King Fahd University of Petroleum & Minerals

DEPARTMENT OF MATHEMATICAL SCIENCES

Technical Report Series

TR 009

September 1980

A Hamiltonian Approach to the Factorization of Matrix Riccati Differential (Difference) Equations

Faris A. Badawi and Anders Lindquist

A HAMILTONIAN APPROACH TO THE FACTORIZATION OF MATRIX RICCATI DIFFERENTIAL (DIFFERENCE) EQUATIONS*

Faris A. Badawi† and Anders Lindquist‡

ABSTRACT. In this paper, the theory of Hamiltonian systems is utilized to obtain a factorization of a special form of matrix Riccati differential (difference) equations. The non-Riccati algorithms of Kailath and his coworkers [1-2] and Lindquist [3-8] will be obtained as corollaries. The basic idea is to consider the Riccati equation as arising from an optimal control problem with which a Hamiltonian function is associated. The factorization of the continuous-time equation is an immediate consequence of the fact that the Hamiltonian is constant along the optimal trajectory. The discrete-time setting, on the other hand, is -as expected- more complicated.

---

* Research supported by Air Force Office of Scientific Research, Air Force Systems Command, under Grant AFOSR-78-3519.

† Department of Mathematical Sciences, University of Petroleum and Minerals, Dhahran, Saudi Arabia.

‡ Department of Mathematics, University of Kentucky, Lexington, Kentucky 40506, USA.

Submitted for presentation in "Numerical Methods for System Engineering Problems" to be held at the University of Kentucky, Lexington, Kentucky, June, 1980. The Proceedings will appear in "Mathematical Programming Study Series".

1. Introduction

Matrix Riccati equations are of considerable importance in many fields of study, among which are network theory [9], control theory [10], stochastic realization theory [11-13] and estimation (filtering [14], smoothing [15,16]). Yet, little attention has been directed towards the study of the relationships that exist between Riccati equations and Hamiltonian systems. If the Riccati equation under consideration is of order (say) n and the matrix to be solved for is symmetric (which is usually the case in applications), one is then required to solve $\frac{n(n+1)}{2}$ first order differential (difference) equations. Lately, some authors have provided some new (non-Riccati) algorithms to solve for the unknown matrix by which the number of equations to be solved is largely reduced. The basic idea of these algorithms is to "factorize" the given Riccati equation. Here, we shall restrict ourselves to the works of Kailath [1-2] and Lindquist [3-8], the first of whom has used a differentiation (differencing) technique to prove these algorithms, while Lindquist has provided a proof based on backward innovations.

In this paper, we shall present a new approach for factorizing the Riccati equation under consideration based on Pontryagin's Maximum Principle. In this way, we shall obtain a new dirivation of the non-Riccati algorithms of the above mentioned authors. Our derivation will shed more light on these algorithms and will provide new links to the

Hamiltonian formulation of the given Riccati equation.

It is worth noting that the Riccati equations considered in this paper are not of the most general type that one might encounter. However, our aim here is to convey the basic ideas of the method, and we are therefore restricting ourselves to the form that usually appears in stochastic realization theory [11,12,13].

2. The Continuous-Time Case

Consider the matrix Riccati differential equation

$$\dot{P} = \Lambda(P) \; ; \qquad P(0) = P_0 \tag{2.1}$$

where the map $\Lambda : R^{n\times n} \to R^{n\times n}$ is defined by

$$\Lambda(P) = FP + PF' + (G - PH')R^{-1}(G - PH')' \tag{2.2}$$

($(')$ is the transpose), the quadruplet $[F, G, H, R]$ is a minimal realization [10] of some positive real matrix function [9] $Z(s)$, F, G, H and the positive definite matrix R are constant matrices of dimensions $n\times n$, $n\times m$, $m\times n$ and $m\times m$ respectively and P is a symmetric nonnegative definite $n\times n$-matrix. Hence (2.1) has a unique bounded solution on $[0, \infty)$ for all P_0.

To obtain the required factorization of (2.1), we shall consider the following control problem. Find a square integrable control function $u(\cdot)$ so as to minimize

$$J(u;t_1,a) = -\frac{1}{2}x'(0)P_0x(0) + \frac{1}{2}\int_0^{t_1}[u(t)'Ru(t) + 2x(t)'G\,u(t)]dt, \tag{2.3}$$

subject to

$$\dot{x}(t) = -F'x(t) - H'u(t)\ ; \quad x(t_1) = a\ . \tag{2.4}$$

Note that since $Z(s)$ is positive real, the function $J(u;.,.)$ is bounded from below [9; pp.231-232]. Hence $\xi := \inf_u J(u;.,.) > -\infty$. Let $\{u_k\ ;\ k \in Z^+\}$ (where $Z^+ = \{0, 1, 2, \cdots\}$) be a control sequence such that $J(u_k;.,.) \to \xi$ as $k \to \infty$. Then, using the parallelogram identity in Hilbert space and a completeness argument, it is seen that there is a control u^0 such that $J(u^0;.,.) = \xi$. In fact, this can also been seen from the proof of

<u>Proposition 2.1</u>. There exists a unique square integrable function u^0 minimizing $J(u;t_1,a)$. Moreover, $J(u^0;t_1,a) = -\frac{1}{2}a'P(t_1)a$, where P is the unique solution of (2.1).

<u>Proof</u>. Consider the function $\frac{1}{2}x(t)'P(t)\,x(t)$. Upon differentiating

this quantity and integrating between 0 and t_1, we obtain

$$J(u;t_1,a) + \frac{1}{2}a'\,P(t_1)a = \frac{1}{2}x(0)'[P(0) - P_0]x(0) + \frac{1}{2}\int_0^{t_1}[\dot{P}(t) - \Lambda(P(t))]$$

$$+ \frac{1}{2}\int_0^{t_1}\|u(t) + R^{-1}G'x(t) - R^{-1}HP(t)x(t)\|_R dt ,$$

where $\Lambda(P)$ is given by (2.2) and $\|x\|_R = x'Rx$. Then the result follows by noting that P satisfies (2.1) and that it can be chosen to make the last term zero. $\square$

The optimal control u^0 can be obtained as a corollary to this proposition. However, since the Hamiltonian function will play an important role in what follows, we shall instead apply the Maximum Principle. First, let x^0 be the solution of (2.4) corresponding to u^0. To apply the Pontryagin Maximum Principle, define the <u>Hamiltonian function</u>

$$\mathcal{H}(t,x(t),u(t),y(t)) = \frac{1}{2}u(t)\,Ru(t) + x'(t)Gu(t)$$

$$+ y(t)'[-F'x(t) - H'u(t)]. \tag{2.5}$$

Then the Maximum Principle requires that for a control u^0 to be optimal, $\mathcal{H}(t,x^0(t), u^0(t), y(t))$ must be minimal i.e.

$$\frac{\partial \mathcal{H}}{\partial u}(t,x^0(t),u^0(t),y(t)) = 0 = Ru^0(t) + G'x^0(t) - Hy(t) , \tag{2.6}$$

where the adjoint function y is given by

$$\frac{\partial H}{\partial x}\,(t,x^0(t),u^0(t),y(t)) = -\dot{y} = Gu^0(t) - Fy(t) \tag{2.6b}$$

with initial condition $y(0) = P_0x^0(0)$. Hence the optimal control is

$$u^0(t) = -R^{-1}[G'x^0(t) - Hy(t)] \; . \tag{2.7}$$

Using (2.4), (2.6) and (2.7), it is easy to see that the 2n-vector $\begin{bmatrix} x^0 \\ y \end{bmatrix}$ satisfies

$$\begin{bmatrix} \dot{x}^0 \\ \dot{y} \end{bmatrix} = \mathcal{F}\begin{bmatrix} x^0 \\ y \end{bmatrix} ; \quad \begin{bmatrix} x^0(t_1) \\ y(0) \end{bmatrix} = \begin{bmatrix} a \\ P_0x^0(0) \end{bmatrix} \; , \tag{2.8}$$

where $\mathcal{F}$ is the Hamiltonian matrix defined by

$$\mathcal{F} = \begin{bmatrix} -(F-GR^{-1}H)' & -H'R^{-1}H \\ GR^{-1}G' & F-GR^{-1}H \end{bmatrix} \tag{2.9}$$

Then

$$x^0(t) = X(t)x_0 \quad \text{and} \quad y(t) = Y(t)x_0 \tag{2.10a}$$

where $x_0 = x^0(0)$ and X and Y are $n \times n$ matrix functions satisfying

$$\begin{bmatrix} \dot{X} \\ \dot{Y} \end{bmatrix} = \mathcal{F}\begin{bmatrix} X \\ Y \end{bmatrix} ; \quad \begin{bmatrix} X(0) \\ Y(0) \end{bmatrix} = \begin{bmatrix} I \\ P_0 \end{bmatrix} \; . \tag{2.10b}$$

We recall the following well-known fact:

Proposition 2.2. Let X and Y be as in (2.10). Then $X(t)$ is non-singular for all t and the matrix function $P(t) := Y(t)\,X(t)^{-1}$ is the unique solution of (2.1).

As a corollary to the above proposition and using (2.10a), it can be seen that

$$y(t) = P(t)\,x^0(t)\ . \tag{2.11}$$

Now, let the function $H^0 : [0, t_1] \to R$ be defined by

$$H^0(t) := H(t, x^0(t), u^0(t), y(t))\ . \tag{2.12}$$

Lemma 2.3. Let H^0 be defined by (2.12). Then

$$H^0(t) = -\frac{1}{2}\,x^0(t)'\dot{P}(t)x^0(t)\ . \tag{2.13}$$

Moreover, $\frac{\partial H^0}{\partial t}(t) = 0$, i.e. H is constant along the optimal trajectory.

Proof. Clearly, for each $t \in [0, t_1]$, the problem to minimize $J(u;t,x^0(t))$ has the optimal solution $\{u^0(s);\ s \in [0,t]\}$; we shall misuse notations somewhat by calling this restricted function u^0 also. Then, by Proposition 2.1, $J(u^0;t,x^0(t)) = -\frac{1}{2}\,x^0(t)'P(t)x^0(t)$. Hence, since $H^0(t) = \frac{dJ}{dt}(u^0;t,x^0(t)) + y(t)'\dot{x}^0(t)$, where $y(t)$ is given by (2.11), (2.13) follows. The fact that $\frac{\partial H^0}{\partial t}(t) = 0$ follows from elementary calculus. □

<u>Lemma 2.4</u>. Let H^0 be defined by (2.12). Then

$$H^0(t) = -\frac{1}{2} x_0' M(t) x_0 , \tag{2.14}$$

where $x_0 = x^0(0)$ and the $n \times n$ matrix function M is defined by

$$M(t) := X(t)'\dot{P}(t)X(t) . \tag{2.15}$$

<u>Proof</u>. The result is an immediate consequence of (2.10a) and (2.13) □

<u>Lemma 2.5</u>. Let M and Λ be given by (2.15) and (2.2) respectively. Then

$$M(t) = \Lambda(P_0) \tag{2.16}$$

(i.e. M is constant)

<u>Proof</u>. Using (2.8) and (2.11), it is easy to see that $\dot{x}^0 = -\Gamma' x^0$, where Γ is the <u>feedback matrix</u> (2.17). Let Ψ be the transition matrix of Γ. Then $x_0 = \Psi(0,t_1)a$. Hence, $H^0(t) = -\frac{1}{2} a'\Psi(0,t_1)'M(t)\Psi(0,t_1)a$, which is constant for all $a \in R^n$. Consequently, $\Psi(0,t_1)'M(t)\Psi(0,t_1)$ is a constant matrix, hence the same is true for M. But, by definition

$M(0) = \dot{P}(0)$, which, in view of (2.1) is the same as $\Lambda(P_0)$; and consequently (2.17) follows. □

<u>Lemma 2.6</u>. Let X be as in (2.10). Then

$$\frac{d}{dt}\,(X(t)')^{-1} = \Gamma(t)(X(t)')^{-1}\ ,$$

where Γ is the feedback matrix defined by

$$\Gamma = F - (G - PH')R^{-1}H\ . \tag{2.17}$$

<u>Proof</u>. Using $\frac{d}{dt}\,(X'^{-1}) = -X'^{-1}\dot{X}X'^{-1}$, (1.45) and $Y(t) = P(t)X(t)$, the results follows. □

We are now ready to state the main result. First observe that, since the $n \times n$-matrix $\Lambda(P_0)$ is symmetric, there exist two constant matrices N and S such that $\Lambda(P_0) = NSN'$, where N is $n \times r$, S is $r \times r$ and r is the rank of $\Lambda(P_0)$. (For example, S can be choosen as the signature matrix although we shall use a different S below.)

<u>Theorem 2.7</u>.([1]). Let P be the unique solution of (2.1). Then

$$\begin{cases} \dot{P}(t) = Q(t)SQ(t)' \quad ; \quad P(0) = P_0 & (2.18a) \\ \dot{Q}(t) = \Gamma(t)Q(t) \quad ; \quad Q(0) = N\ , & (2.18b) \end{cases}$$

where $Q(t) = (X'(t))^{-1}N$ and N and S are given as above.

<u>Proof</u>. From (2.15) and (2.16), we have $\dot{P} = X'^{-1}\Lambda(P_0)X^{-1}$, which by the preceding discussion and Lemma 2.6, is (2.18). □

As an application of the factorization(2.18), consider the problem of determining the <u>Kalman gain</u> K given by

$$K = (G - \Pi H')R^{-1/2} \tag{2.19}$$

where Π is the covariance function of the <u>Kalman-Bucy filter</u> and is given by

$$\dot{\Pi}(t) = \Lambda(\Pi(t)) \quad ; \quad \Pi(0) = 0. \tag{2.20}$$

In that case, $\Pi_0 = 0$ and $\Lambda(0) = GR^{-1}G'$. Hence, choosing $N = GR^{-1/2}$ and $S = I$ in (2.18), we obtain the following non-Riccati algorithm for K

$$\dot{K} = -QQ'H'R^{-1/2} \quad ; \quad K(0) = GR^{-1/2} \tag{2.21a}$$

$$\dot{Q} = (F - KR^{-1/2}H)Q \; ; \quad Q(0) = GR^{-1/2}, \tag{2.21b}$$

which was first obtained independently by Kailath [1], who used the factorization above, and Lindquist [3], who derived it from basic principles using backward innovations.

As another application of interest in stochastic realization theory, consider the case where P_0 is any element of the set $\mathcal{P}$ defined as $\mathcal{P} := \{P = P' > 0 | \Lambda(P) \leq 0\}$. In that case, $\Lambda(P_0) = -B_2B_2'$, where $[F,\ B = (B_1,\ B_2),\ H,\ (R^{1/2},\ 0)]$ is a (wide sense) stochastic realization [11] of the output process. In such a case, we may choose N and S to be B_2 and $-I$ respectively to obtain

$$\dot{P} = -QQ' \quad ; \quad P(0) = P_0 \tag{2.22a}$$

$$\dot{Q} = \Gamma Q \quad ; \quad Q(0) = B_2 \tag{2.22b}$$

Then it can be shown [11] that $P(t) \in \mathcal{P}$ for each $t \in \mathcal{R}$ and therefore $P(t)$ is the state covariance of a realization for which

$$B_1 = (G - P(t)H')R^{-1/2} \quad \text{and} \quad B_2 = Q(t) .$$

Hence, we have the following non-Riccati algorithm generating a family of realizations

$$\dot{B}_1 = B_2B_2'H'R^{-1/2} \quad ; \quad B_1(0) = (B_0)_1 \tag{2.23a}$$

$$\dot{B}_2 = (F - B\ R^{-1/2}H)B_2 \quad ; \quad B_2(0) = (B_0)_2 \tag{2.23b}$$

for a given initial matrix $B_0 = [(B_1)_0, (B_2)_0]$. (Note the parameter t is not time now.)

This algorithm was first presented in [11]. Its discrete-time version will be presented somewhere else.

3. The Discrete-Time Case

In this section, we shall consider the matrix Riccati difference equation

$$P(t+1) - P(t) = A(P(t)) \quad ; \quad P(0) = P_0 \quad , \tag{3.1}$$

where

$$A(P) := -P + FPF' + (G - FPH')R^{-1}(G - FPH')' \quad , \tag{3.2}$$

the quadruplet $[F,G,H,J]$ is a minimal realization [10] of some discrete positive real matrix function [17] $S(z)$, F, G, H and J are constant matrices of dimensions $n \times n$, $n \times m$, $m \times n$ and $m \times m$ respectively, R, which will be denoted by $R(t)$ or $R(P)$ in the sequel, is a nonsingular $m \times m$ matrix defined by $R := J + J' - HPH'$ and P is a symmetric nonnegative $n \times n$-matrix. The aim is to obtain a factorization of this matrix Riccati equation analogous to the one obtained in the previous section. This will facilititate easy comparison with the continuous-time setting. Moreover, as will be seen shortly, the lack of symmetry between the two settings will be illustrated.

Before satating the control problem which gives rise to the above Riccati equation, we note that the matrix $T := J + J'$ is nonsingular since $T = R(t) + HP(t)H'$, and $R(t)$ is assumed to be positive definite for all t. Again the problem is to find a control $u(\cdot)$ which minimizes

$$J(u;t_1,a) = -\frac{1}{2}x(0)'P_0x(0) + \frac{1}{2}\sum_{t=0}^{t_1-1}[u(t+1)'Tu(t+1) + 2x(t+1)'Gu(t+1)] , \tag{3.3}$$

subject to

$$x(t) = F'x(t+1) + H'u(t+1) \quad ; \quad x(t_1) = a . \tag{3.4}$$

As in the continuous-time setting, the assumption that $S(z)$ is positive real insures the boundedness of the functional J and the existence of the optimal control u^0. Also, using an argument similar to that of Proposition 2.1, it is not hard to check that

$$J(u^0;t,x^0(t)) = -\frac{1}{2}x^0(t)P(t)x^0(t) , \tag{3.5}$$

where x^0 is the solution of (3.4) corresponding to u^0 and $J(u;t,x(t))$ is the value function defined by

$$J(u;t,x(t)) = -\frac{1}{2}x(0)'P_0x(0) + \frac{1}{2}\sum_{k=0}^{t-1}[u(k+1)'Tu(k+1) + 2x(k+1)'Gx(k+1)]. \tag{3.6}$$

As in section 2 we are misusing notations somewhat by denoting u^0 re-restricted to $[0, t]$ u^0 also. Again, the optimal control u^0 can be obtained from the derivation of (3.5); however, we shall resort to the Hamiltonian. To exploit the analogy with the continuous-time problem, we shall use the Maximum Principle of [18] withthe Hamiltonian function

$$H(t,x(t+1), u(t+1), y(t)) = \frac{1}{2}u(t+1)'Tu(t+1) + x'(t+1)Gu(t+1) + y(t)'[x(t+1) - F'x(t+1) - H'u(t+1)], \tag{3.7}$$

where $y(t) - y(t+1) = \frac{\partial H}{\partial x(t+1)}(t,x^0(t+1), u^0(t+1), y(t))$, i.e.

$$y(t+1) = Fy(t) - Gu^0(t+1) \quad ; \quad y(0) = P_0x^0(0) , \tag{3.8a}$$

u^0 and x^0 are as above. Note that then

$$x^0(t+1) - x^0(t) = \frac{\partial H}{\partial y(t)}\,(t, x^0(t+1), u^0(t+1), y(t)).$$

Hence, with this formulation, there is a complete analogy between the discrete-and the continuous-time settings just exchanging derivatives for differences.

Now the Maximum Principle states that $H(t, x^0(t+1), u^0(t+1), y(t))$ has a minimum for $u = u^0(t)$ i.e. by differentiation

$$Tu^0(t+1) + G'x^0(t+1) - Hy(t) = 0$$

which implies

$$u^0(t+1) = -T^{-1}[G'x^0(t+1) - Hy(t)]\;. \tag{3.8b}$$

Using relations (3.4) and (3.8), some straight-forward algebraic manipulations yield the result that the 2n-vector $\begin{bmatrix} x^0 \\ y \end{bmatrix}$ satisfies

$$\begin{bmatrix} x^0(t) \\ y(t+1) \end{bmatrix} = F \begin{bmatrix} x^0(t+1) \\ y(t) \end{bmatrix} \;;\; \begin{bmatrix} x^0(t_1) \\ y(0) \end{bmatrix} = \begin{bmatrix} a \\ P_0x^0(0) \end{bmatrix}\,, \tag{3.9}$$

where the matrix F is given by

$$F = \begin{bmatrix} A & H'T^{-1}H \\ GT^{-1}G' & A \end{bmatrix} \tag{3.10a}$$

and

$$A = F - GT^{-1}H. \tag{3.10b}$$

Then, as in the continuous-time setting

$$x^0(t) = X(t)x_0 \tag{3.11a}$$

$$y(t) = Y(t)x_0 \tag{3.11b}$$

where $x_0 = x^0(0)$ and X and Y are $n \times n$ matrix functions satisfying

$$\begin{cases} A'X(t+1) = X(t) - H'T^{-1}HY(t) \quad ; \quad X(0) = I & (3.12a) \\ \;\; Y(t+1) = AY(t) + GT^{-1}G'X(t+1) \quad ; \quad Y(0) = P_0 & (3.12b) \end{cases}$$

To exploit the analogy with the continuous-time setting, the following lemma is needed.

<u>Lemma 3.1</u>. Let P be any $n \times n$ symmetric matrix and let $R := T - HPH'$ be nonsingular. Then, the matrix $[I - H'T^{-1}HP]$ is full rank and its inverse is $[I + H'R^{-1}HP]$.

<u>Proof</u>. It is easy to check that

$$[I - H'T^{-1}HP][I + H'R^{-1}HP] = I.$$

Hence, the two matrices in the left hand side are full rank. □

<u>Lemma 3.2</u>. Let $\begin{bmatrix} X \\ Y \end{bmatrix}$ be the solution of (3.12) and let P be the solution of the Riccati equation (3.1). Then $Y(t) = P(t)X(t)$.

<u>Proof</u>. Replace Y by PX in (3.12b) to obtain

$$P(t+1)X(t+1) = AP(t)X(t) + GT^{-1}G'X(t+1)$$

Then, use Lemma 3.1 to obtain $X(t)$ from (3.12a) with $Y(t)$ set equal to $P(t)X(t)$. Then inserting this into the above equation, we obtain

$$[P(t+1) - P(t) - \Lambda(P(t))]X(t+1) = 0$$

which, in view of (3.1) is an identity. Hence the lemma follows. □

As a first corollary to the above two lemmas, we have

<u>Corollary 3.3</u>. Let X and P be as in Lemma 3.2. Then the matrix $X(t)$ is nonsingular for all $t \in Z^+$. Moreover, the matrix A is nonsingular.

<u>Proof</u>. From (3.12a), we have $A'X(1) = [I - H'T^{-1}HP_0]$, which by Lemma 3.1 is full rank since $R(0) := T - HP_0H'$ isnonsingular.

Hence A and $X(1)$ are also. The result now follows by repeating this argument for $t = 1, 2, \ldots$. □

As another corollary, we obtain the counterpat of Proposition 2.2.

<u>Proposition 3.4</u>. Let $\begin{bmatrix} X \\ Y \end{bmatrix}$ be the solution of the system

$$\begin{bmatrix} X(t+1) \\ Y(t+1) \end{bmatrix} = \begin{bmatrix} A'^{-1} & -A'^{-1}H'T^{-1}H \\ GT^{-1}G'A'^{-1} & A-GT^{-1}G'A'^{-1}H'T^{-1}H \end{bmatrix} \begin{bmatrix} X(t) \\ Y(t) \end{bmatrix}; \begin{bmatrix} X(0) \\ Y(0) \end{bmatrix} = \begin{bmatrix} I \\ P_0 \end{bmatrix} \tag{3.13a}$$

Then $P(t) := Y(t)X(t)^{-1}$ is the solution of (3.1) and

$$Y(t) = P(t)x^0(t). \tag{3.13b}$$

In analogy with the continuous-time setting, define the $n \times n$-matrix

$$M(t) = X(t + 1)'\delta P_{t+1}X(t + 1) \tag{3.14}$$

where $X(t)$ is given by (3.11) and (3.13) and $\delta P_{t+1} := P(t + 1) - P(t)$. Just as in Section 2, we want to express the optimal sequence $\{H^0(t);\ t = 0, 1, \cdots, t_1\}$ defined by

$$H^0(t) := (t, x^0(t+1), u^0(t+1), y(t)) \tag{3.15}$$

in terms of M.

<u>Proposition 3.5</u>. Let H^0 and M be given by (3.15) and(3.14) respectively. Then

$$H^0(t) = -\frac{1}{2}x_0'M(t)x_0 - \frac{1}{2}\,\delta x^0{}'(t+1)P(t)\delta x^0(t+1) \tag{3.16}$$

where $x^0 = x^0(0)$ and $\delta x^0(t+1) = x^0(t+1) - x^0(t)$.

<u>Proof</u>. It is easy to see that

$$H^0(t) = J(u^0;t+1,x^0(t+1)) - J(u^0;t,x^0(t)) + y(t)'(x^0(t+1) - x^0(t)),$$

where $J(u;t,x(t))$ is given by (3.6). Remember that u^0 restricted to $[0, t]$ minimizes $J(u;t,x^0(t))$. Hence, by (3.5) and (3.13b)

$$H^0(t) = \frac{1}{2}\,x^0(t)'P(t)x^0(t) - \frac{1}{2}\,x^0(t+1)'P(t+1)x^0(t+1)$$

$$- x^0(t)'P(t)x^0(t) + x^0(t)'P(t)x^0(t+1)\;,$$

which, upon adding and subtracting the quantity $\frac{1}{2}\,x^0(t+1)'P(t)x^0(t+1)$ and using (3.11) and (3.14), yields (3.16). □

However, unlike the continuous-time setting, the sequence $H^0(t)$ is not constant, nor is the matrix function M. We shall now use the Riccati equation (3.1) to get an alternative expression for the Hamiltonian sequence $H^0(t)$.

<u>Lemma 3.6</u>. Let H^0 and M be given by (3.15) and (3.14). Then

$$H^0(t) = -\frac{1}{2}x_0'[M(t-1) - M(t-1)X(t)^{-1}H'R(t-1)^{-1}H(X(t)')^{-1}M(t-1)]x_0'$$

$$-\frac{1}{2}\delta x^0(t+1)'P(t)\delta x^0(t+1). \tag{3.17}$$

<u>Proof.</u> First, using (3.7), we may write

$$H^0(t) = \frac{1}{2}u^0(t+1)'Tu^0(t+1) + x^0(t+1)'Gu^0(t+1) - y(t)'x^0(t)$$

$$+ y(t)'x^0(t+1).$$

Upon inserting (3.8b) and (3.13b) into this equation, we obtain

$$H^0(t) = -\frac{1}{2}x^0(t)'P(t)x^0(t) - \frac{1}{2}x^0(t+1)'GT^{-1}G'x^0(t+1)$$

$$+\frac{1}{2}x^0(t)'P(t)H'T^{-1}HP(t)x^0(t) - \frac{1}{2}x^0(t)'P(t)x^0(t)$$

$$+ x^0(t)'P(t)x^0(t+1).$$

Next, it is not hard to check that the Riccati equation (3.1) can be reformulated to read

$$P(t) - A[P(t-1) + P(t-1)H'R(t-1)^{-1}HP(t-1)]A' = GF^{-1}G' ,$$

where A is given by (3.10b). Inserting this value for $GF^{-1}G'$ into the above, we get

$$H^0(t) = -\frac{1}{2}\, x^0(t)'P(t)x^0(t) + \frac{1}{2}\, x^0(t+1)'A[P(t-1)$$

$$+ P(t-1)H'R(t-1)^{-1}HP(t)]A'x^0(t+1)$$

$$+ \frac{1}{2}\, x^0(t)'P(t)H'T^{-1}HP(t)x^0(t) - \frac{1}{2}\delta\, x^0(t+1)'P(t)\delta x^0(t+1).$$

Finally, it is not difficult to check, using (3.8b) and (3.3) that

$$x^0(t+1) = A'^{-1}[I - H'T^{-1}HP(t)]x^0(t) \tag{3.18}$$

which inserted into the above equation for $H^0(t)$ yields (3.17). □

The above lemma, together with Proposition 3.5, provide us with a recursion for M.

<u>Lemma 3.7</u>. Let M be given by (3.14). Then M satisfies

$$M(t) = M(t-1) - M(t-1)X(t)^{-1}H'R(t-1)^{-1}H(X(t)')^{-1}M(t-1) \tag{3.19}$$

<u>Proof</u>. Using the same argument as in the proof of Lemma 2.6, this follows from the fact that the <u>feedback matrix</u> defined by (3.21) below is nonsingular for all t, which follows from Lemma 3.8 below. []

Before we state the main results, we shall need the following

<u>Lemma 3.8</u>. Let $X(t)$ be as in (3.11) and (3.12). Then

$$(X(t+1)')^{-1} = \Gamma(t)(X(t)')^{-1} . \tag{3.20}$$

where Γ is the feedback matrix defined by

$$\Gamma = F - (G - FPH')R^{-1}H . \tag{3.21}$$

<u>Proof</u>. Upon applying the matrix inversion lemma to (3.12a), in view of Proposition 3.4, one obtains

$$(X(t+1)')^{-1} = [F + FP(t)H'R(t)^{-1}H - G(T^{-1} + T^{-1}HP(t)H'R(t)^{-1})H](X(t)')^{-1} . \tag{3.22}$$

The last term in (3.22) can be rewritten

$$-G(T^{-1}R(t) + T^{-1}HP(t)H')R(t)^{-1}H(X(t)')^{-1} ,$$

which is $-GR(t)^{-1}H(X(t)')^{-1}$. □

Analogously to the continuous-time case, let $r := \text{rank } M(0)$; then $M(0) = X(1)'\delta P_1 X(1)$ can be written NSN', where N is $n \times r$ and S is $r \times r$.

<u>Theorem 3.9</u>. Let $\{P(t);\ t \in Z^+\}$ be the solution of (3.1) and let N and S be as above. Then P can be determined from the system

$$P(t+1) = P(t) - Q(t)Z(t)Q(t)' \quad ; \quad P(0) = P_0 \tag{3.23a}$$

where the matrix sequences $\{Q(t);\ t \in Z^+\}$, $\{Z(t);\ t \in Z^+\}$ are generated by

$$Q(t+1) = [F - U(t+1)R(t+1)^{-1}H]Q(t) \quad ; \quad Q(0) = \Gamma(0)N \tag{3.23b}$$

$$U(t+1) = U(t) + FQ(t)Z(t)Q(t)'H' \quad ; \quad U(0) = G - FP_0H' \tag{3.23c}$$

$$R(t+1) = R(t) + HQ(t)Z(t)Q(t)'H' \quad ; \quad R(0) = J + J' - HP_0H' \tag{3.23d}$$

$$Z(t+1) = Z(t) + Z(t)Q(t)'H'R(t)^{-1}HQ(t)Z(t) \quad ; \quad Z(0) = -S \tag{3.23e}$$

<u>Proof</u>. Let $Q(t) := (X(t+1)')^{-1}N$ and $U(t) := G - FP(t)H'$. Then by (3.20), $Q(t+1) = \Gamma(t+1)Q(t)$ which, by (3.21) and the definition

of U, yields (3.23b). Next, in view of (3.19) and the fact that $M(0) = NSN'$, it can be easily seen that $M(t) = -NZ(t)N'$. Then (3.23a) follows from (3.14). Finally, (3.23c) and (3.23d) follow from (3.23a) and the definitions of U and R. □

As an application of the factorization (3.23), consider the problem of determining the Kalman gain K given by

$$K = (G - F\Pi H')R(\Pi)^{-\frac{1}{2}} \tag{3.24}$$

where Π is the covariance function of the Kalman-Bucy filter and is given by

$$\Pi(t+1) - \Pi(t) = \Lambda(\pi(t)) \quad ; \quad \Pi(0) = 0 \, . \tag{3.25}$$

In this case $\Pi_0 = 0$ and $\Lambda(0) = GT^{-1}G'$. Hence, we may choose $N = X'(1)GT^{-\frac{1}{2}}$ and $S = I$, in which case,

$$K(t) = U(t)R(t)^{-\frac{1}{2}} \quad ; \quad K(0) = GT^{-\frac{1}{2}} \tag{3.26}$$

where $U(t)$, $R(t)$ are given by (3.23) with initial conditions $Q(0) = GR^{-\frac{1}{2}}$, $U(0) = G$, $R(0) = T$ and $Z(0) = -I$.

This version of the algorithm is the one originally presented by Lindquist [6, 8]. The general case can be found in [2], where the following factorization was used

$$\delta P_{t+1} = \Gamma(t)[\delta P_t - \delta P_t H' R(t-1)^{-1} H P_t]\Gamma(t)'. \tag{3.27}$$

Relation (3.27) can be obtained from (3.19) by inserting (3.14) and noting that $(X(t+1)')^{-1}X(t)' = \Gamma(t)$.

REFERENCES

[1] T. Kailath, Some new algorithms for recursive estimation in constant linear systems, IEEE Trans. Information Theory, IT-19(1973), 750-760.

[2] M. Morf, G.S. Sidhu, and T. Kailath, Some new algorithms for recursive estimation in constant, linear discrete-time systems, IEEE Trans. Automatic Control, AC-19(1974), 315-323.

[3] A Lindquist, Optimal filtering of continuous-time stationary processes by means of the backward innovation process, SIAM J. Control, 4(1974), 747-754.

[4] ―――――, On Fredholm integral equations, Toeplitz equations and Kalman-Bucy filtering, Appl. Math. & Optimization, 1(1975), 355-373.

[5] A. Lindquist, Some reduced order non-Riccati equations for linear least-squares estimation: The stationary single-output case, Int. J. Control, 24(1976), 821-842.

[6] ________, A new algorithm for optimal filtering of discrete-time stationary processes, SIAM J. Control, 12(1974), 736-746.

[7] ________, On integrals in multi-output discrete-time Kalman-Bucy filtering, Math. Prog. Study, 5(1976), 145-168.

[8] ________, Linear least squares estimation of discrete-time stationary processes by means of backward innovation, Proc. International Symp. on Control Theory, Numerical Methods and Computer Systems Modeling, Rocquencourt, France, June 1974, Lecture Notes in Economics and Mathematical Systems, Vol.107 (Springer, Berlin, 1975), 44-63.

[9] B.D.O. Anderson and S. Vongpanitlerd, Network Analysis and Synthesis, Prentice-Hall, 1973.

[10] R.W. Brockett, Finite Dimensional Linear Systems, Wiley, 1970.

[11] A. Lindquist and G. Picci, On the stochastic realization problem, SIAM J. Control, 17(1979), 365-389.

[12] M. Pavon, Stochastic realization and invariant directions of the matrix Riccati equation, SIAM J. Control (to appear).

[13] F. Badawi, Structures and Algorithms in Stochastic Realization Theory and the Smoothing Problem, Ph.D. Thesis, University of Kentucky, Jan. 1980.

[14] R.E. Kalman, A new approach to linear filtering and prediction problems, Trans. ASME, J. Basic Eng'g., Vol.82(1960), 34-45.

[15] F. Badawi, A. Lindquist, and M. Pavon, A stochastic realization approach to the smoothing problem, IEEE Trans. Automatic Control, AC-24(1979), 878-888.

[16] ________________________, On the Mayne-Fraser Formula and stochastic realization theory for nonstationary stochastic systems, Proc. 1979 Conf. Decision and Control, Fort Lauderdale, Florida, December 1979.

[17] L. Hitz and B.D.O. Anderson, Discrete positive real fnctions and their application to system stability, Proc. IEEE, Vol.116, No.1(1969), 153-155.

[18] M.D. Canon, C.D. Collum, Jr., and E. Polak, Theory of Optimal Control and Mathematical Programming, McGraw-Hill Co., New York, 1970.