The Analysis of Random Signals

A Thesis
Presented for the degree of
Doctor of Philosophy
in
Electrical Engineering

in the
University of Canterbury,
Christchurch, New Zealand.

by
E. Chacko
1969
Acknowledgements

The author is grateful to all those who have helped in so many ways in the preparation of this thesis. In particular, the help and encouragement of Dr. J.K. Bargh of the Electrical Engineering Department and Dr. J.J. Deely of the Mathematics Department must be mentioned. Thanks are also due to Miss P. Lovell for her help in the typing of the thesis, and to the Department of External Affairs for their financial support.
# Table of Contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2.0</td>
<td>A Brief Survey of Random Signal Analysis</td>
<td>9</td>
</tr>
<tr>
<td>2.1</td>
<td>Assumptions and Approximations in Random Signal Analysis</td>
<td>10</td>
</tr>
<tr>
<td>2.2</td>
<td>Errors and Assumptions in Data Acquisition</td>
<td>16</td>
</tr>
<tr>
<td>2.2.1</td>
<td>Ensemble</td>
<td>17</td>
</tr>
<tr>
<td>2.2.2</td>
<td>Record Length</td>
<td>18</td>
</tr>
<tr>
<td>2.2.3</td>
<td>Sampling Process</td>
<td>20</td>
</tr>
<tr>
<td>2.2.4</td>
<td>Aliasing</td>
<td>21</td>
</tr>
<tr>
<td>2.2.5</td>
<td>Amplitude Quantization</td>
<td>22</td>
</tr>
<tr>
<td>2.2.6</td>
<td>Noise</td>
<td>23</td>
</tr>
<tr>
<td>2.3</td>
<td>Errors and Assumptions in Digital Computation</td>
<td>23</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Reconstruction</td>
<td>23</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Representation</td>
<td>24</td>
</tr>
<tr>
<td>2.3.3</td>
<td>Determination of Probability Density Function</td>
<td>25</td>
</tr>
<tr>
<td>2.3.4</td>
<td>Determination of Correlation Function</td>
<td>26</td>
</tr>
<tr>
<td>2.3.5</td>
<td>Determination of Power Spectrum</td>
<td>27</td>
</tr>
<tr>
<td>3.0</td>
<td>Introduction</td>
<td>28</td>
</tr>
<tr>
<td>3.1</td>
<td>Basic Concepts and Statement of Problem</td>
<td>31</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>3.11</td>
<td>Basic Concepts</td>
<td>31</td>
</tr>
<tr>
<td>3.12</td>
<td>Problem</td>
<td>39</td>
</tr>
<tr>
<td>3.2</td>
<td>Results</td>
<td>40</td>
</tr>
<tr>
<td>3.3</td>
<td>Details</td>
<td>46</td>
</tr>
<tr>
<td>3.30</td>
<td>Introduction</td>
<td>46</td>
</tr>
<tr>
<td>3.31</td>
<td>The Inequality I</td>
<td>48</td>
</tr>
<tr>
<td>3.32</td>
<td>The Inequality II</td>
<td>52</td>
</tr>
<tr>
<td>3.321</td>
<td>Determination of the Estimates ( \hat{\mu}_j ) of ( \mu_j )</td>
<td>52</td>
</tr>
<tr>
<td>3.322</td>
<td>To Show that ( \hat{\mu} ) and ( \mu ) are Close</td>
<td>68</td>
</tr>
<tr>
<td>3.4</td>
<td>Gaussian Approximation</td>
<td>81</td>
</tr>
<tr>
<td>3.5</td>
<td>Some Examples of Choosing ( K )</td>
<td>84</td>
</tr>
<tr>
<td>3.51</td>
<td>Gaussian Distribution</td>
<td>85</td>
</tr>
<tr>
<td>3.52</td>
<td>Uniform Distribution</td>
<td>87</td>
</tr>
<tr>
<td>3.53</td>
<td>Random Phase Sinewave Distribution</td>
<td>89</td>
</tr>
<tr>
<td>3.54</td>
<td>Rayleigh Distribution</td>
<td>92</td>
</tr>
<tr>
<td>4.1</td>
<td>Review</td>
<td>100</td>
</tr>
<tr>
<td>4.2</td>
<td>Comparison of Some Methods of Determining Power Spectra</td>
<td>101</td>
</tr>
<tr>
<td>4.3</td>
<td>Specific Problems in Determining Power Spectral Functions from Correlation Functions</td>
<td>103</td>
</tr>
<tr>
<td>4.31</td>
<td>Zero and Very Low Frequencies</td>
<td>104</td>
</tr>
<tr>
<td>4.32</td>
<td>Correlation Divisor</td>
<td>106</td>
</tr>
</tbody>
</table>
6.42 The Conversion Unit
   6.421 Amplifier and Filter 165
   6.422 Sample and Hold 166
   6.423 The Analogue to Digital Converter 168
   6.424 Control Signals to ADC 169
   6.425 Interconnections 171

6.43 The Encoder
   6.431 Binary to Binary-Coded-Decimal 173
   6.432 BCD Register and Decoding 175
   6.433 Digit Sequencing for Punching 175
   6.434 Alarm 179

6.44 Divider and Controller
   6.441 Clock Signal Circuit 182
   6.442 Start/Stop Signal Circuit 186
   6.443 Sampling Point Determination 188
   6.444 Operating Procedure 191

6.45 Punching of Data
   6.451 Interposer Switches 198
   6.452 Sequence Numbering of Cards 200
   6.453 End of Punch Signal 200

6.46 Sorting of Data
   6.461 Reading in Data 201
   6.462 Finding the Data Set and Checking 203
   6.463 Sorting and Punching 204
   6.464 870 Panel 206
CHAPTER 1

INTRODUCTION

Man has always been concerned with improving his methods and range of communication. The study of communication as a science may be considered to have started with the advent of electrical signalling around the early 19th century. Relatively little progress was made, however, till the beginning of the 20th century, when the thermionic valve was invented. During the First World War, communication techniques developed very rapidly. Between the two world wars, communication systems began to be studied from a general point of view and, particularly during the Second World War, the general theory of communication or "cybernetics" as it was called by Wiener, developed very considerably both in technical complexity of systems and mathematical rigour of analysis. Over the last two decades, the development of communication theory has been at an ever increasing pace in depth as well as scope, as testified by the flood of books and papers on the many aspects of the subject.

Communication theory may be classified into three main categories, although there is considerable overlap between
the three branches.

1) Non statistical detection of wanted signals (i.e. the "vehicle" of conveying information) from unwanted signals or "noise".

2) Statistical detection of signals.

3) Coding of information.

The author is mainly concerned with the second branch. In particular, some of the problems encountered in the analysis of signals using a digital computer are investigated in this thesis.

In many situations, a description of signals from a statistical point of view is necessary and the theory of random processes may be applied in such cases. A complete statistical description of the random signals involved in a physical system may be a prohibitive task. However, in many cases, a limited amount of information about the signals may be sufficient to characterise the system. Thus the correlation functions or power spectra of input and output signals of a linear system give all the information necessary to define the system, subject to the input signal having appreciable frequency content over the range of frequencies to which the system can respond [32]. For this reason the power spectrum of a signal is of particular interest.

Most systems however are nonlinear, or linear only over
a restricted range of input signal amplitudes. The amplitude probability distributions of the input and output signals may be necessary in such cases. The first order probability density function is the simplest of these and though generally insufficient on its own, gives useful information about the signal.

The processing of signals is being increasingly carried out on digital computers. This means however that, when the signals are in analogue form, they have to be converted to digital form. Thus the question arises as to how many, how often, and how accurately readings have to be taken. The best way to operate on the digital data to yield desired answers also needs considering.

The thesis project was started as part of a long term project concerned with analysis of signals from industrial plants in order to determine plant characteristics from the signals. At the time the project was started, the only facility available in the department for accurate recording of data with frequency components of more than a few hundred Hz, was an Ampex SP-300 magnetic tape recorder. Moreover, because of the limited funds available, the purchasing of any elaborate equipment for analogue to digital conversion and direct digital recording was not possible. Digital computing facilities in the University were limited to an
IBM 1620 with a card reader and typewriter as the only input devices. Modifications to the computer to accept data directly from an analogue to digital converter were not possible since the computer was already being heavily used, as well as because of IBM servicing contract arrangements. Thus the only practical method of direct communication with the computer was via punched cards. An IBM 870 documenting system was also available, capable of converting from paper tape to punched cards.

It was therefore decided that the project would involve the building of a data acquisition system, capable of transferring analogue data on magnetic tape to punched cards. The system would be a temporary solution to the data acquisition problem and the design of it was to compromise between economy of time and money. Any equipment bought was preferably to be such as to enable it to be reused, should an on line conversion system become available. Further, since it was expected to convert a considerable amount of data from analogue to digital form, conversion should be as automatic as possible.

Thus the system developed has a basic analogue-to-digital converter (ADC) of eight bit capacity, a control unit, and a modified card punch. The signal to be converted is recorded on magnetic tape and formed into an endless loop. The control unit initiates conversions at a speed of about
two per second, a rate restricted by the card punch. A set of readings are taken for each revolution of tape. Successive sets of readings per revolution of tape are effectively displaced in time from the previous set by an amount equal to the actual sampling interval required. The basic ADC itself, which was the main expense in the system, is capable of up to 50,000 conversions a second and is therefore capable of reuse as part of an online data conversion system handling frequencies up to 25kHz. The card punch was made available by the Computing Centre of the university.

The thesis starts with a review of some of the relevant aspects of random signals analysis in Chapter 2. Chapter 3 is concerned with the determination of probability density functions (p.d.f.s) of random processes. In particular, the author has investigated the question of obtaining the p.d.f. of a random process when only a single observation from the process of interest is available. The situation envisaged is one where the signal is stored in a computer in the form of a set of points time-sampled according to Nyquist requirements - that is, sampled in such a way that the signal may be reconstructed in the computer. Although some work on p.d.f.s can be found in the literature, there does not seem to be any method of dealing with this situation. The author here finds conditions on a process under which, in the given
situation, the p.d.f. of the process can be found and gives methods of determining the p.d.f. and the errors that can arise in doing so.

It is shown that under fairly general conditions the p.d.f. of a process can in fact be found from a finite length of a single observation from the process. Relationships between the error and signal length are found under various assumptions on the behaviour of the process. For the purpose of representing the p.d.f., it is divided into a number of channels and the probability of the signal being in each channel is found. Thus the question of the number of channels necessary and the method of reconstruction that can be used are considered. Further to evaluate the probabilities within the channels a number of values of the stored signal have to be evaluated. The number of points that are needed to obtain the probabilities with necessary accuracies are found. These results can be combined and thus it is shown that in many cases, it is possible to determine the p.d.f. of a process in the stated situation. Although the results have been developed with analysis using a digital computer in mind, they are generally applicable to analogue and hybrid methods of analysis.

Much has been written on the subject of power spectrum estimation. In Chapter 4, a comparison of some methods of
power spectrum determination is given. The distortion of the spectrum by zero and low frequencies, the question of smoothing the spectrum, the choice of "windows" for smoothing, and the presence of negative values of spectrum are some of the problems met with in the digital determination of the power spectrum from the correlation function. These questions are discussed in Chapter 4 and the author comments on suggested solutions in some cases. The author also considers the problem from an examination of properties of Fourier transform pairs and from the point of view of interpolation. These considerations are seen to be useful in explaining some of the above problems and also enable some simple error bounds to be determined.

The important subject of interpolation or reconstruction of signals in a computer is considered in Chapter 5. Some theory and a brief survey is given there. It is shown, for example, how from the general theory of interpolation using entire functions, Shannon's sampling theorem and a sampling theorem due to Jagerman and Fogel are obtained. Results of various authors on bounds for several interpolation formulae are given and some of these results are modified to give lower error bounds. Finally, an interpolation formula, which fits an entire function to a set of numbers that are the areas between intervals of the function to be reconstructed, is
derived. This formula is found to be useful in p.d.f.
determination. An error bound for this interpolation formula
is also obtained.

Chapter 6 describes the data conversion system developed.
The system outlined above has performed satisfactorily in
converting a large amount of analogue data to punched cards
for another thesis project [76].

In Chapter 7, some of the more interesting circuits used
in the data conversion system (e.g. the sample and hold and
the clock-filter circuit) are examined in more detail, along
with some of the problems particular to the system. The
sample and hold developed is described since it is a cheap
and simple, but at the same time fast and accurate device.
The clock-filter circuit is used to compensate for clock
drop-outs from the tape.

The thesis concludes with Chapter 8 which gives a
summary of the main results and also indicates some possible
directions for further research.
CHAPTER 2

A BRIEF SURVEY OF RANDOM SIGNAL ANALYSIS

2.0 Introduction

For a long time only deterministic methods were used in the analysis of electrical signals. Complicated signals were analysed by the Fourier series method if periodic, or the Fourier integral method if transient in nature. A signal was approximated to fall into either category. In communication systems, particularly, this was not very satisfactory. A communication channel may be subjected to a large number of different signals corresponding to many possible messages. Each of these messages has its own probability of occurrence. Thus the input to the channel is an ensemble of functions and a statistical approach to the problem is necessary. Often however, to simplify design of the system and to easily determine the output of the system the desired information about the input is the "frequency content" of the signal. Consequently, the relationship connecting the correlation function in the time domain and the power spectrum in the frequency domain gave a new and powerful tool to communication engineering and related fields.

The methods of correlation and spectral analysis have provided a far more realistic way of representing signals in communication systems. Further, signals, besides being
a member of an ensemble of signals, are often contaminated with unwanted and in many cases random fluctuations—generally classed as "noise" or "random noise". This may be due to indescribably complex (and hence "random") fluctuation in loading, interactions between stages in the system, "shot" or "Johnson" noise in electrical parts of the system or a vast class of possible sources. It is thus extremely useful to have a means of dealing with these situations. The method of random signal analysis covers most of these cases.

2.1 Assumptions and Approximations in Random Signal Analysis

It is normally assumed that a "random" signal \( x(t) = X(\omega,t) \) is a "sample function" from an "ensemble" \( \{X(\omega,t), t \in \mathbb{R}\} \) (see e.g. [32,35]). The ensemble may be thought of as the collection of all the possible outcomes of an experiment. This implies the existence of a probability law governing the occurrence of the events. The sample function is any member of this collection.

A complete description of any but the simplest random process would be a prohibitive task. A process is usually described by only a few of its parameters; the first and perhaps the second order probability density functions, the mean, the variance and the correlation functions are the common ones. Often only one member of the ensemble is available for analysis and in any case it is normally inconvenient
to analyse more than one sample function. Generally the
time origin is considered to be arbitrary, and the statistics
of the process are assumed to depend only on the differences
in time and not on the actual time itself. Further to this
assumption of "stationarity" it is normally assumed that the
process is ergodic, i.e., the time averages are taken to be
the same as ensemble averages. Ergodicity justifies the use
of a single time record from the process of interest. Whereas
there are a large number of cases where these assumptions are
valid, sometimes the simplification afforded by such assumptions
is their main justification. A study of the underlying
mechanism of the process may reveal whether the process could
be considered stationary and perhaps ergodic.

The computational problems involved in the analysis of
an arbitrarily distributed sample function can be complicated
and possibly unsolvable. The assumption is often made that
the process is gaussian. The gaussian distribution is
classified by its first and second moments alone, so that
the computations are greatly simplified. The "central limit
theorem" shows that under fairly general conditions the joint
effect of a large number of random processes of arbitrary
distribution tends to give a gaussian distribution. Further,
linear transformations on gaussian distributions result in
gaussian distributions.

However, signals are often subjected to non-linear
transformations. Moreover, the central limit theorem is valid only for contributions from a large number of events. When there are only a small number of random effects contributing to the final process, the assumption of a gaussian distribution is not necessarily valid.

In cases where the random process is gaussian, the autocorrelation function characterizes the process. Further if the process is stationary, the correlation function is independent of time. In this case, the power spectrum completely characterizes the random process.

It is difficult to make an accurate analysis of a signal about which nothing is known. Often the assumptions of stationarity and ergodicity and, perhaps, of having a gaussian distribution can be made about the signal, or a modification of it. For example, the effect of deterministic or transient signals superimposed on a gaussian signal is to cause the amplitude distribution to deviate from a normal one. Similarly removing low frequency trends may make a signal stationary (See e.g. [9]).

Even when the analysis can be based on a single record, only a finite length of signal can be used. When ergodicity is assumed the mean $\bar{x}$, variance $\sigma_x^2$, correlation function $R(\tau)$ are with probability one, given by 2.1, 2.2 and 2.3 respectively. The power spectrum and correlation function
are related by 2.4. (See e.g. [8,32,62,72]).

\[ \bar{x} = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t) dt \]  
2.1

\[ \sigma_x^2 = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} \{x(t) - \bar{x}\}^2 dt \]  
2.2

\[ R(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)x(t+\tau) dt \]  
2.3

\[ S(w) = \int_{-\infty}^{\infty} R(\tau)e^{-jw\tau} d\tau \]  
2.4

In practice since the signal length T is finite, estimates of the mean \( \bar{x}_T \), variance \( \sigma_{x_T}^2 \), correlation function \( R_T(\tau) \) and spectrum \( S_T(w) \) only may be obtained. The estimators commonly used are

\[ \bar{x}_T = \frac{1}{T} \int_{0}^{T} x(t) dt \]  
2.5

\[ \sigma_{x_T}^2 = \frac{1}{T} \int_{0}^{T} \{x(t) - \bar{x}_T\}^2 dt \]  
2.6

\[ R_T(\tau) = \frac{1}{T} \int_{0}^{T-|\tau|} x(t) x(t+|\tau|) dt, \quad -T < \tau < T \]  
2.7

or
Different portions of the signal will in general give different values for the estimates. Thus the question of the "goodness" of the estimates arises. Usually the merits of an estimate are judged by means of the following: (See e.g. [30,38,108])

1. Bias, \( b = \) (actual value) - (mean value of estimates)

2. Consistency. An estimate is consistent in its context if the estimate tends to the true value as \( T \to \infty \).

3. Efficiency. An estimate \( \theta \) is more efficient than another \( \theta' \) if \( \theta \) has a lower variance.

4. Sufficiency. An estimator is sufficient if it utilizes all the available information.

The estimators 2.5 and 2.6 are unbiased and consistent. The estimator 2.7 is biased but as shown by Parzen [73] is more efficient than 2.7' which is unbiased.

Though 2.8 is an asymptotically unbiased estimator, (i.e. \( b \to 0 \) as \( T \to \infty \)) it is not consistent except for \( S(w) = 0 \).
Parzen [72,73] and Bartlet [6] have shown how consistent estimates may be obtained. The method amounts to "smoothing" the "raw" estimate 2.8 to give a smoothed spectrum, $S'_T(w)$ with some suitable kernel or "window" $K(w,y)$

$$S'_T(w) = \int_{-\infty}^{\infty} S_T(y)K(w,y)dy$$

However, the reduction in estimate variance is at the expense of resolution $\Delta w$ of the spectrum since the smoothed spectral estimate $S'_T(w)$ is a weighted average of the estimate $S_T(w)$ around $w$, i.e., the window has the effect of "smudging" the spectrum. A correspondingly greater length of record will be necessary to get the same resolution.

Increasing the record length to a suitable value will give as accurate an estimate as desired. In the case of spectral estimates, an "uncertainty principle" exists, given by $T\Delta w = K$, $K$ a constant. Using this, the record length for a given desired resolution can be computed. The uncertainty constant, $K$ depends on the method of analysis. For example, if the amplitude information is thrown away so that the spectrum estimates are from the signs ($\pm$) of the signal, a much longer record would be needed than when the actual values are used. One is limited in the length of record (or number of samples in the digital case) that one can examine due to storage or computation time limitations.
Analogue and digital techniques are used for the analysis of signals. Analogue methods are generally faster but analogue equipment tends to be more specific in nature, more limited in accuracy and require more maintenance.

Digital methods are rapidly becoming more and more used. Their main disadvantages are lack of speed of computation and difficulties with data acquisition, i.e., the process of converting the signal from its original form to a form suitable for direct input to the computer. However, they have the great advantage of versatility, only software changes being necessary to alter the processing carried out on the data. Faster computers and better methods of computing and computer organization are making digital methods of comparable speed to analogue methods. Much progress has also been made in recent years in the theory and practice of data acquisition. As a digital system stores a finite number of discrete values only, it is necessary to optimize the information content for a given amount of data points. The problems of sampling and quantizing the signals present a whole new class of problems.

2.2 Errors and Assumptions in Data Acquisition

When signal analysis is considered in practice, various assumptions and approximations are necessary. In the first place, a suitable signal or set of signals has to be chosen, and this is not always a trivial problem. If the signal is
in analogue form, it has to be converted to a form that is suitable as an input for a digital computer in those cases where a computer is to be used for analysis. Then, since only a finite length of a signal or a finite number of signals from an ensemble can be examined, the statistical properties of the sample will need to be considered.

2.21 Ensemble

Certain assumptions, such as stationarity and ergodicity, are often made about the process of interest. A sample record is said to be self-stationary if its properties do not vary, with any statistical significance, when different portions of sufficient length of the record are analysed [9,50]. Frequently no distinction is made between self-stationarity and stationarity, and often it is self-stationarity only that is tested for. It is also often assumed that a process is ergodic if it is found to be self-stationary. A non-stationary record can sometimes be treated as stationary (and ergodic) by removing slow trends or selecting and restricting attention to sections that are stationary. When stationarity cannot be assumed, the parameters (e.g. mean) become functions of time.

In the remainder of this chapter it is assumed that any process \(- \{X(\omega,t), t \in \mathbb{R}\}\) referred to is ergodic and consequently it is sufficient to deal with a single member
$X(\omega_0, t) = x(t)$ of the ensemble.

2.22 Record Length

In general the random variable $x(t)$ exists for all time $t$. Even if $x(t)$ is known to be identically zero for all but a finite length of time $T_d$, it may be necessary, or desirable, to use only a smaller portion $T$ of the signal since

1) It may not be possible, convenient, or economical to acquire more of the signal.

2) Even when larger lengths of the $x(t)$ are available, processing cost, facilities or time may prevent the use of all of the data.

Thus it is usual to have to "estimate" the properties of the sample function $x(t)$ from a finite length $T$ of it. Since $x(t)$ is assumed stationary, the time origin may be taken as being at the start of the available signal. Hence

$$x_T(t) = x(t) \quad t \in [0,T]$$

$$= 0 \quad t \notin [0,T]$$

is available, which is essentially a statistical sample from $x(t)$ and the procedure will be one of statistical parameter estimation for the process. Each parameter determined will be a random variable with its associated distribution. The "goodness" of an estimate is a property of this distribution and it is of interest to know the form or relevant features of the distribution.
Briefly, a "good" estimate would be one that gives a sufficiently accurate value for the parameter every time the estimation is done under similar conditions. A longer sample length $T_L$ will give better results, but will involve more effort (see e.g. [24]).

The length of sample $T$ required will depend on the desired accuracy, the method of computation, the parameter being estimated, the "cost" in terms of money, time and computational effort as well as on the distribution functions of the random process.

The errors introduced in the data acquisition procedure will affect the accuracy of the parameters and hence in turn affect the necessary sample length.

Thus the determination of the sample length is not simple or straightforward and various approximations are made, or rules of thumb used, in deciding on a value for $T$. In some cases, however, these approximations lead to large and impractical values and a compromise may have to be made [9,13,60].

For example, in determining the power spectrum [9] with a maximum relative error $e$, the number of points is

$$N = \frac{4f_c}{e \pi B}$$

where $f_c$ is the highest frequency present in the signal
and $B_e$ is the desired resolution. If $f_c = 1,000\text{Hz}$, to get a resolution of $10\text{Hz}$ with $\epsilon = 0.01$ requires $N = 4 \times 10^6$. If only 4,000 points can be used then as $B_e \epsilon^2 = 1$ the frequency resolution is only $100\text{Hz}$ even allowing for a $10\%$ error.

2.23 Sampling Process

The sampling process, or time quantization, may be represented as in figure 2.1. Values of the input $x(t)$ are read off at times $\{t_n\}$ resulting in a sequence of values $\{x(t_n)\}$ at the output. Alternatively, particularly for on-line or hybrid systems, the output may be thought of as being the input modulated with a series of unit impulses at times $\{t_n\}$. Usually the sampling intervals $(t_{i+1} - t_i)$ are a constant $h$ and periodic sampling with period $h$ results.

Although the nominal sampling period is $h$, there may be "jitter", i.e. slight random variations in the intervals $(t_{i+1} - t_i)$. The sampling times are thus $t_n = nh + \zeta_n$, where $\zeta_n$ is a random variable. These problems have been discussed in the literature [3,20,70,85].

It may happen that there is a failure to sample at some of the times $t_n$. In this case two possibilities may arise:
The fact that a sample is missing is not known, but the probability of a miss is known or is estimated.

2) The position of the missing sample is known.

Case 1 does not seem to have been dealt with anywhere.

Case 2 has been treated [70] along with the situation where there is a partial error in the reading due to loss of some of the bits corresponding to the digital value.

2.24 Aliasing

Although physical functions are frequency limited, it is not always economical to sample at the "Nyquist rate", that is, at a rate \(2f_c\), where \(f_c\) is the highest frequency present in the signal. However if the sampling rate \(2f_s\) is below the Nyquist rate, "aliasing" occurs. This phenomenon is the folding back about \(f_s\) of the energy of the frequencies above \(f_s\) to distort the signal.

Often not all frequencies present are of interest, but to avoid aliasing sampling should be at or above the Nyquist rate. If low, as well as high, frequencies are present, a very large number of samples would result. For example, if the low frequency to high frequency ratio is 1:1,000, and if 10 cycles of the lowest frequency are needed, at least 20,000 samples will have to be taken. The common remedy is to use filters whenever possible.
The error introduced by the presence of frequencies above $f_s$ is termed "aliasing error" and expressions for this type of error has been worked out by various authors (e.g. [42,70,88, 89]).

2.25 Amplitude Quantization

A finite time is necessary for analogue to digital conversion. During this time, the analogue voltage is generally changing. A "sample and hold" device is frequently employed to overcome this problem although it may be possible to avoid this complication and expense. For example, in a staircase type of analogue to digital converter, it is possible to work out the actual time of sampling $t_{na}$ corresponding to the value of the output $V_o$, knowing the nominal sampling time $t_n$, and the clock period $\delta t$. If the quantization interval is $q$ then the number of clock periods $N$ is given by $N = V_o/q$.

Thus

$$t_{na} = t_n + N\delta t$$

This suggestion, however, while saving the expense of a sample-and-hold, gives unequal sampling intervals.
The discrete digital values normally have a quantization error $\varepsilon$ associated with them. $\varepsilon$ may be made as small as desired by increasing the number of levels of quantization, but this is at the expense of an increase in cost and in conversion time. There is no point, in any case, in increasing resolution beyond the stage where it is small in comparison to any other sources of error (e.g. transducer errors) or beyond a stage where the quantization error has a negligible effect on the final accuracy. Quantization error has been dealt with by various authors (e.g. [42, 67, 70, 80, 89, 105]).

2.26 Noise

Often, when the signal to be quantized has to pass through a channel (e.g. tape recorder), noise is added contaminating the signal before it is even quantized. The effect depends very much on the nature of the noise (see e.g. [7, 12, 22]).

2.3 Errors and Assumptions in Digital Computation

2.31 Reconstruction

In order to obtain values of the signal other than those stored in a computer, interpolation is necessary.

The cardinal function $\text{Sinc}(x) = (\sin(x)/x)$ may be used to interpolate between any of the points sampled at or above the Nyquist rate [68, 84].
\[ x(t) = \lim_{N \to \infty} \sum_{-N}^{N} x(nT) \text{Sinc}(t-nT) \quad 2.10 \]

provided \( \phi(w) = \int_{-\infty}^{\infty} x(t)e^{-jwt}dt \quad w \in \left[ -\frac{\pi}{T}, \frac{\pi}{T} \right] \quad 2.11 \)

\[ = 0, \quad \text{otherwise} \]

That is if

1) The sampled points \( x(nT) \) are error free
2) The \( x(t) \) has no frequencies above \( \frac{\pi}{T} \)
3) An infinite number of terms in the sum are used, then \( x(t) \) is exactly recoverable for all \( t \).

Violation of one or more of these conditions gives rise to errors in \( x(t) \). The effect of violation of conditions one, two or three separately have been considered by various authors (e.g. [42]). The effect of all three combined does not appear to have been studied. Nearly always, in practice, all three conditions are violated.

Weighting functions other than the Sinc function have been considered. (e.g. [47,71,89]. See also Chapter 5.)

2.32 Representation

An alternate way to store signals is to store their
Fourier coefficient with respect to some orthonormal set of functions. The trigonometric functions are often chosen, but other orthonormal set of functions may give a lower truncation error for a given number of terms, and sometimes facilitate computation (see e.g. [33]).

2.33 Determination of Probability Density Function

The fraction of time \( T_X / T \) that a signal is within an amplitude window of width \( h \) about the value \( X \) is estimated as the value of the probability of occurrence of \( x(t) \) in the interval \([X - \frac{h}{2}, X + \frac{h}{2}]\) [9,60]

\[
F_T(X,h) = \frac{T_X}{T} = P[X - \frac{h}{2} \leq x(t) \leq X + \frac{h}{2}]
\]

The estimated value of \( f(x) \), the probability density function (p.d.f.) at \( X \) is

\[
\hat{F_T}(X) = F_T(X,h)/h
\]

\( \hat{F_T}(X) \) is in general biased and inconsistent. The actual value at \( X \) is given by

\[
f(X) = \lim_{T \to \infty} \lim_{h \to 0} \frac{F_T(X,h)}{h} = \lim_{t \to \infty} \lim_{h \to 0} \frac{T_X}{Th}
\]

The optimum window width for a given signal clearly depends on the distribution. The larger \( h \) is, the more "smudged" or "averaged" the value of \( \hat{F_T}(X) \) will be at \( X \). The bias increases with \( h \), but the variance decreases so
that a compromise between the two will be needed.

Various analogue (e.g. [4, 16, 23, 56]) and digital (e.g. [60]) methods of determining p.d.f.s have been described in the literature, and Chapter 3 is concerned with the determination of probability density functions.

2.3.4 Determination of Correlation Functions

Correlation functions are normally estimated using the equations given in (2.7) or (2.7').

In the discrete case the formulae reduce to summations,

\[ R_M(m) = \frac{1}{M-m} \sum_{n=1}^{M-m} x(n)x(n+m) \quad 2.15 \]

or

\[ R'_M(m) = \frac{1}{M} \sum_{n=1}^{M-m} x(n)x(n+m) \quad 2.15' \]

In addition to statistical errors, inaccuracies in data will cause further error. The errors in the data affect the final correlation values. The correlation function of "white noise", for example, is an impulse at zero lag so that the effect of additional white noise contaminating the signal, as well as the effect of quantization noise, which may reasonably be considered as white noise (Bennet [10]), is only detectable near the origin.
Correlation functions have been determined using only zero crossings, and with various coarseness of quantizations. Usually simplicity of computation is at the cost of requiring a greater number of points (see [60]).

Correlation Functions are discussed further in Chapter 4.

2.35 Determination of Power Spectrum

The normal analogue method of spectral analysis is to pass the signal through a narrow band-pass filter and measure the output on an averaging device [9,13,60]. The estimate is generally biased though asymptotically consistent. In theory, an exact value can be obtained with diminishingly small bandwidth and infinitely long record length.

In the digital case the commonest method used to be to determine the spectrum as the Fourier transform of the correlation function. The correlation function is normally modified to minimize end effects [9]. It is also feasible to determine the spectrum by a modified "periodogram" method which uses a weighted average of the raw spectrum [53]. Another way is to simulate the analogue method [99]. Recently the Cooley Tukey algorithm has enabled a more direct method of determining power spectrum[29].

The determination of power spectrum is further dealt with in Chapter 4.
CHAPTER 3

DETERMINATION OF PROBABILITY DENSITY FUNCTIONS

3.0 Introduction

In practice the engineer is often confronted with a sample from a signal which may take one of a number of possible forms. This sample normally consists of an observation of a finite length $T$ of the signal. From this length $T$ the engineer must make predictions about other possible forms of the signal; for example, what values the signal amplitudes are likely to have when the signal is confronted again. Formally this amounts to saying that one member of a random process has been observed for a finite length of time $T$, and the probability density function (p.d.f.) of the process is desired, assuming of course that the p.d.f. exists.

Knowledge of the p.d.f. is particularly useful in design considerations because it allows the designer to assign probabilities to other observations and it gives a profile of the range of the signal. For example, suppose one wants to know the probability of finding the signal in a small interval of width $dx$; this probability is given by $f(x)dx$ where $f(x)$ is the p.d.f. of the process. Or when one is designing an amplifier, it is important to ensure that there is only a small probability of the signal taking values outside the
linear range $A$ of the amplifier. This probability is given by $1 - \int_{A} f(x) \, dx$.

Although the p.d.f. is the derivative of the cumulative distribution function (c.d.f.), it is generally easier and quicker to determine certain properties of the signal by looking at the p.d.f. Such things as bimodality, symmetry, range of possible values, preferred values, are readily obtained from the p.d.f. Measurement of probabilities from a c.d.f. involves taking differences and consequently can give large errors when the probabilities are small. Furthermore differentiating the c.d.f. numerically to obtain the p.d.f. could induce large errors. (See e.g. [60]).

For a general random signal, the p.d.f. is usually unknown and a precise determination of the p.d.f. from data is physically impossible (except in trivial cases) since one can only observe the signal for a finite length of time $T$. This raises then the following questions:

1) How does one use the observed sample to obtain an estimate of the unknown p.d.f.?

and 2) If it is possible to do 1), how large does $T$ have to be to obtain a given precision for the estimate?

These questions have received some attention in the literature. (See e.g. [2, 14, 21, 31, 69, 74, 82, 92, 96]). For
example, [14,21,82,92] have indicated how the p.d.f. may be found in the form of a sum of elements from an orthogonal set of functions. However, in these cases it is generally assumed that a sequence of independent points are available from the process of interest. What is often available in practice (particularly when analysis is carried out using a digital computer) is a sequence of points obtained from a finite length of the signal by a deterministic method and it is not always clear how and if an independent set of points may be obtained. Moreover, when a graphical picture of the p.d.f. is needed, a direct numerical representation of the p.d.f. is preferrable, rather than a representation as a sum of various functions. Also, these methods do not answer the important practical questions such as how many separate values of the p.d.f. are needed to faithfully represent the p.d.f. and how the length T of the observed signal is related to the precision of the estimate. These questions are dealt with in this chapter. The results herein obtained have, in general, wider application than those found in the literature. (See remarks x and xiii, page 78.)

In section 3.1 the basic concepts, definitions and the mathematical model are given. Section 3.2 briefly describes the results and indicates the method of proof. The mathematical details because they are rather tedious and frequently removed from the basic results, have been separated into
sections 3.3 - 3.5.

3.4 Basic Concepts and Statement of Problem

3.4.1 Basic Concepts

A more detailed description of the concepts given here can be found in various books (e.g. Loeve [64], Doob [35], Halmos [45], Dubes [36]), but the essential definitions are included here for clarity and completeness.

A totally finite measure space \((\Omega, \Theta, P)\) for which
\[
P(\Omega) = 1
\]
is called a probability space, and \(P\) is called a probability measure. \(\Theta\) is a \(\sigma\)-ring of measurable subsets or events, of \(\Omega\). A real valued measurable function defined on \((\Omega, \Theta, P)\) is called a random variable (r.v.). The integral
\[
\int_{\Omega} X dP
\]
of a r.v. \(X\) is called the expectation of \(X\) and is denoted by \(E[X]\).

The real valued function \(F_X(x)\) defined by
\[
F_X(x) = P\{\omega : X(\omega) \leq x\}, \ x \in \mathbb{R}
\]
is called the cumulative distribution function (c.d.f.). The c.d.f. is nondecreasing and continuous from the left on \(\mathbb{R}\) with
\[
F_X(-\infty) = 0 \text{ and } F_X(\infty) = 1.
\]
If \(X\) is a random vector defined by
\[
X = (X_1, X_2, \ldots, X_N)
\]
where \(X_1, X_2, \ldots, X_N\) are r.v.s on \((\Omega, \Theta, P)\), then the joint \(N\) dimensional c.d.f. of \(X\) is given by:
\[ F_X(x) = P \left\{ \omega: \bigcap_{i=1}^{N} [X_i \leq x_i] \right\} \]

where \( x = (x_1, x_2, \ldots, x_N) \in \mathbb{R}^N = \prod_{i=1}^{N} \mathbb{R}_i \).

In general, for any index set \( \Gamma \), the family of r.v.s \( \{X(\omega, t), t \in \Gamma\} \) is called a random or stochastic process. The function \( X(\omega, t) \) with \( \omega \) fixed and \( t \) varying over \( \Gamma \) is called a sample function from the process.

Let \( \{X(\omega, t), t \in \Gamma\} \) be an arbitrary random process, and let \( (t_1, t_2, \ldots, t_N) \) be \( N \) distinct elements of \( \Gamma \), where \( N \) is any positive integer. Then the \( N \)th order c.d.f. for the process at \( (t_1, t_2, \ldots, t_N) \) is the \( N \)th order c.d.f. of the vector \( \{X(\cdot, t_1), X(\cdot, t_2), \ldots, X(\cdot, t_N)\} \) given by

\[ F_X(x; t_1, t_2, \ldots, t_N) = P \left\{ \omega: \bigcap_{i=1}^{N} [X(\omega, t_i) \leq x_i] \right\}. \]

Let \( u \) be such that \( t_i + u \in \Gamma \), \( i = 1, 2, \ldots, N \). Then if \( F_X(x; t_1, t_2, \ldots, t_N) = F_X(x; t_1+u, t_2+u, \ldots, t_N+u) \) for all \( (t_1, t_2, \ldots, t_N) \) and \( u \) as defined above, the process is said to be strictly stationary. (Dubes [36] p.335). In particular, if \( F_X(x; t_1, t_2) = F_X(x; t_1+u, t_2+u) \) for all \( t_1, t_2, t_1+u, t_2+u \in \Gamma \), then the process is called wide-sense stationary, or second-order stationary. Second order stationarity implies first
order stationarity.

If \( F_X(x; t) \) has a derivative

\[
f_X(x; t) = \frac{\partial}{\partial x} F_X(x; t)
\]
everywhere in \( x \), \( f_X(x; t) \) is called the **first order p.d.f.** of the process. The \( N \) dimensional p.d.f. can similarly be defined.

Since the main interest here is in first order p.d.f.s, it will be implicitly assumed that all p.d.f.s are of first order unless otherwise stated.

Closely associated with the c.d.f. \( F(x) \) of a process is the characteristic function \( \phi(u) \) related to the c.d.f. by the Fourier Stieltjes transform as follows:

\[
\phi(u) = \int_{-\infty}^{\infty} e^{jux} dF(x)
\]

If the p.d.f. exists, then

\[
\phi(u) = \int_{-\infty}^{\infty} e^{jux} f(x) dx.
\]

The mathematical model given above, of course, describes a large variety of problems. In the present case, the main concern is with applications to communication systems, and in particular, to control systems. Hence it will be more convenient to adopt a vocabulary which is suggestive of
these applications.

The following situation can be considered as a typical illustration of how the mathematical model can be applied to control systems. In the study of control systems, one is often confronted with a signal (say $X(\omega,t), t \in \mathbb{R}$ where $\omega$ is a parameter) from a plant or process. This signal is one of many possible signals that could have resulted; i.e. one of the whole population \{$X(\omega,t), t \in \mathbb{R}$\} where $\omega \in \mathbb{R}$. Moreover, signals with certain characteristics are more likely to occur than some others; i.e. there exists a probability measure $P$ over $\mathbb{R}$. Thus this situation can be thought of in the framework of a stochastic process as suggested.

In this case then, a signal is simply a sample function; the parameter $t$ being time, and the parameter index set $\Gamma$ being the real line $\mathbb{R}$.

When referring to a fixed $\omega$, it will be convenient to use a simpler notation $x(t)$ for $X(\omega,t)$. $x(t)$ will be said to be generated by the process \{\(X(\omega,t), t \in \mathbb{R}\)\}.

In this context, a stationary process can be roughly thought of as generating signals whose randomness is not changed over time. This is not to say that the signal values at two distinct times are independent (stochastically), but it is to say that the distributions of the signal values at each time individually are not functions of time.
In this chapter, all processes are assumed to be wide-sense stationary, and for convenience, $F_X(x; t)$ will be written $F_X(x)$ with corresponding p.d.f. $f_X(x)$. Also, when the process in question is clearly indicated, the subscript $X$ will be dropped.

A further assumption which is usually satisfied in practice is that any generating process has finite memory: which is to say that the r.v.s $x(t)$ and $x(t+u)$ are stochastically independent when $u$ is large. Formally, the property will be defined as follows:

**Definition 3.1-1:** A process $\{X(\omega, t), t \in \mathbb{R}\}$ is said to have finite memory of length $U_0$ if $x(t_1)$ and $x(t_2)$ are stochastically independent when $t_1$ and $t_2$ are any real numbers for which $|t_2 - t_1| > U_0$.

In the study of stochastic processes, the concept of the correlation function is found to be useful. The correlation function $R(s, t)$ of the random process $\{X(\omega, t), t \in \mathbb{R}\}$ is defined by:

$$R(s, t) = \int_{\Omega} X(\omega, t)X(\omega, s)d\Omega$$

$$= E\{x(t) x(s)\}.$$

If the process is wide-sense stationary, the correlation function is a function of the difference $(s-t)$ only. Thus
for a wide-sense stationary process,

$$R(u) = \mathbb{E}[x(t)x(t+u)].$$

If in addition the process \{X(\omega,t), t \in \mathbb{R}\} has finite memory of length \(U_0\), then

$$R(u) = \mathbb{E}^2[x(t)] \text{ for } u > U_0.$$  

Closely related to the correlation function is the spectral distribution function, \(\Psi(w)\). To define the spectral distribution function, the following procedure may be followed. Firstly, any wide-sense stationary process \(\{X(\omega,t), t \in \mathbb{R}\}\) has a spectral representation, that is, there exists a function \(\Psi(\omega,w)\) such that

$$X(\cdot,t) = \int_{-\infty}^{\infty} e^{2\pi j tw} d\Psi(\cdot,w), \quad 3.1-1$$

provided the process satisfies the continuity hypothesis (Doob [35] § XI.3); namely that

$$\lim_{t-s \to 0} \mathbb{E}[|x(t) - x(s)|^2] = 0.$$  

It can be seen that the continuity hypothesis will be satisfied if \(x(t)\) is continuous and bounded. Signals from physical systems are continuous and bounded and hence the continuity hypothesis is not a restrictive one.

The correlation function \(R(u)\) of a process for which the continuity hypothesis is true, is continuous with
\[ R(u) = R(-u), \quad \text{and there exists a function } \psi(w) \text{ such that} \]
\[ R(u) = \int_{-\infty}^{\infty} e^{2\pi ju} \psi(w) dw, \]

in which case \( \psi(w) \) is called the spectral distribution function of the process. It is monotone non-decreasing and bounded. If it is also absolutely continuous, its derivative \( \psi'(w) \) is called the spectral density function or spectrum of the process.

The process \( \{\psi(\omega,w), \omega \in \mathbb{R}\} \) generated by the above definition (3.1-1) is related to \( \psi(w) \) by the equation \( \mathbb{E}\{|\psi(w_1) - \psi(w_2)|^2\} = \psi(w_1) - \psi(w_2) \) for \( w_1, w_2 \in \mathbb{R} \). This process has orthogonal increments, that is,

\[ \mathbb{E}\{|\psi(w_1) - \psi(w_2)|^2\} < \infty, w_1, w_2 \in \mathbb{R}, \]

and if \( w_1 < w_2 < w_3 < w_4 \) and \( w_1, w_2, w_3, w_4 \in \mathbb{R} \),

\[ \mathbb{E}\{[\psi(w_2) - \psi(w_1)][\psi(w_4) - \psi(w_3)]\} = 0. \]

If the spectral distribution function of the process contains only a finite number of jumps at \( w = w_i, i = 1, 2, \ldots M \), then the spectral representation becomes,

\[ x(t) = \sum_{i=1}^{M} \psi_i e^{2\pi jtw_i} \]

where the \( \psi_i = \psi(w_i^+) - \psi(w_i^-) \) are mutually orthogonal. In
the general case, the spectral representation 3.1-1 may be approximated arbitrarily closely by \( x_e(t) \) as a suitable Riemann-Stieltjes sum,

\[
x_e(t) = \sum_{i=1}^{M} \Psi_i e^{2\pi j t w_i}
\]

where \( \Psi_i = \Psi(w_{i+1}) - \Psi(w_i) \).

Note that since the \( \{\Psi(\omega, w), w \in \mathbb{R}\} \) process has orthogonal increments, the \( \Psi_i \) are mutually orthogonal as above. Thus the \( x_e(t) \) is in the form of a sum of \( M \) mutually orthogonal random variables.

The parameter \( w \) in 3.1-2 and 3.1-3 corresponds in communication systems to frequency. Thus 3.1-3 gives \( x(t) \) approximated as the sum of \( M \) components with frequencies \( w_i, i = 1, 2, \ldots, M \). Note that the choice of these \( w_i \)s is not arbitrary but is determined by the particular process involved. Also if the process is periodic, (i.e. \( x(t) = x(t+nT_p) \) for all \( t, t+nT_p \in \mathbb{R}, n \) being an integer and \( 0<T_p<\infty \), then \( w_i = k_i/T_p \) where \( k_i \) is an integer. Conversely, if \( w_i = k_i/T_p \) then the process is periodic. If the spectral representation or its approximation has a common factor \( w_p \) (or can be approximated to have a common factor \( w_p \)), then the process is periodic (or is approximated to be periodic) with period \( 1/w_p \).
One further association between the physical problem and the mathematical structure needs to be made. Analysis of signals is often made on digital computers. In such cases, the signal is usually stored in the form of a set of discrete points as mentioned earlier. Whereas the signal may theoretically be obtained for any $t$ over the available length of signal $[0,T]$ from the stored points, for computational economy, it is necessary to conserve the number of points observed. Thus a finite number of points, $N$ are selected from the interval $[0,T]$ for analysis. The way that the size of $N$ affects the answer is of interest, since it is desirable to keep $N$ as small as possible.

All $\varepsilon$s (with or without subscripts) used in this chapter will be assumed greater than zero and all $\delta$ as will be taken to be a positive number less than 1.

3.12 Problem

The problem treated in this chapter can now be stated formally.

Suppose one can observe $N$ independent points $x(t_1), x(t_2), \ldots, x(t_N)$ from a signal $x(t)$ over the interval $[0,T]$, where the $x(t)$ is generated by a process with p.d.f. $f(x)$. Can one find a function $\hat{f}(x) = \hat{f}(x,T,N)$ such that for each $1 > \alpha > 0$, $\varepsilon > 0$, there exists a number $N(\varepsilon,\alpha)$ and a length $T(\varepsilon,\alpha)$ so that taking $N \geq N(\varepsilon,\alpha)$ and $T \geq T(\varepsilon,\alpha)$
implies $P\{\sup_{x} |\hat{f}(x) - f(x)| > \varepsilon\} < \alpha$.

In the next section, conditions under which such an estimate exists are given and its form stated.

3.2 Results

In this section, formulae for $N,T$ and $\hat{f}$ for prespecified numbers $\varepsilon$ and $\alpha$ are given; $\varepsilon$ being the maximum allowable error and $(1 - \alpha)$ the level of confidence. The estimates $\hat{f}$ of $f$ will be obtained by estimating a sufficiently close and more convenient version $g$ of $f$, where $g$ is formed from a set of probabilities $m_j$, $j = 1, 2, \ldots, K$, obtained from $f$. Under stated conditions, it will be shown that these estimates indeed meet the requirements stated in the problem.

The $g$ to be used depends upon properties of the unknown density function $f(x)$. In all cases, it will be assumed that $f(x)$ has a support $S$ which is assumed to be an interval $[a, b]$ on the real line and of length $B = (b - a) < \infty$ but in some cases it is sufficient that $S$ be the essential support. $S$ will be called the essential support of $f$ if,

$$\int_{x \notin S} |f(x)| dx < \varepsilon$$

and $S$ will be called the support of $f(x)$ if

$$\int_{x \notin S} |f(x)| dx = 0$$
Let $x_0, x_1, \ldots, x_K$ be a partition of $[a,b]$ into $K$ intervals of equal length. The function $\varphi = \varphi(x) = \varphi(x;m_1, m_2, \ldots, m_K)$ will be formed from $K$ values $m_1, m_2, \ldots, m_K$, where

$$m_j = \int_{x_{j-1}}^{x_j} f(x) \, dx, \quad j = 1, 2, \ldots, K, \quad 3.2-1$$

$$x_j = x + j(B/K)$$

Formulæ for $g$ and $K$ are given below under three smoothness assumptions on $f$. The quantities $\varepsilon_a, \varepsilon_b$ and $\varepsilon_1$ are all prespecified positive numbers with $\varepsilon_1 = \varepsilon_a + \varepsilon_b < \varepsilon$. Also $h = B/K$.

**Ia) Bounded first derivative**

Suppose that $\sup_{x \in S} |f'(x)| \leq D$. Then choosing

$$g(x) = \frac{m_j}{(x_j - x_{j-1})}, \quad x \in [x_{j-1}, x_j], \quad j = 1, 2, \ldots, K \quad 3.2-2$$

$$= 0, \quad x \notin [x_0, x_K]$$

and

$$K \geq \frac{BD}{2 \varepsilon_1}$$

$$\Rightarrow |g(x) - f(x)| < \varepsilon_1.$$
Ib) Bounded second derivative and characteristic function of bounded essential support

Suppose \( \exists A \subseteq \mathbb{R} \) and a real number \( E > 0 \)

\[
\frac{1}{\pi} \int_{|u|>A} |\phi(u)| \, du \leq \varepsilon_a,
\]

and

\[
\text{Sup} |f''(x)| \leq E,
\]

then, choosing

\[
K \geq \max \left\{ \frac{B^2_E}{24\varepsilon_b}, \frac{AB}{\pi} \right\}
\]

and

\[
g(x) = \sum_{j=1}^{K} \left( \frac{m_j}{h} \right) \text{Sic} \left( \frac{\pi}{h} (x' - j'h) \right), \quad x \in S
\]

\[
eq 0, \quad x \notin S
\]

where \( \text{Sic} \, v = \sin \, \frac{v}{v} \)

\[
x' = x - (x_0 + h/2)
\]

and

\[
j' = j - 1;
\]

\( \Rightarrow |g(x) - f(x)| < \varepsilon_1.\)

Ic) Characteristic function of bounded essential support

Suppose \( \exists A \subseteq \mathbb{R} \)

\[
\frac{1}{\pi} \int_{|u|>A} |\phi(u)| \, du \leq \varepsilon_1,
\]

then choosing
and 

\[ g(x) = \frac{\pi}{h} \sum_{j=1}^{\infty} d_j \cos \frac{\pi}{h}(x' - jh), \quad x \in S \]

\[ = 0 \quad , \quad x \notin S \]

where \( d_j = \sum_{i=1}^{j} m_i \) for \( j = 1, 2, \ldots, K-1 \)

\[ = 1 \quad \text{for} \quad j \geq K \]

\[ \cos \nu = (\cos \nu - \sin \nu) / \nu \]

and \( x' = x - x_0 \).

\[ \Rightarrow |g(x) - f(x)| < \varepsilon_1. \]

Proofs of these results are given in Lemmas 1, 2 and 3 respectively, and examples of applications of these results are given in section 3.5.

Assuming that one of Ia, Ib, or Ic above obtains, an estimate of \( g \) then consists of the simultaneous estimation of \( m_1, m_2, \ldots, m_K \). These estimates are denoted by \( \hat{m}_1, \hat{m}_2, \ldots, \hat{m}_K \). Then the estimate \( \hat{f} \) is taken as

\[ \hat{f}(x) = g(x; \hat{m}_1, \hat{m}_2, \ldots, \hat{m}_K). \]

To obtain the estimates \( \hat{m}_j \)'s, a length \( T \) of signal is chosen. From this length, \( N \) points are selected at random
and the values \( x(t_1), x(t_2), \ldots, x(t_N) \) obtained. (See
remark xiv.) Let \( N_j \) be the number of values from
\( \{x(t_1), x(t_2), \ldots, x(t_N)\} \) which fall in \([x_{j-1}, x_j]\). The
estimate \( \hat{m}_j \) is then taken to be

\[
\hat{m}_j = \frac{N_j}{N}, \quad j = 1, 2, \ldots, K.
\]

The formulas for \( N \) and \( T \) to achieve a desired accuracy
\( \varepsilon_e \) with confidence \( 1 - \alpha \) will now be given. Assuming that
the r.v.s \( X(w, t) \) and \( X(w, t+u) \) tend to be independent as \( u \to \infty \),
it is shown that the \( m_j \)'s can be determined with probability
\( 1 \) using a single sample function from the process of interest.
For three particular assumptions on the behaviour of the
signals, expressions relating \( N, T \) and the error \( \sup_j |m_j - \hat{m}_j| \)
are given. In the following let \( \varepsilon_e, \varepsilon_c, \varepsilon_d, \alpha, \alpha_c, \alpha_d \), be such
that \( \varepsilon_e = \varepsilon_c + \varepsilon_d < \varepsilon - \varepsilon_1 \) and \( \alpha = \alpha_c + \alpha_d \).

IIa) Finite memory

Suppose the process is of finite memory \( U_0 \). Let

\[
T_0 = \max \left\{ \frac{U_0 K}{2\alpha_c \varepsilon_c^2}, U_0 \right\}
\]

Then \( T > T_0 \)

\[
=> P\{\sup_j |m_j - \hat{m}_j| > \varepsilon_e\} < \alpha.
\]

IIb) Approximately finite memory

Suppose \( |R(u) - m_j^2| < \alpha_c \varepsilon_c^2 \) for \( |u| > U_1 \), and let
\[ T_0 = U_1/(2\alpha_c \epsilon_c^2) \]

Then \( T \geq T_0 \)

\[ \Rightarrow P\{\sup_j (m_j - \hat{m}_j) > \epsilon_e\} < \alpha, \]

where \( R(u) \) is as defined in 3.3-6.

**IIc) Periodicity**

Suppose the process is periodic or is approximated to be so (see section 3.1 and the comment before Lemma 8), with period \( T_p \), and let

\[ T_0 = T_p/(2\epsilon_c) \]

Then \( T \geq T_0 \)

\[ \Rightarrow P\{\sup_j (m_j - \hat{m}_j) > \epsilon_e\} < \alpha. \]

For each of the above cases, \( N \geq N_0 \) where

\[ N_0 = (k - 1)/(k\alpha_d \epsilon_d^2). \] 3.2-6

If \( N \) is so large that the normal approximation may be used (e.g. if the number in the smallest class is about 100, see section 3.4), then,

\[ N_0 = \left\{ \frac{Z(1 - \alpha_d/2K)}{2\epsilon_d} \right\}^2 \] 3.2-7

where \( Z(1 - \alpha_d/2K) \) is the standard normal ordinate having an area of \( \alpha_d/K \) in the tails.
III) If \( g(x) \) is given as in Ia, Ib or Ic, and \( T_0 \) and \( N_0 \) as in IIa, IIb or IIc, then taking \( T \geq T_0 \), \( N \geq N_0 \) and 
\[
\hat{f}(x) = g(x; \hat{m}_1, \hat{m}_2, \ldots \hat{m}_K)
\]
\[
\Rightarrow \mathbb{P}(\sup_x |\hat{f}(x) - f(x)| > \varepsilon) < \alpha
\]

Details of these results are given in sections 3.3 and 3.4.

Thus given a stationary process \( \{X(\omega, t), t \in \mathbb{R}\} \) satisfying one of IIa, IIb, or IIc, and possessing a p.d.f. \( f(x) \) satisfying one of Ia, Ib, or Ic, it is then possible to estimate the p.d.f. of the process to a desired accuracy and with a given confidence interval by taking \( K, T, \) and \( N \) as indicated. Alternatively, of course, if a \( T \) and/or an \( N \) is given, the accuracy attainable can be calculated from these formulae.

3.3 Details

3.30 Introduction

As explained earlier, in order to estimate the p.d.f. \( f(x) \) by an estimate \( \hat{f}(x) \), the set of probabilities \( m_j, j = 1, 2, \ldots K \) are first estimated. The function \( g(x) \) which is obtained from the \( m_j \)'s is itself an estimate of \( f(x) \). However, precise determination of the \( m_j \)'s is not possible since only a finite length \( T \) of signal is normally
available for analysis. Thus intermediate estimates

\[ \theta_{T,j} = \theta_j \] (see 3.3-7) of \( m_j \) must be obtained. Ideally the

\( \theta_j \)'s can be found by a deterministic method. However, it is

usually more practical to estimate the \( \theta_j \)'s by \( \hat{m}_j \)'s using a

random sampling method as explained in remark xiv. It so

happens that the number of points \( N \) needed to form the

estimates \( \hat{m}_j \)'s from the observed length \( T \) is independent of \( T \).

Hence it is possible to combine the requirements for \( T \) and \( N \)
in a single inequality expression. To prove these results it

will be helpful to break the estimation procedure into two

stages:

1). \( f \) by \( g \)

2). \( g \) by \( \hat{f} \).

What is of interest is to find

\[ P\{ \text{Sup}_{x \in S} |f(x) - \hat{f}(x)| > \varepsilon \} \]

But,

\[ \text{Sup}_{x \in S} |f(x) - \hat{f}(x)| = \text{Sup}_{x \in S} |f(x) - g(x) + g(x) - \hat{f}(x)| \]

\[ \leq \text{Sup}_{x \in S} |f(x) - g(x)| + \text{Sup}_{x \in S} |g(x) - \hat{f}(x)|. \]

It follows that,

\[ P\{ \text{Sup}_{x \in S} |f(x) - \hat{f}(x)| > \varepsilon \} \]

\[ \leq P\{ \text{Sup}_{x \in S} |f(x) - g(x)| > \varepsilon_1 \} \cup \{ \text{Sup}_{x \in S} |g(x) - \hat{f}(x)| > \varepsilon_2 \} \]
\[ \leq \Pr[\sup_{x \in S}|f(x) - g(x)| > \varepsilon_1] + \Pr[\sup_{x \in S}|g(x) - \hat{f}(x)| > \varepsilon_2] \]

when \( \varepsilon_1 + \varepsilon_2 = \varepsilon \).

Thus,

\[ \Pr[\sup_{x \in S}|f(x) - \hat{f}(x)| > \varepsilon] < \alpha \quad \text{provided} \]

\[ \Pr[\sup_{x \in S}|f(x) - g(x)| > \varepsilon_1] < \alpha_1 \]

\[ \Pr[\sup_{x \in S}|g(x) - \hat{f}(x)| > \varepsilon_2] < \alpha_2 \]

where \( \alpha_1, \alpha_2 \) are positive and their sum is \( \alpha < 1 \).

In sections 3.31 and 3.32 it is shown that the estimation procedure given does in fact give the inequalities I and II in 3.3-1.

**3.31 The inequality I**

Lemmas 1–3 give inequality I under different assumptions on \( f(x) \) and for different forms for \( g(x) \).

**Lemma 1**

Let the essential support of \( f(x) \) be the interval \( S = [a, b] \), and let \( (b - a) = B < \infty \). If \( f(x) \) is continuous on \( S \) and \( \sup_{x \in S}|f'(x)| < D \), then with \( g \) as in 3.2-2 and \( K \geq BD/(2\varepsilon_1) \), implies the inequality I in 3.3-1.

**Proof:** By defn of \( m_j \) and expanding \( f \) at a fixed point \( y \in [x_{j-1}, x_j] \), \( m_j \) becomes
\[
m_j = \int_{x_{j-1}}^{x_j} \{f(y) + (x-y)f'(\xi)\} \, dx
\]
where \( \xi \in [x_{j-1}, x_j] \)
\[
= \left\{ f(y) - f'(\xi) \left( y - \frac{x_j + x_{j-1}}{2} \right) \right\} (x_j - x_{j-1})
\]
This implies
\[
|g(y) - f(y)| = \left| f'(\xi) \left( y - \frac{x_j + x_{j-1}}{2} \right) \right|
\leq DB/2K \text{ for all } y
\]
Thus taking
\[
K \geq \frac{DB}{2\varepsilon_1} \Rightarrow |g(y) - f(y)| < \varepsilon_1.
\]

**Lemma 2**

Let the support of \( f(x) \) be \( S \) as in Lemma 1 and let \( \varepsilon_1 = \varepsilon_a + \varepsilon_b \). Suppose \( \exists A \subseteq \mathbb{R} \) and a real number \( B \gg \)
\[
\frac{1}{\pi} \int_{|u| > A} |\phi(u)| \, du \leq \varepsilon_a
\]
and
\[
\sup_{x \in S} |f''(x)| \leq E
\]
Then taking
\[
K \geq \max \left\{ \frac{B^2 E}{24 \varepsilon_b}, \frac{AB}{\pi} \right\}
\]
and \( g(x) \) as in 3.2-3 implies the inequality I in 3.3-1.
Proof:

Let $c_j$ be the point $(x_j + x_{j-1})/2$ and let $b_j$ defined by

$$b_j = m_j/(x_j - x_{j-1}) - f(c_j),$$

be the bias error at $c_j$.

Then, $g(x)$ in 3.2-3 becomes

$$g(x) = \sum_{j=1}^{K} f(c_j) \text{Sic } v_j + \sum_{j=1}^{K} b_j \text{Sic } v_j$$

where $v_j = \{x - (x + \frac{h}{2}) - (j-1)h\} \frac{f(x)}{h}$

and $h = B/K = (x_j - x_{j-1}) \forall j = 1, 2, \ldots K$

By 5.11,

$$\left| \sum_{j=1}^{K} f(c_j) \text{Sic } v_j - f(x) \right| < \varepsilon_a, \quad 3.3-2$$

By expanding $f(x)$ in a Taylor series about $c_j$, $f(x)$ may be written as

$$f(x) = f(c_j) + uf'(c_j) + \frac{u^2}{2} f''(\xi),$$

where $u = (x - c_j)$

and $\xi \in [x_{j-1}, x_j]$.

and using this in the definition of $b_j$ gives
\[ b_j = \frac{h^2}{24} f''(\xi) \leq \frac{h^2 E}{24} \text{ for } \forall x \in S \]

Hence taking \( K \geq \text{Max}\left\{ \frac{BA}{n}, \frac{h^2 E}{24c^2 b} \right\} \)

\[ \Rightarrow |g(x) - f(x)| \leq \left| \sum_{j=1}^{K} f(c_j \text{ sic } v_j - f(x)) \right| + \left| \sum_{j=1}^{K} b_j \text{ sic } v_j \right| \]

\[ \leq \epsilon_a + \epsilon_b = \epsilon_1. \]

\textbf{Lemma 3}

Let the support of \( f(x) \) be \( S \) as in Lemma 1. Suppose \( \exists A \subset R \ni \)

\[ \frac{1}{\pi} \int_{|u|>A} |\phi(u)| du \leq \epsilon_1. \]

Then taking \( K \geq \frac{AB}{\pi} \)

and \( g(x) \) as in 3.2-4, implies the inequality I in 3.3-1.

\textbf{Proof:-}

The proof is a direct application of equations 5.24 and 5.25.

The following theorem summarizes the results of Lemmas 1, 2 and 3.
Theorem 1.

If \( f(x) \) satisfies the requirements of one of Lemmas 1, 2 or 3, then \( K \) and \( g(x) \) given by 3.2-2, 3.2-3, or 3.2-4 respectively implies the inequality I in 3.3-1 for any positive number \( e_1 \).

\[ \Diamond \]

3.32 The inequality II

\( \hat{f} \) is formed from the estimates \( \hat{m}_j \)s of \( m_j \)s in the same way as \( g \) is formed from the \( m_j \)s. That is

\[ \hat{f}(x) = g(x; \hat{m}_1, \hat{m}_2, \ldots, \hat{m}_K). \]

Estimates \( \hat{m}_j \)s only can be used for the \( m_j \)s since the \( m_j \)s cannot be determined exactly. In order to show that \( \hat{f} \) and \( g \) can be made sufficiently close, it is first shown that there are estimates \( \hat{m}_j \)s that are sufficiently close to the \( m_j \)s. Conditions under which these estimates can be found are given. The proof of inequality II is thus in two parts; the first part shows how the \( m_j \)s can be estimated and the second part shows that \( \hat{f} \) and \( g \) are sufficiently close when these estimates are used.

3.321 Determination of the estimations \( \hat{m}_j \)s of \( m_j \)s

To determine the \( m_j \)s, two simultaneous sampling schemes are necessary. This is so since, in the first place, only a finite length \( T \) of signal is used for analysis, and in the second place, only a finite number of points \( N \) are selected.
from this length of signal. (See remark xiv.) Lemmas 4-12 show that this double sampling scheme is in fact possible and indicate how \( T \) and \( N \) are related to the errors in the estimation.

Consider first the first sampling procedure, namely the determination of the length \( T \) necessary for a prespecified accuracy. In particular consider first the estimation of one of the \( m_j \)s, for some \( j = 1,2, \ldots K \), where \( K \) is determined from Lemma 1, 2 or 3. Recall that by definition (3.2-1),

\[
m_j = P\{\omega: X(\omega, t) \in [x_{j-1}, x_j]\}
\]

and that the stationarity assumption on the process \( \Rightarrow \) that this probability is independent of \( t \).

Let

\[
A_j = [x_{j-1}, x_j]
\]

and let \( B_j \subseteq \Omega \) be

\[
B_j = B_j(t) = B(t; x_{j-1}, x_j) = \{\omega: X(\omega, t) \in A_j\}
\]

Define \( z_j(t) \) by

\[
z_j(t) = \begin{cases} 
1 & \text{if } B_j \text{ occurs} \\
0 & \text{otherwise} 
\end{cases}
\]

i.e. \( z_j(t) = I_{B_j}(t) \), the indicator function for the set \( B_j(t) \). For convenience the subscript \( j \) is dropped momentarily.
It is seen that $z(t)$ is a simple random variable defined on $(\Omega, \Theta, P)$, and that

$$E[z(t)] = \int_{\Omega} z(t) dP$$

$$= \int_{B(t)} dP$$

$$= P[B(t)]$$

$$= m. \quad 3.3-5$$

Also $R(\tau) = E[z(t)z(t+\tau)] = \int_{\Omega} z(t)z(t+\tau) dP$

$$= P[B(t) \cap B(t+\tau)] \quad 3.3-6$$

This is independent of $t$ as it depends only on the second order distribution of the process.

Thus $E[z(t)z(t+\tau)]$ may be denoted by $R_z(\tau)$ or $R(\tau)$ for convenience.

Consider the random variable $\theta_T(\omega) = \theta_T = \theta_{T,j}$ defined by

$$\theta_T(\omega) = \frac{1}{T}\int_{0}^{T} z(t) dt. \quad 3.3-7$$

Then using 3.3-5, $E[\theta_T] = \frac{1}{T} \int_{0}^{T} E[z(t)] dt = m. \quad 3.3-8$
and $\text{Var}[\theta_T] = E[(\theta_T - m)^2] = \frac{1}{T^2} \mathbb{E} \left[ \int_0^T \int_0^T z(t)z(s) \, dt \, ds \right] - m^2$

$$= \frac{1}{T^2} \int_0^T \int_0^T E[z(t)z(s)] \, dt \, ds - m^2$$

$$= \frac{1}{T^2} \int_0^T \int_0^T R(t-s) \, dt \, ds - m^2$$

Put $(t-s) = u$

$t = v$

Then the Jacobian is $1$

and the area of integration is changed as shown in the figures. Thus

$$\text{Var}[\theta_T] = \frac{1}{T^2} \left[ \int_{-T}^0 R(u) \, du \int_0^{T-u} v \, dv + \int_0^T R(u) \, du \int_u^T v \, dv \right] - m^2$$

$$= \frac{2}{T^2} \int_0^T \{R(u) - m^2\} (1 - \frac{u}{T}) \, du,$$

3.3-9

since $R(u)$ is symmetric.

The following Lemmas, 4 and 5, which show that the $\theta_T$ can be obtained from a single sample function, can now be proved.
Lemma 4

If \(|R(u) - m^2| \to 0\) as \(|u| \to \infty\), then

\[
p \lim_{T \to \infty} \theta_T = m,
\]

where \(R(u)\) is defined by 3.3-6, \(m\) is defined by 3.2-1 and \(\theta_T\) is defined by 3.3-7.

Proof:–

By 3.3-9, it is seen that the hypothesis on \(R \Rightarrow \text{Var}[\theta_T] \to 0\) as \(T \to \infty\). In view of 3.3-5, application of Tchebycheff's theorem gives the desired result.

\[\diamondsuit\]

The result of Lemma 4 can be strengthened to

Lemma 5

If \(|R(u) - m^2| \leq \frac{C}{|u|^p} \forall u\), then

\[
\lim_{T \to \infty} \theta_T = m \text{ with probability } 1,
\]

provided \(p > 0\), \(0 < C < \infty\) and \(R(u)\), \(m\) and \(\theta_T\) are as in Lemma 4.

Proof:–

Let \(\zeta(t) = z(t) - m\).

Then \(E[\zeta(t)] = 0\)

and \(E[\zeta(t)\zeta(t+u)] = C(u) = R(u) - m^2\)

Further, let \(\eta_T = \eta = (1/T) \int_0^T \zeta(t) dt\)
Then \[ \mathbb{E}[\eta] = 0 \]

and \[ \text{Var}[\eta] = \frac{2}{T} \int_0^T C(u)(1 - \frac{u}{T})du \]

Also, \[ |R(u) - m^2| \leq \frac{C}{|u|^p} \iff |C(u)| \leq \frac{C}{|u|^p} \]

so that \[ \text{Var}[\eta] \leq \frac{2C}{T} \int_0^T u^p(1 - \frac{u}{T})du \]

\[ \leq \frac{C'}{T^p}, \text{ if } 0 < p < 1 \]

or if \( p > 1 \), \[ \text{Var}[\eta] \leq \frac{2C}{T \int_0^1 C(0)du + \int_1^T (\frac{C}{u^p})du} \]

\[ \leq \frac{C''}{T^\gamma}, \text{ where } 0 < \gamma, \]

and where \( C' \) and \( C'' \) are constants.

That is if \( C(u) \leq \text{constant} \frac{1}{|u|^p} \) where \( p > 0 \), then

\[ \text{Var}[\eta] \leq \frac{\kappa}{T^\nu}, \text{ where } \kappa \text{ and } \nu \text{ are positive constants,} \]

\[ \text{and } \kappa \text{ is } < \infty. \]

Now choose \( \beta \) so that \( \nu \beta > 1 \), and form a sequence

\[ \{p^\beta\}, p = 1, 2, \ldots \]

Then for \( T > p^\beta \), \[ \text{Var}[\eta_T] \leq \frac{\kappa}{p^{\nu\beta}}, \text{ therefore, for arbitrary } \epsilon > 0, \]
\[ P\{ |\eta|_p^\beta > \varepsilon \} \leq \frac{\kappa}{p^{\beta \varepsilon^2}} \] and
\[ \sum_{p=1}^{\infty} P\{ |\eta|_p^\beta > \varepsilon \} = \sum_{p=1}^{\infty} \frac{\kappa}{p^{\beta \varepsilon^2}} < \infty \]
so that by the Borel Cantelli Lemma, for \( p \) large enough
\[ \eta|_p^\beta < \varepsilon \text{ a.e. or, } \lim_{p \to \infty} \eta|_p^\beta = 0 \text{ with probability 1.} \]

For the interval between \( p^\beta \) and \( (p+1)^\beta \), note that
\[ E\left\{ \max_{p^\beta < T < (p+1)^\beta} \left| \frac{1}{T} \int_0^T \zeta(t) dt - \frac{1}{T} \int_0^\beta \zeta(t) dt \right|^2 \right\} \]
\[ \leq \frac{1}{p^{2\beta}} E\left[ \int_{p^\beta}^{(p+1)^\beta} |\zeta(t)| dt \right]^2 \]
\[ = \frac{1}{p^{2\beta}} \int_{p^\beta}^{(p+1)^\beta} E[|\zeta(t)\zeta(s)|] dt ds \]
\[ \leq \text{Constant}\left\{ \frac{(p+1)^\beta - p^\beta}{p^\beta} \right\}^2 \]
\[ = \text{Constant}\left\{ \frac{\beta^2}{p^\beta} + \text{higher powers of } \frac{1}{p} \right\} \]
\[ \sim \frac{\text{Constant}}{p^2} \]

' by similar argument to the previous case
\[
\lim_{P \to \infty} \left| \frac{1}{T} \int_0^T \xi(t) dt - \frac{1}{T} \int_0^P \xi(t) dt \right| = 0 \text{ a.e.}
\]

where \( P^\beta < T < (p+1)^\beta \)

\[
\lim_{T \to \infty} \frac{1}{T} \int_0^T \xi(t) dt = \lim_{P \to \infty} \frac{1}{P} \int_0^P \xi(t) dt
\]

\[
= \lim_{P \to \infty} \frac{P^\beta}{T} \eta P^\beta
\]

\[
= 0 \text{ a.e.}
\]

i.e. \( \lim_{T \to \infty} \frac{1}{T} \int_0^T z(t) dt = m, \) with probability 1.

\[
\text{Particular values for } T \text{ are found in Lemmas 6 - 8.}
\]

**Lemma 6**

If \( R(u) \) as defined in 3.3-6 is equal to \( m^2 \) for \( |u| > U_0 \), then

\[
T > T_0 = \max\{U_0/2\alpha c \varepsilon_c^2, U_0\}
\]

\[
\Rightarrow P\{|\bar{\sigma}_T - m| > \varepsilon_c\} < \alpha_c.
\]

**Proof:**

From equation 3.3-9,
\[ \text{Var}[\theta_T] = \frac{2}{T} \int_0^T [R(u) - m^2] (1 - \frac{u}{T}) \, du \quad \text{for } T > U_0 \]

\[ < \frac{2}{T} \int_0^U |R(u) - m^2| \, du \quad 3.3-10 \]

But

\[ |R(u) - m^2| = \left| \int \{z(t) - m\} \{z(t+u) - m\} \, dP \right| \]

\[ \leq \left[ \int |z(t) - m|^2 \, dP \right]^{\frac{1}{2}} \left[ \int |z(t+u) - m|^2 \, dP \right]^{\frac{1}{2}} \]

\[ = \int |z(t) - m|^2 \, dP \]

\[ = m - m^2 \quad \text{since } z(t)^2 = z(t) \]

\[ < \frac{1}{2} \quad \text{since } 0 < m < 1. \]

Hence, from 3.3-10, \[ \text{Var}[\theta_T] < \alpha_c \varepsilon_c^2 \]

when \[ T \geq \text{Max}\{U_0/(2\alpha_c \varepsilon_c^2), U_0\} \]

The desired result follows as a consequence of Tchebycheff’s theorem.

The following Lemma is slightly less restrictive.

**Lemma 7**

If \[ |R(u) - m^2| < \alpha_c \varepsilon_c^2/3 \text{ for } |u| > U_1, \text{ then,} \]
Proof:

\[ T \geq T_0 = \max\{3U_1/(2\alpha_c e_c^2), U_1\} \]

\[ \Rightarrow P\{|\theta_T - m| > \varepsilon_c\} < \alpha_c. \]

\[ \text{Var}[\theta_T] = \frac{2}{T} \int_0^{U_1} (1 - \frac{u}{T})C(u)du + \frac{2}{T} \int_{U_1}^T \frac{\alpha_c e_c^2}{3} (1 - \frac{u}{T})du \]

\[ < \frac{2}{T}(U_1/4 + (T-U_1)\alpha_c e_c^2/3) \]

\[ < U_1/(2T) + 2\alpha_c e_c^2/3 \]

so that for

\[ T \geq \max\{3U_1/(2\alpha_c e_c^2), U_1\}, \]

\[ \text{Var}[\theta_T] < \alpha_c e_c^2. \]

Lemma 8 gives a result that is useful for periodic functions. Recall that (section 3.1) \(x(t)\) may be represented by \(x_c(t)\) or approximated arbitrarily closely by \(x_c(t)\), where

\[ x_c(t) = \sum_{n=1}^M c_n e^{2\pi j n t}. \]

In many cases, the \(w_n\) have a common factor \(w_p\) (see e.g. [32, 61]) and then the signal may be said to be periodic with period \(T_p = 1/w_p\).
Lemma 8

If \( x(t) \) is periodic with period \( T_p \), then taking

\[
T \geq T_0 = T_p \left( \frac{\epsilon_c + 1}{2\epsilon_c} \right)
\]

\[
\Rightarrow P\{|\theta_T - m| > \epsilon_c\} < \alpha_c.
\]

Proof:

From equation 3.3-7,

\[
\theta_T, j = \frac{1}{T} \int_0^T z_j(t) dt.
\]

Let \( J \) be the nearest integer equal to \( T/T_p \). For convenience drop the subscript \( j \) for the moment. Then if \( JT_p > T \),

\[
\theta_T = \frac{1}{T} \int_0^{JT_p} z(t) dt - \frac{1}{T} \int_T^{JT_p} z(t) dt
\]

and

\[
(\theta_T - m) = \frac{JT_p}{T} \epsilon_m - m - \frac{1}{T} \int_T^{JT_p} z(t) dt
\]

so that \((JT_p - T)(m - 1)/T \leq (\theta_T - m) \leq m(JT_p - T)/T\).

Similarly, it is seen that when \( T > JT_p \),

\[
-m(T - JT_p)/T \leq (\theta_T - m) \leq (T - JT_p)(1 - m)/T
\]

and in either case,

\[
|\theta_T - m| < |(JT_p - T)/T| \leq 1/(2J - 1)
\]

\[
< \epsilon_c \quad \text{if } J > (\epsilon_c + 1)/(2\epsilon_c).
\]
For the simultaneous estimation of the $m_j$s, the following result is obtained.

**Lemma 9**

Under the hypothesis of one of Lemmas 6, 7 or 8 $\exists a$ $T_0 \ni$ for $T > T_0$

$$P\left\{ \bigcup_{j=1}^{K} |m_j - \theta_{T,j}| > \varepsilon_c \right\} < \alpha_c.$$

**Proof:**

For $\varepsilon_c > 0$ and $\alpha_c \ni 0 < \alpha_c < 1$, and one of the cases given in Lemmas 6, 7 or 8, $\exists T_0 \ni T > T_0 \Rightarrow$

$$P\left\{ \bigcup_{j=1}^{K} |m_j - \theta_{T,j}| > \varepsilon_c \right\} < \alpha_c / K.$$

Then since

$$P\left\{ \bigcup_{j=1}^{K} |m_j - \theta_{T,j}| > \varepsilon_c \right\} \leq \sum_{j=1}^{K} P\{ |m_j - \theta_{T,j}| > \varepsilon_c \},$$

the result follows. $\square$

Let $t = (t_1, t_2, \ldots, t_N)$ and let

$\Gamma = \{ t: t_i \in [0,T], i = 1, 2, \ldots, N \}$. Let $P'$ be the probability measure on $\Gamma$ determined by the product of $N$ independent uniform distributions on $[0,T]$. Thus the probability space $\{\Gamma, \Lambda, P'\}$ is obtained, where $\Lambda$ is the class of all measurable
sets in $\Gamma$. For convenience denote $t$ by $\gamma$. Thus the selection of $N$ points $t_1, t_2, \ldots, t_N \in [0,T]$ corresponds to the selection of a point $\gamma \in \Gamma$.

It is easily established that the $\hat{m}_j$s as defined by 3.2-1 has a multinomial distribution. (See section 3.4.) Then

$$E_T[\hat{m}_j(\gamma, \omega)] = \theta_{T,j}(\omega)$$

and

$$\text{Var}_T[\hat{m}_j(\gamma, \omega)] = \frac{\theta_{T,j}(\omega)(1-\theta_{T,j}(\omega))}{N}.$$ 

Lemmas 10 and 11 give $N_0$ necessary for a given accuracy.

**Lemma 10**

Given an $\varepsilon_d$ and an $\alpha_d \exists$ an $N_0 \geq \frac{\alpha_d}{\varepsilon_d}, \forall N \geq N_0 = N_d(\varepsilon_d, \alpha_d)$,

$$P\left[|\theta_{T,j} - \hat{m}_j| > \varepsilon_d\right] < \alpha_d,$$

where $\theta_{T,j}$ is given by 3.3-7 and $\hat{m}_j$ is given by 3.2-5.

**Proof:**

From the statistics of $\hat{m}_j$ (section 3.4),

$$E_T[\theta_{T,j} - \hat{m}_j] = 0,$$

$$\text{Var}_T[\theta_{T,j} - \hat{m}_j] = \frac{\theta_{T,j}(1-\theta_{T,j})}{N}.$$ 

Since $0 < \theta_{T,j} < 1$,

$$\text{Var}_T[|\theta_{T,j} - \hat{m}_j|] < 1/(4N).$$

The result follows using Tchebycheff's theorem, with
Note that the important thing here is that $N_0$ is independent of $\theta_{T,j}$'s and hence of $T$.

The following Lemma deals with the simultaneous estimation of all the $\theta_{T,j}$'s.

**Lemma 11**

Given an $\epsilon_d, \alpha_d \exists$ an $N_0 \gg$, $\forall N > N_0 = N_0(\epsilon_d, \alpha_d)$,

$$P'\left\{ \bigcup_{j=1}^{K} \left[ |\theta_{T,j} \hat{\hat{n}}_j| > \epsilon_d \right] \right\} < \alpha_d,$$

where $\theta_{T,j}, \hat{\hat{n}}_j$, are as in Lemma 10 and $K$ is determined as in Lemmas 1, 2 or 3.

**Proof:**

$$P'\left\{ \bigcup_{j=1}^{K} \left[ |\theta_{T,j} \hat{\hat{n}}_j| > \epsilon_d \right] \right\}$$

$$\leq \sum_{j=1}^{K} P'\left\{ |\theta_{T,j} \hat{\hat{n}}_j| > \epsilon_d \right\}$$

$$\leq \sum_{j=1}^{K} \frac{\theta_{T,j} (1 - \theta_{T,j})}{N_0 \epsilon_d^2}$$

$$\leq \frac{(K - 1)}{(K \epsilon_d^2)}$$

since
\[
\sum_{j=1}^{K} \left( \frac{1}{K} - \theta_{T,j} \right)^2 = \frac{1}{K} - \frac{2}{K} \sum_{j=1}^{K} \theta_{T,j} + \sum_{j=1}^{K} \theta_{T,j}^2 = \sum_{j=1}^{K} \theta_{T,j}^2 - \frac{1}{K} > 0.
\]

The Lemma follows with \( N_0 = (K - 1)/(K \alpha \varepsilon_\alpha^2). \)  

Consider now the product probability space \( \{\Omega \times \Gamma, \Theta \times \Lambda, P \times P'\} \) and denote \( P \times P' \) by \( P^* \). What is of interest is to find the probability of error in estimating the \( m_j \)'s by the \( \hat{m}_j \)'s. This probability is

\[
P^*\{\text{Sup}_j |m_j - \hat{m}_j| > \varepsilon_c\}.
\]

Lemmas 9 and 11 can be combined to give the desired result, and this is done in Lemma 12.

**Lemma 12**

Given a process satisfying the hypothesis of one of Lemmas 6, 7 or 8, an \( \varepsilon_e \) and an \( \alpha_\alpha \) then \( \exists \ a \ T_0 \) and an

\( N_0 \Rightarrow N > N_0 \) and \( T \geq T_0 \),

\[
\Rightarrow P^*\{\text{Sup}_j |\hat{m}_j - m_j| > \varepsilon_e\} < \alpha_\alpha,
\]

where \( \hat{m}_j \) is defined by 3.2-5, and \( m_j \) is defined by 3.2-1.
Proof:-

Let $\varepsilon_c + \varepsilon_d = \varepsilon_e$, and let $\alpha_c + \alpha_d = \alpha_2$. Then note that

$$P^*\{\sup_j |\hat{m}_j - m_j| > \varepsilon_c\}$$

$$< P^*\{[\sup_j |\theta_{T, j} - m_j| > \varepsilon_c] \cup [\sup_j |\theta_{T, j} - \hat{m}_j| > \varepsilon_d]\}$$

$$< P^*\{\sup_j |\theta_{T, j} - m_j| > \varepsilon_c\} + P^*\{\sup_j |\theta_{T, j} - \hat{m}_j| > \varepsilon_d\}.$$ 

Consider the first term on the right hand side. Denote the set in $\{\}$ by $F$ and the $\omega$ and $\gamma$ sections by $F_\omega$ and $F_\gamma$ respectively. Then,

$$F_\omega = \begin{cases} \Gamma & \text{if } \omega \in F_\gamma \\ \text{null set} & \text{if } \omega \notin F_\gamma \end{cases}$$

and (Halmos [45] p.144),

$$P^*\{F_\omega\} = \int_{\Omega} P'[F_\omega]\,dP(\omega)$$

$$= \int_{\Omega} I_{F_\gamma}(\omega)\,dP(\omega).$$

Note that

$$F_\gamma = \{\omega: \sup_j |m_j - \theta_{T, j}| > \varepsilon_c\}$$

so that, using Lemma 9,

$$P^*\{F\} = P\{F_\gamma\} < \alpha_2.$$

Consider now the second term on the right hand side. Denote the set in $\{\}$ by $G$ and the $\omega$ and $\gamma$ sections by $G_\omega$ and $G_\gamma$.
respectively. Then
\[ G_\omega = \{ \gamma: \sup_j d_{T_j}(\omega) - \hat{m}_j(\gamma, \omega) > \epsilon_d \} \]
and thus, using Lemma 11,
\[ P^*\{G\} = \int_\Omega P^\prime\{G_\omega\} dP(\omega) \]
\[ < \alpha_d \int_\Omega dP(\omega) \]
\[ = \alpha_d. \]
Hence the result, provided \( T_0 \) is as in Lemma 9 and \( N_0 \) is as in Lemma 11.

3.322 To show that \( \hat{f} \) and \( g \) are close

Inequality II can now be proved, and this is done in the following Theorem.

Theorem 2.

If \( g(x) \) is defined as in one of
i) 3.2-2
ii) 3.2-3
or iii) 3.2-4,
and \( \hat{f}(x) \) is defined by
\[ \hat{f}(x) = g(x; \hat{m}_1, \hat{m}_2, \ldots, \hat{m}_K), \]
then
\[ P^*\{\sup_{x \in S} |\hat{f}(x) - g(x)| > \epsilon_2\} < \alpha_2. \]
Proof: -

i) \(|\hat{\phi}(x) - g(x)| = \left| \frac{\hat{m}_j}{h} - \frac{\hat{m}_j}{h} \right| x \in [x_{j-1}, x_j], j = 1, \ldots, K \leq \epsilon_x x \in S.\]

\(= \epsilon_2.\)

The desired result follows from Lemma 12.

ii) Let \(\hat{\epsilon}_j\) be the error in estimating \(m_j.\) Then, with the notation of Lemma 2, \(\hat{\epsilon}_j\)

\[\hat{f}(x) = \sum_{j=1}^{K} \frac{\hat{m}_j}{h} \text{Sic}(v_j)\]

\[= \sum_{j=1}^{K} \frac{m_j + \hat{\epsilon}_j}{h} \text{Sic}(v_j)\]

\[= g(x) + \sum_{j=1}^{K} \frac{\hat{\epsilon}_j}{h} \text{Sic}(v_j).\]

Therefore,

\[|\hat{f}(x) - g(x)| \leq \sum_{j=1}^{K} \frac{\hat{\epsilon}_j}{h} \text{Sic}(v_j)|\]

\[\leq (2\epsilon_x/n) \sum_{j=1}^{K/2} \frac{1}{jh} \] where \(\epsilon_x = \sup \hat{\epsilon}_j\)

\[\leq \frac{2}{\sqrt{n}} \epsilon_x \log_2(K/2) = \epsilon_2.\]
and hence, by Lemma 12, the result.

\[ \hat{f}(x) = \frac{\pi}{h} \sum_{j=1}^{\infty} \hat{a}_j \cos(c_j) \]

\[ = \frac{\pi}{h} \sum_{j=1}^{\infty} (d_j + \eta_j) \cos(c_j) \]

\[ = g(x) + \frac{\pi}{h} \sum_{j=1}^{\infty} \eta_j \cos(c_j) \]

where \( \eta_j = \sum_{i=1}^{j} \varepsilon_i, j = 1, 2, \ldots K-1 \), and \( \eta_j = 0 \) for \( j \geq K \)

Since \( d_j = 1 \) for \( j \geq K \).

Thus,

\[ |\hat{f}(x) - g(x)| = \frac{\pi}{h} \sum_{j=1}^{K-1} \eta_j |\cos(c_j)|. \]

The right hand side is,

\[ \left| \frac{\pi}{h} \left( \sum_{i=1}^{1} \varepsilon_1 \cos(c_1) + \sum_{i=1}^{2} \varepsilon_1 \cos(c_2) + \sum_{i=1}^{K-1} \varepsilon_i \cos(c_{K-1}) \right) \right| \]

\[ = \left| \frac{\pi}{h} \left( \sum_{j=1}^{1} \varepsilon_1 \cos(c_j) + \sum_{j=2}^{K-1} \varepsilon_{K-1} \cos(c_{K-1}) \right) \right| \]

\[ \leq K \varepsilon_2 / (2h) \]

\[ = \varepsilon_2 \]
and again the result follows from Lemma 12.

The overall result is now stated in the following Theorem.

Theorem 3
Suppose \( \{X(\omega, t), t \in \mathbb{R}\} \) is a second order stationary random process satisfying the hypothesis of Theorem 2 and has a p.d.f. \( f(x) \) satisfying the hypothesis of Theorem 1. Then, given an \( \varepsilon > 0 \) and a positive \( \alpha \) less than 1, \( \exists \) an estimate \( \hat{f}(x) = \hat{f}(x; N, T) \), a number \( N_0 \), and a length \( T_0 \), \( \forall N > N_0 \) and \( T > T_0 \),

\[
P*\{\sup_{x \in \mathbb{S}} |\hat{f}(x) - f(x)| > \varepsilon\} < \alpha.
\]

Proof:-

Theorems 1 and 2 and the formula 3.3-1 with \( \alpha_1 = 0 \) implies the result.

Remarks

i) In Lemma 1, since \( K \) depends on \( f'(x) \) and \( B \) only and \( f'(\rho x) = \frac{1}{\rho} f'(x) \), \( K \) will be unchanged even if the points \( a \), \( b \), are shifted provided \( \varepsilon_1 \) is expressed as a percentage of \( f \).

ii) From considerations of properties of Fourier transform pairs, (see e.g. [30,107,108]), it is seen that many processes satisfy the requirement
\[ \int_{|u|>A} |\phi(u)| \, du \to 0 \text{ as } A \to \infty. \] 3.3-11

In particular, if the nth derivative \( f^n(x) \) exists and \( |f^n(x)| \) is integrable, then [30],

\[ |\phi(u)| < \text{constant}/(|u|^n). \]

Also, in many cases, p.d.f.s are expandable in terms of the Hermite functions \( Y_n(x) \) (see e.g. [30,78]), and in these cases the condition 3.3-11 is satisfied. This can be seen by noting that

\[ Y_n(x) = H_n(x)e^{-x^2/2} \]

where

\[ H_n(x) = (-1)^n e^{x^2} \left( \frac{d^n}{dx^n} e^{-x^2} \right), \]

and that the Fourier transform of the Hermite functions are the same Hermite functions apart from a constant (see e.g. [107]). See also section 3.5.

iii) In certain cases, \( f(x) \) is "band limited" to \( A \) (written \( f(x) \in B[-A,A] \)), which means that

\[ \phi(u) = 0 \text{ for } |u| > A, \]

\( A \) being a constant. In such cases, it is possible to obtain error free reconstruction from either values of \( f(x) \) at points every \( \frac{\pi}{A} \) apart, or areas of the p.d.f. within class intervals of width \( \frac{\pi}{A} \) (see chapter 5).
iv) When \( f(x) \) is not band limited to \( A \), but is reconstructed using either equation 5.8 or 5.24 with the incorrect assumption that \( f(x) \in B[-A,A] \), then an upper bound \( \varepsilon_u \) for the error in reconstruction due to "aliasing" is (see chapter 5)

\[
\varepsilon_u = \int_{|u|>A} |\phi(u)| \, du.
\]

v) In Lemma 2, note that the condition \( |f''(x)| \leq E \), \( \Rightarrow \) that if \( \varepsilon > 0 \), then for any \( A_E \geq 2BE/(\pi \varepsilon) \),

\[
\frac{1}{\pi} \int_{|u|>A_E} |\phi(u)| \, du < \varepsilon.
\]

To show this, note that

\[
\phi(u) = \int_{-\infty}^{\infty} e^{jux} f(x) \, dx
\]

\[
= -\frac{1}{u} \int_{-\infty}^{\infty} e^{jux} f''(x) \, dx,
\]

(as is seen by integrating by parts twice), and hence,

\[
\frac{1}{\pi} \int_{|u|>A_E} |\phi(u)| \, du
\]

\[
= \frac{1}{\pi} \int_{|u|>A_E} \left| \frac{1}{u} \int_{-\infty}^{\infty} e^{jux} f''(x) \, dx \right| \, du
\]

\[
\leq \frac{1}{\pi} \int_{|u|>A_E} \frac{BE}{(u^2)} \, du
\]
vi) In Lemma 2 if the assumption of support is replaced by essential support for \( f(x) \), then there is an additional error term to be considered, which however is small, as shown below,

\[
\left| \sum_{j=1}^{K} f(c_j) \text{Sic}(v_j) - f(x) \right|
\]

\[
= \left| \sum_{j=-\infty}^{\infty} f(c_j) \text{Sic}(v_j) - \sum_{j=-\infty}^{0} f(c_j) \text{Sic}(v_j) - \sum_{j=K+1}^{\infty} f(c_j) \text{Sic}(v_j) - f(x) \right|
\]

\[
\leq \left| \sum_{j=-\infty}^{\infty} f(c_j) \text{Sic}(v_j) - f(x) \right| + \sum_{0}^{\infty} + \sum_{K+1}^{\infty}
\]

\[
\leq \varepsilon_a + \varepsilon_s
\]

where \( \varepsilon_s \) is the additional error, and

\[
\varepsilon_s \leq \int_{x \notin S} f(x) \, dx.
\]

vii) In Lemmas 2 and 3 the simple transformation of the r.v. \( x \) into \( y \) by

\[
y = \rho x + \gamma
\]
has the effect upon the error and $K$ as follows:

If $K$ is unaltered the error is replaced by $\varepsilon_1/\rho$. Note that the fractional error is unchanged.

To prove this let $f_2(y)$ be the p.d.f. of the r.v. $y$ and let $\phi_2(u)$ be its characteristic function. Then

$$\phi_2(u) = e^{ju} \phi(\rho u)$$

so that

$$\frac{1}{\pi} \int_{|u| > A} |\phi(u)| \, du < \varepsilon_a \Rightarrow \frac{1}{\pi} \int_{|u| > A/|\rho|} |\phi_2(u)| \, du < \varepsilon_a/|\rho| \quad 3.3-12$$

Furthermore,

$$f_2(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-juy} \phi_2(u) \, du$$

which on substituting for $\phi_2(u)$ and $y$ in the integrand gives

$$f_2(y) = f(x) \rho^{-a} \quad 3.3-13$$

which in turn gives

$$f''_2(y) = f''(x)/\rho^a \quad 3.3-14$$

The equations 3.3-12 to 3.3-14 prove the assertion in the remark.

viii) In cases in which the p.d.f. is not continuous, (i.e. the c.d.f. has a discontinuity of the first kind), the p.d.f. cannot be accurately determined by these methods. However, some idea of the p.d.f. can still be obtained. In
the case of the method in Lemma 1 for example, the errors will be in the interval where the discontinuity exists. In the other cases, the errors in parts other than the discontinuous points will be slightly increased, although the main errors will still be around the discontinuous points (See examples in section 3.5).

ix) As an example of Lemma 8, consider the experimental determination of the p.d.f. of a random phase sinewave process. "Bias" can arise if incomplete periods are present in the length of the signal being analysed.

Thus the distribution of A in fig. 3.2 may be determined to any desired precision with the methods given here. But in the case of B in the figure, the p.d.f. will be greater for positive x than for negative x. A case such as C will also be biased, but clearly, in this case, the effect of the incomplete period is considerably less.
The bias is seen to be due to the incomplete portion of a cycle in the signal, and is the fraction of incomplete half cycles over the total length of signal.

x) The error as given by Lemma 8 can be decreased by "tailoring" the signal at the ends to an extent corresponding to about a quarter of a cycle of the lowest frequency present in the signal. For example, the signal

\[ x_T(t) = \begin{cases} 
  x(t) & t \in [0,T] \\
  0 & \text{otherwise}
\end{cases} \]

may be multiplied by a window \( K_T(t) \) to give

\[ y_T(t) = K_T(t)x_T(t) \]

where

\[ K_T(t) = \begin{cases} 
  \kappa(t,\gamma) & t \in [0,\gamma] \\
  1 & t \in [\gamma, T-\gamma] \\
  \kappa(t+\gamma-T,\gamma) & t \in [T-\gamma,T] \\
  0 & t \notin [0,T]
\end{cases} \]

and \( \gamma = T_P/4, \ 0 \leq \kappa(t,\gamma) \leq 1, \kappa(0,\gamma) = 0 \) and \( \kappa(\gamma,\gamma) = 1 \).

For example \( \kappa(t,\gamma) \) may be \((1 - \cos \frac{\pi t}{\gamma})/2\).

The tailoring of the ends will tend to increase the bias due to some frequency components and decrease those due to others - that is, even out the bias effects.

The error given by Lemma 8 will be greater than that
obtained when correction for end effects are made. Even so, Lemma 8 gives an idea of the error to be expected, or signal length needed for a given error. Lemma 8 is derived under fairly general conditions and is less restrictive than those suggested by [9] or [2] and is easier to apply than [2]. The signal length requirement is also generally lower than required by [9] which, moreover is only an intuitive result.

Note that the signal need not be periodic but only needs to repeat its statistical characteristics, i.e. each adjacent length $T_p$ is to be statistically "similar" in the sense that $T_x$ be constant for each interval.

xi) The assumption of bounded support in Lemmas 1-3 is not a particularly restrictive one since any physical system would necessarily have only a finite amplitude range.

xii) The assumption $|R(u) - m^2| \to 0$ as $|u| \to \infty$ in Lemma 2 is equivalent to saying that the events $B(t)$ and $B(t+u)$ tend to be independent for large $|u|$, since

$$R(u) = P[B(t) \cap B(t+u)] = m^2$$

if $B(t)$ and $B(t+u)$ are independent. This is a reasonable assumption since it is only saying that the present is independent of the distant past (and or future).

xiii) Although analysis in a computer has been assumed
throughout, the results are applicable in parts to analysis with analogue and hybrid equipment. For example, the result for signal length can be used for analogue analysis. Also, in a hybrid system, it may be possible to directly obtain the samples by reading off the values at times $t(i)$, $i = 1, 2, \ldots N$. Here $t(i)$, $i = 1, 2, \ldots N$ is the rearrangement of the times \{t, $i = 1, 2, \ldots N$\} corresponding to a $\gamma \in \Gamma$ in Lemma 11 such that $t(i) \leq t(i+1)$ for all $i = 1, 2, \ldots N-1$. Reconstruction of the $f(x)$ may be accomplished by passing a sequence of pulses of amplitudes $(m \cdot N)$ at appropriate times through a low-pass filter.

xiv) As mentioned earlier, the equation 3.3-7 is tedious to evaluate directly. This is seen to be so since the values of $x(t)$ has to be evaluated for a very large number of points to determine whether $z(t)$ in 3.3-7 is 0 or 1. However it is not necessary to distinguish between every possible amplitude value of $x(t)$ since the signals are always frequency limited (and hence continuous) and thus the amplitudes of neighbouring points cannot vary greatly from each other. In particular the points stored in the computer may be assumed to have been sampled at at least the Nyquist rate (see chapter 2), which means that the largest possible amplitude range between successive points is the maximum range of the signal. Thus the extreme case is when a maximum amplitude sinusoidal com-
ponent at the maximum frequency is present. If the original points are spaced at intervals of h seconds, this component may be written as

\[ x(t) = B \cos \left( \frac{\pi t}{h} \right) \]

and the maximum rate of change is

\[ \frac{u_B}{h} \]

at \( t = h/2 \). To ensure that the amplitude range \( x(t_{i+1}) - x(t_i) \) is less than \( B/K \),

\[ \frac{B}{2K} > B \sin \left( \frac{\pi \delta}{2h} \right) \]

where \( \delta = (t_{i+1} - t_i) \). For small \( \delta \), this may be approximated to

\[ \left( \frac{\pi \delta}{2h} \right) < 1/(2K), \]

that is

\[ \delta < \frac{h}{\pi K} \]

Thus, to ensure every possible case, it is necessary to subdivide \( h \) into at least \((\pi K)\) further divisions. For example, if \( K=100 \) and there are 6000 points stored, the number of points to be evaluated is about \( 2 \times 10^6 \). On the other hand, as is shown in section 3.4, 10,000 points randomly sampled will give a maximum error of (0.02) with a confidence level of 99%.
3.4 Gaussian Approximation

Let a sample of size $N$ be drawn, and let the $N$ samples be separated into $K$ classes. Let the number falling in the $j$th class be $N_j$. Thus

$$\sum_{j=1}^{K} N_j = N$$

and $\hat{m}_j$ is taken to be $N_j/N$.

The probability distribution of the $\hat{m}_j$s is easily shown to be (see e.g. [24, 38, 108]) the multinomial distribution. That is

$$p(\hat{m}_1, \hat{m}_2, \ldots, \hat{m}_K) = \frac{N!}{(N_1!)(N_2!)...(N_K!)} \prod_{j} \theta_j^{N_j}$$

where $\theta_j = E[\hat{m}_j]$.

Thus $E[\hat{m}_j] = \theta_j$ and $\text{Var}[\hat{m}_j] = \frac{\theta_j Q_j}{N}$ where $Q_j = (1 - \theta_j)$.

If however $N$ is large so that the distribution of $\hat{m}_j$s can be taken to be Gaussian (see e.g. [24, 81]), then with

$$Z_j = (\hat{m}_j - \theta_j)/\sqrt{\theta_j Q_j/N},$$

$$P[|Z_j| > z(1 - \alpha_d/2)] < \alpha_d,$$

where $z(1 - \alpha_d/2)$ is the ordinate having an area of $\alpha_d$ in the tails. From this it follows that for
\[ N > N_0 = \frac{1}{2} \left( \frac{Z(1 - \alpha_d/2)}{\varepsilon_d} \right)^2 \geq \frac{Z(1 - \alpha_d/2)^2}{\varepsilon_d^2} \geq \theta_j Q_j, \]

\[ P\{|\hat{m}_j - \theta_j| > \varepsilon_d\} < \alpha_d. \]

Moreover when \( K \) simultaneous estimations are made, if

\[ N > N_0 = \frac{1}{2} \left( \frac{Z(1 - \alpha_d/2K)}{\varepsilon_d} \right)^2 \]

where \( Z(1 - \alpha_d/2K) \) is the ordinate having an area of \( \alpha_d/K \) in the tails of the normalized Gaussian distribution, then

\[ P\{\text{at least one error} > \varepsilon_d\} \]

\[ = \sum_{1}^{K} P\{\text{each error} > \varepsilon_d\} \]

\[ < \sum_{1}^{K} \alpha_d/K \]

\[ = \alpha_d. \]

For example, let \( \varepsilon_d = 0.1, \alpha_d = 0.05 \) and \( K = 100. \)

Then \( [1] \) \( Z(1 - \alpha_d/2K) = 3.5 \), so that

\[ N_0 = 0.99 \times 100 \times 20 = 1980 \] without the Gaussian approximation and

\[ N_0 = \frac{1}{2} \times 3.5 \times 3.5 \times 100 \approx 300 \] with the Gaussian approximation.

For another example let \( \varepsilon_d = 0.02 \) and \( \alpha_d = 0.01 \) with \( K \) as above. Here \( Z(1 - \alpha_d/2K) = 3.9 \) giving

\[ N_0 \approx 250,000 \] without the Gaussian approximation.
<table>
<thead>
<tr>
<th>α</th>
<th>ε</th>
<th>0.1</th>
<th>0.05</th>
<th>0.02</th>
<th>0.01</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>1.65</td>
<td>70</td>
<td>270</td>
<td>1,700</td>
<td>6,800</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1,000</td>
<td>4,000</td>
<td>25,000</td>
<td>100,000</td>
</tr>
<tr>
<td>0.05</td>
<td>2.0</td>
<td>100</td>
<td>400</td>
<td>2,500</td>
<td>10,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2,000</td>
<td>8,000</td>
<td>50,000</td>
<td>200,000</td>
</tr>
<tr>
<td>0.01</td>
<td>2.57</td>
<td>165</td>
<td>660</td>
<td>4,200</td>
<td>16,500</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10,000</td>
<td>40,000</td>
<td>25×10⁴</td>
<td>10⁸</td>
</tr>
<tr>
<td>0.005</td>
<td>2.81</td>
<td>200</td>
<td>790</td>
<td>4,950</td>
<td>19,800</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0025</td>
<td>3.03</td>
<td>230</td>
<td>920</td>
<td>5,750</td>
<td>23,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.001</td>
<td>3.3</td>
<td>250</td>
<td>1,000</td>
<td>6,300</td>
<td>25,200</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0005</td>
<td>3.5</td>
<td>300</td>
<td>1,200</td>
<td>7,600</td>
<td>36,500</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0001</td>
<td>3.9</td>
<td>380</td>
<td>1,500</td>
<td>9,500</td>
<td>38,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Gaussian Approximation: (K=1), upper figures. The low values of α are used for large K. e.g., if K=100, α=0.0001 gives Kα=0.01.

Without Gaussian Approximation: Lower figures.

**TABLE 3.1. No. Of SAMPLES N FOR ERROR<ε AND CONFIDENCE>(1-α).**
Some common p.d.f.s are here examined for their channel requirements. The p.d.f.s are chosen for their widely differing shapes. The Gaussian distribution $f_g$ is smooth and symmetrical. The uniform distribution $f_u$ has a sharp discontinuity at each end. The "sinusoidal distribution" $f_s$ (the p.d.f. of a random phase sinewave) is asymptotic and discontinuous at the ends. The Rayleigh distribution $f_r$ is smooth to the right of the origin, but it is asymmetrical and has a discontinuous derivative at the origin. None of these are band limited functions (i.e. $f_g, f_u, f_s, f_r \in B[-A,A]$ for any finite A), so that error free reconstruction is not possible. In fact the error bound as given by 3.3-12 is infinite for $f_u$ and $f_s$, though very large errors will only result at the discontinuities, particularly where the jumps are large. (See figs. 3.3-3.6.) However, in practice p.d.f.s will not normally exhibit sharp jumps. This is seen to be particularly so when it is noted that the distribution of the sum $z = x + y$ of independent random variables $x$ and $y$ is the convolution of the separate distributions of $x$ and $y$. The characteristic function of $z$ is the product of the characteristic functions of $x$ and $y$. Since in practice, signals are contaminated with noise which is often Gaussian
(see e.g. [78]) and the characteristic function of the Gaussian distribution is of the form \( \exp(-k^2u^2) \), characteristic functions of practical signals will tend to satisfy the requirements of Lemmas 1 to 3 (see section 3.53). In any case, the examples will give an idea of the number of channels that one should take when determining p.d.f.s, particularly when the conditions of Lemmas 1 to 3 are not completely satisfied.

The numerical values used in the following calculations have been taken largely from [1] and also from [30]. The subscript for \( \varepsilon_1 \) in Lemmas 1 to 3 will be dropped for convenience.

### 3.51 Gaussian distribution

Values for \( K \) for \( \varepsilon = 0.05 \times \text{Max} |f_g(x)| \) are calculated below for the three separate cases in Lemmas 1 to 3.

#### 3.51.1 Method of Lemma 1

\[
 f_g(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-x^2/(2\sigma^2)} .
\]

Then,

\[
 f'_g(x) = -\frac{x}{\sigma^2} f_g(x)
\]

and

\[
 f''_g(x) = \frac{1}{\sigma^2} f_g(x)\{(\bar{x}/\sigma)^2 - 1\}
\]

\[ D = \text{Max} |f'_g(x)| \div 0.24/\sigma^2 , \]

\[ \varepsilon = 0.05f_g(0) \div 0.02/\sigma , \]
so that \( K > \frac{BD}{2\varepsilon} > 30. \)

### 3.5.12 Method of Lemma 2

Using results of Lemma 2, and in view of remark vii, let

\[
f_g(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \quad (i.e. \mu = 0, \sigma = 1)
\]

Then

\[
\phi(u) = e^{-u^2/2},
\]

\[
B = \max |f''_g(x)| = f'_g(0) = 0.4,
\]

\[
B > 5,
\]

and

\[
\varepsilon = 0.02005.
\]

\[
\frac{B^2 E}{24\varepsilon} = 1/(2.4\varepsilon_b).
\]

Take

\[
\varepsilon_\alpha = 0.00005 \text{ and } \varepsilon_\beta = 0.02,
\]

giving

\[
K > 22.
\]

### 3.5.13 Method of Lemma 3

Finally using the result of Lemma 3, gives

\[
K > 5.
\]

Note however that the hypothesis that the support of \( f(x) \) is finite is not strictly correct in this case.
3.52 Uniform distribution

\[ f_U(x) = \begin{cases} \frac{1}{B} & x \in [-B/2, B/2] \\ 0 & \text{otherwise.} \end{cases} \]

3.52.1 Method of Lemma 1

The method does not strictly apply since there is a discontinuity at + and - B/2. Clearly, the smaller the intervals, the more accurately the discontinuous points are located. The errors are largely where the discontinuity falls.

3.52.2 Method of Lemma 2

The characteristic function is [1]

\[ \frac{\sin(Bu/2)/(Bu/2)}{\sin(z)/z} = \frac{\sin(z)/z}{\sin(z)/(n\pi+\pi)} \]

where \( z = Bu/2 \)

Consider

\[ s_n = \int_{n\pi}^{(n+1)\pi} |\sin(z)/z| \, dz. \]

\[ s_n > \int_{n\pi}^{(n+1)\pi} \sin(z)/(n\pi+\pi) \, dz \]

\[ = 2/(n\pi+\pi). \]
Hence,
\[ \int_{A}^{L} \left| \frac{\sin(z)}{z} \right| dz \]
\[ > \sum_{r=n}^{1} \frac{1}{(r+1)} \]

where \( n, l \) are integers \( \geq \pi n \) \( \geq A \) and \( 1 \pi \leq L \), and thus
\( I_{L} \to \infty \) as \( L \to \infty \) \( \forall \) finite \( A \).

Thus the method of Lemma 2 does not strictly apply in the sense that the error bound in Lemma 2 is \( \infty \). However, as mentioned in remark viii), and as illustrated by fig. 3.4, some idea of the shape is still obtainable.

3.523 Method of Lemma 3.

Similar remarks to that in section 3.522 above apply.

As a modification to this case, consider the distribution shown at right.

\[ D = \frac{4}{B^2} \]
\[ \phi(u) = \left\{ \frac{\sin(uB/4)}{(uB/4)} \right\}^2 \]
\[ < \frac{16}{(B^2u^2)}. \]

\( E \) is zero except at \( x = 0 \) where it is undefined. However, the bias there is small. Hence in this case, the requirements of Lemmas 1, 2 or 3 may be considered to be satisfied.
3.53 Random phase sinewave distribution

\[
\begin{align*}
    f_S(x) &= \begin{cases} 
        1/\{\pi(v^2 - x^2)^\frac{1}{2} \} & |x| \leq v \\
        0 & |x| > v 
    \end{cases} \\
    f_S'(x) &= \frac{x}{\pi(v^2 - x^2)^\frac{3}{2}}
\end{align*}
\]

and

\[
\frac{Bf_S'(x)}{f_S(x)} = \frac{2vx}{v^2 - x^2}
\]

which is unbounded as \( x \) approaches \( v \). Also,

\[
\phi(u) = \int_{-v}^{v} \frac{e^{jux}}{\pi(v^2 - x^2)^{\frac{1}{2}}} \, dx.
\]

Putting \( x = v\cos(\theta) \) into the above gives

\[
\phi(u) = \frac{1}{\pi} \int_{0}^{\pi} e^{jvu\cos(\theta)} \, d\theta
\]

\[
= J_0(vu),
\]

where \( J_0(\cdot) \) is the Bessel function of order zero \([1,109]\).

Since \( |J_0| \) is of the order of \( 1/\sqrt{x} \),

\[
\int_{|u|>A} |J_0(vu)| \, du
\]

is clearly not bounded for finite \( A \).

Thus none of the methods in the Lemmas 1 to 3 strictly apply.
Consider now the following modification. Define an

\[ f(x) = \begin{cases} 
  \frac{k}{\sqrt{v^2 - x^2}} & |x| < \lambda < v \\
  0 & \text{otherwise.}
\end{cases} \]

Here \( k \) is a normalizing constant.

Then, \( D = \lambda \sqrt{\pi(v^2 - \lambda^2)} \) so that Lemma 1 will give a finite value for \( K \). But,

\[ \phi(u) = 2k \int_0^\lambda \frac{\cos(ux)}{(v^2 - x^2)^{1/2}} \, dx \]

\[ = \left( \frac{2k}{v} \right) \int_0^\lambda \left[ 1 + \frac{1}{2} \left( \frac{x}{v} \right)^2 + \frac{3}{8} \left( \frac{x}{v} \right)^4 + \ldots \right] \cos(ux) \, dx \]

This becomes, with \( y = ux \) and integrating term by term

(see e.g. [90, 109])

\[ \frac{k}{uv} \sin(\lambda u) \left[ 1 + \frac{1}{2} \left( \frac{1}{u} \right)^2 \right] \left[ (\lambda u)^2 \sin(\lambda u) + 2\lambda u \cos(\lambda u) - \sin(\lambda u) \right] + \frac{3}{8} \left( \frac{1}{u} \right)^2 \left[ (\lambda u)^4 \sin(\lambda u) + 4(\lambda u)^3 + \ldots \right] \]

\[ = \frac{k}{uv} \sin(\lambda u) \left[ 1 + \frac{1}{2} \left( \frac{\lambda}{v} \right)^2 + \frac{3}{8} \left( \frac{\lambda}{v} \right)^4 + \frac{5}{16} \left( \frac{\lambda}{v} \right)^6 + \ldots \right] + O \left( \frac{1}{u^7} \right). \]

Thus for large \( u \) and \( \lambda < v \),

\[ \phi(u) \approx \frac{\lambda k}{\sqrt{v^2 - \lambda^2}} \, \{ \sin(\lambda u)/(\lambda u) \}. \]

That is, \( |\phi(u)| \) decreases as \( \sin(z)/z \), as would be expected, since the behaviour of \( \phi(u) \) for large \( u \) is governed by the discontinuities (see e.g. Wiener [107]) and for \( \lambda < v \), the
discontinuities are similar to the case of the rectangular distribution. Thus similar remarks to those for the rectangular distribution apply as regards Lemmas 2 and 3.

Assume now that there is say 10% of independent Gaussian noise added to the sinusoid - that is, the random phase sinewave process and noise with p.d.f.

\[ f_y(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/(2\sigma^2)} \quad \text{where } \sigma = 0.1v/\sqrt{2} \]

is added together to give a process with p.d.f. \( f_z(z) \) where \( z = x + y \). Then since

\[ \phi_y(u) = e^{-\sigma^2u^2/2} , \]

the characteristic function of the summed process, \( \phi_z(u) \) will be

\[ \phi_z(u) = J_0(\nu u)e^{-\sigma^2u^2/2} , \]

so that

\[ |\phi_z(u)| = |J_0(\nu u)|e^{-\sigma^2u^2/2} \leq e^{-\sigma^2u^2/2} \]

reducing it to the Gaussian case.

Also, since

\[ f_z(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_z(u)e^{-jux}du , \]

\[ |f_z'(x)| = \left| \frac{d}{dx} \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_z(u)e^{-jux}du \right| \]
\[
\frac{1}{2\pi} \int_{-\infty}^{\infty} |J_0(yu)| |u| e^{-\sigma^2 u^2/2} du \\
\leq \frac{1}{\pi\nu^2},
\]
and,
\[
|f''(x)| \leq \left\{ \frac{1}{2\pi} \right\} \int_{-\infty}^{\infty} u^2 \exp(-\sigma^2 u^2/2) du \\
= \frac{1}{\sigma^2 \sqrt{2\pi}}.
\]
Thus all the conditions of Lemmas 1, 2 or 3 are effectively satisfied.

3.54 Rayleigh Distribution \cite{68,78}

\[
f_r(x) = \begin{cases} 
\frac{x}{\sigma^2} e^{-x^2/2\sigma^2} & x \geq 0 \\
0 & x < 0
\end{cases}
\]

\[
f'_r(x) = \left( \frac{1}{\sigma^2} - \frac{x^2}{\sigma^4} \right) e^{-x^2/2\sigma^2}
\]

\[
f''_r(x) = \left( \frac{x^3}{\sigma^6} - \frac{3x}{\sigma^4} \right) e^{-x^2/2\sigma^2}
\]

\[
f'''_r(x) = \left( \frac{5x^2}{\sigma^6} - \frac{x^4}{\sigma^8} - \frac{3x}{\sigma^4} \right) e^{-x^2/2\sigma^2}
\]

Max|\(f(x)\)| $\leq$ 0.6065/$\sigma$

Max|\(f'(x)\)| $\leq$ 1/$\sigma^2$

Max|\(f''(x)\)| $\leq$ 1.5/$\sigma^3$

\[
\psi(u) = \int_{0}^{\infty} \frac{x}{\sigma^2} e^{-x^2/2\sigma^2} e^{jux} dx
\]
\[ \begin{align*}
&= ju \int_0^\infty e^{-x^2/2\sigma^2} e^{jux} dx - \int_0^\infty (ju - \frac{x}{\sigma^2}) e^{jux-x^2/2\sigma^2} dx \\
&= I_1 + I_2 + 1
\end{align*} \]

where
\[ I_1 = ju \int_0^\infty \cos(ux)e^{-x^2/2\sigma^2} dx \]
and
\[ I_2 = -u \int_0^\infty \sin(ux)e^{-x^2/2\sigma^2} dx \]

From [1]
\[ I_1 = ju\sqrt{\frac{\pi}{2}} e^{-\sigma^2 u^2/2} \]
and
\[ I_2 = -u\sqrt{2}\sigma D(u\sigma/\sqrt{2}) \]
where \( D(v) = e^{-v^2} \int_0^v e^x dx \)

is the Dawson Integral.

Thus, with \( v = (u\sqrt{2}/\sigma) \),
\[ \phi(u\sqrt{2}/\sigma) = \{1 - 2vD(v)\} + j\sqrt{\pi}v e^{-v^2} \]

Note that for \( A > 1 \)
\[ I_L(A) = \int_A^L |\{2ve^{-v^2} \int_0^v (e^{x^2} - \frac{1}{v}) dx\}| dv \]
\[ = \int_A^L \{2ve^{-v^2} \int_0^v e^{x^2} dx - 1\} dv \]
\[ = I_a + I_b - (L - A) \]

where
\[ I_a = \int_A^L 2ve^{-v^2} \int_0^A e^{x^2} dx dv \]
\[ = D(A)\{1 - e^{A^2-L^2}\} \]
and
\[ I_d = \int_{A}^{L} e^{-x^2} \int_{A}^{V} e^{x^2} \, dx \, dv \]
\[ = \int_{A}^{L} e^{x^2} \int_{x}^{L} 2ve^{-v^2} \, dv \]
\[ = (L - A) - e^{-L^2} \int_{A}^{L} e^{x^2} \, dx \]
\[ = (L - A) - D(L) + D(A)e^{A^2-L^2}. \]

Thus \( I_L(A) \rightarrow D(A) \) as \( L \rightarrow \infty \), and hence
\[ \int_{A}^{\infty} |\phi(u)| \, du \leq \frac{\sqrt{2}}{\sigma} \left\{ D\left(\frac{\theta A}{\sqrt{2}}\right) + \frac{\sqrt{\pi}}{\epsilon} e^{-\sigma^2 A^2/2}\right\}. \]

Values of \( K \) are again evaluated for
\[ \epsilon = 0.05 \ \max |f_n(x)| \leq 0.0303/\sigma. \]

3.541 Method of Lemma 1

\[ B \leq 3, \]
\[ K \geq \frac{DB}{\epsilon} \leq 100. \]

3.542 Method of Lemma 2

\[ \int_{A}^{\infty} |\phi(u)| \, du \leq \frac{\sqrt{2}}{\sigma} D\left(\frac{\theta A}{\sqrt{2}}\right) \]
\[ \leq \frac{1}{\sigma^2 A} \ \text{for large} \ A \]

So that
\[ \frac{AB}{\sigma^2} \leq \frac{3}{\sigma \epsilon A^2} \]
\[ \frac{B^2E}{24e_b} = \frac{4.5}{80e_b} \]

with \( e_a = 0.02/\sigma \) and \( e_b = 0.0103/\sigma \),

\[ K \approx 50. \]

3.543 Method of Lemma 3

\[ K \geq \frac{AB}{\pi} \approx 30. \]
Ze = Zero line of error.
Zo = Zero line of output.

Error as % of \( f_g(0) \).

10 Channels; Max. error \( \approx 0.35 \) percent.

40 Channels; Max. error \( \approx 0.03 \) percent.

**FIGure 3.3. GAUSSIAN DISTRIBUTION, \( f_{g_x} \)**
Ze = Zero line of error.
Zo = Zero line of output. 

Input Uniform over \([11, 211]\). 

\[ e_1 \leq 50 \text{ percent error, (Max.)} \]
\[ e_2 \leq 11 \text{ percent error.} \]

10 Channels.

\[ e_1 \leq 50 \text{ percent error, (Max.; outside } [11, 211]) \]
\[ e \leq 11 \text{ percent error, (Max. in } [11, 211]) \]

40 Channels.

FIG. 3.4. UNIFORM DISTRIBUTION, \( f_u \).
$e_1, e_2, e_3, e_4, e_5, e_6, e_7$ are 98, 80, 27, 10, 2, 3, 5 percent errors expressed as percent of actual value.

10 Channels.

$Ze = \text{Zero line of error.}$
$Zo = \text{Zero line of output.}$

Input asymptotes at 11 and 211.

$e_1, e_2, e_3, e_4, e_5, e_6, e_7$ are 960, 350, 100, 11, 11, 8, 2 percent errors expressed as percent of actual value.

40 Channels.

**Fig. 3.5. Distribution of random phase sine wave, $f_s$.**
e₁ = 30, e₂ = 3, and e₃ = 1 percent of (fₐ Max.).

10 Channels.

Ze = Zero line of error.
Zo = Zero line of output.
Input: fₐ(0) is at 11.

e₁ = 4, e₂ = 5 (Max.), e₃ = 1.5 and e₄ = 3 percent of (fₐ Max.).

40 Channels.

FIG. 3.6. RAYLEIGH DISTRIBUTION, fₐ.
CHAPTER 4
CORRELATION AND POWER SPECTRUM

4.1 Review

The importance of power spectral density determination was first shown by Wiener [106]. Spectral analysis has since been found to be an invaluable tool in signal processing although there are many problems associated with the practical determination of power spectra. Many of these were studied in detail by Blackman and Tukey [13]. Some of the early work on the subject includes publications by Tukey [95], Phillips [52], and Solodonikov [86].

A number of books [8,32,62,63] on random signal analysis were published about 1955-1960 along with some more rigorous mathematical texts [5,35,43,75]. Some recent publications on the more practical problems of power spectral measurements (as well as on data handling in general) are the books by Blackman [12], Bendat and Piersol [9], Korn [60], and the papers [40,41,54,58,79,94,110].

Practically all of the above approaches (not counting analogue schemes) considered obtaining the spectrum as the Fourier transform of the correlation function. Simulation of analogue methods have also been attempted [44,99]. Other methods include the complex demodulation method and
more recently, with the aid of the fast Fourier transform, spectral averaging of sections of the signal ([11, 18, 25, 27, 28, 100] particularly [11]).

4.2 Comparison of Some Methods of Determining Power Spectra

In this section a brief comparison is made of some of the methods of determining power spectra.

(a) Periodogram (Schuster) Analysis.

In the Periodogram method the spectral components $S_M(w)$ are estimated as

$$S_M(w) = \frac{1}{M} \sum_{n=0}^{M} x(t_n) e^{jwt_n}$$

For random signals, however, this is quite unsatisfactory as the variance of the estimates can be extremely large and does not decrease to zero with the number of samples $M$ [6, 53].

i.e. \( \lim_{M \to \infty} \text{Var} [S_M(w)] = S(w) \)

where $S(w)$ is the true value of the spectrum at frequency $w$.

(b) Fourier Transform of Correlation Function.

For a long time, the spectrum has been determined as the Fourier transform of the correlation function as proved by
the Wiener-Kintchine relationship [63]. Bartlet [6] suggested averaging a number of periodograms of segments of the series to get a consistent estimate, and this in fact works out to be the same as determining the spectrum from a weighted correlation function. Parzen [73] showed how consistent estimates could be obtained by taking a "spectral average" of the calculated spectrum.

(c) Filter Simulation (simulation of the analogue process).

Digital simulation of the analogue methods would seem an "obvious" method of determining the power spectrum. However, the digital simulations are tedious and have become practical only since the advent of high speed computers. Even then, the correlation function method is usually advantageous. When it is desired to have the frequency resolution a function of the frequency or when specific frequencies only are required the simulation methods have an advantage [44,99].

(d) Complex Demodulation.

In the complex demodulation method using a digital computer, the signal is multiplied by $e^{-j\omega nt}$ and passed through a simulated low pass filter. Multiplying by $e^{-j\omega nt}$ shifts the spectrum so that the frequency $\omega_n$ is at the origin and passing through a low pass filter measures the power around the frequency origin. This method has
the advantage that a narrow band, low-pass filter is easier to simulate than a moving narrow band, band-pass filter.

(e) Fast Fourier Transform.

The Cooley Tukey algorithm [29] has enabled Fourier transforms to be calculated at considerably higher speeds than before. Thus it is feasible, and in fact quicker [100], to use the "original" method of averaging periodograms (modified to correct for "end effects"). It is also possible to get the correlation function by transforming the resultant spectrum.

The correlation method, however, gives a "preview" of the spectrum and is useful as an aid in determining if a meaningful spectrum can be obtained. The direct method may give erroneous results. The correlation method is not that much slower, and the programming is also simpler resulting in smaller storage requirements and programming ease.

4.3 Specific Problems in Determining Spectral Distributions From Correlation Functions

A signal \( x(t) \) can usually be approximated by the sum of a number of components as follows:

\[
x(t) = y(t) + a_0 + a_\delta \sin \delta t + \sum_{i=1}^{M} a_i \cos(w_i t + \phi_i)
\]
where \( y(t) \) is a function with a smooth spectrum, \( a_0 \) is a "d.c." component, \( a_\delta \) \( \sin \delta t \) represents a function with spectral components low compared to those of interest and the 
\[
\sum_{i=1}^{M} a_i \cos (w_i t + \phi_i)
\]
represents a finite number of periodic components. The terms other than \( y(t) \) would contribute to irregularities in the spectrum of \( x(t) \).

A spectral analysis of \( x(t) \) in the presence of these other terms can give quite meaningless and deceptive results. This and related problems are considered in the next sections.

4.31 Zero and Very Low Frequencies

Zero and very low frequencies (or slow trends) present in data have to be removed before computation or their presence compensated for. The "dc" and "near dc" "spikes" dominate the low frequency region of the spectrum as well as distort the remaining parts of the spectrum owing to the side lobes of any windows used (see section 4.33) \[13,79\].

Removal of the mean is easily achieved, but subtraction of any slow trend is not so straightforward. Generally a correction \( C \) is subtracted from the correlation estimate. Blackman and Tukey \[13\] use

\[
C = \frac{3}{16} \left( 1 - \frac{1}{n^2} - \frac{2r}{n} - \frac{2r^2}{n^2} \right) (\bar{x}^+ - \bar{x}^-)^8
\] \hspace{1cm} 4.4
where \( n \) = number of data points
\[ r = \text{value of lag} \]
\[ \bar{x}^- = \text{mean of the first third of signal} \]
\[ = \frac{3}{n} \sum_{i=1}^{\frac{n}{3}} x_i \]
and \[ \bar{x}^+ = \text{mean of the last third of signal} \]
\[ = \frac{3}{n} \sum_{i=\frac{2n}{3}}^{n} x_i \]

This is based on removing a straight line from the original data points. Other schemes for subtracting least square straight line are also given in [13]. These corrections, for both "dc" and "near dc", are made to the correlation function.

The author found however, that it was preferable to make corrections before computing the correlation function. In theory, the constant value to which the true correlation function tends is the square of the mean value. In the calculated correlation function this is not strictly true, since

\[ \text{FIG. 4.1.} \]
the number of points \((n-r)\) used for calculating the function varies for each lag value \(r\). A mean calculated using a certain number of points \(n_1 = n - r_1\), for lag \(r_1\), can cause a bias for lags \(r + r_1\). A similar observation was made in the case of the slow trend correction. In some cases the correction actually introduced a "slow-trend component" into the signal. In the case of periodic or nearly periodic signals, for instance, the "third means" \(\overline{x}^-\) and \(\overline{x}^+\) may be biased due to incomplete periods in each third (see fig.4.2) so that although the signal does not have a slow trend, the correction formula would apply a "correction" which would now appear as a slow trend.

**FIG. 4.2.**

4.32 Correlation Divisor

The time correlation function is defined as

\[
R(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t) x(t+\tau) dt
\]

For practical numerical computations it is approximated as a finite sum in one or other of the following forms [72]
\begin{align*}
R(m) &= \frac{1}{M-m} \sum_{i=1}^{M-m} x_i x_{i+m} \quad 4.6 \\
R_b(m) &= \frac{1}{M} \sum_{i=1}^{M-m} x_i x_{i+m} \quad 4.7
\end{align*}

Parzen in [72] advocates the use of equation 4.7. However, the author found that in some cases a linear trend resulted, if correction for the mean was carried out after correlation since the cause of error mentioned in the previous section is even more pronounced in this case.

If corrections are done before correlation; the effect is to have the correlation function decay more rapidly thus causing a smoothing of the spectrum (see section 4.4). This is to be preferred.

\(R(m_1)\), for a large lag \(m_1\), will have \(M-m_1\) terms used in its computation and will hence be less reliable than a value \(R(m_2)\) where \(m_2 < m_1\). Consequently less reliance can be placed on the values in the tail. The use of \(M\) for the divisor in 4.7 means that less significance is in fact placed on the values of \(R(m)\) for the larger lags. Thus
4.7 is a better choice all round.

4.33 Windows

The Wiener-Kintchine relationship [63] gives the relationship between the correlation function \( R(\tau) \) and the power spectrum \( S(w) \), and for the discrete case this may be written as

\[
S(w_j) = \frac{1}{\pi} \left[ 1 + 2 \sum_{k=1}^{\infty} R(\tau_k) \cos w_j \tau_k \right]
\]

4.8

In practice only a finite number of values \( R(\tau_k) \) are available, for \( k = 1, 2, \ldots, M \) say. In this case an approximate value \( S_M(w_j) \) is obtained for \( S(w_j) \), given by the relationship [53]

\[
S_M(w_j) = \int_0^{2\pi} S(y) K(w-y) \, dy
\]

where

\[
K(w-y) = \frac{1}{2} \mu(w-y) + \mu(w+y)
\]

4.9

and

\[
\mu(y) = \begin{bmatrix} \sin(M+\frac{3}{2})y \\ \sin \frac{1}{2} y \end{bmatrix}
\]

K(w-y) is generally termed a kernel but it has been given the name of "window" by Blackman and Tukey [23]. The estimate \( S_M(w_j) \) is a "smudged" or "averaged" value of \( S(w) \) about \( w = w_j \).

The variances of the above estimates are large for
finite sums. Thus the spectral estimation problem is concerned with designing "windows" to decrease the variability of the estimates. As might be expected, a decrease in variability is at the cost of a decrease in resolution or an increase in "smudging". Various kernels have been proposed but all of them are based on "cut and try" methods. Basically the problem may be looked upon as being one of maximising the area of the kernel function $\mu(w)$ within the desired resolution (or keeping the sidebands at a minimum) and at the same time maximizing the area of $\eta(\tau)$ (the Fourier transform of $\mu(w)$) over the available length of signal.

The optimum requirements vary depending on the spectrum to be estimated. Thus, if the spectrum is peaked, a greater resolution is required than when a relatively flat spectrum is being estimated.

The author feels that it is better to correct for window error in the time domain before transforming as it
is easier to multiply the correlation function with the necessary window. The process of either "Hanning" or "Hamming" used by Blackman and Tukey [13] in the frequency domain is only an approximation to the window.

4.34 Negative Values of Spectrum

One of the problems of using a digital computer for estimating the spectrum of signals is that, on obtaining the spectrum by means of equation 4.8, one often finds negative values [18]. As shown below, this is related to the problem of missing out the peaks, or failing to obtain the actual peak values present in a spectrum.

One method suggested for overcoming this is prewhitening and/or smoothing [13]. Often this only reduces the effect.

As correlation estimates are time limited, the effect on the spectrum is convolution with a Sin(w)/w "window". If sharp peaks are present in \( S_M(w) \), adjacent points can take on negative values. Thus in fig.4.5, a peak at \( w_2 \) would give a strong negative contribution to the estimation of \( S_M(w_1) \) at \( w_1 \). This effect can be minimized by smoothing which amounts to convolving with a window which
has smaller or no negative side-lobes. It is better to filter out the frequency $w_2$ and thus have a smoother spectrum to start with.

If $w_2$ falls at one of the frequencies being estimated, since all the zeros of the $\text{Si}(w)$ function fall at other $w$ values, no error should be introduced. However the peaks are not sharply defined in practice. The best solution seems to be to filter out any peaks as nearly as possible without distorting the remaining spectrum too severely.

The fast Fourier transform method does not produce this particular effect, as there the spectrum is taken to be the sum of the squares of the Fourier coefficients and is always non-negative.

The presence of negative spectra indicates that little reliance can be placed on these values and that there is the likelihood of a peak somewhere.

4.4 Error Bounds

The author here considers the method of determining the power spectrum from the correlation function and
examines some of the problems mentioned in section 4.3 by consideration of Fourier transform pairs.

The fact that the spectrum and the correlation function are Fourier transform pairs means that certain information is available about one from the other. For example, it is possible to decide from the correlation function if sufficient lags $\tau$ have been taken to accurately determine the power spectrum.

Let $S(w)$ be "smooth", i.e. let the $n$th derivative $S^n(w)$ of $S(w)$ exist and be integrable. Then (c.f. Cramer [30], and section 3.3),

$$R(\tau) \leq \frac{K}{|\tau|^n}$$

where $K$ is a constant.

Conversely, if $R(\tau)$ is smooth, then

$$|S(w)| \leq \frac{K'}{|w|^n}$$

That is, the smoothness of the spectrum $S(w)$ is reflected in the behaviour of the correlation function $R(\tau)$ for large values of $\tau$ and conversely, the smoothness of the correlation function is indicative of the behaviour of the spectrum for large values of $w$.

Let $|R(\tau)|$ be small for $\tau \geq \tau_0$

i.e. $S(w) \in B_\varepsilon[-\tau_0, \tau_0]$
Then $S(w)$ is smooth and the "sampling" requirement (see Chapter 5) is $\Delta w = \frac{\pi}{T_0}$ which would be the case if all of $R(\tau)$ is transformed.

Some simple error bounds may be developed for the determination of the spectrum from the correlation function. From 4.8, assuming $(\tau_{k+1} - \tau_k) = h$, a constant

$$
S(w_j) = \frac{1}{\pi} \left[ 1 + 2 \sum_{k=1}^{M} R(\tau_k) \cos(w_j k) \right] + \frac{2}{\pi} \sum_{M+1}^{\infty} R(\tau_k) \cos(w_j k)
$$

$$
= S_M(w_j) + \frac{2}{\pi} \sum_{M+1}^{\infty} R(\tau_k) \cos(w_j k)
$$

$$
e_S(w_j) = S(w_j) - S_M(w_j) = \frac{2}{\pi} \sum_{M+1}^{\infty} R(\tau_k) \cos(w_j k)
$$

$$
|e_S(w_j)| \leq \frac{2}{\pi} \sum_{M+1}^{\infty} |R(\tau_k)|
$$

This is the error in determining the points
\( S(w_j), j = 0,1, \ldots M \) but not the error in \( S(w) \) at \( w = w_0 \neq w_j, j = 0,1, \ldots M \). For example, if the zeros of \( R(\tau_k) \) lie at the points \( \tau_k \) then \( e_S(w_j) = 0 \) by 4.10, but \( e_S(w) \neq 0 \) necessarily for \( w \neq w_j, j = 0,1, \ldots M \) (See 4.11).

Values of \( S(w_\theta) \) may be obtained from the set of values 
\[
\{S(w_j), j = 0,1, \ldots M\}
\]
by interpolation. If the points \( S(w_j) \) were error free or had negligible error, an error bound could be obtained using the result 5.11
\[
e_{\text{finite}}(w) \leq \frac{2}{\pi} \int_{T_e}^{\infty} |R(\tau)| d\tau
\]
where \( T_e = Mn \).

The actual errors may be greater since the estimates \( S(w_j) \) will not be error free when \( R(\tau_k) \neq 0, \tau_k > T \), except for special cases as in the example above. Further, if \( R(\tau) \in \mathbb{B}[-W,W] \), then it is necessary that \( h \leq \frac{W}{W} \) to prevent aliasing errors.

Presence of "peaks" or nearly periodic components means that \( R(\tau) \) does not tend to zero or only tends to zero very slowly as \( \tau \) increases. According to the above expression, the error in determining \( S(w) \) will be unbounded or large and hence, it is necessary to remove periodic components from the data. It is seen also that weighting down the tail end
of $R(\tau)$ corresponds to smoothing the spectrum. On the other hand, if it is known that the spectrum exists only for a small interval, the true correlation function must be smooth. Large irregularities would indicate the presence of errors and it may be advantageous to smooth the data or the correlation function.
CHAPTER 5
INTERPOLATION

5.0 Introduction

In numerical work, a continuous function can only be represented by a finite sequence of numbers. These numbers will normally be chosen so that they "represent" the function in the sense that any desired values of the function can be obtained from them. Broadly speaking, the process of obtaining these desired values is termed interpolation.

5.1 Theory and Review

Interpolation is thus concerned with finding a function \( f \in \mathcal{F} \), to agree on a given set of points \( \{w_i\} \), in the sense that they satisfy some given conditions \( L_{i,j}(f) \). It is possible, for example, to fit a unique polynomial of degree \( n \) given \( (n+1) \) distinct points (e.g. see [33,91,103]). In a sense, it is the converse of the sampling process referred to in section 2.23.

In general if \( X \) is a linear function space of dimension \( n \), and \( L_i, (i = 1,2,...,n) \) are \( n \) elements of \( X^* \), then the general problem of finite interpolation may be stated as

\[
L_i(X) = w_i, \ i = 1,2,...,n \ \text{or} \ i \in [1,n]
\]

5.1

where \( w_i \) are a given set of arbitrary elements, the \( L_i \) are
linear functionals defined on \( X \) and \( X^* \) is the algebraic conjugate space of \( X \), (i.e. the linear space formed by the set of linear functionals \( L_i \) defined on \( X \), along with appropriate rules for sums and scalar products). In this chapter, \( i \in [j,k] \) will be taken to mean \( i = j,j+1,\ldots,k-1,k \), when \( [j,k] \) is an indexing set.

The general problem has a solution if the \( L_i \) form an independent set in \( X^* \).

\[
\text{i.e. } \sum_{i=1}^{n} \alpha_i L_i = 0 \Rightarrow \alpha_i = 0, \ i \in [1,n]
\]

If \( x_i, i \in [1,n] \), are independent in \( X \), then the condition for solution may be expressed as the non-vanishing of the "generalized Gram's determinant" \( |L_i(x_j)| \)

\[
\text{i.e. } \ |L_i(x_j)| \neq 0
\]

Some common interpolation problems (see e.g. [33,48,91,102]) are:

(a) Polynomial interpolation at discrete points

Here \( X = P_n \), the class of all polynomials of degree \( \leq n \).

\[
L_i(f) = f(z_i), \ i \in [0,n] \quad \text{where } f \text{ is the function that is to be interpolated at the } (n+1) \text{ points } z_0, z_1, \ldots, z_n.
\]
(b) General Hermite Polynomial Interpolation

For this $X = P_N$ and

$$L_{m_0+i}(f) = f^i(z_0) \quad i = 0, 1, \ldots M, \quad m_0 = 1$$

$$L_{m_1+i}(f) = f^i(z_1) \quad i = 0, 1, \ldots M_1, \quad m_1 = M_0 + m_0 + 1$$

$$\vdots$$

$$L_{m_n+i}(f) = f^i(z_n) \quad i = 0, 1, \ldots M_n, \quad m_n = M_{n-1} + m_{n-1} + 1$$

for distinct $z_i$ and where $N = M_0 + M_1 + \ldots M_n + n$.

The $M_i$ are integers $\geq 0$. (Note there are $N+1$ "conditions") and $f^i(z_i)$ is the $i$th differential at $z_i$.

(c) General Taylor Interpolation

Here $X$ consists of $n+1$ linearly independent functions,

$$\phi_i(z) \quad i \in [0, n] \quad \text{analytic at } z_0$$

and $L_1(f) = f^1(z_0) \quad i \in [0, n]$

and

(d) Interpolation by Fourier Series

For this $X = T_n$, the trigonometric polynomials of degree $n$.

$$L_{2k}(f) = \int_{-\pi}^{\pi} f(x)\cos(kx)dx, \quad k \in [0, n]$$

$$L_{2k-1}(f) = \int_{-\pi}^{\pi} f(x)\sin(kx)dx, \quad k \in [1, n]$$
The classical interpolation formula is that due to
Lagrange. This may be generalized as shown below. In
particular, it is shown how two common interpolation formulae
may be obtained from the generalised form.

In terms of Lagrange polynomials, the polynomial \( p_n(z) \)
which passes through \((n+1)\) distinct points \( w_i, i \in [0,n] \) is
given by

\[
p_n(z) = \sum_{k=0}^{n} w_k l_k(z)
\]

where \( l_k(z) \) is the Lagrange polynomial of order \( n \) given by

\[
l_k(z) = \frac{w(z)}{(z-z_k)w'(z_k)}, w(z) = \prod_{i=0}^{n} (z-z_i)
\]

and \( w'(z) = \frac{dw(z)}{dz} \).

An expression for the general Hermite interpolation
problem may be obtained. In the particular case of the value
of the function and its first derivative at each point,
(osculatory Hermite) with \( w(z) \) and \( l_k(z) \) as before
\[ P_{2n-1}(z) = \sum_{k=1}^{n} f(z_k) \left[ 1 - \frac{w''(z_k)}{w'(z_k)} (z-z_k) \right] l_k^2(z) \]

\[ + \sum_{k=1}^{n} f'(z_k) (z-z_k) l_k^2(z) \]

It will be noted that the Lagrange interpolation function has zeros at the given points \( z_i \). In general when there are \( M_i \) values at each \( z_i \), the interpolation function needs to have \( M_i \) zeros at \( z_i \). The interpolation formulae can be generalized in most cases to fit an entire function to an infinite set of points. In particular, Guichard's theorem [33] states that, given a sequence of distinct numbers \( z_i \), \( i \in [0, \infty] \) such that

\[ \lim_{n \to \infty} z_n = \infty \text{ and } \{w_i\}, \ i \in [0, \infty] \]

is an arbitrary sequence of values, then there is an entire function \( f(z) \) such that \( f(z_i) = w_i \), \( i \in [0, \infty] \).

By the Weierstrass product theorem [33,65], it is always possible to construct an entire function with the necessary number of zeros at the sequence of points \( z_i \). Also it is possible (Mittag-Leffler's Partial Fraction Theorem [33,65]) to construct a meromorphic function with principle part at \( z_i \) being

...
\[ \sum_{j=1}^{n_1} \frac{a_{i,j}}{(z-z_i)^j} \] where \( a_{i,j} \) are arbitrary.

If now a function
\[ f_1(z) = g(z)h(z) \]
is constructed, so that \( g(z) \) has simple zeros at the points \( z_i \) and \( h(z) \) has simple poles with principle part at \( z_i \) given by
\[ \frac{w_i}{g'(z_i)(z-z_i)} \]
then it is seen that \( f_1(z) \) is entire and \( f_1(z_i) = w_i \).

Similarly, let \( g(z) \) have 2 zeros at each \( z_i \) (i.e. \( g(z_i) = g'(z_i) = 0 \) but \( g''(z_i) \neq 0 \)) and \( h(z) \) have 2 poles with principal part at \( z_i \) given by
\[ \frac{2f'(z_i)-kf(z_i)}{g''(z_i)(z-z_i)} + \frac{2f(z_i)}{g''(z_i)(z-z_i)^2} \]
where \( k = \frac{g'''(z_i)}{3g''(z_i)} \)

Then \( f_2(z) \), given by \( f_2(z) = g(z)h(z) \), is an entire function.

Expanding \( g(z) \) in a Taylor series about \( z_i \), with
\[ a_i = f'(z_i) - kf(z_i), \]
\[ f_2(z) = \left[ g(z_1) + g'(z_1)(z-z_1) + \frac{g''(z_1)}{2}(z-z_1)^2 + \text{terms in} \right. \\
\left. (z-z_1)^3 \text{ etc} \right] \times \left[ \frac{2a_i}{g''(z_1)(z-z_1)} + \frac{2f(z_i)}{g''(z_1)(z-z_1)^2} + r(z) \right] \]

where \( r(z) \) is analytic at \( z_1 \).

Thus \( f_2(z) = f(z_1) + (z-z_1) \) [terms without poles at \( z_1 \)].

so that \( f_2(z) = f(z_1) \) at \( z = z_1 \).

Similarly \( f_1'(z) = f'(z_1) \) at \( z = z_1 \).

With \( g(z) = \sin z \), the expression for \( h(z) \) becomes

\[ \sum_{i=0}^{\infty} \frac{w_i}{(z-z_1) \cos z_1} \text{ and thus } f_1(z) = \sin z \sum_{i=0}^{\infty} \frac{w_i}{(z-z_1) \cos z_1} \]

where \( z_1 \) are the zeros of \( \sin z \). Let \( z = 2\pi Wt \) and since \( z_1 = \pm n\pi \),

\[ t_1 = \pm \frac{n}{2W} \]

Let \( w_i = f(t_1) \). Then

\[ f_1(t) = \sum_{n=0}^{\infty} f \left( \frac{\pm n}{2W} \right) \frac{\sin(2\pi Wt \pm n\pi)}{(2\pi Wt - n\pi)} \]
\[ f_1(t) = \sum_{n=-\infty}^{\infty} f\left(\frac{n}{2W}\right) \frac{\sin(2\pi W t - n\pi)}{(2\pi W t - n\pi)} \]

This is the "sampling theorem" commonly attributed to Shannon [83,84]. It was first dealt with by E.T. Whittaker [101] and later by J.M. Whittaker [104].

Using \( \sin^2 z \) for the case of \( g(z) \) with a double zero,

\[
\begin{align*}
g(z) &= \sin^2 z \\
g'(z) &= 2\sin z \cos z = 0 \text{ at } z = z_1, \text{ the zeros of } g(z) \\
g''(z) &= 2\cos^2 z - 2\sin^2 z = 2\cos^2 z_1 \text{ at } z = z_1 \\
g'''(z) &= -8\cos z \sin z = 0 \text{ at } z = z_1
\end{align*}
\]

Thus

\[
h(z) = \sum_{n=0}^{\infty} \left( \frac{f(z_1)}{(z-z_1)^2 \cos^2 z_1} + \frac{f'(z_1)}{(z-z_1) \cos^2 z_1} \right)
\]

giving

\[
\begin{align*}
f_2(z) &= \sum_{n=0}^{\infty} \frac{\sin^2 z}{(z-z_1)^2 \cos^2 z_1} \left\{ f(z_1) + (z-z_1)f'(z_1) \right\} \\
&= \sum_{n=0}^{\infty} \frac{\sin(z-z_1)^2}{(z-z_1)^2} \left\{ f(z_1) + (z-z_1)f'(z_1) \right\}
\end{align*}
\]

which yields the "derivative sampling theorem" which was first derived by Jagerman and Fogel, [51] in a less straight
It is seen that the method can be extended for sampling of derivatives of any order by similar techniques.

5.2 Errors

A number of errors can result in an interpolation scheme, particularly if the term "interpolation scheme" is taken to include the whole process. This may include obtaining samples from a source, possibly contaminated with noise, by means of a practical time sampling device. Subsequent interpolation, using these sampled values may be achieved by means of a piece of hardware or a numerical formula. The sources of error may be listed as follows:

1. Noise mixed with the signal.
2. Quantization.
3. Time sampling.
4. Imperfections of sampling mechanisms.
5. Imperfections in storage or transmission of the digital values.
6. Use of approximate interpolation formulae.
7. Use of finite sums in infinite series type interpolation formulae.

1. Often the available signal is mixed with "noise". This contamination may be due to the characteristics of the
channel the signals have to pass through, (e.g. magnetic tape in the case of recorded signals) or due to some operations to which the signals have been subjected. Sometimes the "values" may be information computed from signal values, e.g. determination of the shape of the p.d.f. envelope as in Chapter 3.

The way to correct for this is to filter out the noise components. This assumes a knowledge of some particulars of the signal and noise statistics (see e.g. [17,87]). In general it is not possible to completely separate the noise from the signal but varying degrees of improvement may be possible.

2. In most cases when interpolation is considered, the available data points have been subjected to quantization errors (e.g. amplitude quantization in analogue to digital conversion systems, round-off errors in finite-word length digital systems).

Commonly, the quantization errors are treated as uniformly distributed random noise with zero mean and maximum amplitude $= \frac{q}{2}$ where $q$ is the quantization interval (see e.g. [10,55,89,105]).

3. Interpolation is the problem of selecting a function $\hat{f} \in \mathcal{F}$ in such a way as to satisfy a given set of conditions. Thus $\hat{f}$ approximates a function $f \in \mathcal{F}_1$, from which the set of
conditions $L_1(f)$ were obtained. For example, in first order derivative sampling, the value of the function $f(z)$ at a series of points $z_i$, $i \in [0,\infty]$ along with the value of the derivative of the function $f'(z_i)$ are read. $f(z)$ is approximated by a function $\tilde{f}(z)$ so that

$$\tilde{f}(z) = f(z)$$

and $\tilde{f}'(z) = f'(z)$ at $z = z_i$. It is not necessary that $\tilde{f}_1 = \tilde{f}$ and $f$ need only agree as regards the conditions $L_1$. It is usually of interest to know how well the $\tilde{f}$ approximates $f$, i.e. the behaviour of the error

$$e(t) = |\tilde{f}(t) - f(t)|$$

Indeed it may be desirable to forgo the requirement $L_1(\tilde{f}) = L_1(f)$ if by doing so, an $\tilde{f}$ can be found to give lower errors overall.

The case $f \in \mathcal{B}$ or $f \in \mathcal{B}_e$, are often the signals of interest in electrical systems. In such cases information is necessary at at least twice the rate of the highest frequency content in the signal $f$ to uniquely determine $f$. This is true for low pass signals, but for band-pass signals a modification to the statement is usually necessary (see e.g. [19, 59, 68]). When there are frequencies above the Nyquist frequency $f_s = 2/T_s$, where $T_s$ is the effective
sampling interval, erroneous reconstruction results. The spectrum of the sampled signal is band limited to 
\( \left( -\frac{f_s}{2}, \frac{f_s}{2} \right) \) and the energy of the signal at higher frequencies is "folded-back" and added to the spectrum of the sampled signals. This is known as aliasing.

Error due to this source has been considered by various authors (see the review article [89], expressions later derived in essentially the same form in [70,88], and also, in more rigorous form, [19] and [97]). An upper bound \( \epsilon_u \) for the error in assuming \( f \in B[-W,W] \) is given by

\[
\epsilon_u = \frac{1}{\pi} \int_{|w| \geq W} |\phi(w)| \, dw \tag{5.11}
\]

where \( \phi(w) \) is the Fourier transform of the signal.

The use of prefilters to reduce aliasing has been considered by various authors (see e.g. [87,89]).

4. Imperfections of sampling mechanisms cause errors in the values \( f(nT) \). The case of jitter has been investigated by Shapiro and Silverman [85] (in connection with alias free sampling), Blakrishnan [3], Brown [20] and also by Papoulis [70] (see also [89]). In some instances samples are lost and this case is treated in [70].

5. If n binary bits represent the digital values of the signal, a change in any of the bits may occur giving rise to errors and this case is examined in [70].
6. Using the cardinal series (equation 5.8) corresponds to using an ideal low pass filter, which is difficult to approximate physically. The cardinal series converges to the signal \( f(t) \) if band limited requirements are satisfied. However other interpolation or approximation methods can be adopted, but the sum

\[
 f_N(t) = \sum_{n=-N}^{N} f(nT)K(t,nT)
\]

need not necessarily converge to \( f(t) \). The kernel \( K(t,nT) \) may be chosen because of its ease of implementation or because \( |f_N(t) - f(t)| \) is smaller for a given \( N \) than the cardinal series case [47]. Various kernels have been considered (e.g. [47,98]).

7. In practice, it is only possible to sum a finite number of terms in any formula involving an infinite series as in the case of equations 5.8 and 5.9. Thus a "truncation" error results. This has been treated by various authors (e.g. [33,47,70,89,93,111]. See also section 5.3).

In general, all or a number of the possible sources of errors may be present. Very little work has been carried out regarding the combined effects. Bennet [10] has examined the spectra of quantized and sampled signals. Also
in [55] and [42] the errors in reconstruction due to combined effects of quantization and (time) sampling are considered. Tufts and Johnson [93] have considered recovering random waveforms from a finite number of samples, where the samples are assumed contaminated with noise.

5.3 Truncation Errors

In practice, only a finite number of terms \( N = N_1 + N_2 \) can be used in any formula giving \( f(z) \) as the sum of an infinite series. Thus instead of

\[
f(z) = \sum_{n=-\infty}^{\infty} f(nT)K(n,z)
\]

as, for example, in equation 5.8 and 5.9, a finite sum

\[
f_N(z) = \sum_{n=-N_1}^{N_2} f(nT)K(n,z)
\]

is formed giving rise to an error

\[
e(t) = f(z) - f_N(z)
\]
called the truncation error.

5.31 Results by Various Authors

Most of the work [47, 89, 111] in this connection has been to determine the errors by evaluation of a contour integral [65, 77, 90].
The sampling theorem can be expressed as a product of an analytic function with zeros at $z_i$ and a meromorphic function. This suggests the use of the residue theorem (see also Jagerman and Fogel [37]). As shown earlier (equation 5.5):

$$f(z) = g(z)h(z)$$

where $g(z)$ has the appropriate number of zeros at $z_i$ and $h(z)$ has the same number of poles. Writing

$$f(z) = g(z) = \sum_i \left\{ \frac{a_{i,1}}{z-z_i} + \frac{a_{i,2}}{(z-z_i)^2} + \cdots + \frac{a_{i,k_i}}{(z-z_i)^{k_i}} \right\}$$

suggests writing

$$e_N(z) = \frac{1}{2\pi i} \oint_{C_N} \frac{f(\xi)/g(\xi)}{(\xi-z)} d\xi$$

where $C_N$ includes $N$ poles. It is shown in [37] for 2 cases that $e_N \to 0$ as $N \to \infty$ and the residues give the sampling theorem.

Further, as in the case of $g(z) = \text{Sin}(z)$,

$$e(t') = f(t') - \sum_{n=-N_1}^{N_2} f(nT)\text{Sin}(t'-nT) \quad 5.12$$

and since $f(\xi)/g(\xi)$ is analytic within a contour such as $C'$.
\[ f(z) = \int_{C} \frac{f(\xi)}{g(\xi) (\xi - z)} \, d\xi \]

which is seen to be expressible as

\[ \int_{C''} \frac{f(\xi)}{(\xi - z)} - \sum_{n=-N_1}^{N_2} \frac{(-1)^n f(nT)}{2\pi i (nT - z)} \]

so that the error is given by

\[ e(t') = \frac{E(t')}{2\pi i} \int_{C''} \frac{f(\xi)}{(\xi - t')g(\xi)} \, d\xi \]

This expression for the error for polynomial interpolation is shown in [33]. It is also shown to be correct for particular cases in [47] and [111].

Bounds on the errors can then be evaluated by evaluating the contour integrals under various constraints on \( f(z) \).

Band limited functions are entire functions of order 1 and type \( W \) (Boas [15]) where \( W \) is the cut-off angular frequency of the function. Using properties of entire functions, various bounds for \( f(z) \) can be obtained. In Helms and Thomas [47] and also Yao and Thomas [111] a bound for the error in terms of \( \max |f(z)| = M \) is obtained as
\[ e(t') \leq \frac{2M|\sin(Wt')|}{\pi^2(1-r)} \left( \frac{1}{N_1} + \frac{1}{N_2} \right) = B_1 \]  
5.14

In [47] and

\[ e(t') \leq \frac{M|\sin Wt'|}{2\pi \cos \omega W/2} \left( \frac{1}{N_1} + \frac{1}{N_2} \right) = B_2 \]  
5.15

in [111],

where \( 0 < r < 1 \) is a "guard" band; i.e. \( f(z) \) is assumed band

limited to \( rW \) radians since the bound for \( r \to 1 \) approaches \( \infty \).

In [47] "self truncating" expansions for \( f(t) \) are also

considered given by

\[ f(t) = \Sigma f(nT) [\text{sinc}(\frac{\omega u}{m})]^m \text{sinc} u \]

where \( T = \pi/W \)

\( q = 1 - r \)

and \( u = W(t-nT) \)

and an optimum value for \( m \) in terms of \( r \) is obtained.

In [111], a bound in terms of the energy of the signal

is also obtained as

\[ e(t) \leq \frac{2E(\frac{rW}{\pi})^2 |\sin Wt|}{\pi^2(1-r)} \left( \frac{1}{N_1} + \frac{1}{N_2} \right) = B_3 \]  
5.16

where \( E \) is the total energy content of the signal.

Papoulis [70] gives a result (communicated to him by

Jagerman)
It is seen that in all the results the error becomes unbounded when \( t' \) approaches the end point. Also in the first three cases, the bound \( \rightarrow \infty \) as \( r \rightarrow 1 \).

Here the author obtains realistic bounds for these cases by a method somewhat similar to that in [70]. Further, the bound \( B_3 \) is shown to be more easily derived and in fact a lower bound is obtained for this case.

Let \( f(t) \) be band limited to \( W \) radians/second. Then

\[
f(t) = \sum_{n=-\infty}^{\infty} f(nT) \frac{\sin(W(t-nT))}{W(t-nT)}, \quad T = \frac{W}{W}
\]

If the sum is formed from \( N_L \) terms to the left of \( t = 0 \) and \( N_R \) terms to the right,

\[
f_N(t) = \sum_{n=-N_L}^{N_R} f(nT) \sin(W(t-nT))
\]
so that the error

\[ e(t) = \sum_{-\infty}^{-(N_L+1)} f(nT) \sin t(nT) + \sum_{N_R+1}^{\infty} f(nT) \sin t(nT) \]

\[ = \frac{\sin Wt}{W} \{S_1 + S_2\} \]

where

\[ S_1 = \sum_{-\infty}^{-(N_L+1)} \frac{f(nT)(-1)^n}{t-nT} \]

and

\[ S_2 = \sum_{N_R+1}^{\infty} \frac{f(nT)(-1)^n}{t-nT} \]

Using Schwartz inequality [39] for the sums,

\[ S_1 < \left( \sum_{-\infty}^{-(N_L+1)} \frac{1}{(t-nT)^2} \right)^{\frac{1}{2}} \left( \sum_{-\infty}^{-(N_L+1)} f^2(nT) \right)^{\frac{1}{2}} \]

\[ S_2^2 < \frac{K}{T^2} \sum_{N_R+1}^{\infty} \frac{1}{(m+nT)^2} \]

where

\[ K = \left( \sum_{-\infty}^{-N_L+1} f^2(nT) \right) \]
Consider \( S = \sum_{m=N+1}^{\infty} \frac{1}{(m+k)^2} \). Put \( n = m - N \)

\[
S = \sum_{1}^{\infty} \frac{1}{(n+N+k)^2} \leq \sum_{1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}
\]

(see [49]) for \((N+k) \geq 0\).

Thus an upper bound for \( e(t) \) is

\[
e(t) \leq \left( \frac{\left| \sin \omega t \right|}{\omega} \right) \left( \frac{1}{T} \right) \left[ \sum_{-\infty}^{\infty} f^2(nT) \right]^{\frac{1}{2}} \left[ \sum_{N+1}^{\infty} f^2(nT) \right]^{\frac{1}{2}} \frac{\pi}{\sqrt{6}} = \mathcal{B}_S
\]

for \(-N_LT \leq t \leq N_RT\)

or, for

\[-N_L \leq \frac{t}{T} \leq -N_L +1.\]

\[
S_2 \leq K' \sum_{N+1}^{\infty} \frac{1}{(N_L+n)^2}
\]

\[
\leq K' \sum_{N+1}^{\infty} \frac{1}{m^2}
\]

\[
\leq \frac{1}{N}
\]

where,

\[
m = n + N_L
\]

\[
N = N_L + N_R
\]
giving, \[ |e(t)| < \frac{\sin W t}{W} \cdot \frac{1}{T} \pi \sqrt{\frac{K}{N} + \frac{K'}{N}} \]
or when, \( N - 1 < \frac{t}{T} \leq N \)

This approach is of course only valid if the sequences \( \{f(nT)\}_{n=1}^{\infty} \) and \( \{f(-nT)\}_{n=1}^{\infty} \in \ell_2 \)

(See e.g. [39]).

If now \( f(t) \) is band limited to \( W \) rad/sec, then \( g(t) = f^2(t) \) has the Fourier transform \( \phi_g(w) = \phi_f(w)\phi_f(w) \) where \( \phi_f(w) \) is the Fourier transform of \( f(t) \) and * denotes convolution. Thus \( f^2(t) \) is band limited to \( 2W \) rad/sec.

Consequently, with \( T' = \frac{\pi}{2W} = \frac{T}{2} \)

\[
g(t) = f^2(t) = \sum_{n=-\infty}^{\infty} f^2(nT') \frac{\sin 2W(t-nT')}{2W(t-nT')} \\
= \sum_{n=-\infty}^{\infty} g(nT') \sin 2W(t-nT')
\]

and \( \phi_g(w) = \sum_{n=-\infty}^{\infty} g(nT')e^{jnT'w} \phi_f(2W) \)
where \( \Phi_R(2w) \) is the F.T. of \( \Sigma_i(2Wt) \).

With \( w = 0, \)

\[
\phi_g(0) = \sum_{-\infty}^{\infty} T'g(nT').
\]

Also

\[
\phi_g(0) = \int_{-\infty}^{\infty} g(t)dt
\]

Thus

\[
E^2 = \int_{-\infty}^{\infty} f^2(t)dt = \sum_{-\infty}^{\infty} T'f^2(nT') = T'\sum_{-\infty}^{\infty} f^2(2nT')
\]

\[
= T'2\sum_{-\infty}^{\infty} f^2(nT)
\]

i.e.

\[
\sum_{-\infty}^{\infty} f^2(nT) < \frac{2}{T}E^2
\]

Thus the upper bound may also be expressed in terms of the energy of the signal, \( E^2 \) giving
\[
e(t) = \left(\frac{\sin Wt}{W}\right) \left(\frac{1}{\sqrt{T}}\right) \left(\frac{2B^2}{T}\right)^{\frac{1}{2}} + \left(\frac{2B^2}{T}\right)^{\frac{1}{2}} \frac{W}{\sqrt{6}}
\]

\[
= \frac{\sin Wt}{W} \left(\frac{1}{\sqrt{T}}\right) (2\sqrt{2}) \frac{B}{\sqrt{6}}
\]

\[
= \left(\frac{1}{31}\right)^{\frac{1}{2}} E \sin Wt = B_s,
\]

5.20

or for the case when \(t\) is not close to the ends,

\[
e(t) = \frac{\sin Wt}{\sqrt{Wt}} \left\{ \left(\frac{1}{N_L T + t}\right)^{\frac{1}{2}} + \left(\frac{1}{N_R T - t}\right)^{\frac{1}{2}} \right\} \frac{\sqrt{2E}}{\sqrt{T}}
\]

\[
= \frac{\sqrt{2E} \sin Wt}{\pi} \left\{ \left(\frac{1}{N_L T + t}\right)^{\frac{1}{2}} + \left(\frac{1}{N_R T - t}\right)^{\frac{1}{2}} \right\} = B_7
\]

5.21

The Fourier transform relationship gives

\[f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_f(w) e^{-j\omega t} dw\]

and using the Schwartz inequality

\[|f(t)| \leq \left\{ \frac{1}{\pi W^2} \left[ \int |\phi_f(w)|^2 dw \right] \left[ \int |e^{-j\omega t}|^2 dw \right] \right\}^{\frac{1}{2}}\]

so that if \(f(t)\) is band limited to \(W\) radius/second and \(\phi_f(w) \in L_2[-W,W]\)

\[M = \text{Max} |f(t)| \leq \left\{ \frac{1}{\pi W^2} \int_{-W}^{W} |\phi_f(w)|^2 dw \right\}^{\frac{1}{2}} \int_{-W}^{W} \frac{W}{\pi W^2} \frac{W}{W} \frac{W}{W} = \left(\frac{W}{\pi W^2}\right)^{\frac{1}{2}} = E \sqrt{\frac{W}{\pi}} \quad (\text{see}[70])\]
Thus a relationship between the maximum \( M \) and energy \( E^2 \) is obtained in terms of the spectrum band. Using this relationship, the bound \( B_3 \) is readily obtained from \( B_1 \).

\[
e(t) \leq \frac{2M\left|\sin W t\right|}{\pi^2 (1-r)} \quad \text{where} \quad K = \left(\frac{1}{N_1} + \frac{1}{N_2}\right)
\]

\[
\leq \frac{2BN(rW/\pi)|\sin W t|}{\pi^2 (1-r)} = B_3 \quad \text{as} \quad f(t) \quad \text{is band limited to} \quad rW.
\]

Indeed a lower bound is obtained in this case using \( B_2 \) and the above relationship

\[
e(t) \leq \frac{M\left|\sin W t\right|}{2\pi \cos(rW/2)^K}
\]

\[
\leq \frac{BN(rW/\pi)|\sin W t|}{2 \cos(rW/2)} = B_3
\]

which is smaller than 3, since

\[
\frac{B_3}{B_3} = \frac{2}{\pi^2 (1-r)} \quad 2\pi \cos\left(\frac{rW}{2}\right) = \frac{4}{\pi} \left(1-(rW/2)^2\right) \ldots
\]

\[
> 1
\]
$B_1$ by Helms and Thomas.

$B_2$ by Yao and Thomas.

$B_3$ by Helms and Thomas.

$B_4$ by Papoulis.

$B_5$ to $B_8$ by the author.
5.33 Comparison of Bounds

The graphs 5.1, 5.2 and 5.3 compare the bounds $B_1$, $B_2$, $B_3$, $B_4$ with the author's bounds $B_5$, $B_6$, $B_7$ and $B_8$.

Graph 1 compares bounds $B_1$ and $B_2$

Graph 2 compares the bound $B_3$ with $B_4$. As stated the bound $B_6$ gives realistic values for $S = t/T$ approaching $-N_L$ or $N_R$, whereas $B_4 \to \infty$

Graph 3 shows that $B_8$ is smaller than $B_3$ for all values of $r$. For $r$ approaching 1, both bounds $\to \infty$. $B_6$ or $B_7$ (provided $S$ is not close to $-N_L$ or $N_R$) give better results when $r \to 1$.

5.4 Interpolation Using Areas

In some cases (e.g. determining p.d.f.) the information available is the area between ordinates. Thus a set of values $\{C_i\}$ are available, where

\[ C_i = \int_{x_i}^{x_{i+1}} f(x) \, dx \]

and it is required to find an interpolation formula for $f(x)$.

There is no known method for doing this and the author
here shows the existence of a polynomial solution to the problem. Further, the author finds an entire function solution to the problem.

5.4 Polynomials

The functional \( L \) is in this case defined by

\[
L_1(f) = \int_{x_1}^{x_{i+1}} f(x)dx.
\]

\( L \) is linear since

\[
L_1(f + g) = \int_{x_1}^{x_{i+1}} (f + g) = \int_{x_1}^{x_{i+1}} f + \int_{x_1}^{x_{i+1}} g = L_1(f) + L_1(g)
\]

and

\[
L_1(\alpha f) = \int_{x_1}^{x_{i+1}} \alpha f = \alpha \int_{x_1}^{x_{i+1}} f = \alpha L_1(f)
\]

Consider \( X = \mathbb{P}_n \). (See section 5.1). Then \((n+1)\) independent elements are \( X_0 = x^0, X_1 = x^1, X_2 = x^2, \ldots X_n = x^n \).

Given \((n+2)\) points \((x_0, x_1, \ldots, x_{n+1})\), the question of interest is whether there is a function \( p(x) \in \mathbb{P}_n \) such that \( L_1[p(x)] = c_1 \).
Thus the generalized Gram's determinant is

\[ |L_1(X_j)| = \left| \frac{x_{i+1}^{j+1} - x_i^{j+1}}{j+1} \right| \]

i.e.

\[
\begin{vmatrix}
  x_1 - x_0 & x_2^2 - x_0^2 & \cdots & x_n^{n+1} - x_0^{n+1} \\
  1 & 2 & \cdots & n+1 \\
  x_2 - x_1 & x_3^2 - x_1^2 & \cdots & \vdots \\
  1 & 2 & \cdots & \vdots \\
  \vdots & \vdots & \ddots & \vdots \\
  x_n^{n+1} - x_n & \cdots & \vdots & 1 \\
  1 & \cdots & \cdots & 1
\end{vmatrix}
\]

\[ = \frac{1}{(n+1)!} \begin{vmatrix}
  x_1 - x_0 & x_2^2 - x_0^2 & \cdots & x_n^{n+1} - x_0^{n+1} \\
  x_2 - x_1 & \cdots & \cdots & \vdots \\
  \vdots & \ddots & \vdots & \vdots \\
  x_n^{n+1} - x_n & \cdots & \cdots & 1
\end{vmatrix}
\]

\[ = \frac{1}{(n+1)!} \frac{D}{D}
\]

Consider D.

If \( x_1 = x_0 \), the first row is all zeros \( \cdots (x_1 - x_0) \) is a factor. Similarly, it is seen that
\[ \prod_{i=1}^{n+1} (x_i - x_{i-1}) \]

is a factor. Further, if \( x_\circ = x_0 \), rows 1 and 2 are the same except for sign. Thus \((x_\circ - x_0)\) is a factor. Similarly, putting \( x_i = x_{i-2} \) makes the \( i \)th row the same as the \((i-1)\)th row except for sign, for \( i \in [2, n+1] \).

\[ \prod_{i=3}^{n+1} (x_i - x_{i-3}) \]

is also a factor, since for any \( i \in [3, n+1] \), putting \( x_i = x_{i-3} \) and adding row \((i-2)\) to row \((i-1)\) makes rows \((i)\) and \((i-1)\) equal but for sign.

In the general case, \((x_i - x_{i-r})\) are seen to be factors for \( r \in [1, n+1] \) and \( i \in [r, n+1] \), since

- step 1, adding row \((i+1-r)\) to row \((i+2-r)\)
- step 2, adding the new row \((i+1-r)\) to row \((i+3-r)\)

etc. to

- step \( r-2 \), adding the new row \((i-2)\) to row \((i-1)\)
- step \( r-1 \), adding the new row \((i-1)\) to row \((i)\)

and putting \((x_i = x_{i-r})\) makes the \( i \)th row zero.

Thus

\[ D = k \prod_{r=1}^{n+1} \prod_{i=r}^{n+1} (x_i - x_{i-r}). \]

Order of \( D = \sum_{r=1}^{n+1} r \), same as that of product above.
Thus $k$ is constant. On comparing coefficients $k = 1$, giving

$$|L_1(X_j)| = \frac{1}{(n+1)^n} \prod_{r=1}^{n+1} \prod_{s=r}^{n+1} (x_s - x_{s-r})$$

\[\therefore \quad |L_1(X_j)| \neq 0 \iff x_i \neq x_j \text{ for } i \neq j\]

where $i, j \in [0, n+1]$

Hence by 5.2, it is possible to find a function $p(z) \in P_n \geq L_1(p) = C_i, i \in [0, n]$ where the $C_i$ are an arbitrary set of given values.

5.42 Entire Function

Further, an entire function $f(z)$ is found such that, for an arbitrary set of values $w_i, i = 0, 1, \ldots$ and distinct points $z_i, i = 0, 1, \ldots$ such that $\lim_{n \to \infty} z_n = \infty$ the function $f(z)$ satisfies the linear functional relationships

$$L_n(f) = \int_{z_n}^{z_{n+1}} f(z)dz = w_n \quad n \in [0, \infty]$$

By Guichard's theorem it is possible to find an entire function $F(z)$ with prescribed values at the $z_i$.

$$d_n = \sum_{i=0}^{n-1} w_i \quad n \in [1, \infty]$$
and let \( d_0 = 0 \)

Now let \( F(z) = G(z)H(z) \) as before with \( G(z) \) having simple zeros at \( z_i \), and \( H(z) \) having simple poles at \( z_i \) with principal part \( \frac{d_i}{G'(z_i)} \frac{1}{(z-z_i)^n} \) at these poles so that \( F(z_i) = d_i \)

The requirement for \( G(z) \) and \( H(z) \) is that

\[
\sum_{i=1}^{\infty} \frac{1}{(z_i)^n}
\]

should converge for some integer \( n \), which is easily met [33,65].

Let \( f(z) = F'(z) = G'(z)H(z) + H'(z)G(z) \)

Then \( f(z) \) is the required function.

1. \( f(z) \) is entire since it is the derivative of an entire function.

2. \( L_i(f) = \int_{z_i}^{z_{i+1}} f(z)dz = F(z_{i+1}) - F(z_i) = w_i \)

Note, assuming \( d_0 = 0 \Rightarrow \int_{-\infty}^{z_0} f(z)dz \) is assumed to be zero, i.e. \( F(z_0) = 0 \).

In particular let

\[
G(z) = z \prod_{n=1}^{\infty} \left[ 1 - \left( \frac{z}{n} \right)^2 \right] = \sin z
\]
and

\[ H(z) = \sum_{n=0}^{\infty} \frac{d_n}{(z-z_n)\cos z_n} \]

since \( z_n = n\pi \) and \( \sum_{1}^{\infty} \frac{1}{(n\pi)^2} = \frac{1}{6} < \infty \), \( \{z_n\} \) satisfies the necessary conditions.

\[ G'(z) = \cos z \]

\[ H'(z) = \sum_{n=0}^{\infty} \frac{-d_n}{(z-z_n)^2\cos(z_n)} \]

which is holomorphic everywhere except at \( z = z_i \) where there are poles of order 2.

Hence

\[ f(z) = \sum_{i=0}^{\infty} \frac{d_i}{(z-z_i)} \left( \cos z - \frac{\sin z}{z-z_i} \right) \frac{1}{\cos z_i} \]

\[ = \sum_{i=0}^{\infty} d_i K(z,i) \]

where

\[ K(z,i) = \frac{\cos z - \sin z}{(z-z_i)\cos z_i} \]

To check that \( f(z) \) is entire, it is only necessary to check at the points \( z_i \).
\begin{align*}
K(z,i) &= \frac{\cos(z-z_1) - \frac{\sin(z-z_1)}{z-z_1}}{u} \\
&= \cosu - \sinu \\
\text{where } u &= z - z_1 \quad \text{and } \sinu = \frac{\sin u}{u} \\
K(z,i) &= \left\{ (1 - \frac{u^2}{2} + \ldots) - (1 - \frac{u^2}{6} + \ldots) \right\} / u \\
&= \left\{ -\frac{u^2}{3} + O(u^4) \right\} / u \\
\therefore \lim_{z \to z_i} K(z,i) &= 0 \\
|K(z,i)| &= \left| \frac{\cosu - \sinu}{u} \right| \leq \frac{1}{|u|} \left( 1 + \frac{1}{|u|} \right) = \frac{1}{|u|} + O\left( \frac{1}{u^2} \right) \\
\text{So that } K(z,i) \text{ decreases as } \frac{1}{|z-z_1|} \text{ and it can be seen that it decays at a similar rate as the } \frac{\sin x}{x} \text{ interpolation function. The rate of decay is important in that it determines the relevance of the distant points and hence affects the errors in considering only a finite sum.} \\
\text{It is easy to show that} \\
f(z) \in B[-\pi,\pi] \\
\text{by making use of the fact that} \\
F(z) \in B[-A,A] \\
\Rightarrow f(z) = F'(z) \in B[-A,A] \\
\text{proved below. If } \phi_P(w), \text{ the Fourier transform of } F(z), \text{ is}
continuous, i.e. \( F(z) \) has no periodic components, then

\[
F(z) = \int_{-\infty}^{\infty} \phi_p(w) e^{-jwz} dw
\]

\[
f(z) = F'(z) = \frac{d}{dz} \int_{-\infty}^{\infty} \phi_p(w) e^{-jwz} dw
\]

\[
= \int_{-\infty}^{\infty} \{-jw\phi_p(w)\} e^{-jwz} dw
\]

\[
\therefore \quad |\phi_p(w)| = |w\phi_p(w)| = 0 \quad \text{for } w \notin [-A, A]
\]

If \( F(z) \) has periodic components but may be written as \( F(z) = F_1(z) + F_2(z) \) where \( F_1(z) \) has a continuous spectrum and \( F_2(z) \) is periodic

then

\[
F_2(z) = \sum_{i=0}^{m} a_i e^{jw_0 iz}
\]

where \( w_0 = \frac{\pi}{T} \)

so that \( f_2(z) = F_2'(z) \in B[-A, A] \)

Hence \( f(z) = (F_1 + F_2)' = F_1' + F_2' \in B [-A, A] \)

Finally it is shown that the sum on the right hand side of 5.24 does converge to \( f(t) \). Consider a function \( f(t) \) such that
\[
\int_{-\infty}^{0} f(t) dt = 0
\]

Then if \( f(t) \in B[-W,W], \)

\[
f(t) = W \sum_{n=1}^{\infty} d_n g(t-nT)
\]

where \( T = \frac{\pi}{W} \)

\[
d_n = \sum_{i=1}^{n} C_i \quad \text{where} \quad C_i = \int_{(i-1)T}^{iT} f(t) dt, \quad i = 1, 2, \ldots n
\]

\[
= \int_{0}^{nT} f(t) dt
\]

\[
g(u) = \frac{\cos(Wu) - \frac{\sin(Wu)}{(Wu)}}{(Wu)}
\]

To prove this, first it will be shown that

a)

\[
\phi(w) = (jw e^{jTw})_P = W \sum_{n=0}^{\infty} e^{jwnT} g(t-nT)
\]

and that
b) 

\[ f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_p(w) e^{j tw} dw \]

\[ \Rightarrow W \sum_{n=1}^{\infty} d_n g(t-nT) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi(w) \frac{\phi_p(w)}{jw} dw \]

where \( \phi_p(w) \) is the periodic extension of \( \phi(w) \) defined on \([-W, W]\) over to the line \((-\infty, \infty)\) with period \(2W\), \( |\phi_p(w)| \) is assumed to be integrable, \( t \in \mathbb{R} \) and \( T < \infty \).

**Proof of (a)**

Let \( a_n \) be the nth coefficient of the Fourier expansion of \( \phi(w) \)

i.e. \( a_n = \frac{1}{2W} \int_{-W}^{W} (jwe^{j tw} e^{-jwnT} dw \]

\[ = \frac{1}{2W} \frac{d}{dt} \int_{-W}^{W} e^{jw(t-nT)} dw \]

\[ = \frac{d}{dt} \frac{\sin(t-nT)W}{(t-nT)W} \]

\[ = Wg(t-nT) \]

.. by Fourier series theory,
\[ \phi(w) = \sum_{n=0}^{\infty} e^{jwnT} g(t-nT) \quad \text{a.e.} \quad \text{(See e.g. [90])} \]

which proves (a).

Proof of (b)

Since \( \lim_{N \to \infty} W \sum_{n=0}^{N} e^{jwnT} g(t-nT) = (jwe^{jwt})_p \)

for any \( \epsilon > 0 \), \( \exists \) an \( N_0 \) \( \Rightarrow \forall \ N > N_0 \)

\[ \left| W \sum_{n=0}^{N} e^{jwnT} g(t-nT) - (jwe^{jwt})_p \right| < \epsilon \]

\[ \therefore \left| W \sum_{n=0}^{N} e^{jwnT} g(t-nT) \right| \]

\[ = \left| W \sum_{n=0}^{N} e^{jwnT} g(t-nT) - (jwe^{jwt})_p + (jwe^{jwt})_p \right| \]

\[ < \epsilon + |jwe^{jwt}| \]

\[ = \epsilon + |w| \]

and, clearly, for finite \( N \), \( W \sum_{n=0}^{N} e^{jwnT} g(t-nT) \) is finite so that
\[ \left| \sum_{n=0}^{N} e^{jwnT} g(t-nT) \frac{\phi_p(w)}{jw} \right| \leq M \left| \frac{\phi_p(w)}{w} \right| < \infty \forall N > 0. \]

Further, by (a)

\[
\lim_{N \to \infty} \sum_{n=0}^{N} e^{jwnT} g(t-nT) \frac{\phi_p(w)}{jw} = \phi(w) \frac{\phi_p(w)}{jw}
\]

\[ \therefore (\text{Lebesgue bounded convergence theorem [45]}) \]

\[
\lim_{N \to \infty} \sum_{n=0}^{N} \left[ \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\phi_p(w)}{jw} e^{jwnT} \right] g(t-nT) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi(w) \frac{\phi_p(w)}{jw}
\]

Note that \( d_n = \int_{0}^{nT} f(t) dt \)

\[
= F(nT) - F(0)
\]

\[
= F(nT) \quad \text{where} \quad F(t) = \int_{-\infty}^{t} f(u) du
\]

so that if \( F(t) = \int_{-\infty}^{\infty} \phi_F(w) e^{jwt} dw \)

then \( f(t) = \int_{-\infty}^{\infty} jw\phi_F(w) e^{jwt} dw \).

That is \( \phi_F(w) = \phi_f(w)/(jw) \)

Thus,

\[
\frac{1}{jw} \int_{-\infty}^{\infty} \phi_f(w) e^{jwnT} = F(nT) = d_n,
\]
and hence,

\[
\lim_{N \to \infty} W \sum_{n=0}^{N} d_n g(t-nT) = \frac{1}{2\pi} \int_{-\infty}^{\infty} (j\omega e^{j\omega t}) \phi_f(\omega) d\omega
\]

If now \( \phi_f(\omega) = 0 \) a.e. for \( |\omega| > W, \omega \in \mathbb{R}, \)
then

\[
\lim_{N \to \infty} W \sum_{n=0}^{N} d_n g(t-nT) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega t} \phi_f(\omega) d\omega
\]

\[= f(t) \text{ a.e.} \]

since \( \phi(\omega) = j\omega e^{j\omega t} \text{ for } \omega \in [-W,W] \)
Thus 5.24 is proved.

Further, the aliasing error can be found for the case when \( f(t) \notin \mathbb{B}[-W,W] \)
Let \( f(t) = f_a(t) + f_b(t) \)
where \( f_a(t) = \int_{|\omega|<W} \phi_f(\omega) e^{j\omega t} d\omega \)
and \( f_b(t) = \int_{|\omega|>W} \phi_f(\omega) e^{j\omega t} d\omega \)
Consider \( W \sum_{n=0}^{\infty} d_n g(t-nT) \)
\[ F(nT) = d_n = \int_{0}^{nT} \{ f_a(t) + f_b(t) \} dt \]

\[ = F_a(nT) + F_b(nT) \text{ assuming } F_a(0) = 0 \]

\[ \therefore e(t) = f(t) - W \sum_{n=0}^{\infty} d_n g(t-nT) \]

\[ = f_a(t) - W \sum_{n=0}^{\infty} F_a(nT) g(t-nT) + f_b(t) - W \sum_{n=0}^{\infty} F_b(nT) g(t-nT) \]

\[ = f_b(t) - W \sum_{n=0}^{\infty} F_b(nT) g(t-nT) \]

\[ = \frac{1}{2\pi} \int_{|w|>W} \phi_f(w) e^{jwt} dw - \frac{1}{2\pi} \int_{|w|>W} (jwe^{jwt}) \frac{\psi_f(w)}{jw} dw \]

\[ |e(t)| < \frac{1}{2\pi} \int_{|w|>W} |\phi_f(w)| dw + \frac{1}{2\pi} \int_{|w|>W} |\phi_f(w)| dw \]

\[ = \frac{1}{\pi} \int_{|w|>W} |\phi_f(w)| dw \]

Some examples of the use of the above formula are given in Fig. 5.6 along with a comparison of truncation errors for the above case and the cardinal function case. The convergence is seen to be at a similar rate.
FIG. 5.6-A. SINUSOID.

Ze = Zero line of error.

FIG. 5.6-B. SUM OF A NUMBER OF RANDOM SINUSOIDS.
Ze = Zero line of error. Max. error $\leq 3\%$.

**FIG. 5.6-C. GAUSSIAN DISTRIBUTION FUNCTION.**

**FIG. 5.6-D. DECAY OF TAILS OF $\sin(x)$ AND $\cos(x)$ FUNCTIONS.**
CHAPTER 6

AN INTERMEDIATE SPEED DATA CONVERSION SYSTEM USING
A SLOW SPEED CARD PUNCH

6.1 Introduction

At the commencement of this project, no facilities were available within the Department for electrical analogue to digital conversion. Therefore, a system has been developed for the conversion of analogue signals to digital data, with effective sampling rates of up to 5,000 samples a second whilst utilising a slow speed card punch. Signals are recorded on magnetic tape. Before conversion, the tape is formed into an endless loop. During conversion samples are taken at a rate limited by the card punch to approximately two per second. During each revolution of tape, a set of samples is taken; the number in the set is determined by the length of the loop and the tape speed. Each set of samples from successive revolutions of the tape is displaced in time from the previous set by an interval corresponding to the effective sampling rate required. The resultant output was initially rearranged into the correct chronological sequence using an IBM 1620 computer. To avoid having to rewrite the program when the IBM 1620 was superseded by an IBM 360/44, the 1620 simulator is used now.
6.2 General Description

Fig. 6.2 is a schematic of the system. The signal to be analysed and a clock signal are recorded on separate tracks of a magnetic tape. The tape is formed into an endless loop and two control pulses are recorded on a third track to mark the start and end of the portion of the tape to be analysed.

On replaying, the clock and control signals are fed into a timing and control circuit which determines the sampling positions. The signal goes through a filter and amplifier to a sample and hold circuit and is then converted to digital form using an 8 bit analogue to digital converter (ADC). The binary output of the ADC is converted to binary coded decimal form and then decoded to give an output in decimal form. The decimal output is presented in the correct order (hundreds, tens and unit digits) to a modified IBM 026 card punch by a digit sequencer. An end of punch pulse from the card punch advances the digit sequencer. An "eleven" punching is used to identify each field of three digits and each set of samples. Silicon-Controlled-Rectifiers are used to energize the punch interposer magnets. An alarm circuit turns off the system should one of a number of errors occur in the punching procedure. A sorting program also checks the data for various possible errors.
6.3 Recording

6.3.1 Signal Sources:

A study of industrial plants involves the measurement of a wide range of physical parameters. For analysis, it is usually necessary to have the quantities to be studied in electrical form, and a wide range of amplitudes and frequencies normally have to be coped with. For off-line analysis these signals have to be recorded as well and the limitations of the transducers and the recorder used need to be considered.

The only suitable recorder available for storing signals in electrical form was an Ampex SP300 instrumentation tape recorder. This has four tape speeds and four recording channels. Each channel is provided with both direct recording and frequency modulation modes. At the maximum tape speed of 15"/sec, the bandwidth available in direct mode is 50Hz to 40kHz and in the frequency modulation mode from 0 to 2.5kHz. For an amplitude accuracy of 1%, frequency modulation must be used. Also most signals from industrial plant contain frequency components below 50Hz which rules out the use of the direct recording mode.

The tape recorder FM units have an input impedance of 10 kilohms and require a 1VRMS recording signal. Field experience has shown that buffer amplifiers are nearly always required and should have the following characteristics:
1. Input impedance greater than 1MΩ.
2. Differential input.
3. Gain of up to 60 dB.
4. Frequency response; 0 - 2.5kHz.
5. Output impedance less than 1kΩ (single ended).
6. Provision for DC offsets.
7. Provision for AC coupling.
8. Preferably, ability to have inputs at high voltages relative to ground.
9. For high voltages, a series of attenuators giving up to 60·db attenuation.
10. Overall accuracy better than 1%.

6.32 Clock:

A clock frequency is recorded simultaneously with the plant signal. The signal then has its own time reference and errors due to tape stretch and speed variations in recording and playback phases are minimised. The shortest time increments required govern the choice of the clock frequency, e.g. a frequency of at least 10 kHz is needed to enable sampling up to rates of 10,000 a second. The frequency stability of the clock has to be better than one in \((\pi/2 \times 10^3)\) to obtain an accuracy of .1% in the amplitude of a 2.5 kHz signal.

The clock signal is recorded in direct mode on to the
tape-recorder. A square wave is used as it is the zero crossings of the clock signal which are important.

6.3.3 Start/Stop Signals:

A control signal is recorded on the third track of the tape to mark the actual portion of the tape to be analysed. A positive going step denotes the start of the signal and a negative going step denotes the end.

The recording of the control signal is best made after the formation of the endless loop so that the portion of the tape to be used can be conveniently selected. The joint and adjacent regions are left unused. This eliminates noise due to the effect of the tape joint moving over the head or round other tape guides.

6.4 Replaying

6.4.1 General:

When the tape is replayed, the main sections involved may be classified as the divider and controller, the converter, the encoder and the data-punching circuits.

The controller determines the sampling instants. At these instants scaled and quantized values of the input signal are determined by the converter. The encoder codes these and presents them in a suitable form to the card punch for punching as a three digit number.
FIG.6.421. FILTER AND AMPLIFIER.

FIG.6.422. SAMPLE AND HOLD.
6.42 The Conversion Unit:

The analogue input is amplified and filtered as necessary. It is then sampled by a sample and hold circuit and converted to an eight bit binary number by the ADC. A set of indicator lights display the output which is also available as a set of voltage levels. (See Fig. 6.425).

6.421 Amplifier and Filter (Fig. 6.421)

The output of the tape recorder is about 3V peak-to-peak. The ADC unit requires a signal in the range 0 to -10V for full output. Thus a gain of about 3.4 and a -5V offset is required for a signal with zero mean. Higher frequencies than those desired are filtered to simplify sampling requirements.

A Nexus SQ-10a operational amplifier is used to provide the gain and filtering is done by switching a capacitor in the feed-back circuit. An offset of up to -10V is available. This enables a certain amount of cancelling of DC bias in inputs and is also useful for setting up the ADC as will be explained. The feed-back resistor is chosen to be a large value so that low cut off frequencies can be obtained; the lowest being about 5Hz when using a 1μF capacitance. External capacitors may be used and there is an off position for the filter. The Nexus amplifiers are well protected against accidental overload of inputs and outputs so that no special precautions are needed.
They have open loop gains of greater than 1,000 up to 2.5kHz and so are suitable in this regard too. They are capable of an output swing of over ±10V.

To utilize the maximum range of the ADC unit the input needs to be between 0 and -10 volts. This introduces the problem of ensuring that this range is fully utilised without exceeding either limit. An ordinary oscilloscope is not sufficiently accurate as the resolution of the ADC is 40mV per bit or about .4% in terms of an accuracy requirement on the oscilloscope. A "peak detector" has been developed which has a dead band for most of the 0 to -10V range and then a sensitivity of about 3 bits full-scale.

To detect randomly occurring peaks of short duration, it has a very high on to off time constant ratio. A 40μSec pulse will give a good deflection on the meter. (See Chapter 7).

6.422 Sample and Hold:

As the ADC takes about 20μSecs to complete a conversion, it is necessary to hold the value of the input signal over this time. A 2.5kHz signal (10V p-p) could vary by about 300mV in 20μSecs.

The schematic of the sample and hold is shown in Fig. 6.422. In the sampling position, S2 is open and S1 closed. The storage capacitor voltage is such that the input and
**FIG.6.423-A. OUTPUT REGISTER STATES IN A 3 STAGE SUCCESSIVE APPROXIMATION CONVERTER.**

<table>
<thead>
<tr>
<th>Set</th>
<th>Comp</th>
<th>Reset</th>
<th>Set</th>
<th>Comp</th>
<th>Reset</th>
<th>Set</th>
<th>Comp</th>
<th>Reset</th>
<th>OP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>010</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>101</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>110</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>111</td>
</tr>
</tbody>
</table>

---

**FIG.6.423-B. SUCCESSIVE APPROXIMATION CONVERTER.**

Analogue input → Comparator → Control Gates and Clock → Start Pulse

Digital to Analogue → OP Register → Control Register → End of Conversion
output voltages of the system are within $1/(\text{open loop gain of the feed-back system})$. In the hold condition, S1 opens isolating the storage capacitor; the output voltage is that of the input voltage at the instant S1 opened, minus the leakage which is negligible over $20\mu$Secs. S2 closes as S1 opens so as not to drive the first amplifier into saturation.

When a conversion is required, a start pulse turns S1 off and S2 on. When the conversion is complete, the end of conversion pulse from the ADC is used to turn S2 off and S1 on again allowing the output to follow the input. The settling time is short enough to enable sampling at up to 10,000/Sec.

6.423 The Analogue to Digital Converter:

The ADC used is an 8 bit "successive approximation" type sold in kit-form by Digital Equipment Co-operation Limited. The method of operation and a schematic of the converter are given in Figs. 6.423-A and 6.423-B. See also [34].

The input voltage is repeatedly approximated, the possible range of an ADC step being halved each time. In the first step, the output register is set to half range to see if the input is greater or less than it. Then quarter or three quarter range is tried depending on the outcome of the first test. The process is repeated until the least
significant bit is reached. Fig. 6.423-A shows all possible sequence steps in a 3 bit ADC.

The start pulse sets a one in the most significant bit of the output register and the first position in the control register. The ADC clock frequency is such that the comparator has settled before the arrival of the next clock pulse. At this time the first bit in the output register is retained or removed depending on whether the input was greater or lesser than the output of the digital to analogue converter. At the same time a one is placed in the next bit of the output register and also the bit in the control register is shifted one position. The control register keeps track of the step in progress. When all the bits have been tested the end of conversion pulse, from the last bit in the control register, stops the action.

Input range to the ADC is 0V to -10V. The time per step is about 2μSecs, making the total conversion time, including settling times, just under 20μSecs for 8 bits.

6.424 Control Signals to ADC:

In order to provide the correct start pulse for the ADC, and also to enable manual operation of the ADC, a single shot multivibrator followed by a pulse shaper is used, the output of this becoming the start pulse. The unit converts a 6V positive going pulse of a few microseconds rise time to
**FIG. 6.425. THE CONVERSION UNIT.**

Analog

- **IP**
  - 2.8V P-P
  - Offset
  - Max ±14V

Amplifier, Filter, Offset.

Sample and Hold.

Peak Detection

Sample and Hold

resume sampling

start holding

Pulse Shapen

start ADC

ADC complete

Analogue to Digital Converter.

Lamp Indicators, Buffer Amplifiers.

OP

8 Binary Bits.

IP +6V

Max rate

10000/sec

OP(6V)
the required 3V pulses of short duration and rise time. Eight lamp indicators are used to show the states of the output registers and the lamp drivers also serve as buffer amplifiers between the output register and terminals. The output is a -15V level or a 0V level for "on" or "off" conditions respectively. The end of conversion pulse output of the ADC is shaped into a 6V positive going pulse before reaching the output terminal.

6.425 Interconnections:

The converter unit interconnections are shown in Fig. 6.425. The amplified and filtered analogue input is continuously presented to the sample and hold unit. The output of the sample and hold is normally following the input. On arrival of a start pulse, however, the sample and hold output is held constant. The pulse shaper delay permits the ADC input to settle. The ADC makes the conversion and the end of conversion pulse is shaped and used to enable the sampling process to be repeated and also appears as an output. The result is displayed on lamps and appears as an 8 bit output on the binary output terminals.

6.43 The Encoder:

The input to the encoder is an 8 bit binary number and the output is a 3 digit decimal number between 1 and 256. The digits are presented sequentially, highest power first,
FIG. 6.431-A. BINARY TO BCD CONVERTER.

FIG. 6.431-B. SYNCHRONIZING UNIT.

FIG. 6.431-C. WAVEFORMS.
on a ten line output cable and successive digits appear
when the digit sequencer is pulsed. An eleventh output
line is also used. An elevens punching above the hundredth
position denotes the start of a sample and an elevens
punching above the unit digit indicates the start of a new
set of samples from a successive tape revolution. An alarm
circuit checks for certain possible errors.

6.431 Binary to Binary-Coded-Decimal Conversion

The organisation of the binary to binary coded decimal
converter is shown in Fig. 6.431-A. The 8 binary inputs are
normally isolated from the 8 bit register by a set of
electronic switches. The start command fires a single shot
which closes the switches so that the binary register is set
to the complement of the input. The inverse of the set pulse
is synchronized with an internal clock and triggers a flip
flop which opens a gate letting the clock pulses into both
the binary and BCD registers. When sufficient clock pulses
have been added to the binary register to make the value
stored zero, the binary register gives out a pulse. This
is used to disconnect the clock. The BCD register has
counted the same number of clock pulses and hence has the
required value stored in it. Synchronisation of the clock
and start pulses is necessary because the start pulse occurs
randomly with respect to the clock and ambiguities could
FIG. 6.432-A. BCD REGISTER AND DECODING.

FIG. 6.432-B. WAVEFORMS.
result in the absence of synchronism.

Fig. 6.431-B illustrates the synchronizing action and Fig. 6.431-C gives the associated waveforms.

6.432. BCD Register and Decoding:

The code chosen for the BCD register was a symmetrical 1242 code as this simplifies the decoding to decimal. Coding is achieved by gating rather than feedback as the former is generally considered more reliable [66].

For the decimal decoding only 28 gate inputs are required as opposed to 30 for most of the other types of codes. A maximum of three input gates are required per digit decoded and this simplifies gate design.

Fig. 6.432-A shows the 1242 symmetrical coding decade register and the arrangement of decoding used. The waveforms are shown in Fig. 6.432-B.

6.433. Digit Sequencing for Punching:

As the card punch only punches one digit at a time, a sequencer is used to present the digits one at a time.

There are four positions for the punch sequencer; hundreds, tens, and units digits and the wait positions. The sequencer is normally in the wait position till the end of binary to BCD conversion pulse steps it to the hundreds position. The "end of punch" signal from the card punch steps the sequencer on to the tens position, then to the unit
FIG. 6.433-A. DIGIT SEQUENCING FOR PUNCHING.

A
B
C
D
E
F
G
H

End of B to BCD.
Wait.
100s
Blanking.
End of Punch.
10s
Units
Punch 111. Line 1.
H1
H2

End of Punch.
Line 1.

FIG. 6.433-B. WAVEFORMS.
Start of Rev. A
Punch Units D
C
100s. OP line. F

FIG. 6.433-C. CONTROL DIGIT.

Start Punch.

And

Not

Reset

Wait (Sequencer)

Sequencer to Wait Position

Driver.

To TR Remote Control.

To 026 Keyboard Lock

To External Alarm.

FIG. 6.434. ALARM CIRCUIT.
position and finally back to the wait position when the three digits have been punched.

As only one set of lines are provided for the digits 0 to 9, the sequencer must select the power of ten to be punched. Only one of the lines in the set will carry a level on it for each power of ten. The drivers are blanked off for a few milliseconds and then switched on. The appropriate line in the set 0 to 9 and possibly the elevens punching line carry pulses to the card punch. After punching is complete, the card punch returns an "end of punch" signal which advances the sequencer. This enables the next digit to be punched. The blanking is necessary because the drive capacitor on the punch switch must be discharged to -6V before the switch can operate again. (see Fig. 6.433-A). If there was no blanking, consecutive punchings of the same digit would not be possible.

An eleven punching over the hundreds digit is used as a control digit to denote the start of each three digit numbers and one over the unit position is used to denote the first number in the set of readings from each revolution of tape. A flip-flop opens a gate when a new tape revolution starts. This allows the "punch unit" pulse to be sent along the eleven line, as well as to reset the flip-flop and thus close the gate till the start of the next tape revolution. The "punch 100" pulse always energizes the eleven line. (see Figs. 6.433-C and 6.433-D).
The four positions for the sequencer are obtained from two cascaded flip-flops, decoded to give the four states sequentially. The common line to the hundreds, tens or units digits circuits at the digit switch is appropriately energized, permitting the active line in this digit circuit to be connected to the driving circuit. The pulses that increment the sequencer also trigger a single shot which blanks out the drivers to the punch for 5m Secs. Besides assuring that the drive capacitors have been discharged, this also gives ample time for the inputs to the drivers to settle. This is particularly necessary when two punchings (the eleven and a digit) are being made simultaneously.

The end of punch pulses from the punch are gated so that those due to duplicating or sequence numbering do not advance the sequencer. Punching can only be started from the encoder when it is ready to punch.

6.434. Alarm:

An alarm circuit turns off the tape recorder, by means of the remote control switch, if punching is attempted at a time when the card punch is not ready to punch a number. Thus the alarm operates if

a) punching of numbers is attempted at too high a rate.

The longest punch cycle occurs when the last columns on a card have to be punched and then a new card has to be
registered before the next number can be punched. Thus, even though the system may punch successfully at the middle of a card, the punching rate may be too high.

b) stray pulses or other malfunctions cause the system to attempt to punch a number. This protection is not infallible if the wait period is longer than one punch cycle.

c) the card punch runs out of cards or if the card feed mechanism fails. The first cause can be used to automatically stop the system at the end of a run if the appropriate number of cards are placed in the card feeder.

It is possible to stop the system in the event of a card-jam at the punch station, using the fact that the card punch will not duplicate when there is no card present at the read station.

A pulse arriving at the sequencer from the encoder when it is not waiting to punch is directed to a flip-flop which in turn energizes a relay. The relay turns off the remote control tape recorder switch. A voltage level, or contacts on the relay are available for external use. The alarm is reset externally. One of the pairs of contacts of the relay is also used to lock or unlock the keyboard of the card punch. This enables the card punch to be used for setting up but locks it when the conversion system is in use so that the automatic punching cannot be inadvertently interrupted from the
FIG. 6.44-A. DIVIDER AND CONTROLLER.

FIG. 6.44-B. WAVEFORMS.
keyboard. (see Fig. 6.434).

6.44. Divider and Controller:

A schematic of the circuit is shown in Fig. 6.44-A along with associated waveforms in Fig. 6.44-B.

The clock signal recorded on one of the tape channels is used to control the timing for the sampling point determination. As the clock signal has been recorded simultaneously with the signal being processed, correct time relationship is ensured, irrespective of tape stretch and tape transport speed fluctuations.

The relevant parts of the signal are marked by start and stop pulses recorded on a third track of the tape. The sampling points are all referred to the start pulse. During each revolution of tape, punchings are made at approximately half second intervals until the stop position on the tape is reached. The next set of samples is displaced from the first set by $T_s$ seconds where $1/T_s$ is the desired sampling rate. When all the sub-intervals between successive punchings in the first revolution have been covered, a complete set of data has been obtained. Thus if $1/T_s$ is 256 and punching is approximately every half second, the tape makes about 128 revolutions before a complete set of data is obtained.

6.441. Clock Signal Circuit:

The clock pulses are first shaped by a Schmidt trigger
FIG. 6. Clock Signal Circuit.
and amplified to the standard 6V level needed for the system logic. (see Fig. 6.441). However, due to imperfections in tape properties and dust accumulated on the heads during playback, some of the recorded clock pulses may be lost. To overcome this a "pulse filter" is used. This is an astable fly wheel oscillator whose nominal frequency is automatically controlled by the incoming clock frequency. Normally it acts as a locked oscillator, locking on to the clock pulses to maintain the correct timing. To achieve this the basic frequency of the oscillator is set just below the clock frequency. Should any clock pulses be missed, however, the flywheel will take over and fill in the pulses at approximately the correct time. Provided not more than a few clock pulses are missing, the oscillator will lock on to the clock again without error. (See Chapter 7). Though the position of insertion of the pulse may not be strictly accurate, this does not cause an error unless a sample is being taken at that point, in which case there will be a slight jitter at the sample point.

The pulse filter has a range of from 1-4kHz. A three position switch enables either

1) the filter to be switched off (e.g. for direct use at other frequencies), or
2) the filter to be connected into the input to the divider chain if the clock frequency is in the range 1-4kHz.
or 3) the filter to be coupled in after two stages of binary division if the clock frequency is between 4-16kHz.

Depending on the sampling rate needed, the clock pulses may go through a number of stages of binary division. The clock pulse is then taken to a gate which connects it to one of three positions -

(A) to an off position,
(B) on to a counter in the comparator circuit, or
(C) back to further stages of binary division.

At the beginning of a new tape revolution when the start pulse occurs the gate is set to position (B), and the clock pulses are counted by a "displacement" counter. The contents of this register are compared with a "displacement" store which records the number of revolutions the tape has made. This number corresponds to the displacement of the first sample from the start position. When the two registers are equal, the clock pulses are diverted back to the binary dividers by the comparator switching the gate to position (C). The clock pulses then are divided down to produce the nominal \( \frac{1}{2} \text{ sec} \) timing pulses. As there are a total of 13 dividers, the maximum usable clock frequency is 16kHz. For lower clock frequencies the appropriate number of the dividers may be patched out of circuit. The input to the gate is tapped off from the divider chain at a point between the filtered clock input and the output to the ADC. The higher the sampling
rate, the closer the tapping point is made to the input end, as this corresponds to shorter sampling intervals.

When the stop pulse is generated the gate is set to the off position (A). As well the displacement counter and the portion of the divider chain after the gate are reset; the counter to zero and the divider chain to contain all 1's. This latter feature is necessary so that the first sample is taken as soon as the comparator sets the gate to position (C).

A display lamp lights up, at the times samples are taken, to enable a visual check to be obtained on the punching speed while setting up.

6.442. Start/Stop Signal Circuit:

The start-stop signal is recorded on the third tape channel. Due to drop out on the tape the signal has to be filtered. The start position has to be accurately defined to minimize jitter, but the stop pulse needs only to be positioned within a punch interval. As the tendency of the tape-recorder in the FM mode is to have only negative going interference pulses, a positive going level is used for the start pulse. The negative going pulses are smoothed, and a negative level present for at least 100 mSecs is taken to be the stop pulse.

The start pulse triggers a flip-flop FF1 (Fig. 6.442) to state A from state B. This opens a gate letting the clock pulses into the comparator. The other output of FF1 is used.
A = Displacement Counter.
B = Divider Chain.
C = Displacement Store.
D = Reset.

**FIG.6.442. START/STOP AND CLOCK GATE.**

- Reset at start signal.
- Clock from gate.
- OP to equal gate.
- Increment at start signal.

**FIG.6.443. DISPLACEMENT CIRCUIT.**

Note: Divider 5 is omitted.

**FIG.6.444. PATCH PANEL.**
to increment the displacement store count when the flip-flop is switched back again. The comparator "equal" pulse also triggers another flip-flop FF2, from state B to state A. This opens the gate that connects the clock pulses back to the dividers. As FF1 is already back to state B, the clock pulses are prevented from going to the comparator. The stop pulse sets FF2 back to state B, preventing the clock signal from reaching the dividers. The other output of FF2 is used to initiate the pulse to reset the dividers and the displacement counter. (see Fig. 6.44-B for waveforms).

6.443. Sampling Point Determination:

The number of revolutions of the loop of tape is stored in the displacement store register and this determines the displacement from the start point before sampling commences. The store is incremented by one for each tape revolution and sampling starts one unit of time $T_S$ later than the previous revolution. The unit of time is determined by the clock frequency and the number of divider stages before the signal is taken to the gate. If the clock period is $T_C$ and there are $n$ stages of division before the input to the gate, then

$$T_S = 2^n T_C$$

assuming for the moment that there is no tape-recorder speed reduction. (Note that $1/T_S$ is the sampling rate.)

At the start of a run the gate connects the clock to the
displacement counter and the contents of this counter are compared with the contents of the displacement store. If the store has an initial value \( N \) stored in it, then the counter records \( N \) counts before an output is obtained from the comparator. Thus the first sample is taken at a time \( N T_s \) seconds after the start pulse.

As explained in section 6.44.1, the gate now connects the clock to the remainder of the divider chain. If the divider chain has a total of \( m \) stages patched in, then samples will be taken at intervals of \( T_P = 2^m T_c \) seconds.

The maximum speed for straight punching on the card punch is 20 characters a second, i.e., 50 mSecs per character. However, the card register cycle takes 150 mSecs so that the absolute minimum value for \( T_p \) is 300 mSecs. In practice this value for \( T_p \) was found to be extending the machine somewhat and a slightly longer interval, say 400 mSecs, is preferred. Thus \( T_p > 0.4 \) sec.

Sampling at \( T_p \) second intervals continues till the stop pulse is encountered. The clock pulses are isolated at the gate, the counter and dividers are reset and the system waits until the next start pulse is received.

The store now has a value \( (N + 1) \) stored in it and hence sampling is delayed \( (N + 1)T_s \) seconds from the start point; that is \( T_s \) seconds from the first of the previous set of samples. In each set, sample points are \( T_P \) seconds apart and
and hence $T_s$ seconds from the corresponding sample points in the previous set.

At the $i^{th}$ tape revolution the store will have $(N + i)$ stored in it and sampling starts $iT_s$ seconds from the first sample point.

Samples will start overlapping previously obtained ones when

$$iT_s = T_p$$

substituting $T_s = 2^mT_c$ and $T_p = 2^MT_c$

$$i = 2^m/2^n$$

$$= 2^{(m-n)}$$

At this point all the positions on the tape will have been sampled. However, further points may have to be taken, in order to get a "complete" set - complete in the sense of the sorter program.

The maximum value that $m$ can take is equal to the total number of divider stages available, i.e., 13. The displacement counter and store have ten binary stages and hence, allowing for the fact that a few samples have to overlap, the maximum number of revolutions for a complete set is $2^9 = 512$. Thus $m - n < 9$.

The maximum clock frequency that can be used is 20.5kHz since there are only 13 divider stages. The maximum sampling rate of 1280 samples per second is independent of the clock
frequency, but depends on the punching rate. (i.e. clock frequency has to be \( \leq 2^{13}/T_p \) and sampling rate is \( \leq 2^9/T_p \).)

However, a maximum time scaling of 8 is available on the tape-recorder enabling effective sampling rates of up to 10,240 samples a second.

Fig. 6.444 shows the patch panel wired for \( n=3 \) and \( m=12 \).

6.444. Operating Procedure:

Let the desired sampling rate be at least \( f_{sd} \) samples per second, where

\[
f_{sd} \leq 10,240.
\]

If \( f_{sd} \leq 1,280 \) then no speed change on the tape-recorder is necessary. In general let \( k \) stages of binary speed reduction be needed. Then

\[
2^k \leq f_{sd}/1,280
\]

Let the clock frequency originally recorded be \( f_0 \). The clock frequency input to the system = \( f_c = f_0/2^k \). The number of stages \( n \), before the gate is given by,

\[
f_s = f_c/2^n \geq f_{sd}
\]

where \( f_s \) is the actual sampling rate.

\[
\therefore \ n \leq \log_2(f_c/f_{sd})
\]

The total number of divider stages \( m \), is given by,

\[
f_c/(2^m) \leq 1/T_p.
\]

\[
m \geq \log_2(T_p f_c).
\]
If the signal is \( T_r \) seconds long between start and stop pulses at the original recording speed, then the time between the start and stop pulses when replaying is

\[
T'_r = T_r 2^k \text{ seconds}
\]

and the number of punchings per revolution, \( N_{pr} \), is the greatest integer < \( T'_r/T_p \).

Let

\[
(T'_r - T_p N_{pr}) = \delta T
\]

and let the number of sampling increments \( (T_S \times 2^k) \) that can be contained in \( \delta T \) be \( j \);

i.e. \( j = \lfloor \text{largest integer} \leq \delta T/(T_S 2^k) \rfloor \)

Then for \( j \) revolutions, there will be \( N_{pr} \) punchers per revolution. After this there will be \( (N_{pr} - 1) \) punches per revolution for a further \( M \) revolutions where

\[
M = 2^{(m-n)}
\]

and the \( N_d = \lfloor (N_{pr} - 1) \times M \rfloor \) sample points cover the tape at the basic sampling interval of \( T_S \) seconds. After this there will be groups of \( (N_{pr} - 2) \) punchers per revolution.

As there are 26 sample points punched on a card the number of cards needed may be estimated as

\[
N_c \geq \frac{N_d}{26}.
\]

For the sorter program, it is necessary to have at least two sets of data at \( (N_{pr}) \) punchings per revolution, \( M \) sets of
(N_{pr} - 1) punchings per revolution and at least one set of
(N_{pr} - 2) punchings per revolution. This enables the program
to check that there is in fact a complete set as well as to
find the set without having to be given any further instruction.

It will be noticed that up to a maximum of T_p seconds of
tape may be unused out of the record length T'_r and this should
be allowed for.

It is possible to start at the beginning of a complete
set by using the displacement store display lights and the
sample-point trigger display light when T_r and f_o are known.
The value j is calculated and the displacement store value N
is set to just under j. Punching will then start with N_{pr}
punches per revolution but will change to (N_{pr} - 1) after a few
revolutions. A saving in time and cards of possibly up to
100% can be achieved in this way.

Punching is stopped when the groups of punchings become
(N_{pr} - 2). This may be automatically achieved, as explained,
by putting the appropriate number of cards in the card feeder.

Consider the case where a signal is to be sampled at
least 3,000 times a second, and the recorded clock frequency
is 8,000 Hz.

\[ i.e. \quad f_{sd} = 3,000 \]
\[ \text{and} \quad f_o = 8,000 \]
$2^k > \frac{f_{sd}}{1280} \div 2.5$

i.e. $k = 2$ and $f_c = 2,000 \text{ Hz}$

$\frac{f_o}{f_s} \div 2.6$

so that $n = 1$

$m > \log_2(800) \div 9.5$

Take $m = 10$.

$m - n = 9$ which is acceptable and the tape will go round 512 times.

The actual sampling rate $f_s$ is

$\frac{(f_c \times 2^k)}{2^n}$

i.e. $f_s = 4,000$

The punching speed is $1/T_p = f_c/2^m$

$\div 2$ punchings per second.

If $T_r = 2$ secs, then

$N_{pr} = 2^k T_r / T_p$

$\div 16$ per revolution.

The number of data points will be about

$N_{pr} \times N_r \div 6 \times 512 \div 8,000$

The number of cards needed is about

$N_{pr} \times 2^{m-n}/26 \div 300$.

In fact $T'_r = 2^k \times T_r = 8$ and
Thus \( \frac{T_r'}{T_p} = 15.625 \)

Thus \( N_{pr} = 15 \)

\[
\delta T = T_p \left( \frac{T_r'}{T_p} - N_{pr} \right)
\]

\[
= 2^{10} \times 0.625/2000
\]

\[
= 0.32
\]

and \( j = [0.32 \times 1000] = 320. \)

The starting value \( N \) could be set to about 315. The actual number of sample points is \( (N_{pr} - 1)M = 7168. \)

It is probably easier to find the setting for starting just before a set by trial. The store is set to various trial values and the number of flashes on the "sample-light" or the number of punches counted till the transition point is determined. This eliminates estimating \( T_r \) which is advantageous as \( T_r \) has to be known fairly accurately for the calculation.

The displacement store can be reset to zero, or set to any desired number \( N \) using the manual reset and set switches. For convenience, there are three rates of 128 pulses/second, 16 pulses/second and single pulses. (Fig. 6.443). The pulses are derived from a circuit that is a monostable but is switched to be an astable one for either of the above frequencies. Display lights indicate the state of the store at any time.
6.45. Punching of Data:

A general schematic of the punching circuits is shown in Fig. 6.45.

The card punch (IBM 026) punches data sequentially column by column. When any interposer on the card punch is energized the card advances one column and punches the digit or digits corresponding to the interposer or interposers energized. Any combination up to a maximum of four columns may be punched provided the interposers are energized simultaneously.

The punch interposers are normally energized by manual depression of keys on the keyboard. For the purpose of driving them electrically from an external source, a bank of Silicon Controlled Rectifiers have been added.

Card advancing and registering of new cards are automatically carried out by the card punch itself. The punch also has a number of other features, one of them being automatic duplicating of specified columns of the card just completed. This feature is utilized to duplicate two columns on to every card for the purpose of, among other things, keeping a constant format on the 80 columns of a card. Provision has been made to enable sequence numbering of the cards to be carried out if necessary.

When the punch is being used with the converter, the keyboard of the punch is locked by means of a relay energized
FIG. 6.45. PUNCHING CIRCUITS.

FIG. 6.451-A. INTERPOSER SWITCHING.

FIG. 6.451-B. SCR CIRCUIT.

FIG. 6.451-C. TIMING.
from the converter. The interlock only operates when cards are registered. Furthermore, the interlock is made inoperative when the converter alarm is energized. Hence, switching on the alarm enables the card punch to be set up.

6.451. Interposer Switches:

The overall circuit of the interposer switches is shown in Fig. 6.451-A.

A Silicon Controlled Rectifier (SCR) is used to switch in a punch interposer. The SCR circuit is shown in Fig. 6.451-B. The sensitivity of the SCR's can be adjusted by a potentiometer in the input circuit to fire at about 4V. The input pulse has to be a positive going one of at least four volts amplitude. Before going positive it has to have been negative for at least 5 mSecs. This feature is necessary to reduce the tendency to fire spuriously off stray pulses.

The SCR's are de-energized by isolating the common line of the punch circuit from the SCR's common rail shortly after it has been energized. Otherwise the interposer can be re-energized causing multiple punching, particularly if the interposer springs are weak and return is slow. A spare contact on relay R46 of the card punch is found to achieve the above results. (Fig. 6.451-C). Also the relay contact of R46 is wired to the punch common rail via the card lever switch, so that the SCR's cannot be energized if a card is not registered.
FIG. 6.452. SEQUENCE NUMBERING.

FIG. 6.453. END OF PUNCH CIRCUIT.
6.452. Sequence Numbering of Cards:

Three uniselectors are used as the counters for sequence numbering. Two flip-flops, as in the converter system, sequence the digits to be punched. An independent circuit is incorporated in the punch to minimize interconnections as well as to enable the sequence numbering to be used on its own. The end of punch signal is used to step the digit sequencers on only when (1) a "2" punching is present on the program card of the punch and (2) when the sequencer is not in the wait position. When the digit sequencer is in the wait position and a "2" is encountered on the program card, the sequencer advances to the "drive hundreds" position punching the hundreds digit of the sequence number. The end of punch pulses are now permitted to step the sequencer on thus punching the tens and unit digits of the sequence number and reaching the wait position. The unit uniselectors are advanced at this point incrementing the sequence number. Pressing the reset switch sets the sequence number to 000. The corresponding digits on each selector are paralleled and taken to the SCR's. (see Fig. 6.452).

6.453. End of Punch Signal:

The end of punch signal is obtained using the relay R46. When it operates, it shorts to ground the end of punch line which is normally biased to +6 volts. When the relay releases
again, the end-of-punch line returns to +6 volts and this constitutes the end of punch signal. (Fig. 6.453).

6.46. Sorting of Data:

The data as punched is finally sorted into its correct sequence using the sorter program and the 1620 simulator on the IBM 360/44. The program (see Fig. 6.46 for block diagram) reads in a header card and as many data cards as are present, provided the storage of the computer is not exceeded.

The data is then checked to see if there is a complete set of points and that there are no erratic number of punchings per revolution. During sorting the data is checked for discrepancies such as non three digit numbers. The sorted data is punched out with the appropriate headings. The typewriter also types out a diagnosis of the data, giving such information as how many extreme points there were and the range of values.

6.461. Reading in Data:

One header card is read first followed by data cards. The end of one data set is denoted by a blank card after the set, before the header card for the next data. The data cards must have a digit in at least one of columns 4 or 5; a blank in these two columns is taken as being the end of a data set. Any number of data sets may be sorted, one after the other. If the cards contain more data points than there are storage positions, this information is typed out. The remainder of
FIG. 6.46. SORTER PROGRAM.

FIG. 6.462. FINDING THE DATA SET.
the cards in the set are disregarded. The data in storage is then examined unless console switch 1 on the 1620 simulator is on, in which case the data in storage are ignored and a new set of data, if present, processed.

If the cards have sequence numbers in columns 75-77, these columns are not read as data. The absence of an eleven punching in column 75 is taken as indicating that it is a sequence number. If it is an erroneous punching, the error will be detected later, e.g., if the "eleven" punching had failed, this will almost inevitably result in unevenness in the resulting number of punchings per revolution.

6.462. Finding the Data Set and Checking:

A block diagram for this section is shown in Fig. 6.462. When input data has been read, the data points are scanned for start of runs. This is indicated by an eleven punching (which reads into the machine as a "flag") over the low order position of the data points. Every low order position is checked. The first flag is ignored as eleven punchings can arise in the starting procedure. The number of punchings between low order flags is counted. Thus if the first group of data have \( N_{pr} \) punchings between low order flags, this should be followed by a set of \( (N_{pr} - 1) \) punchings between flags. If the number of punchings after \( N_{pr} \) punchings between flags is other than \( (N_{pr} - 1) \) there is something erroneous and a
message including the position of error is typed out. The groups of \((N_{pr} - 1)\) punchings constitute the required set. It should be followed by a group of \((N_{pr} - 2)\) punchings; if there is a group of any other number an error is indicated as above. If on the other hand, there are no groups other than \((N - 1)\) present, a complete set of data does not necessarily exist and this is indicated. Processing of any set terminates in the event of any error and the next set of input data is processed.

6.463. Sorting and Punching:

When the necessary data points have been found, it is relatively easy to sort the data. The sorting process may be visualized as the transposing of a matrix \([x_{i,j}]\) if the data points are appropriately named. The first suffix of the data points \(x_{i,j}\) is to denote the group of punchings the data is to be taken from, corresponding to the number of the tape revolution. The second suffix denotes the position in this group, e.g., the sixth punching from the third set of data would be \(x_{3,6}\). Thus data appear on the cards as \(x_{1,1}, x_{1,2}, x_{1,3}, \ldots x_{2,1}, x_{2,2}\) etc. After sorting it will appear as \(x_{1,1}, x_{2,1}, x_{3,1}, \ldots x_{1,2}, x_{2,2}\) etc.

Each number is checked to see if it is a three digit number and if it is an extreme valued (i.e. 001 or 256) number. The number of extremes are counted and the maximum and minimum values, excluding the extremes, are also found.
Remove the 3 leads marked as at left if no sequence no.'s

FIG. 6, 464: 570-PANEL WIRING.
Further, the number of punchings per revolution and the number of revolutions as found by the sorter are typed out for comparison with expected results. The portion of the input deck used is also worked out.

The present program can only sort as many data points as the 1620 simulator storage allows. It is possible, however, to construct programs that are multipass; n passes are required for sorting data of size \((n-1)\) times storage. The first pass would determine the amount of data and leave subsequent passes to sort and punch the data. The inconvenience of "multipass" is not noticeable if a disk is available, apart from the additional processing time.

6.464. 870 Panel:

A panel has been wired for the IBM 870 data documenting system which effectively presents a "sorted" listing of the data from the original deck of cards.

Very briefly, the low order flag is used to return carriage on the typewriter, so that the second data in each set appears first on a line. Reading down vertically gives the correct sequence of the data. This has been found useful particularly in testing as "sorting" takes place irrespective of whether there is a proper set of data or not.
6.5. Performance

The ADC system has performed satisfactorily and has been used extensively in the conversion of a large number of signals, (Peek [76]). The Figs. 6.5-A, B, C and D show some of the data or processed form of the data converted by the system.
Input to ADC system.
100 Hz b/w signal from noise generator.

**FIG. 6.5-A.**

Output of ADC system.
Input above.

**FIG. 6.5-B.**
FIG. 6.5-C. PROBABILITY DENSITY FUNCTION OF A GAUSSIAN SIGNAL.

Zero Lag.

FIG. 6.5-D. CORRELATION FUNCTION OF A RANDOM SIGNAL.
This chapter gives details of some of the units referred to in the previous chapter.

7.1. Sample and Hold Unit

As the accuracy of the ADC is about 0.5%, it is desirable to have the sample and hold accurate to better than 0.5% - say 0.1%. Sample and hold units described in the literature (e.g. [46] and [57]) were found to be too expensive, not sufficiently accurate or to involve critical adjustments and components.

A sample and hold unit has been developed which overcomes these objections. This is a feedback type in which the overall characteristics are governed principally by the choice of the DC amplifier incorporated in it.

The circuit finally used is shown in figure 7.1. A Nexus SQ-10a operational amplifier is used. It has an open-loop DC gain of about 100,000 and a gain greater than 1,000 up to 2.5kHz, although with the 1,000pf capacitive loading necessary, it was initially found to be unstable.

The voltage on the capacitor is sensed by the output buffer amplifier and unity feedback is applied from this output. In the hold position the capacitor is isolated by S2 and a low gain feedback loop is connected across the high
**FIG. 7.1. SAMPLE AND HOLD CIRCUIT.**

Hewlett Packard Function Generator
Set to Ext., Har.

Audio Frequency Oscillator.

\[ f_0 = A \sin(wt) \]

\[ f_1 = A' \sin(nwt + \phi) \]

G.R. Pulse Generator.
Set to External.
Pulse width = Hold time.

Follow
Hold

**SAMPLE AND HOLD.**

Tektronix Type 545 'Scope.
Type W Plug-in.

**FIG. 7.11. TESTING CIRCUIT.**
gain amplifier. This prevents the amplifier saturating. It is necessary to prevent saturation of the amplifier since the recovery time once it has saturated is considerable.

P-channel field effect transistors (FET'S) are used for switching and they are switched by means of a flip-flop which is in turn interrogated by the "sample" pulse or the "end of conversion" pulse, as necessary. The FET used has an "on" resistance of the order of 300 megohm. However, the junction capacitance alters between the "on" and the "off" states thus changing the charge held in the hold capacitor.

The above reason was the main factor in determining the minimum value of the hold-capacitance. The maximum value of the hold-capacitance is determined by the high frequency response desired, as well as by the stability of the driving amplifier when feeding the capacitive load. The variation in the capacitance change is the crucial thing, as any permanent change is merely an offset in the output. Leakage from the capacitors should also be considered in determining the minimum capacitance. In this case, the leakage was essentially through the FET switch. As the time constant is 0.3 secs, leakage has a negligible effect seeing that the conversion time of the converter is 20μsecs.

The buffer amplifier was made from a MOS-FET linearized by connecting a transistor across the drain and source, thus
providing feed-back and a constant current source for the
FET. The leakage is negligible and linearity good.

The switches operate within 60μsecs, and moreover as
the output is normally following the signal, the value held
is that at the instant of sampling. The settling time is
about 2μsecs when S₂ is turned off and about 20μsecs when S₂
is turned on.

7.11. Testing:

The circuit was tested for DC and AC performance.

1. DC. Voltages corresponding to the normal range of the
circuit were applied to the input and the output examined
using a Tektronix storage oscilloscope with a type W plug in.
The plug in has an accurate offsetting device which enables
measurements to be made to about .01% (1mV resolution in 10V).
The output of the sample and hold was found to be accurate to
within a few millivolts over the 0 to 10V range, with a
permanent offset of about 20mV.

2. AC steady state. The difference between the input and
output in the follow mode was measured using the type W
plug-in and the error found to be within 20mV peak to peak
up to 2.5kHz.

3. AC sampling. Using the circuit shown in Fig. 7.11 the
dynamic performance of the device was tested over a range
of frequencies and for varying positions on the wave-form.

The function generator gives an output synchronized to a multiple of the input frequency, which is the testing frequency. The phase of the output with respect to the input can be varied. This output drives a pulse generator which gives the necessary pulses for operating the sample and hold switches. The pulse width controls the hold time. Synchronized sampling of up to a few samples per cycle was achieved with the point of sampling being continuously variable.

7.2. Filtering

7.2.1. Problems Arising From Using a Magnetic Tape Loop:

A number of problems can be attributed directly to the fact that a magnetic tape recorder is used for recording and storing the signals. These are:-

1. Noise in the signal channel.

   Noise in the signal channel can be minimized by filtering the output of the tape-recorded at the maximum frequency of interest.

2. Drop-outs in the clock channel.

   A "clock filter" has been made to reduce this source of error.

3. Noise and stray pulses in the start/stop channel.

   A start-stop pulse "filter" has been developed
to overcome this source of error.

These faults may be mainly attributed to (see [7] and [26]).

a. The nature of the magnetic recording process. The finiteness of the magnetic oxide particle size and inherent noise associated with magnetic and electronic circuits create a certain amount of noise. A further source of noise is wow and flutter in the transport mechanism which shifts the frequency and hence gives a noise voltage in the F.M. mode of operation. The error due to noise is generally low and is of the order of 1% for the Ampex SP300 recorder.

b. Imperfections in the tape materials resulting in portions of tape having gaps in the oxide coating or portions with unduly rough surfaces. The signal does not get recorded faithfully because of the absence of the oxides or separation of the tape from the head.

c. Dust and oxide accumulation on the tape and particularly on the heads. In extreme cases, especially with tape loops, oxide may "cake" on the guides and wear a groove on the tape. Regular cleaning of the recorder will minimize this. In this respect, a felt pad to clean the tape just before it comes to the recording head is of great value. However, it does not seem to
Unless otherwise stated, all resistors in kΩ, all capacitors in pf.
be possible to eliminate this effect completely.

d. Imperfections in any tape joints. It is not possible to eliminate joints if an endless loop has to be used. The tape gets a jerk as the joint passes the tape guides and this gives rise to voltage spikes being picked up on the heads. To minimize this effect regions of the tape near the joint are not used directly.

e. Crinkling of tape. This stresses the importance of careful handling of tape; even a slight crinkle causes "drop-outs".

7.22. Clock Filtering:

Since the clock signal is essential for timing in the converter system, any interruptions due to drop-outs or stray pulses will cause errors. To overcome these effects a clock filter has been developed. The requirements of the filter are:

1. The clock frequency should be followed precisely when the clock signals are present. Any normal variations in frequency, due to wow or flutter or mains frequency variations for example, should be followed.

2. When there are clock pulses missing, they should be inserted as close as possible to the true position.

3. The circuit should be capable of working auto-
matically over a range of frequencies so that external switching is minimized.

The circuit devised is shown schematically in Fig. 7.22. The clock frequency from the tape recorder is integrated to get a signal whose amplitude is inversely proportional to the frequency of the input. This voltage is rectified and smoothed in a parallel circuit (R₂C) charged through a series diode and resistor R₁. The parallel circuit charges to the peak value of the voltage, provided that this has been steady for a time corresponding to the time constant R₁C. This prevents drop outs being interpreted as a low frequency clock signal.

The output voltage of the parallel circuit is changed to a corresponding current, which feeds a current controlled oscillator. The current magnitude and therefore the output frequency are adjustable. Thus an output following the input frequency is obtained over a range of frequencies.

The output frequency f-Δ of the oscillator, when it is not synchronized, is adjusted to be just below the input frequency f. Normally the oscillator is synchronized to the input frequency so that the system operates as a locked oscillator. However, in the absence of the input signal, the oscillator free runs at a frequency which slowly drops from f to f-Δ. Thus a pulse is filled in at approximately
the correct place.

$f - \Delta$ is made as close to $f$ as the normal variations in the input frequency allow. The difference is kept small so that a number of pulses can be filled in, should this be necessary.

7.23 Start/Stop Filtering:

The start/stop signal used for identifying the portion of the tape to be analysed is ideally a voltage level changing only at the start and stop points. However, due to causes referred to in section 7.24, the ideal signal is not recovered on playback and some form of filtering is necessary.

The start position is fairly critical as all the timing is derived from this point. For this reason the start signal is taken as a negative to positive going transition, as the negative level is relatively immune from interference. The positive level is filtered by means of a "diode filter" as shown in Fig. 7.23. A positive voltage level quickly charges the capacitor. Any leakage from the capacitor due to negative going excursions of short duration caused by dropouts is restored at the end of the dropout so that the voltage on the capacitor is held close to the maximum. When the voltage is held negative for a sufficient length of time, the capacitor discharges sufficiently to enable a Schmidt trigger to fire. This firing defines the stop point.
FIG. 7.23. START/STOP FILTER.

Comp. = Comparator.
Sense. = Sensitivity Control.

FIG. 7.3. PEAK DETECTOR.
The capacitance is adjusted to have a discharge time constant of about 100msecs, which is sufficient to overcome normal drop out effects.

However, the device has draw-backs. If positive pulses occur while the signal level is supposed to be held negative, or "stray drop outs" of appreciable duration occur just before the negative going change of level causing the capacitor to start discharging early, there can be some erroneous effects. By "stray drop outs" is meant drop outs not permanently on tape, but due to dust or oxide accumulation. Positive spikes sometimes seem to appear when the tape joint "jerks" past the guides; this effect has been removed by blanking the start/stop pulse for the time the tape joint is moving past the guides. Blanking times of about 1, 2, 4 and 8 secs are available for the various speeds.

7.3 Peak Detector Unit (Fig. 7.3).

The input range of the ADC is 0 to -10V. In order to utilize the full capacity of the ADC, it is necessary to have the analogue voltage scaled (and offset if necessary) to have its most positive value close to and less than 0V as well as its most negative value close to and greater than -10V. If the amplitude is smaller, the effective quantization level of the signal increases. If the range is exceeded, clipping occurs. Even using an oscilloscope
\[\text{D} = \text{Driver}, \ \text{E} = \text{Exclusive or.}\]

**FIG. 7.4. DECIMAL TO IBM-CODE CONVERSION.**
it is difficult to ensure that the signals have been amplified to the correct range particularly with random signals. To help in this respect the peak-detector unit has been developed.

The peak detector consists of two meters to indicate the maximum and minimum points. To achieve the necessary accuracy, the meters only operate when the signal is close to the peak values they are set to detect, at which time their sensitivity is 120mV full scale. Further, to enable reading of relatively short peaks, the meter circuits have a large discharge to charge time constant ratio.

The range of values about which the peak can be detected is -3V to +1.5V for the positive peak detector and -7V to -11V for the negative peak detector.

A single cycle of a 2.5kHz sinewave can be detected within 250mV of the static calibration point.

7.4. Adaptor for a Paper Tape Punch

In anticipation of a paper tape punch being purchased by the Department of Electrical Engineering, a conversion unit for operating a tape punch in IBM code was designed and is shown in Fig. 7.4.

Some of the details of the circuit will depend on the particular punch. The main purpose of the diagram is to show
the coding and parity check circuits as well as the general schematic.

The decimal inputs are coded by a system of gates to give the 1248 IBM BCD code. The parity check circuit consists of a number of comparators. The outputs to the punch are compared in pairs; the resulting outputs are again compared in pairs so that an equal output is obtained should there be an even number of bits energized, in which case the check bit is energized.

The time for punching is determined by sensing if any of the units, tens or hundreds lines has a digit to be punched, in which case the appropriate power of ten terminal has a pulse on it. This triggers a single shot delay which enables the input, comparators and gates to settle. Punching then takes place by energizing the punch drivers for a sufficient length of time determined by means of another single shot. The digits corresponding to the input digit, as well as the X (flag digit) if present, are thus energized.
CHAPTER 8

CONCLUSIONS

In this concluding chapter, the main results of the thesis are summarized and some possible avenues of further research are indicated.

In Chapter 3, it was shown that, given a process \( \{X(\omega, t), t \in \mathbb{R}\} \) satisfying certain second order properties and having a first order p.d.f., \( f \) satisfying certain smoothness assumptions, then it is possible to estimate \( f \) to within a desired accuracy from a finite length \( T \) of a single sample function \( x(t) = X(\omega_0, t) \) from the process. These results are summarized in greater detail in section 3.2.

The results should easily extend to the determination of second order p.d.f.s, but a corresponding increase in complexity would be expected. For example, the determination of \( \theta_{T,j} \) in equation 3.3-7 will now correspond to the determination of \( \theta_{T,j,k} \), the estimate of the probability that the signal takes on two given values at two given times; that is, with \( A_j \) and \( B_j(t) \) as before, \( \theta_{T,j,k} \) is the estimate of the probability \( m_{j,k} = P[B_j(t) \cap B_k(t + \tau)] \). \( \theta_{T,j,k} \) will again be given by an equation such as 3.3-7, except that \( z_j(t) \) becomes \( z_{j,k}(t) = z_j(t)z_k(t + \tau) \) and the signal is assumed known for a length of time \( T + \tau \) rather than \( T \). The expectation and
variance of $\theta_{T,j,k}$ will depend on the second and fourth order properties of the process. The Lemmas 1 to 3 would be replaced with corresponding Lemmas dealing with interpolation in two dimensions. The question of the choice of $\tau$ to represent the family of second order p.d.f.s $f(x_1,x_2,\tau)$ may also need to be considered.

Some other interesting problems that may prove worthwhile investigating are:

1) The relationship between the concepts of ergodicity and "finite memory" as defined in definition 3.1-1. Whereas ergodicity is the commonly used concept, the author feels that finite memory is intuitively more meaningful.

2) The relationship between (i) the random variables $X(\omega,t)$ and $X(\omega,t+\tau)$ being independent and (ii) the correlation function $R(\tau) = \mathbb{E}[X(\omega,t)X(\omega,t+\tau)]$ being zero. Although $R(\tau) = 0$ merely says that the random variables are uncorrelated, one might expect that $R(\tau)$ being zero for some interval $[a,b)$, particularly if $b = \infty$, implies something more than just zero correlation.

3) The method of determining $\hat{\tau}_{j}$, the estimate of $\theta_{T,j}$ in Lemmas 10 and 11. $\hat{\tau}_{j}$ is found there by means of simple random sampling. Improved results would of course be obtained if more sophisticated sampling methods are employed.
4) Extension of some existing papers. As mentioned in Chapter 3, a number of papers have indicated how the p.d.f. may be determined in the form of a sum of orthogonal functions, using a sequence of independent points from the process. However, it is not clear from these papers when and how this sequence may be obtained from a finite length of a single sample function. The earlier results of Chapter 3 may be used to show that under certain stated conditions, it is possible to construct a process with a p.d.f. that can be made close to that of the required process by choice of T. It is then possible to obtain an independent sequence of points from this constructed process. Thus it is likely that the results of the papers referred to above can be extended to determining the p.d.f. of a process in the form of the sum of a set of orthogonal functions, given a sufficient length of a single sample function from the process.

In Chapter 4, some of the more common methods of estimating the power spectrum of a random process were briefly compared. Further, the problems often encountered in finding the power spectrum using the correlation function were discussed, along with some comments on suggested solutions to these problems. Some of the results are, and may remain, based on experience and it is hoped that the author's comments will be of use to those attempting to determine
power spectrum using the correlation function. The author also hopes to extend the results on the simple bounds found in section 4.4.

In Chapter 5, the general theory of interpolation was briefly presented and its relationship to some existing interpolation formulae was shown. This was followed by a discussion on errors arising in interpolation in a wide sense - the errors in acquiring the values used in any interpolation formulae and the errors such as truncation errors arising in practical approximations to the formulae. This discussion is in the form of a review article and is intended to make one aware of the kind of problems that one is likely to meet with and to point out ways of going about overcoming these problems. In particular, some results on truncation errors by various authors have been given and these modified to give lower bounds. Finally the author deals with the problem of interpolation using areas, that is, instead of using values of the function at regularly spaced intervals, areas between these intervals are used. First it is shown that there exists a solution using polynomial interpolation. However, no solution has yet been found in this case and the finding of a solution could provide an interesting extension. A solution to the problem in terms of entire functions is found, and this should prove more useful in electrical
engineering uses since a band-limited signal is entire. The author also proves that if the function \( f \) being interpolated is band-limited, then the interpolation formula can be made to converge to \( f \). If \( f \) is not band-limited, the aliasing error that occurs is also found, but further work could be done as regards truncation errors and errors caused by inaccuracies in the data points. The latter error has been calculated for a special case in Chapter 3.

The hardware described in Chapters 6 and 7 has performed satisfactorily as an interim solution to the data acquisition problem. However, with improved facilities in the department, it is now being modified to convert data more directly.
List of References

31. DANIELS, H.E. "SADDLE POINT APPROXIMATION IN STATISTICS." *ANNALS OF MATHEMATICAL STATISTICS,* V25, NO DECEMBER 1954.
34. DIGITAL EQUIPMENT CORPORATION. "THE DIGITAL EQUIPMENT HANDBOOK." 1966-67 EDITION.
35. DODD, J.J. "STOCHASTIC PROCESSES." JOHN WILEY, 1953.
38. FREUND, R.J. "MATHEMATICAL STATISTICS." PRENTICE HALL, 1962.
40. GOODMAN, N.R. "SOME COMMENTS ON SPECTRAL ANALYSIS OF TIME SERIES." *TECHNOMETRICS,* V3, PP.221-228, MAY 1961.
43. GRENDEL, R. AND ROZENBLATT, M. "STOCHASTIC ANALYSIS OF STATIONARY TIME SERIES." JOHN WILEY, 1957.
44. GUASS, E.J. "ESTIMATION OF POWER SPECTRAL DENSITY BY FILTERS." *JOURNAL OF THE ASSOCIATION FOR COMPUTING MACHINERY,* VII, NO. 1, JANUARY 1964.
45. HAKOS, P.R. "MEASURE THEORY." VAN NOSTRAND, 1963.
47. HELMHOLTZ, H. AND THOMAS, J.B. "TRUNCATION ERROR OF SAMPLING-THEOREM EXPANSIONS." BELL TELEPHONE SYSTEM TECHNICAL PUBLICATIONS, MONOGRAFIES 4106.
50. HALE, P.R. "MEASURE THEORY." VAN NOSTRAND, 1963.
52. HELMHOLTZ, H. AND THOMAS, J.B. "TRUNCATION ERROR OF SAMPLING-THEOREM EXPANSIONS." BELL TELEPHONE SYSTEM TECHNICAL PUBLICATIONS, MONOGRAFIES 4106.
55. HALE, P.R. "MEASURE THEORY." VAN NOSTRAND, 1963.


81. Sampford, M. 'Sampling Theory.' *Oliver and Boyd,* 1962.


86. Slud, H. V. V. 'Introduction to the Theory of Statistical Dynamics.' *Dover,* 1967.


98. THOMAS, J.B. AND LIU, R. 'ERROR PROBLEMS IN SAMPLING REPRESENTATIONS.' IEEE INTERNATIONAL CONVENTION RECORD. PT. 5 V12 PP69-2771960.
100. TODD, J. (ED.) 'SURVEY OF NUMERICAL ANALYSIS.' MCGRAW-HILL, 1962.
101. TSYPKIN, Y.A. 'USE OF THE STOCHASTIC APPROXIMATION METHOD IN ESTIMATING UNKNOWN DISTRIBUTION DENSITIES FROM OBSERVATIONS.' AUTOMATIKA.
111. WHITTAKER, J.H. 'THE FOURIER INTEGRAL AND CERTAIN OF ITS APPLICATIONS.' DODGE, 1933.
112. WHITTAKER, E.J. 'THE FOURIER INTEGRAL AND CERTAIN OF ITS APPLICATIONS.' DOVER, 1933.
113. WHITTAKER, J.M. 'INTERPOLATORY FUNCTION THEORY.' CAMBRIDGE TRACTS IN MATHEMATICS AND MATHEMATICAL PHYSICS N33. CAMBRIDGE UNIVERSITY PRESS, 1935.
114. WHYNOTH, A. 'A STUDY OF ROUGH AMPLITUDE QUANTIZATION BY MEANS OF NYQUIST SAMPLING THEORY.' IEEE TRANS. ON CIRCUIT THEORY VCT-3 N4 PP266-276 DECEMBER 1956.
115. WIENER, N. 'EXTRAPOLATION, INTERPOLATION AND SMOOTHING OF STATIONARY TIME SERIES.' M.I.T. PRESS (PAPERBACK EDITION, 1964)
116. WIENER, N. 'THE FOURIER INTEGRAL AND CERTAIN OF ITS APPLICATIONS.' DOVER, 1933.
118. WITTAKER, E.T. AND WATSON, G.N. 'A COURSE OF MODERN ANALYSIS.' 4TH EDITION CAMBRIDGE, 1933.
List of Special Symbols

The following is a list of special symbols used in the thesis. The meaning of the other symbols used should be clear from the context.

\( \in \) \quad \text{Belongs to}

\( \rightarrow \) \quad \text{Such that}

\( \forall \) \quad \text{For all}

\( \exists \) \quad \text{There exists}

\( \S \) \quad \text{Section}

\( \diamond \) \quad \text{End of proof}

\( j \) \quad \sqrt{-1}

\( R \) \quad \text{The set of real numbers (the real line)}

\( a \implies b \) \quad \text{a implies b}

\( t, \tau \) \quad \text{Time variable}

\( f \) \quad \text{Frequency}

\( w \) \quad \text{Angular frequency}

\( T, T_a, \text{ etc.} \) \quad \text{Fixed period of time}

\([a,b],(a,b],
\quad [a,b),(a,b)\) \quad \text{Intervals: closed, left open, right open, open}

\( |a_{i,j}| \) \quad \text{Determinant with elements } a_{i,j}

\( [a_{i,j}] \) \quad \text{Matrix with elements } a_{i,j}

\( \{a_n\} \) \quad \text{Sequence of points or numbers, } a_1, a_2, ..., a_n, ...

\( f \in L_p(a,b) \text{ or } L_p \)

\( \text{if } a = b = \infty \) \quad \text{if } a \neq b

\( \int_a^b f(x) \, dx \text{ exists} \)
\[ \{s_n\} \in \ell_p \quad \{s_n\} \Rightarrow \sum_{n=1}^{\infty} s_n^p \text{ exists} \]

\[ \phi(u) = \phi_f(u) \quad \text{Fourier transform of } f(t) \]

\[ B[-A,A] \quad \text{Set of functions band limited to } A, \text{ i.e.} \]

\[ \{f \Rightarrow \phi_f(u) = 0 \text{ for } |u| > A, \text{ a constant}\} \]

\[ B_e[-A,A] \quad \text{Set of functions essentially band-limited to } A. \{f \Rightarrow \int_0^{|\phi(u)|du} \text{ is small} \} \]

\[ |u| > A \]

\[ B,B_e \quad \text{B[-A,A] and } B_e[-A,A] \text{ when } A \text{ is arbitrary} \]

\[ f^*(x) \quad \text{Conjugate function of } f(x) \]

\[ x(t) = X(\omega,t) \quad \text{Random signal, sample function} \]

\[ X(\omega,t_0) = x, (t \text{ fixed}) \quad \text{Random variable (r.v.)} \]

\[ \{X(\omega,t), t \in \mathbb{R}\} \quad \text{Random process, (ensemble of functions)} \]

\[ \overline{x} = E[x] \quad \text{Mean of (a r.v.) } x \]

\[ \text{Var}[x] = \sigma_x^2 \quad \text{Variance of } x \]

\[ R(\tau) = R_X(\tau) \quad \text{Correlation function (continuous) of } x \]

\[ R(m) \quad \text{" " discrete " "} \]

\[ S(w) \quad \text{Power spectrum} \]

\[ x_T(t) = \text{Truncated form of } x(t) \]

\[ R_T(\tau), S_T(w), \overline{x}_T \text{ etc.} \quad \text{Estimates from finite length } T \text{ of } x(t) \]

\[ R_M(\tau), S_M(w) \text{ etc.} \quad \text{Estimates from finite number of terms } M \]

\[ K(w,t) \quad \text{Kernel or window} \]

\[ \text{Sinc}(x) \quad \text{Sin}(x)/x \]

\[ f(x) \quad \text{Arbitrary function, probability density function (p.d.f.) in chapter 3} \]

\[ f'(x), f''(x), f'''(x), \ldots, f^n(x) \quad \text{Derivatives of } f(x); 1\text{st, 2nd, 3rd, nth.} \]
\[ P(B_j) \] Probability measure of a set \( B_j \)

\[ (\Omega, \mathcal{S}, P) \] Probability space

\[ E[X] = \int X \, dP \] Expectation of \( X \)

\[ F(x), F_X \text{ etc.} \] Cumulative distribution function (c.d.f.)

\[ \hat{f}, \hat{m} \text{ etc.} \] Estimates of \( f, m \) etc.

\[ \cos(v), \text{ etc.} \] Arbitrary positive numbers \( > 0 \) (errors)

\[ \epsilon, \epsilon_a \text{ etc.} \] Arbitrary positive no. \( \geq 0 \) and \( < 1 \)

\[ \alpha, \alpha_a \text{ etc.} \] Support or essential support of \( f \)

\[ S \] No. of intervals into which \( S \) is partitioned

\[ A_j \] The interval \( [x_{j-1}, x_j] \subset S \)

\[ B_j(t) = B_j \] The set \( \{ \omega : X(\omega, t) \in A_j \} \)

\[ z_j(t) = z_j \] Indicator function of \( B_j(t) \)

\[ m_j \] \( P(B_j) \)

\[ \theta_{T,j} \] Finite length estimate of \( m_j \)

\[ \delta \int_0^T z_j(t) \, dt \]

\[ D \] \( \sup_{x \in S} |f'(x)| \)

\[ E \] \( \sup_{x \in S} |f''(x)| \)

\[ Z(1 - \alpha/2) \] Ordinate of \( \{ 1/\sqrt{2\pi} \} \exp(-x^2/2) \) with area of \( \alpha \) in the tails

\[ \sim \] Of the order of

\[ \mathcal{F} \] Set of functions

\[ L_{1,j}(f) \] Linear functional

\[ i \in [j, k] \] \( i = j, j+1, j+2, \ldots, k-2, k-1, k \)

\[ (\cdot)_P \] Periodic extension