Report No. 1102

REPORT OF THE NASA SEMINAR ON PILOT-VEHICLE SYSTEMS IDENTIFICATION

held at

The Ames Research Center, May 1963

Jerome I. Elkind

Contract NASw-668

Submitted to:

National Aeronautics and Space Administration
Washington, D. C.
Attention: Dr. T. L. K. Smull

TABLE OF CONTENTS

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

On May 13, 1963 through May 15, 1963 a seminar on pilot-vehicle system identification problems was held at the NASA Ames Research Center. The seminar focused on the comparison of methods for determining human pilot dynamic response characteristics. The paradigm of multiple linear regression analysis was used as the basis for comparison of the analysis techniques.

The meeting was arranged by Mr. Melvin A. Sadoff of the Ames Research Center. I was responsible for the technical content of the seminar and gave most of the lectures. These were based in large part on work done under Contracts NASw-185 and NASw-668 which were monitored by Mr. Sadoff with the support of Mr. Robert W. Taylor of the Electronics and Control Program Office of OART, NASA Headquarters. Mr. D. T. McRuer of Systems Technology Inc., Dr. George A. Bekey of Space Technology Laboratory, and Mr. James Adams of the NASA Langley Research Center contributed greatly to the meeting by discussing the analysis techniques that they have been using for measurement of human pilot dynamics and presenting the results they have obtained using these techniques. Appendix A contains the outline of the seminar.

The seminar was attended by personnel from three NASA Research Centers - Ames, Langley, and Flight Research -- NASA Headquarters, Air Force Systems Division, Systems Technology Inc., Space Technology Laboratory, RIAS, and ACF. A list of attendees is given in Appendix B.

During the seminar I promised to write up my notes on the material discussed and distribute them to the attendees. This took longer than expected, but the task finally was completed in December 1963. The work required to put this material in final form was supported by the Air Force under Contract No. AF33(657)-8224 and will be reported in a forthcoming Air Force Report ASD-TDR-63-618, "Further Studies of Multiple Regression Analysis of Human Pilot Dynamic Response: A Comparison of Analysis Techniques and Evaluation of Time-Varying Measurements." The second chapter of that report, "A Comparison of Techniques for Determining Human Pilot Describing Functions with Stochastic Inputs," is an elaboration of the NASA Seminar material. That chapter is reproduced as the second chapter of this seminar report and should serve as the written notes for the seminar.

The seminar meetings were informal. The discussion was lively and helped to clarify many issues concerning the identification of human pilot characteristics. The format and organization of the seminar should be useful as a prototype for other similar meetings, particularly those devoted to intensive discussions of a specific topic by people actively working on problems being discussed.

CHAPTER II

A COMPARISON OF TECHNIQUES FOR DETERMINING HUMAN PILOT DESCRIBING FUNCTIONS WITH STOCHASTIC INPUTS

I. INTRODUCTION

In recent years a variety of techniques have been developed for the identification of the dynamic characteristics of linear systems and for the determination of describing functions for non-linear and time-varying systems when the input signal to these systems is a random process. Among the techniques that have been applied to identification of human pilot dynamic response characteristics the most important have been cross-correlation analysis (refs. 4,5), cross-power density spectral analysis (refs, 6-8) differential equation coefficient methods (refs. 9-11), and orthogonalized exponential function analysis (refs. 1-3).

As we shall show in this chapter, all of these analysis techniques can be considered to be special cases of or approximations to multiple linear regression analysis (ref. 12, 13). A study of regression analysis and in particular of the statistical properties of the measurements obtained from such analysis is central to the understanding of and to the comparison of these techniques.

In the next section we review briefly the principal ideas of regression analysis. We give without proof some of the results obtained in Appendix B relating to the statistical properties of regression measurements that are important for comparison of the techniques. In Section III we describe the different analysis techniques. In Section IV the methods are

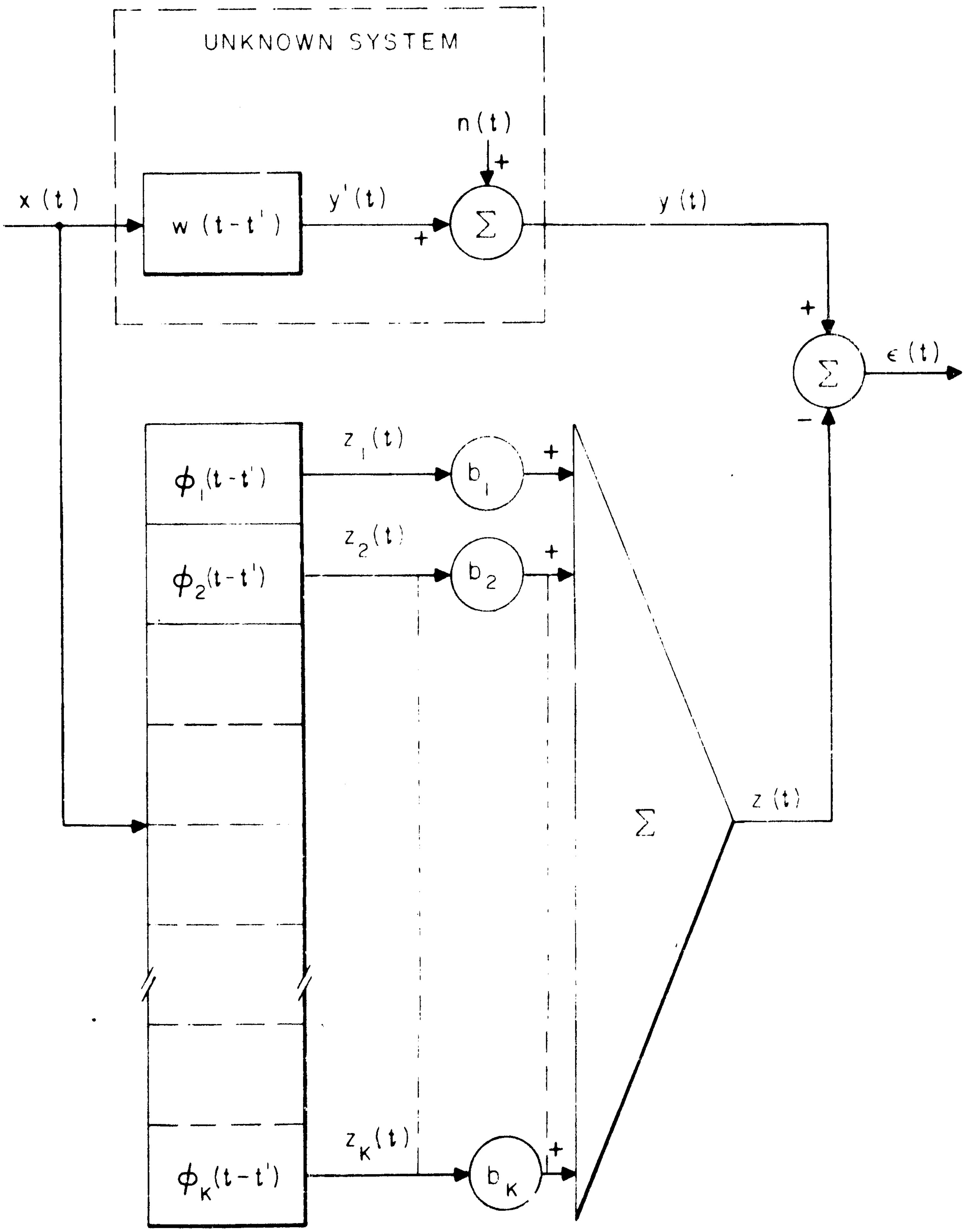

Figure 1 Multiple Regression Analysis Paradigm for Identification of Dynamic Systems

compared with respect to accuracy and precision. In Section V a method for obtaining the coefficients of the differential equation for a system from regression coefficients is described. Finally, we discuss the problem of using only signals circulating within the feedback loop to determine the open-loop describing function of the human operator.

II. MULTIPLE LINEAR REGRESSION ANALYSIS

A. Regression Equations

Multiple linear regression analysis is a procedure for finding the best linear relation between a dependent variable y and a set of independent variables z_j. The z_j can be interpreted as the coordinates or bases of a vector space, and the dependent variable y as a vector in that space. In this context y is represented as the vector sum of the basis vectors, the contribution of each component vector being just the projection of y on that basis vector.

A convenient paradigm for discussing the application of regression analysis to system identification is shown in Fig. 1. An unknown system is to be approximated by linear filter with weighting function w(t-t') to whose output is added a remnant signal n(t). This is the usual describing function method of representing dynamic systems (ref. 14). We will consider only the case in which the system input x(t) is a random signal. In this case the remnant will, in general, also be a random signal. The identification problem is to determine the weighting function w(t-t').

The input-output relation of the unknown system expressed in terms of the convolution integral is

$$y(t) = \int_{-\infty}^{t} w(t-t')\, x(t')\, dt' + n(t) \qquad (2.1)$$

where $y(t)$ is the output including the noise. The weighting function $w(t-t')$ is to be determined from measurements made on the input $x(t)$ and the response $y(t)$. The noise $n(t)$ cannot be measured directly.

A model for the system is constructed from a set of K filters connected in parallel as shown in Fig. 1. The system input $x(t)$ is fed to each of these filters and the filter outputs $z_j(t)$ are weighted by coefficients b_j and summed to form the output of the model $z(t)$.

$$z(t) = \sum_{j=1}^{K} b_j \, z_j(t) = \sum_{j=1}^{K} b_j \int_{-\infty}^{t} \phi_j(t-t') \, x(t') \, dt' \quad (2.2)$$

where $\phi_j(t-t')$ is the weighting function of $j^{\underline{th}}$ filter.

Several different criteria can be used to determine the coefficients b_j. We will consider only the case in which the b_j are determined so that the mean-square difference between the system output $y(t)$ and the model output $z(t)$ is a minimum. The mean-square difference is determined by averaging the square of the difference between $y(t)$ and $z(t)$ over a period of T seconds duration. Doing this we obtain

$$\overline{D^2} = \frac{1}{T} \int_{0}^{T} D^2(t) \, dt = \frac{1}{T} \int_{0}^{T} [y - \sum_{j=1}^{K} b_j \, z_j]^2 \, dt \qquad (2.3)$$

where $D(t)$ is the difference $y(t)-z(t)$ and the bar indicates that the average with respect to time is to be taken. The values of the coefficients that minimize the error can be found by taking the derivatives of Eq. (2.3) with respect to each of the b_j and setting the results equal to zero. Doing this, a set of K equations is obtained.

$$\begin{aligned}
\overline{z_1 z_1}\, b_1 + \overline{z_1 z_2}\, b_2 + \cdots + \overline{z_1 z_K}\, b_K &= \overline{z_1 y} \\
\overline{z_2 z_1}\, b_1 + \overline{z_2 z_2}\, b_2 + \cdots + \overline{z_2 z_K}\, b_K &= \overline{z_2 y} \\
\vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \qquad &\quad \vdots \\
\overline{z_K z_1}\, b_1 + \overline{z_K z_2}\, b_2 + \cdots + \overline{z_K z_K}\, b_K &= \overline{z_K y}
\end{aligned} \tag{2.4}$$

where $\overline{z_i z_j}$ is the sample covariance of z_i and z_j for the period T. In matrix notation the set of equations required to specify all K coefficients is

$$\underline{\underline{L}}\, \underline{b} = \underline{y} \tag{2.5}$$

where $\underline{\underline{L}}$ is the K x K covariance matrix whose elements ℓ_{ij} are $\overline{z_i z_j}$; $\underline{b}$ is the coefficient vector with K elements b_j; $\underline{y}$ is the vector with K elements $\overline{z_i y}$.

In general, the model will not account for all of the output of the system. The part not accounted for, which is called the residual $\epsilon(t)$, is the difference between the system output $y(t)$ and the model output $z(t)$ obtained when the b_j that satisfy Eq. (2.5) are used in the model of Fig. 1. By carrying out the squaring operation on the right side of Eq. (2.3) and substituting $\underline{y}$ from Eq. (2.5) in the result, we obtain for the mean-square residual

$$\overline{\epsilon^2} = \overline{y^2} - \sum_{j=1}^{K} b_j \overline{z_j y} \tag{2.6}$$

If the mod ̄ is well chosen the relation

$$w(t-t') \doteq \sum_{j=1}^{K} b_j \phi_j(t-t') \tag{2.7}$$

will give a good approximation to the system weighting function $w(t-t')$. Thus, by knowing the $\phi_j(t-t')$ and the b_j, an approximation to $w(t-t')$ can be obtained.

Eq. (2.5) provides the best linear relation in a mean-square error sense between system output $y(t)$ and the filter outputs $z_j(t)$. The coefficients b_j are coefficients of regression of $y(t)$ on the $z_j(t)$ (refs. 12, 13). They define a vector in a K-dimensional vector space whose basis vectors are the $z_j(t)$. The coefficients b_j are projections of the vector representing $y(t)$ onto the basis vectors.

B. Expected Values of Regression Coefficients

Although our objective is to obtain an accurate representation for the system weighting function, the identification procedure yields the least mean-square error approximation to the system output. We should not be surprised to find that unless special precautions are taken in the selection of the basis vectors $z_j(t)$ or of the basis functions $\phi_j(t)$ a good representation for system output may not yield a good representation for system weighting function. If a good choice of $\phi_j(t)$ or $z_j(t)$ is made, a few $z_j(t)$ will be required to represent the system output with a high degree of accuracy. If, on the other hand, they are not well chosen, a large number may be required to achieve a good representation.

To see the kind of errors that are likely to occur because the method approximates system output, assume that in the model of Fig. 1 a set of K filters chosen from a complete orthonormal set is used. In this case, the weighting function of the system $w(t-t')$ can be approximated with vanishingly small mean-square error by an infinite number of filters chosen from this set. Thus,

$$w(t-t') = \sum_{j=1}^{\infty} w_j \phi_j(t-t') \tag{2.8}$$

and

$$y(t) = \sum_{j=1}^{\infty} w_j z_j(t) + n(t) \tag{2.9}$$

where w_j is the weight applied to the $j^{\underline{th}}$ filter.

Equations (2.8) and (2.9) are series representations for system weighting function and system output when an infinite number of orthonormal filters is used to represent the system. When a finite number is used, the weight applied to each (the coefficients b_j) will not, in general, equal the w_j in Eqs. (2.8) and (2.9). The values of the coefficients that will be obtained can be found by using Eq. (2.9) to expand the sample covariances that appear on the right side of Eq. (2.5) (note that in Eq. (2.5) $\underline{y}$ is a vector whose components are $\overline{z_i y}$.)

When Eq. (2.9) is substituted for $\underline{y}$ in the covariances $\overline{z_i y}$ of Eq. (2.5), the following matrix equation is obtained

(ref. 1)

$$\underline{L}\,\underline{b} = \underline{L}\,\underline{w} + \underline{L}_{K+1}\,\underline{w}_{K+1} + \underline{n} \qquad (2.10)$$

where $\underline{w}$ is the vector whose elements are the first K coefficients w_j of Eq. (2.9); $\underline{w}_{K+1}$ is the vector whose elements are the coefficients w_j for $j>K$; $\underline{L}_{K+1}$ is the matrix with sample covariances $\overline{z_i z_j}$ for $i\leq K$ and $j>K$; $\underline{n}$ is the vector whose elements are $\overline{z_i n}$ for $i\leq K$.

The solution of Eq. (2.10) is

$$\underline{b} = \underline{w} + \underline{L}^{-1}\,\underline{L}_{K+1}\underline{w}_{K+1} + \underline{L}^{-1}\,\underline{n} \qquad (2.11)$$

or

$$\underline{b} = \underline{w} + \underline{g} + \underline{h} \qquad (2.12)$$

where $\underline{g}$ and $\underline{h}$ are defined by the second and third terms of Eq. (2.11).

Thus each regression coefficient b_j is shown to be the sum of three components. The first component is the corresponding weight w_j that would result from using an infinite sum of orthonormal functions to represent $w(t-t')$ as in Eq. (2.8). The second component g_j is a bias caused by approximating the system with a finite number of orthonormal filters. The third component h_j results from the noise $n(t)$.

By taking the expected values of both sides of Eq. (2.12) we obtain a relation for the expected values of the regression coefficients which we designate β_j. Since the noise is uncorrelated with the filter outputs, the expected value of $\underline{h}$

will be zero and

$$\underline{\beta} = \underline{w} + \underline{\gamma} \tag{2.13}$$

where $\underline{\beta}$ is the vector whose components are β_j and $\underline{\gamma}$ is obtained from $\underline{g}$ in Eqs. (2.11) and (2.12) by replacing the sample covariances in $\underline{L}^{-1}\ \underline{L}_{K+1}$ by their expected values. Thus we find that unless the measurement is performed in a way that makes $\underline{\gamma}$ zero, the estimates of the regression coefficients, the b_j, will be biased in the sense that the expected value β_j will not be equal to the true value w_j. The bias arises because we have used only a finite number of basis functions to represent the system characteristics and because we have found a least mean-square error approximation to the system output rather than to the system weighting function.

The bias can be reduced or eliminated two ways. One method is to choose the basis functions so that the first K of them match the system weighting function w(t) exactly. This will result in $\underline{w}_{K+1}$ being identically zero, thus making $\underline{\gamma}$ zero and yielding a $\underline{\beta}$ that will be equal to the true value $\underline{w}$. The second method is to select filter outputs $z_j(t)$ that are orthogonal. If this is done $\underline{\Lambda}_{K+1}$, the expected value of $\underline{L}_{K+1}$, will be zero, making $\underline{\gamma}$ zero and yielding an unbiased $\underline{\beta}$. Orthogonalization of filter outputs can be achieved either by selecting a filter set whose outputs will be orthogonal for all input signals (such as very narrow bandpassed filters with non-overlapping pass bands), or by tailoring the input signal to the filters so that orthogonal outputs are obtained. This last condition is obtained if the input is white noise and the filters have orthogonal weighting functions. If the input is not white noise, it frequently can be prefiltered before it is fed to the set of filters and thereby "whitened" so that the effective input signal to these filters will be approximately white noise.

Actually, one need not go so far as to choose a basis set that can approximate the system characteristics exactly, or to force the filter outputs to be orthogonal. It is not difficult to obtain an approximation to the system characteristics that is a very good one in the sense that a few filters will approximate the system behavior with very small error. In such a case, the elements of $\underline{w}_{K+1}$ will be small, thereby making the bias small and giving a good approximation to the system weighting function.

C. Statistical Properties of Regression Coefficients

If the input x(t), the output y(t) and the residual $\epsilon(t)$ have a normal distribution with zero mean, the regression coefficients obtained from a particular sample of the input signal will be normally distributed with an expected value $\underline{\beta}$ given by Eq. (2.13) and a variance that can be shown (ref. 1) to be

$$\sigma^2_{bj|s} = \frac{\sigma^2_\epsilon}{N\, s^2_{ju}} \tag{2.14}$$

where σ^2_ϵ is the variance of the residual, N is the number of independent samples of the residual obtained from the T-second long sample used to compute b_j. As such it is the number of degrees of freedom of the residual. s^2_{ju} is the sample variance of the part of the output of the j^{th} filter that is independent of the other filter outputs. $1/s^2_{ju}$ is the j^{th} diagonal term of $\underline{L}^{-1}$ and can be computed from $\underline{L}$. The variance $\sigma^2_{bj|s}$ depends upon that particular value of s^2_{ju} obtained in a measurement.

To find the expected variance for all values of s^2_{ju} we take the expected value of $\sigma^2_{bj|s}$. By taking advantage of the fact that the distribution of s^2_{ju} is proportion to χ^2 (see Appendix

B) we obtain

$$\sigma_{bj}^2 = \frac{\sigma_\epsilon^2}{N\sigma_{ju}^2} \frac{M}{M-(K+1)} \tag{2.15}$$

where σ_{ju}^2 is the variance of the part of the output of the j^{th} filter that is uncorrelated with the other filter outputs, M is the number of degrees of freedom in the part of $z_j(t)$ uncorrelated with the other filter outputs obtained in the T-second long sample, and K is the number of filters used in the model.

In any measurement it is desirable to have as stable a set of measures as possible. This can be achieved by making the ratio σ_{bj}/β_j as small as possible. A small "sigma-to-mean" ratio can be obtained by using a small number of filters that approximates the system well so that each β_j will tend to be large thereby tending to make σ_{bj}/β_j small. The σ_{bj} can be made small by (1) choosing the filters so that they account for almost all of y(t) that is correlated with the input thereby making σ_ϵ^2 small, (2) making N large, which can be accomplished by making the sample length T large (if this is possible)*, (3) choosing filters whose outputs are uncorrelated so that σ_{ju}^2 is large, or (4) choosing the filter outputs to have as large a bandwidth as possible, thereby making M large. Using a small number of filters will also help reduce σ_{bj}^2 by making K small. This will help keep the term M/[M-(K+1)] in Eq. (2.15) close to unity.

In addition to the relative variability of the coefficients, one must also be concerned with the possibility that the b_j will be biased. A bias will occur when $\underline{\gamma}$ in Eq. (2.13) is not zero.

---

*As shown in Appeddix B, N is approximately 2WT where W is the effective bandwidth of the residual in cycles per second.

If the b_j are biased, then the representation of the system characteristics will be in error. For spectral and orthonormal exponential function analysis, the bias can be made small by choosing filters that represent $w(t-t')$ with very small error or by making the filter outputs orthogonal over the interval T. This will not be true in the case of one of the differential equation coefficient methods in which a bias due to the remnant noise cannot be removed.

III. DESCRIPTION OF ANALYSIS TECHNIQUES

In this section four common methods for measuring system dynamics are discussed from the viewpoint of multiple regression analysis. The four methods are cross-correlation analysis, cross-spectral analysis, orthogonalized exponential function analysis, and differential equation coefficient methods.

A. Cross-Correlation Analysis

The basic relation for cross-correlation analysis of a system is

$$
\begin{aligned}
R_{xy}(\tau) &= \int_{-\infty}^{\tau} w(\tau-t')\, R_{xx}(t')\, dt' \\
&= \int_{0}^{\infty} w(t')\, R_{xx}(\tau-t')\, dt'
\end{aligned}
\qquad \tau \geq 0 \qquad (2.16)
$$

where R_{xy} and R_{xx} are, respectively, the cross-correlation function of system input and output and the autocorrelation of system input (ref. 4). Equation (2.16) is obtained from Eq. (2.1) by correlating both sides of the equation with $x(t)$. The remnant noise $n(t)$ does not appear in Eq. (2.16) because it is uncorrelated with the input $x(t)$.

The integral in Eq. (2.16) can be approximated by a summation.

$$R_{xy}(n\Delta\tau) \doteq \sum_{j=0}^{\infty} w(j\Delta\tau)\, R_{xx}[(n-j)\Delta\tau]\, \Delta\tau \qquad n \geq 0$$

$$\overline{x(t-n\Delta\tau)\, y(t)} \doteq \sum_{j=0}^{\infty} w(j\Delta\tau)\, \overline{x[t-(n-j)\Delta\tau]x[t]}\, \Delta\tau \qquad (2.17)$$

$$\doteq \sum_{j=0}^{\infty} w(j\Delta\tau)\, \overline{x(t-n\Delta\tau)x(t-j\Delta\tau)}\, \Delta\tau$$

where $\Delta\tau$ is the increment in τ used in the computation of the correlation functions. The relations between input and output implied by Eq. (2.17) may be represented by the multiple regression paradigm as shown in Fig. 2 (ref. 15). The filters used to represent the system have impulse repsonses that are pure time delays whose outputs are the same as the input except for the delay in time of $n\Delta\tau$ seconds. The impulse response of these filters is a delta function $\delta(t-n\Delta\tau)$, a unit impulse at $t=n\Delta\tau$. The coefficients b_j are equal to $w(j\Delta\tau)\Delta\tau$ as is apparent from Eq. (2.17).

In terms of the filter outputs shown in Fig. 2 the regression equations, Eq. (2.4), can be written

$$\begin{array}{ccccccccc}
\overline{z_0 z_0}\, b_0 & + & \overline{z_0 z_1}\, b_1 & + & \cdots\cdots & + & \overline{z_0 z_K}\, b_K & = & \overline{z_0 y} \\
\overline{z_1 z_0}\, b_0 & + & \overline{z_1 z_1}\, b_1 & + & \cdots\cdots & + & \overline{z_1 z_K}\, b_K & = & \overline{z_1 y} \\
\vdots & & \vdots & & & & \vdots & & \vdots \\
\overline{z_K z_0}\, b_0 & + & \overline{z_K z_1}\, b_1 & + & \cdots\cdots & + & \overline{z_K z_K}\, b_K & = & \overline{z_K y}
\end{array} \qquad (2.18)$$

or

$$\underline{L}\, \underline{b} = \underline{y} \qquad (2.19)$$

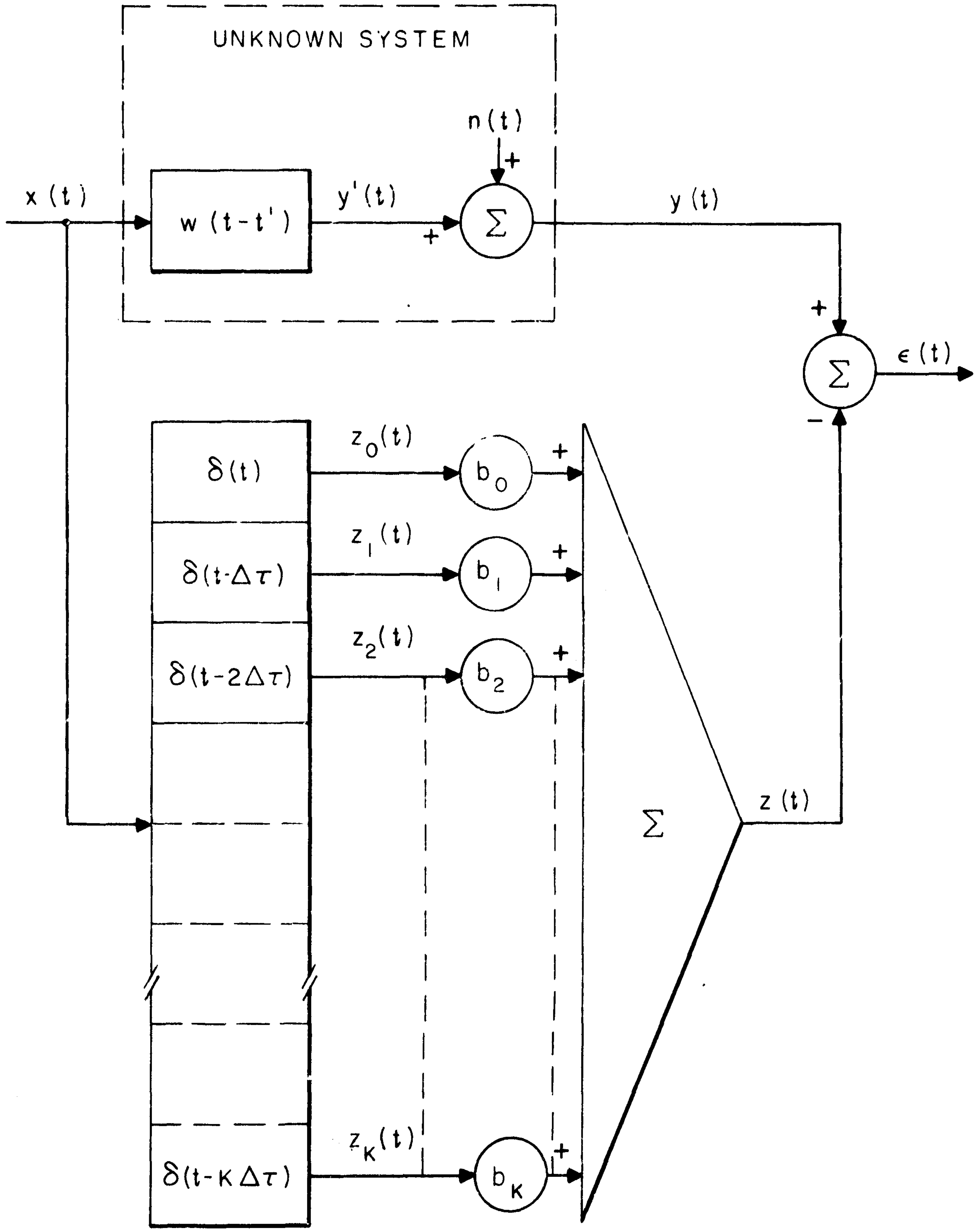

Figure 2 Cross Correlation Analysis

where $\underline{L}$, $\underline{b}$ and $\underline{y}$ are defined in a manner corresponding to that used in Eq. (2.5).

Each row of Eq. (2.18) is equivalent to Eq. (2.17) for a particular value of n. Except for the special case of a white noise input, the filter outputs $z_j(t)$ will not, in general, be mutually orthogonal. Therefore, the complete set of equations of Eq. (2.18) must be solved for the coefficient b_j. When this is done a representation for the system weighting function is obtained

$$w(t) = \sum_{j=0}^{K} b_j \; \delta(t-j\Delta\tau) \tag{2.20}$$

where $\delta(t-j\Delta\tau)$, the impulse response of the j^{th} filter is a unit impulse occurring at $t=n\Delta\tau$.

B. Cross-Spectral Analysis

The basic relation for cross-spectral analysis of system characteristics is

$$W(\omega) = \frac{S_{xy}(\omega)}{S_{xx}(\omega)} \tag{2.21}$$

where $W(\omega)$ is the transfer function of the system being measured and $S_{xy}(\omega)$ and $S_{xx}(\omega)$ are, respectively, the cross-power density spectrum of input and output and the power density spectrum of the input. Equation (2.21) is obtained by taking the Fourier transform of Eq. (2.16) (ref. 4).

1. Conventional Spectral Analysis

We can use the variation of the paradigm of multiple regression analysis in Fig. 3 to represent the relation betwen the signals at frequency ω_j implied by Eq. (2.21). $\Phi_{Rj}(\omega)$ and $\Phi_{Ij}(\omega)$ are the transfer functions of two narrow bandpass filters with center frequency ω_j. The amplitude characteristics of the filters are identical. The phase characteristics are also identical except that the phase $\Phi_{Ij}(\omega)$ is advanced 90 degrees relative to $\Phi_{Rj}(\omega)$.

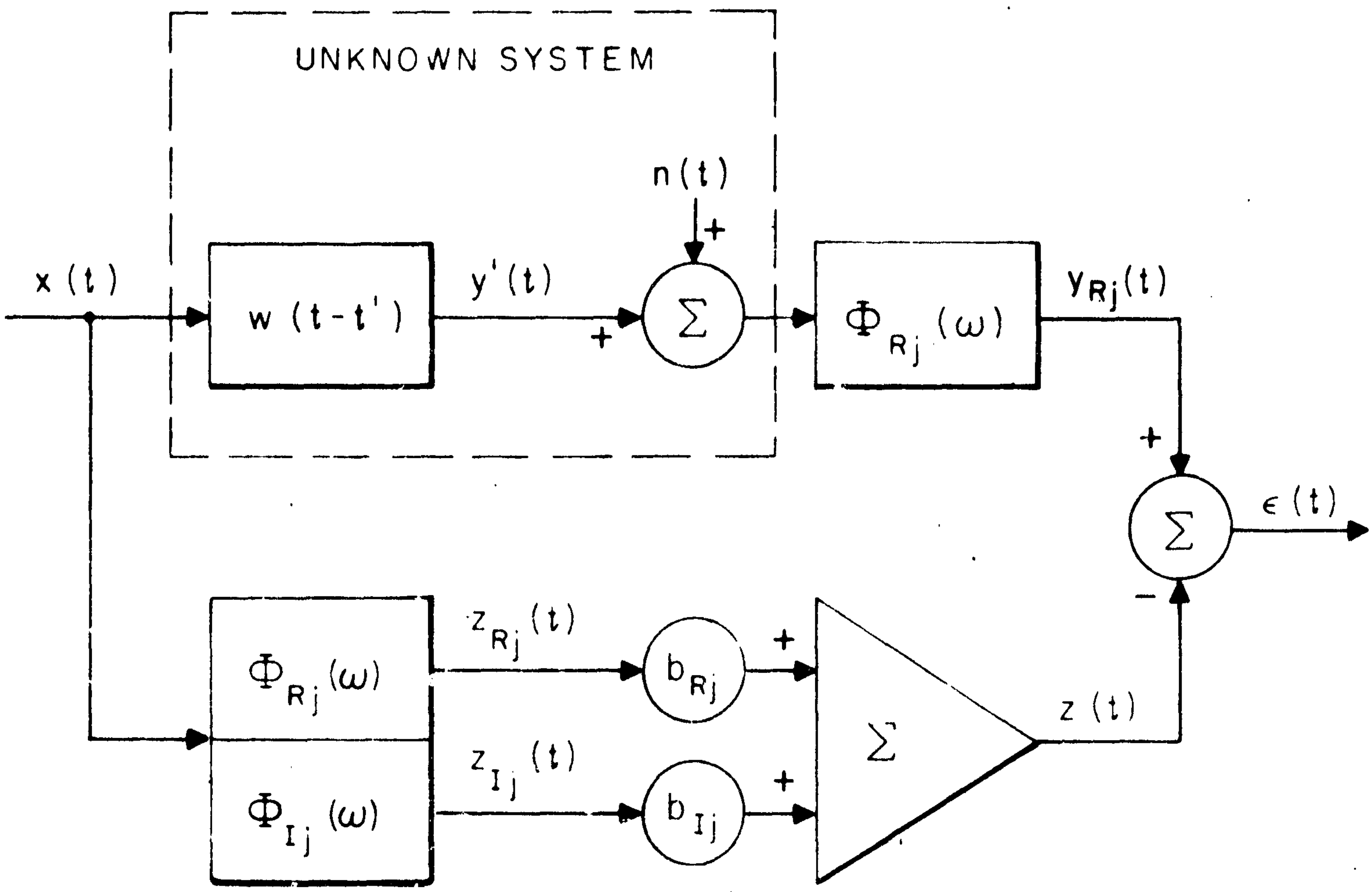

Figure 3 Conventional Spectral Analysis

Note that the system output y(t) is filtered by $\Phi_{Rj}(\omega)$. The regression coefficients b_{Rj} and b_{Ij} of Fig. 3 are determined so that this filtered output $y_{Rj}(t)$ is approximated with least mean-square error. The regression equations obtained for the configuration of Fig. 3 are

$$\begin{aligned} \overline{z_{Rj}^2}\, b_{Rj} + \overline{z_{Rj} z_{Ij}}\, b_{Ij} &= \overline{z_{Rj} y_{Rj}} \\ \overline{z_{Ij} z_{Rj}}\, b_{Rj} + \overline{z_{Ij}^2}\, b_{Ij} &= \overline{z_{Ij} y_{Rj}} \end{aligned} \tag{2.22}$$

Since the relative phase shift between the two filters is 90 degrees, the two off-diagonal terms $\overline{z_{Rj} z_{Ij}}$ and $\overline{z_{Ij} z_{Rj}}$ in Eq. (2.22) will be approximately zero if the sample length T is sufficiently long. Also, since the two filters have identical amplitude response $\overline{z_{Rj}^2}$ will be equal to $\overline{z_{Ij}^2}$. If all these conditions are true, Eq. (2.22) reduces to

$$\begin{aligned} \overline{z_j^2}\, b_{Rj} &= \overline{z_{Rj} y_{Rj}} \\ \overline{z_j^2}\, b_{Ij} &= \overline{z_{Ij} y_{Rj}} \end{aligned} \tag{2.23}$$

where $\overline{z_j^2}$ is used to represent $\overline{z_{Rj}^2}$ and $\overline{z_{Ij}^2}$.

By shifting the filters in frequency and performing the same filtering operations on input and output at a number of frequencies, several sets of equations of the form of Eq. (2.23) will be obtained, one for each frequency. The combined equations will be

$$
\begin{aligned}
\overline{z_1^2}\, b_{R1} &= \overline{z_{R1} y_{R1}} \\
\overline{z_1^2}\, b_{I1} &= \overline{z_{I1} y_{R1}} \\
\overline{z_2^2}\, b_{R2} &= \overline{z_{R2} y_{R2}} \\
&\vdots \\
\overline{z_K^2}\, b_{RK} &= \overline{z_{RK} y_{RK}} \\
\overline{z_K^2}\, b_{IK} &= \overline{z_{IK} y_{RK}}
\end{aligned}
\tag{2.24}
$$

We now show that if the $\Phi_{Rj}(\omega)$ and $\Phi_{Ij}(\omega)$ are very narrow bandpass filters with sharp cut-off, the b_{Rj} and b_{Ij} will be approximately equal to the real and imaginary parts of the system describing function $W(\omega)$. Take one equation from the set of Eq. (2.24) and write the averages in terms of the appropriate integrals. Assume that the sample length T is large and that the analysis filters have non-overlapping passbands.

$$
b_{Rj} \left| \lim_{T \to \infty} \frac{1}{T} \int_0^T z_j^2 \, dt \right| = \lim_{T \to \infty} \frac{1}{T} \int_0^T z_{Rj} y_{Rj} \, dt \tag{2.25}
$$

A similar expression can be written for b_{Ij}.

We note that

$$
\lim_{T \to \infty} \frac{1}{T} \int_0^T z_j^2 \, dt = R_{jj}(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{jj}(\omega) \, d\omega
$$

and (2.26)

$$
\begin{aligned}
\lim_{T \to \infty} \frac{1}{T} \int_0^T z_{Rj} y_{Rj} \, dt &= R_{z_{Rj} y_{Rj}}(0) \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{Re}[S_{z_{Rj} y_{Rj}}(\omega)] \, d\omega
\end{aligned}
$$

where $R_{jj}(0)$ and $R_{z_{Rj}y_{Rj}}(0)$ are the autocorrelation function of $z_{Rj}(t)$ and the cross-correlation function of $z_{Rj}(t)$ and $y_{Rj}(t)$, respectively, for $\tau=0$. $S_{jj}(\omega)$ and $Re[S_{z_{Rj}y_{Rj}}(\omega)]$ are the power density spectrum and the real part of the cross-power density spectrum corresponding to these correlation functions. Equation (2.26) may be used to rewrite Eq. (2.25) as

$$b_{Rj} \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{jj}(\omega)\, d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} Re[S_{z_{Rj}y_{Rj}}(\omega)]\, d\omega \qquad (2.27)$$

However,

$$S_{jj}(\omega) = |\Phi_{Rj}(\omega)|^2 \; S_{xx}(\omega)$$

and (2.28)

$$S_{z_{Rj}y_{Rj}}(\omega) = |\Phi_{Rj}(\omega)|^2 \; S_{xy}(\omega)$$

where $S_{xx}(\omega)$ and $S_{xy}(\omega)$ are the power density and cross-power density spectra of the signal $x(t)$ and the signals $x(t)$ and $y(t)$, respectively.

In terms of these two spectra Eq. (2.27) can be written

$$b_{Rj} \frac{1}{2\pi} \int_{-\infty}^{\infty} |\Phi_{Rj}(\omega)|^2 \; S_{xx}(\omega)\, d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\Phi_{Rj}(\omega)|^2 \; Re[S_{xy}(\omega)]\, d\omega \quad (2.29)$$

Solving for b_{Rj}

$$b_{Rj} = \frac{\int_{-\infty}^{\infty} |\Phi_{Rj}(\omega)|^2 \; Re[S_{xy}(\omega)]\, d\omega}{\int_{-\infty}^{\infty} |\Phi_{Rj}(\omega)|^2 \; S_{xx}(\omega)\, d\omega}$$

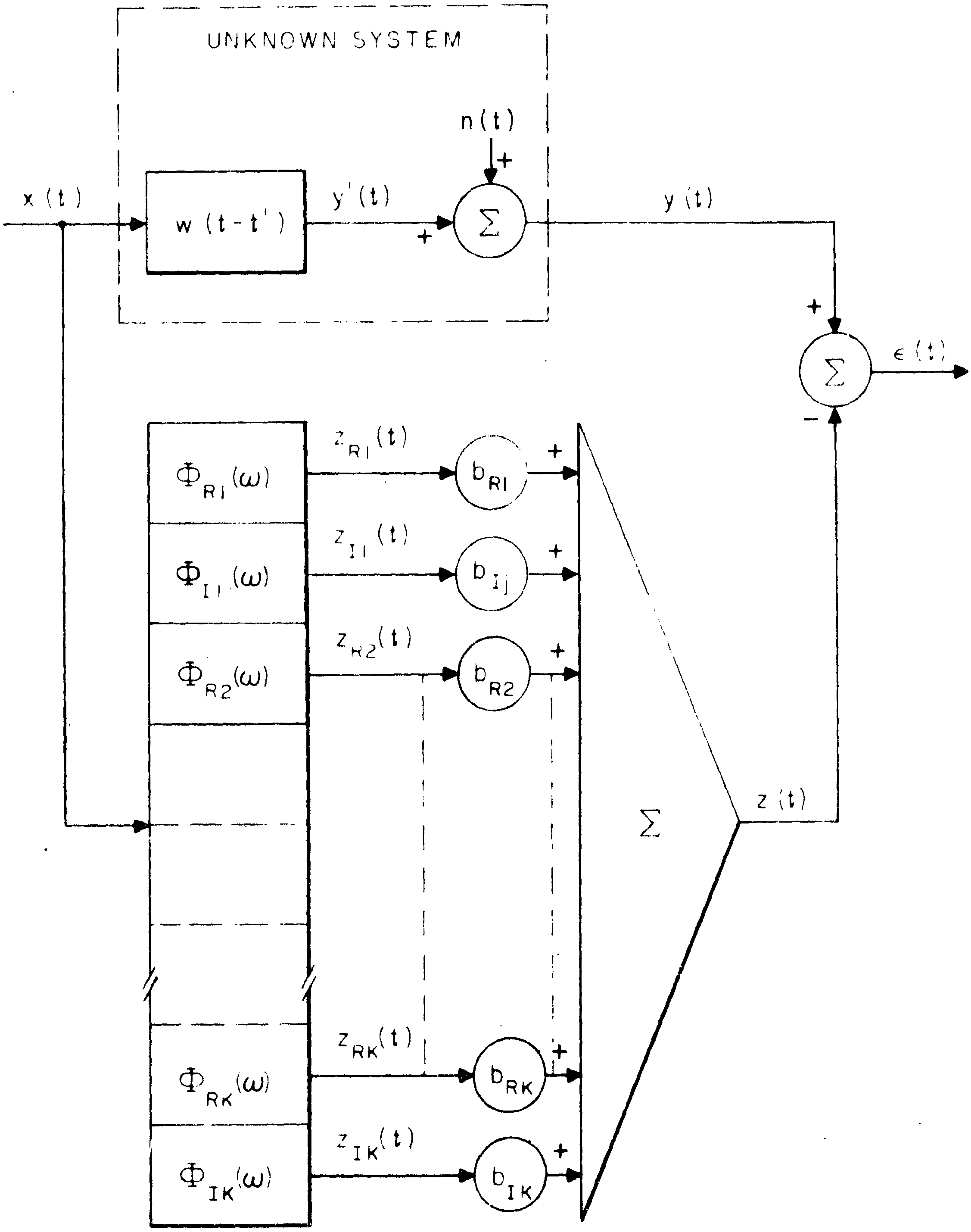

Figure 4 Direct Regression Method of Spectral Analysis

Similarly,

$$b_{Ij} = \frac{\int_{-\infty}^{\infty} |\Phi_{Rj}(\omega)|^2 \, \mathrm{Im}[S_{xy}(\omega)] \, d\omega}{\int_{-\infty}^{\infty} |\Phi_{Rj}(\omega)|^2 \, S_{xx}(\omega) \, d\omega} \tag{2.30}$$

where $\mathrm{Im}[S_{xy}(\omega)]$ is the imaginary part of $S_{xy}(\omega)$.

If $\Phi_{Rj}(\omega)$ is a very narrow bandpass filter and S_{xx} and S_{xy} are relatively constant in the passband, and equal to $S_{xx}(\omega_j)$ and $S_{xy}(\omega_j)$,

$$\begin{aligned} b_{Rj} &\doteq \frac{\mathrm{Re}[S_{xy}(\omega_j)]}{S_{xx}(\omega_j)} = \mathrm{Re}[W(\omega_j)] \\ b_{Ij} &\doteq \frac{\mathrm{Im}[S_{xy}(\omega_j)]}{S_{xx}(\omega_j)} = \mathrm{Im}[W(\omega_j)] \end{aligned} \tag{2.31}$$

Thus, for the case of long sample lengths or orthogonal outputs and very narrow bandpass filters, it is clear that the approximate equations Eq. (2.24) yield coefficients equal to the real and imaginary parts of the system transfer function. If these conditions are not satisfied, the more complete Eq. (2.22) will have to be used.

2. A Direct Regression Method of Spectral Analysis

Consider next the form of spectral analysis illustrated in Fig. 4. A set of K pairs of narrow bandpass filters,

$\Phi_{Rj}(\omega)$ and $\Phi_{Ij}(\omega)$, of the same type as used in Fig. 3, are connected in parallel, excited by the input signal $x(t)$, and used to approximate the entire system output $y(t)$. In terms of this set of filters, the regression equations of Eq. (2.4) become

$$
\begin{array}{l}
\overline{z_{R1}z_{R1}}\, b_{R1} + \overline{z_{R1}z_{I1}}\, b_{I1} + \cdots\cdots + \overline{z_{R1}z_{RK}}\, b_{RK} + \overline{z_{R1}z_{IK}}\, b_{IK} = \overline{z_{R1}y} \\
\overline{z_{I1}z_{R1}}\, b_{R1} + \overline{z_{I1}z_{I1}}\, b_{I1} + \cdots\cdots + \overline{z_{I1}z_{RK}}\, b_{RK} + \overline{z_{I1}z_{IK}}\, b_{IK} = \overline{z_{I1}y} \\
\qquad \vdots \\
\overline{z_{RK}z_{R1}}\, b_{R1} + \overline{z_{RK}z_{I1}}\, b_{I1} + \cdots\cdots + \overline{z_{RK}z_{RK}}\, b_{RK} + \overline{z_{RK}z_{IK}}\, b_{IK} = \overline{z_{RK}y} \\
\overline{z_{IK}z_{R1}}\, b_{R1} + \overline{z_{IK}z_{I1}}\, b_{I1} + \cdots\cdots + \overline{z_{IK}z_{RK}}\, b_{RK} + \overline{z_{IK}z_{IK}}\, b_{IK} = \overline{z_{IK}y}
\end{array}
\tag{2.32}
$$

or

$$\underline{L}\,\underline{b} = \underline{y} \tag{2.33}$$

where $\underline{L}$, $\underline{b}$ and $\underline{y}$ are defined in a manner similar to that used for Eq. (2.5).

The model of Fig. 4 indicates that for the direct regression method of spectral analysis the system weighting function is represented by a weighted sum of filters.

$$w(t-t') \doteq \sum_{j=1}^{K} [b_{Rj}\, \phi_{Rj}(t-t') + b_{Ij}\, \phi_{Ij}(t-t')]$$

and (2.34)

$$W(\omega) \doteq \sum_{j=1}^{K} [b_{Rj}\, \Phi_{Rj}(\omega) + b_{Ij}\, \Phi_{Ij}(\omega)]$$

The approximations of Eq. (2.34) are valid whether or not the filter outputs are orthogonal. However, if they are orthogonal Eq. (2.32) can be simplified greatly.

First assume that the filter outputs are mutually orthogonal over the sample length T. The non-diagonal terms of Eq. (2.32) will be zero and the following reduced set of equations is obtained.

$$\begin{aligned}
\overline{z_{R1}^2}\, b_{R1} &= \overline{z_{R1} y} \\
\overline{z_{I1}^2}\, b_{I1} &= \overline{z_{I1} y} \\
\overline{z_{R2}^2}\, b_{R2} &= \overline{z_{R2} y} \\
&\vdots \\
\overline{z_{RK}^2}\, b_{RK} &= \overline{z_{RK} y} \\
\overline{z_{IK}^2}\, b_{IK} &= \overline{z_{IK} y}
\end{aligned} \qquad (2.35)$$

For filters having orthogonal outputs it is not necessary to connect the entire set of 2K filters in parallel and excite them simultaneously. The measurement could be made one filter at a time.

It should be noted that the coefficients b_{Rj} and b_{Ij} are not easily interpreted as the real and imaginary components of the system describing function as was the case for the conventional spectral analysis. Rather, the series expressions of Eq. (2.34) which involve these coefficients as well as the analysis filter characteristics must be used to represent the system describing function.

3. Spectral Analysis Using Sinusoidal Components of Input

If the input signal $x(t)$ is composed of the sum of K sinusoids as is frequently the case in studies of human pilot dynamics (refs. 7, 16, and 17) it is not necessary to filter this signal by a set of bandpass filters as in Fig. 4. Each component of the signal can be taken directly and used as one of the $z_{Rj}(t)$ signals. By shifting the phase of each of these signals by 90 degrees the $z_{Ij}(t)$ signals can be obtained. Since each of the $z_{Rj}(t)$ signals is just a sinusoid, its mean square will be the power of that sinusoidal component of the input. Thus, for long sample lengths T,

$$\overline{z_{Rj}^2} = \overline{z_{Ij}^2} = S_{xx}(\omega_j) \tag{2.35}$$

Similarly, the product of $z_{Rj}(t)$ and $y(t)$ will be the real part of the cross-power spectrum of $x(t)$ and $y(t)$ at frequency ω_j. That is,

$$\overline{z_{Rj}y} = \mathrm{Re}[S_{xy}(\omega_j)]$$

and (2.37)

$$\overline{z_{Ij}y} = \mathrm{Im}[S_{xy}(\omega_j)]$$

Substituting Eqs. (2.36) and (2.37) in Eq. (2.35) we obtain

$$b_{Rj} = \frac{\mathrm{Re}[S_{xy}(\omega_j)]}{S_{xx}(\omega_j)} = \mathrm{Re}[W(\omega_j)]$$

(2.38)

$$b_{Ij} = \frac{\mathrm{Im}[S_{xy}(\omega_j)]}{S_{xx}(\omega_j)} = \mathrm{Im}[W(\omega_j)]$$

Thus the coefficients b_{Rj} and b_{Ij} determined using this method are the real and imaginary components of the system describing function.

For short sample lengths the sinusoidal components of the input may not be mutually orthogonal. In this case the complete set of equations of Eq. (2.32) must be used to find the b_{Rj} and b_{Ij}. These coefficients will still be the real and imaginary parts of the system describing function as indicated in Eq. (2.38).

C. Orthogonalized Exponential Function Analysis

The third method that we shall discuss is based directly upon the multiple regression analysis but employs a special set of orthogonalized exponential filters in the model or representation for the system in Fig. 1.

It is noted that all lumped parameter systems have weighting functions that are the sum of damped exponential functions. Therefore, it seems reasonable that an efficient way of representing such systems is to use functions that are also damped exponentials. If such functions are used, it should be possible to approximate such systems to within a specified error with a relatively small number of functions.

Kautz and Huggins (refs. 18 and 19) have suggested a set of exponential functions that are orthonormal and that are easy to construct. These are a natural set to use for representation of systems whose impulse responses are also damped exponentials. These functions have transfer functions of the form

$$
\begin{aligned}
\Phi_1(s) &= \frac{\sqrt{-2s_1}}{(s-s_1)} \\
\Phi_2(s) &= \frac{\sqrt{-2s_2}}{(s-s_2)} \; \frac{(s+s_1)}{(s-s_1)} \\
\Phi_k(s) &= \frac{\sqrt{-2s_k}}{(s-s_k)} \prod_{j=1}^{k-1} \frac{(s+s_j)}{(s-s_j)}
\end{aligned}
\tag{2.39}
$$

where the poles s_j are negative.

If one uses such a set of filters in the model of Fig. 1 the entire covariance matrix of Eq. (2.5) is computed and solved for the regression coefficients b_j. In such a case the system weighting function can be approximated by the weighted sum of the weighting functions of each of the filters as in Eq. (2.7):

$$
w(t-t') \doteq \sum_{j=1}^{K} b_j \phi_j(t-t') \tag{2.7}
$$

This method of representing system dynamics is similar to the spectral analysis technique shown in Fig. 4 except that orthogonalized exponential filters are used instead of narrow bandpass filters. The use of orthogonalized exponential filters permits representation of system characteristics with fewer filters than with narrow bandpass filters because exponential filters can be made to resemble more closely the dynamics of the system being measured (ref. 1).

D. Differential Equation Coefficient Methods

Differential equation coefficient methods lead directly to estimates of the coefficients of the differential equation for the system being measured. A differential equation for the system to be measured is assumed. In general, this will be of the form

$$a_N \frac{d^N y'}{dt^N} + \ldots + a_n \frac{d^n y'}{dt^n} + \ldots + a_1 \frac{dy'}{dt} + y' = c_M \frac{d^M x}{dt^M} + \ldots + c_m \frac{d^m x}{dt^m} + \ldots + c_1 \frac{dx}{dt} + c_o x \qquad (2.40)$$

where $y'(t)$ is the output of the linear filter before the addition of the remnant noise (see Fig. 1), N and M determine the order of the equation and are chosen in advance. The coefficients a_n and c_m are to be determined. Any one of the coefficients may be chosen in advance. For convenience we assume a_o to be unity.

We simplify notation by letting

$$A(p)y' = \sum_{n=o}^{N} a_n p^n y' = \sum_{n=o}^{N} a_n \frac{d^n y'}{dt^n}$$

and (2.41)

$$C(p)x = \sum_{m=o}^{M} c_m p^m x = \sum_{m=o}^{M} c_m \frac{d^m x}{dt^m}$$

where $p = d/dt$.

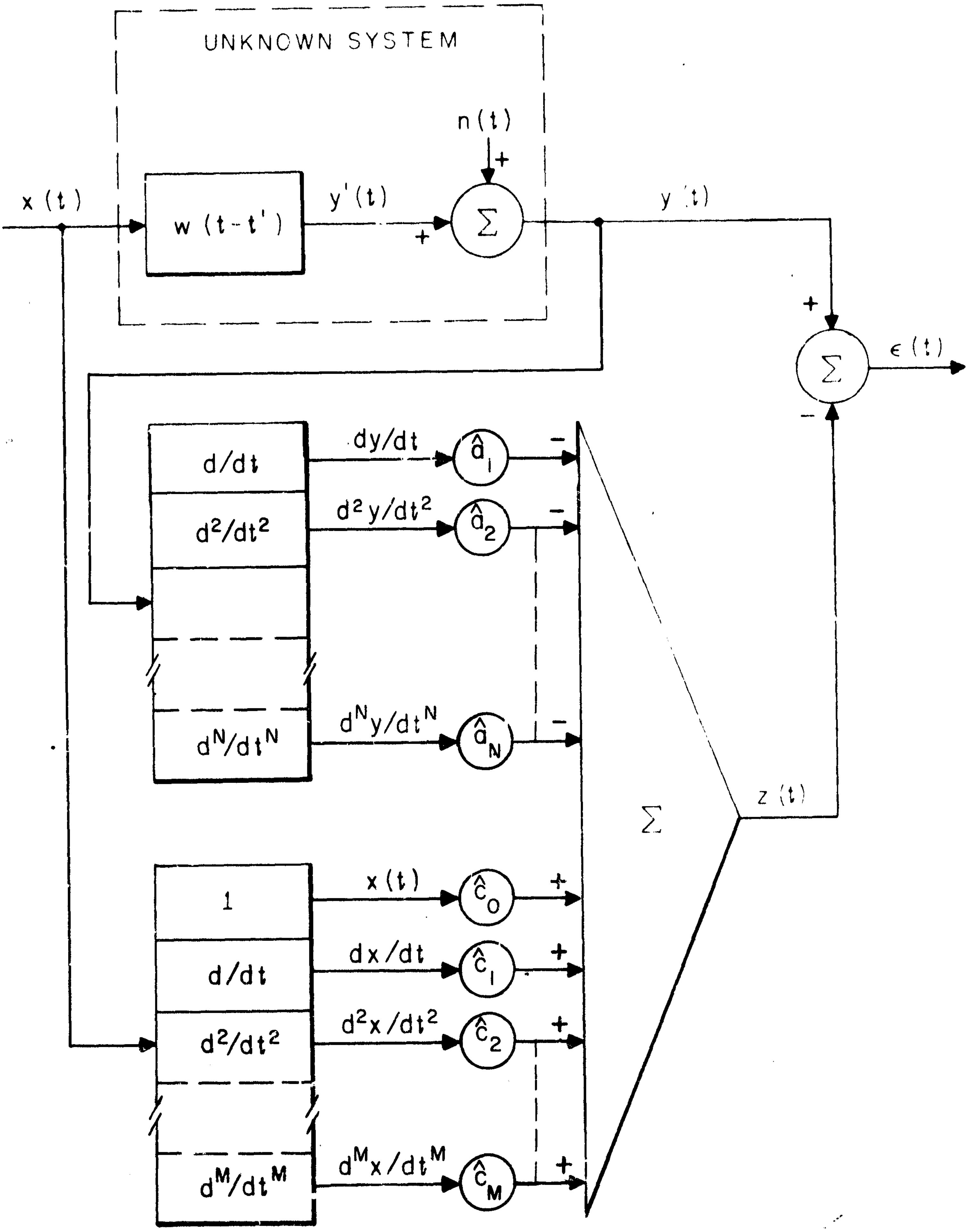

Figure 5 Equation Error Differential Equation Method

In these terms Eq. (2.40) becomes

$$A(p)y' = C(p)x \tag{2.42}$$

We would like to be able to represent the input-output relations of Eq. (2.42) by the multiple regression paradigm of Fig. 1. Unfortunately, this is not possible since the signal $y'(t)$ is not available to us.

Two methods have been used to circumvent the problem caused by the unavailability of $y'(t)$. The first, which is called the equation error method, is illustrated in Fig. 5. The output $y(t)$ is taken as an approximation to $y'(t)$ and the coefficients of the equation

$$A(p)y = C(p)x \tag{2.43}$$

are found using multiple regression techniques. The second method, which is called the output error method, is illustrated in Fig. 6. A linear filter with transfer function

$$M(p) = \frac{C(p)}{A(p)} \tag{2.44}$$

is simulated. The coefficients of the numerator and denominator are found using iterative techniques or the method of steepest descent. (refs. 11 and 22.) Linear regression methods cannot be used except in the special case when $A(p)$ is known in advance of measurement.

1. Equation Error Method

Equation (2.43) is represented in terms of the multiple regression paradigm of Fig. 1 by a set of differentiators operating on both input and output, as shown in Fig. 5. The coefficients $\hat{a}_n$ and $\hat{c}_m$ are to be chosen to minimize the mean-square difference between the weighted sum of y(t) and its derivatives, and the weighted sum of x(t) and its derivatives. This difference is called the equation error and represents the difference between the left and right sides of Eq. (2.43). Since the coefficient of y(t), a_o, was assumed to be unity, minimization of the equation error is equivalent to minimization of the mean-square difference between y(t) and the weighted sum of the outputs of all the differentiators that operate on x(t) and y(t).

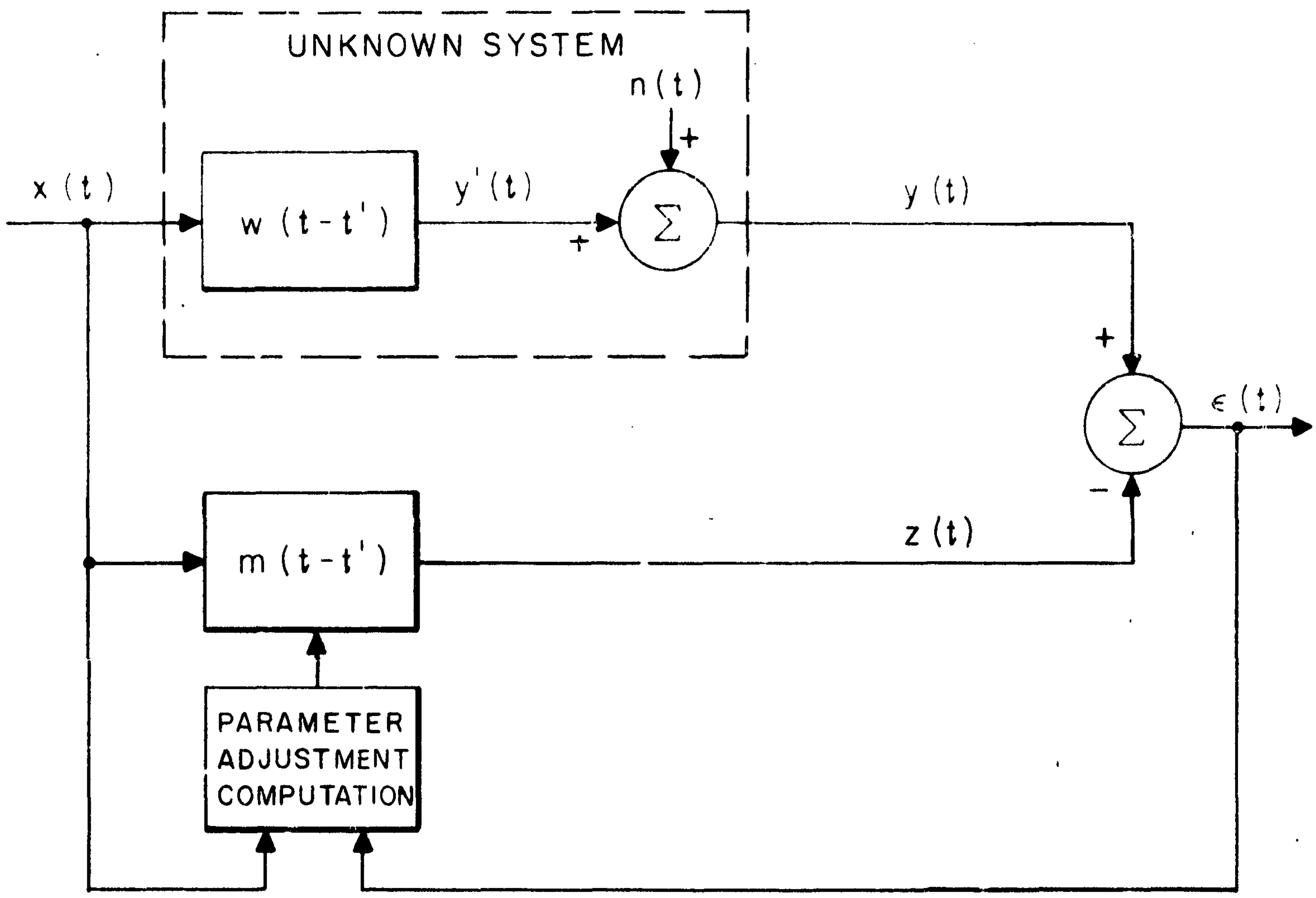

Figure 6 Output Error Differential Equation Method

The regression equations for the configuration of Fig. 5 are

$$
\begin{array}{lllll}
-\sum_{n=1}^{N} \hat{a}_n \overline{(p^n y)(p\ y)} & + & \sum_{m=o}^{M} \hat{c}_m \overline{(p^m x)(p\ y)} & = & \overline{y(p\ y)} \\
-\sum_{n=1}^{N} \hat{a}_n \overline{(p^n y)(p^2 y)} & + & \sum_{m=o}^{M} \hat{c}_m \overline{(p^m x)(p^2 y)} & = & \overline{y(p^2 y)} \\
\vdots & & \vdots & & \vdots \\
-\sum_{n=1}^{N} \hat{a}_n \overline{(p^n y)(p^N y)} & + & \sum_{m=o}^{M} \hat{c}_m \overline{(p^m x)(p^N y)} & = & \overline{y(p^N y)} \\
-\sum_{n=1}^{N} \hat{a}_n \overline{(p^n y)(x)} & + & \sum_{m=o}^{M} \hat{c}_m \overline{(p^m x)(x)} & = & \overline{y\ x} \\
-\sum_{n=1}^{N} \hat{a}_n \overline{(p^n y)(p\ x)} & + & \sum_{m=o}^{M} \hat{c}_m \overline{(p^m x)(p\ x)} & = & \overline{y(p\ x)} \\
\vdots & & \vdots & & \vdots \\
-\sum_{n=1}^{N} \hat{a}_n \overline{(p^n y)(p^M x)} & + & \sum_{m=o}^{M} \hat{c}_m \overline{(p^m x)(p^M x)} & = & \overline{y(p^M x)}
\end{array}
\tag{2.45}
$$

where $\hat{a}_n$ and $\hat{c}_m$ are estimates of a_n and c_m in Eq. (2.40).

In matrix notation this becomes

$$
\begin{array}{lllll}
-\underline{L}_y \hat{\underline{a}} & + & \underline{L}_{xy} \hat{\underline{c}} & = & \underline{Y}_y \\
-\underline{L}^t_{xy} \hat{\underline{a}} & + & \underline{L}_x \hat{\underline{c}} & = & \underline{Y}_x
\end{array}
\tag{2.46}
$$

or more generally

$$
\underline{L}\ \underline{b} = \underline{Y} \tag{2.47}
$$

where $\underline{L}_y$ is the covariance matrix of the derivatives of y and $\underline{L}_x$ the covariance matrix of x and its derivatives. $\underline{L}_{xy}$ is the covariance matrix of x and its derivatives and those of y. $\underline{L}^t_{xy}$ is the transpose of $\underline{L}_{xy}$. $\hat{\underline{a}}$ and $\hat{\underline{c}}$ are vectors with elements $\hat{a}_n$ and $\hat{c}_m$, $\underline{Y}_y$ is the vector of covariances of y and its derivatives, and $\underline{Y}_x$ is the vector of covariances of y with x and its derivatives.

Since

$$y(t) = y'(t) + n(t) \tag{2.48}$$

Equation (2.46) can be written in terms of y'(t) and n(t).

$$\begin{aligned} -[\underline{L}_{y'}+\underline{L}_n+\underline{L}_{y'n}+\underline{L}^t_{y'n}]\,\hat{\underline{a}} + [\underline{L}_{xy'}+\underline{L}_{xn}]\,\hat{\underline{c}} &= \underline{Y}'_{y'}+\underline{Y}'_n+\underline{n}_{y'}+\underline{n}_n \\ -[\underline{L}^t_{xy'}+\underline{L}^t_{xn}]\,\hat{\underline{a}} + \underline{L}_x\,\hat{\underline{c}} &= \underline{Y}'_x+\underline{n}_x \end{aligned} \tag{2.49}$$

where $\underline{L}_n$ is the covariance matrix of the noise derivatives, $\underline{L}_{y'n}$ is the covariance matrix of the derivatives of y' and n, $\underline{Y}'_n$ is the vector covariance of y' and noise derivatives, $\underline{n}_{y'}$ is the vector covariance of n and derivatives of y', $\underline{n}_n$ is the vector covariance of n and its derivatives, and $\underline{n}_x$ is the vector covariance of n and x and its derivatives. The other quantities are defined either as before or else with y' serving instead of y.

The remnant noise n(t) is uncorrelated with y'(t) or x(t), so for simplicity we might assume that all terms involving covariances of the noise or its derivatives with y'(t) and x(t) or their derivatives to be zero. Doing this we obtain

$$
\begin{aligned}
-[\underline{L}_{y'} + \underline{L}_n]\, \hat{\underline{a}} + \underline{L}_{xy'}\, \hat{\underline{c}} &= \underline{Y}'_{y'} + \underline{n}_n \\
-\underline{L}^t_{xy'}\, \hat{\underline{a}} \qquad + \underline{L}_x\, \hat{\underline{c}} &= \underline{Y}'_x
\end{aligned}
\tag{2.50}
$$

When the remnant noise is zero the regression equations reduce to a form that gives correct (unbiased) estimates of the coefficient of the differential equation, Eq. (2.40), if the assumed differential equation is of the same order as the actual equation. Under such conditions Eq. (2.50) becomes

$$
\begin{aligned}
-\underline{L}_{y'}\, \hat{\underline{a}} + \underline{L}_{xy'}\, \hat{\underline{c}} &= \underline{Y}'_{y'} \\
-\underline{L}^t_{xy'}\, \hat{\underline{a}} + \underline{L}_x\, \hat{\underline{c}} &= \underline{Y}'_x
\end{aligned}
\tag{2.51}
$$

If the remnant noise is not zero the estimates of a_n and c_m obtained from Eqs. (2.49) or (2.50) will be biased and not equal to their correct values. Surber (ref. 21) points out that this bias will be small if the ratio of noise power to signal power (power in $y'(t)$) is less than 1 to 4 per cent and a sufficiently long sample is used in the computation, conditions not frequently encountered in analysis of human operator dynamics.

There are a number of ways in which the coefficients of Eq. (2.43) can be determined, but all methods give equivalent results insofar as their long sample length performance is concerned. Direct solution of the matrix Eq. (2.46) will give the least mean-square error solution. Other methods such as those based on the method of steepest descent (ref. 11), or iteration, cannot do better than direct solution of Eq. (2.43). The difficulty encountered in building differentiators leads to some modification in the error criteria and in the form of the equations that are solved. However, the procedure that we have described will indeed yield the least mean-square error approximation to system output for the interval T and no other method can do better. Thus, we see that the equation error method for finding the coefficients of the differential equation for a system is a direct application of multiple regression analysis.

2. Output Error Method

The second differential equation method, the output error method, is illustrated in Fig. 6. A linear filter with weighting function $m(t-t')$ is used as a model for the system to be identified. It operates on the input signal $x(t)$ and produces the response $z(t)$. The parameters of this model are variable and are adjusted so that the filter output $z(t)$ matches the system output $y(t)$ with least mean-square error. The model characteristics can be represented by a differential equation of the form of Eq. (2.40). The coefficients of this equation are the parameters to be adjusted. When a good match is obtained between the outputs, the coefficients of the differential equation for the model are estimates of the coefficients of the differential equation representation for the unknown system.

There are several algorithms that one can use to adjust the model parameters (refs. 10, 11, 21, 22, 23). We will not discuss in detail any of these methods, but will concentrate on the properties of the estimates of the coefficients of Eq. (2.40) that are obtained from the technique illustrated in Fig. 6. Two properties of these estimates are of interest: their expected values and their variability. Consider first the expected values of the estimates of the coefficients obtained from the method.

As the first step in our investigation of the properties of estimates of differential equation coefficients, we assume a structure for the model in Fig. 6, that is we must specify the order of the differential equation representation for the system. Next, note that the model operates only on the input $x(t)$, which is assumed to be uncorrelated with the remnant noise $n(t)$. Therefore the output of the model will also be uncorrelated with the noise. Since the noise is uncorrelated with the model input and output, it does not affect the expected values of the estimates of the differential equation coefficients, and the noise may, for the present, be taken to be zero. With the noise zero direct observation of $y'(t)$, the system output prior to the addition of the noise, is possible. If the structure of the model is sufficiently similar to that of the system, the model output $z(t)$ will be very nearly equal to $y'(t)$ and the differential equation representing the model,

$$\sum_{n=0}^{N} \hat{a}_n p^n z = \sum_{m=0}^{M} \hat{c}_m p^m x \qquad (2.52)$$

may be replaced by

$$\sum_{n=0}^{N} \hat{a}_n p^n y' = \sum_{m=0}^{M} \hat{c}_m p^m x \qquad (2.53)$$

Equation (2.52) is of the same form as Eq. (2.40) which represents the system. N and M are the assumed orders of the derivatives of $z(t)$ and $x(t)$ in the model. The coefficients of Eq. (2.52), $\hat{a}_n$ and $\hat{c}_m$, are estimates of the coefficients of Eq. (2.40), a_n and c_m. Equation (2.53) is obtained by substituting $y'(t)$ for $z(t)$ in Eq. (2.52). It will represent the model accurately if $z(t)$ is very nearly equal to $y'(t)$.

Equation (2.53) is in a form suitable for solution by the equation error method discussed above. The coefficients $\hat{a}_n$ and $\hat{c}_m$ are solutions to the regression equations of Eq. (2.51). In the notation of Eq. (2.5), Eq. (2.51) may be written

$$\underline{L}\ \underline{b} = \underline{y}' \tag{2.54}$$

where $\underline{L}$ is the combined covariance matrix which is partitioned in Eq. (2.51) into $\underline{L}_y$, $\underline{L}_{xy}$, $\underline{L}^t_{xy}$ and $\underline{L}_x$. $\underline{b}$ is the vector of coefficients $\hat{a}_n$ and $\hat{c}_m$, and $\underline{y}'$ is the covariance vector found by combining $\underline{y}'_y$, and $\underline{y}'_x$.

Following the methods used to derive Eq. (2.10), the right side of Eq. (2.54) may be written in terms of the true coefficients by substituting for $y'(t)$ in the vector $\underline{y}'$ the expression

$$y'(t) = \sum_{m=o}^{M} c_m\, p^m x - \sum_{n=1}^{N} a_n\, p^n y \tag{2.55}$$

Doing this

$$\underline{L}\ \underline{b} = \underline{L}\ \underline{w} + \underline{L}_{K+1}\ \underline{w}_{K+1} \tag{2.56}$$

or

$$\underline{b} = \underline{w} + \underline{L}^{-1}\ \underline{L}_{K+1}\ \underline{w}_{K+1} \tag{2.57}$$

where $\underline{w}$ is the vector of coefficients a_n and c_m from Eq. (2.40) $\underline{L}_{K+1}$ is the covariance matrix of the derivatives of $y(t)$ and $x(t)$ that appear in the equation for the system, Eq. (2.40), but not in the model, Eq. (2.52), and $\underline{w}_{K+1}$ are the coefficients of the system that are not represented in the model. The notation in Eqs. (2.56) and (2.57) is the same as that used in Eq. (2.11)

The expected values of the estimates of the coefficients β_j are obtained by replacing the covariances in Eq. (2.57) by their expected values.

$$\underline{\beta} = \underline{w} + \underline{\gamma} \tag{2.58}$$

where $\underline{\beta}$ is the vector whose components are β_j and $\underline{\gamma}$ is obtained from the second term on the right of Eq. (2.57) by replacing the covariances by their expected values. $\underline{\gamma}$ is a bias term that is caused by not using a model whose structure is the same as the system.

Now consider the variability of the estimates $\hat{a}_n$ and $\hat{c}_m$. Assume that by some means the expected values of these estimates have been found. We designate these expected values a_{no} and c_{mo}, respectively. If we no longer assume the remnant n(t) to be zero, and if these values of coefficients are used in the model as initial conditions for a series of measurements of system coefficients, the estimates obtained will fluctuate about their original expected values. The fluctuation is caused by the remnant noise n(t).

To a first approximation the output of the model can be written

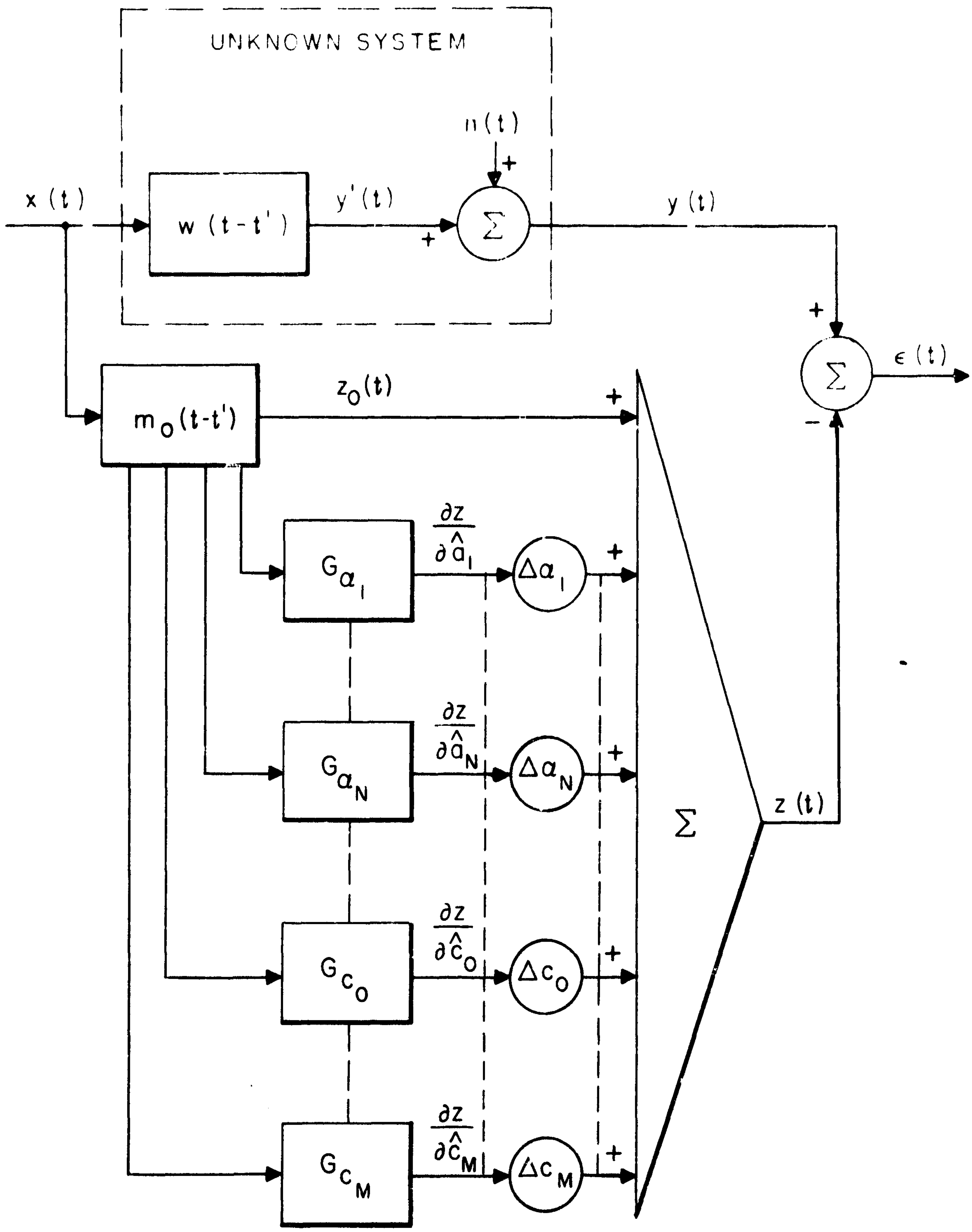

Figure 7 Incremental Analysis of Output Error Differential Equation Method

$$z(t) \doteq z_o(t) + \left(\frac{\partial z}{\partial \hat{a}_1}\right)_{a_{1o}} \Delta a_1 + \left(\frac{\partial z}{\partial \hat{a}_2}\right)_{a_{2o}} \Delta a_2 + \ldots$$

$$+ \left(\frac{\partial z}{\partial \hat{a}_N}\right)_{a_{No}} \Delta a_N + \left(\frac{\partial z}{\partial \hat{c}_o}\right)_{c_{oo}} \Delta c_o + \left(\frac{\partial z}{\partial \hat{c}_1}\right)_{c_{1o}} \Delta c_1 + \ldots + \left(\frac{\partial z}{\partial \hat{c}_M}\right)_{c_{Mo}} \Delta c_M \qquad (2.59)$$

or

$$z(t) \doteq z_o(t) + \sum_{n=1}^{N} \left(\frac{\partial z}{\partial \hat{a}_n}\right)_{a_{no}} \Delta a_n + \sum_{m=o}^{M} \left(\frac{\partial z}{\partial \hat{c}_m}\right)_{c_{mo}} \Delta c_M$$

where $z_o(t)$ is the model output when the expected values of the estimates of the coefficients are used, and Δa_n and Δc_m are increments in the model coefficients caused by the noise. They are to be determined. The partial derivative $\partial z/\partial \hat{a}_n$ and $\partial z/\partial \hat{c}_m$ are called parameter influence coefficients and are evaluated at a_{no} and c_{mo}.

Equation (2.59) is in a form suitable for representation in terms of the multiple regression paradigm. Such a representation is in Fig. 7. The parameter influence coefficients, $\partial z/\partial \hat{a}_n$ and $\partial z/\partial \hat{c}_m$, can be obtained by a linear operation on the signals in the computer simulation of the model (refs. 10, 11, 21, 22). The linear operators are $G_{a1} \ldots G_{cM}$. Their outputs, the parameter influence coefficients, are weighted by the increments in the parameters, Δa_n and Δc_m, and summed with $z_o(t)$ to form the model output $z(t)$.

The parameter increments that minimize the mean-square difference between y(t) and z(t) are found from a set of regression equations which, in this case are

$$
\begin{aligned}
&\sum_{n=1}^{N} \overline{\frac{\partial z}{\partial \hat{a}_n}\frac{\partial z}{\partial \hat{a}_1}}\,\Delta a_n + \sum_{m=o}^{M} \overline{\frac{\partial z}{\partial \hat{c}_m}\frac{\partial z}{\partial \hat{a}_1}}\,\Delta c_m = \overline{(y-z_o)\frac{\partial z}{\partial \hat{a}_1}} \\
&\qquad \vdots \\
&\sum_{n=1}^{N} \overline{\frac{\partial z}{\partial \hat{a}_n}\frac{\partial z}{\partial \hat{a}_N}}\,\Delta a_n + \sum_{m=o}^{M} \overline{\frac{\partial z}{\partial \hat{c}_m}\frac{\partial z}{\partial \hat{a}_N}}\,\Delta c_m = \overline{(y-z_o)\frac{\partial z}{\partial \hat{a}_N}} \\
&\sum_{n=1}^{N} \overline{\frac{\partial z}{\partial \hat{a}_n}\frac{\partial z}{\partial \hat{c}_o}}\,\Delta a_n + \sum_{m=o}^{M} \overline{\frac{\partial z}{\partial \hat{c}_m}\frac{\partial z}{\partial \hat{c}_o}}\,\Delta c_m = \overline{(y-z_o)\frac{\partial z}{\partial \hat{c}_o}} \\
&\qquad \vdots \\
&\sum_{n=1}^{N} \overline{\frac{\partial z}{\partial \hat{a}_n}\frac{\partial z}{\partial \hat{c}_M}}\,\Delta a_n + \sum_{m=o}^{M} \overline{\frac{\partial z}{\partial \hat{c}_m}\frac{\partial z}{\partial \hat{c}_M}}\,\Delta c_m = \overline{(y-z_o)\frac{\partial z}{\partial \hat{c}_M}}
\end{aligned}
\qquad (2.60)
$$

In matrix notation this becomes

$$
\begin{aligned}
\underline{L}_{\partial a}\underline{\Delta a} + \underline{L}_{\partial a \partial c}\underline{\Delta c} &= \underline{\Delta z}_{\partial a} \\
\underline{L}^{t}_{\partial a \partial c}\underline{\Delta a} + \underline{L}_{\partial c}\underline{\Delta c} &= \underline{\Delta z}_{\partial c}
\end{aligned}
\qquad (2.61)
$$

where $\underline{L}_{\partial a}$ is the covariance matrix of parameter influence coefficients of $\hat{a}_n$, $\underline{L}_{\partial a \partial c}$ is the covariance matrix of parameter influence coefficients of $\hat{a}_n$ and $\hat{c}_m$, $\underline{L}_{\partial c}$ is the covariance matrix of parameter influence coefficients of $\hat{c}_m$. $\underline{\Delta z}_{\partial a}$ is the covariance vector of Δz ($\Delta z = y - z_o$) and the $\partial z/\partial\hat{a}_n$ and $\underline{\Delta z}_{\partial c}$ is the covariance vector of Δz and the $\partial z/\partial\hat{c}_m$. $\underline{\Delta a}$ and $\underline{\Delta c}$ are vectors of the coefficient increments Δa_n and Δc_m. If, indeed, $z_o(t) = y'(t)$, Δz will be equal to the remnant noise $n(t)$ and the terms on the right side of Eq. (2.61) can be replaced $\underline{n}_{\partial a}$ and $\underline{n}_{\partial c}$, covariance vectors of $n(t)$ and $\partial z/\partial \hat{a}_n$ and $\partial z/\partial \hat{c}_m$, respectively.

For simplicity we write

$$\underline{L}_{\partial b}\,\underline{\Delta b} = \underline{\Delta z} \tag{2.62}$$

where $\underline{L}_{\partial b}$ is the combined covariance matrix of parameter influence coefficients, $\underline{\Delta b}$ is the combined coefficient increment vector, and $\underline{\Delta z}$ is the combined covariance vector of $\underline{\Delta z}_{\partial a}$ and $\underline{\Delta z}_{\partial c}$ in Eq. (2.61)

Since Δb_j is the variation of b_j about its expected value, its variance will be the variance of b_j. Equation (2.14) and Eq. (2.15) can be applied directly to determine the variance. From Eq. (2.14) we obtain

$$\sigma^2_{bj|s} \doteq \frac{\sigma^2_\epsilon}{Ns^2_{\partial ju}}$$

and (2.63)

$$\sigma^2_{bj} \doteq \frac{\sigma^2_\epsilon}{N\sigma^2_{\partial ju}}\;\frac{M}{M-(K+1)}$$

where $\sigma^2_{bj|s}$ is the variance of b_j for a particular sample of input signal and $s^2_{\partial ju}$ is the sample variance of the part of the $j^{\underline{th}}$ parameter influence coefficient that is uncorrelated with the other parameter influence coefficients and $\sigma^2_{\partial ju}$ is its expected value.

The filters whose outputs are the parameter influence coefficients are closely related to the differential equation for the model, Eq. (2.52) (refs. 10, 11, 21, 22). When the partial derivative of Eq. (2.52) with respect to $\hat{a}_j$ or $\hat{c}_j$ is taken, we obtain

$$\sum_{n=0}^{N} \hat{a}_n p^n \frac{\partial z}{\partial \hat{a}_j} = - p^j z$$

and (2.64)

$$\sum_{n=0}^{N} \hat{a}_n p^n \frac{\partial z}{\partial \hat{c}_j} = p^j x$$

Thus the parameter influence coefficients $\partial z/\partial \hat{a}_j$ and $\partial z/\partial \hat{c}_j$ can be obtained by deriving a filter whose differential equation is the left side of Eq. (2.52) (which is identical to homogeneous part of the differential equation for the model) by $-p^j z$ or $p^j x$, respectively.

Equations (2.61) and (2.59) provide a basis for an algorithm for adjusting the coefficients of the model. The parameter influence coefficients are obtained using filters whose differential equations are of the form of Eq. (2.64). The regression equation, Eq. (2.61), is solved for the increments in coefficients Δa_n and Δc_m. The model coefficients are changed by an amount proportional to these increments and the process is repeated either with the same sample of input and system output signals or with a succeeding sample until stable values of $\hat{a}_n$ and $\hat{c}_m$ are obtained.

Alternatively, one can make use of the fact that the coefficients are to be adjusted so that the mean-square difference $y(t)-z(t)$ is minimized. This difference is

$$\overline{D^2} = \frac{1}{T} \int_0^T [y-z]^2 \, dt = \overline{[y-z]^2} \qquad (2.65)$$

The partial derivative of $\overline{D^2}$ with respect to a coefficient $\hat{a}_j$ is

$$\frac{\partial \overline{D^2}}{\partial \hat{a}_j} = \frac{1}{T} \int_0^T \frac{\partial}{\partial \hat{a}_j} \overline{[y-z]^2}\, dt$$

$$= -\frac{1}{T} \int_0^T \overline{2[y-z] \frac{\partial z}{\partial \hat{a}_j}}\, dt \qquad (2.66)$$

$$= -2\, \overline{D \frac{\partial z}{\partial \hat{a}_j}}$$

A similar expression is obtained for $\hat{c}_j$. If we make the rate of change of $\hat{a}_j$ proportional to $\overline{D \partial z/\partial \hat{a}_j}$, the coefficients will converge along lines of steepest descent to their expected values. Thus

$$\frac{\partial \hat{a}_j}{\partial t} = k\, \overline{D \frac{\partial z}{\partial \hat{a}_j}} \qquad (2.67)$$

This method is easily implemented on an analog computer by multiplying the parameter influence coefficient $\partial z/\partial \hat{a}_j$ (the outputs of the filter $G_{a_j}(s)$ in Fig. 7) by the error $[y(t)-z(t)]$, averaging the product over an interval T and using the result to establish the rate of change of $\hat{a}_j$. (refs. 10, 22). Feeding $\overline{D \partial z/\partial \hat{a}_j}$ to an integrator that drives a servomultiplier on whose shaft is a potentiometer representing $\hat{a}_j$ will establish the appropriate rate of change of $\hat{a}_j$.

IV. COMPARISON OF ANALYSIS TECHNIQUES

The accuracy of the estimates of the system's characteristics and the variability of the estimates are the two factors we consider in a comparative evaluation of the analysis techniques. Two limiting conditions of measurement are of interest: (1) the long sample situation for which, in effect, the signal sample length available is unlimited; and (2) the short sample situation.

A. Long Sample Measurements

If the available sample length T is virtually unlimited, the variability of the measured regression coefficients σ^2_{bj} can be made as small as desired. We need be concerned only with the bias in the expected values of the coefficients given in a general form by Eq. (2.13).

$$\underline{\beta} = \underline{w} + \underline{\gamma} \qquad (2.13)$$

where γ, the bias term is

$$\underline{\gamma} = \underline{\lambda}^{-1}\, \underline{\lambda}_{K+1}\, \underline{w}_{K+1} \qquad (2.68)$$

$\underline{\lambda}$ and $\underline{\lambda}_{K+1}$ are the expected values of the covariance matrices $\underline{L}$ and $\underline{L}_{K+1}$. It will be remembered that $\underline{w}_{K+1}$ is the vector whose elements are the coefficients of the complete representation for the system weighting function $w(t-t')$, (Eq. (2.8)), that are not among the K coefficients used in the model for the system (Eq. (2.7)). $\underline{\lambda}_{K+1}$ is the covariance matrix whose elements λ_{ij} are the expected covariances of the K outputs of the filters included in the model with the outputs of the filters not included in the model ($i \leq K$, $j > K$). If the first K coefficients provide an exact representation of the system weighting function, or if the covariances of the filter outputs in $\underline{\lambda}_{K+1}$ are zero (the outputs are orthogonal), the bias will be zero.

1. <u>Cross-Correlation Analysis</u>

In Fig. 2 we see that the model is composed of a set of time delays whose outputs are identical to the input except for the shift in time. From results in Appendix B, Eq. (B.14), we know that samples of the input signal $x(t)$ will be approximately independent (covariance zero) when they are separated in time by $1/2W_x$ seconds, where W_x is the effective bandwidth of the signal in cycles per second. Therefore, the covariance of any two filter outputs will be approximately zero when

$$\Delta\tau \approx 1/2W_x \tag{2.69}$$

If the model is constructed by selecting the first K equally spaced time delays (delay increment $\Delta\tau$) from an infinite set of such time delays, and if $\Delta\tau \approx 1/2W_x$, the omission of the $(K+1)^{\underline{st}}$ and all higher delays of the same set from the model will not bias the estimates of the first K coefficients. Even if the input bandwidth is not sufficiently large to give unbiased estimates, it should be noted that the covariance of the outputs of the $(K+1)^{\underline{st}}$ and that of the $K^{\underline{th}}$ delay will generally be larger than the covariances of the $(K+1)^{\underline{st}}$ and other outputs in the model whose separation in time is greater. Consequently, the bias introduced by omitting the $(K+1)^{\underline{st}}$ delay will affect the $K^{\underline{th}}$ coefficient more than the other coefficients. In this way the effects of truncating the representation of the system at the $K^{\underline{th}}$ delay tend to be localized to higher terms of the representation.

However, to represent the system weighting function exactly requires an infinite number of delays whose separation in time is infinitesimal. The bias introduced by the omission of these delays, which lie between those used on the model, will be zero only when "white" noise or impulse inputs are used. Otherwise, the measured coefficients will be biased by their omission to the extent that the expected values of b_i will be a weighted average of the system weighting function $w(t-t')$ over an interval of time. If $\Delta\tau = 1/2W_x$, the bias of each coefficient will be determined principally by $w(t-t')$ in a region $\Delta\tau$ seconds long centered about the delay time τ appropriate to that coefficient.

To obtain a reasonably accurate representation for system weighting function particularly when a priori knowledge of the system is not too good, requires a model composed of approximately ten delays extending from zero to some maximum time, τ_{max}. τ_{max} is related to the effective bandwidth of the system and a convenient value for τ_{max} may be obtained from the relation

$$\tau_{max} = 1/2W_w \tag{2.70}$$

where W_w is the effective bandwidth of the system. Using this result and Eq. (2.69) we obtain $W_x \triangleq 10W_w$. Thus to obtain relatively unbiased estimates of the coefficients and still have reasonable resolution in the representation for the system, input bandwidth should be about ten times the system bandwidth. If some biasing of the higher order coefficients is tolerable, this input bandwidth requirement may be relaxed to the extent that a ratio of input bandwidth to system bandwidth of 2.5 to 5.0 may be acceptable.

2. Cross-Spectral Analysis

For long sample lengths all three variations of the spectral analysis method can be made to give unbiased estimates of system transfer function with very little difficulty. For the conventional and the regression spectral analysis methods, Figs. 3 and

4, respectively, the bandpass filters must be designed so that their outputs are orthogonal. To make the filter outputs orthogonal for arbitrary input signals requires that the skirts of the filters be very steep and adjacent filters be non-overlapping. This can usually be accomplished approximately. For the sinusoidal spectral analysis method the sinusoidal components of the input signal will be orthogonal automatically.

When the orthogonality conditions are satisfied the bias $\underline{\gamma}$ in Eq. (2.13) will be zero. Moreover, the covariance matrix $\underline{L}$ will be diagonal and the reduced sets of equations, Eqs. (2.24) and (2.35), will be equivalent to the complete set of regression equations as in Eq. (2.32).

3. Orthogonalized Exponential Function Analysis

In the application of orthogonalized exponential function analysis to long sample measurements the regression coefficients will be biased if the filter outputs are not orthogonal and if $\underline{w}_{K+1}$ in Eq. (2.11) is not zero. The bias can be made zero for all inputs by choosing a model that can match the system weighting function exactly. Except in unusual circumstances this is difficult to do. Another method of reducing the bias is to prefilter the input or to tailor the analysis filters to a particular input so that the filter outputs are approximately orthogonal. Usually, it is not necessary to take these precautions. It is not difficult to design the analysis filters so that they match the system weighting function with very small error and still obtain filter outputs that are reasonably close to being orthogonal. In such a case, the bias error will be small.

4. Differential Equation Simulation Method

With the equation error method, Fig. 5, there are two sources of bias: (1) the conventional bias, represented by γ in

Eq. (2.13), caused when the model does not match $w(t-t')$ exactly and the filter outputs are not orthogonal, and (2) the remnant noise which introduces another bias in measured coefficients as shown in Eq. (2.50).

To reduce the bias $\underline{\gamma}$ a judicious choice of model must be made. The filter outputs in general will not be orthogonal and the only way to make $\underline{\gamma}$ zero is to insure that $\underline{w}_{K+1}$ in Eq. (2.11) is zero. This vector can be made zero by including in the model all of the terms of the differential equation whose coefficients are not zero in the system. This means that the assumed form for the system must be at least of as high an order both in the numerator and denominator of its transfer function as the system being measured. If it is not, the measured coefficients will not be the correct coefficients of the differential equation of the system.

However, the bias resulting from the remnant noise in Eq. (2.50) cannot be removed when there is remnant noise present, except in a few special circumstances. For human dynamics measurements the remnant may be an appreciable fraction of the system output. Hence, measurements made using this technique in the presence of such remnant or output noise will lead to incorrect or biased values for the coefficients of the differential equation. This would appear to be a very serious disadvantage of this method.

The output error method does not suffer from bias caused by remnant noise. The only bias results from an inappropriate choice of model. Equation (2.58) can be used to compute the bias for any particular measurement problem. However, if the sample length is very long there is little reason for not making the model of sufficient order so that any terms omitted will be small and the bias unimportant.

B. Short Sample Measurements

We consider the type of short sample measurement in which a long sample of input signal $x(t)$ is fed to the model and the coefficients of the model are adjusted to achieve a good match for short samples of these signals of the model output $z(t)$ and the system output. Thus at the beginning of each segment the model will have initial conditions approximately equal to those of the system.

If the sample length T is short, both bias and variability must be considered in the evaluation of the analysis techniques. The discussion of bias for the long sample case applies also to the short sample case, so the discussion here will be limited to the question of variability.

Equation (2.15) for the expected variance of σ^2_{bj} provides the basis for the discussion of variability

$$\sigma^2_{bj} = \frac{\sigma^2_{\epsilon}}{N\sigma^2_{ju}} \; \frac{M}{M-(K+1)} \tag{2.15}$$

In addition to the expected variance σ^2_{bj} another measure of variability useful for comparing systems is the relative variability σ_{bj}/β_j.

If two measurement techniques are matched in the sense that they both approximate the system output with the same residual error, one may still be preferred to the other because the coefficients obtained using it have smaller variability. This may result for several reasons. The analysis filter outputs may be more nearly orthogonal for one method making σ_{ju}^2 in Eq. (2.15) larger thereby reducing σ_{bj}^2. The number of degrees of freedom in the analysis filter outputs, M, obtained from a sample T-seconds long may be larger with one method. If measurements are being attempted with very short samples so that M is of the same order as (K+1), the term [M/M-(K+1)] in Eq. (2.15) will have a strong influence on the variance σ_{bj}^2. In such a case, increasing in M may lead to a considerable reduction in σ_{bj}^2. In the same way if one measurement technique requires a smaller number of analysis filters, all other factors being kept equal, σ_{bj}^2 will be smaller because [M/M-(K+1)] will be small. This effect is particularly important for very short samples.

The use of a small number of filters will also have an important effect on the relative variability σ_{bj}/β_j. When the model approximates the system closely

$$\sum_{j=1}^{K} \beta_j^2 \doteq \int_0^{\infty} [w(t)]^2 \, dt \qquad (2.71)$$

Equation (2.71) is a generalization of Parseval's theorem for the equivalence of the integral over frequency of the power density spectrum of a signal and the integral over time of the square of that signal (ref. 4). Because the sum of the squares of the regression coefficients is approximately equal to the integral square of the system weighting function and therefore is independent of the number of filters K, reducing the K on the average will make the coefficients β_j larger. This in turn will tend to reduce the relative variability σ_{bj}/β_j.

Making statements that will be true in general about the relative advantages of the analysis techniques for short sample measurements is difficult because the results obtained with each technique depend greatly upon the nature of the measurement situation, that is, upon that characteristic of the system being measured, the input signal, and the extent to which the analysis filters chosen for the measurement match the system characteristics. For any specific measurement situation, Eq. (2.15) together with the regression equations given in Section III can be used to estimate the variability of the regression coefficient and hence serve as a basis for selecting the most appropriate analysis technique.

In the following discussion we consider each of the analysis techniques and point out some of the major factors that must be considered in applying these techniques to short sample measurements. Some comparisons are made for typical measurement situations.

1. <u>Cross-Correlation Analysis</u>

Equations (2.18) or (2.19) are the basic relations for the cross-correlation analysis technique which is illustrated in Fig. 2. Assuming σ_ϵ^2 and N are maintained constant for all analysis techniques, the variance σ_{bj}^2 in Eq. (2.15) will be determined by the number of filters K, by σ_{ju}^2 (the uncorrelated part of the output of each filter), and by the number of degrees of freedom of each filter output M.

To obtain reasonably complete definition of the system weighting function, typically, about ten points or ordinates of the weighting function should be obtained. This means that about ten filters (time delays) will be required in the model for the system, and that K will equal 10.

Determination of M and σ^2_{ju} is more complicated. If the effective input bandwidth W_x is large compared to $1/\Delta\tau$ ($\Delta\tau$ is the delay increment), the filter outputs will be orthogonal and σ^2_{ju} will be equal to σ^2_x, the input variance. M will be approximately $2W_xT$. Reasonable values of σ^2_{bj} will be obtained for M equal to twenty (twice K) or for $T = 10/W_x$. For values of M smaller than 20 the term $[M/M-(K+1)]$ in Eq. (2.15) will dominate the behavior of σ^2_{bj} and lead to large variance.

For the case in which the input bandwidth is smaller than $1/\Delta\tau$, a not unusual circumstance, the filter outputs will not be orthogonal and σ^2_{ju} as well as M will tend to be reduced thus making σ^2_{bj} larger. In this case both M and σ^2_{ju} can be predicted from the autocorrelation function of the input.

2. Cross-Spectral Analysis

For short samples, the outputs of narrow bandpass filters will not, in general, be orthogonal. Those spectral analysis techniques that assume such orthogonality will give estimates of the regression coefficients that have excessive variability. The reduced set of regression equations obtained by dropping the off-diagonal terms from the covariance matrix such as Eqs. (2.23) and (2.24) for conventional spectral analysis, and Eq. (2.35) for the direct regression, and for the sinusoidal components methods of spectral analysis are not appropriate for the short sample case. The complete equations such as Eqs. (2.22) and (2.32) should be used.

When the conventional spectral analysis method is used with Eq. (2.22), the residual variance σ_{ϵ}^2 and the number of degrees of freedom N in Eq. (2.15) for σ_{bj}^2 relate to the output of the narrow bandpass filter operating on y(t). When Eq. (2.32) is used for the regression or sinusoidal components methods, σ_{ϵ}^2 and N relating to the entire output y(t) should be used in the computation of σ_{bj}^2 from Eq. (2.15). However, if the spectrum of the residual is fairly flat, the ratio σ_{ϵ}^2/N will be approximately the same in both cases.

For the conventional and regression methods when the filters are orthogonal, the expected covariance of any two filter outputs will be zero and σ_{ju}^2 will equal the variance of the $j^{\underline{th}}$ filter output. The number of degrees of freedom of the filter output M is determined by the effective bandwidth of the filter W_j. From Appendix B

$$M = 2W_jT \tag{2.72}$$

Approximately ten pair of bandpass filters typically would be used in the model. Half of these may be assumed to lie in the frequency region below the system bandwidth. Thus, the analysis filter bandwidth is likely to be at least one-fifth the system bandwidth.

The situation is much the same for the sinusoidal components method. The expected variance σ_{bj}^2 given by Eq. (2.15) was derived from $\sigma_{bj|s}^2$ in Eq. (2.14) by taking the expected value of $1/s_{ju}^2$ which was assumed to be distributed as $1/\chi^2$. For sinusoidal signals $1/s_{ju}^2$ will not have this distribution and we might think that Eq. (2.15) would not apply. However, Eq. (2.15) provides a useful approximation to σ_{bj}^2. Consider the following variation of the sinusoidal components method. The sinusoids are added together to form the input signal x(t) which is then fed to a bank of very sharp cut-off narrow bandpass filters. If the center frequencies of the filters are set to the frequencies of the input components and the filter bandwidths are less than the spacing between components, the original sinusoids can be extracted from the composite signal x(t). Since this process is equivalent to that performed on the input signal by the filters employed in the regression method of spectral analysis, the relation for σ_{bj}^2 for that method should hold at least approximately for the sinusoidal components method. Thus

$$\begin{aligned} \sigma_{bj}^2 &\doteq \frac{\sigma_\epsilon^2}{N\sigma_j^2} \frac{M}{M-(K+1)} \\ &\doteq \frac{\sigma_\epsilon^2}{N\sigma_j^2} \frac{2\Delta WT}{2\Delta WT-(K+1)} \end{aligned} \tag{2.73}$$

where the spacing between components ΔW is used to compute the degrees of freedom M.

The number of components used in the sinusoidal method and their spacing is likely to be the same as for the direct regression method, that is, about five components are likely to span the system bandwidth. Thus ΔW will equal about one-fifth the system bandwidth.

If the complete set of regression equations is used for each of the three spectral analysis methods, the estimates of the regression coefficients will have approximately the same expected variance in all three cases. The number of filters used with these methods is likely to be larger than the number used in the cross-correlation method (twenty compared to ten), but the filter outputs will be uncorrelated. Thus, σ^2_{ju} will be larger for the spectral analysis methods than for cross-correlation method. The number of degrees of freedom M for the spectral analysis methods will be less than for the cross-correlation method, where M was determined principally by the input bandwidth. For the spectral methods M is determined by the filter bandwidths which are likely to be a fraction of the input bandwidth.

3. <u>Orthogonalized Exponential Function Analysis</u>

If the poles of the orthogonalized exponential functions, Eq. (2.39), are chosen so that they lie close to or bracket the poles of the weighting function that is being measured, only a few filters will be required to approximate $w(t-t')$ with small error. (See Appendix A.) Typically, four or five filters will be sufficient to approximate $w(t-t')$ with a mean-square error of well under one per cent (ref. 2). Because the number of filters K is small the coefficients β_j will tend to be large and fewer degrees of freedom M are required to keep $[M/M-(K+1)]$ in Eq. (2115) approximately unity.

Since the poles of the analysis filters are chosen to be in the neighborhood of the poles of the system, the bandwidths of the analysis filters will be of the same order of magnitude as the system. For this reason M will be larger than for the spectral analysis method (where narrow bandwidth filters are used) and probably smaller than for the cross-correlation method (where M is determined principally by the input signal).

On the other hand, the outputs of the orthonormal exponential filters will not be orthogonal unless the input bandwidth is much larger than the filter outputs. For input bandwidths of the same order as the system σ^2_{ju} will generally be lower than for the spectral method (for which filters were designed to give orthogonal outputs for all inputs) and larger than the cross-correlation method (for which the filter outputs are highly correlated unless the input bandwidth is very much greater than the system's).

Thus the orthogonalized exponential method is something of a compromise between the cross-correlation and cross-spectral methods. It is better than the correlation method insofar as σ^2_{ju} is concerned and better than the spectral method with respect to M. On the other hand, it is worse than the correlation method with respect to M and worse than the spectral method with respect to σ^2_{ju}.

In one respect, the orthogonalized exponential method is superior to either of these other two methods. The number of K filters required to represent the system accurately generally will be smaller, (by a factor of two to four) for the orthogonalized exponential method than for either of the other two methods. The reason is that the weighting functions of these filters resemble typical system weighting functions. A small K leads to a value of $[M/M-(K+1)]$ in Eq. (2.15) that is closer to unity and to larger values of β_j (because of Eq.(2.71)). As a result σ^2_{bj} will tend to be smaller and, most important, the relative variability σ_{bj}/β_j will be very much smaller for the orthogonalized exponential method.

4. Differential Equation Coefficient Methods

If the equation error differential equation coefficient method is to be applied to short sample measurements, it is essential that the computation of the coefficients of the differential equation be accomplished by the matrix method of Eq. (2.45) or by methods equivalent. Real-time parameter tracking techniques (refs. 9 and 10) require too much time to reach the correct coefficient values and in effect this type of adjustment procedure wastes information about the coefficients contained in the signals. These methods may be used with good success if the same sample of data is fed repetitively to the parameter tracking device and the coefficients adjusted on the basis of repeated analysis of the same data. Such a method at best will do as well as (but not better than) the matrix methods discussed here. The same difficulty exists with the output error method. Real-time parameter tracking (ref. 10) is wasteful of information, and the same data should be analyzed repetitively until the asymptotic values of the coefficients are obtained. Matrix methods are not available for the output error method.

The variance of σ^2_{bj} of the coefficients obtained with the equation error method is given by Eq. (2.15). The variance obtained with the output error method is given by Eq. (2.63). The differential equation representation for system characteristics is a fairly efficient representation in the sense that the number of coefficients required to represent the system is equal to the number of terms in its differential equation and can be made as small as permitted by the equation. However, the correlation amongst the filter outputs will be large. For the equation error method the output of the $j^{\underline{th}}$ differentiation will be orthogonal to the outputs of the $(j+1)^{\underline{st}}$ and $(j-1)^{\underline{st}}$ differentiators, but it will not be orthogonal to the outputs of the

$(j+2)^{\underline{nd}}$ and $(j-2)^{\underline{nd}}$ differentiators. Moreover, the derivatives of the input $x(t)$ and those of the output $y(t)$ will be correlated because of the relation between these quantities imposed by the system equation itself, Eq. (2.40). As a result σ^2_{ju}, the part of the $j^{\underline{th}}$ filter output that is uncorrelated with all the others, will tend to be small.

For the output error method, σ^2_{ju} will also tend to be small. As can be seen from Eq. (2.64) the filters used to extract the parameter influence coefficients (Fig. 7) all have the same transfer functions but are excited by different derivatives of the model output or system input. The filter outputs will be correlated because of the correlation among the signals driving the filters.

For the equation error method, the number of degrees of freedom M in the outputs of the differentiators operating on the input $x(t)$ will be at least as large as the number of degrees of freedom in $x(t)$ and will, therefore, be determined by the input bandwidth. For the outputs of the differentiators operating on $y(t)$, M will be determined largely by the system bandwidth unless the bandwidth of $x(t)$ is much less than that of the system.

For the output error method M will be determined by the system bandwidth or the input bandwidth, whichever is smaller. The filters in Fig. 7 have transfer functions equal to the denominator of the transfer function of the model, which should resemble closely the denominator of the system transfer function.

The differential equation methods have many of the same advantages as the orthonormal exponential function methods. M is likely to be greater than for the spectral methods and probably less than for the cross-correlation method. σ^2_{ju} is likely to be less than that obtained with the spectral methods, but greater than that with the cross-correlation method.

Finally, since system characteristics can be represented by only a few filters, β will tend to be large and the relative variability will be lower than with either the spectral or the cross-correlation methods.

V. DETERMINING THE COEFFICIENTS OF THE SYSTEM DIFFERENTIAL EQUATION FROM THE ORTHONORMAL EXPONENTIAL FUNCTION ANALYSIS METHOD

One of the principal objectives of a system identification procedure is to obtain an analytic expression for the transfer function of the system. We now show how the coefficients of an assumed differential equation representation for the system can be computed from the regression coefficients obtained from the orthonormal exponential function analysis method.

The expected values of the regression coefficients obtained when an orthonormal set of functions is used to represent the unknown system are

$$\beta_j = \int_0^{\infty} w(t)\,\overline{\phi_j}(t)\,dt = \int_{c-j\infty}^{c+j\infty} W(s)\,\overline{\Phi_j}(-s)\,\frac{ds}{2\pi j} \qquad (2.74)$$

where $w(t)$ is the system impulse response and $\phi_j(t)$ the impulse response of the $j^{\underline{th}}$ filter used to represent $w(t)$ in the regression paradigm at Fig. 1. The bar indicates that the complex conjugate is to be taken.

If the $\phi_j(t)$ are orthogonalized exponential functions, they will have transfer functions of the form

$$\phi_1(s) = \frac{\sqrt{-2s_1}}{(s-s_1)}$$

$$\phi_k(s) = \frac{\sqrt{-2s_k}}{(s-s_k)} \prod_{j=1}^{k-1} \frac{(s+s_j)}{(s-s_j)} \tag{2.39}$$

where the s_j are negative real.

Substituting $\Phi_j(s)$ into Eq. (2.74) and integrating over the right half-plane we obtain

$$\beta_1 = \sqrt{-2s_1} \quad W(-s_1)$$

$$\beta_2 = \sqrt{-2s_2} \left[W(-s_2) \frac{(s_2+s_1)}{(s_2-s_1)} + W(-s_1) \frac{(2s_1)}{(s_1-s_2)} \right]$$

$$\beta_k = \sqrt{-2s_k} \left[\sum_{j=1}^{k} \left[W(-s_j) \frac{\prod_{i=1}^{k-1} (s_j+s_i)}{\prod_{\substack{i=1 \\ i\neq j}}^{k} (s_j-s_i)} \right] \right]$$

It should be noted that the β_j are functions of the true transfer function of the system at $s = -s_1, -s_2, \ldots -s_k$ (the s_k are the poles of the basis functions). Thus, whereas

$$W(s) \doteq \sum_{j=1}^{K} \beta_j \, \Phi_j(s) \tag{2.76}$$

is an approximation to $W(s)$, the β_j are functions of the true or exact $W(s)$. That is

$$W(-s_k) = \sum_{j=1}^{K} \beta_j \, \Phi_j(-s_k) \tag{2.77}$$

We can take advantage of this result to find the coefficients of the differential equation for $W(s)$. To do this we must assume a form for the differential equation such as

$$\hat{a}_N \frac{d^N y'}{dt^N} + \dots\dots + \hat{a}_1 \frac{dy'}{dt} + y' = \hat{c}_M \frac{d^M x}{dt^M} + \dots\dots + \hat{c}_1 \frac{dx}{dt} + \hat{c}_o x \tag{2.78}$$

The Laplace transform of this equation is

$$(\hat{a}_N s^N + \dots\dots \hat{a}_1 s+1) \; Y'(s) = (\hat{c}_M s^M + \dots\dots \hat{c}_1 s+\hat{c}_o) \; X(s) \tag{2.79}$$

and the assumed system transfer function is

$$\hat{W}(s) = \frac{Y'(s)}{X(s)} = \frac{(\hat{c}_M s^M + \dots\dots \hat{c}_1 s+\hat{c}_o)}{(\hat{a}_N s^N + \dots\dots \hat{a}_1 s+1)} \tag{2.80}$$

where $\hat{W}(s)$ is the estimate of the system transfer function and $\hat{a}_N$ and $\hat{c}_M$ are the estimates of the coefficients.

By letting $s = -s_1, -s_2 \dots -s_k$ a set of simultaneous equations involving $\hat{W}(-s_j)$ and β_j can be written which when solved will give the coefficients $\hat{a}_n$ and $\hat{c}_m$. These equations are

$$\beta_1 = \sqrt{-2s_1}\ \hat{W}(-s_1) = \sqrt{-2s_1}\ \frac{\sum_{m=0}^{M} \hat{c}_m(-s_1)^m}{\sum_{n=0}^{N} \hat{a}_n(-s_1)^n}$$

$$\beta_k = \sqrt{-2s_k}\ \sum_{j=1}^{k} \left[\frac{\sum_{m=0}^{M} \hat{c}_m(-s_j)^m}{\sum_{n=0}^{N} \hat{a}_n(-s_j)^n} \quad \frac{\prod_{i=1}^{k-1} (s_j+s_i)}{\prod_{\substack{i=1 \\ i\neq j}}^{k} (s_j-s_i)} \right] \tag{2.81}$$

Estimates of β_j are the b_j determined by multiple regression analysis. These estimates are substituted in the set of equations, Eq. (2.81) and the equations are solved for the $\hat{a}_n$ and $\hat{c}_m$. At the frequencies $-s_1, -s_2, \ldots -s_k$ Eq. (2.81) will provide the most accurate estimates $\hat{a}_n$ and $\hat{c}_m$ since at these frequencies $W(s)$ is most accurately determined.

VI. A COMPARISON BETWEEN OPEN AND CLOSED-LOOP MEASUREMENTS OF DYNAMIC SYSTEMS

A. Introduction

Consider the manual control system whose block diagram is in Fig. 8. The system input $i(t)$, error $e(t)$, pilot's output $c(t)$, and system output $o(t)$ can be observed and recorded. The remnant $n_c(t)$ and the linear part of the pilot's output $c'(t)$ cannot be obtained directly. The pilot's describing function $Y_p(s)$ is to be determined.

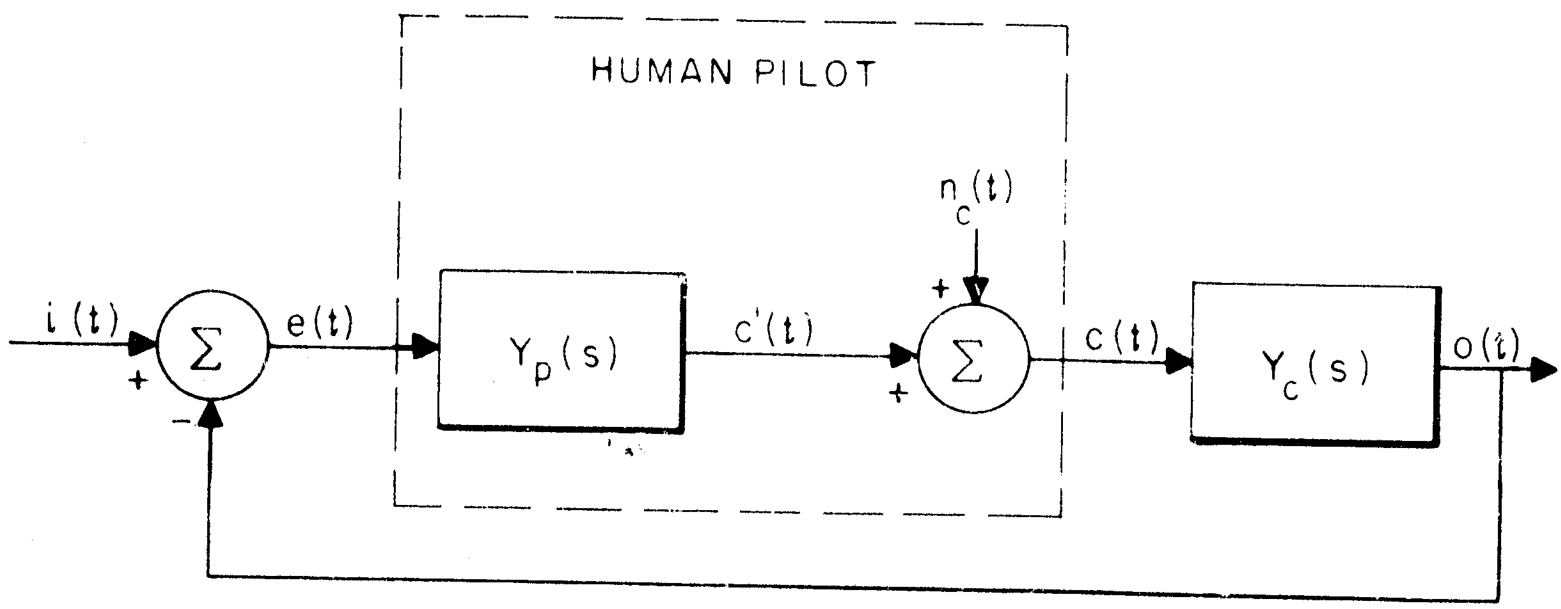

Figure 8 Block Diagram of Typical Manual Control System

It would seem from examination of Fig. 8 that the describing function $Y_p(s)$ could be determined either by determining the response of the pilot to the input forcing function $i(t)$, or by determining his response to the error signal $e(t)$. If the input signal is a random process, and one chose to use the first method, $Y_p(\omega)$ could be computed from the following relation

$$Y_p(\omega) = \frac{S_{ic}(\omega)}{S_{ie}(\omega)} \tag{2.82}$$

where $S_{ic}(\omega)$ is the cross-power density spectrum of input and pilot's output and $S_{ie}(\omega)$ is the cross-power density spectrum of input and error (ref. 14). In effect, Eq. (2.82) defines $Y_p(\omega)$ to be the transfer function relating the part of the pilot's output that is correlated with the input to the part of the error that is correlated with the input.

The second method of computation is based on finding the operation that the system performs on the error in order to produce the output. One is tempted to assume that the describing function $Y_p(\omega)^*$ obtained from the following relation

$$Y_p(\omega)^* = \frac{S_{ec}(\omega)}{S_{ee}(\omega)} \tag{2.83}$$

is equal to $Y_p(\omega)$ obtained from Eq. (2.82). In Eq. (2.83) $S_{ec}(\omega)$ is the cross-power density spectrum between the error and the pilot's output and $S_{ee}(\omega)$ is the power density spectrum of the error. As Graham and McRuer (ref. 14) point out, Eq. (2.83) will <u>not</u> in general give the same result for $Y_p(\omega)$ as does Eq. (2.82).

B. Comparison of the Two Describing Functions

The difference between $Y_p(\omega)$ and $Y_p(\omega)^*$ is made apparent by writing the $S_{ec}(\omega)$ in terms of the input and remnant power density spectra. From the relations inherent in the block diagram of Fig. 8, we see that

$$\begin{aligned} S_{ec}(\omega) &= S_{ec'}(\omega) + S_{en}(\omega) \\ &= Y_p(\omega)\; S_{ee}(\omega) + S_{en_c}(\omega) \end{aligned} \tag{2.84}$$

where $S_{ec'}(\omega)$ is the cross-power spectrum between error and the linear part of the pilot's output. Note that the cross-power density spectrum of the error and the pilot's remnant $S_{en_c}(\omega)$, is not zero since part of the error is caused by the remnant.

Substituting Eq. (2.84) for $S_{ec}(\omega)$ in Eq. (2.83), we obtain

$$Y_p(\omega)^* = Y_p(\omega) + \frac{S_{en_c}(\omega)}{S_{ee}(\omega)} \tag{2.85}$$

Thus it is clear that unless $S_{en_c}(\omega)$ is zero $Y_p(\omega)^* \neq Y_p(\omega)$ and a measurement of $Y_p(\omega)^*$ made using $e(t)$ as the input will not give the same result as a measurement made using $i(t)$.

$S_{en_c}(\omega)$ may be expanded in terms of the input and remnant noise by noting that

$$E(\omega) = \frac{I(\omega)}{1+Y_pY_c(\omega)} - \frac{N_c(\omega)Y_c(\omega)}{1+Y_pY_c(\omega)} \tag{2.86}$$

where $I(\omega)$ and $N_c(\omega)$ are the Fourier transforms of the input and the pilot's remnant.

Using this relation we may write

$$S_{en_c}(\omega) = \frac{S_{in_c}(\omega)}{1+Y_pY_c(-\omega)} - \frac{S_{n_cn_c}(\omega)\ \overline{Y}_c(-\omega)}{1+Y_pY_c(-\omega)} \tag{2.87}$$

where the bar indicates the complex conjugate. Since the input and remnant are uncorrelated $S_{in_c}(\omega)=o$ and

$$S_{en_c}(\omega) = -\frac{\overline{Y}_c(-\omega)\ S_{n_cn_c}(\omega)}{1+Y_pY_c(-\omega)} \tag{2.88}$$

Substitute Eq. (2.88) into Eq. (2.85)

$$\begin{aligned} Y_p(\omega)^* &= Y_p(\omega) - \frac{\overline{Y}_c(-\omega)\ S_{n_cn_c}(\omega)}{[1+\overline{Y_pY_c}(-\omega)]\ S_{ee}(\omega)} \\ &= Y_p(\omega) - \frac{[1+Y_pY_c]\ |Y_c|^2}{Y_c|1+Y_pY_c|^2}\ \frac{S_{n_cn_c}(\omega)}{S_{ee}(\omega)} \end{aligned} \tag{2.89}$$

Note that the spectrum of the part of the error that is not linearly correlated with the input, $S_{n_en_e}(\omega)$, is

$$S_{n_en_e} = \frac{|Y_c|^2 S_{n_cn_c}}{|1+Y_cY_p|^2} \tag{2.90}$$

Using this result in Eq. (2.89), we obtain

$$Y_p(\omega)^* \quad = \quad Y_p(\omega) - \frac{[1+Y_pY_c(\omega)]}{Y_c(\omega)} \; \frac{S_{n_en_e}(\omega)}{S_{ee}(\omega)} \tag{2.91}$$

The ratio $S_{n_en_e}(\omega)/S_{ee}(\omega)$ is the fraction of the error power at frequency ω that is not linearly correlated with the input. If $\rho_e^2(\omega)$ is used to denote the fraction of the error power that is linearly correlated with the input,

$$Y_p(\omega)^* \quad = \quad Y_p(\omega) - \frac{[1+Y_pY_c(\omega)] \; [1-\rho_e^2(\omega)]}{Y_c(\omega)} \tag{2.92}$$

$$= \quad \rho_e^2(\omega) \; Y_p(\omega) - [1-\rho_e^2(\omega)] \; \frac{1}{Y_c(\omega)}$$

From Eq. (2.92) we see that $Y_p(\omega)^*$ at each frequency ω is equal to $Y_p(\omega)$ attenuated minus a factor. In systems where tracking performance is very good, so that the linear part of the system output follows the input very closely, a large part of the error may result from remnant noise and $\rho_e^2(\omega)$ may be small. In such a case, the difference between $Y_p(\omega)^*$ and $Y_p(\omega)$ will be large.

If the remnant is zero $\rho_e^2(\omega)$ will be unity and $Y_p(\omega)^*$ will equal $Y_p(\omega)$. In multiple regression analysis the residual plays the same role as the remnant noise insofar as its effects on the measured coefficients are concerned. Thus, if the model accounts for all of the system output, a perfect measurement will be made when the feedback error signal is used as the input to the model.

If the controlled element dynamics are unity, $Y_c(s)=1$,

$$1-\rho_e^2(\omega) = \frac{S_{n_e n_e}(\omega)}{S_{ee}(\omega)} = \frac{S_{n_c n_c}(\omega)}{S_{ii}(\omega)+S_{n_c n_c}(\omega)} \tag{2.93}$$

and

$$Y_p(\omega)^* = \frac{S_{ii}(\omega)}{S_{ii}(\omega)+S_{n_c n_c}(\omega)} Y_p(\omega) - \frac{S_{n_c n_c}(\omega)}{S_{ii}(\omega)+S_{n_c n_c}(\omega)} \tag{2.94}$$

If the input signal is zero and the only forcing function input to the system is the noise $n_c(t)$, $\rho_e^2(\omega)$ will be zero and from Eq. (2.92)

$$Y_p(\omega)^* = -\frac{1}{Y_c(\omega)} \tag{2.95}$$

Thus, without an input signal the only transfer relationship that one can measure is the reciprocal of the feedback path. As Graham and McRuer (ref. 14) point out, if $n_c(t)$ is not additive noise, but represents the results of non-linearities in the pilot's characteristics, indeed $Y_p(\omega)$ must be equal to $-1/Y_c(\omega)$ if there is to be a circulating signal in the loop. Such a circulating signal is an example of a limit cycle oscillation. On the other hand, if $n_c(t)$ is additive noise that cannot be isolated or measured, then $Y_p(\omega)$ need not be equal to $-1/Y_c(\omega)$ for there to be circulating signals and $Y_p(\omega)^*$ will not be a correct estimate of the human operator's describing function.

One final point, the use of regression analysis techniques in which a model composed of physically realizable filters operates on the pilot's input or the system's input to match the pilot's output leads to estimates of the describing functions that will be physically realizable. We have not introduced this constraint on the nature of the describing functions in the development above. In some circumstances imposing the realizability constraint will change the results obtained.

C. Application to Human Operator Measurements

In Fig. 9 are shown Bode plots of human operator open-loop dynamic response characteristics that were obtained in three ways:

1. Conventional spectral analysis of closed-loop yielding an estimate of $Y_p(\omega)$.

2. Orthonormal exponential analysis of closed-loop yielding an estimate of $Y_p(\omega)$.

3. Orthonormal exponential analysis of open-loop yielding an estimate of $Y_p(\omega)^*$.

These data were obtained in a simple compensatory manual control system in which $Y_c(s)=1$ and the input was an approximation to white noise passed through one RC low-pass filter with 3db frequency at 1.5 radians/sec. The poles of the orthogonalized exponential filters used in the analysis are shown in the figure. Note that the open-loop amplitude ratio measurements $Y_p(\omega)^*$ are 2 - 3db below the results obtained from analysis of closed-loop characteristics $Y_p(\omega)$. This result is in agreement with the previous development of the point that remnant will serve to attenuate the measured system characteristics.

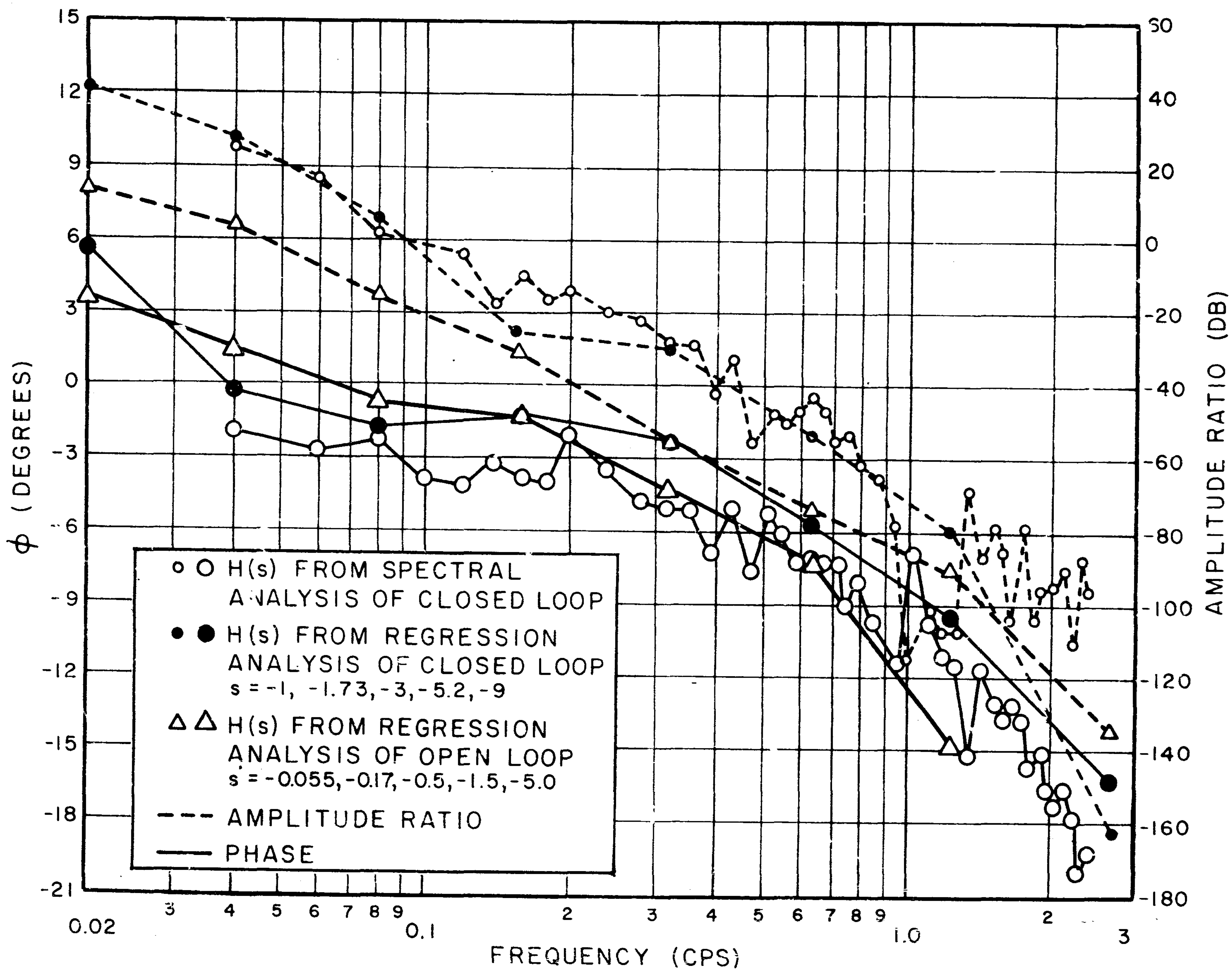

Figure 9 Comparison of Closed and Open Loop Measurements of Human Operator Describing Function

D. Summary

We have analysed the kinds of errors that result from making direct open-loop measurements of system elements. The magnitude of the error depends directly upon the extent to which the input to an element contains components that cannot be accounted for by a linear operation upon the input signal to that element. If a component is a perfect linear operator, there is no difficulty in making direct open-loop measurements of its characteristics. The input signal need not be isolated and recorded, nor need it appear in the place normally assigned to the input forcing function. Any kind of input or random disturbance occurring outside of the limits of the component being measured will suffice to excite the system so that measurements can be made. If a component is either non-linear, or has noise added to its output so that not all of its response can be accounted for by a linear operation upon its input, then an open-loop measurement is likely to be in error. The more nearly linear a device, the more accurate will be the estimate of $Y_p(\omega)$ obtained from open-loop measurements. When dealing with non-linear or time-varying devices, or devices having relatively high noise components in their output, one should exercise caution in interpreting the results of open-loop measurements.

VII. CONCLUSIONS

In this chapter we have shown how the paradigm of multiple regression analysis in Fig. 1 serves as a basis for comparing all of the commonly used techniques for identifying human pilot dynamic response characteristics: cross-correlation analysis, cross-spectral analysis, orthogonalized exponential analysis and differential equation coefficient methods. Expressions for the expected values and the variances of the measures obtained using these techniques have been derived.

Although the characteristics of the measurement situation have to be specified in detail in order to make accurate comparisons of the methods, a number of generalizations can be made. For the long sample case all methods except the equation error differential equation coefficient method will yield equivalent results if the model for the system is sufficiently complete. If the model is not able to match the system weighting function with high accuracy and if the outputs of the filters are not orthogonal, the measurements will be biased. It is not difficult to design the measurement procedure so that the bias will be small. The equation error differential equation method will give biased estimates of the coefficients if the pilot's output contains remnant noise. This bias cannot be removed except in a few special cases and therefore this method should be avoided.

For the short sample case all of the methods can be used provided the complete set of regression equations are used to find the regression coefficients. The normal procedure followed in spectral analysis of assuming that the off-diagonal covariances are zero will lead to estimates of the regression coefficients that will have excessively large variance. These approximate methods should not be used for short sample measurements. Similarly, differential equation methods in which the coefficients are adjusted by a parameter tracking technique operating continuously upon the signals lead to estimates having excessive variance. To apply differential equation methods to short samples the same sample of signals should be fed to the model repetitively until the coefficients have reached their asymptotic values. The orthogonalized exponential and differential equation methods require fewer coefficients to represent typical systems and therefore tend to give estimates having smaller variances and smaller relative variability than the cross-correlation and cross-spectral methods.

Once the regression coefficients have been determined from the orthogonalized exponential analysis method, the coefficients of an assumed differential equation for the system can be determined by solving a set of equations.

When direct open-loop measurements of human pilot describing function are to be made using the error signal as the input to the model, the describing function obtained will, in general, be different from that obtained from closed-loop measurements (after transformation of the closed-loop transfer function to its equivalent open-loop) in which the system input is used as the input to the model. The difference increases as $\rho_e^2(\omega)$, the fraction of the error power that is correlated with the input, decreases. Care should be exercised in interpreting results of open-loop measurements when $\rho_e^2(\omega)$ is much less than unity.

REFERENCES

1. Elkind, J. I. and D. M. Green, "Measurement of Time-Varying and Non-Linear Dynamic Characteristics of Human Pilots," ASD-TR-61-225, Aeronautical Systems Division, Wright-Patterson Air Force Base, Ohio, May 1961.

2. Elkind, J. I., E. A. Starr, D. Lucille Darley and D. M. Green, "Implementation of a Technique for Determining Time-Invariant and Time-Variant Dynamic Characteristics of Human Pilots," NASA TN D-1897, National Aeronautics and Space Administration, Washington, D. C., May 1963.

3. Elkind, J. I., D. M. Green and E. A. Starr, "Application of Multiple Regression Analysis to Identification of Time-Varying Linear Systems," IEEE Transactions on Automatic Control, pp. 160-162, Vol. AC-8, April 1963.

4. Lee, Y. W., Statistical Theory of Communication, John Wiley and Sons, New York, 1960.

5. Goodman, T. P. and J. S. Reswick, "Determination of System Characteristics from Normal Operating Records," Transactions ASME, February 1956.

6. Elkind, J. I., "Characteristics of Simple Manual Control Systems," TR 111, Lincoln Laboratory, Lexington, Massachusetts, April 1956.

7. Seltzer, L. J. and D. T. McRuer, "Survey of Cross-Spectral Analyzers," WADC TR 59-241, Wright Air Development Division, Wright-Patterson Air Force Base, Ohio, December 1959.

8. McRuer, D. T. and E. S. Krendel, "Dynamic Response of Human Operators," WADC-TR-56-524, Wright Air Development Center, Wright-Patterson Air Force Base, Ohio, October 1957.

9. Ornstein, G. N., "Applications of a Technique for the Automatic Analog Determination of Human Response Equation Parameters," Report No. NA61H-1, North American Aviation, Inc., Columbus, Ohio, January 1961.

10. Adams, J. J., "A Simplified Method for Measuring Human Transfer Functions," NASA TN D-1782, National Aeronautics and Space Administration, Washington, D. C., April 1963.

11. Meissinger, H. F., "The Use of Parameter Influence Coefficients in Computer Analysis of Dynamic Systems," Proceedings of the Western Joint Computer Conference, May 1960.

12. Hoel, P. G., Introduction to Mathematical Statistics, John Wiley and Sons, New York, 1947.

13. Cramer, H., Mathematical Methods of Statistics, Princeton University Press, Princeton, New Jersey, 1961.

14. Graham, D. and D. T. McRuer, Analysis of Nonlinear Central Systems, John Wiley and Sons, New York, 1961.

15. Levinson, N. L., "The Wiener RMS Error Criterion in Filter Design and Prediction," Appendix to Extrapolation, Interpolation and Smoothing of Stationary Time Series, by N. Wiener, John Wiley and Sons, New York, 1950.

16. Hall, I. A. M., "Effects of Controlled Element on the Human Operator," Aeronautical Engineering Lab. Rept. 389; Princeton University, Princeton, New Jersey, 1957.

17. Russell, L., "Characteristics of the Human as a Linear Servo-Element," M. S. Thesis, Elect. Engrg. Dept., Massachusetts Institute of Technology, Cambridge, Massachusetts, May 1951.

18. Huggins, W. H., "Representation and Analysis of Signals, Part I: The Use of Orthogonalized Exponentials," AFCRC TR 57-357, Air Force Cambridge Research Center, Bedford, Massachusetts, September 1957.

19. Kautz, W. H., "Transient Synthesis in the Time Domain," IRE Transactions on Circuit Theory, Vol. CT-1, pp. 29-39, September 1954.

20. Young, Tzay Y., "Representation and Analysis of Signals, Part X. Signal Theory and Electrocardiography," John Hopkins University, Baltimore, Maryland, 15 May 1962.

21. Surber, W. H., "On the Identification of Time-Varying Systems," Problems in Communication and Control Systems, Princeton University Conference Proceedings, March 1962.

22. Osburn, P. V., "Investigation of a Method of Adaptive Control," Sc.D. Thesis, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, Massachusetts, January 1962.

23. Kirschner, H. J., "Hill Climbing Methods for the Optimization of Multi-Parameter Noise Disturbed Systems," Proceedings of the Joint Automatic Control Conference, June 1962.

APPENDIX A

NASA SEMINAR

PILOT-VEHICLE IDENTIFICATION PROBLEMS

PRELIMINARY COURSE OUTLINE

SESSION I

INTRODUCTION

A. Statement of the human operator identification problem

B. Review of fundamentals and notation

a. Fourier and Laplace transforms

b. System representations: differential equations, weighting functions, convolution, transfer functions.

c. Harmonic Analysis: correlation functions and and power spectra

d. Simple statistical distributions.

SESSION II

Multiple Regression Analysis of Time-Invariant Dynamic Systems

A. Orthogonal functions: the importance of damped exponentials.

B. Basic measurement technique

C. Representation of correlation functions

D. Selection of filters

E. Examples - Human Operator

SESSION III

Regression Analysis of Time-Varying Systems:

A. Statistical properties of regression co-efficients

B. Estimation of degrees of freedom

C. Effect of initial condition

D. Examples - Human Operator

SESSIONS IV and V

Comparison with Other Measurement Techniques

A. Spectral analysis

B. Parameter tracking

1) G. Bekey

2) J. Adams

C. Direct open versus closed-loop measurements

D. Finding co-efficients of differential equations

SESSION VI

Applications to H-O Measurement and Signal Analysis

A. McRuer: Problems encountered in human pilot measurements: training, stability, linearity

B. Summary

APPENDIX B

List of Attendees for Seminar on System Identification Problems

NASA

Ames Research Center:

E. C. Stewart
S. C. Brown

M. D. White
W. E. McNeill
C. T. Snyder
A. C. Marcy

M. Sadoff
T. E. Wempe
J. D. Stewart
J. C. Howard
R. M. Patton

B. Y. Creer
G. H. Hardy

NASA Headquarters:

R. W. Taylor

Flight Research Center:

L. Taylor
E. C. Holleman

Langley Research Center:

J. Adams
R. Saucer

Systems Technology, Inc.:

D. T. McRuer
R. Magdaleno

Space Technology Laboratories:

G. Bekey
H. Meissinger

Aeronautical System Division (Flight Control Laboratory):

R. J. Wasicko

Martin (Research Institute for Advanced Studies):

F. Muckler
R. Obermayer

Bolt Beranek and Newman Inc:

J. I. Elkind

ACF

W. Stokes