

# Spectral Factorization of Time-Varying Covariance Functions

BRIAN D. O. ANDERSON, MEMBER, IEEE, JOHN B. MOORE, MEMBER, IEEE, AND SONNY G. LOO, MEMBER, IEEE

**Abstract**—The determination of the state-space equations of a time-varying finite-dimensional linear system with a prescribed output covariance matrix is considered when the system is excited by Gaussian white-noise inputs. It is shown that a symmetric state covariance matrix provides the key link between the state-space equations of a system and the system output covariance matrix. Furthermore, such a matrix satisfies a linear matrix differential equation if the state-space equations of the system are known, and a matrix Riccati equation if the output covariance matrix of the system is given. Existence results are given for the Riccati equation solution, and discussion of asymptotic solutions of the differential equations is also included.

## I. INTRODUCTION

THE CONCEPT of spectral factorization has become increasingly more important since Wiener's original work on optimal filtering, and has appeared in such diverse areas as network theory, Lyapunov stability and optimal control. A simple statement of the spectral factorization problem, but certainly not the only one, is the following.

Suppose that a linear system is driven by white Gaussian noise and that the covariance of the output is known; state the equations that describe the system.

In its more common form, the problem is confined to stationary situations; the systems under consideration must be time-invariant and excitation by white noise must have commenced infinitely far back in the past. In order that the output covariance be well-defined under these excitation conditions, the system must be asymptotically stable, and, as a consequence, initial conditions, be they stochastic or deterministic, play no part in affecting the output covariance.

It is also more common than not for the system under consideration to be finite-dimensional; this implies that the system possesses a transfer function matrix rational in the Laplace transform variable  $s$ , and that the Laplace transform of the output covariance should also have this property.

When the output covariance is a scalar rather than a matrix, and possesses a rational Laplace transform, specification of a system which will generate this covariance is a simple matter. In the matrix case, the prob-

lem becomes much harder, and several extended algorithms are available for solving the spectral factorization problem [1]–[4].

The spectral factorization problem is still of interest if either the time-invariance or the finite-dimensionality constraints are relaxed. In this paper, finite-dimensionality is retained, but the systems considered are permitted to be time-varying and the covariances permitted to be nonstationary.

We believe that it is important to regard the situation where the input noise is applied from infinitely far back in time as a special case of a more general situation, namely, one where the noise is applied from some initial finite time  $t_0$  and the system is in some initial state, possibly nondeterministic. It seems that this viewpoint may be necessary in some applications, while the  $t_0 = -\infty$  case can be recovered from the finite  $t_0$  case by a limiting operation.

The time-varying spectral factorization problem as applied to finite-dimensional systems has received attention in several places. Darlington [5] has achieved results for very limited classes of systems; Batkov [6] has attempted a recursive method of solution, which is alleged by Stear [7] to break down; Stear offers a reformulation of the problem as one requiring the solution of simultaneous nonlinear integral equations. Both Stear and Zadeh [8] provide good summaries of the work on the problem. Zadeh's formulation of the spectral factorization problem is to seek a scalar function  $h(t, \tau)$ , satisfying

$$\int_{-\infty}^{+\infty} h(t, \lambda) h(\tau, \lambda) d\lambda = R_v(t, \tau) \quad (1)$$

with  $h(t, \tau) = 0$  for  $t < \tau$  and  $R_v(t, \tau) = R_v(\tau, t)$  known. This amounts to seeking the impulse response of a system which, when driven by white noise, has output covariance  $R_v(t, \tau)$ . Zadeh thus does not seek to characterize the system via its state-space equations, in contrast to the technique to be described here.

Both the requirement of solving an equation such as (1), or an equation of the form

$$\int_0^T R_v(t, \tau) a(\tau) d\tau = m(t) \quad 0 \leq t \leq T \quad (2)$$

where  $m(\cdot)$  and  $R_v(\cdot, \cdot)$  are known and  $a(\cdot)$  is unknown, and the requirement of constructing whitening filters are generated by a number of communication theory problems; examples include the detection of signals in

noise, and least squares filtering, prediction and smoothing. Quite commonly the time-varying nature of the problem arises because the stochastic processes considered, even though stationary, may only be of interest, or indeed measurable, over a finite amount of time. As for time-invariant problems, the key to solving such problems lies in the spectral factorization of  $R_v(t, s)$ .

A discussion of such problems for  $R_v(t, s)$  of the form  $f(t)g(s)1(s-t) + f(s)g(t)1(t-s)$  may be found in [9]. The Kalman-Bucy filter [10] is essentially a whitening filter [11], [12] for a process with covariance  $R_v(t, \tau) = R_n(t, \tau) + R_s(t, \tau)$ ; here  $R_n(t, \tau)$  is a delta function and is the covariance of white noise added to a signal  $s$ , which is generated by passing white noise into a linear system of known form, so that the output covariance of this system is  $R_s(t, \tau)$ . Since it is critical to the construction of the Kalman-Bucy filter that the structure of a system generating  $R_s(t, \tau)$  be known, linear least squares estimation via the Kalman-Bucy filter is not a complete generalization of the corresponding time-invariant (Wiener) problem.

Reference [13], which is an application of the present paper, considers the problem of estimating a signal  $s$  whose covariance  $R_s(t, \tau)$  is known [without knowledge of a system generating  $R_s(t, \tau)$ ] given the measurements  $y$  of  $s$ , (consisting of  $s$  with additive white noise), when  $y$  is not available over  $(-\infty, \infty)$ , or  $R_s(t, \tau)$  is not stationary, or the additive noise is not stationary.

An advantage of solving the time-varying spectral factorization problem by specifying the generating system with state-space equations is that the subsequent specification of a whitening filter is much easier, in fact, no inversion of an integral kernel is required, and the parameters of a whitening filter can be given in terms of those of the generating system with calculations no more difficult than matrix inversion [14].

In Reference [15], abstract procedures for whitening filter construction are given, and the emphasis given in these procedures to the state vector suggests that state-space equations are the logical system descriptions to work with to achieve greatest simplicity.

In the material that follows, we first review the procedure for passing from the description of a system via state-space equations together with its input statistics to the covariance of its output. Emphasis is given to the calculation of the covariance  $P(t)$  of  $x(t)$ , where  $x(t)$ , is the state of the system at time  $t$ . This is because such a covariance constitutes the key link between the state-space equations of a system and the system output covariance, both when passing from the system to the covariance, and when passing from the covariance to the system.

It turns out that, given a description of the system,  $P(t)$  satisfies a linear differential equation, and the system output covariance can be simply written in terms of  $P(t)$  and the matrices appearing in the state-space equations of the system. Conversely, given a knowledge of the

covariance of the system output,  $P(t)$  satisfies a nonlinear differential equation (actually a matrix Riccati equation), and the system can be written in terms of  $P(t)$  and matrices appearing in the prescribed covariance. The existence of solutions of the nonlinear differential equation for  $P(t)$  is also discussed.

Attention is paid to both finite initial time problems and to problems where the initial time  $t_0$  approaches  $-\infty$ . The procedures of course work for the time-invariant spectral factorization problem, and have recently been discussed in the literature [16].

The layout of the paper is as follows. Section II reviews earlier work on covariances, with emphasis on passing from a system described in state-space form to the covariance of the output [17], [18]. Section III considers the inverse problem of obtaining a system description from a covariance, and extends the earlier results of one of the authors [19]. Section IV reviews some results of optimal control [20], [21]. Section V establishes the existence of solution to the nonlinear differential equations appearing in Section III by using the optimal control results and the following observation.

If  $R(t, \tau)$  is a covariance matrix on the interval  $[T_1, T_2]$  in the sense that

$$\int_{T_1}^{T_2} \int_{T_1}^{T_2} u'(t) R(t, \tau) u(\tau) dt d\tau \geq 0, \quad (3)$$

for all  $u(\cdot)$ , then  $\hat{R}(t, \tau)$  is also a covariance on  $[-T_2, -T_1]$  where

$$\hat{R}(t, \tau) = R(-t, -\tau). \quad (4)$$

Section VI is an example illustrating some of the theory.

Since the original submission of this paper, a number of further references have come to light [30]–[36]. Reference [30] provides a scheme for solving (1) when  $R_v(t, \tau)$  is the sum of a delta function and a completely continuous kernel. The theory, using a number of results concerning Fredholm integral equations, was applied by Geesey in [31] to the finite-dimensional problem considered here, and he obtained essentially the same results as this paper, with different existence proofs. Reference [32] contains further results along the same lines, including a large number of results dealing with a singular version of the spectral factorization problem (see also [28]). The solution of (2) is considered in [33]. Connections between the infinite-dimensional solution [30] of (1) and the finite-dimensional solution are touched upon briefly in an important paper of Schumitzky [34]. That the solution of (1) is closely related to generalizing the Kalman filter to make it a smoother is a point of view taken by Kailath [35], for whose valuable insights we are immensely grateful. Kailath also drew our attention to the existence of [36], which contains results indispensable in a consideration of singular problems; Geesey has made use of these results in [31]. Finally, unpublished work of Silverman and Anderson deals with the factoring of matrix rather than scalar covariances, including singular problems.

## II. OUTPUT COVARIANCE MATRICES OF LINEAR SYSTEMS

Consider a linear finite-dimensional dynamical system described by the equations

$$\dot{x}(t) = F(t)x(t) + G(t)u(t) \quad (5a)$$

$$y(t) = H'(t)x(t) + J(t)u(t) \quad (5b)$$

where  $x(\cdot)$ ,  $u(\cdot)$ , and  $y(\cdot)$  are the state, input, and output vectors of dimensions  $n \times l$ ,  $r \times l$ , and  $m \times l$ , respectively, and  $F(\cdot)$ ,  $G(\cdot)$ ,  $H(\cdot)$ , and  $J(\cdot)$  are continuous matrices of dimensions  $n \times n$ ,  $n \times r$ ,  $n \times m$ , and  $m \times r$ , respectively. The superscript prime denotes matrix transposition. An alternative description for the system (5) is the impulse-response matrix

$$S(t, \tau) = H'(t)\Phi(t, \tau)G(\tau)1(t - \tau) + J(t)\delta(t - \tau) \quad (6)$$

where  $\delta(t)$  is the Dirac delta function,  $1(t)$  is the unit step function and  $\Phi(t, \tau)$  is the state transition matrix defined by

$$\frac{d}{dt}\Phi(t, \tau) = F(t)\Phi(t, \tau) \quad \Phi(\tau, \tau) = I_n \quad (7)$$

where  $I_n$  is an  $n$ -dimensional identity matrix.

It is well known that the state vector of the system is given by [15]

$$x(t) = \Phi(t, t_0)x(t_0) + \int_{t_0}^t \Phi(t, \sigma)G(\sigma)u(\sigma) d\sigma. \quad (8)$$

Let the initial state  $x(t_0)$  be a random variable. Without loss of generality, it will be assumed that  $x(t_0)$  has zero mean and symmetric covariance matrix  $P_0$ . Note that if  $x(t_0)$  is deterministic, we may view this situation as being a particular case of the stochastic situation. Let the input  $u(t)$  be Gaussian white noise with zero mean and covariance matrix

$$R_u(t, \tau) = I_r \delta(t - \tau). \quad (9)$$

Furthermore, assume that  $x(t_0)$  is statistically independent of  $u(t)$  for all  $t$  so that the expected value of the product of  $x(t_0)$  and  $u(t)$  is equal to zero.

Direct calculation based on (8) and (9) yields

$$E[x(t)x'(\tau)] = \Phi(t, \tau)P(\tau)1(t - \tau) + P(t)\Phi'(\tau, t)1(\tau - t) \quad (10)$$

where

$$\begin{aligned} P(t) &= E[x(t)x'(t)] \\ &= \Phi(t, t_0)P_0\Phi'(t, t_0) \\ &\quad + \int_{t_0}^t \Phi(t, \sigma)G(\sigma)G'(\sigma)\Phi'(t, \sigma) d\sigma. \end{aligned} \quad (11)$$

Equations (10) and (11) hold for  $t$  and  $\tau \geq t_0$ .

When (9) and (10) are combined with (5b) there obtains

$$\begin{aligned} R_v(t, \tau) &= H'(t)\Phi(t, \tau)K(\tau)1(t - \tau) \\ &\quad + K'(t)\Phi'(\tau, t)H(\tau)1(\tau - t) \\ &\quad + J(t)J'(\tau)\delta(t - \tau) \end{aligned} \quad (12)$$

where

$$K(t) = P(t)H(t) + G(t)J'(t). \quad (13)$$

Of course,  $R_v(t, \tau)$  is only defined for  $t, \tau \geq t_0$ .

The calculation of  $R_v$  from  $F$ ,  $G$ ,  $H$ , and  $J$  (or equivalently,  $\Phi$ ,  $G$ ,  $H$ , and  $J$ ) can be viewed as a calculation in the first instance of  $P(t)$ , the covariance of  $x(t)$ , followed by a calculation of  $R_v(t, \tau)$ . The latter follows with no difficulty from  $\Phi$ ,  $G$ ,  $H$ ,  $J$ , and  $P$ , while  $P$  is given by (11). It may also be verified that the following equation defines  $P$ .

$$\dot{P} = PF' + FP + GG' \quad (14)$$

with initial condition  $P(t_0) = P_0$ .

By defining

$$A'(t) = H'(t)\Phi(t, t_1) \quad (15a)$$

$$B(\tau) = \Phi(t_1, \tau)K(\tau) \quad (15b)$$

$$C(t) = J(t)J'(t) \quad (15c)$$

where  $t_1 \geq t_0$  and is arbitrary but fixed, the covariance  $R_v(t, \tau)$  becomes simply

$$\begin{aligned} R_v(t, \tau) &= A'(t)B(\tau)1(t - \tau) \\ &\quad + B'(t)A(\tau)1(\tau - t) + C(t)\delta(t - \tau). \end{aligned} \quad (16)$$

This form highlights the fact that  $R_v(t, \tau)$  results from passing white noise into a *finite-dimensional* system: the multipliers of  $1(t - \tau)$  and  $1(\tau - t)$  are separable functions of  $t$  and  $\tau$ . Finally, note that the original continuity assumptions on  $F(\cdot)$ ,  $G(\cdot)$ ,  $H(\cdot)$ , and  $J(\cdot)$  guarantee that  $A(\cdot)$ ,  $B(\cdot)$ , and  $C(\cdot)$  are continuous.

## III. SPECTRAL FACTORIZATION

The converse situation in which the output covariance matrix is assumed known and the system has to be found will now be considered for covariances of the form  $R_v(t, \tau)$  in (16). Note that a covariance of this form *could* result by passing white noise into a linear finite-dimensional system with no feedthrough term, and adding white noise to the output of this system. This would require  $A'(t)B(\tau)1(t - \tau) + B'(t)A(\tau)1(\tau - t)$  to be a covariance in its own right, and one could pose the problem of finding a system to generate this covariance. Noting though that even if  $R_v(t, \tau)$  in (16) is a covariance,  $A'(t)B(\tau)1(t - \tau) + B'(t)A(\tau)1(\tau - t)$  need not be, we shall seek a system generating  $R_v(t, \tau)$  in the form of (5a)–(5c).

It is evident that the known data are the three matrices  $A(\cdot)$ ,  $B(\cdot)$ , and  $C(\cdot)$ ; it will be assumed that the matrices are continuous and without loss of generality, that  $A(\cdot)$  and  $B(\cdot)$  have a number of rows which is minimal (i.e., linearly independent) over the time interval for which  $R_v(t, \tau)$  is defined. (If this is not initially the case,  $A(\cdot)$  and  $B(\cdot)$  may be replaced by matrices with the minimality property, with the replacement leaving  $R_v(t, \tau)$  unaltered.) The spectral factorization problem requires the determination of  $F$ ,  $G$ ,  $H$ ,  $J$ , and  $P$ , such that (12) and (13) hold subject to the constraint of (14).

It is immediately evident that there can be no unique solution, because  $R_v(t, \tau)$  contains no information about the coordinate basis used for describing the state vector. This means (see [22]) that  $F(t)$  is totally unspecified except for its dimension and a continuity constraint, and an arbitrary choice for  $F(t)$  may be made. However some choices may be more natural than others; thus if  $R_v(t, \tau)$  is stationary, so that  $R_v(t, \tau) = R_v(t - \tau)$ , then  $F$  would normally be chosen such that  $F$ ,  $K$ , and  $H$  in (12) were constant. Again, it seems more natural to choose  $F$  to be bounded, and  $K$  and  $H$  also if possible (see [23]). The system theory literature dealing with state-space realizations of prescribed impulse responses, e.g., [15], [23], and [24] constitutes a good guide to techniques and reasons for selecting a particular  $F$ . The fact that there is no unique solution to the spectral factorization problem in the time-varying case is not especially surprising; even in the time-invariant case, there is an infinity of systems which will generate a prescribed stationary covariance with, in the scalar case, the transfer functions differing by an arbitrary all-pass factor. If a unique solution to the spectral factorization problem is desired, it becomes necessary to specify some extra constraint on the generating system; for example, usually in the time-invariant case the generating system is required to be minimum phase (thus guaranteeing stability of the associated whitening filter), and of minimum order.

The spectral factorization problem now becomes one of finding matrices  $P$ ,  $G$ ,  $H$ , and  $J$ , given matrices  $A$ ,  $B$ ,  $C$ , and  $F$ , (or equivalently  $\Phi$ ), such that (13)–(15) hold. The initial condition for (14) should be of the type  $P(t_0) = P_0$ , with  $P_0$  some nonnegative definite symmetric matrix. The time  $t_0$  should be such that the system (5) need only be defined for  $t \geq t_0$ .

A constructive procedure will now be given that is valid provided that  $C(t)$  is nonsingular for values of  $t$  in the range of interest. In the case where  $C(t)$  is singular, the covariance (16) can be associated in a natural way with singular problems of optimal control [26] and with singular problems in detection theory [27]. We shall consider the singular spectral factorization problem elsewhere [28] as different techniques are required.

With the assumption that  $C^{-1}$  exists, we proceed as follows. The matrices  $H(\cdot)$  and  $K(\cdot)$  follow immediately from (15a) and (15b); we take  $J(t)$  as  $C^{1/2}(t)$ , the unique positive definite symmetric square root of  $C$ . The matrix  $P(t)$  is the solution at time  $t$  of the matrix Riccati equation

$$\begin{aligned} \dot{P} = & P(F' - HC^{-1}K') + (F - KC^{-1}H')P \\ & + PHC^{-1}H'P + KC^{-1}K' \end{aligned} \quad (17)$$

defined for  $t \geq t_0$ , with initial condition  $P(t_0) = P_0$ , with  $P_0$  nonnegative definite symmetric. Finally,

$$G(t) = [K(t) - P(t)H(t)]C^{-1/2}(t). \quad (18)$$

Leaving aside the question of existence of solutions to (17), to be discussed later, let us check other aspects of the correctness of the procedure. Clearly, (15a)–(15c)

and (13) all hold. It remains to check that (14) holds. To see this observe that (17) may be written as

$$\dot{P} = PF' + FP + (K - PH)C^{-1}(K - PH)' \quad (19)$$

which, on using (18), is observed to be (14). The initial condition for (17), viz.  $P(t_0) = P_0$ , forces the same initial condition on (14).

By taking different initial conditions for the Riccati equation (17), and assuming the existence of solutions to (17) for these different initial conditions, different  $P$  and thus  $G$  matrices result. Thus there is a whole class of systems, differing in their  $G$  matrix and initial state covariance matrix, which have the same output covariance. The contribution in the output covariance of the non-deterministic nature of the initial state is in some way traded off against the contribution of the input white noise.

For arbitrary  $P_0$ , standard results from differential equation theory guarantee the existence of a unique solution to (17) in the neighborhood of  $t_0$  because  $F$ ,  $H$ , etc., are continuous and finite-valued. However a more sophisticated approach is required to demonstrate that there is no escape time for the solution, i.e., that the solution exists outside of a (small) neighborhood of  $t_0$ .

Section V is devoted to showing that provided the covariance  $R_v(t, \tau)$  satisfies one of two meaningful but simple restrictions (listed as A2) and A3) in that section), a solution to (17) exists for the initial condition  $P(t_0) = 0$  over the entire interval for which  $R_v(t, \tau)$  is a covariance. It follows from the continuity of the coefficients in (17) that the solutions depend continuously on the initial condition; consequently (17) also has solutions outside of a small neighborhood of  $t_0$  for nonzero, sufficiently small, nonnegative definite symmetric  $P_0$ .

We can now show that solutions cannot normally exist for arbitrary  $P_0$ . Let us proceed on the basis that a spectral factorization has been carried out for a prescribed  $R_v(t, \tau)$  and a certain nonnegative definite symmetric  $P_0$ . We shall show that with  $P_0 \neq 0$ , the existence of a generating system for  $R_v(t, \tau)$  implies that  $R_v(t, \tau)$  satisfies a condition over and above being a covariance. It will become evident that not all covariances  $R_v(t, \tau)$  can satisfy this property.

Accordingly, suppose (5a) and (5b) have been determined as a generating system, with  $E[x(t_0)x'(t_0)] = P_0$ . Equation (11) shows that

$$P(t) = P_1(t) + P_2(t) \quad (20)$$

where for  $t \geq t_0$ ,  $P_1(t)$  and  $P_2(t)$  are both nonnegative definite symmetric, being given by

$$P_1(t) = \Phi(t, t_0)P_0\Phi'(t, t_0) \quad (21a)$$

$$P_2(t) = \int_{t_0}^t \Phi(t, \sigma)G(\sigma)G'(\sigma)\Phi'(t, \sigma) d\sigma. \quad (21b)$$

Further,  $R_v(t, \tau)$  can be expressed as

$$R_v(t, \tau) = R_{v1}(t, \tau) + R_{v2}(t, \tau), \quad (22)$$

where  $R_{v1}$  is the covariance that would result if the input

noise were removed [equivalently  $G = 0$ ], and is given by replacing  $K(t)$  in (12) by  $K_1(t) = P_1(t)H(t)$ ;  $R_{v2}$  is the covariance that would result if  $P_0$  were zero, and is given by replacing  $K(t)$  in (12) by  $K_2(t) = P_2(t)H(t) + G(t)J'(t)$  [see (13)].

The covariance  $R_{v1}$  is actually

$$\begin{aligned} R_{v1}(t, \tau) &= H'(t)\Phi(t, t_0)P_0\Phi'(\tau, t_0)H(\tau) \\ &\quad \cdot [1(t - \tau) + 1(\tau - t)] \\ &= H'(t)\Phi(t, t_0)P_0\Phi'(\tau, t_0)H(\tau). \end{aligned} \quad (23)$$

Now return to the implementation of the spectral factorization procedure. If a  $P_0 \neq 0$  is assumed,  $R_{v1}(t, \tau)$  can immediately be written down, since  $\Phi$  and  $H$  are both known. Then the operator  $R_{v2}$  can be written down as  $R_v - R_{v1}$ . However, if  $P_0$  is chosen arbitrarily,  $R_v - R_{v1}$  cannot be guaranteed to be a covariance. Indeed for large enough  $P_0$ , one would not expect  $R_v - R_{v1}$  to be a covariance. We then observe the following. Suppose  $P_0$  is such that the constructive procedure works; then a system generating  $R_v(t, \tau)$  exists with this  $P_0$ , and therefore  $R_{v2} = R_v - R_{v1}$  is a covariance. On the other hand, if  $P_0$  were such that  $R_v - R_{v1}$  was a priori not a covariance, the constructive procedure would have to fail. The conclusion is therefore that a necessary condition for (17) to have a solution outside of a neighborhood of  $t_0$  with initial condition  $P_0$  is that  $R_v(t, \tau) - R_{v1}(t, \tau)$  [with  $R_{v1}$  given by (23)] be a covariance. Section V shows that, if in addition condition A2) or A3) of Section V holds when applied to  $R_{v2}$ , then the constructive procedure will work.

Practical computation procedures for solving (19) with arbitrary initial conditions are discussed in [29].

#### IV. REVIEW OF OPTIMAL CONTROL RESULT

In [20] the minimization of the performance index

$$V(x_0, u, \hat{t}_0, \hat{t}_1) = \int_{\hat{t}_0}^{\hat{t}_1} (u'\hat{C}u + 2x'\hat{K}u) dt \quad (24)$$

for the system

$$\dot{x} = \hat{F}x + \hat{H}u \quad (25)$$

is considered with finite  $\hat{t}_0$  and  $\hat{t}_1$ . The transition matrix associated with (25) will be denoted by  $\hat{\Phi}$ .

In connection with the solution of this problem, it is necessary to lay down a number of assumptions as follows.

- $\hat{A}1)$   $\hat{F}$ ,  $\hat{H}$ ,  $\hat{K}$ , and  $\hat{C}$  are continuous and finite-valued, and  $\hat{C}$  is positive definite symmetric on  $[\hat{t}_0, \hat{t}_1]$ .
- $\hat{A}2)$  There exists a time  $T_0$  and extensions of  $\hat{F}$ ,  $\hat{H}$ ,  $\hat{K}$ , and  $\hat{C}$  defined on  $[\hat{t}_0 - T_0, \hat{t}_0]$  such that the extended  $\hat{F}$ , etc., are continuous and

$$\begin{aligned} \hat{R}(t, \tau) &= \hat{C}(t) \delta(t - \tau) + \hat{K}'(t)\hat{\Phi}(t, \tau)\hat{H}(\tau)1 \\ &\quad \cdot (t - \tau) + \hat{H}'(t)\hat{\Phi}(\tau, t)\hat{K}(\tau)1(\tau - t) \end{aligned} \quad (26)$$

is a covariance on  $[\hat{t}_0 - T_0, \hat{t}_1]$ ; simultaneously all states at  $\hat{t}_0$  are reachable. This restricts  $\hat{F}$  and  $\hat{H}$  in a certain way (see [24]).

- $\hat{A}3)$   $\hat{R}(t, \tau) - \eta I \delta(t - \tau)$  is a covariance on  $[\hat{t}_0, \hat{t}_1]$  for some positive  $\eta$ , i.e.,  $\hat{R}(t, \tau)$  is positive definite rather than merely nonnegative definite.

We comment that  $\hat{A}3)$  is shown to imply  $\hat{A}2)$  in [20]. The reverse implication may also be demonstrated. Note also that condition  $\hat{A}2)$  may well hold as a result of  $\hat{R}(t, \tau)$  being a covariance on a larger interval than  $[\hat{t}_0, \hat{t}_1]$ , say  $[\hat{t}_0 - T_0, \hat{t}_0]$ , while the reachability condition reduces to requiring simply that  $F$  be of minimal dimension on  $[\hat{t}_0 - T_0, \hat{t}_0]$  (see [24]). Moreover, if the dimension of  $F$  is minimal over  $[\hat{t}_0 - T_0, \hat{t}_1]$ , it will be often minimal over  $[\hat{t}_0 - T_0, \hat{t}_0]$ , for example, if  $\hat{R}(t, \tau)$  is time invariant, or if  $\hat{F}$  and  $\hat{H}$  are periodic or analytic.

It is shown in connection with the derivation of a control law minimizing (24) that if  $\hat{A}1)$  and either  $\hat{A}2)$  or  $\hat{A}3)$  hold, the following Riccati equation has a solution defined on  $[\hat{t}_0, \hat{t}_1]$ .

$$\begin{aligned} -\dot{\hat{P}} &= \hat{P}(\hat{F} - \hat{K}\hat{C}^{-1}\hat{H}') + (\hat{F}' - \hat{H}\hat{C}^{-1}\hat{K}')\hat{P} \\ &\quad - \hat{P}\hat{K}\hat{C}^{-1}\hat{K}'\hat{P} - \hat{H}\hat{C}^{-1}\hat{H}' \end{aligned} \quad (27)$$

with initial condition  $\hat{P}(\hat{t}_1) = 0$ .

Reference [21] discusses problems where in (24)  $\hat{t}_0$  may be  $-\infty$ , or  $\hat{t}_1$  may be  $+\infty$ , or both. The specific extra assumptions required, and the conclusions therefrom, will not be mentioned here.

#### V. EXISTENCE THEORY

To make connection between the optimal control results and the material of Section III, we use the matrices of Section III to define an optimal control problem by

$$\hat{F}(t) = F'(-t) \quad (28a)$$

$$\hat{H}(t) = H(-t) \quad (28b)$$

$$\hat{K}(t) = K(-t) \quad (28c)$$

$$\hat{C}(t) = C(-t). \quad (28d)$$

Associated with this optimal control problem are matrices  $\hat{\Phi}(t, \tau)$ ,  $\hat{R}(t, \tau)$ ,  $\hat{P}(t)$ , times  $\hat{t}_0$  and  $\hat{t}_1$  and conditions A1)–A3). It is not hard to show that

$$\hat{\Phi}(t, \tau) = -\Phi^{-1}(-t, -\tau). \quad (29)$$

It then follows on comparing the definition of  $\hat{R}(t, \tau)$  [see (26)], and  $R_v(t, \tau)$  [see (12)], that

$$\hat{R}(t, \tau) = R_v(-t, -\tau). \quad (30)$$

As remarked in the Introduction, if  $R_v(t, \tau)$  is a covariance on  $(t_0, t_1)$ , then  $R_v(-t, -\tau)$  or  $\hat{R}(t, \tau)$  is a covariance on  $(-t_1, -t_0)$ . This means that under appropriate conditions there exists a matrix  $\hat{P}(\cdot)$  satisfying (27). Make the definition

$$P(t) = -\hat{P}(-t) \quad (31)$$

and then it may be checked that, if  $\hat{P}(\cdot)$  satisfies (27) in  $\hat{t}_0 \leq t \leq \hat{t}_1$  with initial condition  $\hat{P}(\hat{t}_1) = 0$ , then  $P(\cdot)$  satisfies (17) in  $-\hat{t}_1 \leq t \leq -\hat{t}_0$  with initial condition  $P(-\hat{t}_1) = 0$ . We therefore make the identifications  $\hat{t}_0 = -\hat{t}_1$

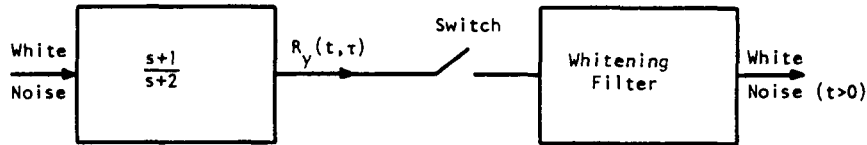


Fig. 1. System discussed in example.

and  $t_1 = -\hat{t}_0$ . Evidently, the existence of a solution to (17) depends on certain of the following conditions being satisfied (these represent conditions equivalent to  $\hat{A}1$ – $\hat{A}3$ ), the equivalences being derived using (28)–(31)).

- A1)  $F$ ,  $H$ ,  $K$  and  $C$  are continuous and finite-valued, and  $C$  is positive definite symmetric on  $[t_0, t_1]$ .
- A2) There exists a time  $T_1$  and extensions of  $F$ ,  $H$ ,  $K$  and  $C$  defined on  $[t_1, t_1 + T_1]$  such that the extended  $F$ , etc., are continuous, and  $R_v(t, \tau)$  given by (12) is a covariance on  $[t_0, t_1 + T_1]$ ; simultaneously all states of the system  $\dot{x} = Fx$  at time  $t_1$  must be observable from the output  $H'x$  over  $[t_1, t_1 + T_1]$ .
- A3)  $R_v(t, \tau) - \eta I \delta(t - \tau)$  is a covariance on  $[t_0, t_1]$  for some positive  $\eta$ , i.e.,  $R(t, \tau)$  is positive definite rather than merely nonnegative definite.

Since A1) has been assumed throughout it follows that the constructive procedure of Section III is valid if either A2) or A3) holds.

One can regard A2) as giving a condition for the existence of a solution to the Riccati equation over the interval  $[t_0, t_1 - T_1]$  for some  $T_1$ , if  $R_v(t, \tau)$  is a covariance on  $[t_0, t_1]$ , and if  $[F, H]$  is an observable pair over  $[t_1 - T_1, t_1]$ . Since  $A(\cdot)$  and  $B(\cdot)$  in  $R_v(t, \tau)$  have been assumed to have a minimal number of rows over  $[t_0, t_1]$ , it may well occur that this minimum number is the same as that which would apply if  $A(\cdot)$  and  $B(\cdot)$  were defined only over  $[t_1 - T_1, t_1]$ . Then (see [24]), the observability condition will be satisfied.

Condition A3) will hold if  $R_v(t, \tau)$  is actually generated by passing white noise into a linear finite-dimensional system without direct feedthrough and then white noise (independent of the input noise) is added at the output of the system, to make  $y$  the sum of the system output and the additional added noise.

Let us now consider cases where  $t_0 = -\infty$ , or  $t_1 = +\infty$ , or both. As for finite-time interval problems, it is possible to dualize the appropriate optimal control results of [21]. In carrying out the dualizing procedure, it is helpful to note that (29) implies that  $\hat{F}$  is uniformly asymptotically stable if and only if  $F$  is uniformly asymptotically stable. The main conclusions follow.

The constructive procedure with initial condition  $\lim_{t \rightarrow -\infty} P(t_0) = 0$  is valid if  $F$ ,  $H$ ,  $K$ ,  $C$ , and  $C^{-1}$  are continuous and finite-valued with  $C = C' > 0$ , and either  $R_v(t, \tau) - \eta I \delta(t - \tau)$  is a covariance on  $(-\infty, t_1)$ ,  $F$  is uniformly asymptotically stable,  $F$  and  $H$  are bounded, or  $F$ ,  $H$ ,  $K$ , and  $C$  can be chosen on  $(t_1, t_1 + T_1)$  for some  $T_1$  so that  $R_v(t, \tau)$  is a covariance on  $(-\infty, t_1)$  and  $F$ ,  $H$  possesses the observability property defined in A3).

*Note:* As remarked in Section III, in passing from  $R_v(t, \tau)$  to a generating system, the  $F$  may be selected arbitrarily. In particular, it may be selected to be uniformly asymptotically stable (see [23] for approaches to ensuring that simultaneously  $F$  and  $H$  are bounded, etc.).

With  $F$  uniformly asymptotically stable and  $F$  and  $H$  bounded, it is clear intuitively (and may be checked by letting  $t_0$  approach  $-\infty$  in some of the Section II formulas), that the initial condition  $P(t_0) = P_0$  as  $t_0 \rightarrow -\infty$  has vanishingly small effect on  $P(t)$  and  $R_v(t, \tau)$ . This means that a boundary condition of the form  $\lim_{t \rightarrow -\infty} P(t) = P_0$ , where  $P_0$  is an arbitrary nonnegative definite symmetric matrix, will yield the same  $P(t)$ , and thus the same system generating  $R_v(t, \tau)$ .

To distinguish the  $P(t)$  resulting from this limiting operation, we shall call it  $\Pi(t)$ . A further parallel with the optimal control results is: with  $\Pi(\cdot)$  existing, with  $F$  uniformly asymptotically stable and  $F$  and  $H$  bounded,  $\dot{x} = [F - HC^{-1}(K' - H'\Pi)]x$  is asymptotically stable. (This may be shown to define the zero-input response of a whitening filter for  $R_v(t, \tau)$  (see [14]). Its stability is analogous to having a minimum phase generating system for  $R_v(t, \tau)$  in the scalar, time-invariant spectral factorization problem.)

Other optimal control results may be used to give comparatively uninteresting bounds on  $P(t)$ ,  $\Pi(t)$  and  $\Pi^{-1}(t)$ .

## VI. EXAMPLE

We consider the problem of constructing a whitening filter for a Gaussian process with covariance

$$R_v(t, \tau) = \delta(t - \tau) - \frac{3}{4}e^{-2t}e^{2\tau}1(t - \tau) - \frac{3}{4}e^{2t}e^{-2\tau}1(\tau - t). \quad (32)$$

The filter is to commence operation at time zero, rather than time  $-\infty$ . (Fig. 1.) Note that  $R_v(t, \tau)$  may be computed to be the output covariance of a single-input, single-output time-invariant system driven by white noise of covariance  $\delta(t - \tau)$ , the transfer function of the system being  $(s + 1)/(s + 2)$ . Note also that  $R_v(t, \tau)$  cannot be represented as the sum of a white-noise term and another covariance not involving a white-noise term, because

$$-\frac{3}{4}e^{-2t}e^{2\tau}1(t - \tau) - \frac{3}{4}e^{2t}e^{-2\tau}1(\tau - t)$$

is not a covariance.

One might be tempted to use as a whitening filter the system

$$\dot{x}_w = -x_w + u_w \quad (33a)$$

$$y_w = x_w + u_w, \quad (33b)$$

which has a transfer function  $(s + 2)/(s + 1)$ . We shall check however, that no matter what initial value of  $E[x_w(0)x'_w(0)]$  is taken, (33) will not suffice to whiten the covariance (32). We shall then define a time-varying filter which will carry out the whitening function effectively, by using the earlier theory.

Let us represent the system with transfer function  $(s + 1)/(s + 2)$  by the state-space equations,

$$\dot{x} = -2x + u, \quad (34a)$$

$$y = -x + u. \quad (34b)$$

It is easy to establish that with  $u$  white noise applied from time  $-\infty$ ,  $E[x(0)x'(0)] = \frac{1}{4}$ . Using the fact that  $u_w$ , the input of the whitening filter, is the same as  $y$ , the output of the generating system, for  $t \geq 0$ , we have

$$\dot{x} = -2x + u \quad (35a)$$

$$\dot{x}_w = -x_w - x + u \quad (35b)$$

$$y_w = x_w - x + u, \quad (35c)$$

from which,

$$\frac{d}{dt}(x_w - x) = -(x_w - x) \quad (35d)$$

with these equations holding for  $t \geq 0$ . It is immediately evident from (35c), that  $y_w$  will only be white noise (i.e.,  $u$ ) if  $x_w = x$  for all  $t$ . From (35d), this requires  $x_w = x$  at time zero. The conclusion is that unless the states of the whitening filter and the generating system can be matched at time zero, true whitening is impossible. However, situations may readily be envisaged where the matching cannot be done for physical reasons; we shall now define a whitening filter that will operate even in these situations.

We shall obtain in state-space form a system generating  $R_v(t, \tau)$ , with the property that  $E[x(0)x'(0)] = 0$ , i.e.,  $P(0) = 0$  in the notations of earlier sections. In terms of the symbols of Section III;  $F = -2$ ,  $K = 1$ ,  $H = -\frac{3}{4}$ ,  $C = 1$ , and (17) becomes

$$\dot{P} = \frac{9}{16}P^2 - \frac{5}{2}P + 1. \quad (36)$$

The solution of this equation satisfying  $P(0) = 0$  is

$$P(t) = \frac{4[e^{2t} - 1]}{9e^{2t} - 1}. \quad (37)$$

The  $G$  matrix of the generating system, from (18), is

$$G(t) = \frac{4[3e^{2t} - 1]}{9e^{2t} - 1}, \quad (38)$$

and for  $t \geq 0$ , the covariance of (32) may be generated by

$$\dot{x} = -2x + \frac{4(3e^{2t} - 1)}{9e^{2t} - 1}u \quad (39a)$$

$$y = -\frac{3}{4}x + u \quad (39b)$$

with  $E[x(0)x'(0)] = 0$ , i.e.,  $x(0) = 0$  and  $u$  white noise of covariance  $\delta(t - \tau)$ .

The following system will now be shown to be a

whitening filter. (The procedure for deducing the system from (39) is given in [14].)

$$\dot{x}_w = \left[ -2 + \frac{3(3e^{2t} - 1)}{9e^{2t} - 1} \right] x_w + \frac{4(3e^{2t} - 1)}{9e^{2t} - 1} u_w \quad (40a)$$

$$y_w = \frac{3}{4}x_w + u_w, \quad (40b)$$

with  $x_w = 0$ . Identifying  $u_w$  and  $y$ , it follows from (39b) and (40a) that

$$\dot{x}_w = \left[ -2 + \frac{3(3e^{2t} - 1)}{9e^{2t} - 1} \right] x_w - \frac{3(3e^{2t} - 1)}{9e^{2t} - 1} x + \frac{4(3e^{2t} - 1)}{9e^{2t} - 1} u. \quad (41)$$

Now subtracting this from (39a), there obtains

$$\frac{d}{dt}(x - x_w) = \left[ -2 + \frac{3(3e^{2t} - 1)}{9e^{2t} - 1} \right] (x - x_w). \quad (42)$$

Since  $x(0) = x_w(0) = 0$ ,  $x = x_w$  for all  $t$ , and from (39b) and (40b), immediately  $y_w = u$ , i.e.,  $y_w$  is white noise.

Note that, as expected, the whitening filter behaves as  $t \rightarrow \infty$  like (33). Note also that correct operation of the whitening filter is independent of the actual system used to generate the covariance (32). We have constructed one such system, namely (39), merely for the purpose of defining the whitening filter.

#### REFERENCES

- [1] D. C. Youla, "On the factorization of rational matrices," *IRE Trans. Information Theory*, vol. IT-7, pp. 172-189, July 1961.
- [2] R. W. Brockett, "Spectral factorization of rational matrices," *IEEE Trans. Information Theory* (to be published).
- [3] M. C. Davis, "Factoring the spectral matrix," *IEEE Trans. Automatic Control*, vol. AC-8, pp. 296-305, October 1963.
- [4] B. D. O. Anderson, "An algebraic solution to the spectral factorization problem," *IEEE Trans. Automatic Control*, vol. AC-12, pp. 410-414, August 1967.
- [5] S. Darlington, "Nonstationary smoothing and prediction using network theory concepts," *Trans. 1959 Internatl. Symp. on Circuit and Information Theory* (Los Angeles, Calif.), pp. 1-13.
- [6] A. Batkov, "Generalization of the shaping filter method to include nonstationary random processes," *Automation and Remote Control*, vol. 20, pp. 1049-1062, August 1959.
- [7] E. B. Stear, "Shaping filters for stochastic processes," in *Modern Control Systems Theory*. New York: McGraw-Hill, 1965.
- [8] L. A. Zadeh, "Time-varying networks, I," *Proc. IRE*, vol. 49, pp. 1488-1503, October 1961.
- [9] T. Kailath, "Some integral equations with 'nonrational' kernels," *IEEE Trans. Information Theory*, vol. IT-12, pp. 442-447, October 1966.
- [10] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Trans. ASME, J. Basic Engrg.*, vol. 83D, pp. 95-108, March 1961.
- [11] L. D. Collins, "Realizable whitening filters and state-variable realizations," *Proc. IEEE* (Correspondence), vol. 56, pp. 100-101, January 1968.
- [12] B. D. O. Anderson and J. B. Moore, "State estimation via the whitening filter," *Proc. Joint Automatic Control Conf.* (Ann Arbor, Mich., June 1968).
- [13] —, "Solution of a time-varying Wiener filtering problem," *Electron. Letters*, vol. 3, pp. 562-563, December 1967.
- [14] J. B. Moore and B. D. O. Anderson, "Whitening filters: A state-space viewpoint," Dept. of Electrical Engineering, University of Newcastle, Newcastle, N. S. W., Australia, Tech. Rept. EE-6707, August 1967.
- [15] L. A. Zadeh and C. A. Desoer, *Linear System Theory*. New York: McGraw-Hill, 1963.
- [16] J. B. Moore and B. D. O. Anderson, "The simulation of stationary stochastic processes," *Proc. IEE* (London), vol. 115, February 1968.
- [17] R. E. Kalman, "Linear stochastic filtering theory-reappraisal and outlook," *Proc. Brooklyn Polytechnic Symp. on System Theory* (New York, 1965), pp. 197-205.

- [18] R. W. Newcomb and B. D. O. Anderson, "The description of coloured noise in dynamical systems," *Internatl. J. Control*, vol. 3, pp. 221-227, 1966.
- [19] B. D. O. Anderson, "Time-varying spectral factorization," Stanford Electronics Labs., Stanford, Calif., Rept. SEL-66-107 (TR 6560-8), October 1966.
- [20] J. B. Moore and B. D. O. Anderson, "Extensions of quadratic minimization theory, I: Finite time results," *Internatl. J. Control*, vol. 7, pp. 465-472, May 1968.
- [21] B. D. O. Anderson and J. B. Moore, "Extensions of quadratic minimization theory, II: Infinite time results," *Internatl. J. Control*, vol. 7, pp. 473-480, May 1968.
- [22] L. M. Silverman and B. D. O. Anderson, "Controllability, observability and stability of linear systems," *J. SIAM on Control*, vol. 6, no. 1, pp. 121-130, 1968.
- [23] L. M. Silverman, "Synthesis of impulse response matrices by internally stable and passive realizations," *IEEE Trans. Circuit Theory*, vol. CT-15, pp. 238-245, September 1968.
- [24] L. Weiss and R. E. Kalman, "Contributions to linear system theory," *Internatl. J. Engrg. Sci.*, vol. 3, pp. 141-171, 1965.
- [25] R. E. Kalman, "Mathematical description of linear dynamical systems," *J. SIAM on Control*, ser. A, vol. 1, no. 2, pp. 152-192, 1963.
- [26] M. Athans and P. L. Falb, *Optimal Control*. New York: McGraw-Hill, 1966.
- [27] T. Kailath, "Some results in singular detection," *Inform. and Control*, vol. 9, pp. 130-152, April 1966.
- [28] J. B. Moore and B. D. O. Anderson, "Spectral factorization of time-varying covariance functions: The singular case," Dept. of Elec. Engrg., University of Newcastle, Newcastle, N. S. W., Australia, Tech. Rept. EE-6811, June 1968.
- [29] R. E. Kalman and T. S. Englar, "A user's manual for ASPC," RIAS Rept. for NASA, Baltimore, Md., 1965.
- [30] I. C. Gohberg and M. E. Krein, "On the factorization of operators in Hilbert spaces," *Am. Math. Soc. Trans.*, vol. 51, pp. 155-188, 1966.
- [31] R. Geesey, "Canonical representations of second-order processes with applications," Ph.D. dissertation, Dept. of Elec. Engrg., Stanford University, Stanford, Calif., December 1968.
- [32] L. H. Brandenburg, "Shaping filter models for nonstationary random processes," Ph.D. dissertation, Dept. of Elec. Engrg., Columbia University, New York, N. Y., June 1968.
- [33] J. B. Moore, "Solution of integral equations occurring in information theory using Riccati equations," *Proc. 2nd Hawaii Internatl. Conf. on System Science* (January 1969), pp. 787-790.
- [34] A. Schumitzky, "On the equivalence between matrix Riccati equations and Fredholm resolvents," *J. Computer and System Sci.*, vol. 2, pp. 76-87, December 1968.
- [35] T. Kailath, "Application of a resolvent identity to a linear smoothing problem," *Proc. 1968 IFAC Symp. on System Dynamics and Automatic Control in Basic Industries* (Sydney, Australia, August 1968), pp. 145-147.
- [36] L. A. Shepp, "Radon-Nikodym derivatives of Gaussian measure," *Ann. Math. Stat.*, vol. 37, pp. 321-354, April 1966.