PREDICTION AND AUTOMATIC CONTROL WITH SPECIAL REFERENCE TO THE ELECTRICITY SUPPLY INDUSTRY

M.Sc.Thesis

by

Vishnu Raghunath Kulkarni

The University of Aston in Birmingham

November 1967.

S U M M A R Y:

The principal mathematical concepts and techniques of Automatic Control theory are briefly reviewed and those relating to probability concepts and prediction in particular, are discussed more fully.

A detailed account is given of the two most significant prediction theories in the field viz. the Wiener theory and the exponentially weighted method.

The application of prediction theory generally to a complex industrial situation and especially to the Electricity Supply Industry is explored and an experimentally based investigation leading to a computer programme is described. A closed-loop feedback model incorporating prediction is proposed as a basis essential for any systematic study of the relation of prediction theory to the practical problem of automatic contol of an optimal or near optimal character in the Electricity Supply Industry.

An appreciation is made of the general relevance of prediction theory as it now exists, to the theoretical problem of optimization and to the practical problem of effective automatic control in the electricity supply industry as understood from examination of recent Conference Reports.

ACKNOWLEDGEMENTS

This work was done under the supervision of Mr.J.A.Formby and I wish to express my sincere gratitude for his guidance and advice which were extremely helpful in the preparation of this Thesis.

My sincere thanks are also due to Mr.J.Gritton of CEGB for his help in obtaining the data and also to Mr.R.Kochache of the Electrical Engineering Department for his useful comments.

C O N T E N T S.

C O N T E N T S (contd)

C O N T E N T S (contd)

CHAPTER 1.

GENERAL INTRODUCTION.

1.1 PREDICTION.

The main task of a mathematician, employed for any Scientific or Industrial work, is to make predictions. As such what a mathematician tries to do is to construct a mathematical model of a situation from the experimentally-determined data which could be used to predict the likely future conditions. Mathematical models, (as W.Weaver puts it) are a set of assumptions plus the resulting body of pure theory which apply with strict accuracy to an idealised physical system so that the theory of the idealised system will "explain", or at least organise and simplify the real phenomena. These models consist of tabulations, graphs, equations and sets of equations, and the main advantages of the mathematical models are that -

(a) they express the real relations in a highly condensed and useful way,

(b) they are usually suitable for numerical calculation,

(c) they save a great deal of experimental work, calculating what would happen under certain assumed conditions.

Early scientists after carrying out simple experiments, formulated the results into the precise mathematical models. This is how, Robert Hooke after using different weights, and hence different forces to stretch elastic materials, formulated his experience to produce the mathematical model in which the

1.1 contd.

extension was precisely proportional to the force. One can see how Hooke's law is extremely useful for predicting the actual elongation of bars of metals under tension, provided the forces and extensions are not too great. Similarly, the knowledge of Newton's laws of motion enables one to predict the paths of the planets with extreme accuracy.

Ideally then, the mathematician's job in any Scientific or Industrial work should be to -

(i) Identify a problem,

(ii) Formulate the problem in the mathematical terms,

(iii) Solve the mathematical problem and

(iv) Interprete the solution.

In general, in actual practice, one finds the first two stages to be extremely difficult. Stage three, which is the solution of the problem, is greatly facilitated by the use of Digital computers.

A mathematician needs to compare his results with the results of further experimental measurements. The following schematic and a very representative diagram of the process of prediction is due to O.G.Sutton, which can be interpreted as -

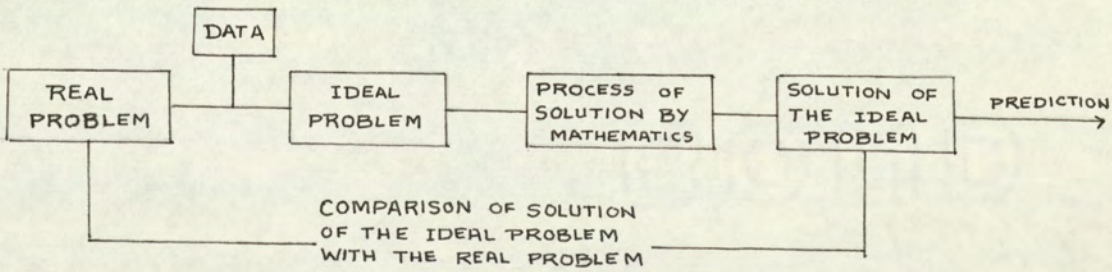


Fig. 1-1

1.1 contd.

one first specifies - a real problem which is then idealised such that it bears a strong resemblance to the real problem and to which the existing mathematical techniques can be applied. The solution of the ideal problem is then compared with the measurements made in the real problem.

At this point, it would be convenient to distinguish between the two kinds of models, viz. the deterministic and the stochastic. The models, such as those arising from Hooke's Law and Newton's laws of motion do not contain random variables and are termed deterministic. The deterministic models give definite predictions. A model containing random variables is called Stochastic and the future of a stochastic process is only partly determined by the past values of the variables. We shall confine our discussion mainly to stochastic processes and shall work with them in the following chapters.

Let us now turn our attention to the scientific theory of prediction which has its origin in the techniques developed for detecting hidden periodicities in a time series. It can be seen that, in sciences such as meteorology and astronomy, the observed phenomena may be represented by sums resembling a trigonometric series of the type -

$$a_1 \cos(\omega_1 t+\phi_1) + a_2 \cos(\omega_2 t+\phi_2) + a_3 \cos(\omega_3 t+\phi_3) + \ldots$$

1.1 contd.

Here, although the sum is not a periodic function of t, the elements constituting the sum are periodic, the periods being $\frac{2\pi}{n_1}$, $\frac{2\pi}{n_2}$, etc. So, if we are measuring the height of the tide at any instant at a given place, we need to take account of the effects due to the constituent tides such as, Semi-diurnal (with a period of half a day), Diurnal (with a period of one day), Fortnightly etc., each of which produces its own effect independently of others. The actual height then is the sum of these effects, and can be represented by an expression of the form

$$y = a_0 + a_1 \cos(n_1 t + \varepsilon_1) + a_2 \cos(n_2 t + \varepsilon_2) + a_3 \cos(n_3 t + \varepsilon_3) + \ldots$$

where each term of the series corresponds to the constituent tide. The amplitudes $a_0, a_1, a_2, a_3, \ldots$ and the phases $\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots$ can be determined experimentally, which makes it possible to predict the height of a tide at any future time, possibly with only a small error arising due to certain random elements such as wind.

In many other cases, however, the periods are quite unknown and the main task of the mathematician becomes one of discovering the periods $\frac{2\pi}{n_1}$, $\frac{2\pi}{n_2}$, etc., of the constituent terms by experiment and analysis. Lagrange's work published in 1774, appears to be the first mention of an analytical method for determining hidden periodicities. More recently, the periodogram

1.1 contd.

method for detecting hidden periodicities was developed by Sir Arthur Schuster. A periodogram is essentially a plot of the correlation ratio, and the recognition of peaks in the periodogram is the means by which the hidden periodicities are discovered.

The most modern theory of prediction and smoothing originated by N.Wiener is based on the extension of the idea of determination of periodicities in the random time-series. Wiener makes an extensive use of autocorrelation functions and the most important property of an autocorrelation function is the means with which one can discover hidden periodicities.

1.2 AUTOMATIC CONTROL.

The idea of automatic control dates back to the great contributions of James Watt to the development of steam engines in the late 18th century. These steam engines provided power to do the jobs which were previously done with human labour and also made power available, in a flexible form, to do various other jobs. A great amount of research, that followed the development of steam engines, was mainly in the field of controlling of the power-driven machines; however, the first analysis of the speed Governor which is the first real automatic controller did not appear until 1868 when Maxwell published his paper "ON GOVERNORS".

During the first world war period, Minorsky developed

1.2 contd.

the concept of a system which would automatically maintain a ship in a prescribed course. A signal, the difference between the desired course and the actual course, actuated a mechanism which continuously reset the ship's rudder to maintain the proper course. Minorsky in his paper published in 1922, considered problems of the reduction of steady state errors and of increasing the stability of a closed-loop control system. The present-day techniques of linear analysis have grown from the original work of H.Nyquist in the field of electronic amplifiers. Nyquist's famous criteria of the stability of Feedback control systems, found almost immediate and important applications in various fields.

A great deal of development in Automatic control occurred during the second world war. By that time, the previous anti-aircraft devices proved inadequate for the greatly increased speed achieved by the aircraft and this gave a great impetus to research into the effective means of adequate detection and tracking. The development of many allied techniques followed around the same period.

With the progress in the field of automatic control, there arose a need to design many complex control systems. The study of such complex systems is greatly faciliated by the science of Cybernetics, born soon after the war. This science is mainly a result of the research done by N.Wiener and

1.2 contd.

Arturo Rosenbleuth aimed at understanding certain neurological phenomena. It was realised that, in the workings of a human body, there are already superb examples of control systems, the designs of which are far more sophisticated than anything that has been achieved by a man. For example, the body temperature remains almost constant even when there are large and sudden changes in the surrounding temperature. Similarly, the eye rapidly changes its optical characteristics with the changing distances. The benefit of the study of Cybernetics reflects in the field of machine control, particularly feedback processes and the understanding of feedback mechanisms and malfunctions which occur in the diseases of man. At present, the main research work in Cybernetics is directed towards understanding the mechanism of the human brain, and toward the alternate communication paths which may permit the deaf or blind to "hear" or "see".

The field of automatic control is now extended to the various branches of science and Engineering. The basic essentials of automatic control, however, are common to all these branches; these are - i) a CONTROLLED CONDITION, e.g. temperature, ii) a MEASURING UNIT for measuring the value of the controlled condition, iii) a REGULATING UNIT which is an apparatus capable of effecting a change in the controlled condition and iv) a CONTROLLING UNIT

1.2 contd.

which is a means for operating the regulating unit in response to the measuring unit. The closed-loop is achieved through the controller and the basic arrangement of an automatic control system is as depicted in the figure below. However, in practice, the actual examples vary considerably in their layout and construction.

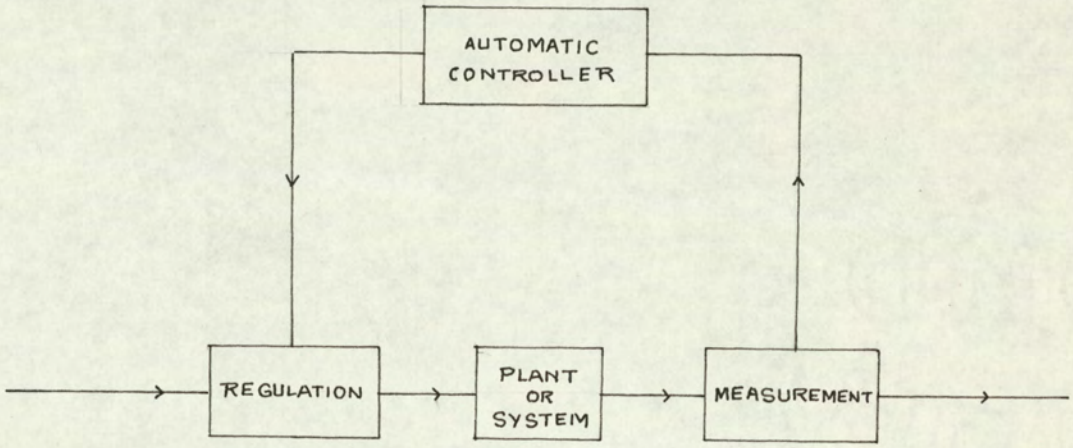


Fig.1-2. Basic arrangement of Automatic Control System.

1.3 CLASSIFICATION OF THE CHAPTERS:

After reviewing some of the basic concepts and definitions necessary to understand Wiener's theory in chapter 2, the Wiener theory of Linear least square smoothing and prediction is reviewed in chapter 3. Chapter 4 contains a survey of the more recently developed techniques of Exponentially weighted prediction and the other methods. The purpose of chapter 5 is to consider the applications of the various prediction

1.3 contd.

techniques reviewed, in the light of a specific problem. The problem considered is one of Electricity demand prediction in an area, up to several hours ahead.

The sixth chapter is on "Automatic control". The first section of the chapter contains a review of the techniques which form a basis for the modern automatic control theory, and in Section II, a specific problem of automatic control in the Electricity supply is considered, where a proposed general system for power-flow control is depicted by means of a block diagram and is discussed to some extent. General discussion and conclusion appears in the last chapter.

CHAPTER 2.

SOME CONCEPTS AND DEFINITIONS.

2.1 RANDOM PROCESSES:

One can cite a wide range of natural phenomena which constitute random processes. For example, meteorological phenomena such as, temperature variations, wind velocity, light intensity etc., and certain economic fluctuations, thermal noise in electric circuits etc., are all random processes. The main characteristic of a random process is that, it is very difficult to decide the future course, just from the knowledge of the past behaviour of a single random record. The most one can do is to draw reasonable conclusions about the probable spread in future values from estimates of mean and mean square values derived from an ensemble or large collections of records. Consider, for instance, the daily time records of certain meteorological phenomena at a particular place. This would exhibit random fluctuations about an average value. What one could do is to calculate probabilities of exceeding certain arbitrary values from all of the records, and in this way estimate how future records might behave.

Definition: A random process, also called time-series or a stochastic process, is an ensemble or assembly (i.e. large collections of records) comprised of functions of time $\{x(t)\}$, such that the ensemble can be characterised through the statistical properties.

Now, consider the ensemble $\{x(t)\}$ of the time functions

2.1 contd.

which constitutes a random process. We shall denote by $P_1(x,t)dx$, the probability that x will lie between x and x+dx at time t. The function $P_1(x,t)$ is the first probability distribution from which the ensemble average $\overline{x(t)}$ can be calculated -

$$\overline{x(t)} = \int_{-\infty}^{\infty} xP_1(x,t)dx \qquad (2\text{-}1.1)$$

This formula can be generalised to give mean-squared value of x(t) and more generally the nth moment of x(t).

$\overline{x(t)}^2$ = the mean squared value of x(t)

$$= \int_{-\infty}^{\infty} x^2 P_1(x,t)dx \qquad (2\text{-}1.2)$$

and $\overline{x(t)}^n$ = nth moment of x(t)

$$= \int_{-\infty}^{\infty} x^n P_1(x,t)dx \qquad (2\text{-}1.3)$$

Also, the probability that the total number of pairs in which x occurs in the range x_1 to x_1+dx_1 at t_1 and in the range x_2 to x_2+dx_2 at t_2 can be written as $P_2(x_1,t_1;x_2,t_2)dx_1dx_2$. The function $P_2(x_1,t_1;x_2,t_2)$ being the second probability distribution. From the second probability distribution, the correlation function $C_{xx}(t_1,t_2)$ can be calculated -

2.1 contd.

$$C_{xx}(t_1, t_2) = \overline{x(t_1)x(t_2)}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 P_2(x_1, t_1; x_2, t_2)\, dx_1\, dx_2 ; \quad (2\text{-}1.4)$$

This correlation function is also called the AUTOCORRELATION FUNCTION OF the random process $\{x(t)\}$ at the times t_1 and t_2. It can be noted that for $t_1 = t_2$, the correlation function becomes the ensemble average $\overline{x(t)}^2$, which is the mean square (ensemble) value of the random process at $t = t_1$.

Suppose there are two random processes $\{x(t)\}$ and $\{y(t)\}$. Then the function

$$C_{xy}(t_1, t_2) = \overline{x(t_1)y(t_2)}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, y\, P_2'(x, t_1;\, y, t_2)\, dx\, dy \qquad (2\text{-}1.5)$$

is called the CROSS CORRELATION FUNCTION. $P_2'(x, t_1; y, t_2)$ being the second joint probability distribution function.

The above random processes are of the general type and are called NON-STATIONARY. The ensemble averages are always used to determine their statistical properties.

Properties of stationary random processes:

One of the main assumptions on which the Wiener theory (next chapter) rests is that the input time series, represented by signal $s(t)$ and noise $n(t)$ are stationary random

2.1 contd.

processes and for this reason it will be desirable to discuss some of the important properties of stationarity.

A random function is called a stationary random function, if all the statistical characteristics of the function are time independent. Thus, we can denote by $P_1(x)$ the first probability distribution for a stationary random function, i.e. $P_1(x)dx$ is the probability of finding x between x and x+dx and let x(t) be a random function. Then for a stationary random function, the ensemble or assembly average and the time average are equal, i.e.

$$\overline{x(t)} = \int_{-\infty}^{\infty} x\, P_1(x)dx = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\, dt; \qquad (2\text{-}1.6)$$

where the very large record on a single member of the ensemble is cut into lengths T, and T is large.

As above, the formula for assembly average can be generalised to the arbitrary powers of x, giving moments. Thus

$$M_n = \overline{x^n} = \int_{-\infty}^{\infty} x^n\, P_1(x)dx \qquad (2\text{-}1.7)$$

M_n denotes the nth moment of the first probability distribution.

It is interesting to see that more and more information about $P_1(x)$ can be obtained as more and more moments are known. From the first and second moments, one can compute the variance

2.1 contd.

σ^2 and thus the standard deviation σ, which is a measure of the probability distribution $P_1(x)$ about the average value $\overline{x(t)}$,

$$\sigma^2 = \overline{(x - \overline{x})^2} = \int_{-\infty}^{\infty} (x-\overline{x})^2 \, P_1(x)dx = \overline{x^2} - (\overline{x})^2; \quad (2\text{-}1.8)$$

The third moment gives a measure of skewness of the probability distribution. In some cases it is possible to determine the distribution from the knowledge of moments. For example, if

$$M_{2k+1} = 0 \qquad \text{for } k = 0,1,2, \ldots$$

$$M_{2k} = 1.3.5 \ldots (2k-1)\, \sigma^{2k}$$

Then the first probability distribution is a Gaussian or normal distribution -

$$P_1(x) = \frac{1}{\sigma \sqrt{2\pi}} \, e^{-x^2/2\sigma^2} \qquad (2\text{-}1.9)$$

For a stationary random function, the second probability distribution P_2 will depend only on the time interval $\tau = t_2 - t_1$, rather than on t_1 and t_2 separately. Hence, for a stationary random function, the probability of finding a pair of values between x_1+dx_1 and between x_2+dx_2 at an interval of time equal to τ is $P_2(x_1, x_2; \tau)dx_1 dx_2$. The function $P_2(x_1, x_2; \tau)$ is the second distribution function.

The correlation function can now be written as -

2.1 contd.

$$C_{xx}(\tau) = \overline{x_1 x_2} = \overline{x(t)\ x(t+\tau)}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 P_2(x_1, x_2; \tau)\, dx\, dx \qquad (2\text{-}1.10)$$

In the case of a stationary random function this can also be obtained by the time averaging i.e.

$$C_{xx}(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\ x\ (t+\tau)dt \qquad (2\text{-}1.11)$$

$C_{xx}(\tau)$ measures the inter-relationship of the x's measured at two different time instants and is called the autocorrelation function of the stationary random process. This inter-relation weakens with the incrdase of τ and for large τ, $x(t)$ and $x(t+\tau)$ will be independent of each other. The second probability distribution then becomes the product of $P_1(x_1)$ and $P_2(x_2)$. Thus for large τ

$$C_{xx}(\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 P_1(x_1)\ P_2(x_2) dx_1 dx_2 = (\bar{x})^2 \qquad (2\text{-}1.12)$$

For $\tau = 0$, equation (2-1.11) reduces to

$$C_{xx}(0) = \overline{x^2} \qquad (2\text{-}1.13)$$

NOTE: If the value of the autocorrelation function eq.(2-1.11) and the mean value given by

2.1 contd.

$$\bar{x} = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\, dt$$

do not differ when computed over different sample functions, the random process is said to be ergodic. In actual practice, random data representing stationary physical phenomena are generally ergodic, and hence throughout our discussion of the stationary random processes we shall assume ergodicity.

Autocorrelation and hidden periodicity:

Let

$$x(t) = a \sin(\omega t + \phi) \qquad \omega = 2\pi f \qquad (2\text{-}1.14)$$

be a single record of a stationary random process where a and ϕ are constants. Then the autocorrelation function (from eq.2-1.11)

$$C(\tau, x) = \lim_{T \to \infty} \frac{a^2}{T} \int_{-T/2}^{T/2} \sin(\omega t + \phi) \sin[\omega(t+\tau) + \phi]\, dt$$

$$= a^2 \cos \omega\tau \qquad (2\text{-}1.15)$$

It can be seen that $C(\tau, x)$ is independent of the phase angle ϕ, an even function of τ with a maximum at $\tau = 0$, and periodic in τ with the same frequency as originally present in $x(t)$.

Similarly, if $x(t)$ is of the form

$$x(t) = \frac{a_0}{2} + \sum_{1}^{N} (a_n \operatorname{Cos} \omega_n t + b \operatorname{Sin} \omega_n t) \quad ; \omega_n = 2\pi f_n$$

2.1 contd.

$$= c_0 + \sum_{1}^{N} c_n \cos(\omega_n t - \theta_n) \qquad (2\text{-}1.16)$$

where $c_0 = \frac{a_0}{2}$, $c_n^2 = a_n^2 + b_n^2$, $\theta_n = \tan^{-1}(b_n/a_n)$.

The quantities c_n and θ_n represent amplitude and phase factors associated with each frequency f_n in $x(t)$;

From (2-1.11)

$$C(\tau,x) = c_0^2 + \sum_{1}^{N} c_n^2 \cos(\omega_n \tau) \qquad (2\text{-}1.17)$$

$C(\tau,x)$ is again independent of phase relations in $x(t)$, an even function of τ with a maximum at $t = 0$, and containing all of the previous periodicities of $x(t)$.

Thus, if a random function contains such hidden periodicities, then the calculation of $C(\tau,x)$ for large values of τ should uncover them, since otherwise, $C(\tau,x)$ would be expected to vanish for large τ.

2.2 INTEGRAL TRANSFORMS:

(i) Laplace transform:-

If $f(t)$ is a function of the time variable t, where $t > 0$, the Laplace transform of $f(t)$ denoted by $F(s)$ is given by

2.2 contd.

(i) contd.

$$\mathcal{L}[f(t)] = F(s) = \int_0^\infty e^{-st} f(t)\, dt \qquad (2\text{-}2.1)$$

where s is a complex variable, $(s = \sigma + i\omega)$, with a positive real part.

When F(s) is known, f(t) can be calculated by the inverse Laplace transform:

$$\mathcal{L}^{-1}[F(s)] = f(t) = \frac{1}{2\pi i}\int_{\nu-i\infty}^{\nu+i\infty} e^{st} F(s)\, ds \qquad (2\text{-}2.2)$$

The actual evaluation of f(t) from this formula, can be done, by deforming the path of integration according to the singularities of F(s). Laplace transformation is found suitable for the problems, where the function f(t) is defined by a differential equation with specific initial conditions.

Certain fundamental properties:-

If $f_1(t)$ and $f_2(t)$ are of exponential type, and c_1 and c_2 are any two complex numbers, then

$$\mathcal{L}(c_1 f_1 + c_2 f_2) = c_1 \mathcal{L}(f_1) + c_2 \mathcal{L}(f_2) \qquad (2\text{-}2.3)$$

Also, if $f_1(t) = \mathcal{L}^{-1}[F_1(s)]$ and $f_2(t) = \mathcal{L}^{-1}[F_2(s)]$, then

$$\mathcal{L}^{-1}(c_1 F_1 + c_2 F_2) = c_1 \mathcal{L}^{-1}(F_1) + c_2 \mathcal{L}^{-1}(F_2)$$

$$= c_1 f_1(t) + c_2 f_2(t) \qquad (2\text{-}2.4)$$

One most important theorem concerns the

2.2 contd.

(i) contd.

convolution (or Faltung) of two functions: If $f(t)$ and $g(t)$ are the two given functions, then the convolution is defined by the relation -

$$f(t) * g(t) = \int_0^t f(u)\, g(t-u)\, du \qquad (2\text{-}2.5)$$

The Laplace transform of the convolution is the product of the Laplace transforms of the functions i.e.

$$\mathcal{L}\,[f(t) * g(t)] = F(s)\, G(s) \qquad (2\text{-}2.6)$$

(ii) Fourier transforms:-

The Fourier transforms are defined by the relations -

$$F(\omega) = \mathcal{F}(f) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)\, e^{-i\omega t}\, dt \qquad (2\text{-}2.7)$$

$$f(t) = \mathcal{F}(F) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega)\, e^{i\omega t}\, d\omega \qquad (2\text{-}2.8)$$

Here, the convolution is defined by the expression

$$f * g = \int_{-\infty}^{\infty} f(u)\, g(t-u)\, du \qquad (2\text{-}2.9)$$

Also, if $F(\omega)$ and $G(\omega)$ denote the Fourier transforms of $f(t)$ and $g(t)$ respectively, then

2.2 contd.

(ii) contd.

$$\left(\frac{1}{\sqrt{2\pi}}\right) \mathcal{F}(f * g) = F(\omega)\ G(\omega) \qquad (2\text{-}2.10)$$

Fourier Sine and Cosine transforms are given by the relations -

$$F(\omega) = \sqrt{\frac{2}{\pi}} \int_0^\infty f(t)\ \mathrm{Sin}(\omega t)\ dt \qquad (2\text{-}2.11)$$

$$F(\omega) = \sqrt{\frac{2}{\pi}} \int_0^\infty f(t)\ \mathrm{Cos}(\omega t)\ dt \qquad (2\text{-}2.12)$$

and the inverse of these transforms are given by -

$$f(t) = \sqrt{\frac{2}{\pi}} \int_0^\infty F(\omega)\ \mathrm{Sin}(\omega t)\ d\omega \qquad (2\text{-}2.13)$$

$$f(t) = \sqrt{\frac{2}{\pi}} \int_0^\infty F(\omega)\ \mathrm{Cos}(\omega t)\ d\omega \qquad (2\text{-}2.14)$$

Probability distribution of sums:-

Consider a function $f(t)$ as a sum of a signal $s(t)$ and noise $n(t)$ where $s(t)$ and $n(t)$ are independent random variables. We have

$$f(t) = s(t) + n(t)$$

Now, let $p(f)$, $p_1(s)$ and $p_2(n)$ be the respective probability density functions of $\{f(t)\}$, $\{s(t)\}$ and $\{n(t)\}$. Then by convolution

2.2 contd.

$$p(f) = \int_{-\infty}^{\infty} p_1(s)\, p_2(f - s)\, ds \qquad (2\text{-}2.15)$$

$p(f)$ can also be calculated from the Fourier transforms of the probability densities. For example, if $Y(i\omega)$, $Y_1(i\omega)$ and $Y_2(i\omega)$ denote the Fourier transforms of $p(f)$, $p_1(s)$ and $p_2(n)$ respectively, then we have

$$Y(i\omega) = \int_{-\infty}^{\infty} e^{-i\omega f}\, p(f)\, df$$

$$Y_1(i\omega) = \int_{-\infty}^{\infty} e^{-i\omega s}\, p_1(s)\, ds$$

$$Y_2(i\omega) = \int_{-\infty}^{\infty} e^{-i\omega n}\, p_2(n)\, dn,$$

from which it can be shown that

$$Y(i\omega) = Y_1(i\omega)\, Y_2(i\omega)$$

The inverse transform of the last relation gives $p(f)$. Thus

$$p(f) = \mathcal{F}[Y(i\omega)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t}\, Y_1(i\omega)\, Y_2(i\omega)\, ds; \qquad (2\text{-}2.16)$$

2.3 LINEAR SYSTEMS:

A physical system is said to be linear when the equations which govern it are linear differential equations

2.3 contd.

i.e. the terms of the equation contain only the first powers of the dependent variables or its time derivatives. When the coefficients of the differential equation are constants independent of time, the system is called constant coefficient or fixed parameter system, and when the coefficients are functions of time it is called the system with variable coefficients.

If the terms contain higher powers of the dependent variable or cross products of the dependent variable and its derivatives, the differential equation is said to be non-linear and the system which is governed by such an equation - a non-linear system.

In actual practice, there are no perfect linear systems, because any physical system when analysed in great detail is always non-linear. It is only under some assumptions that a system may be correctly represented by a linear scheme. It is found that a large number of engineering systems fall into the category of linear systems with constant parameters.

Some properties:

For a system denoted by 'h', the relationship between all the input signals x(t) and all the output signals y(t) is of the type -

$$y(t) = h\,[x(t)] \qquad (2\text{-}3.1)$$

2.3 contd.

and the following properties of additivity and homogeneity follow -

$$h[x_1(t) + x_2(t)] = h[x_1(t)] + h[x_2(t)] \qquad (2\text{-}3.2)$$

$$h[a\,x(t)] = a\,h[x(t)] \qquad (2\text{-}3.3)$$

where a is a constant. For a time independent or constant parameter system -

$$y(t + \tau) = h[x(t + \tau)] \qquad (2\text{-}3.4)$$

Consider, that an input $x(t) = e^{i\omega t}$, is applied to a linear constant parameter, infinite operating system. Using (2-3.4), the output signal is

$$\begin{aligned} y(t + \tau) &= h[e^{i\omega(t+\tau)}] \\ &= h[e^{i\omega t}]\, e^{i\omega\tau} \\ &= y(t)\, e^{i\omega\tau} \end{aligned} \qquad (2\text{-}3.5)$$

In particular, for $t = 0$,

$$y(\tau) = y(0)\, e^{i\omega\tau} \qquad (2\text{-}3.6)$$

is the response of a given system to the input signal $e^{i\omega\tau}$.

Now, let $x(t)$ be a signal for which a Fourier transform $X(i\omega)$ exists, i.e.

$$X(i\omega) = \mathfrak{F}[x(t)] = \int_{-\infty}^{\infty} e^{i\omega t}\, x(t)\, dt \qquad (2\text{-}3.7)$$

and then

$$x(t) = \mathfrak{F}^{-1}[X(i\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(i\omega) e^{i\omega t} d\omega \qquad (2\text{-}3.8)$$

2.3 contd.

From (2-3.1) we have

$$\mathcal{F}[y(t)] = Y(i\omega) = X(i\omega)\ H(i\omega) \qquad (2\text{-}3.9)$$

the inverse of which gives -

$$y(t) = \mathcal{F}[X(i\omega)\ H(i\omega)]$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} X(i\omega)\ H(i\omega)\ e^{i\omega t}\ d\omega \qquad (2\text{-}3.10)$$

The response or the output of a linear system (or filter) subjected to a unit impulse signal $x(t) = \delta(t)$ as an input, is of particular significance. This unit impulse function $\delta(t)$ is called the Dirac-delta function and is mathematically defined by

$$\delta(-t) = \delta(t) = 0 \quad \text{for } t \neq 0$$

$$\delta(t) \to \infty \quad \text{for } t = 0 \qquad (2\text{-}3.11)$$

$$\int_{-\infty}^{\infty} \delta(t)dt = 1 \text{ and } \int_{0}^{\infty} \delta(t)\ dt = 1/2 \qquad (2\text{-}3.12)$$

For the function $f(t)$ continuous at $t = t_o$, the Dirac-delta function has four characteristic properties -

$$\int_{-\infty}^{\infty} f(t)\ \delta(t-t_o)dt = f(t_o) = \int_{-\infty}^{\infty} f(t)\ \delta(t_o-t)dt; \qquad (2\text{-}3.13)$$

$$\int_{-\infty}^{\infty} f(t)\ \delta(t-t_o)dt = 1/2\ f(a_o);\ \text{if } t_o=t_1 \text{ or } t_o=t_2; \qquad (2\text{-}3.14)$$

$$\int_{-\infty}^{\infty} \delta(t-t_o)dt = 1 \qquad (2\text{-}3.15)$$

2.3 contd.

$$\int_{-\infty}^{t_o} \delta(t-t_o)dt = \int_{t_o}^{\infty} \delta(t-t_o)dt = 1/2 \qquad (2\text{-}3.16)$$

Example.

$$\int_{-\infty}^{\infty} f(t)\ \delta(at-b)dt = \int_{-\infty}^{\infty} f(u/a)\ \delta(u-b)\ du/a = 1/a\ f(b/a)$$

Thus, in case of a unit impulse input signal to a linear system

$$[x(t)] = X(i\omega) = \mathcal{F}[\delta(t)] = \int_{-\infty}^{\infty} \delta(t)\ e^{-i\omega t}\ dt = 1; \quad (2\text{-}3.17)$$

From equations (2-3.9) and (2-3.17),

$$Y(i\omega) = 1.\ H(i\omega) \qquad (2\text{-}3.18)$$

and from (2-3.10) and (2-3.18)

$$y_\delta(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} H(i\omega)\ e^{i\omega t}\ d\omega \qquad (2\text{-}3.19)$$

The last relation is an inverse Fourier transform of $H(i\omega)$ and hence the following relations hold

$$h(t) = y_\delta(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} H(i\omega)\ e^{i\omega t}\ d\omega \qquad (2\text{-}3.20)$$

$$H(i\omega) = \int_{-\infty}^{\infty} h(t)\ e^{-i\omega t}\ dt \qquad (2\text{-}3.21)$$

If the unit impulse response $h(t) = 0$ for $t < 0$,

2.3 contd.

the linear system is said to be physically realisable.

Response of a linear system:

Using relations (2-3.10) and (2-3.21), we can write

$$y(t) = \int_{-\infty}^{\infty} h(\tau)\, d\tau \cdot \frac{1}{2\pi} \int_{-\infty}^{\infty} X(i\omega)\, e^{i\omega(t-\tau)}\, d\omega$$

The second integral is the inverse Fourier transform of $X(i\omega)$ evaluated at $t-\tau$, which is written as $x(t-\tau)$ and hence

$$y(t) = \int_{-\infty}^{\infty} h(\tau)\, x(t-\tau) d\tau \; ; \; h(\tau) = 0, \; \tau < 0$$

$$= \int_{0}^{\infty} h(\tau)\, x(t-\tau) d\tau \qquad (2\text{-}3.22)$$

The equivalent form is

$$y(t) = \int_{-\infty}^{\infty} x(\tau)\, h(t-\tau) d\tau$$

$$= \int_{-\infty}^{t} x(\tau)\, h(t-\tau) d\tau \qquad (2\text{-}3.23)$$

Dirac-delta function and Wiener's prediction theory and filter problem.

Wiener's prediction theory and the filter problem will be discussed to some extent in the next chapter. If we consider the problem of determining the optimum constant

2.3 contd.

parameter linear system for stationary random processes and infinite operating times, then in pure filtering cases, the desired output is the signal input $s(t)$ and hence

$$s(t) = \int_{-\infty}^{\infty} \delta(\tau)\ s(t-\tau)\,d\tau \qquad (2\text{-}3.24)$$

where as above, $\delta(\tau)$ is a Dirac-delta (or unit impulse) function defined by

$$\delta(\tau) = 0 \quad \text{for } \tau \neq 0$$

$$\int_{-t}^{t} \delta(\tau)\,d\tau = 1 \quad \text{for any } t > 0$$

In case of prediction, the desired output is the future value of the signal input, so that with $\alpha > 0$

$$s(t+\alpha) = \int_{-\infty}^{\infty} \delta(\tau+\alpha)\ s(t-\tau)\,d\tau \qquad (2\text{-}3.25)$$

So, the only change needed to go from a filtering problem to a prediction problem is to replace $\delta(\tau)$ by $\delta(\tau+\alpha)$. Also when $\alpha = 0$, the problem becomes one of filtering without prediction. The solution to a practical example of a combined filter-prediction problem will appear at the end of the next chapter.

2.4 POWER SPECTRUM OF A STATIONARY RANDOM FUNCTION:

Consider the autocorrelation function given by

2.4 contd.

$$C_{xx}(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\, x(t+\tau)\, dt$$

This can be reduced to the form -

$$C_{xx}(\tau) = \int_0^\infty \phi(\omega)\, \mathrm{Cos}(\omega\tau)\, d\omega \qquad (2\text{-}4.1)$$

$$= \frac{1}{2} \int_{-\infty}^{\infty} \phi(\omega)\, e^{i\omega\tau}\, d\omega \qquad (2\text{-}4.2)$$

and by the inversion formula

$$\phi(\omega) = \frac{2}{\pi} \int_0^\infty C_{xx}(\tau)\, \mathrm{Cos}(\omega\tau)\, d\tau \qquad (2\text{-}4.3)$$

$\phi(\omega)$ in this case is called the power spectrum of a stationary random function. The last two relations are Wiener-Khinchine relations.

Example: If the correlation function is given by a Gaussian curve,

$$C(\tau) = C(0)\, e^{-a^2\tau^2},$$

then the power spectrum is

$$\phi(\omega) = \frac{2}{\pi} C(0) \int_0^\infty \mathrm{Cos}(\omega\tau)\, e^{-a^2\tau^2}\, d\tau \qquad (2\text{-}4.4)$$

for $a \to \infty$, the power spectrum becomes constant and independent of frequency; $C(\tau)$ becomes zero i.e. x's are not correlated

2.4 contd.

at all. This most chaotic random function is called "white noise".

Relationship between the power spectrum of the input and the power spectrum of the output:

Let $x(t)$ be the stationary random input to a linear system and $y(t)$ be its output. If $h(t)$ is the response of the linear system to a unit impulse at $t = 0$, then the output in terms of the response function is -

$$y(t) = \int_{-\infty}^{t} x(\tau)\, h(t-\tau)\, d\tau \qquad (2\text{-}4.5)$$

The upper limit can be extended to $+\infty$ in which case,

$$y(t) = \int_{-\infty}^{\infty} x(\tau)\, h(t-\tau)\, d\tau = \int_{-\infty}^{\infty} x(t-u)\, h(u)\, du; \qquad (2\text{-}4.6)$$

If $\bar{F}(s)$ denotes the transfer function of the linear system, it is the Laplace transform of $h(t)$ and thus

$$F(i\omega) = \int_{0}^{\infty} e^{i\omega u}\, h(u)\, du \qquad (2\text{-}4.7)$$

Now, if $\phi(\omega)$ and $g(\omega)$ are the power spectra of input and output respectively, then it can be proved that -

$$g(\omega) = |F(i\omega)|^2\, \phi(\omega) \qquad (2\text{-}4.8)$$

which gives the power spectrum of the output in terms of the power spectrum of the input and the frequency response of a linear system.

2.5 MEASUREMENT OF THE STATISTICAL FUNCTIONS:

(i) The probability density functions:

For the measurement of a probability density function (p.d.f.), let us consider the sample time history record x(t) as illustrated in the figure below.

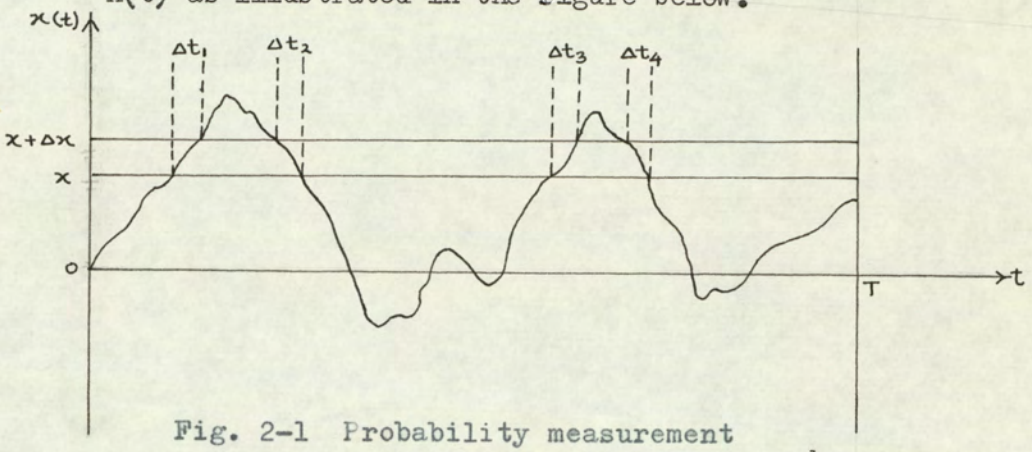


Fig. 2-1 Probability measurement

If say T_x is the total amount of time $\left(T_x = \sum_{i=1}^{k} \Delta t_i\right)$, for which x(t) takes values between x and x + Δx, and if T is the total observation time, the probability that x(t) assumes a value within the range x and x + Δx may be obtained by taking the ratio of T_x/T. With T approaching infinity, T_x/T approaches an exact probability description, i.e.

$$\text{Prob}\,[x < x(t) \leq x + \Delta x] = \lim_{T \to \infty} \frac{T_x}{T} \qquad (2\text{-}5.1)$$

The first order probability density function $p_1(x)$, for small Δx can be defined by the relation

2.5 contd.

(i) contd.

$$\text{Prob}\,[x < x(t) \leqslant x + \Delta x] \simeq p_1(x)\,\Delta x \qquad (2\text{-}5.2)$$

and more precisely,

$$p_1(x) = \lim_{\Delta x \to 0} \frac{\text{Prob}\,[x < x(t) \leqslant x + \Delta x]}{\Delta x}$$

$$= \lim_{\Delta x \to 0} \lim_{T \to \infty} \frac{1}{T}\left(\frac{T_x}{\Delta x}\right) \qquad (2\text{-}5.3)$$

(ii) Autocorrelation function:-

The autocorrelation function for random data describes the general dependence of the values of the data at one time on the values at another time. Let us consider the sample time history record as in the fig.(2-2). Then an estimate for the autocorrelation between the values of $x(t)$ at time t and $t+\tau$ may be obtained by taking the product of the two values and averaging over the observation time T. This average will approach an exact autocorrelation function as $T \to \infty$ i.e.

$$C_{xx}(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\, x(t+\tau)\, dt$$

2.5 contd.

(ii) contd.

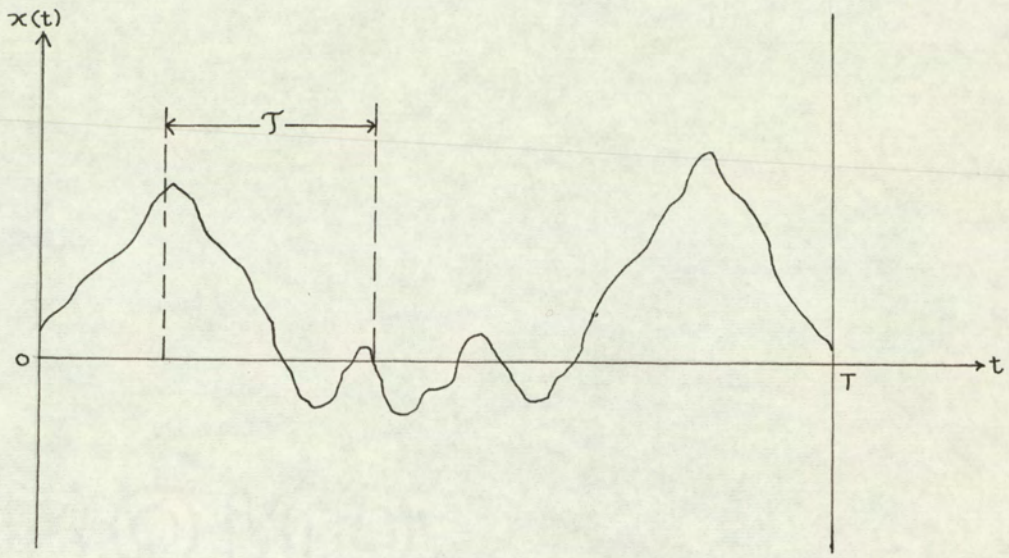


Fig. 2-2 Autocorrelation measurement.

$C_{xx}(\tau)$ has a maximum value at $\tau = 0$, and

$$C_{xx}(-\tau) = C_{xx}(\tau) \qquad (2\text{-}5.4)$$

$$C_{xx}(0) \geq |C_{xx}(\tau)| \text{ for all } \tau \qquad (2\text{-}5.5)$$

Also, the mean and the mean square values of x(t) in terms of the autocorrelation function are given by the relations

$$\overline{x(t)} = \sqrt{C_{xx}(\infty)} \qquad (2\text{-}5.6)$$

and $$\overline{x(t)^2} = C_{xx}(0) \qquad (2\text{-}5.7)$$

The major application of an autocorrelation function measurement of physical data is to establish the influence of values at any time over values at a future

2.5 contd.

(ii) contd.

time. The autocorrelation function for random data diminishes to zero for large time displacement whereas for the deterministic data it persists over all time displacements. Thus, autocorrelation measurement provides a powerful tool for distinguishing deterministic data from random data.

(iii) The joint probability density functions:-

The prbability that two random sample records will simultaneously assume values within some defined pair of ranges at any instant of time is given by their joint probability density function. For instance, considering the pair of time history records $x(t)$ and $y(t)$ as illustrated in fig.(2-3), the probability that $x(t)$ assumes a value within the range between x and $x + \Delta x$ while $y(t)$ simultaneously assumes a value within the range between y and $y + \Delta y$ is given by taking the ratio $T_{x,y}/T$, where $T_{x,y}$ is the total time for which $x(t)$ and $y(t)$ fall simultaneously inside the ranges $(x, x+\Delta x)$ and $(y, y+\Delta y)$, respectively during an observation time T. The ratio $T_{x,y}/T$ will approach an exact probability description as $T \to \infty$ i.e.

2.5 contd.

(iii) contd.

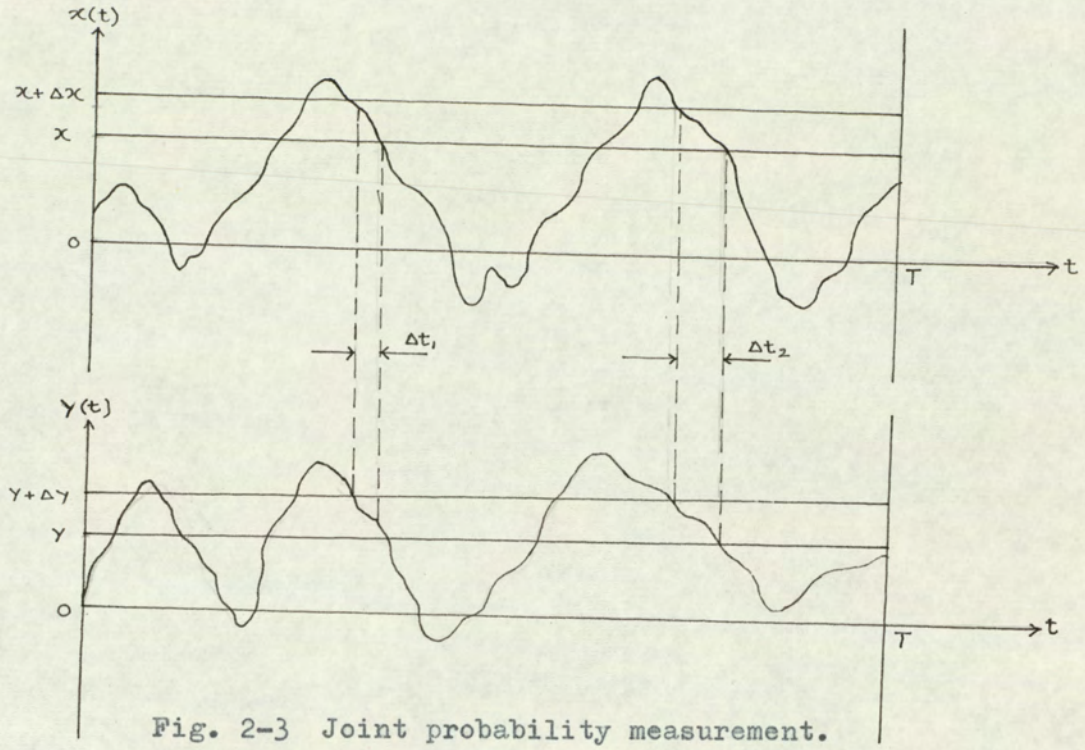


Fig. 2-3 Joint probability measurement.

$$\text{Prob}\,[x < x(t) \leq x + \Delta x;\ y < y(t) \leq y+\Delta y] = \lim_{T \to \infty} \frac{T_{x,y}}{T} \qquad (2\text{-}5.8)$$

For small values of Δx and Δy, the joint p.d.f. is given by

$$\text{Prob}\,[x < x(t) \leq x+\Delta x;\ y < y(t) \leq y+\Delta y] \simeq P(x,y)\ \Delta x\ \Delta y$$

or more precisely,

$$P(x,y) = \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \frac{\text{Prob}\,[x < x(t) \leq x+\Delta x;\ y < y(t) \leq y+\Delta y]}{(\Delta x)\ (\Delta y)}$$

$$= \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \lim_{T \to \infty} \frac{1}{T} \left[\frac{T_{x,y}}{(\Delta x)\ (\Delta y)} \right] \qquad (2\text{-}5.9)$$

(iv) Cross-correlation function:-

The general dependence of the values of one set of random data on the other can be described by the Cross-

2.5 contd.

(iv) contd.

correlation function for two sets of data. If we consider the pair of time history records x(t) and y(t) as illustrated in the figure below, an estimate of the Cross-correlation function of the values of x(t) at time t and y(t) at time $t + \tau$ may be obtained by taking the average

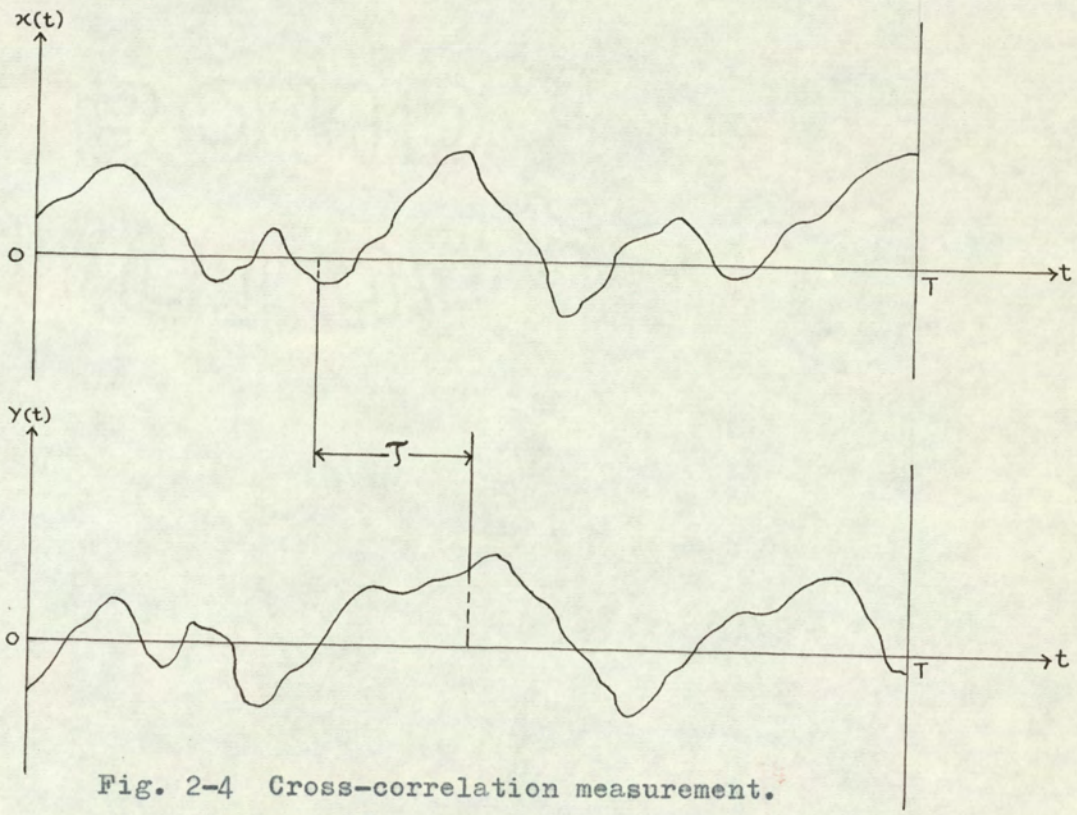


Fig. 2-4 Cross-correlation measurement.

product for the two values over the observation time T, as is done for autocorrelation functions. With $T \to \infty$, the resulting average product will approach an exact Cross-correlation function i.e.

2.5 contd.

(iv) contd.

$$C_{xy}(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\, y(t+\tau)\, dt.$$

When $C_{xy}(\tau) = 0$, $x(t)$ and $y(t)$ are said to be uncorrelated.

CHAPTER 3.

LINEAR LEAST SQUARE SMOOTHING AND PREDICTION THEORY.

This theory was originated by Wiener and Kolmogoroff during the second world war. Wiener's theory in its original form, involves rather difficult mathematical treatment and hence is difficult for Engineers and Scientists who may not possess such a deep understanding of mathematical techniques. However, Engineers soon realised its usefulness in the systematic designing of filters and have tried to make it more practicable by generalising, interpreting, extending and modifying the original W-K theory. Amongst the principal contributors were Bode and Shannon, Zadeh and Ragazzini, Booton, Darlington, Lanning and Battin, Bendat.

The W-K theory was formulated for the purpose of optimal separation of a signal which has been perturbed by the addition of noise or random process, by the use of linear filter. The input data is a time function $f(t)$ which is a combination of a signal $s(t)$ and noise $n(t)$ i.e.

$$f(t) = s(t) + n(t)$$

where $s(t)$ and $n(t)$ are typical members drawn from the ensembles of these functions, which have certain known statistical characteristics. The central problem then is one of the derivation of an estimate $f_p(t)$ of $s(t)$ in the form,

$$f_p(t) = s(t+\alpha)$$

where for $\alpha > 0$, $f_p(t)$ gives a smoothed prediction of $s(t)$.

This prediction will involve some errors in which case,

$$f_p(t) = s(t+\alpha) + e(t)$$

e(t) being the error term. Then, the objective of the W-K theory is to determine specific conditions and operations which will minimise e(t).

The assumptions which limit the range of application of the W-K theory are:

(a) that the time series represented by the signal s(t) and the noise n(t) are stationary;

(b) that the optimum system is characterised as having a minimum mean square ensemble system error, and

(c) that the system is a constant parameter linear device.

3.1 Wiener's theory of Prediction:-

Wiener's theory may be considered to be derived from the techniques for detecting hidden periodicities in a time series, using the method of the periodogram, as developed by Sir Arthur Schuster. The full appreciation of the theory, however, demands a knowledge of several special mathematical fields, including Lebesgue integration, the Stieltje's integral, Fourier transforms, Generalised Harmonic analysis, Orthogonal functions (in particular, Laguerre functions), Integral equations and the theory of functions of a complex variable.

Wiener's work, as one can see, was mainly directed towards predicting the path of an enemy aircraft and for the

3.1 contd.

purposes of fire control. Since the war, it has been modified and extended in various ways so as to constitute a major tool of considerable potential value for research into effective methods of prediction, both in engineering and economics.

It will be sufficient to discuss the theory for a single time series. The problem here is - for a given function $f(t)$, find an operator which when operating upon its past and present values yields $f(t+\alpha)$. For $\alpha > 0$, $f(t+\alpha)$ gives a smoothed prediction of $f(t)$. As stated above, it is required that the operator should be linear, invariant with respect to the choice of the origin of time and dependent only on the past and present values of $f(t)$. It is assumed that the prediction should be the "best PREDICTABLE" in the "mean square" sense, that is to say, if $f_p(t)$ is the predicted value of $f(t+\alpha)$, then

$$\lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} [f(t+\alpha) - f_p(t)]^2 \, dt \qquad (3\text{-}1.1)$$

is to be a minimum.

Wiener's theory assumes the existence of a function $K(\tau)$, such that the Stieltje's integral

$$\int_0^\infty f(t-\tau) \, dK(\tau)$$

3.1 contd.

is a predictable value, $f_p(t)$, and that $K(\tau)$ can be so determined as to minimise the expression (3-1.1). This expression can thus be regarded as a measure of the extent to which the operator $K(\tau)$ fails to predict the value of $f(t+\alpha)$.

The concepts chiefly employed in the solution of the problems are -

(a) the autocorrelation function

$$C(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} f(t+\tau)\, f(t)\, dt$$

which is analogous to the autocorrelation coefficient

$$\lim_{N \to \infty} \frac{1}{2N+1} \sum_{k=-N}^{N} x_{k+j}\, x_k$$

in its normalised form, and

(b) The Fourier transform pais $f(t)$ and $g(\omega)$ related by the equations

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(\omega)\, e^{i\omega t}\, d\omega \quad \text{and} \quad g(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)\, e^{-i\omega t}\, dt$$

The condition for the minimum of the expression (3-1.1) may be obtained by adding to $K(\tau)$ the expression $t(\delta K(\tau))$, differentiating with respect to t, equating the derivative to zero, and then letting t tend to zero. (This

3.1 contd.

technique belongs to the calculus of variations).

This leads to the condition

$$C(\alpha+\tau) = \int_0^\infty C(\tau-t)\, dK(t), \text{ for } \tau > 0 \qquad (3\text{-}1.2)$$

Thus, if the autocorrelation function is known, then $K(t)$ can be determined by solving the integral equation (3-1.2). Wiener's method of solving this equation is to introduce the Fourier transform $\phi(\omega)$ of $C(t)$, that is to say,

$$C(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(\omega)\, e^{i\omega t}\, d\omega \quad \text{and}$$

$$\phi(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} C(t)\, e^{-i\omega t}\, dt$$

If we know the poles and zeros (for definition see ch.6) of $\phi(\omega)$ regarded as a function of the complex variable ω, then, in the case where $\phi(\omega)$ is an even rational function, it can be expressed in the form

$$\phi(\omega) = \frac{A^2 \prod_{k=1}^{n} (\omega-\omega_k)(\omega-\bar{\omega}_k)}{\prod_{j=1}^{m} (\omega-\omega_j)(\omega-\bar{\omega}_j)} \qquad (3\text{-}1.3)$$

where $\bar{\omega}_k$ is the complex conjugate of ω_k

3.1 contd.

$\bar{\omega}_j$ is the complex conjugate of ω_j

and A is real, and the imaginary parts of each ω_k and ω_j are positive. Under these conditions $\phi(\omega)$ can be factorised, that is to say,

$$\phi(\omega) = \psi_1(\omega)\ \psi_2(\omega) \quad \text{where}$$

$$\psi_1(\omega) = \frac{A \prod_1^n (\omega - \omega_k)}{\prod_1^m (\omega - \omega_j)} \quad \text{and} \quad \psi_2(\omega) = \frac{A \prod_1^n (\omega - \bar{\omega}_k)}{\prod_1^m (\omega - \bar{\omega}_j)}$$

From the function $\psi_1(\omega)$, the inverse $\psi(t)$ can be obtained by the relationship

$$\psi(t) = \lim_{B \to \infty} \frac{1}{\sqrt{2\pi}} \int_{-B}^{B} \psi_1(\omega)\ e^{i\omega t}\ d\omega \qquad (3\text{-}1.4)$$

Wiener then shows that $K(\tau)$ can be obtained by solving the equation

$$\int_{-\infty}^{t} \psi(t-\tau)\ dK(\tau) = \psi(t+\alpha) \qquad (3\text{-}1.5)$$

To solve (3-1.5), we write

$$\int_0^{\infty} e^{i\omega t}\ dK(t) = k(\omega)$$

and, taking Fourier transforms of both sides of (3-1.5),

$$\int_0^{\infty} \psi(\alpha+t) e^{-i\omega t}\ dt = \psi_1(\omega) k(\omega)$$

3.1 contd.

so that

$$k(\omega) = \int_0^\infty \psi(\alpha+t)\ e^{-i\omega t}\ dt/\psi_1(\omega)$$

or

$$k(\omega) = \frac{1}{\sqrt{2\pi}\ \psi_1(\omega)} \int_0^\infty e^{-i\omega t}\ dt \int_{-\infty}^\infty \psi_1(\omega)\ e^{i\omega(t+\alpha)}\ d\omega$$

knowing $k(\omega)$, the solution is easily completed.

The complete process may be illustrated by a simple example. Let us assume that the autocorrelation function $C(\tau)$ of the time series $f(t)$ has a Fourier transform

$$\phi(\omega) = \frac{1}{1+\omega^2}$$

Now,

$$\frac{1}{1+\omega^2} = \frac{1}{\omega - i} \cdot \frac{1}{\omega + i}$$

so that we identify $\psi_1(\omega) = \frac{1}{\omega - i}$. Hence

$$k(\omega) = \frac{\omega - i}{\sqrt{2\pi}} \int_0^\infty e^{-i\omega t}\ dt \int_{-\infty}^\infty \frac{e^{i\omega(t+\alpha)}}{\omega - i}\ d\omega$$

Now applying the calculus of residues,

$$\int_{-\infty}^\infty \frac{e^{i\omega(t+\alpha)}}{\omega - i}\ d\omega = 2\pi i \cdot \text{residue at } \omega = i.$$

To determine this residue, we write

3.1 contd.

$$\frac{e^{i\omega(t+\alpha)}}{\omega-i} = \frac{e^{i(t+\alpha)(\omega-i+i)}}{\omega-i} = \frac{e^{i(t+\alpha)(\omega-i)} \cdot e^{i(t+\alpha)i}}{\omega-i}$$

$$= \frac{e^{-(t+\alpha)}}{\omega-i}\left[e^{i(t+\alpha)(\omega-i)} \right]$$

Hence the residue is $e^{-(t+\alpha)}$. Therefore,

$$k(\omega) = \frac{\omega-i}{\sqrt{2\pi}} \int_0^\infty e^{-i\omega t}\, 2\pi i e^{-(t+\alpha)}\, dt$$

$$= i(\omega-i)e^{-\alpha} \cdot \sqrt{2\pi} \int_0^\infty e^{-i\omega t}\, e^{-t}\, dt$$

$$= i(\omega-i)\, e^{-\alpha} \cdot \sqrt{2\pi} \int_0^\infty e^{-t(1+i\omega)}\, dt$$

$$= i(\omega-i)e^{-\alpha} \cdot \sqrt{2\pi} \left[\frac{-1}{1+i\omega}\, e^{-t(1+i\omega)} \right]_0^\infty$$

$$= e^{-\alpha}\sqrt{2\pi}$$

Now, the predicted value of $f(t+\alpha)$ is

$$\int_0^\infty f(t-\tau)\, dK(\tau) \text{ where}$$

$$\frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-i\omega t}\, dK(t) = e^{-\alpha}$$

At this point it is convenient to introduce a new function $h(t)$ defined by $dK(t) = h(t)\, dt$. So that we have

3.1 contd.

$$\frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-i\omega t} h(t)\, dt = e^{-\alpha}$$

and the predicted value of $f(t+\alpha)$ is

$$\int_0^\infty f(t-\tau)\, h(\tau)\, d\tau \qquad (A)$$

Now, the expression (A) is a convolution integral. Its transform is

$$F(\omega)\, H(\omega)$$

where $F(\omega)$ is the Fourier transform of $f(t)$ and $H(\omega)$ is the Fourier transform of $h(\tau)$. Hence the predicted value of $f(t+\alpha)$ is such that its Fourier transform is

$$F(\omega)\, H(\omega),$$

But, $H(\omega) = e^{-\alpha}$ and hence, the predicted value of $f(t+\alpha)$ is the inverse transform of $F(\omega)\, e^{-\alpha}$, which is $f(t)\, e^{-\alpha}$.

In this example, we have taken as our starting point the function

$$\phi(\omega) = \frac{1}{1+\omega^2}$$

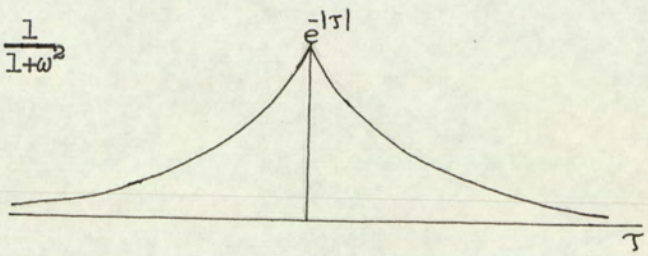
This, in fact, is the Fourier transform of a function of time equal to

$$\sqrt{\frac{\pi}{2}} \cdot e^{-|t|}$$

Thus the normalised autocorrelation function $e^{-|\tau|}$ leads to the transform

3.1 contd.

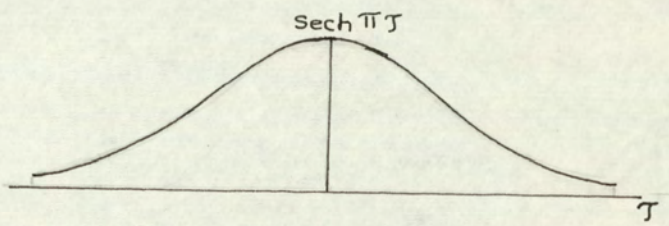
$$\phi(\omega) = \sqrt{\frac{2}{\pi}} \cdot \frac{1}{1+\omega^2}$$



The graph of $e^{-|\tau|}$ has properties consistent with those of an autocorrelation function.

Another example, starting with an autocorrelation function of the form of Sech $\pi\tau$, leads to the Fourier transform

$$\sqrt{\frac{2}{\pi}} \cdot \frac{1}{1+\omega^4}$$



In each of these two examples, the prediction problem can be solved by the analytical procedure indicated.

In practice, the autocorrelation function must be determined empirically. Having obtained it, we must then see whether it can be represented with sufficient accuracy by a linear combination of functions, such as

$$\frac{1}{1+\omega^2}, \quad \frac{1}{1+\omega^4}$$

and other rational fraction functions of ω. In so far as this can be done, the Wiener method leads to a result.

Wiener gives some further examples of the function $\phi(\omega)$, namely

3.1 contd.

$$\frac{1}{(1+\omega^2)^2}\ , \quad \frac{1+\omega^2}{1+\omega^4} \quad \text{and} \quad \frac{1}{(1+\frac{\omega^2}{n})^n}$$

which is an approximation to $e^{-\omega^2}$, the transform of $e^{-\tau^2}$. He also hints at the possible expression of the auto-correlation function in such a way that orthogonal functions, and, especially Laguerre functions, may be used, but he does not enlarge upon this.

3.2 A GENERAL THEORY OF LINEAR PREDICTION AND FILTERING.

In the generalised form, the collections of input signal $\{s(t)\}$ and noise $\{n(t)\}$ are considered to be non-stationary and the linear system to be time varying with its characteristic weighting function $h(t,\tau)$, denoting the response of the time-varying system at time t to a unit impulse applied at time $t-\tau$. If for an arbitrary input time, the input combination

$$f(t) = s(t) + n(t)$$

passes through this time-varying system for a finite sampling time T, then the actual output response is given by

$$g(t) = \int_0^T h(t,\tau)\ f(t-\tau)d\tau \qquad (3\text{-}2.1)$$

Now, if the desired output is $s(t+\alpha)$, then the system error e(t) is given by

3.2 contd.

$$e(t) = \int_0^T h(t,\tau)\, f(t-\tau)\, d\tau - s(t+\alpha) \qquad (3\text{-}2.2)$$

The weighting function $h(t,\tau)$ is to be determined such that the mean square error is minimum. From (3-2.2)

$$e^2(t) = \left[\int_0^T h(t,\tau)\, f(t-\tau) d\tau = s(t+\alpha) \right] \left[\int_0^T h(t,\tau') f(t-\tau') d\tau' - s(t+\alpha) \right]$$

$$= \int_0^T \int_0^T h(t,\tau)\, h(t,\tau')\, f(t-\tau) f(t-\tau') d\tau\, d\tau'$$

$$- s(t+\alpha) \left[\int_0^T h(t,\tau)\, f(t-\tau)\, d\tau \right.$$

$$\left. + \int_0^T h(t,\tau')\, f(t-\tau')\, d\tau' \right] + s(t+\alpha)\, s(t+\alpha); \qquad (3\text{-}2.3)$$

This is because, the value of a definite integral is independent of the variable of integration τ or τ'. (3-2.3) is then simplified to

$$e^2(t) = \int_0^T \int_0^T h(t,\tau)\, h(t,\tau')\, f(t-\tau)\, f(t-\tau') d\tau\, d\tau'$$

$$- 2 \int_0^T h(t,\tau')\, f(t-\tau')\, s(t+\alpha)\, d\tau' + s(t+\alpha)\, s(t+\alpha) \qquad (3\text{-}2.4)$$

By averaging (3-2.4) we get

3.2 contd.

$$\overline{e^2(t)} = \int_0^T \int_0^T h(t,\tau)\, h(t,\tau')\, \overline{f(t-\tau)\, f(t-\tau')}\, d\tau\, d\tau'$$

$$- 2 \int_0^T h(t,\tau')\, \overline{f(t-\tau')\, s(t+\alpha)}\, d\tau' + \overline{s(t+\alpha)\, s(t+\alpha)}$$

But, $\overline{f(t-\tau)\, f(t-\tau')}$ = the autocorrelation function $C_{ff}(t-\tau, t-\tau')$

$\overline{f(t-\tau')s(t+\alpha)}$ = the cross-correlation function $C_{fs}(t-\tau', t-\alpha)$

$\overline{s(t+\alpha)\, s(t+\alpha)}$ = the autocorrelation function $C_{ss}(t+\alpha,\ t+\alpha)$

and therefore,

$$\overline{e^2(t)} = \int_0^T \int_0^T h(t,\tau)\, h(t,\tau')\, C_{ff}(t-\tau,\ t-\tau')\, d\tau\, d\tau'$$

$$- 2 \int_0^T h(t,\tau')\, C_{fs}(t-\tau',\ s+\alpha)\, d\tau' + C_{ss}(s+\alpha, s+\alpha) \quad (3\text{-}2.5)$$

Consider now, that the response function $h(t,\tau)$ is varied by $\eta h(t,\tau)$, η being an arbitrary real constant. Then the corresponding variation to the first order in $\overline{e^2(t)}$ is given by

$$\overline{\eta e^2(t)} = 2 \int_0^T \eta h(t,\tau') \left\{ \int_0^T h(t,\tau)\, C_{ff}(t-\tau,\ t-\tau')\, d\tau \right.$$

$$\left. - C_{fs}(t-\tau',\ t+\alpha) \right\} d\tau'$$

This variation must vanish, if $h(t,\tau)$ is the response function of the optimum filter. It will be so, when

3.2 contd.

$$\int_0^T h(t,\tau)\, C_{ff}(t-\tau, t-\tau')\, d\tau = C_{fs}(t-\tau', t+\alpha); \quad (0 \leqslant \tau' \leqslant T). \quad (3\text{-}2.6)$$

(3-2.6) is the modified form of the Winer-Hopf equation. Substituting from (3-2.6) in (3-2.5) we get

$$\overline{e^2(t)}_{min} = C_{ss}(t+\alpha,\ t+\alpha) - \int_0^T h(t,\tau)\, C_{fs}(t-\tau, t+\alpha)\, d\tau'; \quad (3\text{-}2.7)$$

The main problem in the applicability of this theory is the solving of the appropriate integral equation (3-2.6) using particular correlation functions involved.

3.3 BODE AND SHANNON'S INTERPRETATION OF WIENER'S THEORY

In Winer's theory, as opposed to the general theory, the problem is one of determining the optimum constant parameter linear system (or filter), for stationary random inputs and infinite operating times.

Let the system be characterised by its weighting function $h(\tau)$ and the total input to it be the sum of independent signal and noise terms. Now, if $Y(i\omega)$ is the transfer function of the linear system then at a particular frequency ω_1, the contribution to the error due to noise is given by

3.3 contd.

$$|Y(i\omega_1)|^2 \phi_n(\omega_1) \qquad (3\text{-A})$$

where $\phi_n(\omega_1)$ is the average power due to noise at frequency ω_1 and

$$Y(i\omega_1) = \int_{-\infty}^{\infty} h(\tau)\, e^{i\omega\tau}\, d\tau$$

Truxal. P.463.
Exp. arises because desired output is input x timedelay.

The contribution to the error due to a signal is given by

$$|Y(i\omega_1) - e^{i\alpha\omega_1}|^2 \phi_s(\omega_1) \qquad (3\text{-B})$$

where $\phi_s(\omega_1)$ denotes the average power due to the signal at frequency ω_1. Here, the component of frequency ω_1 is advanced in phase by $\alpha\omega_1$.

At frequency ω_1, the total mean square error will be the sum of (3-A) and (3-B) i.e.

$$E\omega_1 = |Y(i\omega_1)|^2 \phi_n(\omega_1) + |Y(i\omega_1) - e^{i\alpha\omega_1}|^2 \phi_s(\omega_1); \quad (3\text{-}3.1)$$

and for all frequencies, the total mean square error will be given by

$$E = \int_{-\infty}^{\infty} \left[|Y(i\omega)|^2 \phi_n(\omega) + |Y(i\omega) - e^{i\alpha\omega}|^2 \phi_s(\omega) \right] d\omega \quad (3\text{-}3.2)$$

E is to be minimised with a proper choice of $Y(i\omega)$. Also, $Y(i\omega)$ should be physically realisable.

The condition that $Y(i\omega)$ should be a physically realisable transfer function, creates the main difficulty in minimising E. The condition of physical realisability

3.3 contd.

of the filter is thus waived temporarily and all choices of $Y(i\omega)$ are allowed with the result that $h(\tau)$ is not required to be zero for $\tau < 0$. However, this physically impossible condition must be corrected later. We are assuming that the entire function $f(t) = s(t) + n(t)$ from $t = -\infty$ to $t = +\infty$ is available for use in prediction. The minimisation of (3-3.2) over all possible $Y(i\omega)$ yields a "theoretical" frequence response function.

In (3-3.2) suppose -

$$Y(i\omega) = C(\omega)\, e^{iB(\omega)}$$

where $C(\omega)$ and $B(\omega)$ are real functions of ω. Then

$$E = \int_{-\infty}^{\infty} \left[C^2(\omega)\, \phi_n(\omega) + \left\{ C^2(\omega) + 1 - 2C(\omega)\, \mathrm{Cos}(B(\omega) - \alpha\omega) \right\} \phi_s(\omega) \right] d\omega$$

It will be clear from this, that in order to minimise E, the best choice of $B(\omega)$ is $B(\omega) = \alpha\omega$, which maximises $\mathrm{Cos}(B(\omega) - \alpha\omega)$ and then

$$E = \int_{-\infty}^{\infty} \left[C^2(\omega) \left\{ \phi_s(\omega) + \phi_n(\omega) \right\} - 2C(\omega)\, \phi_s(\omega) + \phi_s(\omega) \right] d\omega$$

Completing the square in $C(\omega)$ by adding and subtracting,

$$\frac{\phi_s{}^2(\omega)}{\phi_s(\omega) + \phi_n(\omega)},$$

we get

3.3 contd.

$$E = \int_{-\infty}^{\infty} \left[C^2(\omega)\left\{ \phi_s(\omega)+\phi_n(\omega) \right\} - 2C(\omega)\ \phi_s(\omega) + \frac{\phi_s^{\,2}(\omega)}{\phi_s(\omega)+\phi_n(\omega)} \right.$$

$$\left. - \frac{\phi_s^{\,2}(\omega)}{\phi_s(\omega)+\phi_n(\omega)} + \phi_s(\omega) \right] d\omega$$

$$= \int_{-\infty}^{\infty} \left\{ \left[C(\omega)\ \sqrt{\phi_s(\omega)+\phi_n(\omega)} - \frac{\phi_s(\omega)}{\sqrt{\phi_s(\omega)+\phi_n(\omega)}} \right]^2 + \frac{\phi_s(\omega)\ \phi_n(\omega)}{\phi_s(\omega)+\phi_n(\omega)} \right\} d\omega$$

The bracketed term is the square of a real number and therefore positive or zero. From the last expression, for minimum E, the bracketed term must be zero i.e.

$$C(\omega) = \frac{\phi_s(\omega)}{\phi_s(\omega)+\phi_n(\omega)} \qquad (3\text{-}3.3)$$

and hence

$$Y(i\omega) = \frac{\phi_s(\omega)}{\phi_s(\omega)+\phi_n(\omega)}\ .e^{i\alpha\omega} \qquad (3\text{-}3.4)$$

The mean square error then becomes

$$E = \int_{-\infty}^{\infty} \frac{\phi_s(\omega)\ \phi_n(\omega)}{\phi_s(\omega)+\phi_n(\omega)}\, d\omega \qquad (3\text{-}3.5)$$

and the best weighting function in this case -

$$h(t) = \mathcal{F}^{-1}[Y(i\omega)]$$

3.3 contd.

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\phi_s(\omega)}{\phi_s(\omega)+\phi_n(\omega)} \cdot e^{i\omega(t+\alpha)} d\omega \qquad (3\text{-}3.6)$$

Example:

Consider a signal s(t) with spectral density

$$\phi_s(\omega) = \frac{1}{1+\omega^2}$$

The signal is additively mixed with white noise, which is charqcterised by a constant power spectral density

$$\phi_n(\omega) = N, \text{ a constant}$$

Let $\alpha = 0$, then for an optimum filter -

$$Y(i\omega) = \frac{\phi_s(\omega)}{\phi_s(\omega)+N} = \frac{\frac{1}{1+\omega^2}}{\frac{1}{1+\omega^2} + N}$$

$$= \frac{1}{1+N(1+\omega^2)} = \frac{\frac{1}{N}}{(1+\frac{1}{N}) + \omega^2}$$

The corresponding weighting function will be

$$h(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y(i\omega)\, e^{i\omega t}\, d\omega$$

$$= \frac{1}{2[(N(N+1)]^{\frac{1}{2}}} \cdot \exp\left[- |t| \left\{ \frac{N+1}{N} \right\}^{\frac{1}{2}} \right]$$

However, this particular weighting function h(t) is not physically realisable.

In general, h(t) in (3-3.6) will extend from $t = -\infty$ to $t = +\infty$ and does not represent the impulsive response of a physical filter. However, it is a perfectly

3.3 contd.

good weighting function and if one could wait until all the function $s(t) + n(t)$ is available, it will be a proper one to apply in estimating $s(t+\alpha)$. This means that the weighting function $h(\tau)$, can be obtained in physical filter if sufficient delay is allowed so that $h(\tau)$ is substantially zero for the future. Although $Y(i\omega)$ in (3-3.4) is non-physical, $Y(i\omega)e^{-iB\omega}$ will be physical, or nearly so, if B is taken sufficiently large.

The optimum physically realisable filter is obtained by first designing a shaping filter (see next example), for input $s(t)+n(t)$. Let $Y_1(i\omega)$ be the transfer function of this filter and $Y_1^{-1}(i\omega)$ be its inverse. If we pass $s(t)+n(t)$ through $Y_1^{-1}(i\omega)$, the output will be a white noise. Both $Y_1(i\omega)$ and $Y_1^{-1}(i\omega)$ are physically realisable. The knowledge of the input and output of this filter are equivalent. The best linear operation on the output will give the same prediction as the corresponding best linear operation on the input $s(t)+n(t)$.

The following steps are followed in constructing the optimum physically realisable filter

(a) Construct

$$Y_2(i\omega) = Y(i\omega)\, Y_1(i\omega) \qquad (3\text{-}3.7)$$

where $Y(i\omega)$ is given by (3-3.4).

For a prediction interval α, this can be written as

3.3 contd.

(a) contd.

$$Y_2(i\omega) = e^{i\omega\alpha}\, Y_2^*(i\omega) \qquad (3\text{-}3.8)$$

where $Y_2^*(i\omega)$ denotes the sum of the partial fractions.

(b) Denote by $Y_2^*(i\omega)$, the sum of those terms of $Y_2^*(i\omega)$ which are physically realisable.

(c) Calculate

$$h_3(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} Y_3^*(i\omega)\, e^{i\omega(t+\alpha)}\, d\omega \qquad (3\text{-}3.9)$$

Then $h_3(t)$ for $t > 0$ represents the optimum physically realisable weighting function to use on the white noise output created by the shaping filter $Y_1^{-1}(i\omega)$.

(d) The corresponding transfer function for the white noise is given by

$$Y_3(i\omega) = \int_0^{\infty} h_3(t)\, e^{-i\,\omega t}\, dt \qquad (3\text{-}3.10)$$

(e) Lastly, the transfer function of the optimum filter is given by

$$Y_4(i\omega) = Y_3(i\omega)\, Y_1^{-1}(i\omega) \qquad (3\text{-}3.11)$$

Thus, the optimum least square prediction is obtained by passing $s(t)+n(t)$ through a filter whose transfer function is given by

$$Y_4(i\omega) = Y_3(i\omega)\, Y_1^{-1}(i\omega).$$

One can best illustrate this method by an example:

3.3 contd.

Example:

Consider again, the signal s(t) with power spectral density function

$$\phi_s(\omega) = \frac{1}{1 + \omega^2}$$

and assume the noise is white which is characterised by a constant power spectrum

$$\phi_n(\omega) = N, \text{ a constant.}$$

The problem is to find a transfer function of the optimum constructible filter for a prediction time interval α.

From (3-3.4), the optimum theoretical frequency response function is given by

$$Y(i\omega) = \frac{\phi_s(\omega)}{\phi_s(\omega) + \phi_n(\omega)} \cdot e^{i\omega\alpha} = \frac{e^{i\omega\alpha}}{(1+N) + N\omega^2} \qquad (3\text{-}3.12)$$

The shaping filter is determined from

$$|Y(i\omega)|^2 = \phi_s(\omega) + \phi_n(\omega) = \frac{(1+N) + N\omega^2}{1+\omega^2} \qquad (3\text{-}3.13)$$

and hence

$$Y_1(i\omega) = \frac{(1+N)^{\frac{1}{2}} + iN^{\frac{1}{2}}\omega}{1 + i\omega} \qquad (3\text{-}3.14)$$

From equation (3-3.7)

$$Y_2(i\omega) = Y(i\omega)\, Y_1(i\omega)$$

$$= \frac{e^{i\omega\alpha}}{(1+N) + N\omega^2} \cdot \frac{(1+N)^{\frac{1}{2}} + iN^{\frac{1}{2}}\omega}{1 + i\omega}$$

3.3 contd.

$$= \frac{e^{i\omega\alpha}}{[(1+N)^{\frac{1}{2}} - iN\,\tfrac{1}{2}\omega](1+i\omega)}$$

$$= e^{i\omega\alpha}\left[\frac{A}{(1+N)^{\frac{1}{2}} - iN^{\frac{1}{2}}\omega} + \frac{B}{1+i\omega}\right] \qquad (3\text{-}3.15)$$

$$= e^{i\omega\alpha}\; Y_2^*(i\omega) \qquad (3\text{-}3.16)$$

where

$$A = \frac{N^{\frac{1}{2}}}{N^{\frac{1}{2}} + (1+N)^{\frac{1}{2}}} = \frac{1}{1+\mu} \quad \text{by taking } \mu = \frac{(1+N)^{\frac{1}{2}}}{N^{\frac{1}{2}}}$$

and $$B = \frac{1}{N^{\frac{1}{2}} + (1+N)^{\frac{1}{2}}} = \frac{1}{N^{\frac{1}{2}}(1+\mu)} \qquad (3\text{-}3.17)$$

Only the second term of $Y_2^*(i\omega)$ is physically realisable and hence using (3-3.9)

$$h_3(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{B}{1+i\omega}\; e^{i\omega(t+\alpha)}\, d\omega$$

$$= Be^{-(t+\alpha)} \qquad \text{for } t > 0 \qquad (3\text{-}3.18)$$

and then

$$Y_3(i\omega) = \int_0^{\infty} B\, e^{-(t+\alpha)} \cdot e^{-i\omega t}\, dt$$

$$= \frac{B\, e^{-\alpha}}{1+i\omega} \qquad (3\text{-}3.19)$$

Thus, the optimum constructible filter will have the transfer function given by

3.3 contd.

$$Y_4(i\omega) = Y_3(i\omega)\, Y_1^{-1}(i\omega) = \frac{B\, e^{-\alpha}}{(1+N)^{\frac{1}{2}}+iN^{\frac{1}{2}}\omega} \qquad (3\text{-}3.20)$$

Substituting for B from (3-3.17)

$$Y_4(i\omega) = \frac{(N(1+\mu))^{-1}}{\mu+i\omega} \cdot e^{-\alpha} .$$

DISCUSSION

In this chapter we have reviewed the linear prediction and filter theory as developed by Winer, Bode and Shannon and the general theory developed by Booton and Bendat. The theory is of particular importance in designing optimum engineering systems, either for projecting of past information into the future, or in the recovery of desired signals that are distorted by random noise disturbances. One can see that the applications of the theory exist not only in communication fields but also in meteorological forecasting and economic analysis.

As already stated, the Wiener's theory is based on the three main assumptions and before using it one must carefully consider each of them with regard to the particular smoothing or prediction problem involved.

The general theory provides a method for optimisation over the class of time-variable linear system with a non-stationary input and the more realistic finite sampling time instead of the smaller subclass of constant-parameter linear system with station-

ary input and the non-existent infinite sampling periods. However, one of the limitations of the theory, which still reamins is that when we minimise the mean-square error, we are, in effect, weighting large errors more heavily than the small errors.

CHAPTER 4

EXPONENTIALLY WEIGHTED PREDICTION AND OTHER METHODS.

The first mention of the technique of exponential smoothing seems to be by C.C.Holt in 1957. Several different writers have discussed methods of exponential smoothing since then with different applications of the technique in mind. A readable account is due to R.G.Brown (1959) and a description of a more ambitious scheme incorporating seasonals by Winters (1960). The theoretical papers by Professor Bernard (1959) and Box and Jenkins (1962) throw considerable light on the validity of the technique. D.R.Cox (1961) in his theoretical paper considers exponential weighting when no provision is made for trend.

In this chapter, we shall briefly review some of the methods of exponentially weighted prediction and certain other methods.

4.1 Exponentially weighted moving averages (e.w.m.a.):

Probably one of the earliest and simplest methods of prediction of future values in any time series is by moving averages. It can be seen that moving averages do not have adaptive qualities, which are essential for many of the prediction problems. It does not account for trend, seasonal variations etc. and hence is seldom used.

Given ... x(t-1), x(t), a stochastic process in discrete time, a more realistic approach to estimate

4.1 contd.

$\hat{x}_e(t)$ - a value for the period (t+1) - would be to use the formula

$$\hat{x}_e(t) = a_1 x(t) + a_2 x(t-1) + a_3 x(t-2) + \ldots \qquad (4\text{-}1.1)$$

where $a_i > a_i + 1$ and

$\Sigma a_i = 1$; a_i's are called weighting factors.

However, Holt suggested a method of exponentially weighted moving averages (e.w.m.a.) which is a special case of (4-1.1). Here a_i form a geometric progression, a_2 being the constant fraction of a_1, a_3 of a_2 and so on.

i.e. $\frac{a_2}{a_1} = \frac{a_3}{a_2} = \frac{a_4}{a_3} = \ldots = \text{const.}\rho$

and $\rho < 1$. This reduces (4-1.1) to the form

$$\hat{x}_e(t) = a_1 x(t) + (1-a_1)\,\hat{x}_e(t-1) \qquad (4\text{-}1.2)$$

$$= \hat{x}_e(t-1) + a_1[x(t) - \hat{x}_e(t-1)] \qquad (4\text{-}1.3)$$

where a_1 is the exponentially weighted factor and $0 \leq a_1 \leq 1$.

Substituting for $\hat{x}_e(t-1)$ from (4-1.2) in (4-1.3) and repeating it further it can be shown that in general, the value of the process at time (t+h) is given by the predictor $\hat{x}_e(t,h;a)$ where

$$\hat{x}_e(t, h;a) = a \sum_{n=0}^{\infty} (1-a)^n x(t-n) \qquad (4\text{-}1.4)$$

$$= a\,x(t) + (1-a)\,\hat{x}_e(t-1, h;a) \qquad (4\text{-}1.5)$$

4.1 contd.

for $|a| < 1$.

That is to say the estimated value for the period (t+1) is equal to the estimated value for the period (t) plus the portion of the error between the immediately previous value and the immediately previous estimate.

The main problem here is that of choosing an appropriate value for the portion of error a_1 which is also known as "smoothing constant". However, one can start with some arbitrary value, say $a_1 = 0.4$ and observe the central tendency. a_1 should be increased to follow the central tendency more closely and decreased to suppress more of the random variations. It is found that for slowly increasing or decreasing trends, a value of a_1 equal to 0.1 or 0.2 gives a useful prediction and for a sharply increasing or decreasing trend, a much higher value, say $a_1 = 0.8$ or 0.9 should be chosen.

Though the problem of choosing a smoothing constant is great, the e.w.m.a. has a tremendous advantage over many other methods for the obvious reason of being extremely easy to compute. Also, the method is extremely valuable when large number of predictions are required. The predictor is adaptive to any change in mean, but a disadvantage is that these predictors are not optimal for stationary time series.

4.1 contd.

i) Allowance for trend:-

Trend is a rate of change. For increasing or decreasing values a second variable τ_t is introduced, denoting the estimated trend at the period t. So the value for the period (t+1), predicted at the end of time t, will now be

$$\hat{x}_e(t) + \tau_t$$

In general, the prediction at the end of the period t, for the period (t+h), is given by

$$\hat{x}_e(t,h) = \hat{x}_e(t) + h\,\tau_t \qquad (4\text{-}1.6)$$

Using (4-1.6) we can modify (4-1.2) and (4-1.3) to take account of the trend. Thus

$$\hat{x}_e(t) = a_1 x(t) + (1-a_1)[\hat{x}_e(t-1) + \tau_{t-1}] \qquad (4\text{-}1.7)$$

Let the error be given by

$$e_t = x(t) - [\hat{x}_e(t-1) + \tau_{t-1}] \qquad (4\text{-}1.8)$$

Then we can write

$$\hat{x}_e(t) = [\hat{x}_e(t-1) + \tau_{t-1}] + a_1 e_t \qquad (4\text{-}1.9)$$

If it is found that the trend itself changes gradually, τ_t can be derived from τ_{t-1} by adding to it a quantity $b_1 e_t$, where b_1 is a small positive function, i.e.

$$\tau_t = \tau_{t-1} + b_1 e_t \qquad (4\text{-}1.10)$$

ii) Allowance for seasonal variation:-

If there is a reason to expect seasonal variation,

4.1 contd.

ii) contd.

allowance must be made for that. Consider the seasonal variations about the current mean to be $s_1, s_2, s_3, \ldots$ Then the prediction for the period (t+h) made at the end of period t will be given by

$$\hat{x}_e(t,h) = \hat{x}_e(t) + h\,r_t + s_{t+h} \qquad (4\text{-}1.11)$$

also, (4-1.8) will then become

$$e_t = x_t - [\,\hat{x}_e(t-1) + r_{t-1} + s_t] \qquad (4\text{-}1.12)$$

The seasonal pattern is usually expressed as proportions of the moving average, in which case (4-1.6) becomes

$$\hat{x}_e(t,h) = [\hat{x}_e(t) + hr_t]\,C_{t+h} \qquad (4\text{-}1.13)$$

where C_i denote the proportional seasonal constants and (4-1.8) takes the form

$$e_t = x(t) - [\hat{x}_e(t-1) + r_{t-1}]C_t \qquad (4\text{-}1.14)$$

4.2 Box and Jenkins' method:

While investigating the adaptive optimisation problem, Box and Jenkins were led to the problems of adaptive control which in turn, lead to the problems of prediction and smoothing of a time series. (In their paper (Ref. 8) they have suggested another variation on Holt's method. The formula suggested is

4.2 contd.

$$\hat{Z}_{p+1} = Z_p + \nu_{-1}\Delta e_p + \nu_0 e_p + \nu_1 S^{\ell} e_p$$

which can be interpreted in our notations as

$$\hat{x}_e(t) = \hat{x}_e(t-1) + C_{-1}\,(e_t - e_{t-1}) + C_0 e_t + C_1 \sum_{p=0}^{t} e_p \qquad (4\text{-}2.1)$$

where as before

$$e_t = x(t) - \hat{x}_e(t-1)$$

For the discrete processes the terms $C_{-1}(e_t - e_{t-1})$, $C_0 e_t$ and $e_1 \sum_{p=0}^{t} e_p$ correspond respectively to first difference, proportional and cumulative terms. The control is obtained through these three, in which case they are referred to difference control, proportional control and cumulative control.

From (4-2.1) it can be seen that we constantly need to take into account the difference between the two most recent errors $(e_t - e_{t-1})$, the latest error in prediction e_t, and the cumulative sum $\sum_{p=0}^{t} e_p$ of the previous errors in prediction. However, it can be observed that when $C_{-1} = 0$, $C_0 = a_1$ and $C_1 = b_1$, the predictions obtained by the Box and Jenkins' method are identical with those obtained by Holt's method.

4.2 contd.

Assuming $\hat{x}_e(t)$ and τ_t to be the best unbiased estimates of $x(t)$ and τ, the following relationship follows -

$$\hat{x}_e(t,1) = \hat{x}_e(t) + C_1 \sum_{p=o}^{t} e_p + \tfrac{1}{2}\left\{ C_o - \sqrt{C_o^2 - 4C_1C_{-1}} \right\} e_t \qquad (4\text{-}2.2)$$

and for h steps ahead

$$\hat{x}_e(t,h) = \hat{x}_e[(t-1),h] + C_1 \sum_{p=o}^{t} e_p + \tfrac{1}{2}\left\{ C_o - \sqrt{C_o^2 - 4C_1C_{-1}} \right\} e_t \qquad (4\text{-}2.3)$$

Generally the term $\frac{1}{2}\{C_o - \sqrt{C_o^2 - 4C_1C_{-1}}\}$ is negligible.

The following predictors are said to be optimal.

i) For one step ahead

$$\hat{x}_e(t) = \hat{x}_e(t-1) + C_{-1}(e_t - e_{t-1}) + C_o e_t + C_1 \sum_{p=o}^{t} e_p$$

ii) For two steps ahead

$$\hat{x}_e(t,1) = \hat{x}_e(t) + C_1 \sum_{p=o}^{t} e_p - C_{-1} e_t$$

iii) and for $h > 2$

$$\hat{x}_e(t,h) = \hat{x}_e(t-1,h) + C_1 \sum_{p=o}^{t} e_p$$

4.3 Double smoothing:-

R.G.Brown suggested double smoothing for the linear growth models. Double smoothing is achieved through a parameter β which is thought of as the rate at which an observation loses its importance every period. In this case -

$$\hat{x}_e(t) = \hat{x}_e(t-1) + \tau_{t-1} + (1-\beta^2)\, e_t \qquad (4\text{-}3.1)$$

and

$$\tau_t = \tau_{t-1} + (1-\beta)^2\, e_t \qquad (4\text{-}3.2)$$

This is a particular case of Holt's method.

All these methods are quite easy for computation and are extensively used in industry and economics. It can be shown, that the predictors obtained by e.w.m.a. are the same as would be obtained by Wiener theory for a process with rational spectral density -

$$g(z) = \sigma^2 \left|\frac{Q(z)}{P(z)}\right|^2$$

where $\frac{Q(z)}{P(z)}$ is a polynomial ratio and that for $P(z) = 1-\alpha z$ and $Q(z) = 1-\beta z$, the Wiener predictor obeys the relation

$$\hat{x}_e(t,h) = (\alpha-\beta)\alpha^h x(t) + \beta\, \hat{x}_e(t,h-1)$$

which is similar to (4-1.5).

4.4 Prediction of a process by its characteristic modes:-

Farmer suggested a method for prediction of a process

4.4 contd.

by its characteristic modes. Characteristic modes arise from the physical nature of the sourcd of a particular process and generally for stochastic processes, sample functions can be expressed in terms of a small number of such characteristic modes. Wiener predictor is not found appropriate because, the calculation of autocorrelation function involves a large number of sample functions and more so for the reason that it is modes that characterise the process rather than the correlation function.

Let $x_m(t)$, $m = 1,2, \ldots, M$, be the M sample function of a stochastic process. The first mode can be specified by seeking a function $Z_1(t)$, a scaling factor $\lambda_1^{\frac{1}{2}}$ and a set of coefficients a_{m1}, in such a manner that $\lambda_1^{\frac{1}{2}} a_{m_1} Z_1(t)$ approximates to $x_m(t)$, in a least square sense, over the interval (0,T). Similarly seeking functions, scaling factors and the set of coefficients for second, third and so on modes, we can write

$$x(t) = \lambda_1^{\frac{1}{2}} a_{m_1} Z_1(t) + \lambda_2^{\frac{1}{2}} a_{m_2} Z_2(t) + \lambda_3^{\frac{1}{2}} a_{m_3} z_3(t) + \ldots \quad (4\text{-}4.1)$$

where

$$\lambda_k Z_k(t) = \int_0^R R(t,\tau) Z_k(t)\, d\tau \quad (4\text{-}4.2)$$

and

4.4 contd.

$$a_{mk} = \lambda_k^{\frac{1}{2}} \int_0^T Z_k(t) x_m(t)\, dt \qquad (4\text{-}4.3)$$

$R(t,\tau)$ is a correlation function i.e.

$$R(t,\tau) = \frac{1}{M} \sum_{m=1}^{M} x_m(t) x_m(\tau) \qquad (4\text{-}4.4)$$

It has been established that the autocorrelation $R(\tau,\tau')$, can be expanded in terms of the characteristic modes of a process, as under

$$R(\tau,\tau') = \sum_{k=1}^{\infty} \lambda_k Z_k(\tau) Z_k(\tau') \qquad (4\text{-}4.5)$$

$\lambda_k > 0$ and that if the process is expanded to k terms as a combination of its characteristic modes, then the time integrated square error is given by

$$ET = \int_0^T (R(\tau,\tau')d\tau' - \sum_{k=1}^{K} \lambda_k \qquad (4\text{-}4.6)$$

For achieving a specified accuracy, the number of terms K can be derived by using (4-4.6).

If we know in advance, the values of a sample function $x(t)$ of a non-stationary process in the interval $(0, T_o)$, then the prediction problem becomes one of estimating values of $x(t)$ in the interval (T_o, T).

If C_k is a set of coefficients, then to the specified

4.4 contd.

accuracy, the sample function $x(t)$ can be expressed in the form

$$x(t) = \sum_{1}^{K} C_k Z_k(t) \; ; \qquad 0 < t < T \tag{4-4.7}$$

As this is valid over the whole interval $(0,T)$ it must be valid over (T_o,T) which is included in $(0,T)$. This method automatically treats the sample function $x(t)$ as a combination of its characteristic modes, and hence the prediction problem is reduced to one of determining the constants C_k.

It is assumed that the values of $x(t)$ are known over the interval $(0,T_o)$. One can therefore choose the coefficients C_k such that (4-4.7) is the best mean square approximation to $x(t)$ over $(0,T_o)$.

Thus the expansion error

$$e(t) = x(t) - \sum_{k=1}^{K} C_k Z_k(t)$$

gives the integrated square error

$$E = \int_{0}^{T_o} \left\{ x(t) - \sum_{k=1}^{K} C_k Z_k(t) \right\}^2 dt \tag{4-4.8}$$

Also, E is minimum when

4.4 contd.

$$\frac{\partial E}{\partial c_k} = 0 \; ; \quad k = 1,2, \ldots K$$

This gives

$$\sum_{k'=1}^{K} C_{k'} \int_{o}^{T_o} Z_k(\tau) Z_{k'}(\tau) d\tau = \int_{o}^{T_o} x(\tau) Z_k(\tau) d\tau; \quad k = 1,2,\ldots K \quad (4\text{-}4.9)$$

This is a set of simultaneous equations and determines C_k uniquely. $x(t)$ can now be predicted by applying (4-4.7) over the time interval (T_o, T).

4.5 Winter's Method:-

A seasonal pattern around a trend movement is a common feature in many economic time series. One can assume an additive seasonal pattern and a linear trend for this purpose (as in Sec.4-1), but Winter has pointed out that the multiplicative seasonal pattern and an exponential rather than a linear trend will be more realistic. By taking logarithms, however, the exponential trend can be transformed into a linear one and simultaneously the multiplicative seasonal pattern into a linear one.

Consider again the relation -

$$\hat{x}_e(t) = A\, x(t) + (1-A)\, \hat{x}_e(t-1)$$

Let the periodicity of the seasonal effect be L. If the period

4.5 contd.

is a month, L would be 12 months. Winter suggests the predictor

$$\hat{x}_e(t) = A \frac{x(t)}{f_{t-L}} + (1-A)\, \hat{x}_e(t-1) \; ; \; 0 \leqslant A \leqslant 1 \qquad (4\text{-}5.1)$$

where

$$f_t = B \frac{x(t)}{\hat{x}_e(t)} + (1-B)\, f_{t-L} \; ; \qquad 0 \leqslant B \leqslant 1 \qquad (4\text{-}5.2)$$

which represents the current estimate of the seasonal factor for the period t. The new estimate, f_t, is again a weighted sum of the current estimate, $\frac{x(t)}{\hat{x}_e(t)}$, and the previous estimate f_{t-L}; The expected value for the following period is given by

$$\hat{x}_e(t,1) = \hat{x}_e(t)\, f_{t-L+1} \qquad (4\text{-}5.3)$$

More generally, for T periods ahead

$$\hat{x}_e(t,T) = \hat{x}_e(t)\, f_{t-L+T}, \qquad T \leqslant L^s \qquad (4\text{-}5.4)$$

Incorporating trend effect:-

Here (4-5.1) becomes -

$$\hat{x}_e(t) = A \frac{x(t)}{f_{t-L}} + (1-A) \left[\hat{x}_e(t-1) + R_{t-1} \right]; \qquad (4\text{-}5.5)$$

where R_{t-1} is the most recent estimate of the additive trend factor. The trend estimate can be revised by the relation

4.5 contd.

$$R_t = C \left[\hat{x}_e(t) - \hat{x}_e(t-1) \right] + (1-c)\, R_{t-1} \qquad (4\text{-}5.6)$$

and the predictor becomes

$$\hat{x}_e(t,T) = \left[\hat{x}_e(t) + T\, R_t \right] f_{t-L+T} \; ; \; T = 1,2, \ldots t \qquad (4\text{-}5.7)$$

We shall consider Winter's model, in detail, in the next chapter.

4.6 Conclusion:-

The exponentially weighted prediction methods are easier from the computational point of view as compared to the Wiener's theory, where, a calculation of the autocorrelation function involves a large number of sample functions. But, although the above exponentially weighted prediction methods explain precisely how the predictors are formulated, they are far from explicit with respect to the justification of the method. For this reason, the optimal values of the parameters have to be determined empirically. The complete model is discussed in the next chapter.

CHAPTER 5

APPLICATIONS

In this chapter, we shall discuss the preceding theories in the light of a specific problem. The problem considered is one of electricity demand prediction in an area, up to several hours ahead. The load predictions are necessary for the purposes of ordering generating plant, loading of plant and checking the security of power flow in advance. As we shall see in the next chapter, if Automatic Control is to be applied to a power flow system then load prediction by means of a digital computer becomes a necessity.

5.1 WIENER'S THEORY:

One is always tempted to apply Winer's theory to a prediction problem as far as possible. But the two kinds of problem - (a) obtaining data: Stationariness, spectra, correlation between the useful part of the inputs (message) and that due to noise etc., which necessitates a minute study both in theory and by experiment of the environment of the system and (b) the realisation of the synthetic transfer function obtained by calculation, which is not necessarily possible (and even if it is, it may not be without approximations) - make its application difficult in actual situations. However, if one decides to apply it one can surely expect some useful information.

Also, in the generalised form it is seen [equation (3-2.6)] that the autocorrelation function $C_{ff}(t-\tau, t-\tau')$, specifies

5.1 contd.

the response function $h(t,\tau)$ of the Wiener filter. Clearly, the application of the generalised predictor depends on the possibility of constructing a representative and realistic autocorrelation function, which should in practice, be physically realisable with a finite number of sample functions. Then, if the autocorrelation function is formed from the M sample functions $f_m(t)$, $m = 1,2, \ldots, M$,

$$C_{ff}(t-\tau, t-\tau') = \frac{1}{M} \sum_{m=1}^{M} f_m(t-\tau)\, f_m(t-\tau').$$

The sample functions $f_m(t)$ have to be determined by performing M sets of measurement under identical conditions. However, in actual practice, due to certain uncontrollable variations in one or more parameters of the process, the number of sample functions which one could measure under identical conditions is severely limited. For example, if $f_m(t)$ were to represent the electricity demand curves, the number of sample functions to be used in calculating the autocorrelation function will be limited by the seasonal trend of the load and by the economic growth in demand.

Moreover, in the problem of electricity demand prediction, the sample functions can be expressed in terms of a small number of characteristic modes [see Sec (4-4) and also the next section], in which case, the calculation of a

5.1 contd.

autocorrelation function is neither necessary nor desirable and hence the Wiener predictor is not found useful.

5.2 Prediction by the Characteristic modes:

The method of prediction by characteristic modes is quite useful for the problem of electricity demand prediction. Following the method outlined in Sec.(4-4), the calculation of weighting coefficients etc. can be performed on a computer. For its application to the problem of electricity demand prediction, we divide each load curve into part-day periods of several hours' duration. The prediction problem is one of estimating the load values over the interval (T_o, T), from the values over the interval $(0, T_o)$. The load during each of these periods depends on many factors of which the meteorological ones have the greatest influence. So the load x_{mn} for the mth period and at the nth instant is expressible in the form

$$x_{mn} = \alpha_n + \beta_n f_1(T_m) + \nu_n f_2(L_m) + \delta_n f_3(W_m) + \ldots \quad (5\text{-}2.1)$$

where $f_1(T_m)$, $f_2(L_m)$, $f_3(W_m)$ are respectively the functions of temperature T_m, Light intensity L_m, Wind velocity W_m. The quantity α_n represents the base load and the factors β_n, ν_n, δ_n, allow for the varying importance of weather parameters with the time of day. As it becomes obvious

5.2 contd.

from (5-2.1) each load vector is linearly dependent on the vectors α,β,ν,δ, ... ; however, if the load is expanded to K terms as a combination of its characteristic modes, the quantity x_{mn} takes the form -

$$x_{mn} = \sum_{k=1}^{K} C_{mk} Z_{kn}$$

The mode vectors Z_k minimise the expansion error and if the mode vectors are replaced by K linearly independent combinations then the error is unchanged. Here the modes describe the basic trends of the load under average weather conditions for the period of the records whereas the weighting coefficients C_{mk} depend on the meteorological parameters relevant to the mth period.

5.3 TESTING AN ADAPTIVE PREDICTION (OR EXPONENTIALLY WEIGHTED PREDICTION) MODEL:

Two weeks' hourly electricity demand data, for this purpose, have been obtained. The data consists of 240 readings for 10 days, (Monday to Friday - 5 days a week). The total data is divided into two parts of 120 readings each, the first of which is used to develop initial values and the latter part is used to try out the method by pretending that the future is unknown.

We can use the following symbols:

5.3 contd.

$\hat{D}_{t,T}$ = Predicted demand for the period t + T, after incorporating the trend and seasonal effects.

$\bar{D}_t$ = the estimate of demand for the period t + 1, made in period t.

D_t = actual demand for the period t

W_e = smoothing factor; $0 \leq W_e \leq 1$

W_s = seasonal factor ; $0 \leq W_s \leq 1$

W_a = trend factor ; $0 \leq W_a \leq 1$

A = additive trend adjustment

S = multiplicative seasonal factor

L = the cyclic period

$e_{t,T}$ = the prediction error for the period t + T.

Input data:

For the input data sheet with following data see Appendix A.

$\bar{D}o$ = Initial estimate of demand,

Ao = Initial estimate of A_{t-1}, when t = 1,

H = The maximum value of t in the pretest data (=120).

L = The cyclic period (=24).

S = Total number of periods for which the data is available (=240).

D_s = Actual demand in each of s periods

5.3 contd.

So to S_{-2s} = Initial values of S_{t-L}

Formulae used:

$$\bar{D}_t = W_e \frac{D_t}{S_{t-L}} + (1-W_e)(\bar{D}_{t-1} + A_{t-1}) \qquad (5\text{-}3.1)$$

$$S_t = W_s \frac{D_t}{\bar{D}_t} + (1-W_s)S_{t-L} \qquad (5\text{-}3.2)$$

$$A_t = W_a(\bar{D}_t - \bar{D}_{t-1}) + (1-W_a)\, A_{t-1} \qquad (5\text{-}3.3)$$

$$\hat{D}_{t,T} = (\bar{D}_t + T\, A_t)S_{t-L+T} \qquad (5\text{-}3.4)$$

$$e_{t,T} = \hat{D}_{t,T} - D_{t+T} \qquad (5\text{-}3.5)$$

$$\varepsilon_t = \sum_{t=1}^{t} e^2_{t,T} \qquad (5\text{-}3.6)$$

Steps in Computation:

1. Make use of equation (5-3.1) to find $\bar{D}_t$
2. Use (5-3.2) to find S_t and store as S_{t-L}
3. Use (5-3.3) to find A_t and store as A_{t-1}
4. When $t \geq H$, use (5-3.4) to find predicted values
5. Use (5-3.5) to find $e_{t,T}$
6. Square and acummulate error values
7. Print Standard Deviation along side the combination of values of W_e, W_s, W_a.
8. Repeat steps 1 to 7 for the sequence of values of W_e, W_s, W_a.

5.3 contd.

9. Select the optimum set of W_e, W_s, W_a by visual inspection i.e. the set for which the S.D. is minimum (see Appendix C)

10. Repeat 1 to 7 using optimal set and printing the $\hat{D}_{t,T}$ values.

See Appendix B and Appendix D for the computer programme and results respectively.

Note on initial values:-

If $D_{1i}, D_{2i}, \ldots, D_{5i}, (i = 1,2, \ldots, 24)$, denote the demand during the 1st, 2nd, and so on days, the total demand for days is - (For data see appendix A);

$$\sum_{i=1}^{24} D_{1i} = 5429.7 \qquad \sum_{1}^{24} D_{2i} = 5935.5$$

$$\sum_{1}^{24} D_{3i} = 5564.6 \qquad \sum_{1}^{24} D_{4i} = 5657.1$$

$$\sum_{1}^{24} D_{5i} = 5640.7$$

giving

$\bar{D}_1 = 226.2$, $\bar{D}_2 = 247.3$, $\bar{D}_3 = 231.8$,

$\bar{D}_4 = 235.7$, $\bar{D}_5 = 235.0$.

To know the trend we calculate -

ELECTRICITY DEMAND TIME SERIES USED FOR MODEL-BUILDING

TABLE 5-1: $S = \frac{D_t}{\bar{D}}$

Hour t	Day D_1	D_2	D_3	D_4	D_5	Total	Average	Adjusted Averages
1	0.46	0.60	0.64	0.60	0.65	2.95	0.59	0.60
2	0.43	0.57	0.59	0.55	0.61	2.75	0.55	0.56
3	0.41	0.53	0.56	0.51	0.57	2.58	0.51	0.52
4	0.42	0.56	0.59	0.55	0.60	2.72	0.54	0.55
5	0.43	0.56	0.59	0.57	0.60	2.75	0.55	0.56
6	0.47	0.60	0.62	0.59	0.63	2.91	0.58	0.59
7	0.70	0.75	0.77	0.75	0.79	3.76	0.75	0.76
8	1.07	1.02	1.05	1.05	1.10	5.29	1.05	1.06
9	1.133	1.27	1.28	1.27	1.32	6.47	1.29	1.31
10	1.42	1.26	1.30	1.26	1.28	6.52	1.30	1.32
11	1.36	1.24	1.22	1.22	1.23	6.27	1.25	1.27
12	1.29	1.26	1.27	1.23	1.24	6.29	1.25	1.27
13	1.22	1.23	1.15	1.19	1.21	6.00	1.20	1.21
14	1.24	1.13	1.21	1.19	1.21	5.98	1.19	1.20
15	1.24	1.13	1.21	1.16	1.18	5.92	1.18	1.19
16	1.23	1.13	1.10	1.12	1.18	5.76	1.15	1.16
17	1.30	1.19	1.27	1.24	1.21	6.21	1.24	1.25
18	1.31	1.27	1.27	1.26	1.24	6.35	1.27	1.28
19	1.20	1.16	1.17	1.19	1.18	5.90	1.18	1.19
20	1.19	1.12	1.13	1.16	1.17	5.77	1.15	1.16
21	1.16	1.09	1.09	1.11	1.10	5.55	1.11	1.12
22	1.17	1.08	1.08	1.15	1.01	5.49	1.09	1.10
23	1.05	0.99	0.99	1.05	0.85	4.93	0.98	0.99
24	0.80	0.75	0.74	0.84	0.73	3.86	0.77	0.78
TOTAL	23.90	23.49	23.89	23.81	23.89	118.98	23.72	24.00

5.3 contd.

$$A_o = \frac{\Sigma D_{5i} - \Sigma D_{1i}}{4 \times 24} = 2.2$$

This being small we can neglect the trend factor and thus calculate the S factor from the relation

$$S = \frac{D_t}{\overline{D}}$$

This factor is calculated for each observation over a period of 5 days. The rowwise and the columnwise totals are taken and the average S ratio is then calculated. Table 5-1 gives the S values.

In case, the trend factor is significant, the S factor is given by

$$\frac{D_t}{\overline{D} + A(t-12)} \quad ; \quad (t = 1,2,\ldots,24).$$

5.4 ERRORS IN PREDICTION:

To improve the reliability of a prediction model, one will need some criteria to check the results and decide on any changes in the parameter values resulting from the changing circumstances. This can be achieved through appropriate control limits for the errors in prediction. We have

$$e_{t,T} = D_{t+T} - \hat{D}_{t,T}$$

From the Central Limit Theorem, one expects both the predictions

5.4 contd.

and the errors in prediction to be normally distributed. It is convenient to employ the mean absolute deviation η_e (which is proportional to the Standard Deviation of the error distribution σ_e), rather than work with σ_e which involves computations of squares of roots. Denoting the probability density function by f(D), we have

$$\sigma^2 = \int_{-\infty}^{\infty} (D-\bar{D})^2 \, f(D) \, dD \qquad (5\text{-}4.1)$$

and

$$\eta = \int_{-\infty}^{\infty} |D - \bar{D}| \, f(D) \, dD \qquad (5\text{-}4.2)$$

and for a normal distribution -

$$\eta = \sqrt{2/\pi}\,\sigma = 0.7979\sigma \qquad (5\text{-}4.3)$$

Thus for calculations, one can substitute

$$\sigma = \frac{\eta}{0.7979}$$

The current estimate of η is obtained by single smoothing the error series

$$\eta_t = \alpha|e_t| + (1-\alpha)\,\eta_{t-1} \qquad (5\text{-}4.4)$$

Again, an a priori estimate of η_0 is necessary to begin the smoothing operation.

The confidence limits then take the form

5.4 contd.

$$C.L. = \hat{D}_{t,T} \pm \frac{K \eta_t}{0.7979} \qquad (5\text{-}4.5)$$

with K setting the level of confidence. By referring to a Normal Distribution table, one can find out the probability of the observation D_{t+T} falling outside the limits.

In case, the mean prediction error is zero, the sum of the prediction errors at any time can be expected to be close to zero.

$$E_t = e_t + E_{t-1} \qquad (5\text{-}4.6)$$

R.G.Brown has shown that the variance of the sum of the prediction errors σ_E^2 is given by

$$\sigma_E^2 = \frac{\sigma_D^2}{1 - (1-\alpha)^{2n}} \qquad (5\text{-}4.7)$$

where n stands for the degree of smoothing (n = 1 for a constant model, 2 for a linear model etc.). By considering the mean absolute deviation of the error series

$$\sigma_D^2 = \frac{\eta^2 \pi(2-\alpha)}{4} \qquad (5\text{-}.48)$$

From (5-.47) and (5-4.8)

$$\sigma_{E_t} = \frac{\eta_t}{2} \sqrt{\frac{\pi(2-\alpha)}{1 - (1-\alpha)^{2n}}}$$

and then the control limits for the error sum become

5.4 contd.

$$C.L. = E_t \pm K \frac{\eta_t}{2} \sqrt{\frac{\pi(2-\alpha)}{1-(1-\alpha)^{2n}}} \qquad (5\text{-}4.9)$$

K being the normal deviate.

DISCUSSION:

Here we have discussed the models for load prediction in the Electricity Supply Industry and also seen a method to establish the control limits for the prediction errors. The exponentially weighted prediction model as developed by Winter mainly for sales forecasting has been tested against the data for the electricity demand in a particular area. In this model the set of parameter values for which the S.D. is minimum gives the optimum prediction. For comparison purposes, the results obtained with a slight change in parameter values (viz. $W_e = 0.90$, $W_a = 0.00$, $W_s = 0.90$ instead of $W_e = 0.85$, $W_a = 0.00$, $W_s = 0.95$ as in optimum case), have been listed in Appendix E.

To make further comparison of the results possible, one can use the predictor given by equation (4-1.3) i.e.

$$\hat{x}_e(t) = \hat{x}_e(t-1) + a_1[x(t) - \hat{x}_e(t-1)]$$

$$= \hat{x}_e(t-1) + a_1 e_t$$

The 24 hourly moving averages for the purpose are calculated by using data for the first five days (see Appendix F) and the seasonal variations in the last column are then adjusted such

5.4 contd.

that their sum is zero. To find the predicted values, the following table (with actual results appearing in Appendix G) is used.

PREDICTED (1)	ACTUAL (2)	A - P (3)	$\hat{x}_e(t) = \hat{x}_e(t-1)+a_1 e_t$ (4)	SEASONAL ADJUSTMENTS (5)	P_1 (6)
			235		156
156	117	-39	200	-79	111
111	108	- 3	197	-89	99
99	107	8	⋮	-98	⋮
⋮	⋮	⋮			

We begin with the revised value of $\hat{x}_e(t)$ for the last value in column (5) Appendix F viz. 235. First the prediction one period ahead P_1 is obtained by adding to this figure of 235 the seasonal adjustment for the next periods i.e. 235 - 79 = 156 and so on. The P_1 values are copied in column (1). When the actual demand for that period is known (117), this is inserted in column (2) and the difference between the actual and predicted values (117 - 156 = -39), is inserted in column (3). The new value is obtained by adding to the previous value of $\hat{x}_e(t-1)$, 0.9 times the difference e_t, (i.e. $\hat{x}_e(t) = 235+0.9(-39)=200$ to the nearest whole number. P_1 values are obtained by adding

5.4 contd.

the appropriate seasonal adjustments to the new value of $\hat{x}_e(t)$ and the process is repeated for each reading.

CHAPTER 6

AUTOMATIC CONTROL.

A great deal of theoretical and practical work in the field of Automatic Control systems dates back to the second world war and the industrial use of such systems has increased ever since. This is because of the complexity involved in the modern Industrial processes which demand faster and more accurate control systems. The applications of automatic control systems in Process industries, manufacturing, the steering and operation of ships and aircraft, and in modern weapons systems have increased tremendously and one could see that the future possibilities of such applications greatly depend on the developments in the field of digital computers and also the new science of Cybernetics (see Sec.1-2). We shall summarise below, some of the important features of Digital Computers.

DIGITAL COMPUTERS:- The general idea of automatic control through digital computers, also dates from soon after the war, but apart from some small machines, the first really large digital computer came into operation in 1951-52. In the early stages, these were used only for the scientific and engineering calculations, but now, one can see that the majority of the machines are used for business applications and it is most likely that a substantial number of digital computers will soon find their use in control applications. The main purpose of this chapter is to explore some of the possibilities of such applications, with the Electricity Supply Industry primarily in mind.

A digital computer consists of a large number of

vacuum tubes or transistors and though the machine is very simple conceptually, the physical realisation involves an exceedingly complex network. The programme of elementary operations is performed by the control unit which is the central organ of the machine. From the users point of view, the complexity lies in the programme rather than in the machine and for a particular application the writing of a programme can become a long and costly undertaking.

Ordinarily, a digital computer prints its results or punches them on cards or paper tapes, but with special output devices, it could do other things such as plotting of curves (viz. Autocorrelation curves etc.) or giving an alarm to the programmer indicating that a certain anticipated situation has occurred. The control operations such as switching electric motors on or off, adjusting hydraulic valves etc., can also be performed with special output arrangements and similarly, information directly from the plant can be received with special input connections.

Let us now turn to the subject matter of this chapter. For convenience, we shall divide the following discussion into two sections. The first one covering briefly, some concepts, criteria and mathematical techniques which form the basis of the modern Automatic control theory and the second one dealing with a specific problem of automatic control in the Electricity Supply Industry, where we shall discuss a closed loop general system which could, in future, become capable of controlling the power flow automatically.

6.2 contd.

is at s_1, it is always possible to find a small circular region around s_1 such that no other singularities lie within this region. Any rational algebraic function can be written as a ratio of the polynomials

$$\frac{(s-a)(s-b) \quad \dots \quad (s-m)}{(s-a_1)(s-b_1) \quad \dots \quad (s-m_1)} .$$

All singularities of the rational algebraic functions are poles. In the above form, the poles ard at $a_1, b_1, \dots, m_1$. If n of the denominator factors are identical, one of the poles is of order n.

6.3 SYSTEMS WITH FEEDBACK AND STABILITY CRITERIA:-

In a Feedback system the output is fed back into the input part of the system in such a way that it will affect its own value. Thus a feedback path, as well as a forward path exists within the system, resulting in what is referred to as a closed-loop system.

Feedback is an essential element in the operation of an automatic control system and the performance of a feedback control system is determined by its accuracy and stability. An unstable feedback cannot perform any useful control function, for which, it is very important to be able to determine whether the system with feedback is stable. One can discuss stability in terms of the transfer functions represented by blocks.

In the following control system with single feedback loop,

6.3 contd.

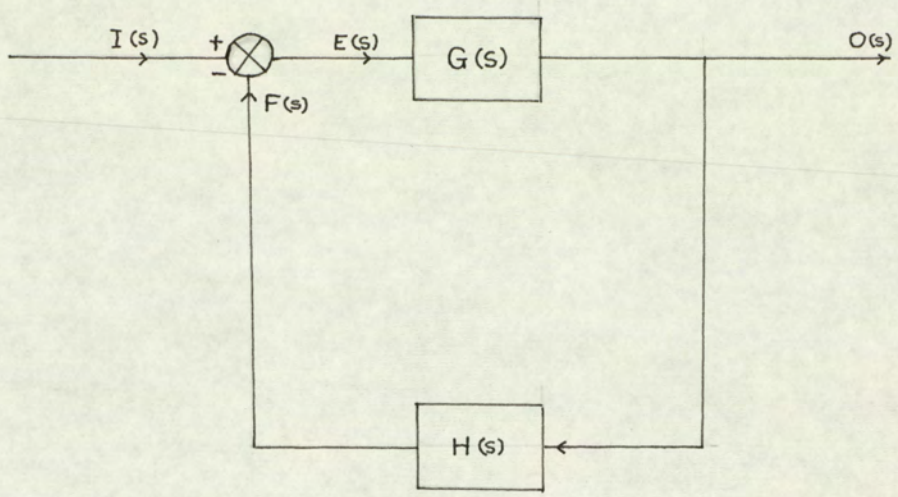


Fig. 6-1 Feedback system.

where I(s), O(s), F(s) and E(s) represent the Laplace transforms of the input function i(t), the output function o(t), the feedback function f(t), and the error signal e(t) respectively, it can be shown that -

$$\frac{O(s)}{I(s)} = \frac{G(s)}{1+G(s)H(s)} \qquad (6\text{-}3.1)$$

This is the ratio of the transferred output to the transferred input and is called the system transfer function. Generally G(s) and H(s) are the function of rational polynomials in s and in that case $\frac{O(s)}{I(s)}$ can be written as

$$G_o(s) = \frac{O(s)}{I(s)} = \frac{a_m s^m + a_{m-1} s^{m-1} + \dots + a_1 s + a_o}{b_n s^n + b_{n-1} s^{n-1} + \dots + b_1 s + b_o} \qquad (6\text{-}3.2)$$

If the unit step function U(t) is applied to the

6.3 contd.

system we have, $I(s) = \mathcal{L}[U(t)] = \frac{1}{s}$ and then

$$O(s) = \frac{1}{s} G_o(s) = \frac{a_m s^m + a_{m-1} s^{m-1} + \ldots + a_1 s + a_o}{s(b_n s^n + b_{n-1} s^{n-1} + \ldots + b_1 s + b_o)} = \frac{A(s)}{B(s)}, \text{ say.}$$

The last expression is reducible to the form

$$O(s) = \frac{1}{b_n} \left[\frac{K_o}{s} + \frac{K_1}{s-s_1} + \frac{K_2}{s-s_2} + \ldots + \frac{K_n}{s-s_n} \right] \quad (6\text{-}3.3)$$

where the constants $K_o, K_1, K_2, \ldots, K_n$ are given by the Heaviside's formula

$$K_k = b_n \left\{ (s-s_k) \frac{A(s)}{B(s)} \right\}_{s=s_k} \quad (6\text{-}3.4)$$

Also by using Heaviside's expansion formula -

$$\mathcal{L}^{-1} \left[\frac{A(s)}{B(s)} \right] = \mathcal{L}^{-1} \left\{ \frac{1}{b_n} \sum \frac{K_k}{s-s_k} \right\} = \sum_{k=1}^{n} \left\{ (s-s_k) \frac{A(s)}{B(s)} \right\}_{s=s_k} e^{s_k t} \quad (6\text{-}3.5)$$

One can show that the output or response function is given by

$$o(t) = \mathcal{L}^{-1}[O(s)] = \frac{1}{b_n} \left[K_o + K_1 e^{s_1 t} + K_2 e^{s_2 t} + \ldots + K_n e^{s_n t} \right] U(t) \quad (6\text{-}3.6)$$

The first step in this equation represents the steady state response, whereas the exponential terms are the transient responses, $s_1, s_2, s_3, \ldots, s_n$ represent roots of the polynomial in the denominator of the overall transfer function $G_o(s)$. Corresponding to the roots $s_k = \sigma_k + i\,\omega_k$, the output function is

6.3 contd.

$$\frac{K_k}{b_n} e^{(\sigma_k + i\,\omega_k)t} = \left[\frac{K_k}{b_n} e^{\sigma_k t}\right] e^{i\,\omega_k t}$$

and one can see that for $\sigma_k > 0$, the amplitude of this term increases with time and for $\sigma_k < 0$ the amplitude decreases with time. Therefore, positive σ_k causes an unstable situation and can not be tolerated. Also, $\sigma_k = 0$ will cause sustained oscillations of constant amplitude resulting in instability.

The roots of the polynomial equation are called zeros of the polynomial. $s_1, s_2, \ldots, s_k, \ldots, s_n$, are zeros of the denominator of $G_o(s)$ or zeros of the polynomial $1 + G(s)H(s)$. Thus the necessary and sufficient condition for the stability of a feedback system is that all the zeros of $1 + G(s)H(s)$ have negative real parts.

Alternatively, $s_1, s_2, s_3, \ldots, s_n$ also represent the poles of $G_o(s)$ in which case, the necessary and sufficient condition for the stability of a feedback system is that the poles of its overall transfer function have negative real parts.

Very often, the solving of the equation $1 + G(s)H(s) = 0$, for all its roots, involves a lot of complications. However, Routh-Hurwitz and Nyquist criteria have been designed to determine stability without actually finding the roots. We shall describe them briefly.

(i) Routh-Hurwitz criterion:-

6.3 contd.

(i) contd.

This is based on the characteristic equation of the type -

$$A_n s^n + A_{n-1} s^{n-1} + A_{n-2} s^{n-2} + \ldots + A_1 s + A_0 = 0$$

and it considers the determinant

$$\begin{vmatrix} A_{n-1} & A_n & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ A_{n-3} & A_{n-2} & A_{n-1} & A_n & 0 & 0 & 0 & 0 & 0 \\ A_{n-5} & A_{n-4} & A_{n-3} & A_{n-2} & A_{n-1} & A_n & 0 & 0 & 0 \\ A_{n-7} & A_{n-6} & A_{n-5} & A_{n-4} & A_{n-3} & A_{n-2} & A_{n-1} & A_n & \\ - & - & - & - & - & - & - & - & A_0 \end{vmatrix}$$

formed by the coefficients of the characteristic equation.

The Routh-Hurwitz criterion states that there will be no positive real parts to the roots, if this determinant is greater than zero and all the determinants formed by successive elimination of the right hand column and the bottom row are also greater than zero.

The chief merit of this criterion is its simplicity, but it conveys information only on stability. It does not describe the transient behaviour of the system and hence is of little use for system synthesis.

(ii) Nyquist Criterion:-

The Nyquist stability criterion is more difficult

6.3 contd.

(ii) contd.

to apply than the Routh-Hurwitz criterion but its following features make it preferable for the stability analysis of a feedback control system:

(a) It provides the same amount of information as does the Routh Hurwitz criterion and in addition gives the indication of the degree of stability of a stable system.

(b) It suggests ways for improving the system stability, if necessary.

(c) The Nyquist locus gives information concerning the frequency response of the system.

The assumptions underlying the Nyquist's theory are:

(A) that the system must be governed by a system of linear differential equations with constant coefficients,

(B) that the limit of G(s)H(s) must approach a constant or zero as $s \to \infty$, i.e. G(s)H(s) must be a proper fraction which means that the power of s in the denominator of G(s)H(s) must be equal to or greater than the power of s in the numerator. (This is always true in practical cases),

(C) that there are no poles of G(s)H(s) in the positive real half of the complex plane.

6.3 contd.

We have seen that, for a stable system, no roots of the characteristic equation should have positive real parts, i.e. there should not be any zeros of $1 + G(s)H(s)$ in the positive real half of the complex plane.

Now the poles of $1+G(s)H(s)$ are the same as those of $G(s)H(s)$ and the zeros of $1+G(s)H(s)$ are the roots of the characteristic equation, whose location is our main concern in the stability study. Considering assumptions (B) and (C), it can be shown that, as ω is varied from $-\infty$ to $+\infty$, the number of times that the vector $1+G(i\omega)H(i\omega)$ goes round the origin in the clockwise direction is equal to the number of zeros of $1+G(s)H(s)$ in the positive real half of the complex plane. Also vector of $1+G(i\omega)H(i\omega)$ going round the origin is the same as vector $G(i\omega)H(i\omega)$ going round the point $(-1,i0)$.

The Nyquist Criterion states that: A feedback system is stable only if the frequency response locus of the system transfer function does not enclose the point $(-1+i0)$. As the locus is traversed in the direction of increasing ω, i.e. from 0 to ∞, $-\infty$ to 0 and along the contour connecting the $+0$ and -0 extremities, the $-1+i0$ point must always be on the left for a stable system.

The GH-plot and the Nyquist diagram supply information on more than the state of stability of a system. The Nyquist diagram in the complex plane is a good indication of the

6.3 contd.

transient response of a system, and suggests how system design may be improved.

In case of instability, the system will have to be changed in such a way that the modified GH-plot does not encircle the point (-1,i0) as ω changes from $+\infty$ to zero. In other words, the point (-1,i0) must lie to the right of the curve traced as ω varies

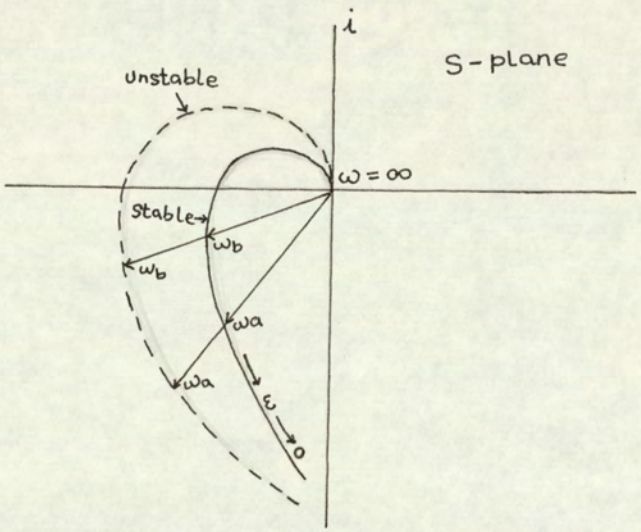


Fig. 6-2 Nyquist diagram.

from $+\infty$ to zero. This can be achieved by reducing the steady state gain of the system, which has the effect of reducing the magnitude of vector $G(i\omega)H(i\omega)$, while keeping it at the same phase angle. The effect of this on the Nyquist diagram is as shown in the fig.(6-2).

6.4 RANDOM FUNCTIONS IN AUTOMATIC CONTROL:

Wiener's theory of optimum filters and Prediction

6.4 contd.

theory, along with Shannon's information theory, have demonstrated the use of random functions in system calculations. The practical use of the theory of random functions is increasing and will increase much faster as more and more engineers and scientists become familiar with random functions and the calculation of probabilities. One could relate the main reason for the increase of activity in this field to the immense development in the field of automatic computers, which has made it possible to calculate correlations and spectra or to solve numerically the resulting integral equations.

Some of the properties and uses of random functions have been demonstrated in the earlier chapters.

6.5 METHODS OF OPTIMISATION:-

In recent years there has been a remarkable growth of interest in problems of systems optimisation and of optimal control. As a result of this interest, have sprung up various methods useful for different situations. We shall summerise some of them.

Phillips method:-

In servo-systems one requires that the input $x(t)$ and the output $y(t)$ of the system should remain as close as possible i.e. the error

$$e(t) = x(t) - y(t)$$

6.5 contd.

should be minimal. However, it becomes difficult when the minimization involves the difference between two random functions of time. Most often e(t), in this case, is described by its mean square value -

$$\overline{e^2} = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} e^2(t)\, dt \qquad (6\text{-}5.1)$$

and the best servo system will be one for which $\overline{e^2}$ is minimum. This is the mean square criterion.

Now if H(s) is the transfer function of error to input of a system whose probable input is a stationary random variable of frequency spectrum $\phi(\omega)$, then

$$\overline{e^2} = \int_{-\infty}^{\infty} \phi(\omega)\, |H(i\omega)|^2\, d\omega \qquad (6\text{-}5.2)$$

Also, if the linear servo is subject to two non-correlated inputs (a) the command s with frequency spectrum ϕ_s and (b) the disturbance d with spectrum ϕ_d, then it can be shown that

$$\overline{e^2} = \int_{-\infty}^{\infty} [\phi_s |H_1|^2 + \phi_d |H_2|^2]\, d\omega \qquad (6\text{-}5.3)$$

This calculation determines the servo which minimises $\overline{e^2}$, with greatest weight to the most probable inputs or to those frequencies for which $\phi(\omega)$ is large.

6.5 contd.

However, in actual practice, one finds the mean square error as a function of the parameters $\alpha, \beta, \ldots,$ which define the possible regulation of the servo system

$$\overline{e^2} = f(\alpha, \beta, \ldots,)$$

in which case, one has to solve the equations obtained by partial differentiation of $\overline{e^2}$

$$\frac{\overline{\partial e^2}}{\partial \alpha} = 0, \quad \frac{\overline{\partial e^2}}{\partial \beta} = 0, \ldots, \text{ etc.}$$

with the increase of degree of the equations, the job of solving these equations becomes tedius and the use of computers becomes essential.

Wiener's method:-

This has already been dealt to some extent earlier. Here the system is defined by its impulse response h(t) and the problem becomes that of seeking the form of h(t) which will minimise the quantity -

$$\overline{e^2} = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} [g(t) - f(t)]^2 \, dt \qquad (6\text{-}5.4)$$

where g(t) and f(t) are the ideal and effective outputs respectively.

Minimisation of $\overline{e^2}$ involves the calculus of variations which leads to the Wiener-Hopf equation -

6.5 contd.

$$C_{gf} - \int_{-\infty}^{\infty} h(\tau') \, C_{ff}(\tau-\tau') d\tau' = 0 \qquad (6-5.5)$$

for all the positive values of τ.

Linear Programming:-

The Linear programming techniques and their use in optimisation is much publicised since the war. The theory depends upon the fact that in actual practice, variables do not take an unrestricted range of values, but are subject to a variety of practical constraints. In linear programming the focus is on optimisation under inequality constraints. Mainly, the Simplex and the Transportation techniques are used in solving the problems. The power of Linear Programming, is greatly increased by the duality relations discovered and published by Kuhn and Tucker which provide direct contact betwedn Linear programming and other important theories such as Games and Statistical decisions.

Dynamic Programming:-

Dynamic Programming is a new method of optimisation and is due to R.Bellman of Rand Corporation. It is maily for the purpose of treating multistage decision processes and those other processes which could be interpreted in this fashion.

(i) Multistage Processes:-

6.5 contd.

(i) contd.

Let p be a point in a set of space P and let the function T(p) denote a transformation with the property that the transformed point $p_1 = T(p)$ belongs to P whenever p does. If a sequence of point is generated in the following fashion -

$$p_1 = T(p),\ p_2 = T(p_1),\ \ldots,\ p_{N+1} = T(p_N),\ldots \quad (6\text{-}5.6)$$

the set of points so generated is called a multistage process and the point 'p' a state variable.

(ii) Recurrence Relations.

Let us now associate a scalar function $h(p,p_1,\ldots,p_N,\ldots)$, with the above multistage process and consider h to have either of the following forms -

$$g(p) + g(p_1) + \ldots + g(p_N) + k(p_N)$$

$$g(p,p_1) + g(p_1,p_2) + \ldots$$

$$\max_i\ g(p_i) \quad (6\text{-}5.7)$$

To illustrate the method, let us begin with the function

$$h(p,p_1,p_2,\ldots,p_N) = g(p) + g(p_1) + \ldots + g(p_N) + k(p_N)$$

which is associated with a finite multistage process. Since the point p_i are determined by means of the recurrence relation

6.5 contd.

(ii) contd.

$$p_{i+1} = T(p_i) \ , \quad i = 0,1,2, \ \ldots$$

it follows that h is a function of p and N and hence we can write

$$h(p,p_1, \ \ldots,p_N) = f_N(p)$$

where N = 0,1,2, ... and all p in P. A simple recurrence relation can now be derived for this sequence of functions, namely,

$$\begin{aligned} f_N(p) &= h(p) + [h(p_1)+ h(p_2)+ \ \ldots + h(p_N)+ k(p_N)] \\ &= h(p) + f_{N-1}(p_1) \\ &= h(p) + f_{N-1}[T(p)] \ , \quad N \geq 1 \end{aligned} \tag{6-5.8}$$

also $f_0(p) = h(p) + k(p)$.

Similarly, for

$$h = g(p,p_1) + g(p_1,p_2) + \ \ldots + g(p_{N-1},p_N), \tag{6-5.9}$$

writing $f_N(p) = g(p,p_1) + g(p_1,p_2) + \ .. + g(p_{N-1},p_N)$

we get

$$\begin{aligned} f_N(p) &= g(p,p_1) + [g(p_1,p_2)+ \ \ldots + g(p_{N-1},p_N)] \\ &= g(p,p_1) + f_{N-1}(p_1) \\ &= g[p,T(p)] \ + f_{N-1} \ [T(p)] \end{aligned} \tag{6-5.10}$$

and finally,

$$\begin{aligned} f(p) &= \max_i g(p_i) \\ &= \max \ [g(p), \ g(p_1), \ g(p_2), \ \ldots] \\ &= \max \ [g(p), \ \max_{i \geq 1} \ g(p_i)] \\ &= \max \ [g(p), \ f(T(p))] \end{aligned} \tag{6-5.11}$$

6.5 contd.

(iii) Multistage Decision Processes:-

Suppose that at each stage we have a choice of transformations to employ. Then

$$p_{n+1} = T(p_n, q_{n+1}) \qquad (6\text{-}5.12)$$

where q_n is an element of a set Q, and $T(p,q)$ is a transformation with the property that $T(p,q)$ belong to P whenever p belongs to P and q to Q.

Now starting in stage p, select q_1 such that the selection of q_1 is equivalent to the choice of a transformation $T(p,q_1)$ i.e. decision is equated to transformation. Beginning in the new state p_1, select q_2 and so on, so that the value of q that is selected at each stage depends upon the current state of the system.

To determine the choice of q_i, a criterion function

$$Q = Q(p, p_1, \ldots, p_N, \ldots, q_1, q_2, \ldots, q_N, \ldots) \qquad (6\text{-}5.13)$$

is constructed which is to be maximised. A set of decisions $q_{N+1}(p_N)$ is called a policy and a set which maximises Q, an optimum policy. Suppose, Q is of the form

$$h(p,q_1) + h(p_1,q_2) + \ldots + h(p_{N-1},q_N) + g(p_N) \qquad (6\text{-}5.14)$$

we can then write

$$f_N(p) = \max_{\{q\}} [h(p,q_1)+h(p_1 q_2)+ \ldots +h(p_{N-1},q_N)+g(p_N)] \qquad (6\text{-}5.15$$

for $N \geq 1$ and p in P, the recurrence relation for

6.5 contd.

(iii) contd.

$\{f_N(p)\}$ can be obtained as under

$$f_N(p) = \max_{q_1} \max_{q_2} \dots \max_{q_N} [h(p,q_1)+h(p_1,q_2)+\dots+h(p_{N-1},q_N)+g(p_N)]$$

$$= \max_{q_1} \left\{ h(p,q_1) + \max_{q_2} \max_{q_3} \dots \max_{q_N} [h(p_1,q_2)+\dots+g(p_N)] \right\}$$

$$= \max_{q_1} [h(p,q_1) + f_{N-1}(p_1)]$$

$$= \max_{q_1} [h(p,q_1)+f_{N-1}(T(p,q_1))] \qquad (6\text{-}5.16)$$

for $N \geq 2$, with $f_1(p) = \max_{q_1} h(p,q_1)$.

This reduces the multidimensional maximisation problem to a sequence of lower-dimensional problem.

(iv) Principle of Optimality:-

An optimal policy has a property that whatever the initial decision and the initial state are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

To illustrate the above ideas, consider the following problem in control theory. To improve the performance of a physical system which is not behaving in a desirable fashion, we exert some additional forces - control forces. The basic problem of control is to strike a balance

6.5 contd.

(iv) contd.

between the cost of control and the cost of deviation of the system from its ideal performance.

Suppose that the physical system is described at time t, $t = 0,1,2,\ldots$, by the scalar quantity x_t, and the equation governing this quantity is

$$x_{t+1} = h(x_t, y_t), \quad x_0 = a \tag{6-5.17}$$

where y_t is the control variable and 'a' the initial state of the system. Here, we want to choose the value of y_t so as to keep the system as closely as possible to the constant state c. Let $k(x_t-c)$ represent the cost of the deviation at the t th stage and $g(y_t)$ be the cost of control. The aim is to minimise the expression

$$\sum_{t=0}^{N} [k(x_t-c) + g(y_t)] \tag{6-5.18}$$

If $f_N(a)$ denotes the minimum value, for $N = 0,1,2,\ldots$, $-\infty < a < \infty$. Then

$$f_0(a) = \min_y [k(a-c) + g(y)] \tag{6-5.19}$$

$$\text{and } f_N(a) = \min_y [k(a-c) + g(y) + f_{N-1}(h(a,y))] \tag{6-5.20}$$

For a case where there is no free choice of the control variable at each stage, but there is a constraint

6.5 contd.

(iv) contd.

of the type $|y_t| \leq k$, the recurrence relation (6-5.20) will become

$$f_N(a) = \min_{|y| \leq k} [k(a-c) + g(y) + f_{N-1}(h(a,y))] \quad (6\text{-}5.21)$$

Games Theory:-

The Games Theory first came to the general notice after the publication of "The theory of Games and Economic behaviour" by Von Neuman and Morgenstern in 1944. In an extensive form, a game is described by a set of players and a set of rules which specify (1) the choices of action that are open to each player under all possible circumstances and (2) Each player's payoff at the end of any "play". Games are classified in terms of (a) the number of players (b) the number of moves (c) the nature of the payoff and (d) other characteristics of the rules. A game is said to be finite or infinite according as the total number of alternatives are finite or infinite.

In the beginning, the attention was focussed on the two-person zero sum games, finite or infinite; but now the emphasis has shifted to general games where there are still many important unsolved problems. More recently, the theory of games is being used with success in the problems of automatic

6.5 contd.

control. Random disturbances play a major role in many control problems and in problems such as designing an automatic pilot, the game will be between the random disturbances on one side and the designer on the other.

6.6 ADAPTIVE CONTROL SYSTEMS:-

An adpative system is designed to modify itself in the face of a new environment, so as to optimise its performance. In actual practice, one does not have the knowledge of the plant or the environmental changes, and under such circumstances, one has to call on the controller to compute or identify the characteristics of the plant while the system is in normal operation. The controller must then make a decision concerning the way in which the system should be adjusted so as to improve the operation with respect to a defined performance index and then certain signals or parameters must undergo modification to accomplish this result. Thus, adaptivity involves the functions of (i) identification (ii) decision and (iii) modification.

The general adaptive control system, as represented by blocks is shown in figure below:

6.6 contd.

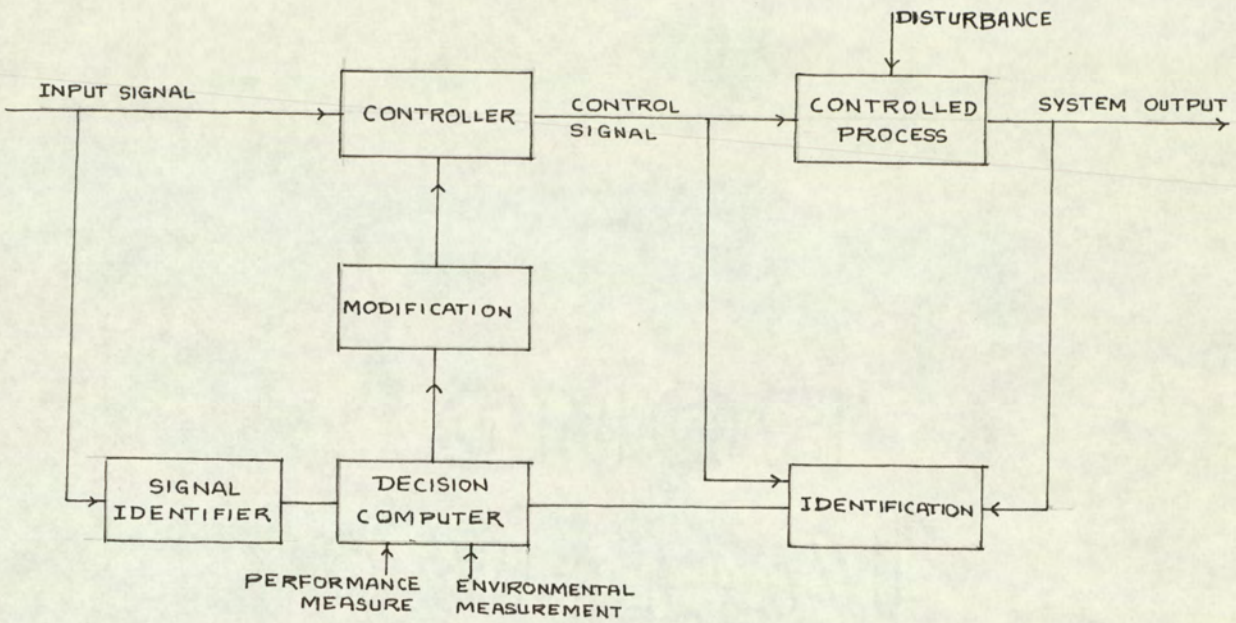


Fig. 6-3 Basic adaptive system.

(i) Identification of the characteristics of a system:-

This is closely related to the evaluation of system parameters. The usual approach to the problem is to introduce a chosen signal into the input of a system and then to observe and measure the output. This can be done in two ways - either an off-line test or an on-line test, - the latter being more prominent in the process industries where the characteristics of large and expensive systems are required. In this case the problem of parameter evaluation is more complex. More often, the system is in operation for all the 24 hours and it would be prohibitively expensive to take it off-line in order to

6.6 contd.

(i) contd.

carry out experiments. Furthermore, if the experiment needs a disturbance signal at the input, this disturbance has to be applied together with the normal operating input and its amplitude must be small, so that the system is not disturbed too far from its optimum operating condition. Another reason for the amplitude of the disturbance signal to be small is to ensure that the characteristics so obtained, describe a system which could be considered to be linear about its operating condition. (This is important, because most of the theories of feedback control, stability of systems etc., are only valid for linear systems).

The usual well established methods used for the evaluation are

(a) Sinusoidal testing

(b) A step response testing

(c) Impulse response testing

A more recent approach is concerned with the use of a particular type of disturbance signal or "forcing function" for the system evaluation - the random signal. We shall describe this briefly.

A random signal $x(t)$, from a noise generator, is applied as excitation to the input of the system and the

6.6 contd.

(i) contd.

Cross-correlation between the signal and the resulting output $y(t)$ is obtained by analogue computing methods. With a white noise input, the Cross-correlation function so computed is an amplitude/Time curve which, it can be shown, is the impulse response of the system multiplied by a constant. This curve, therefore, characterises the system.

For understanding the theoretical basis, it is important to remember the following relationships:

(a) the convolution integral which relates the output signal $y(t)$ with the input signal $x(t)$

$$y(t) = \int_0^\infty h(\tau)\ x(t-\tau)\,d\tau \qquad (6\text{-}6.1)$$

where $h(t)$ is the system impulse response and τ the time delay between the application of an instantaneous input signal and the measurement of the output.

(b) The Auto-correlation function of $x(t)$ is

$$C_{xx}(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\ x(t+\tau)\,dt \qquad (6\text{-}6.2)$$

(c) The Cross-correlation function of $y(t)$

6.6 contd.

(c) contd.

$$C_{xy}(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\, y(t+\tau)\, dt \qquad (6\text{-}6.3)$$

when $x(t) = y(t)$ the cross-correlation function reduces to the autocorrelation function.

Since the impulse response must be zero before the input is applied, $h(t) = 0$ for $t < 0$ and the lower limit of the convolution integral can be extended to infinity so that

$$y(t) = \int_{-\infty}^{\infty} h(s)\, x(t-s)\, ds \qquad (6\text{-}6.4)$$

and then

$$C_{xy}(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t) \left\{ \int_{-\infty}^{\infty} h(s)\, x(t+\tau-x)\, ds \right\} dt$$

$$= \int_{-\infty}^{\infty} h(s) \left\{ \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\, x(t+\tau-s)\, dt \right\} ds$$

$$= \int_{-\infty}^{\infty} h(s)\, C_{xx}(\tau-s)\, ds \qquad (6\text{-}6.5)$$

Comparing (6-6.5) with the convolution integral, it can be seen that if a signal whose autocorrelation function is $C_{xx}(\tau)$ is applied to a system with impulse response $h(t)$,

6.6 contd.

(c) contd.

the cross-corrleation function of the input and output signal is equal to the time response of the system when subjected to an input signal $C_{xx}(t)$.

In particular, if the input signal is white noise, its power density spectrum is flat i.e.

$$\phi_{xx}(\omega) = 2\pi K \qquad (6\text{-}6.6)$$

The relationship between the autocorrelation function and the power density spectrum is the Fourier transform -

$$C_{xx}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_{xx}(\omega)\, e^{i\omega\tau}\, d\omega \qquad (6\text{-}6.7)$$

which becomes

$$C_{xx}(\tau) = K \int_{-\infty}^{\infty} e^{i\omega\tau}\, d\omega$$

$$= K\, \delta(\tau) \qquad (6\text{-}6.8)$$

$\delta(\tau)$ being the Dirac-delta function. Thus (6-6.5) becomes

$$C_{xy}(\tau) = K \int_{-\infty}^{\infty} h(s)\, \delta(\tau - s)\, ds$$

$$= K\, h(\tau) \qquad (6\text{-}6.9)$$

i.e. for a white noise input, the Cross-correlation function of the input and the output is a constant times

6.6 contd.

(c) contd.

the impulse response. (Note that any noise whose power density spectrum is flat over a frequency range much greater than the bandwidth of the system may be considered as being white noise).

The main advantage of this technique over the usual ones is that the experiment can be performed while the system is operating in its normal mode, thus making it necessary to disconnect the system from its associated components. This is possible since the noise excitation energy is spread over a wide frequency range, with a resulting low noise intensity that will not affect the normal operation of the plant and its controls.

The main disadvantage of the method is the long time required to obtain an accurate cross-correlation function, (infinite in the ideal case). This is overcome by generating what is called a periodic white noise. This type of noise would have to have the same type of autocorrelation function as white noise (i.e. an impulse) which would be repeated with a period T/2. i.e.

$$C_{xx}(\tau) = \frac{2}{T} \int_{0}^{T/2} x(t)\, x(t+\tau)\, dt \qquad (6\text{-}6.10)$$

6.6 contd.

(c) contd.

because of its periodicity, or for argument $(\tau-s)$,

$$C_{xx}(\tau-s) = \frac{2}{T}\int_0^{T/2} x(t)\, x(t+\tau-s)\, dt \qquad (6\text{-}6.11)$$

and then

$$C_{xy}(\tau = \int_{-\infty}^{\infty} h(s)\left\{\frac{2}{T}\int_0^{T/2} x(t)\, x(t+\tau-s)dt\right\} ds;$$

$$= \frac{2}{T}\int_0^{T/2} x(t)\left\{\int_{-\infty}^{\infty} h(s)\, x(t+\tau-s)\, ds\right\} dt$$

$$= \frac{2}{T}\int_0^{T/2} x(t)\, y(t+\tau)dt \qquad (6\text{-}6.12)$$

Therefore, by the use of such periodic white noise, the cross-correlation function may be computed to its full accuracy by integration over one period of noise only.

Because of the periodicity of this noise, the cross-correlation function may now be written as

$$C_{xy}(\tau) = K\left\{h(\tau) + h(T/2+\tau) + h(T+\tau)+\ldots\right\}; \qquad (6\text{-}6.13)$$

but if it is so arranged that the impulse response decays to zero in a time less than T, (6-6.13) becomes

$$C_{xy}(\tau) = Kh(\tau) \qquad (6\text{-}6.14)$$

as before.

6.6 contd.

(ii) The decision problem:-

This deals with the development and specification of analytical methods, which help in evaluating system performance and from which a strategy for achieving optimsation can be evolved. The most common method for system evaluation is the use of performance index which is a functional relationship involving system characteristics in such a manner that optimum performance may be determined from it. Performance index, thus, can be considered to be any criteria for measuring system quality.

A Note on performance index:-

Consider the general control system,

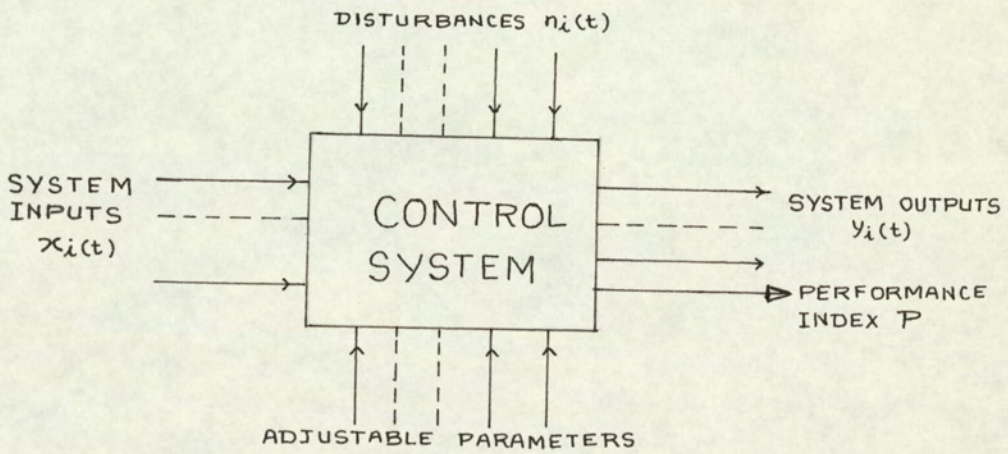


Fig. 6-4

where the variables $x_i(t)$ and $y_i(t)$ are the normal input and output of the system. The signals $n_i(t)$ are unknown disturbances. P is a measure of performance of the system,

6.6 contd.

(ii) contd.

which may be the accuracy of control, efficiency or profit. There may be practical limitations on the maximum control, power available or restrictions on the quality of the product. The different and possibly conflicting requirements (.e.g. maximum profit with a minimum guaranteed product quality), can normally be combined into a single figure-of-merit or performance index.

In general, P will be a function of the system variables, that is the system inputs, the disturbances and the settings of the parameters x_i of the system, some of which are adjustable. The performance has to be controlled by adjusting these parameters and the maximum value of P can usually be found for a particular setting of the parameters.

(iii) The Modification problem:-

This is concerned with the actual adjustment of the system so as to optimize its performance. The general process is called the control signal modification, since the object is to modify error or control signal to achieve optimum performance. This is achieved through either of the two methods - (a) Parameter adjustment and (b) Control system synthesis. In parameter adjustment fig.(a), the basic

6.6 contd.

(iii) contd.

configuration of the conventional control system is maintained intact, and the means of automatically adjusting the parameters toward the optimum is superimposed. In control signal modification fig.(b), the conventional configuration is entirely abandoned, and the whole control function is assumed by the adaptive computer.

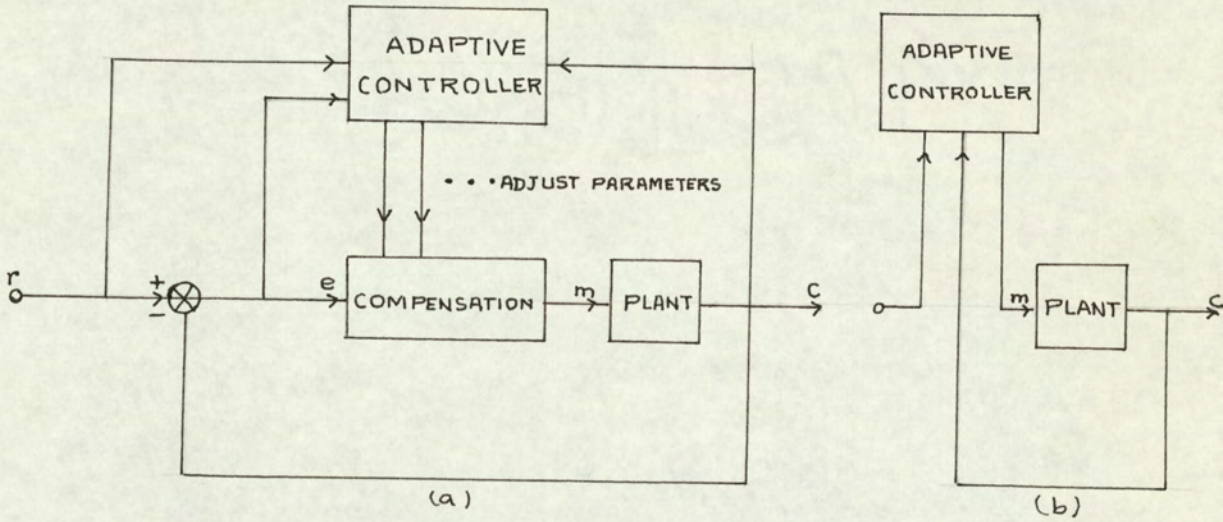


Fig. 6-5 Control-signal modification. (a) Parameter adjustment. (b) Control signal synthesis.

The field of adaptive control is not yet developed and is still concerned with the control of relatively simple processes. With the increasing use of adaptivity in future, one can foresee digital computer forming a basis for the design of the controller functions of identification, decision and modification.

6.6 contd.

(iii) contd.

To summarise, an adaptive control is useful under either of the two conditions: (1) when the process characteristics are unknown and the designer does not know the environmental factors which affect the process dynamics, and (2) when process dynamics changed markedly and unpredictably with time or environmental conditions. Adative control is a logical broadening of the familiar concept of feedback control.

SECTION II.

6·7 AUTOMATIC CONTROL IN ELECTRICITY SUPPLY:

In an electricity supply industry, the generated power, the power demand, and the power flow, continually change randomly. There are about 200 power stations and over 600 grid substations in England and Wales, which are interconnected by some,8,500 miles of lines and cables. There are also some tens of thousands of circuit breakers and switches. This complex system has to be manipulated to meet all the demands at the Area Boards, in the most economical way. The manipulation techniques are based on past experience and although the modern equipment and techniques have brought about new facilities, up to now it has not been done automatically. The use of analogue and digital computers is being developed to provide the control engineer with the information

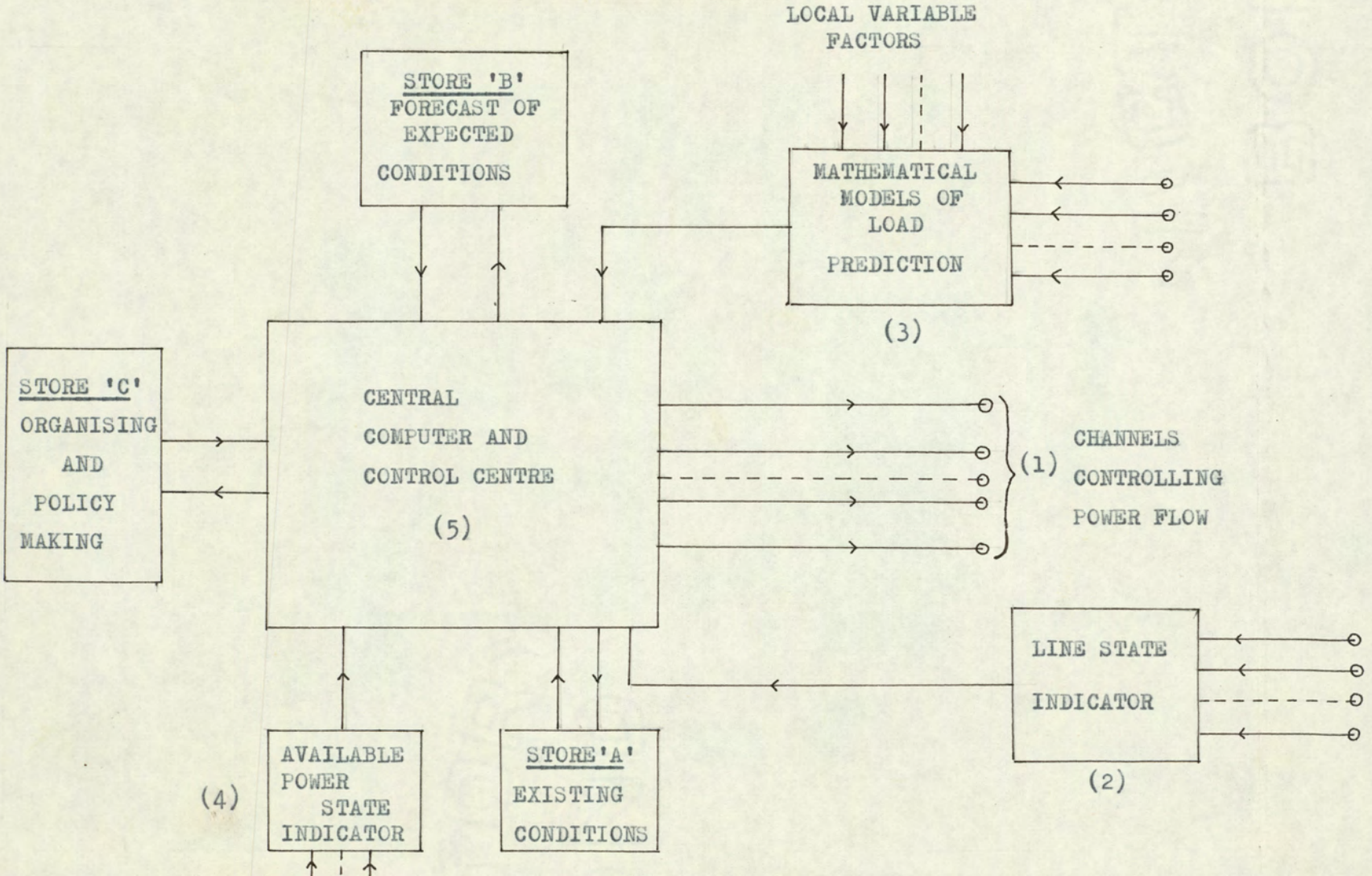
6.7 contd.

he needs to operate the system. The final decision, however, is still the engineer's, and some degree of personal judgement on the part of a control engineer is required for this. So, in this situation, the true closed-loop system should mean the replacement of the engineer by means of an equipment, which should be flexible enough to cope with the problems and this cannot be achieved without the combination of an automatic equipment of high organising ability.

To reflect on the literature on the present state of automatic control in electricity supply, which is obtainable from the various conference publications, one can conceive a closed-loop general system which may in future become capable of controlling the power-flow automatically. This general system for power-flow control can be best described by means of a block diagram. In the block diagram appearing on the next page, we have indicated by

1) The channels which control the power flow, i.e. they control each switchgear, trip relay etc., in the flow system.
2) The line-state indicator - which indicates the state of the power flow and the state of each interconnection or trip (i.e. whether on or off).
3) The prediction models - these models should represent each area of concern and should give us predictions on the expected electricity demand in each area. There should also

A GENERAL SYSTEM FOR POWER FLOW CONTROL



6.7 contd.

3) contd.

be a feedback from a monitor which should observe the consumption of power in each area; this is used to modify the prediction. There are also other variable parameters in the models which can be set to compensate for local conditions viz. weather changes etc.

4) Available power state indicator - this should provide information about the power generated in each station and monitor the output of each substation.

5) Central computer - Actually, at present, there is a central computer in London where they collect some basic data, but what we need is a computer or a set of computers working in conjunction, which should receive information about the state of the power flow, line and expected demand and evaluate existing conditions. There will be a continuous data logging and updating of information and two separate stores may be used for this - store 'A' for logging of existing conditions and store 'B' for expected conditions which should include not only the load forecasts, but also the information about maintenance, diversions and other expected and planned changes.

Now, as one can see, up to here the system would provide a very efficient backing and information source for the control engineer. So part of the function of the central computer and the store 'C' should be to close the loop and

6.7 contd.

assume the functions of the engineer. The store should act in part as the experience of the engineer and special programmes should be provided to act after evaluation of the existing conditions received from the rest of the equipment and planning the policy and then giving command to the controlling channels. This, as one will realise, is the most difficult part of the system and a lot of experience on the use of Games and Decision theory will have to be utilised before a really workable model can be made to cope with all the complexities and emergencies that might occur. The physical realisation of such a system should need a great amount of effort on the part of mathematicians working in coordination with the control engineers.

CHAPTER 7

GENERAL DISCUSSION

In the previous chapters, a survey of the present day important methods of prediction has been carried out and also their usefulness in case of a particular problem of Electricity demand prediction has been discussed. The models of exponentially weighted prediction have been tested against the data obtained for a particular area. One feels that many more tests will have to be carried out before one could decide on a predictor giving an optimum prediction. These models form an important part of the proposed general system for power flow control as discussed in Section II of Chapter 6. The physical realization of such a system will involve a great amount of research by mathematicians and control engineers. From a mathematician's point of view, the main area for research will be in determination of the optimality of such a system and for this reason we shall confine the following discussion to the concept of an optimum system.

The term "Control System" generally brings to mind a servomechanism or a simple temperature regulator. Such a device is characterised by the fact that it detects when the output differs from the desired value and feeds back to the input a signal proportional to this difference, which in turn causes the output to change in the direction tending to decrease this difference. Eventually then, the output is made to approximate the desired

value to an arbitrary tolerance. Such elementary application of the feedback principle is an adequate guide to the design of the simple control systems, but in more complex control problems a much deeper theory is needed.

In the more general control situations the problem is still one of adjusting the inputs (adjustment based on the measurement of system outputs and the knowledge of the environment interacting with the system) so as to optimize a performance criterion. In this case, it is no longer obvious when the optimum has been reached and what changes in the several inputs will result in an improvement.

In practical situations, the significance of methods of optimization reviewed in the last chapter is of a distinctly different order. One must realise that it is not possible to introduce into optimization calculations all the variables significant in a real problem, and that optimizing the equations is not necessarily equivalent to finding the best solution to a technical problem (the influence of those parameters not taken into account by the mathematical optimization could be significant in the neighbourhood of optimum). Thus the greatest actual difficulties lie in the definition of the parameters which control the system and, to a lesser extent, the precise definition of the system itself or, alternatively, making its realization conform to a calculated expression. The difficulty is practical

because it is sometimes hard to predict precisely enough the probable behaviour of the parameters before the system is in existence. On the theoretical plane, however, the methods of optimization are of considerable interest. Starting from the data of a control system they allow, in fact, an estimate to be made of the best performance that one could in theory hope to obtain from it and also could give an indication in what directions efforts at improvement have the best chance of paying off.

Thus the optimization of a system whatever the system may be, involves several steps of analysis, the principal ones being -

(i) Identifcation of system parameters:-

This has been discussed to some extent in Section 6-6. It is an extremely important operation and varies in difficulty according to the system. Ideally, one must identify a system to the n^{th} degree for the modern optimizing techniques to be rigorously applied. However, as systems are often complex (involve nonlinearities etc.) an absolute identification becomes impossible. Thus a good approximation to a system has to be made and verified by the use of different models and computer techniques.

(ii) The Performance Index:-

This is a function of the control parameters, which, when

(ii) contd.

brought to an extremum minimizes the undesirable elements in the system, such as error, cost, speed etc. In case of a complex index, it is difficult to attain an idealised extremum assumed, since the importance of the individual parameters has to be weighted accordingly.

(iii) Defining the Optimum policy:-

The optimum policy is one which produces an extremum of the index.

In an industry, one sets up a system mainly by using his judgement based on the past experience gained by trial and error plus intuition. The modern approach has to surmount the difficulty in identifying the system parameters and in attaching appropriate weights to each relevant parameter of the performance index. As one will appreciate, a lot of ingenuity and intuition is involved in devising models which behave in a similar fashion to a system so that its study can guide in the optimization of the system itself.

APPENDIX A

226.2	0	0	1	120	24	240	
106.0	97.7	94.3	95.7	98.1	107.4	159.5	243.3
301.8	321.3	306.9	293.2	277.3	281.8	281.3	280.4
295.7	296.9	273.4	269.4	263.1	265.4	237.9	181.9
148.8	141.3	132.6	139.2	140.9	148.4	185.7	253.2
315.6	313.0	307.6	311.9	306.6	304.3	305.0	311.3
315.4	314.3	289.0	278.9	271.0	267.4	246.4	187.7
148.9	138.9	130.6	137.8	138.1	145.0	179.2	244.2
296.8	301.9	284.1	294.9	268.7	281.1	281.1	257.0
295.2	295.8	272.2	263.7	254.5	251.2	231.0	172.7
142.8	131.3	120.9	131.4	135.1	140.2	178.1	249.3
301.2	299.0	289.7	291.3	282.4	281.8	275.7	265.7
292.9	298.8	281.3	275.1	263.0	271.8	248.4	199.9
154.3	143.8	135.0	142.0	142.8	148.5	187.5	258.5
311.3	302.2	290.5	292.8	286.5	285.2	279.4	278.5
284.7	293.0	278.2	275.7	259.1	237.4	200.5	173.3
116.8	108.2	106.6	107.3	110.9	124.9	184.1	285.9
359.6	368.8	353.6	349.5	336.8	338.1	339.1	331.5
295.7	337.1	316.7	314.6	308.6	311.9	278.5	206.3
173.0	161.7	154.3	161.4	163.9	173.3	222.5	318.9
362.2	357.1	341.7	348.3	343.6	346.5	344.5	331.6
349.0	357.8	326.0	313.7	308.1	314.9	281.3	213.3
173.8	161.9	156.4	158.9	159.9	170.3	217.7	306.4
360.4	354.5	343.9	351.8	336.6	333.3	323.5	325.1
343.6	346.1	318.6	313.6	302.8	308.7	263.1	205.6
166.0	156.2	148.7	156.4	159.6	168.5	201.5	300.8
340.8	344.6	323.8	333.3	315.2	305.5	306.2	307.6
325.1	340.5	316.5	309.5	299.9	311.2	275.7	239.6
170.8	157.3	150.3	157.5	159.2	169.2	214.1	309.8
354.6	340.0	324.3	324.3	317.4	315.5	318.6	315.1
320.4	338.9	325.3	318.5	302.0	294.9	266.3	208.1
0.60	0.56	0.52	0.55	0.56	0.59	0.76	1.06
1.31	1.32	1.27	1.27	1.21	1.20	1.19	1.16
1.25	1.28	1.19	1.16	1.12	1.10	0.99	0.78

APPENDIX B

```
ADAPTIVE PREDICTION'

BEGIN REAL DBAR,AA'
      INTEGER SERIES,T,M.L,N'
      READ DBAR,AA,SERIES,T,M.L,N'
BEGIN ARRAY DEMAND(1 :N+T-1),S(-L+1 :N),FORS,ERROR(M+T:N+T-1)'
      REAL ESTS1,ESTS2,ACCABS,ACCSQ,ACCERR,SIGMA,
           INWE,STWE,FIWE,INWA,STWA,
           FIWA,INWS,STWS,FIWS,WE,WA,WS,SD,A'
      INTEGER I,J.K,PAR1'
      FOR I :=(M+T) STEP 1 UNTIL (N+T-1) DO
               FORS(I) :=ERROR(I) :=0'
      FOR I :=-L+1 STEP 1 UNTIL N DO
               S(I) :=0'
      FOR I :=1 STEP 1 UNTIL N+T-1 DO
               READ DEMAND(I)'
      FOR I :=-L+1 STEP 1 UNTIL 0 DO
               READ S(I)'
      WAIT' COMMENT LOAD DATA TAPE 2'
      READ PAR1,INWE,STWE,FIWE,INWA,STWA,FIWA,INWS,STWS,FIWS'
      FOR WE:=INWE STEP STWE UNTIL FIWE DO
      FOR WA:=INWA STEP STWA UNTIL FIWA DO
      FOR WS:=INWS STEP STWS UNTIL FIWS DO
      BEGIN
      ACCABS:=ACCSQ:=SD:=ACCERR:=SIGMA:=0'
      ESTS1:=DBAR' A:=AA'
      FOR I:=1,I+1 WHITE I LESS N DO
      BEGIN
      ESTS2:=WE*DEMAND(I)/S(I-L)+(1-WE)*(ESTS1+A)'
      S(I):=WS*DEMAND(I)/ESTS2+(1-WS)*S(I-L)'
      A:=WA*(ESTS2-ESTS1)+(1-WA)*A'
      ESTS1:=ESTS2'
      IF I CHEQ M THEN
      BEGIN
      FORS(I+T):=(ESTS2+T*A)*S(I-L+T)'
      ERROR(I+T):=FORS(I+T)-DEMAND(I+T)'
      ACCABS:=ACCABS+ABS(ERROR(I+T))'
      ACCERR:=ACCERR+ERROR(I+T)'
      ACCSQ:=ACCSQ+ERROR(I+T)**2'
      END'
      END'
      SD:=SQRT(ACCSQ/(N-M-1))'
      SIGMA:=SQRT(((N-M)*ACCSQ-ACCERR**2)/(N-M)**2)'
      IF PAR1=0 THEN
     PRINT ££L??,SERIES,SAMELINE,ALIGNED(1,3),WE,WA,WS,
              SCALED(5),PREFIX(££S1??),
              ACCABS,SD,ACCERR,SIGMA
```

CONTD.

```
        ELSE IF PAR1=1 THEN
        BEGIN
        BEGIN
        FOR I :=N, I+1 WHILE I LESS N+1 DO
        BEGIN
        ESTS2:=WE*DEMAND(I)/S(I-L)+(1-WE)*(ESTS1+A)'
        S(I):=WS*DEMAND(I)/ESTS2+(1-WS)*S(I-L)'
        A:=WA*(ESTS2-ESTS1)+(1-WA)*A'
        ESTS1:=ESTS2'
        END'
        END'
        PRINT ££L?OPTIMUM FORECAST SERIES?,££S??,SAMELINE,SERIES,
              ££L2?ESTS2=?ALIGNED(9,0),SAMELINE,ESTS2,££S8?A=?,A,
              ££L2?WE=?,ALIGNED(1,3),SAMELINE,WE,££S4?WA?,WA,££S4?WS?,WS,
              ££L4S7?I?,££S8?FORS(I)?,££S8?ERROR(I)?'
        FOR I :=M+T STEP 1 UNTIL N+T-1 DO
        PRINT ££L??,I,SCALED(5),SAMELINE,££S4??,FORS(I),££S5??,ERROR(I)'
        FOR I :=N-L STEP 1 UNTIL N DO
        PRINT ££L??,I,SAMELINE,ALIGNED(2,5),S(I)'
        END ELSE STOP'
END'
END'
END PROGRAM'
```

SMOOTHING CONSTANT SAMPLES USED IN OPTIMISING FORECASTS.

APPENDIX C.

W_e	W_a	W_s	S.D.
0.850	0.000	0.850	1.4066
0.850	0.000	0.900	1.4052
0.850	0.000	0.950	1.4042
0.850	0.050	0.850	1.4687
0.850	0.050	0.900	1.4671
0.850	0.050	0.950	1.4658
0.900	0.000	0.850	1.4173
0.900	0.000	0.900	1.4149
0.900	0.000	0.950	1.4127
0.900	0.050	0.850	1.4813
0.900	0.050	0.900	1.4786
0.900	0.050	0.950	1.4761
0.950	0.000	0.850	1.4442
0.950	0.000	0.900	1.4422
0.950	0.000	0.850	1.4442
0.950	0.000	0.900	1.4422
0.950	0.000	0.950	1.4402
0.950	0.050	0.850	1.5115
0.950	0.050	0.900	1.5093
0.950	0.050	0.950	1.5071

TIME SERIES PREDICTION USING OPTIMUM SMOOTHING (S.D. = 1.4042)

APPENDIX D

ESTS2= 266 A = 0

WE= 0.850 WA= 0.000 WS= 0.950

I	PRE(I)	ERROR(I)
121	131.36	14.56
122	110.68	2.48
123	101.98	- 4.62
124	111.92	4.62
125	109.61	- 1.29
126	116.76	- 8.14
127	159.59	- 24.51
128	252.93	- 32.97
129	344.76	- 14.84
130	360.04	- 8.76
131	355.13	1.53
132	354.83	5.33
133	333.61	- 3.19
134	335.22	- 2.88
135	332.86	- 6.24
136	329.49	- 2.01
137	354.45	58.75
138	310.92	- 26.18
139	331.26	- 5.44
140	308.61	- 5.99
141	301.64	- 6.96
142	302.93	- 9.87
143	279.57	1.07
144	221.21	14.91
145	155.42	- 17.58
146	160.76	- 0.94

I	PRE(I)	ERROR(I)
147	153.19	- 1.11
148	160.89	- 0.51
149	165.02	1.12
150	174.37	1.07
151	226.05	3.55
152	311.85	- 7.05
153	385.88	23.68
154	367.56	10.46
155	345.22	3.52
156	342.66	- 5.64
157	332.06	- 11.54
158	340.71	- 5.79
159	341.19	- 3.31
160	334.59	2.99
161	345.29	- 3.71
162	369.93	12.13
163	333.07	7.07
164	319.62	5.92
165	302.65	- 5.45
166	302.93	- 11.97
167	280.55	- 0.75
168	221.06	7.76
169	163.86	- 9.94
170	160.37	- 1.53
171	153.33	- 3.07
172	162.68	3.78
173	162.89	2.99
174	170.43	0.13
175	221.65	3.95
176	306.88	0.48
177	267.49	7.09

I	PRE(I)	ERROR(I)
178	365.11	10.61
179	343.67	- 0.23
180	345.60	- 6.20
181	336.16	- 0.44
182	334.58	1.28
183	328.87	5.37
184	314.60	- 10.50
185	337.37	- 6.23
186	361.50	15.40
187	323.24	4.64
188	312.15	- 1.45
189	303.06	0.26
190	299.42	- 9.28
191	273.97	10.87
192	206.97	1.37
193	159.36	- 6.64
194	152.53	- 3.67
195	147.83	- 0.87
196	154.04	- 2.36
197	159.51	- 0.09
198	170.06	1.56
199	219.04	17.54
200	287.65	- 13.15
201	357.40	16.60
202	346.25	1.65
203	334.27	10.47
204	327.80	- 5.50
205	317.80	2.60
206	313.53	8.03
207	301.91	- 4.29
208	298.48	- 9.12
209	318.69	- 6.41

I	PRE(I)	ERROR(I)
210	338.92	- 1.58
211	317.03	0.53
212	310.35	0.85
213	299.20	- 0.70
214	297.72	- 13.48
215	272.85	- 2.85
216	216.27	- 23.33
217	184.04	13.24
218	159.34	2.04
219	149.31	- 0.99
220	155.89	- 1.61
221	160.42	1.22
222	169.60	0.40
223	217.32	3.22
224	308.05	- 1.75
225	365.36	10.76
226	361.53	21.53
227	331.40	7.10
228	330.08	5.78
229	309.72	- 7.68
230	313.38	- 2.12
231	312.04	- 6.56
232	310.96	- 4.14
233	326.81	6.41
234	335.30	- 3.60
235	314.97	- 10.33
236	317.33	- 1.17
237	307.82	5.82
238	302.55	7.65
239	260.03	- 6.27
240	211.11	3.01

APPENDIX E

$W_e = 0.900$ $W_a = 0.000$ $W_s = 0.900$

TIME SERIES PREDICTION USING NEAR OPTIMUM SMOOTHING. (S.D. = 1.4149).

Period	Predicted Values	Error
121	132	15
122	111	2
123	102	-5
124	112	5
125	110	-1
126	117	-8
127	160	-24
128	254	-32
129	347	-12
130	361	-7
131	355	1
132	354	5
133	333	-3
134	335	-3
135	333	-6
136	330	-2
137	355	60
138	308	-29
139	312	-5
140	309	-6
141	302	-6
142	303	-9
143	280	1
144	220	14
145	156	-17
146	162	-0
147	152	-2
148	161	-0
149	165	1
150	174	1
151	225	2
152	311	-8
153	388	26
154	367	10
155	345	3
156	342	-6
157	332	-12
158	341	-5
159	342	-3
160	335	4
161	350	1
162	366	8
163	333	7
164	319	5

APPENDIX E (contd)

Period	Predicted Values	Error
165	303	-5
166	303	-12
167	281	0
168	221	8
169	164	-10
170	162	-0
171	153	-4
172	163	4
173	163	3
174	170	-0
175	221	3
176	305	-1
177	370	10
178	365	11
179	343	-1
180	345	-7
181	336	-1
182	335	1
183	329	6
184	315	-10
185	342	-2
186	359	13
187	322	4
188	312	-2
189	303	-0
190	299	-10
191	275	12
192	207	1
193	159	-7
194	154	-2
195	147	-1
196	155	-2
197	160	-0
198	170	1
199	218	17
200	285	-16
201	361	20
202	347	2
203	333	10
204	326	-7
205	317	2
206	313	8
207	302	-4
208	298	-9
209	322	-3

APPENDIX E (contd)

Period	Predicted Values	Error
210	339	-2
211	317	0
212	310	0
213	299	-1
214	297	-15
215	275	-1
216	217	-23
217	184	13
218	160	2
219	149	-1
220	156	-1
221	161	1
222	169	-0
223	218	3
224	304	-5
225	369	15
226	362	22
227	330	6
228	328	4
229	309	-8
230	314	-1
231	312	-7
232	311	-4
233	330	10
234	335	-4
235	315	-11
236	317	-1
237	308	6
238	301	6
239	261	-5
240	211	3

APPENDIX F

DETERMINATION OF SEASONAL ADJUSTMENT (FOR USE IN APPENDIX G).

Actual Demand	Cumulative Sum	Cumulative Sum written after one Cycle	24 × M.A.	M.A.	Difference
106	106				
98	204				
94	298				
96	394				
98	492				
107	599				
160	759				
243	1002				
302	1304				
321	1625				
307	1932				
293	2225				
277	2502				
282	2784				
281	3065				
280	3345				
296	3641				
297	3938				
273	4211				
269	4480				
263	4743				
265	5008				
238	5246				
182	5428		5428	226	-44
149	5577	106	5471	228	-79
141	5718	204	5514	230	-89
133	5851	298	5553	231	-98
139	5990	394	5596	233	-94

141	6131	492	5639	235	-94
148	6279	599	5680	237	-89
186	6465	759	5706	238	-52
253	6718	1002	5716	238	15
316	7034	1304	5730	239	77
313	7347	1625	5722	238	75
308	7655	1932	5723	238	70
312	7967	2225	5742	239	73
307	8274	2502	5772	240	67
304	8578	2784	5794	241	63
305	8883	3065	5818	242	63
311	9194	3345	5849	243	68
315	9509	3641	5867	244	71
314	9823	3938	5885	245	69
289	10112	4211	5901	246	43
279	10391	4480	5911	246	33
271	10662	4743	5919	246	25
267	10929	5008	5921	246	21
246	11175	5246	5929	247	-1
188	11363	5428	5935	247	-59
149	11512	5577	5935	247	-98
139	11651	5718	5933	247	-108
131	11782	5851	5931	247	-116
138	11920	5990	5930	247	-109
138	12058	6131	5927	247	-109
145	12203	6279	5924	247	-102
179	12382	6465	5917	247	-68
244	12626	6718	5908	246	-2
297	12923	7034	5889	245	52
302	13225	7347	5878	245	57
284	13509	7655	5854	244	40
295	13804	7967	5837	243	52
269	14073	8274	5799	242	27

281	14354	8578	5776	241	40
281	14636	8883	5753	240	41
257	14893	9194	5699	237	20
295	15188	9509	5679	237	58
296	15484	9823	5661	236	60
272	15756	10112	5644	235	37
264	16020	10391	5629	234	30
255	16275	10662	5613	234	21
251	16526	10929	5597	233	18
231	16767	11175	5582	232	-1
173	16930	11363	5567	232	-59
143	17073	11512	5561	231	-88
131	17204	11651	5553	231	-100
121	17325	11782	5543	231	-110
131	17456	11920	5536	231	-100
135	17591	12058	5533	230	-95
140	17731	12203	5528	230	-90
178	17909	12382	5527	230	-52
249	18158	12626	5532	230	19
301	18459	12923	5536	230	71
299	18758	13225	5533	230	69
290	19048	13509	5539	230	60
291	19339	13804	5535	230	61
282	19621	14073	5548	231	51
282	19903	14345	5558	232	50
276	20179	14636	5543	230	46
266	20445	14893	5552	231	35
293	20738	15188	5550	231	62
299	21037	15484	5553	231	68
281	21318	15756	5562	232	49
275	21593	16020	5573	232	43
263	21856	16275	5581	232	31
272	22128	16526	5602	233	39

248	22376	16757	5619	234	14
200	22576	16930	5646	235	-35
154	22730	17073	5657	235	-81
144	22874	17204	5670	236	-92
135	23009	17325	5684	237	-102
142	23151	17456	5695	237	-95
143	23294	17591	5703	238	-95
149	23443	17331	5712	238	-99
188	23631	17909	5722	238	-50
259	23890	18158	5732	239	20
311	24201	18459	5742	239	72
302	24503	18758	5745	239	63
291	24794	19048	5746	239	52
293	25087	19331	5756	240	53
287	25374	19621	5753	240	47
285	25659	19903	5756	240	45
279	25938	20179	5759	240	39
279	26217	20445	5772	240	39
285	26502	20738	5764	240	45
293	26795	21037	5758	240	53
278	27073	21318	5755	240	38
276	27349	21593	5756	240	36
259	27608	21856	5752	240	19
237	27845	22128	5717	238	-1
201	28046	22376	5670	236	-35
173	28219	22576	5643	235	-62

APPENDIX G

TIME SERIES PREDICTION USING HOLT'S SIMPLE SMOOTHING. (S.D.=?)

	PREDICTED	ACTUAL	A-P	$\hat{x}_e(t)=\hat{x}_e(t-1) + 0.9\ e_t$	SEASONAL VARIATIONS	P_1
				235		156
121	156	117	-39	200	- 79	111
122	111	108	- 3	197	- 89	99
123	99	107	8	199	- 98	105
124	105	107	2	201	- 94	107
125	107	111	4	205	- 94	116
126	116	125	9	213	- 89	161
127	161	184	23	234	- 52	250
128	250	286	36	266	16	344
129	344	360	16	280	78	355
130	355	369	14	293	75	364
131	364	354	-10	284	71	357
132	357	350	- 7	278	73	345
133	345	337	- 8	271	67	334
134	334	338	4	275	63	339
135	339	339	0	275	64	343
136	343	332	-13	263	68	335
137	335	296	-39	228	72	297
138	297	337	40	264	69	308
139	308	317	9	272	44	306
140	306	315	9	280	34	306
141	306	309	3	283	26	305
142	305	312	7	289	22	288
143	288	279	- 9	281	- 1	222
144	222	206	-16	267	- 59	169
145	169	173	4	271	- 98	163
146	163	162	- 1	270	-108	154
147	154	154	0	270	-116	161
148	161	161	0	270	-109	161
149	161	164	3	273	-109	171

150	171	173	2	275	-102	207
151	207	223	16	289	- 68	287
152	287	319	32	319	- 2	371
153	371	362	- 9	311	52	368
154	368	357	-11	301	57	342
155	342	342	0	301	41	353
156	353	348	- 5	296	52	324
157	324	344	20	314	28	354
158	354	347	- 7	308	40	349
159	349	345	- 4	304	41	325
160	325	332	7	310	21	368
161	368	349	-19	293	58	354
162	354	358	4	297	61	334
163	334	326	- 8	290	37	321
164	321	314	- 7	284	31	306
165	306	308	2	286	22	304
166	304	315	11	296	18	295
167	295	281	-14	283	- 1	224
168	224	213	-11	273	- 59	185
169	185	174	-11	263	- 88	163
170	163	162	- 1	262	-100	152
171	152	156	4	266	-110	166
172	166	159	- 7	260	-100	165
173	165	160	- 5	255	- 95	165
174	165	170	5	260	- 90	208
175	208	218	10	269	- 52	288
176	288	306	28	294	19	366
177	366	360	- 6	289	72	358
178	358	355	- 3	286	69	347
179	347	344	- 3	283	61	345
180	345	352	7	289	, 62	341
181	341	337	- 4	285	52	336
182	336	333	- 3	282	51	328
183	328	324	- 4	278	46	314

184	314	325	11	288	36	351
185	351	344	- 7	282	63	350
186	350	346	- 4	278	68	327
187	327	319	- 8	271	49	315
188	315	314	- 1	270	44	302
189	302	303	1	271	32	310
190	310	309	- 1	270	39	285
191	285	263	-22	250	15	215
192	215	206	- 9	242	- 35	161
193	161	166	5	247	- 81	155
194	155	156	1	248	- 92	146
195	146	149	3	251	-102	156
196	156	156	0	251	- 95	156
197	156	160	4	255	- 95	156
198	156	169	13	267	- 99	217
199	217	202	-15	253	- 50	274
200	274	301	27	277	21	350
201	350	341	- 9	269	73	332
202	332	345	13	281	63	334
203	334	324	-10	272	53	326
204	326	333	7	278	54	326
205	326	315	-11	268	48	314
206	314	306	- 5	263	46	302
207	302	306	4	267	39	306
208	306	308	2	269	39	315
209	315	325	10	278	46	332
210	332	341	9	286	54	324
211	324	317	- 7	280	38	317
212	317	310	- 7	274	37	293
213	293	300	7	280	19	279
214	279	311	32	309	- 1	274
215	274	276	2	311	- 35	249
216	249	240	- 9	303	- 62	

REFERENCES.

1) BARTLETT, M.S., "An introduction to stochastic processes" (Cambridge Uni.Press, 1966).

2) BELLMAN, R. and KALABA, R., "Dynamic programming and modern control theory" (Academic press, 1965).

3) BENDAT, J.S., "Principles and applications of random noise theory" (Wiley, 1958).

4) BENDAT, J.S., "A general theory of linear prediction and filtering" (J.Soc.Indust.Appl.Math., Vol.4, No.3, Sept., 1956, pp.131-151).

5) BENDAT, J.S. and PIERSOL, A.G. "Measurement and analysis of random data" (Wiley, 1966).

6) BLACKMAN, R.B. and TUKEY, J.W., "The measurement of power spectra" (Dover, 1958).

7) BODE, H.W. and SHANNON, C.E., "A simplified derivation of linear least square smoothing and prediction theory" (Proc.IRE vol.38, 1950, pp.417-25).

8) BOX, G.E.P. and JENKINS, G.M., "Some Statistical aspects of adaptive optimization and control" (J.R.Stat.Soc.B,24, 1962, pp.297-331).

9) BROWN, R.G., "Statistical forecasting for inventory control" (New York McGraw-Hill, 1959)

10) BROWN, R.G., "Smoothing forecasting and prediction of discrete time series" (Prentice Hall, 1963).

11) CHENG, D.K., "Analysis of linear systems" (Addison-Wesley,1959)

REFERENCES (contd)

12) CLARK, R.N., "Introduction to automatic control systems" (Wiley, 1966).

13) COX, D.R., "Prediction by exponentially weighted moving averages and related methods" (J.R.Stat.Soc.B, vol 23, No.1, 1963, pp.414-22)

14) CRANK, J., "Mathematics and Industry" (Oxford Uni.Press, 1962).

15) CUNDY and ROLLETT, "Mathematical models" (second edition, 1961).

16) DAVIES, W.D.T., "Random signal testing for evaluating system dynamics" (Wireless world, August, 1966, pp.407-12).

17) DEUTSCH, R., "Estimation theory" (Prentice Hall, 1965)

18) DRAVID, V.N., "An appraisal of servo-system analytical methods" (Control, vol 9, 1965, pp.432-35)

19) FAGAN, T.L., "Exponentially smoothing for prediction of reliability growth" (Trans.ASQC Annual Technical Conference, 1965, pp.353-60).

20) FARMER, E.D., "A method of prediction for non-stationary processes and its application to the problem of load estimation" (Proc. IFAC, 1963,pp.47-54)

21) FARRINGTON, G.H., "Elements of automatic control" (Chapman and Hall, 1951).

22) FORMBY, J.A., "An introduction to the mathematical formulation of self organising systems"(SPON, 1965)

REFERENCES (contd)

23) GRITTON, J., "Short term demand forecasting for Ice cream products" (M.Sc. dissertation, Dept. I.A., Uni. of Aston, B'ham, 1966).

24) HARRISON, P.J., "Short-term sales forecasting" (J.Appl.Stat.C, vol 14, No.1, pp.102-139).

25) HRONES, Dr.J.A., "On automatic control" (Proc. of the Symposium on automatic control, Jan., 1960, pp.1-12)

26) ICI Monograph No.2 "Short term forecasting" (Oliver and Boyd, 1964)

27) IEE Conference Publication, "Manchester conference" (No.16, part 1, March, 1966).

28) LANNING, J. and BATTIN, L., "Random processes in automatic control" (McGraw-Hill, New York, 1956).

29) MACHOL, R.E., "Systems engineering Handbook" (McGraw Hill, 165)

30) MACMILLAN, R.H., and others, "Progress in automatic control" (vol 1, Heywood, 1962).

31) MUIR, A., "Automatic sales forecasting" (Computer Journal, Vol.1, 1958, pp.113-16).

32) PASK, G., "An approach to Cybernetics" (Hutchinson, 1961)

33) PITMAN, R.J.G., "Automatic control systems explained" (Macmillan 1966).

34) ROBSON, J.D., "An introduction to random vibration" (Am. Elsevier, 1963).

35) SCOTT, M.R., BROOKS, A., LEE, A.W. RAMSAY, H.B., "The use of mathematics in the Electrical Industry" (Pitman, 1966)

36) SUTTON, O.G., "Mathematics in action" (London Bell, 1954)

37) TRUXAL, J.G., "Adaptive control - A survey" (Proc.IFAC, 1963, pp.386-92).

38) TSIEN, H.S., "Engineering Cybernetics" (McGraw-Hill, 1954)

39) WEAVER, W., "Lady luck - the theory of probability" (Heinemann, 1964).

40) WHITAKER, E.T., and ROBINSON, G., "The calculus of observations" (second ed., Blackie, 1926).

41) WHITTLE, P., "Prediction and regulation by linear least square methods" (EUP, 1963).

42) WIENER, N., "Cybernetics" (second ed., New York Wiley, 1961)

43) WIENER, N., "Extrapolation interpolation and smoothing of stationary time series" (New York Wiley, 1950).

44) WINTER, P., "Forecasting sales by exponentially weighted moving averages" (Management science, Vol. 6, No.3, 1960, pp324-42).

45) ZADEH, L.A., and RAGAZZINI, J.R. "Optimum filters for the prediction of signals in noise" (Proc. IRE, vol.40, 1952, pp.1223-31)

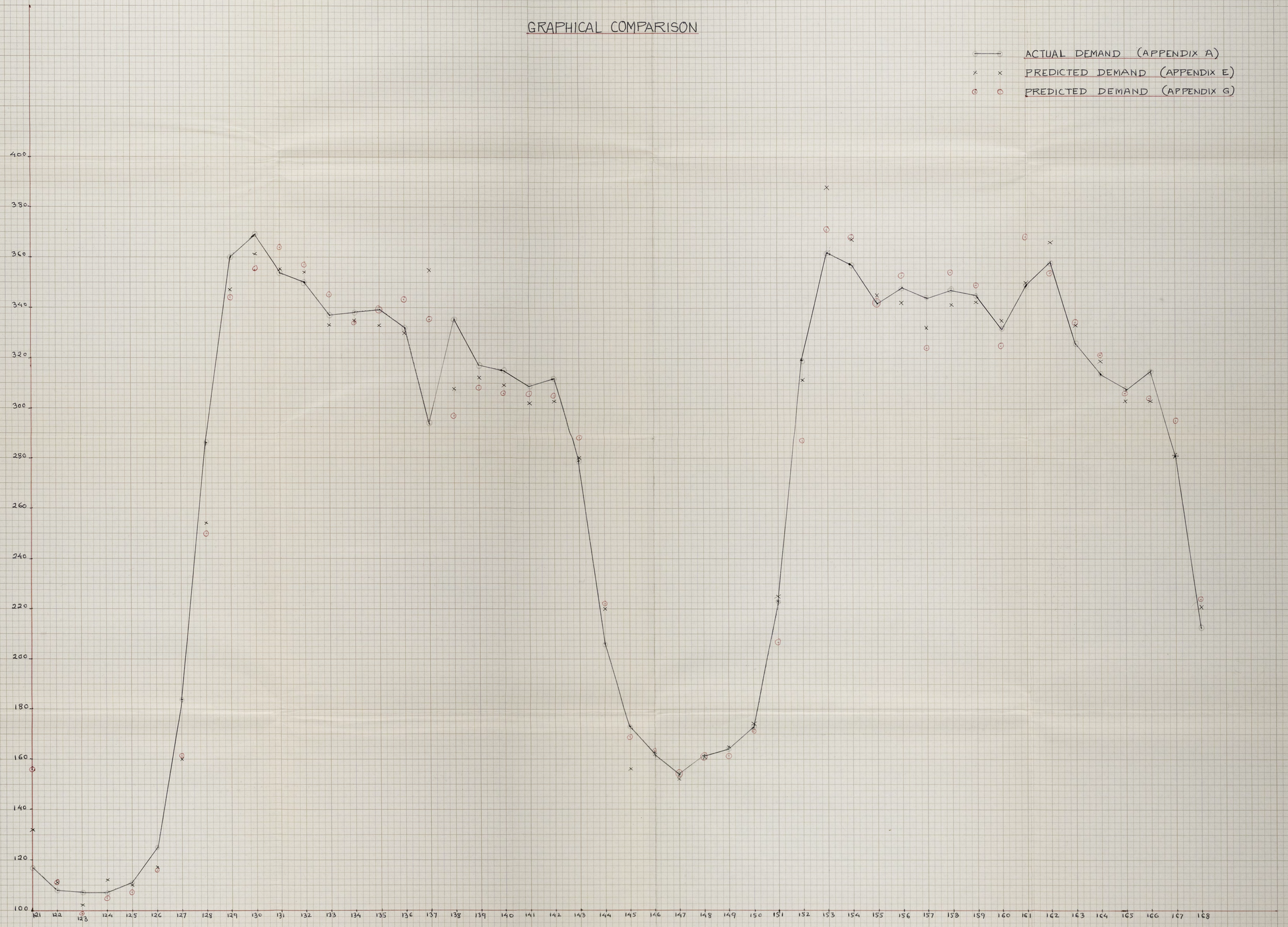
46) ZAHRADNIK, R.L., and ARCHER, D.H., "Process model building" (Instruments and control systems, Aug. 1965, vol. 38, pp.155-58)

GRAPHICAL COMPARISON

ACTUAL DEMAND (APPENDIX A)

PREDICTED DEMAND (APPENDIX E)

PREDICTED DEMAND (APPENDIX G)



GRAPHICAL COMPARISON