Reprinted by permission from IEEE TRANSACTIONS ON AUTOMATIC CONTROL
Vol. AC-19, No. 6, December 1974, pp. 680-692

Spectral Factorization of a Finite-Dimensional Nonstationary Matrix Covariance

BRIAN D. O. ANDERSON, SENIOR MEMBER, IEEE, AND PETER J. MOYLAN, MEMBER, IEEE

***Abstract*—For a given nonstationary matrix covariance with a finite-dimensionality property that is the time-varying generalization of the rational power spectrum matrix property, we show how to find a linear finite-dimensional system driven by white noise with output covariance equal to the prescribed covariance.**

I. Introduction

THE covariance factorization problem, or time-varying spectral factorization problem, can be stated as follows. Suppose there is given a covariance $\mathcal{R}(\cdot,\cdot)$ defined and positive definite in a certain region, say $[0, t_1] \times [0, t_1]$. What linear system, with white noise input, has an output with covariance $\mathcal{R}(\cdot,\cdot)$?

In this paper, we consider the finite-dimensional version of this problem, we allow $\mathcal{R}(\cdot,\cdot)$ to be a matrix, and we allow $\mathcal{R}(\cdot,\cdot)$ to be nonstationary.

The history of such problems is interesting. For stationary matrix $\mathcal{R}(\cdot,\cdot)$, frequency domain procedures based on factorization of the power spectrum matrix have been available for some time, see e.g., [1]–[4]. State-space viewpoints of the finite-dimensional stationary problem, of a nature allowing possible modification for the nonstationary case, are discussed in [5]–[7]; these viewpoints make use, at least indirectly, of the positive real lemma, enunciated in its original form by Kalman [8] and Yakubovic [9].

When one moves to consider the time-varying case, it soon becomes clear that the factorization problem is much easier for nonsingular covariances (those comprising a sum of a nonsingular white noise component and a continuous process component) than for singular covariances (those without the nonsingular white noise component). (The terminology nonsingular/singular is drawn from dual control problems, incidentally.)

For the infinite-dimensional case, Gohberg and Krein have solved the nonsingular problem, [10]; their solution is based upon solving an infinite number of Fredholm equations. In the finite-dimensional case, an approach based on use of the Riccati equation was suggested in [11] and developed in more polished form in [12] and [13]. Results tying together the material of [11]–[13] with the Fredholm equation approach and the Wiener–Hopf equation appear in [14] and [15]. Reference [15a] is also relevant.

Manuscript received January 15, 1974; revised July 29, 1974. This work was supported by the Australian Research Grants Committee.
The authors are with the Department of Electrical Engineering, University of Newcastle, Newcastle, New South Wales, Australia.

In the nonsingular problem, there is no essential difference in dealing with matrix $\mathcal{R}(\cdot,\cdot)$ and scalar $\mathcal{R}(\cdot,\cdot)$. This is, however, not so for the singular problem, and earlier results have been confined to scalar $\mathcal{R}(\cdot,\cdot)$ (although, actually, a limited class of matrix $\mathcal{R}(\cdot,\cdot)$ can be treated by a trivial extension). Results for scalar $\mathcal{R}(\cdot,\cdot)$ were again suggested in [11], and these were extended in a 1968 technical report subsequently appearing as [16]. In June 1968, there also appeared the thesis of Brandenburg [17] containing many similar ideas, and in December 1968 the thesis of Geesey [14]. Much of [17] was subsequently reported in the literature [18], but unfortunately [18] does not cover one of the most interesting ideas of [17], to the effect that a singular factorization problem of given (state-space) dimension, can be reduced to a nonsingular problem of *lower* dimensions than the original. (Solution of the singular problem via transformation to a nonsingular problem was initiated in [11]; it is the dimensionality reduction of [17] which is the interesting and novel idea.) This idea is also developed in [14], which considers too at length the *invertibility* of the system solving the spectral factorization problem. For the nonsingular case, the invertibility problem is easily settled, see [13] and [19], but Kailath and Geesey were the first to explicitly seek such solutions (innovations representations), and to note that some of the systems solving the singular problem in [16] and [17] were in fact invertible.

Other work on the time-varying problem can be found in [20] and [21], approaching the problem, respectively, as one requiring factorization of differential operators, and one requiring solution of a nonlinear integral equation.

As noted earlier, in this paper we consider the nonstationary finite-dimensional matrix problem. We also allow $\mathcal{R}(\cdot,\cdot)$ to be singular.

Our method of approach differs from any employed in treatments to this point of the scalar problem. However, we do make use of the notion of reducing the state-space dimension where possible, though not in the same way as Brandenburg [17]. In general terms, we relate the problem for a singular $r \times r$ covariance with associated state-space dimension n to the problem for an associated $r' \times r'$, not necessarily nonsingular, covariance with associated state-space dimension n'. One has $r' \le r$, $n' \le n$, *with at least one inequality holding*, and one can continue a series of such dimension reductions until either a nonsingular (and thus solvable) problem is encountered, or a dimension shrinks to zero, leaving a trivial problem. The dimensionality reduction is critical, since without it there is no guarantee that the algorithm will terminate. (Note that in contrast to [17], one does not delay the implementation of a dimensionality reduction until a nonsingular problem is encountered.) A further property of the procedure is that it is straightforward to obtain an invertible system as the solution of the factorization problem.

If one attempts to apply the scalar singular procedure of [16] in the matrix case, one encounters a difficulty as soon as different numbers of differentiations of the various components of the vector process are required to produce white noise. One can seek to take this into account with the procedures employed by Bryson and Johansen [22] in their study of Kalman filtering problems, but one encounters still a further difficulty: it may be that it is never possible to obtain a nonsingular problem via differentiation, nor at the same time a totally singular problem.[1] Indeed, one has a new class of problems, termed in the control literature partially singular, which does not appear in the scalar case. Nevertheless, unpublished work of Silverman and Anderson, and a thesis of Powell [34], achieves a solution this way, and Moore has indicated in a private communication how substantial simplification can be made using the ideas of Goh [23], [24] developed for singular control problems. The method of this paper however seems to be simpler.

In developing the ideas of this paper, we have benefited substantially from a survey of linear-quadratic variational problems, see [25]. Combining the ideas of [25] with some of this paper, we have discovered new control results (and network theory results) which we shall report separately.

The plan of the paper is as follows. In Section II, we give a precise formulation of the problem, and note the two key alternative assumptions under which a solution may be obtained. In Sections III and IV using each assumption in turn, we solve the problem. Section V contains some remarks on the stationary problem, and Section VI contains concluding remarks.

II. Formulation of the Problem

Suppose there is given a two-variable function $\mathcal{R}(\cdot,\cdot)$ defined on $[0,t_1] \times [0,t_1]$, which can be written in the form

$$\mathcal{R}(t,\tau) = R(t)\delta(t-\tau) + H'(t)\Phi(t,\tau)K(\tau)1(t-\tau) + K'(t)\Phi'(\tau,t)H(\tau)1(\tau-t). \quad (1)$$

Here, $\delta(\cdot)$ and $1(\cdot)$ denote the delta function and unit step function, respectively, $\Phi(\cdot,\cdot)$ is the $n \times n$ transition matrix associated with some equation $\dot{x} = F(t)x$, and the superscript prime denotes matrix transposition. The matrices $H(\cdot)$ and $K(\cdot)$ are $n \times r$, $R(\cdot)$ is $r \times r$, and $F(\cdot)$, $H(\cdot)$, $K(\cdot)$, and $R(\cdot)$ are all assumed to have entries differentiable as many times as are required in the algorithms to follow.

The covariance of the output of any linear finite-dimensional system excited by white noise will have the form (1). It is well known that the covariance has a nonnegativity property, viz.,

$$\int_0^{t_1}\int_0^{t_1} u'(t)\mathcal{R}(t,\tau)u(\tau)\,dtd\tau \geq 0 \qquad \text{for all continuous } u(\cdot).$$

We shall say that a system *realizing* $\mathcal{R}(\cdot,\cdot)$ is a system such that when driven by white noise and with an appropriate random initial condition, the output covariance over $[0,t_1]$ is $\mathcal{R}(\cdot,\cdot)$. It therefore clearly makes sense, given an $\mathcal{R}(\cdot,\cdot)$ of the form of (1) and with the nonnegativity property, to search for a system realizing $\mathcal{R}(\cdot,\cdot)$.

In searching for a system realizing $\mathcal{R}(\cdot,\cdot)$, it is convenient to restrict a search for such systems to ones with an impulse response $J(t)\delta(t-\tau) + H'(t)\Phi(t,\tau)G(\tau)1(t-\tau)$, so that it is the matrices $G(\cdot)$ and $J(\cdot)$, *together with the system initial condition*, which define the system. The state-space equations of such a system are $\dot{x} = F(t)x + G(t)u$, $y = H'(t)x + J(t)u$, and $u(\cdot)$ is assumed to be unit intensity white noise, i.e., $E[u(t)u'(\tau)] = I\delta(t-\tau)$. The initial condition on the system is a random one, requiring $E[x(0)x'(0)] = P_0$ for some nonnegative definite symmetric P_0, and $x(0)$ is independent of $u(\cdot)$. In order that such a system realize $\mathcal{R}(\cdot,\cdot)$, it is necessary and sufficient [13] that the following equations be satisfied on $[0,t_1]$ for some nonnegative definite symmetric $P(t)$:

$$\begin{aligned} \dot{P} &= PF' + FP + GG' \qquad P(0) = P_0 \\ PH &= K - GJ' \\ JJ' &= R. \end{aligned} \quad (2)$$

The matrix $P(t)$ is actually $E[x(t)x'(t)]$.

If the system is known, but the covariance (1) is not, the latter may easily be computed using (2). However, if the covariance is known in the sense that F, H, K, and R are given, the converse problem of finding the system, or finding P_0, $G(\cdot)$, and $J(\cdot)$ such that (2) holds, is harder; the paper addresses precisely this converse problem.

To solve the converse problem, there are two classes of assumptions we can make. First, we can assume the existence of $P_0 = P_0' \geq 0$, $G(\cdot)$, and $J(\cdot)$ satisfying (2) for the prescribed F, H, K, and $J(\cdot)$ without assuming that we know the particular values. Then we can attempt to compute the values of P_0, $G(\cdot)$, and $J(\cdot)$, or indeed any other triple satisfying (2), i.e., we start with the following assumption.

Assumption 1: There exists a system defined by matrices $F(\cdot)$, $G(\cdot)$, $H(\cdot)$, $J(\cdot)$, and P_0, with the actual values of $G(\cdot)$, $J(\cdot)$, and P_0 unknown, which realizes the covariance (1).

This approach is followed in Section III. Clearly, it is to an extent unsatisfactory, since it leaves unanswered the question of what properties of *the covariance alone* imply existence of a system realizing the covariance. Now it turns out that mere nonnegativity in the earlier described sense of a two-variable $\mathcal{R}(t,\tau)$ of the form (1) is *not* quite adequate to guarantee this existence. Besides

[1] For example, the given covariance matrix could be the direct sum of a nonsingular covariance and a covariance of the type $\sum_{i=1}^{N} \phi_i(t)\phi_i(s)$ for smooth $\phi_i(\cdot)$.

various technical assumptions detailed as the need arises, the following assumption is also required.

Assumption 2: $\mathcal{R}(\cdot,\cdot)$ has the following extendability property: if $\mathcal{R}(\cdot,\cdot)$ possesses the nonnegativity property on $[0,t_1]$, then there exist definitions of $F(\cdot)$, $H(\cdot)$, $K(\cdot)$, and $\mathcal{R}(\cdot)$ on $(t_1,t_1+\epsilon]$, for some $\epsilon > 0$, such that $R(\cdot,\cdot)$ possesses the nonnegativity property on $[0,t_1+\epsilon]$ and such that $\int_{t_1}^{t_1+\epsilon} \Phi'(\tau,t_1)H(\tau)H'(\tau)\Phi(\tau,t_1)\,d\tau$ is nonsingular.

Note that this assumption is one on the covariance alone; note also that it will necessarily be satisfied by any covariance defined on $[0,t_1]$ which has a system realizing it on $[0,t_1]$—for any definition for the system of $F(\cdot)$, $G(\cdot)$, $H(\cdot)$, and $J(\cdot)$ on $(t_1,t_1+\epsilon]$ preserving continuity and such that the nonsingularity of the observability integral is fulfilled will define an extension for the covariance, in which $K(\cdot)$ and $R(\cdot)$ on $(t_1,t_1+\epsilon]$ are computed via (2). For these reasons, the assumption is well justified.

A realization procedure based on this assumption is presented in Section IV. The computations are almost identical with those of Section III. The justification of the computational procedure is however more complex, and herein is the reason for first giving the procedure of Section III.

As noted in the introduction, in the nonsingular case, one can fairly easily solve the nonsingular problem. One proceeds as follows. Define a matrix $\Pi_m(\cdot)$ by

$$\dot{\Pi}_m = \Pi_m F' + F\Pi_m + (\Pi_m H - K)R^{-1}(\Pi_m H - K)' \qquad \Pi_m(0) = 0. \quad (3)$$

Then the identifications $G = (K - \Pi_m H)R^{-\frac{1}{2}}$ and $J = R^{\frac{1}{2}}$ ensure satisfaction of (2) with $P(t) = \Pi_m(t)$; in particular, $P_0 = 0$.

The technical question arises of ensuring that (3) has no escape time, i.e., ensuring that $\Pi_m(t)$ exists on $[0,t_1]$. This is guaranteed either by Assumptions 1 or 2. In case Assumption 2 holds, one can show, as in Appendix I, that $\Pi_m(t)$ is bounded above and below for all $t \in [0,t_1]$. This eliminates the possibility of an escape time. In case Assumption 2 holds, it follows, see Appendix II, that $\mathcal{R}(\cdot,\cdot)$ is positive definite on $[0,t_1]$, i.e., that

$$\int_0^{t_1}\int_0^{t_1} u'(t)\mathcal{R}(t,\tau)u(\tau)\,dt d\tau \geq \eta \int_0^{t_1} u'(t)u(t)\,dt \quad (4)$$

for some $\eta > 0$ and all $u(\cdot)$; this condition guarantees existence of $\Pi_m(\cdot)$ by a theorem of [35], modulo a straightforward time reversal. For completeness, a proof is also contained in Appendix II.

Notice that the equivalence of Assumption 2 and the positive definite property is valid only for nonsingular $R(t)$; attempts to tackle singular realization problems via imposition of a positive definite property on $\mathcal{R}(\cdot,\cdot)$ (as opposed to a nonsingular covariance possibly derived in the course of solving the realization problem) are intrinsically attempts at solving too restrictive a problem. Put another way, Assumption 2 rather than positive definiteness (or even a demand that $\int_0^{t_1}\int_0^{t_1} u'(t)\mathcal{R}(t,\tau) u(\tau)\,dt d\tau > 0$ for all continuous $u(\cdot)$ not identically zero) is the most natural condition encompassing nonsingular and singular problems.

In case R is singular, the approach based on (3) fails, and the realization problem is much harder. For zero R, it can sometimes be the case that application of the scalar covariance procedures of [16] will solve the problem. (This view is espoused in [26].) Our solution procedure makes no such assumption.

Another solution procedure solving the same problem has come to our notice since preparation of the first draft of this paper, [34]. In order to compare the two procedures, we shall defer comment on [34] until the details of our procedure have been described in the next two sections.

In preparation for the next two sections, we now note the following points.

1) We shall have occasion to change the state-space coordinate basis; this of course has no effect on the given covariance, so that the essence of the spectral factorization problem is unchanged.

2) We shall have occasion to transform the vector process $y(t)$ of covariance $\mathcal{R}(\cdot,\cdot)$. Thus if $S(t)$ is a nonsingular $r \times r$ matrix of continuous entries, $\hat{y}(t) = S(t)y(t)$ has covariance $\hat{\mathcal{R}}(t,\tau) = S(t)\mathcal{R}(t,\tau)S'(\tau)$. Again, *the essence of the problem is unchanged.*

3) Our solutions to the realization problem will actually demand further assumptions than Assumptions 1 or 2; these extra assumptions are ones requiring differentiability of the entries and constancy of the rank of various matrices, and, physically, seem to amount to disallowing structural changes in the system realizing $\mathcal{R}(\cdot,\cdot)$. The differentiability and constancy of rank assumptions will be explicitly listed as assumptions when they are needed.

III. Realization Given Model Existence

We start with (1) and the assumption that there exists some $G(\cdot)$, $J(\cdot)$ and P_0 (and therefore $P(t)$) such that (2) holds. *What these latter matrices actually are is unknown;* in fact, we do not even know the number of columns of G and J. For convenience, let us rewrite this assumption as: there exists a nonnegative definite symmetric $P(t)$ defined on $[0,t_1]$ such that

$$M(P) = \begin{bmatrix} \dot{P} - PF' - FP & PH - K \\ H'P - K' & R \end{bmatrix} \geq 0. \quad (5)$$

(Observe that (5) implies the existence of $G(\cdot)$ and $J(\cdot)$ satisfying (2) and conversely.)

Let us adopt the convention that $P(\cdot)$ will denote the matrix whose existence is abstractly known but whose value is not known, and $\Pi(\cdot)$ will denote a nonnegative definite matrix whose value we shall find, and which satisfies (5) with $P(\cdot)$ replaced by $\Pi(\cdot)$. Then finding a nonnegative $\Pi(\cdot)$ satisfying (5) is equivalent to solving the realization problem of finding $G(\cdot)$, $J(\cdot)$, and a P_0 satisfying (2), [compare (2) and (5)].

As noted in the previous section, with $R(\cdot)$ nonsingular, the problem of finding a $\Pi(\cdot)$ is easily solved. Define $\Pi_m(\cdot)$ by (3), and, as shown in Appendix I, there is no escape time; one easily checks that $M(\Pi_m) \geq 0$. Further, as noted in Appendix I, the $\Pi_m(\cdot)$ defined by (3) is minimum amongst all matrices $\Pi(\cdot)$ for which $\Pi(t) = \Pi'(t) \geq 0$, i.e., $\Pi_m(t) \leq \Pi(t)$ for all t and all $\Pi(\cdot)$.

To tackle the case of singular $R(t)$, we shall apply one or both of two sorts of reduction steps to $M(P)$. One step involves reduction of the dimension r or $R(t)$ and is effected with the aid of output transformations. The second step involves reduction of the dimension n of $F(\cdot)$. Application of these reduction procedures leads ultimately to either a problem with output dimension of 0 (i.e., no process remains to be realized), a problem with state dimension of zero (the process has no dynamics), or a problem with nonsingular R matrix (then a known procedure applies). A flow diagram summarizing the whole procedure is given in Fig. 1, and can be examined in conjunction with the detailed description of the procedure.

Step 1—(Output Transformation): Make the following assumption.

Assumption 3: The entries of $R(\cdot)$ are continuously differentiable k times, for some $k \geq 1$, and $R(t)$ has constant rank ρ on $[0,t_1]$.

Then there exists a nonsingular S with entries k times continuously differentiable such that

$$\hat{R} = SRS' = \begin{bmatrix} \hat{R}_0 & 0 \\ 0 & 0_{p_1 \times p_1} \end{bmatrix} \tag{6}$$

with $\hat{R}_0$ a nonsingular matrix (here, $p_1 = r - p$). Set $\hat{H} = HS'$ and $\hat{K} = KS'$. [To see that $S(\cdot)$ exists, notice that by the Lagrange method [27] we can write $R(t) = V(t)V'(t)$ with $V(t)$ square, of rank ρ, and with entries k times continuously differentiable. Then by Dolezal's theorem, [36], there exists a nonsingular $S'(t)$, with entries k times continuously differentiable, such that $V'(t)S'(t) = [V_1'(t) \; 0]$ for $V_1'(t)$ with ρ columns. Then set $\hat{R}_0(t) = V_1(t)V_1'(t)$.]

The physical interpretation of the transformation is as described near the end of the last section; in lieu of examining a process $y(t)$ of covariance $\mathcal{R}(t,\tau)$, we examine a process $\hat{y}(t) = S(t)y(t)$ of covariance $S(t)\mathcal{R}(t,\tau)S'(\tau)$. The last p_1 entries of $\hat{y}(t)$ do not contain a nonsingular white noise component.

Now drop the superscript hat.

Step 2—(Further Output Transformation and Output Dimension Reduction): Partition $H(t)$ as $[H_1(t) \; H_2(t)]$, with $H_2(t)$ of dimension $n \times p_1$. Make the following assumption.

Assumption 4: $H_2(t)$ has constant rank $p \leq p_1$ on $[0,t_1]$. If $p = p_1$, pass to Step 3. Otherwise, let $S_0(t)$ be a nonsingular $p_1 \times p_1$ matrix with entries as differentiable as the entries of $H_2(t)$ and such that $H_2(t)S_0'(t) = [\hat{H}_2(t) \; 0]$, with $\hat{H}_2(t)$ having p columns. (Note that Doležal's theorem guarantees that $S_0(t)$ exists.) Set $S(t) = I \oplus S_0(t)$ with the unit matrix of dimension $(r - p_1)$, and define $\hat{R}(t) = S(t)R(t)S'(t)$, $\hat{H}(t) = H(t)S'(t)$, $K(t) = \hat{K}(t)S'(t)$. This yields, dropping the superscript hat again,

$$R(t) = \begin{bmatrix} R_0(t) & 0 \\ 0 & 0_{p_1 \times p_1} \end{bmatrix} \qquad H(t) = [H_1(t) \;\; H_2(t) \;\; 0_{n \times (p_1 - p_1)}]$$

$$K(t) = [K_1(t) \quad K_2(t) \quad K_3(t)].$$

Here, $K(t)$ is partitioned like $H(t)$. Now observe that (5) forces $K_3(t) = 0$. (Any vector whose entries are 0 except for the last $(p - p_1)$ is in the nullspace of $R(t)$, and so for (5) to hold must be in the nullspace of $PH - K$. It is in the nullspace of H, and so must be in the nullspace of K.) This conclusion of course uses the fact that $P(t)$ exists, but does not use its value, which anyway is unknown.

The physical interpretation is that there is an output transformation whose effect is restricted to those components of the process not containing a white noise component. After the transformation, the last $(p_1 - p)$ entries of the vector process of covariance $\mathcal{R}(\cdot,\cdot)$ are zero almost everywhere, and, accordingly may be dropped from consideration.

Now define

$$\hat{R} = \begin{bmatrix} R_0 & 0 \\ 0 & 0_{p \times p} \end{bmatrix} \quad \hat{H} = [H_1 \; H_2] \quad \hat{K} = [K_1 \quad K_2]. \tag{7}$$

Also define the matrix $\hat{M}$ in an obvious fashion. *The same P that guarantees* (5) *will guarantee* $\hat{M}(P) \geq 0$. Further, if we demonstrate the existence of and compute a Π such that $\hat{M}(\Pi) \geq 0$, then $M(\Pi) \geq 0$.

If now $p = 0$, realization is immediate because R_0 is nonsingular, and if not, we proceed to Step 3, bearing in mind that $H_2(t)$ has rank p. Drop the superscript hat again, and redefine r to be the new dimension of R.

Step 3—(State-Space Coordinate Basis Changes and State-Space Dimension Reduction): Select a coordinate basis change matrix $T(t)$, nonsingular and with entries as differentiable as those of $H_2(t)$, such that

$$\hat{H} = (T^{-1})'H = \begin{bmatrix} H_{11} & 0 \\ H_{21} & I_{p \times p} \end{bmatrix}. \tag{8}$$

(Again, we appeal to Doležal's theorem.) Also define $\hat{F} = TFT^{-1} + \dot{T}T^{-1}$, $\hat{K} = TK$, $\hat{H} = (T^{-1})'H$. The matrix P (and Π) transforms according to $\hat{P} = TPT'$ and then $\hat{M}(\hat{P}) = (T \oplus I_r)M(P)(T' \oplus I_r) \geq 0$. (Notice that for $\hat{F}$ to have entries which are k times continuously differentiable, T must have entries which are $(k + 1)$ times continuously differentiable, because $\dot{T}$ occurs in the formula for $\hat{F}$.) Now drop the superscript hats.

Partition P as

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{12}' & P_{22} \end{bmatrix} \tag{9}$$

(and Π similarly), with P_{11} of dimension $(n - p) \times (n - p)$ and P_{22} of dimension $p \times p$, and partition K similarly to H. The last p columns of $PH - K$ must be zero, which means that

$$\begin{bmatrix} P_{12} \\ P_{22} \end{bmatrix} - \begin{bmatrix} K_{12} \\ K_{22} \end{bmatrix} = 0. \tag{10}$$

Till this point, the actual value of $P(t)$ has been unknown. However, this equation identifies the matrices P_{12} and P_{22}

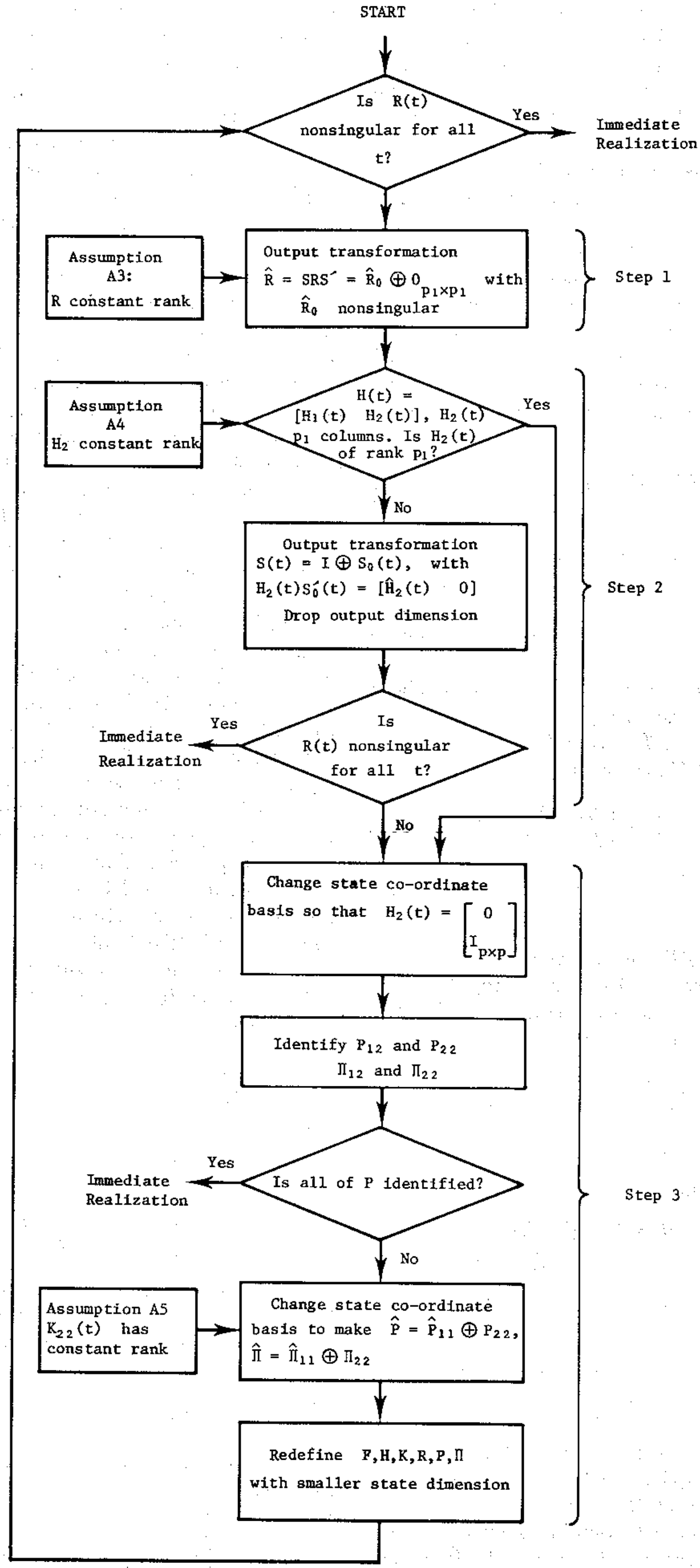
START
Is R(t) nonsingular for all t?
Yes
Immediate Realization
Assumption A3: R constant rank
Output transformation R̂ = SRS′ = R̂₀ ⊕ 0 p₁×p₁ with R̂₀ nonsingular
Step 1
Assumption A4 H₂ constant rank
H(t) = [H₁(t) H₂(t)], H₂(t) p₁ columns. Is H₂(t) of rank p₁?
Yes
No
Output transformation S(t) = I ⊕ S₀(t), with H₂(t)S₀′(t) = [Ĥ₂(t) 0] Drop output dimension
Step 2
Is R(t) nonsingular for all t?
Yes
Immediate Realization
No
Change state co-ordinate basis so that H₂(t) = [0; I pxp]
Identify P₁₂ and P₂₂ Π₁₂ and Π₂₂
Is all of P identified?
Yes
Immediate Realization
No
Step 3
Assumption A5 K₂₂(t) has constant rank
Change state co-ordinate basis to make P̂ = P̂₁₁ ⊕ P₂₂, Π̂ = Π̂₁₁ ⊕ Π₂₂
Redefine F,H,K,R,P,Π with smaller state dimension

Fig. 1.

associated with *any and every* system realizing $\mathcal{R}(\cdot,\cdot)$; in particular then, we must have also $\Pi_{12} = K_{12}$ and $\Pi_{22} = K_{22}$, with K_{22} symmetric and nonnegative definite because Π and P have this property. We must still find $\Pi_{11}(t)$, and we have the knowledge that some $P_{11}(t)$ exists for which $M(P) \geq 0$.

It may be the case that the various dimensions arising in Step 2 lead to $H_1(t)$ having no columns, and accordingly P_{11} and P_{12} evanesce. In this case, the realization procedure terminates because all entries of P and Π are identified. Suppose therefore this is not so.

Now because $P(t) \geq 0$ for all t, it must be the case that $K_{22}(t)\alpha = 0$ for some t and α implies $K_{12}(t)\alpha = 0$, and so $K_{12}'(t) = K_{22}(t)K_{22}^{\#}(t)K_{12}'(t)$. Here, the superscript # denotes the pseudoinverse. Now (in order to block-diagonalize $P(t)$ and $\Pi(t)$) define a further coordinate basis change matrix

$$T(t) = \begin{bmatrix} I & -K_{12}(t)K_{22}^{\#}(t) \\ 0 & I \end{bmatrix}. \tag{11}$$

Obviously, $T(t)$ is nonsingular; to ensure that it has entries inheriting the differentiability of $K_{12}(\cdot)$ and $K_{22}(\cdot)$, we assume the following.

Assumption 5: The matrix $K_{22}(t)$ has constant rank on $[0,T]$.

(This assumption ensures that entries of $K_{22}^{\#}(t)$ inherit the differentiability of entries of $K_{22}(t)$.)

Now set $\hat{F} = TFT^{-1} + \dot{T}T^{-1}$, $\hat{K} = TK$, $\hat{H} = (T^{-1})'H$, $\hat{P} = TPT'$, and $\hat{\Pi} = T\Pi T'$. In particular,

$$\hat{K} = \begin{bmatrix} K_{11}(t) & 0 \\ K_{21}(t) & K_{22}(t) \end{bmatrix} \qquad \hat{H} = \begin{bmatrix} H_{11} & 0 \\ \hat{H}_{21} & I \end{bmatrix}$$

$$\hat{P} = \begin{bmatrix} \hat{P}_{11} & 0 \\ 0 & P_{22} \end{bmatrix} \tag{12}$$

with $\hat{P}_{11} = P_{11} - P_{12}P_{22}^{\#}P_{12}'$ for example. (In computing $\hat{P}$, (10) is used.) Likewise, $\hat{\Pi}$ is $\Pi_{11} \oplus K_{22}$. Drop the superscript hats again, and consider the inequality $M(P) \geq 0$, recalling that $P_{12} = K_{12}$ and $P_{22} = K_{22}$:

$$M(P) = \begin{bmatrix} \dot{P}_{11} - P_{11}F_{11}' - F_{11}P_{11} & -P_{11}F_{21}' - F_{12}K_{22} & P_{11}H_{11} - K_{11} & 0 \\ -F_{21}P_{11} - K_{22}F_{12}' & \dot{K}_{22} - K_{22}F_{22}' - F_{22}K_{22} & K_{22}H_{21} - K_{21} & 0 \\ H_{11}'P_{11} - K_{11}' & H_{21}'K_{22} - K_{21}' & R_0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \geq 0. \tag{13}$$

Make the following definitions:

$$\hat{P} = P_{11} \qquad \hat{F} = F_{11} \qquad \hat{H} = [-F_{21}' \quad H_{11}]$$

$$\hat{K} = [F_{12}K_{22} \quad K_{11}]$$

$$\hat{R} = \begin{bmatrix} \dot{K}_{22} - K_{22}F_{22}' - F_{22}K_{22} & K_{22}H_{21} - K_{21} \\ H_{21}'K_{22} - K_{21}' & R_0 \end{bmatrix}. \tag{14}$$

Then $\hat{M}(\hat{P})$ is precisely the top left part of $M(P)$ which is not identically 0. Obviously, $\hat{M}(\hat{P}) \geq 0$ is equivalent to $M(P) \geq 0$, every part of $\hat{F}$, $\hat{H}$, $\hat{K}$, and $\hat{R}$ is known, and P_{11} is unknown, although the fact of its existence is known. The search for a Π satisfying $M(\Pi) \geq 0$ is equivalent to the search for a $\hat{\Pi}$ such that $\hat{M}(\hat{\Pi}) \geq 0$.

The crucial point is the reduction in the effective state-space dimension achieved at this point. If $\hat{R}$ is nonsingular, $\hat{\Pi}$ can be obtained immediately. If not, then the stage is set for further application(s) of Steps 1–3, and, provided that the assumptions corresponding to Assumptions 3–5 are satisfied, one is guaranteed that the process must end in one of three ways. Either one is left with a process of zero dimension to realize, or a process with no dynamics, or a process with nonsingular R. In each case, determination of Π and thus realization is immediate.

Here are some other points.

1) Suppose one is after the minimum Π, call it Π_m, for which the original $M(\Pi)$ in (5) is nonnegative definite. (That there is a minimum for the singular problem is not at once obvious. However if the realization algorithm goes through, we can construct Π_m, as explained below.) In case $R(t)$ is nonsingular for all t, Lemma 1 gives $\Pi_m(t)$. Otherwise, because the various reductions of Steps 1–3 preserve the ordering of Π matrices, in the sense that if $\Pi_1(\cdot)$ and $\Pi_2(\cdot)$ are two solutions of $M(\Pi) \geq 0$ and if $\hat{\Pi}_1(\cdot)$ and $\hat{\Pi}_2(\cdot)$ are the corresponding solutions of $\hat{M}(\hat{\Pi}) \geq 0$ derived after application of one or a number of Steps 1–3, then $\Pi_1(t) \geq \Pi_2(t)$ if and only if $\hat{\Pi}_1(t) \geq \hat{\Pi}_2(t)$. (This is easily seen, for the only transformations of the Π are congruency transformations.) Consequently, if successive applications of Steps 1–3 lead to a nonsingular situation, for which a minimum $\hat{\Pi}_m(t)$ can be computed, this reflects back to a minimum $\Pi_m(t)$ for the original inequality $M(\Pi) \geq 0$. If successive applications of Steps 1–3 do not lead to a nonsingular situation, but do lead to a realization, it must be the case that all of the matrix Π of any realization is uniquely identified, i.e., $\Pi_m(t) = \Pi(t)$ for all realizations. A realization associated with minimum $\Pi(\cdot)$ has, as we shall see, an invertibility property. This is known for the nonsingular case; see [14], [15], [38], and [31] for results connecting minimality and invertibility properties. The fact that $\Pi_m(t)$ defined in (3) defines, for the nonsingular case, an invertible realization first appeared in [19], see also [12] and [13].

2) Suppose $\mathcal{R}(\cdot,\cdot)$ is singular, and define Π_ϵ by (3) with R replaced by $R + \epsilon I$ for $\epsilon > 0$. If P satisfies (5), then P also satisfies

$$M_\epsilon(P) = \begin{bmatrix} \dot{P} - PF' - FP & PH - K \\ H'P - K' & R + \epsilon I \end{bmatrix} \geq 0$$

and then the argument of Appendix I yields $0 \leq \Pi_\epsilon(t) \leq P(t)$. An analysis of the differential equation for Π_ϵ will show that Π_ϵ increases as $\epsilon \to 0$. In view of the upper bound $P(t)$ on $\Pi_\epsilon(t)$, it is clear that $\lim_{\epsilon\to 0} \Pi_\epsilon(t) = \bar{\Pi}(t)$ exists, with $\bar{\Pi}(t) \leq P(t)$. Now for each $\epsilon > 0$, $M_\epsilon(\Pi_\epsilon) \geq 0$; one cannot necessarily take the limit to conclude that $M(\bar{\Pi}) \geq 0$ for $\bar{\Pi}(\cdot)$ is not guaranteed differentiable. However, the Helly convergence theorem [37] does guarantee that

$$dN(\bar{\Pi}) = \begin{bmatrix} d\bar{\Pi} - (\bar{\Pi}F' + F\bar{\Pi})dt & (\bar{\Pi}H - K)dt \\ (\bar{\Pi}H - K)'dt & R\,dt \end{bmatrix} \tag{15}$$

is a Stieltjes measure with the nonnegativity property $\int_0^{t_1} w'(t)dN(\bar{\Pi})w(t) \geq 0$ for all continuous $w(t)$. The three steps of the algorithm just presented can all be applied to the measure, and in the event that the various constancy of rank and differentiability assumptions are valid, one can identify $\bar{\Pi}(t)$ as $\Pi_m(t)$. Therefore, the same constancy of rank and differentiability assumptions which enable the algorithm to be carried through will also guarantee that when $\bar{\Pi}(\cdot)$ is formed by the above procedure it will be differentiable and, accordingly, yield a solution of the realization problem. The limiting procedure from the computational point of view is not attractive.

3) As noted in [25], (5) when associated with the covariance property becomes a time-varying analog of the Kalman–Yakubovic equations [8],[9]. As such, it can be applied to problems such as network synthesis, see e.g., [29], where, incidentally, nonsingularity of $\Pi(\cdot)$ becomes important. More importantly though, we note that with obvious changes, the material of this section applies to the problem of realizing a stationary covariance, with constant F, H, K, and R, via a time-invariant system. The various state-space and output transformations all become time-invariant, and $\dot{\Pi}$ becomes zero. It is known that the time-invariant problem with nonsingular R is much easier to solve, see [5], [30], and the material here provides a systematic way of reducing a singular problem to a nonsingular problem.

4) Treatments of the scalar covariance singular problem due to Brandenburg [17], and subsequently Geesey [14] have relied on converting the singular problem to a nonsingular problem of the same state-space dimension, as in [16], but then showing that the Riccati equation associated with the nonsingular problem could be replaced by a Riccati equation of lower dimension. In essence, the method here carries out this sort of reduction at each step (rather than waiting until a nonsingular problem is encountered), and, moreover, separates this reduction from any requirement of nonsingularity. Also, it is this reduction at each step which guarantees termination of the algorithm; earlier scalar singular results needed a separate proof of termination.

A generalization of the approach of [14],[16],[17] to the scalar problem has been developed in the thesis of Powell [34] for the matrix covariance singular problem, and it is worthwhile to note some similarities and differences between Powell's and our methods. The more important ones seem to be as follows.

a) In both methods, an assumption that there exists some system realizing the prescribed covariance, coupled with differentiability and constancy of rank assumptions, will allow the algorithms to be carried out. On the other hand, Powell does not use our Assumption 2 (our extendability property for the original covariance). Rather, he demands positive definiteness of a nonsingular covariance derived in the course of the algorithm.

b) Powell's algorithm contains a sequence of output transformations and differentiations interlaced; ours interlaces output transformations and state transformations (which involve differentiability of the coordinate basis change matrix, and to this extent involve differentiations). Both procedures require various constancy of rank and differentiability assumptions, but since the procedures diverge after the first step, it is hard to see whether the assumptions are equivalent.

c) In Powell's procedure, there is no reduction of the size of the Π matrix of interest as one proceeds through the algorithm, though as a final step, one can achieve a reduction. Our method may involve a sequence of reductions through the course of the algorithm.

d) As shown in the next section, our method always allows the construction of an invertible system realizing a prescribed covariance, with the inverse actually computable in the course of the algorithm. Powell's method does not always lead to an inverse system, although it appears that this is due to a failure of the inversion algorithm rather than the derivation of a noninvertible realization. (In this connection, it should be noted that Powell's definition of invertibility is slightly different from ours.)

When our algorithm is specialized to the scalar covariance case, it is the issue raised in b) above which again tends to distinguish it from the algorithm of [16].

5) Suppose one knows a system (2) which realizes a certain covariance, and suppose one wants a system associated with $\Pi_m(t)$ realizing the same covariance. (Such a system, as we shall see, has an invertibility property.) This problem can be simply formulated as follows; one seeks the minimum nonnegative definite Π such that

$$\begin{bmatrix} \dot{\Pi} - \Pi F' - F\Pi & \Pi H - (PH + GJ') \\ H'\Pi - (H'P + JG') & JJ' \end{bmatrix} \geq 0$$

or, equivalently, with $Z = \Pi - P \leq 0$, one seeks the minimum Z for which

$$\begin{bmatrix} \dot{Z} - ZF' - FZ + GG' & ZH - GJ' \\ H'Z - JG' & JJ' \end{bmatrix} \geq 0.$$

6) It is clear from the algorithm given that all solutions $\Pi(\cdot)$ satisfying the original $M(\Pi) \geq 0$ of (5) are uniquely determined up to that part satisfying a condition $\hat{M}(\hat{\Pi})$ involving a nonsingular $\hat{R}$:

$$\hat{M}(\hat{\Pi}) = \begin{bmatrix} \dot{\hat{\Pi}} - \hat{\Pi}\hat{F}' - \hat{F}\hat{\Pi} & \hat{\Pi}\hat{H} - \hat{K} \\ \hat{H}'\hat{\Pi} - \hat{K}' & \hat{R} \end{bmatrix} \geq 0.$$

We have recalled that the minimum $\hat{\Pi}(\cdot)$ satisfying this inequality is given by (3), and the question arises of what other $\hat{\Pi}(\cdot)$ satisfy the inequality. Set $Q(t) = \hat{\Pi}(t) - \hat{\Pi}_m(t)$. Then one can show that $\hat{M}(\hat{\Pi}) \geq 0$ is equivalent to

$$\dot{Q} - Q\bar{F}' - \bar{F}Q - Q\hat{H}\hat{R}^{-1}\hat{H}'Q \geq 0 \qquad (16)$$

where $\bar{F} = \hat{F} + (\hat{\Pi}_{\min}\hat{H} - K)R^{-1}\hat{H}'$.

Steady-state versions of (16) are studied in [7], [30], where all possible Q are characterized. It would take us too far afield to provide the time-varying generalizations, some of which are straightforward to obtain. (The general idea is that (16), through not having a constant term, is feasible to deal with. For example, if Q is invertible, (16) yields a *linear* differential inequality in Q^{-1}.)

7) Let us note that a sufficient condition guaranteeing that $\Pi(t')$ is positive definite for some fixed t', where $\Pi(\cdot)$ satisfies (5), is precisely the complete controllability condition

$$\int_0^{t'} \Phi(t',s)K(s)K'(s)\Phi'(t',s)\,ds > 0.$$

(Nonsingularity of $\Pi(\cdot)$ can be important in applications, see e.g., [29].) To see this, first simplify the problem by selecting a coordinate basis in which $F \equiv 0$. Then suppose $\Pi(t')\alpha = 0$ for some α. Because $M(\Pi) \geq 0$, $\dot{\Pi} \geq 0$ and so $\dot{\Pi}(t)\alpha = 0$ for $0 \leq t \leq t'$. Then $\Pi(t)\alpha = 0$ for $0 \leq t \leq t'$ and so, again since $M(\Pi) \geq 0$, $\alpha'[\Pi(t)H(t) - K(t)] = -\alpha'K(t) = 0$ for $0 \leq t \leq t'$. This means the controllability condition fails.

IV. Realization Given the Extendability Property

We start with (1) and Assumption 2. In very broad terms, the strategy is still the same, i.e., we carry out output transformations that may reduce the dimension of the process whose covariance is to be realized, and state-space coordinate basis transformations which allow reduction of the "degree" (dimension of $\Phi(\cdot,\cdot)$ matrix) of the covariance to be realized. These latter transformations are such that if the lower degree covariance is realizable, the higher degree one is realizable.

In case $R(t)$ is nonsingular, realization follows with the aid of $\Pi_m(t)$ defined by (3), as earlier explained. So we concentrate on singular $R(\cdot)$. As before, we shall have the various constancy of rank assumptions, and *we shall actually compute the same matrix* $\Pi(\cdot)$ solving the realization problem. However, *the existence argument*, or the validation of the computation procedure, *is different.*

The procedure is as follows.

Step 1: This is identical with that of the last section. Assumption 3 is used again.

Step 2: This is identical up to the point where we have

$$R(t) = \begin{bmatrix} R_0(t) & 0 \\ 0 & 0_{p_1\times p_1} \end{bmatrix} \qquad \begin{array}{l} H(t) = [H_1(t) \;\; H_2(t) \;\; 0_{n\times(p_1-p)}] \\ K(t) = [K_1(t) \;\; K_2(t) \;\; K_3(t)] \end{array}$$

with $H_2(t)$ of rank p and of p columns. (Note that Assumption 4 is used again.) We need to conclude that we can set $K_3(t) = 0$. Let $y(t) = [y_1'(t)\; y_2'(t)\; y_3'(t)]'$ be the process with covariance $\mathcal{R}(t,\tau)$, partitioned as $H(\cdot)$. Then

$$E[y_3(\tau)y_3'(\tau)] = [0]\Phi(\tau,\tau)[K_3(\tau)] = 0$$

so $y_3(\tau) = 0$ almost everywhere.[2] Therefore, for $t > \tau$,

$$0 = E[y(t)y_3'(\tau)] = \begin{bmatrix} H_1'(t) \\ H_2'(t) \\ 0 \end{bmatrix} \Phi(t,\tau)K_3(\tau)$$
$$= H'(t)\Phi(t,\tau)K_3(t).$$

Consequently, the covariance is unaltered if we set $K_3(\tau) = 0$ for all τ. This leaves the problem of realizing

$$\mathcal{R}(t,\tau) = \begin{bmatrix} R_0 & 0 \\ 0 & 0_{p\times p} \end{bmatrix}\delta(t-\tau) + \begin{bmatrix} H_1'(t) \\ H_2'(t) \end{bmatrix}\Phi(t,\tau)[K_1(\tau) \cdot K_2(\tau)]1(t-\tau) + \cdots \quad (17)$$

where $H_2(t)$ has p columns and rank p. (To this realization, one adjoins $y_3(t) \equiv 0$ to obtain a realization of the original covariance.) If p is zero, this leaves a nonsingular problem, and we are through. Assume then $p \neq 0$.

Step 3: Proceed as earlier to change the coordinate basis so that

$$H(t) = \begin{bmatrix} H_{11}(t) & 0 \\ H_{21}(t) & I_{p\times p} \end{bmatrix}. \quad (18)$$

Next, we show (by a different technique from that used earlier) that K_{22} is symmetric and that we can, at least after an allowable adjustment of K_{12}, assume $\mathfrak{N}[K_{22}] \subset \mathfrak{N}[K_{12}]$. First, let $y(\cdot) = [y_1'(\cdot)\; y_2'(\cdot)]'$ be the process[3] with covariance (17). Observe that

$$E[y_2(t)y_2'(t)] = [0 \;\; I_{p\times p}]\Phi(t,t)\begin{bmatrix} K_{12}(t) \\ K_{22}(t) \end{bmatrix} = K_{22}(t)$$

from which the symmetry and nonnegativity of $K_{22}(t)$ is immediate. Invoking Assumption 5 as earlier, let $S(t)$ be a nonsingular matrix such that $K_{22}(t)S'(t) = [\bar{K}_{22}(t) \;\; 0]$ with $\bar{K}_{22}(t)$ of full rank. Define $\bar{K}_{12}(t)$ and $\tilde{K}_{12}(t)$ by $K_{12}(t)S'(t) = [\bar{K}_{12}(t) \;\; \tilde{K}_{12}(t)]$. Our task is evidently to show that we can take $\tilde{K}_{12}(t) = 0$. Set $\hat{y}(t) = [I \oplus S(t)]y(t)$, and observe that $\hat{y}(t) = [\hat{y}_1'(t) \;\; \hat{y}_2'(t)]'$ where $\hat{y}_2(t)$ has the form $[\bar{y}_2'(t) \;\; 0]$. Then for $t \geq \tau$,

$$E[\hat{y}(t)\hat{y}_2'(\tau)] = S(t)H'(t)\Phi(t,\tau)\begin{bmatrix} K_{12}(\tau) \\ K_{22}(\tau) \end{bmatrix}S'(\tau)$$
$$= S(t)H'(t)\Phi(t,\tau)\begin{bmatrix} \bar{K}_{12}(\tau) & \tilde{K}_{12}(\tau) \\ \bar{K}_{22}(\tau) & 0 \end{bmatrix}.$$

Taking note of the form of $\hat{y}_2'(\tau)$, it follows that $H'(t)\Phi(t,\tau)\tilde{K}_{12}(\tau) = 0$ for $t \geq \tau$. Evidently, the covariance $E[y(t)y'(\tau)]$ would be unaffected if we replace $\tilde{K}_{12}(\tau)$ by 0 for $\tau \leq t$, and in particular for $\tau = t$.

Now define the state-space coordinate basis change of (11), to obtain $\hat{F}$, $\hat{H}$, and $\hat{K}$ with the latter two matrices as in (12). At this stage, dropping the superscript hat,

$$\mathcal{R}(t,\tau) = \begin{bmatrix} R_0 & 0 \\ 0 & 0_{p\times p} \end{bmatrix}\delta(t-\tau) + \begin{bmatrix} H_{11}'(t) & H_{21}'(t) \\ 0 & I \end{bmatrix}\Phi(t,\tau) \cdot \begin{bmatrix} K_{11}(\tau) & 0 \\ K_{21}(\tau) & K_{22}(\tau) \end{bmatrix}1(t-\tau) + \cdots. \quad (19)$$

Digression—Invertibility: A system $\mathcal{S}$ realizing $\mathcal{R}(t,\tau)$ is

[2] It is a moot point whether we should introduce the process $y(\cdot)$ without a further assumption, since at this point we do not know that there is a process $y(\cdot)$ for which $\mathcal{R}(t,\tau) = E[y(t)y'(\tau)]$. However, it is convenient, and here and later, it shortens an argument not involving $y(\cdot)$ which would be roughly as follows. Partition $u(t) = [u_1'(t)\; u_2'(t)\; u_3'(t)]'$ like $H(\cdot)$ and observe that $\int_0^{t_1}\int_0^{t_1} u'(t)\mathcal{R}(t,\tau)u(\tau)dt\,d\tau =$ [terms involving $u_1(\cdot)$ and $u_2(\cdot)$ but not $u_3(\cdot)$]

$$+ \int_0^{t_1} [u_1'(t)\; u_2'(t)]\begin{bmatrix} H_1'(t) \\ H_2'(t) \end{bmatrix}\int_0^t \Phi(t,\tau)K_3(\tau)u_3(\tau)d\tau \geq 0$$

and since $u(\cdot)$ is arbitrary, for nonnegativity one requires

$$\begin{bmatrix} H_1'(t) \\ H_2'(t) \end{bmatrix}\Phi(t,\tau)K_3(\tau) = 0.$$

[3] As before, one could avoid the introduction of $y(\cdot)$ if desired.

termed invertible if from measurements on the output of the system one can obtain causally the system input, which is white noise, and the system initial state. Differentiation may be involved; causality is essential. Observe that up to this point the algorithm has proceeded by nonsingular transformations on the output vector, and by change of the coordinate basis for the state-space of a realizing system. *This means that if the covariance* (19) *has an invertible realization, so does the covariance* (1). Further, the relation between a causal inverse for a system realizing (19) and one realizing (1) is straightforward to obtain from the output transformation matrices. In the remainder of Step 3, which we now describe, we shall introduce another covariance, such that if it possesses an invertible realization, so does (19).

Define a process $\hat{y} = [\hat{y}_1' \quad \hat{y}_2']'$ by

$$\hat{y}_1(t) = y_1(t) - H_{11}'(t)\int_0^t \Psi(t,\sigma)F_{12}(\sigma)y_2(\sigma)d\sigma - H_{21}'(t)y_2(t)$$

$$\hat{y}_2(t) = \dot{y}_2(t) - F_{22}(t)y_2(t) - F_{21}(t)\int_0^t \Psi(t,\sigma)\cdot F_{12}(\sigma)y_2(\sigma)d\sigma \quad (20)$$

where $\Psi(\cdot,\cdot)$ is the transition matrix associated with $\dot{x} = F_{11}(t)x$. Notice that the covariance of $y_2(t)$, being $\Phi_{22}(t,\tau)K_{22}(\tau)1(t-\tau) + \cdots$ is differentiable with respect to t and τ to yield a covariance of $\dot{y}_2$ which possibly contains a delta function term, but no worse. This means that $\dot{y}_2(t)$ and then $\hat{y}_2(t)$ is as well defined as $y_1(t)$, in that it may contain a white noise component but nothing worse.

It is immediate from (20) that $\hat{y}_1(\cdot)$, $\hat{y}_2(\cdot)$ depend causally on $y_1(\cdot)$ and $y_2(\cdot)$: a differential equation description is provided by

$$\begin{aligned} \dot{w} &= F_{11}w + F_{12}y_2 \qquad w(0) = 0. \\ \hat{y}_1 &= y_1 - H_{11}'w - H_{21}'y_2 \\ \hat{y}_2 &= \dot{y}_2 - F_{22}y_2 - F_{21}w. \end{aligned} \quad (21)$$

These equations can also be rearranged in the following way:

$$\begin{bmatrix} \dot{w} \\ \dot{y}_2 \end{bmatrix} = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix}\begin{bmatrix} w \\ y_2 \end{bmatrix} + \begin{bmatrix} 0 \\ I \end{bmatrix}\hat{y}_2 \qquad w(0) = 0$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} H_{11}' & H_{21}' \\ 0 & I \end{bmatrix}\begin{bmatrix} w \\ y_2 \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix}\hat{y}_1. \quad (22)$$

This rearrangement shows that $y_1(\cdot)$ and $y_2(\cdot)$ are obtainable, causally, from $\hat{y}_1(\cdot)$, $\hat{y}_2(\cdot)$, and $y_2(0)$. Consequently, if an invertible realization can be found for $E[\hat{y}(t)\hat{y}'(\tau)]$, one has, with the aid of (22), an invertible realization for $E[y(t)y'(\tau)]$. Further, since (21) are inverse to (22), one can construct the inverse for $\mathcal{R}(\cdot,\cdot)$ by following (21) with an inverse for $\hat{\mathcal{R}}(\cdot,\cdot)$.

A lengthy formal calculation shows that the associated covariance of $\hat{y}(t)$ is

$$\hat{\mathcal{R}}(t,\tau) = \begin{bmatrix} R_0 & H_{21}'K_{22} - K_{21}' \\ K_{22}H_{21} - K_{21} & \dot{K}_{22} - K_{22}F_{22}' - F_{22}K_{22} \end{bmatrix}\delta(t-\tau) + \begin{bmatrix} H_{11}'(t) \\ -F_{21}(t) \end{bmatrix} \times \Psi(t,\tau)[K_{11}(\tau) \cdot F_{12}(\tau)K_{22}(\tau)]1(t-\tau) + \cdots. \quad (23)$$

The covariance $\hat{\mathcal{R}}(t,\tau)$ of (23) is, with a simple reordering, the same as the covariance defined in (14), which is the covariance resulting after reduction of the state-space dimension in Step 3 of the earlier method. In the last section, the determination of $\mathcal{R}(\cdot,\cdot)$ essentially finished the procedure. Here, too, we are almost done: supposing for the moment a realization of $\hat{\mathcal{R}}(t,\tau)$ is available, one cascades with this realization the linear system of (22), taking for the initial state covariance of (22)

$$E\left\{\begin{bmatrix} w(0) \\ y_2(0) \end{bmatrix}[w'(0) \quad y_2'(0)]\right\} = \begin{bmatrix} 0 & 0 \\ 0 & K_{22}(0) \end{bmatrix}.$$

A messy calculation shows then that the output of (22) has covariance (19). Another way of putting the point is to say that if one has a system realizing $\hat{\mathcal{R}}(\cdot,\cdot)$ in (23), i.e., matrices $\hat{G}(\cdot)$ and $\hat{J}(\cdot)$ and a nonnegative definite $\hat{\Pi}_0$ such that

$$\begin{aligned} \dot{\hat{\Pi}} &= \hat{\Pi}F_{11} + F_{11}'\hat{\Pi} + \hat{G}\hat{G}' \quad \hat{\Pi}(0) = \hat{\Pi}_0 \\ \hat{\Pi}[H_{11} \quad -F_{21}'] &= [K_{11} \quad F_{12}K_{22}] - \hat{G}\hat{J}' \qquad (24) \\ \hat{J}\hat{J}' &= \begin{bmatrix} R_0 & H_{21}'K_{22} - K_{21}' \\ K_{22}H_{21} - K_{21} & \dot{K}_{22} - K_{22}F_{22}' - F_{22}K_{22} \end{bmatrix} \end{aligned}$$

then a system realizing $\mathcal{R}(\cdot,\cdot)$ is defined by matrices $G(\cdot)$, $J(\cdot)$, and Π_0 given by

$$G = \begin{bmatrix} \hat{G} \\ -[0_{p\times(r-p)} \quad I_{p\times p}]\hat{J} \end{bmatrix}$$

$$J = \begin{bmatrix} I_{(r-p)\times(r-p)} & 0 \\ 0 & 0_{p\times p} \end{bmatrix}\hat{J} \quad \Pi_0 = \begin{bmatrix} \hat{\Pi}_0 & 0 \\ 0 & K_{22}(0) \end{bmatrix} \quad (25)$$

and these matrices together with the matrix

$$\Pi = \begin{bmatrix} \hat{\Pi} & 0 \\ 0 & K_{22} \end{bmatrix} \quad (26)$$

satisfy the realization equations (2).

The essentials of Step 3 of the realization process are now complete. As for the earlier procedure, the problem of realizing the original $\mathcal{R}(\cdot,\cdot)$ is reduced by Step 3 to that of realizing $\hat{\mathcal{R}}(\cdot,\cdot)$ of lower degree. Reapplication of Steps 1–3 will cause further degree, and possibly output dimension, reduction until either a nonsingular covariance is encountered, or one of zero degree or zero output dimension. There are however two caveats. First, the various constancy of rank and differentiability assumptions need to be fulfilled. Second, Assumption 2 needs to be retained for the various covariances arising successively in the procedure. That the extensibility property, other than perhaps the observability part of it, is retained is immediately clear.

The extension of F and H on $(t_1, t_1 + \epsilon]$ allow via (21) extension of the domain of definitions of $\hat{\mathcal{R}}(t,\tau)$, and indeed its nonnegativity. To see that the observability property is retained, one can use the following lemma.

Lemma 1: With

$$H(t) = \begin{bmatrix} H_{11}(t) & 0 \\ H_{21}(t) & I_{p\times p} \end{bmatrix} \text{ and } F(t) = \begin{bmatrix} F_{11}(t) & F_{12}(t) \\ F_{21}(t) & F_{22}(t) \end{bmatrix}$$

with $F_{22}(t)$ of dimension $p \times p$, then Assumption 2 implies

$$\int_{t_1}^{t+\epsilon} \Psi'(t,t_1)[H_{11}(t) \quad -F_{21}'(t)] \begin{bmatrix} H_{11}'(t) \\ -F_{21}(t) \end{bmatrix} \Psi(t,t_1)\, dt > 0$$

for all s in $[t_1, t_1 + \epsilon]$, with $\Psi(\cdot,\cdot)$ the transition matrix of $\dot{x}_1 = F_{11}(t)x_1$.

Proof: Suppose the result is false. Then there exists some nonzero $x_1(t_1)$ such that the solution $x_1(t)$ of $\dot{x}_1(t) = F_{11}x_1(t)$ satisfies $H_{11}'(t)x_1(t) = 0$ and $-F_{21}(t)x_1(t) = 0$ on $[t_1, t_1 + \epsilon]$. Then the solution $x(t)$ of $\dot{x} = Fx$ with $x(t_1) = [x_1'(t_1) \quad 0]'$ is evidently $x(t) = [x_1'(t) \quad 0]'$ and satisfies $H'(t)x(t) = 0$. This contradicts Assumption 2.

For nonsingular covariances, it is known that an invertible realization is defined by the minimum $\Pi(t)$, viz., $\Pi_m(t)$. Let us now show how this notion extends to singular problems. The algorithm of Section III shows that the nonuniqueness in choice of $\Pi(t)$ satisfying $M(\Pi) \geq 0$ can all be referred to the nonsingular problem derived in the course of the algorithm, and that $\Pi_m(t)$ for the original problem is given by a minimum $\Pi(t)$ for the nonsingular problem. The material of this section shows that an invertible realization for the nonsingular problem yields an invertible realization for the singular problem. Putting these ideas together with the known nonsingular problem result, it follows that for singular problems too, invertible realizations are associated with minimum $\Pi_m(t)$.

The question arises as to how an inverse system can be built. That for a nonsingular covariance is easily obtained [13], [19]; one would precede this by a cascade of various nondynamic nonsingular transformations, corresponding to Steps 1 and 2 and the bulk of Step 3, together with dynamic systems of the form (21), as noted earlier. It is worth noting that the buildup of the inverse system actually proceeds in parallel with the algorithm for computing $\Pi(t)$. It would be valid, in fact, to view the algorithm as a method for constructing a whitening filter, with a realization of the original covariance matrix being obtained as a byproduct.

We also have a very quick formal solution to the singular filtering problem. Consider the system

$$\begin{aligned} \dot{x} &= Fx + Gu \qquad\qquad E[x(0)x'(0)] = P_0 \qquad (27) \\ y &= H'x + J_1u + J_2v \end{aligned}$$

where $u(\cdot)$ and $v(\cdot)$ are unit intensity Gaussian white noise processes, and $u(\cdot)$, $v(\cdot)$, and $x(0)$ are mutually independent and of zero mean. Assume that (27) is completely observable, in the sense that $H'(t)\Phi(t,\tau)x_0 = 0$ for all $t \geq \tau$ implies $x_0 = 0$. (If this is not the case, a coordinate basis change will separate out the unobservable part of $x(\cdot)$; measurements $y(\cdot)$ are of course useless for estimating this part of $x(\cdot)$, and for filtering purposes, we can confine attention to the observable part.) Then $E[x(t)y'(\tau)] = \Phi(t,\tau)K(\tau)$ for $\tau < t$ where $K = \Pi H + GJ'$ and Π is $E[x(t)x'(t)]$. Suppose the following system with appropriate known initial conditions is a causally invertible realization of $E[y(t)y'(\tau)]$:

$$\begin{aligned} \dot{\hat{x}} &= F\hat{x} + \hat{G}\hat{u} \\ y &= H'\hat{x} + \hat{J}\hat{u}. \end{aligned} \qquad (28)$$

One must then have $E[\hat{x}(t)y'(\tau)] = \Phi(t,\tau)K(\tau)$ for $t > \tau$; for suppose that $E[\hat{x}(t)y'(\tau)] = \Phi(t,\tau)\bar{K}(\tau)$ for $t > \tau$, this being the only possible form on account of (27). Then, for $t > \tau$, $E[y(t)y'(\tau)] = H'(t)E[\hat{x}(t)y'(\tau)] = H'(t)\Phi(t,\tau)\bar{K}(\tau)$. But also, this quantity is $H'(t)\Phi(t,\tau)K(\tau)$. By complete observability, $K(\tau) = \bar{K}(\tau)$.

It follows that $\hat{x}(t) = E[x(t)|y(\tau), \tau < t]$ because, first, $E\{[x(t) - \hat{x}(t)]'y(\tau)\} = 0$ for $\tau < t$, and second, $\hat{x}(t)$ is a function of $\hat{u}(\tau)$, $\tau < t$ and therefore of $y(\tau)$, $\tau < t$ by the causal invertibility. The filtering error is easily seen to be $\Pi - \hat{\Pi}$. The role of $\hat{x}(\cdot)$ in the scalar singular problem—actually in smoothing as well as in filtering problems—has been illuminated in Geesey [14]; for the nonsingular case see [32], [33], [38]. In practice, it is not even necessary to construct the realization (28). As we have seen, the inverse of (28) is actually obtainable directly from the given covariance, and examination of the steps leading up to (21) easily shows that components of $\hat{x}$ may actually be identified as linear combinations of the states of the inverse system.

V. Stationary Covariances

The algorithm of the previous sections applies without change to the realization (over a finite interval) of

$$\begin{aligned} \mathcal{R}(t,\tau) = R\delta(t-\tau) &+ H'e^{F(t-\tau)}K\, 1(t-\tau) \\ &+ K'e^{F'(\tau-t)}H\, 1(\tau - t) \qquad (29) \end{aligned}$$

where F, H, K, and R are constant matrices, and $\mathcal{R}(t,0)$ has a Fourier transform which is nonnegative definite Hermitian for all values of its argument. This approach will, however, lead to time-varying G and J in the realization, and to this extent is unsatisfactory. To obtain a more practical solution, it is desirable to consider the problem of realization over a semiinfinite interval; that is, we allow the system realizing $\mathcal{R}(\cdot,\cdot)$ to start at time $-\infty$.

In this case it is useful (although not entirely essential) to make the following assumption.

Assumption 6: The pair $[F, H]$ is completely observable. This immediately implies that Assumption 2 is satisfied, so that the procedure of Section IV may be carried out. Note the following.

1) Assumptions 3–5 (requiring constancy of rank and differentiability of certain quantities) are always satisfied in the stationary case.

2) At no point in these steps of the algorithm preceding

the obtaining of a nonsingular R matrix do time-varying matrices appear.

3) By Lemma 1, Assumption 6 continues to be satisfied at every stage of the algorithm.

Finally, then, a nonsingular covariance defined by matrices $\hat{F}$, $\hat{H}$, $\hat{K}$, and $\hat{R}$ remains to be realized. Now let

$$\hat{P} = \lim_{t_0 \to -\infty} \Pi(t,t_0) = \lim_{t \to \infty} \Pi(t,t_0)$$

where

$$\dot{\Pi} = \Pi\hat{F}' + \hat{F}\Pi + (\Pi\hat{H} - \hat{K})\hat{R}^{-1}(\Pi\hat{H} - \hat{K}) \qquad \Pi(t_0,t_0) = 0.$$

The limit exists by Assumption 6 as shown in e.g., [30]. Of course, $\hat{P}$ is constant and satisfies an algebraic Riccati equation. Defining $\hat{G} = (\hat{K} - \hat{P}\hat{H})\hat{R}^{-\frac{1}{2}}$ and $\hat{J} = \hat{R}^{\frac{1}{2}}$ essentially completes the procedure. Alternatively, the approach of Section III may be used, since we can now justify Assumption 1.

Example: Since the algorithm is suited more to an efficient computer implementation than to hand calculation, a detailed example is difficult to present. However, the simple example below illustrates some of the more important points.

Consider the covariance (29) with

$$F = [-1] \quad H = [-1 \quad 1]$$

$$K = [0 \quad 2] \quad R = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Then the calculations of Section III proceed as follows.

Step 1: $r = 2$, $p_1 = 1$. No basis change is necessary.

Step 2: $H_2 = [1]$, which is already of full rank. Again, no change is necessary.

Step 3: In the notation of Section III, we have $K_{21} = [0]$, $K_{22} = [2]$, and K_{11} and K_{12} have zero rows. Finally, then

$$P = K_{22} = [2].$$

Note that no Riccati equation needed to be solved for this example, since P became completely determined in Step 3.

Now

$$M(P) = \begin{bmatrix} 4 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} [2 \quad 1 \quad 0].$$

So we identify $G = [2]$ and $J = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

VI. Conclusion

We have presented a procedure for realizing singular, finite-dimensional matrix covariance functions, which at the same time provides a new, and simpler, approach to the scalar, singular covariance factorization problem. It is easy to construct a causally invertible realization, and we have shown that the associated state covariance matrix is the minimum matrix at every time t over the set of such matrices associated with all realizations of the covariance.

To be sure, we do require some technical constancy-of-rank and differentiability assumptions for the ideas to go through. But it should be noted that existing treatments of the scalar singular problem [14], [16], [17] also require such assumptions.

As earlier commented, there are some interesting connections to singular control problems which we shall expound separately; here, one is interested in the dual of the Π matrix associated with the invertible realization, but one is also interested in determining the optimal controls and singular strips. Connection to time-varying network synthesis problems can be found in [29]; the network parallel of the step involving reduction of the state-space dimension of the covariance to be synthesized is the series or shunt extraction of inductor or capacitor elements.

Appendix I

Suppose there exists a nonnegative definite symmetric P_0 together with $G(\cdot)$, $J(\cdot)$ defined on $[0,t_1]$ such that

$$\dot{P} = PF' + FP + GG' \qquad P(0) = P_0$$

$$PH = K - GJ'$$

$$JJ' = R.$$

Observe that these equations imply

$$M(P) = \begin{bmatrix} \dot{P} - PF' - FP & PH - K \\ (PH - K)' & R \end{bmatrix} = \begin{bmatrix} G \\ -J \end{bmatrix} \cdot [G' \quad -J'] \geq 0$$

and so

$$\dot{P} - PF' - FP - (PH - K)R^{-1}(PH - K)' \geq 0.$$

Now use the definition (3) of Π_m; set $Z = P - \Pi_m$ to obtain

$$\dot{Z} - Z[F' + HR^{-1}(H'\Pi_m - K')] - [F + (\Pi_m H - K)R^{-1} \cdot H']Z - ZHR^{-1}H'Z \geq 0 \qquad Z(0) = P_0 \geq 0.$$

It is immediate that $Z(t) \geq 0$ and so $\Pi_m(t) \leq P(t)$. Also, the definition of $\Pi_m(\cdot)$ implies $0 \leq \Pi_m(t)$.

Notice also that the matrix $\Pi_m(\cdot)$ defined by (3) has been shown by the above argument to be minimum amongst all those matrices $\Pi(\cdot) \geq 0$ satisfying $M(\Pi) \geq 0$; by minimum, we mean for all t and such $\Pi(\cdot)$, $\Pi_m(t) \leq \Pi(t)$.

An alternative approach to establishing that $0 \leq \Pi_m(t) \leq P(t)$ is available using the results of [38]. Provided one establishes the existence of an innovations representation realizing $\mathcal{R}(\cdot,\cdot)$ independently of the Riccati equation solution bounding procedure above, one can show that Π_m is actually $E[\hat{x}(t)\hat{x}'(t)]$, where $\hat{x}(t)$ is both the state of this innovations representation and the mean of $x(t)$ conditioned on measurement of a sample function of the process with covariance $\mathcal{R}(\cdot,\cdot)$. Then $0 \leq \Pi_m(t) \leq P(t)$ is immediate. The better technique to be used for establishing this inequality is a function of the background of the reader.

The interpretation of $\Pi_m(t)$ as $E[\hat{x}(t)\hat{x}'(t)]$ also establishes its minimality as a solution of (5).

Appendix II

Our task is to prove the following result. Let $\mathcal{R}(\cdot,\cdot)$ be defined via

$$\mathcal{R}(t,\tau) = R(t)\delta(t-\tau) + H'(t)\Phi(t,\tau)K(\tau)1(t-\tau) + K'(t)\Phi'(\tau,t)H(\tau)1(\tau-t) \quad (1)$$

with $R(t)$ nonsingular for all t. Then the following three statements are equivalent.

1) $\mathcal{R}(t,\tau) - \eta\delta(t-\tau)$ for some $\eta > 0$ is nonnegative on $[0,t_1]$.

2) $\int_0^{t_1}\int_0^{t_1} u'(t)\mathcal{R}(t,\tau)u(\tau)\, dtd\tau = 0$ and $\mathcal{R}(\cdot,\cdot)$ nonnegative on $[0,t_1]$ implies $u(t) \equiv 0$ for $u(\cdot)$ continuous.

3) $\mathcal{R}(\cdot,\cdot)$ has the extendability property described in Assumption 2.

To show the equivalence, we shall use the following lemma.

Lemma: Let $\mathcal{R}(\cdot,\cdot)$ be as defined in (1), and let $\Psi(t,\tau)$ be the transition matrix defined by

$$\frac{\partial}{\partial t}\Psi(t,\tau) = \begin{bmatrix} F - KR^{-1}H' & -KR^{-1}K' \\ HR^{-1}H' & -F' + HR^{-1}K' \end{bmatrix}\Psi(t,\tau)$$

$$\Psi(\tau,\tau) = I$$

and partition Ψ conformably. Then $\int_0^T \mathcal{R}(t,\tau)u(\tau)d\tau = 0$, $t \in [0,T]$, for some T and $u(\cdot)$ not identically zero if and only if $\Psi_{22}(T,0)$ is singular; further, for such a $u(\cdot)$, $\int_0^T \Phi(T,\tau)K(\tau)u(\tau)d\tau \neq 0$.

Proof: Suppose $\int_0^T \mathcal{R}(t,\tau)u(\tau)d\tau = 0$ for $t \in [0,T]$. Set $x(t) = \int_0^t \Phi(t,\tau)K(\tau)u(\tau)d\tau$, so that $\dot{x} = Fx + Ku$, $x(0) = 0$, and $p(t) = \int_t^T \Phi'(\tau,t)H(\tau)u(\tau)d\tau$, so that $\dot{p} = -F'p - Hu$, $p(T) = 0$. It follows that $R(t)u(t) + H'(t)x(t) + K'(t)p(t) = 0$, so that

$$\begin{bmatrix} \dot{x} \\ \dot{p} \end{bmatrix} = \begin{bmatrix} F - KR^{-1}H' & -KR^{-1}K \\ HR^{-1}H' & -F' + HR^{-1}K' \end{bmatrix}\begin{bmatrix} x \\ p \end{bmatrix}. \quad \text{(A1)}$$

Recalling that $x(0) = 0$, we have $p(t) = \Psi_{22}(t,0)p(0)$ and so $p(T) = 0 = \Psi_{22}(T,0)p(0)$. Notice that $p(0) \neq 0$, for otherwise $x(t)$, $p(t)$ and so $u(t)$ would be zero for all t.

Conversely, with $\Psi_{22}(T,0)$ singular, choose $p(0) \neq 0$ so that $\Psi_{22}(T,0)p(0) = 0$. Set $x(t) = \Psi_{22}(t,0)p(0)$, $p(t) = \Psi_{22}(t,0)p(0)$, and verify that $u(t) = -R^{-1}(t)H'(t)x(t) - R^{-1}(t)K'(t)p(t)$ yields $\int_0^T \mathcal{R}(t,\tau)u(\tau)d\tau = 0$ by reversal of the earlier argument.

Because $p(T) = 0$, and $u(t)$ is not identically zero $x(T) = \int_0^T \Phi(T,\tau)K(\tau)u(\tau)d\tau \neq 0$.

Corollary: If $R(t) > 0$, condition 2 holds if and only if $\Psi_{22}(t,0)$ is nonsingular for all $t \in [0,t_1]$.

The proof is immediate, and now we can show the three conditions are equivalent.

$1 \Rightarrow 2$: is immediate.

$2 \Rightarrow 3$: by the corollary, $\Psi_{22}(t,0)$ is nonsingular for all $t \in [0,t_1]$. Hence there exists an extension of $F(\cdot)$, $H(\cdot)$, $K(\cdot)$, and $R(\cdot)$ on $(t_1, t_1+\epsilon]$ preserving continuity and nonsingularity of $\Psi_{22}(t,0)$. By the corollary, $\mathcal{R}(\cdot,\cdot)$ defined over $[0, t_1+\epsilon]$ fulfills condition 2, and so satisfies the extendability property.

$2 \Rightarrow 1$: By condition 2, $\Psi(t,0)$ is nonsingular for all $t \in [0,t_1]$. Since $\Psi(\cdot, 0)$ depends continuously on $R(t)$, there exists a suitably small η so that $\Psi(t,0)$ computed with $R(t)$ replaced by $R(t) - \eta I$ is nonsingular for all $t \in [0,t_1]$, and so that $R(t) - \eta I$ is positive definite for all $t \in [0,t_1]$. The result follows by the corollary.

$3 \Rightarrow 2$: Suppose $\mathcal{R}$ is extendable, but that condition 2 is not fulfilled. Let $\bar{u}$ be such that $\int_0^{t_1} \mathcal{R}(t,\tau)\bar{u}(\tau)d\tau = 0$ for $t \in [0,t_1]$. For the moment, let a continuous $u(\cdot)$ be arbitrary on $(t_1, t_1+\epsilon]$ and equal to $k\bar{u}$ for some constant k to be specified on $[0,t_1]$. Then

$$\begin{aligned} 0 &\leq \int_0^{t_1+\epsilon}\int_0^{t_1+\epsilon} u'(t)\mathcal{R}(t,\tau)u(\tau)\, dtd\tau \\ &= \int_{t_1}^{t_1+\epsilon}\int_{t_1}^{t_1+\epsilon} u'(t)\mathcal{R}(t,\tau)u(\tau)\, dtd\tau \\ &\quad + 2k\int_t^{t_1+\epsilon} u'(t)\, dt \int_0^{t_1} \mathcal{R}(t,\tau)\bar{u}(\tau)d\tau \\ &= \int_{t_1}^{t_1+\epsilon}\int_{t_1}^{t_1+\epsilon} u'(t)\mathcal{R}(t,\tau)u(\tau)\, dtd\tau \\ &\quad + 2k\int_{t_1}^{t_1+\epsilon} u'(t)H'(t)\Phi(t,t_1)\, dt \int_0^{t_1} \Phi(t_1,\tau)K(\tau)\bar{u}(\tau)d\tau. \end{aligned}$$

By the Lemma, $\int_0^{t_1} \Phi(t_1,\tau)K(\tau)\bar{u}(\tau)d\tau = x(t_1) \neq 0$. Choose $u(t) = H'(t)\Phi(t,t_1)x(t_1)$ on $(t_1, t_1+\epsilon]$, and observe that for large negative k a contradiction is obtained. (Admittedly, $u(\cdot)$ is discontinuous, but it can be appropriately approximated by continuous functions, and a contradiction still is obtained.)

Note the critical use of the observability part of the extendability definition in the above argument, to ensure that the second term on the right of the last equality is guaranteed nonzero.

Finally, let us see why satisfaction of conditions 1, 2, or 3 guarantees that the solution of the Riccati equation (3) exists on $[0,t_1]$. Manipulation will show that the quantity $\Psi_{12}(t,0)\Psi_{22}^{-1}(t,0)$ satisfies (3), including the boundary condition. By a standard uniqueness theorem, one must have $\Pi_m(t) = \Psi_{12}(t,0)\Psi_{22}^{-1}(t,0)$, and then the Corollary yields existence of $\Pi_m(t)$ on $[0,t_1]$.

References

[1] V. A. Yakubovic, "The solution of certain matrix inequalities in automatic control theory," *Dokl. Akad. Nauk SSSR*, vol. 143, pp. 1304–1307, 1962.

[2] R. E. Kalman, "Lyapunov functions for the problem of Lur'e in automatic control," in *Proc. Nat. Acad. Sci.*, vol. 49, Feb. 1963, pp. 201–205.

[3] ——, "On the new characterization of linear passive systems," in *Proc. 1st Allerton Conf. Circuit and System Theory*, University of Illinois, Urbana, Ill., Nov. 1963, pp. 456–470.

[4] B. D. O. Anderson, "A system theory criterion for positive real matrices," *SIAM J. Contr.*, vol. 5, pp. 171–182, May 1967.

[5] ——, "An algebraic solution to the spectral factorization problem," *IEEE Trans. Automat. Contr.*, vol. AC-12, pp. 410–414, Aug. 1967.

[6] J. B. Moore and B. D. O. Anderson, "The simulation of stationary stochastic processes," in *Proc. Inst. Elec. Eng.* (London), vol. 115, pp. 337–339, Feb. 1968.

[7] B. D. O. Anderson, "The inverse problem of stationary covariance generation," *J. Statist. Phys.*, vol. 1, pp. 133–147, 1969.

[8] R. E. Kalman, "On a new characterization of linear passive systems," in *Proc. 1st Allerton Conf. Circuit and System Theory*, Univ. of Illinois, Urbana, Ill., Nov. 1963, pp. 456–470.

[9] V. A. Yakubovic, "The solution of certain matrix inequalities in automatic control theory," *Dokl. Akad. Nauk SSSR*, vol. 143, pp. 1304-1307, 1962.

[10] I. C. Gohberg and M. G. Krein, "On the factorization of operators in Hilbert spaces," *Amer. Math. Soc. Transl.*, vol. 51, pp. 155-188, 1966.

[11] B. D. O. Anderson, "Time-varying spectral factorization," Stanford Electron. Labs., Stanford, Calif., Rep. SEL-66-107, (TR 6560-8), Oct. 1966.

[12] J. B. Moore, B. D. O. Anderson, and S. G. Loo, "Generation of prescribed nonstationary covariances," in *Proc. Hawaii Int. Conf. System Science*, Jan. 1968, pp. 710-713.

[13] B. D. O. Anderson, J. B. Moore, and S. G. Loo, "Spectral factorization of time-varying covariance functions," *IEEE Trans. Inform. Theory*, vol. IT-15, pp. 550-557, Sept. 1969.

[14] T. Kailath and R. Geesey, "An innovations approach to least-squares estimation, Part V: Innovations representations and recursive estimation in colored noise," *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 435-453, Oct. 1973. (See also R. Geesey, Ph.D. dissertation, Stanford Univ., Stanford, Calif., Dec. 1968.)

[15] R. A. Geesey and T. Kailath, "Applications of the canonical representation to estimation and detection in coloured noise," in *Proc. Symp. Computer Processing in Communications*, Polytech. Inst. Brooklyn, Brooklyn, N. Y., 1969, pp. 435-457.

[15a] T. Kailath, "Fredholm resolvents, Wiener-Hopf equations, and Riccati differential equations," *IEEE Trans. Inform. Theory*, vol. IT-15, pp. 665-672, Nov. 1969.

[16] J. B. Moore and B. D. O. Anderson, "Spectral factorization of time-varying covariance functions: The singular case," *Math. Syst. Theory*, vol. 4, pp. 10-23, 1970.

[17] L. H. Brandenburg, "Shaping filter models for nonstationary random processes," Ph.D. dissertation, Dep. Elec. Eng., Columbia Univ., New York, N. Y., June 1968.

[18] L. H. Brandenburg and H. E. Meadows, "Shaping filter representation of nonstationary coloured noise," *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 26-31, Jan. 1971.

[19] B. D. O. Anderson and J. B. Moore, "State-space descriptions of inverse and whitening filters," in *Proc. Hawaii Int. Conf. System Science*, Jan. 1968, pp. 714-717.

[20] A. Batkov, "Generalization of the shaping filter method to include nonstationary random processes," *Automat. Remote Contr.* (USSR), vol. 20, pp. 1049-1062, Aug. 1959.

[21] E. B. Stear, "Shaping filters for stochastic processes," in *Modern Control Systems Theory*. New York: McGraw-Hill, 1965.

[22] A. E. Bryson and B. E. Johansen, "Linear filtering for time-varying systems using measurements containing coloured noise," *IEEE Trans. Automat. Contr.*, vol. AC-10, pp. 4-10, Jan. 1965.

[23] B. S. Goh, "Necessary conditions for singular extremals involving multiple control variables," *SIAM J. Contr.*, vol. 4, pp. 716-731, 1966.

[24] ——, "The second variation for the singular Bolza problem," *SIAM J. Contr.*, vol. 4, pp. 309-325, 1966.

[25] D. H. Jacobson, "Totally singular quadratic minimization problems," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 651-658, Dec. 1971.

[26] H. L. Van Trees, *Detection, Estimation and Modulation Theory—Part II; Nonlinear Modulation Theory*. New York: Wiley, 1970.

[27] F. R. Gantmacher, *The Theory of Matrices*. New York: Chelsea, 1959.

[28] B. D. O. Anderson, "Partially singular linear-quadratic control problems," *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 407-409, Aug. 1973.

[29] B. D. O. Anderson and P. J. Moylan, "Synthesis of linear time-varying passive networks," *IEEE Trans. Circuits Syst.*, vol. CAS-21, pp. 678-687, Sept. 1974.

[30] B. D. O. Anderson and S. Vongpanitlerd, *Network Analysis and Synthesis—A Modern Systems Theory Approach*. Englewood Cliffs, N. J.: Prentice-Hall, 1973.

[31] B. D. O. Anderson and T. Kailath, "The choice of signal-process models in Kalman-Bucy filtering," *J. Math. Anal. Appl.*, vol. 35, pp. 659-668, Sept. 1971.

[32] B. D. O. Anderson and J. B. Moore, "State estimation via the whitening filter," in *Proc. Joint Automatic Control Conf.*, Ann Arbor, Mich., June 1968, pp. 123-129.

[33] ——, "The Kalman-Bucy filter as a true time-varying Wiener filter," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-1, pp. 119-127, Apr. 1971.

[34] S. R. Powell, "Modeling of nonstationary stochastic processes with application to image restoration," Ph.D. dissertation, Univ. of Southern California, Los Angeles, Calif., Sept. 1973.

[35] J. B. Moore and B. D. O. Anderson, "Extensions of quadratic minimization theory, I: Finite time results," *Int. J. Contr.*, vol. 7, pp. 465-472, 1968.

[36] V. Doležal, "The existence of a continuous basic of a certain linear subspace of E_r which depends on a parameter," *Cas. Pro. Pest. Mat.*, vol. 89, pp. 466-468, 1964.

[37] G. E. Shilov and B. L. Gurevich, *Integral, Measure and Derivative: A Unified Approach*. Englewood Cliffs, N. J.: Prentice-Hall, 1966.

[38] T. Kailath and R. A. Geesey, "An innovations approach to least squares estimation—Part IV: Recursive estimation given lumped covariance functions," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 720-727, Dec. 1971.

[39] B. D. O. Anderson and J. B. Moore, "Extensions of quadratic minimization theory, II: Infinite time results," *Int. J. Contr.*, vol. 7, pp. 473-480, 1968.

Brian D. O. Anderson (S'62-M'66-SM'74) was born in Sydney, Australia in 1941. He received the B.S. degree in pure mathematics and electrical engineering from the University of Sydney, Sydney, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, Calif.

Before joining the University of Newcastle, Newcastle, New South Wales, Australia, in 1967, where he is a Professor and Head of the Department of Electrical Engineering, he was a faculty member of the Department of Electrical Engineering at Stanford University and Staff Consultant at Vidar Corporation, Mountain View, Calif. For the first six months of 1970 he was a Visiting Professor at Southern Methodist University in Dallas, Tex., and from August 1973 through January 1974, he was a Visiting Professor in the Department of Computer and Information Science, University of Massachusetts. His research interests are in control, communication systems, and networks. He is coauthor of two books, *Linear Optimal Control* with J. B. Moore, and *Network Analysis and Synthesis* with S. Vongpanitlerd, both published by Prentice-Hall in 1971 and 1973, respectively.

Peter J. Moylan (M'73) was born in Hamilton, Victoria, Australia, on February 24, 1948. He received the B.E. degree in electrical engineering from the University of Melbourne, Melbourne, Australia, and the M.E. and Ph.D. degrees from the University of Newcastle, Newcastle, New South Wales, Australia.

Since 1972 he has been a Lecturer in Electrical Engineering at the University of Newcastle. His current research interest are in the fields of control of nonlinear systems, pattern recognition, and electrical networks.