On Faddeev-Leverrier's Method for the Computation of the Characteristic Polynomial of a Matrix and of Eigenvectors

Gilbert Helmberg *, **Gerhard Veltkamp** ** and **Peter Wagner** *

* Institut für Mathematik und Geometrie, Universität Innsbruck,
Technikerstr. 13, A-6020 Innsbruck, Austria

** Faculteit Wiskunde en Informatica, Technische Universiteit Eindhoven
Den Dolech 2, 5612 AZ Eindhoven, Netherlands

Version 12: 8. 4. 1992

Abstract. Faddeev's method of computing the eigenvalues and eigenvectors of a matrix is presented and completed so as also to cover the case of multiple zeros of the characteristic equation.

Let A be an n–by–n matrix with coefficients in a field K of zero characteristic. The characteristic polynomial of A may be obtained by a method proposed by D. K. Faddeev [3, §47; **4**]. This is a modification of a method of U. J. J. Leverrier (1840) which, according to A. S. Householder [8, p. 172], has been "rediscovered and improved" in the late forties also by J. M. Souriau [9] and J. S. Frame [5]. Although this method is not optimal from the point of view of numerical computation, it is rather elegant and easy to realize in a computer program. It consists of the following steps (I denotes the identity matrix):

$$A_1 := A; \tag{1}$$

$$a_k := -\operatorname{trace}(A_k)/k; \qquad (1 \le k \le n) \tag{2}$$

$$A_{k+1} := A(A_k + a_k I); \qquad (1 \le k \le n-1). \tag{3}$$

Putting $a_0 = 1$ one then has

$$\Phi(\lambda) := \det(\lambda I - A) = \sum_{k=0}^{n} a_k \lambda^{n-k}. \tag{4}$$

An additional definition at the same time (and under a condition to be mentioned below) also furnishes eigenvectors to given eigenvalues: Let

$$\overline{A}_0 := I; \tag{5}$$

$$\overline{A}_k := A_k + a_k I; \qquad (1 \le k \le n-1) \tag{6}$$

and define

$$C(\lambda) := \sum_{k=0}^{n-1} \lambda^{n-1-k} \overline{A}_k. \tag{7}$$

Then

$$A \cdot C(\lambda) = \lambda C(\lambda) - \Phi(\lambda) I,$$

and therefore, if λ_0 is an eigenvalue of A, the non-zero columns of $C(\lambda_0)$ are corresponding eigenvectors. Faddeev [3] mentions that $C(\lambda_0) \neq 0$ if all eigenvalues of A are distinct.

Faddeev's proof as well as that of Leverrier, Householder and Gantmacher [7] use the connexion between the coefficients of a polynomial and the power sums of its roots. The proofs of Souriau and Frame (see also [2] and [6]) exploit the concept of the adjoint of the matrix $\lambda I - A$, and we shall repeat it very briefly for the reason of completeness. (A third proof using a companion matrix has been given in [1].) The purpose of this paper is to state then necessary and sufficient conditions for $C(\lambda_0) \neq 0$ and to show how the derivatives of C at λ_0 still may serve to find eigenvectors and generalized eigenvectors of A if $C(\lambda_0)$ vanishes.

Theorem 1. *Let*

$$\Phi(\lambda) := det(\lambda I - A) = \sum_{k=0}^{n} a_k \lambda^{n-k} \tag{4}$$

be the characteristic polynomial of A *and define the matrices* $\overline{A}_k$ $(0 \leq k \leq n-1)$, A_k $(1 \leq k \leq n)$ *and* $C(\lambda)$ *by*

$$\overline{A}_0 := I; \tag{5}$$

$$\overline{A}_k := A_k + a_k I; \qquad (1 \leq k \leq n-1) \tag{6}$$

$$A_1 := A; \tag{1}$$

$$A_{k+1} := A(A_k + a_k I); \qquad (1 \leq k \leq n-1) \tag{3}$$

and

$$C(\lambda) := \sum_{k=0}^{n-1} \lambda^{n-1-k} \overline{A}_k. \tag{7}$$

Then

a) $C(\lambda) = \operatorname{adj}(\lambda I - A)$ *for* $\lambda \in K$;

b) $a_k = \operatorname{trace}(\overline{A}_k)/(n-k)$ *for* $0 \leq k \leq n-1$;

c) $a_k = -\operatorname{trace}(A_k)/k$ *for* $1 \leq k \leq n$.

Proof

a) Let the matrices B_k $(0 \le k \le n-1)$ be defined by

$$\operatorname{adj}(\lambda I - A) = \sum_{k=0}^{n-1} \lambda^{n-1-k} B_k.$$

Then we have

$$\begin{aligned} \Phi(\lambda) I &= (\lambda I - A) \cdot \operatorname{adj}(\lambda I - A), \\ \sum_{k=0}^{n} a_k \lambda^{n-k} I &= (\lambda I - A) \sum_{k=0}^{n-1} \lambda^{n-1-k} B_k \\ &= B_0 \lambda^n + \sum_{k=1}^{n-1} \lambda^{n-k}(B_k - AB_{k-1}) - AB_{n-1}, \end{aligned}$$

and therefore

$$\begin{aligned} a_0 I &= B_0, \\ a_k I &= B_k - AB_{k-1} \qquad (1 \le k \le n-1), \\ a_n I &= -AB_{n-1}. \end{aligned} \tag{8}$$

We conclude

$$B_0 = I = \overline{A}_0,$$

and, by induction on k (with inductive hypothesis $B_{k-1} = \overline{A}_{k-1}$),

$$B_k = AB_{k-1} + a_k I = A_k + a_k I = A\overline{A}_{k-1} + a_k I = \overline{A}_k \qquad (1 \le k \le n-1).$$

This shows

$$\operatorname{adj}(\lambda I - A) = \sum_{k=0}^{n-1} \lambda^{n-1-k} \overline{A}_k = C(\lambda).$$

b) Let A_{ii} be the $(n-1)$–by–$(n-1)$ matrix obtained by deleting the i 'th row and the i 'th column in the matrix A. The coefficient of λ in $\Phi(\lambda) = \det(\lambda I - A)$ is

$$a_{n-1} = \sum_{i=1}^{n} \det\bigl(-A_{ii}\bigr) = \operatorname{trace}\bigl(\operatorname{adj}(-A)\bigr).$$

By the same token, we see that in the polynomial

$$\Phi(\lambda + \mu) = \det\bigl((\lambda + \mu)I - A\bigr) = \sum_{k=0}^{n} a_k (\lambda + \mu)^{n-k}$$

the coefficient of μ is

$$\begin{aligned}\sum_{k=0}^{n-1} a_k(n-k)\lambda^{n-1-k} &= \operatorname{trace}\bigl(\operatorname{adj}(\lambda I - A)\bigr)\\ &= \operatorname{trace}\bigl(C(\lambda)\bigr) \qquad\qquad \text{by a)}\\ &= \sum_{k=0}^{n-1} \lambda^{n-1-k} \operatorname{trace}\bigl(\overline{A}_k\bigr).\end{aligned}$$

c) Since $A_k = \overline{A}_k - a_k I$, we get

$$\operatorname{trace}(A_k) = \operatorname{trace}(\overline{A}_k) - na_k = (n-k)a_k - na_k = -ka_k \quad (1 \le k \le n-1).$$

Finally, from (8), we get

$$na_n = -\operatorname{trace}(A_n). \quad \square$$

Remarks. 1) The equation deduced in b) of the proof can be put in the following form

$$\frac{d}{d\lambda}\det(\lambda I - A) = \operatorname{trace}\bigl(\operatorname{adj}(\lambda I - A)\bigr).$$

2) The characterization of $C(\lambda)$ according to Theorem 1 a) allows us to describe the condition $C(\lambda_0) \neq 0$ more completey. Namely, by standard matrix theory, we obtain the following equivalences:

$C(\lambda_0) \neq 0 \Longleftrightarrow \operatorname{rank}\bigl(C(\lambda_0)\bigr) = 1 \Longleftrightarrow \operatorname{rank}(\lambda_0 I - A) = n - 1$
$\Longleftrightarrow \dim\ker(\lambda_0 I - A) = 1 \Longleftrightarrow$ Jordan's canonical form of A (over the algebraic closure $\overline{K}$ of K) contains one sole Jordan cell with the eigenvalue λ_0.

All these assertions just express the fact that λ_0 is an eigenvalue of A of geometric multiplicity 1.

Some additional information concerning the relation between the matrix $C(\lambda)$ and the generalized eigenspaces of A may be found in the literature:

Gantmacher [7, p. 92] shows that, if λ_0 is an eigenvalue of A and if the elements of $C(\lambda)$ have a largest common factor $(\lambda - \lambda_0)^m$ $(m \ge 0)$, then the non-zero columns of the $m-$th derivative $C^{(m)}(\lambda_0)$ are eigenvectors. Householder [8, p. 168] proves that, if λ_0 is an eigenvalue, "$C(\lambda_0)$ cannot vanish unless λ_0 is at least a double root". He states without proof that, if λ_0 has multiplicity 2, then, if $C(\lambda_0) \neq 0$, the non-zero columns of $C(\lambda_0)$ are eigenvectors and the non-zero columns of $C'(\lambda_0)$ are generalized eigenvectors. He continues: "If, however, $C(\lambda_0) = 0$, then $C'(\lambda_0)$ can be shown to have rank 2, and any nonvanishing column is a proper vector. For roots of higher multiplicity the situation is analogous." In an exercise the reader is invited to prove these assertions.

In the rest of this paper, we try to give a more complete account of the relation between the spans of $C(\lambda_0)$ and of its derivatives $C^{(m)}(\lambda_0)$ at an eigenvalue λ_0, on the one hand, and the generalized eigenspace of A for this eigenvalue, on the other hand. We assume in the sequel, without loss of generality, that K is algebraically closed and that A is given in Jordan's canonical form. By $\operatorname{im}(B)$ we denote the vector space spanned by the columns of a matrix B or, what is the same, the image space of the linear mapping defined by B. If λ_0 has multiplicity l and index s (i. e. the largest Jordan cell corresponding to λ_0 has format $s \times s$), then (see Corollaries b), h) below) $\operatorname{im} C^{(l-1)}(\lambda_0) = \ker(\lambda_0 I - A)^s$, the space of all generalized eigenvectors to λ_0. For $1 \le t \le s-1$, however, we have

$$\ker(\lambda_0 I - A)^t \supseteq \operatorname{im} C^{(l-1-s+t)}(\lambda_0)$$

with equality in place of inclusion only in special cases, to be precisely specified in Theorem 3 below. Still, the dimensions of the spaces $\ker(\lambda_0 I - A)^t$ and the Jordan structure of A, related to λ_0, can be found from the ranks of the matrices $C^{(l-1-s+t)}(\lambda_0)$.

It should be appreciated that the derivatives $C^{(k)}(\lambda_0)$ can easily be found by performing Faddeev's algorithm with the matrices $\tilde{A} = A - \lambda_0 I$ in place of A, which yields by (7) and Theorem 1 a)

$$C^{(k)}(\lambda_0) = k!\overline{\tilde{A}}_{n-1-k}, \qquad 0 \le k \le n-1.$$

Theorem 2. *Let $\lambda_0 \in K$ be an eigenvalue of A of algebraic multiplicity l. Denote by $G = \ker(\lambda_0 I - A)^l$ the subspace of K^n consisting of the generalized eigenvectors with eigenvalue λ_0 (i. e. the vectors annihilated by some power of the matrix $\lambda_0 I - A$) and let $H = \operatorname{im}(\lambda_0 I - A)^l$. Let $M^{(k)} = \operatorname{im} C^{(k)}(\lambda_0)$ for $k \ge 0$. Then*

$$\begin{aligned} &M^{(k)} = (\lambda_0 I - A)^{l-1-k} G \qquad \text{for } 0 \le k \le l-1, \\ &H = \ker C^{(l-1)}(\lambda_0) \subseteq M^{(l)}, \\ &K^n = G \oplus H. \end{aligned}$$

Proof

Let $V = K^n = \bigoplus_{h=1}^{q} V_h$ where $V_h \cong K^{s_h}$ is an A–invariant subspace of V in which A is represented by an $s_h \times s_h$–Jordan cell

$$A_h = \left.\begin{pmatrix} \lambda_h & 0 & \dots & & 0 \\ 1 & \lambda_h & & & \\ 0 & 1 & \lambda_h & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 & \lambda_h \end{pmatrix}\right\} s_h \ .$$

Suppose $\lambda_h = \lambda_0$ for $1 \le h \le p$ and $\lambda_h \ne \lambda_0$ for $p < h \le q$, so $l = \sum_{h=1}^{p} s_h$ and $G = \bigoplus_{h=1}^{p} V_h$.

In V_h the matrix $C(\lambda) = \operatorname{adj}(\lambda I - A)$ is represented by the $s_h \times s_h$ matrix (also referred to as cell $C_h(\lambda)$) of the form

$$\begin{pmatrix} (\lambda-\lambda_h)^{s_h-1} & 0 & \dots & 0 \\ (\lambda-\lambda_h)^{s_h-2} & (\lambda-\lambda_h)^{s_h-1} & & \vdots \\ \vdots & & & 0 \\ 1 & \dots & (\lambda-\lambda_h)^{s_h-2} & (\lambda-\lambda_h)^{s_h-1} \end{pmatrix} \prod_{k\ne h} (\lambda-\lambda_k)^{s_k}.$$

If $c_{ij}(\lambda)$ is a non-zero coefficient of $C_h(\lambda)$, then the multiplicity of λ_0 as a zero of $c_{ij}(\lambda)$ is $\sum_{k=1}^{p} s_k = l$ if $h > p$, and it is

$$\sum_{\substack{k=1\\k\ne h}}^{p} s_k + s_h - 1 - (i-j) = l - 1 - (i-j) \quad \text{if } h \le p.$$

Therefore, if $k < l$, then $C^{(k)}(\lambda_0)$ has non-zero elements in the diagonal given by $k = l - 1 - (i-j)$, i. e. $i - j = l - k - 1$, where this diagonal meets a cell C_h $(1 \le h \le p)$, and all elements outside the cells C_h $(1 \le h \le p)$ and all elements in these cells but above the diagonal $i - j = l - k - 1$ are zero. (We abuse the term "cell C_h" also for the location of $C_h(\lambda)$ as a $s_h \times s_h$ submatrix of the $n \times n$ matrix $C(\lambda)$.) Consequently, $M^{(k)}$ is spanned by those basis vectors which are numbered by the row-indices corresponding to the intersection of this diagonal with the cells C_h $(1 \le h \le p)$. On the other hand, within the cells C_h $(1 \le h \le p)$, the only non-zero elements in $(\lambda_0 I - A)^{l-k-1}$ are 1's in the same diagonal, and therefore

$$(\lambda_0 I - A)^{l-k-1} G = \bigoplus_{h=1}^{p} (\lambda_0 I - A_h)^{l-k-1} V_h$$

spans exactly the same subspace of K^n. Furthermore, $C_h^{(l-1)}(\lambda_0)$, $1 \le h \le p$, is non-singular on G and vanishes on $\bigoplus_{h=p+1}^{q} V_h = H$. Hence $H = \ker C^{(l-1)}(\lambda_0)$. Similarly, for $p < h \le q$, the matrix $C_h^{(l)}(\lambda_0)$ is non-singular. This implies that $M^{(l)} \supseteq H$. The last assertion follows from the decomposition $K^n = \bigoplus_{h=1}^{p} V_h \oplus \bigoplus_{h=p+1}^{q} V_h$. $\square$

The assertions contained in the following corollary and in the remarks thereupon are all quite immediate consequences of Theorem 2 and especially of the equation

$$M^{(k)} = (\lambda_0 I - A)^{l-1-k} \ker(\lambda_0 I - A)^l, \qquad 0 \le k \le l-1.$$

We state these facts in great detail for the only reason to give a description as complete as possible of the situation under scrutiny, and without any pretensions about originality or depth.

Recall that a function $k \longmapsto f(k)$ on $\mathbb{Z}$ is called weakly convex if $k \longmapsto (f(k) - f(k-1))$ is non-decreasing.

Corollary. *Under the assumptions of Theorem 2 and with the additional notations* $E = \ker(\lambda_0 I - A)$, $m_{-2} = m_{-1} = 0$, $m_k = \dim M^{(k)}$ *for* $0 \le k \le l-1$, *the following assertions hold:*

a) for $1 \le k \le l-1$ *one has* $M^{(k-1)} = (\lambda_0 I - A)M^{(k)}$;

b) $M^{(0)} \subseteq M^{(1)} \subseteq \cdots \subseteq M^{(l-1)} = G$;

c) $0 \le m_0 \le m_1 \le \cdots \le m_{l-1} = l$;

d) for $0 \le k \le l-1$ *one has*

$$m_k - m_{k-1} = \dim(E \cap M^{(k)})$$

and this equals the number of Jordan cells in A *belonging to* λ_0 *and of size at least* $l-k$; *the function* $k \longmapsto m_k$, $-1 \le k \le l-1$, *is weakly convex;*

e) let p_k *denote the number of Jordan cells in* A *belonging to* λ_0 *and of size* k $(1 \le k \le l)$; *then* $p_k = m_{l-k} - 2m_{l-k-1} + m_{l-k-2}$;

f) for $1 \le t \le l$ *one has*

$$m_{l-1} - m_{l-1-t} = \dim\ker(\lambda_0 I - A)^t;$$

in particular, $m_{l-1} - m_{l-2} = \dim E$ *and this equals the number of Jordan cells in* A *belonging to* λ_0;

g) if $m_0 > 0$ *then* $M^{(0)} = E$ *and* $m_k = k+1$ *for* $0 \le k \le l-1$; *in this case there is just one Jordan cell belonging to* λ_0;

h) let k_1 $(0 \le k_1 \le l-1)$ *be the smallest index* k *for which* $m_k > 0$, *i.e.*

$$0 = m_{-1} = \cdots = m_{k_1-1} < m_{k_1} < \cdots < m_{l-1} = l;$$

if s *denotes the size of the largest Jordan cell in* A *belonging to* λ_0, *then* $s = l - k_1$ *and* $G = \ker(\lambda_0 I - A)^s$, $H = \operatorname{im}(\lambda_0 I - A)^s$;

i) for $0 \le t \le s$ *one has*

$$\begin{aligned} M^{(k_1-1+t)} &= (\lambda_0 I - A)^{s-t}\ker(\lambda_0 I - A)^s \\ &= \operatorname{im}(\lambda_0 I - A)^{s-t} \cap \ker(\lambda_0 I - A)^t; \end{aligned}$$

in particular, $(\lambda_0 I - A)^t M^{(k_1-1+t)} = \{0\}$, $M^{(k_1)} \subseteq E$;

j) for $1 \le t \le s$ *one has*

$$(\lambda_0 I - A)^{t-1} M^{(k_1-1+t)} = M^{(k_1)} \ne \{0\};$$

k) if $l = n$ *then* $M^{(l-1)} = K^n$ *and* $M^{(l)} = \{0\}$;

l) if $l < n$ *then*

$$\begin{aligned} &(\lambda_0 I - A)M^{(l)} \not\subseteq M^{(l-1)}, \\ &(\lambda_0 I - A)^t M^{(l)} = H \neq \{0\} \qquad \textit{for } t \geq s, \\ &K^n = M^{(l-1)} \bigoplus (\lambda_0 I - A)^s M^{(l)}. \end{aligned}$$

Proof

Assertion a) follows directly from Theorem 2. Since $(\lambda_0 I - A)G \subseteq G$, Theorem 2 implies $M^{(k-1)} \subseteq M^{(k)}$ for $1 \leq k \leq l-1$. This proves b); c) follows since $\dim G = l$. The first statement of d) follows from a) since $E \cap M^{(k)}$ is the kernel of the mapping $(\lambda_0 I - A)$ restricted to $M^{(k)}$. Since $M^{(k)} = (\lambda_0 I - A)^{l-1-k}G$, $\dim(E \cap M^{(k)})$ is the maximal number of linearly independent eigenvectors $\mathbf{x}$ with eigenvalue λ_0 which may be written as $\mathbf{x} = (\lambda_0 I - A)^{l-1-k}\mathbf{y}$, $\mathbf{y} \in G$, i. e. which belong to a subspace V_h $(h \leq p)$ corresponding to a Jordan cell of size at least $l-k$. Furthermore, b) implies that the sequence $\{m_k - m_{k-1}\}_{k=0}^{l-1}$ is non-decreasing. Assertion e) follows from d) since $(m_{l-k} - m_{l-k-1}) - (m_{l-k-1} - m_{l-k-2})$ equals the number of Jordan cells of size $l-k$. Assertion f) follows from $M^{(l-1-t)} = (\lambda_0 I - A)^t M^{(l-1)}$. To prove g) we first note that, by the Remark 2 above, $m_0 > 0$ implies $M^{(0)} = E$ and $m_0 = \dim E = 1$; then b) and d) imply $m_k = k+1$ $(0 \leq k \leq l-1)$.

If s denotes the largest size of a Jordan cell in A belonging to the eigenvalue λ_0, then $G = \ker(\lambda_0 I - A)^s$ and $\ker(\lambda_0 I - A)^{s-1}$ is a proper subspace of G. By Theorem 2 this implies $M^{(l-s-1)} = \{0\} \neq M^{(l-s)}$, so $k_1 = l-s$ and h) is proven. The first line of i) is a reformulation of Theorem 2, taking account of h). The second line follows from the first line since, generally, $P\ker(QP) = \mathrm{im}(P) \cap \ker(Q)$. j) follows from i) taking account of the definition of k_1.

If $l = n$, then $M^{(l)} = M^{(n)} = \{0\}$, since $C(\lambda)$ is a polynomial of degree $n-1$ (comp. (7)), and $M^{(l-1)} = G = K^n$. This proves k). If $l < n$, then the subspace H is non-trivial by Theorem 2. Since $\ker(\lambda_0 I - A) \subseteq G$ and $G \cap H = \{0\}$, the mapping $(\lambda_0 I - A)$ acts as an automorphism on H. Again by Theorem 2

$$(\lambda_0 I - A)M^l \supseteq (\lambda_0 I - A)H = H,$$

so certainly

$$(\lambda_0 I - A)M^l \not\subseteq M^{(l-1)} = G.$$

Finally, for $t \geq s$

$$(\lambda I - A)^t M^l \subseteq \mathrm{im}(\lambda_0 I - A)^s = H = (\lambda_0 I - A)^t H \subseteq (\lambda_0 I - A)^t M^{(l)}.$$

This shows $(\lambda_0 I - A)^t M^{(l)} = H$ and completes the proof. $\square$

Remarks. 1) If $k_1 = l-1 \quad (s=1)$, then one has

$$M^{(0)} = \cdots = M^{(l-2)} = \{0\}, \qquad M^{(l-1)} = \ker(\lambda_0 I - A)$$

with $\dim M^{(l-1)} = l$. This is the case considered in the quotation of Householder. In this case, all Jordan cells corresponding to λ_0 have size 1.

2) In Theorem 2 and its Corollary, the value of the multiplicity l of the eigenvalue λ_0 is presupposed to be known from the beginning e. g. by decomposing the characteristic polynomial given in Theorem 1. Alternatively, l can be characterized in terms of the numbers m_k or the spaces $M^{(k)}$. If $m_0 = 0$, it follows from Corollary d) that

$$l = \min\{k : 1 \le k \text{ and } m_{k-1} = k\}.$$

Corollary g) shows that this is not true if $m_0 > 0$. Other characterizations are based on the distinction between the spaces $M^{(l)}$ and $M^{(k)}$ for $k < l$, as expressed by Corollaries a), k) and l). It is easily verified that

$$l = \max\{m : 1 \le m \le n \text{ and } M^{(k-1)} \subseteq \ker(\lambda_0 I - A)^k \text{ for } 1 \le k \le m\}.$$

With knowledge of k_1, which eventually can be obtained from the definition, viz.

$$k_1 = \min\{k : 0 \le k \text{ and } M^{(k)} \ne \{0\}\},$$

the last characterization by Corollaries i) and l) can be sharpened to

$$l = \max\{m : 1 \le m \le n \text{ and } M^{(k-1)} \subseteq \ker(\lambda_0 I - A)^{k-k_1} \text{ for } k_1 \le k \le m\}.$$

Finally, observe that also (setting $M^{(-1)} = \{0\}$)

$$l = \max\{m : 1 \le m \le n \text{ and } (\lambda_0 I - A)M^{(k-1)} \subseteq M^{(k-2)} \text{ for } 1 \le k \le m\}.$$

Theorem 3. *Let A, λ_0, l, E, G, $M^{(k)}$, m_k, k_1, s be defined as in Theorem 2 and its Corollary. Then the following assertions are equivalent:*

a) $M^{(k_1)} = E$;

b) $m_{k_1} = \dim E$;

c) $m_k - m_{k-1}$ *is independent of* k *for* $k_1 \le k \le l-1$;

d) $m_k - m_{k-1} = \dim E$ *for* $k_1 \le k \le l-1$;

e) $M^{(k)} = \ker(\lambda_0 I - A)^{k-k_1+1}$ *for some* k, $k_1 \le k < l-1$;

f) $M^{(k)} = \ker(\lambda_0 I - A)^{k-k_1+1}$ *for* $k_1 \le k \le l-1$;

g) all the Jordan cells of A belonging to the eigenvalue λ_0 have the same size.

Proof

Since $M^{(k_1)} \subseteq E$ by i) of the Corollary, the equivalence of the assertions a) and b) is obvious. Furthermore, $M^{(k_1)} = (\lambda_0 I - A)^{s-1} G$ equals E if and only if all the Jordan cells of A belonging to the eigenvalue λ_0 have the same size s. Therefore, the assertions a) and g) are equivalent. Obviously, d) implies b) as well as c) and, conversely, c) implies d) by Corollary f). Also, assertion a) which is equivalent to b) implies d) by Corollaries b) and d). Since $f) \Rightarrow e)$ is trivial, the proof will be complete upon showing that $c) \Rightarrow f)$ and $e) \Rightarrow c)$. Let $k = k_1 - 1 + t \quad (1 \leq t \leq s)$. Then we have to deal with the relations

$$M^{(k_1-1+t)} = \ker(\lambda_0 I - A)^t. \tag{9}$$

Since, by Corollary i), $M^{(k_1-1+t)} \subseteq \ker(\lambda_0 I - A)^t$, (9) is true iff $m_{k_1-1+t} = \dim\ker(\lambda_0 I - A)^t$. The latter dimension equals $m_{l-1} - m_{l-1-t}$ by Corollary f) and hence, if we define

$$\phi(t) := m_{k_1-1+t}, \qquad 0 \leq t \leq s,$$

then $\phi(0) = 0$ and (9) proves to be equivalent with

$$\phi(t) + \phi(s-t) = \phi(0) + \phi(s). \tag{10}$$

Now c) means that ϕ is linear, which fact clearly implies (10) and thus (9) for $0 \leq t \leq s$. Hence $c) \Rightarrow f)$ is proven. In order to show $e) \Rightarrow c)$ we infer from Corollary d) that ϕ is weakly convex in the interval $0 \leq t \leq s$. Hence there are but two possibilities: either ϕ is linear or ϕ is strictly sub-linear. In the latter case (10) does not hold for any t with $0 < t < s$, i. e., e) is false. Consequently, e) implies the linearity of ϕ and hence c). $\square$

Appendix

Finally, we want to present a proof of Theorem 2 which does not use the Jordan decomposition.

If $\Phi(\lambda) = \det(\lambda I - A)$ as above, then we obtain from the identity $\Phi(\lambda) I = (\lambda I - A) C(\lambda)$ by differentiation that

$$\Phi^{(k)}(\lambda) I = (\lambda I - A) C^{(k)}(\lambda) + k C^{(k-1)}(\lambda) \quad \text{for } k \geq 1.$$

(This equation is a paraphrase of the recursion formulae (3) and (6) with the matrix $A - \lambda I$ in place of A.) If $\lambda_0 \in K^n$ is an eigenvalue of multiplicity l, then $\Phi^{(k)}(\lambda_0) = 0$ for $0 \leq k \leq l-1$ and hence

$$C^{(k-1)}(\lambda_0) = -\frac{1}{k}(\lambda_0 I - A) C^{(k)}(\lambda_0), \quad M^{(k-1)} = (\lambda_0 I - A) M^{(k)}, \quad 1 \leq k \leq l-1.$$

Next we show $M^{(l-1)} = G$. Note that

$$(\lambda_0 I - A)^l C^{(l-1)}(\lambda_0) = (-1)^{l-1}(l-1)!(\lambda_0 I - A)C(\lambda_0) = (-1)^{l-1}(l-1)!\Phi(\lambda_0)I = 0,$$

and hence $M^{(l-1)} = \operatorname{im} C^{(l-1)}(\lambda_0) \subseteq G$. If we denote $\ker C^{(l-1)}(\lambda_0)$ by N, then

$$\dim M^{(l-1)} + \dim N = n = \dim G + \dim H,$$

and hence we can conclude that $M^{(l-1)} = G$ and $N = H$ once we have proven that $N \subseteq H$. To this end, restrict the linear mapping

$$\Phi^{(l)}(\lambda_0)I = (\lambda_0 I - A)C^{(l)}(\lambda_0) + lC^{(l-1)}(\lambda_0)$$

to N, in which case the last term vanishes. Since $\Phi^{(l)}(\lambda_0) \neq 0$ it follows that

$$N = (\lambda_0 I - A)C^{(l)}(\lambda_0)N \subseteq (\lambda_0 I - A)N \subseteq N.$$

(Notice that the derivatives of $C(\lambda)$ commute with one another. Hence $C^{(l)}(\lambda_0)N \subseteq N$. Moreover, these derivatives also commute with A and this implies $(\lambda_0 I - A)N \subseteq N$.) Consequently, $N = (\lambda_0 I - A)^l N \subseteq H$. This finishes the proof of $M^{(l-1)} = G$ and $N = H$.

Incidentally, it has been shown that $(\lambda_0 I - A)^l H = H$, which implies $G \cap H = \ker(\lambda_0 I - A)^l \cap H = \{0\}$. Taking account of dimensions, $G \bigoplus H = K^n$ follows.

Finally, since $C^{(l)}(\lambda_0)(\lambda_0 I - A)N = (\lambda_0 I - A)C^{(l)}(\lambda_0)N = N$, it also follows that $M^{(l)} = \operatorname{im} C^{(l)}(\lambda_0) \supseteq N = H$. $\square$

The first author gratefully acknowledges the facilities of a research fellowship offered by the Department of Mathematics and Computing Sciences at the University of Technology in Eindhoven.

References

1. S. Barnett, *Leverrier's algorithm: a new proof and extensions*, SIAM J. Matrix Anal. Appl. **10** (1989), 551–556.
2. P. S. Dwyer, *Linear Computation*, Wiley, New York, 1951.
3. D. K. Faddeev and W. N. Faddeeva, *Computational Methods of Linear Algebra*, Freeman, San Francisco, 1963.
4. D. K. Faddeev and I. S. Sominskii, *Collection of Problems in Higher Algebra (in Russian)*, Gosudarstv. Izdat. Tehn-Teor. Lit., Moscow, 1952.
5. J. S. Frame, *A simple recursion formula for inverting a matrix (abstract)*, Bull. AMS **55** (1949), 1045.
6. J. S. Frame, *Matrix functions and applications, Part IV - Matrix functions and constituent matrices*, IEEE Spectrum **1** (1964), 123–131.
7. F. R. Gantmacher, *Matrix Theory, Vol.1*, Chelsea, New York, 1959.
8. A. S. Householder, *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York, 1964.
9. J. M. Souriau, *Une méthode pour la décomposition spectrale et l'inversion des matrices*, C. R. Acad. Sci. **227** (1948), 1010–1011.