Normal Forms of Symplectic Pencils and the Discrete-Time Algebraic Riccati Equation*

Harald K. Wimmer
Mathematisches Institut
Universität Würzburg
D-8700 Würzburg, Germany

Dedicated to Professor Wolfgang Hahn on the occasion of his eightieth birthday.

Submitted by Thomas J. Laffey

ABSTRACT

The solution of the discrete-time algebraic Riccati equation leads to symplectic pencils of matrices. Normal forms of such pencils under symplectic equivalence are determined. Special attention is given to characteristic roots of modulus 1 and their corresponding elementary divisors and inertial invariants.

1. INTRODUCTION: THE ALGEBRAIC RICCATI EQUATION

The motivation for our study comes from the linear-quadratic regulator problem for discrete-time systems (see e.g. [1, 9, 10]). Consider the linear system

$$x(k+1) = Fx(k) + Bu(k), \qquad x(0) = x_0,$$

$$y(k) = Cx(k), \qquad k = 0,1,2,\ldots,$$

*Research supported by the Stimulation Programme of the Commission of the European Communities (grant SC1/0126-C) and by Deutsche Forschungsgemeinschaft (grant KN 164/3-1).

655 Avenue of the Americas, New York, NY 10010 0024-3795/91/$3.50

together with the cost functional

$$V = \frac{1}{2} \sum_{k=0}^{\infty} \left[y^*(k) T y(k) + u^*(k) R u(k) \right]$$

where F, B, C, T, and R are complex matrices whose dimensions are consistent with $x(k) \in \mathbb{C}^n$, $u(k) \in \mathbb{C}^p$, $y(k) \in \mathbb{C}^q$, and where R is positive definite ($R > 0$) and T is positive semidefinite ($T \geqslant 0$). Define

$$\Gamma = BR^{-1}B^* \quad \text{and} \quad Q = C^*TC,$$

and let G and D be matrices such that

$$\Gamma = GG^* \quad \text{and} \quad Q = P^*P.$$

If the pair (F, B) is stabilizable and if $\binom{F}{P}$ is detectable, then there exists an optimal control $\tilde{u}(k)$ which minimizes V (see [10]). The control law can be derived from the discrete maximum principle [17, 9]. The hamiltonian formalism applied to

$$H(k) = \tfrac{1}{2}\left[x(k)^* C^* T C x(k) + u^*(k) R u(k) \right] + p^*(k+1) x(k+1)] \tag{1.1}$$

leads to a system of linear difference equations which couples the state $x(k)$ with the adjoint state $p(k)$, namely

$$x(k+1) = Fx(k) - BR^{-1}B^* p(k+1),$$

$$p(k) = C^*TCx(k) + F^* p(k+1). \tag{1.2}$$

Under the given assumption we have $p(k) = Xx(k)$. The matrix X is a solution of the *discrete-time algebraic Riccati equation*

$$X - F^*XF + F^*XB(R + B^*XB)^{-1} B^*XF - Q = 0. \tag{1.3}$$

The optimal control $\tilde{u}$ is of the form

$$\tilde{u}(k) = -(R + B^*\tilde{X}B)^{-1} B^*\tilde{X}Fx(k),$$

where $\tilde{X}$ is the unique positive semidefinite solution of (1.3). The minimal cost $\tilde{V}$ is given by

$$\tilde{V} = \tfrac{1}{2}x(0)^*\tilde{X}x(0).$$

Furthermore, the closed-loop system

$$x(k+1) = \left[F - B(R + B^*\tilde{X}B)^{-1}B^*\tilde{X}F\right]x(k)$$

is asymptotically stable, or equivalently, all the eigenvalues of the matrix

$$F_{\tilde{X}} = F - B(R + B^*\tilde{X}B)^{-1}B^*\tilde{X}F$$

lie in the open unit disc.

To the linear first-order difference equation (1.2) corresponds a matrix pencil. We rewrite (1.2) as

$$\begin{pmatrix} F & 0 \\ -Q & I \end{pmatrix}\begin{pmatrix} x(k) \\ p(k) \end{pmatrix} - \begin{pmatrix} I & \Gamma \\ 0 & F^* \end{pmatrix}\begin{pmatrix} x(k+1) \\ p(k+1) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and put

$$M = \begin{pmatrix} F & 0 \\ -Q & I \end{pmatrix} \quad \text{and} \quad L = \begin{pmatrix} I & \Gamma \\ 0 & F^* \end{pmatrix}. \tag{1.4}$$

It is the pencil $M - zL$ and its (finite and infinite) elementary divisors and Jordan chains which determines the solution space of (1.2). From Pappas, Laub, and Sandell [16], de Souza, Gevers, and Goodwin [6], Mehrmann [13, 14], and other authors we know that the pencil $M - zL$ also plays a crucial role in the study of the algebraic Riccati equation (1.3).

In this first section we focus on Equation (1.3) and put together the basic facts on the pencil $M - zL$. We shall see how certain assumptions on elementary divisors of $M - zL$ or on factorizations of $\det(M - zL)$ are essential for the existence of solutions of (1.3). Only hermitian matrices which satisfy (1.3) will be considered as solutions of (1.3).

LEMMA 1.1.

(a) *Suppose U, V, and* Λ *are complex* $n \times n$ *matrices which satisfy*

$$\begin{pmatrix} F & 0 \\ -Q & I \end{pmatrix}\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} I & \Gamma \\ 0 & F^* \end{pmatrix}\begin{pmatrix} U \\ V \end{pmatrix}\Lambda. \tag{1.5}$$

If

(α) *U is nonsingular,*
(β) VU^{-1} *is hermitian, and*
(γ) $I + G^*(VU^{-1})G$ *is nonsingular,*

then

$$X = VU^{-1}$$

is a solution of (1.3).

(b) *If X is a solution of* (1.3), *then* (1.5) *holds with*

$$\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} I \\ X \end{pmatrix} \tag{1.6}$$

and

$$\Lambda = (I + \Gamma X)^{-1}F. \tag{1.7}$$

Furthermore we have

$$(M - zL)\begin{pmatrix} I & 0 \\ X & I \end{pmatrix} = W\left[\begin{pmatrix} \Lambda & 0 \\ 0 & I \end{pmatrix} - z\begin{pmatrix} I & D \\ 0 & \Lambda^* \end{pmatrix}\right], \tag{1.8}$$

where $D = D^* = (I + \Gamma X)^{-1}\Gamma$ *and*

$$W = \begin{pmatrix} I + \Gamma X & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} I & 0 \\ F^*X & I \end{pmatrix}.$$

Proof. Let us first note a useful matrix identity (see e.g. [7, p. 3]) which involves $\Gamma = GG^*$. We have

$$(I + \Gamma X)^{-1} = I - G(I + G^*XG)^{-1}G^*X. \tag{1.9}$$

Part (a): If U is nonsingular, then (1.5) is equivalent to

$$\begin{pmatrix} F & 0 \\ -Q & I \end{pmatrix}\begin{pmatrix} I \\ X \end{pmatrix} = \begin{pmatrix} I & \Gamma \\ 0 & F^* \end{pmatrix}\begin{pmatrix} I \\ X \end{pmatrix}\tilde{\Lambda}, \tag{1.10}$$

where $X = VU^{-1}$ and $\tilde{\Lambda} = U\Lambda U^{-1}$. Assumption (γ) implies that $I + \Gamma X$ is nonsingular. From $F = (I + \Gamma X)\tilde{\Lambda}$ and $-Q + X = F^*X\tilde{\Lambda}$ follows

$$-Q + X = F^*X(I + \Gamma X)^{-1}F. \tag{1.11}$$

According to (1.9) the equations (1.3) and (1.11) are the same.

Part (b): If X is a solution of (1.3), then we have

$$-Q + X = F^*X\left[I - G(I + G^*XG)^{-1}G^*X\right]F.$$

Using (1.9) again, we see that (1.11) holds and (1.10) is satisfied with $\tilde{\Lambda} = (I + \Gamma X)^{-1}F$. It is easy to verify (1.8). ■

Let J be the $2n \times 2n$ matrix given by

$$J = J_n = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}.$$

As Q and Γ in (1.4) are hermitian, we have

$$MJM^* = LJL^*. \tag{1.12}$$

A nonsingular pencil $M - zL$ with the property (1.12) is called *symplectic*. If L is nonsingular, then $M - zL$ is equivalent to $L^{-1}M - zI$, and $S = L^{-1}M$ satisfies $SJS^* = J$. By a slight abuse of terminology we call such an S a *(complex) symplectic* matrix and define the subgroup Sp of $\mathrm{Gl}(2n, \mathbb{C})$ by

$$\mathrm{Sp} = \{S \mid SJS^* = J\}.$$

LEMMA 1.2. *Let P and R be two nonsingular $2n \times 2n$ matrices. For all symplectic pencils $M - zL$ the equivalent pencil $P(M - zL)R$ is also symplectic if and only if cR is a symplectic matrix for some $c \in \mathbb{C}$.*

Proof. If $S \in \mathrm{Sp}$, then $I - zS$ is a symplectic pencil. Assume that the pencil $P(I - zS)R$ is symplectic for all $S \in \mathrm{Sp}$. Then $R^{-1}SR \in \mathrm{Sp}$ for all

$S \in \mathrm{Sp}$, i.e., R is in the normalizer of Sp in $\mathrm{Gl}(2n, \mathbb{C})$. Hence [8] we have $cR \in \mathrm{Sp}$ for some nonzero number c. The "if" part of the lemma is obvious. ■

DEFINITION. Two symplectic pencils $M - zL$ and $M' - zL'$ are said to be *symplectically equivalent* if there is a nonsingular matrix P and a symplectic matrix R such that

$$M' - zL' = P(M - zL)R.$$

Note that

$$R = \begin{pmatrix} I & 0 \\ X & I \end{pmatrix}$$

is symplectic if and only if X is hermitian. Hence the pencils $M - zL$ and

$$\begin{pmatrix} \Lambda & 0 \\ 0 & I \end{pmatrix} - z \begin{pmatrix} I & D \\ 0 & \Lambda^* \end{pmatrix}$$

in (1.8) are symplectically equivalent.

The main feature of symplectic pencils is a pairing of characteristic roots λ and $\bar{\lambda}^{-1}$ and of corresponding elementary divisors. [We say λ is a characteristic root if $\det(M - \lambda L) = 0$.] The following result can be found in [16] and [6]. We give a proof which does not use Jordan chains.

THEOREM 1.3 (See [16, 6]). *Let $M - zL$ be a symplectic pencil. If $\lambda \neq 0$ is a characteristic root and $(z - \lambda)^k$ is a corresponding elementary divisor, then $(z - \bar{\lambda}^{-1})^k$ is also an elementary divisor of $M - zL$. To each elementary divisor of the form z^m corresponds an infinite elementary divisor of degree m.*

Proof. Let S and T be nonsingular such that

$$S(M - zL) = \begin{pmatrix} A - zI & 0 \\ 0 & I - zN \end{pmatrix} T, \tag{1.13}$$

where N is nilpotent. Then $A - zI$ contains the finite and $I - zN$ the infinite elementary divisors of $M - zL$. As $S(M - zL)$ is also a symplectic pencil, it is no loss of generality to assume $S = I$. Let T be partitioned according to (1.13) as

$$T = \begin{pmatrix} T_e \\ T_\infty \end{pmatrix}.$$

Then

$$M = \begin{pmatrix} AT_e \\ T_\infty \end{pmatrix} \quad \text{and} \quad L = \begin{pmatrix} T_e \\ NT_\infty \end{pmatrix},$$

and we have

$$MJM^* = \begin{pmatrix} AT_eJM^* \\ T_\infty JM^* \end{pmatrix}$$

and

$$LJL^* = \begin{pmatrix} T_eJL^* \\ NT_\infty JL^* \end{pmatrix}.$$

Now (1.12) implies

$$AT_eJM^* = T_eJL^*, \qquad T_\infty JM^* = NT_\infty JL^*, \tag{1.14}$$

and (1.14) together with the trivial identities

$$IT_eJM^* = T_eJM^*, \qquad T_\infty JL^* = IT_\infty JL^*$$

yields

$$(A - zI)T_eJM^* = T_eJ(L^* - zM^*)$$

and

$$T_\infty J(L^* - zM^*) = (I - zN)T_\infty JL^*.$$

Hence

$$\begin{pmatrix} A - zI & 0 \\ 0 & I - zN \end{pmatrix} \begin{pmatrix} T_e JM^* \\ T_\infty JL^* \end{pmatrix} = TJ(L^* - zM^*),$$

and the pencil $M - zL$ is equivalent to $(zM - L)^*$. ■

NOTATION. Let $q(z) = \prod(\lambda_\nu - z)$ be a complex polynomial with $q(0) \neq 0$. Put $\tilde{q}(z) = \prod(1 - \bar{\lambda}_\nu z)$. We call $f(z) = z^m q(z)$ an *unmixed polynomial* if all the common zeros λ of q and $\tilde{q}$ satisfy $|\lambda| = 1$. In particular, a polynomial f is unmixed if all its zeros lie in the closed unit disc. A solution X of (1.3) will be called unmixed if the characteristic polynomial of its associated closed-loop matrix

$$F_X = (I + \Gamma X)^{-1} F = F - G(I + G^*XG)^{-1} G^*XF$$

is unmixed. Finally, we say that

$$\det(M - zL) = cz^m q(z)\tilde{q}(z), \quad c \in \mathbb{C} \tag{1.15}$$

is an *unmixed factorization* if q is unmixed.

In the case where the symplectic pencil $M - zL$ has no characteristic roots with modulus 1, the preceding theorem tells us that $M - zL$ is equivalent to a pencil of the form

$$\begin{pmatrix} \Lambda - zI & 0 \\ 0 & I - z\Lambda^* \end{pmatrix}$$

and Λ can be chosen in such a way that $\det(\Lambda - zI)$ and $\det(I - z\Lambda^*)$ have no zeros in common. Then

$$\det(M - zL) = \det(\Lambda - zI)\det(I - z\Lambda^*)$$

is an unmixed factorization. The situation with no unimodular characteristic roots arises when the pair (F, B) is stabilizable and $\binom{F}{P}$ is detectable, i.e. when

$$\operatorname{rank}(F - \lambda I, B) = \operatorname{rank}\begin{pmatrix} F - \lambda I \\ P \end{pmatrix} = n$$

for all λ with $|\lambda| \geqslant 1$ [16].

In the sequel the characteristic roots α of $M - zL$ with $|\alpha| = 1$ will require special attention. We introduce the following condition.

(E) All the elementary divisors of $M - zL$ which belong to characteristic roots α with $|\alpha| = 1$ have even degree.

Obviously (E) is sufficient for the existence of an unmixed factorization of $\det(M - zL)$. Why the condition (E) should be important for the algebraic Riccati equation (1.3) will become clear from the next theorem and Lemma 1.1.

THEOREM 1.4. *If* (E) *holds and* (1.15) *is a given unmixed factorization, then there are nonsingular matrices K and R such that*

$$(M - zL)R = K\begin{pmatrix} \Lambda - zI & -zD \\ 0 & I - z\Lambda^* \end{pmatrix}, \qquad D = D^*, \tag{1.16}$$

and

$$\det(\Lambda - zI) = (-1)^m z^m q(z).$$

If R is partitioned into $n \times n$ blocks

$$R = \begin{pmatrix} U & \cdot \\ V & \cdot \end{pmatrix}, \tag{1.17}$$

then

$$M\begin{pmatrix} U \\ V \end{pmatrix} = L\begin{pmatrix} U \\ V \end{pmatrix}\Lambda \tag{1.18}$$

and $\begin{pmatrix} U \\ V \end{pmatrix}$ *has full rank.*

Proof. For (1.16) it is enough to consider a $2k \times 2k$ Jordan block

$$B = B_{2k}(\alpha) = \begin{pmatrix} \alpha & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \alpha \end{pmatrix}$$

with $|\alpha| = 1$. Then $B - zI$ is equivalent to

$$\begin{pmatrix} B_k(\alpha) - zI & zE \\ 0 & I - zB_k(\alpha)^* \end{pmatrix},$$

where $E = \text{diag}(0,\dots,0,1)$. In order to verify (1.18) let

$$K = \begin{pmatrix} K_1 & \cdot \\ \cdot & \cdot \end{pmatrix}$$

be partitioned according to (1.17). Then (1.16) implies

$$M\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} K_1 \\ \cdot \end{pmatrix}\Lambda \quad \text{and} \quad L\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} K_1 \\ \cdot \end{pmatrix}.$$

Hence we have (1.18). ∎

The matrix R in (1.16) and (1.17) yields a solution of (1.3) if the blocks U and V satisfy the three conditions (α)–(γ) of Lemma 1.1. It turns out in the next lemma that we need not be concerned about the nonsingularity of $I + \Gamma X$. Furthermore, if R in (1.16) can be chosen to be symplectic and if U^{-1} exists, then $X = VU^{-1}$ is necessarily hermitian.

LEMMA 1.5. *Let K and R be two nonsingular matrices such that* (1.16) *holds, and let R be partitioned as in* (1.17). *If U is nonsingular, then $I + \Gamma(VU^{-1})$ is nonsingular.*

Proof. Put $X = VU^{-1}$, $W = \text{block diag}(U^{-1}, U^*)$, and $\hat{K} = KW$. Then

$$(M - zL)\begin{pmatrix} I & \cdot \\ X & \cdot \end{pmatrix} = \hat{K}\begin{pmatrix} \tilde{\Lambda} - zI & -z\tilde{D} \\ 0 & I - z\tilde{\Lambda}^* \end{pmatrix},$$

where $\tilde{\Lambda} = U\Lambda U^{-1}$ and $\tilde{D} = UDU^*$. Hence we can assume in (1.16) that R is given as

$$R = \begin{pmatrix} I & S \\ X & T \end{pmatrix}.$$

From

$$(M-zL)R\begin{pmatrix} I & -S \\ 0 & I \end{pmatrix} = K\begin{pmatrix} I & -\Lambda S \\ 0 & I \end{pmatrix}\begin{pmatrix} I & \Lambda S \\ 0 & I \end{pmatrix}$$
$$\times\begin{pmatrix} \Lambda - zI & -zD \\ 0 & I - z\Lambda^* \end{pmatrix}\begin{pmatrix} I & -S \\ 0 & I \end{pmatrix}$$

follows

$$(M-zL)\begin{pmatrix} I & 0 \\ X & \tilde{T} \end{pmatrix} = C\begin{pmatrix} \Lambda - zI & -z(D+\Lambda S\Lambda^* - S) \\ 0 & I - z\Lambda^* \end{pmatrix},$$

where

$$C = \begin{pmatrix} C_1 & C_{12} \\ \cdot & \cdot \end{pmatrix}$$

is nonsingular, and $\tilde{T} = T - XS$. Therefore

$$\begin{pmatrix} F & 0 \\ -Q & I \end{pmatrix}\begin{pmatrix} I & 0 \\ X & \tilde{T} \end{pmatrix} = \begin{pmatrix} C_1 & C_{12} \\ \cdot & \cdot \end{pmatrix}\begin{pmatrix} \Lambda & 0 \\ 0 & I \end{pmatrix} \tag{1.19}$$

and

$$\begin{pmatrix} I & \Gamma \\ 0 & F^* \end{pmatrix}\begin{pmatrix} I & 0 \\ X & \tilde{T} \end{pmatrix} = \begin{pmatrix} C_1 & C_{12} \\ \cdot & \cdot \end{pmatrix}\begin{pmatrix} I & \cdot \\ 0 & \Lambda^* \end{pmatrix}. \tag{1.20}$$

From (1.19) we obtain $C_{12} = 0$, and (1.20) implies $C_1 = I + \Gamma X$. Thus

$$C = \begin{pmatrix} I + \Gamma X & 0 \\ \cdot & \cdot \end{pmatrix},$$

and $I + \Gamma X$ is nonsingular. ∎

LEMMA 1.6. *Let*

$$R = \begin{pmatrix} U & S \\ V & T \end{pmatrix}$$

be symplectic. If U is nonsingular then $X = VU^{-1}$ is hermitian.

Proof. If R is symplectic then $R^{-1} = -JR^*J$. Therefore

$$R^{-1} = \begin{pmatrix} T^* & -S^* \\ -V^* & U^* \end{pmatrix}$$

and we have

$$-V^*U + U^*V = 0,$$

and if U is nonsingular then

$$-(VU^{-1})^* + VU^{-1} = 0.$$

■

The fact that there exists a matrix R which satisfies (1.16) and which is symplectic will follow from the results of the next section, where we set out to determine normal forms of symplectic pencils. We shall assume that condition (E) holds. It will be shown that under symplectic equivalence a decomposition into blocks of the form

$$\begin{pmatrix} \Lambda - zI & -zD \\ 0 & I - z\Lambda^* \end{pmatrix}, \qquad D = D^*,$$

can be achieved, where

$$\Lambda = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}.$$

In the case of an elementary divisor $(z-\alpha)^{2m}$, $|\alpha| = 1$, we have $D = \pm\operatorname{diag}(0,\ldots,0,1)$; in the case of a pair of elementary divisors $(z-\lambda)^k$, $(1-\bar{\lambda}^{-1}z)^k$, we have $D = 0$.

2. A NORMAL FORM UNDER SYMPLECTIC EQUIVALENCE

As a first step a given symplectic pencil is reduced to two blocks $S - zI$ with $S \in \mathrm{Sp}$ and $M' - zL'$ where $\det(M' - zL') = (-1)^m z^m$.

THEOREM 2.1. *Let* $M - zL$ *be a symplectic pencil such that* $\det(M - zL) = z^m h(z)$, $h(0) \neq 0$. *Then there are matrices* P *and* R *such that*

$$P(M - zL) = \operatorname{blockdiag}(S - zI, N - zI, I - zN^*)\, R \tag{2.1}$$

and

$$RJR^* = \operatorname{blockdiag}(J_{n-2m}, J_{2m}), \tag{2.2}$$

and $N \in \mathbb{C}^{m \times m}$ *is nilpotent and* S *is symplectic. The matrix* S *is determined up to symplectic similarity, the matrix* N *up to similarity.*

Proof. According to Theorem 1.3 the pencil $M - zL$ is equivalent to $\operatorname{blockdiag}(A - zI, N - zI, I - zN^*)$ with a nilpotent $m \times m$ matrix N and a nonsingular matrix A. Since for any nonsingular P the two pencils $M - zL$ and $P(M - zL)$ are symplectically equivalent, it is no loss of generality to assume

$$M - zL = \operatorname{blockdiag}(A - zI, N - zI, I - zN^*)\, H. \tag{2.3}$$

Put $\Pi = HJH^*$. Then $\Pi = -\Pi^*$ and

$$\begin{aligned} &\operatorname{blockdiag}(A, N, I)\, \Pi\, \operatorname{blockdiag}(A^*, N^*, I) \\ &\qquad = \operatorname{blockdiag}(I, I, N^*)\, \Pi\, \operatorname{blockdiag}(I, I, N). \end{aligned} \tag{2.4}$$

Let $\Pi = (\Pi_{ij})$, $i, j = 1,2,3$, be partitioned conforming to (2.3) so that (2.4) becomes

$$\begin{pmatrix} A\Pi_{11} & A\Pi_{12}N^* & A\Pi_{13} \\ \cdot & N\Pi_{22}N^* & N\Pi_{23} \\ \cdot & \cdot & \Pi_{33} \end{pmatrix} = \begin{pmatrix} \Pi_{11} & \Pi_{12} & \Pi_{13}N \\ \cdot & \Pi_{22} & \Pi_{23}N \\ \cdot & \cdot & N^*\Pi_{33}N \end{pmatrix}.$$

It is not difficult to show that

$$\Pi = \begin{pmatrix} \Pi_{11} & 0 & 0 \\ 0 & 0 & \Pi_{23} \\ 0 & -\Pi_{23}^* & 0 \end{pmatrix}.$$

For example, consider the equation $A\Pi_{13} = \Pi_{13}N$. Since A and N have no common eigenvalues, we have $\Pi_{13} = 0$. The nonzero blocks in Π satisfy

$$A\Pi_{11}A^* = \Pi_{11} = -\Pi_{11}^* \tag{2.5}$$

and

$$N\Pi_{23} = \Pi_{23}N. \tag{2.6}$$

Now (2.6) implies

$$\operatorname{blockdiag}\left(I, \Pi_{23}^{-1}, I\right)(M - zL) = \operatorname{blockdiag}(A - zI, N - zI, I - zN^*)\,\tilde{H} \tag{2.7}$$

with $\tilde{H} = \operatorname{blockdiag}(I, \Pi_{23}^{-1}, I)\,H$ and

$$\tilde{H}J\tilde{H}^* = \operatorname{blockdiag}(\Pi_{11}, J_m). \tag{2.8}$$

As usual, the inertia $\operatorname{In} D$ of a hermitian matrix D is the triple of nonnegative integers containing the numbers of positive, negative, and zero eigenvalues of D. Since iJ is congruent to $-iJ$, we have $\operatorname{In}(iJ) = (n, n, 0)$, and (2.8) implies $\operatorname{In}(i\Pi_{11}) = (n - m, n - m, 0)$. Hence

$$G\Pi_{11}G^* = J_{n-m} \tag{2.9}$$

for some nonsingular G. Put $S = GAG^{-1}$. From (2.7) we obtain

$$\operatorname{blockdiag}\left(G, \Pi_{23}^{-1}, I\right)(M - zL) = \operatorname{blockdiag}(S - zI, N - zI, I - zN^*)\,R$$

with $R = \operatorname{blockdiag}(G, I, I)\,\tilde{H}$. Because of (2.5) the matrix S is symplectic, and it is obvious that (2.2) holds.

To prove the last statement of the theorem let us consider a relation

$$P_1(M - zL) = \operatorname{blockdiag}(S_1 - zI, N_1 - zI, I - zN_1^*)\,R_1$$

with properties like those of (2.1). Put $X = R_1R^{-1}$ and $Y = P_1P^{-1}$. Then X and Y are nonsingular and

$$\operatorname{blockdiag}(S_1 - zI, N_1 - zI, I - zN_1^*)\,X$$

$$= Y \operatorname{blockdiag}(S - zI, N - zI, I - zN^*).$$

Since equivalent pencils have the same elementary divisors, the matrices S_1 and S are similar, as are N_1 and N. Arguments like those used above for Π show that

$$X = Y = \operatorname{blockdiag}(X_1, X_2, X_2^{-*}).$$

Hence $S_1X_1 = X_1S$. From

$$RJR^* = R_1JR_1^* = \operatorname{blockdiag}(J_{n-2m}, J_{2m})$$

we obtain $X_1J_{n-2m}X_1^* = J_{n-2m}$. Thus X_1 is a symplectic similarity of S_1 and S. ■

On the right-hand side of (2.1) two distinct pencils appear, namely $S - zI$, $S \in \mathrm{Sp}$, and

$$M' - zL' = \begin{pmatrix} N & 0 \\ 0 & I \end{pmatrix} - z\begin{pmatrix} I & 0 \\ 0 & N^* \end{pmatrix}, \tag{2.10}$$

where $N^m = 0$. We call the matrix S a *symplectic component* and N a *nilpotent component* of $M - zL$. The last part of the proof of Theorem 2.1 with the matrix $X = \operatorname{blockdiag}(X_1, X_2, X_2^{-*})$ shows that two symplectic pencils are symplectically equivalent if and only if their symplectic components are symplectically similar and their nilpotent components are similar.

The pencil (2.10) can easily be reduced further. Let T be a matrix which transforms N into Jordan form,

$$TNT^{-1} = \operatorname{blockdiag}(N_1, \dots, N_r) = \tilde{N},$$

where the Jordan blocks N_i have size $m_i \times m_i$. The matrix $\Sigma = \operatorname{blockdiag}(T^{-1}, T^*)$ is symplectic, and

$$\Sigma^{-1}(M' - zL')\Sigma = \begin{pmatrix} \tilde{N} - zI & 0 \\ 0 & I - z\tilde{N}^* \end{pmatrix}.$$

Hence there exists a matrix R' such that

$$R'J(R')^* = \operatorname{blockdiag}(J_{m_1}, \dots, J_{m_r})$$

and

$$\Sigma^{-1}(M' - zL')R' = \operatorname{block\,diag}(\Omega_1, \ldots, \Omega_r), \tag{2.11}$$

where

$$\Omega_i = \begin{pmatrix} N_i - zI & 0 \\ 0 & I - zN_i^* \end{pmatrix}, \qquad i = 1, \ldots, r. \tag{2.12}$$

The reduction of the pencil $S - zI$ leads to the problem of finding a normal form of a symplectic matrix S under symplectic similarity. Because $S(iJ)S^* = iJ$, one can regard S as an isometry of a space with an indefinite inner product given by the hermitian matrix iJ and refer to [12] or [15]. One could also refer to papers of Cikunov [4, 5] and of Laub and Meyer [11], which deal with real symplectic matrices. The results in [12] are too general for our purposes, since in our case iJ induces an indefinite inner product with signature zero. On the other hand [4, 5, 10] are also not directly applicable in our context, as we are dealing with complex spaces. The approach which is best suited for our purposes is described in [2] and [3]. Although Ciampi [2, 3] studies hamiltonian matrices, only small modifications are necessary to derive results on symplectic matrices along the lines of [2]. Therefore we shall present the basic result of Theorem 2.2 without proof. The case of eigenvalues α of S with $|\alpha| = 1$ will be discussed separately. An appendix (Section 4) describes the inertial invariants introduced in (2.18) in a coordinate free setting.

The notation M^{-*} in (2.15) means $(M^*)^{-1}$.

THEOREM 2.2a. *Let $S \in \mathrm{Sp}$ be a symplectic matrix with elementary divisors $(z - \lambda_i)^{k_i}, (z - \bar{\lambda}_i^{-1})^{k_i}$, $|\lambda_i| \neq 1$, $i = 1, \ldots, r$, and $(z - \alpha_j)^{2m_j}$, $|\alpha_j| = 1$, $j = 1, \ldots, s$. Put*

$$\tilde{J} = \operatorname{block\,diag}(J_{k_1}, \ldots, J_{k_r}, J_{m_1}, \ldots, J_{m_s}). \tag{2.13}$$

Then there exists a matrix R which satisfies

$$R\tilde{J}R^* = \tilde{J}$$

and which transforms S into

$$RSR^{-1} = \operatorname{block\,diag}(C_1, \ldots, C_r, A_1, \ldots, A_s). \tag{2.14}$$

The matrices C_i correspond to elementary divisors $(z-\lambda_i)^{k_i}, (z-\bar{\lambda}_i^{-1})^{k_i}$ and are of the form

$$C_i = \begin{pmatrix} \lambda_i I + N & 0 \\ 0 & (\lambda_i I + N)^{-*} \end{pmatrix}_{2k_i \times 2k_i}. \tag{2.15}$$

The $2m_j \times 2m_j$ matrices A_j are symplectic; they correspond to the elementary divisors $(z-\alpha_j)^{2m_j}$.

REMARK 2.3. *The pencil*

$$\begin{pmatrix} \lambda I + N & 0 \\ 0 & (\lambda I + N)^{-*} \end{pmatrix} - zI \tag{2.16}$$

is symplectically equivalent to

$$\begin{pmatrix} (\lambda I + N) - zI & 0 \\ 0 & I - z(\lambda I + N)^* \end{pmatrix}. \tag{2.17}$$

Proof. Multiply (2.16) on the left by blockdiag$(I, (\lambda I + N)^*)$. ∎

Let the rows of the matrix R in Theorem 2.2 be partitioned conforming to $\tilde{J}$ in (2.13) as

$$R = (R_1', \dots, R_r', R_1, \dots, R_s)^T,$$

and let $g_j \in$ rowspan R_j be such that

$$g_j(S - \alpha_j I)^{2m_j - 1} \neq 0.$$

Put

$$a_j = \alpha_j^{2m_j - 1} g_j J \left[g_j (S - \alpha_j I)^{2m_j - 1} \right]^*. \tag{2.18}$$

It is shown in the appendix that $a_j \in \mathbb{R}$ and $a_j \neq 0$. Set $\epsilon_j = \operatorname{sign} a_j$. The numbers $\epsilon_j \in \{1, -1\}$, $j = 1, \dots, s$, which are associated to the elementary divisors $(z-\alpha_j)^{2m_j}$ are called the *inertial invariants* of S. Up to ordering of its entries the tuple $(\epsilon_1, \dots, \epsilon_s)$ is independent of the choice of R.

THEOREM 2.2b. *Two symplectic matrices are symplectically similar if and only if they are similar and have the same inertial invariants.*

Because of Theorem 2.1 the preceding result can be extended to pencils. Let us define the inertial invariants of a symplectic pencil to be those of its symplectic component.

THEOREM 2.3. *Two symplectic pencils are symplectically equivalent if and only if they have the same (finite and infinite) elementary divisors and the same inertial invariants.*

We now focus on a $2m \times 2m$ matrix A which is assumed to be symplectic with a single elementary divisor $(z-\alpha)^{2m}$, $|\alpha|=1$, and inertial invariant ϵ. By a Cayley transformation we pass to hamiltonian matrices. Put

$$B=(A-\alpha I)(A+\alpha I)^{-1}; \tag{2.19}$$

then $BJ=-JB^*$, i.e., B is hamiltonian, and B is nilpotent with minimal polynomial z^{2m}. Let y be a row vector such that $yB^{2m-1}\neq 0$. Then

$$2^{2m-1}yJ(yB^{2m-1})^* = \alpha^{2m-1}yJ\left[y(S-\alpha I)^{2m-1}\right]^* = \epsilon|a|$$

for some nonzero $a\in\mathbb{R}$. It is known (see e.g. [11, 18]) that there exists a symplectic matrix R such that

$$RBR^{-1}=\begin{pmatrix} N & \epsilon\Delta \\ 0 & -N^* \end{pmatrix} =: \tilde{B},$$

where

$$N=\begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix} \tag{2.20}$$

and

$$\Delta=\operatorname{diag}(0,\ldots,0,1).$$

From (2.19) we obtain $A = \alpha(I+B)(I-B)^{-1}$. Hence

$$RAR^{-1} = \alpha(I+\tilde{B})(I-\tilde{B})^{-1} =: \tilde{A}.$$

An easy calculation shows that

$$\tilde{A} = \begin{pmatrix} \alpha(I+N)(I-N)^{-1} & 2\alpha\epsilon(I-N)^{-1}\Delta(I+N^*)^{-1} \\ 0 & \alpha(I-N^*)(I+N^*)^{-1} \end{pmatrix}.$$

Let T be a nonsingular matrix such that

$$T\alpha(I+N)(I-N)^{-1}T^{-1} = \alpha I + N. \tag{2.21}$$

Put

$$P = \operatorname{block\,diag}(T, T^{-*}). \tag{2.22}$$

Then P is symplectic and

$$P\tilde{A}P^{-1} = \begin{pmatrix} \alpha I + N & \epsilon V \\ 0 & (\alpha I + N)^{-*} \end{pmatrix},$$

where

$$V = 2T(I-N)^{-1}\Delta(I-N)^{-*}T^*(\alpha I+N)^{-*}. \tag{2.23}$$

Since $\Delta = (0,\ldots,0,1)^T(0,\ldots,0,1)$, we can write V as a dyadic product $V = vv^*$ with

$$v = (v_0,\ldots,v_{r-1})^T = \sqrt{2}\,T(I-N)^{-1}(0,\ldots,0,1)^T.$$

Note that $v_{r-1} \neq 0$. Otherwise both $(1,0,\ldots,0)^T$ and $(0,\ldots,0,1)^T$ would be right eigenvectors of $P\tilde{A}P^{-1}$. But rank $(\tilde{A} - \alpha I) \leqslant 2m-2$ is impossible, as A is nonderogatory. Because $v_{r-1} \neq 0$, there exists a nonsingular matrix $F = \sum f_i N^i$ such that

$$Fv = (0,\ldots,0,1)^T.$$

Since F commutes with N, we can replace T by FT in (2.21), (2.22), (2.23), which yields $V = \Delta(\alpha I + N)^{-*}$.

THEOREM 2.4. *Let A be a symplectic $2m \times 2m$ matrix such that $(z-\alpha)^{2m}$, $|\alpha| = 1$, is the only elementary divisor of A, and ϵ is its inertial invariant. Then there exists a symplectic matrix R such that*

$$RAR^{-1} = \begin{pmatrix} \alpha I + N & \epsilon\Delta(\alpha I + N)^{-*} \\ 0 & (\alpha I + N)^{-*} \end{pmatrix} =: A_0,$$

where N is an upper triangular Jordan block and $\Delta = \operatorname{diag}(0, \dots, 0, 1)$. The pencil $A - zI$ is symplectically equivalent to

$$\begin{pmatrix} (\alpha I + N) - zI & z\epsilon\Delta \\ 0 & I - z(\alpha I + N)^* \end{pmatrix}. \tag{2.24}$$

Proof. It is easy to verify that the product

$$\begin{pmatrix} I & -\epsilon\Delta \\ 0 & (\alpha I + N)^* \end{pmatrix}(A_0 - zI)$$

is equal to (2.24). ■

We sum up the information contained in (2.11) and (2.12), in (2.15) and (2.17), and in (2.24).

THEOREM 2.5. *Let $M - zL$ be a symplectic pencil. Assume that condition (E) is satisfied; let $(z-\lambda_i)^{k_i}, (z - \bar{\lambda}_i^{-1})^{k_i}$, $|\lambda_i| \neq 1$, $i = 1, \dots, r$, and $(z - \lambda_i)^{2k_i}$, $|\lambda_i| = 1$, $i = r+1, \dots, m$, be the elementary divisors of $M - zL$, and let $\epsilon_{r+1}, \dots, \epsilon_m$ be its inertial invariants. Put*

$$\tilde{J} = \operatorname{block\,diag}(J_{k_1}, \dots, J_{k_m})$$

and let

$$H_i = \begin{pmatrix} \lambda_i I + N & 0 \\ 0 & I \end{pmatrix} - z\begin{pmatrix} I & -\epsilon_i\Delta_i \\ 0 & (\lambda_i I + N)^* \end{pmatrix}$$

be of size $2k_i \times 2k_i$ *with*

$$\Delta_i = \begin{cases} 0 & \textit{if} \quad |\lambda_i| \neq 1, \\ \operatorname{diag}(0,\dots,0,1) & \textit{if} \quad |\lambda_i| = 1, \end{cases}$$

where the matrix N *is a nilpotent Jordan block as in* (2.20). *Then there exists a nonsingular matrix* P *and a matrix* R *which satisfies* $R\tilde{J}R^* = \tilde{J}$ *such that*

$$P(M - zL)R = \operatorname{block\,diag}(H_1,\dots,H_m).$$

A rearrangement of the blocks which make up the matrices H_i, R, and P yields a sharper version of Theorem 1.4. In view of Lemma 1.6 and of condition (β) of Lemma 1.1, such a refinement is an important tool in our approach to solve the algebraic Riccati equation (1.3). Based on the results of this article and on those of [19], a subsequent paper will deal with the existence of stabilizing and unmixed solutions of (1.3).

THEOREM 2.6. *Let* $M - zL$ *be a symplectic pencil and let* $\det(M - zL) = cz^m q(z)\tilde{q}(z)$ *be a given unmixed factorization. If* (E) *holds, then* $M - zL$ *is symplectically equivalent to a pencil of the form*

$$\begin{pmatrix} \Lambda - zI & -zD \\ 0 & I - z\Lambda^* \end{pmatrix}, \qquad D = D^*,$$

where $\det(\Lambda - zI) = (-1)^m z^m q(z)$.

3. THE RICCATI PENCIL

In this section we return to the $2n \times 2n$ pencil

$$M - zL = \begin{pmatrix} F & 0 \\ Q & I \end{pmatrix} - z\begin{pmatrix} I & \Gamma \\ 0 & F^* \end{pmatrix} \tag{3.1}$$

which is related to the algebrai[illegible] equation (1.3). The assumptions are $\Gamma = GG^* \geqslant 0$ and $Q \geqslant 0$. Let [illegible] 1) a *Riccati pencil*. We are going to discuss some features related [illegible] odular characteristic roots of $M - zL$.

LEMMA 3.1. *Suppose* (1.3) *has a solution X. Then* $\lambda \neq 0$ *is a characteristic root of the Riccati pencil* (3.1) *if and only if* λ *or* $\bar{\lambda}^{-1}$ *is an eigenvalue of*

$$F_X = (I + \Gamma X)^{-1} F = F - G(I + G^* X G)^{-1} G^* X F. \tag{3.2}$$

If α *is a characteristic root with* $|\alpha| = 1$ *then* α *is also an eigenvalue of* F.

Proof. If X is a solution, then (1.8) and (1.7) yield

$$\det(M - zL) = c \det(F_X - zI) \det(I - zF_X^*).$$

Hence $|\alpha| = 1$ and $\det(M - \alpha L) = 0$ imply that α is an eigenvalue of F_X. It is well known (see e.g. [6]) that X satisfies the discrete-time Lyapunov equation

$$X - F_X^* X F_X = Q + F^* X G (I + G^* X G)^{-2} G^* X F.$$

Let y be an eigenvector of F_X such that $F_X y = \alpha y$, $y \neq 0$, $|\alpha| = 1$. Then

$$0 = y^*(X - F_X^* X F_X) y = y^* Q y + y^* F^* X G (I + G^* X G)^{-2} G^* X F y.$$

Thus $Qy = 0$ and $G^* X F y = 0$, and from (3.2) we obtain $F_X y = Fy = \alpha y$. ∎

In Lemma 3.2 we state a result for matrices of rational functions. It will be used in the special case of the pencil

$$\begin{pmatrix} \Lambda - zI & -zD \\ 0 & I - z\Lambda^* \end{pmatrix}. \tag{3.3}$$

Let A be the ring of those complex rational function which have no pole at α. The local ring A is a principal ideal domain with maximal ideal $(z - \alpha)$. If F and S are two $n \times n$ matrices over A which have the same elementary divisors, we write $F \overset{\alpha}{\sim} S$. In other words, $UFV = S$ for some matrices U and V which are invertible in $A^{n \times n}$.

For a complex polynomial $f(z) =$ [illegible] define $\bar{f}$ by $\bar{f}(z) = \sum \bar{f}_\nu z^\nu$, and for $h = f/g \in \mathbb{C}(z)$ put $\bar{h} = \bar{f}/\bar{g}$. If $H = (h_{ij}(z))$ is an $n \times n$ matrix of rational functions, then $\hat{H}$ is defined by $\hat{H} = (\bar{h}_{ji}(z^{-1}))$.

LEMMA 3.2. *Let $\alpha \in \mathbb{C}$ be such that $|\alpha| = 1$. Assume that $F \in A^{n\times n}$ and $\Gamma \in A^{n\times n}$ have the properties*

(i) $F \overset{\alpha}{\sim} \operatorname{diag}(1,\dots,1,(z-\alpha)^{k_1},\dots,(z-\alpha)^{k_r})$ *and* $0 < k_1 \leqslant \cdots \leqslant k_r$,
(ii) $\Gamma = \hat{\Gamma}$,
(iii) $\Gamma(\alpha) \leqslant 0$.

Put

$$P = \begin{pmatrix} F & z\Gamma \\ 0 & -z\hat{F} \end{pmatrix}.$$

Then we have

$$P \overset{\alpha}{\sim} \operatorname{diag}\big(1,\dots,1,(z-\alpha)^{2k_1},\dots,(z-\alpha)^{2k_r}\big)$$

if and only if

$$\operatorname{rank}\big(F(z), \quad z\Gamma(z)\big) = n \qquad \textit{for} \quad z = \alpha. \tag{3.4}$$

Proof. If U and V are invertible in $A^{n\times n}$ such that $UFV = S = \operatorname{diag}(1,\dots,1,(z-\alpha)^{k_1},\dots,(z-\alpha)^{k_r})$ then

$$\begin{pmatrix} U & 0 \\ 0 & \hat{V} \end{pmatrix} P \begin{pmatrix} V & 0 \\ 0 & \hat{U} \end{pmatrix} = \begin{pmatrix} S & zU\Gamma\hat{U} \\ 0 & -z\hat{S} \end{pmatrix}.$$

For S and $U\Gamma\hat{U}$ conditions of the form (i)–(iii) are satisfied. For $z = \alpha$ we have $\operatorname{rank}(F(z), z\Gamma(z)) = \operatorname{rank}(S(z), zU(z)\Gamma(z)\hat{U}(z))$. Hence we can assume without loss of generality $F = \operatorname{block diag}(I, F_2)$, where

$$F_2(z) = \operatorname{diag}\big((z-\alpha)^{k_1},\dots,(z-\alpha)^{k_r}\big).$$

Let Γ be partitioned conforming to F:

$$\Gamma = \begin{pmatrix} \Gamma_1 & \Gamma_{12} \\ \Gamma_{21} & \Gamma_2 \end{pmatrix}.$$

Then $P \overset{\alpha}{\sim} \operatorname{block\,diag}(I, P_2)$ with

$$P_2 = \begin{pmatrix} F_2 & z\Gamma_2 \\ 0 & -z\hat{F}_2 \end{pmatrix}.$$

We want to show that (3.4) is equivalent to $\Gamma_2(\alpha) < 0$. Because $F_2(\alpha) = 0$, it is obvious that (3.4) holds if and only if

$$\operatorname{rank}\big(z\Gamma_{21}(z), \quad z\Gamma_2(z)\big) = r \qquad \text{for} \quad z = \alpha. \tag{3.5}$$

Suppose $\Gamma_2(\alpha)$ is singular and $\Gamma_2(\alpha)b = 0$, $b \neq 0$. Then

$$(0 \quad b^*)\Gamma(\alpha)\begin{pmatrix} 0 \\ b \end{pmatrix} = b^*\Gamma_2(\alpha)b = 0,$$

and (iii) implies

$$(0 \quad b^*)\Gamma(\alpha) = b^*\big(\Gamma_{21}(\alpha), \quad \alpha\Gamma_2(\alpha)\big) = 0,$$

which is incompatible with (3.5). On the other hand, if $\Gamma_2(\alpha)$ is nonsingular, then we have (3.5). Therefore it is no loss of generality to continue under the assumption that $r = n$ and $F = F_2 = \operatorname{diag}((z-\alpha)^{k_1}, \dots, (z-\alpha)^{k_n})$, $0 < k_1 \leqslant \cdots \leqslant k_n$. Let us consider first the hypothesis that (3.4) or equivalently $\Gamma(\alpha) < 0$ holds. Then $\Gamma \overset{\alpha}{\sim} \Gamma^{-1} \overset{\alpha}{\sim} I$, and elementary α-equivalent transformations yield

$$P \overset{\alpha}{\sim} \begin{pmatrix} 0 & I \\ \hat{F}\Gamma^{-1}F & 0 \end{pmatrix}.$$

Because $\Gamma(\alpha) < 0$, all principal minors of $T = \hat{F}\Gamma^{-1}F$ are nonzero. Since F is diagonal, the gcd of all $\nu \times \nu$ minors of T is $(z-\alpha)^{e_\nu}$, and $e_\nu = \sum_{i=1}^{\nu} 2k_i$. Therefore $T \overset{\alpha}{\sim} \hat{F}F$ and

$$P \overset{\alpha}{\sim} \operatorname{diag}\big(1, \dots, 1, (z-\alpha)^{2k_1}, \dots, (z-\alpha)^{2k_n}\big). \tag{3.6}$$

Conversely, assume now that (3.6) holds. Since n elementary divisors of P are equal to 1, there must be an $n \times n$ submatrix H of P such that $\det H(\alpha) \neq 0$. But $F(\alpha) = 0$; therefore $H = z\Gamma$. From (iii) follows $\Gamma(\alpha) < 0$.

■

Condition (E) is closely related to the B-controllability of eigenvalues α of F. Recall that $\Gamma = BR^{-1}B^*$ and $R > 0$. Hence

$$\operatorname{rank}(\,F - \alpha I, \quad \Gamma\,) = \operatorname{rank}(\,F - \alpha I, \quad B\,).$$

THEOREM 3.3. *Let* (1.3) *have a positive semidefinite solution X, and let $(z-\alpha)^{k_i}$, $i = 1, \dots, r$, be the elementary divisors of $F_X = (I + \Gamma X)^{-1}F$ which belong to an eigenvalue α with $|\alpha| = 1$. Then*

$$M - zL \overset{\alpha}{\sim} \operatorname{diag}\big(1, \dots, 1, (z-\alpha)^{2k_1}, \dots, (z-\alpha)^{2k_r}\big) \tag{3.7}$$

is equivalent to

$$\operatorname{rank}(\,F - \alpha I, \quad \Gamma\,) = n. \tag{3.8}$$

If (3.8) *holds for all α with $|\alpha| = 1$, then the Riccati pencil* (3.1) *satisfies condition* (E), *and all inertial invariants of* (3.1) *are equal to* -1.

Proof. According to (1.8) the pencil $M - zL$ is symplectically equivalent to

$$\begin{pmatrix} \Lambda - zI & -zD \\ 0 & I - z\Lambda^* \end{pmatrix}, \tag{3.3}$$

where $\Lambda = F_X$ and $D = D^* = (I + \Gamma X)^{-1}\Gamma \geqslant 0$. Lemma 3.2 applied to the pencil (3.3) shows that (3.7) is equivalent to

$$\operatorname{rank}(\,\Lambda - \alpha I, \quad -\alpha D\,) = n.$$

The definitions of $\Lambda = F_X$ and of D yield

$$\operatorname{rank}(\Lambda - \alpha I, \quad -\alpha D) = \operatorname{rank}(\,F - \alpha I, \quad \Gamma\,).$$

If (3.7) holds for all α with $|\alpha| = 1$, then we have (E). Using arguments which lead to Theorem 2.5, one can assume $\Lambda = \operatorname{blockdiag}(\Lambda_1, \dots, \Lambda_m)$ with $\Lambda_i = \lambda_i I + N$ and $D = \operatorname{blockdiag}(D_1, \dots, D_m)$ where $D_i \neq 0$ if $|\lambda_i| \neq 1$. Then $D \geqslant 0$ implies $\epsilon_i = -1$ for all inertial invariants. ■

4. APPENDIX: INERTIAL INVARIANTS

Let V be a complex $2n$-dimensional vector space, and let $\omega: V\times V\to\mathbb{C}$ be an indefinite inner product on V which has the following properties: For all $x, x_1, x_2, y\in V$ and for all $c_1, c_2\in\mathbb{C}$ we have

(1) $\omega(c_1x_1+c_2x_2, y)=c_1\omega(x_1,y)+c_2\omega(x_2,y)$,
(2) $\omega(y,x)=-\overline{\omega(x,y)}$,
(3) $\omega(\cdot,\cdot)$ is nondegenerate, i.e., if $\omega(x,y_0)=0$ for all $x\in V$ then $y_0=0$.

A subspace U of V is called *isotropic* if $\omega(u,w)=0$ for all $u,w\in U$. If

(4) $V=U_1\oplus U_2$ and U_1, U_2 are isotropic,

then $i\omega$ is a hermitian form on V with signature 0. We call the pair (V,ω) an *s_0-space* if (1)–(4) are satisfied.

For our definition the Lagrange splitting in (4) is essential. In the case of real symplectic spaces (which involve nondegenerate skew symmetric bilinear forms) condition (4) is a consequence of (1)–(3).

Let

$$B=(b_1,\dots,b_{2n}) \tag{4.1}$$

be a basis of V, and let

$$\Omega=\left(\omega(b_i,b_j)\right),\qquad i,j=1,\dots,2n, \tag{4.2}$$

be the matrix of the inner product corresponding to B. Then $\Omega=-\Omega^*$ and Ω is nonsingular. Condition (4) is equivalent to the existence of a *symplectic basis* $\Sigma=(s_1,\dots,s_{2n})$ for which $(\omega(s_i,s_j))=J$ holds.

A subspace U of V is called nondegenerate [an s_0-subspace] if $\omega|_U$ is nondegenerate [$(U,\omega|_U)$ is an s_0-subspace]. In the following it will often be convenient to denote the inner product by $[\cdot,\cdot]$, so that $[x,y]=\omega(x,y)$.

Let σ be an isometry of V, i.e. an endomorphism of V which satisfies $[x,y]=[\sigma x,\sigma y]$ for all $x,y\in V$, and let B and Ω be as in (4.1) and in (4.2). If $S=(s_{\kappa\nu})$ is the matrix representation of σ with respect to the basis B such that $\sigma(b_\kappa)=\sum_{\nu=1}^{2n}s_{\kappa\nu}b_\nu$, $\kappa=1,\dots,2n$, then $S\Omega S^*=\Omega$. In the case of a symplectic basis (with $\Omega=J$) the matrix S is symplectic.

The decomposition (4.3) below leads to the block diagonal form (2.14).

THEOREM 4.1. *Let (V, ω) be an s_0-space, and let σ be an isometry of V with elementary divisors $(z-\lambda_i)^{k_i}, (z-\bar{\lambda}_i^{-1})^{k_i}$, $|\lambda_i| \neq 1$, $i = 1, \ldots, r$, and $(z-\alpha_j)^{t_j}$, $|\alpha_j| = 1$, $j = 1, \ldots, s$. Then there exists an orthogonal decomposition of V into σ-invariant subspaces*

$$V = W_1 \oplus \cdots \oplus W_r \oplus U_1 \oplus \cdots \oplus U_s \tag{4.3}$$

with the following properties: W_i is an s_0-space, $W_i = W_{i1} \oplus W_{i2}$ with $\dim W_{i1} = \dim W_{i2} = k_i$, *both W_{i1} and W_{i2} are isotropic and σ-invariant, $\sigma|_{W_{i1}}$ and $\sigma|_{W_{i2}}$ have minimal polynomial $(z-\lambda_i)^{k_i}$ and $(z-\bar{\lambda}_i^{-1})^{k_i}$ respectively, the subspaces U_j are nondegenerate,* $\dim U_j = t_j$, *and $(z-\alpha_j)^{t_j}$ is the minimal polynomial of $\sigma|_{U_j}$.*

Let $|\alpha| = 1$ be an eigenvalue of α, and assume that $(z-\alpha)^t$ appears p times in the sequence of elementary divisors of σ. Let

$$U = U_1^{\alpha} \oplus \cdots \oplus U_p^{\alpha}$$

be the direct sum of those p subspaces in (4.3) which correspond to $(z-\alpha)^t$.

LEMMA 4.2. *Put $\bar{U} = U/(\sigma - \alpha)U$, and define*

$$\gamma(\bar{x}, \bar{y}) = \alpha^{t-1}\left[x, (\sigma - \alpha)^{t-1} y\right]$$

for $\bar{x} = x + (\sigma - \alpha)U$, $\bar{y} = y + (\sigma - \alpha)U$, $x, y \in U$. Then $\gamma : \bar{U} \times \bar{U} \to \mathbb{C}$ is a well-defined nondegenerate form on $\bar{U}$ which satisfies

$$\gamma(c_1\bar{x}_1 + c_2\bar{x}_2, \bar{y}) = c_1\gamma(\bar{x}_1, \bar{y}) + c_2\gamma(\bar{x}_2, \bar{y})$$

and

$$\gamma(\bar{y}, \bar{x}) = (-1)^t \overline{\gamma(\bar{x}, \bar{y})}. \tag{4.4}$$

Proof. Since σ is an isometry, we have

$$[x, (\sigma - \alpha) y] = [(\sigma^{-1} - \bar{\alpha}) x, y],$$

and $|\alpha| = 1$ yields

$$[x,(\sigma-\alpha)y] = -\bar{\alpha}[(\sigma-\alpha)x,\sigma y].$$

Hence

$$\left[(\sigma-\alpha)^i x,(\sigma-\alpha)^k y\right] = 0 \qquad \text{if} \quad i+k \geqslant t, \tag{4.5}$$

and $x,y \in U$. From (4.5) it follows that γ is well defined. In order to show that γ is nondegenerate, let σ and ω be restricted to U and to $U \times U$, respectively. Let $\bar{v} = v + (\sigma-\alpha)U$ be such that $\gamma(\bar{u},\bar{v}) = 0$ for all $\bar{u} \in \overline{U}$. Then $[u,(\sigma-\alpha)^{t-1}v] = 0$ for all $u \in U$. Since U is nondegenerate, we have $(\sigma-\alpha)^{t-1}v = 0$. Therefore $v \in \mathrm{Ker}(\sigma-\alpha)^{t-1} = \mathrm{Im}(\sigma-\alpha) = (\sigma-\alpha)U$ and $\bar{v} = \bar{0}$. The property (4.4) can be verified as follows:

$$\begin{aligned}
\gamma(\bar{y},\bar{x}) &= \alpha^{t-1}\left[y,(\sigma-\alpha)^{t-1}x\right] \\
&= \alpha^{t-1}(-\bar{\alpha})^{t-1}\left[(\sigma-\alpha)^{t-1}y,\sigma^{t-1}x\right] \\
&= (-1)^{t-1}\left[(\sigma-\alpha)^{t-1}y,(\sigma-\alpha+\alpha)^{t-1}x\right] \\
&= (-1)^{t-1}\left[(\sigma-\alpha)^{t-1}y,\alpha^{t-1}x\right] = (-1)^t\bar{\alpha}^{t-1}\,\overline{\left[x,(\sigma-\alpha)^{t-1}y\right]} \\
&= (-1)^t\,\overline{\gamma(\bar{x},\bar{y})}\,.
\end{aligned}$$

■

Only the case where the degree of $(z-\alpha)^t$ is even will be of interest here.

Lemma 4.3. *Let Y be a nondegenerate σ-invariant subspace of V such that* $\dim Y = 2m$ *and* $(z-\alpha)^{2m}$, $|\alpha| = 1$, *is the only elementary divisor of* $\sigma|_Y$. *Then Y is an s_0-subspace of V.*

Proof. Since $(z-\alpha)^{2m}$ is the minimal polynomial of σ on Y, there exists a vector $x_0 \in Y$ such that $(\sigma-\alpha)^{2m-1}x_0 \neq 0$. Define $x_i = (\sigma-\alpha)^i x_0$, $i = 0,\dots,2m-1$, and put

$$\Omega = ([x_i,x_k]), \qquad i,k = 0,\dots,2m-1.$$

From $[x_i, x_k] = [(\sigma - \alpha)^i x_0, (\sigma - \alpha)^k x_0]$ and (4.5) it follows that $[x_i, x_k] = 0$ if $i + k \geqslant 2m$. We know that $\Omega = -\Omega^*$. Hence

$$\Omega = \begin{pmatrix} M & Q^* \\ -Q & 0 \end{pmatrix},$$

where $M = -M^*$ and Q is nonsingular. Put

$$T = \begin{pmatrix} I & \frac{1}{2}MQ^{-1} \\ 0 & I \end{pmatrix}.$$

Then

$$T\Omega T^* = \begin{pmatrix} 0 & Q^* \\ -Q & 0 \end{pmatrix},$$

which shows that Y is an s_0-subspace. ■

If $t = 2m$, then the form γ in (4.4) is hermitian. Its inertia is completely determined by U. The inertial invariants introduced via (2.18) are now derived from a matrix representation of γ.

LEMMA 4.5. *Let U, $\overline{U}$, and γ be as in Lemma* 4.3, *and assume* $t = 2m$. *Let $g_i \in U_i^\alpha$, $i = 1, \dots, p$, be vectors such that $(\sigma - \alpha)^{2m-1} g_i \neq 0$. Put $\bar{g}_i = g_i + (\sigma - \alpha)U$. Then*

$$\gamma(\bar{g}_i, \bar{g}_i) = \alpha^{2m-1}\left[g_i, (\sigma - \alpha)^{2m-1} g_i\right] = \epsilon_i |c_i|^{-2}, \tag{4.6}$$

where $\epsilon_i \in \{1, -1\}$ and $c_i \neq 0$. Furthermore, $\overline{G} = (c_1 \bar{g}_1, \dots, c_p \bar{g}_p)$ is a basis of $\overline{U}$, and the matrix of γ with respect to $\overline{G}$ is given by $\operatorname{diag}(\epsilon_1, \dots, \epsilon_p)$.

Proof. Since γ is a nondegenerate hermitian form on $U_i^\alpha / (\sigma - \alpha) U_i^\alpha$ we see that sign ϵ_i in (4.6) is independent of the choice of g_i. ■

REFERENCES

1 J. Ackermann, *Sampled-Data Control Systems*, Springer-Verlag, Berlin, 1985.
2 A. Ciampi, Classical Hamiltonian Linear Systems, Queen's Papers in Pure and Appl. Math. 31, Queen's Univ., Kingston, Ont., 1972.
3 A. Ciampi, Classification of Hamiltonian linear systems, *Indiana Univ. Math. J.* 23:513–526 (1973).

4 I. K. Cikunov, On the structure of isometric transformations of a symplectic or orthogonal vector space (in Russian), *Ukrain. Mat. Zh.* 18:79–93 (1966).
5 I. K. Cikunov, On a class of isometric transformations of a symplectic or orthogonal vector space (in Russian), *Ukrain. Mat. Zh.* 18:122–127 (1966).
6 C. E. de Souza, M. R. Gevers, and G. C. Goodwin, Riccati equations in optimal filtering of nonstabilizable systems having singular state transition matrices, *IEEE Trans. Automat. Control* 31:831–838 (1986).
7 G. H. Golub and C. F. van Loan, *Matrix Computations*, North Oxford Academic, Oxford, 1983.
8 G. Köhler, private communication.
9 V. Kučera, The discrete Riccati equation of optimal control, *Kybernetica* 8:430–442 (1972).
10 H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*, Wiley-Interscience, New York, 1972.
11 A. J. Laub and K. Meyer, Canonical forms for symplectic and hamiltonian matrices, *Celestial Mech.* 9:213–328 (1974).
12 A. I. Malcev, *Foundations of Linear Algebra*, Freeman, San Francisco, 1963.
13 V. Mehrmann, The Linear Quadratic Control Problem: Theory and Numerical Algorithms, Habilitationsschrift, Univ. Bielefeld, 1987.
14 V. Mehrmann, Linear algebra problems in linear quadratic optimal control, *Linear Algebra Appl.* 121:291–331 (1989).
15 J. Milnor, On isometries of inner product spaces, *Invent. Math.* 8:83–97 (1969).
16 T. Pappas, A. J. Laub, and N. R. Sandell, On the numerical solution of the discrete-time algebraic Riccati equation, *IEEE Trans. Automat. Control* 25:631–641 (1980).
17 A. P. Sage, *Optimal Systems Control*, Prentice-Hall, Englewood Cliffs, N.J. 1968.
18 H. K. Wimmer, The algebraic Riccati equation without complete controllability, *SIAM J. Algebraic Discrete Methods* 3:1–12 (1982).
19 H. K. Wimmer, Stabilizing and unmixed solutions of the discrete time algebraic Riccati equation, in *Workshop on The Riccati Equation in Control, Systems, and Signals, Lecture Notes* (S. Bittanti, Ed.), Pitagora Editrice, Bologna, 1989, pp. 95–98.

Received 5 February 1990; final manuscript accepted 16 August 1990