1968

Recursive filtering with a distortionless constraint

Marvin Lee Bakker
Iowa State University

BAKKER, Marvin Lee, 1942-
RECURSIVE FILTERING WITH A DISTORTIONLESS CONSTRAINT.

Iowa State University, Ph.D., 1968
Engineering, electrical

University Microfilms, Inc., Ann Arbor, Michigan

RECURSIVE FILTERING WITH A DISTORTIONLESS CONSTRAINT

by

Marvin Lee Bakker

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of

The Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major Subject: Electrical Engineering

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State University
Of Science and Technology
Ames, Iowa

1968

TABLE OF CONTENTS

I. INTRODUCTION

This thesis is concerned with extending the concept of "distortionless" filtering to certain discrete-data situations, in particular, to the situations to which the usual Kalman filter theory is applicable. A brief derivation of Kalman's recursive equations is given in Section III, and then in Section IV a distortionless constraint is applied. However, before Sections III and IV are presented, a very brief review of the areas of classical and Wiener filter theory and an extensive review of distortionless filtering are given in this section and Section II, respectively. This background material is mainly related to continuous-data situations.

To be specific, consider the simple filtering situation shown in Figure 1. The filter input is a signal $x(t)$ plus noise $n(t)$, and the object of the filter is to operate on this input in such a way that the output $\hat{x}(t)$ is a good approximation to the true signal value $x(t)$. Although nonlinear filters are sometimes used, only linear filters are considered here; and so it is appropriate to represent the filter by a transfer function $Y(s)$.

Now in classical filter theory, it is assumed that the desired frequency response of the filter is known; i.e., $Y(s)$ is given. The purpose of the theory is to help the filter designer choose an appropriate electrical network configuration to yield the given frequency response. For

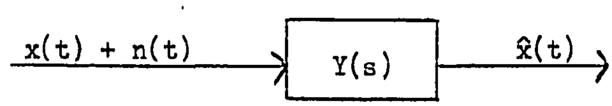

Figure 1. A simple continuous-data filtering situation

example, if $x(t)$ is a low-frequency signal and $n(t)$ is mainly high-frequency noise, then the desired frequency response is obviously that of a low-pass filter. The classical theory gives various ways to implement the filter, depending on such things as whether $x(t)$ is a voltage or a current, whether capacitors are more readily available than inductors, and so forth.

It can be seen that if the frequency spectra of the signal and the noise are known and do not overlap, then the filter can be chosen, at least theoretically, so that its output is exactly $x(t)$. In other words, for this case it is relatively easy to choose $Y(s)$ so that the filter blocks all the noise without in any way distorting the signal. However, in the case where the signal and noise are both "noise-like" in character and there is an overlap in their frequency spectra, it is no longer obvious what filter frequency response would be best. In fact, it can be seen that in this case even an ideal filter will corrupt or distort the signal somewhat in process of attenuating the noise; and so in general a compromise must be made which permits some signal distortion in order to attenuate more of the noise. Choosing the frequency response that gives the best compromise then becomes the first step of the filter-design problem, and it is this part of the problem to which Wiener filter theory is applicable. After the desired transfer function is obtained, however, the actual implementation is still frequently done by using classical filter theory.

In Wiener filter theory, an appropriate problem statement for a situation like that shown in Figure 1 is as follows: given the spectral density functions of $x(t)$ and $n(t)$, determine the "optimum" transfer function Y(s). The optimization criterion is minimization of the mean-square error, where the error e(t) is defined as the difference between the output and the true signal; i.e.,

$$e(t) = \hat{x}(t) - x(t) \tag{1.1}$$

Discussion of the restrictions placed on Y(s) and on the situations to which Wiener's work may be applied can be found, for example, in Bendat (2) or Brown and Nilsson (5).

The point of the previous discussion that is most important to this thesis is that using Wiener's theory in a situation like that of Figure 1 usually yields a transfer function that distorts the signal. Hence, even if the input noise happens to be zero for all time, the output is not exactly equal to the signal. (It should be noted that the word "happens" is used here, and in similar statements later, to imply that the filter is designed in anticipation of some nonzero noise.) However, another situation that sometimes occurs is that the signal is available from two independent, noisy sources. In this case, even when nothing is known about the signal, it is possible to attenuate the noise (the extent of attenuation depending on the particular situation) and yet satisfy a distortionless constraint such that the filter output is exactly equal to the signal if both input noises happen to be zero. A rather thorough review of distortionless filtering is given in the next section, and so no more will be said about it at this time.

It will be seen in Section III that Kalman's theory involves a set of recursive matrix equations, implementation of which usually requires a digital computer. Thus a Kalman filter is quite different from the electrical network that results from the use of classical filter theory. The reason that the term filter is frequently given to Kalman's technique is that it may be considered to be a discrete analogue of a multi-dimensional Wiener filter. The term "estimator" is often an appropriate equivalent for the term filter and is sometimes used in connection with Kalman's work. Although the term estimator will not be used in this thesis, filter outputs will generally be called estimates and errors such as that defined by 1.1 will be called estimation errors.

The somewhat sketchy background given by this introductory section is supplemented by the next two sections, after which the major part of the thesis will be presented.

II. A REVIEW OF DISTORTIONLESS FILTERING TECHNIQUES

Distortionless filters have been used since the early 1950's in continuous-data situations; and more recently, some work has been done to extend the concept to discrete-data situations. But before any specific comments are made about the previous work in this area, it should be pointed out that the name distortionless filter is not often used. A term that has been used more frequently is "complementary" filter, and quite frequently no special name has been given to the filter at all. The significance of the names distortionless and complementary will be seen later. As for the situations where no special name has been used, one reason probably is that the filter was not designed specifically to satisfy a distortionless constraint, but instead just happened to be distortionless.

An example of the use of a complementary (as they call it) filter is provided by References (1), (9), and (11). All three papers were written in connection with an instrument landing system for aircraft which includes a rather specific filter. The filter combines three related signals such that under certain conditions the output would be a perfect estimate if the input noise were zero.

In References (2), (5), and (6) distortionless filters appear in Sections 4.6 and 4.7, Chapter 15, and Sections 3.4 and 4.2, respectively. Although neither Bendat (2) nor Darlington (6) call the resulting filter by any special name, Brown and Nilsson (5) mention both the names distortionless and complementary. In all three of these references, the situation that led to a distortionless filter involved estimating a signal (using Wiener filter theory) when two independent, noisy sources are available. The sources might, for example, be measurements of the same variable with two different types of measuring devices or measurements of related variables, such as position and velocity. An interesting point that can be seen from these references is that a distortionless filter may have one of two basic physical configurations, both of which yield the same estimate.

In order to clarify this last statement, as well as some of the previous ones, an example from Benning (3) will be presented here. Consider a situation where there are two inputs available, say $x(t) + n_1(t)$ and $x(t) + n_2(t)$, where it is assumed that $n_1(t)$ and $n_2(t)$ are independent of each other and are realizations of different random processes with known spectral density functions. The filter will be called distortionless if it satisfies either of the following equivalent constraints: 1) the filter has a unity transfer function with respect to the signal, or 2) if the input noise happens to be zero, then the filter yields a perfect estimate of the signal.

Figures 2 and 3 show the physical configurations that will be considered. In Figure 2 the output may be written as follows (after the Laplace transform is taken of each time-domain function):

$$\hat{X}(s) = X(s)\left[Y_1(s) + Y_2(s)\right] + N_1(s)Y_1(s) + N_2(s)Y_2(s) \qquad (2.1)$$

Applying the distortionless constraint here requires that $\hat{X}(s) = X(s)$ when $N_1(s)$ and $N_2(s)$ are zero. It can be seen from 2.1 that this constraint is satisfied for any value of the signal $X(s)$ if the term in brackets is equal to one, in which case $Y_2(s)$ can be written in terms of $Y_1(s)$ as follows:

$$Y_2(s) = 1 - Y_1(s) \qquad (2.2)$$

Thus when the distortionless constraint is applied, Equation 2.2 can be used to write the output in terms of only the transfer function $Y_1(s)$ as

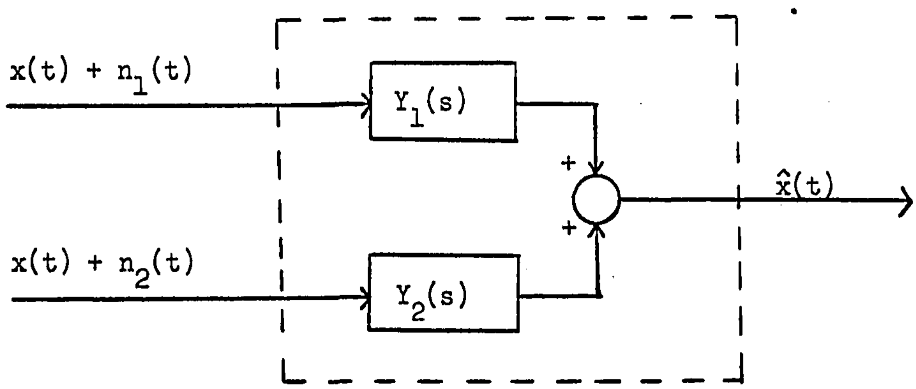

Figure 2. A simple linear system for estimating x(t) directly

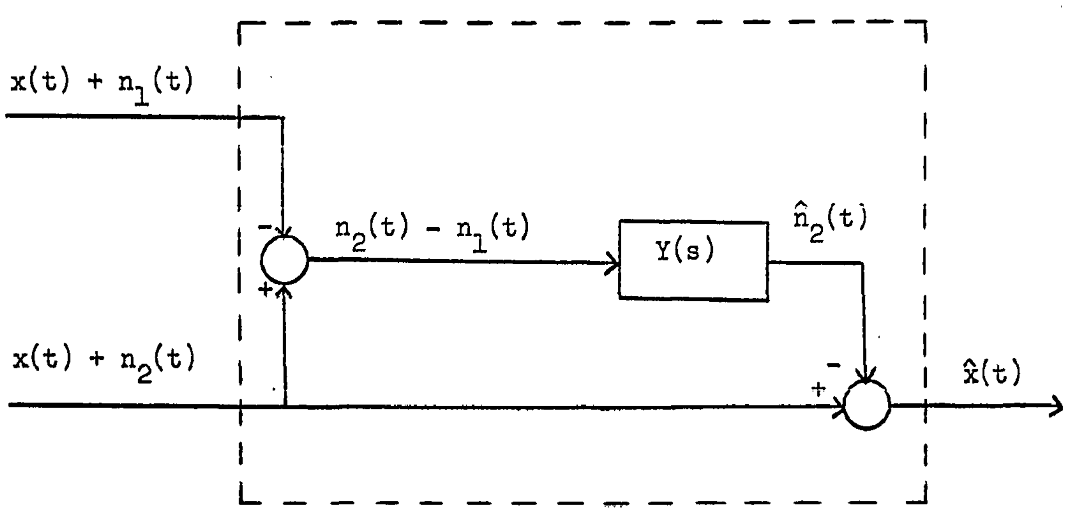

Figure 3. A system equivalent to that shown in Figure 2

follows:

$$\hat{X}(s) = X(s) + N_1(s)Y_1(s) + N_2(s)\left[1 - Y_1(s)\right] \qquad (2.3)$$

It is at this point that the significance of the terms distortionless and complementary can best be seen. From 2.2 it might be said that the second transfer function, $Y_2(s)$, is the "complement" of $Y_1(s)$. If, for example, this type of distortionless filter is extended to the case of three inputs, each with additive noise, the result is that the sum of two of the transfer functions is complementary to the third transfer function (see Brown and Nilsson (5, p. 373)). On the other hand, it can be seen from 2.3 that for any $Y_1(s)$ this filter passes the signal X(s) without "distorting" it. Thus it is seen that there is some basis for the use of each name; nevertheless, the name distortionless will be used almost exclusively throughout the rest of this thesis.

Next, consider the filter configuration shown in Figure 3. The following equation may be written for the output:

$$\hat{X}(s) = X(s) + N_2(s) - \hat{N}_2(s) \qquad (2.4)$$

But it can also be seen from Figure 3 that $\hat{N}_2(s)$, the Laplace transform of the estimate of $n_2(t)$, may be written as

$$\hat{N}_2(s) = \left[N_2(s) - N_1(s)\right]Y(s) \qquad (2.5)$$

This equation can then be used to rewrite 2.4 to give

$$\hat{X}(s) = X(s) + N_1(s)Y(s) + N_2(s)\left[1 - Y(s)\right] \qquad (2.6)$$

It can be seen that 2.6 is identical to 2.3 if $Y(s) = Y_1(s)$. Thus the dashed boxes in Figures 2 and 3 are identical distortionless filters if, in Figure 2, $Y_1(s)$ is replaced by Y(s) and $Y_2(s)$ is replaced by 1 - Y(s).

A third equivalent filter can be obtained by switching the inputs in Figure 3, in which case the intermediate estimate is of $n_1(t)$ and Y(s) corresponds to $Y_2(s)$ in Figure 2 rather than $Y_1(s)$.

Finally, one answer will be given to the important question of what $Y(s)$ to choose. If the optimization criterion is minimization of mean-square error, then $Y(s)$ is the Wiener filter which would estimate $n_2(t)$ from an input of $n_2(t) - n_1(t)$. This seems quite reasonable for the configuration of Figure 3, and in Section 15-2 of Brown and Nilsson (5) the same conclusion is reached for the configuration of Figure 2.

A significant feature of the example being considered here is that no assumptions are made about the characteristics of the signal. The point is that nothing needs to be assumed or known about the signal, since the choice of transfer functions depends entirely on the characteristics of the noise when the filter is forced to satisfy a distortionless constraint. The same idea may be stated in the other direction as follows: if very little is known about the signal characteristics (or if the filter is to be capable of handling a variety of signals), then in situations where two or more independent, noisy sources of the signal are available a good filter to use is one that satisfies a distortionless constraint. On the other hand, a distortionless filter is not the best choice in situations where, for example, the signal is "noise-like", with a known spectral density function. In that case it would be better to use a two-dimensional Wiener filter (see Section 15-3 of Brown and Nilsson (5)); i.e., a configuration like that shown in Figure 2 but with no constraining relationship between $Y_1(s)$ and $Y_2(s)$.

It might be mentioned here that using a distortionless filter is somewhat analogous to using a minimax decision rule in a decision theory problem (see Sections 10-1 and 11-1 of Harman (7)). In decision theory, using a Bayes decision rule is the "best" procedure when "enough" a priori information is available. However, if the a priori information turns out to be inaccurate, then using a Bayes decision rule can lead to very poor results. On the other hand, the minimax decision rule, although not as good as Bayes when accurate a priori information is available, is better than Bayes when the a priori information is poor. In other words, the minimax decision rule is used to avoid the possibility of extremely poor results, and in this sense it may be considered a "safe" procedure to follow. Similarly, a distortionless filter, with its ability to follow even the most abrupt changes in the signal, may be considered a "safe" filter to use.

A brief review of Benning's thesis (3), which essentially suggested the topic being considered in this thesis, will be given to conclude this section. In terms of the example presented earlier, Benning's work might be described as an extension of both the distortionless filter configurations, shown in Figures 2 and 3, to a multi-dimensional case. In particular, the situation considered is one where there are n inputs available, each of which is a linear combination of m signals (where m is less than n) and each of which also contains an additive noise term. The work involves considerable use of Wiener filter theory; in fact, the extension of the second configuration (i.e., Figure 3) yields a filter with two major "blocks", of which the first is a "linear algebraic operator" and the second is a "generalized (n-m)-dimensional Wiener filter". Benning then reasoned that the filter just described could be used in discrete-data situations by simply replacing the Wiener filter by an ordinary Kalman filter, a change that in a real-life situation might involve replacing numerous electrical networks by a digital computer. Unfortunately, extending the first configuration (i.e., Figure 2) to the multi-dimensional discrete-data case, which is essentially the purpose of this thesis, is not so straightforward. In fact, the approach taken in the derivations in Section IV in no way relies on Benning's analogous filter for the continuous-data case.

This review of previous work in the area of distortionless filtering should be sufficient background for an understanding of the constraint applied in Section IV.

III. A REVIEW OF THE KALMAN FILTER EQUATIONS

The main purpose of this section is to obtain the usual Kalman filter equations in the same notation used throughout the rest of the thesis, as well as to present a derivation that can be modified to take into account a distortionless constraint. The notation used here is similar to that of Sorenson (10), and the derivation is a combination of the one presented by Sorenson and the one presented in Brown's unpublished notes (4).

The mathematical model that is used assumes that the state of the system can be described by the following linear, vector difference equation:

$$\underline{x}(k) = \Phi(k,k-1)\underline{x}(k-1) + \underline{w}(k-1) \tag{3.1}$$

where

$\underline{x}(k)$ is the n-dimensional state vector of the system at time t_k

$\Phi(k,k-1)$ is the n by n state transition matrix

$\underline{w}(k-1)$ is an n-dimensional vector of state responses due to white-noise driving functions that occur between t_{k-1} and t_k.

The vector random sequence $\underline{w}(k)$ is assumed to have zero mean, i.e.,

$$E\left[\underline{w}(k)\right] = 0 \qquad \text{for all } k \tag{3.2}$$

Also, it is assumed to have a known covariance matrix as follows:

$$E\left[\underline{w}(k)\underline{w}^T(j)\right] = W(k)\,\delta_{kj} \tag{3.3}$$

where δ_{kj} is the Kronecker delta and $W(k)$ is a symmetric, n by n matrix which is assumed to be nonnegative-definite.

In general, the R.H.S. (right hand side) of 3.1 might include a control term and a deterministic driving function, but these terms are not important here and so will not be included.

In addition to 3.1, the mathematical model includes the following measurement (or output) equation:

$$\underline{y}(k) = H(k)\underline{x}(k) + \underline{v}(k) \tag{3.4}$$

where

$\underline{y}(k)$ is the m-dimensional measurement vector at time t_k

$H(k)$ is the m by n observation matrix

$\underline{v}(k)$ is an m-dimensional vector of measurement noise.

The random sequence $\underline{v}(k)$ also has zero mean and a known covariance matrix as follows:

$$E\left[\underline{v}(k)\right] = 0 \quad \text{for all } k \tag{3.5}$$

$$E\left[\underline{v}(k)\underline{v}^T(j)\right] = V(k)\delta_{kj} \tag{3.6}$$

where $V(k)$ is a symmetric, m by m matrix which is assumed to be non-negative-definite. It is also assumed that $\underline{v}(k)$ is uncorrelated with $\underline{w}(k)$.

The object of the Kalman filter is to estimate $\underline{x}(k)$, the state vector at time t_k, by making "optimum" use of the measurement $\underline{y}(k)$ and the a priori estimate $\underline{\hat{x}}'(k)$. The a priori estimate is really the optimum estimate of $\underline{x}(k)$, given measurement data through time t_{k-1}, and is defined as

$$\underline{\hat{x}}'(k) \overset{Df}{=} \Phi(k,k-1)\underline{\hat{x}}(k-1) \tag{3.7}$$

This is a reasonable definition since, according to 3.2 and 3.3, $\underline{w}(k-1)$ has zero mean and is not correlated with $\underline{w}(k-2)$. The "optimum" estimate $\underline{\hat{x}}(k)$ is defined to be that linear combination of $\underline{y}(k)$ and $\underline{\hat{x}}'(k)$ which minimizes the sum of the mean-square errors associated with estimating each of the state variables. The estimation equation may be written in either of the following two ways:

$$\underline{\hat{x}}(k) = \underline{\hat{x}}'(k) + K(k)\left[\underline{y}(k) - H(k)\underline{\hat{x}}'(k)\right] \tag{3.8}$$

or

$$\underline{\hat{x}}(k) = \left[I - K(k)H(k)\right]\underline{\hat{x}}'(k) + K(k)\underline{y}(k) \tag{3.9}$$

where the n by m matrix K(k) is called the gain (or weighting) matrix.

By analogy with 1.1 for a scalar situation, the estimation error is given here as

$$\underline{e}(k) = \hat{\underline{x}}(k) - \underline{x}(k) \qquad (3.10)$$

Similarly, the error in the a priori estimate is defined as

$$\underline{e}'(k) = \hat{\underline{x}}'(k) - \underline{x}(k) \qquad (3.11)$$

The covariance matrices associated with $\underline{e}(k)$ and $\underline{e}'(k)$ are, respectively,

$$P(k) = E\left[\underline{e}(k)\underline{e}^T(k)\right] \qquad (3.12)$$

$$P'(k) = E\left[\underline{e}'(k)\underline{e}'^T(k)\right] \qquad (3.13)$$

both of which are symmetric, n by n matrices. A general element of P(k), denoted as $p_{ij}(k)$, is equal to $E\left[e_i(k)e_j(k)\right]$; and a general element of P'(k), denoted as $q_{ij}(k)$, is equal to $E\left[e_i'(k)e_j'(k)\right]$.

In Section D.1 of Sorenson (10) it is shown that the Kalman filter gives an unbiased estimate of the state vector if the initial estimate $\hat{\underline{x}}(0)$ is chosen to be equal to $E\left[\underline{x}(0)\right]$; and so it will be assumed here that

$$E\left[\hat{\underline{x}}(k)\right] = E\left[\underline{x}(k)\right] \qquad \text{for all } k \qquad (3.14)$$

From this equation and 3.10 it can be seen that the estimation error has zero mean, from which it follows that the variance of the i^{th} element of $\underline{e}(k)$ is equal to the expected value of the square of $e_i(k)$; i.e.,

$$\text{Var}\left[e_i(k)\right] = E\left[e_i^2(k)\right] \qquad (3.15)$$

Thus, the trace (the sum of the elements along the major diagonal) of P(k) may be written as

$$\text{Tr } P(k) = \sum_{i=1}^{n} E\left[e_i^2(k)\right] \qquad (3.16)$$

According to the definition of "optimum" given prior to 3.8, Tr P(k) may be considered to be a loss function which is to be minimized by the appropriate choice of gain matrix K(k); and so an equation is needed that relates

P(k) and K(k). First, 3.10 may be written as follows, by using $\hat{\underline{x}}(k)$ from 3.9 with $\underline{y}(k)$ replaced by 3.4:

$$\underline{e}(k) = \left[I - K(k)H(k)\right]\underline{e}'(k) + K(k)\underline{v}(k) \tag{3.17}$$

This can then be multiplied by its transpose and the expectation taken to yield P(k). Some terms in the resulting equation are zero from previous assumptions, and so the desired equation relating P(k) and K(k) is

$$P(k) = \left[I - K(k)H(k)\right]P'(k)\left[I - K(k)H(k)\right]^T + K(k)V(k)K^T(k) \tag{3.18}$$

The method used here for determining the optimum gain matrix is to take the partial derivative of the trace of P(k) with respect to K(k) and then to set the result equal to zero as follows:

$$\frac{\partial \mathrm{Tr}\, P(k)}{\partial K(k)} = -2\left[I - K(k)H(k)\right]P'(k)H^T(k) + 2K(k)V(k) = 0 \tag{3.19}$$

Rather than explain the differentiation rules that are used here, it will simply be mentioned that the appropriate identities are listed in Brown's notes (4). It should also be noted that Sorenson (10) derives the same equation for K(k) without doing any differentiation.

Solving 3.19 for the optimum gain matrix yields

$$K(k) = P'(k)H^T(k)\left[H(k)P'(k)H^T(k) + V(k)\right]^{-1} \tag{3.20}$$

where it has obviously been assumed that the matrix in brackets is non-singular. This equation for K(k) can be used in 3.18 to give the following equation for the covariance matrix:

$$P(k) = P'(k) - K(k)\left[H(k)P'(k)H^T(k) + V(k)\right]K^T(k) \tag{3.21}$$

Note that this equation holds only for the optimum K(k), whereas 3.18 holds for any gain matrix.

Both of the last two equations involve P'(k), the error-covariance matrix for the a priori estimate; and so it is necessary to derive an equation for it in terms of known quantities. First a new equation for

$\underline{e}'(k)$ may be obtained by using 3.1 and 3.7 in 3.11 to give

$$\underline{e}'(k) = \Phi(k,k-1)\underline{e}(k-1) - \underline{w}(k-1) \quad (3.22)$$

This equation can then be used in the definition of $P'(k)$ given by 3.13. By using some of the prior definitions and assumptions, one can obtain the following result:

$$P'(k) = \Phi(k,k-1)P(k-1)\Phi^T(k,k-1) + W(k-1) \quad (3.23)$$

This completes the derivation of the Kalman filter equations; and so to summarize, a few comments will be made about the order in which the equations are used. If measurements have been taken through time t_{k-1}, the estimate $\hat{\underline{x}}(k-1)$ has been made, and $P(k-1)$ has been computed, then the next computations can be done in the following order:

1) use 3.7 to obtain the a priori estimate $\hat{\underline{x}}'(k)$ and use 3.23 to compute the associated error-covariance matrix $P'(k)$;
2) compute the optimum gain matrix $K(k)$ by using 3.20 with the $P'(k)$ determined in 1);
3) obtain the new estimate $\hat{\underline{x}}(k)$ by using either 3.8 or 3.9 with the $K(k)$ determined in 2) and the current measurement $\underline{y}(k)$;
4) compute the error-covariance matrix $P(k)$ associated with this estimate by using 3.21 with the $P'(k)$ determined in 1) and the $K(k)$ determined in 2).

The cycle then begins over with $\hat{\underline{x}}(k)$ being used in 3.7 to give the a priori estimate of the state vector at time t_{k+1} and $P(k)$ being used in 3.23 to yield $P'(k+1)$, and so forth.

The concepts of distortionless filtering and recursive (or Kalman) filtering have now been reviewed without any specific connection between the two. The next section will show how a distortionless constraint can be applied in a Kalman-type situation.

IV. DERIVATION OF EQUATIONS FOR DISTORTIONLESS RECURSIVE FILTERS

A. General Case

In terms of the ideas introduced in the first three sections, the object in this section is to apply a distortionless constraint (analogous to the second one listed early in Section II) to a Kalman filtering situation. The results here are an extension of the basic configuration shown in Figure 2, in the sense that the estimate is direct rather than being the result of subtracting out an intermediate estimate of the noise. A rather general situation is considered first and then various special cases are considered. However, before the derivations are presented, two general comments are in order.

First, in order to simplify the notation, the time at which a variable is evaluated will usually be given only when a time other than t_k is considered. Thus, whenever an equation from Section III is used here, the k's will usually be omitted.

Second, many of the equations that follow involve partitioned vectors and matrices, an introduction to which may be found in Hohn (8, p. 33). Essentially, the submatrices that result from partitioning may be treated exactly as elements of matrices as long as the dimensions are compatible for the operations being performed. One requirement that deserves special mention is that in a product of two matrices, say AB, the columns of A must be partitioned in exactly the same way as the rows of B.

The situation to be considered here might be termed a modified Kalman model. The first modification is that each of the state variables is designated as either a "signal" variable or as a "noise" variable. In particular, it is assumed that there are r "signal" variables and that these variables are the first r elements of the state vector $\underline{x}$. Hence, the state vector may be partitioned as follows:

$$\underline{x} = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_r \\ \hline x_{r+1} \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \overset{\text{Df}}{=} \begin{bmatrix} \underline{x}_S \\ \hline \underline{x}_N \end{bmatrix} \tag{4.1}$$

where $\underline{x}_S$ is the r-dimensional "signal" vector and $\underline{x}_N$ is the (n-r)-dimensional "noise" vector. The designation given a particular variable may be somewhat arbitrary, as long as the restrictions that are given later in this section are satisfied. On the other hand, in some situations the choice of "signal" and "noise" vectors is obvious. For example, if the measurement noise is not a white noise sequence but instead is correlated between sampling times, then it is frequently possible to consider the measurement noise to be the output of a "shaping filter". A shaping filter operates on a white noise input in such a way that the output has a given covariance matrix (see Sections F.1 and F.2 of Sorenson (10) for further discussion of shaping filters). In this case, the measurement noise is used to augment the original state vector, and $\underline{v}$ is zero in the measurement equation. The original state vector is then a natural choice as the "signal" vector, with the measurement noise as the "noise" vector. This noise-free measurement model is one of the special cases considered later, but for now $\underline{v}$ is assumed to be nonzero.

In order to lead up to the second basic assumption, recall from 3.1 that the state vector is given by

$$\underline{x}(k) = \Phi(k,k-1)\underline{x}(k-1) + \underline{w}(k-1) \tag{4.2}$$

which will be written here as simply

$$\underline{x} = \Phi\underline{x}(k-1) + \underline{w}(k-1) \tag{4.3}$$

Now the transition matrix Φ can be partitioned into four submatrices as follows:

$$\Phi = \left[\begin{array}{ccc|ccc} \phi_{11} & \cdots & \phi_{1r} & \phi_{1,r+1} & \cdots & \phi_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \phi_{r1} & \cdots & \phi_{rr} & \phi_{r,r+1} & \cdots & \phi_{rn} \\ \hline \phi_{r+1,1} & \cdots & \phi_{r+1,r} & \phi_{r+1,r+1} & \cdots & \phi_{r+1,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \phi_{n1} & \cdots & \phi_{nr} & \phi_{n,r+1} & \cdots & \phi_{nn} \end{array}\right]$$

$$\overset{\text{Df}}{=} \left[\begin{array}{c|c} \Phi_S & \Phi_3 \\ \hline \Phi_4 & \Phi_N \end{array}\right] \tag{4.4}$$

In addition, the first r rows of Φ will be denoted as Φ_1 and the last n-r rows as Φ_2; or in terms of the submatrices defined in 4.4,

$$\Phi_1 = \left[\begin{array}{c|c} \Phi_S & \Phi_3 \end{array}\right] \tag{4.5}$$

$$\Phi_2 = \left[\begin{array}{c|c} \Phi_4 & \Phi_N \end{array}\right] \tag{4.6}$$

Also, $\underline{w}$ may be partitioned into $\underline{w}_S$ and $\underline{w}_N$ in a manner similar to the partitioning of $\underline{x}$ in 4.1; and so 4.3 can be rewritten in the following partitioned form:

$$\left[\begin{array}{c} \underline{x}_S \\ \hline \underline{x}_N \end{array}\right] = \left[\begin{array}{c} \Phi_1 \\ \hline \Phi_2 \end{array}\right]\underline{x}(k-1) + \left[\begin{array}{c} \underline{w}_S(k-1) \\ \hline \underline{w}_N(k-1) \end{array}\right] \tag{4.7}$$

This can be written as two equations, one of which is written as follows when 4.5 and the partitioned form of $\underline{x}(k-1)$ are used:

$$\underline{x}_S = \left[\Phi_S \;\vdots\; \Phi_3\right] \begin{bmatrix} \underline{x}_S(k-1) \\ \underline{x}_N(k-1) \end{bmatrix} + \underline{w}_S(k-1)$$

$$= \Phi_S \underline{x}_S(k-1) + \Phi_3 \underline{x}_N(k-1) + \underline{w}_S(k-1) \qquad (4.8)$$

Similarly, 4.6 can be used in the other equation that can be written from 4.7 to give

$$\underline{x}_N = \Phi_4 \underline{x}_S(k-1) + \Phi_N \underline{x}_N(k-1) + \underline{w}_N(k-1) \qquad (4.9)$$

The assumption is now made that

$$\Phi_4 = 0 \quad \text{for all } k \qquad (4.10)$$

It can be seen from 4.9 that this assumption means that the value of the "noise" vector at time t_k is not to depend on the value of the "signal" vector at time t_{k-1}. Although this restriction may at times prevent a desired defining of "signal" and "noise" variables, it will be seen later that 4.10 must hold if the distortionless constraint used here is to be satisfied.

With the restrictions that have been placed on the situation, the distortionless constraint to be applied here can be defined by direct analogy with the second alternative given in Section II. The constraint may be stated as follows: if both the "noise" vector and the measurement noise happen to be zero for all k, then the filter must yield a perfect estimate of the "signal" vector; i.e., the following equality must hold:

$$\hat{\underline{x}}_S = \underline{x}_S \qquad (4.11)$$

if both $\underline{x}_N$ and $\underline{v}$ happen to be zero for all k.

It will be seen later that this constraint can be satisfied by requiring that the estimate of the state vector be independent of the a priori estimate of the "signal" vector, i.e., that $\hat{\underline{x}}$ be independent of $\hat{\underline{x}}_S'$. The procedure now is to use partitioning to obtain an equation for $\hat{\underline{x}}$ which is in appropriate form for applying the condition just stated.

First, in order to use the partitioned form of $\underline{x}$ in the measurement equation, it is necessary to partition H as follows:

$$H = \begin{bmatrix} h_{11} & \cdots & h_{1r} & h_{1,r+1} & \cdots & h_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ h_{m1} & \cdots & h_{mr} & h_{m,r+1} & \cdots & h_{mn} \end{bmatrix} \overset{\text{Df}}{=} \left[H_S \mid H_N \right] \qquad (4.12)$$

Using this equation and 4.1 in 3.4 yields the following form for the measurement equation:

$$\underline{y} = \left[H_S \mid H_N \right] \begin{bmatrix} \underline{x}_S \\ --- \\ \underline{x}_N \end{bmatrix} + \underline{v}$$

$$= H_S\underline{x}_S + H_N\underline{x}_N + \underline{v} \qquad (4.13)$$

It is also necessary to partition an n^{th} order identity matrix so that it and the partitioned form of H can be used in the estimation equation. The notation of simply I and 0 will be used for identity and zero matrices, respectively, when the dimensions are obvious or unimportant. However, it will sometimes be desirable to denote an i^{th} order identity matrix as $I^{(i)}$ and an i by j zero matrix as $0^{(i,j)}$. Thus, $I^{(n)}$ may be partitioned as follows:

$$I^{(n)} = \begin{bmatrix} I^{(r)} & 0^{(r,n-r)} \\ 0^{(n-r,r)} & I^{(n-r)} \end{bmatrix} \overset{\text{Df}}{=} \left[I_S \mid I_N \right] \qquad (4.14)$$

The estimation equation given by 3.9 can be rewritten here as

$$\hat{\underline{x}} = \left[I^{(n)} - KH \right] \hat{\underline{x}}' + K\underline{y} \qquad (4.15)$$

Using 4.12 and 4.14, one can partition the factor in brackets as follows:

$$\left[I^{(n)} - KH \right] = \left[I_S \mid I_N \right] - K\left[H_S \mid H_N \right]$$

$$= \left[I_S - KH_S \mid I_N - KH_N \right] \qquad (4.16)$$

The a priori estimate $\underline{\hat{x}}'$ can, of course, be partitioned the same way that $\underline{x}$ is in 4.1; and so using 4.16 and a partitioned $\underline{\hat{x}}'$ in 4.15 yields the following equation:

$$\underline{\hat{x}} = \left[I_S - KH_S\right]\underline{\hat{x}}'_S + \left[I_N - KH_N\right]\underline{\hat{x}}'_N + K\underline{y} \tag{4.17}$$

According to the statement below Equation 4.11, the distortionless constraint is satisfied if $\underline{\hat{x}}$ is independent of $\underline{\hat{x}}'_S$. It can be seen from 4.17 that this condition is satisfied if

$$\left[I_S - KH_S\right] = 0 \tag{4.18}$$

and so this will be called the constraint equation. Using 4.18 in 4.17 yields the following equation for the estimate of the state vector made by the distortionless filter:

$$\underline{\hat{x}} = \left[I_N - K^* H_N\right]\underline{\hat{x}}'_N + K^*\underline{y} \tag{4.19}$$

where K^* is the gain matrix that minimizes Tr P while at the same time satisfying 4.18, the constraint equation. But before obtaining an equation for K^*, it will be shown that 4.19 yields a distortionless estimate of the "signal" vector for any gain matrix K that satisfies 4.18. First, K can be partitioned between rows r and r+1 and written as follows:

$$K \equiv \begin{bmatrix} K_S \\ \hline K_N \end{bmatrix} \tag{4.20}$$

Using this and the definition of I_S given by 4.14, one may write the following two equations from the constraint equation:

$$K_S H_S = I^{(r)} \tag{4.21}$$

$$K_N H_S = 0^{(n-r,r)} \tag{4.22}$$

Using 4.20 and the definition of I_N given by 4.14, one can rewrite 4.19 as

$$\begin{bmatrix} \hat{\underline{x}}_S \\ \hat{\underline{x}}_N \end{bmatrix} = \begin{bmatrix} 0 - K_S H_N \\ I - K_N H_N \end{bmatrix} \hat{\underline{x}}'_N + \begin{bmatrix} K_S \\ K_N \end{bmatrix} \underline{y} \qquad (4.23)$$

From this equation the estimate of the "signal" vector can be written as follows, by using 4.13 for $\underline{y}$:

$$\hat{\underline{x}}_S = -K_S H_N \hat{\underline{x}}'_N + K_S \left[H_S \underline{x}_S + H_N \underline{x}_N + \underline{v} \right] \qquad (4.24)$$

This can be rearranged and the factor that multiplies $\underline{x}_S$ can be simplified by using 4.21 so that the following equation results:

$$\hat{\underline{x}}_S = \underline{x}_S + K_S H_N \left(\underline{x}_N - \hat{\underline{x}}'_N \right) + K_S \underline{v} \qquad (4.25)$$

The other equation that can be written from 4.23 is the following (again using 4.13 for $\underline{y}$):

$$\hat{\underline{x}}_N = \left[I - K_N H_N \right] \hat{\underline{x}}'_N + K_N \left[H_S \underline{x}_S + H_N \underline{x}_N + \underline{v} \right] \qquad (4.26)$$

In this equation, the factor multiplying $\underline{x}_S$ is zero (according to 4.22); and so 4.26 can be rewritten as follows:

$$\hat{\underline{x}}_N = \left[I - K_N H_N \right] \hat{\underline{x}}'_N + K_N H_N \underline{x}_N + K_N \underline{v} \qquad (4.27)$$

Now recall from 3.7 that the a priori estimate is given by

$$\hat{\underline{x}}' = \Phi \hat{\underline{x}}(k-1) \qquad (4.28)$$

Then by using the partitioned form of Φ given by 4.4, one can write the following equation for $\hat{\underline{x}}_N$ by direct analogy with 4.9:

$$\hat{\underline{x}}'_N = \Phi_4 \hat{\underline{x}}_S(k-1) + \Phi_N \hat{\underline{x}}_N(k-1) \qquad (4.29)$$

At this point the significance of the assumption that Φ_4 is zero (see 4.10) can be seen. First, 4.29 simplifies to give

$$\hat{\underline{x}}'_N = \Phi_N \hat{\underline{x}}_N(k-1) \qquad (4.30)$$

If $\underline{x}_N$ and $\underline{v}$ both happen to be zero for all k, then 4.27 becomes

$$\hat{\underline{x}}_N = \left[I - K_N H_N\right] \hat{\underline{x}}'_N \tag{4.31}$$

Using 4.31 evaluated at time t_{k-1} in 4.30 yields the following equation:

$$\hat{\underline{x}}'_N = \Phi_N \left[I - K_N(k-1)H_N(k-1)\right] \hat{\underline{x}}'_N(k-1) \tag{4.32}$$

from which it can be seen that $\hat{\underline{x}}'_N$ is zero for all k if it is zero for time t_0 (for $\underline{x}_N$ and $\underline{v}$ being zero for all k). Fortunately, it is reasonable for $\hat{\underline{x}}'_N(0)$ to be zero, since the state variables can usually be defined such that the initial state has zero mean. Thus, if $\underline{x}_N$ and $\underline{v}$ are zero for all k, then $\hat{\underline{x}}'_N$ is also zero for all k. It can be seen from 4.25 that in this situation the estimate of the "signal" vector is perfect, and it can also be seen from 4.27 that the noise is correctly estimated to be zero. Thus the requirement that $\hat{\underline{x}}$ be independent of $\hat{\underline{x}}'_S$ is sufficient to cause the filter to yield a distortionless estimate of $\underline{x}_S$ (if Φ_4 is zero), and so the gain matrix will be required to satisfy the constraint equation given by 4.18.

The next step is to determine the K^* which minimizes Tr P while at the same time satisfying the constraint equation. In similar scalar situations, the usual procedure is to use the Lagrange multiplier technique; and so an extension of that technique will be used here. First, it should be noted that the constraint equation could be written as nr scalar equations or as r column-vector equations. In particular, if superscripts are used to denote the columns of I_S and H_S, then 4.18 implies that

$$I_S^{\,i} - KH_S^{\,i} = 0 \qquad \text{for } i = 1,\ldots,r \tag{4.33}$$

Now let $\underline{\lambda}_i$ be an arbitrary, n-dimensional column vector (for i = 1, ...,r) and define θ as follows:

$$\theta = \text{Tr } P + \sum_{i=1}^{r} \underline{\lambda}_i^{T} \left[I_S^{\,i} - KH_S^{\,i}\right] \tag{4.34}$$

Note that θ is a scalar which includes the necessary nr Lagrange multipliers. The next step is to take the partial derivative of θ with respect to K as follows:

$$\frac{\partial\theta}{\partial K} = \frac{\partial \mathrm{Tr}\, P}{\partial K} - \sum_{i=1}^{r} \frac{\partial}{\partial K} \underline{\lambda}_i^T K H_S^i \tag{4.35}$$

The first term here is given by 3.19 and the general term in the summation is simply $\underline{\lambda}_i \left(H_S^i\right)^T$. Or since the transpose of the i^{th} column of H_S is equal to the i^{th} row of H_S^T, denoted as $\left(H_S^T\right)_i$, 4.35 may be written as

$$\frac{\partial\theta}{\partial K} = -2\left[I^{(n)} - KH\right]P'H^T + 2KV - \sum_{i=1}^{r} \underline{\lambda}_i \left(H_S^T\right)_i \tag{4.36}$$

Now if an n by r matrix Λ is defined such that its i^{th} column is $\underline{\lambda}_i$, then the product of ΛH_S^T may be written as follows:

$$\Lambda H_S^T = \left[\underline{\lambda}_1 \,\vdots\, \cdots \,\vdots\, \underline{\lambda}_r\right] \begin{bmatrix} \left(H_S^T\right)_1 \\ \hline \vdots \\ \hline \left(H_S^T\right)_r \end{bmatrix} = \underline{\lambda}_1 \left(H_S^T\right)_1 + \cdots + \underline{\lambda}_r \left(H_S^T\right)_r \tag{4.37}$$

It is clear that the right side of 4.37 can be written as a summation that is identical to the summation in 4.36, and so 4.36 can be written as

$$\frac{\partial\theta}{\partial K} = -2\left[I^{(n)} - KH\right]P'H^T + 2KV - \Lambda H_S^T \tag{4.38}$$

Setting this partial derivative equal to zero and rearranging the result yields the following equation:

$$K^*\left(HP'H^T + V\right) = P'H^T + \tfrac{1}{2}\Lambda H_S^T \tag{4.39}$$

If, as in the usual Kalman filter derivation, $\left[HP'H^T + V\right]$ is assumed to be of rank m and hence invertible, then both sides of 4.39 can be postmultiplied by this inverse to give

$$K^* = \left[P'H^T + \tfrac{1}{2}\Lambda H_S{}^T\right]\left(HP'H^T + V\right)^{-1} \qquad (4.40)$$

In order to eliminate the Lagrange multipliers from 4.40, it is necessary to use the constraint equation, which may be rewritten as

$$K^*H_S = I_S \qquad (4.41)$$

Thus, both sides of 4.40 can be postmultiplied by H_S and the result can be set equal to I_S according to 4.41. This latter equation can be rearranged into the following form:

$$\tfrac{1}{2}\Lambda H_S{}^T\left(HP'H^T + V\right)^{-1}H_S = I_S - P'H^T\left(HP'H^T + V\right)^{-1}H_S \qquad (4.42)$$

It will now be assumed that $\left[H_S{}^T\left(HP'H^T + V\right)^{-1}H_S\right]$ is of rank r and hence invertible. This assumption requires that H_S be of rank r and that m, the order of $\left(HP'H^T + V\right)^{-1}$, be greater than or equal to r (see Hohn (8, p. 103)). In other words, at each time t_k there must be at least as many measurements (elements of $\underline{y}$) as there are "signal" variables; and, in addition, the measurements must include r linearly independent combinations of the "signal" variables. With this assumption, both sides of 4.42 may be postmultiplied by the appropriate inverse to give an expression for $\tfrac{1}{2}\Lambda$, an expression which can then be substituted into 4.40 to give the following equation for the gain matrix of the distortionless filter:

$$K^* = \left\{P'H^T + \left[I_S - P'H^T\left(HP'H^T + V\right)^{-1}H_S\right]\left[H_S{}^T\left(HP'H^T + V\right)^{-1}H_S\right]^{-1}H_S{}^T\right\}\left(HP'H^T + V\right)^{-1} \qquad (4.43)$$

This can be written in various other ways, one of which requires partitioning P' in the same way that Φ is in 4.4 and using that form of P', the partitioned form of H^T, and the definition of I_S. The result can (after several steps that are omitted here) be written as the following two equations:

$$K_S^* = P_3'H_N^T\left(HP'H^T + V\right)^{-1}\left\{I - H_S\left[H_S^T\left(HP'H^T + V\right)^{-1}H_S\right]^{-1}H_S^T\left(HP'H^T + V\right)^{-1}\right\} + \left[H_S^T\left(HP'H^T + V\right)^{-1}H_S\right]^{-1}H_S^T\left(HP'H^T + V\right)^{-1} \tag{4.44}$$

$$K_N^* = P_N'H_N^T\left(HP'H^T + V\right)^{-1}\left\{I - H_S\left[H_S^T\left(HP'H^T + V\right)^{-1}H_S\right]^{-1}H_S^T \left(HP'H^T + V\right)^{-1}\right\} \tag{4.45}$$

Another equation for K^* can be obtained by solving the original two equations in a different way. First, setting the partial derivative given by 4.38 equal to zero yields the following equation:

$$-2\left[I^{(n)} - K^*H\right]P'H^T + 2K^*V - \Lambda H_S^T = 0 \tag{4.46}$$

If P_1' and P_2' are defined to be submatrices of P' analogous to the submatrices of Φ defined by 4.5 and 4.6, respectively, then the partitioned form of $\left[I^{(n)} - K^*H\right]$ suggested by 4.16 can be used to write the first term of 4.46 as

$$-2\left[I^{(n)} - K^*H\right]P'H^T = -2\left\{\left[I_S - K^*H_S\right]P_1' + \left[I_N - K^*H_N\right]P_2'\right\}H^T \tag{4.47}$$

But since $\left[I_S - K^*H_S\right]$ is constrained to be zero, 4.47 can be simplified and then used to rewrite 4.46 as

$$-2\left[I_N - K^*H_N\right]P_2'H^T + 2K^*V - \Lambda H_S^T = 0 \tag{4.48}$$

which can be rearranged into the following form:

$$K^*\left[H_NP_2'H^T + V\right] = I_NP_2'H^T + \tfrac{1}{2}\Lambda H_S^T \tag{4.49}$$

This equation and the constraint equation can now be solved for K^* by the same procedure used in solving 4.39 and 4.41 for K^*. The only different assumption that is necessary here is that $\left[H_N P_2' H^T + V\right]$ be invertible. Although the result obtained is not identical to 4.43, it is of the following similar form:

$$K^* = \left\{ I_N P_2' H^T + \left[I_S - I_N P_2' H^T \left(H_N P_2' H^T + V \right)^{-1} H_S \right] \left[H_S^T \left(H_N P_2' H^T + V \right)^{-1} H_S \right]^{-1} H_S^T \right\} \left(H_N P_2' H^T + V \right)^{-1} \qquad (4.50)$$

This equation can be written as two equations in a manner similar to that which led to 4.44 and 4.45. First, P_2' can be partitioned in the same way that Φ_2 is in 4.6. This form of P_2', the partitioned form of H, and the definitions of I_N and I_S can then be used in 4.50 to yield the following two equations:

$$K_S^* = \left[H_S^T \left(H_N P_2' H^T + V \right)^{-1} H_S \right]^{-1} H_S^T \left(H_N P_2' H^T + V \right)^{-1} \qquad (4.51)$$

$$K_N^* = P_N' H_N^T \left(H_N P_2' H^T + V \right)^{-1} \left\{ I - H_S \left[H_S^T \left(H_N P_2' H^T + V \right)^{-1} H_S \right]^{-1} H_S^T \left(H_N P_2' H^T + V \right)^{-1} \right\} \qquad (4.52)$$

It can be seen that this equation for K_N^* can be written in terms of K_S^* as follows:

$$K_N^* = P_N' H_N^T \left(H_N P_2' H^T + V \right)^{-1} \left[I - H_S K_S^* \right] \qquad (4.53)$$

Even though there is little resemblance between 4.51 and 4.44, the two equations should yield identical results, as should 4.53 and 4.45. It seems clear that the computation time involved will generally be shorter if 4.51 and 4.53 are used. In fact, the main reason for presenting 4.44 and 4.45 is that they are easily modified to fit the first special case considered in part B.

The error-covariance matrix P associated with the distortionless estimate can be obtained from 3.18, since it holds for any gain matrix. However, the constraint equation can be used to simplify 3.18. If P' is partitioned in the same way that Φ is in 4.4, then that form of P' and the partitioned form of $\left[I - KH\right]$ given by 4.16 can be used to write 3.18 as

$$P = \left[I_S - K^* H_S \;\vdots\; I_N - K^* H_N\right] \begin{bmatrix} P'_S & P'_3 \\ P'_4 & P'_N \end{bmatrix} \begin{bmatrix} \left(I_S - K^* H_S\right)^T \\ \left(I_N - K^* H_N\right)^T \end{bmatrix} + K^* V K^{*T} \qquad (4.54)$$

But since the constraint equation requires that $\left[I_S - K^* H_S\right]$ be zero, 4.54 becomes

$$P = \left[I_N - K^* H_N\right] P'_N \left[I_N - K^* H_N\right]^T + K^* V K^{*T} \qquad (4.55)$$

If 4.43 or 4.44 and 4.45 are used for computing the gain matrix, then the covariance matrix P' associated with the error in the a priori estimate is given by 3.23, which is rewritten here as

$$P' = \Phi P(k-1) \Phi^T + W(k-1) \qquad (4.56)$$

However, if the gain matrix is computed from 4.50 or 4.51 and 4.53, then only the last n-r rows of P' (i.e., P'_2) are needed. If W(k-1) is also partitioned such that its last n-r rows are denoted as $W_2(k-1)$, then the following equation for P'_2 can be written from 4.56:

$$P'_2 = \Phi_2 P(k-1) \Phi^T + W_2(k-1) \qquad (4.57)$$

This equation will be left as is, even though more partitioning and expanding could be done (and some simplification does occur because Φ_4 is zero).

This completes the derivation of equations for a distortionless recursive filter to correspond with the usual Kalman filter equations. It might be mentioned that the computations for the distortionless filter are done in the same order that was suggested at the end of Section III for the Kalman filter.

Before special cases are considered, an "intuitive" method for implementing the distortionless filter will be explained. It was noted in Section III that the elements along the major diagonal of P are the variances of the estimation errors. Similarly, the elements along the major diagonal of P' are the variances of the errors in the a priori estimates. It seems reasonable that if one of these elements, say q_{ii}, is very large, then the a priori estimate $\hat{x}_i'$ of the corresponding state variable should receive very little weight in determining the new estimate $\underline{\hat{x}}$. Looking at the situation in the other direction suggests that $\underline{\hat{x}}$ could be made independent of $\hat{x}_i'$ by replacing q_{ii} by some large value in the usual Kalman equation for the gain matrix. Thus if $q_{11},\ldots,q_{rr}$ are replaced by some large value in the usual gain matrix equation, then the resulting estimate of $\underline{x}$ should be independent of $\hat{x}_1',\ldots,\hat{x}_r'$ (i.e., $\underline{\hat{x}}_s'$); and so the resulting filter should yield the same results as those obtained by using 4.43 or 4.50 for K^*.

The only justification for this method is that it does seem to be intuitively satisfying and that it checks with the previously obtained equations for K^* in the particular cases that have been tried. The second example in Section V is worked by both methods to illustrate their equivalence in the situation presented there.

B. Special Cases

Special cases will now be considered, beginning with the noise-free measurement model; i.e., the situation where $\underline{v}(k)$ is known to be zero for all k. Hence, V(k) is also zero for all k. The distortionless constraint can then be stated as follows: if the "noise" vector happens to be zero for all k, then the filter must yield a perfect estimate of the "signal" vector. The procedure here is like that used in the general case, and so only the differences will be mentioned. First, the expanded form of the measurement equation given by 4.13 simplifies to

$$\underline{y} = H_S\underline{x}_S + H_N\underline{x}_N \tag{4.58}$$

The constraint equation and estimation equation given by 4.18 and 4.19, respectively, are the same here except that the $\underline{y}$ in 4.19 is given by 4.58 rather than 4.13. Thus the equation (corresponding to 4.25) for the estimate of the "signal" vector becomes

$$\hat{\underline{x}}_S = \underline{x}_S + K_S H_N(\hat{\underline{x}}_N - \hat{\underline{x}}_N') \tag{4.59}$$

The covariance matrix P is given by 3.18 without the second term, and so the modification of equations such as 4.36 and 4.38 that is necessary here is to simply drop the terms that have V as a factor. Thus 4.39 becomes

$$K^* HP'H^T = P'H^T + \tfrac{1}{2}\Lambda H_S{}^T \tag{4.60}$$

It is now assumed that $HP'H^T$ is of rank m and hence invertible. This assumption requires that H be of rank m, that P' be of rank at least as great as m, and hence that m be less than or equal to n. Then both sides of 4.60 can be multiplied by the appropriate inverse to give

$$K^* = \left[P'H^T + \tfrac{1}{2}\Lambda H_S{}^T\right]\left(HP'H^T\right)^{-1} \tag{4.61}$$

This equation and the constraint equation (which is unchanged here) can be solved for K^* by direct analogy with the derivation for the more general situation. The same requirements that arise following 4.42 must also be met here. The gain matrix for this first special case is given by the following equation, which is 4.43 with V replaced by zero:

$$K^* = \left\{ P'H^T + \left[I_S - P'H^T \left(HP'H^T \right)^{-1} H_S \right] \left[H_S{}^T \left(HP'H^T \right)^{-1} H_S \right]^{-1} H_S{}^T \right\} \left(HP'H^T \right)^{-1} \tag{4.62}$$

This can be partitioned and written as two equations which are the same as 4.44 and 4.45 with V replaced by zero. Essentially, the equations for the general case are modified by setting $\underline{v}$ and V equal to zero when the model indicates no measurement noise.

The next special case con· .dered is that of r=n; in other words, the situation where all the state variables are designated as "signal" variables. The covariance matrix V is assumed to be nonzero; and the distortionless constraint is that if $\underline{v}$ happens to be zero, then a perfect estimate of the state vector $\underline{x}$ must be obtained.

It can be seen from 4.15 that the equation that corresponds to the previous constraint equation (given by 4.18) is

$$I - K^* H = 0 \tag{4.63}$$

When this constraint is satisfied, 4.15 simplifies to give the following estimation equation:

$$\hat{\underline{x}} = K^* \underline{y} \tag{4.64}$$

where $\underline{x}$ could also be denoted as $\underline{x}_S$. Again the derivation is like that used in the general situation, and only the changes will be pointed out. In 4.39 and 4.40, the $H_S{}^T$ is replaced by H^T. Instead of postmultiplying 4.40 by H_S, it is postmultiplied by H; and the result is set equal to I according to 4.63. Hence the equation that corresponds to 4.42 is

$$\tfrac{1}{2}\Lambda H^T \left(HP'H^T + V \right)^{-1} H = I - P'H^T \left(HP'H^T + V \right)^{-1} H \tag{4.65}$$

It will now be assumed that the matrix product that multiplies $\tfrac{1}{2}\Lambda$ in 4.65 is of rank n (which equals r). This assumption requires that m be greater than or equal to n and that H be of rank n. In other words, the measurement vector must include n linearly independent cor... tions of the state variables. Postmultiplying both sides of 4.65 by the appropriate inverse yields

$$\tfrac{1}{2}\Lambda = \left[H^T\left(HP'H^T + V\right)^{-1}H\right]^{-1} - P' \tag{4.66}$$

Using this expression in the equation for K^* that corresponds to 4.40 yields an equation that can be written in the following form:

$$K^* = \left[H^T\left(HP'H^T + V\right)^{-1}H\right]^{-1}H^T\left(HP'H^T + V\right)^{-1} \tag{4.67}$$

Recall that the measurement equation is

$$\underline{y} = H\underline{x} + \underline{v} \tag{4.68}$$

In the special case where H is square and invertible, a distortionless estimate of $\underline{x}$ can be obtained by letting K^* be H^{-1}. Using this form of K^* and 4.68, one can rewrite 4.64 as

$$\underline{\hat{x}} = \underline{x} + H^{-1}\underline{v} \tag{4.69}$$

Fortunately, when H is invertible, 4.67 reduces to $K^* = H^{-1}$. Thus it might be said that, in this situation, 4.67 yields the "obvious" answer, a result that could also be obtained by using just the constraint equation.

Finally, the expression for the error-covariance matrix simplifies to

$$P = K^*VK^{*T} \tag{4.70}$$

when the constraint equation given by 4.63 holds.

It should be noted that the estimation equation (given by 4.64) for this r=n situation no longer has a recursive nature; i.e., each estimate depends only on the measurement taken at that time and does not depend on previous measurements in any way.

The last special case considered involves a very simple change from the general case, namely, that r=m. This means that the number of "signal" variables is exactly equal to the dimension of the measurement vector $\underline{y}$ and that H_S is square. Also, corresponding to the assumption following 4.42, H_S will be assumed to be invertible. No new derivation is necessary here, but considerable simplification of 4.51 and 4.53 is possible. First, the following equation can be written (see Hohn (8, p. 95)):

$$\left[H_S{}^T\left(H_N P_2' H^T + V\right)^{-1} H_S\right]^{-1} = H_S{}^{-1}\left(H_N P_2' H^T + V\right)\left(H_S{}^T\right)^{-1} \qquad (4.71)$$

Using 4.71 in 4.51 gives

$$K_S^* = H_S{}^{-1}\left(H_N P_2' H^T + V\right)\left(H_S{}^T\right)^{-1} H_S{}^T\left(H_N P_2' H^T + V\right)^{-1}$$
$$= H_S{}^{-1} \qquad (4.72)$$

Using this result in 4.53 gives simply

$$K_N^* = 0 \qquad (4.73)$$

It can be seen from 4.25 and 4.27 that the estimates of $\underline{x}_S$ and $\underline{x}_N$ in this case become

$$\hat{\underline{x}}_S = \underline{x}_S + H_S{}^{-1} H_N\left(\underline{x}_N - \hat{\underline{x}}_N'\right) + H_S{}^{-1}\underline{v} \qquad (4.74)$$

$$\hat{\underline{x}}_N = \hat{\underline{x}}_N' \qquad (4.75)$$

Since the initial "noise" vector is estimated to be zero, it can be seen from 4.30 and 4.75 that $\hat{\underline{x}}_N$ and $\hat{\underline{x}}_N'$ will be zero for all k. Thus 4.74 becomes

$$\hat{\underline{x}}_S = \underline{x}_S + H_S{}^{-1}\left(H_N \underline{x}_N + \underline{v}\right) \qquad (4.76)$$

from which it is clear that the distortionless constraint is being satisfied.

The results of this section are summarized in Section VI, after first being illustrated by the two examples presented in Section V.

V. EXAMPLES

Both examples considered here are discrete-data versions of the two-input problem introduced in Section II. The object of the filter is to make a distortionless estimate of the signal, called s(k) here to distinguish it from the state vector, using the two noisy measurements

$$y_1(k) = s(k) + n_1(k) \tag{5.1}$$

$$y_2(k) = s(k) + n_2(k) \tag{5.2}$$

For the first example, n_1 and n_2 are assumed to be uncorrelated measurement noises with variances v_{11} and v_{22}, respectively. The covariance matrix V can then be written as

$$V = \begin{bmatrix} v_{11} & 0 \\ 0 & v_{22} \end{bmatrix} \tag{5.3}$$

The state vector has only one component, s(k), which will be denoted as x_1 (where the time notation has been dropped), and so the measurement equation is

$$\underline{y} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} x_1 + \begin{bmatrix} n_1 \\ n_2 \end{bmatrix} \tag{5.4}$$

This example clearly corresponds to the second special case considered in part B of the last section, the dimensions being n=r=1 and m=2. Hence, the covariance matrix associated with the a priori estimate has only one element and will be written as

$$P' = q_{11} \tag{5.5}$$

The gain matrix can be computed by using 4.67, from which it can be seen that the inverse of the following matrix is required:

$$HP'H^T + V = \begin{bmatrix} q_{11} + v_{11} & q_{11} \\ q_{11} & q_{11} + v_{22} \end{bmatrix} \tag{5.6}$$

where H, of course, is the vector multiplying x_1 in 5.4. The inverse of the matrix may be written as follows:

$$\left(HP'H^T + V\right)^{-1} = \frac{1}{q_{11}(v_{11} + v_{22}) + v_{11}v_{22}} \begin{bmatrix} q_{11} + v_{22} & -q_{11} \\ -q_{11} & q_{11} + v_{11} \end{bmatrix} \tag{5.7}$$

Premultiplying both sides here by H^T gives

$$H^T\left(HP'H^T + V\right)^{-1} = \frac{1}{q_{11}(v_{11} + v_{22}) + v_{11}v_{22}} \begin{bmatrix} v_{22} & v_{11} \end{bmatrix} \tag{5.8}$$

Postmultiplying both sides of this by H yields a scalar whose inverse is

$$\left[H^T\left(HP'H^T + V\right)H\right]^{-1} = \frac{q_{11}(v_{11} + v_{22}) + v_{11}v_{22}}{v_{11} + v_{22}} \tag{5.9}$$

It can be seen from 4.67 that K^* is then obtained by premultiplying 5.8 by 5.9 to give

$$K^* = \frac{1}{v_{11} + v_{22}} \begin{bmatrix} v_{22} & v_{11} \end{bmatrix} \tag{5.10}$$

Using this gain matrix in the estimation equation given by 4.64 yields the following equation:

$$\hat{x}_1 = \frac{1}{v_{11} + v_{22}} \left(v_{22}y_1 + v_{11}y_2\right) \tag{5.11}$$

It might be noted here that the two measurements are weighted by complementary amounts. Also, it can be seen that the result obtained might have been reached by requiring the sum of the two weighting terms to be equal to one and then choosing the weighting such that the measurement with the greatest noise-variance is given proportionately the least weight.

Using 5.4 in 5.11 gives the following equation:

$$\hat{x}_1 = x_1 + \frac{1}{v_{11} + v_{22}} \left(v_{22}n_1 + v_{11}n_2\right) \tag{5.12}$$

This equation is quite similar to 2.3 for the analogous continuous-data situation.

The error-covariance matrix associated with the estimate given above can be computed by using 4.70. Premultiplying and postmultiplying V by, respectively, K^* and its transpose yields

$$P = \frac{v_{11}v_{22}}{v_{11} + v_{22}} \tag{5.13}$$

It is of some interest to compare the results obtained for a distortionless filter with those obtained using the usual Kalman equations. First, it can be seen from 3.20 that the K for this example is found by premultiplying 5.7 by $P'H^T$. The result can be written as

$$K = \frac{1}{v_{11} + v_{22} + v_{11}v_{22}/q_{11}} \begin{bmatrix} v_{22} & v_{11} \end{bmatrix} \tag{5.14}$$

The error-covariance matrix can then be computed by using 3.21, which yields the following equation (after several intermediate steps):

$$P = \frac{v_{11}v_{22}}{v_{11} + v_{22} + v_{11}v_{22}/q_{11}} \tag{5.15}$$

A comparison of 5.14 and 5.15 with 5.10 and 5.13 indicates that how much affect the distortionless constraint has in this example depends on the size of q_{11} relative to $v_{11}v_{22}$. It can be seen that a relatively uncertain a priori estimate (i.e., large q_{11}) results in a gain matrix that is essentially the same as K^*, as the intuitive method suggested. As would be expected, the error-covariance matrix which results by using the usual Kalman equations is better (i.e., smaller) than when a constraint is applied.

In the second example, $n_1(k)$ and $n_2(k)$ are assumed to be Markov and are modeled as the output of a shaping filter. The situation thus involves noise-free measurements, each of which is the sum of two state variables. In particular, the three-dimensional state vector is defined as

$$\underline{x}(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} \overset{Df}{=} \begin{bmatrix} s(k) \\ n_1(k) \\ n_2(k) \end{bmatrix} \tag{5.16}$$

In terms of these state variables, the measurement equation (written from 5.1 and 5.2 without the time notation) is given by

$$\underline{y} = \begin{bmatrix} x_1 + x_2 \\ x_1 + x_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \underline{x} \tag{5.17}$$

This is an example of the first special case considered in part B of Section IV, with dimensions r=1, m=2, and n=3. Thus, m is less than n; and it can also be seen that H is of rank m. Finally, P' is assumed to be of rank at least 2 at each value of k. Thus, the requirements are satisfied so that 4.62 can be used to compute K^*.

The necessary partitioning of H and $I^{(3)}$ is given by the following two equations:

$$H = \left[\begin{array}{c|cc} 1 & 1 & 0 \\ 1 & 0 & 1 \end{array}\right] = \left[H_S \;\vdots\; H_N\right] \tag{5.18}$$

$$I^{(3)} = \left[\begin{array}{c|cc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] = \left[I_S \;\vdots\; I_N\right] \tag{5.19}$$

The covariance matrix P' can be written as

$$P' = \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{bmatrix} \tag{5.20}$$

Some of the intermediate steps that lead to the final result will be noted. The first matrix product of interest is

$$HP'H^T = \begin{bmatrix} q_{11} + 2q_{12} + q_{22} & q_{11} + q_{12} + q_{13} + q_{23} \\ q_{11} + q_{12} + q_{13} + q_{23} & q_{11} + 2q_{13} + q_{33} \end{bmatrix} \tag{5.21}$$

The inverse of this matrix can be postmultiplied by H_S to give

$$\left(HP'H^T\right)^{-1} H_S = \frac{1}{\left|HP'H^T\right|} \begin{bmatrix} q_{13} + q_{33} - q_{12} - q_{23} \\ q_{12} + q_{22} - q_{13} - q_{23} \end{bmatrix} \tag{5.22}$$

where the indicated determinant is given by the following equation:

$$\begin{aligned} \left|HP'H^T\right| = {} & q_{11}q_{22} + q_{11}q_{33} + q_{22}q_{33} + 2q_{12}q_{13} + 2q_{12}q_{33} + 2q_{13}q_{22} \\ & - q_{12}^2 - q_{13}^2 - q_{23}^2 - 2q_{11}q_{23} - 2q_{12}q_{23} - 2q_{13}q_{23} \end{aligned} \tag{5.23}$$

Premultiplying 5.22 by H_S^T yields a scalar with the following inverse:

$$\left[H_S^T \left(HP'H^T\right)^{-1} H_S\right]^{-1} = \frac{\left|HP'H^T\right|}{q_{22} - 2q_{23} + q_{33}} \tag{5.24}$$

The first bracketed factor in 4.62 is given by

$$\left[I_S - P'H^T\left(HP'H^T\right)^{-1} H_S\right] = \frac{q_{12}q_{33} + q_{13}q_{22} + q_{22}q_{33} - q_{12}q_{23} - q_{13}q_{23} - q_{23}^2}{\left|HP'H^T\right|} \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} \tag{5.25}$$

This can be postmultiplied by $\left[H_S^T \left(HP'H^T\right)^{-1} H_S\right]^{-1} H_S^T$ (using 5.24) to give the following expression:

$$\frac{q_{12}q_{33} + q_{13}q_{22} + q_{22}q_{33} - q_{12}q_{23} - q_{13}q_{23} - q_{23}^2}{q_{22} - 2q_{23} + q_{33}} \begin{bmatrix} 1 & 1 \\ -1 & -1 \\ -1 & -1 \end{bmatrix} \tag{5.26}$$

The factor in braces in 4.62 is then obtained by adding 5.26 to $P'H^T$. The result will be denoted as follows:

$$\left\{ \cdot \right\} \overset{Df}{=} \frac{1}{q_{22} - 2q_{23} + q_{33}} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \tag{5.27}$$

Finally, it can be seen from 4.62 that K^* is given by postmultiplying 5.27 by $\left(HP'H^T\right)^{-1}$. Again the result is given a shortened notation; i.e.,

$$K^* = \frac{1}{\left(q_{22} - 2q_{23} + q_{33}\right)\left|HP'H^T\right|} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} \tag{5.28}$$

It turns out that each of the b_{ij}'s here has $\left|HP'H^T\right|$ as a factor, and the final form of the equation for K^* becomes

$$K^* = \frac{1}{q_{22} - 2q_{23} + q_{33}} \begin{bmatrix} q_{33} - q_{23} & q_{22} - q_{23} \\ q_{22} - q_{23} & q_{23} - q_{22} \\ q_{23} - q_{33} & q_{33} = q_{23} \end{bmatrix} \tag{5.29}$$

This same example will now be worked by the intuitive method. As the previous section indicated, the result will check with 5.29. The first step is to compute the usual Kalman gain matrix. Since V is zero here, K is given by 3.20 with V omitted; and so the equation to be used is simply

$$K = P'H^T\left(HP'H^T\right)^{-1} \tag{5.30}$$

Although not previously written out, this expression could be computed as an intermediate step leading to 5.25. At any rate, using the particular matrices of this example in 5.30 yields

$$K = \frac{1}{\left|HP'H^T\right|} \begin{bmatrix} \begin{pmatrix} q_{11}q_{33} + q_{12}q_{13} + q_{12}q_{33} \\ -q_{11}q_{23} - q_{13}^2 - q_{13}q_{23} \end{pmatrix} & \begin{pmatrix} q_{11}q_{22} + q_{12}q_{13} + q_{13}q_{22} \\ -q_{11}q_{23} - q_{12}^2 - q_{12}q_{23} \end{pmatrix} \\ \begin{pmatrix} q_{11}q_{22} + q_{12}q_{13} + q_{12}q_{33} \\ + 2q_{13}q_{22} + q_{22}q_{33} - q_{11}q_{23} \\ -q_{12}^2 - 2q_{12}q_{23} - q_{13}q_{23} \\ -q_{23}^2 \end{pmatrix} & \begin{pmatrix} q_{11}q_{23} + q_{12}^2 + q_{12}q_{23} \\ -q_{11}q_{22} - q_{12}q_{13} - q_{13}q_{22} \end{pmatrix} \\ \begin{pmatrix} q_{11}q_{23} + q_{13}^2 + q_{13}q_{23} \\ -q_{11}q_{33} - q_{12}q_{13} - q_{12}q_{33} \end{pmatrix} & \begin{pmatrix} q_{11}q_{33} + q_{12}q_{13} + 2q_{12}q_{23} \\ +q_{13}q_{22} + q_{22}q_{33} - q_{11}q_{23} \\ -q_{12}q_{23} - q_{13}^2 - 2q_{13}q_{23} \\ -q_{23}^2 \end{pmatrix} \end{bmatrix} \tag{5.31}$$

An interesting, important point here is that the q_{11}^2 terms cancel each other in the computation of K. This is important because K^* is to be obtained by letting q_{11} be very large in 5.31, and so the highest power of q_{11} occurring in the elements of the matrix of 5.31 should be the same as the highest power of q_{11} occurring in $\left|HP'H^T\right|$ (given by 5.23). As q_{11} becomes large in 5.31, the terms containing no q_{11} become negligible; and then q_{11} cancels out of the rest of the terms so that K^* is given by

$$K^{*} = \frac{1}{q_{22} - 2q_{23} + q_{33}} \begin{bmatrix} q_{33} - q_{23} & q_{22} - q_{23} \\ q_{22} - q_{23} & q_{23} - q_{22} \\ q_{23} - q_{33} & q_{33} - q_{23} \end{bmatrix} \tag{5.32}$$

which does check with 5.29. Presumably, computer implementation of the intuitive method could be accomplished in this example by replacing q_{11} by, say, ten (or maybe one hundred) times the largest element in the original P' matrix.

VI. SUMMARY

The purpose of this thesis has been to extend the basic distortionless filter configuration shown in Figure 2 to discrete-data situations, in particular, to situations to which the usual Kalman filter theory is applicable. As might have been expected, however, it has been necessary to place some restrictions on the Kalman-type situations to which a distortionless constraint may be applied. The restrictions, which, for emphasis, will be repeated here in the summary, mainly involve the transition matrix Φ, the relative dimensions of the measurement vector and the partitioned parts of the state vector, and the rank of the measurement matrix or partitioned parts of it.

First, whenever there are both a "signal" vector and a "noise" vector, the Φ_4 part of the transition matrix must be zero (see 4.10). Because of this assumption, it is not permissible to designate a particular state variable as a "noise" variable if its value at time t_k depends on the value of a "signal" variable at a time prior to t_k.

Another significant assumption was made following 4.42. That assumption requires H_S to be of rank r and also requires the following inequality to hold:

$$r \leq m \tag{6.1}$$

In other words, there must be at least as many elements in the measurement vector as there are "signal" variables.

Although various equations are given for the gain matrix in the general case, the most useful equations are 4.51 and 4.53. These equations are sufficiently straightforward that computing K^* should not require an unreasonable amount of computer time.

In the special case where $\underline{v}$ is known to be zero, the additional requirements (see below 4.60) are that H be of rank m and that P' have rank at least m, which then requires the following inequality to hold:

$$m \leq n \tag{6.2}$$

This simply means that the dimension of $\underline{y}$ must not be greater than the dimension of $\underline{x}$, a condition which is nearly always satisfied when the model calls for $\underline{v}$ to be zero. The equation for K^* in the noise-free measurement situation is given by 4.62.

Now in the special case where all the state variables are considered to be "signal" variables (i.e., when r=n), the inequality that corresponds to 6.1 is

$$n \leq m \tag{6.3}$$

This inequality is essentially the opposite of the one given by 6.2 and means that there must be at least as many elements in $\underline{y}$ as there are state variables. Although this situation is not a common Kalman-type situation, it is a natural carry-over from the distortionless filter idea presented in Section II, as can be seen from the first example in Section V. The gain matrix is given by 4.67 for this r=n case.

Finally, in the last special case considered, the only thing that is different from the general case is that H_S is assumed to be square and invertible. This assumption leads to considerable simplification of the equations for the gain matrix, as can be seen from 4.72 and 4.73.

It should generally be possible to obtain the distortionless filter gain matrix by using the intuitive approach explained in Section IV (shortly after 4.57) and illustrated in Section V.

In conclusion, a distortionless constraint can be applied in various situations, but certain requirements must be met in each case. The situation in which a distortionless recursive filter is most apt to be useful is the noise-free measurement case, especially the case where the actual measurement noise is considered to be the output of a shaping filter. In this case, the original state vector is augmented by the measurement noise and $\underline{v}$ is zero. If the "new" state variables are designated as the "noise" variables, then Φ_4 is zero and the first major requirement is automatically satisfied.

VII. LITERATURE CITED

1. Anderson, W. G. and Fritze, E. H. Instrument approach system steering computer. Institute of Radio Engineers Proceedings 41: 219-228. 1953.

2. Bendat, Julius S. Principles and applications of random noise theory. New York, N.Y., John Wiley and Sons, Inc. 1958.

3. Benning, Roger David. Multiple-input distortionless filters for estimating signals corrupted by noise. Unpublished Ph.D. thesis. Ames, Iowa, Library, Iowa State University of Science and Technology. 1966.

4. Brown, R. G. Kalman filter notes. Unpublished mimeographed notes for Electrical Engineering Course 675. Ames, Iowa, Department of Electrical Engineering, Iowa State University of Science and Technology. 1967.

5. Brown, Robert Grover and Nilsson, James William. Introduction to linear systems analysis. New York, N.Y., John Wiley and Sons, Inc. c1962.

6. Darlington, Sidney. Linear least-squares smoothing and prediction, with applications. Bell System Technical Journal 37: 1221-1294. 1958.

7. Harman, Willis W. Principles of the statistical theory of communication. New York, N.Y., McGraw-Hill Book Company, Inc. c1963.

8. Hohn, Franz E. Elementary matrix algebra. 2nd ed. New York, N.Y., The Macmillan Company. c1964.

9. Rector, Jack D. and Hattendorf, Edwin R. All-weather landing systems for automatic and manual control. Unpublished paper presented at Institute of Navigation meeting, Cedar Rapids, Iowa, June, 1966. Cedar Rapids, Iowa, Avionics and Applied Science Division, Collins Radio Company, <u>ca</u>. 1966.

10. Sorenson, H. W. Kalman filtering techniques. Advances in Control Systems: Theory and Applications 3: 219-292. 1966.

11. Stability analysis of complementary filter beam guidance computer. Collins Radio Company, Cedar Rapids [Iowa] Division Technical Report CTR 237. 1962.

VIII. ACKNOWLEDGEMENTS

The author wishes to express his sincere appreciation for the assistance given him by his major professor, Dr. R. G. Brown, who made many helpful suggestions and comments during all phases of the work presented in this thesis. Special mention should be made of the fact that Dr. Brown suggested the thesis topic and also suggested the intuitive approach for obtaining the gain matrix K^*.

The author also wishes to thank the National Science Foundation for the financial support that it provided during three years of his graduate program.