Linear Algebra and its Applications 634 (2022) 37–56

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

On a formula of Thompson and McEnteggert for the adjugate matrix

Kenier Castillo [a,1], Ion Zaballa [b,*,2]

[a] *CMUC, Department of Mathematics, University of Coimbra, 3001-501 Coimbra, Portugal*
[b] *Departamento de Matemática Aplicada y EIO, Universidad del País Vasco (UPV/EHU), Apdo, 644, 48080 Bilbao, Spain*

A R T I C L E I N F O

Article history:
Received 24 February 2021
Accepted 21 October 2021
Available online 28 October 2021
Submitted by F. Dopico

MSC:
15A18
15A15

Keywords:
Adjugate
Eigenvalues
Eigenvectors
Elementary divisors
Rank-one matrices

A B S T R A C T

For an eigenvalue λ_0 of a Hermitian matrix A, the formula of Thompson and McEnteggert gives an explicit expression of the adjugate of $\lambda_0 I - A$, $\operatorname{Adj}(\lambda_0 I - A)$, in terms of eigenvectors of A for λ_0 and all its eigenvalues. In this paper Thompson-McEnteggert's formula is generalized to include any matrix with entries in an arbitrary field. In addition, for any nonsingular matrix A, a formula for the elementary divisors of $\operatorname{Adj}(A)$ is provided in terms of those of A. Finally, a generalization of the eigenvalue-eigenvector identity and three applications of the Thompson-McEnteggert's formula are presented.

* Corresponding author.
E-mail addresses: kenier@mat.uc.pt (K. Castillo), ion.zaballa@ehu.eus (I. Zaballa).

[1] Partially supported by the Centre for Mathematics of the University of Coimbra - UIDB/00324/2020, funded by the Portuguese Government through FCT/MCTES.
[2] Partially supported by "Ministerio de Economía, Industria y Competitividad (MINECO)" of Spain and "Fondo Europeo de Desarrollo Regional (FEDER)" of EU through grants MTM2017-83624-P and MTM2017-90682-REDT, and by UPV/EHU through grant GIU16/42.

https://doi.org/10.1016/j.laa.2021.10.021

1. Introduction

Let $\mathcal{R}$ be a commutative ring with identity. Following [17, Ch. 30], for a polynomial $p(\lambda) = \sum_{k=0}^{n} p_k \lambda^k \in \mathcal{R}[\lambda]$ its derivative is $p'(\lambda) = \sum_{k=1}^{n} k p_k \lambda^{k-1}$. Recall that if $X \in \mathcal{R}^{n\times n}$ is a square matrix of order n with entries in $\mathcal{R}$ and $M_{ij}(X)$ is the minor obtained from X by deleting the ith row and jth column then the *adjugate* of X, $\mathrm{Adj}(X)$, is the matrix whose (i,j) entry is $(-1)^{i+j}M_{ji}(X)$; that is,

$$\mathrm{Adj}(X) = \left[(-1)^{i+j}M_{ji}(X)\right]_{1\le i,j\le n}.$$

Formula (1) below, from now on TM formula, was proved, with $w = v$ and the normalization $w^*v = 1$, for a Hermitian matrix $A \in \mathbb{C}^{n\times n}$ by Thompson and McEnteggert (see [34, pp. 212-213]). Inspection of the proof shows that the formula also holds for normal matrices over $\mathbb{C}$ (see [29]). With the same arguments we can go further. Recently, Denton, Parke, Tao, and Zhang pointed out that the TM formula has an extension to a non-normal matrix $A \in \mathbb{R}^{n\times n}$, so long as it is diagonalizable (see [13, Rem. 4]). Even more, as shown in Remark 5 of [13] it holds for matrices over commutative rings (see [18] for an informal proof). A more detailed proof of this result will be given in Section 2. However, for matrices over fields (or over integral domains) with repeated eigenvalues, (1) does not provide meaningful information (see Remark 2.4). We will exhibit in Section 2 a generalization of the TM formula which holds for matrices over arbitrary fields with repeated eigenvalues. This new TM formula will be used to generalize the so-called *eigenvector-eigenvalue identity* (see (23)) for non-diagonalizable matrices over arbitrary fields. In addition we will provide a complete characterization of the similarity invariants of $\mathrm{Adj}(A)$ in terms of those of A, generalizing a result about the eigenvalues and the minimal polynomial in [19]. Then in Section 3 three additional consequences of the TM formula will be analysed.

2. The TM formula and its generalization

Let $A \in \mathcal{R}^{n\times n}$ be a square matrix of order n with entries in $\mathcal{R}$. An element $\lambda_0 \in \mathcal{R}$ is said to be an *eigenvalue* of A if $Ax = \lambda_0 x$ for some nonzero vector $x \in \mathcal{R}^{n\times 1}$ ([7, Def. 17.1]). This vector is said to be a *right eigenvector* of A for (or associated with) λ_0. The *left eigenvectors* of A for λ_0 are the right eigenvectors for λ_0 of A^T, the transpose of A. That is to say, $y \in \mathcal{R}^{n\times 1}$ is a left eigenvector of A for λ_0 if $y^T A = \lambda_0 y^T$.[3] The characteristic polynomial of A is $p_A(\lambda) = \det(\lambda I_n - A)$ and λ_0 is an eigenvalue of A if and only if $p_A(\lambda_0)$ is a zero divisors of $\mathcal{R}$ ([7, Lem. 17.2]).

[3] If $\mathcal{R} = \mathbb{C}$ is the field of complex numbers, the left eigenvectors are the right eigenvectors of $A^* = \bar{A}^T$, the conjugate transpose of A. That is, $y \in \mathcal{R}^{n\times 1}$ is a left eigenvector of $A \in \mathbb{C}^{n\times n}$ for λ_0 if $y^*A = \lambda_0 y^*$. Since in this section we will work with matrices over arbitrary commutative rings or fields we will adopt the "transpose notation" in the understanding that for complex matrices ^{T} must be replaced by *.

The following result, in a slightly different form, was proved by D. Grinberg in [18].

Theorem 2.1. *Let $A \in \mathcal{R}^{n\times n}$ and let $\lambda_0 \in \mathcal{R}$ be an eigenvalue of A. Let v, $w \in \mathcal{R}^{n\times 1}$ be a right and a left eigenvector, respectively, of A for λ_0. Then*

$$w^T v \operatorname{Adj}(\lambda_0 I_n - A) = p'_A(\lambda_0) v w^T. \tag{1}$$

The proof in [18] is based on the following Lemma which is interesting in its own right.

Lemma 2.2. *Let $A \in \mathcal{R}^{n\times n}$ be a matrix and let $w \in \mathcal{R}^{n\times 1}$ be a left eigenvector of A for the eigenvalue 0. For $j = 1, \dots, n$, let $(\operatorname{Adj} A)_j$ be the jth column of $\operatorname{Adj}(A)$. Then, for all $i, j = 1, \dots, n$,*

$$w_i(\operatorname{Adj} A)_j = w_j(\operatorname{Adj} A)_i, \tag{2}$$

where $w = \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix}^T$.

This is Lemma 3 of [18]. The author himself considers the proof to be informal. So a detailed proof of Lemma 2.2, following Grinberg's ideas,[4] is given next for reader's convenience. Note that when $\mathcal{R}$ is a principal domain one-line proof can be given, by simply observing that each row of the adjugate is a multiple of w^T over the field of fractions of $\mathcal{R}$.

Proof of Lemma 2.2. Let us take $i, j \in \{1, \dots, n\}$ and assume that $i \neq j$; otherwise, there is nothing to prove. We assume also, without loss of generality, that $i < j$. Let $w = \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix}^T$ and, for $k = 1, \dots, n$, let a_k be the kth row of A. Define $B \in \mathcal{R}^{n\times n}$ to be the matrix whose kth row, b_k, is equal to a_k if $k \neq i, j$ and $b_k = w_k a_k$ if $k = i, j$. A simple computation shows that $w_i(\operatorname{Adj} A)_j = (\operatorname{Adj} B)_j$ and $w_j(\operatorname{Adj} A)_i = (\operatorname{Adj} B)_i$. We claim that $(\operatorname{Adj} B)_j = (\operatorname{Adj} B)_i$. This would prove the lemma.

It follows from $w^T A = 0$ that

$$\sum_{k=1}^{n} w_k a_k = 0$$

and so

$$b_i + b_j = -\sum_{k=1, k\neq i,j}^{n} w_k b_k. \tag{3}$$

Let

[4] Grinberg's permission was granted to include the proofs of this Lemma and Theorem 2.1.

$$
P = \begin{array}{rl}
 & \begin{matrix} & & \phantom{-w_{i-1}} & i & \phantom{-w_{i+1}} & & \phantom{-w_{j-1}} & j & \phantom{-w_{j+1}} & & \end{matrix} \\
\begin{matrix} \\ \\ \\ i \\ \\ \\ \\ j \\ \\ \\ \\ \end{matrix} &
\begin{bmatrix}
1 & & & & & & & & & & \\
 & \ddots & & & & & & & & & \\
 & & 1 & & & & & & & & \\
-w_1 & \cdots & -w_{i-1} & -1 & -w_{i+1} & \cdots & -w_{j-1} & 0 & -w_{j+1} & \cdots & -w_n \\
 & & & & 1 & & & & & & \\
 & & & & & \ddots & & & & & \\
 & & & & & & 1 & & & & \\
-w_1 & \cdots & -w_{i-1} & 0 & -w_{i+1} & \cdots & -w_{j-1} & -1 & -w_{j+1} & \cdots & -w_n \\
 & & & & & & & & 1 & & \\
 & & & & & & & & & \ddots & \\
 & & & & & & & & & & 1
\end{bmatrix}
\end{array}.
$$

This matrix is invertible in $\mathcal{R}$ (its determinant is 1) and by (3),

$$
\widetilde{B} = PB = \begin{bmatrix} b_1^T & \cdots & b_{i-1}^T & b_j^T & b_{i+1}^T & \cdots & b_{j-1}^T & b_i^T & b_{j+1}^T & \cdots & b_n^T \end{bmatrix}^T .
$$

Then, $\operatorname{Adj}(\widetilde{B}) = \operatorname{Adj}(B) \operatorname{Adj}(P)$ and, since P is invertible, $\operatorname{Adj}(P) = (\det P)P^{-1} = P^{-1}$. Hence $\operatorname{Adj}(B) = \operatorname{Adj}(\widetilde{B})P$ and for $k = 1, \dots, n$

$$
(\operatorname{Adj} B)_{ki} = \sum_{\ell=1}^{n} (\operatorname{Adj} \widetilde{B})_{k\ell} P_{\ell i}.
$$

But in the ith column of P the only nonzero entry is -1 in position (i, i). Thus, $(\operatorname{Adj} B)_{ki} = -(\operatorname{Adj} \widetilde{B})_{ki}$. We recall now that $\widetilde{B}$ is the matrix B with rows ith and jth interchanged and that $M_{ij}(X)$ is the minor of X obtained by deleting the ith row and jth column of X. Thus, if $b_\ell(k)$ is the ℓth row of B with the kth component removed; i.e. $b_\ell(k) = [b_{\ell\, 1} \quad \cdots \quad b_{\ell\, k-1} \quad b_{\ell\, k+1} \quad \cdots \quad b_{\ell\, n}]$, we have

$$
M_{ik}(\widetilde{B}) = \det \begin{bmatrix} b_1(k) \\ \vdots \\ b_{i-1}(k) \\ b_{i+1}(k) \\ \vdots \\ b_{j-1}(k) \\ b_i(k) \\ b_{j+1}(k) \\ \vdots \\ b_n(k) \end{bmatrix} = (-1)^{j-i-1} \det \begin{bmatrix} b_1(k) \\ \vdots \\ b_{i-1}(k) \\ b_i(k) \\ b_{i+1}(k) \\ \vdots \\ b_{j-1}(k) \\ b_{j+1}(k) \\ \vdots \\ b_n(k) \end{bmatrix} = (-1)^{j-i-1} M_{jk}(B).
$$

Therefore

$$\begin{aligned}(\operatorname{Adj} B)_{ki} &= -(\operatorname{Adj} \widetilde{B})_{ki} = (-1)^{k+i+1} M_{ik}(\widetilde{B}) \\ &= (-1)^{k+i+1}(-1)^{j-i-1} M_{jk}(B) \\ &= (-1)^{k+j} M_{jk}(B) = (\operatorname{Adj} B)_{kj},\end{aligned}$$

as claimed. □

There is a “row version” of Lemma 2.2 which can be proved along the same lines.

Lemma 2.3. *Let $A \in \mathcal{R}^{n\times n}$ be a matrix and let $v \in \mathcal{R}^{n\times 1}$ be a right eigenvector of A for the eigenvalue 0. For $j = 1, \dots, n$ let $(\operatorname{Adj} A)^j$ be the jth row of $\operatorname{Adj}(A)$. Then, for all $i, j = 1, \dots, n$,*

$$v_i(\operatorname{Adj} A)^j = v_j(\operatorname{Adj} A)^i, \tag{4}$$

where $v = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}^T$.

The proof of Theorem 2.1 which follows is very much that of Grinberg in [18]. It is included for completion and reader’s convenience.

Proof of Theorem 2.1. Let $B = \lambda_0 I_n - A$ and $p_B(\lambda) = \det(\lambda I_n - B)$ its characteristic polynomial. Then $p_B(\lambda) = \lambda^n + \sum_{k=1}^{n}(-1)^k c_k \lambda^{n-k}$ where, for $k = 0, \dots, n$, $c_k = \sum_{1\le i_1<\dots<i_k\le n} \det B(i_1 : i_k, i_1 : i_k)$, and $B(i_1 : i_k, i_1 : i_k) = [b_{i_j,i_\ell}]_{1\le j,\ell\le k}$ is the principal submatrix of B formed by the rows and columns $i_1, \dots, i_k$. In particular, $c_{n-1} = \sum_{j=1}^{n} M_{jj}(B)$ where $M_{jj}(B)$ is the principal minor of B obtained by deleting the jth row and column. Thus $p_B'(0) = (-1)^{n-1}\sum_{j=1}^{n} M_{jj}(B)$.

On the other hand, $p_B(\lambda) = \det(\lambda I_n - B) = \det(\lambda I_n - \lambda_0 I_n + A) = (-1)^n \det((\lambda_0 - \lambda)I_n - A) = (-1)^n p_A(\lambda_0 - \lambda)$. It thus follows from the chain rule that

$$p_A'(\lambda_0) = (-1)^{n+1} p_B'(0) = \sum_{j=1}^{n} M_{jj}(B).$$

Hence, proving (1) is equivalent to proving

$$w^T v \operatorname{Adj}(B) = \sum_{j=1}^{n} M_{jj}(B) v w^T. \tag{5}$$

It follows from $Av = \lambda_0 v$ and $w^T A = \lambda_0 w^T$ that $Bv = 0$ and $w^T B = 0$, respectively. So we can apply to B properties (2) and (4). If w_k and v_k are the kth components of w and

v, respectively, we get from (2) $w_k(\operatorname{Adj} B)_{ij} = w_j(\operatorname{Adj} B)_{ik}$ for all $i, j, k \in \{1, \dots, n\}$. Then $v_k w_k(\operatorname{Adj} B)_{ij} = w_j v_k(\operatorname{Adj} B)_{ik}$ and by (4), $v_k(\operatorname{Adj} B)_{ik} = v_i(\operatorname{Adj} B)_{kk}$. Hence,

$$v_k w_k(\operatorname{Adj} B)_{ij} = v_i w_j(\operatorname{Adj} B)_{kk}, \quad i, j, k = 1, \dots, n.$$

Summing over k and taking into account that $(\operatorname{Adj} B)_{kk} = M_{kk}(B)$, we get

$$w^T v(\operatorname{Adj} B)_{ij} = \sum_{k=1}^{n} M_{kk}(B) v_i w_j, \quad i, j = 1, \dots n.$$

This is equivalent to (5) and the theorem follows. $\square$

Remark 2.4. Assume that $\mathcal{R}$ is an integral domain and note that in this case the rank of $A \in \mathcal{R}^{m\times n}$, $\operatorname{rank}(A)$, is the rank of A computed as a matrix over the field of fractions of $\mathcal{R}$. It is an interesting consequence of (1) that $w^T v = 0$ implies $p'_A(\lambda_0) = 0$. The converse is not true in general. For example, if $A = \lambda_0 I_2$ then $v = [1 \quad 0]^T$ satisfies both $Av = \lambda_0 v$ and $v^T A = \lambda_0 v^T$, but $v^T v = 1$ and $p'_A(\lambda_0) = 0$. However, if $p'_A(\lambda_0) = 0$ and $\operatorname{rank}(\lambda_0 I_n - A) = n-1$ then, necessarily, $w^T v = 0$ because $\operatorname{Adj}(\lambda_0 I_n - A)$ is not the zero matrix. In particular, when $\mathcal{R}$ is a field, it follows from (1) that if $w^T v = 0$ then λ_0 is an eigenvalue of algebraic multiplicity at least 2. On the other hand, it is easily checked that if λ_0 is an eigenvalue of algebraic multiplicity bigger that 1 and geometric multiplicity 1 (see below the definitions of these two notions) then $w^T v = 0$ for any right and left eigenvectors, v and w respectively, of A for λ_0. This is the case, for example, of $A = \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right]$. For this matrix, the TM formula (1) does not provide any substantial information about $\operatorname{Adj}(\lambda_0 I_n - A)$ because, in this case, $w^T v = 0$ and $p'_A(\lambda_0) = 0$. Thus, the TM formula (1) is relevant for matrices with simple eigenvalues. $\square$

Our next goal is to provide a generalization of the TM formula (1) which is meaningful for nondiagonalizable matrices over fields. We will use the following notation. $\mathbb{F}$ will denote an arbitrary field. If $A \in \mathbb{F}^{n\times n}$ then $p_1(\lambda), \dots, p_r(\lambda)$ will be its (possibly repeated) elementary divisors in $\mathbb{F}$; that is the elementary divisors of its characteristic matrix $\lambda I_n - A$ as a matrix polynomial ([16, Ch. VI, Sec. 3]). These are powers of monic irreducible polynomials of $\mathbb{F}[\lambda]$ (the ring of polynomials with coefficients in $\mathbb{F}$). We will assume that for $j = 1, \dots, r$,

$$p_j(\lambda) = \lambda^{d_j} + a_{j1}\lambda^{d_j-1} + a_{j2}\lambda^{d_j-2} + \cdots + a_{jd_j-1}\lambda + a_{jd_j}. \tag{6}$$

Let Δ_A denote the determinant of A and $\Lambda(A)$ the set of eigenvalues (the spectrum) of A in, perhaps, an extension field, $\widetilde{\mathbb{F}}$, of $\mathbb{F}$. Thus $\lambda_0 \in \Lambda(A)$ if and only if it is a root in $\widetilde{\mathbb{F}}$ of $p_j(\lambda)$ for some $j \in \{1, 2, \dots, r\}$. In particular, $p_A(\lambda) = \prod_{j=1}^{r} p_j(\lambda)$ is the characteristic polynomial of A.

As usual, if $\lambda_0 \in \Lambda(A)$ then its *algebraic multiplicity* ($\operatorname{ma}(\lambda_0)$) is the multiplicity of λ_0 as a root of $p_A(\lambda)$, and its *geometric multiplicity* ($\operatorname{mg}(\lambda_0)$) is the number of irreducible polynomials $p_j(\lambda)$ with λ_0 as a root. Note that if $0 \in \Lambda(A)$ and $p_j(0) = 0$ then $p_j(\lambda) = \lambda^{d_j}$ because $p_j(\lambda)$ is a power of an irreducible polynomial.

We review now some basic notions and results about the structure of linear operators on n-dimensional vector spaces. Our main reference is [16, Ch. VII]. First, if $0 \in \Lambda(A)$ and $p_j(0) = 0$ then (see [16, Ch. VII, Th. 8]), associated to the elementary divisor $p_j(\lambda) = \lambda^{d_j}$, there is a cyclic (or Krylov) subspace $\mathfrak{I}_j$ of dimension d_j whose minimal polynomial is $p_j(\lambda)$; i.e., $p_j(A)x = 0$ for all $x \in \mathfrak{I}_j$ and there is no polynomial of degree less that d_j satisfying this property. Then $\mathfrak{I}_j$ admits a Krylov basis; that is, $\mathfrak{I}_j =< x, Ax, \dots, A^{d_j-1}x >$ for some vector $x \in \mathbb{F}^{n\times 1}$ whose minimal polynomial is $p_j(\lambda) = \lambda^{d_j}$; i.e. $A^{d_j}x = 0$ and $A^k x \neq 0$ for $k < d_j$. Such a vector is said to be a *generating vector* of $\mathfrak{I}_j$. Now, if $p_j(\lambda)$ is the only elementary divisor which is multiple of λ (that is, $\mathrm{mg}(0) = 1$) then, for $k = 1, \dots, d_j$, $\ker A^k =< A^{d_j-k}x, A^{d_j-k+1}x, \dots, A^{d_j-1}x >$. In particular, $\ker A^{d_j} =< x, Ax, \dots, A^{d_j-1}x >= \mathfrak{I}_j$ and $A^{d_j-1}x$ is a right eigenvector of A for the eigenvalue 0. Conversely, if a right eigenvector u of A for the eigenvalue 0 is given and $\mathrm{mg}(0) = 1$ then we can always construct a Krylov basis of $\ker A^{d_j}$ by solving the linear systems $Ax_{d_j-2} = u$ and $Ax_j = x_{j+1}$, $j = 1, \dots, d_j - 3$ (see [16, Ch. VII, Sec. 7]). In fact, as already seen, $\ker A^{d_j}$ always admits a Krylov basis $< v, Av, \dots, A^{d_j-1}v >$ where $A^{d_j-1}v$ is an eigenvector of A for the eigenvalue 0. Since $\mathrm{mg}(0) = 1$, $u = \alpha A^{d_j-1}v$ for some nonzero scalar $\alpha \in \mathbb{F}$. Setting $x_k = \alpha A^k v$ for $k = 0, 1, \dots, d_j - 2$, we get $Ax_k = x_{k+1}$ and so $< x_1, \dots, x_{d_j-2}, u >$ is a Krylov basis of $\ker A^{d_j}$ as claimed. The relationship between any two Krylov basis for $\ker A^{d_j}$ is analysed below in Lemma 2.5. If u is an eigenvector of A for the eigenvalue 0 and $< x_1, \dots, x_{d_j-2}, u >$ is a Krylov basis for $\ker A^{d_j}$, then we will say that x_1 is a *generating vector of* $\ker A^{d_j}$ *for the eigenvector* u.

As mentioned above the following lemma gives the relationship between two Krylov bases of the same subspace. Its proof is straightforward. Recall that a Toeplitz matrix is a square matrix whose entries (i, j) and $(i + 1, j + 1)$ coincide for all possible values of i and j.

Lemma 2.5. *Let* $0 \in \Lambda(A)$ *be an eigenvalue with* $\mathrm{ma}(0) = k$ *and* $\mathrm{mg}(0) = 1$ *and let* $x, y \in \mathbb{F}^{n\times n}$ *be generating vectors of the Krylov subspace* $\ker A^k$. *Then there is an invertible lower triangular Toeplitz matrix* $X \in \mathbb{F}^{k\times k}$ *such that*

$$\begin{bmatrix} y & Ay & \cdots & A^{k-1}y \end{bmatrix} = \begin{bmatrix} x & Ax & \cdots & A^{k-1}x \end{bmatrix} X. \tag{7}$$

Note that since A and A^T are similar matrices, the above results apply for left Krylov subspaces and left generating vectors.

Item (ii) of the following theorem is an elementary result that is included for completion. Item (i) is a generalization of [19, Th. 2].

Theorem 2.6. *Let* $A \in \mathbb{F}^{n\times n}$ *and let the polynomials* $p_j(\lambda)$ *of* (6) *be its elementary divisors,* $j = 1, \dots, r$.

(i) *If* $0 \notin \Lambda(A)$ *then the elementary divisors of* $\mathrm{Adj}(A)$ *are* $q_1(\lambda), \dots, q_r(\lambda)$ *where for* $j = 1, \dots, r$,

$$q_j(\lambda) = \lambda^{d_j} + \Delta_A \frac{a_{jd_j-1}}{a_{jd_j}} \lambda^{d_j-1} + \cdots + \Delta_A^{d_j-1} \frac{a_{j1}}{a_{jd_j}} \lambda + \Delta_A^{d_j} \frac{1}{a_{jd_j}}. \tag{8}$$

(ii) *If* $0 \in \Lambda(A)$ *and* $\mathrm{mg}(0) \geq 2$ *then* $\mathrm{Adj}(A) = 0$.

(iii) *Let* $0 \in \Lambda(A)$, $\mathrm{mg}(0) = 1$ *and assume that* $p_k(0) = 0$. *Let* $u, v \in \mathbb{F}^{n\times 1}$ *be arbitrary right and left eigenvectors of* A, *respectively, for the eigenvalue* 0 *and let* $x, y \in \mathbb{F}^{n\times 1}$ *be arbitrary right and left generating vectors of* $\ker A^{d_k}$ *and* $\ker(A^T)^{d_k}$ *for the eigenvectors* u *and* v, *respectively. Then* $y^T A^{d_k-1} x \neq 0$ *and*

$$\mathrm{Adj}(A) = (-1)^{n-1} \prod_{j=1, j\neq k}^{r} p_j(0) \frac{uv^T}{y^T A^{d_k-1} x}. \tag{9}$$

Proof. For $j = 1, \dots, r$, let the companion matrix of $p_j(\lambda)$ be

$$C_j = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_{jd_j} \\ 1 & 0 & \cdots & 0 & -a_{jd_j-1} \\ 0 & 1 & \cdots & 0 & -a_{jd_j-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{j1} \end{bmatrix}. \tag{10}$$

Then (see [16, Ch. VII, Sec. 5]) there is an invertible matrix $S \in \mathbb{F}^{n\times n}$ such that

$$C = S^{-1}AS = \bigoplus_{j=1}^{r} C_j. \tag{11}$$

An explicit computation shows that

$$\mathrm{Adj}(C_j) = (-1)^{d_j} \begin{bmatrix} -a_{jd_j-1} & a_{jd_j} & 0 & \cdots & 0 \\ -a_{jd_j-2} & 0 & a_{jd_j} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_{k1} & 0 & 0 & \cdots & a_{jd_j} \\ -1 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Bearing in mind that $\det C_j = (-1)^{d_j} a_{jd_j}$, we obtain $\mathrm{Adj}(C) = \oplus_{j=1}^{r} L_j$ where, for $j = 1, \dots, r$,

$$L_j = (-1)^n \prod_{i=1, i\neq j}^{r} a_{id_i} \begin{bmatrix} -a_{jd_j-1} & a_{jd_j} & 0 & \cdots & 0 \\ -a_{jd_j-2} & 0 & a_{jd_j} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_{k1} & 0 & 0 & \cdots & a_{jd_j} \\ -1 & 0 & 0 & \cdots & 0 \end{bmatrix}. \tag{12}$$

Therefore, from (11) we get

$$\mathrm{Adj}(A) = S\left(\bigoplus_{j=1}^{r} L_j\right) S^{-1}. \tag{13}$$

(i) Assume that $0 \notin \Lambda(A)$. This means that $a_{jd_j} \neq 0$ for all $j = 1, \dots, r$ and we can write

$$L_j = \det A \begin{bmatrix} -\frac{a_{jd_j-1}}{a_{jd_j}} & 1 & 0 & \cdots & 0 \\ -\frac{a_{jd_j-2}}{a_{jd_j}} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\frac{a_{j1}}{a_{jd_j}} & 0 & 0 & \cdots & 1 \\ -\frac{1}{a_{jd_j}} & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Taking into account the definition of $q_j(\lambda)$ of (8),

$$\begin{aligned} &\det(\lambda I_{d_j} - L_j) \\ &= \Delta_A^{d_j}\left(\frac{\lambda^{d_j}}{\Delta_A^{d_j}} + \frac{a_{jd_j-1}}{a_{jd_j}}\frac{\lambda^{d_j-1}}{\Delta_A^{d_j-1}} + \cdots + \frac{a_{j1}}{a_{jd_j}}\frac{\lambda}{\Delta_A} + \frac{1}{a_{jd_j}}\right) = q_j(\lambda). \end{aligned}$$

Let us see that $q_j(\lambda)$ is a power of an irreducible polynomial in $\mathbb{F}[\lambda]$. In fact, put

$$s_j(\lambda) = \lambda^{d_j} p_j\left(\frac{1}{\lambda}\right) = a_{jd_j}\lambda^{d_j} + a_{jd_j-1}\lambda^{d_j-1} + \cdots + a_{j1}\lambda + 1.$$

This polynomial is sometimes called the *reversal polynomial* of $p_j(\lambda)$ (see, for example, [23]). Since $p_j(\lambda)$ is an elementary divisor of A in $\mathbb{F}$, it is a power of an irreducible polynomial of $\mathbb{F}[\lambda]$. By [1, Lemma 4.4], $s_j(\lambda)$ is also a power of an irreducible polynomial. Now, it is not difficult to see that $q_j(\lambda) = \frac{1}{a_{jd_j}} s\left(\frac{\lambda}{\Delta_A}\right)$ is a power of an irreducible polynomial too. As a consequence, $q_1(\lambda), q_2(\lambda), \dots, q_r(\lambda)$ are the elementary divisors of $\mathrm{Adj}(C) = \oplus_{j=1}^r L_j$. Since this and $\mathrm{Adj}(A)$ are similar matrices (cf. (13)), $q_1(\lambda)$, $q_2(\lambda), \dots, q_r(\lambda)$ are the elementary divisors of $\mathrm{Adj}(A)$. This proves (i).

(ii) If $\mathrm{mg}(0) \geq 2$, then $\mathrm{rank}(A) = \mathrm{rank}(C) \leq n-2$. Hence all minors of A of order $n-1$ are equal to zero and so $\mathrm{Adj}(A) = 0$.

(iii) Assume now that $\mathrm{mg}(0) = 1$ and let $p_k(\lambda) = \lambda^{d_k}$ be the only elementary divisor of A with 0 as a root. Then, in (6), $a_{kj} = 0$ for $j = 1, \dots, d_k$. By (10) and (12), $C_k = \left[\begin{smallmatrix} 0 & 0 \\ I_{d_k-1} & 0 \end{smallmatrix}\right]$ and

$$\begin{aligned} L_k &= (-1)^{n-1} \prod_{j=1, j\neq k}^{r} a_{jd_j} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \\ &= (-1)^{n-1} \prod_{j=1, j\neq k}^{r} a_{jd_j} e_{d_k} e_1^T, \end{aligned} \tag{14}$$

respectively. Also, it follows from $a_{kd_k} = 0$ that $L_j = 0$ for $j = 1, \ldots, r$, $j \neq k$. Recall now that $S^{-1}AS = C = \oplus_{j=1}^{r} C_j$ and split S and S^{-1} accordingly:

$$S = \begin{bmatrix} S_1 & S_2 & \cdots & S_r \end{bmatrix}, \quad S^{-1} = \begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_r \end{bmatrix},$$

with $S_j \in \mathbb{F}^{n\times d_j}$ and $T_j \in \mathbb{F}^{d_j \times n}$, $j = 1, \ldots, r$. Then

$$AS_k = S_k C_k, \quad T_k A = C_k T_k. \tag{15}$$

For $i = 1, \ldots, d_k$ let s_{ki} and t_{ki}^T be the i-th column and row of S_k and T_k, respectively:

$$S_k = \begin{bmatrix} s_{k1} & s_{k2} & \cdots & s_{kd_k} \end{bmatrix}, \quad T_k = \begin{bmatrix} t_{k1}^T \\ t_{k2}^T \\ \vdots \\ t_{kd_k}^T \end{bmatrix}.$$

Bearing in mind that $\mathrm{Adj}(A) = S(\oplus_{j=1}^{r} L_j)S^{-1}$ (cf. (13)), the representation of L_k as a rank-one matrix of (14) and that $L_j = 0$ for $j \neq k$, we get

$$\mathrm{Adj}(A) = S_k L_k T_k = (-1)^{n-1} \left(\prod_{j=1, j\neq k}^{r} a_{jd_j} \right) s_{kd_k} t_{k1}^T. \tag{16}$$

Now, it follows from (15) that

$$\begin{aligned} & s_{kj} = As_{kj-1}, \qquad && t_{kj-1}^T = t_{kj}^T A, \qquad && j = 2, 3, \ldots, d_k, \\ & As_{kd_k} = 0, && t_{k1}^T A = 0. \end{aligned}$$

Henceforth, s_{kd_k} and t_{k1}^T are right and left eigenvectors of A for the eigenvalue 0, s_{k1} is a generating vector of $\ker A^{d_k} =< s_{k1}, As_{k1}, \ldots, A^{d_k-1} s_{k1} >=< s_{k1}, s_{k2}, \ldots, s_{kd_k} >$ and t_{kd_k} is a generating vector of $\ker(A^T)^{d_k} =< t_{kd_k}, A^T t_{kd_k}, \ldots, (A^T)^{d_k-1} t_{kd_k} >=< t_{kd_k}, t_{kd_k-1}, \ldots, t_{k1} >$. Thus (16) is an explicit rank-one representation of $\operatorname{Adj}(A)$ in terms of a right and a left eigenvectors of A for the eigenvalue zero. Actually this representation depends on a particular normalization of the vectors which span the cyclic subspaces $\ker A^{d_k}$ and $\ker(A^T)^{d_k}$. Specifically, $T_k S_k = I_{d_k}$. However, we are looking for a more general representation in terms of arbitrary right and left eigenvectors for which such a normalization may not hold. Let us assume that $u, v \in \mathbb{F}^{n\times 1}$ are arbitrary right and left eigenvectors of A for the eigenvalue 0 and let $x, y \in \mathbb{F}^{n\times 1}$ be arbitrary right and left generating vectors of $\ker A^{d_k}$ and $\ker(A^T)^{d_k}$ for u and v, respectively. Then $\{x, Ax, \dots, A^{d_k-2}x, u\}$ and $\{y, A^T y, \dots, (A^T)^{d_k-2}y, v\}$ are Krylov bases of $\ker A^{d_k}$ and $\ker(A^T)^{d_k}$, respectively. By Lemma 2.5, there are invertible lower triangular Toeplitz matrices X and Y such that

$$\begin{bmatrix} x & Ax & \cdots & A^{d_k-2}x & u \end{bmatrix} = \begin{bmatrix} s_{k1} & s_{k2} & \cdots & s_{kd_k-1} & s_{kd_k} \end{bmatrix} X$$

and

$$\begin{bmatrix} y & A^T y & \cdots & (A^T)^{d_k-2}y & v \end{bmatrix} = \begin{bmatrix} t_{kd_k} & t_{kd_k-1} & \cdots & t_{k2} & t_{k1} \end{bmatrix} Y.$$

Equivalently,

$$\begin{bmatrix} v & (A^T)^{d_k-2}y & \cdots & A^T y & y \end{bmatrix} = \begin{bmatrix} t_{k1} & t_{k2} & \cdots & t_{kd_k-1} & t_{kd_k} \end{bmatrix} Y^T.$$

Let

$$X = \begin{bmatrix} \alpha_1 & & & & \\ \alpha_2 & \alpha_1 & & & \\ \vdots & \vdots & \ddots & & \\ \alpha_{d_k-1} & \alpha_{d_k-2} & \cdots & \alpha_1 & \\ \alpha_{d_k} & \alpha_{d_k-1} & \cdots & \alpha_2 & \alpha_1 \end{bmatrix}, \quad Y^T = \begin{bmatrix} \beta_1 & \beta_2 & \cdots & \beta_{d_k-1} & \beta_{d_k} \\ & \beta_1 & \cdots & \beta_{d_k-2} & \beta_{d_k-1} \\ & & \ddots & \vdots & \vdots \\ & & & \beta_1 & \beta_2 \\ & & & & \beta_1 \end{bmatrix}$$

and note that $u = \alpha_1 s_{kd_k}$ and $v = \beta_1 t_{k1}$ with $\alpha_1 \neq 0$ and $\beta_1 \neq 0$. It follows from (16) that

$$\operatorname{Adj}(A) = (-1)^{n-1} \left(\prod_{j=1, j\neq k}^{r} a_{jd_j} \right) \frac{uv^T}{\alpha_1 \beta_1}. \tag{17}$$

Since $T_k = \begin{bmatrix} t_{k1} & t_{k2} & \cdots & t_{kd_k-1} & t_{kd_k} \end{bmatrix}^T$, $S_k = \begin{bmatrix} s_{k1} & s_{k2} & \cdots & s_{kd_k-1} & s_{kd_k} \end{bmatrix}$ and $T_k S_k = I_{d_k}$,

$$\begin{bmatrix} v^T \\ y^T A^{d_k-2} \\ \vdots \\ y^T A \\ y^T \end{bmatrix} \begin{bmatrix} x & Ax & \cdots & A^{d_k-2}x & u \end{bmatrix} = Y T_k S_k X = YX.$$

But YX is a lower triangular matrix whose diagonal elements are all equal to $\alpha_1\beta_1$. Thus, for $j = 1, \ldots, d_k$, $\alpha_1\beta_1 = v^T x = y^T u = y^T A^{d_k-1}x$. Since $\alpha_1 \neq 0$ and $\beta_1 \neq 0$, $y^T A^{d_k-1}x \neq 0$ as claimed. Now, from (17)

$$\operatorname{Adj}(A) = (-1)^{n-1} \left(\prod_{j=1, j\neq k}^{r} a_{jd_j} \right) \frac{uv^T}{y^T A^{d_k-1}x}. \tag{18}$$

Finally, $p_A(\lambda) = \prod_{j=1}^{r} p_j(\lambda) = \lambda^{d_k} \prod_{j=1, j\neq k}^{r} p_j(\lambda)$ with $p_j(0) = a_{jd_j} \neq 0$ for $j \neq k$. Therefore (18) is equivalent to (9) and the theorem follows. $\square$

As a first consequence of Theorem 2.6 we present a generalization of the formula for the eigenvalues of the adjugate matrix (see [19]).

Corollary 2.7. *Let $A \in \mathbb{F}^{n\times n}$ be a nonsingular matrix. Let $\lambda_0 \in \Lambda(A)$ and let $m_1 \geq \ldots \geq m_s$ be its partial multiplicities (i.e., the sizes of the Jordan blocks associated to λ_0 in any Jordan form of A in, perhaps, an extension field $\widetilde{\mathbb{F}}$). Then $\frac{\Delta_A}{\lambda_0}$ is an eigenvalue of* $\operatorname{Adj}(A)$ *with $m_1 \geq \ldots \geq m_s$ as partial multiplicities.*

Proof. The elementary divisors of A for the eigenvalue λ_0 in $\widetilde{\mathbb{F}}(\lambda)$ are $(\lambda - \lambda_0)^{m_1}, \ldots, (\lambda - \lambda_0)^{m_s}$. Then, it follows from item (i) of Theorem 2.6 (see (8)) that $\left(\lambda - \frac{\Delta_A}{\lambda_0}\right)^{m_1}, \ldots, \left(\lambda - \frac{\Delta_A}{\lambda_0}\right)^{m_s}$ are the corresponding elementary divisors of $\operatorname{Adj}(A)$. $\square$

Assume now that $\lambda_0 \in \Lambda(A) \cap \mathbb{F}$ and $m_1 \geq \ldots \geq m_s$. If $s > 1$ then $\operatorname{rank}(\lambda_0 I_n - A) \leq n - 2$ and so $\operatorname{Adj}(\lambda_0 I_n - A) = 0$. For $s = 1$ we have the following result.

Corollary 2.8. *Let $A \in \mathbb{F}^{n\times n}$ and let the polynomials $p_j(\lambda)$ of (6) be its elementary divisors, $j = 1, \ldots, r$. Assume that $\lambda_0 \in \Lambda(A) \cap \mathbb{F}$ is an eigenvalue of A such that* $\operatorname{mg}(\lambda_0) = 1$ *and $p_k(\lambda_0) = 0$. Let $u, v \in \mathbb{F}^{n\times 1}$ be arbitrary right and left eigenvectors of A for λ_0 and let $x, y \in \mathbb{F}^{n\times 1}$ be right and left generating vectors of* $\ker(\lambda_0 I_n - A)^{d_k}$ *and* $\ker((\lambda_0 I_n - A)^T)^{d_k}$ *for the eigenvectors u and v, respectively. Then*

$$\operatorname{Adj}(\lambda_0 I_n - A) = (-1)^{d_k+1} \prod_{j=1, j\neq k}^{r} p_j(\lambda_0) \frac{uv^T}{y^T(\lambda_0 I_n - A)^{d_k-1}x}. \tag{19}$$

Proof. Put $B = \lambda_0 I_n - A$. Then $0 \in \Lambda(B)$, u and v are right and left eigenvectors of B for the eigenvalue 0, $\operatorname{mg}(0) = 1$ and $\operatorname{ma}(0) = d_k$ are the geometric and algebraic multiplicities of this eigenvalue and x and y are right and left generating vectors of $\ker B^{d_k}$ and $\ker(B^T)^{d_k}$ for the eigenvectors u and v. Also, for $j = 1, \ldots, r$, (recall that we are taken the elementary divisors to be monic polynomials) $q_j(\lambda) = (-1)^{d_j} p_j(\lambda_0 - \lambda)$ are the elementary divisors of B. We get from $p_k(\lambda) = (\lambda - \lambda_0)^{d_k}$ that $q_k(\lambda) = \lambda^{d_k}$. By Theorem 2.6,

$$\mathrm{Adj}(\lambda_0 I_n - A) = \mathrm{Adj}(B) = (-1)^{n-1} \prod_{j=1, j\neq k}^{r} q_j(0) \frac{uv^T}{y^T(\lambda_0 I_n - A)^{d_k-1}x}.$$

Therefore (19) follows from $q_j(0) = (-1)^{d_j} p_j(\lambda_0)$ and the fact that $d_1 + \cdots + d_r = n$. □

The following result is an immediate consequence of Corollary 2.8.

Corollary 2.9. *Let $A \in \mathbb{F}^{n\times n}$ and let $\Lambda(A) = \{\lambda_1, \ldots, \lambda_r\}$ be its spectrum. Assume that $\Lambda(A) \subset \mathbb{F}$ and let m_j and g_j be the algebraic and geometric multiplicities of A for the eigenvalue λ_j, $j = 1, \ldots, r$.*

(i) If $g_k > 1$ for some $k \in \{1, \ldots, r\}$ then $\mathrm{Adj}(\lambda_k I - A) = 0$, and
(ii) if $g_k = 1$ for some $k \in \{1, \ldots, r\}$ then

$$\mathrm{Adj}(\lambda_k I - A) = (-1)^{m_k+1} \prod_{j=1,\, j\neq k}^{r} (\lambda_k - \lambda_j)^{m_j} \frac{u_k v_k^T}{y_k^T A^{m_k-1} x_k}, \tag{20}$$

where u_k and v_k are right and left eigenvectors of A for λ_k and x_k and y_k are right and left generating vectors of $\ker(\lambda_k I_n - A)^{m_k}$ and $\ker((\lambda_k I_n - A)^T)^{m_k}$ for the eigenvectors u_k and v_k, respectively.

Remark 2.10. When $d_k = 1$ in Corollary 2.8, (19) becomes

$$\mathrm{Adj}(\lambda_k I_n - A) = \prod_{j=1, j\neq k}^{r} p_j(\lambda_0) \frac{uv^T}{v^T u}$$

because, in this case $x = u$ and $y = v$. Since $p'_A(\lambda_k) = \prod_{j=1,j\neq k}^{r} p_j(\lambda_k)$ we conclude that (19) generalizes (1) for matrices over fields. Derivatives can be also used in (19) to produce an expression similar to that of (1), but we must "pay a price". In fact, if $p_A^{(d_k)}(\lambda_k)$ denotes the d_k derivative of $p_A(\lambda)$ and $\mathrm{ma}(\lambda_k) = d_k > 1$ then $\prod_{j=1,j\neq k}^{r} p_j(\lambda_0) = \frac{1}{d_k!} p_A^{(d_k)}(\lambda_k)$ when this expression makes sense; that is to say, provided that $d_k! \neq 0$. If $\mathbb{F}$ is required to be a field of characteristic zero, this is always guaranteed. In other words, if $\mathbb{F}$ is a field of characteristic zero then, under the hypothesis of Corollary 2.8, (19) is equivalent to

$$\mathrm{Adj}(\lambda_0 I_n - A) = \frac{(-1)^{d_k+1}}{d_k!} p_A^{(d_k)}(\lambda_0) \frac{uv^T}{y^T(\lambda_0 I_n - A)^{d_k-1}x}, \tag{21}$$

which, formally, looks like the natural generalization of (1). □

The TM formula (1) can be used to provide an easy proof of the so-called eigenvector-eigenvalue identity (see [13, Sec. 2.1]). In fact, under the hypothesis of Theorem 2.1, it follows from (1) that $w^T v[\mathrm{Adj}(\lambda_0 I_n - A)]_{jj} = p'_A(\lambda_0) v_j w_j$, $j = 1, \ldots, n$.

Hence, if M_j is the submatrix of A obtained by removing its jth row and column, then $p_{M_j}(\lambda_0) = \det(\lambda_0 I_{n-1} - M_j) = [\operatorname{Adj}(\lambda_0 I_n - A)]_{jj}$ Therefore

$$(w^T v)\, p_{M_j}(\lambda_0) = p'_A(\lambda_0)\, v_j w_j, \quad j = 1, \dots, n. \tag{22}$$

In particular, if $A \in \mathbb{C}^{n\times n}$ is Hermitian, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ are its eigenvalues and, for $i = 1, \dots, n$, $v_i = \begin{bmatrix} v_{i1} & v_{i2} & \cdots & v_{in} \end{bmatrix}^T$ is a unitary right and left eigenvector of A for λ_i; that is $Av_i = \lambda_i v_i$, $v_i^* A = \lambda_i v_i^*$ and $v_i^* v_i = 1$; (recall that we must change *transpose* by *conjugate transpose* in the complex case) then

$$|v_{ij}|^2 p'_A(\lambda_i) = p_{M_j}(\lambda_i), \quad i, j = 1, \dots, n.$$

Equivalently, if $\mu_{j1} \geq \mu_{j2} \geq \cdots \geq \mu_{jn-1}$ are the eigenvalues of M_j,

$$|v_{ij}|^2 \prod_{k=1, k\neq i}^{n} (\lambda_i - \lambda_k) = \prod_{k=1}^{n} (\lambda_i - \mu_{jk}) \quad i, j = 1, \dots, n. \tag{23}$$

This is the classical eigenvector-eigenvalue identity (see [13, Thm. 1]).

As mentioned in Remark 2.4, if $\mathbb{F}$ is a field and $A \in \mathbb{F}^{n\times n}$ then (22) is meaningful if λ_0 is a simple eigenvalue. If λ_0 is defective and its geometric multiplicity is bigger than 1 then (22) becomes a trivial identity because, in this case, $\operatorname{Adj}(\lambda_0 I_n - A) = 0$ (item (ii) of Theorem 2.6) and so $p_{M_j}(\lambda_0) = \det(\lambda_0 I_{n-1} - M_j) = 0$. However, if λ_0 is defective and its geometric multiplicity is 1, then (19) can be used to obtain a generalization of the eigenvector-eigenvalue identity. In fact, one readily gets from (19):

$$p_{M_j}(\lambda_0) = (-1)^{d_k+1} \prod_{j=1, j\neq k}^{r} p_j(\lambda_0)\, \frac{u_j v_j^T}{y^T(\lambda_0 I_n - A)^{d_k-1} x} \tag{24}$$

where $p_1(\lambda), \dots, p_r(\lambda)$ of (6) are the elementary divisors of A with $p_k(\lambda) = (\lambda - \lambda_0)^{d_k}$ and u_j, v_j are the jth components of u, v respectively.

Moreover, if both $p_A(\lambda)$ and $p_{M_j}(\lambda)$ split in $\mathbb{F}$ then, with the notation of Corollary 2.9, the following identity follows from (20) for the non-repeated eigenvalues $\{\mu_{j1}, \dots, \mu_{js_j}\}$ of M_j and provided that $\operatorname{mg}(\lambda_i) = 1$ for $i = 1, \dots, r$:

$$\prod_{k=1}^{s_k} (\lambda_i - \mu_{jk})^{q_{jk}} = (-1)^{d_i+1} \frac{u_{ij} v_{ij}}{y_i^T A^{d_i-1} x_i} \prod_{k=1, k\neq i}^{r} (\lambda_i - \lambda_k)^{m_k}, \quad j = 1, \dots, n, \tag{25}$$

where $u_i = \begin{bmatrix} u_{i1} & \cdots & u_{in} \end{bmatrix}^T$, $v_i = \begin{bmatrix} v_{i1} & \cdots & v_{in} \end{bmatrix}^T$, and q_{jk} is the algebraic multiplicity of μ_{jk}, $k = 1, \dots, s_j$ and $j = 1, \dots, n$.

In the following section two additional applications will be presented.

3. Two additional consequences of the TM formula

The well-known formula (26) below gives the derivative of a simple eigenvalue of a matrix depending on a (real or complex) parameter. The investigation about the eigenvalue sensitivity of matrices depending on one or several parameters can be traced back to the work of Jacobi ([20]). However a systematic study of the perturbation theory of the eigenvalue problem starts with the books of Rellich (1953), Wilkinson (1965) and Kato (1966), as well as the papers by Lancaster [21], Osborne and Michaelson [28], Fox and Kapoor [15], Crossley and Porter [10] (see also [32] and the references therein). Since then this topic has become classical as evidenced by an extensive literature including books and papers addressed to mathematicians and a broad spectrum of scientist and engineers. In addition to the above early references, a short, and by no means exhaustive, list of books could include [4, p. 463], [25, Ch. 8, Sec. 9], [11, Sec. 4.2] or [22, pp. 134-135].

In proving (26), one first must prove, of course, that the eigenvalues smoothly depend on the parameter. It is also a common practice to prove or assume (see [24], [14, Ch. 11, Th. 2] and the referred books), the existence of eigenvectors which depend smoothly on the parameter. It is worth-remarking that in the proof by Lancaster in [21] only the existence of eigenvectors continuously depending on the parameter is required. We propose a simple and alternative proof of (26) where no assumption is made on the right and left eigenvector functions.

Let $D_\epsilon(z_0)$ be the open disc of radius $\epsilon > 0$ with center z_0. For the following result $\mathbb{F}$ will be either the field of real numbers $\mathbb{R}$ or of the complex numbers $\mathbb{C}$. Recall our convention that $v \in \mathbb{C}^{n\times 1}$ is a left eigenvector of $A \in \mathbb{C}^{n\times n}$ for an eigenvalue z_0 if $v^*A = z_0 v^*$ where $v^* = \bar{v}^T$ is the transpose conjugate of v.

Proposition 3.1. *Let $A(\omega) \in \mathbb{F}^{n\times n}$ be a square matrix-valued function whose entries are analytic at $\omega_0 \in \mathbb{C}$. Let z_0 be a simple eigenvalue of $A(\omega_0)$. Then there exist $\epsilon > 0$ and $\delta > 0$ so that $z : D_\epsilon(\omega_0) \to D_\delta(z_0)$ is the unique eigenvalue of $A(\omega)$ with $z(\omega) \in D_\delta(z_0)$ for each $\omega \in D_\epsilon(\omega_0)$. Moreover, z is analytic on $D_\epsilon(\omega_0)$ and*

$$z'(\omega) = \frac{v(\omega)^* A'(\omega) u(\omega)}{v(\omega)^* u(\omega)}, \quad (26)$$

where, for $w \in D_\epsilon(\omega_0)$, $u(\omega)$ and $v(\omega)$ are arbitrary right and left eigenvector, respectively, of A for $z(\omega)$.

Proof. Since z_0 is a simple root of $p(z,\omega) = \det(zI - A(\omega))$, by the analytic implicit function theorem, we have, in addition to the first part of the result, that

$$z'(\omega) = -\frac{\dfrac{\partial p}{\partial \omega}(z(\omega),\omega)}{\dfrac{\partial p}{\partial z}(z(\omega),\omega)}.$$

By the Jacobi formula for the derivative of the determinant and TM formula (1), we have (note that since $z(\omega)$ is a simple eigenvalue, $v(\omega)^* u(\omega) \neq 0$ for any right and left eigenvectors $u(\omega)$ and $v(\omega)$)

$$\begin{aligned}\frac{\partial p}{\partial z}(z(\omega), \omega) &= \operatorname{tr}(\operatorname{Adj}(z(\omega)\, I - A(\omega))\\ &= p'(z(\omega), \omega)\\ \frac{\partial p}{\partial \omega}(z(\omega), \omega) &= -\operatorname{tr}(\operatorname{Adj}(z(\omega)\, I - A(\omega))A'(\omega))\\ &= -p'(z(\omega), \omega)\frac{v(\omega)^* A'(\omega) u(\omega)}{v(\omega)^* u(\omega)},\end{aligned}$$

and the result follows. □

Remark 3.2.

(a) The same conclusion can be drawn in Proposition 3.1 if A is a complex or real matrix-valued differentiable function of a real variable. In the first case, we would need a non-standard version of the implicit function theorem like the one in [3, Theorem 2.4]. In the second case the standard implicit function theorem is enough.

(b) It is shown in [2] that the existence of eigenvectors smoothly depending on the parameter can be easily obtained from the properties of the adjugate matrix. In fact, since $z(\omega)$ is a simple eigenvalue of $A(\omega)$ for each $\omega \in D_\epsilon(\omega_0)$, $\operatorname{rank}(z(\omega)I_n - A(\omega)) = n-1$ and so by the TM formula, $\operatorname{rank}\operatorname{Adj}(z(\omega)I_n - A(\omega)) = 1$ (see Remark 2.4). Now, $\operatorname{Adj}(z(\omega)I_n - A(\omega))$ is a differentiable matrix function of $\omega \in D_\epsilon(\omega_0)$ and $(z(\omega)I_n - A(\omega))(\operatorname{Adj}(z(\omega)I_n - A(\omega))) = (\operatorname{Adj}(z(\omega)I_n - A(\omega)))(z(\omega)I_n - A(\omega)) = \det(z(\omega)I_n - A(\omega))I_n = 0$. Henceforth, all nonzero columns and rows of $\operatorname{Adj}(z(\omega)I_n - A(\omega)$, which are all proportional, are right and left eigenvectors of $A(\omega)$ for $z(\omega)$, respectively. □

The second application is related to the problem of characterizing the admissible eigenstructures and, more generally, the similarity orbits of the rank-one updated matrices. There is a vast literature on this problem. A non-exhaustive list of publications is [33,30,35,27,6,26,8,5] and the references therein. It is a consequence of Theorem 2 in [33] (see also [27] and [26]) that if λ_0 is an eigenvalue of $A \in \mathbb{F}^{n\times n}$ with geometric multiplicity 1 and $\operatorname{rank}(B - A) = 1$ then λ_0 may or may not be an eigenvalue of $B \in \mathbb{F}^{n\times n}$. It is then proved in [27, Th. 2.1] and [26, Th. 2.3] (see also the references therein) that in the complex case, generically, λ_0 is not an eigenvalue of B. That is to say, there is a Zariski open set $\Omega \subset \mathbb{C}^n \times \mathbb{C}^n$ such that for all $(x, y) \in \Omega$, λ_0 is not an eigenvalue of $A + xy^T$. With the help of the TM formula we can be a little more precise about the set Ω. From now on, $\mathbb{F}$ will be again an arbitrary field. For $\mathbb{F} = \mathbb{C}$, the result which follows can be obtained as a consequence of [27, Th. 2.1].

Proposition 3.3. *Let $A \in \mathbb{F}^{n\times n}$ and let λ_0 be an eigenvalue of A in, perhaps, an extension field $\widetilde{\mathbb{F}}$. Assume that the geometric multiplicity of λ_0 is 1 and its algebraic multiplicity is m. Let $u_0, v_0 \in \mathbb{F}^{n\times 1}$ be right and left eigenvectors of A for λ_0. If $x, y \in \mathbb{F}^{n\times 1}$ then λ_0 is an eigenvalue of $A + xy^T$ if and only if $y^T u_0 = 0$ or $v_0^T x = 0$.*

Proof. Let $B = A+xy^T$. Then $\lambda I_n - A = \lambda I_n - B - xy^T$. Taking into account that $\lambda I_n - B$ is invertible in $\mathbb{F}(s)^{n\times n}$, where $\mathbb{F}(s)$ the field of rational functions, we can use the formula of the determinant of updated rank-one matrices ($\det(L + xy^T) = (1 + y^T L^{-1}x)\det L$ provided that L is invertible) to get

$$p_B(\lambda) = p_A(\lambda) + p_A(\lambda)y^T(\lambda I_n - A)^{-1}x = p_A(\lambda) + y^T \operatorname{Adj}(\lambda I_n - A)x.$$

In particular,

$$p_B(\lambda_0) = y^T \operatorname{Adj}(\lambda_0 I_n - A)x. \tag{27}$$

It follows from (19) that (recall that if w, and z are right and left generating vectors of $\ker(\lambda_0 I_n - A)^m$ and $\ker((\lambda_0 I_n - A)^T)^m$ for the eigenvectors u_0 and v_0, respectively, then $z^T(\lambda_0 I_n - A)^{m-1}w \neq 0$)

$$p_B(\lambda_0) = (-1)^{m+1} \prod_{j=1, j\neq k}^{r} p_j(\lambda_0) \, \frac{uv^T}{z^T(\lambda_0 I_n - A)^{m-1}w},$$

where $p_1(\lambda), \dots, p_r(\lambda)$ of (6) are the elementary divisors of A with $p_k(\lambda) = (\lambda - \lambda_0)^m$. Since $\prod_{j=1,j\neq k}^{r} p_j(\lambda_0) \neq 0$, the Proposition follows. □

Remark 3.4.

(i) Note that, by (27) and item (ii) of Theorem 2.6, if the geometric multiplicity of λ_0 as eigenvalue of A is 2 then $\operatorname{Adj}(\lambda_0 I_n - A) = 0$ and so, λ_0 is necessarily an eigenvalue of $A + xy^T$. This is an easy consequence of the interlacing inequalities of [33, Th. 2] (and also of [27, Th. 2.1], for example). However, proving that those interlacing inequalities are necessary conditions that the invariant polynomials of A and $A + xy^T$ must satisfy is by no means a trivial matter.

(ii) An easy alternative proof of Proposition 3.3, which does not use the TM formula, was offered by one of the anonymous referees: Setting $C = A - \lambda_0 I_n$, we can assume that u_0 and v_0 are right and left eigenvectors of C, respectively, for the eigenvalue 0. The "if" part is obvious: if $y^T u_0 = 0$ then u_0 is a right eigenvector of $C + xy^T$ for the eigenvalue 0. And if $v_0^T x = 0$ then v_0 is a left eigenvector of $C + xy^T$ for that eigenvalue. Assume now that v is a right eigenvector of $C + xy^T$ for the eigenvalue 0. Then $Cv = -xy^T v$ and, since v_0 is a left eigenvector of C for the eigenvalue 0, $v_0^T Cv = 0$. Hence $v_0^T xy^T v = 0$ implying that $v_0^T x = 0$ or $y^T v = 0$. If $v_0^T x = 0$ we are done and, otherwise, $Cv = -xy^T v = 0$. This means that v is a right eigenvector of C for the eigenvalue 0. It follows from $\mathrm{mg}(0) = 1$ that all right eigenvectors are proportional and so $y^T u_0 = 0$. $\square$

The eigenvalues of rank-one updated matrices are at the core of the *divide and conquer algorithm* to compute the eigenvalues of real symmetric or complex hermitian matrices (see, for example, [12, Sec. 5.3.3], [31, Sec. 2.1]). At each step of the algorithm a diagonal matrix $D = D_1 \oplus D_2$ and a vector $u \in \mathbb{C}^{n\times 1}$ are given such that the eigenvalues and eigenvectors of $D + uu^*$ are to be computed. In order the algorithm to run smoothly, it is required, among other things, that the diagonal elements of D are all distinct. Thus, a so-called *deflation* process must be carried out. This amounts to check at each step the presence of repeated eigenvalues and, if so, remove and save them. The result that follows is related to the problem of detecting repeated eigenvalues but for much more general matrices over fields. For $\mathbb{F} = \mathbb{C}$ it can be obtained from [9, Lem. 2.1].

Proposition 3.5. *Let $A = A_1 \oplus A_2$ with $A_i \in \mathbb{F}^{n_i\times n_i}$, $i = 1, 2$. Let $x, y \in \mathbb{F}^{n\times 1}$ and split $B = A + xy^T = [B_{ij}]_{ij=1,2}$ into 2×2 blocks such that $B_{ii} \in \mathbb{F}^{n_i\times n_i}$, $i = 1, 2$. Assume also that the eigenvalues of A_1 and A_2 have geometric multiplicity equal to 1 and $\Lambda(A_1) \cap \Lambda(B_{11}) = \Lambda(A_2) \cap \Lambda(B_{22}) = \emptyset$. Then*

$$\Lambda(A_1) \cap \Lambda(A_2) = \Lambda(B) \cap \Lambda(A_1) = \Lambda(B) \cap \Lambda(A_2).$$

Proof. If $\lambda_0 \in \Lambda(A_1) \cap \Lambda(A_2)$ then λ_0, as eigenvalue of A, has geometric multiplicity 2. By Remark 3.4, $\lambda_0 \in \Lambda(B) \cap \Lambda(A_1) \cap \Lambda(A_2)$. Assume that $\lambda_0 \in \Lambda(B) \cap \Lambda(A_1)$ but $\lambda_0 \notin \Lambda(A_2)$. Let us see that this assumption leads to a contradiction. Let $u_0, v_0 \in \mathbb{F}^{n_1\times 1}$ be a right and a left eigenvectors of A_1, respectively. Then $w_0 = \begin{bmatrix} u_0^T & 0 \end{bmatrix}^T \in \mathbb{F}^{n\times 1}$ and $z_0 = \begin{bmatrix} w_0^T & 0 \end{bmatrix}^T \in \mathbb{F}^{n\times 1}$ are right and left eigenvectors of A, respectively, for λ_0. Since $\lambda_0 \notin \Lambda(A_2)$, the geometric multiplicity of λ_0 as eigenvalue of A is 1. Then, by Proposition 3.3, $y^T w_0 = 0$ or $z_0^T x = 0$ because $\lambda_0 \in \Lambda(B)$. Let us assume that $y^T w_0 = 0$, on the contrary we would proceed similarly with $z_0^T x = 0$. If we put $y = \begin{bmatrix} y_1^T & y_2^T \end{bmatrix}^T$ and $x = \begin{bmatrix} x_1^T & x_2^T \end{bmatrix}^T$, with $x_1, y_1 \in \mathbb{F}^{n_1\times 1}$, then $y_1^T u_0 = 0$ and $B_{11} = A_{11} + x_1 y_1^T$. It follows from Proposition 3.3 that $\lambda_0 \in \Lambda(B_{11})$, contradicting the hypothesis $\Lambda(A_1) \cap \Lambda(B_{11}) = \emptyset$. That $\Lambda(B) \cap \Lambda(A_2) \subset \Lambda(A_1) \cap \Lambda(A_2)$ is proved similarly. $\square$

Remark 3.6.

(i) Note that, with the notation of the proof of Proposition 3.5, $B_{11} = A_1 + x_1 y_1^T$ and $B_{22} = A_2 + x_2 y_2^T$. Then, according to Proposition 3.3, $\lambda_0 \notin \Lambda(B_{11})$ unless $(y_1^T u_0)(v_0^T x_1) = 0$. Hence, the hypothesis $\Lambda(A_1) \cap \Lambda(B_{11}) = \emptyset$ is a generic property, and so is $\Lambda(A_2) \cap \Lambda(B_{22}) = \emptyset$.

(ii) Consider Proposition 3.5 over $\mathbb{C}$. If A and B are both Hermitian or unitary, then $\Lambda(B)\backslash\big(\Lambda(A_1)\cap\Lambda(A_2)\big)$ and $\Lambda(A_1)\cup\big(\Lambda(A_2)\backslash(\Lambda(A_1)\cap\Lambda(A_2))\big)$ strictly interlace on the real line or the unit circle, respectively (see, for example, [31, Th. 2.1, Sec. 2]). $\square$

Declaration of competing interest

There is no competing interest.

Acknowledgement

Thanks are given to anonymous referees whose comments and suggestions helped us to improve the results and presentation of the paper.

References

[1] A. Amparan, S. Marcaida, I. Zaballa, On the structure invariants of proper rational matrices with prescribed finite poles, Linear Multilinear Algebra 61 (11) (2013) 1464–1486.
[2] A.L. Andrew, K.-W.E. Chu, P. Lancaster, Derivatives of eigenvalues and eigenvectors of matrix functions, SIAM J. Matrix Anal. Appl. 14 (4) (1993) 903–926.
[3] M.S. Ashbaugh, E.M. Harrell II, Perturbation theory for shape resonances and large barrier potentials, Commun. Math. Phys. 83 (2) (1982) 151–170.
[4] F.V. Atkinson, Discrete and Continuous Boundary Problems, Mathematics in Science and Engineering, vol. 8, Academic Press, New York-London, 1964.
[5] I. Baragaña, The number of distinct eigenvalues of a regular pencil and of a square matrix after rank perturbation, Linear Algebra Appl. 588 (2020) 101–121.
[6] M.A. Beitia, I. de Hoyos, I. Zaballa, The change of the Jordan structure under one row perturbations, Linear Algebra Appl. 401 (2005) 119–134.
[7] W.C. Brown, Matrices over Commutative Rings, Marcel Dekker Inc., New York, 1993.
[8] R. Bru, R. Cantó, A.M. Urbano, Eigenstructure of rank one updated matrices, Linear Algebra Appl. 485 (2015) 372–391.
[9] K. Castillo, J. Petronilho, Refined interlacing properties for zeros of paraorthonormal polynomials on the unit circle, Proc. Am. Math. Soc. 146 (8) (2018) 3285–3294.
[10] T.R. Crossley, B. Porter, Eigenvalue and eigenvector sensitivities in linear system theory, Int. J. Control 10 (1969) 163–170.
[11] D. Hinrichsen, A.J. Pritchard, Mathematical System Theory I. Modelling, State Space Analysis, Stability and Robustness, Springer, Berlin, 2005.
[12] J.W. Demmel, Applied Numerical Linear Algebra, SIAM, Philadelphia, 1997.
[13] P.B. Denton, S.J. Parke, T. Tao, X. Zhang, Eigenvectors from eigenvalues: a survey of a basic identity in linear algebra, Bull. Am. Math. Soc. (2021).
[14] L.C. Evans, Partial Differential Equations, second edition, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 2010.
[15] R.L. Fox, M.P. Kapoor, Rate of change of eigenvalues and eigenvectors, AIAA J. 6 (1968) 2426–2429.
[16] F.R. Gantmacher, The Theory of Matrices, AMS Chelsea Publishing, Providence, Rhode Island, 1988.
[17] R. Godement, Cours d'algèbre, Hermann Éditeurs, Paris, 2005.
[18] D. Grinberg, Eigenvectors from eigenvalues: a survey of a basic identity in linear algebra — what's new, https://terrytao.wordpress.com/2019/12/03/eigenvectors-from-eigenvalues-a-survey-of-a-basic-identity-in-linear-algebra/#comment-531597, 2019.
[19] R.D. Hill, E.E. Underwood, On the matrix adjoint (adjugate), SIAM J. Algebraic Discrete Methods 6 (4) (1985) 731–737.
[20] C.G.J. Jacobi, Über ein leichtes verfahren die in der theorie der säcularstörungen vorkommenden gleichungen numerisch aufzulösen, J. Reine Angew. Math. 1846 (30) (1846) 51–94.
[21] P. Lancaster, On eigenvalues of matrices dependent on a parameter, Numer. Math. 6 (1964) 377–387.

[22] P.D. Lax, Linear Algebra and Its Applications, second edition, Pure and Applied Mathematics (Hoboken), Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, 2007.
[23] D.S. Mackey, N. Mackey, C. Mehl, V. Mehrmann, Vector spaces of linearizations for matrix polynomials, SIAM J. Matrix Anal. Appl. 28 (4) (2006) 971–1004.
[24] J.R. Magnus, On differentiating eigenvalues and eigenvectors, Econom. Theory 1 (1985) 179–191.
[25] J.R. Magnus, H. Neudecker, Matrix Differential Calculus with Applications in Statistics and Econometrics, John Wiley & Sons, Chichester, 1988.
[26] C. Mehl, V. Mehrmann, A.C.M. Ran, L. Rodman, Eigenvalue perturbation theory of classes of structured matrices under generic structured rank one perturbations, Linear Algebra Appl. 435 (3) (2011) 687–716.
[27] J. Moro, F.M. Dopico, Low rank perturbation of Jordan structure, SIAM J. Matrix Anal. Appl. 25 (2) (2003) 495–506.
[28] M.R. Osborne, S. Michaelson, The numerical solution of eigenvalue problems in which the eigenvalue appears nonlinearly, with an application to differential equations, Comput. J. 7 (1964) 66–71.
[29] D.S. Scott, How to make the Lanczos algorithm converge slowly, Math. Comput. 33 (1979) 239–247.
[30] F.C. Silva, The rank of the difference of matrices with prescribed similarity classes, Linear Multilinear Algebra 24 (1) (1988) 51–58.
[31] G.W. Stewart, Matrix Algorithms, Volume II: Eigensystems, SIAM, Philadelphia, 2001.
[32] J.G. Su, Multiple eigenvalue sensitivity analysis, Linear Algebra Appl. 137 (4) (1990) 183–211.
[33] R.C. Thompson, Invariant factors under rank one perturbations, Can. J. Math. 32 (1) (1980) 240–245.
[34] R.C. Thompson, P. McEnteggert, Principal submatrices. II: the upper and lower quadratic inequalities, Linear Algebra Appl. 1 (1968) 211–243.
[35] I. Zaballa, Pole assignment and additive perturbations of fixed rank, SIAM J. Matrix Anal. Appl. 12 (1) (1991) 16–23.