

Triangularizing Quadratic Matrix Polynomials

Tisseur, Françoise and Zaballa, Ion

2012

MIMS EPrint: **2012.30**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

Reports available from: <http://eprints.maths.manchester.ac.uk/>

And by contacting: The MIMS Secretary
School of Mathematics
The University of Manchester
Manchester, M13 9PL, UK

ISSN 1749-9097

TRIANGULARIZING QUADRATIC MATRIX POLYNOMIALS*

FRANÇOISE TISSEUR[†] AND ION ZABALLA[‡]

Abstract. We show that any regular quadratic matrix polynomial can be reduced to an upper triangular quadratic matrix polynomial over the complex numbers preserving the finite and infinite elementary divisors. We characterize the real quadratic matrix polynomials that are triangularizable over the real numbers and show that those that are not triangularizable over the real numbers are quasi-triangularizable with diagonal blocks of sizes 1×1 and 2×2 . We also derive complex and real Schur-like theorems for linearizations of quadratic matrix polynomials with nonsingular leading coefficients. In particular, we show that for any monic linearization $\lambda I + A$ of an $n \times n$ quadratic matrix polynomial, there exists a nonsingular matrix defined in terms of n orthonormal vectors that transforms A to a companion linearization of a (quasi)-triangular quadratic matrix polynomial. This provides the foundation for designing numerical algorithms for the reduction of quadratic matrix polynomials to upper (quasi)-triangular form.

Key words. triangularization, triangular, quasi-triangular, companion linearization, equivalence, quadratic eigenvalue problem, Schur theorem

AMS subject classifications. 15A18, 15A22, 65F15

1. Introduction. There is no analogue of the generalized Schur decomposition for quadratic matrix polynomials in the sense that $Q(\lambda) = A_2\lambda^2 + A_1\lambda + A_0$ cannot, in general, be reduced to triangular form $T(\lambda) = T_2\lambda^2 + T_1\lambda + T_0$ by unitary or strict equivalence transformations. For eigenvalue and frequency response computations, it is common practice to convert $Q(\lambda)$ to a linear polynomial $L(\lambda) = A\lambda + B$ of twice the dimension, a process called linearization, and then triangularize $L(\lambda)$ in place of $Q(\lambda)$ (see the survey paper [12]).

It turns out that if unimodular transformations are used in place of strict equivalence transformations, any monic quadratic matrix polynomial is equivalent to a triangular monic quadratic matrix polynomial over the complex numbers. In other words, there exist matrix polynomials $U(\lambda)$ and $V(\lambda)$ with nonzero constant determinant such that

$$U(\lambda)Q(\lambda)V(\lambda) = T(\lambda),$$

where $T(\lambda)$ is a monic triangular quadratic and $Q(\lambda)$ and $T(\lambda)$ have the same eigenstructure. We say that $Q(\lambda)$ is *triangularizable*. This apparently little-known fact appears in the proof of a theorem on the inverse problem for linearization [9, Thm. 1.7]. Using the notion of strong equivalence so as to preserve the elementary divisors at infinity, we prove in Section 3 that *any* regular (i.e., $\det Q(\lambda) \not\equiv 0$) not necessarily monic quadratic matrix polynomial is triangularizable over the complex numbers. The proof is constructive.

Over the reals however, not all quadratic matrix polynomials are triangularizable. We characterize those that are triangularizable over the real numbers and show in

*Version of February 24, 2012.

[†]School of Mathematics, The University of Manchester, Manchester, M13 9PL, UK (ftisseur@ma.man.ac.uk, <http://www.ma.man.ac.uk/~ftisseur/>). This author was supported by Engineering and Physical Sciences Research Council grant EP/I005293 and a Fellowship from the Leverhulme Trust.

[‡]Departamento de Matemática Aplicada y EIO, Euskal Herriko Unibertsitatea (UPV/EHU), Apdo. Correos 644, 48080 Bilbao, Spain (ion.zaballa@ehu.es). This author was supported by the Dirección General de Investigación, Proyecto de Investigación MTM2010-19356-C02-01 and Gobierno Vasco, GIC10/IT-361-10 and UPV/EHU UFI11/52.

Section 4 that the real quadratic matrix polynomials that are not triangularizable over the real numbers are quasi-triangularizable with diagonal blocks of size 2×2 and 1×1 . In this case, a characterization of the elementary divisors of the 2×2 diagonal blocks as well as the minimal number of them is provided.

Finally, we investigate in Section 5 how the subspaces version and matrix version of Schur's theorem extends to any monic linearization $\lambda I + A$ of a quadratic matrix polynomial $Q(\lambda)$ with nonsingular leading coefficient. In particular, we show that there is a $2n \times n$ matrix U with orthonormal columns such that $S = [U \ AU]$ is nonsingular and $S^{-1}AS = \begin{bmatrix} 0 & -T_0 \\ I & -T_1 \end{bmatrix}$ is a companion linearization of a triangular monic quadratic matrix polynomial $T(\lambda) = I_n\lambda^2 + T_1\lambda + T_0$ equivalent to $Q(\lambda)$. This result is useful for the design of algorithms that reduce $Q(\lambda)$ to triangular form [10]. Based on a geometric interpretation of the real Schur form of any real square matrix, a Schur-like theorem for real quadratic matrix polynomials with nonsingular leading coefficient is also exhibited.

2. Preliminaries. The eigenstructure of a matrix polynomial comprises the eigenvalues and their partial multiplicities or, equivalently, the elementary divisors (or invariant factors) of the matrix polynomial, including those at infinity. Two matrix polynomials with the same finite elementary divisors are said to be *equivalent*. The equivalence of matrix polynomials can be characterized by the action of the unimodular group as follows. As mentioned in the introduction, a matrix polynomial is called *unimodular* if it is square and its determinant is a nonzero constant polynomial. These matrices form a non-commutative group (the group of units of the ring of square matrices polynomials of fixed size) and two matrix polynomials $A(\lambda)$ and $B(\lambda)$ are equivalent if and only if there are unimodular matrices $U(\lambda)$ and $V(\lambda)$ such that $B(\lambda) = U(\lambda)A(\lambda)V(\lambda)$. If, in addition, $A(\lambda)$ and $B(\lambda)$ have the same elementary divisors at infinity they are called *strongly equivalent*. We say that a regular quadratic matrix polynomial $Q(\lambda)$ with nonsingular leading coefficient A_2 is *triangularizable (diagonalizable)* if it is equivalent to a triangular (diagonal, respectively) monic quadratic matrix polynomial. When $\det A_2 = 0$, $Q(\lambda)$ is *triangularizable (diagonalizable)* if it is *strongly equivalent* to a triangular (diagonal, respectively) quadratic matrix polynomial whose diagonal entries are monic and of degree at most 2: their roots are the finite eigenvalues of $Q(\lambda)$.

Any $m \times n$ matrix polynomial $A(\lambda)$ with coefficients in an arbitrary field \mathbb{F} , is equivalent to a diagonal matrix polynomial (of different degree, in general) called the *Smith form* of $A(\lambda)$. That is to say, there are unimodular matrices $U(\lambda) \in \mathbb{F}[\lambda]^{m \times m}$, and $V(\lambda) \in \mathbb{F}[\lambda]^{n \times n}$, such that

$$U(\lambda)A(\lambda)V(\lambda) = D(\lambda) = \begin{bmatrix} \text{diag}(\alpha_1(\lambda), \dots, \alpha_r(\lambda)) & 0 \\ 0 & 0 \end{bmatrix},$$

where $r = \text{rank } A(\lambda)$ and $\alpha_1(\lambda) | \dots | \alpha_r(\lambda)$ are monic polynomials. Here, “|” stands for divisibility, so that $\alpha_j(\lambda)$ is divisible by $\alpha_{j-1}(\lambda)$. These polynomials are the *invariant factors* of $A(\lambda)$ and are uniquely determined by $A(\lambda)$.

In what follows, \mathbb{F} denotes either the field of complex numbers \mathbb{C} or the field of real numbers \mathbb{R} . Since $Q(\lambda) \in \mathbb{F}[\lambda]^{n \times n}$ is regular, $\text{rank } Q(\lambda) = n$ and its Smith form is $\text{diag}(\alpha_1(\lambda), \dots, \alpha_n(\lambda))$ with $\alpha_1(\lambda) | \dots | \alpha_n(\lambda)$. The invariant factors of $Q(\lambda)$ can be

decomposed into irreducible factors over \mathbb{F} as follows [7, Chap. VI, §3]:

$$\begin{aligned}\alpha_n(\lambda) &= \phi_1(\lambda)^{m_{11}} \cdots \phi_s(\lambda)^{m_{s1}}, \\ \alpha_{n-1}(\lambda) &= \phi_1(\lambda)^{m_{12}} \cdots \phi_s(\lambda)^{m_{s2}}, \\ &\vdots \\ \alpha_1(\lambda) &= \phi_1(\lambda)^{m_{1n}} \cdots \phi_s(\lambda)^{m_{sn}},\end{aligned}\tag{2.1}$$

where $\phi_i(\lambda)$, $i = 1 : s$ are distinct monic polynomials irreducible over $\mathbb{F}[\lambda]$, and

$$m_{i1} \geq m_{i2} \geq \cdots \geq m_{in} \geq 0, \quad i = 1 : s.\tag{2.2}$$

The factors $\phi_i(\lambda)^{m_{ij}}$ with $m_{ij} > 0$ are the *elementary divisors* of $Q(\lambda)$. Notice that when $\mathbb{F} = \mathbb{C}$, $\phi_i(\lambda) = \lambda - \lambda_i$ and when $\mathbb{F} = \mathbb{R}$, $\phi_i(\lambda) = \lambda - \lambda_i$ if $\lambda_i \in \mathbb{R}$ and $\phi_i(\lambda) = \lambda^2 - 2\operatorname{Re}(\lambda_i)\lambda + |\lambda_i|^2$ for the complex eigenvalues $\lambda_i, \bar{\lambda}_i$. We will use in sections 3 and 4 the procedure in (2.1) to construct the invariant factors of a matrix polynomial from its elementary divisors.

The elementary divisors at infinity of $Q(\lambda)$ can be defined in several ways (see for example [6], [13]). The following definition is the most convenient for our development: the elementary divisors of $Q(\lambda)$ at infinity are those of $\operatorname{rev} Q(\lambda)$ at 0, where

$$\operatorname{rev} Q(\lambda) = \lambda^2 Q(\lambda^{-1}) = A_0 \lambda^2 + A_1 \lambda + A_2$$

is the reversal of $Q(\lambda)$. Notice that $\operatorname{rev} \operatorname{rev} Q(\lambda) = Q(\lambda)$.

We also need the following result, which is a particular case of [4, Thm. 5.2] and [9, Thm. 1.7]. Here $\deg(\cdot)$ denotes “the degree of”.

LEMMA 2.1. *Let \mathbb{F} be an arbitrary field and let $\alpha_1(\lambda)|\alpha_2(\lambda)|\cdots|\alpha_n(\lambda)$ be monic polynomials with coefficients in \mathbb{F} . There exists a quadratic $Q(\lambda) \in \mathbb{F}[\lambda]^{n \times n}$ with nonsingular leading coefficient and $\alpha_1(\lambda)|\alpha_2(\lambda)|\cdots|\alpha_n(\lambda)$ as invariant factors if and only if $\sum_{i=1}^n \deg(\alpha_i(\lambda)) = 2n$.*

Note that Lemma 2.1 is also a consequence of Rosenbrock’s theorem on pole placement (see [1, Thm. 1.1], [11]). The next technical result will be used in Section 4.

PROPOSITION 2.2. *Let $T(\lambda) = (t_{ij}(\lambda)) \in \mathbb{F}[\lambda]^{n \times n}$ be an upper triangular regular matrix polynomial. There are upper triangular unimodular matrices $U(\lambda)$ and $V(\lambda)$ with unit diagonal entries such that if $\tilde{T}(\lambda) = U(\lambda)T(\lambda)V(\lambda) = (\tilde{t}_{ij}(\lambda))$ then $t_{ii}(\lambda) = \tilde{t}_{ii}(\lambda)$ for $i = 1 : n$ and $\deg(\tilde{t}_{ij}(\lambda)) < \deg(\gcd(t_{ii}(\lambda), t_{jj}(\lambda)))$ for $1 \leq i < j \leq n$.*

Proof. We describe a procedure to obtain $\tilde{T}(\lambda)$. Let $j = 2$, $i = 1$.

step 1 Compute $d_{ij}(\lambda) = \gcd(t_{ii}(\lambda), t_{jj}(\lambda))$.

step 2 Compute the Euclidean division $t_{ij}(\lambda) = q_{ij}(\lambda)d_{ij}(\lambda) + \tilde{t}_{ij}(\lambda)$.

step 3 Compute polynomials $x_{ij}(\lambda)$, $y_{ij}(\lambda)$ such that $x_{ij}(\lambda)t_{ii}(\lambda) + y_{ij}(\lambda)t_{jj}(\lambda) = -q_{ij}(\lambda)d_{ij}(\lambda)$.

step 4 Add column i multiplied by $x_{ij}(\lambda)$ to column j and row j multiplied by $y_{ij}(\lambda)$ to row i .

step 5 If $i > 1$ then $i = i - 1$ else $j = j + 1$, $i = j - 1$. If $j \leq n$ go to step 1 else stop.

□

3. Triangularizable quadratic matrix polynomials. As mentioned in the introduction, the proof of [9, Thm. 1.7] shows that any complex matrix polynomial of degree ℓ with nonsingular leading coefficient is equivalent to a monic triangular matrix polynomial of the same degree. We recall the procedure for $\ell = 2$.

step 1 Compute the invariant factors $\alpha_1(\lambda)|\alpha_2(\lambda)|\cdots|\alpha_n(\lambda)$ of the given matrix polynomial and let $D(\lambda) = \operatorname{diag}(\alpha_1(\lambda), \dots, \alpha_n(\lambda))$.

We only have to multiply rows 3 and 4 by -1 to produce a monic quadratic triangular matrix equivalent to $Q(\lambda)$. \square

For a given $Q(\lambda)$, the above construction produces different triangular matrices depending on the choice of polynomials $s(\lambda)$ and $q(\lambda)$. A natural question is how much freedom one may expect to have in choosing the diagonal elements in the triangularization process. An answer to this question is provided by the following result.

THEOREM 3.2 ([5]). *Let $\alpha_1(\lambda) \cdots \alpha_n(\lambda)$ and $\delta_1(\lambda), \dots, \delta_n(\lambda)$ be monic polynomials with coefficients in an arbitrary field \mathbb{F} . Then there exists a triangular matrix polynomial in $\mathbb{F}[\lambda]^{n \times n}$ with diagonal entries $\delta_1(\lambda), \dots, \delta_n(\lambda)$ and $\alpha_1(\lambda), \dots, \alpha_n(\lambda)$ as invariant factors if and only if $\prod_{j=1}^n \alpha_j(\lambda) = \prod_{j=1}^n \delta_j(\lambda)$ and*

$$\prod_{j=1}^k \alpha_j(\lambda) \mid \gcd \left\{ \prod_{j=1}^k \delta_{i_j}(\lambda), 1 \leq i_1 < \dots < i_k \leq n \right\}, \quad 1 \leq k \leq n-1.$$

In Example 3.1, the condition $\alpha_1(\lambda)\alpha_2(\lambda) \mid \gcd(\delta_1(\lambda)\delta_2(\lambda), \delta_1(\lambda)\delta_3(\lambda), \delta_2(\lambda)\delta_3(\lambda))$ imposes strong restrictions on the choice of the diagonal elements. Namely, one of them must be $\lambda^2 + 1$, another $(\lambda - 1)(\lambda - i)$ and the third one must be $(\lambda - 1)(\lambda + i)$. Provided that these are the three diagonal elements, they may be placed in any diagonal position.

The Gohberg, Lancaster and Rodman triangularization procedure cannot be applied when $Q(\lambda)$ has a singular leading coefficient. We will show that any regular quadratic matrix polynomial is strongly equivalent over $\mathbb{C}[\lambda]$ to a triangular quadratic matrix polynomial with monic diagonal polynomials, and that real quadratics with finite eigenvalues all real are triangularizable over $\mathbb{R}[\lambda]$.

THEOREM 3.3. *Let $Q(\lambda) \in \mathbb{F}[\lambda]^{n \times n}$ be a regular quadratic matrix polynomial. If $\det Q(\lambda)$ admits a prime factorization in linear factors over $\mathbb{F}[\lambda]$ then $Q(\lambda)$ is triangularizable.*

Proof. We have seen that when $Q(\lambda)$ has nonsingular leading coefficient it is triangularizable. In particular, if $\mathbb{F} = \mathbb{R}$ and $\det Q(\lambda)$ has a prime factorization in linear factors over \mathbb{R} then Gohberg, Lancaster and Rodman's procedure produces a real triangular quadratic matrix polynomial.

We now consider the case where the leading coefficient of $Q(\lambda)$ is singular, i.e., $\det(A_2) = 0$. If $\det(A_0) \neq 0$ then $\text{rev } Q(\lambda) = \lambda^2 A_0 + \lambda A_1 + A_2$, has nonsingular leading coefficient and the elementary divisors of $Q(\lambda)$ and $\text{rev } Q(\lambda)$ are related as follows: $(\lambda - \lambda_0)^{m_0}$ is an elementary divisor of $Q(\lambda)$ if and only if $(\lambda - 1/\lambda_0)^{m_0}$ is an elementary divisor of $\text{rev } Q(\lambda)$. In particular, the elementary divisors at infinity of $Q(\lambda)$ are the elementary divisors at 0 of $\text{rev } Q(\lambda)$. Hence, if $\det Q(\lambda)$ admits a prime factorization in linear factors over $\mathbb{F}[\lambda]$ so does $\text{rev } Q(\lambda)$. Hence $\text{rev } Q(\lambda)$ is equivalent to an upper triangular quadratic matrix polynomial $\text{rev } T(\lambda)$ whose reversal $T(\lambda)$ is equivalent to $Q(\lambda)$. The quadratics $Q(\lambda)$ and $T(\lambda)$ have the same elementary divisors (finite and at infinity), i.e., they are strongly equivalent.

If A_2 and A_0 are both singular then since $Q(\lambda)$ is regular, there exists $\sigma \in \mathbb{F}$ such that $Q(\sigma)$ is nonsingular. Let $\mu = \lambda - \sigma$ and define

$$Q_\sigma(\mu) := Q(\lambda) = Q(\mu + \sigma) = A_2\mu^2 + Q'(\sigma)\mu + Q(\sigma)$$

whose trailing coefficient is nonsingular. If $\det Q(\lambda)$ admits a prime factorization in linear factors over $\mathbb{F}[\lambda]$ so does $\det Q_\sigma(\mu)$ and the elementary divisors of $Q_\sigma(\mu)$ are those of $Q(\lambda)$ shifted by σ . In particular, both matrices have the same elementary

divisors at infinity. Hence, as shown above, $Q_\sigma(\mu)$ is strongly equivalent to an upper triangular quadratic matrix polynomial $T_\sigma(\mu)$, where

$$T_\sigma(\mu) = T_{\sigma 2}\mu^2 + T_{\sigma 1}\mu + T_{\sigma 0} = T_2\lambda^2 + T_1\lambda + T_0 =: T(\lambda)$$

with $T_2 = T_{\sigma 2}$, $T_1 = T_{\sigma 1} - 2\sigma T_{\sigma 2}$, $T_0 = T_{\sigma 0} + \sigma^2 T_2 - \sigma T_1$ upper triangular.

Finally, in order to produce monic polynomials in the diagonal we proceed as follows: let a_i be the leading coefficient of the polynomial in position (i, i) of $T(\lambda)$. Put $D = \text{diag}(a_1, \dots, a_n)$. Then $D^{-1}T(\lambda)$ has diagonal monic polynomials and the same elementary divisors, finite and at infinity, as $T(\lambda)$. \square

The following result is a direct consequence of Theorem 3.3.

COROLLARY 3.4. *Let $Q(\lambda) \in \mathbb{F}[\lambda]^{n \times n}$ be a quadratic matrix polynomial with nonsingular leading coefficient. If $\det Q(\lambda)$ admits a prime factorization in linear factors over $\mathbb{F}[\lambda]$ then $Q(\lambda)$ is equivalent to an upper triangular quadratic matrix polynomial whose strictly upper triangular part is linear in λ at most.*

Proof. By Theorem 3.3, $Q(\lambda)$ is equivalent to a triangular quadratic matrix polynomial $T(\lambda) = \lambda^2 T_2 + \lambda T_1 + T_0$. We can use Proposition 2.2 or observe that T_2 is nonsingular because $Q(\lambda)$ has nonsingular leading coefficient. Hence $Q(\lambda)$ is equivalent to the monic triangular quadratic $T_2^{-1}T(\lambda) = \lambda^2 I + \lambda T_2^{-1}T_1 + T_2^{-1}T_0$. \square

We have shown in Example 3.1 how to reduce a quadratic matrix polynomial with nonsingular leading coefficient to triangular form. Such quadratics have no elementary divisors at infinity. In the following example we consider the opposite situation: the given matrix only has elementary divisors at infinity, i.e., it is unimodular.

EXAMPLE 3.5. Consider the following quadratic matrix polynomial

$$Q(\lambda) = \begin{bmatrix} \lambda^2 + 1 & \lambda & \lambda^2 - \lambda \\ \lambda^2 + \lambda & \lambda + 1 & \lambda^2 \\ \lambda^2 & \lambda & \lambda^2 - \lambda + 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \lambda^2 + \begin{bmatrix} 0 & 1 & -1 \\ 1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \lambda + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Since $\det Q(\lambda) = 1$, $Q(\lambda)$ is unimodular. The invariant factors of $\text{rev } Q(\lambda)$ are $\alpha_1(\lambda) = 1$, $\alpha_2(\lambda) = \lambda^2$ and $\alpha_3(\lambda) = \lambda^4$. This means that the exponents of the elementary divisors of $Q(\lambda)$ at infinity are 2 and 4. We now use the Gohberg, Lancaster, Rodman procedure to construct a triangular quadratic matrix polynomial $\text{rev } T(\lambda) = \lambda^2 I_3 + e_1 e_3^T$ with $\alpha_j(\lambda)$, $j = 1: 3$ as invariant factors. Here e_j denotes the j th column of the 3×3 identity matrix. Then, its reversal

$$T(\lambda) = \begin{bmatrix} 1 & 0 & \lambda^2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is triangular and strongly equivalent to $Q(\lambda)$. \square

We mentioned that as a consequence of Theorem 3.2 the only possible diagonal elements of any triangular matrix polynomial equivalent to $Q(\lambda)$ in (3.1) are $\lambda^2 + 1$, $(\lambda - i)(\lambda - 1)$ and $(\lambda + i)(\lambda - 1)$. Hence $Q(\lambda)$ cannot be triangularizable over $\mathbb{R}[\lambda]$. In fact, the reduction process may fail for matrices over arbitrary fields due to the possible existence of irreducible polynomials of any degree. The real quadratic matrix polynomials that are triangularizable over $\mathbb{R}[\lambda]$ are characterized in the following theorem for which we need additional notation.

Let the elementary divisors of an $n \times n$ real quadratic matrix polynomial $Q(\lambda)$ be

$$\begin{aligned} (\lambda^2 + a_i\lambda + b_i)^{n_{ij}}, & \quad i = 1: s, \quad j = 1: t_i, \\ (\lambda - \lambda_i)^{m_{ij}}, & \quad i = 1: r, \quad j = 1: p_i, \\ \mu^{m_{0j}}, & \quad j = 1: p_0, \end{aligned} \quad (3.2)$$

where $a_i^2 - 4b_i < 0$, i.e., $\lambda^2 + a_i\lambda + b_i$ is irreducible over $\mathbb{R}[\lambda]$, and the $\mu^{m_{0j}}, j = 1: p_0$ are the elementary divisors at infinity. We assume that the partial multiplicities m_{ij}, n_{ij} are ordered as in (2.2). Denote

$$n_c = \sum_{i=1}^s \sum_{j=1}^{t_i} n_{ij}, \quad p = \max_{0 \leq i \leq r} p_i, \quad (3.3)$$

that is, n_c is the sum of the degrees of the elementary divisors that are powers of irreducible polynomials of degree 2 and p is the largest chain of elementary divisors corresponding to any and the same real eigenvalue or to the eigenvalue at infinity. In other words, this eigenvalue has geometric multiplicity p and there is no other real or infinite eigenvalue with larger geometric multiplicity. We are now ready to state the theorem.

THEOREM 3.6. *A quadratic matrix polynomial $Q(\lambda) \in \mathbb{R}[\lambda]^{n \times n}$ is triangularizable over $\mathbb{R}[\lambda]$ if and only if $p \leq n - n_c$.*

Proof. Assume first that the leading coefficient of $Q(\lambda)$ is nonsingular. This means that $p_0 = 0$ and $p = \max_{1 \leq i \leq r} p_i$.

(\Rightarrow) Suppose that $Q(\lambda)$ is upper triangular and that $q_{ii}(\lambda) = \lambda^2 + a_i\lambda + b_i$ and $q_{jj}(\lambda) = (\lambda - \lambda_{j_1})(\lambda - \lambda_{j_2})$ with $a_i^2 - 4b_i < 0$ and $\lambda_{j_1}, \lambda_{j_2} \in \mathbb{R}$ and these two real numbers may be equal. Thus $q_{ii}(\lambda)$ and $q_{jj}(\lambda)$ are coprime polynomials and if $i < j$ there exist polynomials $x_{ij}(\lambda)$ and $y_{ij}(\lambda)$ such that

$$x_{ij}(\lambda)q_{ii}(\lambda) + y_{ij}(\lambda)q_{jj}(\lambda) = q_{ij}(\lambda).$$

Subtract the j th column of $Q(\lambda)$ from its i th column multiplied by $x_{ij}(\lambda)$ and the i th row from the j th row multiplied by $y_{ij}(\lambda)$. Then we obtain a matrix equivalent to $Q(\lambda)$, which is upper triangular and the entry in position (i, j) is zero.

Similarly, if $j < i$ there exist polynomials $x_{ji}(\lambda)$ and $y_{ji}(\lambda)$ such that

$$x_{ji}(\lambda)q_{ii}(\lambda) + y_{ji}(\lambda)q_{jj}(\lambda) = q_{ji}(\lambda).$$

If we subtract the i th column of $Q(\lambda)$ from its j th column multiplied by $y_{ji}(\lambda)$ and the j th row from the i th row multiplied by $x_{ji}(\lambda)$, then we get an upper triangular matrix equivalent to $Q(\lambda)$ with zero in position (j, i) .

After a finite number of such elementary transformations (starting from the bottom of the matrix in order to preserve the already zeroed elements) $Q(\lambda)$ is transformed to an equivalent upper triangular matrix with the property that $q_{ij}(\lambda) = 0$ if either $q_{ii}(\lambda) = \lambda^2 + a_i\lambda + b_i$ and $q_{jj}(\lambda) = (\lambda - \lambda_{j_1})(\lambda - \lambda_{j_2})$ or $q_{ii}(\lambda) = (\lambda - \lambda_{i_1})(\lambda - \lambda_{i_2})$ and $q_{jj}(\lambda) = \lambda^2 + a_j\lambda + b_j$ with $a_i^2 - 4b_i < 0$ and $a_j^2 - 4b_j < 0$. Now, by permuting rows and columns we transform that matrix to block diagonal form

$$T(\lambda) = \begin{matrix} n_c & n-n_c \\ n_c & n-n_c \end{matrix} \begin{bmatrix} T_c(\lambda) & 0 \\ 0 & T_r(\lambda) \end{bmatrix},$$

7

where $T_c(\lambda)$ and $T_r(\lambda)$ are upper triangular and such that the elementary divisors of $T_c(\lambda)$ are of the form $(\lambda^2 + a_i\lambda + b_i)^{n_{ij}}$ with $a_i^2 - 4b_i < 0$ and the elementary divisors of $T_r(\lambda)$ are of the form $(\lambda - \lambda_i)^{m_{ij}}$ with $\lambda_i \in \mathbb{R}$. Since $T_r(\lambda)$ is $(n - n_c) \times (n - n_c)$, the largest chain of elementary divisors corresponding to the same real eigenvalue must not be bigger than $n - n_c$, i.e., $p \leq n - n_c$.

(\Leftarrow) Conversely, take all elementary divisors of $Q(\lambda)$ of the form $(\lambda^2 + a_i\lambda + b_i)^{n_{ij}}$ with $\lambda^2 + a_i\lambda + b_i$ irreducible over $\mathbb{R}[\lambda]$ and construct the upper triangular quadratic matrix polynomials

$$T_{ij}(\lambda) = (\lambda^2 + a_i\lambda + b_i)I_{n_{ij}} + N_{n_{ij}}, \quad 1 \leq j \leq t_i, \quad 1 \leq i \leq s,$$

whose elementary divisors are $(\lambda^2 + a_i\lambda + b_i)^{n_{ij}}$. Here $N_{n_{ij}}$ is an $n_{ij} \times n_{ij}$ Jordan block with eigenvalue zero. Using the notation $\bigoplus_{j=1}^s F_j(\lambda)$ to denote the direct sum of the matrix polynomials $F_1(\lambda), \dots, F_s(\lambda)$ we have that

$$T_c(\lambda) = \bigoplus_{i=1}^s \bigoplus_{j=1}^{t_i} T_{ij}(\lambda)$$

is an upper triangular quadratic matrix polynomial with elementary divisors $(\lambda^2 + a_i\lambda + b_i)^{n_{ij}}$, $1 \leq j \leq t_i$, $1 \leq i \leq s$ (see [7, Thm. 5 Chap. VI]).

Consider now the elementary divisors of $Q(\lambda)$ of the form $(\lambda - \lambda_i)^{m_{ij}}$ with $\lambda_i \in \mathbb{R}$. First of all

$$\sum_{i=1}^r \sum_{j=1}^{p_i} m_{ij} = 2(n - n_c) \geq 2p.$$

Put $n_r = n - n_c$ and consider the process of constructing the invariant factors of a polynomial matrix out of its elementary divisors as described in Section 2. Assuming that $m_{i1} \geq m_{i2} \geq \dots \geq m_{ip_i}$ and bearing in mind that p is the largest chain associated with any real eigenvalue and $p \leq n_r$, we have that

$$\begin{aligned} \gamma_{n_r}(\lambda) &= (\lambda - \lambda_1)^{m_{11}} \dots (\lambda - \lambda_r)^{m_{r1}}, \\ &\vdots \\ \gamma_{n_r-p+1}(\lambda) &= (\lambda - \lambda_1)^{m_{1p}} \dots (\lambda - \lambda_r)^{m_{rp}}, \\ \gamma_{n_r-p}(\lambda) &= 1, \\ &\vdots \\ \gamma_1(\lambda) &= 1, \end{aligned} \tag{3.4}$$

where for notational simplicity we may have included exponents $m_{ij} = 0$. Thus $\gamma_1(\lambda) \cdots \gamma_{n_r}(\lambda)$ are the invariant factors of any $n_r \times n_r$ matrix polynomial with $(\lambda - \lambda_i)^{m_{ij}}$ as elementary divisors. Taking into account that $\gamma_{n_r}(\lambda)$ factorizes into linear factors over $\mathbb{R}[\lambda]$, by Theorem 3.3 there exists an $n_r \times n_r$ upper triangular quadratic matrix polynomial $T_r(\lambda)$ whose elementary divisors are the $(\lambda - \lambda_i)^{m_{ij}}$. Then $T(\lambda) = T_c(\lambda) \oplus T_r(\lambda)$ has $(\lambda^2 + a_i\lambda + b_i)^{n_{ij}}$ and $(\lambda - \lambda_i)^{m_{ij}}$ as elementary divisors and is quadratic. Therefore $Q(\lambda)$ is triangularizable.

Finally, if the leading matrix coefficient of $Q(\lambda)$ is singular (i.e., $p_0 > 0$), and $Q(\lambda)$ is triangular then, as in the proof of Theorem 3.3, there is $\sigma \in \mathbb{R}$ such that $Q(\sigma) \neq 0$. Hence $\text{rev}_\sigma Q(\lambda) = Q\left(\frac{1}{\lambda - \sigma}\right)$ is triangular, it has nonsingular leading coefficient and

its elementary divisors are:

$$\begin{aligned} & \left((\lambda - \sigma)^2 + \frac{a_i}{b_i}(\lambda - \sigma) + \frac{1}{b_i} \right)^{n_{ij}}, & i = 1: s, j = 1: t_i, \\ & \left(\lambda - \frac{1}{\lambda_i - \sigma} \right)^{m_{ij}}, & i = 1: r, j = 1: p_i \\ & \lambda^{m_{0j}}, & j = 1: p_0 \end{aligned}$$

with $\left(\left(\frac{a_i}{b_i} \right)^2 - 4 \frac{1}{b_i} \right) < 0$ and $0 \neq \lambda_i - \sigma \in \mathbb{R}$. Therefore, $p \leq n - n_c$.

Conversely, If $p \leq n - n_c$ then $\text{rev}_\sigma Q(\lambda)$ is triangularizable and so is $Q(\lambda)$. \square

Theorem 3.6 says that if the real eigenvalues and eigenvalues at infinity of a real quadratic matrix polynomial have geometric multiplicity less or equal to $n - n_c$ then the quadratic is triangularizable over the real numbers. The majority of real quadratic eigenvalue problems in the NLEVP collection [2] are triangularizable over $\mathbb{R}[\lambda]$.

EXAMPLE 3.7. As already stated, the matrix $Q(\lambda)$ of Example 3.1 is not triangularizable over $\mathbb{R}[\lambda]$. Let us check that the condition of Theorem 3.6 is not satisfied. The elementary divisors of $Q(\lambda)$ are $\lambda^2 + 1$, $\lambda^2 + 1$, $\lambda - 1$ and $\lambda - 1$ so that $n_c = 2$ and $p = 2$. Hence $n - n_c = 1 < 2 = p$. \square

EXAMPLE 3.8. As an example of a quadratic matrix polynomial with structure at infinity consider

$$Q(\lambda) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \lambda^2 + \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \lambda + \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix},$$

with elementary divisors $\lambda^2 + 1$, $\lambda^2 + 1$, μ and μ . Then $n - n_c = 1 < 2 = p$ and $Q(\lambda)$ is not triangularizable over $\mathbb{R}[\lambda]$. \square

4. Quasi-triangularizable real quadratic matrix polynomials. A natural question, in analogy with the triangularization problem for real square matrices, is whether any quadratic matrix polynomial is equivalent to a quasi-triangular quadratic matrix polynomial with 2×2 and 1×1 diagonal blocks. We show now that the answer is in the affirmative. In order to do so we need the following technical lemma.

LEMMA 4.1. *Let an $n \times n$ real quadratic matrix polynomial with nonsingular leading coefficient have the elementary divisors list in (3.2) with*

$$p_1 \geq p_2 \geq \cdots \geq p_r, \quad t_1 \geq t_2 \geq \cdots \geq t_s. \quad (4.1)$$

Let p and n_c be defined as in (3.3). If $n > 2$ and $p > n - n_c$ then

- (i) $n_c > 0$, $p \geq 2$, and $p_i \leq p - 2$, $i = 2, \dots, r$,
- (ii) $m_{1p} = m_{1(p-1)} = 1$, and if $p = n$ and $n_{1t_1} \geq 2$ then $m_{1(p-2)} = 1$ as well,
- (iii) $\max\{p_2, t_1 - 1, t_2\} \leq \begin{cases} n - 3 & \text{if } n_{1t_1} \geq 2, \\ n - 2 & \text{otherwise.} \end{cases}$

Proof. (i) Recall that the largest chain of elementary divisors corresponding to the same real eigenvalue cannot exceed the size of the matrix polynomial, i.e., $p \leq n$. But $p > n - n_c$ so $n_c > n - p \geq 0$. Also, the assumption $p > n - n_c$ implies that

$$2p - 2 \geq 2n - 2n_c = \sum_{j=1}^{p_i} \sum_{i=1}^r m_{ij}.$$

Bearing in mind that $p_1 = p$ we get

$$p - 2 = 2p - 2 - p \geq \sum_{j=1}^{p_i} \sum_{i=1}^r m_{ij} - p = \sum_{j=1}^p (m_{1j} - 1) + \sum_{j=1}^{p_i} \sum_{i=2}^r m_{ij} =: \Phi. \quad (4.2)$$

Taking into account that $m_{ij} > 0$ we conclude that $\Phi \geq 0$, that is, $p \geq 2$, and $\Phi \geq \sum_{j=1}^{p_i} \sum_{i=2}^r m_{ij} \geq p_i$, $i = 2: r$. Thus property (i) follows at once.

(ii) It also follows from (4.2) that

$$p - 2 \geq \Phi \geq \sum_{j=1}^p (m_{1j} - 1) \geq \#\{j : m_{1j} \geq 2\},$$

where $\#$ stands for “number of elements of”. This shows that there are at least two elementary divisors equal to $\lambda - \lambda_1$, that is, $m_{1p} = m_{1(p-1)} = 1$.

If $p = n$ and $n_{1t_1} \geq 2$ there are at least three elementary divisors equal to $\lambda - \lambda_1$ because if there are only two then

$$2n \geq 2n_c + \sum_{j=1}^p m_{1j} \geq 2n_{1t_1} + 2(n - 2) + 2 \geq 4 + 2n - 4 + 2 = 2n + 2,$$

which is a contradiction.

(iii) By (i), $p_2 \leq n - 2$. If $n_{1t_1} \geq 2$ and $p = n$ then we show that $p_2 \neq n - 2$. Indeed, if $p_2 = n - 2$ then $2n \geq 2n_c + p + p_2 = 2n_c + 2n - 2$ so that $n_c = 1$ implying that $n_{1t_1} = 1$, which is a contradiction. If $n_{1t_1} \geq 2$ and $p \leq n - 1$ then by (i), $p_2 \leq p - 2 \leq n - 3$. So $p_2 \leq n - 3$ if $n_{1t_1} \geq 2$ and $p_2 \leq n - 2$ otherwise.

Next we show that $t_1 \leq n - 1$. If $t_1 = n$ then since by (i) $p \geq 2$ we have that

$$2n_c \geq \sum_{i=1}^{t_1} 2n_{1i} \geq 2t_1 = 2n \geq 2n_c + p \geq 2n_c + 2,$$

which is impossible. Hence $t_1 - 1 \leq n - 2$. Now if $n_{1t_1} \geq 2$ and $t_1 = n - 1$ then since $n_{11} \geq \dots \geq n_{1t_1} \geq 2$, we have that $n_c \geq 2t_1 = 2n - 2$. But $2n \geq 2n_c \geq 4n - 4$, that is, $n \leq 2$ contradicting $n > 2$. Thus when $n_{1t_1} \geq 2$, $t_1 - 1 \leq n - 3$.

Finally, $t_2 \leq t_1 \leq n - 1$ and $t_2 = t_1 = n - 1$ leads to a contradiction because, as $p \geq 2$ and $2n \geq 2n_c + p$, we find that $2n - 2 \geq 2n_c \geq 2(t_1 + t_2) = 4t_1 = 4n - 4$, and so $n \leq 1$ and we are assuming $n > 2$. Thus, if $t_1 = t_2$ then $t_1 = t_2 \leq n - 2$. Now suppose that $n_{1t_1} \geq 2$. If $t_2 = n - 2$ then $t_1 = t_2 = n - 2$ and this leads to a contradiction again. Indeed, since $n_{1j} \geq n_{1t_1} \geq 2$ and $n_{2j} \geq 1$ $j = 1, \dots, t_2$ we have that $n_c \geq 2(n - 2) + (n - 2) = 3n - 6$. But $2n > 2n_c$. Then $2n > 6n - 12$ and so $3 > n$ contradicting that $n > 2$. Hence when $n_{1t_1} \geq 2$, $t_2 \leq n - 3$ and this completes the proof. \square

The reduction of any real quadratic matrix polynomial $Q(\lambda)$ to quasi-triangular form is based on the following result.

LEMMA 4.2. *Let $Q(\lambda) \in \mathbb{R}[\lambda]^{n \times n}$ be quadratic with nonsingular leading coefficient and with the elementary divisors in (3.2), where $p_1 \geq p_2 \geq \dots \geq p_r$ and $t_1 \geq t_2 \geq \dots \geq t_s$. Assume that $p > n - n_c$, where p and n_c are defined in (3.3). Then $Q(\lambda)$ is equivalent to a real monic quadratic matrix polynomial of the form*

$$\begin{array}{cc} & \begin{array}{cc} 2 & n-2 \end{array} \\ \begin{array}{c} 2 \\ n-2 \end{array} & \left[\begin{array}{cc} Q_1(\lambda) & X(\lambda) \\ 0 & Q_2(\lambda) \end{array} \right], \end{array}$$

where

- (i) $Q_1(\lambda)$ has elementary divisors $\lambda - \lambda_1, \lambda - \lambda_1$ and $\lambda^2 + a_1\lambda + b_1$,
- (ii) $Q_2(\lambda)$ has elementary divisors

$$\begin{aligned}
& (\lambda - \lambda_1)^{m_{1j}}, & j = 1: p - 2, \\
& (\lambda - \lambda_i)^{m_{ij}}, & j = 1: p_i, \quad i = 2: r, \\
& (\lambda^2 + a_1\lambda + b_1)^{n_{1j}}, & j = 1: t_1 - 1 \text{ if } t_1 > 1, \\
& (\lambda^2 + a_1\lambda + b_1)^{n_{1t_1} - 1} & \text{if } n_{1t_1} > 1, \\
& (\lambda^2 + a_i\lambda + b_i)^{n_{ij}}, & j = 1: t_i, \quad i = 2: s.
\end{aligned} \tag{4.3}$$

Proof. The case $n = 1$ does not arise because if $n = 1$ then either $p = 0$ or $n_c = 0$, which cannot happen when $p > n - n_c$. Indeed, $n \geq p$ and $n \geq n_c$. Thus $p > n - n_c \geq 0$ and so $n_c > n - p \geq 0$.

If $n = 2$ then it follows from $p > 0$ and $n_c > 0$ that $p = 2$ and $n_c = 1$ so that the elementary divisors of $Q(\lambda)$ are $\lambda - \lambda_1, \lambda - \lambda_1$ and $\lambda^2 + a_1\lambda + b_1$. Hence, if A_2 is the nonsingular leading coefficient of $Q(\lambda)$, then $Q_1(\lambda) = A_2^{-1}Q(\lambda)$ is a matrix with the desired properties.

From now on we assume that $n > 2$. By Lemma 4.1 (ii), $Q(\lambda)$ has at least two elementary divisors equal to $\lambda - \lambda_1$. Let us define

$$S_1(\lambda) = \text{diag}(\lambda - \lambda_1, (\lambda - \lambda_1)(\lambda^2 + a_1\lambda + b_1)). \tag{4.4}$$

We split now the study into three possible cases according as (a) $n_{1t_1} = 1$, (b) $n_{1t_1} \geq 2$ and $p = n$, and (c) $n_{1t_1} \geq 2$ and $p < n$.

(a) $n_{1t_1} = 1$. In this case, we remove $\lambda - \lambda_1$ twice and $\lambda^2 + a_1\lambda + b_1$ from the list of elementary divisors of $Q(\lambda)$ so as to end up with the list of powers of prime polynomials displayed in (4.3). The sum of the degrees of all these polynomials is $2n - 4$. We now show that there is an $(n - 2) \times (n - 2)$ real quadratic matrix polynomial with these polynomials as elementary divisors. In order to apply Lemma 2.1 we use the procedure outlined in (3.4) to construct an invariant factors chain of length $n - 2$. For the procedure to go through, the maximal length of any chain of elementary divisors must be at most $n - 2$. In other words we must show that

$$\max\{p - 2, p_2, \dots, p_r, t_1 - 1, t_2, \dots, t_s\} \leq n - 2. \tag{4.5}$$

That (4.5) holds follows directly from (4.1) and Lemma 4.1 (iii). Hence we can use the procedure in (3.4) with the polynomials in (4.3) to construct an invariant factors chain $\nu_1(\lambda)|\nu_2(\lambda)|\dots|\nu_{n-2}(\lambda)$ of length $n - 2$. Let us define

$$S_2(\lambda) = \text{diag}(\nu_1(\lambda), \nu_2(\lambda), \dots, \nu_{n-2}(\lambda)).$$

Then $Q(\lambda)$ and $\text{diag}(S_1(\lambda), S_2(\lambda))$ are $n \times n$ matrix polynomials with the same elementary divisors, that is, they are equivalent. By Lemma 2.1, $S_1(\lambda)$ and $S_2(\lambda)$ are respectively equivalent to real quadratic matrix polynomials $Q_1(\lambda)$ and $Q_2(\lambda)$. Hence $Q(\lambda)$ is equivalent to $\text{diag}(Q_1(\lambda), Q_2(\lambda))$ and $Q_1(\lambda)$ and $Q_2(\lambda)$ have the desired elementary divisors.

(b) $n_{1t_1} \geq 2, p = n$. By Lemma 4.1 (ii) there are at least three elementary divisors equal to $\lambda - \lambda_1$. Hence, we can remove three copies of $\lambda - \lambda_1$ and $(\lambda^2 + a_1\lambda + b_1)^{n_{1t_1}}$ from the list of elementary divisors of $Q(\lambda)$ and construct

$$S_3(\lambda) = (\lambda - \lambda_1) \oplus \begin{bmatrix} (\lambda - \lambda_1)(\lambda^2 + a_1\lambda + b_1) & \lambda - \lambda_1 \\ 0 & (\lambda - \lambda_1)(\lambda^2 + a_1\lambda + b_1)^{n_{1t_1} - 1} \end{bmatrix}$$

whose elementary divisors are $\lambda - \lambda_1$, $\lambda - \lambda_1$, $\lambda - \lambda_1$ and $(\lambda^2 + a_1\lambda + b_1)^{n_{1t_1}}$. The remaining elementary divisors are

$$\begin{aligned} (\lambda - \lambda_1)^{m_{1j}}, & \quad j = 1: p - 3, \\ (\lambda - \lambda_i)^{m_{ij}}, & \quad j = 1: p_i, \quad i = 2: r, \\ (\lambda^2 + a_1\lambda + b_1)^{n_{1j}}, & \quad j = 1: t_1 - 1 \text{ if } t_1 > 1, \\ (\lambda^2 + a_i\lambda + b_i)^{n_{ij}}, & \quad j = 1: t_i, \quad i = 2: s. \end{aligned} \quad (4.6)$$

On using (4.1) and Lemma 4.1 (iii) we find that

$$\max\{p - 3, p_2, \dots, p_r, t_1 - 1, t_2, \dots, t_s\} \leq n - 3.$$

The procedure in (3.4) allows the construction of an invariant factors chain of length $n - 3$ out of the list of polynomials in (4.6). Let $\nu_1(\lambda)|\nu_2(\lambda)|\dots|\nu_{n-3}(\lambda)$ be such a chain and define

$$S_4(\lambda) = \text{diag}(\nu_1(\lambda), \nu_2(\lambda), \dots, \nu_{n-3}(\lambda)).$$

Hence $Q(\lambda)$ is equivalent to $\text{diag}(S_3(\lambda), S_4(\lambda))$. But

$$\text{diag}(S_3(\lambda), S_4(\lambda)) = \begin{bmatrix} S_1(\lambda) & X(\lambda) \\ 0 & S_5(\lambda) \end{bmatrix},$$

where $S_1(\lambda)$ is as in (4.4) and

$$X(\lambda) = \begin{bmatrix} 0 & 0 \\ (\lambda - \lambda_1) & 0 \end{bmatrix}, \quad S_5(\lambda) = \begin{bmatrix} (\lambda - \lambda_1)(\lambda^2 + a_1\lambda + b_1)^{n_{1t_1}-1} & 0 \\ 0 & S_4(\lambda) \end{bmatrix}.$$

Notice that the elementary divisors of $S_5(\lambda)$ are those in the list (4.6) plus $\lambda - \lambda_1$ and $(\lambda^2 + a_1\lambda + b_1)^{n_{1t_1}-1}$, that is, those of (4.3) when $n_{1t_1} > 1$.

Now, by Lemma 2.1 $S_1(\lambda)$ and $S_4(\lambda)$ are equivalent to a 2×2 and $(n-2) \times (n-2)$ real quadratic matrix polynomials $Q_1(\lambda)$ and $Q_2(\lambda)$, respectively. Therefore $Q(\lambda)$ is equivalent to

$$\begin{bmatrix} Q_1(\lambda) & Y(\lambda) \\ 0 & Q_2(\lambda) \end{bmatrix},$$

where $Y(\lambda)$ may have degree greater than 2. If that is the case and taking into account that necessarily the leading coefficient of $Q_1(\lambda)$ is nonsingular (otherwise the degree of its determinant would be smaller than 4) we can divide $Y(\lambda)$ by $Q_1(\lambda)$ on the left:

$$Y(\lambda) = Q_1(\lambda)W(\lambda) + R(\lambda), \quad \deg(R(\lambda)) \leq 1.$$

Then $Q(\lambda)$ is equivalent to

$$\begin{bmatrix} Q_1(\lambda) & Y(\lambda) \\ 0 & Q_2(\lambda) \end{bmatrix} \begin{bmatrix} I_2 & -W(\lambda) \\ 0 & I_{n-2} \end{bmatrix} = \begin{bmatrix} Q_1(\lambda) & R(\lambda) \\ 0 & Q_2(\lambda) \end{bmatrix},$$

which is a matrix with the desired properties.

(c) $n_{1t_1} \geq 2, p < n$. In this case we remove $\lambda - \lambda_1$ twice and $(\lambda^2 + a_1\lambda + b_1)^{n_{1t_1}}$ from the list of elementary divisors of $Q(\lambda)$ to obtain

$$\begin{aligned} (\lambda - \lambda_1)^{m_{1j}}, & \quad j = 1: p - 2, \\ (\lambda - \lambda_i)^{m_{ij}}, & \quad j = 1: p_i, \quad i = 2: r, \\ (\lambda^2 + a_1\lambda + b_1)^{n_{1j}}, & \quad j = 1: t_1 - 1, \\ (\lambda^2 + a_i\lambda + b_i)^{n_{ij}}, & \quad j = 1: t_i, \quad i = 2: s. \end{aligned} \quad (4.7)$$

divisors as $Q(\lambda)$. Thus $T_2(\lambda)$ vanishes, $T_1(\lambda) = Q_1(\lambda)$, $r = \max\{0, p + n_c - n\} = 1$ and the only linear elementary divisors of $T_1(\lambda)$ are $\lambda - \lambda_1$ twice.

Let us assume now that $n > 2$ and that the theorem holds for any real quadratic matrix polynomial of size at most $n - 1$. As in Lemma 4.2 we can assume without loss of generality that $p = p_1 \geq p_2 \geq \dots \geq p_r$ and $t_1 \geq t_2 \geq \dots \geq t_s$. By this lemma $Q(\lambda)$ is equivalent to a monic real quadratic block-triangular matrix polynomial

$$\tilde{Q}(\lambda) = \begin{bmatrix} Q_1(\lambda) & X(\lambda) \\ 0 & Q_2(\lambda) \end{bmatrix},$$

where $Q_1(\lambda)$ is of size 2×2 with elementary divisors $\lambda - \lambda_1$, $\lambda - \lambda_1$ and $\lambda^2 + a_1\lambda + b_1$, and $Q_2(\lambda)$ is of size $(n - 2) \times (n - 2)$ with elementary divisors listed in (4.3). By Lemma 4.2, the real elementary divisors of $\tilde{Q}(\lambda)$ are of those of $Q_1(\lambda)$ together with those of $Q_2(\lambda)$.

Let \tilde{n}_c be the sum of the exponents of the non real elementary divisors of $Q(\lambda)$ and \tilde{p} be the maximal length of the chains of real elementary divisors of $Q_2(\lambda)$. Then by Lemma 4.1 (i) and (4.3),

$$\tilde{n}_c = n_c - 1, \quad \tilde{p} = \tilde{p}_1 = p - 2.$$

Now if $\tilde{p} = (n - 2) - \tilde{n}_c$ then by Theorem 3.6, $Q_2(\lambda)$ is triangularizable over $\mathbb{R}[\lambda]$ and $Q(\lambda)$ is equivalent to

$$T(\lambda) = \begin{bmatrix} T_1(\lambda) & T_3(\lambda) \\ 0 & T_2(\lambda) \end{bmatrix},$$

where $T_1(\lambda) = Q_1(\lambda)$ and $T_2(\lambda)$ is triangular. In this case $r = p + n_c - n = 1$, $T_1(\lambda)$ is 2×2 and its invariant factors are $(\lambda - \lambda_1)|(\lambda - \lambda_1)(\lambda^2 + a_1\lambda + b_1)$. The real elementary divisors of $T(\lambda)$ are the real elementary divisors of $T_1(\lambda)$ and those of $T_2(\lambda)$. Hence $T(\lambda)$ satisfies all requirements.

If $\tilde{p} > (n - 2) - \tilde{n}_c$ then by the induction hypothesis $Q_2(\lambda)$ is equivalent to

$$\tilde{Q}_2(\lambda) = \begin{bmatrix} T_{21}(\lambda) & Y(\lambda) \\ 0 & T_{22}(\lambda) \end{bmatrix},$$

where $T_{22}(\lambda)$ is triangular and $T_{21}(\lambda)$ is quasi-triangular with $\tilde{r} = \tilde{p} + \tilde{n}_c - (n - 2) = p + n_c - n - 1 = r - 1$ diagonal blocks of size 2×2 having $(\lambda - \lambda_1)|(\lambda - \lambda_1)(\lambda^2 + a_i\lambda + b_i)$, $1 \leq i \leq \tilde{r}$ as invariant factors, with polynomials $\lambda^2 + a_i\lambda + b_i$ irreducible over $\mathbb{R}[\lambda]$ and not necessarily distinct. In addition, the only linear elementary divisors of $T_{21}(\lambda)$ are $\lambda - \lambda_1$, $2\tilde{r}$ times, and the real elementary divisors of $\tilde{Q}_2(\lambda)$ are those of $T_{21}(\lambda)$ and $T_{22}(\lambda)$ together with possible repetitions. Therefore $Q(\lambda)$ is equivalent to

$$T(\lambda) = \begin{bmatrix} Q_1(\lambda) & X_1(\lambda) & X_2(\lambda) \\ 0 & T_{21}(\lambda) & Y(\lambda) \\ 0 & 0 & T_{22}(\lambda) \end{bmatrix}.$$

Let

$$T_1(\lambda) = \begin{bmatrix} Q_1(\lambda) & X_1(\lambda) \\ 0 & T_{21}(\lambda) \end{bmatrix}, \quad T_2(\lambda) = T_{22}(\lambda)$$

and let us analyze the invariant factors of these matrices.

Since the real elementary divisors of $\tilde{Q}(\lambda)$ are those of $Q_1(\lambda)$ and $\tilde{Q}_2(\lambda)$, we conclude that the only linear elementary divisors of $T_1(\lambda)$ are $\lambda - \lambda_1$, $2 + 2\tilde{r} = 2 + 2r - 2 = 2r$ times and the real elementary divisors of $T(\lambda)$ are those of $T_1(\lambda)$ and $T_2(\lambda)$ counting with repetitions. In conclusion, $T(\lambda)$ is the desired matrix.

Using arguments similar to those used at the end of the proofs of Theorems 3.3 and 3.6, we prove that the theorem also holds when the leading matrix coefficient is singular. \square

EXAMPLE 4.4. Let $Q(\lambda)$ be the quadratic of Example 3.8. Its reversal has λ^2 , $\lambda^2, \lambda^2 + 1$ and $\lambda^2 + 1$ as elementary divisors and is equivalent to the quasi-triangular quadratic $\text{rev } T(\lambda) = \begin{bmatrix} \lambda^2 & -\lambda \\ \lambda & \lambda^2 \end{bmatrix} \oplus [\lambda^2 + 1]$. Hence $T(\lambda) = \begin{bmatrix} 1 & -\lambda \\ \lambda & 1 \end{bmatrix} \oplus [\lambda^2 + 1]$ is a quasi-triangular quadratic matrix polynomial strongly equivalent to $Q(\lambda)$ with the properties of Theorem 4.3. \square

REMARK 4.5. In Theorem 4.3 the quasi-triangular quadratic matrix polynomial $T(\lambda)$ is constructed so as to have the 2×2 diagonal blocks in the upper left part. We show that the construction can be modified so that the diagonal blocks of $T(\lambda)$ appear in any desired diagonal position.

1. Using real elementary transformations we can bring each 2×2 diagonal block of $T(\lambda)$ to the equivalent block $\text{diag}((\lambda - \lambda_1), (\lambda - \lambda_1)(\lambda^2 + a_i\lambda + b_i\lambda))$ so that $T(\lambda)$ is equivalent to a real (perhaps not quadratic) triangular matrix.
2. Theorem 3.2 guarantees the existence of a triangular matrix, $\tilde{T}(\lambda)$, equivalent to $T(\lambda)$, with the same diagonal elements of $T(\lambda)$ in any desired order. In particular we can arrange the diagonal elements as we wish but keeping together the blocks $\text{diag}((\lambda - \lambda_1), (\lambda - \lambda_1)(\lambda^2 + a_i\lambda + b_i\lambda))$.
3. Reversing the transformations of item 1, we recover a real monic quadratic block-triangular matrix with the 2×2 and 1×1 diagonal blocks in the desired order.
4. For quadratics with nonsingular leading coefficient, the 2×2 diagonal blocks can be constructed so that their non diagonal entries are of degree less than 2. It may happen however that other off-diagonal elements in the upper left part of the matrix have degree greater than 1. If so, we can use the corresponding diagonal entries or diagonal blocks to reduce their degree below 2. \square

The nonnegative integer $r = \max\{0, p + n_c - n\}$ appearing in Theorem 4.3 is an invariant for the equivalence of quadratic matrix polynomials. We now show that, for a given $n \times n$ quadratic matrix polynomial $Q(\lambda)$, this is the minimum number of 2×2 diagonal blocks that a quasi-triangular quadratic matrix polynomial $T(\lambda)$ equivalent to $Q(\lambda)$ may have. This property plays an important role in the next section. For this purpose we give the following definition.

DEFINITION 4.6. A monic quasi-triangular quadratic $T(\lambda) \in \mathbb{R}[\lambda]^{n \times n}$ with s 2×2 diagonal blocks is called *irreducible* if any monic quasi-triangular quadratic matrix polynomial equivalent to $T(\lambda)$ has at least s 2×2 diagonal blocks.

THEOREM 4.7. Let $Q(\lambda) \in \mathbb{R}[\lambda]^{n \times n}$ be quadratic with nonsingular leading coefficient and, with the notation (3.2)–(3.3), let $r = \max\{0, p + n_c - n\}$. A monic quasi-triangular quadratic matrix polynomial $T(\lambda)$ equivalent to $Q(\lambda)$ is irreducible if and only if the number of its 2×2 diagonal blocks is r .

The proof of this theorem is based on the following lemma.

LEMMA 4.8. *Let*

$$\begin{matrix} & 2s & n-2s \\ \begin{matrix} 2s \\ n-2s \end{matrix} & \begin{bmatrix} T_1(\lambda) & T_3(\lambda) \\ 0 & T_2(\lambda) \end{bmatrix}, & s > 0 \end{matrix}$$

be a monic quadratic matrix polynomial with $T_1(\lambda)$ quasi-triangular with s 2×2 diagonal blocks and $T_2(\lambda)$ triangular. If $T(\lambda)$ is irreducible then

- (i) the invariant factors of the i -th 2×2 diagonal block of $T_1(\lambda)$ are $(\lambda - \lambda_1)|(\lambda - \lambda_1)(\lambda^2 + a_i\lambda + b_i)$, $i = 1: s$, where λ_1 is a real eigenvalue of $T(\lambda)$ associated to a chain of elementary divisors of maximal length and the polynomials $\lambda^2 + a_i\lambda + b_i$ are irreducible over $\mathbb{R}[\lambda]$ and not necessarily distinct,
- (ii) $\lambda - \lambda_1, \dots, \lambda - \lambda_1$ ($2s$ times) are the only linear elementary divisors of $T_1(\lambda)$.
- (iii) The real elementary divisors of $T(\lambda)$ are the real elementary divisors of $T_1(\lambda)$ and $T_2(\lambda)$ together with possible repetitions.

Proof. (i)–(ii). Let $B(\lambda)$ be any 2×2 diagonal block of $T_1(\lambda)$. It has been seen in the proof of Theorem 4.3 that $B(\lambda)$ is irreducible over $\mathbb{R}[\lambda]$ if and only if its elementary divisors are $\lambda - \tilde{\lambda}$ twice and $\lambda^2 + a\lambda + b$, where $\tilde{\lambda}$ is the real eigenvalue of $B(\lambda)$ and $a^2 - 4b < 0$. Hence, if $T(\lambda)$ is irreducible then the elementary divisors of each 2×2 diagonal block must be of the required form. We now show that the real eigenvalues of the 2×2 diagonal blocks are all equal to say, λ_1 . Let $B_1(\lambda)$ and $B_2(\lambda)$ be 2×2 diagonal blocks of $T(\lambda)$ with real eigenvalue λ_1 and λ_2 , respectively. If $\lambda_1 \neq \lambda_2$ then the maximal possible length of the chain of elementary divisors associated to λ_1 and λ_2 in

$$B(\lambda) = \begin{bmatrix} B_1(\lambda) & X(\lambda) \\ 0 & B_2(\lambda) \end{bmatrix}$$

is 2 for any matrix polynomial $X(\lambda)$. For $B(\lambda)$, $n = 4$, $n_c = 2$ and $n - n_c = 2 \geq p$. This means that $B(\lambda)$ is triangularizable for any $X(\lambda)$. Therefore $T(\lambda)$ is not irreducible, a contradiction.

It remains to show λ_1 is the real eigenvalue of $T(\lambda)$ associated to a chain of elementary divisors of maximal length. In fact, assume that this is not the case and let p be the maximal length of the chains of elementary divisors associated to real eigenvalues. These polynomials must be elementary divisors of $T_2(\lambda)$ because λ_1 is the only real eigenvalue of $T_1(\lambda)$ (see [3, Thm. 1]). Now, $T_2(\lambda)$ is triangular of size $n - 2s$ and the sum of the degrees of the quadratic elementary divisors is $n_c - s$. Hence, by Theorem 3.6, $p + n_c - s \leq n - 2s$. Thus $n > p + n_c$ and by Theorem 3.6, $T(\lambda)$ is triangularizable, hence a contradiction.

(iii) As above, the elementary divisors of $T(\lambda)$ associated to real eigenvalues different from λ_1 are those of $T_2(\lambda)$. Hence it is enough to show that the elementary divisors of $T(\lambda)$ associated to λ_1 are $(\lambda - \lambda_1)$ $2s$ times together with the elementary divisors of $T_2(\lambda)$ which are powers of $(\lambda - \lambda_1)$. Let $T_4(\lambda)$ be the submatrix of $T_2(\lambda)$ whose diagonal elements are of the form $(\lambda - \lambda_1)(\lambda - \tilde{\lambda}_i)$ where $\tilde{\lambda}_i$ represents any real eigenvalue of $T(\lambda)$ including, possibly, $\tilde{\lambda}_i = \lambda_1$. Let us assume (see Remark 4.5) that

$$T_2(\lambda) = \begin{bmatrix} T_4(\lambda) & T_7(\lambda) \\ 0 & T_8(\lambda) \end{bmatrix}$$

and write

$$T(\lambda) = \begin{bmatrix} T_1(\lambda) & T_5(\lambda) & T_6(\lambda) \\ 0 & T_4(\lambda) & T_7(\lambda) \\ 0 & 0 & T_8(\lambda) \end{bmatrix}.$$

We claim that

$$\tilde{T}(\lambda) = \begin{bmatrix} T_1(\lambda) & T_5(\lambda) \\ 0 & T_4(\lambda) \end{bmatrix}$$

is equivalent to $\text{diag}(T_1(\lambda), T_4(\lambda))$ so that the elementary divisors of $T(\lambda)$ are those of $T_1(\lambda)$ and $T_2(\lambda)$ together with possible repetitions. So, let us prove our claim.

Let $T_{ii}^{(1)}(\lambda)$ be the i th 2×2 diagonal block of $T_1(\lambda)$ with invariant factors $(\lambda - \lambda_1) | (\lambda - \lambda_1)(\lambda^2 + a_i\lambda + b_i)$. Let $U(\lambda) = \text{diag}(U_1(\lambda), \dots, U_s(\lambda))$ and $V(\lambda) = \text{diag}(V_1(\lambda), \dots, V_n(\lambda))$ be unimodular matrices with $U_i(\lambda), V_i(\lambda)$ such that

$$D_i(\lambda) = U_i(\lambda)T_{ii}^{(1)}(\lambda)V_i(\lambda) = \text{diag}((\lambda - \lambda_1), (\lambda - \lambda_1)(\lambda^2 + a_i\lambda + b_i)), \quad i = 1 : s.$$

Then

$$\tilde{T}_1(\lambda) = U(\lambda)T_1(\lambda)V(\lambda) = \begin{bmatrix} D_1(\lambda) & X_{12}(\lambda) & \cdots & X_{1s}(\lambda) \\ & D_2(\lambda) & \ddots & \vdots \\ & & \ddots & X_{s-1,s}(\lambda) \\ & & & D_s(\lambda) \end{bmatrix}.$$

If $T_4(\lambda) = (t_{ij}^{(4)})$ with $t_{ii}^{(4)}(\lambda) = (\lambda - \lambda_1)(\lambda - \tilde{\lambda}_i)$ then the procedure described in the proof of Proposition 2.2 allows us to construct matrices $Z_1(\lambda), Z_2(\lambda) \in \mathbb{F}[\lambda]^{2s \times q}$ such that

$$\hat{T}(\lambda) = \begin{bmatrix} I_{2s} & Z_1(\lambda) \\ 0 & I_q \end{bmatrix} \tilde{T}(\lambda) \begin{bmatrix} I_{2s} & Z_2(\lambda) \\ 0 & I_q \end{bmatrix} = \begin{bmatrix} \tilde{T}_1(\lambda) & Y(\lambda) \\ 0 & T_4(\lambda) \end{bmatrix}$$

where

$$\deg(Y_{ij}(\lambda)) < \begin{cases} \deg(\gcd(\lambda - \lambda_1, t_{jj}^{(4)}(\lambda))) = \deg(\lambda - \lambda_1) & \text{if } i \text{ is odd,} \\ \deg(\gcd((\lambda - \lambda_1)(\lambda^2 + a_i\lambda + b_i), t_{jj}^{(4)}(\lambda))) = \deg(\lambda - \lambda_1) & \text{if } i \text{ is even,} \end{cases}$$

i.e., the elements of $Y(\lambda)$ are constant polynomials. We aim to show that, actually, $Y(\lambda) = 0$. In fact, consider the submatrix

$$\hat{T}_k(\lambda) = \begin{bmatrix} \lambda - \lambda_1 & 0 & y_{2k,j} \\ 0 & (\lambda - \lambda_1)(\lambda^2 + a_k\lambda + b_k) & y_{2k+1,j} \\ 0 & 0 & (\lambda - \lambda_1)(\lambda - \tilde{\lambda}_j) \end{bmatrix},$$

and let us show that $y_{2k,j} = y_{2k+1,j} = 0$. If $y_{2k,j} \neq 0$ or $y_{2k+1,j} \neq 0$ then since they are constant polynomials, the first invariant factor of $\hat{T}_k(\lambda)$ is equal to 1 so that the geometric multiplicity of λ_1 is at most 2. By Theorem 3.6, $\hat{T}_k(\lambda)$ is triangularizable since $n - n_c = 3 - 1 = 2 \geq p$. Hence there are unimodular matrices $\tilde{U}_k(\lambda), \tilde{V}_k(\lambda)$ such that $\tilde{U}_k(\lambda)\hat{T}_k(\lambda)\tilde{V}_k(\lambda) \in \mathbb{R}[\lambda]^{3 \times 3}$ is upper triangular and quadratic. By Remark 4.5 we can assume for notational simplicity that $k = s$ and $i = 1$. Thus, if $\tilde{U}(\lambda) = \text{diag}(U_1(\lambda)^{-1}, \dots, U_{s-1}^{-1}(\lambda), \tilde{U}_s(\lambda), I_{q-1})$, $\tilde{V}(\lambda) = \text{diag}(V_1(\lambda)^{-1}, \dots, V_{s-1}^{-1}(\lambda), \tilde{V}_s(\lambda), I_{q-1})$ then $\tilde{U}(\lambda)\hat{T}(\lambda)\tilde{V}(\lambda)$ is upper triangular and quadratic with $(s-1) 2 \times 2$ diagonal blocks. Hence $T(\lambda)$ is equivalent to a triangular matrix with less than $s 2 \times 2$ diagonal blocks contradicting that it is irreducible. In conclusion $Y(\lambda) = 0$ and $\tilde{T}(\lambda)$ is equivalent to $\text{diag}(T_1(\lambda), T_4(\lambda))$ as desired. \square

Proof of Theorem 4.7. Let $T(\lambda)$ be an irreducible monic quasi-triangular quadratic matrix polynomial equivalent to $Q(\lambda)$ and let s be the number of 2×2 diagonal blocks of $T(\lambda)$. First of all if $T(\lambda)$ is triangular then $s = 0$ and there is nothing to prove. So, we analyze the case $s > 0$ so that $Q(\lambda)$ is not triangularizable over $\mathbb{R}[\lambda]$ and $r = p + n_c - n > 0$. We show that $s > r$ leads to a contradiction. In fact, if $s > r$ then, by Theorem 4.3, $T(\lambda)$ is equivalent to a quasi-triangular matrix with r diagonal blocks of size 2×2 contradicting that $T(\lambda)$ is irreducible. Therefore $s \leq r$. Now, by Remark 4.5, we can assume that

$$T(\lambda) = \begin{bmatrix} T_1(\lambda) & T_3(\lambda) \\ 0 & T_2(\lambda) \end{bmatrix},$$

where $T_1(\lambda)$ and $T_2(\lambda)$ have the properties of Lemma 4.8. Then, $T_2(\lambda)$ is a triangular matrix of size $(n - 2s) \times (n - 2s)$, the number of its elementary divisors which are powers of $(\lambda - \lambda_1)$ is $p - 2s$ and the sum of the degrees of its elementary divisors which are powers of irreducible quadratic polynomials over $\mathbb{R}[\lambda]$ is $n_c - s$. By Theorem 3.6, $n - 2s \geq (p - 2s) + n_c - s$, that is, $n \geq p + n_c - s$, or equivalently, $s \geq p + n_c - n = r$. Since $s \leq r$ we conclude that $s = r$, as desired.

The converse follows now from the already proved fact that there cannot be monic quasi-triangular quadratic matrices equivalent to $Q(\lambda)$ with less than r diagonal blocks of size 2×2 . \square

We conclude this section by pointing out that, according to Lemma 4.8 and Theorem 4.7, Theorem 4.3 provides a procedure to construct an irreducible block-triangular quadratic matrix polynomial equivalent to $Q(\lambda)$.

5. A Schur-like theorem for quadratic matrix pencils.

5.1. The complex case. The theorem of Schur for complex matrices states that for any $A \in \mathbb{C}^{n \times n}$ there is a unitary matrix $U \in \mathbb{C}^{n \times n}$ such that $U^*AU = T$ is a triangular matrix. Schur's theorem can be rewritten as follows.

THEOREM 5.1. *Let $A \in \mathbb{C}^{n \times n}$. There are subspaces $\mathcal{V}_1, \dots, \mathcal{V}_n$ of \mathbb{C}^n satisfying*

- (i) $\mathbb{C}^n = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \dots \oplus \mathcal{V}_n$,
- (ii) *the subspaces $\mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \dots \oplus \mathcal{V}_k$, $k = 1: n$ are A -invariant,*
- (iii) $\mathcal{V}_k = \langle u_k \rangle$, $k = 1: n$, where u_1, \dots, u_n form an orthonormal system of vectors of \mathbb{C}^n .

We investigate in this section how the matrix version ($U^*AU = T$) and the subspaces version (Theorem 5.1) of Schur's theorem extend to any linearizations of quadratic matrix polynomials with nonsingular leading coefficient. Recall that a pencil $\lambda I - A$ is a monic linearization of $Q(\lambda) \in \mathbb{F}[\lambda]^{n \times n}$ if $A \in \mathbb{F}^{2n \times 2n}$ has the same elementary divisors as $Q(\lambda)$. For example, the left companion matrix

$$A = \begin{bmatrix} 0 & -A_0A_2^{-1} \\ I_n & -A_1A_2^{-1} \end{bmatrix} =: C_L(Q) \quad (5.1)$$

defines a linearization of $Q(\lambda) = A_2\lambda^2 + A_1\lambda + A_0$. Now suppose that $Q(\lambda)$ is equivalent to a quadratic matrix polynomial $T(\lambda)$. Then their left companion matrices are similar, that is, there exists a nonsingular S such that $C_L(Q)S = SC_L(T)$. The similarity S is called a *left companion structure preserving similarity* [8, Def. 2.1].

A consequence of Theorem 3.3 is the following result about the linearizations of triangularizable quadratic matrix polynomials.

THEOREM 5.2. *Let $Q(\lambda) \in \mathbb{C}[\lambda]^{n \times n}$ be a quadratic matrix polynomial with nonsingular leading coefficient and let $\lambda I - A$ be any $2n \times 2n$ monic linearization of $Q(\lambda)$.*

Then there is an invertible matrix $S \in \mathbb{C}^{2n \times 2n}$ such that

$$A_T = S^{-1}AS$$

is a quasi-triangular matrix with 2×2 diagonal blocks that are companion matrices of scalar quadratic polynomials. In particular, if $A = C_L(Q)$ in (5.1) then S is, up to column permutation, a left companion structure preserving similarity.

Proof. By Corollary 3.4, $Q(\lambda)$ is equivalent to an upper triangular quadratic matrix polynomial $T(\lambda) = I\lambda^2 + T_1\lambda + T_0$. Using the permutation matrix,

$$P = [e_1 \ e_{n+1} \ e_2 \ e_{n+2} \ \dots \ e_n \ e_{2n}], \quad (5.2)$$

where e_i is the i th column of the $2n \times 2n$ identity matrix, we have that

$$P^T C_L(T)P = P^T \begin{bmatrix} 0 & -T_1 \\ I_n & -T_0 \end{bmatrix} P =: A_T = (A_{ij})_{1 \leq i, j \leq n},$$

where, with the notation $T_1 = (t_{ij}^{(1)})_{1 \leq i, j \leq n}$ and $T_0 = (t_{ij}^{(0)})_{1 \leq i, j \leq n}$,

$$A_{ii} = \begin{bmatrix} 0 & -t_{ii}^{(0)} \\ 1 & -t_{ii}^{(1)} \end{bmatrix}, \quad A_{ij} = \begin{bmatrix} 0 & -t_{ij}^{(0)} \\ 0 & -t_{ij}^{(1)} \end{bmatrix} \text{ for } i < j, \quad A_{ij} = 0_{2 \times 2} \text{ for } i > j. \quad (5.3)$$

Hence A_T is a quasi-triangular matrix with 2×2 diagonal blocks, which are the companion matrices of the diagonal elements of $T(\lambda)$. The pencil A_T is a linearization of $T(\lambda)$ and, since $T(\lambda)$ is equivalent to $Q(\lambda)$, $\lambda I - A_T$ is also a linearization of $Q(\lambda)$. Therefore A and A_T are similar matrices, and so there is an invertible matrix S such that $A_T = S^{-1}AS$.

Finally, if $A = C_L(Q)$ then $PS^{-1}C_L(Q)SP = PA_T P^T = C_L(T)$, i.e., SP is a left companion structure preserving similarity. \square

The quasi-triangular structure of A_T in Theorem 5.2 reveals the existence of invariant subspaces with respect to any monic linearization $\lambda I - A$ of $Q(\lambda)$. This leads to a Schur-like theorem for quadratic matrix polynomials along the lines of Theorem 5.1.

THEOREM 5.3. *Let $Q(\lambda) \in \mathbb{C}[\lambda]^{n \times n}$ be quadratic with nonsingular leading coefficient and let $\lambda I - A$ be any $2n \times 2n$ monic linearization of $Q(\lambda)$. There are subspaces $\mathcal{V}_1, \dots, \mathcal{V}_n$ of \mathbb{C}^{2n} satisfying*

- (i) $\mathbb{C}^{2n} = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \dots \oplus \mathcal{V}_n$,
- (ii) the subspaces $\mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \dots \oplus \mathcal{V}_k$, $k = 1: n$ are A -invariant,
- (iii) $\dim \mathcal{V}_k = 2$ and $\mathcal{V}_k = \langle u_k, Au_k \rangle$, $k = 1: n$, where the generating vectors u_1, \dots, u_n are linearly independent and can be chosen to form an orthonormal system of vectors of \mathbb{C}^{2n} .

Proof. By Theorem 5.2 there is a nonsingular matrix $S \in \mathbb{C}^{2n \times 2n}$ such that $S^{-1}AS = A_T$, where A_T is quasi-triangular with 2×2 blocks A_{ij} as in (5.3). Let $s_i = Se_i$ and $\mathcal{V}_k = \langle s_{2k-1}, s_{2k} \rangle$, $k = 1, \dots, n$. Then it is plain that $\mathbb{C}^{2n} = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \dots \oplus \mathcal{V}_n$ and since $AS = SA_T$ we have that $\mathcal{V}_1 \oplus \dots \oplus \mathcal{V}_k$ is A -invariant and $s_{2k} = SA_T e_{2k-1} = AS e_{2k-1} = A s_{2k-1}$. Hence there are linearly independent vectors s_1, \dots, s_{2n-1} such that $\{s_1, As_1, \dots, s_{2k-1}, As_{2k-1}\}$ is a basis of $\mathcal{V}_1 \oplus \dots \oplus \mathcal{V}_k$. Note that the \mathcal{V}_k , $k = 1: n$ are Krylov subspaces of dimension 2.

Let P be the permutation in (5.2) and $X \in \mathbb{C}^{2n \times n}$ such that $Xe_k = s_{2k-1}$. Then $SP^T = [X \ AX]$ and $AS = SA_T$ can be rewritten as

$$A \begin{bmatrix} X & AX \end{bmatrix} = \begin{bmatrix} X & AX \end{bmatrix} \begin{bmatrix} 0 & -T_0 \\ I_n & -T_1 \end{bmatrix}, \quad (5.4)$$

where T_i , $i = 0, 1$ are upper triangular matrices. Now let $X = UR$ be a QR factorization of X . Then (5.4) becomes

$$A[U \ AU] = [U \ AU] \begin{bmatrix} R & 0 \\ 0 & R \end{bmatrix} \begin{bmatrix} 0 & -T_0 \\ I_n & -T_1 \end{bmatrix} \begin{bmatrix} R & 0 \\ 0 & R \end{bmatrix}^{-1} = [U \ AU] \begin{bmatrix} 0 & -\tilde{T}_0 \\ I_n & -\tilde{T}_1 \end{bmatrix},$$

where the \tilde{T}_i , $i = 0, 1$ are still upper triangular. If we let $\tilde{S} = [U \ AU]P$ then $A\tilde{S} = \tilde{S}A_{\tilde{T}}$, where $A_{\tilde{T}}$ has the same block structure as A_T . As above, we have that $\mathcal{V}_k = \langle \tilde{s}_{2k-1}, \tilde{s}_{2k} \rangle = \langle \tilde{s}_{2k-1}, A\tilde{s}_{2k-1} \rangle = \langle u_k, Au_k \rangle$ showing that the generating vectors of the Krylov subspaces \mathcal{V}_k can be taken to form an orthonormal system. \square

The next result, which follows from the above proof, is a matrix version of Theorem 5.3.

THEOREM 5.4. *For any $2n \times 2n$ monic linearization $\lambda I - A$ of a quadratic $Q(\lambda) \in \mathbb{C}[\lambda]^{n \times n}$ with nonsingular leading coefficient, there exists $U \in \mathbb{C}^{2n \times n}$ with orthonormal columns such that $[U \ AU]$ is nonsingular and*

$$[U \ AU]^{-1}A[U \ AU] = \begin{bmatrix} 0 & -T_0 \\ I_n & -T_1 \end{bmatrix},$$

where $I_n\lambda^2 + T_1\lambda + T_0$ is upper triangular and equivalent to $Q(\lambda)$.

Theorems 5.3 and 5.4 show that to triangularize a quadratic matrix polynomial, it suffices to construct a set of n orthonormal generating vectors. It is shown in [10] that these generating vectors can be constructed from the Schur vectors of a linearization of $Q(\lambda)$.

5.2. The real case. The extension of a Schur theorem from linear to quadratic matrix polynomials is more involved in the real case. Recall first that if $A \in \mathbb{R}^{n \times n}$ then there exists an orthogonal matrix $U \in \mathbb{R}^{n \times n}$ such that $U^T A U = T$ is a quasi-triangular matrix with 1×1 and 2×2 diagonal blocks corresponding, respectively, to the real and the non-real complex conjugate eigenvalues of A . We aim to write this decomposition as a theorem that, properly generalized, reflects the fact that any real quadratic matrix polynomial can be reduced to quasi-triangular form.

THEOREM 5.5. *Let $A \in \mathbb{R}^{n \times n}$. There is a nonnegative integer $r \leq n/2$ and there are subspaces $\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_{n-r}$ of \mathbb{R}^n satisfying*

- (i) $\mathbb{R}^n = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \dots \oplus \mathcal{V}_{n-r}$,
- (ii) the subspaces $\mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \dots \oplus \mathcal{V}_k$, $k = 1 : n - r$, are A -invariant,
- (iii) $\mathcal{V}_k = \langle u_{2k-1}, u_{2k} \rangle$, $k = 1 : r$ and $\mathcal{V}_{r+k} = \langle u_{2r+k} \rangle$, $k = 1 : n - 2r$, where u_1, \dots, u_n form an orthonormal basis of \mathbb{R}^n ,
- (iv) $r \leq r'$ for any other decomposition of \mathbb{R}^n as a direct sum of subspaces $\mathcal{V}'_1, \mathcal{V}'_2, \dots, \mathcal{V}'_{n-r'}$ satisfying properties (i)-(iii).

Proof. Let r be the number of pairs of non-real complex conjugate eigenvalues of A . Then there is an orthogonal matrix $U \in \mathbb{R}^{n \times n}$ such that

$$U^T A U = T = \begin{matrix} & & 2r & n-2r \\ & & \begin{bmatrix} T_1 & T_3 \\ 0 & T_2 \end{bmatrix} \\ & & n-2r & \end{matrix}, \quad (5.5)$$

where T_1 is quasi-triangular with r 2×2 diagonal blocks containing the pairs of non-real eigenvalues and T_2 is upper triangular. Properties (i)-(iii) follow by defining $\mathcal{V}_k = \langle u_{2k-1}, u_{2k} \rangle$ for $k = 1 : r$ and $\mathcal{V}_{r+k} = \langle u_{2r+k} \rangle$ for $k = 1 : n - 2r$. The minimality

property of r of condition (iv) is also clear. If there is a nonnegative integer $r' < r$ and subspaces $\mathcal{V}'_1, \mathcal{V}'_2, \dots, \mathcal{V}'_{n-r'}$ satisfying properties (i)-(iii) then A has a Schur decomposition with less than r 2×2 diagonal blocks. Hence A would have more than $n - r$ real eigenvalues, which is not possible. \square

REMARK 5.6. We have seen that Theorem 5.5 is a consequence of the existence of a real Schur form of A . The converse is also true. If \mathbb{R}^n admits a decomposition as direct sum of $n - r$ subspaces satisfying (i)-(iii) and we write $U = [u_1 \ \dots \ u_n]$ then $U^T A U$ is quasi-triangular as in (5.5) with r diagonal blocks of size 2×2 . If the eigenvalues of any of these blocks were real then such a block could be further reduced by orthogonal similarity to triangular form. Then \mathbb{R}^n could be decomposed as a direct sum of $n - r'$ subspaces satisfying properties (i)-(iii) and $r' < r$ contradicting the minimality property (iv) of r . Hence, the eigenvalues of the 2×2 blocks are non-real complex conjugate numbers.

When $Q(\lambda)$ is real and not triangularizable over $\mathbb{R}[\lambda]$ we can use Theorem 4.3 to obtain a similar but more general result than Theorem 5.2. In order to simplify the statement of such a theorem we introduce the following notion.

DEFINITION 5.7. Let $Q(\lambda) \in \mathbb{R}[\lambda]^{n \times n}$ be quadratic with nonsingular leading coefficient and, with the notation (3.2)–(3.3), let $r = \max\{0, p + n_c - n\}$ and $s = n - 2r$. A matrix $A \in \mathbb{R}^{2n \times 2n}$ is said to be an irreducible block-triangular controllability linearization of $Q(\lambda)$ if

$$\begin{array}{cc} 4r & 2s \\ 4r & \begin{bmatrix} A_1 & A_3 \\ 0 & A_2 \end{bmatrix}, \\ 2s & \end{array}$$

where $A_1 = (A_{ij})_{1 \leq i, j \leq r}$ has 4×4 blocks A_{ij} such that $A_{ij} = 0_{4 \times 4}$ for $i > j$,

$$A_{ii} = \left[\begin{array}{cc|cc} 0 & -t_{2i-1, 2i-1}^{(0)} & 0 & -t_{2i-1, 2i}^{(0)} \\ 1 & -t_{2i-1, 2i-1}^{(1)} & 0 & -t_{2i-1, 2i}^{(1)} \\ \hline 0 & -t_{2i, 2i-1}^{(0)} & 0 & -t_{2i, 2i}^{(0)} \\ 0 & -t_{2i, 2i-1}^{(1)} & 1 & -t_{2i, 2i}^{(1)} \end{array} \right], \quad A_{ij} = \left[\begin{array}{cc|cc} 0 & -t_{2i-1, 2j-1}^{(0)} & 0 & -t_{2i-1, 2j}^{(0)} \\ 0 & -t_{2i-1, 2j-1}^{(1)} & 0 & -t_{2i-1, 2j}^{(1)} \\ \hline 0 & -t_{2i, 2j-1}^{(0)} & 0 & -t_{2i, 2j}^{(0)} \\ 0 & -t_{2i, 2j-1}^{(1)} & 0 & -t_{2i, 2j}^{(1)} \end{array} \right]$$

for $i < j$, $A_3 = (A_{ij})_{1 \leq i < r < j \leq r+s}$ has 4×2 blocks A_{ij} such that

$$A_{ij} = \begin{bmatrix} 0 & -t_{2i-1, j}^{(0)} \\ 0 & -t_{2i-1, j}^{(1)} \\ 0 & -t_{2i, j}^{(0)} \\ 0 & -t_{2i, j}^{(1)} \end{bmatrix}$$

and $A_2 = (A_{ij})_{r+1 \leq i, j \leq r+s}$ has 2×2 blocks A_{ij} such that $A_{ij} = 0_{2 \times 2}$ for $r+1 \leq j < i \leq r+s$, $A_{ii} = \begin{bmatrix} 0 & -t_{ii}^{(0)} \\ 1 & -t_{ii}^{(1)} \end{bmatrix}$ for $r+1 \leq i \leq r+s$, and $A_{ij} = \begin{bmatrix} 0 & -t_{ij}^{(0)} \\ 0 & -t_{ij}^{(1)} \end{bmatrix}$ for $r+1 \leq i < j \leq r+s$, and

- (i) the 4×4 diagonal blocks A_{ii} have for invariant factors $(\lambda - \lambda_1)|(\lambda - \lambda_1)(\lambda^2 + a_i \lambda + b_i)$, $1 \leq i \leq r$, where λ_1 is a real eigenvalue of $Q(\lambda)$ associated to a chain of elementary divisors of maximal length and the not necessarily distinct scalar polynomials $\lambda^2 + a_i \lambda + b_i$ are irreducible over $\mathbb{R}[\lambda]$,
- (ii) $\lambda - \lambda_1$ ($2r$ times) are the only linear elementary divisors of A_1 ,
- (iii) the real elementary divisors of A are the real elementary divisors of A_1 and A_2 together with possible repetitions.

The proof of the next theorem is similar to that of Theorem 5.2 and we omit it.

THEOREM 5.8. *Let $Q(\lambda) \in \mathbb{R}[\lambda]^{n \times n}$ be quadratic with nonsingular leading coefficient and let $\lambda I - A$ be any $2n \times 2n$ monic linearization of $Q(\lambda)$. Then there is a nonsingular matrix $S \in \mathbb{R}^{2n \times 2n}$ such that $A_T = S^{-1}AS$ is an irreducible block-triangular controllability linearization of $Q(\lambda)$. In particular, if $A = C_L(Q)$ is the left companion matrix of $Q(\lambda)$ in (5.1) then S can be chosen to be, up to column permutation, a left companion structure preserving similarity.*

Notice that when $Q(\lambda)$ is triangularizable, $r = 0$ and Theorem 5.8 reduces to Theorem 5.2.

The next theorem can be thought of as a Schur-like theorem for linearizations of real quadratic matrix polynomials (compared with Theorems 5.3 and 5.5).

THEOREM 5.9. *Let $Q(\lambda) \in \mathbb{R}[\lambda]^{n \times n}$ be quadratic with nonsingular leading coefficient and let $\lambda I - A$ be any $2n \times 2n$ monic linearization of $Q(\lambda)$. Then there is a nonnegative integer $r \leq n/2$ and there are subspaces $\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_{n-r}$ of \mathbb{R}^{2n} satisfying*

- (i) $\mathbb{R}^{2n} = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \dots \oplus \mathcal{V}_{n-r}$,
- (ii) the subspaces $\mathcal{V}_1 \oplus \dots \oplus \mathcal{V}_k$, $k = 1: n - r$ are A -invariant,
- (iii) $\dim \mathcal{V}_k = 4$ with $\mathcal{V}_k = \langle u_{2k-1}, Au_{2k-1}, u_{2k}, Au_{2k} \rangle$ for $k = 1: r$ and $\dim \mathcal{V}_k = 2$ with $\mathcal{V}_k = \langle u_{2r+k}, Au_{2r+k} \rangle$ for $k = 1: n - 2r$, where the generating vectors u_1, \dots, u_n are linearly independent and can be chosen to form an orthonormal system of vectors of \mathbb{R}^{2n} ,
- (iv) $r \leq r'$ for any other decomposition of \mathbb{R}^{2n} as a direct sum of subspaces $\mathcal{V}'_1, \mathcal{V}'_2, \dots, \mathcal{V}'_{n-r'}$ satisfying properties (i)–(iii).

Proof. Let $r = \max\{0, p + n_c - n\}$, where p and n_c are defined in (3.3). By Theorem 5.8, there is a nonsingular matrix $S \in \mathbb{R}^{2n \times 2n}$ such that $S^{-1}AS = A_T$, where $A_T = (A_{ij})_{1 \leq i, j \leq r+s}$ and the blocks A_{ij} have the form described in Theorem 5.8. Let $s_k = Se_k$ and

$$\mathcal{V}_k = \begin{cases} \langle s_{4k-3}, s_{4k-2}, s_{4k-1}, s_{4k} \rangle, & k = 1: r, \\ \langle s_{4r+2k-1}, s_{4r+2k} \rangle, & k = 1: n - 2r. \end{cases}$$

Then it is plain that $\mathbb{R}^{2n} = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \dots \oplus \mathcal{V}_{n-r}$ and since $AS = SA_T$ we have that $\mathcal{V}_1 \oplus \dots \oplus \mathcal{V}_k$ is A -invariant and $s_{2k} = As_{2k-1}$ for $k = 1: n$. Putting $u_k = s_{2k-1}$ property (iii) follows at once. The proof that the generating vectors can be taken orthogonal is analogous to that for the complex case.

It only remains to prove property (iv). Assume that $\mathbb{R}^{2n} = \mathcal{V}'_1 \oplus \mathcal{V}'_2 \oplus \dots \oplus \mathcal{V}'_{n-r'}$ is another decomposition of \mathbb{R}^{2n} as a direct sum of subspaces satisfying properties (i)–(iii) with $r' < r$. Then there is a matrix A'_t similar to A and with the form described in Theorem 5.8 and r' diagonal blocks of size 4×4 . With that matrix A'_t it is easy to construct a real quasi-triangular monic quadratic matrix polynomial, $T'(\lambda)$, with 2×2 and 1×1 diagonal blocks in the diagonal and the number of 2×2 diagonal blocks being $r' < r$. But this is impossible because $Q(\lambda)$ and $T'(\lambda)$ are equivalent and by Theorem 4.7 the minimum number of 2×2 diagonal blocks in any monic block-triangular quadratic matrix polynomial $T(\lambda)$ equivalent to $Q(\lambda)$ is r . \square

We now give an analogue of Theorem 5.4 for real quadratic matrix polynomials.

THEOREM 5.10. *For any monic linearization $\lambda I - A \in \mathbb{R}[\lambda]^{2n \times 2n}$ of a quadratic $Q(\lambda) \in \mathbb{R}[\lambda]^{n \times n}$ with nonsingular leading coefficient, there exists $U \in \mathbb{R}^{2n \times n}$ with orthonormal columns such that $[U \ AU]$ is nonsingular and*

$$[U \ AU]^{-1}A[U \ AU] = \begin{bmatrix} 0 & -T_0 \\ I_n & -T_1 \end{bmatrix},$$

where $I_n\lambda^2 + T_1\lambda + T_0$ is equivalent to $Q(\lambda)$ and upper quasi-triangular with $r = \max\{0, p + n_c - n\}$ 2×2 diagonal blocks, where p and n_c are defined in (3.3).

6. Conclusion. We have shown that every regular quadratic matrix polynomial is triangularizable over $\mathbb{C}[\lambda]$ and have identified the real quadratics that are triangularizable over the real numbers. We have also shown that those quadratics that are not triangularizable over $\mathbb{R}[\lambda]$ are quasi-triangularizable.

We have established a Schur-like theorem for quadratic matrix polynomials, which provides the foundation for the design of algorithms reducing quadratic matrix polynomials to upper triangular quadratic matrix polynomials [10]. Theorem 5.3 shows that such algorithm should aim to compute a set of, possibly orthonormal, generating vectors.

REFERENCES

- [1] A. AMPARAN, S. MARCAIDA, AND I. ZABALLA, *On the existence of linear systems with prescribed invariants for system similarity*, Linear Algebra Appl., 413 (2006), pp. 510–533.
- [2] T. BETCKE, N. J. HIGHAM, V. MEHRMANN, C. SCHRÖDER, AND F. TISSEUR, *NLEVP: A collection of nonlinear eigenvalue problems*, MIMS EPrint 2011.116, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, Dec. 2011.
- [3] D. CARLSON, *Inequalities for the degrees of elementary divisors of modules*, Linear Algebra Appl., 5 (1972), pp. 293–298.
- [4] E. M. DE SÁ, *Imbedding conditions for λ -matrices*, Linear Algebra Appl., 24 (1979), pp. 33–50.
- [5] ———, *On the diagonal of integer matrices*, Czechoslovak Mathematics, 5 (1980), pp. 207–212.
- [6] J. DION AND C. COMMAULT, *Smith-mcmillan factorizations at infinity of rational matrix functions and their control interpretation*, Systems and Control Letters, 1 (1982), pp. 312–320.
- [7] F. R. GANTMACHER, *The Theory of Matrices*, vol. one, Chelsea, New York, 1959.
- [8] S. D. GARVEY, P. LANCASTER, A. A. POPOV, U. PRELLS, AND I. ZABALLA, *Filters connecting isospectral quadratic systems*, To appear in Linear Algebra Appl., (2012).
- [9] I. GOHBERG, P. LANCASTER, AND L. RODMAN, *Matrix Polynomials*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2009. Unabridged republication of book first published by Academic Press in 1982.
- [10] Y. NAKATSUKASA, L. TASLAMAN, F. TISSEUR, AND I. ZABALLA, *A schur approach to triangularize quadratic matrix polynomials*, tech. report, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, 2012. In preparation.
- [11] H. H. ROSENBROCK, *State-space and Multivariable Theory*, Thomas Nelson and Sons, London, 1970.
- [12] F. TISSEUR AND K. MEERBERGEN, *The quadratic eigenvalue problem*, SIAM Rev., 43 (2001), pp. 235–286.
- [13] A. I. G. VARDULAKIS, *Linear multivariable control*, Wiley, New York, 1991.