\lambda, \mathbb{A}_I)(I)$ we set the following definition. Observe that the differential appears on the left to emphasize the fact that in $\mathcal{M}_n$ the scalar multiplication is performed from the left.

DEFINITION 3.21. Let $\mathcal{A} \in \mathcal{M}_n$. For f analytic in a domain Ω containing $\sigma_{\mathcal{C}}(\mathcal{A})$, set

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

where Γ is a smooth contour surrounding $\sigma_{\mathcal{C}}(\mathcal{A})$ in Ω.

Alternatively, we could use $f(\mathcal{A}) = f(\mathbb{A}_I)(I)$ to define the element $f(\mathcal{A}) \in \mathcal{M}_n$. Because $\operatorname{span}\{I, f(\mathcal{A}), f^2(\mathcal{A}), \ldots\} \subset \mathbf{P}(\mathcal{A})$ holds, the degree of $f(\mathcal{A})$ is at most the degree of $\mathcal{A}$.

THEOREM 3.22. *Let $\mathcal{A} \in \mathcal{M}_n$. If p is a polynomial with real coefficients, then*

$$\sigma_{\mathcal{C}}(p(\mathcal{A})) = p(\sigma_{\mathcal{C}}(\mathcal{A})).$$

Proof. Since $\sigma_{\mathcal{C}}(\mathcal{A}) = \sigma(\mathbb{A}_I)$, we have $p(\sigma_{\mathcal{C}}(\mathcal{A})) = p(\sigma(\mathbb{A}_I)) = \sigma(p(\mathbb{A}_I))$. Hence it remains to prove that $\sigma(p(\mathbb{A}_I)) = \sigma_{\mathcal{C}}(p(\mathcal{A}))$.

Since p has real coefficients, with $\mathcal{B} = p(\mathcal{A})$ there holds $\mathbb{B} = p(\mathbb{A})$. If $\operatorname{span}\{I, p(\mathcal{A}), p^2(\mathcal{A}), p^3(\mathcal{A}), \ldots\} = \mathbf{P}(\mathcal{A})$, then $p(\mathbb{A}_I) = p(\mathbb{A})_I = \mathbb{B}_I$. Hence λ is a proper value of $\mathcal{B}$ if and only if it is an eigenvalue of $p(\mathbb{A}_I)$. Therefore $\sigma(p(\mathbb{A}_I)) = \sigma_{\mathcal{C}}(p(\mathcal{A}))$.

If we have $\operatorname{span}\{I, p(\mathcal{A}), p^2(\mathcal{A}), p^3(\mathcal{A}), \ldots\} \subset \mathbf{P}(\mathcal{A})$, then $\sigma(p(\mathbb{A})_I) = \sigma(\mathbb{B}_I) \subset \sigma(p(\mathbb{A}_I))$. However, since I is a cyclic vector for $\mathbb{A}_I$, we have $\sigma(p(\mathbb{A}_I)) \subset \sigma(p(\mathbb{A})_I)$. Therefore $\sigma(p(\mathbb{A}_I)) = \sigma_{\mathcal{C}}(p(\mathcal{A}))$. ∎

For instance, if $\mathcal{A}$ is invertible, then we can find a polynomial p with real coefficients such that $\mathcal{A}^{-1} = p(\mathcal{A})$.

With this theorem we introduce a decomposition of $\mathcal{A}$ that in the $\mathbb{C}$-linear case would separate, generically, the Jordan block structure of the operator. To this end, take a proper value $\lambda_j \in \sigma_{\mathcal{C}}(\mathcal{A})$ and form the respective pair $\{\lambda_j, \overline{\lambda_j}\}$. Naturally, if λ_j is real, then the pair reduces to a point. For each such pair (or a real point) choose the monic polynomial l_j of the least degree that attains the value one both at λ_j and at $\overline{\lambda_j}$, and is zero elsewhere on $\sigma(\psi(\mathcal{A}))$. Hence $q_j(z) = z l_j(z)$

has real coefficients. If k is the number of such polynomials, then we have

$$\mathcal{A} = \sum_{j=1}^{k} \mathcal{A}_j \quad \text{with} \quad \mathcal{A}_j = \frac{1}{2\pi \mathrm{i}} \int_{\Gamma} \mathrm{d}\zeta\, q_j(\zeta) \mathcal{C}(\zeta, \mathcal{A}). \tag{3.12}$$

By the construction we have $\sigma_{\mathcal{C}}(\mathcal{A}_j) \subset \{\lambda_j, \overline{\lambda_j}, 0\}$ such that $k \leqslant 2n$. If there are no real proper values, then $k \leqslant n$.

EXAMPLE 3.23. Observe how the decomposition (3.12) splits a circlet $\alpha + \beta\tau$ with two separate real proper values. For instance, with $\alpha = 1 + 3\mathrm{i}$ and $\beta = 5$ we have $\sigma_{\mathcal{C}}(\mathcal{A}) = \{-3, 5\}$ and $\alpha + \beta\tau = \mathcal{A}_1 + \mathcal{A}_2 = \frac{3}{8}(-4 + 3\mathrm{i} + 5\tau) + \frac{5}{8}(4 + 3\mathrm{i} + 5\tau)$. Otherwise a circlet is not split.

With l_j defined above we obtain an idempotent $\mathbb{R}$-linear operator $\mathcal{P}_j = \frac{1}{2\pi\mathrm{i}} \int_{\Gamma} \mathrm{d}\zeta\, l_j(\zeta)\mathcal{C}(\zeta, \mathcal{A})$ satisfying $\mathcal{P}_j^2 = \mathcal{P}_j$ and $\mathcal{A}\mathcal{P}_j = \mathcal{P}_j\mathcal{A}$.

A particular class of $\mathbb{R}$-linear operators is obtained as follows. If $f(\mathcal{A}) = p(\mathcal{A})$ for some polynomial p with real coefficients, then $f(\mathcal{A})$ is representable by a real polynomial. This is of interest since we obviously have $\sigma_{\mathcal{C}}(f(\mathcal{A})) = p(\sigma_{\mathcal{C}}(\mathcal{A}))$. For instance, there holds $\sigma_{\mathcal{C}}(\mathrm{e}^{\mathcal{A}}) = \mathrm{e}^{\sigma_{\mathcal{C}}(\mathcal{A})}$. Furthermore, if $f(\mathcal{A})$ is representable by a real polynomial, then it can also be represented with a right polynomial.

EXAMPLE 3.24. For the exponential we have $\mathrm{e}^{\mathcal{A}} = \sum_{k=0}^{\infty} \frac{1}{k!}\mathcal{A}^k = \sum_{k=0}^{\infty} \mathcal{A}^k \frac{1}{k!}$ and therefore $\mathrm{e}^{\mathcal{A}}$ can also be represented with a right polynomial in $\mathcal{A}$. The degree of the right polynomial needed is at most $\deg(\mathcal{A}^*)$.

For $\mathcal{A}$ having sufficiently small norm, the logarithm $\log(I + \mathcal{A})$ yields another important example of a function of $\mathcal{A}$ that is representable by a real polynomial.

4. LOCAL ASPECTS AND REPRESENTATION BY POLYNOMIALS

We end this paper by making some very preliminary observations on applying functions of $\mathbb{R}$-linear operators to vectors.

In spite of the many similarities with the $\mathbb{C}$-linear theory, Kaplansky's theorem does not hold in the $\mathbb{R}$-linear case. To see this, consider the following example.

EXAMPLE 4.1. For a simple example that is readily generalized, consider $\mathcal{A} = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right] \tau$ on $\mathbb{C}^2$. For any $b \in \mathbb{C}^2$ the vectors b, $\mathcal{A}b$ and $\mathcal{A}^2 b$ are linearly dependent while $q(z) = z^4 - 5z^2 + 4$ is the minimal polynomial of $\mathcal{A}$.

This type of local aspects have a number of intriguing consequences. For one thing, repeatedly applying $\mathcal{A} \in \mathcal{M}_n$ to a vector is not a valid approach to find the minimal polynomial of $\mathcal{A}$. For the matrix case, see [6].

To apply an analytic function f of $\mathcal{A}$ to a vector $b \in \mathbb{C}^n$, the respective Krylov subspace defined as

$$\mathcal{K}(\mathcal{A}; b) = \operatorname{span}\{b, \mathcal{A}b, \mathcal{A}^2 b, \mathcal{A}^3 b, \dots\} \tag{4.1}$$

is of importance by the fact that $f(\mathcal{A})b$ is representable by a polynomial in $\mathcal{A}$ applied to b. Its dimension is bounded from above by left-rank$(\mathcal{A}) + 1$, which should be contrasted with Proposition 3.2.

For an illustration, since for any invertible $\mathcal{A}$ we have $\mathcal{A}^{-1} = p(\mathcal{A})$ for some polynomial p, the iterates (4.1) are of use in solving $\mathbb{R}$-linear systems of equations with iterative methods, i.e., the solution to $\mathcal{A}x = b$ can be represented as $x = q(\mathcal{A})b$ for a polynomial q. In practice approximations only with low degree polynomials can be constructed; see [3].

For the proper values of $\mathcal{A}$ it is of interest to compute the monic polynomial p of the least degree satisfying $p(\mathcal{A})b = 0$. Unlike in the $\mathbb{C}$-linear case, it is not clear how the zeros of p are related to $\sigma_{\mathcal{C}}(\mathcal{A})$. This remains as an open problem.

For one more familiar example, consider applying the exponential function.

EXAMPLE 4.2. The exponential of $\mathcal{A}$ defined by the power series (3.10) can be given as $\mathrm{e}^{\mathcal{A}t} = \sum_{k=0}^{\deg(\mathcal{A})} \beta_k(t)\mathcal{A}^k$ with some functions β_k for $k = 0, 1, \dots, \deg(\mathcal{A})$. Applying this to a vector $b \in \mathbb{C}^n$ changes dramatically the coefficients in the representation of the least degree because $\mathrm{e}^{\mathcal{A}t} b = \sum_{k=0}^{\dim(\mathcal{K}(\mathcal{A};b))-1} \beta_k(b; t)\mathcal{A}^k b$. This is a more compressed way of representing $\mathrm{e}^{\phi(\mathcal{A})t}\left[\begin{smallmatrix}\mathrm{Re}\, b\\ \mathrm{Im}\, b\end{smallmatrix}\right]$ since generically $\dim(\mathcal{K}(\mathcal{A}; b)) = \frac{1}{2}\dim(\mathcal{K}(\phi(\mathcal{A}); \left[\begin{smallmatrix}\mathrm{Re}\, b\\ \mathrm{Im}\, b\end{smallmatrix}\right]))$. An interesting problem, though beyond the scope of this paper, is that of computing the functions $\beta_k(b; \cdot)$.

By considering the right polynomials in $\mathcal{A}$, we set the respective $\mathbb{R}$-linear manifold of $\mathbb{C}^n$ that consists of the vectors representable as

$$\sum_{j=0}^{k} \mathcal{A}^j c_j b \quad \text{with } c_0, c_1, \dots, c_k \in \mathbb{C}. \tag{4.2}$$

All those functions f of $\mathcal{A}$ that are representable by a real polynomial allow $f(A)b$ to be written alternatively in the form (4.2).

CONCLUSIONS

The eigenvalues and the proper values are two natural spectral sets associated with an $\mathbb{R}$-linear operator $\mathcal{A}$. In their complement the resolvent and cosolvent operators are defined, respectively. The proper values of $\mathcal{A}$, given by the zeros of its minimal polynomial, also determine the spectrum of the real matrix representation of $\mathcal{A}$. This interpreted conversely, the eigenvalues of real matrices can be found with the proper values. Analytic functions of $\mathcal{A}$ can be defined with the cosolvent operator. An application of this leads to a decomposition of $\mathbb{R}$-linear operators.

REFERENCES

[1] B. AUPETIT, *A Primer on Spectral Theory*, Springer-Verlag, New-York 1991.

[2] T. EIROLA, Krylov integrators for Hamiltonian systems, talk given at the *Workshop on Exponential Integrators (Innsbruck, October 20–23 2004)*, http://techmath.uibk.ac.at/numbau/alex/events/conference2004.html.

[3] T. EIROLA, M. HUHTANEN, J. VON PFALER, Solution methods for $\mathbb{R}$-linear problems in $\mathbb{C}^n$, *SIAM J. Matrix Anal. Appl.* **25**(2004), 804–828.

[4] Y.P. HONG, R.A. HORN, On simultaneous reduction of families of matrices to triangular or diagonal form by unitary congruences, *Linear and Multilin. Algebra* **17**(1985), 271–288.

[5] Y.P. HONG, R.A. HORN, A canonical form for matrices under consimilarity, *Linear Algebra Appl.* **102**(1988), 143–168.

[6] A.S. HOUSEHOLDER, *The Theory of Matrices in Numerical Analysis*, Blaisdell Publ. Co. Ginn and Co., New York-Toronto-London 1974.

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

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

[9] M. HUHTANEN, O. NEVANLINNA, Real linear matrix analysis, *Banach Center Publ.* **75**(2007), 171–189.

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

MARKO HUHTANEN, INSTITUTE OF MATHEMATICS, HELSINKI UNIVERSITY OF TECHNOLOGY, BOX 1100, FIN-02015, FINLAND
E-mail address: Marko.Huhtanen@tkk.fi

OLAVI NEVANLINNA, INSTITUTE OF MATHEMATICS, HELSINKI UNIVERSITY OF TECHNOLOGY, BOX 1100, FIN-02015, FINLAND
E-mail address: Olavi.Nevanlinna@tkk.fi

Received July 22, 2005.