Spectral Factorization of Time-varying Covariance Functions: The Singular Case

by

JOHN B. MOORE*
and
BRIAN D. O. ANDERSON*

Department of Electrical Engineering
University of Newcastle
New South Wales, 2308
Australia

ABSTRACT

If a known linear system is excited by Gaussian white noise, the calculation of the output covariance of the system is relatively straightforward. This paper considers the harder converse problem, that of passing from a known covariance to a system which will generate it. The problem is solved for covariances $R_y(t, \tau)$ with $|R_y(t, t)| < \infty$ for all t and such that the y-process is Gauss-Markov, i.e., it may be obtained as the output of a linear finite-dimensional system excited by white noise.

1. Introduction

A simple statement of the spectral factorization problem 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.

This problem has been solved in a number of ways for the case when the system is finite-dimensional and time-invariant [1, 2, 3] and has a wide area of application. For the more general case when the system is finite-dimensional, time-varying (with the time-variation such that no actual changes in system structure occur as time evolves) and with the ouput containing a white noise component, the spectral factorization problem has recently been solved [4]. The various theorems involved have also found application in areas of whitening filter theory [5], state-estimation theory [6] and impedance synthesis [7].

This paper complements [4] by considering time-varying spectral factorization results for the case when the specified output covariance, call it $R_y(t, \tau)$, does not contain a white noise component (so that $R_y(t, \tau)$ is finite for all t). This situation will be referred to as the singular case. We comment that the solution of the singular problem is more difficult than that of the nonsingular problem (where

*Work supported by Australian Research Grants Committee; MS submitted June 1968.

$R_y(t, \tau)$ contains a $\delta(t-\tau)$ or white noise part) and is in fact dependent, at least in the developments that follow, on the nonsingular spectral factorization results.

Since [4] gives a detailed discussion of the time-varying spectral factorization problem, its history, and its applications, this information will not be repeated here. In the material that follows, we first review relevant Riccati differential equation theory and various analysis results for systems driven by white noise. The time-varying spectral factorization (nonsingular case) results are then briefly reviewed. Using the results of the introductory sections and further results developed as required, we solve first the spectral factorization problem for the most straightforward case corresponding to a situation where single differentiation of the system output introduces white noise, and then introduce extensions to cover the more general cases when multiple differentiations of the system output are required to obtain white noise.

2. Riccati Equation Theory

In the time-varying spectral factorization procedures of [4] and this paper, Riccati matrix differential equations of the form

$$\dot{P} = P\left(F' - \frac{h_1 k_1'}{j_1^2}\right) + \left(F - \frac{k_1 h_1'}{j_1^2}\right)P + \frac{P h_1 h_1 P}{j_1^2} + \frac{k_1 k_1'}{j_1^2} \tag{1}$$

with a non-negative definite symmetric initial condition $P(t_0)$ are encountered. Both F and P are $n \times n$ matrices, k_1 and h_1 are n-vectors and j_1 is a nonzero scalar and all quantities may be time-varying. Associated with F is its transition matrix $\Phi(t, \tau)$ defined from

$$\frac{d}{dt}\Phi(t, \tau) = F(t)\,\Phi(t, \tau); \qquad \Phi(\tau, \tau) = I. \tag{2}$$

The physical significance of the quantities in (1) will be discussed later as required; meanwhile some results concerning the existence of solutions of (1) given in [4] and based on optimal control results [8, 9] are now summarized.

We first define $R_{y1}(t, \tau)$ and $R_{y1}(t, \tau)|_{(u=0)}$ as

$$R_{y1}(t, \tau) = h_1'(t)\Phi(t, \tau)k_1(\tau)1(t-\tau) + k_1'(t)\Phi'(\tau, t)h_1(\tau)1(\tau-t) + j_1^2(t)\delta(t-\tau) \tag{3}$$

$$R_{y1}(t, \tau)|_{(u=0)} = h_1'(t)\Phi(t, t_0)P(t_0)\Phi'(\tau, t_0)h_1(\tau) \tag{4}$$

where $1(t)$ is the unit step function and $\delta(t)$ is the Dirac delta function. Interpretations of $R_{y1}(t, \tau)$, $R_{y1}(t, \tau)|_{(u=0)}$ and $[R_{y1}(t, \tau) - R_{y1}(t, \tau)|_{(u=0)}]$ will be reviewed in Section 3.

Sufficient conditions for the solution $P(\cdot)$ of (1) to be well defined over $[t_0, t_1]$ are that (A1) and *either* (A2) *or* (A3) be satisfied where:

(A1) $F(\cdot)$, $k_1(\cdot)$, $h_1(\cdot)$ and $j_1(\cdot)$ are finite-valued and continuous with $j_1(t)$ nonzero for all t.

(A2) $[R_{y1}(t, \tau) - R_{y1}(t, \tau)|_{(u=0)} - \eta\delta(t-\tau)]$ is a covariance for some positive η. (For the limiting situations where $t_0 \to -\infty$ and/or $t_1 \to \infty$, the condition (A4) given below must also be satisfied.)

(A3) There exist a time T_1 and extensions of R, h_1, k_1 and j_1 defined on $[t_1, t_1+T_1]$ such that (A1) is satisfied on $[t_1, t_1+T_1]$ and $[R_{y1}(t, \tau)-R_{y1}(t, \tau)|_{(u=0)}]$ given from (3) and (4) 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 an output $y = h_1'x$ over $[t_1, t_1+T_1]$.

(A4) $F(\cdot)$ is uniformly asymptotically stable and $F(\cdot)$ and $H(\cdot)$ are bounded (for the case $[t_0, t_1]$ nonfinite).

3. Analysis Results

Consider the single-input, single-output system having state-space equations

$$\dot{x} = Fx+gu; \qquad y_1 = h_1'x+j_1u \tag{5}$$

(where g is an n-vector), with the initial state $x(t_0)$ being a Gaussian random variable having zero mean and a covariance matrix $P(t_0)$. Suppose also that the system (5) is driven by white noise of zero mean and a covariance $\delta(t-\tau)$.

For the case when $P(t_0)$, F, h_1, j_1 in (5) and some n-vector k_1 are such that (A1) and *either* (A2) *or* (A3) are satisfied, then the solution of (1) will be well defined. This means that if we were to set

$$g = \frac{k_1-Ph_1}{j_1}, \tag{6}$$

then the vector $g(t)$ would be well defined for all $t \in [t_0, t_1]$.

In [4], analysis results may be found which apply to the system (5) with the g vector given by (6). These yield:

$$E[x(t)x'(t)] = P(t), \qquad t \geq t_0, \tag{7}$$

where $P(\cdot)$ is the solution of (1); moreover, the covariance of y is precisely $R_{y1}(t, \tau)$ in (3). Furthermore, if for the system (5), u is set to zero but the initial state covariance is left unchanged, it is straightforward to compute that the output covariance is now the quantity we have called $R_{y1}(t, \tau)|_{(u=0)}$ in equation (4). It is then easy to see that $[R_{y1}(t, \tau)-R_{y1}(t, \tau)|_{(u=0)}]$ would be the output covariance of (5) if the initial state vector $x(t_0)$ is set to zero (i.e., $P(t_0)$ is replaced by the zero matrix) while the input u is again white noise of zero mean and covariance $\delta(t-\tau)$.

Further applications of the analysis result in [4] yield that the output covariance of the system

$$\dot{x} = Fx+gu, \qquad y = h'x \tag{8}$$

with an initial state covariance matrix $P(t_0)$ is

$$R_y(t, \tau) = h'(t)\Phi(t, \tau)P(\tau)h(\tau)1(t-\tau)+h'(t)P(t)\Phi'(\tau, t)h(\tau)1(\tau-t) \tag{9}$$

where g is still given as in (6), h is an n-vector and $P(\cdot)$ is the solution of (1) with initial condition $P(t_0)$.

The above analysis results will be used in Sections (4) and (5).

4. Synthesis Results. Nonsingular Case

A re-interpretation of some of the results discussed in the two preceding sections gives the solution to the time-varying spectral factorization problem for the nonsingular case (see also [4]). These results are now given for later reference.

THEOREM 1. *If $R_{y1}(t, \tau)$ is specified as in* (3) (*i.e.*, $h_1(\cdot)$, $k_1(\cdot)$, $j_1(\cdot)$, $\Phi(\cdot, \cdot)$ *and thus* $F(\cdot)$ *are given*), *then for each* $P(t_0)$ *chosen such that* (1) *has a well defined solution* $P(\cdot)$ (*or such that* (A1) *and either* (A2) *or* (A3) *are satisfied*), *there is a system defined by the quadruple* $F(\cdot)$, $g(\cdot)$, $h_1(\cdot)$, $j_1(\cdot)$ *having the form of* (5) (*with* $g(\cdot)$ *given in terms of* $P(\cdot)$; *see* (6)) *and having the following properties: with an initial state covariance* $P(t_0)$ *and a white noise zero mean input having a covariance* $\delta(t-\tau)$, *the system state covariance* $E[x(t)x'(t)]$ *is* $P(t)$, *the solution of* (1), *and the output covariance is the specified covariance* $R_{y1}(t, \tau)$.

We note that if the covariance $R_{y1}(t, \tau)$ is specified in the following form

$$R_{y1}(t, \tau) = A'(t)B(\tau)1(t-\tau)+B'(t)A(\tau)1(\tau-t)+j_1^2(t)\delta(t-\tau), \tag{10}$$

then an $F(\cdot)$, $h_1(\cdot)$ and $k_1(\cdot)$ may be determined from $A(\cdot)$ and $B(\cdot)$ as discussed in [4].

5. Synthesis Results. Singular Case

The spectral factorization problem is now considered for the case when the specified covariance is

$$R_y(t, \tau) = h'(t)\Phi(t, \tau)k(\tau)1(t-\tau)+k'(t)\Phi'(\tau, t)h(\tau)1(\tau-t), \tag{11}$$

where h and k are n-vectors. It is assumed that $R_y(t, \tau)$ is differentiable in the sense that the $\partial^2 R_y(t, \tau)/\partial t\partial\tau$ exists. Since $R_y(t, \tau)$ is a covariance and $\partial^2 R_y(t, \tau)/\partial t\partial\tau$ exists then $\partial^2 R_y(t, \tau)/\partial t\partial\tau$ also is a covariance [10]. Explicit calculation yields

$$\begin{aligned} \frac{\partial^2 R_y(t, \tau)}{\partial t\partial\tau} &= [\dot{k}'(t)h(t)-2k'(t)F(t)h(t)-\dot{h}'(t)k(t)]\delta(t-\tau) \\ &\quad +[\dot{h}'(t)+h'(t)F(t)]\Phi(t, \tau)[\dot{k}(\tau)-F(\tau)k(\tau)]1(t-\tau) \\ &\quad +[\dot{k}'(t)-k'(t)F'(t)]\Phi'(\tau, t)[\dot{h}(\tau)+F'(\tau)h(\tau)]1(\tau-t). \end{aligned} \tag{12}$$

With the identifications

$$h_1 = \dot{h}+F'h; \qquad k_1 = \dot{k}-Fk; \qquad j_1 = \sqrt{k_1'h-h_1'k}, \tag{13}$$

the covariance $\partial^2 R_y(t, \tau)/\partial t\partial\tau$ becomes identical with the covariance $R_{y1}(t, \tau)$ of (3).

We now state and prove the key lemma.

LEMMA 1. *Consider the case when $R_y(t, \tau)$ is specified as in* (11) ($h(\cdot)$, $k(\cdot)$, $\Phi(\cdot, \cdot)$ *and thus* $F(\cdot)$ *are given*) *over an interval* $[t_0, t_1]$, *and* $\partial^2 R_y(t, \tau)/\partial t\partial\tau$ *exists and is written in the form* (3) *with the identifications* (13) *holding and with* (A1) *satisfied. Then a necessary and sufficient condition for the solution P of* (1) (*assumed to be well defined*) *to satisfy*

$$Ph = k \tag{14}$$

for all $t \in [t_0, t_1]$ *is that the non-negative definite symmetric initial condition* $P(t_0)$ *for* (1) *satisfy*

$$P(t_0)h(t_0) = k(t_0). \tag{15}$$

Furthermore, if (15) *holds, the system* (8) (*see also* (6)) *with an initial state covariance* $P(t_0)$ *when driven by white noise having a covariance* $\delta(t-\tau)$ *has as its state covariance* $E[x(t)x'(t)] = P(t)$ *the solution of* (1), *and as its output covariance the specified covariance* (11).

Proof. If (14) holds, (15) obviously holds. For the converse, suppose that $P(t)$ is the solution of (1) with non-negative definite symmetric initial condition $P(t_0)$ satisfying (15). Some elementary manipulations using (1), (6), (13) and (14) yield

$$\frac{d}{dt}(k-Ph) = \left(F-\frac{gh_1'}{j}\right)(k-Ph); \qquad (k-Ph)|_{t=t_0} = 0. \tag{16}$$

This means that $Ph = k$ for all $t \in [t_0, t_1]$.

The output covariance of the system (8) is given from the analysis results of the previous section as $R_y(t, \tau)$ in (9). When (15) holds, so does (14), and then (9) rewritten using the substitution $Ph = k$ becomes the specified covariance $R_y(t, \tau)$ given in (11). This establishes the lemma.

Following on from the previous lemma, we give two further lemmas which are useful in constructing a non-negative definite symmetric $P(t_0)$ satisfying (15) and such that the solution of (1) will be well defined. When such a $P(t_0)$ is found, then the problem of passing from the covariance (11) to the system (8) with output covariance equal to (11) is solved. The particular $P(t_0)$ constructed is the minimal non-negative definite symmetric $P(t_0)$, written $P_m(t_0)$, which satisfies (15) and has the property that $[P(t_0)-P_m(t_0)]$ is non-negative definite symmetric for all non-negative definite symmetric $P(t_0)$ satisfying (15). To see that such a $P_m(t_0)$ exists, we have

LEMMA 2. *Suppose we are given n-vectors* $h(t_0)$ *and* $k(t_0)$ *for which there exists at least one non-negative definite symmetric matrix* $P(t_0)$ *for which* (15) *holds. Then a non-negative definite symmetric* $P_m(t_0)$, *minimal in the sense above, exists such that* $P_m(t_0)h(t_0) = k(t_0)$. *Moreover* $P_m(t_0) = 0$ *if* $h'(t_0)k(t_0) = 0$ *and otherwise*

$$P_m(t_0) = k(t_0)[k'(t_0)h(t_0)]^{-1}k'(t_0). \tag{17}$$

Proof. For the case $h'(t_0)k(t_0) = 0$, we have that for any $P(t_0)$ satisfying (15), $h'(t_0)P(t_0)h(t_0) = 0$ and thus $P(t_0)h(t_0) = 0$, i.e., $k(t_0) = 0$. Then clearly $P_m(t_0) = 0$ has the required properties.

For the case $h'(t_0)k(t_0) \neq 0$, it is readily checked that $P_m(t_0)$ given by (17) satisfies (15). Consider now an arbitrary n-vector z resolved into the sum of a vector parallel to $k(t_0)$ and a vector in the manifold orthogonal to $k(t_0)$, i.e.,

$$z = \alpha k(t_0) + M\beta \tag{18}$$

$$M'k(t_0) = 0, \tag{19}$$

where α is a scalar, β is an $(n-1)$-vector and M is an $n\times(n-1)$ matrix of rank $(n-1)$ whose columns form a basis in the manifold orthogonal to $k(t_0)$. The vector $h(t_0)$ may also be resolved in a similar manner as

$$h(t_0) = \gamma k(t_0)+M\delta, \tag{20}$$

where γ is a scalar and δ an $(n-1)$-vector. Now $\gamma \neq 0$, since otherwise $h'(t_0)k(t_0) = 0$. This means that z may be written as

$$z = \hat{\alpha}h(t_0)+M\hat{\beta} \tag{21}$$

where $\hat{\alpha} = \alpha/\gamma$ and $\hat{\beta} = \beta-\alpha\delta/\gamma$. Let $P(t_0)$ be any non-negative definite symmetric matrix satisfying (15); then from (17) and 21),

$$\begin{aligned} z'[P(t_0)-P_m(t_0)]z &= [h'(t_0)\hat{\alpha}+\hat{\beta}'M']\{P(t_0) \\ &\quad -k(t_0)[k'(t_0)h(t_0)]^{-1}k'(t_0)\}[\hat{\alpha}h(t_0)+M\hat{\beta}] \\ &= \hat{\beta}'M'P(t_0)M\hat{\beta}. \end{aligned} \tag{22}$$

The second equality follows when we expand and use (15) and (19). Since $P(t_0)$ is non-negative definite symmetric, we may conclude that $[P(t_0)-P_m(t_0)]$ is non-negative definite symmetric and thus $P_m(t_0)$ given by (17) is the required minimal $P(t_0)$ satisfying (15).

A further result, established in [11], relates the existence of solutions to the Riccati equation (1) with differing initial conditions: as a consequence of the non-negativity in (1) of h_1h_1'/j_1^2, the "coefficient" of the term involving P quadratically, the existence of a solution to (1) with a symmetric initial condition $P_1(t_0)$ implies the existence of a solution for any symmetric initial condition $P_2(t_0)$ for which $P_1(t_0)-P_2(t_0)$ is non-negative definite. This result immediately establishes:

LEMMA 3. *If there is one non-negative definite symmetric $P(t_0)$, call it $P_1(t_0)$, satisfying* (15) *and such that the solution of* (1) *with initial condition $P_1(t_0)$ is well defined, then the solution of* (1) *with initial condition $P_m(t_0)$ as defined above will be well defined.*

Lemma 1 may now be modified using the results of Lemmas 2 and 3 to yield:

THEOREM 2. *Suppose a covariance $R_y(t,\tau)$ is specified in* (11) *($h(\cdot)$, $k(\cdot)$, $\Phi(\cdot\,.\,\cdot)$ and thus $F(\cdot)$ are given) over an interval $[t_0, t_1]$; suppose $R_y(t,\tau)$ is differentiable with $R_{y1}(t,\tau) = \partial^2 R_y(t,\tau)/\partial t\partial\tau$ given by* (3), *with the identifications* (13) *holding and condition* (A1) *satisfied. If it is known that the solution of* (1) *is well defined for some (unknown) non-negative definite symmetric initial condition $P(t_0)$ which also satisfies $P(t_0)h(t_0) = k(t_0)$, then an initial condition $P_m(t_0)$ may be chosen as zero for the case $h'(t_0)k(t_0) = 0$ and as* (17) *for the case $h'(t_0)k(t_0) \neq 0$, and the solution of* (1) *with this initial condition $P_m(t_0)$ will be well defined. Moreover, if the system* (8) *(see also* (6)) *resulting from this solution has an initial state covariance $P_m(t_0)$, and the system is driven by white noise having a covariance $\delta(t-\tau)$, the system will have as its output covariance the specified covariance* (11).

This theorem provides a solution to the spectral factorization problem under the following conditions:

(a) The prescribed $R_y(t, \tau)$ is known to have resulted from some system with the F matrix and h vector as predicted from $R_y(t, \tau)$, or:

(a′) The prescribed $R_y(t, \tau)$ is known to have resulted from some system, *and* in (11), the pair F, k is completely reachable at every time t and the pair F, h completely observable at every time t. (Condition (a′) implies condition (a), because the constraints on F, k and h guarantee definition of the state-vector of a system generating $R_y(t, \tau)$ to within an arbitrary coordinate basis change [12], and the existence of solutions to the Riccati equations associated with $R_y(t, \tau)$ is a coordinate free property; see [13].)

(b) The $\delta(t-\tau)$ term in $\partial^2 R_y(t, \tau)/\partial t \partial \tau$ is identically nonzero.

The reasoning used to see that these two conditions guarantee solvability of the spectral factorization problem is as follows. By (b), the Riccati equation (1) can be formed since $j_1^2(t)$ in (1), which is the $\delta(t-\tau)$ term in $\partial^2 R_y(t, \tau)/\partial t \partial \tau$ (see (13)) is everywhere nonzero. By (a), there is some non-negative definite symmetric $P(t_0)$ for which $P(t_0)h(t_0) = k(t_0)$ and which serves as an initial condition for (1)—otherwise there could be no system generating $R_y(t, \tau)$. The ability to form the Riccati equation and its solvability are the two conditions set out in the theorem which guarantee the constructability of a system generating $R_y(t, \tau)$.

The physical interpretation of condition (b) is that a system generating $R_y(t, \tau)$ must have at least one integration in each feedforward path between input and output, and the sum of all path gains through paths consisting of precisely one integration must be nonzero. If this sum is zero (or if there is no path with only one integration), then $j_1(t) = 0$ for all t; this situation will be considered in the next section. A situation where $j_1(t)$ is zero for some t and nonzero for other t is ruled out on the grounds that this would imply a structural change of the underlying system differential equation. Admittedly one can conceive of a time-varying system where such structural changes occur; but the theory here cannot cope with such difficulties.

6. More General Results

In the previous section, the spectral factorization constructive procedure required that the derivative $\partial^2 R_y(t, \tau)/\partial t \partial \tau$ of the specified covariance $R_y(t, \tau)$ include a term $j_1^2(t)\delta(t-\tau)$ with $j_1(t)$ nonzero for all t. We now consider the more general case where we require the mth differentiation of $R_y(t, \tau)$ with respect to t and τ to yield a covariance $R_{ym}(t, \tau) = \partial^{2m} R_y(t, \tau)/\partial t^m \partial \tau^m$ having the form

$$R_{ym}(t, \tau) = h_m'(t)\Phi(t, \tau)k_m(\tau)1(t-\tau) + k_m'(t)\Phi'(\tau, t)h_m(\tau)1(\tau - t) + j_m^2(t)\delta(t-\tau) \tag{23}$$

where for $i = 1, 2, \cdots, m$

$$h_i = \dot{h}_{i-1} + F'h_{i-1}; \quad k_i = \dot{k}_{i-1} - Fk_{i-1}; \quad j_i = \sqrt{k_i'h_{i-1} - h_i'k_{i-1}}. \tag{24}$$

(Note: $h_0 \equiv h$, $k_0 \equiv k$.)
We further require that

(A5) $F(\cdot)$, $k_m(\cdot)$, $h_m(\cdot)$ and $j_m(\cdot)$ are finite valued and continuous with $j_m(t)$ nonzero for all t and $j_i(t) = 0$ $(i = 1, 2, \cdots, m-1)$ for all t.

We shall find that this situation can be handled in principle like that considered earlier ($m = 1$), but the algebra becomes much more involved.

A linear system excited by white noise with an output covariance possessing the above properties must be such that the sum of all path gains through paths including precisely m series integrations is nonzero, and the sums of all path gains through paths including precisely $1, 2, \cdots, m-1$ integrations are zero. If $j_i(\cdot) = 0$, $i = 1, 2, \cdots, m-1$ and $j_m(t)$ is zero for some t, nonzero for other t, this corresponds to the earlier disallowed situation where there are structural changes in the linear system differential equation.

LEMMA 4. *Let m be such that $R_{ym}(t, \tau)$ exists and has the form of* (23), (24) *with* (A5) *satisfied. Then*

$$k_p' h_q = k_q' h_p \tag{25}$$

for $0 \le p \le m-1$, $0 \le q \le m-1$ and for $p = m$, $0 \le q \le m-2$.

Proof. We can assume without loss of generality that $p-q \ge 0$. Clearly the result (25) holds for $p-q = 0$ trivially, and for $p-q = 1$ for $0 \le p \le m-1$ and $0 \le q \le m-1$ by the fact that $j_i(\cdot) = 0$, $i = (1, 2, \cdots, m-1)$ (see (24)). Assume that (25) holds for $p-q = 0, 1, 2, \cdots, r$ $(0 \le p \le m-1, 0 \le q \le m-1)$; we shall show by induction that if $r \ge 1$, (25) holds for $p-q = r+1$, $(0 \le p \le m-1, 0 \le q \le m-1)$. Now

$$k_{q+r}' h_q = k_q' h_{q+r} \tag{26a}$$

$$k_{q+r}' h_{q+1} = k_{q+1}' h_{q+r} \tag{26b}$$

and we may assume that $q+r+1 \le m-1$. Differentiating (26a) and using (24) and (26b), we obtain

$$k_{q+r+1}' h_q = k_q' h_{q+r+1}, \tag{27}$$

and it now becomes clear that (25) may be established for $0 \le p \le m-1$ and $0 \le q \le m-1$ using induction.

Now differentiate (25) with $p = m-1$, $q = 0, 1, \cdots, m-2$. We obtain

$$k_m' h_q + k_{m-1}' h_{q+1} = k_{q+1}' h_{m-1} + k_q' h_m.$$

Now $k_{m-1}' h_{q+1} = k_{q+1}' h_{m-1}$ since $q \le m-2$, and the desired result follows.

We now define $H_m = [h, h_1, h_2, \cdots, h_{m-1}]$ and $K_m = [k, k_1, k_2, \cdots, k_{m-1}]$ and establish a more general form for Lemma 1.

LEMMA 5. *Consider the case when $R_y(t, \tau)$ is specified as in* (11) *($h(\cdot)$, $k(\cdot)$, $(\cdot,\cdot)$ and thus $F(\cdot)$ are given) over an interval $[t_0, t_1]$ and $\partial^{2m} R_y(t, \tau)/\partial t^m \partial \tau^m$ exists and is written in the form* (23) *with the identifications* (24) *holding and* (A5) *satisfied. Then necessary and sufficient conditions for the solution P of the Riccati differential equation*

$$\dot{P} = P\left(F' - \frac{h_m k_m'}{j_m^2}\right) + \left(F - \frac{k_m h_m'}{j_m^2}\right)P + \frac{P h_m h_m' P}{j_m^2} + \frac{k_m k_m'}{j_m^2} \tag{28}$$

with a non-negative definite symmetric initial condition $P(t_0)$ to satisfy $Ph = k$ for *all $t \in [t_0, t_1]$ are that the solution of* (28) *with initial condition $P(t_0)$ be well defined and that $P(t_0)$ satisfy*

$$P(t_0)H_m(t_0) = K_m(t_0). \tag{29}$$

For the case when $Ph = k$ is satisfied for all t, the system

$$\dot{x} = Fx+gu; \qquad y = h'x \tag{30}$$

with initial state covariance $P(t_0)$ and

$$g = \frac{k_m-Ph_m}{j_m}, \tag{31}$$

when driven by white noise having a covariance $\delta(t-\tau)$, has as its state covariance $E[x(t)x'(t)]$ the solution of (28) *and as its output covariance the specified covariance* (11).

Proof. Consider the derivative of (k_i-Ph_i) where $i = 0, 1, 2, \cdots, m-1$:

$$\begin{aligned}\frac{d}{dt}(k_i-Ph_i) &= \dot{k}_i-\dot{P}h_i-P\dot{h}_i \\ &= k_{i+1}+Fk_i-FPh_i-Ph_{i+1}+\frac{Ph_mk'_mh_i}{j_m^2}+\frac{k_mh'_mPh_i}{j_m^2} \\ &\quad -\frac{k_mk'_mh_i}{j_m^2}-\frac{Ph_mh'_mPh_i}{j_m^2} \qquad \text{(using (24) and (28)).}\end{aligned} \tag{32}$$

Application of Lemma 4 ($k'_mh_i = h'_mk_i$ for $0 \le i \le m-2$) and (31) yields

$$\frac{d}{dt}(k_i-Ph_i) = \left(F-\frac{gh'_m}{j_m^2}\right)(k_i-Ph_i)+(k_{i+1}-Ph_{i+1}) \quad \text{for } 0 \le i \le m-2. \tag{33}$$

For the case $i = m-1$, the relation $j_m^2 = k'_mh_{m-1}-h'_mk_{m-1}$ yields, from (32),

$$\frac{d}{dt}(k_{m-1}-Ph_{m-1}) = \left(F-\frac{gh'_m}{j_m^2}\right)(k_{m-1}-Ph_{m-1}). \tag{34}$$

We shall now show using (33) and (34) that a necessary and sufficient condition for $P(t)h(t) = k(t)$ to hold for all t is that (29) hold. Necessity follows by observing from (33) in turn that $Ph = k, \cdots, Ph_{m-1} = k_{m-1}$ for all t. Thus these equations hold for t_0 and (29) holds. Conversely with (29) holding, (34) yields $Ph_{m-1} = k_{m-1}$ for all t, and then (29) and (33) yield in sequence $Ph_{m-2} = k_{m-2}$, $\cdots$, $Ph = k$ for all t.

The second part of the lemma is a straightforward generalization of the result in Lemma 1.

We now give a constructive procedure for determining a particular $P(t_0)$ satisfying (29); this is a generalization of the results of Lemma 3.

LEMMA 6. *Given the $n \times m$ matrices $H_m(t_0)$ and $K_m(t_0)$ such that* (29) *is satisfied for some non-negative definite symmetric $P(t_0)$, then $H'_m(t_0)K_m(t_0)$ is non-*

negative definite symmetric, and a non-negative definite solution of (29) *is provided by*

(35) $$P_m(t_0) = 0 \quad \text{if} \quad H_m'(t_0)K(t_0) = 0$$

and otherwise by

(36) $$P_m(t_0) = K_m(t_0)[H_m(t_0)K_m(t_0)]^{\#} K_m'(t_0),$$

where $^{\#}$ *denotes a pseudo-inverse, i.e., if* $H_m'(t_0)K_m(t_0)$ *is nonsingular, this is the ordinary inverse, and if* $H_m'(t_0)K_m(t_0)$ *is singular, the pseudo-inverse is defined as follows. With* V *any orthogonal matrix such that*

(37) $$H_m'(t_0)K_m(t_0) = V' \begin{bmatrix} \Lambda_s & 0 \\ 0 & 0 \end{bmatrix} V$$

where Λ_s *is a diagonal nonsingular matrix, then*

(38) $$[H_m'(t_0)K_m(t_0)]^{\#} = V' \begin{bmatrix} \Lambda_s^{-1} & 0 \\ 0 & 0 \end{bmatrix} V.$$

Moreover, $P(t_0)-P_m(t_0)$ *is non-negative definite for all non-negative definite symmetric* $P(t_0)$ *satisfying* (29).

Proof. Let $P(t_0)$ be a non-negative definite symmetric matrix satisfying (29). Then $H_m'(t_0)P(t_0)H_m'(t_0) = H_m'(t_0)K_m(t_0)$ is non-negative definite symmetric.

If $H_m'(t_0)K_m(t_0) = 0$, then $P(t_0)\ H_m(t_0) = 0$, i.e., $K_m(t_0) = 0$, and it is clear that $P_m(t_0)$ as specified by (35) has all the desired properties.

If $H_m'(t_0)K_m(t_0)$ is nonsingular, it is straightforward to verify that $P_m(t_0)$ as given by (36) satisfies (29).

If $H_m'(t_0)K_m(t_0)$ is singular and nonzero, we define

(39) $$[H_{m1} : H_{m2}] = H_m(t_0)V'; \qquad [K_{m1} : K_{m2}] = K_m(t_0)\ V'.$$

This means that (37) may be written as

(40) $$\begin{bmatrix} H_{m1}'K_{m1} & H_{m1}'K_{m2} \\ H_{m2}'K_{m1} & H_{m2}'K_{m2} \end{bmatrix} = \begin{bmatrix} \Lambda_s & 0 \\ 0 & 0 \end{bmatrix}$$

and thus $H_{m2}'K_{m2} = 0$.

Since there is some non-negative definite $P(t_0)$ satisfying (29), $0 = H_{m2}'K_{m2} = H_{m2}'P(t_0)H_{m2}$ implies

(41) $$P(t_0)H_{m2} = K_{m2} = 0.$$

Moreover,

$$\begin{aligned}
P_m(t_0)H_m(t_0) &= K_m(t_0)[H_m'(t_0)K_m(t_0)]^{\#} K_m'(t_0)H_m(t_0) && \text{(using (36))} \\
&= K_m(t_0)V' \begin{bmatrix} \Lambda_s^{-1} & 0 \\ 0 & 0 \end{bmatrix} VK_m'(t_0)H_m(t_0) && \text{(using (38))} \\
&= [K_{m1} : K_{m2}] \begin{bmatrix} \Lambda_s^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} K_{m1}' \\ K_{m2}' \end{bmatrix} [H_{m1}VH_{m2}V] && \text{(using (39))} \\
&= K_{m1}\Lambda_s^{-1}K_{m1}'[H_{m1}V : H_{m2}V] \\
&= [K_{m1}V : 0] && \text{(using (40) transposed)} \\
&= K_m(t_0) && \text{(using (39) and (41)).}
\end{aligned}$$

It remains to be shown that $P(t_0)-P_m(t_0)$ is non-negative definite. The proof is a generalization of that given in Lemma 3.

Let z be an arbitrary n-vector. Observing that $H'_{m1}K_{m1} = \Lambda_s$ and that Λ_s is nonsingular, we see that the $n\times s$ matrix K_{m1} has rank s. Let M be an $n\times(n-s)$-matrix whose columns span the manifold orthogonal to that spanned by the colums of K_{m1}. Then

$$z = K_{m1}\alpha+M\beta; \qquad M'K_{m1} = 0 \tag{42}$$

for some s-vector α and $(n-s)$-vector β. Moreover,

$$H_{m1} = K_{m1}C+MD \tag{43}$$

for some $s\times s$ matrix C and $(n-s)\times s$ matrix D; the matrix C is nonsingular because multiplication of (43) on the left by K'_{m1} gives

$$K'_{m1}H_{m1} = K'_{m1}K_{m1}C \tag{44}$$

and both $K'_{m1}H_{m1}$ and $K'_{m1}K_{m1}$ have full rank, viz. s. Hence we may write

$$z = H_{m1}\gamma+M\delta \tag{45}$$

for some s-vector γ and $(m-s)$-vector δ. Now let $P(t_0)$ be an arbitrary matrix satisfying (29). It follows that $P(t_0)H_{m1} = K_{m1}$. Moreover,

$$\begin{aligned} &z'[P(t_0)-P_m(t_0)]z \\ &= (\gamma'H'_{m1}+\delta'M')[P(t_0)-K_{m1}(K'_{m1}H_{m1})^{-1}K'_{m1}](H_{m1}\gamma+M\delta) \\ &= (\gamma'H'_{m1}+\delta'M')P(t_0)(H_{m1}\gamma+M\delta)-\gamma'H'_{m1}K_{m1}\gamma \\ &= \delta'M'P(t_0)M\delta. \end{aligned} \tag{46}$$

The second equality follows from $M'K_{m1} = 0$, and the final equality from $P(t_0)H_{m1} = K_{m1}$ and $M'K_{m1} = 0$. We conclude that $[P(t_0)-P_m(t_0)]$ is non-negative definite and this establishes the lemma.

The results of Lemmas 2, 5, and 6 may now be applied to yield the main result of this section.

THEOREM 3. *Consider the case when $R_y(t, \tau)$ is specified as in* (11), *($h(\cdot)$, $K(\cdot)$, $\Phi(\cdot,\cdot)$ and thus $F(\cdot)$ are given) over an interval $[t_0, t_1]$, and $R_y(t, \tau)$ is differentiable with $R_{ym}(t, \tau) = \partial^{2m}R_y(t, \tau)/\partial t^m\partial\tau^m$ given as in* (23) *with the identifications* (24) *holding and* (A5) *satisfied. Then if it is known that the solution of* (28) *is well defined for some non-negative definite symmetric initial condition, the initial condition $P_m(t_0)$ given by* (35) *or* (36) *is such that the associated solution of* (28) *is well defined. Moreover, if the system* (30) *and* (31) *resulting from this solution has an initial state covariance $P_m(t_0)$ and is driven by white noise having a covariance $\delta(t-\tau)$, the system will have as its state covariance the solution of* (28) *and as its output covariance the specified covariance* (11).

The same remarks, mutatis mutandis, as were made following Theorem 2 may now be made.

Theorem 3 thus essentially completes the solution of the spectral factorization problem, save for the few remarks following on the significance of m.

It is clear that the search for a nonzero $j_m(t)$ implied by the successive differentiations of a prescribed covariance need not be pursued past $m = n$, and if $j_i(t)$ is zero for $i = 0, 1, \cdots, n$ there is no generating system with feedthrough from input to output. This implies that any covariance associated with the system output can only arise from a nonzero initial state-covariance, and the g-vector of a generating system should be taken as zero. The F matrix and h vector are immediately known from the covariance. A prescribed covariance as in (11) must be capable of being written as in (4), and the identification of $P(t_0)$ in (4) when given the form (11) is straightforward.

7. Example

We consider the generation of

(47) $$R(t, \tau) = a(t)b(\tau)1(t-\tau)+b(t)a(\tau)1(\tau-t)$$

for $0 \leq t,\ \tau \leq T$ where a and b are continuous functions, uniquely determined from $R(t, \tau)$ to within an arbitrary constant, by a system of the form

(48) $$\dot{x} = g(t)u \qquad y = a(t)x.$$

Here, the system F matrix is taken as zero; u is of course white noise, and the function $g(t)$ is required together with an initial state covariance for (48).

We form

(49) $$\frac{\partial^2}{\partial t \partial \tau} R(t, \tau) = [\dot{b}(t)a(t)-b(t)\dot{a}(t)]\delta(t-\tau) + \dot{a}(t)\dot{b}(\tau)1(t-\tau)+\dot{b}(t)\dot{a}(\tau)1(\tau-t).$$

Notice that

$$\frac{d}{dt}\left[\frac{b(t)}{a(t)}\right] = [\dot{b}(t)a(t)-b(t)\dot{a}(t)]/a^2(t)$$

and positivity of the coefficient of $\delta(t-\tau)$ corresponds to $b(t)/a(t)$ being strictly increasing, a condition claimed by Doob [14] to be necessary for (47) to be a covariance. That the condition is not necessary follows by noting that if $b(t) = p_0 a(t)$ with constant $p_0 > 0$, then (48) with $g \equiv 0$ and $E[x^2(0)] = p_0$ yields a system generating $R(t, \tau)$.

Suppose now that $\dot{b}(t)a(t)-b(t)\dot{a}(t) > 0$ for all t. Then according to the preceding theory, we form

(50) $$\dot{p} = \frac{(p\dot{a}-\dot{b})^2}{\dot{b}a-b\dot{a}} \qquad p_0 = \frac{b(0)}{a(0)}.$$

The solution of this equation is $p(t) = b(t)/a(t)$. (From the earlier theory it is a consequence of (50) that $p(t)a(t) = b(t)$. Because $a(t)$ and $b(t)$ are scalars, $p(t)$ can be regarded as following from this relation rather than (50); substitution in (50) will of course verify the solution.) The function $g(t)$ is given (see (6)) by

(51) $$g(t) = \frac{p\dot{a}-\dot{b}}{\sqrt{\dot{b}a-b\dot{a}}} = -\frac{\sqrt{\dot{b}a-b\dot{a}}}{a}.$$

8. Conclusion

In this paper, the spectral factorization problem has been solved for linear systems with the following constraints:

(a) The systems are finite-dimensional, with at least one integration in every feedforward path between input and output;
(b) The systems are single-output, and as a consequence of the synthesis procedure are single-input;
(c) No structural changes are allowed in the differential equations of underlying systems.

The natural question arises as to whether any of these assumptions can be removed. There appears to be no straightforward way of extending the ideas of this paper to cope with infinite dimensional systems; indeed, the gap between the difficulties of solving infinite dimensional and finite dimensional problems would have to parallel the gap for the corresponding time-invariant problems; for infinite dimensional problems, sophisticated results of complex variable theory are required while for finite dimensional problems, polynomial factorization will suffice.

The extension of the ideas to multiple-output systems is, by contrast, comparatively straightforward. The main idea is again to use Riccati equations, and again differentiation of a prescribed $R_y(t, \tau)$ is needed in order to generate a $\delta(t-\tau)$ term. Because $R_y(t, \tau)$ is now a matrix, so is the coefficient of the $\delta(t-\tau)$ term, and for the Riccati theory to work, this matrix must be nonsingular. This implies that integers $m_1, m_2, \cdots, m_r$ must be selected, where $R_y(t, \tau)$ is $r \times r$, such that the matrix with $i-j$ term

$$\frac{\partial^{m_i+m_j}}{\partial t^{m_i} \partial \tau^{m_j}} (R_y(t, \tau))_{i,j}$$

has a nonsingular matrix coefficient of the $\delta(t-\tau)$ term. The matrix

$$\frac{\partial^{m_i+m_j}}{\partial t^{m_i} \partial \tau^{m_j}} (R_y(t, \tau))_{i,j}$$

is the covariance of the set

$$\frac{d^{m_i} y_i}{dt^{m_i}} \; (i = 1, 2, \cdots, r).$$

As will be appreciated, the definition of $P_m(t_0)$ becomes considerably more complex, though in principle the same, as for the single output case.

A spectral factorization procedure involving structural changes would appear to be possible if these changes occurred at discrete instants of time. It would be necessary to solve Riccati equations over the time interval between two structural changes, and somehow match boundary conditions for the equations at the end of these intervals.

REFERENCES

[1] D. C. YOULA, On the factorization of rational matrices, *IRE Transactions on Information Theory*, **IT-7**, 3 (1961), 172–189.

[2] R. W. BROCKETT, Spectral factorization of rational matrices, *IEEE Transactions on Information Theory*, to appear.

[3] M. C. DAVIS, Factoring the spectral matrix, *IEEE Transactions on Automatic Control*, **AC-8**, 4 (1963), 296–305.

[4] B. D. O. ANDERSON, J. B. MOORE and S. G. LOO, Spectral factorization of time-varying covariance functions, *IEEE Transactions on Information Theory*, **IT-15**, 5 (1969), 550–557.

[5] B. D. O. ANDERSON and J. B. MOORE, State-space descriptions of inverse and whitening filters, *Proceedings of Hawaii International Conference on System Sciences*, pp. 714–717, 1968.

[6] B. D. O. ANDERSON and J. B. MOORE, State estimation via the whitening filter, *Proceedings of JACC*, 1968, pp. 123–129.

[7] B. D. O. ANDERSON and J. B. MOORE, Procedures for time-varying impedance synthesis, *Proceedings of the Eleventh Midwest Symposium on Circuit Theory*, May 1968, pp. 17–26.

[8] J. B. MOORE and B. D. O. ANDERSON, Extensions of quadratic minimization theory I: Finite time results, *International Journal of Control* **7**, 5 (1968), 465–472.

[9] B. D. O. ANDERSON and J. B. MOORE, Extensions of quadratic minimization theory, II: Infinite time results, *International Journal of Control* **7**, 5 (1968), 473–480.

[10] A. PAPOULIS, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, 1965.

[11] B. D. O. ANDERSON and J. B. MOORE, Convergence properties of Riccati equation solutions, *IEEE Trans. on Automatic Control*, **AC-13**, 6 (1968), 732–733.

[12] R. E. KALMAN, P. L. FALB and M. A. ARBIB, *Topics in Mathematical Systems Theory*, McGraw-Hill, 1969.

[13] M. ATHANS, The relationship of alternate state-space representations in linear filtering problems, *IEEE Trans. on Automatic Control*, **AC-12**, 6 (1967), 775–776.

[14] J. L. DOOB, *Stochastic Processes*, Wiley, New York, 1953.

(*Received 7 January 1969*)