

Matrix quadratic equations

W.A. Coppel

Matrix quadratic equations have found the most diverse applications. The present article gives a connected account of their theory, and contains some new results and new proofs of known results.

1. Introduction

An autonomous linear hamiltonian system of differential equations has the form

$$(1) \quad x' = J^{-1}Hx ,$$

where $x \in R^{2n}$ and

$$J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} , \quad H = \begin{pmatrix} A & B^* \\ B & C \end{pmatrix}$$

are constant $2n \times 2n$ real matrices which are respectively skew-symmetric, symmetric.¹ It is well-known, and easily established, that there is a close connection between the solutions of (1) and the solutions of the $n \times n$ matrix Riccati equation

$$(2) \quad R[W] \equiv W' + A + WB + B^*W + WCW = 0 .$$

The constant solutions of (2) are just the solutions of the matrix quadratic equation

$$(3) \quad Q[W] \equiv A + WB + B^*W + WCW = 0 .$$

Largely as a result of this connection, the matrix quadratic equation (3) has found applications in many different fields; for example, optimal

Received 4 February 1974.

¹ The conjugate transpose of a matrix is denoted by an asterisk.

control, stability theory, filtering theory, network theory and differential games. In these applications it is the symmetric solutions of (3) which are of interest. The object of the present article is to give a connected account of the theory of the equation (3). References to previous work are made at various places in the text.

The following identities are easily verified. If W_1 and W_2 are solutions of (3) then their difference $D = W_2 - W_1$ satisfies

$$(4) \quad D(B + CW_1) + (B^* + W_2^*C)D = 0,$$

$$(5) \quad D(B + CW_1) + (B^* + W_1^*C)D + DCD = 0.$$

Conversely, if W_1 is a solution of (3) and if D satisfies (4) or (5), then W_2 is also a solution of (3).

The coefficient matrix

$$(6) \quad M = J^{-1}H = \begin{pmatrix} B & C \\ -A & -B^* \end{pmatrix}$$

has the property

$$(7) \quad JM = -M^*J.$$

It follows that the characteristic polynomial

$$(8) \quad \varphi(s) = \det(sI - M)$$

is an even function of s . If W is a solution of (3) then

$$(9) \quad \varphi(s) = \det(sI - B - CW)\det(sI + B^* + WC),$$

since

$$M \begin{pmatrix} I & 0 \\ W & I \end{pmatrix} = \begin{pmatrix} I & 0 \\ W & I \end{pmatrix} \begin{pmatrix} B + CW & C \\ 0 & -B^* - WC \end{pmatrix}.$$

Finally we recall some standard definitions. A matrix is said to be *stable* if all its eigenvalues have negative real part. The inequality $W_2 \geq W_1$ between two symmetric matrices W_1, W_2 means that $W_2 - W_1$ is non-negative definite. If B is an $n \times n$ matrix and C an $n \times m$ matrix the ordered pair (B, C) is said to be *controllable* if the $n \times m$

block matrix $(C \ BC \ \dots \ B^{n-1}C)$ has rank n . For further discussion of this concept see, for example, Kalman *et al* [3]. If the pair (B, C) is controllable then for any $m \times n$ matrix W the pair $(B+CW, C)$ is also controllable. If $C = C^* \leq 0$ then (B, C) is controllable if and only if

$$(10) \quad \Omega \equiv - \int_0^T e^{-tB} C e^{-tB^*} dt > 0$$

for some, and hence every, $T > 0$.

LEMMA 1. *If (B, C) is controllable and $C \leq 0$ then $\hat{B} = B + C\Omega^{-1}$ is stable.*

Proof. We have

$$\begin{aligned} B\Omega + \Omega B^* &= \int_0^T \left[\frac{d}{dt} e^{-tB} C e^{-tB^*} \right] dt \\ &= e^{-TB} C e^{-TB^*} - C, \end{aligned}$$

and hence

$$\hat{B}\Omega + \Omega \hat{B}^* = e^{-TB} C e^{-TB^*} + C.$$

Let λ be an eigenvalue of \hat{B}^* and ζ a corresponding eigenvector. Then

$$\begin{aligned} (\lambda + \bar{\lambda}) \zeta^* \Omega \zeta &= \zeta^* (\hat{B}\Omega + \Omega \hat{B}^*) \zeta \\ &= \zeta^* (e^{-TB} C e^{-TB^*} + C) \zeta \\ &\leq 0. \end{aligned}$$

Since $\Omega > 0$ it follows that $\text{Re } \lambda \leq 0$. Moreover, if equality holds then $C\zeta = 0$. From the definition of \hat{B} this implies $B^*\zeta = \lambda\zeta$. Hence

$$\Omega \zeta = - \int_0^T e^{-tB} C \zeta e^{-\lambda t} dt = 0.$$

Since $\zeta \neq 0$, this is a contradiction. Hence \hat{B} is stable.

A somewhat different proof of Lemma 1 is given by Lukes [7].

2. Extreme solutions

Throughout this section it will be assumed that $C \leq 0$ and that (B, C) is controllable. Moreover by a solution of (3) we will always mean a *symmetric* solution, and we will assume that (3) has at least one solution

\tilde{W} .

THEOREM 1. *The matrix quadratic equation (3) has maximal and minimal solutions W_+ and W_- ; that is, every solution \tilde{W} satisfies*

$$W_- \leq \tilde{W} \leq W_+ .$$

All eigenvalues of $B + CW_+$ have non-positive real part and all eigenvalues of $B + CW_-$ have non-negative real part.

We give two proofs. The first uses the relation with the Riccati equation (2). Let \tilde{W} be a solution of (3) and put

$$\tilde{B} = B + C\tilde{W} ,$$

$$U(t) = \int_0^t e^{\tau\tilde{B}} C e^{\tau\tilde{B}^*} d\tau .$$

Since $C \leq 0$ and (\tilde{B}, C) is controllable, the symmetric matrix $U(t)$ is a strictly decreasing function of t . In particular, $U(t) > 0$ for $t < 0$. As $t \rightarrow -\infty$, $U^{-1}(t)$ decreases and is bounded below by 0 . Thus $\lim_{t \rightarrow -\infty} U^{-1}(t)$ exists and is non-negative.

The function $U(t)$ is the solution of the linear differential equation

$$(11) \quad U' = \tilde{B}U + U\tilde{B}^* + C$$

which vanishes at $t = 0$. For any $T > 0$ put

$$W_T(t) = \tilde{W} + U^{-1}(t-T) .$$

Using (11), it is easily verified that $W_T(t)$ is a solution of (2) on the interval $[0, T)$. The representation

$$W_T(t) = [\tilde{W}U(t-T) + I]U^{-1}(t-T)$$

shows that $W_T(t)$ is non-singular on a small interval $[T-\delta, T)$ and

$W_T^{-1}(t) \rightarrow 0$ as $t \rightarrow T$. If $V_T(t)$ is the solution of the 'inverse'

Riccati equation

$$V' - C - VB^* - BV - VAV = 0$$

such that $V_T(T) = 0$ it follows that $V_T(t)$ is non-singular and

$V_T^{-1}(t) = W_T(t)$ on $[T-\delta, T)$. This shows that $W_T(t)$ can also be defined independently of \tilde{W} . The existence of \tilde{W} simply guarantees the existence of $W_T(t)$ over the interval $[0, T)$. As $T \rightarrow \infty$, $W_T(t)$ converges uniformly on compact t -intervals to the constant matrix

$$(12) \quad W_+ = \tilde{W} + \lim_{t \rightarrow -\infty} U^{-1}(t).$$

Thus W_+ is a constant solution of the Riccati equation (2); that is, it is a solution of the quadratic equation (3). Moreover $W_+ \geq \tilde{W}$ for any solution \tilde{W} of (3). Since (\tilde{B}, C) is controllable it follows from Lemma 1 that

$$B + CW_T(0) = \tilde{B} - C \left[\int_0^T e^{-\tau \tilde{B}} C e^{-\tau \tilde{B}^*} d\tau \right]^{-1}$$

is a stable matrix. Hence, letting $T \rightarrow \infty$, all eigenvalues of $B + CW_+$ have non-positive real part.

By replacing W by $-W$ we see that (3) also has the minimal solution

$$(13) \quad W_- = \tilde{W} + \lim_{t \rightarrow \infty} U^{-1}(t),$$

and all eigenvalues of $B + CW_-$ have non-negative real part.

The second proof provides a practical algorithm for the determination of the maximal and minimal solutions. It is based on the identity

$$(14) \quad W(B+CW) + (B+CW)^*W - WCW = W(B+C\hat{W}) + (B+C\hat{W})^*W - \hat{W}C\hat{W} + (W-\hat{W})C(W-\hat{W}),$$

valid for arbitrary symmetric matrices C, W, \hat{W} , and on the following simple

LEMMA 2. Suppose $C \leq 0$ and the linear equation

$$(15) \quad WB + B^*W = C$$

has a solution $W \geq 0$. If B has an eigenvalue $\lambda = \mu + i\nu$ with $\mu \geq 0$ and if ζ is a corresponding eigenvector then $C\zeta = 0$.

In fact from (15) we obtain at once

$$\zeta^* C \zeta = 2\mu \zeta^* W \zeta .$$

Since the left side is non-positive and the right side is non-negative they must both be zero. Since $C \leq 0$ this implies $C \zeta = 0$.

We now continue with the second proof of Theorem 1. By Lemma 1 there exists a symmetric matrix W_0 such that $B + CW_0$ is stable. We define a sequence $\{W_\nu\}$ of symmetric matrices inductively by the linear equations

$$(16) \quad W_{\nu+1}(B + CW_\nu) + (B + CW_\nu)^* W_{\nu+1} = W_\nu CW_\nu - A \quad (\nu \geq 0) .$$

Suppose W_ν has been defined and $B + CW_\nu$ is stable. Then (16) has a unique and symmetric solution $W_{\nu+1}$. By the identity (14), with $W = \tilde{W}$ and $\hat{W} = W_\nu$,

$$(17) \quad -A = \tilde{W}(B + CW_\nu) + (B + CW_\nu)^* \tilde{W} - W_\nu CW_\nu + (W_\nu - \tilde{W})C(W_\nu - \tilde{W}) .$$

Adding this to (16) we get

$$(18) \quad (W_{\nu+1} - \tilde{W})(B + CW_\nu) + (B + CW_\nu)^*(W_{\nu+1} - \tilde{W}) = (W_\nu - \tilde{W})C(W_\nu - \tilde{W}) .$$

Therefore

$$\begin{aligned} W_{\nu+1} - \tilde{W} &= - \int_0^\infty e^{t(B + CW_\nu)^*} (W_\nu - \tilde{W})C(W_\nu - \tilde{W}) e^{-t(B + CW_\nu)} dt \\ &\geq 0 . \end{aligned}$$

By the identity (14) also, with $W = W_{\nu+1}$ and $\hat{W} = W_\nu$,

$$(19) \quad W_{\nu+1}(B + CW_{\nu+1}) + (B + CW_{\nu+1})^* W_{\nu+1} - W_{\nu+1} CW_{\nu+1} = -A + (W_{\nu+1} - W_\nu)C(W_{\nu+1} - W_\nu) .$$

Adding (17), with ν replaced by $\nu + 1$, we get

$$(20) \quad (W_{\nu+1} - \tilde{W})(B + CW_{\nu+1}) + (B + CW_{\nu+1})^*(W_{\nu+1} - \tilde{W}) = (W_{\nu+1} - W_\nu)C(W_{\nu+1} - W_\nu) + (\tilde{W} - W_{\nu+1})C(\tilde{W} - W_{\nu+1}) .$$

Suppose $B + CW_{\nu+1}$ had an eigenvalue λ with non-negative real part, and let ζ be a corresponding eigenvector. Then, by Lemma 2,

$$C(W_{\nu+1} - W_{\nu})\zeta = C(\tilde{W} - W_{\nu+1})\zeta = 0 .$$

Therefore

$$(B + CW_{\nu})\zeta = (B + CW_{\nu+1})\zeta = \lambda\zeta .$$

Since $B + CW_{\nu}$ is stable, this is a contradiction. Thus $B + CW_{\nu+1}$ is stable.

Subtracting (16) from (19) with ν replaced by $\nu - 1$ we get

$$(21) \quad (W_{\nu} - W_{\nu+1})(B + CW_{\nu}) + (B + CW_{\nu})^*(W_{\nu} - W_{\nu+1}) = (W_{\nu-1} - W_{\nu})C(W_{\nu-1} - W_{\nu}) \quad (\nu \geq 1) .$$

Therefore

$$W_{\nu} - W_{\nu+1} = - \int_0^{\infty} e^{t(B + CW_{\nu})^*} (W_{\nu-1} - W_{\nu})C(W_{\nu-1} - W_{\nu}) e^{t(B + CW_{\nu})} dt \geq 0 .$$

Thus, after the first term, the sequence $\{W_{\nu}\}$ is monotonic decreasing and bounded below by \tilde{W} . It therefore converges to a limit W_+ . Letting $\nu \rightarrow \infty$ in (16) we see that W_+ is a solution of (3). Since $W_+ \geq \tilde{W}$ for any solution \tilde{W} of (3) it is in fact a maximal solution. Since $B + CW_+ = \lim(B + CW_{\nu})$, all eigenvalues of $B + CW_+$ have non-positive real part.

The existence of extreme solutions seems to have been first established by Reid [11], to whom the formulae (12), (13) are also due (Reid [12]). The algorithm used in the second proof was introduced in a special case by Kleinman [4] and further developed by Wonham [16]. Its applicability in the general case is mentioned by Willems [15].

Our next result relates two arbitrary solutions of (3).

THEOREM 2. *Let W_1, W_2 be any two solutions of (3) and set*

$$D = W_2 - W_1, \quad B_1 = B + CW_1, \quad B_2 = B + CW_2 .$$

Let N denote the nullspace of D and let² m, p, q be the dimensions of

² $m, p,$ and q are the number of zero, positive, and negative eigenvalues of D .

the largest subspaces invariant under D on which $D = 0$, $D > 0$, $D < 0$.

Then N is invariant under B_1 and $B_1 = B_2$ on N . Let $\lambda_1, \dots, \lambda_m$ be the eigenvalues of the restriction of B_1 to N and let $\lambda_{m+1}, \dots, \lambda_n$ be the remaining eigenvalues of B_1 . Then the eigenvalues of B_2 are $\lambda_1, \dots, \lambda_m, -\lambda_{m+1}, \dots, -\lambda_n$. Moreover among the $p + q$ eigenvalues $\lambda_{m+1}, \dots, \lambda_n$ exactly p have positive real part and q have negative real part.

Proof. We can assume $p + q > 0$, since the assertion is trivial for $D = 0$. The difference D is a solution of the quadratic equation

$$(22) \quad DB_1 + B_1^*D + DCD = 0.$$

Moreover (B_1, C) is also controllable. If $\xi \in N$ then, on postmultiplying (22) by ξ , we obtain $DB_1\xi = 0$. Thus N is invariant under B_1 . Since $B_2 = B_1 + CD$ it is obvious that $B_1 = B_2$ on N .

By replacing D, B_1, C by $T^*DT, T^{-1}B_1T, T^{-1}CT^*^{-1}$, which does not alter the hypotheses, we can assume that

$$D = \begin{pmatrix} 0 & 0 \\ 0 & D_1 \end{pmatrix},$$

where D_1 is non-singular. The nullspace N now consists of all vectors whose last $p + q$ coordinates vanish. Let the corresponding decompositions of B_1 and C be

$$B_1 = \begin{pmatrix} B_{00} & B_{01} \\ 0 & B_{11} \end{pmatrix}, \quad C = \begin{pmatrix} C_{00} & C_{01} \\ C_{01}^* & C_{11} \end{pmatrix}.$$

Then, putting $G = D_1^{-1}$, we have

$$(23) \quad B_{11}G + GB_{11}^* + C_{11} = 0.$$

Moreover

$$B_2 = B_1 + CD = \begin{pmatrix} B_{00} & B_{01} + C_{01}D_1 \\ 0 & B_{11} + C_{11}D_1 \end{pmatrix}.$$

Since $B_{11} + C_{11}D_1 = -GB_{11}^*G^{-1}$, this proves the assertion concerning the eigenvalues of B_2 .

We are going to show now that B_{11} has no pure imaginary eigenvalues. Suppose on the contrary that $B_{11}^*\zeta_1 = i\mu\zeta_1$, where μ is real and $\zeta_1 \neq 0$. Then from (23) we obtain

$$\zeta_1^*C_{11}\zeta_1 = 0.$$

Since $C \leq 0$ this implies $C_{11}\zeta_1 = 0$, and also $C_{01}\zeta_1 = 0$. Thus if we put

$$\zeta = \begin{pmatrix} 0 \\ \zeta_1 \end{pmatrix}$$

then

$$\zeta^*B_1 = -i\mu\zeta^*, \quad \zeta^*C = 0.$$

Hence

$$\zeta^*(C BC \dots B^{n-1}C) = 0.$$

Since $\zeta \neq 0$, this contradicts the hypothesis of controllability.

Put

$$B_{11}(\epsilon) = B_{11} + \epsilon G.$$

Then, by (23), $B_{11}(\epsilon)$ is a solution of the linear equation

$$B_{11}(\epsilon)G + GB_{11}^*(\epsilon) + C_{11} - 2\epsilon G^2 = 0.$$

By a generalisation of Ljapunov's Lemma (see Ostrowski and Schneider [10], Lancaster [6]) it follows that for $\epsilon > 0$, $B_{11}(\epsilon)$ has no pure imaginary eigenvalues and further that it has p eigenvalues with positive real part, q eigenvalues with negative real part. Since B_{11} has no pure

imaginary eigenvalues, the same holds by continuity for it.

COROLLARIES. (i) If \tilde{W} is a solution of (3) such that all eigenvalues of $B + C\tilde{W}$ have nonpositive real part then $\tilde{W} = W_+$. (Take $W_2 = W_+$, $W_1 = \tilde{W}$.)

(ii) If \tilde{W} is any solution of (3) then the nullspace V_0 of $W_+ - W_-$ is invariant under $B + C\tilde{W}$ and the restriction of $B + C\tilde{W}$ to V_0 has only pure imaginary eigenvalues. No other eigenvalues of $B + C\tilde{W}$ are pure imaginary.

Proof. If $\xi \in V_0$ then $W_+\xi = \tilde{W}\xi = W_-\xi$, since

$$0 \leq \xi^*(W_+ - \tilde{W})\xi \leq \xi^*(W_+ - W_-)\xi = 0.$$

Since V_0 is invariant under $B + CW_+$ it is equally invariant under $B + C\tilde{W}$. By Theorem 1 the restriction of $B + CW_+$ to V_0 has only pure imaginary eigenvalues. By Theorem 2, $B + CW_+$ has no other pure imaginary eigenvalue. If $B + C\tilde{W}$ had an additional pure imaginary eigenvalue then, by Theorem 2 again, it would be associated with the nullspace of $W_+ - \tilde{W}$ and therefore be an eigenvalue of $B + CW_+$.

The second corollary has been stated without proof by Willems [15].

It will now be shown that all solutions of (3) can be expressed in terms of the extreme solutions. Put

$$\Delta = W_+ - W_-, \quad B_+ = B + CW_+, \quad B_- = B + CW_-.$$

By (4) we have

$$(24) \quad B_+\Delta = -\Delta B_-.$$

The nullspace V_0 of Δ is invariant under B_+ and, by what we have just proved, there is a unique subspace V_+ invariant under B_+ such that the whole space $V = R^n$ satisfies

$$(25) \quad V = V_0 \dot{+} V_+.$$

All eigenvalues of the restriction of B_+ to V_0 are pure imaginary and all eigenvalues of the restriction of B_+ to V_+ have negative real part.

LEMMA 3. *Let V_1 be a subspace invariant under B_+ . If $V_0 \cap V_1 = 0$ then $V_1 \subseteq V_+$.*

Proof. Suppose $\xi \neq 0 \in V_1$. Then $\xi = \xi_0 + \xi_+$, where $\xi_0 \in V_0$ and $\xi_+ \in V_+$. Since $V_0 \cap V_1 = 0$ we must have $\xi_+ \neq 0$. Let $p(s)$ be the polynomial of least positive degree such that $p(B_+)\xi_+ = 0$. Then $p(s)$ has only roots with negative real part and

$$p(B_+)\xi = p(B_+)\xi_0.$$

Since V_0 is invariant under B_+ and $V_0 \cap V_1 = 0$ it follows that $p(B_+)\xi_0 = 0$. Therefore $\xi_0 = 0$, and $V_1 \subseteq V_+$.

THEOREM 3. *Let V_1 be an arbitrary subspace of V_+ invariant under B_+ . If V_2 is the subspace of all vectors ξ such that $\Delta\xi$ is orthogonal to V_1 then*

$$V = V_1 \dot{+} V_2.$$

Let P be the corresponding projection of V onto V_1 . Then

$$(26) \quad W = W_+P + W_-(I-P)$$

is a solution of the quadratic equation (3). Moreover all solutions of (3) are obtained in this way, and the correspondence between V_1 and W is one-to-one.

Proof. Evidently $V_0 \subseteq V_2$. If $\xi \in V_1 \cap V_2$ then $\xi^*\Delta\xi = 0$. Since $\Delta \geq 0$ it follows that $\Delta\xi = 0$. Since $V_0 \cap V_1 = 0$ this implies $\xi = 0$. Thus $V_1 \cap V_2 = 0$. We have also

$$\dim \Delta V = n - \dim V_0,$$

$$\dim V_1^\perp = n - \dim V_1,$$

and hence

$$\dim(\Delta V \cap V_1^\perp) \geq n - \dim V_0 - \dim V_1 .$$

Therefore

$$\dim V_2 = \dim V_0 + \dim(\Delta V \cap V_1^\perp) \geq n - \dim V_1 .$$

It follows that

$$V_1 \dot{+} V_2 = V .$$

If $\xi_1 \in V_1$ and $\xi_2 \in V_2$ then, by (24),

$$\xi_1^* \Delta B_- \xi_2 = -\xi_1^* B_+^* \Delta \xi_2 = 0 ,$$

since V_1 is invariant under B_+ . Thus V_2 is invariant under B_- .

Hence the projection P satisfies

$$(27) \quad PB_+P = B_+P$$

and

$$(I-P)B_-(I-P) = B_-(I-P) ;$$

that is,

$$(28) \quad PB_-P = PB_- .$$

Since ΔV_2 is orthogonal to V_1 we have also

$$P^* \Delta (I-P) = 0 ;$$

that is,

$$P^* \Delta = P^* \Delta P .$$

Since the right side is symmetric this gives

$$(29) \quad P^* \Delta = \Delta P .$$

By (27) and (28),

$$\begin{aligned} (I-P)B_+ &= (I-P)B_+(I-P) \\ &= (I-P)(B_- + C\Delta)(I-P) \\ &= B_-(I-P) + (I-P)C\Delta(I-P) . \end{aligned}$$

Therefore, by (24),

$$\Delta(I-P)B_+ + B_+^*\Delta(I-P) = \Delta(I-P)C\Delta(I-P) .$$

If we define W by (26) then

$$D = W_+ - W = \Delta(I-P) .$$

By (29), D is symmetric, and hence W is also. Moreover W is a solution of the quadratic equation (3) if D is a solution of the quadratic equation

$$DB_+ + B_+^*D - DCD = 0 .$$

But we have just shown that this is the case.

Furthermore

$$W\xi = W_+\xi \text{ for } \xi \in V_1, \quad W\xi = W_-\xi \text{ for } \xi \in V_2 .$$

Therefore V_1 and V_2 are invariant under $B + CW$ and all eigenvalues of the restriction of $B + CW$ to V_1, V_2 have respectively negative, non-negative real parts. This proves that V_1 is uniquely determined by W .

Conversely, let W be any solution of (3). Then we have a unique direct decomposition

$$V = V_1 \dot{+} V_2 ,$$

where V_1 and V_2 are invariant under $B + CW$ and all eigenvalues of the restriction of $B + CW$ to V_1, V_2 have respectively negative, non-negative real parts. It follows from Theorem 2 that

$$W\xi = W_+\xi \text{ for } \xi \in V_1, \quad W\xi = W_-\xi \text{ for } \xi \in V_2 .$$

Thus if P is the corresponding projection of V onto V_1 then W satisfies (26). Moreover V_1 is invariant under B_+ and $V_0 \cap V_1 = 0$. Therefore, by Lemma 3, $V_1 \subseteq V_+$. Since $W_+ - W = \Delta(I-P)$ is symmetric, (29) holds. Therefore, since P is a projection,

$$P^*\Delta = P^*\Delta P ;$$

that is,

$$P^*\Delta(I-P) = 0 .$$

Thus ΔV_2 is orthogonal to V_1 . Since $V = V_1 \dot{+} V_2$ it follows from the first part of the proof that V_2 is the subspace of all ξ such that $\Delta\xi$ is orthogonal to V_1 . This completes the proof.

In particular, if $\Delta > 0$ then $V_0 = 0$, V_1 is an arbitrary invariant subspace of B_+ , and $V_2 = \Delta^{-1}V_1^\perp$. Theorem 3 was proved for this case by Willems [15].

In the statement and proof of the following result we use the notation of Theorem 3.

THEOREM 4. *Let W, \tilde{W} be solutions of (3) corresponding to the invariant subspaces V_1, \tilde{V}_1 of B_+ . Then $W \geq \tilde{W}$ if and only if $\tilde{V}_1 \subseteq V_1$.*

Proof. Suppose first that $\tilde{V}_1 \subseteq V_1$. Then, by Theorem 3, $V_2 \subseteq \tilde{V}_2$. On \tilde{V}_1 we have $W\xi = W_+\xi = \tilde{W}\xi$. On V_2 we have $W\xi = W_-\xi = \tilde{W}\xi$. Therefore $\tilde{V}_1 \dot{+} V_2$ is contained in the nullspace N of $W - \tilde{W}$ and is invariant under $B + CW$. The remaining eigenvalues of $B + CW$ have negative real part, whereas the remaining eigenvalues of $B + C\tilde{W}$ have non-negative real part. Hence $N = \tilde{V}_1 \dot{+} V_2$. It follows from Theorem 2 that $W \geq \tilde{W}$.

Conversely, suppose $W \geq \tilde{W}$. The nullspace N of their difference is invariant under $B + C\tilde{W}$, and hence has a unique direct decomposition

$$N = N_+ \dot{+} N_- ,$$

where N_+, N_- are invariant under $B + C\tilde{W}$ and all eigenvalues of the restriction of $B + C\tilde{W}$ to N_+, N_- have respectively negative, non-negative real part. By the proof of Lemma 3 we have $N_+ \subseteq V_1$, $N_+ \subseteq \tilde{V}_1$. By Theorem 2 the eigenvalues of $B + C\tilde{W}$ not belonging to N have positive real part. Therefore $\dim N_+ = \dim \tilde{V}_1$. Thus $N_+ = \tilde{V}_1$ and $\tilde{V}_1 \subseteq V_1$.

It follows at once that the set of all solutions of (3) is a complete lattice. This is rather remarkable, since the set of all symmetric $n \times n$

matrices is not a lattice if $n > 1$.

3. Existence of solutions

In this section we drop the hypotheses and convention imposed at the beginning of Section 2. If \tilde{W} is a solution of (3), not assumed symmetric, such that $B + C\tilde{W}$ is stable then the characteristic polynomial

$$\psi(s) = \det(sI - B - C\tilde{W})$$

is relatively prime to $\psi(-s)$. Since, by (9), $\psi(s)$ divides $\varphi(s) = \det(sI - M)$, and $\varphi(s)$ is even, it follows that

$$(30) \quad \varphi(s) = (-1)^n \psi(s)\psi(-s).$$

On the other hand it is clear that $\varphi(s)$ admits a factorisation (30), where the (real) polynomial $\psi(s)$ is monic and relatively prime to $\psi(-s)$, if and only if M has no pure imaginary eigenvalue. Moreover it is possible to choose $\psi(s)$ so that all its roots lie in the left half-plane.

THEOREM 5. *Suppose $\varphi(s)$ has a factorisation (30), where $\psi(s)$ is monic and relatively prime to $\psi(-s)$. Then there exists a non-singular matrix X such that*

$$(31)^3 \quad X^* J X = J,$$

$$(32) \quad M X = X \begin{pmatrix} T & 0 \\ 0 & -T^* \end{pmatrix},$$

$$(33) \quad \det(sI - T) = \psi(s).$$

Proof. There exist polynomials $\chi_1(s), \chi_2(s)$ such that

$$(34) \quad 1 = \chi_1(s)\psi(s) + \chi_2(s)\psi(-s).$$

Then for any vector $x \in R^{2n}$,

$$(35) \quad x = \chi_1(M)\psi(M)x + \chi_2(M)\psi(-M)x.$$

Let V_1 denote the subspace of all vectors x such that $\psi(M)x = 0$, and

let V_2 denote the subspace of all vectors x such that $\psi(-M)x = 0$.

³ A matrix X satisfying (31) is said to be *symplectic*.

From (35) we obtain at once

$$V_1 \cap V_2 = 0 ,$$

$$V_1 + V_2 = R^{2n} .$$

Moreover V_1 and V_2 are invariant under M .

Let x_1, \dots, x_m be a basis for V_1 and x_{m+1}, \dots, x_{2n} a basis for V_2 , and let \tilde{X} be the non-singular matrix whose k -th column is x_k ($k = 1, \dots, 2n$) . Then

$$M\tilde{X} = \tilde{X} \begin{pmatrix} T & 0 \\ 0 & S \end{pmatrix} ,$$

where T is an $m \times m$ matrix, S is a $(2n-m) \times (2n-m)$ matrix, and $\psi(T) = 0$, $\psi(-S) = 0$. Thus the minimal polynomial of T divides $\psi(s)$. Since every prime factor of the characteristic polynomial of a matrix divides its minimal polynomial, it follows that $\det(sI-T)$ is relatively prime to $\psi(-s)$. On the other hand, $\det(sI-T)$ divides $\varphi(s) = (-1)^n \psi(s)\psi(-s)$. Therefore $\det(sI-T)$ divides $\psi(s)$. Thus $m > 0$ implies $m \leq n$. Similarly $m < 2n$ implies $m \geq n$. Hence $m = n$ and (33) holds.

It follows from (34) that $\psi(-T)$ and $\psi(S)$ are non-singular. Also, by (7),

$$(36) \quad \tilde{X}^* J \tilde{X} \begin{pmatrix} T & 0 \\ 0 & S \end{pmatrix} = \tilde{X}^* J M \tilde{X} = -\tilde{X}^* M^* J \tilde{X} = \begin{pmatrix} -T^* & 0 \\ 0 & -S^* \end{pmatrix} \tilde{X}^* J \tilde{X} .$$

Since $\psi(T) = 0$, $\psi(-S) = 0$ it follows that

$$(37) \quad \tilde{X}^* J \tilde{X} \begin{pmatrix} 0 & 0 \\ 0 & \psi(S) \end{pmatrix} = \begin{pmatrix} \psi(-T)^* & 0 \\ 0 & 0 \end{pmatrix} \tilde{X}^* J \tilde{X} .$$

Write, in partitioned form,

$$\tilde{X}^*J\tilde{X} = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix},$$

where $U_{12}^* = -U_{21}$. Then (37) implies that $U_{11} = 0$, $U_{22} = 0$. Since $\tilde{X}^*J\tilde{X}$ is non-singular it follows that U_{12} is non-singular. From (36) we now obtain

$$U_{12}S = -T^*U_{12}.$$

Put

$$X = \tilde{X} \begin{pmatrix} I & 0 \\ 0 & -U_{12}^{-1} \end{pmatrix}.$$

Then (31) and (32) hold. This completes the proof.

Write in partitioned form

$$X = \begin{pmatrix} Y_1 & Y_2 \\ Z_1 & Z_2 \end{pmatrix}.$$

Then (31) is equivalent to

$$(38) \quad \begin{aligned} Z_1^*Y_1 &= Y_1^*Z_1, \\ Z_2^*Y_2 &= Y_2^*Z_2, \\ Z_2^*Y_1 - Y_2^*Z_1 &= I, \end{aligned}$$

and (32) implies

$$(39) \quad \begin{aligned} BY_1 + CZ_1 &= Y_1^T, \\ -AY_1 - B^*Z_1 &= Z_1^T. \end{aligned}$$

Suppose Y_1 is non-singular, and put $W_1 = Z_1Y_1^{-1}$. Then, by (38)₁, W_1 is symmetric. By (39),

$$B + CW_1 = Y_1^TY_1^{-1},$$

$$A + W_1 B = -Y_1^{*-1} T^* Z_1^* .$$

Hence $\det(sI - B - CW_1) = \psi(s)$ and

$$\begin{aligned} Q[W_1] &= A + W_1 B + (B^* + W_1 C) W_1 \\ &= -Y_1^{*-1} T^* Z_1^* + \left(Y_1^{*-1} T^* Y_1^* \right) W_1 \\ &= 0 . \end{aligned}$$

Let W be any solution of (3) such that $\psi(s) = \det(sI - B - CW)$. Then

$$M \begin{pmatrix} I \\ W \end{pmatrix} = \begin{pmatrix} I \\ W \end{pmatrix} (B + CW)$$

and hence

$$\psi(M) \begin{pmatrix} I \\ W \end{pmatrix} = \begin{pmatrix} I \\ W \end{pmatrix} \psi(B + CW) = 0 .$$

Consider in general the linear equation

$$(40) \quad \psi(M) \begin{pmatrix} I \\ W \end{pmatrix} = 0 .$$

Writing $\psi(M) = \psi(M)X.X^{-1}$, where

$$\psi(M)X = X \begin{pmatrix} 0 & 0 \\ 0 & \psi(-T)^* \end{pmatrix}$$

and

$$X^{-1} = J^{-1} X^* J = \begin{pmatrix} Z_2^* & -Y_2^* \\ -Z_1^* & Y_1^* \end{pmatrix} ,$$

and recalling that $\psi(-T)$ is non-singular, we see that (40) is equivalent to

$$W^* Y_1 = Z_1 .$$

Since X is non-singular there is no vector $\xi \neq 0$ such that $Y_1 \xi = Z_1 \xi = 0$. It follows that (40) is soluble only if Y_1 is non-singular, and it then has the unique solution $W = W_1$.

Thus the equation (3) has a solution W_1 such that $\psi(s) = \det(sI - B - CW_1)$ if and only if Y_1 is non-singular. When such a solution exists, it is unique and symmetric. Similarly the equation (3) has a solution W_2 such that $(-1)^n \psi(-s) = \det(sI - B - CW_2)$ if and only if Y_2 is non-singular. Moreover W_2 is then the unique solution of the linear equation

$$(-W I)\psi(M) = 0 ,$$

with the same coefficient matrix as (40). If W_1 and W_2 both exist then $W_2 - W_1$ is non-singular, since by (38)3,

$$W_2 - W_1 = Y_2^*{}^{-1} Y_1^{-1} .$$

Some results in this direction have been stated by Roth [14] and Bucy and Joseph [1]. However Theorem 5, which underpins the method, is absent in these references.

THEOREM 6. *Suppose $\varphi(s)$ has a factorisation (30), where $\psi(s)$ is monic and relatively prime to $\psi(-s)$. Further suppose $C \leq 0$, $C \neq 0$. Thus we can write $C = -ER^{-1}E^*$, where E is an $n \times m$ matrix ($m \leq n$) and $R = R^* > 0$ is an $m \times m$ matrix. Then either*

- (i) *for any $m \times n$ matrix F , $\det(sI - B - EF)$ has a fixed factor of positive degree in common with $\psi(-s)$, or*
- (ii) *the quadratic equation (3) has a unique, symmetric solution W_1 such that $\det(sI - B - CW_1) = \psi(s)$.*

Similarly, either

- (i)' *for any $m \times n$ matrix F , $\det(sI - B - EF)$ has a fixed factor of positive degree in common with $\psi(s)$, or*
- (ii)' *the quadratic equation (3) has a unique, symmetric solution W_2 such that $\det(sI - B - CW_2) = (-1)^n \psi(-s)$.*

Both (ii) and (ii)' hold if and only if (B, E) is controllable; that is, if and only if (B, C) is controllable.

Proof. Let

$$X = \begin{pmatrix} Y_1 & Y_2 \\ Z_1 & Z_2 \end{pmatrix}$$

be the matrix whose existence was established in Theorem 5. We have already seen that if Y_1 is non-singular then (ii) holds.

Suppose Y_1 is singular, and let $\xi \in R^n$ be any vector such that $Y_1\xi = 0$. From (39)₁ and (38)₁ we obtain

$$\begin{aligned} \xi^*Z_1^*BY_1\xi + \xi^*Z_1^*CZ_1\xi &= \xi^*Z_1^*Y_1T\xi \\ &= \xi^*Y_1^*Z_1T\xi . \end{aligned}$$

Since $Y_1\xi = 0$ it follows that $\xi^*Z_1^*CZ_1\xi = 0$. Therefore, since $R > 0$,

$$(41) \quad E^*Z_1\xi = 0 .$$

Putting $\eta = Z_1\xi$, we now obtain from (39),

$$(42) \quad \begin{aligned} Y_1T\xi &= 0 , \\ -B^*\eta &= Z_1T\xi . \end{aligned}$$

This shows that we can replace ξ by $T\xi$ in the preceding argument. Thus for any non-negative integer k ,

$$(43) \quad \begin{aligned} E^*Z_1T^k\xi &= 0 , \\ Y_1T^k\xi &= 0 . \end{aligned}$$

For any $m \times n$ matrix F ,

$$\begin{aligned} \eta^*(B+EF) &= -\xi^*T^*Z_1^* + \xi^*Z_1^*EF \\ &= -\xi^*T^*Z_1^* . \end{aligned}$$

By induction on k , using (43) and (39)₂, it is easily shown that

$$\eta^*(B+EF)^k = (-1)^k \xi^*(T^*)^k Z_1^* .$$

Hence for any polynomial $p(s)$,

$$\eta^*p(-B-EF) = \xi^*p(T^*)Z_1^* .$$

If $Z_1 p(T)\xi = 0$ then $p(T)\xi = 0$, since $Y_1 p(T)\xi = 0$ and X is non-singular.

Fix a particular $\xi \neq 0$ in the nullspace of Y_1 and let $\omega(s)$ be the monic polynomial of least positive degree such that $\omega(T)\xi = 0$. Then $\omega(s)$ divides the minimal polynomial of T , and hence also $\psi(s)$. Furthermore $\omega(s)$ is the monic polynomial of least degree such that

$$\eta^*\omega(-B-EF) = 0 .$$

Thus $\omega(-s)$ divides the minimal polynomial, and also the characteristic polynomial, of $B + EF$. This proves (i). Since

$$B + CW_1 = B + E\left(-R^{-1}E^*W_1\right) ,$$

and $\psi(s)$ is relatively prime to $\psi(-s)$, it is clear that (i) and (ii) cannot hold simultaneously.

By replacing $\psi(s)$ by $(-1)^n\psi(-s)$ and Y_1, Z_1 by Y_2, Z_2 we obtain (i)' and (ii)'. Suppose now that (B, E) is controllable. Then by a well-known property of controllability (see, for example, Heymann [2]), neither (i) nor (i)' holds. Therefore both (ii) and (ii)' hold. Conversely, suppose that (B, E) is not controllable. Then there exists a non-singular $n \times n$ matrix P such that

$$PBP^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{pmatrix} , \quad PE = \begin{pmatrix} E_1 \\ 0 \end{pmatrix} ,$$

where $\dim B_{22} \geq 1$. Write

$$F = \begin{pmatrix} F_1 & F_2 \end{pmatrix} P .$$

Then

$$P(B+EF)P^{-1} = \begin{pmatrix} B_{11} + E_1 F_1 & B_{12} + E_1 F_2 \\ 0 & B_{22} \end{pmatrix} ,$$

and hence

$$\det(sI - B - EF) = \det(sI - B_{11} - E_1 F_1) \det(sI - B_{22}) .$$

If both (ii) and (ii)' hold then, taking $F = -R^{-1}E^*W_i$ ($i = 1, 2$), we see that $\det(sI - B_{22})$ divides both $\psi(s)$ and $\psi(-s)$. Since they are relatively prime and $\det(sI - B_{22})$ is of positive degree this is a contradiction.

COROLLARY. *Suppose $C \leq 0$. Then the following statements are equivalent:*

- (i) (B, C) is controllable and M has no pure imaginary eigenvalues;
- (ii) the equation (3) has a solution W_1 such that $B + CW_1$ is stable and a solution W_2 such that $-(B + CW_2)$ is stable.

This Corollary has been established by Molinari [8], [9] by a rather long argument involving spectral factorisation of matrices of rational functions. The key idea in the proof of Theorem 6, namely, the derivation of (41) and (42), is due to Kučera [5]. However Kučera treats only a special case and uses the Jordan canonical form.

We next use Theorem 5 to discuss the applicability of an algorithm introduced by Roberts [13].

LEMMA 4. *Let A be a linear transformation of the vector space V which has no pure imaginary eigenvalues, and let $V = V_- \dot{+} V_+$ be the direct decomposition of V into subspaces invariant under A such that the eigenvalues of the restriction of A to V_- , V_+ have respectively negative, positive real parts. Set*

$$(44) \quad A_0 = A, \quad A_{\nu+1} = \left(A_{\nu} + A_{\nu}^{-1} \right) / 2 \quad (\nu \geq 0) .$$

Then the sequence $\{A_{\nu}\}$ is defined for all ν and $A_{\nu} \rightarrow A_{\omega}$ as $\nu \rightarrow \infty$, where A_{ω} is the linear transformation defined by

$$(45) \quad A_{\omega} \xi = -\xi \quad \text{if } \xi \in V_- , \quad A_{\omega} \xi = \xi \quad \text{if } \xi \in V_+ .$$

Proof. Since the effect of changing A into $-A$ is obvious we may suppose that A is stable. If A has the eigenvalues λ_k ($k = 1, \dots, n$) then

$$\tilde{A} = (A+A^{-1})/2$$

has the eigenvalues

$$\tilde{\lambda}_k = (\lambda_k + \lambda_k^{-1})/2 \quad (k = 1, \dots, n) .$$

Hence \tilde{A} is also stable. It follows that $A - I$ and $\tilde{A} - I$ are non-singular and

$$(\tilde{A}+I)(\tilde{A}-I)^{-1} = [(A+I)(A-I)^{-1}]^2 .$$

Therefore the sequence $\{A_\nu\}$ is defined for all ν and

$$(A_\nu+I)(A_\nu-I)^{-1} = [(A+I)(A-I)^{-1}]^{2^\nu} .$$

Since all eigenvalues of $(A+I)(A-I)^{-1}$ have absolute value less than 1 this shows that $(A_\nu+I)(A_\nu-I)^{-1} \rightarrow 0$, and hence $A_\nu \rightarrow -I$.

It is easily seen that if A has at least one pure imaginary eigenvalue then the sequence $\{A_\nu\}$ is either not defined for all ν or $\lim_{\nu \rightarrow \infty} A_\nu$ does not exist.

Suppose now that the matrix M in (6) has no pure imaginary eigenvalues. Then, by Theorem 5, there exists a symplectic matrix X such that

$$M = X \begin{pmatrix} T & 0 \\ 0 & -T^* \end{pmatrix} X^{-1} ,$$

where T is stable. Hence

$$M_\omega = X \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix} X^{-1} .$$

Writing

$$M_{\omega} = \begin{pmatrix} \hat{B} & \hat{C} \\ -\hat{A} & -\hat{B}^* \end{pmatrix}, \quad X = \begin{pmatrix} Y_1 & Y_2 \\ Z_1 & Z_2 \end{pmatrix},$$

we get

$$\begin{aligned} \hat{A} &= Z_1 Z_2^* + Z_2 Z_1^* = 2Z_2 Z_1^* = 2Z_1 Z_2^*, \\ \hat{B} &= -Y_1 Z_2^* - Y_2 Z_1^* = -2Y_2 Z_1^* - I = -2Y_1 Z_2^* + I, \\ \hat{C} &= Y_1 Y_2^* + Y_2 Y_1^* = 2Y_2 Y_1^* = 2Y_1 Y_2^*. \end{aligned}$$

If Y_1 and Y_2 are non-singular then \hat{C} is non-singular and the linear equations

$$(46) \quad \hat{C}W = -I - \hat{B}, \quad \hat{C}W = I - \hat{B}$$

have the unique solutions $W_1 = Z_1 Y_1^{-1}$, $W_2 = Z_2 Y_2^{-1}$. Thus whenever (3) has both stabilizing and antistabilizing solutions they can be found by Roberts' algorithm. This result is independent of whether $C \leq 0$, and consequently should be useful in the applications to differential games.

References

- [1] Richard S. Bucy, Peter D. Joseph, *Filtering for stochastic processes with applications to guidance* (Interscience Tracts in Pure and Applied Mathematics, 23. Interscience [John Wiley & Sons], New York, London, Sydney, Toronto, 1968).
- [2] Michael Heymann, "Comments 'On pole assignment in multi-input controllable linear systems' ", *IEEE Trans. Automatic Control* 13 (1968), 748-749.
- [3] R.E. Kalman, P.L. Falb, M.A. Arbib, *Topics in mathematical system theory* (McGraw-Hill, New York; Toronto, Ontario; London; 1969).
- [4] David L. Kleinman, "On an iterative technique for Riccati equation computations", *IEEE Trans. Automatic Control* 13 (1968), 114-115.
- [5] Vladimír Kučera, "A contribution to matrix quadratic equations", *IEEE Trans. Automatic Control* 17 (1972), 344-347.

- [6] Peter Lancaster, "Explicit solutions of linear matrix equations", *SIAM Rev.* 12 (1970), 544-566.
- [7] Dahlard L. Lukes, "Stabilizability and optimal control", *Funkcial. Ekvac.* 11 (1968), 39-50.
- [8] B.P. Molinari, "The stabilizing solution of the algebraic Riccati equation", *SIAM J. Control* 11 (1973), 262-271.
- [9] B.P. Molinari, "Equivalence relations for the algebraic Riccati equation", *SIAM J. Control* 11 (1973), 272-285.
- [10] Alexander Ostrowski and Hans Schneider, "Some theorems on the inertia of general matrices", *J. Math. Anal. Appl.* 4 (1962), 72-84.
- [11] William T. Reid, "Riccati matrix differential equations and non-oscillation criteria for associated linear differential systems", *Pacific J. Math.* 13 (1963), 665-685.
- [12] William T. Reid, "A matrix equation related to a non-oscillation criterion and Liapunov stability", *Quart. Appl. Math.* 23 (1965), 83-87.
- [13] J.D. Roberts, "Linear model reduction and solution of the algebraic Riccati equation by use of the sign function", Dep. Eng. Univ. Cambridge, Cambridge, England, Rep. CUED/B Control/TR 13, 1971.
- [14] William E. Roth, "On the matrix equation $X^2 + AX + XB + C = 0$ ", *Proc. Amer. Math. Soc.* 1 (1950), 586-589.
- [15] Jan C. Willems, "Least squares stationary optimal control and the algebraic Riccati equation", *IEEE Trans. Automatic Control* 16 (1971), 621-634.
- [16] W.M. Wonham, "On a matrix Riccati equation of stochastic control", *SIAM J. Control* 6 (1968), 681-697; "Erratum: On a matrix Riccati equation of stochastic control", *SIAM J. Control* 7 (1969), 365.

Department of Mathematics,
Institute of Advanced Studies,
Australian National University,
Canberra, ACT.