FACTORIZATION OF MATRIX POLYNOMIALS WITH SYMMETRIES

A.C.M RAN* AND L. RODMAN†

Abstract. An $n \times n$ matrix polynomial $L(\lambda)$ (with real or complex coefficients) is called selfadjoint if $L(\lambda) = (L(\overline{\lambda}))^*$, and symmetric if $L(\lambda) = (L(\pm\lambda))^T$. Factorizations of selfadjoint and symmetric matrix polynomials of the form $L(\lambda) = (M(\overline{\lambda}))^* DM(\lambda)$ or $L(\lambda) = (M(\pm\lambda))^T DM(\lambda)$ are studied, where D is a constant matrix and $M(\lambda)$ is a matrix polynomial. In particular, the minimal possible size of D is described in terms of the elementary divisors of $L(\lambda)$ and (sometimes) signature of the Hermitian values of $L(\lambda)$.

1. Introduction.

Let $L(\lambda) = \sum_{j=0}^{\ell} \lambda^j A_j$ be a matrix polynomial, where $A_j \quad (j = 0, \ldots, \ell)$ are complex $n \times n$ matrices and λ is a complex parameter. The polynomial $L(\lambda)$ is called *selfadjoint* if $L(\lambda) = (L(\overline{\lambda}))^*$ for all $\lambda \in \mathbb{C}$.

Factorizations of the form

$$L(\lambda) = (M(\overline{\lambda}))^* D\ M(\lambda), \tag{1.1}$$

where $D = D^*$ is a constant matrix (not necessarily of the same size as $L(\lambda)$ and $M(\lambda)$ is a matrix polynomial, have been studied in the literature, under various additional hypotheses (see [Ja, Co, GLR1, GLR2]). The study of factorizations (1.1) is motivated by several applied problems, such as filtering (Chapter 9 of [AM]). Factorizations of matrix polynomial $L(\lambda)$ having other types of symmetries, such as $L(\lambda) = (L(-\lambda))^T$ or $L(\lambda) = (L(\lambda))^T$ have been studied in the literature as well (see, e.g., [Lyu1, Lyu2]). For such polynomials, it is natural to seek factorizations of type

$$L(\lambda) = (M(\varepsilon\lambda))^T D\ M(\lambda), \tag{1.2}$$

where $D = D^T$ is a constant matrix (not necessarily of the same size as $L(\lambda)), M(\lambda)$ is a matrix polynomial, and $\varepsilon = 1$ or $\varepsilon = -1$, as appropriate.

In this paper we identify the minimal possible size of the matrix D in factorization of types (1.1) and (1.2), where $L(\lambda)$ has the appropriate symmetry. The cases when $L(\lambda)$ has complex coefficients or real coefficients are studied (if $L(\lambda)$ is assumed to be real, then in (1.1) and (1.2) $M(\lambda)$ and D are assumed to be real as well). Our result concerning

*Faculteit Wiskunde en Informatica, Vrije Universiteit, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands

†Department of Mathematics, College of William and Mary Williamsburg, VA 23187–8795, USA. Partially supported by NSF Grant DMS–9000839 and by the NSF International Cooperation Grant with the Netherlands.

factorization (1.1) is a generalization of the main result of [GLR2] where only the case of constant signature was considered under the additional hypothesis that $\det L(\lambda) \not\equiv 0$

We present also (in Section 2) general factorization results in an abstract framework, for matrix polynomials over a field having suitable symmetries. These results, although independently interesting, play an auxiliary role in this paper, serving as essential ingredients in the proofs of the main results given in Sections 3–6.

The following notation will be used throughout the paper. Standard notation $\mathbb{R}(\mathbb{C})$ to denote the real (complex) field, and I_k for the $k \times k$ unit matrix. A^T (resp. (A^*) stands for the transpose (resp. conjugate transpose) of a matrix A, and $(A^T)^{-1}$ (resp. $(A^*)^{-1}$) is abbreviated to A^{-T} (resp. A^{-*}). Block diagonal matrix with the blocks $Z_1, \dots, Z_m$ on the main diagonal will be denoted $Z_1 \oplus \cdots \oplus Z_m$ or diag $(Z_1, \dots, Z_m)$. For a Hermitian $n \times n$ matrix X, let $\nu_+(X)$ (resp $\nu_-(X)$, or $\nu_0(X)$) be the number of positive (resp. negative, or zero) eigenvalues of X counted with multiplicities. Thus,

$$\nu_+(X) + \nu_-(X) + \nu_0(X) = n.$$

Given a matrix polynomial $L(\lambda)$ over $\mathbb{C}$, its *general rank* $r(L)$ is defined by

$$r(L) = \max_{\lambda_0 \in \mathbb{C}} \{\operatorname{rank} L(\lambda_0)\}.$$

It coincides (when $F = \mathbb{C}$) with the notion of general rank introduced and used in Section 2 for matrix polynomial over a field F. The points $\lambda_0 \in \mathbb{C}$ for which $\operatorname{rank} L(\lambda_0) = r(L)$ will be called *regular* points of $L(\lambda)$; all other points $\lambda_0 \in \mathbb{C}$ will be called *singular*. Clearly, the set of singular points is finite (or possibly empty). An $n \times n$ matrix polynomial $L(\lambda)$ is called *regular* if $r(L) = n$, or, equivalently, if $\det L(\lambda) \not\equiv 0$.

Acknowledgement. The problem concerning the minimal possible size of D in factorizations (1.1) for complex selfadjoint matrix polynomials has been posed by
Prof. I. Gohberg.

2. Symmetrix matrix polynomials over general field. Let F be a (commutative) field, and let $F[\lambda]$ be the ring of polynomials over F in one variable λ. Matrices $L(\lambda)$ with entries in $F[\lambda]$ will be called *matrix polynomials* (over F). It is well-known (see, e.g., [M]) that every $m \times n$ matrix polynomial $L(\lambda)$ admits a representation (called the *Smith form*)

$$L(\lambda) = E(\lambda) \operatorname{diag}(d_1(\lambda), d_2(\lambda), \dots d_r(0), 0, \dots, 0) F(\lambda) \tag{2.1}$$

where $E(\lambda)$ and $F(\lambda)$ are matrix polynomials with sizes $m \times m$ and $n \times n$, respectively, and having constant non-zero determinants, and $d_1(\lambda), \dots, d_r(\lambda)$ are monic scalar polynomials (over F) such that $d_i(\lambda)$ divides $d_{i+1}(\lambda)$ $(i = 1, \dots, r-1)$. The polynomials $d_i(\lambda)$ are called *invariant polynomials* of $L(\lambda)$; these polynomials, as well as their number r, are uniquely determined by $L(\lambda)$: $r \times r$ is the maximal size of a square submatrix in $L(\lambda)$ with determinant not identically zero, and for $i = 1, \dots, r$ the product $d_1(\lambda) \dots d_i(\lambda)$ is the greatest common divisor of the determinants of all $i \times i$ submatrices in $L(\lambda)$.

The number r will be called the *general rank* of $L(\lambda)$ and denoted $r(L)$.

In this section we will study factorizations of symmetric matrix polynomials, and the Smith form will be our main tool.

From now on we assume that the characteristic of F is different from 2. For a given automorphism σ of F such that $\sigma^2 =$ identity; and for fixed $\varepsilon = \pm 1$ consider the following transformation: for $a(\lambda) = \sum a_j \lambda^j \in F[\lambda]$ let

$$a_*(\lambda) = \sum \sigma(a_j)(\varepsilon\lambda)^j = \sum \sigma(a_j)\varepsilon^j \lambda^j \in F[\lambda]. \tag{2.2}$$

For an $m \times n$ matrix polynomial $X(\lambda) = [x_{ij}(\lambda)]_{i=1,j=1}^{m,n}$ over F define

$$X_*(\lambda) = [\widetilde{x}_{ij}(\lambda)]_{i=1,j=1}^{n,m}$$

where $\widetilde{x}_{ij}(\lambda) = [x_{ji}(\lambda)]_*$. We have:

1) $[X(\lambda)Y(\lambda)]_* = Y(\lambda)_* X(\lambda)_*$

2) $[X(\lambda)]_{**} = X(\lambda)$

3) $[x(\lambda)X(\lambda) + y(\lambda)Y(\lambda)]_* = x_* X_* + y_* Y_*$ for scalar polynomials $x(\lambda)$ and $y(\lambda)$.

4) if $\det X(\lambda) \equiv \text{const} . \neq 0$, then

$$[X_*(\lambda)]^{-1} = ([X(\lambda)]^{-1})_*.$$

These rules will be used often in the sequel.

An $m \times n$ matrix polynomial $L(\lambda)$ will be called *generally invertible* if all its invariant polynomials are constant 1. The terminology is justified by the fact that $L(\lambda)$ is generally invertible if and only if $L(\lambda)$ has a generalized inverse, i.e. matrix polynomial $N(\lambda)$ such that $N(\lambda)L(\lambda)N(\lambda) = N(\lambda)$ and $L(\lambda)N(\lambda)L(\lambda) = L(\lambda)$ (this fact is easily proved using the Smith form). A matrix polynomial $L(\lambda)$ will be called *right* (resp. *left*) *invertible* if there exists a matrix polynomial $N(\lambda)$ such that $L(\lambda)N(\lambda) \equiv I$ (resp. $N(\lambda)L(\lambda) \equiv I$).

We now state one of the main factorization results of this section.

THEOREM 2.1. *Let $L(\lambda)$ be an $n \times n$ generally invertible matrix polynomial such that*

$$L(\lambda) = L_*(\lambda), \tag{2.3}$$

and let r be the general rank of $L(\lambda)$. Then $L(\lambda)$ can be factorized in the form

$$L(\lambda) = M_*(\lambda)DM(\lambda), \tag{2.4}$$

where $M(\lambda)$ is an $r \times n$ right invertible matrix polynomial and D is an $r \times r$ constant matrix such that $D = D_*$.

Conversely if (2.4) holds for an $r \times n$ right invertible matrix polynomial $M(\lambda)$ and a constant matrix $D = D_*$, then $L(\lambda)$ satisfies (2.3), is generally invertible and has general rank r.

Proof. The converse statement is easy: Indeed, if

$$M(\lambda) = \widetilde{E}(\lambda)[I \quad 0]\widetilde{F}(\lambda)$$

is the Smith form for $M(\lambda)$, then

$$\begin{aligned} L(\lambda) &= \widetilde{F}_*(\lambda) \begin{bmatrix} I \\ 0 \end{bmatrix} \widetilde{E}_*(\lambda) D \widetilde{E}(\lambda)[I \quad 0]\widetilde{F}(\lambda) \\ &= \widetilde{F}_*(\lambda) \begin{bmatrix} \widetilde{E}_*(\lambda) & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \widetilde{E}(\lambda) & 0 \\ 0 & I \end{bmatrix} \widetilde{F}(\lambda), \end{aligned}$$

so by uniqueness of the Smith form $L(\lambda)$ is generally invertible and has general rank r. The verification of (2.3) is trivial.

We now prove the direct statement.

Observe that the proof is easily reduced to the case when $r = n$, i.e., $\det L(\lambda) \equiv \text{const} . \neq 0$. Indeed, let $L(\lambda) = E(\lambda)D(\lambda)F(\lambda)$ be the Smith form of $L(\lambda)$, and let $\widetilde{L}(\lambda) = (F_*(\lambda))^{-1}L(\lambda)F(\lambda)^{-1}$. Clearly, $\widetilde{L}(\lambda)$ is a matrix polynomial, $\widetilde{L}(\lambda) = \widetilde{L}_*(\lambda)$, and because of the equality $\widetilde{L}(\lambda) = (F_*(\lambda))^{-1}E(\lambda)D(\lambda)$ the last $n-r$ rows and columns of $\widetilde{L}(\lambda)$ are zeros. Obviously, it will suffice to prove the direct statement for the $r \times r$ matrix polynomial $N(\lambda)$ formed by the first r rows and columns of $\widetilde{L}(\lambda)$. As $\det N(\lambda) \equiv \text{const} . \neq 0$, the required reduction is accomplished.

We assume from now on that $\det L(\lambda) \equiv \text{const} . \neq 0$. In this case the direct statement follows from Theorem 3 in [Lyu1] (see also [Lyu2]). We outline an alternative procedure developed in [Co]. As in [Co] or [GLR2] (Section 4) we prove that there exists an $n \times n$ matrix polynomial $X(\lambda)$ with $\det X(\lambda) \equiv \text{const} . \neq 0$ such that for the matrix polynomial

$$A(\lambda) := X_*(\lambda)L(\lambda)X(\lambda) = [\alpha_{ij}(\lambda)]_{i,j=1}^n$$

either $\alpha_{11} \equiv 0$ or $A(\lambda)$ is diagonally dominant (i.e. for $j = 1, \ldots, n$, the degree of $\alpha_{jj}(\lambda)$ is bigger than the degrees of all non-zero entries in the j-th row and the j-th column in $A(\lambda)$). Because of this fact, without loss of generality we can assume that either $L(\lambda)$ is diagonally dominant or the (1,1) entry in $L(\lambda)$ is identically zero. If $L(\lambda)$ is diagonally dominant, then it must be constant, and we are done. So let

$$L = \begin{bmatrix} 0 & a_* \\ a & A_1 \end{bmatrix}, \; L^{-1} = \begin{bmatrix} \gamma & c_* \\ c & C_1 \end{bmatrix}$$

where $A_1 = A_{1*}$ and $C_1 = C_{1*}$ are $(n-1) \times (n-1)$ matrix polynomials. Now put

$$y = \frac{1}{2}\left(1 - c_* A_1 c\right) \; ; \; x = -A_1 c - ay.$$

$$Y = \begin{bmatrix} y & x_* \\ c & I \end{bmatrix}.$$

A calculation shows that

$$\begin{bmatrix} 1 & -x_* \\ 0 & I \end{bmatrix} Y \begin{bmatrix} 1 & 0 \\ -c & I \end{bmatrix} \equiv I \ ,$$

and so $\det Y = 1$. Another straightforward calculation shows that

$$Y_* L Y = \begin{bmatrix} 1 & 0 \\ 0 & L_0 \end{bmatrix} \ ,$$

where $L_0(\lambda)$ is an $(n-1) \times (n-1)$ matrix polynomial. Thus, we have reduced the size of L by one and can complete the proof by induction on n. □

As the proof of Theorem 2.1 shows, the constant matrix D can be taken diagonal.

If $L(\lambda)$ is not generally invertible, then easy examples show that the representation (2.4) (with D having the size equal to the general rank of $L(\lambda)$) is not always possible, even if we omit the requirement that $M(\lambda)$ is right invertible. We can, however, obtain a factorization result for not generally invertible $L(\lambda)$ if we allow D to be a polynomial (with special properties). To state and prove this result we need the concept of elementary divisors. Let $L(\lambda)$ be an $m \times n$ matrix polynomial with invariant polynomials $d_1(\lambda), \ldots, d_r(\lambda)$, where $d_s(\lambda), d_{s+1}(\lambda), \ldots, d_r(\lambda)$ are non-constant (if $L(\lambda)$ is generally invertible, then we say that $L(\lambda)$ has no elementary divisors). Factor:

$$d_i(\lambda) = (f_{i1}(\lambda))^{\alpha_{i1}} (f_{i2}(\lambda))^{\alpha_{i2}} \cdots (f_{ik_i}(\lambda))^{\alpha_{ik_i}}; \qquad i = s, s+1, \ldots, r \tag{2.5}$$

where $f_{i1}(\lambda), \ldots, f_{i,k_i}(\lambda)$ are irreducible and pairwise relatively prime non-constant monic scalar polynomials (over F). The collection of factors $(f_{ij}(\lambda))^{\alpha_{ij}}$ $(j = 1, \ldots, k_i; i = s, s+1, \ldots, r)$, where each factor is repeated as many times as it occurs in (2.5), is called *elementary divisors* of $L(\lambda)$, and the positive integer α_{ij} is called the *order* of the elementary divisor $(f_{ij}(\lambda))^{a_{ij}}$. Because of the divisibility relations among invariant polynomials, the collection of elementary divisors of $L(\lambda)$ determines the invariant polynomials uniquely, and therefore is invariant under the transformations

$$L(\lambda) \to E(\lambda) L(\lambda) F(\lambda), \quad \text{where} \quad \det E(\lambda) \equiv \text{const}\,. \neq 0, \ \det F(\lambda) \equiv \text{const}\,. \neq 0.$$

THEOREM 2.2. *Let $L(\lambda)$ be an $n \times n$ matrix polynomial such that*

$$L(\lambda) = L_*(\lambda),$$

and let r be the general rank of $L(\lambda)$. Further, let $\{f_1(\lambda)^{\alpha_1}, \dots, f_q(\lambda)^{\alpha_q}\}$ be the collection of elementary divisors of $L(\lambda)$. Then $L(\lambda)$ admits a factorization

$$L(\lambda) = M_*(\lambda)D(\lambda)M(\lambda),$$

where $M(\lambda)$ is an $r \times n$ matrix polynomial, and $D(\lambda) = D_*(\lambda)$ is an $r \times r$ matrix polynomial; moreover, the collection of elementary divisors of $D(\lambda)$ is $\{f_j(\lambda) : j \in J\}$, where the subset J of $\{1, 2, \dots, q\}$ consists precisely of those indices j for which $f_j = \varepsilon^{\text{degree } f_j} f_{j*}$ and α_j is odd.

Recall that $\varepsilon = \pm 1$ is taken from (2.2).

Proof. As in the proof of Theorem 2.1, we can assume that $r = n$. Let $L(\lambda) = E(\lambda)D_1(\lambda)F(\lambda)$ be the Smith form $L(\lambda)$, where the invariant polynomials $d_1(\lambda), \dots d_r(\lambda)$ are on the main diagonal of $D_1(\lambda)$. Because $L = L_*$, and by the uniqueness of invariant polynomials we have in fact $d_{i*} = \varepsilon^{\text{degree } d_i} d_i \quad (i = 1, \dots, r)$. The factor $\varepsilon^{\text{degree } d_i}$ appears because d_i is monic (this is part of the definition of invariant polynomials) while the leading coefficient of d_{i*} is $\varepsilon^{\text{degree } d_i}$. In the sequel it will be convenient to denote $f_+ = \varepsilon^{\text{degree } f} f_*$ for a scalar polynomial f. Thus, $d_{i+} = d_i$. Observe that f_+ is monic if f is monic and that $(f_1 f_2)_+ = f_{1+} f_{2+}$ for all pairs of polynomials f_1, f_2.

Replacing L by $F_*^{-1} L F^{-1}$, we can further assume without loss of generality that the i-th column of L is divisible by $d_i(\lambda) \quad (i = 1, \dots, n)$. By symmetry, the i-th row of L is divisible by $d_{i*}(\lambda)$. Let the nonconstant invariant polynomials of $L(\lambda)$ be $d_s(\lambda), \dots d_r(\lambda)$ and factor them as in (2.5). Then

$$d_i = \prod_{j=1}^{k_i} (f_{ij})^{\alpha_{ij}} = \prod_{j=1}^{k_i} (f_{ij+})^{\alpha_{ij}} \qquad (i = s, s+1, \dots r),$$

and by the uniqueness of decomposition (2.5) we obtain that the set $\{f_{i1}, \dots, f_{ik_i}\}$ must consist of selfsymmetric polynomials ($f_{ij} = f_{ij+}$) and/or of pairs of mutually symmetric polynomials $f_{ij_1} = f_{ij_2+}$; in this case necessarily $\alpha_{ij_1} = \alpha_{ij_2}$. Say, $f_{i1}, \dots, f_{i,p_i}$ are selfsymmetric and

$$f_{i,p_i+1} = (f_{i,p_i+2})_+, \dots, f_{i,p_i+2q_i-1} = (f_{i,p_i+2q_i})_+;$$

here $p_i + 2q_i = k_i$. Let i_0 be the smallest index such that $\alpha_{i_0 j} > 1$ for some $j \in \{1, \dots, p_i\}$; say, $\alpha_{i_0 1} > 1$ (if no such i_0 exists, we put

$$h_i = \prod_{j=1}^{q_i} (f_{i,p_i+2j})^{\alpha_{i,p_i+2j}} \qquad (i = s, s+1, \dots, r)).$$

Define

$$h_i = \prod_{j=1}^{q_i} (f_{i,p_i+2j})^{\alpha_{i,p_i+2j}} \qquad i = s, s+1, \dots i_0 - 1,$$

$$h_i = f_{i1} \prod_{j=1}^{q_i} (f_{i,p_i+2j})^{\alpha_{i,p_i+2j}} \qquad i = i_0, i_0+1, \dots, r,$$

where we assume that the elementary divisors are numbered so that $f_{i1} = f_{i_0 1}$ for $i = i_0+1, \dots, r$. To make the subsequent formulas more uniform we define also $h_i \equiv 1$ for $i = 1, \dots, s-1$. The divisibility relations between the d_i's imply that whenever $f_{i_1 j_1} = f_{i_2 j_2}$, where $i_1 < i_2$, then $\alpha_{i_1 j_1} \le \alpha_{i_2 j_2}$. Consequently, we obtain that h_i divides h_{i+1} $(i = 1, \dots, r-1)$.

The formulas (2.5) lead to the factorization

$$d_i = h_{i*} g_i h_i \qquad (i = 1, \dots, r), \tag{2.6}$$

where $g_i \equiv 1$ for $i = 1, \dots, s-1$;

$$g_i = \pm \prod_{j=1}^{p_i} (f_{ij})^{\alpha_{ij}} \qquad \text{for} \qquad i = s, \dots, i_0 - 1;$$

$$g_i = \pm f_{i1}^{\alpha_{ij}-2} \prod_{j=2}^{p_i} (f_{ij})^{\alpha_{ij}} \qquad \text{for} \qquad i = i_0, i_0+1, \dots, r.$$

(the sign $+$ or $-$ in g_i is chosen so that g_i is monic). Clearly, g_i divides g_{i+1} for $i = 1, \dots, r-1$.

In view of (2.6) we now have a factorization

$$L(\lambda) = \operatorname{diag}(1, \dots, 1, h_{s*}(\lambda), \dots, h_{r*}(\lambda)) \widetilde{L}(\lambda) \operatorname{diag}(1, \dots, 1, h_s(\lambda), \dots, h_r(\lambda)) \tag{2.7}$$

for some matrix polynomial $\widetilde{L} = \widetilde{L}_*$. Denote by $\widetilde{d}_1(\lambda), \dots, \widetilde{d}_r(\lambda)$ the invariant polynomials of $\widetilde{L}(\lambda)$. Equality (2.7), together with the Binet–Cauchy formula for determinants of submatrices in the product of several matrices implies the following: Every determinant of $j \times j$ submatrix in $L(\lambda)$ is a linear combination (with polynomial coefficients) of the determinants of $j \times j$ submatrices in $\widetilde{L}(\lambda)$ when the determinants are multiplied by $\prod_{i=1}^{j} (h_{i*}(\lambda) h_i(\lambda))$. It follows that $d_1(\lambda) \dots d_j(\lambda)$ divides $\widetilde{d}_1(\lambda) \dots \widetilde{d}_j(\lambda) \prod_{i=1}^{j} (h_{i*}(\lambda) h_i(\lambda))$ for

$j = 1, \dots, r$). The equality (2.6) now shows that $g_1(\lambda) \dots g_j(\lambda)$ divides $\widetilde{d}_1(\lambda) \dots \widetilde{d}_j(\lambda)$. On the other hand, for the (i,j)-th entries $\widetilde{p}_{ij}$ of $\widetilde{L}$ and p_{ij} of L, respectively, we obtain

$$\widetilde{p}_{ij} = h_{i*}^{-1} p_{ij} h_j^{-1} = h_{i*}^{-1} p_{ij} d_i^{-1} h_{j*} g_j,$$

and since (assuming $i \leq j$) both $p_{ij} d_i^{-1}$ and $h_{i*}^{-1} h_{j*}$ are polynomials, $\widetilde{p}_{ij}$ is divisible by g_j. Also, $\widetilde{p}_{ij}$ is divisible by g_i (because g_i divides g_j if $i \leq j$). By the symmetry of $\widetilde{L}$, we obtain that $\widetilde{p}_{ij}$ is divisible by $g_{\max(i,j)}$. Therefore, the determinant of every $j \times j$ submatrix of $\widetilde{L}$ is divisible by $g_1 \dots g_j$. Consequently, $\widetilde{d}_1 \dots \widetilde{d}_j$ divides $g_1 \dots g_j$. Comparing with the previously obtained opposite divisibility relation, we conclude that $g_1, \dots, g_r$ are in fact the invariant polynomials of $\widetilde{L}$.

Repeat the procedure given above with L replaced by $\widetilde{L}$, and so on, until (after a finite number of steps) we obtain a matrix polynomial $D(\lambda)$ with the properties required in Theorem 2.2. □

3. Factorization of selfadjoint matrix polynomials on the real axis. In this section we consider matrix polynomials $L(\lambda)$ over $\mathbb{C}$ with the following property:

$$L(\lambda) = (L(\overline{\lambda}))^* \quad , \quad \lambda \in \mathbb{C}.$$

Such polynomials will be called *selfadjoint*.

THEOREM 3.1. *Let $L(\lambda)$ be a selfadjoint $n \times n$ matrix polynomial. Then $L(\lambda)$ admits a factorization*

$$L(\lambda) = (M(\overline{\lambda}))^* \ D \ M(\lambda), \tag{3.1}$$

where D is an $m \times m$ constant Hermitian matrix and $M(\lambda)$ an $m \times n$ matrix polynomial if and only if

$$m \geq m_0, \tag{3.2}$$

where

$$m_0 = \max_{\lambda \in \mathbb{R}} \nu_+(L(\lambda)) + \max_{\lambda \in \mathbb{R}} \nu_-(L(\lambda)). \tag{3.3}$$

Moreover, in all factorizations (3.1) having the minimal size $m_0 \times m_0$ of D, the matrix D is uniquely determined up to congruence: D has $\max \nu_+(L(\lambda))$ positive eigenvalues and $\max \nu_-(L(\lambda))$ negative eigenvalues (multiplicities counted).

We can say more about the spectral properties of the factor $M(\lambda)$ in (3.1). A set Λ of non-real numbers is called a *c-set* (with respect to a selfadjoint matrix polynomial $L(\lambda)$) if Λ is a maximal (by inclusion) set of non-real singular points of $L(\lambda)$ with the property that $\lambda_0 \in \Lambda \Longrightarrow \overline{\lambda}_0 \notin \Lambda$. (The case when a c-set is empty is not excluded.) The concept of c-set was introduced and used in [GLR1, GLR3]. It turns out that, given $L(\lambda)$ as in Theorem 3.1, and given a c-set Λ, there exists a factorization (3.1) where D is $m_0 \times m_0$

and where the set of non-real singular points of $M(\lambda)$ coincides with Λ. This statement follows as a by-product of the proof (given below) of Theorem 3.1.

Theorem 3.1 admits an alternative formulation. An $n \times n$ matrix polynomial $M(\lambda)$ will be called *elementary* if $r(M) = 1$ and $M(\lambda)$ is positive semidefinite for all real λ. It is not difficult to see (this fact is actually a particular case of Theorem 3.1) that $M(\lambda)$ is elementary if and only if $M(\lambda)$ is of the form $M(\lambda) = x(\lambda)(x(\overline{\lambda}))^*$, where $x(\lambda) \not\equiv 0$ is an $n \times 1$ column polynomial. One can consider elementary matrix polynomials as building blocks for selfadjoint matrix polynomial, in the same spirit as the constant rank 1 positive semidefinite matrices are building blocks for constant Hermitian matrices:

THEOREM 3.2. *Any selfadjoint $n \times n$ matrix polynomial $L(\lambda)$ admits a representation*

$$L(\lambda) = \sum_{j=1}^{m} \varepsilon_j M_j(\lambda), \tag{3.4}$$

where $\varepsilon_j = \pm 1$ and $M_j(\lambda)$ are elementary matrices. The number m of terms in (3.4) is greater than or equal to m_0, where m_0 is given by (3.3), and if $m = m_0$, then exactly $\max \nu_+(L(\lambda))$ *of ε_j's are equal to +1 and exactly* $\max \nu_- L(\lambda))$ *of ε_j's are equal to* -1.

To obtain Theorem 3.2 from Theorem 3.1, assume (without loss of generality) that in (3.1) D is a diagonal matrix with ± 1's on the main diagonal. Then let $M_j(\lambda) = (x_j(\overline{\lambda}))^* x_j(\lambda)$, where $x_j(\lambda)$ is the j'th row of $M(\lambda)$, to produce the formula (3.4).

COROLLARY 3.3. *Any selfadjoint $n \times n$ matrix polynomial admits a factorization (3.1), or a representation (3.4), where $m \leq 2n$.*

There are selfadjoint matrix polynomials, for example, $L(\lambda) = \lambda I$, for which there do not exist representations (3.1) or (3.4) with $m < 2n$.

The rest of this section will be devoted to the proof of Theorem 3.1.

We start with the easy direction. Let be given a factorization (3.1), and let λ_0 be a real point for which

$$\nu_+(L(\lambda_0)) = \max_{\lambda \in \mathbb{R}} \nu_+(L(\lambda)).$$

As $L(\lambda_0) = Y^* D Y$, where $Y = M(\lambda_0)$, the Hermitian matrix D must have at least $\nu_+(L(\lambda_0))$ positive eigenvalues. Analogously, D must have at least $\nu_-(L(\lambda_1))$ negative eigenvalues, where $\lambda_1 \in \mathbb{R}$ is chosen so that

$$\nu_-(L(\lambda_1)) = \max_{\lambda \in \mathbb{R}} \nu_-(L(\lambda)).$$

We obtain therefore the inequality (3.2). It is also clear that in any factorization (3.1), where D is $m_0 \times m_0$, the Hermitian matrix D is unique up to congruence.

It remains to show that a given selfadjoint matrix polynomial $L(\lambda)$ admits a factorization (3.1) with $m_0 \times m_0$ the size of D. This is the difficult part and we need some preliminaries. Note that $L(\lambda)$ is selfadjoint if and only if $L = L_*$, where the transformation $a \to a_*$ is defined as in Section 2, with $F = \mathbb{C}$, $\sigma(x) = \overline{x}$ ($x \in \mathbb{C}$), and $\varepsilon = 1$. Nevertheless, here the general results of Section 2 will not be used because the preliminary results we need (such as Proposition 3.4 below) are already available in the literature (it should be noted however that the result of Theorem 2.2 plays an essential role in the proof of Proposition 3.4).

First observe that there exists an $n \times n$ matrix polynomial $N(\lambda)$ with constant non-zero determinant such that

$$L(\lambda) = (N(\overline{\lambda}))^* \begin{bmatrix} L_0(\lambda) & 0 \\ 0 & 0 \end{bmatrix} N(\lambda), \tag{3.5}$$

where $L_0(\lambda)$ is a selfadjoint $k \times k$ matrix polynomial, $k = r(L)$. See, e.g., Theorem 32.4 in [M], where (3.5) is proved for symmetric matrices over principal ideal rings, with $(N(\overline{\lambda}))^*$ replaced by $N(\lambda)^T$; the same proof works to produce (3.5); also, (3.5) can be obtained without difficulties from the Smith form of $L(\lambda)$ (see Section 2). Because of (3.5) we can (and will) assume from the very beginning that the general rank of L is equal to n, i.e. $\det L(\lambda) \not\equiv 0$.

Our next observation is that the result of Theorem 3.1 is known in the case $L(\lambda)$ has *constant signature*, i.e. $\nu_+(L(\lambda))$, and therefore also $\nu_-(L(\lambda))$ and $\nu_0(L(\lambda))$, is constant for all real regular points λ:

PROPOSITION 3.4. *([GLR2]) Let $L(\lambda)$ be a selfadjoint $n \times n$ matrix polynomial such that*

$$\max_{\lambda \in \mathbb{R}} \nu_+(L(\lambda)) + \max_{\lambda \in \mathbb{R}} \nu_-(L(\lambda)) = n.$$

(necessarily $\det L(\lambda) \not\equiv 0$). Then $L(\lambda)$ admits a factorization (3.1) with $n \times n$ the size of D.

We will prove the following lemma:

LEMMA 3.5. *Let $L(\lambda)$ be a selfadjoint $n \times n$ matrix polynomial with $\det L(\lambda) \not\equiv 0$, and let $m_0 \geq n$ be defined by (2.3). Then there exists an $m_0 \times m_0$ selfadjoint matrix polynomial $\widetilde{L}(\lambda)$ such that*

$$\widetilde{L}(\lambda) = \begin{bmatrix} L(\lambda) & 0 \\ 0 & * \end{bmatrix}$$

and such that

$$\max_{\lambda \in \mathbb{R}} \nu_+(\widetilde{L}(\lambda)) + \max_{\lambda \in \mathbb{R}} \nu_-(\widetilde{L}(\lambda)) = m_0, \tag{3.6}$$

or equivalently $\widetilde{L}$ is regular and has constant signature.

Proof. By Rellich's Theorem [R] (see also [GLR3]) the eigenvalues $\mu_1(\lambda), \dots, \mu_n(\lambda)$ of $L(\lambda)$ for λ real can be enumerated so that $\mu_1(\lambda), \dots, \mu_n(\lambda)$ are real analytic functions of the real variable λ. Clearly, $\lambda_0 \in \mathbb{R}$ is singular if and only if λ_0 is a zero of at least one of the analytic functions $\mu_1(\lambda), \dots, \mu_n(\lambda)$. Let $\lambda_0 \in \mathbb{R}$ be singular, and let

$$\Omega(\lambda_0) = \{1 \le j \le n \mid \mu_j(\lambda_0) = 0\}.$$

For every $j \in \Omega(\lambda_0)$ let m_j be the multiplicity of λ_0 as a zero of $\mu_j(\lambda)$, and let ε_j be the sign of the non-zero real number $[\mu_j^{(m_j)}(\lambda)]|_{\lambda=\lambda_0}$. (We suppress the dependence of m_j and ε_j on λ_0 in the notation.) Define the integer $q(\lambda_0)$ by

$$\begin{aligned} q(\lambda_0) = \{ \ \# \text{ of indices } j \in \Omega(\lambda_0) \text{ such that } m_j \text{ is odd and } \varepsilon_j = 1\} \\ - \{\# \text{ of indices } j \in \Omega(\lambda_0) \text{ such that } m_j \text{ is odd and } \varepsilon_j = -1\}. \end{aligned}$$

From the definition of $q(\lambda_0)$ it is clear that

$$\nu_+(L(\lambda_0 + \varepsilon)) - \nu_+(L(\lambda_0 - \varepsilon)) = q(\lambda_0) \tag{3.7}$$

for all sufficiently small $\varepsilon > 0$. It is easy to see that

$$m_0 := \max_{\lambda \in \mathbb{R}} \nu_+(L(\lambda)) + \max_{\lambda \in \mathbb{R}} \nu_-(L(\lambda)) = n + \max |\nu_+(L(\lambda_1)) - \nu_+(L(\lambda_2))| \tag{3.8}$$

where the maximum is taken over all regular real points λ_1 and λ_2. Also, as it follows from (3.7),

$$\max |\nu_+(L(\lambda_1)) - \nu_+(L(\lambda_2))| = \max_{\lambda_1 < \lambda_2} \left| \sum_{\lambda_1 < \lambda_0 < \lambda_2} q(\lambda_0) \right| \tag{3.9}$$

where the summation in the right hand side of (3.9) is over all singular points λ_0 in the interval $\lambda_1 < \lambda_0 < \lambda_2$.

Denote the right-hand side of (3.9) by p. We now construct p scalar real polynomials $r_1(\lambda), \dots, r_p(\lambda)$ with the following properties:

(i) all zeros of $r_j(\lambda)$ $(j = 1, \dots, p)$ are real and simple and belong to the set S of real singular points λ_0 of $L(\lambda)$ for which $q(\lambda_0) \neq 0$;

(ii) for every $\lambda \in S$ exactly $|q(\lambda_0)|$ polynomials among $r_1(\lambda), \dots, r_p(\lambda)$ have λ_0 as their zeros; and for each $r_j(\lambda)$ such that $r_j(\lambda_0) = 0$ we have $q(\lambda_0) r_j'(\lambda_0) < 0$.

The definition of p ensures that such polynomials $r_1(\lambda), \dots r_p(\lambda)$ can indeed be constructed. Let

$$\widetilde{L}(\lambda) = \operatorname{diag}(L(\lambda), r_1(\lambda), \dots, r_p(\lambda)).$$

By the property (ii), an in view of the qualities (3.8), (3.9), it is easy to see that the number of positive eigenvalues of $\widetilde{L}(\lambda)$ is constant for every real λ which is a regular point for $L(\lambda)$. The equality of (3.6) therefore follows. □

Now we can easily finish the proof of Theorem 3.1. Indeed, given a selfadjoint matrix polynomial $L(\lambda)$ with $\det L(\lambda) \not\equiv 0$, construct $\widetilde{L}(\lambda)$ as in Lemma 3.5 and apply Proposition 3.4 to $\widetilde{L}(\lambda)$:

$$\widetilde{L}(\lambda) = (N(\overline{\lambda}))^* D\ N(\lambda),$$

where D is a constant $m_0 \times m_0$ Hermitian matrix. Then (3.1) holds for $M(\lambda)$ formed by first n columns of $N(\lambda)$. □

4. Factorization of real symmetric matrix polynomials. Let $L(\lambda) = \sum_{j=0}^{\ell} \lambda^j A_j$ be a *real symmetric* matrix polynomial, i.e. A_j ($j = 0, \ldots, \ell$) are real symmetric $n \times n$ matrices. For such polynomials $L(\lambda)$ we consider factorizations

$$L(\lambda) = (M(\lambda))^T DM(\lambda), \tag{4.1}$$

where D is a constant real symmetric $m \times m$ matrix and $M(\lambda)$ is a matrix polynomial with real coefficients.

It will be convenient to state the next theorem in terms of elementary divisors (see Section 2 for definitions of the concepts related to elementary divisors).

THEOREM 4.1. *Let $L(\lambda)$ be a real symmetric $n \times n$ matrix polynomial, and assume that the elementary divisors of $L(\lambda)$ which are powers of irreducible quadratic polynomials (over $\mathbb{R}$) all have even orders. Then $L(\lambda)$ admits a factorization (4.1) if and only if $m \geq m_0$, where m_0 is defined by (3.3). Moreover, in factorization (4.1) with the minimal possible size of D, the matrix D is uniquely determined up to congruence and has exactly $\max_{\lambda \in \mathbb{R}} \nu_+(L(\lambda))$ positive eigenvalues and exactly $\max_{\lambda \in \mathbb{R}} \nu_-(L(\lambda))$ negative eigenvalues, multiplicities counted. Alternatively, $L(\lambda)$ admits a representation*

$$L(\lambda) = \sum_{j=1}^{m} \varepsilon_j M_j(\lambda), \tag{4.2}$$

where $M_j(\lambda)$ are real elementary matrices and $\varepsilon_j = \pm 1$, and if and only if $m \geq m_0$. In case $m = m_0$, the number of $+1$'s (resp -1's) among the $\varepsilon_1, \ldots, \varepsilon_m$ is exactly $\max_{\lambda \in \mathbb{R}} \nu_+(L(\lambda)$ (resp $\max_{\lambda \in \mathbb{R}} \nu_-(L(\lambda)))$.

In particular, Theorem 4.1 applies if all singular points of $L(\lambda)$ are real.

Proof. The "only if" part (easy direction) is proved as in the proof of Theorem 3.1. Also, we can easily reduce the proof to the case when $\det L(\lambda) \not\equiv 0$. Using Theorem 2.2 (with $F = \mathbb{R}$, $\sigma =$ identity, $\varepsilon = 1$), we further can assume that all elementary divisors of $L(\lambda)$ are first degree polynomials (necessarily with real roots). From now on the proof proceeds in the same way as that of Theorem 3.1. The role of Proposition 3.4 is played by Proposition 4.2 below. □

PROPOSITION 4.2. *Let $L(\lambda)$ be a real symmetry $n \times n$ matrix polynomial with all elementary divisors first degree polynomials. Assume further that*

$$\max_{\lambda \in \mathbb{R}} \nu_+(L(\lambda)) + \max_{\lambda \in \mathbb{R}} \nu_-(L(\lambda)) = n.$$

Then $L(\lambda)$ admits a factorization (4.1) with $n \times n$ the size of D.

Proposition 4.2 can be proved by repeating the arguments leading to the proof of Theorem 1 in [GLR2]. We omit the details.

If the hypothesis on the orders of elementary divisors of $L(\lambda)$ is omitted in Theorem 4.1, easy scalar examples (for example, $L(\lambda) = \lambda^2 + 1$) show that the result of Theorem 4.1 is generally not valid. Scalar examples show also that, in this case the matrices D of minimal size in factorizations (4.1) are not necessarily congruent to each other:

$$\lambda^2 + 1 = [\lambda \quad 1] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda \\ 1 \end{bmatrix} = \left[\frac{1}{2}(\lambda^2 + 2), \frac{1}{2}\ \lambda^2\right] \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{2}(\lambda^2 + 2) \\ \frac{1}{2}\ \lambda^2 \end{bmatrix}.$$

We have, however, and upper bound on the minimal rise of D:

THEOREM 4.3. *Let $L(\lambda)$ be a real symmetric $n \times n$ matrix polynomial, and let m_0 be defined by (3.3). Then for every $m \geq 2\min(m_0, n) L(\lambda)$ admits a factorization (4.1).*

Proof. Assume first that $m_0 \leq n$. By Theorem 3.1, we have

$$L(\lambda) = (M(\overline{\lambda}))^* D\ M(\lambda), \tag{4.3}$$

where D is $m_0 \times m_0$ constant Hermitian matrix (which can can be chosen to be real without loss of generality), and $M(\lambda)$ is a complex matrix polynomial. Write $M(\lambda) = M_1(\lambda) + iM_2(\lambda)$, where $M_1(\lambda)$ and $M_2(\lambda)$ are real matrix polynomials. Then, separating the real part in (4.3) we obtain

$$L(\lambda) = [M_1(\lambda)^T \qquad M_2(\lambda)^T] \begin{bmatrix} D & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} M_1(\lambda) \\ M_2(\lambda) \end{bmatrix},$$

which is the desired factorization (with $m = 2m_0$).

If $m_0 > n$, use the simple equality:

$$4L(\lambda) = [L(\lambda) + I \qquad L(\lambda) - I] \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} L(\lambda) + I \\ L(\lambda) - I \end{bmatrix}. \quad \square$$

5. Factorization of symmetric real polynomials on the imaginary axis. In this section we consider the case of $n \times n$ matrix polynomials $L(\lambda)$ such that $L(\lambda) = (L(-\lambda))^T$ and $L(\lambda)$ is real for real λ. Note that such a polynomial is selfadjoint on the imaginary axis, i.e., $L(\lambda) = (L(\lambda))^*$ for $\lambda \in i\mathbb{R}$. An immediate consequence of Theorem 3.1 (applied to $L(i\lambda)$) is that such matrix polynomial admits a factorization

$$L(\lambda) = (M(-\lambda))^T D \, M(\lambda) \tag{5.1}$$

for a complex $m \times m$ Hermitian matrix D and a complex $m \times n$ matrix polynomial $M(\lambda)$ if and only if

$$m \geq \max_{\lambda \in i\mathbb{R}} \nu_+(L(\lambda)) + \max_{\lambda \in i\mathbb{R}} \nu_-(L(\lambda)).$$

We shall show in this section that D and $M(\lambda)$ can be taken real. Note that here is a contrast with the situation of Section 4, where an analogous factorization of a real symmetric matrix polynomial having real factors is not always possible.

First we shall deal with the case when $L(\lambda)$ is regular and has constant signature (on the imaginary axis), after which the general case is reduced to the case of constant signature.

So in view of Theorem 2.2 (with $F = \mathbb{R}$, $\sigma =$ identity, $\varepsilon = -1$) we can restrict our attention to matrix polynomials having only elementary divisors of the form λ or $\lambda^2 + \lambda_0^2$ where λ_0 is real and nonzero. Next, we deal with the case when L is regular and has constant signature.

THEOREM 5.1. *Suppose $L(\lambda)$ is a real regular $n \times n$ matrix polynomial satisfying $L(\lambda) = L(-\lambda)^T$ and having constant signature on the imaginary axis: $\nu_+(L(\lambda))$ is constant for all regular points $\lambda \in i\mathbb{R}$. Then L admits a factorization*

$$L(\lambda) = (M(-\lambda))^T D \, M(\lambda)$$

where M is an $n \times n$ matrix polynomial with real coefficients and D is an $n \times n$ constant real matrix.

Proof. Again by Theorem 2.2 we may assume L has only pure imaginary eigenvalues and all elementary divisors are linear (in the sense of $\mathbb{C}$). First we deal with the case when $\lambda^2 + \lambda_0^2$, $\lambda_0 \in \mathbb{R} \backslash \{0\}$ is an elementary divisor of L. Using the Smith form of $L(\lambda)$, write

$$L(\lambda) = E(\lambda) \operatorname{diag}((\lambda^2 + \lambda_0^2)p_1(\lambda), \ldots, (\lambda^2 + \lambda_0^2)p_q(\lambda), p_{q+1}(\lambda), \ldots, p_n(\lambda))F(\lambda),$$

where $p_j(\lambda)$ $(j = 1, \ldots, n)$ are real monic scalar polynomials and $E(\lambda)$, $F(\lambda)$ are real $n \times n$ matrix polynomials with $\det E(\lambda) \equiv \text{const} . \neq 0$, $\det F(\lambda) \equiv \text{const} . \neq 0$. Then we have for

$$\widehat{L}(\lambda) := F(-\lambda)^{-T} L(\lambda) F(\lambda)^{-1} :$$

$$\widehat{L}(\lambda) = \begin{bmatrix} (\lambda^2 + \lambda_0^2)\widehat{B}_{11}(\lambda) & (\lambda^2 + \lambda_0^2)\widehat{B}_{12}(\lambda) \\ (\lambda^2 + \lambda_0^2)\widehat{B}_{21}(\lambda) & \widehat{A}_{22}(\lambda) \end{bmatrix} \tag{5.2}$$

where $\widehat{B}_{11}$ is a $q \times q$ matrix polynomial. Moreover $\widehat{B}_{11}$ $(\pm i\lambda_0)$ must be invertible, as otherwise $\det \widehat{L}(\lambda)$ and hence also $\det L(\lambda)$ would be divisible by $(\lambda^2 + \lambda_0^2)^{q+1}$. Note: if L has constant signature, so has $\widehat{L}$, which means $\widehat{L}(i\lambda)$ for $\lambda \in \mathbb{R}$ is a Hermitian matrix having constant signature for all real λ except at a finite number of points. Using Rellich's theorem [R], we can write

$$\widehat{L}(i\lambda) = (U(\lambda))^* \operatorname{diag}(\mu_1(\lambda), \dots, \mu_n(\lambda)) U(\lambda) \quad , \quad \lambda \in \mathbb{R}$$

where $U(\lambda)$ is unitary valued and analytic, and $\mu_j(\lambda)$ is analytic and real. The functions $\mu_j(\lambda)$ have simple zeros only as $\widehat{L}$ has only linear elementary divisors (over $\mathbb{C}$). Without loss of generality we may assume $\mu_1(\lambda_0) = \cdots = \mu_q(\lambda_0) = 0$. By Lemma 6 in [GLR2] we have that q is even, and exactly half of the numbers $\mu_j'(\lambda_0)(j = 1, \dots, q)$ is positive, the other half if negative. Let u_j be the j-th column of $(U(\lambda_0))^*$. Then one calculates

$$\langle \widehat{L}'(i\lambda_0) u_j, u_i \rangle = \langle D'(\lambda_0) e_j, e_i \rangle,$$

where $D(\lambda) = \operatorname{diag}(\mu_1(\lambda), \dots, \mu_n(\lambda))$. Note that $u_1, \dots, u_q$ span Ker $L(i\lambda_0)$. So on Ker $L(i\lambda_0)$ the quadratic form given by $\widehat{L}'(i\lambda_0)$ has $\frac{q}{2}$ positive squares and $\frac{q}{2}$ negative squares. Now by (5.2) Ker $L(i\lambda_0)$ is span $\{e_1, \dots, e_q\}$, and for $x, y \in$ Ker $L(i\lambda_0)$ we have

$$\frac{d}{d\lambda} \langle L(i\lambda)x, y \rangle \rfloor_{\lambda = \lambda_0} = -2\lambda_0 \langle \widehat{B}_{11}(i\lambda_0)x, y \rangle.$$

Therefore we conclude that there is an invertible matrix V such that

$$\widehat{B}_{11}(i\lambda_0) = V^* \left\{ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \oplus \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \right\} V, \tag{5.3}$$

where the block $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ is repeated $\frac{q}{2}$ times. Moreover, a simple argument shows that V can be taken such that it has a real determinant.

Now we will first state and prove a lemma, after which we shall return to the proof of Theorem 5.1.

LEMMA 5.2. *Let W be a complex invertible $n \times n$ matrix with real determinant, and let λ_0 be a non-zero real number. Then there exists a real $n \times n$ matrix polynomial $M(\lambda)$ with constant determinant such that $M(i\lambda_0) = W$.*

Proof. We can decompose W as a product of elementary matrices:

$$W = W_1 \cdot W_2 \cdot \dots W_k, \tag{5.4}$$

where each W_j is either triangular with ones on the diagonal and exactly one non-zero off-diagonal entry, or W_j is a diagonal invertible matrix. Multiplying each diagonal w_j by a suitable complex number α_j so that $\det(\alpha_j W_j)$ is real, we can assume without loss of generality that $\det W_j$ is real ($j = 1, \ldots, k$); here we use the hypothesis that $\det W$ is real. Furthermore, by writing

$$W_j = \operatorname{diag}(c_{j1}, \ldots, c_{jn}) = \operatorname{diag}(c_{j1}, \overline{c}_{j1}, 1, \ldots, 1) \cdot \operatorname{diag}(1, c_{j2}\overline{c}_{j1}^{-1}, \overline{c}_{j2}c_{j1}^{-1}, 1, \ldots 1) \cdot$$
$$\cdots \operatorname{diag}(1, 1, \ldots, p_{jn}, c_{jn})$$

(here $p_{jn} \in \mathbb{C} - \{0\}$), we can assume that every diagonal matrix W_j in (5.4) has real non-zero determinant and at most two diagonal entries different from 1 (located in adjacent positions). Clearly, it will suffice to construct a polynomial $M(\lambda)$ as required such that $M(i\lambda_0) = W_j$ (for a fixed j). If W_j is triangular, let

$$M(\lambda) = \frac{1}{2}(W_j + \overline{W}_j) + \frac{\lambda}{2\lambda_0} \cdot \frac{1}{i}(W_j - \overline{W}_j).$$

If $W_j = \operatorname{diag}(1, \ldots, 1, d_1, d_2, 1, \ldots, 1)$ with d_1, d_2 real than the constant $M(\lambda) \equiv W_j$ will do. Finally, if $W_j = \operatorname{diag}(1, \ldots, 1, d_1, d_2, 1, \ldots, 1)$ with $d_1 d_2 \in \mathbb{R}$ but $d_1 \notin \mathbb{R}$, then put

$$M(\lambda) = \operatorname{diag}\left(1, \ldots, 1, \begin{bmatrix} p(\lambda) & (\lambda^2 + \lambda_0^2)r \\ \lambda^2 + \lambda_0^2 & q(\lambda) \end{bmatrix}, 1, \ldots, 1\right). \tag{5.5}$$

Here

$$p(\lambda) = d_{1R} + \lambda\lambda_0^{-1} d_{1I}, \qquad r = d_1 d_2 (\lambda_0^2 + (\lambda_0 d_{1R} d_{1I}^{-1})^2)^{-2}$$
$$q(\lambda) = p(\lambda)^{-1}[r\lambda^4 + 2r\lambda^2\lambda_0^2 + r\lambda_0^4 + d_1 d_2],$$

where d_{1R} (resp. d_{1I}) stands for the real (resp. imaginary) part of d_1. (The 2×2 block $\begin{bmatrix} p(\lambda) & (\lambda^2 + \lambda_0^2)r \\ \lambda^2 + \lambda_0^2 & q(\lambda) \end{bmatrix}$ is in the same position in $M(\lambda)$ as the position of $\begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}$ is in W_j.) It is easy to verify that $q(\lambda)$ is in fact a real polynomial, $M(\lambda)$ (defined by (5.5)) is a real matrix polynomial with constant non-zero determinant, and $M(i\lambda_0) = W_j$. □

Now let us return to the proof of Theorem 5.1. Let V be as in (5.3), and choose $M(\lambda)$, a real $q \times q$ polynomial with constant non-zero determinant such that $M(i\lambda_0) = V$. This is possible by Lemma 5.2 (recall $\det V$ is real). Now we may replace $\widehat{L}(\lambda)$ by

$$\widetilde{L}(\lambda) = \begin{bmatrix} M(-\lambda)^{-T} & 0 \\ 0 & I \end{bmatrix} \widehat{L}(\lambda) \begin{bmatrix} M(\lambda)^{-1} & 0 \\ 0 & I \end{bmatrix}.$$

As M has constant non-zero determinant, $\widetilde{L}$ is a matrix polynomial, and we may write

$$\widetilde{L}(\lambda) = \begin{bmatrix} (\lambda^2+\lambda_0^2)B_{11}(\lambda) & (\lambda^2+\lambda_0^2)B_{12}(\lambda) \\ (\lambda^2+\lambda_0^2)B_{21}(\lambda) & A_{22}(\lambda) \end{bmatrix}$$

where

$$B_{11}(i\lambda_0) = \mathrm{diag}(1,-1,\ldots,1,-1).$$

Now put

$$K(\lambda) = \begin{bmatrix} \lambda & -\lambda_0 \\ \lambda_0 & \lambda \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} \lambda & -\lambda_0 \\ \lambda_0 & \lambda \end{bmatrix} \oplus I_{n-q}$$

where the leading block is repeated $\dfrac{q}{2}$ times.

We shall show that $N(\lambda) = K(-\lambda)^{-T}\widetilde{L}(\lambda)K(\lambda)^{-1}$ is a polynomial. Note that it may have a pole only at $\pm i\lambda_0$, and it suffices to show it has no pole at either one of these points. Moreover, any pole of N must appear in its leading $q\times q$ block. This leading $q \times q$ block equals

$$(\lambda^2+\lambda_0^2)^{-1}\left(\begin{bmatrix} -\lambda & -\lambda_0 \\ \lambda_0 & -\lambda \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} -\lambda & -\lambda_0 \\ \lambda_0 & -\lambda \end{bmatrix}\right) B_{11}(\lambda) \left(\begin{bmatrix} \lambda & \lambda_0 \\ -\lambda_0 & \lambda \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} \lambda & \lambda_0 \\ -\lambda_0 & \lambda \end{bmatrix}\right)$$

as an easy computation shows. Now at $\lambda = i\lambda_0$ we have that

$$\begin{bmatrix} -i\lambda_0 & -\lambda_0 \\ \lambda_0 & -i\lambda_0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} i\lambda_0 & \lambda_0 \\ -\lambda_0 & i\lambda_0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Recalling that $B_{11}(i\lambda_0) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ we see that $N(\lambda)$ is a polynomial. Taking determinants we see that $N(\lambda)$ has no eigenvalue at $\pm i\lambda_0$. Applying the same argument at each non-zero, singular point of L we reduce the proof of Theorem 5.1 to the case when zero is the only possible singular point of $L(\lambda)$. However, for that case a similar argument shows that $L(\lambda)$ admits a representation $L(\lambda) = (K(-\lambda))^T\ N(\lambda)\ K(\lambda)$ with $K(\lambda)$ a real matrix polynomial and $N(\lambda)$ a real matrix polynomial without singular points (cf. the proof of Proposition 3.4 given in [GLR2]).

So we have reduced to the case where L has no singular points. In this case the result follows from Theorem 2.1. □

Next we state the main result of this section.

THEOREM 5.3. *Let $L(\lambda) = (L(-\lambda))^T$ be an $n \times n$ matrix polynomial with real coefficients. Then there is a real $m \times m$ matrix $D = D^T$ and a $m \times n$ matrix polynomial $M(\lambda)$ with real coefficients such that*

$$L(\lambda) = (M(-\lambda))^T D\ M(\lambda) \tag{5.6}$$

if and only if

$$m \geq m_0 := \max_{\lambda \in i\mathbb{R}} \nu_+(L(\lambda)) + \max_{\lambda \in i\mathbb{R}} \nu_-(L(\lambda)). \tag{5.7}$$

Moreover, when $m = m_0$, the matrix D is unique up to congruence by a real orthogonal matrix.

Analogously to Theorem 3.1, the polynomial $M(\lambda)$ in Theorem 5.3 can be chosen with additional spectral properties. Given a polynomial $L(\lambda)$ as in Theorem 5.3, a set Λ of numbers with non-zero real parts will be called a *d-set* (with respect to $L(\lambda)$) if Λ is a maximal set of singular points of $L(\lambda)$ with non-zero real parts having the property that $\lambda \in \Lambda \Longrightarrow \overline{\lambda} \in \Lambda, -\overline{\lambda} \notin \Lambda$. (A d-set may be empty.) It turns out that under the hypotheses of Theorem 5.3, for every given d-set Λ there exists a factorization (5.6) where $D = D^T$ is $m_0 \times m_0$ and where the set of non-real singular points of $M(\lambda)$ coincides with Λ. This follows as a by-product of the proof of Theorem 5.3 (including Theorem 5.1 and Theorem 2.2 with $F = \mathbb{R}$, $\sigma =$ identity, $\varepsilon = -1$).

Proof. The uniqueness of D is verified as in the proof of Theorem 3.1, as well as the fact that $m \geq m_0$ is necessary for the existence of real $D = D^T$ and $M(\lambda)$ such that (5.6) is satisfied is seen as in the proof of Theorem 3.1. It remains to prove sufficiency. We may reduce to the regular case again as in Section 3. In case L has constant signature we are finished, using Theorem 5.1. In case L does not have constant signature on the imaginary axis, it will be shown that there exists a real $m_0 \times m_0$ matrix polynomial $\widetilde{L}(\lambda)$ such that $\widetilde{L}(\lambda) = \widetilde{L}(-\lambda)^T$ and

$$\widetilde{L}(\lambda) = \begin{bmatrix} L(\lambda) & 0 \\ 0 & * \end{bmatrix},$$

while $\widetilde{L}$ is regular and has constant signature on the imaginary axis. Indeed, as $L(i\lambda)$ is selfadjoint for real λ we can write (using Rellich's theorem [R], also [GLR3])

$$L(i\lambda) = (U(\lambda))^* \operatorname{diag}(\mu_1(\lambda), \ldots, \mu_n(\lambda))U(\lambda)$$

where $\mu_j(\lambda)$ is analytic and real valued and $U(\lambda)$ is analytic and unitary. Since $L(i\lambda) = \overline{L(-i\lambda)}$, $\lambda \in \mathbb{R}$, the matrices $L(i\lambda)$ and $L(-i\lambda)$ have the same eigenvalues, and therefore for every point $\lambda_0 \in \mathbb{R}$ there is a permutation σ on $\{1, \ldots, n\}$ such that

$$\mu_i(-\lambda) = \mu_{\sigma(i)}(\lambda) \qquad (i = 1, \ldots, n) \tag{5.8}$$

in a neighborhood of λ_0.

Let $\lambda_0 \in \mathbb{R}$ be such that $\pm i\lambda_0$ are singular points of $L(\lambda)$. Define $\Omega(\lambda_0)$ and $q(\lambda_0)$ as in the proof of Lemma 3.5. It follows from (5.8) that $q(\lambda_0) = -q(-\lambda_0)$; in particular,

$q(0) = 0$ (if $\lambda_0 = 0$ is a singular point of $L(\lambda)$). Furthermore (analogously to (3.8) and (3.9))

$$m_0 = \max_{\lambda>0} \nu_+(L(i\lambda)) + \max_{\lambda>0} \nu_-(L(i\lambda)) =$$
$$= n + \max |\nu_+(L(i\lambda_1)) - \nu_+(L(i\lambda_2))|,$$

and

$$\max |\nu_+(L(i\lambda_1)) - \nu_+ L(i\lambda_2))| =$$
$$= \max_{\lambda_1<\lambda_2} \left| \sum_{\lambda_1<\lambda_0<\lambda_2} q(\lambda_0) \right| \tag{5.9}$$

where the summation is over all $\lambda_0 \in (\lambda_1, \lambda_2)$ such that $i\lambda_0$ is a singular point of L. (here the real numbers λ_1 and λ_2 are such that $i\lambda_1$ and $i\lambda_2$ are regular points of L). Denote the number (5.9) by p, as in the proof of Lemma 3.5. Now construct p polynomials $r_1(\lambda), \dots, r_p(\lambda)$ with real coefficients having the following properties:

(i) $r_j(\lambda) = r_j(-\lambda)$ is real for $\lambda \in \mathbb{R}$,

(ii) all zeros of r_j are pure imaginary, non-zero numbers and belong to the set S of pure imaginary singular points λ_0 of L for which $q(\lambda_0) \neq 0$,

(iii) for every $\lambda_0 \in S$ exactly $|q(\lambda_0)|$ polynomials among $r_1, \dots, r_p$ have λ_0 as a zero and for each r_j having λ_0 as a zero we have

$$q(\lambda_0) \frac{d}{d\lambda} r_j(i\lambda) \Big|_{\lambda=\lambda_0} < 0 \ .$$

(Note that because of (i) and $q(-\lambda_0) = -q(\lambda_0)$ condition (iii) is satisfied at $-\lambda_0$ if it is satisfied at λ_0.) Put

$$\widetilde{L}(\lambda) = L(\lambda) \oplus \mathrm{diag}(r_1(\lambda), \dots, r_p(\lambda)).$$

Then $\widetilde{L}(\lambda)$ is regular and has constant signature on the imaginary axis as desired. Thus by Theorem 5.1 $\widetilde{L}$ admits a factorization

$$\widetilde{L}(\lambda) = (N(-\lambda))^T \ D \ N(\lambda)$$

with D an $m_0 \times m_0$ real matrix and $N(\lambda)$ an $m_0 \times n$ real matrix polynomial. Taking for $M(\lambda)$ the matrix polynomial formed by the first n columns of N now finishes the proof. $\square$

Analogously to Theorem 3.2, the result of Theorem 5.3 can be put in terms of additive representations of $L(\lambda)$ via elementary matrix polynomials. Here, a real $n \times n$ matrix polynomial $M(\lambda)$ will be called *elementary* if $r(M) = 1$ and $M(\lambda)$ is positive semidefinite Hermitian for all $\lambda \in i\mathbb{R}$.

THEOREM 5.4. *Let $L(\lambda)$ be as in Theorem 5.3. Then $L(\lambda)$ admits a representation*

$$L(\lambda) = \sum_{j=1}^{q} \varepsilon_j M_j(\lambda) \tag{5.10}$$

where $\varepsilon_j = \pm 1$ and $M_j(\lambda)$ are elementary matrix polynomials, if and only if $q \geq m_0$, where m_0 is defined by (5.7). Moreover, when $q = m_0$, exactly $\max\limits_{\lambda_0 \in i\mathbb{R}} \nu_+(L(\lambda_0))$ of the ε_j's in (5.10) are equal to $+1$, and exactly $\max\limits_{\lambda \in i\mathbb{R}} \nu_-(L(\lambda))$ of them are equal to -1.

We omit the easy derivation of Theorem 5.4 from Theorem 5.3.

6. Factorization of complex symmetric polynomials. In this section we consider $n \times n$ matrix polynomials $L(\lambda)$ with complex coefficients having the symmetry

$$L(\lambda) = (L(\varepsilon\lambda))^T \tag{6.1}$$

where $\varepsilon = \pm 1$ is fixed, and their factorizations of the form

$$L(\lambda) = (M(\varepsilon\lambda))^T D M(\lambda) \tag{6.2}$$

where $M(\lambda)$ is an $m \times n$ matrix polynomial (over $\mathbb{C}$), and D is a constant complex symmetric matrix. Observe that every $m \times m$ complex symmetric matrix D can be factored as $D = V^T V$ for some complex matrix V (see, e.g., p. 159 in [HJ]). Therefore, we may assume that $D = I$ in (6.2).

Here (in contrast with Sections 3–5) signatures of Hermitian matrices do not play a role.

We start with the case $\varepsilon = -1$.

THEOREM 6.1. *Let $L(\lambda)$ be an $n \times n$ matrix polynomial satisfying (6.1), where $\varepsilon = -1$. Then the minimal size m for which $L(\lambda)$ admits a factorization*

$$L(\lambda) = (M(-\lambda))^T M(\lambda) \tag{6.3}$$

with an $m \times n$ matrix polynomial $M(\lambda)$ is equal to the general rank r of $L(\lambda)$.

Proof. We use the same ideas as in the proofs of results in the previous sections. Therefore, the proof of Theorem 6.1 will be presented with less detail.

Clearly, a factorization (6.3) is impossible if $m < r$. Therefore we have to prove only that such a factorization exists for $m = r$. We can (and will) assume that in fact $n = r$, i.e., $\det L(\lambda) \not\equiv 0$.

Apply Theorem 2.2 with $F = \mathbb{C}$, $\sigma =$ identity, $\varepsilon = -1$. Since the only irreducible monic complex polynomial f satisfying $f = \varepsilon^{\text{degree}\, f} f_*$ is $f(\lambda) = \lambda$, by Theorem 2.2 we can assume (replacing $L(\lambda)$ by $D(\lambda)$) that the elementary divisors of $L(\lambda)$ are $\lambda, \lambda, \dots, \lambda$ (k times). Here k is necessarily even. Indeed, the property $L(\lambda) = (L(-\lambda))^T$ ensures that $\det L(\lambda) = (\text{const}\,.)\lambda^k$ is an even function.

If $k = 0$, i.e. $\det L(\lambda) \equiv \text{const}\,. \neq 0$, and application of Theorem 2.1 gives the desired result. Suppose therefore that $k > 0$. Using the Smith form of $L(\lambda)$ write

$$L(\lambda) = E(\lambda) \begin{bmatrix} \lambda I_k & 0 \\ 0 & I_{n-k} \end{bmatrix} F(\lambda),$$

where $E(\lambda)$ and $F(\lambda)$ are $n \times n$ matrix polynomials with constant non-zero determinants. Replacing $L(\lambda)$ by $(F(-\lambda))^{-T} L(\lambda) F(\lambda)^{-1}$ we can assume that the first k columns (and, by symmetry, also the first k rows) of $L(\lambda)$ are divisible by λ. Thus:

$$L(\lambda) = \begin{bmatrix} \lambda L_1(\lambda) & \lambda L_2(\lambda) \\ -\lambda (L_2(-\lambda))^T & L_3(\lambda) \end{bmatrix}$$

where the matrix polynomials L_1, L_2 and L_3 are $k \times k$, $k \times (n-k)$ and $(n-k) \times (n-k)$, respectively. Moreover, $-L_1(-\lambda) = (L_1(\lambda))^T$ and $L_3(-\lambda) = (L_3(\lambda))^T$. We claim that $L_1(0)$ is invertible. Indeed, if $L_1(0)$ were not invertible, then $\det \begin{bmatrix} L_1(\lambda) & L_2(\lambda) \\ -\lambda(L_2(-\lambda))^T & L_3(\lambda) \end{bmatrix}$ would be divisible by λ, and consequently $\det L(\lambda) = \lambda^k \det \begin{bmatrix} L_1(\lambda) & L_2(\lambda) \\ -\lambda L_2(-\lambda))^T & L_3(\lambda) \end{bmatrix}$ would be divisible by λ^{k+1}, an impossibility. Now $L_1(0)$ is skew-symmetric, and therefore admits a factorization

$$L_1(0) = Q^T \left(\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right) Q$$

for some invertible matrix Q. Let

$$M_1(\lambda) = \left(\begin{bmatrix} -\lambda & 0 \\ 0 & 1 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} -\lambda & 0 \\ 0 & 1 \end{bmatrix} \oplus I_{n-k}^* \right) \begin{bmatrix} Q & 0 \\ 0 & I_{n-k} \end{bmatrix} \tag{6.4}$$

$\left(\text{the summand } \begin{bmatrix} -\lambda & 0 \\ 0 & 1 \end{bmatrix} \text{ is repeated } \frac{k}{2} \text{ times}\right)$. Then

$$L(\lambda) = (M_1(-\lambda))^T \widetilde{L}(\lambda) M_1(\lambda)$$

for some matrix polynomial $\widetilde{L}(\lambda)$ such that $\widetilde{L}(-\lambda) = (\widetilde{L}(\lambda))^T$ and $\det \widetilde{L}(\lambda) \equiv \text{const}\,. \neq 0$. The only thing not immediate here is the claim that $\widetilde{L}(\lambda)$ is indeed a polynomial. But the only point in $\mathbb{C}$ where $\widetilde{L}(\lambda)$ could conceivably have a pole is $\lambda_0 = 0$. We have

$$Z(\lambda) := (M_1(-\lambda))^{-T} L(\lambda) M_1(\lambda)^{-1} = \left(\begin{bmatrix} \lambda^{-1} & 0 \\ 0 & 1 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} \lambda^{-1} & 0 \\ 0 & 1 \end{bmatrix} \oplus I_{n-k} \right) \begin{bmatrix} Q^{-T} & 0 \\ 0 & I_{n-k} \end{bmatrix}$$

$$\cdot \begin{bmatrix} \lambda L_1(\lambda) & \lambda L_2(\lambda) \\ -\lambda L_2(-\lambda))^T & L_3(\lambda) \end{bmatrix} \begin{bmatrix} Q^{-1} & 0 \\ 0 & I_{n-k} \end{bmatrix} \left(\begin{bmatrix} -\lambda^{-1} & 0 \\ 0 & 1 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} -\lambda^{-1} & 0 \\ 0 & 1 \end{bmatrix} \oplus I_{n-k} \right).$$

Clearly, $\lambda Z(\lambda)$ is analytic at $\lambda_0 = 0$, and the coefficient of λ^{-1} in the Laurent series of $Z(\lambda)$ in a neighborhood of $\lambda_0 = 0$ is

$$\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \oplus 0\right) \begin{bmatrix} Q^{-T} & 0 \\ 0 & I_{n-k} \end{bmatrix} \begin{bmatrix} L_1(0) & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} Q^{-1} & 0 \\ 0 & I_{n-k} \end{bmatrix} \cdot$$
$$\cdot \left(\begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} \oplus 0\right) =$$
$$= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} \oplus 0 = 0 \, .$$

To finish the proof, it remains to apply Theorem 2.1 to $Z(\lambda)$. □

Finally, we consider matrix polynomials $L(\lambda)$ having symmetry (6.1) with $\varepsilon = 1$.

THEOREM 6.2. *Let $L(\lambda) = L(\lambda)^T$ be an $n \times n$ matrix polynomial over $\mathbb{C}$, and let r be the general rank of $L(\lambda)$. Then $L(\lambda)$ admits a factorization*

$$L(\lambda) = (M(\lambda))^T M(\lambda) \tag{6.5}$$

for some $m \times n$ matrix polynomial $M(\lambda)$ if and only if $m \geq r$ (resp. $m \geq r+1$) in case the product of invariant polynomials of $L(\lambda)$ is a square of some complex polynomial (resp. is not a square of any complex polynomial).

Proof. Again, we omit many details here. We will assume that $r = n$. If $L(\lambda)$ admits factorization (6.5) with $M(\lambda)$ $n \times n$, then $\det L(\lambda) = (\det M(\lambda))^2$, and so the product of invariant polynomials of $L(\lambda)$ must be a square as well. This implies the "only if" part.

To prove the "if" part, first of all observe that it suffices to consider only the case when $\det L(\lambda)$ is the square of a polynomial (if it is not, replace $L(\lambda)$ by $\begin{bmatrix} L(\lambda) & 0 \\ 0 & f(\lambda) \end{bmatrix}$, where $f(\lambda)$ is a judiciously chosen scalar polynomial so that $f(\lambda) \det L(\lambda)$ is a square). By Theorem 2.2 we may assume that all elementary divisors of $L(\lambda)$ are first degree polynomials. Since $\det L(\lambda)$ is a square, the number $k = k(a)$ of elementary divisors $\lambda - a, \ldots, \lambda - a$ of $L(\lambda)$ (where $a \in \mathbb{C}$ is fixed) is even. As in the proof of Theorem 6.1, we can further assume that

$$L(\lambda) = \begin{bmatrix} (\lambda - a)L_1(\lambda) & (\lambda - a)L_2(\lambda) \\ (\lambda - a)(L_2(\lambda))^T & L_3(\lambda) \end{bmatrix}$$

for some matrix polynomials $L_1(\lambda) = L_1(\lambda)^T$, $L_2(\lambda)$ and $L_3(\lambda) = L_3(\lambda)^T$ of sizes $k \times k$, $k \times (n-k)$ and $(n-k) \times (n-k)$, respectively. Moreover, $L_1(a)$ is invertible and symmetric, and therefore

$$L_1(a) = Q^T \left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right) Q$$

for some invertible matrix q (the direct summand $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is repeated here $\frac{k}{2}$ times). As in the proof of Theorem 6.1, we verify that

$$L(\lambda) = (M_1(\lambda - a))^T \widetilde{L}(\lambda) M_1(\lambda - a)$$

where $M_1(\lambda)$ is defined by (6.4), and $\widetilde{L}(\lambda)$ is a matrix polynomial such that $\widetilde{L}(\lambda) = (\widetilde{L}(\lambda))^T$ and $\widetilde{L}(\lambda)$ has no elementary divisors of the form $\lambda - a$. Apply the above procedure to $\widetilde{L}(\lambda)$ in place of $L(\lambda)$, using elementary divisors $\lambda - b, \ldots, \lambda - b$ of $\widetilde{L}(\lambda)$ for some $b \in \mathbb{C}$, and so on, until a matrix polynomial $L_1(\lambda) = (L_1(\lambda))^T$ with $\det L_1(\lambda) \equiv \text{const}. \neq 0$ is obtained. Now apply Theorem 2.1 to get the desired factorization of $L - 1(\lambda)$. □

Theorems 6.1 and 6.2 can be recast in terms of elementary matrices (analogously to Theorem 3.2). An $n \times n$ matrix polynomial $M(\lambda)$ (over $\mathbb{C}$) is called ε-elementary if

$$M(\lambda) = (x(\varepsilon\lambda))^t x(\lambda)$$

for some $1 \times n$ row polynomial $x(\lambda) \neq 0$ (here $\varepsilon = \pm 1$ is fixed).

THEOREM 6.3. *Let $L(\lambda)$ be an $n \times n$ matrix polynomial satisfying (6.1), and let r be the general rank of $L(\lambda)$. Then $L(\lambda)$ can be written as sum of r ε-elementary matrices, unless $\varepsilon = 1$ and the product of elementary divisors of $L(\lambda)$ is not a square of any polynomial. In this latter case $L(\lambda)$ can be written as sum of $r + 1$ 1-elementary matrices, and cannot be represented as sum of any r 1-elementary matrices.*

REFERENCES

[AM] B.D.O. Anderson and J.B. Moore, *Optimal Filtering*, Prentice Hall, Englewood Cliffs, 1979.

[Co] W.A. Coppel, *Linear Systems*, Notes in Pure Mathematics 6 (1972), Australian National University, Canberra.

[GLR1] I. Gohberg, P. Lancaster and L. Rodman, *Spectral analysis of selfadjoint matrix polynomials*, Annals of Math. 112 (1980), 33–71.

[GLR2] I. Gohberg, P. Lancaster and L. Rodman, *Factorization of selfadjoint matrix polynomials with constant signature*, Linear and Multilinear Algebra 11 (1982), 209–224.

[GLR3] I. Gohberg, P. Lancaster and L. Rodman, *Matrix Polynomials*, Academic Press, New York, etc., 1982.

[HJ] R.A. Horn, C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, 1990.

[Ja] V.A. Jakubovič, *Factorization of symmetric matrix polynomials*, Soviet Math. Doklady 11 (1970) 1261–1264.

[Lyu1] B.D. Lyubachevskii, *Factorization of symmetric matrices with elements from a ring with involution I*, Siberian Math. Journal 14 (1973), 233–246.

[Lyu2] B.D. Lyubachevskii, *Factorization of symmetric matrices with elements from a ring with involution II*, Siberian Math. Journal 14 (1973), 423–433.

[M] C.C. Mac Duffee, *The Theory of Matrices*, Chelsea Publ. Company, New York, 1946.

[R] R. Rellich, *Perturbation Theory for Eigenvalue Problems*, Gordon & Breach, New York, 1969.