TIME-VARYING SPECTRAL FACTORIZATION*

B. D. O. Anderson

October 1966

Stanford Electronics Laboratories
Stanford University
Stanford, California

*Work supported by the Office of Naval Research under Nonr-225(83).

ABSTRACT

The problem is posed of obtaining the spectral factorization of the covariance of the output of a linear, finite-dimensional, time-varying system driven by white noise. It is shown that a factorization can be found, though possibly not globally, by solving a matrix Riccati equation.

CONTENTS

I. INTRODUCTION

Problems in electrical engineering involving the concept of spectral factorization have proved to be legion. Directly or indirectly, the idea appears in network synthesis, optimal control, optimal estimation and filtering, and Lyapunov stability studies. In its simplest form, the problem of spectral factorization can be stated as follows:

> Given a linear system (possibly time-varying, possibly distributed, possibly multiple-input and output) and given the covariance of the system output when its input is excited by white gaussian noise; find the impulse response of the system.

The great majority of results available deal with time-invariant finite dimensional systems, i.e. those for which the output covariance can be characterized by its Laplace transform, which is a matrix of rational functions of a complex variable s. (Note that it is a non-trivial result of this theory that given a covariance which has a rational transform, there always exists a corresponding finite-dimensional system.)

Procedures for actually carrying out spectral factorization are of practical interest; three approaches are detailed in [1], [2], and [3] for the finite-dimensional time-invariant system case.

The aim of this paper is to present the theory which applies to the finite-dimensional time-varying case. For ease of presentation, attention is restricted to covariances which are scalars rather than matrices; there appears to be no inherent difficulty in extending to the matrix situation if required.

This problem has not escaped the attention of other authors. Darlington [4] has presented a formulation of the problem, and achieved results for very limited classes of systems, which in no way could be called general. Batkov [5] has attempted a recursive method of solution which according to Stear [6] breaks down. Stear himself gives a reformulation of the problem as one involving nonlinear simultaneous integral equations.

It would be a fair comment to say that initial conditions receive too scant attention in any of these treatments. If one conceived of a state-variable representation of the finite-dimensional system, it is necessary to take one of two possible views; either at some finite time t_o, the initial state is specified deterministically or statistically, or that at times under consideration the effect of initial states has died out. This last view requires the imposition of an asymptotic stability requirement, in effect backward in time. A time-varying system asymptotically stable in the usual sense, i.e. forward from some time t_o, may not, however, be asymptotically stable in the sense required to neglect initial conditions (though, of course, a time-invariant system will); thus, one cannot blindly carry over simple ideas of the time-invariant theory to the time-varying theory without perhaps making overly restrictive assumptions.

II. PROBLEM FORMULATION

Consider a linear system described by its impulse response function $w(t,\tau)$. (For simplicity, we shall neglect the associated matrix problem.) If white gaussian noise is applied at the input of the system, i.e. if we excite the system with a random variable whose covariance is $\delta(t-\tau)$, then the covariance of the system output is given by elementary calculations [4], [5], [6], and [7] as

$$R(t,\tau) = \int_{-\infty}^{+\infty} w(t,\lambda)\, w(\tau,\lambda)\, d\lambda \tag{1}$$

always assuming, of course, that the integral exists, and neglecting for the moment problems associated with initial conditions.

The following theorem characterizes covariance functions, $R(t,\tau)$:

Theorem 1. A given function of two variables $R(t,\tau)$ is a covariance function if and only if it is a nonnegative definite kernel, i.e. for all functions $i(\cdot)$ for which the integral exists, the relation

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} R(t,\tau)\, i(t)\, i(\tau)\, dt d\tau \geq 0 \tag{2}$$

holds.

It should be noted that this theorem does not guarantee that if (2) holds, R should be the output covariance of a linear system excited by white noise, but merely that there should be a gaussian stochastic process possessing $R(t,\tau)$ as covariance. See [8] for a proof of the nontrivial 'if' part of the theorem. The 'only if' part is easy to demonstrate.

At this point the general spectral factorization problem can be presented:

Problem 1.

Given a function $R(t,\tau)$ and given that (2) holds for all $i(\cdot)$ for which the integral on the left hand side exists, determine if a relation of the form (1) exists, and if it does, find a $w(t,\tau)$ satisfying (1).

Without apology, we shall pass from this problem to an easier one imposing the finite dimensionality constraint. We recall, see [9], that a linear, time-varying finite-dimensional system possesses an impulse response of the form

$$w(t,\tau) = \varphi'(t)\ \psi(\tau)\ 1(t-\tau) \tag{3}$$

where $\varphi(\cdot)$ and $\psi(\cdot)$ are vector functions of time, the prime denotes matrix transposition, and $1(t)$ is the unit step function, zero for negative values of t, and unity for positive values.

Besides an impulse response of the form (3), we shall also be concerned with systems with an impulse response of the form

$$w(t,\tau) = \mu(t)\ \delta(t-\tau) + \varphi'(t)\ \psi(\tau)\ 1(t-\tau) \tag{4}$$

Here $\mu(\cdot)$ is a scalar function of time, and $\delta(t-\tau)$ is the unit impulse.

The covariance of the output of a system of the form of (3) and (4) can be computed in terms of the relevant w and an assumed initial distribution on the state variables at time t_o. This is for convenience taken with a zero mean. A system with impulse response as in (3) then leads to a covariance of the form:

$$R(t,\tau) = a'(t)\ b(\tau)\ 1(t-\tau) + b'(t)\ a(\tau)\ 1(\tau-t) \tag{5}$$

and a system with impulse response as in (4) leads to a covariance of the form

$$R(t,\tau) = \mu^2(t)\ \delta(t-\tau) + a'(t)\ b(\tau)\ 1(t-\tau) + b'(t)\ a(\tau)\ 1(\tau-t) \tag{6}$$

Of course, it is not true that arbitrary $a(\cdot)$, $b(\cdot)$, and $\mu(\cdot)$ determine a covariance of the form (5) or (6); references [10] and [11] spell out restrictions imposed on $a(\cdot)$, $b(\cdot)$, and $\mu(\cdot)$ by virtue of (5) and (6) constituting the output covariance of a finite-dimensional dynamical system. Two problems unanswered by these references, however, are as follows:

Problem 2.

If $R(t,\tau)$ is a covariance, i.e. (2) holds for all $i(\cdot)$, and if R has the form of (5) or (6), is there a finite-dimensional system such that $R(t,\tau)$ is the covariance of its output, with suitable excitation and initial conditions?

Problem 3.

Assuming the answer to Problem 2 is yes, how may one find the system, or its impulse response?

We shall show later that with certain reasonable conditions on the vector functions, $a(\cdot)$ and $b(\cdot)$, and the scalar function, $\mu(\cdot)$, that both (4) and (5) are the covariance over some finite interval of a linear finite-dimensional system. Whether, however, a linear finite-dimensional system can yield the prescribed covariance globally is another question.

III. THE FORM OF A COVARIANCE

Consider a prescribed $w(t,\tau)$ given by

$$w(t,\tau) = \varphi'(t)\ \psi(\tau)\ 1(t-\tau) \tag{3}$$

Then w is the impulse response of the following system

$$\dot{x} = \psi(t)\ u(t) \tag{7}$$

$$y = \varphi'(t)\ x(t) \tag{8}$$

If at time t_o the state variable x is a random variable with mean zero and covariance the symmetric matrix $P(t_o)$, and if the input is white noise with covariance $\delta(t-\tau)$, calculations detailed in [10] yield

$$R(t,\tau) = E[y(t)y'(\tau)] = \varphi'(t)\left[P(t_o) + \int_{t_o}^{\tau} \psi(\lambda)\psi'(\lambda)d\lambda\right]\varphi(\tau)\ 1(t-\tau)$$

$$+\ \varphi'(t)\left[P(t_o) + \int_{t_o}^{t} \psi(\lambda)\psi'(\lambda)d\lambda\right]\varphi(\tau)\ 1(\tau-t) \tag{9}$$

By defining the symmetric matrix

$$P(t) = P(t_o) + \int_{t_o}^{t} \psi(\lambda)\ \psi'(\lambda)\ d\lambda \tag{10}$$

$R(t,\tau)$ is defined for $t,\tau \geq t_o$ by

$$R(t,\tau) = \varphi'(t)\ P(\tau)\ \varphi(\tau)\ 1(t-\tau) + \varphi'(t)\ P(t)\ \varphi(\tau)\ 1(\tau-t) \tag{11}$$

Notice that $P(\cdot)$ will be nonsingular for all $t > t_o$ if $\psi(\cdot)$ consists of linearly independent functions over an arbitrary interval, or if $P(t_o)$ is nonsingular. Then (11) could be written

$$R(t,\tau) = \psi'(t)\ P^{-1}(t)\ \psi(\tau)\ 1(t-\tau) + \psi'(t)\ P^{-1}(\tau)\ \psi(\tau)\ 1(\tau-t) \quad (12)$$

Notice also that $P \geq 0$, $\dot{P} \geq 0$, (in the sense that P and $\dot{P}$ are non-negative definite).

For the case where

$$w(t,\tau) = \mu(t)\ \delta(t-\tau) + \varphi'(t)\ \psi(\tau)\ 1(t-\tau) \quad (4)$$

we consider the system

$$\dot{x} = \psi(t)\ u(t) \quad (13)$$

$$y = \varphi'(t)\ x(t) + \mu(t)\ u(t) \quad (14)$$

With $E[x(t_o)\ x'(t_o)] = P(t_o)$ as before, the covariance of the output due to white noise input can be found as in [8] to be (for $t,\tau \geq t_o$)

$$R(t,\tau) = \mu^2(t)\ \delta(t-\tau) + \varphi'(t)\ [P(\tau)\ \varphi(\tau) + \mu(\tau)\ \psi(\tau)]\ 1(t-\tau)$$

$$+ [\varphi'(t)\ P(t) + \psi'(t)\ \mu(t)]\ \varphi(\tau)\ 1(\tau-t) \quad (15)$$

where again

$$P(t) = P(t_o) + \int_{t_o}^{t} \psi(\lambda)\ \psi'(\lambda)\ d\lambda \quad (10)$$

Here again, P will be nonsingular (for $t \geq t_o$) provided reasonable conditions are satisfied by $\psi(\bullet)$, or if $P(t_o)$ is nonsingular.

IV. PROPERTIES OF FINITE-DIMENSIONAL SYSTEM COVARIANCES

In general, we may be given a covariance function of two variables in the form

$$R(t,\tau) = a'(t)\ b(\tau)\ 1(t-\tau) + b'(t)\ a(\tau)\ 1(\tau-t) \qquad (5)$$

and asked to find an associated system; from the preceding section, evidently this is equivalent to finding a matrix $P(\cdot)$ such that

$$P(t) = P(t_o) + \int_{t_o}^{t} \psi(\lambda)\ \psi'(\lambda)\ d\lambda \qquad t \geq t_o \qquad (10)$$

for some $\psi(\cdot)$ and some symmetric nonnegative definite $P(t_o)$, and such that

$$P(t)\ a(t) = b(t) \qquad (16)$$

Equations (10) and (16) imply that R has the form of Eq. (9) from which a spectral factorization immediately follows by Eqs. (7) and (8).

For the case

$$R(t,\tau) = \mu^2(t)\ \delta(t-\tau) + a'(t)\ b(\tau)\ 1(t-\tau) + b'(t)\ a(\tau)\ 1(\tau-t) \qquad (6)$$

spectral factorization is equivalent to finding a matrix $P(\cdot)$ such that

$$P(t) = P(t_o) + \int_{t_o}^{t} \psi(\lambda)\ \psi'(\lambda)\ d\lambda \qquad (10)$$

for some $\psi(\cdot)$ and some symmetric nonnegative definite $P(t_o)$, and such that

$$P(t)\ a(t) + \mu(t)\ \psi(t) = b(t) \qquad (17)$$

Since $P(t)$ represents, in both cases, the key to the spectral factorization, we now examine ways of determining it. It turns out that $P(\cdot)$ can be represented as the solution of a Riccati equation, a different equation naturally applying in each case. The presence of a delta function in R actually permits us to determine the relevant equation a little more easily than for the case when the delta function is absent, and accordingly we cover this first.

We shall make the following key assumption:

$$\mu(t) \neq 0 \qquad \text{for all } t \tag{18}$$

This is equivalent to forbidding a structural change in the system. Assuming we have found P to satisfy (10) and (17) it follows that

$$\dot{P} = \psi\,\psi' \qquad \text{from (10)}$$

$$= \left[\frac{Pa - b}{\mu}\right]\left[\frac{Pa - b}{\mu}\right]' \qquad \text{from (17)} \tag{19}$$

together with a specification of P at time t_o as a certain symmetric matrix.

We have thus demonstrated the following result:

<u>Theorem 2.</u> Let $R(t,\tau)$, given by (6) and (15), be the covariance associated with the system (13), (14). Further, suppose (18) holds, and P is defined by (10). Then P satisfies the Riccati differential equation

$$\dot{P} = \left[\frac{Pa - b}{\mu}\right]\left[\frac{Pa - b}{\mu}\right]' \tag{19}$$

Turning attention to the other case, we observe from (16) that

$$\dot{P}a + P\dot{a} = \dot{b} \tag{20}$$

whence from (10)

$$\psi \, \psi' a + P\dot{a} = \dot{b}$$

Premultiplying by a' leads to

$$(\psi' a)^2 = a'\dot{b} - b'\dot{a} \tag{21}$$

At this stage, it might well be asked why the quantity on the right hand side of (20) should be nonnegative. This follows from the non-negativity of R; for suppose we examine the sequence

$$\iint R(t,\tau) \; i_n(t) \; i_n(\tau) \; dt \; d\tau$$

where $i_n(\cdot)$ approaches $\dot{\delta}(t-\tau)$, the derivative of the delta function. It is not hard to calculate that the integrals converge to precisely the right hand sice of (21), which must then be nonnegative. We shall assume though that the right hand side of (21) is strictly positive,

$$a'\dot{b} - b'\dot{a} \neq 0 \qquad \text{for all} \quad t \tag{22}$$

and note that techniques suggested in [6] can be used to accomodate the contrary situation.

Equation (20) then becomes

$$\psi \sqrt{(a'\dot{b} - b'\dot{a})} + P\dot{a} = \dot{b}$$

or

$$\psi = - \frac{P\dot{a} - \dot{b}}{\sqrt{a'\dot{b} - b'\dot{a}}} \tag{23}$$

Hence

$$\psi\psi' = \dot{P} = \frac{(P\dot{a} - \dot{b})(P\dot{a} - \dot{b})'}{a'\dot{b} - b'\dot{a}} \tag{24}$$

with P satisfying an initial condition requiring $P(t_o)$ to be symmetric nonnegative definite and

$$P(t_o)\ a(t_o) = b(t_o) \tag{25}$$

We have thus demonstrated the following

<u>Theorem 3.</u> Let $R(t,\tau)$, given by (5) and (9), be the output covariance associated with the system (7) and (8). Further, suppose (22) holds. Then P satisfies the Riccati differential equation

$$\dot{P} = \frac{(P\dot{a} - \dot{b})(P\dot{a} - \dot{b})'}{a'\dot{b} - b'\dot{a}} \tag{24}$$

and its symmetric nonnegative definite initial value $P(t_o)$ satisfies

$$P(t_o)\ a(t_o) = b(t_o) \tag{25}$$

Theorems 2 and 3 state necessary conditions of $R(t,\tau)$, in terms of properties of the matrix P. In the next section we show that these conditions are also sufficient to guarantee a local spectral factorization.

V. SPECTRAL FACTORIZATION

The aim here is to pass from

$$R(t,\tau) = a'(t)\ b(\tau)\ 1(t-\tau) + b(t)\ a(\tau)\ 1(\tau-t) \tag{5}$$

or

$$R(t,\tau) = \mu^2(t)\ \delta(t-\tau) + a'(t)\ b(\tau)\ 1(t-\tau) + b'(t)\ a(\tau)\ 1(\tau-t) \tag{6}$$

to a P matrix which will define a spectral factorization. Any such P matrix satisfies the relevant Riccati differential equation by Theorem 2 or 3. This suggests that <u>we use the Riccati equations to define P</u>, as explained in the following theorem:

<u>Theorem 4.</u> Let $R(t,\tau)$ be a covariance of the form of Eq. (6) with $\mu > 0$ for all t. Let $P_*(t_o)$ be an arbitrary symmetric nonnegative definite matrix and let $P_*(t)$ be the solution of

$$\dot{P}_* = \frac{(P_* a - b)(P_* a - b)'}{\mu^2} \tag{19}$$

with value at t_o equal to $P_*(t_o)$. Then the linear system

$$\dot{x} = -\frac{P_* a - b}{\mu} \tag{26}$$

$$y = a'x + \mu u \tag{27}$$

with initial condition $P_*(t_o) = E[x(t_o)\ x'(t_o)]$ and input noise with $E[u(t)\ u'(\tau)] = \delta(t-\tau)$ has covariance $R(t,\tau)$ for values of t and τ within the interval for which the system exists.

Proof of Theorem 4.

First, note that a solution of (19) exists in some finite interval extending to the right of t_o. There is, of course, no guarantee that the solution interval extends to $t = \infty$. Furthermore, this solution is symmetric since $P(t_o)$ is symmetric and $\dot{P}$ is symmetric. The proof itself is simple: direct computation using (26) and (27) yields [see (13), (14), and (15)]

$$R(t,\tau) = \mu^2(t)\ \delta(t-\tau) + a'(t)\left[P(\tau)a(\tau) - \mu(\tau)\ \frac{P_*(\tau)a(\tau) - b(\tau)}{\mu(\tau)}\right] 1(t-\tau)$$

$$+ \left[a'(t)P(t) - \mu(t)\ \frac{a'(t)P_*(t) - b'(t)}{\mu(t)}\right] a(\tau)\ 1(\tau-t) \qquad (28)$$

where

$$P(t) = P_*(t_o) + \int_{t_o}^{t} \frac{[P_*(\lambda)a(\lambda) - b(\lambda)][P_*(\lambda)a(\lambda) - b(\lambda)]'}{\mu^2}\, d\lambda$$

$$= P_*(t_o) + \int_{t_o}^{t} \dot{P}_*(\lambda)\ d\lambda \qquad \text{by (19)}$$

$$= P_*(t)$$

Replacing $P_*(\cdot)$ in (28) by $P(\cdot)$ yields immediately

$$R(t,\tau) = \mu^2(t)\ \delta(t-\tau) + a'(t)\ b(\tau)\ 1(t-\tau) + b'(t)\ a(\tau)\ 1(\tau-t) \qquad (6)$$

as required. This completes the proof.

When the $\delta(t-\tau)$ term is absent in $R(t,\tau)$, the corresponding theorem statement becomes more involved:

Theorem 5. Let $R(t,\tau)$ be a covariance of the form of Eq. (5) with $a'\dot{b} - b'\dot{a} > 0$ for all t. Let $P_*(t_o)$ be a symmetric nonnegative definite matrix which satisfies

$$P_*(t_o)\ a(t_o) = b(t_o) \tag{25}$$

and let P_* be the solution of

$$\dot{P}_* = \frac{(P_*\dot{a} - \dot{b})(P_*\dot{a} - \dot{b})'}{a'\dot{b} - b'\dot{a}} \tag{24}$$

Then the linear system

$$\dot{x} = \frac{P_*\dot{a} - \dot{b}}{\sqrt{a'\dot{b} - b'\dot{a}}} \tag{29}$$

$$y = a'x$$

with initial condition $E[x(t_o)\ x'(t_o)] = P_*(t_o)$ and input noise with $E[u(t)\ u'(\tau)] = \delta(t-\tau)$ has covariance R.

Proof of Theorem 5.

As before, P_* is symmetric because $\dot{P}_*$ and $P_*(t_o)$ are symmetric. Using Eqs. (7), (8), (10), and (11), it follows that

$$R(t,\tau) = a'(t)\ P(\tau)\ a(\tau)\ 1(t-\tau) + a'(t)\ P(t)\ a(\tau)\ 1(\tau-t)$$

where

$$P(t) = P_*(t_o) + \int_{t_o}^{t} \frac{(P_*\dot{a} - \dot{b})(P_*\dot{a} - \dot{b})'}{a'\dot{b} - b'\dot{a}}\, d\lambda$$

$$= P_*(t) \qquad \text{from (24)}$$

Thus,

$$R(t,\tau) = a'(t)\ P_*(\tau)\ a(\tau) + a'(t)\ P_*(t)\ a(\tau)\ 1(\tau-t) \tag{31}$$

It remains to be shown that $P_*(t)\ a(t) = b(t)$ for all t for which $P_*(t)$ is defined. To show this, we note

$$\begin{aligned} \frac{d}{dt}\,[P_*(t)\ a(t) - b(t)] &= \dot{P}_* a + P_* \dot{a} - \dot{b} \\ &= \frac{(P_*\dot{a} - \dot{b})(P_*\dot{a} - \dot{b})' a}{a'\dot{b} - b'\dot{a}} + P_*\dot{a} - \dot{b} \\ &= \left[\frac{P_*\dot{a}\dot{a}' - \dot{b}\dot{a}'}{a'\dot{b} - b'\dot{a}}\right](P_* a - b) \end{aligned} \tag{32}$$

by elementary manipulations. Setting $q(t) = P_* a - b$, and $F(t) = (P_*\dot{a}\dot{a}' - \dot{b}\dot{a}')(a'\dot{b} - b'\dot{a})^{-1}$, (32) becomes

$$\dot{q} = Fq \tag{33}$$

with initial condition from (25) as $q(t_o) = 0$. Immediately $q(t) = 0$ for all t, and thus (31) can be rewritten as (5). This completes the proof.

The question of whether the nonnegativity of $R(t,\tau)$ guarantees the existence of global solutions to the Riccati equations in Theorems 4 and 5 cannot yet be answered. As shown in Theorems 2 and 3, nothing but a global solution will do. In Theorem 4, it is quite evident that the nonnegativity of $R(t,\tau)$ is not fully exploited because Theorem 4 places no special requirements on the vector functions $a(t)$ and $b(t)$ at all; consequently, this nonnegativity would presumably have to be used in an at present unobvious way to demonstrate existence of a global solution to the Riccati equation.

REFERENCES

1. D. C. Youla, "On the Factorization of Rational Matrices," IRE Transactions on Information Theory, Vol. IT-7, No. 3, July 1961, pp. 172-189.

2. M. C. Davis, "Factoring the Spectral Matrix," IEEE Transactions on Automatic Control, Vol. AC-8, No. 4, October 1963, pp. 296-305.

3. B. D. O. Anderson, "Spectral Factorization by Algebra," Report TR 6560-6, Stanford Electronics Laboratories, Stanford, California, September 1966.

4. S. Darlington, "Nonstationary Smoothing and Prediction Using Network Theory Concepts," Transactions of the 1959 International Symposium on Circuit and Information Theory, Los Angeles, California, pp. 1-13.

5. A. Batkov, "Generalization of the Shaping Filter Method to Include Nonstationary Random Processes," Automation and Remote Control, Vol. 20, No. 8, August 1959, pp. 1049-1062.

6. E. B. Stear, "Shaping Filters for Stochastic Processes," Modern Control Systems Theory, McGraw-Hill, New York, 1965.

7. W. B. Davenport and W. L. Root, "An Introduction to the Theory of Random Signals and Noise," McGraw-Hill, New York, 1958.

8. J. L. Doob, "Stochastic Processes," John Wiley, New York, 1953.

9. R. E. Kalman, "Mathematical Description of Linear Dynamical Systems," Journal of the Society for Industrial and Applied Mathematics Control, Series A, Vol. 1, No. 2, 1963, pp. 152-192.

10. R. E. Kalman, "Linear Stochastic Filtering Theory-Reappraisal and Outlook," Proceedings of the Brooklyn Polytechnic Symposium on System Theory, New York, 1965, pp. 197-205.

11. R. W. Newcomb and B. D. O. Anderson, "The Description of Coloured Noise in Dynamical Systems," International Journal of Control, 1966, to appear.