Report R-210

COMPUTER PROGRAM SYNTHESIS BASED ON STATISTICAL COMMUNICATION THEORY

Submitted to the

OFFICE OF NAVAL RESEARCH
Under Contract N5ori-06002
Project NR 232-001

by
Abraham Katz

DIGITAL COMPUTER LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Cambridge 39, Massachusetts

DIC 6782

September 28, 1952

(Thesis Date: September 28, 1951)

FOREWORD

Because it presents information of interest to those studying the application of digital computers to the real-time control of physical systems, this revised thesis investigation is now being issued as an R-series report by the Digital Computer Laboratory at M.I.T.

In such servomechanisms applications, it may frequently be desirable, when computer capacity is available, that the computer be programmed to abstract or otherwise modify the intelligence from a noisy input signal. In this mode of operation, the computer simulates the conventional filter or compensating network. This report concerns itself with methods for optimizing, in a mean square error sense, computer programs which can effect such a filtering process.

The author is indebted to Professor W.K. Linvill for his supervision of this research, and to the Digital Computer Laboratory for the free use of its facilities and for the interest and advice of many of its personnel.

ABSTRACT

Since the use of the large-scale calculating machine as an element in a servomechanism is being actively studied, procedures must be devised for the specification of computer programs which will enable the computer to abstract intelligence from an incoming signal sequence. Here we treat the synthesis of programs (i.e., discrete filters) which process the input sequence in such a manner as to obtain the best possible approximation, in a least square error sense, to a specified function of that sequence. The synthesis procedures are based on equations analogous to the Wiener-Hopf equation. By means of these equations it is possible to specify the optimum least-squares filter whether it be linear or not, time-varying or not. In any case, the optimum filter is obtained by solving a set of simultaneous linear algebraic equations.

TABLE OF CONTENTS

Page

FOREWORD.. ii

ABSTRACT.. iii

LIST OF ILLUSTRATIONS... v

CHAPTER I INTRODUCTION.. 1

1.01 Historical Note.. 1

1.02 Definition of Problem.. 5

1.03 Discussion of Proposed Method of Solution.................... 7

CHAPTER II THEORETICAL ANALYSIS................................... 11

2.0 GENERAL REMARKS... 11

2.01 Synthesis Procedure for Linear Time-Invariant Programs....... 18

2.02 Synthesis Procedure for Linear Time-Varying Programs......... 24

2.03 Synthesis Procedure for Non-Linear Time-Invariant Programs... 31

2.04 Synthesis Procedure for Non-Linear Time-Varying Programs..... 42

2.1 QUANTIZATION NOISE AND NUMERICAL COMPUTATIONS.................. 45

2.11 Analysis of Noise Caused by the Encoding Process............. 46

2.12 Effect of Computational Errors in Discrete Filters........... 53

CHAPTER III EXPERIMENTAL ANALYSIS................................. 54

3.01 Prediction in the Presence of Quantization Noise of Functions Related to Straight-Line Flight.......... 54

3.02 Experimental Results... 62

CHAPTER IV CONCLUSIONS AND SUGGESTIONS FOR FURTHER STUDY.......... 68

APPENDIX.. 72

BIBLIOGRAPHY.. 79

LIST OF ILLUSTRATIONS

Figure	Number	Title	Page
2.00-1	A-37328	Block Diagrams of Linear Constant-Coefficient Systems	12
2.00-2	A-37327	Block Diagrams of Linear Time-Varying Systems	16
2.01-1	A-51953	Theoretical System for Comparing Desired Output with Actual Output	19
2.02-1	A-51954	Filtering System for a Signal Ensemble Whose Statistics Vary Slowly Compared to the Time Constants of the Controlled System	29
2.02-2	A-51946	Approximation of the Weighting Coefficients as Functions of the Time Variable, k	30
2.11-1a	A-50255	Sampling Device	47
2.11-1b	A-50255	Equivalent Sampling Device	47
2.11-1c	A-50255	Sampled Signal	47
2.11-2	A-50256	Quantizing Characteristic	49
2.11-3-A	A-50256	A Quantized Signal and the Corresponding Error	49
3.01-1	A-50259	Hypothetical Air Traffic Control System	57
3.01-2	A-51056	Trajectory for Target Flying Straight-Line Constant-Velocity Course	59
3.02-1	A-51052-2	Frequency of Occurrence of Error for a Two-Element Predictor as a Function of Error Amplitude	65
3.02-2	A-51053-2	Frequency of Occurrence of Error for a Three-Element Predictor as a Function of Error Amplitude	66
3.02-3	A-50951-2	Frequency of Occurrence of Error for a Four-Element Predictor as a Function of Error Amplitude	67

CHAPTER I

INTRODUCTION

1.01 Historical Note

In recent years large-scale digital computers have been brought to that stage of development where their application to the control of dynamic systems is practically realizable. Computer applications, in general, may be divided roughly into the following categories:

a) real-time - that in which the computer must cope with a dynamically changing problem, and in which the computed results directly affect the evolution of the problem (e.g. - chemical process control, simulation), and

b) non-real time - that in which the computer copes with a static or quasi-static problem, and in which the computed results have little or no effect on the evolution of the problem (e.g. - scientific computations, economic analysis).

A real time application to a complex control problem frequently implies a computer having a very high operating speed as well as rather extensive storage facilities. One such computer, Whirlwind I, is being currently developed by the M.I.T. Digital Computer Laboratory. It should be noted that the use of the digital computer in control systems is but a logical extension of servomechanisms applications.

Assuming that the computer and its associated conversion apparatus are properly designed with respect to the particular control system at hand, the question which then arises is the formulation of computer programs which will accomplish the desired results. Viewed in proper perspective, the digital computer can be visualized as a docile slave with a very strong back but a rather feeble mind. Although plagued by a small "memory" and by a limited repertoire of arithmetic and logical "thought processes", this slave is nevertheless capable of consulting this "memory" and executing these processes at truly fantastic speeds. It remains for the applications engineer to instruct the slave in the manner in which "he" is to perform "his" duties.

Generally these computer programs are formulated in terms of difference equations (i.e. - time domain synthesis) with the computer solving the equations by means of arithmetic operations and modifying its course of actions by means of logical operations. In a recent doctoral thesis, Salzer[13] presents a method for the coding of programs which depends on frequency domain synthesis. Since the computer program is essentially a data-processing filter, one would logically expect that many of the methods of analog filter theory are extensible to program synthesis. This thesis constitutes an attempt at program synthesis based on the concepts of Wiener-Lee theory.[1,6,7] In a sense, it is an effort at endowing the digital computer with a higher order of intelligence -- one capable of interpreting a message in terms of its statistical distributions in time.

In his classic monograph, "Extrapolation, Interpolation, and Smoothing of Stationary Time Series", N. Wiener[1] demonstrated the basic unity of the previously unrelated philosophies underlying the investigation of time series in statistics and of message-transmission in communication engineering. This study parallels to a great extent the somewhat earlier work of A. Kolmogoroff[2] and thus represents another instance of "multiple discovery", that curious historical phenomenon which has recurred so frequently in the development of science. Although each of these men was concerned with a slightly different problem, both sought, through an extension of the theory of stochastic processes, to establish the bases for the optimum prediction of time series.

Time series may be defined as sequences, discrete or continuous, of quantitative data assigned to specific moments of time. Since the majority of types of time-variations encountered, both in statistics and in communications, are not of the regular functional type in which the function $f(t)$ can be represented exactly by a mathematical function of t, they can be studied only with respect to the statistics of their distributions in time. From this it follows that the separation of the true data from the noise in any set of sequences must logically be preceded by a study of the statistical characteristics of the set. Since the concepts of statistics are based on large collections or ensembles of events, we remark that the performances of the filters considered herein are to be optimized over an ensemble of possible signals.

By making certain restrictive assumptions as to the properties of the ensemble, Wiener was able to simplify the specification of the optimum filter. The basic concept of his theory was that communication signals were to be regarded as stationary, ergodic time series. A stationary ensemble is one whose statistical properties do not vary with time -- that is, the statistical regularities of the past will hold in the future. An ergodic ensemble is a special case of a stationary ensemble for which any member function is statistically representative of any other member function. For the ergodic ensemble one can, in computing statistical parameters, replace ensemble averages by time averages. This replacement permits a considerable saving in time and effort since the former averaging process is much more difficult than the latter. Furthermore, the assumption of stationarity permits the design of the optimum filter in the form of a constant-coefficient system.

If the concept of treating messages as stationary time series is accepted, the problem which then arises is -- what statistical parameters are of importance in the design of optimum filters? In this, as in most physical problems, the solution is more readily obtained if the limitation of linearity is imposed on the types of operations to be used in the filtering. With this limitation on the characteristics of the filters, it is immediately apparent that the only statistical parameters which need be considered are the signal power spectra -- or their time domain equivalents, the linear correlation functions. It should be noted that if the limitation of linearity is not imposed, then the statistical parameters of importance are both the linear and the higher order nonlinear correlation functions. This point will be treated in somewhat greater detail later (c.f. Chap. II).

Prior to the publication of Wiener's monograph, the problem of filter design was generally handled by the classical methods of (a) prescribing either the desired frequency response or the desired transient response, and (b) synthesizing a network which approximates this response. Not only did Wiener suggest an entirely different approach to the synthesis of linear filters, but he also provided the basis for the development of a theory of information.[3,4,5] Much work has since been done in the interpretation[6,7,8] and extension[9,10,11] of the Wiener theory and in the mechanization[12] of the rather laborious procedure of computing correlation functions.

It should be emphasized that this optimization theory need not be restricted to stationary ensembles nor the filtering system to linear operations. Booton[23] has generalized the theory to include time-varying linear systems with nonstationary statistical inputs. We shall show that, for the "discrete filter", one can specify the optimum system with the same ease regardless of whether the restrictions of linearity and stationarity are imposed or not.

1.02 Definition of Problem

Having sketched the origin of the statistical communication theory, we now consider the extension of its techniques to the synthesis of programs capable of dealing with a specific kind of noise. The noise in question is that arising from the conversion of a continuous time series to one which is discrete.

Information that is to be processed by a digital computer must generally be made available to it in the form of signals which are discretized both in time of occurrence and in magnitude. Thus, in a control application, a continuously varying signal which is related to the controlled variable of the system may be sampled and quantized by means of encoding devices to give one of a finite set of discrete magnitudes for each distinct time interval. Note that the term "sampling" is used for the process of discretizing in time, "quantizing" is used for that in amplitude, and "encoding" is used for the process (involving sampling and quantizing) whereby a continuous time series is converted to one which is discretized in time and amplitude.

Whether the original input signal is noise-free or not, it is distorted in the process of being encoded. This distortion is caused primarily by the fact that the quantizer can reproduce only approximately an instantaneous sample by assigning to it the value of the nearest quantizing level. To attain an accuracy greater than one part in a thousand requires a rather complex electronic system. The result is that this analog-to-digital conversion by the encoder produces data which are good to, say, three decimal figures. When these sampled quantized data are subsequently used in the computation of a correction to the controlled variable of the system, the inevitable accumulation of round-off errors soon destroys the usefulness of the computation. Round-off error stems from the necessity of operating on data which approximate all real numbers by rational numbers with a finite number of digits.

The problem at hand is to synthesize discrete filters (i.e., computer programs) which, when supplied with a noisy data input, will yield the best possible approximation, in a least mean-square error sense, to the message or some function thereof. As an example -- if we assume that quantization noise is the predominant noise component, then, for the register length of the Whirlwind I computer, one might be called upon to specify that system which will process incoming 8-digit data in such a way as to obtain the best possible approximation (in the specified sense) to a 15-digit value of some given function of the input data.

1.03 Discussion of Proposed Method of Solution

Having formulated our problem and indicated the technique that will be used to obtain a solution, we now consider the limitations implicit in the assumption that the criterion of performance shall be the minimization of the mean square error over the ensemble. Whether this is indeed the sense in which a system should be optimized is by no means certain. In establishing suitable performance criteria, engineers are concerned with questions of values rather than questions of fact. The absence of valid criteria which lead to solvable problems requires that the engineer make some arbitrary decision as to the "criterion of goodness" of a system.

In formulating his theory of statistical prediction, Wiener was confronted by this problem and arbitrarily decided to use the criterion of least square error. Of all quantities which lend themselves to an easy minimization, the most natural are those which are inherently positive because they are squares of some simple real expression or sums of such squares. Wiener was well aware that this criterion had serious faults. Firstly, it puts an over-emphasis on those points where the predicted and actual values differ by a large amount; and secondly, it gives a considerable weight to small errors occurring with great frequency over a long interval of time. Thus we find that this criterion is overly sensitive at both ends of the scale of magnitude of error, and slights intermediate values of error.

If one examines this criterion objectively, one finds that, in addition to having the considerable virtue of leading to a mathematically manageable problem, it frequently leads to rather good results. A well-known fact is that the application of the Wiener-Lee theory to the filtering of a Gaussian signal leads to a design which is absolutely optimal. Zadeh[14] has suggested that a more appropriate criterion of design would be the maximization of the probability that the error at a prespecified time, t_o, be less than some prescribed tolerance. The maximization procedure leads to a set of equations whose solution in practice must be carried out by trial and error. If the nature of the system is such that certain conditions are satisfied, then the probability criterion yields the same values for the design constants as does the least square error criterion. These conditions require that either (a) the systematic error as well as the geometric mean of the tolerance and the systematic error at time $t = t_o$ be small in comparison with the r.m.s. value of the random error, or (b) the systematic error be large in comparison with the random error, and the tolerance be approximately equal to the systematic error. Floyd[15] has shown that the criterion of maximization of a probability density is the equivalent of the criterion of least squares when the bias errors for an ensemble of signals have a normal distribution about zero. Stutt,[11] after making an experimental study of "optimum" filters, concluded that least square error network specifications are generally sensible and lead to designs which are relatively non-critical.

The crux of the problem of defining suitable performance criteria lies in the fact that there is little understanding as to what properties of a message are actually utilized by the ultimate receptors and therefore should be retained in the output of the filter. It is obvious that the stationary time series which the filter should handle is a function of the destination. In the absence of information concerning the message-utilizing properties of the ultimate receptors we must content ourselves with an overall system which is not optimal, even in the least square error sense.

The extreme simplicity of the least squares procedure can be illustrated by the following example:

Let it be required that we construct a linear, constant-coefficient filter for which

$$i_k = \text{input at time } t_k$$

$$d_k = \text{desired output at } t = t_k$$

$$o_k = \text{actual output at } t = t_k$$

$$= \sum_{n=0}^{M} A_n i_{k-n}$$

where A_n = coefficients by which input data are to be weighted

and $\epsilon_k = d_k - o_k = \text{error at } t = t_k$

The mean square error taken over all time is then $\overline{\epsilon_k^2}$ and can be minimized by a proper choice of the weighting coefficients, A_n. The conditions for a minimum are

$$\frac{\partial \overline{\epsilon_k^2}}{\partial A_n} = 0 \qquad \text{for } n = 0, 1, \ldots, N.$$

But

$$\overline{\frac{\partial \epsilon_k^2}{\partial A_n}} = \overline{\frac{\partial \epsilon_k^2}{\partial A_n}} = \overline{\epsilon_k \frac{\partial \epsilon_k}{\partial A_n}} = 0$$

or

$$\overline{\epsilon_k\, i_{k-n}} = 0 \qquad \text{for } n = 0, 1, \ldots M.$$

This is the discrete form of the Wiener-Hopf equation for the linear constant-coefficient filter. Expanding this equation, we find that the optimum set of coefficients is determined by the equations

$$\sum_{n=o}^{M} A_n \quad \overline{i_{k-n} i_{k-m}} = \overline{i_k \, d_{k-m}}$$

$$\text{for } m = o, 1, \ldots, M$$

The reader will find that the above described procedure or variations thereof appear throughout this report. Understanding of this simple example enables one to specify rigorously an optimum design for computer programs which are to abstract information from a noisy input data sequence.

CHAPTER II

THEORETICAL ANALYSIS

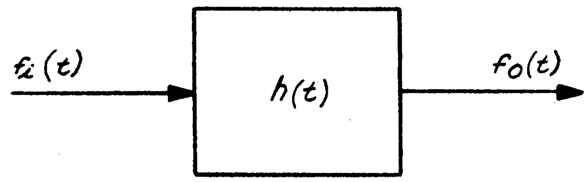
2.0 General Remarks

Before considering the derivation of procedures for specifying classes of optimum filters, we first define the terms "discrete filter" and "continuous filter", and then discuss certain analogies between these types of filters. A "discrete filter" is a transmission device which, when supplied with data from an external source at specified equally spaced moments, furnishes at the same moments output data which depend on past input data and possibly on past output data and time. The input and output signals are discrete time series, and the filter itself is characterized by a sequence (or sequences) of numerical weights. In contrast to this is the conventional "continuous filter" which is characterized by physical elements such as resistance, inductance, and capacitance, and whose input and output are usually continuous time series.

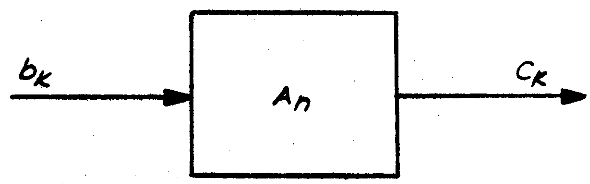
A linear filter, be it discrete or continuous, is one for which there is a linear relation between input and output. This relation may be expressed by recourse to the superposition principle. Confining our attention for the moment to linear systems, we remark that many of the concepts of conventional filter theory have their analogies in the discrete domain. For the constant-coefficient continuous system (Fig. 2.00-1a), the input-output relation is defined by the superposition integral:

$$f_o(t) = \int_{-\infty}^{t} h(t - \tau) f_i(\tau) d\tau \qquad (2.00\text{-}1)$$

$$= \int_{0}^{\infty} h(\tau) f_i(t - \tau) d\tau \qquad (2.00\text{-}1a)$$



a. CONTINUOUS FILTER



b. DISCRETE FILTER

FIG. 2.00-1

BLOCK DIAGRAMS OF LINEAR CONSTANT-COEFFICIENT SYSTEMS

A-37328

where $f_i(t)$ = input signal

$f_o(t)$ = actual output signal

$h(t)$ = response of system to a unit impulse applied at $t = 0$. For a physically realizable system, $h(t) = 0$ for $t < 0$.

To facilitate the understanding of linear systems, one frequently studies the equivalent frequency domain representation obtained by Laplace transforming equation 2.00-1a. Starting from the definition

$$F_o(s) = \int_0^{\infty} f(t)e^{-st}\, dt$$

we obtain

$$F_o(s) = H(s)F_i(s) \tag{2.00-2}$$

where $H(s)$ = System transfer function

Analogous equations defining the corresponding discrete system are easily derived. Before proceeding with the derivation we note that filters in which both input and past output data are weighted are equivalent to those in which an infinite number of input data are weighted. In this report we shall consider only those filters in which a finite number of input data only are weighted.

For the discrete system (Fig. 2.00-1b) in which the coefficients (or elements) of the weighting sequence are invariant with time, the input-output relation is defined by the superposition summation:

$$c_k = \sum_{n=0}^{M} A_n\, b_{k-n} \tag{2.00-3}$$

where $k = 0, 1, 2, \ldots$

= discrete time variable identifying sample datum at $t = t_o + kT_r$

T_r = sampling interval

$b_k = b(kT_r) = b(t_o + kT_r)$
= input sequence

c_k = actual output sequence

A_n = sequence of weighting coefficients analogous to the impulse response of the continuous system.

Since the "memory" of this class of filters has been limited to M+1 samples, we might denote this as a "finite-memory" filter to distinguish it from the class in which both input and past output data are weighted.

A frequency domain representation of this system is readily obtained in the following manner. We define

$$C(e^{-sT_r}) = \sum_{k=o}^{\infty} c_k \, e^{-ksT_r}$$

then, inserting this in equation 2.00-3,

$$C(e^{-sT_r}) = \sum_{k=o}^{\infty} \sum_{n=o}^{M} A_n B_{k-n} e^{-ksT_r}$$

$$= \sum_{n=o}^{M} A_n \, e^{-nsT_r} \sum_{k=o}^{\infty} b_{k-n} \, e^{-(k-n)sT_r}$$

$$C(e^{-sT_r}) = \alpha(e^{-sT_r}) \; B(e^{-sT_r}) \qquad (2.00\text{-}4)$$

where $\alpha(e^{-sT_r}) = \sum_{n=o}^{M} A_n e^{-nsT_r}$

is the system transfer function defined by Salzer.[13]

The analogies between the continuous and the discrete systems, which are now apparent, are set forth explicitly below.

Continuous System	Discrete System
1. Defined by	1. Defined by
a) Differential equation	a) Difference equation
b) Superposition integral	b) Superposition summation
2. Transfer function of constant-coefficient system depends on s.	2. Transfer function of constant-coefficient system depends on e^{-sT_r}.
3. Impulse response or weighting function defined by h(t) for the time-invariant case.	3. Weighting sequence defined by A_n for the time-invariant case.

For the case of the linear time-varying system, the input-output relation of the continuous filter (Fig. 2.00-2a) is defined by the more general form of the superposition integral

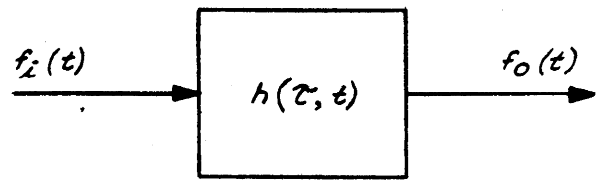
$$f_o(t) = \int_{-\infty}^{t} h(\tau, t) f_i(\tau) d\tau \qquad (2.00\text{-}5)$$

$$= \int_{0}^{\infty} h(t - \sigma, \sigma) f_i(\sigma) d\sigma \qquad (2.00\text{-}5a)$$

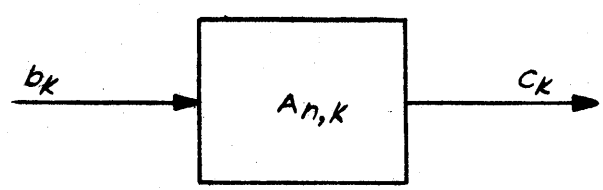
The corresponding discrete system (Fig. 2.00-2b) is analogously defined by the superposition summation

$$c_k = \sum_{n=0}^{M} A_{n,k}\, b_{k-n} \qquad (2.00\text{-}6)$$

$$k = 0, 1, 2, \ldots$$



a. CONTINUOUS FILTER



b. DISCRETE FILTER

FIG. 2.00-2
BLOCK DIAGRAMS OF LINEAR TIME—VARYING SYSTEMS

A-37327

From the foregoing discussion it is evident that many of the techniques of continuous filter theory may be logically extended to discrete systems. Much work has been done by Salzer in utilizing frequency domain methods in describing and in establishing criteria for stability of linear computer programs (linear discrete filters). We shall here concern ourselves with the discrete version of the statistical approach developed by Wiener,[1] Levinson,[16] Kolmogoroff,[2] and others.

Turning our attention now to non-linear systems, we note that in such systems, there is a non-linear relation between input and output. In other words the characteristics of a non-linear filter depend on the signal, and possibly on time. If we restrict our investigation to "finite-memory" systems, we see that a fairly general non-linear program may be described by the equation

$$c_k = \sum_{n=0}^{M} A_{n,k}\, b_{k-n} + \sum_{m=0}^{N} B_{m,k}\, b_{k-m}^2 + \sum_{p=0}^{P} C_{p,k}\, b_{k-p}^3 + \cdots \qquad (2.00\text{-}7)$$

$$k = 0, 1, 2, \ldots$$

If the performance of the program is time-invariant, then

$$A_{n,k} = A_n$$

$$B_{m,k} = B_m$$

$$\vdots$$

In the following sections we shall develop procedures for specifying the optimum (mean-square error) discrete finite-memory filter whether it be linear or not, and whether its input be stationary or not. It should be noted that the specification of the corresponding infinite-memory filters can be derived by methods entirely analogous to those to be set forth below. However, an approximation is needed to make the problem manageable, so that the resulting design is not strictly an optimum.

2.01 Synthesis Procedure for Linear Time-Invariant Programs

The linear time-invariant program is characterized by the input-output relation

$$c_k = \sum_{n=0}^{M} A_n \, b_{k-n} \qquad (2.01\text{-}1)$$

$$k = 0, 1, 2, \ldots$$

Since the actual filter output $\underline{c_k}$ may differ from the ideal, or desired, output $\underline{a_k}$, an error ϵ_k will be present (see Fig. 2.01-1).

Levinson[16] has developed a simple computational procedure for discrete filter design which is applicable to the problem at hand. This procedure is, essentially, classical least squares with the additional specifications of linearity of the filter and stationarity of the input sequence. Recapitulating (with some changes in notation) part of Levinson's procedure, we now determine the nature of the linear program which, with input b_k, will have an output as close as possible to the desired output a_k. An error quantity

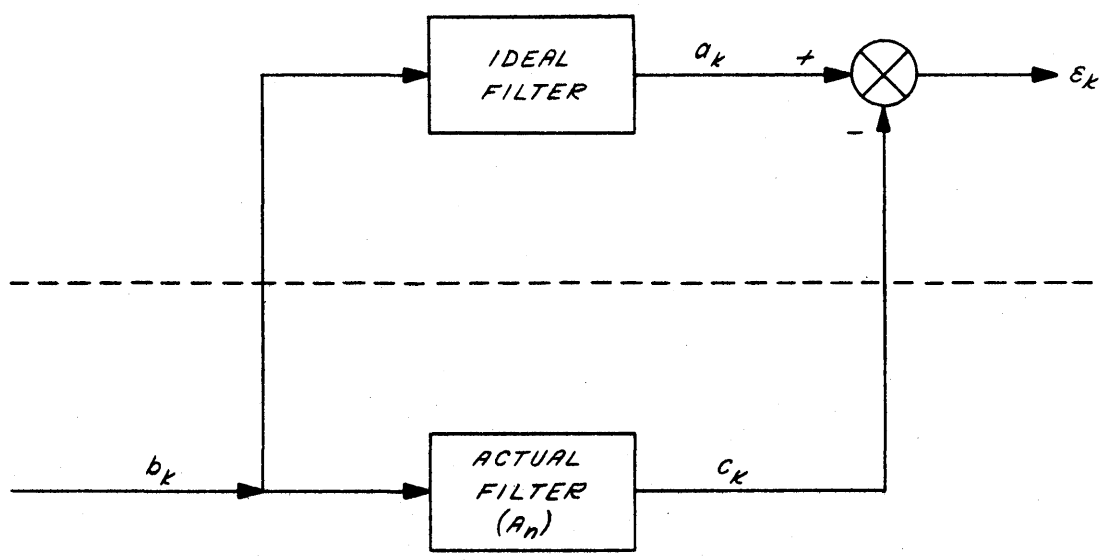


FIG. 2.01-1

THEORETICAL SYSTEM FOR COMPARING DESIRED OUTPUT WITH ACTUAL OUTPUT

A-51953

$$\epsilon_k = a_k - c_k$$

$$= a_k - \sum_{n=o}^{M} A_n b_{k-n} \qquad (2.01\text{-}2)$$

may therefore be defined. Our problem is to derive a sequence of weights A_n such that we minimize

$$I_{1c} = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{k=-N}^{N} (a_k - \sum_{n=o}^{M} A_n\, b_{k-n})^2 \qquad (2.01\text{-}3)$$

= Mean square error for linear constant-coefficient filter

or

$$I_{1c} = \lim_{N \to \infty} \frac{1}{2N+1} \sum_k a_k^2 - 2 \sum_n A_n \lim_{N \to \infty} \frac{1}{2N+1} \sum_k a_k\, b_{k-n}$$

$$+ \sum_n \sum_m A_n A_m \lim_{N \to \infty} \frac{1}{2N+1} \sum_k b_{k-n}\, b_{k-m} \qquad (2.01\text{-}4)$$

Thus far the derivation differs in no respect from classical least squares. If the structure of the series is such that stationarity (or at least quasi-stationarity) is assured, we may now define the auto- and cross-correlation functions:

$$R_{aa}(k) = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{h=-N}^{N} a_h\, a_{h-k}$$

$$R_{ba}(k) = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{h=-N}^{N} a_h\, b_{h-k}$$

Equation (2.01-4) assumes a simpler form when we introduce the correlation functions.

$$I_{1c} = R_{aa}(0) - 2\sum_{n=o}^{M} A_n R_{ba}(n) + \sum_{n=o}^{M}\sum_{m=o}^{M} A_n A_m R_{bb}(m-n) \qquad (2.01\text{-}5)$$

Using the techniques of the calculus of variations, we determine those values of A_n which minimize I_{1c} by setting

$$\frac{\partial I_{1c}}{\partial A_k} = 0 \qquad \text{for } k = 0, 1, \ldots, M.$$

This yields the Wiener-Hopf equation for the linear finite-memory filter:

$$\sum_{n=o}^{M} A_n R_{bb}(k-n) = R_{ba}(k) \qquad (2.01\text{-}6)$$

where $k = 0, 1, \ldots, M$.

For a predicting filter, the Wiener-Hopf equation is of the form:

$$\sum_{n=o}^{M} A_n R_{bb}(k-n) = R_{ba}(k+s) \qquad (2.01\text{-}7)$$

where s = integral number greater than zero denoting the number of sampling periods in the future by which the prediction is made.

$k = 0, 1, \ldots, M$.

The minimum value of I_{1c} for prediction is then

$$\left[I_{1c}\right]_{min} = R_{aa}(0) - \sum_{n=o}^{M} A_n R_{ba}(n+s) \qquad (2.01\text{-}8)$$

Levinson has shown that increasing M always improves the quality of the filtering.

In order to synthesize a linear finite-memory program as a sequence of $M+1$ weights, A_n, one need merely solve the Wiener-Hopf equations (2.01-6) or (2.01-7). In treating the concept of stability for this class of filters, we follow Hurewicz[17] by defining a filter as stable if to a bounded input there always corresponds a bounded output. A necessary and sufficient condition for stability is absolute convergence of the weighting sequence (A_n), that is

$$\sum_{n=0}^{M} \left|A_n\right| < \infty$$

Since we treat only those filters with finite "memories", the only condition required for stability is that all A_n be finite.

Just as the stability of continuous linear systems can be studied in the frequency domain, so can one make a similar study for discrete linear systems. It will be found, in general, that the class of filters under discussion always have a transfer function in the form of a polynomial in e^{-sT_r}.

Example of Derivation Procedure

It is instructive to follow through the derivation for the simple case of prediction when M = 2 in order to observe the effects of imposing the specification of stationarity

$$\epsilon_k = a_{k+s} - (A_o b_k + A_1 b_{k-1} + A_2 b_{k-2})$$

$$I_{1c} = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{k=-N}^{N} \epsilon_k^2$$

$$I_{1c} = \lim_{N \to \infty} \frac{1}{2N+1} \Big\{ \sum a_{k+s}^2 + A_o^2 \sum b_k^2 + A_1^2 \sum b_{k-1}^2 + A_2^2 \sum b_{k-2}^2$$

$$-2A_o \sum a_{k+s} b_k - 2A_1 \sum a_{k+s} b_{k-1} - 2A_2 \sum a_{k+s} b_{k-2}$$

$$+2A_o A_1 \sum b_k b_{k-1} + 2A_o A_2 \sum b_k b_{k-2} + 2A_1 A_2 \sum b_{k-1} b_{k-2} \Big\}$$

If the sequence b_k is stationary, then

$$R_{bb}(0) = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{k=-N}^{N} b_k^2 = \lim_{N \to \infty} \frac{1}{2N+1} \sum_k b_{k-1}^2 = \lim_{N \to \infty} \frac{1}{2N+1} \sum_k b_{k-2}^2$$

$$R_{bb}(1) = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{k=-N}^{N} b_k b_{k-1} = \lim_{N \to \infty} \frac{1}{2N+1} \sum_k b_{k-1} b_{k-2} \text{ .}$$

Introducing the correlation functions, we have

$$I_{1c} = R_{aa}(0) + (A_o^2 + A_1^2 + A_2^2)\, R_{bb}(0) + (2A_oA_1 + 2A_1A_2)\, R_{bb}(1)$$

$$+ 2A_oA_2R_{bb}(2) - 2A_oR_{ba}(s) - 2A_1R_{ba}(1+s) - 2A_2R_{ba}(2+s)$$

Imposing the conditions for the minimization of I_{1c} we obtain the system of equations

$$\frac{\partial I_{1c}}{\partial A_o} = 0$$

$$\frac{\partial I_{1c}}{\partial A_1} = 0$$

$$\frac{\partial I_{1c}}{\partial A_2} = 0$$

$$A_o R_{bb}(0) + A_1 R_{bb}(1) + A_2 R_{bb}(2) = R_{ba}(s)$$

$$A_o R_{bb}(1) + A_1 R_{bb}(0) + A_2 R_{bb}(1) = R_{ba}(1+s)$$

$$A_o R_{bb}(2) + A_1 R_{bb}(1) + A_2 R_{bb}(0) = R_{ba}(2+s)$$

If stationarity of the input can not be assumed, then one can not define meaningful correlation functions on a time-averaging basis. We now consider the approach to be followed when the input is nonstationary.

2.02 Synthesis Procedure for Linear Time-Varying Program

An optimum filter for the case of the non-stationary ensemble can be specified in a manner quite similar to that of the stationary, ergodic ensemble. If the statistical characteristics of an ensemble are time-varying, then a filter whose performance is optimized on the basis of these characteristics will, in general, be time-varying. Thus the elements of the optimum filter will be a function of the sampling instant, j, at which the processing of the input data is to occur.

Let $a_{r,k}$ = desired output datum for the rth member of the ensemble at the kth sampling interval.

$b_{r,k}$ = raw input datum for the same member at the kth sampling interval.

$A_{n,k}$ = coefficient by which $b_{r,k-n}$ is to be weighted at the kth sampling interval. Note that the same $A_{n,k}$ applies to each of the r members of the ensemble. The linear combination of the weighted input data then forms the actual output of the filter for the rth member at the kth instant.

$$c_{r,k} = \text{actual output for } \underline{r}\text{th member at } \underline{k}\text{th instant.}$$

$$= \sum_{n=0}^{M} A_{n,k}\, b_{r,k-n} \qquad (2.02\text{-}1)$$

We now define an error quantity

$$\epsilon_{r,k} = a_{r,k} - c_{r,k}$$
$$= a_{r,k} - \sum_{n=0}^{M} A_{n,k}\, b_{r,k-n} \qquad (2.02\text{-}2)$$

and seek to minimize the mean square of this error, $I_{1v}(k)$, over the ensemble of signals.

$$I_{1v}(k) = \lim_{N \to \infty} \sum_{r=-N}^{N} p(r)\, \epsilon^2_{r,k} \qquad (2.02\text{-}3)$$

where $p(r)$ = probability of the $\underline{r}$th member.

Substituting the expression for the error quantity into equation 2.02-3 and expanding, we obtain

$$I_{1v}(k) = \lim_{N \to \infty} \sum_{r=-N}^{N} p(r) \left[a_{r,k} - \sum_{n=0}^{M} A_{n,k}\, b_{r,k-n} \right]^2$$

$$= \lim_{N \to \infty} \sum_{r=-N}^{N} p(r) a^2_{r,k}$$

$$- 2 \sum_{n} A_{n,k} \lim_{N \to \infty} \sum_{r=-N}^{N} p(r) a_{r,k}\, b_{r,k-n} \qquad (2.02\text{-}4)$$

$$+ \sum_{n} \sum_{m} A_{n,k}\, A_{m,k} \lim_{N \to \infty} \sum_{r=-N}^{N} p(r) b_{r,k-n} b_{r,k-m}$$

Since the signal ensemble is not ergodic, we can no longer justify the equating of time-averages with ensemble-averages. We therefore introduce the concept of the ensemble-averaged linear correlation functions. The autocorrelation is defined as

$$R_{aa}(k,k-n) = \lim_{N \to \infty} \sum_{r=-N}^{N} p(r) a_{r,k} a_{r,k-n}$$

and the cross-correlation as

$$R_{ba}(k,k-n) = \lim_{N \to \infty} \sum_{r=-N}^{N} p(r) a_{r,k} b_{r,k-n}$$

In the absence of specific information as to the distribution of the members of the ensemble, one might assume that the occurrence of all members is equally likely. Physical consideration may frequently justify this assumption. On this basis the mean square error for the linear time-varying filter becomes

$$I_{1v}(k) = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{r=-N}^{N} \epsilon_{r,k}^2 \qquad (2.02\text{-}3a)$$

into which we insert, after expanding, the corresponding correlation functions

$$R_{aa}(k,k-n) = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{r=-N}^{N} a_{r,k} a_{r,k-n}$$

and

$$R_{ba}(k,k-n) = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{r=-N}^{N} a_{r,k} b_{r,k-n} \quad .$$

The resulting equation, irrespective of the probability distribution p(r), is

$$I_{1v}(k) = R_{aa}(k,k) - 2 \sum_{n} A_{n,k} R_{ba}(k,k-n) + \sum_{n} \sum_{m} A_{n,k} A_{m,k} R_{bb}(k-n,k-m) \qquad (2.02\text{-}5)$$

Now minimizing with respect to the $A_{n,k}$ coefficients

$$\frac{\partial I_{1v}(k)}{\partial A_{j,k}} = 0 \qquad \text{for } j = 0, 1, \ldots M$$

we obtain the synthesis equations for the linear time-varying filter

$$\sum_{n=0}^{M} A_{n,k} R_{bb}(k-n,k-j) = R_{ba}(k,k-j) \qquad (2.02\text{-}6)$$

$$\text{for } j = 0, 1, \ldots M.$$

For a predicting filter, the synthesis equations are

$$\sum_{n=0}^{M} A_{n,k} R_{bb}(k-n,k-j) = R_{ba}(k+s, k-j) \qquad (2.02\text{-}7)$$

$$\text{for } j = 0, 1, \ldots M.$$

and $\underline{s}$ is as previously defined

It should be noted that the above-described procedures yield a specification for an optimum filter at a particular instant -- the $\underline{k}$th sampling instant. The approach which one takes in specifying a time-varying discrete filter depends on how rapidly the statistical characteristics of the signal ensemble are varying compared to the time constants of the controlled system.

If the variation is slow compared to these time constants, then one can solve the filtering problem on a quasi-static basis. The solution involves a filtering system consisting of a set of optimum filters (each operating over a certain time interval) and a device for switching from one filter to another. The switching device might be an electromechanical relay network, or a conditional subprogram instruction to the computer (such as the cp(-)x instruction in the Whirlwind I code). The switch might be actuated by information as to elapsed time and/or as to the quality of performance of the output of the filtering system. Such a system, involving P separate discrete filters, is shown in Figure 2.02-1. There it is assumed that each filter operates well over a range of $\pm$ H sampling intervals about some sampling instant (at which it is optimum). The number of filters, P, would obviously depend on what tolerance in quality of performance were permitted, and on the price one were willing to pay in storage facilities, program complexity, and computation time.

For the ensemble whose statistical structure varies too rapidly for effective filtering on a quasi-static basis, we need merely extend this approach to the limit and provide a different filter at each sampling instant. Rather than evaluate and store the large number of sets of weights directly, we might, as before, synthesize P separate discrete filters. After plotting each of the $A_{n,k}$ weights as a function of the time variable k, we approximate numerically by a smooth curve (see Figure 2.02-2) each of the M+1 discrete sequences. The desired set of weights at any sampling instant is then obtained by interpolation. The storage requirements are reduced since we now store only the coefficients specifying the functional approximation of any $A_{i,k}$ as a function of k. Thus, if Lagrangian interpolation were used, the set of M+1 polynomial approximations would yield the optimum filter at each of the original P sampling instants, and an approximation to the optimum filter at any other instant. In this case,

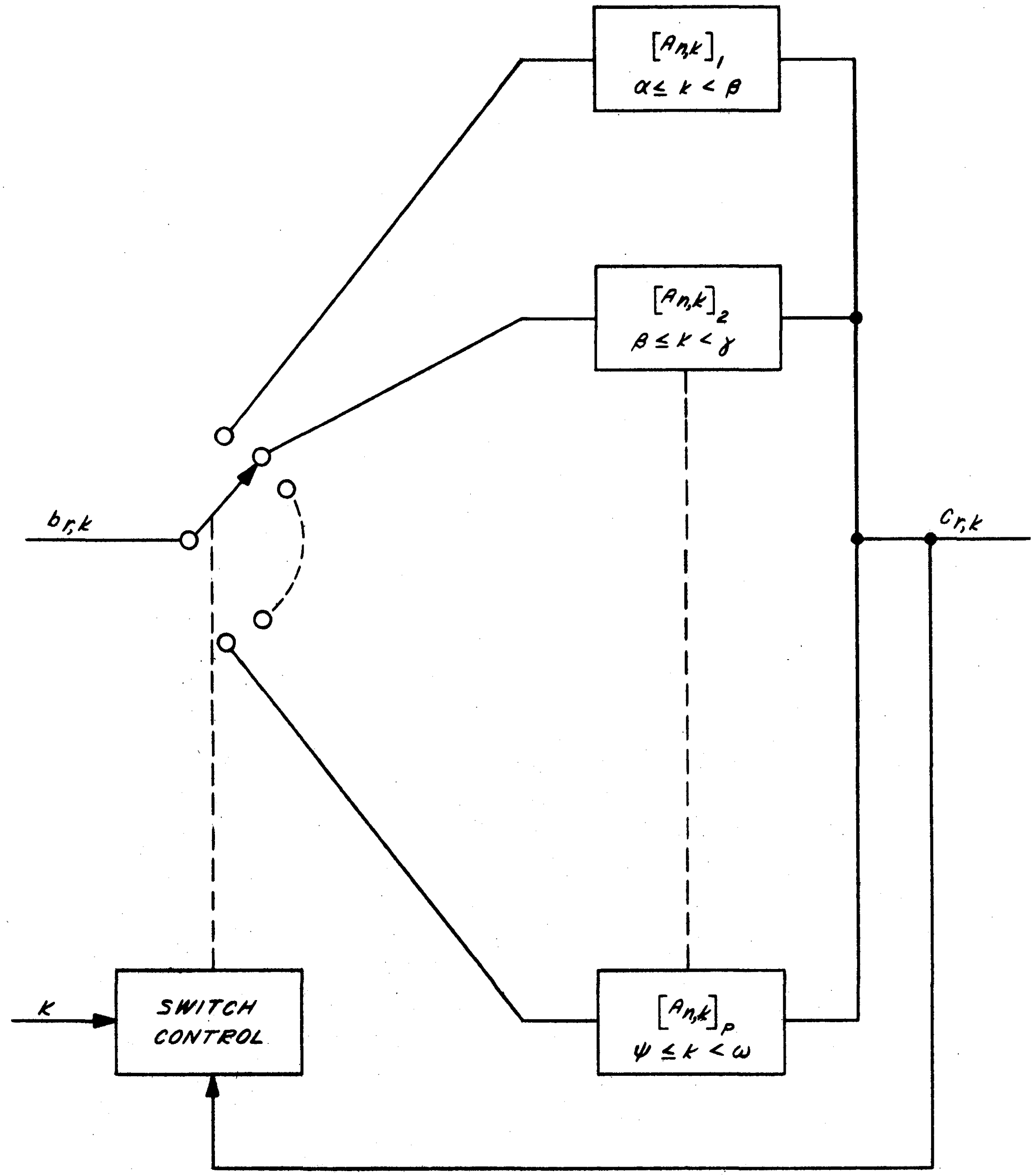


FIG. 2.02-1

FILTERING SYSTEM FOR A SIGNAL ENSEMBLE
WHOSE STATISTICS VARY SLOWLY COMPARED TO
THE TIME CONSTANTS OF THE CONTROLLED SYSTEM

A-51954

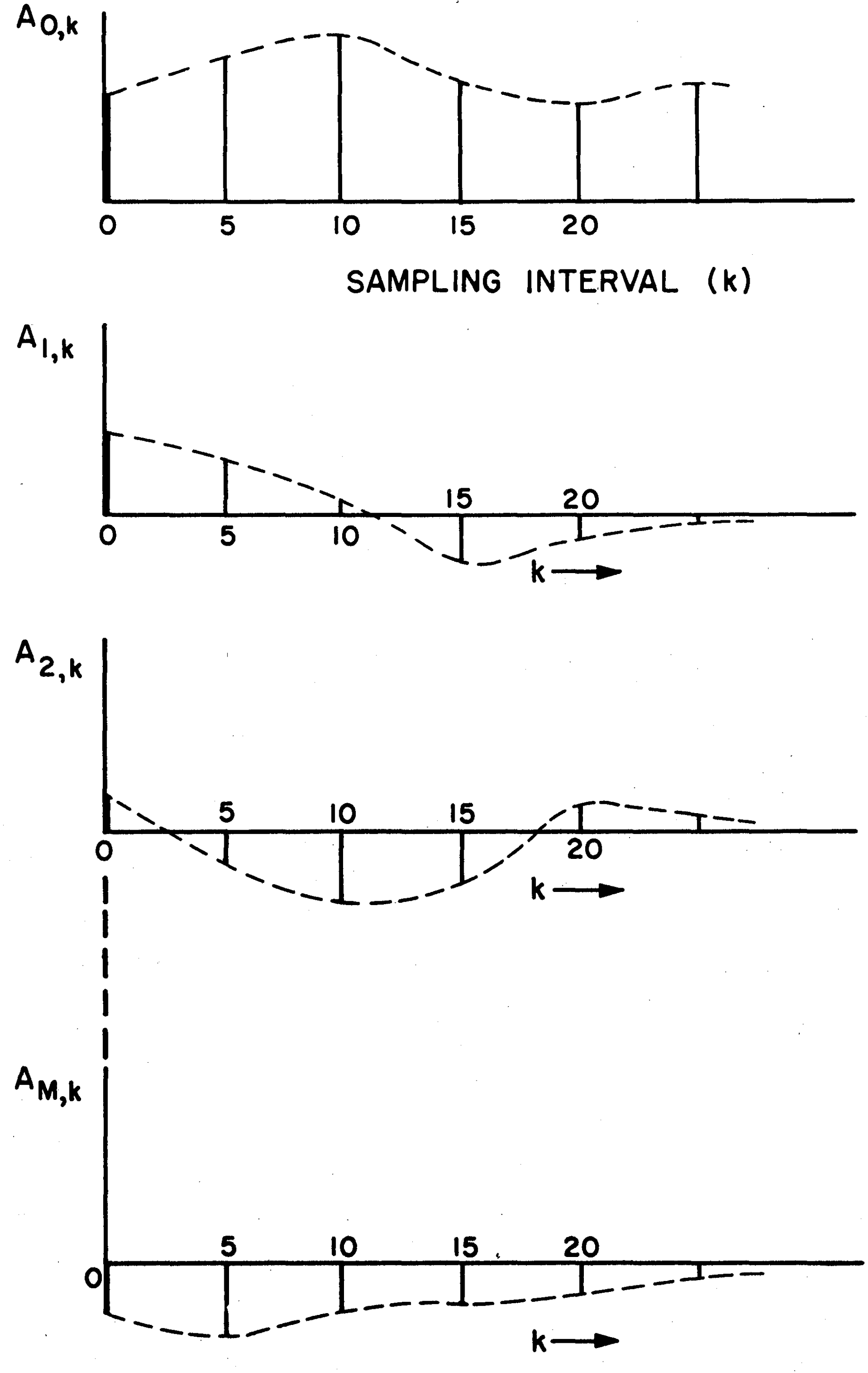


FIG. 2.02-2

APPROXIMATION OF THE WEIGHTING COEFFICIENTS AS FUNCTIONS OF THE TIME VARIABLE k

A-51946

$$A_{0,k} = \alpha_{00} + \alpha_{01} k + \alpha_{02} k^2 + \dots \alpha_{0f} k^f$$

$$A_{1,k} = \alpha_{10} + \alpha_{11} k + \alpha_{12} k^2 + \dots \alpha_{1g} k^g$$

$$\vdots \qquad \vdots$$

$$A_{M,k} = \alpha_{M0} + \alpha_{M1} k + \alpha_{M2} k^2 + \dots \alpha_{Mh} k^h$$

where the degrees of the polynomial (f, g, ... h) need not necessarily be the same. Since the device controlling the position of the switch would under these circumstances be time-actuated, we must necessarily have insight into the manner in which the statistical structure of the ensemble is varying in time.

2.03 Synthesis Procedure for Non-Linear Time-Invariant Programs

We consider now that class of discrete filters for which there is a non-linear, time-invariant relationship between the output data and a finite number of input data. In particular, we treat the simple case defined by

$$C_k = \sum_{n=0}^{M} A_n b_{k-n} + \sum_{p=0}^{Q} B_p b^2_{k-p} \qquad (2.03\text{-}1)$$

Accordingly, we define an error quantity

$$\epsilon_k = a_k - \sum_{n=0}^{M} A_n b_{k-n} - \sum_{p=0}^{Q} B_p b^2_{k-p} \qquad (2.03\text{-}2)$$

and proceed to derive sequences A_n and B_p such that

$$I_{nc} = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{k=-N}^{N} \epsilon_k^2$$

= mean square error for a non-linear constant-coefficient filter

is a minimum.

$$\begin{aligned} I_{nc} = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{k} a_k^2 &+ \sum_{n=o}^{M} \sum_{m=o}^{M} A_n A_m \lim_{N \to \infty} \frac{1}{2N+1} \sum_{k} b_{k-n} b_{k-m} \\ &+ \sum_{p=o}^{Q} \sum_{q=o}^{Q} B_p B_q \lim_{N \to \infty} \frac{1}{2N+1} \sum_{k} b_{k-p}^2 \, b_{k-q}^2 \\ &- 2 \sum_{n=o}^{M} A_n \lim_{N \to \infty} \frac{1}{2N+1} \sum_{k} a_k \, b_{k-n} \\ &- 2 \sum_{p=o}^{Q} B_p \lim_{N \to \infty} \frac{1}{2N+1} \sum_{k} a_k \, b_{k-p}^2 \\ &+ 2 \sum_{n=o}^{M} \sum_{p=o}^{Q} A_n B_p \lim_{N \to \infty} \frac{1}{2N+1} \sum_{k} b_{k-n} \, b_{k-p}^2 \end{aligned} \qquad (2.03\text{-}3)$$

If stationarity can be assumed, the concept of time-averaged correlation functions may be introduced again and extended further. In addition to the previously defined linear functions, we shall have occasion to utilize the higher order non-linear functions.

$$R_{bbb}(k_1;k_2) = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{h=-N}^{N} b_h \, b_{h-k_1} \, b_{h-k_2}$$

$$R_{bba}(k_1;k_2) = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{h=-N}^{N} a_b\, b_{h-k_1}\, b_{h-k_2}$$

$$R_{bbbb}(k_1;k_2;k_3) = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{h=-N}^{N} b_h b_{h-k_1} b_{h-k_2} b_{h-k_3}$$

Because we shall always be dealing with powers of the input sequence, some of the shifts, k_i, will be equal, and hence the complete generality of the above defined functions is unnecessary. These higher order functions may under such circumstances be redefined in terms of linear correlations between the powers of the input. For the case treated here, considerable simplification is obtained if we let

$$g_{k-n} = b_{k-n}^2$$

Then

$$\begin{aligned} I_{nc} = R_{aa}(0) &+ \sum_n \sum_m A_n A_m\, R_{bb}(m-n) \\ &+ \sum_p \sum_q B_p B_q\, R_{gg}(q-p) \\ &- 2\sum_n A_n R_{ba}(n) - 2\sum_p B_p\, R_{ga}(p) \\ &+ 2\sum_n \sum_p A_n B_p\, R_{bg}(n-p) \end{aligned} \qquad (2.03\text{-}4)$$

Minimization of I_{nc} requires that

$$\frac{\partial I_{nc}}{\partial A_h} = 0 \qquad h = 0, 1, \ldots M$$

$$\frac{\partial I_{nc}}{\partial B_j} = 0 \qquad j = 0, 1, \ldots Q$$

From these relations we obtain the Wiener-Hopf equations for the non-linear constant-coefficient filter:

$$\sum_{n=o}^{M} A_n R_{bb}(h-n) + \sum_{p=o}^{Q} B_p R_{bg}(h-p) = R_{ba}(h)$$

$$\sum_{n=o}^{M} A_n R_{bg}(n-j) + \sum_{p=o}^{Q} B_p R_{gg}(p-j) = R_{ga}(j) \qquad (2.03\text{-}5)$$

$$h = 0, 1, \ldots M.$$

$$j = 0, 1, \ldots Q.$$

For a predicting filter, the Wiener-Hopf equations are of the form

$$\sum_{n=o}^{M} A_n R_{bb}(h-n) + \sum_{p=o}^{Q} B_p R_{bg}(h-p) = R_{ba}(h+s)$$

$$\sum_{n=o}^{M} A_n R_{bg}(n-j) + \sum_{p=o}^{Q} B_p R_{gg}(p-j) = R_{ga}(j+s) \qquad (2.03\text{-}6)$$

$$h = 0, 1, \ldots M.$$

$$j = 0, 1, \ldots Q.$$

The minimum value for I_{nc} for prediction is then

$$\left[I_{nc}\right]_{min} = R_{aa}(0) - \sum_{n=o}^{M} A_n R_{ba}(n+s) - \sum_{p=o}^{Q} B_p R_{ga}(p+s)$$

(2.03-7)

Because of the non-linearity of the filters, conventional frequency domain analysis is not applicable. Note that the extension of the above derivation procedure to any order of non-linearity is rather easily made.

Illustrative Example 2.03-1

As an illustration of a case in which a non-linear weighting sequence is superior to one which is linear, we consider the problem involved in predicting flux linkages, $\lambda(t, i)$, in an electrical circuit composed of an ideal capacitance C and an iron core inductance L. The prediction of future values of flux linkages is predicated upon a knowledge of past values of $\lambda(t, i)$ and of the physical mechanism governing the variation of $\lambda(t, i)$. Because of the fact that the core is subject to magnetic saturation, the flux linkages depend on the amplitude of the current i as well as on time t. The condition of equilibrium of electromotive forces in the circuit (Kirchoff's voltage law) gives

$$\frac{d\lambda}{dt} + \frac{1}{C}\int_0^t i\,dt = 0$$

For the linearized version (an approximation of zero order) of the problem, we have the familiar relation

$$\frac{d\lambda}{dt} = L\frac{di}{dt}$$

As a first approximation however, we can assume that the condition of saturation can be expressed by the equation

$$i = A\lambda + B\lambda^3$$

Substituting this expression in the original equation and differentiating, we have

$$\frac{d^2\lambda}{dt^2} + \frac{A\lambda + B\lambda^3}{C} = 0$$

We now approximate this equation describing a non-linear conservative system by means of a difference equation.

$$\frac{\lambda_{n+1} - 2\lambda_n + \lambda_{n-1}}{h^2} + \frac{A}{C}\lambda_n + \frac{B}{C}\lambda_n^3 = 0$$

or

$$\lambda_{n+1} = C_o\lambda_n + C_1\lambda_{n-1} + D_o\lambda_n^3 \qquad (2.03\text{-}9)$$

where

$$C_o = (2 - \frac{Ah^2}{C})$$

$$C_1 = -1$$

$$D_o = -\frac{Bh^2}{C}$$

Information as to the values of A and B, and hence as to C_o and D_o, is contained in the normal magnetization curve for the inductor.

Let us assume that our problem is the following:

A sequence of data relating to the time-variation of flux linkages in the inductor is to actuate a control system. These data, having been experimentally observed by instruments capable of measuring only to within ± 0.01 flux linkages, are thereby contaminated by quantization noise. Not only shall our programs operate so as to reduce the effects of noise, but, in addition, we shall require that they are to improve the over-all system response by acting as lead or predicting networks. We shall further assume that, because of limited storage facilities, we can allocate only two registers to retain the weighting coefficients of a program.

Our problem, in essence, is to devise an optimum two-element predicting filter for the processing of quantized λ_k data.

Physical insight into the inductor-capacitor network has provided us with a mechanism for generating future values of the λ_k sequence from past data -- equation 2.03-9. By utilizing the knowledge of the mechanism for the formation of the λ_k sequence, the designer should be able to predict more intelligently future values of this sequence. In the absence of quantization noise, one would logically use the recursion equation 2.03-9 in his prediction. The introduction of noise, however, produces a sequence of perturbed data, b_k, for which the recursion equation is no longer valid. To predict future values of λ_k from present and past values of the perturbed sequence, one must introduce a smoothing mechanism (e.g., the least squares procedure). If, in specifying the characteristics of the filter, the designer were to constrain the system so that it performs only linear operations, he would not use all of the available knowledge, and the performance of the filter would consequently be poorer.

To test the validity of this reasoning, we shall determine, by means of classical least squares, a two-element predictor of each type -- linear and non-linear.

Suppose, for a given test signal, we use two different fluxmeters as measuring devices. One, capable of indicating ± 0.0001 flux linkages is primarily suited for laboratory work; the other, capable of indicating ± 0.01 flux linkages is rugged enough for control purposes. The former supplies a λ_k sequence; the latter, a b_k sequence. For this test signal, these sequences happen to be the following:

k	λ_k	b_k
0	0.0000	0.00
1	0.8000	0.80
2	0.8512	0.85
3	1.0729	1.07
4	1.3666	1.37
5	1.8364	1.84
6	2.7290	2.73
7	5.1288	5.13
8	19.1660	19.17

Linear Predictor

We define

$$\epsilon_k = \lambda_{k+s} - (A_o b_k + A_1 b_{k-1})$$

and seek to minimize

$$S_1 = \sum_{k=o}^{8} (\lambda_{k+s} - A_o b_k - A_1 b_{k-1})^2$$

The minimization procedure leads to the system of equations

$$A_o \sum_k b_k^2 + A_1 \sum_k b_k b_{k-1} = \sum_k \lambda_{k+s} b_k$$

$$A_o \sum_k b_k b_{k-1} + A_1 \sum_k b_{k-1}^2 = \sum_k \lambda_{k+s} b_{k-1}$$

For s = 1, we have

$$(409.0286)\ A_o + (122.9464)\ A_1 = 122.9156$$

$$(122.9464)\ A_o + (409.0286)\ A_1 = 69.4838$$

From which

$$A_o = 0.27422$$

$$A_1 = 0.08745$$

<u>Non-Linear Predictor</u>

We define

$$\epsilon_k = \lambda_{k+s} - (A_o b_k + B_o b_k^3)$$

and seek to minimize

$$S_{nl} = \sum_{k=o}^{8} (\lambda_{k+s} - A_o b_k - B_o b_k^3)^2$$

The minimization procedure leads to the system of equations

$$A_o \sum_k b_k^2 + B_o \sum_k b_k^4 = \sum_k \lambda_{k+s} b_k$$

$$A_o \sum_k b_k^4 + B_o \sum_k b_k^6 = \sum_k \lambda_{k+s} b_k^3$$

For s = 1, we have

$$(409.0286)\ A_o + (135\ 813.4437)\ B_o = 122.9156$$

$$(135\ 813.4437)\ A_o + (49\ 647\ 362.50)\ B_o = 2716.3631$$

From which

$$A_o = 3.07939$$

$$B_o = -0.00837$$

We now test these filters by attempting to predict λ_{k+1} when given b_k and b_{k-1} for the linear filter or b_k for the non-linear filter.

We define

λ_{kp} = predicted value for λ_k based on quantized data

$$\epsilon_k = \lambda_k - \lambda_{kp}$$

The results are summarized below:

Table 2.03-1 - Comparison of Performance of a Linear and a Non-Linear Predictor

λ_k	Linear Predictor		Non-Linear Predictor	
	λ_{kp}	ϵ_k	λ_{kp}	ϵ_k
λ_2	0.2194	0.6318	2.4592	-1.6080
λ_3	0.3030	0.7699	2.6123	-1.5394
λ_4	0.3677	0.9989	3.2847	-1.9181
λ_5	0.4693	1.3671	4.1972	-2.3608
λ_6	0.6244	2.1046	5.6139	-2.8849
λ_7	0.9095	4.2193	8.2364	-3.1076
λ_8	1.6455	17.5205	14.6673	4.4987

It is not surprising to find the performance of the non-linear predictor to be superior to that of the other since the data sequence is derived from a non-linear generating mechanism. This superiority is made evident by a comparison of the sums of the squares of the errors in prediction for the two filters. Thus, for the linear predictor

$$S_l = 333.0584$$

and for the non-linear predictor

$$S_{nl} = 52.4260$$

In spite of the fact that the non-linear predictor used a smaller "memory" (one datum), its performance was definitely superior. Clearly, the availability of a priori information as to the nature of the system enabled us to design a better predictor.

In order to visualize more clearly the mechanism of predicting a perturbed sequence, one might resolve it into the mechanism of predicting in the absence of noise and that of smoothing a perturbed sequence. It is not our purpose to imply that these are not interrelated processes, but rather to suggest this artificial separation as an aid to the imagination. It is then seen that the former mechanism should approximate that of the recursion equation as closely as possible, while the latter should be as effective as possible in removing the undesirable perturbations. This reasoning provides us with an insight into the manner in which the particular type of non-linearity should be chosen. We may conclude that, when the mechanism for generating the true data sequence is known or suspected to be of a particular type of non-linearity, one should design the appropriate non-linear filter.

2.04 Synthesis Procedure for Non-linear Time-Varying Programs

The last procedure to be discussed is that in which the filters process the non-stationarity input data in such manner as to establish a non-linear, time-varying relationship between the output and a finite number of input data. Since the extension to higher orders of non-linearity is obvious, we again treat the simple case defined by

$$c_{r,k} = \sum_{n=o}^{M} A_{n,k} b_{r,k-n} + \sum_{p=o}^{Q} B_{p,k} b^2_{r,k-p} \qquad (2.04\text{-}1)$$

The error quantity is then

$$\epsilon_{r,k} = a_{r,k} - \sum_{n=o}^{M} A_{n,k} b_{r,k-n} - \sum_{p=o}^{Q} B_{p,k} b^2_{r,k-p} \qquad (2.04\text{-}2)$$

and if we substitute

$$g_{r,k-p} = b^2_{r,k-p}$$

we obtain as the ensemble-averaged mean square error, $I_{nv}(k)$, for the non-linear filter

$$I_{nv}(k) = \lim_{N \to \infty} \sum_{r=-N}^{r=N} P(r) \epsilon^2_{r,k}$$

If the assumption of equal likelihood of occurrence of the member signals is justified, then we may write

$$I_{nv}(k) = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{r=-N}^{N} a^2_{r,k} \qquad (2.04\text{-}3)$$

$$+ \sum_{n=o}^{M} \sum_{m=o}^{M} A_{n,k} A_{m,k} \lim_{N \to \infty} \frac{1}{2N+1} \sum_{r} b_{r,k-n} b_{r,k-m}$$

$$+ \sum_{p=o}^{Q} \sum_{q=o}^{Q} B_{p,k} B_{q,k} \lim_{N \to \infty} \frac{1}{2N+1} \sum_{r} g_{r,k-p} g_{r,k-q}$$

$$- 2 \sum_{n=o}^{M} A_{n,k} \lim_{N \to \infty} \frac{1}{2N+1} \sum_{r} a_{r,k}\, b_{r,k-n} \qquad (2.04\text{-}3)$$

$$- 2 \sum_{p=o}^{Q} B_{p,k} \lim_{N \to \infty} \frac{1}{2N+1} \sum_{r} a_{r,k}\, g_{r,k-p}$$

$$+ 2 \sum_{n=o}^{M} \sum_{p=o}^{Q} A_{n,k} B_{p,k} \lim_{N \to \infty} \frac{1}{2N+1} \sum_{r} g_{r,k-p} b_{r,k-n}$$

Substituting the ensemble-averaged correlations

$$I_{nv}(k) = R_{aa}(k,k) + \sum_{n} \sum_{m} A_{n,k} A_{m,k} R_{bb}(k-n,k-m)$$

$$+ \sum_{p} \sum_{q} B_{p,k} B_{q,k} R_{gg}(k-p,k-q)$$

$$-2 \sum_{n} A_{n,k} R_{ba}(k,k-n) \qquad (2.04\text{-}4)$$

$$-2 \sum_{p} B_{p,k} R_{ga}(k,k-p)$$

$$+2 \sum_{n} \sum_{p} A_{n,k} B_{p,k} R_{bg}(k-p,k-n)$$

Minimizing $I_{nv}(k)$ with respect to $A_{h,k}$ and $B_{j,k}$ we obtain the synthesis equations

$$\sum_{n=o}^{M} A_{n,k}R_{bb}(k-n,k-h) + \sum_{p=o}^{Q} B_{p,k}R_{bg}(k-p,k-h) = R_{ba}(k,k-h)$$

$$\sum_{n=o}^{M} A_{n,k}R_{bg}(k-j,k-n) + \sum_{p=o}^{Q} B_{p,k}R_{gg}(k-j,k-p) = R_{ga}(k,k-j) \qquad (2.04\text{-}5)$$

$$h = 0, 1, \ldots M$$

$$j = 0, 1, \ldots Q$$

For a predicting filter, the equations are

$$\sum_{n=o}^{M} A_{n,k}R_{bb}(k-n,k-h) + \sum_{p=o}^{Q} B_{p,k}R_{bg}(k-p,k-h) = R_{ba}(k+s,k-h)$$

$$\sum_{n=o}^{M} A_{n,k}R_{bg}(k-j,k-n) + \sum_{p=o}^{Q} B_{p,k}R_{gg}(k-j,k-p) = R_{ga}(k+s,k-j) \qquad (2.04\text{-}6)$$

$$h = 0, 1, \ldots M$$

$$j = 0, 1, \ldots Q$$

The minimum value for $I_{nv}(k)$ for prediction is then

$$\left[I_{nv}(k)\right]_{min} = R_{aa}(k,k) - \sum_{n=o}^{M} A_{n,k}R_{ba}(k+s,k-n) - \sum_{p=o}^{Q} B_{p,k}R_{ga}(k+s,k-p) \qquad (2.04\text{-}7)$$

2.1 QUANTIZATION NOISE AND NUMERICAL COMPUTATIONS

Although variables encountered in control systems are frequently continuous, a computer such as Whirlwind I requires incoming data in the form of discrete quantities. The digital control of a continuous variable consequently involves a conversion from continuous data to discrete (i.e., encoding) when raw information is supplied to the computer, and from discrete to continuous (i.e., decoding) when the processed information is used to actuate the control mechanisms. It should be noted that the conversion problem is not peculiar to real-time computer applications, but rather that it is rendered more difficult by the necessity for obtaining, within a very limited time, results which are usable in controlling a dynamic system.

Not only does the original input signal to the encoder contain noise from sources both external and internal to the control system, but the very process of encoding further corrupts the signal so that filtering of the output sequence is necessary if we are to extract the true message or some function thereof. As has been previously indicated, this distortion results primarily from the quantizing process and manifests itself, in the time domain, as a limiting of the the number of digits by which the variable may be represented. When these sharply limited data are used in numerical computations, we frequently find that a previously stable program will yield either a divergent or an oscillating solution.

In the following sections we shall discuss the characteristics of quantization noise and how they affect the ever-present round-off error in numerical procedures.

2.11 Analysis of Noise Caused by the Encoding Process

Since later sections of this report deal with the design of linear programs capable of smoothing and predicting future values of an encoded sequence, it is appropriate that we now examine the encoding process in some detail. Encoding may be defined as the process whereby a continuous time series is converted to a discrete time series, the elements of which can assume only a finite set of discrete amplitudes. This conversion involves the separate, but commutative, operations of sampling and quantizing.

Since the signal distortion caused by sampling can frequently be made negligibly small by proper design of the encoder, we shall make only a few brief remarks about this operation. Referring to Figure 2.11-1, we see that the sampling device can be represented by a switch rotating at a constant angular frequency ω_r, followed by a holding circuit and an amplifier. For a continuous signal input f(t), the sampler provides at its output a sequence of pulses, f_k. All pulses are of equal duration T, and successive pulses T_r seconds apart. Insofar as information content is concerned, there is no distortion provided only that the sampling frequency is at least twice the highest signal frequency.

Linvill[21] has shown that the sampling operation may be visualized as the process of modulating a continuous signal by an infinite train of unit impulses. The resulting sampled signal has a spectrum which contains the original frequency components as well as all harmonics of these components. So long as the condition on the sampling rate is met, there will be no overlap of the spectra and hence no distortion.

Within a finite range of amplitude variation, a continuous signal, as well as its samples, can assume an infinite number of amplitude

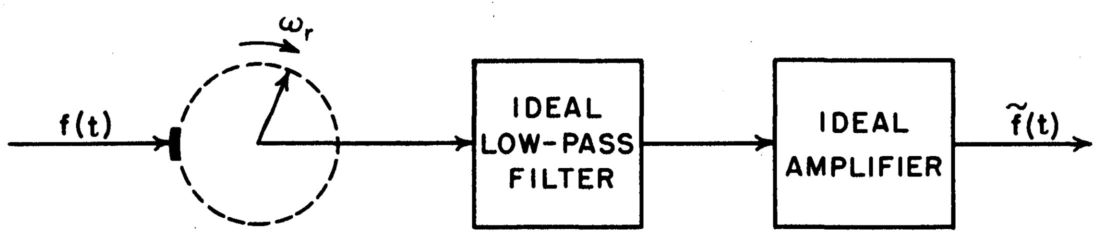


FIGURE 2.11-1a - SAMPLING DEVICE

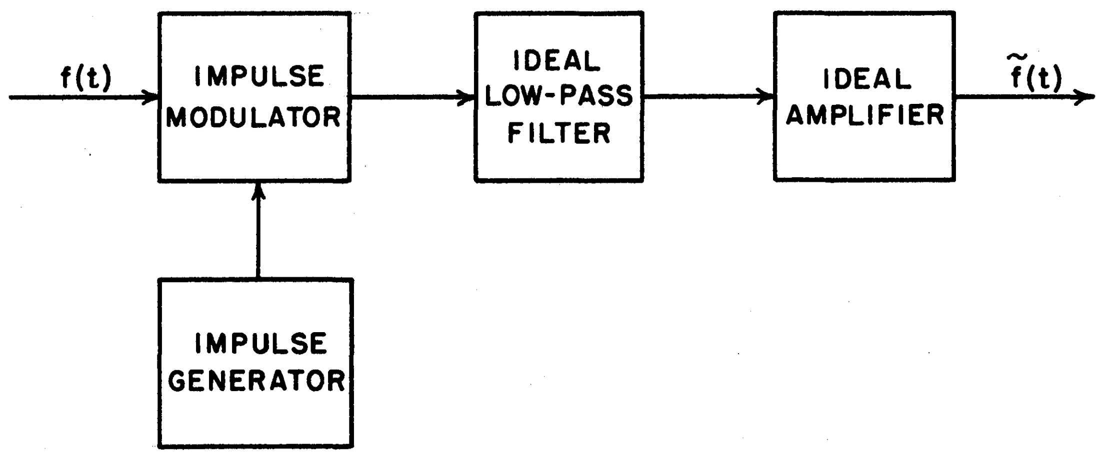


FIGURE 2.11-1b - EQUIVALENT SAMPLING DEVICE

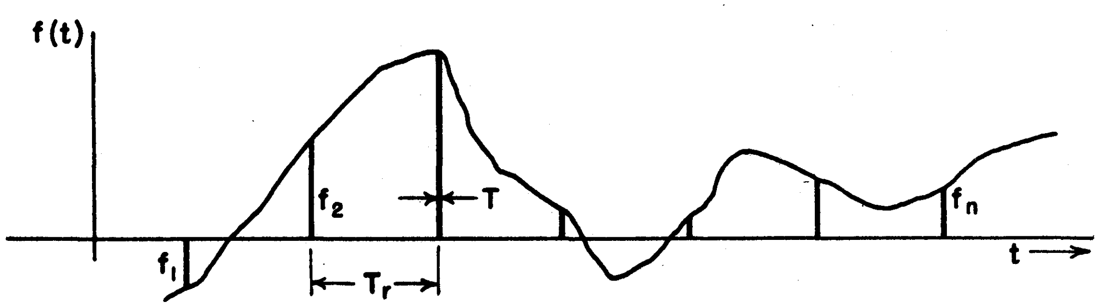


FIGURE 2.11-1c - SAMPLED SIGNAL

A-50255

levels. However, it may be neither possible nor necessary to transmit the exact amplitudes of these samples because of various limitations imposed by the transmission device or by the ultimate receptor. In such cases it is permissible to represent and to transmit all levels within a certain amplitude range by one discrete amplitude level. This means that the original signal is to be replaced by a wave constructed of quantized values selected on a minimum error basis from the discrete set available. Clearly if one assigns the quantum values with sufficiently close spacing one can make the quantized wave indistinguishable from the original signal.

The quantizing process may be visualized as being the result of operating on the signal with a "staircase transducer", a device having the instantaneous output-input characteristic shown in Figure 2.11-2. When a smoothly varying signal is the input, the output remains constant while the input varies within the boundaries of a tread, and changes abruptly by one full step when the signal crosses the boundary. A quantized signal wave and the corresponding error wave are shown in Figure 2.11-3.

The quantization error is, then, the inherent amount of distortion resulting from the fact that the output of the encoder is limited to a finite set of amplitude levels while its input occupies the same amplitude range in a continuous manner. The maximum instantaneous value of the error is half of one step and the total range of variation is from minus half a step to plus half a step. Only if the input is known as an exact function of time can one find an explicit relationship between it and the corresponding error. Otherwise, one must resort to a statistical description of the error q(t) since all that is known, in general, is that

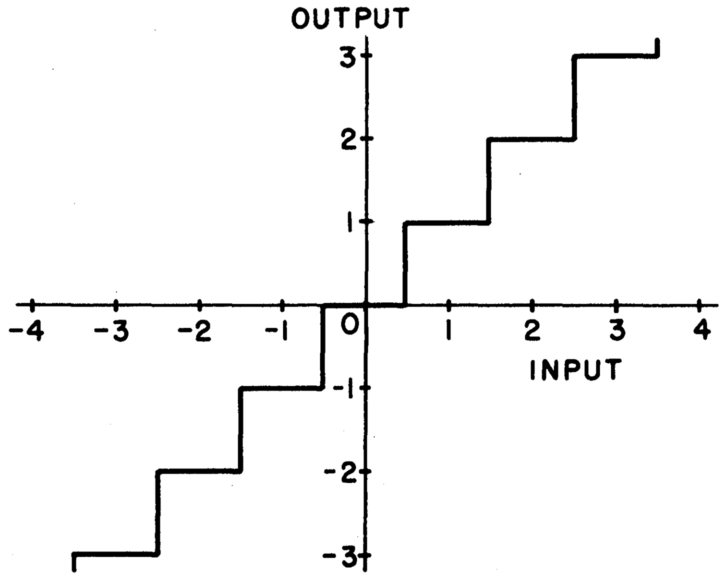


FIGURE 2.11-2-QUANTIZING CHARACTERISTIC

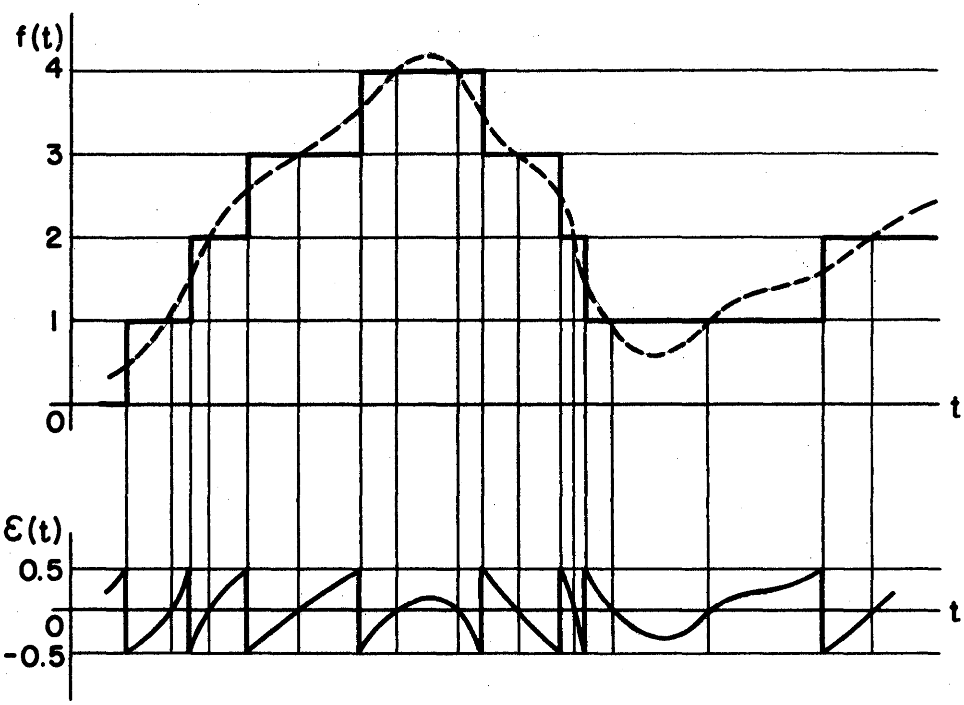


FIGURE 2.11-3-A QUANTIZED SIGNAL AND THE CORRESPONDING ERROR

A-50256

it is a function of amplitude of the quanta, α_i, and the signal function itself.

$$q(t) = F\left\{\alpha_i, f(t)\right\} .$$

One can be somewhat more explicit when the quanta are equal and write

$$q(t) = f(t) - \alpha k(t)$$

where k(t) is a discrete variable taking on a range of integral values in such a manner as to render $|q(t)|$ a minimum.

$$k(t) = \begin{cases} -n & p(-n) \\ -n+1 & p(-n+1) \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ -1 & p(-1) \\ 0 & p(0) \\ +1 & p(+1) \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ n-1 & p(n-1) \\ n & p(n) \end{cases}$$

and p(h) = probability that f(t) lies in $\left\{\alpha(h - 1/2), \alpha(h + 1/2)\right\}$. In the above formulation it is assumed that the signal is bounded in amplitude by $(-n\alpha, n\alpha)$ so that there are 2n+1 quantizing levels. Unless the probability density distributions of f(t) and k(t) are known or certain simplifying assumptions are made, one can deduce little of value from the foregoing analysis.

If one assumes, as did Mayer,[20] that all errors are equally likely in the range (-1/2, +1/2), then considerable simplification results. Let the quantizer have s discrete amplitude levels and accordingly s steps α_i, which need not be equal, and p(i) be the probability of the ith level. Then one can show that the total quantization noise power will be

$$N_q = \frac{1}{12} \sum_i p(i) \alpha_i^2$$

With equal steps α, one obtains

$$N_q = \frac{\alpha^2}{12}$$

There is, of course, no reason why the quantization need be done on the basis of uniform spacing of the levels. Panter and Dite[22] have shown that by taking the statistical properties of the signal into consideration, the distortion introduced in a PCM system due to quantization can be minimized by a proper level distribution which is a function of the amplitude density distribution of the signal. Non-uniform quantization may be accomplished by first compressing the signal, then uniformly quantizing the modified signal. One of the more common forms of compression is the logarithmic one, where the levels are crowded near the origin and spaced farther apart near the peaks. Panter and Dite have also shown that, with logarithmic compression, the distortion is largely independent of the statistical properties of the signal.

Besides studying quantization error from the statistical point of view, one can also investigate the power spectra of quantized signals. Such an investigation for both uniform and non-uniform quantization was made by Bennett.[19] The signal used was one having its energy uniformly distributed throughout a definite frequency band and with the phases of the components randomly distributed. Anticipating binary coding, Bennett determined the power spectra for this signal quantized to several different numbers of binary digits. As might be expected, not all of the distortion fell within the signal band. The spectra of distortion resulting from the uniform quantization of a random noise signal showed that

(a) the fewer the number of binary digits to which the signal was quantized, the greater was the noise power (a corroboration of Mayer's and of Bennett's analyses),

(b) the fewer the number of digits, the richer was the spectrum in low-frequency components,

(c) the greater the number of digits, the flatter was the spectrum over a wider range, but with a smaller maximum density.

By increasing the number of digits (or quantizing levels) indefinitely, one obtains the flat spectrum of "white" noise -- a spectrum which is that of the continuous input signal.

For the case of non-uniform quantization Bennett found that the error spectrum out of the linear quantizer is virtually the same whether or not the signal input is compressed. The advantage of non-uniform quantization appears to lie in the fact that finer divisions are available for weak signals. For a given number of total steps this means that coarser quantization applies near the peaks of large signals, but the larger absolute errors are tolerable here because thay are small relative to the larger signal values.

Having discussed some of the characteristics of quantization noise, we now consider its effect on control systems. It becomes immediately apparent that the presence of such noise may seriously affect the steady-state performance of those systems having large time constants (i.e., systems characterized either by a good response over a narrow range about zero frequency or by a slowly decaying impulse response). This stems from the fact that a major portion of the noise power is concentrated at those low frequencies for which slow systems have a good response. Since many of the systems in which the digital computer will exercise a control function are characterized by large time constants, we must devise techniques for coping with the problem of quantization noise when it is an important noise component. One such technique, dealt with herein, involves the use of filters (which obviously need not be statistical) which are designed with specific reference to the characteristics of this type of noise.

2.12 Effect of Computational Errors in Discrete Filters

In Section 2.0 it was indicated that weighting sequences can be derived from ordinary differential equations by approximating the derivatives by their corresponding divided differences. When this is done, one obtains recursion formulae by means of which one can approximate the solutions of specific differential equations by successive extrapolations. It is apparent, however, that in neglecting the higher divided differences in equations so derived one has committed an error. Each extrapolation will therefore entail this truncation error which, if uncorrected, will tend to accumulate with successvie extrapolations until eventually the results of the computations are rendered useless.

The truncation error is not the only source of uncertainty to which numerical procedures are subject. The necessity of rounding off each extrapolant evaluated with the aid of machines of limited register length will provide another source of accidental error which tends to impair the accuracy of the solution. In any numerical work using approximate formulae and numbers subject to round-off, both errors will co-exist independently. Since their presence cannot be avoided, one generally chooses his formulae and interval h in such a way that, together with data to a sufficient number of digits, the ultimate solution is obtained to the preassigned degree of accuracy.

If one makes certain assumptions as to the manner in which each of these errors propagate one finds that, for N successive extrapolations, the total round-off error will grow more slowly with increasing N than does the truncation error. By decreasing h (sampling more frequently), one reduces the truncation error. However, this necessitates more extrapolations and hence a greater round-off error. One therefore programs his work in such a way that the two errors are equal at the end of the computations.

Although the least-squares sequences obtainable through Wiener-Lee synthesis are not related to any specific differential equations, there is, nevertheless, a truncation error whenever we let M assume a finite value. Levinson has shown that a sequence based on M+1 weights always does a better job of filtering (in the specified sense) than does a sequence based on M weights. However, one rapidly reaches a point of diminishing returns in that the improvement in filtering resulting from the additional weights does not warrant the labor of computing the weights. Furthermore, the round-off error increases with increasing M so that, in general, a short weighting sequence is desirable.

CHAPTER III

EXPERIMENTAL ANALYSIS

3.01 Prediction in the Presence of Quantization Noise of Functions Related to Straight-Line Flight

To lend substance to this investigation, we propose to use as an input signal that generated by an aircraft as it flies on a prescribed path across a polar grid. The resulting encoded sequences of the polar variables, $r(t_k)$ and $\theta(t_k)$, are to be processed by the computer to give the future position of the aircraft. Each of the variables will be treated as a simple time series, although it is possible to treat them together as multiple time series.

Such sequences might well arise in an air traffic control system where the incoming aircraft follow a definite time and space pattern in their approach to the airstrip. Since the use of digital computers in such systems is being actively contemplated, it is pertinent that a study be made of programs which will permit optimum processing of the information by the computer.

It may be argued that, if the path of flight has been completely specified by some geometrical curve, why undertake the labor of determining a sequence of weights for statistical prediction. Wiener[1] himself had the following to say about this aspect:

> "Statistical prediction is essentially a method of refining a prediction which would be perfect by itself in an idealized case but which is corrupted by statistical errors, either in the observed quantity itself or in the observation. Geometrical facts must be predicted geometrically and analytical facts analytically, leaving only statistical facts to be predicted statistically."

Our problem deals with the statistical prediction of a time series derivable from a geometrical fact but corrupted or altered by other time series introduced by the encoding mechanism. We desire to know what the uncontaminated time series will do at some future instant. The problem has been idealized to the extent of assuming that (a) the pilot is capable of flying a geometrical course in spite of air currents and other disturbances, and (b) the errors inherent in radar tracking of aircraft are negligible compared to the quantization error. The validity of these assumptions will be examined later.

In tackling the over-all problem of air traffic control, one might logically hypothesize a control system such as that shown in Figure 3.01-1. The equipment lying within the broken-line boundary may be considered as part of the digital computer.

The operation of this hypothetical system may be described as follows:

a) the detection system monitors the position of the aircraft and supplies information as to the path variables of range and bearing with respect to the airstrip,

b) the continuous, time-varying signal related to either of the variables is fed to the encoder which samples and quantizes it, thus furnishing numerical data for the computer,

c) the prediction program processes this data in such manner as to yield the best possible predicted value for an epoch 15 seconds in the future,

A-50259-1

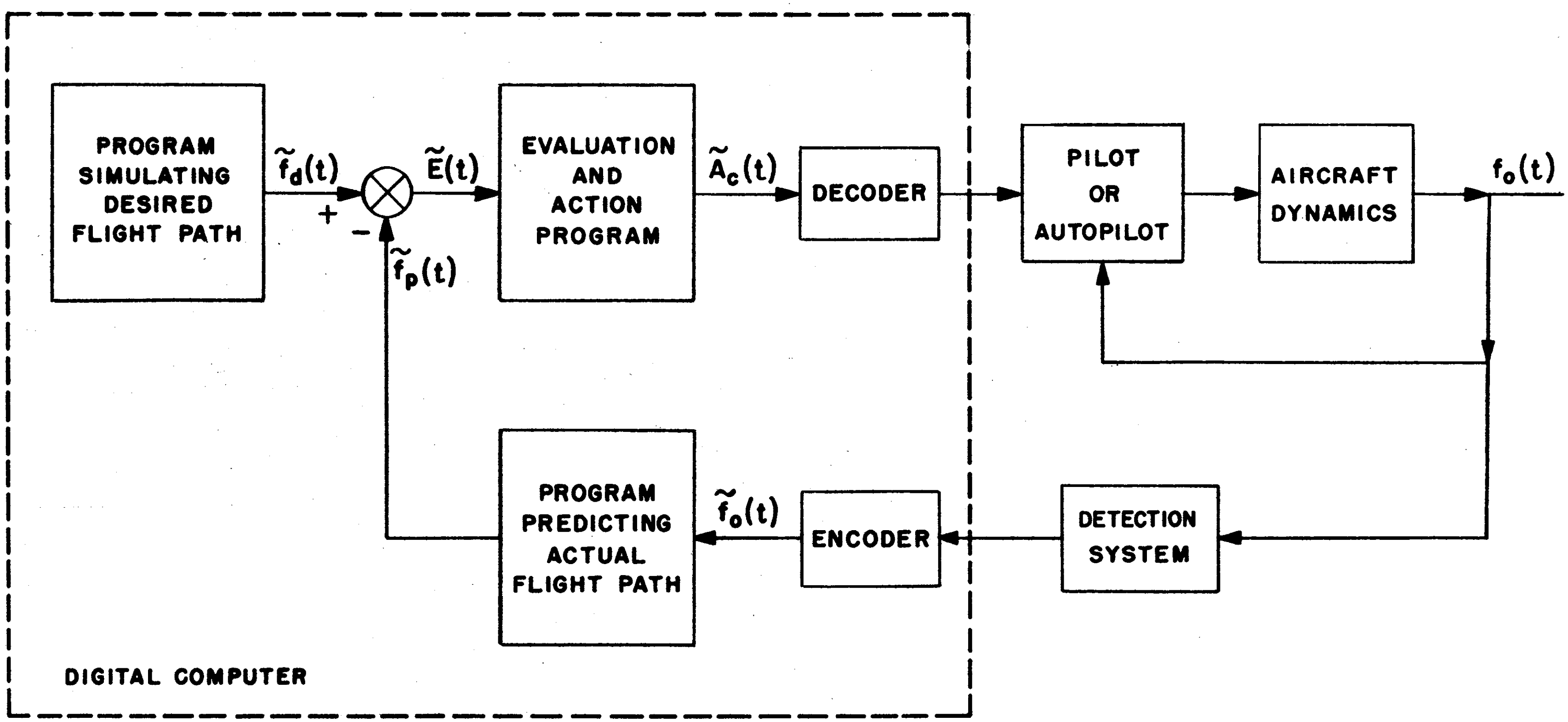


FIGURE 3.01-1-HYPOTHETICAL AIR TRAFFIC CONTROL SYSTEM

d) the actual and desired future values of the path variable are compared to yield an error quantity, $\tilde{E}(t)$. This error is evaluated and appropriate action is initiated to signal the pilot of any discrepancy in his flight path.

It is assumed that the continuous output signal, $f_o(t)$, will be sampled regularly at intervals of fifteen seconds and that a corrective signal, $\tilde{A}_c(t)$, will be supplied to the aircraft control system at the same instants. The various quantities indicated on the diagram may be defined as follows:

$f_o(t)$ = present value of path variable

$\tilde{f}_o(t)$ = sampled quantized present value of path variable

$\tilde{f}_p(t)$ = predicted value of path variable at an epoch fifteen seconds in the future

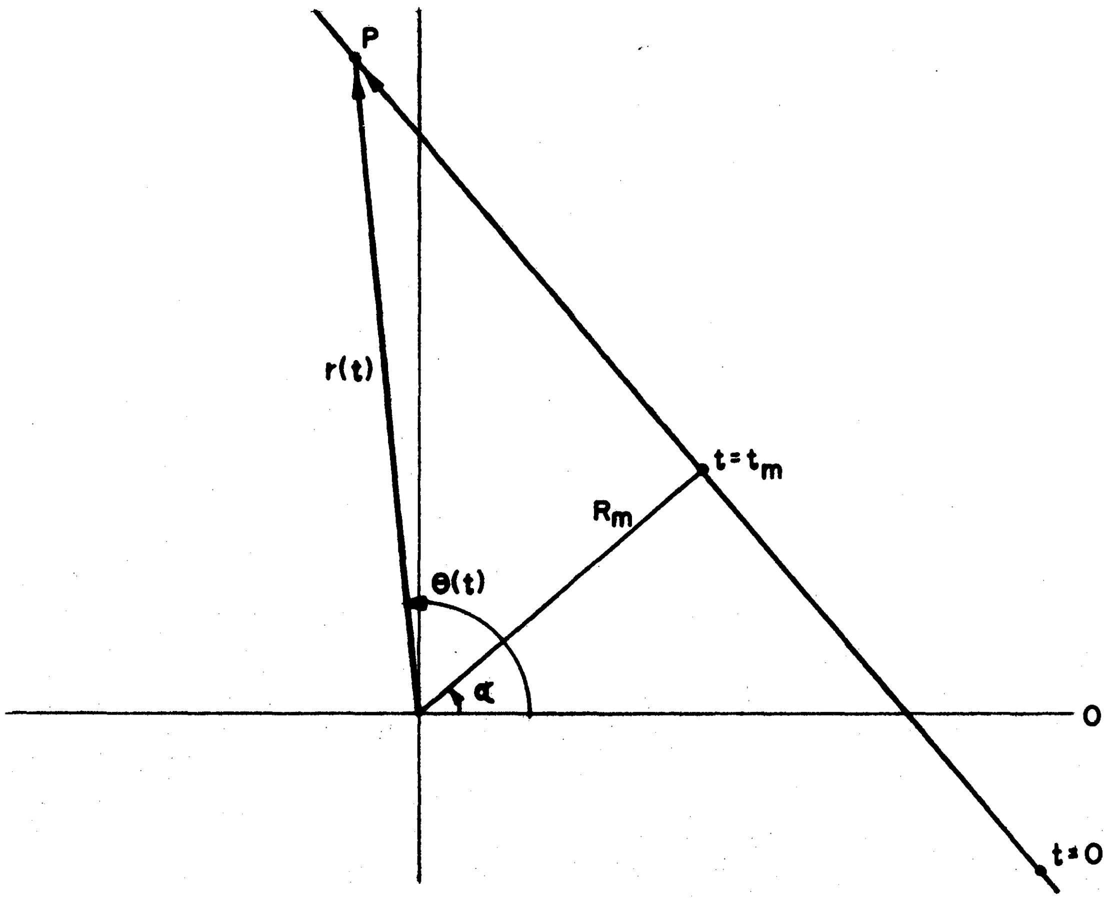
$\tilde{f}_d(t)$ = desired value of path variable at an epoch fifteen seconds in the future

$\tilde{E}(t) = \tilde{f}_d(t) - \tilde{f}_p(t)$

$\tilde{A}_c(t)$ = corrective action signal

Note that all these time series are discrete with the exception of $f_o(t)$.

As a particular $f_o(t)$ we shall use that generated by an aircraft flying a constant-velocity, constant-altitude straight line course which does not pass over the origin of the polar grid. This hypothetical mathematical model is shown in Figure 3.01-2 as well as the equations which



$$\theta(t) = \alpha + \tan^{-1}\left\{\frac{V_h}{R_m}(t - t_m)\right\} \qquad (3.01-1)$$

$$r(t) = R_m \sec(\theta - \alpha) \qquad (3.01-2)$$

WHERE V_h = HORIZONTAL VELOCITY

R_m = RANGE OF CROSSOVER

t_m = TIME CORRESPONDING TO R_m

FIGURE 3.01-2

TRAJECTORY FOR TARGET FLYING
STRAIGHT-LINE CONSTANT-VELOCITY COURSE

A-51056-1

specify the kinematics of the system. Clearly these equations define geometrical sequences which are definitely non-stationary. Furthermore, the nature of the quantization noise is completely specified when $f_o(t)$ is known explicitly as a function of time. The logical basis for the application of statistical techniques, however, lies in the fact that the quantities $\underline{V_h}$, $\underline{R_m}$, and $\underline{\alpha}$ are random variables.

The reader might logically question the necessity of going so far afield in search of an input signal to which to apply statistical prediction. It is a well known fact that, even if it were mandatory that he do so, the pilot is incapable of flying a precise geometrical course because of air perturbations and because of his inherent shortcomings as an element in a control loop. However, we again appeal to the physical context of the problem and point out that, in a two-dimensional control system such as ours, the only motion permissible for the aircraft is a coordinated turn (i.e., no slip, constant altitude). For such turns the aircraft dynamics are characterized by a first order lag in which the time constant is relatively long (about 0.5 seconds). In view of this fact, we may approximate the actual course of the aircraft by a series of straight lines. Thus, if we are able to predict a constant-velocity straight line course, we may be able to predict one in which the aircraft executes slow maneuvers.

The experimental work discussed herein is devoted entirely to the synthesis and evaluation of predicting filters based on equation 3.01-1. In this case it is evident that the ensemble of signals to be processed is a collection of arctangents. Examination of this equation shows that the random variable $\underline{\alpha}$ establishes the d-c level of the signal and, by reason of the circular symmetry of the system, may be ignored in our analysis. It should be noted that ignoring $\underline{\alpha}$ implies that we concern ourselves primarily with the time-varying part of any member of the signal ensemble.

In the absence of specific information regarding any existing air traffic control systems, we are forced to make certain assumptions concerning the control system and the statistical nature of the signals. These assumptions are:

a) the angle encoder is capable of distinguishing 256 levels (eight binary digits) in the signal,

b) quantization noise is the major component of corruption present in the signal,

c) all velocities in the interval $V_{min} \leq V_h \leq V_{max}$ are equally likely,

d) all minimum ranges in the interval $R_a \leq R_m \leq R_b$ are equally likely,

e) the maximum range of interest is R_{max}, and it is at this range that the aircraft are first detected.

f) the continuous signal, $\theta(t)$, is to be sampled regularly at intervals of 15 seconds.

Since the angular jitter in radar noise may at times be of the same order of magnitude as the angular quantum (about 1.4 degrees), it is questionable whether one is justified in assuming that quantizing noise is the major noise component. To simplify the analysis, however, we shall assume that the control system is relatively free of noise.

3.02 Experimental Results

The experimental results presented here are concerned with the design and evaluation of several linear time-varying filters whose inputs are the ensemble of arctangents derivable from equation 3.01-1 and whose outputs are to be future values of the signals. If, in that equation, the quantity α is ignored and assumptions e and f (from Section 3.01) are made, then equation 3.01-1 can be manipulated to obtain

$$\theta(n,j) = \tan^{-1}\left\{\left(\frac{j}{240}\right)\left(\frac{V_h}{R_m}\right)_n - \sqrt{\left(\frac{R_{max}}{R_m}\right)_n^2 - 1}\right\} \qquad (3.02\text{-}1)$$

where n denotes the particular member of the ensemble and j denotes the particular sampling instant. By means of this defining equation, we can determine the value of any member of the ensemble at any sampling instant for any choice of horizontal velocity, V_h, and minimum range, R_m. The input sequences to the filters are thus angular data derived from quantizing equation (3.02-1), and the desired output sequences are angular data corresponding to unquantized predicted values at an instant one sampling interval in the future.

The synthesis equations for the optimum least-squares predictor are given by equation (2.02-7)

$$\sum_{n=0}^{M} A_{n,k}\, R_{bb}(k-n,k-j) = R_{ba}(k+s,k-j)$$

$$\text{for } j = 0, 1, \ldots M$$

$$\text{and } s = 1$$

As has been previously indicated, this system of equations yields the optimum weighting sequence for the kth sampling instant. Since the statistical structure of this ensemble varies rather slowly, it was decided to synthesize a set of three filters, each of which operates over a certain range of the discrete time variable, k. In order to verify experimentally the fact that increasing the number of elements in the filter improves the performance, we have designed three such sets of filters -- for M = 1, 2, and 3.

The task of designing a filter is seen to be two-fold. First, one must calculate the correlation functions; and second, solve the system of simultaneous equations (2.02-7). The latter is relatively simple and can, for small M, be done by hand computation if necessary. The computation of the ensemble-averaged correlations, however, is a formidable task, even by automatic methods. It is rendered manageable, in our case, by the specification of the explicit form of the signals by equation (3.02-1) and of the probability density distributions by assumptions c and d. Under these circumstances the task can be mechanized by coding a computer program for the computation of the autocorrelation, $R_{bb}(k-n,k-j)$, and the crosscorrelation, $R_{ba}(k+1,k-j)$, for appropriate values of the arguments.

Having computed the correlations and inverted the matrix equation, we can then evaluate the performance of each of these three sets of filters by observing its performance on various sample signals. The filter is supplied with quantized data

$$\tilde{\theta}_{n,j},\ \tilde{\theta}_{n,j-1},\ \ldots,\ \tilde{\theta}_{n,j-M}\ .$$

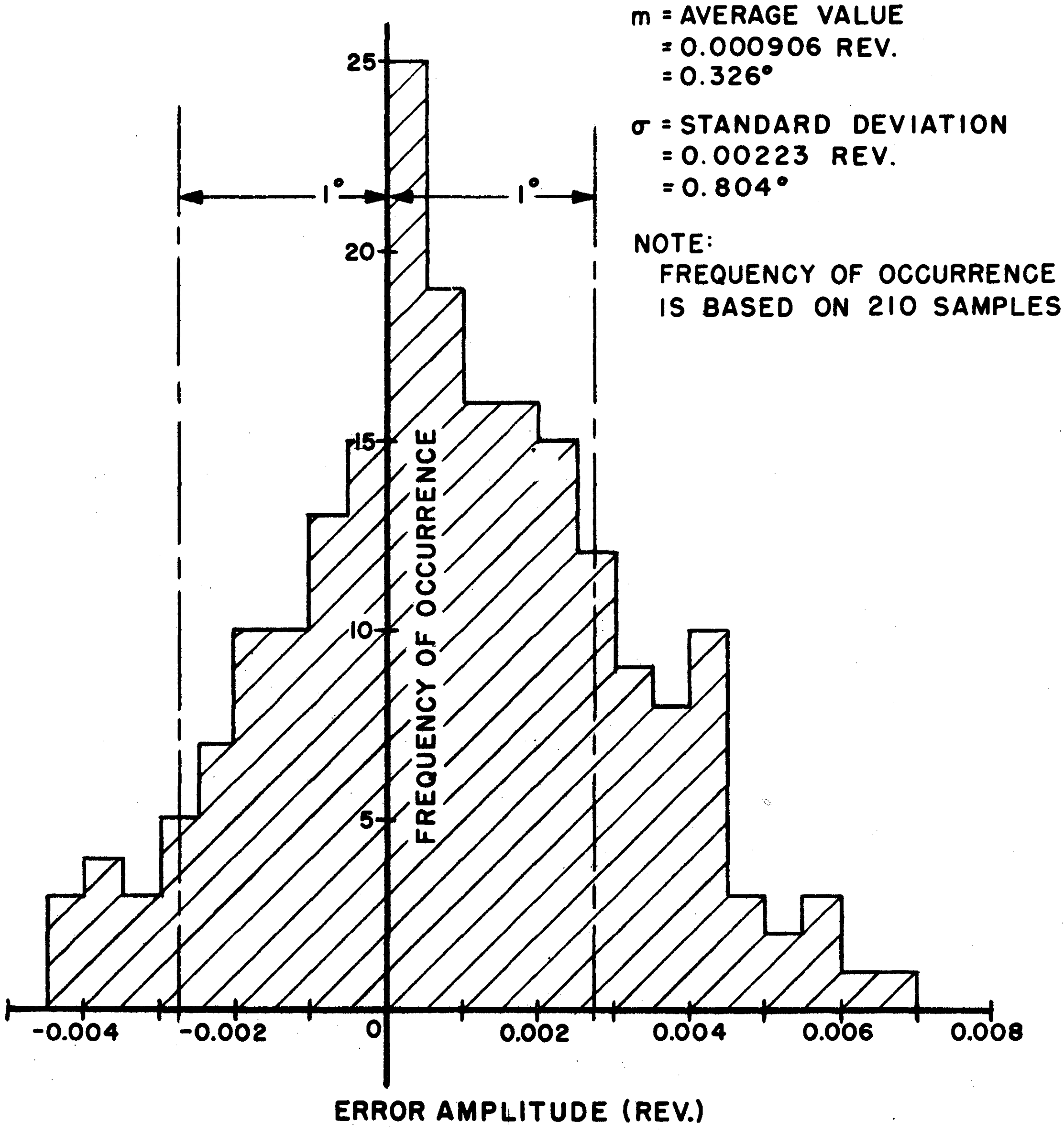


FIG 3.02-1

FREQUENCY OF OCCURRENCE OF ERROR
FOR A TWO-ELEMENT PREDICTOR
AS A FUNCTION OF ERROR AMPLITUDE

A-51052-2

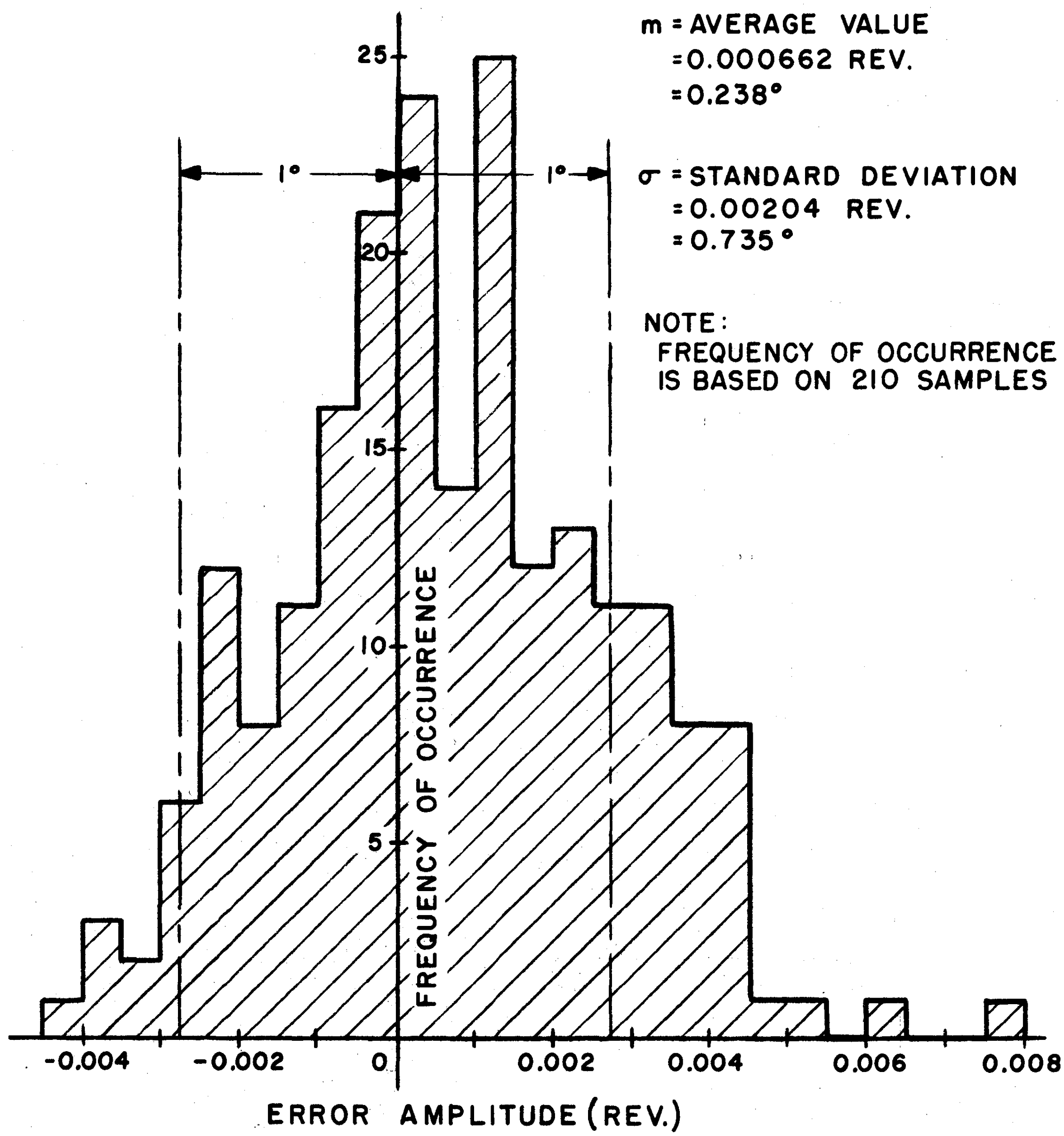


FIG. 3.02-2

FREQUENCY OF OCCURENCE OF ERROR
FOR A THREE-ELEMENT PREDICTOR
AS A FUNCTION OF ERROR AMPLITUDE

A-51053-2

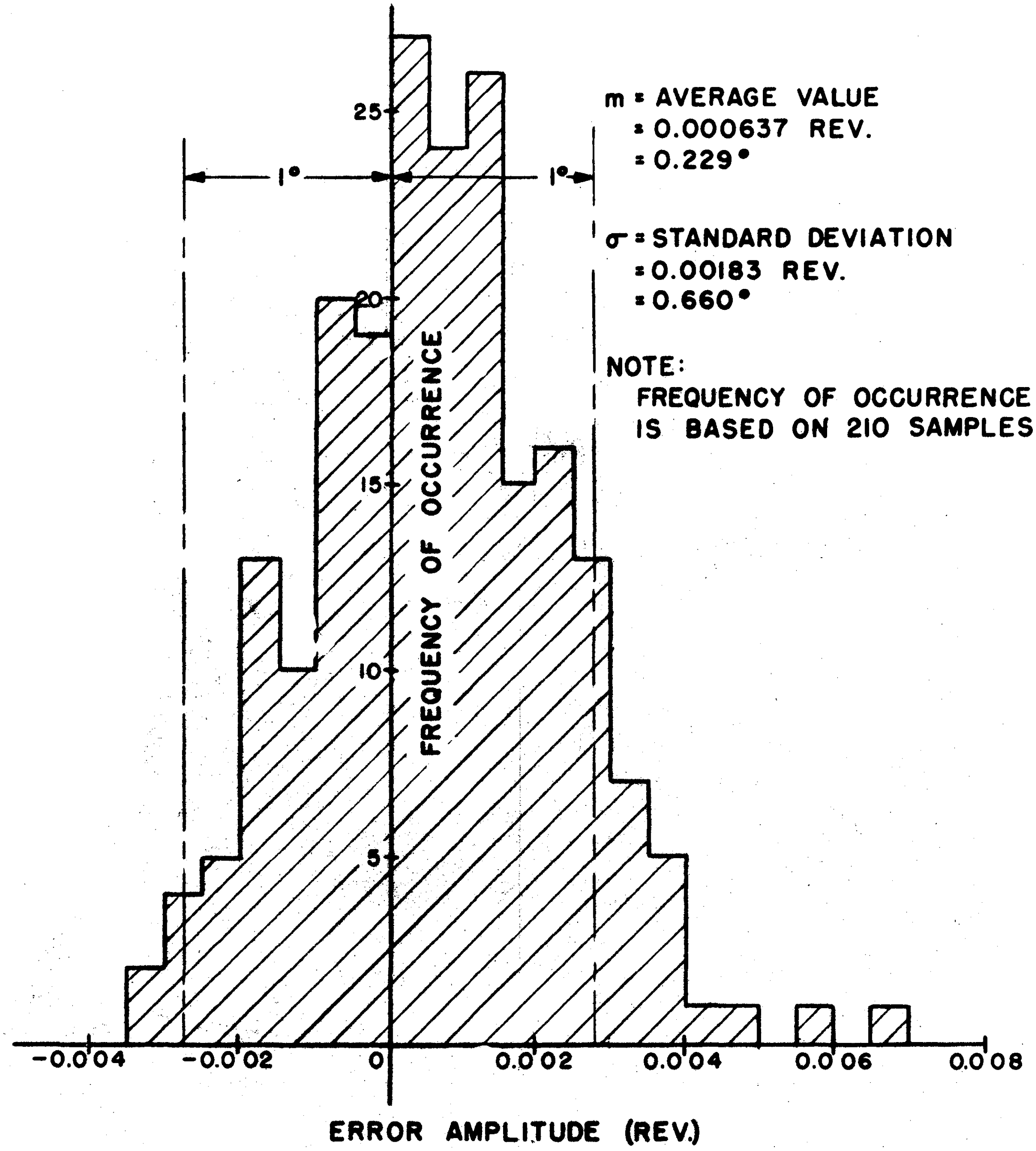


FIG 3.02-3

FREQUENCY OF OCCURRENCE OF ERROR FOR A FOUR-ELEMENT PREDICTOR AS A FUNCTION OF ERROR AMPLITUDE

A-50951-2

It will then weight and combine these data to yield an output

$$\theta_{n,j+1} = A_{0,j}\,\tilde{\theta}_{m,j} + A_{1,j}\,\tilde{\theta}_{n,j-1} + \dots + A_{M,j}\,\tilde{\theta}_{n,j-M}$$

which is the best possible approximation, in the least square error sense, to the true predicted value, $\theta_{n,j+1}$.

The foregoing synthesis and evaluation was made, and the results are presented below. The accompanying graphs show the distribution of frequency of error in prediction as a function of error in prediction for each of the three sets of filters. Some additional figures which give a measure of the quality of performance are given in the following table.

Type of Predictor	Average Error	Standard Deviation	Mean Square Error	Percent of Samples Having less than 1° Error	Maximum Error
2-element	0.326°	0.804°	0.750 deg²	75%	2.46°
3-element	0.238°	0.735°	0.598 deg²	79%	2.57°
4-element	0.229°	0.660°	0.434 deg²	89%	2.50

These results are based, for each set of predictors, on 210 sample signals drawn from an ensemble which included signals in addition to those in the original ensemble. Note that no special significance is to be attached to the tolerance value of one degree. This is merely an arbitrary basis for comparing these filters with each other and with the quantizing unit in azimuth of 1.4 degrees. It should be noted that there is a distinct improvement in performance as the number of elements in the weighting sequence is increased. Additional experiments made for signals with arbitrary non-zero α show that these filters perform nearly as well for these as for members of the original ensemble so long as we do not attempt to predict across the discontinuity at $\theta = \pm 180^{\circ}$. This, however, is not a fault inherent in statistical filters, but results from the fact that the signal is a multi-valued function. The predicting filter can be redesigned to cope with this discontinuity.

The experimental results summarized above indicate that rather good performance may be expected from relatively simple filters. Although improved performance can be expected from more complex filters, the net increment in improvement may not justify the additional computational labor and storage facilities.

It should be noted that our synthesis procedures are not restricted to predicting filters only, but can be employed to derive any of the compensating filters which are so frequently used to improve the performance of a servomechanism.

CHAPTER IV

CONCLUSIONS AND SUGGESTIONS FOR FURTHER STUDY

As a result of the theoretical analysis embodied in this investigation, the statistical communication theory developed by Wiener and Lee has been extended to the synthesis of real time computer programs. The accompanying experimental design of certain linear predictors has established the validity of this synthesis procedure. The extreme flexibility of the digital computer, however, permits, with equal facility, the design and application of either linear or non-linear programs (i.e., discrete filters).

Although the extension of this theory to the problem of discrete filtering may be considered a step in the reduction of the art of the computer programming for certain applications to an exact science, much work remains to be done in the further utilization, in the discrete domain, of the concepts formulated by Wiener and Lee. Some of the more promising subjects for investigation involving a union of statistical communication theory and digital computer practice are the following:

(a) the development of discrete filters capable of dealing with multiple time series. Such filters might be especially useful in industrial or chemical process control where it is desired that the digital computer control the behavior of the several interdependent variables which determine the quality or quantity of the end product, and

(b) the investigation of methods for determining what types of non-linearities, if any, should be incorporated in a filter for a specified ensemble of signals.

The above mentioned union of statistical theory and computer practice offers certain other advantages which should not be over-looked. Whereas the design of an analog network by the Wiener-Lee theory involves the solution of an integral equation, the corresponding design of a discrete filter involves only the solution of a set of linear simultaneous equations. Not only can the digital computer be programmed to evaluate any order correlation function (subject to limitations of storage facilities), but it can be programmed to perform the matrix inversion required for the determination of the weighting sequence. Regardless of whether the filter is to be linear or not, our problem always involves a set of linear equations.

Because of the comparative virginity of the field of discrete filter synthesis it was not possible to make any conclusive comparisons between the performance of this class of statistical filters and those of other classes of filters. We are, however, justified in concluding that the synthesis procedures developed in this investigation lead to sensibly designed filters which are entirely aware of the characteristics of the noise (which is always present in any real signal) as well as those of the message. We may further conclude that, when the generating mechanism for the ensemble of signals is known to be non-linear in nature, the performance of an appropriately designed non-linear filter is definitely superior to that of the linear filter. We note that the specific kind of non-linearity to be incorporated into the filter is of considerable importance.

The problem of quantization noise is of such importance in real time computer applications that a study of its statistical characteristics is certainly merited. These characteristics are a function of the particular ensemble of signals and of the encoding mechanism. Although the experimental designs of this investigation show that our filters are capable of dealing effectively with the noise resulting from uniform quantization of the signal, the use of such filters does not necessarily represent the optimum solution. An alternative and possibly more satisfactory solution involves both a non-uniform quantizing of the signal and a statistical filtering of the signal.

In conclusion, we may remark that the extension of statistical communication theory to the synthesis of digital computer programs has provided us with a logical means for endowing the computer with a higher order of intelligence.

Signed. Abraham Katz
Abraham Katz

Approved. Robert R. Everett
Robert R. Everett

APPENDIX

Statistical Communication Theory

As a ready reference for the reader, we propose to include here a very brief summary of the essential features of the Wiener-Kolmogoroff theory. For more complete information, he should consult references 1, 2, 6, and 7.

In Section 1.01 it was pointed out that the basic concept of this theory is that communication signals are to be treated as stationary time series. The structure of such signals is frequently so complex as to render it irresolvable in terms of summations of periodic or aperiodic components. In fact, if the signals are to convey any new information to the receptor, then they must be characterized by some elements of randomness in that they are at least partially unpredictable in advance by the receptor. Hence it is seen that we are frequently concerned with stationary random time series which may be either continuous or discrete.

For random, continuous phenomena having statistical properties which are stationary, one may define the linear correlation functions as

$$\phi_{aa}(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} f_a(t) f_a(t+\tau) dt$$

$$= \text{autocorrelation of } f_a(t)$$

and

$$\phi_{ab}(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} f_a(t) f_b(t+\tau) dt$$

$$= \text{crosscorrelation of } f_a(t) \text{ with } f_b(t).$$

These functions have the properties that

$$\phi_{aa}(0) \geqslant \phi_{aa}(\tau)$$

$$\phi_{aa}(\tau) = \phi_{aa}(-\tau)$$

$$\phi_{ab}(\tau) = \phi_{ba}(-\tau)$$

The definition of the autocorrelation function indicates a process of multiplying the function continuously by its value at a later time, τ, and averaging. The even function, $\phi_{aa}(\tau)$, is a maximum at $\tau = 0$ and is equal to the square of the rms value of $f_a(t)$. If $f_a(t)$ is nonperiodic, $\phi_{aa}(\tau)$ approaches the square of the average of $f_a(t)$ as τ increases; if $f_a(t)$ is periodic, $\phi_{aa}(\tau)$ has the same period. Any linearly additive component of $f_a(t)$ produces its own linearly additive component of $\phi_{aa}(\tau)$. A composite time function can be separated into linearly additive time functions and the autocorrelation of each added linearly to give the composite autocorrelation function. Furthermore, these linear functions, $\phi_{aa}(\tau)$ and $\phi_{ab}(\tau)$, and their respective power density spectra are determinable one from the other by a Fourier transformation. Thus,

$$\phi_{aa}(\tau) = \int_{-\infty}^{\infty} \Phi_{aa}(\omega) \cos \omega\tau \, d\omega$$

$$\Phi_{aa}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_{aa}(\tau) \cos \omega\tau \, d\tau$$

and

$$\phi_{ab}(\tau) = \int_{-\infty}^{\infty} \Phi_{ab}(\omega) \, e^{j\omega\tau} \, d\omega$$

$$\Phi_{ab}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_{ab}(\tau) \, e^{-j\omega\tau} d\tau$$

Again, linearly additive components of the linear correlation functions can be transformed separately and the separate transforms added to give the composite power spectra.

This extension of Fourier theories to the harmonic analysis of random processes through the medium of the linear correlation functions provides us with a powerful tool for the synthesis of linear networks which are optimal in a mean square error sense. For this class of filters, the expression for the measure of error is

$$\mathcal{E} = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \left[f_o(t) - f_d(t)\right]^2 dt$$

where $f_o(t)$ = actual output signal

$f_d(t)$ = desired output signal

when $f_i(t)$ is the input signal. Minimization of the error expression subject to the condition of linearity of the filtering mechanism yields the Wiener-Hopf equation which relates the impulse response of the optimum linear system to the statistical characteristics of the signal. Expressed in terms of time domain synthesis, this equation requires that

$$\phi_{id}(\tau) = \int_{-\infty}^{\infty} h(\sigma)\, \phi_{ii}(\tau-\sigma)\, d\sigma \quad \text{for } \tau \geq 0$$

where

$\phi_{ii}(\tau)$ = autocorrelation of the input signal

$\phi_{id}(\tau)$ = crosscorrelation between the input and the desired output signals.

From the foregoing we see that the linear correlations are entirely adequate for the specification of the linear system which minimizes the square of the error. An entirely equivalent equation in terms of frequency domain synthesis requires that

$$\Phi_{id}(\omega) = H(\omega)\ \Phi_{ii}(\omega)$$

An "optimum" filter designed for one member of an ensemble of signals will be equally effective in the processing of any other member of the ensemble having the same linear correlations.

By a logical extension of these ideas, one can define an infinite number of higher order correlation functions. Thus, for a stationary random signal

$$\phi_{aaa}(\tau_1;\tau_2) = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} f_a(t)f_a(t+\tau_1)f_a(t+\tau_2)dt$$

$$\phi_{aab}(\tau_1;\tau_2) = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} f_a(t)f_a(t+\tau_1)f_b(t+\tau_2)dt$$

$$\phi_{aaaa}(\tau_1;\tau_2;\tau_3) = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} f_a(t)f_a(t+\tau_1)f_a(t+\tau_2)f_a(t+\tau_3)dt$$

and so forth. Since the shifts τ_j are independent of each other, they may be visualized as orthogonal axes in the hyperspace in which the correlation functions are defined. Just as two dimensions are required for the geometrical representation of the linear functions, so are $n+1$ dimensions required for an nth-order correlation function.

Since it is conceivable that the reader has had little experience with the higher order correlations, we include a few simple examples to serve as illustrations. Accordingly we choose certain of those non-random functions of time which permit a direct integration procedure. For the non-random functions, the linear (and the non-linear) correlation functions are defined somewhat differently. Thus, for periodic functions

$$\phi_{aa}(\tau) = \frac{1}{T}\int_{o}^{T} f_a(t)f_a(t+\tau)dt$$

$$\phi_{ab}(\tau) = \frac{1}{T}\int_{o}^{T} f_a(t)f_b(t+\tau)dt$$

and for aperiodic functions

$$\phi_{aa}(\tau) = \int_{-\infty}^{\infty} f_a(t)f_a(t+\tau)dt$$

$$\phi_{ab}(\tau) = \int_{-\infty}^{\infty} f_a(t)f_b(t+\tau)dt$$

As in the case for correlations related to random sequences, there are unique Fourier transform pairs which relate the power (or energy) density spectra with the appropriate correlation functions.

Consider now the following examples of second-order autocorrelations:

(a) Periodic Signal

$$\phi_{aaa}(\tau_1;\tau_2) = \frac{1}{T}\int_{o}^{T} f_a(t)f_a(t+\tau_1)f_a(t+\tau_2)dt$$

Let $f_a(t) = A\cos(\omega t + \theta)$

$$\phi_{aaa}(\tau_1;\tau_2) = \frac{A^3}{T}\int_{o}^{T} \cos(\omega t+\theta)\cos(\omega t+\theta+\omega\tau_1)\cos(\omega t+\theta+\omega\tau_2)dt$$

$$= \frac{A^3}{4T}\int_{o}^{T} \Big\{\cos(\omega t+\theta+\omega\tau_1) + \cos(\omega t+\theta+\omega\tau_2+\omega\tau_1)$$

$$+ \cos(\omega t+\theta+\omega\tau_2-\omega\tau_2)+\cos(3\omega t+3\theta+\omega\tau_1+\omega\tau_2)\Big\}\, dt$$

$$\phi_{aaa}(\tau_1;\tau_2) = 0$$

It can be shown, in general that the second-order autocorrelation for periodic functions is everywhere zero.

(b) Aperiodic Signal

$$\phi_{aaa}(\tau_1;\tau_2) = \int_{-\infty}^{\infty} f_a(t)f_a(t+\tau_1)f_a(t+\tau_2)dt$$

$$\text{Let } f_a(t) = \begin{cases} E\ e^{-at} & 0 \leq t < \infty \\ 0 & -\infty < t < 0 \end{cases}$$

$$\phi_{aaa}(\tau_1;\tau_2) = E^3 \int_0^{\infty} e^{-at}\ e^{-a(t+|\tau_1|)}\ e^{-a(t+|\tau_2|)}\ dt$$

$$= E^3\ e^{-a(|\tau_1|+|\tau_2|)} \int_0^{\infty} e^{-3at}\ dt$$

$$\phi_{aaa}(\tau_1;\tau_2) = \frac{E^3}{3a}\ e^{-a(|\tau_1|+|\tau_2|)}$$

In general, the autocorrelation of any order of an exponential signal of the form given is also exponential in form.

Parallel to this theory of statistical analysis for continuous phenomena there runs a theory of discrete phenomena. These discrete phenomena may occur naturally or whenever a continuous time sequence is discretized. In the discrete case the function $f(t)$ of the continuous parameter t is replaced by the function f_k of the parameter k, which varies by discrete steps. Similarly the functions $\phi_{aa}(\tau)$ will be replaced by the discrete set of autocorrelation coefficients,

$$R_{aa}(k) = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{h=-N}^{N} a_h a_{h-k}$$

The analog of our previous function $\Phi_{aa}(\omega)$ will be

$$S_{aa}(\omega) = \frac{1}{2\pi} \sum_{-\infty}^{\infty} R_{aa}(k) e^{-jk\omega}$$

of period 2π. Likewise one can derive by a procedure entirely analogous to the continuous case the Wiener-Hopf equation for the linear discrete filter.

$$\sum_{n=0}^{M} G_n R_{bb}(k-n) = R_{ba}(k)$$

where $k = 0, 1, \ldots, M$.

The corresponding equation in terms of frequency domain synthesis is

$$S_{ba}(\omega) = g(\omega)\ S_{bb}(\omega)$$

where

$$g(\omega) = \sum_{-\infty}^{\infty} G_k\ e^{-jk\omega}$$

is the transfer function of the discrete filter.

The foregoing equations describing the linear discrete filter which is optimal in a least square error sense have served as the bases for the extension of the Wiener-Lee theory to the synthesis of digital computer programs.

BIBLIOGRAPHY

1. Wiener, N., The Extrapolation, Interpolation and Smoothing of Stationary Time Series, (New York: Wiley and Sons, 1949).
2. Kolmogoroff, A.,"Interpolation and Extrapolation of Stationaren Zufalligen Folgen," Bull. Acad. Sci. (URSS) Ser. Math 5, (1941), pp. 3-14.
3. Shannon, C.E., "A Mathematical Theory of Communication," Bell Sys. Tech. Jour., 27, I, July 1948; II, October 1948.
4. Fano, R.M., R.L.E. Technical Report, No. 65, M.I.T. (1949).
5. Fano, R.M., R.L.E. Technical Report, No. 149, M.I.T. (1950).
6. Lee, Y.W., R.L.E. Technical Report, No. 181, M.I.T. (1950).
7. Lee, Y.W., Course 6.563 Class Notes, M.I.T., Unpublished.
8. Bode, H.W. and Shannon, C.E., "A Simplified Derivation of Linear Least Square Smoothing and Prediction Theory," Proceedings of the I.R.E., Vol. 38, No. 4, April 1950.
9. Zadeh, L.A., "An Extension of Wiener's Theory of Prediction," Report SPD 243, Job No. 8390, August 1949, Special Products Division, M.W. Kellogg Co.
10. Singleton, H.E., R.L.E. Technical Report, No. 160, M.I.T. (1950).
11. Stutt, C.A., R.L.E. Technical Report, No. 182, M.I.T. (1951).
12. Cheatham, T.P., Jr., R.L.E. Technical Report, N. 122, M.I.T. (1951).
13. Salzer, J.M., D.Sc. Thesis, M.I.T.,(1951).
14. Ragazzini, J.R. and Zadeh, L.A., "Probability Criterion for the Design of Servomechanisms," Jour. App. Phys. Vol. 20 (Feb. 1949).
15. Floyd, G.F., Jr. D.Sc. Thesis, M.I.T. (1949).

16. Levinson, N., "The Wiener Root Mean Square Error Criterion in Filter Design and Prediction," Jour. Math. Phys., Vol. 25, No. 4 (Jan. 1947).

17. James, H.M.. Nichols, N.B., and Phillips, R.S., Theory of Servo-mechanisms, (New York: McGraw-Hill Co., New York, 1947).

18. Milne, W.E., Numerical Calculus, (Princeton, N.J.: Princeton University Press, 1949).

19. Bennett, W.R., "Spectra of Quantized Signals," Bell Sys. Tech. Jour., Vol. 27, No. 3, (July 1948).

20. Advances in Electronics edited by L. Marton, Vol. III, (New York: Academic Press Inc., 1951).

21. Linvill, W.K., D.Sc. Thesis, M.I.T., (1949).

22. Panter, P.F. and Dite, W., "Quantized Distortion in Pulse-Count Modulation with Nonuniform Spacing of Levels," Proceedings of the I.R.E., Vol. 39, No. 1, January 1951.

23. Booton, R.C., "An Optimization Theory for Time-Varying Linear Systems With Nonstationary Statistical Inputs," Project Meteor Report #72, M.I.T., (July 1951).