A NEW APPROACH TO THE ANALYSIS OF THE THIRD HEART SOUND

A thesis completing the requirements for the degree of

MASTER OF SCIENCE

in

The Department of Applied Mathematics,
The University of Adelaide, South Australia.

by

Gary John Ewing, B.App.Sc.(Biophysics)

October 1988
Contents

List of figures ................................................................. iv
List of tables ................................................................. vi
Declaration ................................................................. vii
Acknowledgements ............................................................ viii
Summary ................................................................. ix
Publications and Abstracts .................................................. x

Chapter 1 GENERAL INTRODUCTION ...................................... 1

1.1 Basic Introduction to the Heart and the Heart Sounds....................... 1

1.2 Heart Sounds and The Cardiac Cycle.......................................... 2

1.2.1 First Heart Sound....................................................... 6

1.2.2 Second Heart Sound..................................................... 6

1.2.3 Third Heart Sound....................................................... 7

1.2.4 Fourth Heart Sound...................................................... 8

1.3 Diagnostic Implications of the Third Heart Sound........................... 8

1.4 Scope of the Research....................................................... 10

Chapter 2 THE INSTRUMENTATION AND DATA COLLECTION ................. 11

2.1 Phonocardiography an Extension of Auscultation............................. 11

2.2 Phonocardiogram Data Aquisition............................................ 12
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3</td>
<td>Data Storage.</td>
<td>15</td>
</tr>
<tr>
<td>2.4</td>
<td>Introduction to Echocardiography.</td>
<td>17</td>
</tr>
<tr>
<td>2.5</td>
<td>2-Dimensional Echocardiography.</td>
<td>19</td>
</tr>
<tr>
<td>2.6</td>
<td>Echocardiographic Measurements.</td>
<td>25</td>
</tr>
<tr>
<td>3</td>
<td>ANALYSIS OF THE THIRD HEART SOUND BY THE FAST FOURIER TRANSFORM METHOD.</td>
<td>27</td>
</tr>
<tr>
<td>3.1</td>
<td>The Fast Fourier Transform.</td>
<td>27</td>
</tr>
<tr>
<td>3.2</td>
<td>The Fast Fourier Transform in Heart Sound Analysis.</td>
<td>28</td>
</tr>
<tr>
<td>3.3</td>
<td>Derived Spectral Parameters.</td>
<td>31</td>
</tr>
<tr>
<td>3.4</td>
<td>Correlation of Spectral and Echocardiographic Data.</td>
<td>32</td>
</tr>
<tr>
<td>3.5</td>
<td>Discussion.</td>
<td>36</td>
</tr>
<tr>
<td>3.6</td>
<td>Limitations.</td>
<td>38</td>
</tr>
<tr>
<td>4</td>
<td>MAXIMUM ENTROPY SPECTRAL ANALYSIS</td>
<td>41</td>
</tr>
<tr>
<td>4.1</td>
<td>Entropy</td>
<td>43</td>
</tr>
<tr>
<td>4.2</td>
<td>Linear Digital Filtering</td>
<td>43</td>
</tr>
<tr>
<td>4.3</td>
<td>Weiner Filtering</td>
<td>45</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Mean Square Error.</td>
<td>46</td>
</tr>
<tr>
<td>4.3.2</td>
<td>The Prediction Error Filter.</td>
<td>48</td>
</tr>
<tr>
<td>4.4</td>
<td>Maximum Entropy Spectral Analysis.</td>
<td>51</td>
</tr>
<tr>
<td>4.5</td>
<td>Algorithm to Solve the Prediction Error Filter Equations.</td>
<td>55</td>
</tr>
<tr>
<td>4.5.1</td>
<td>Order Selection.</td>
<td>58</td>
</tr>
<tr>
<td>4.5.2</td>
<td>Computational Load.</td>
<td>59</td>
</tr>
<tr>
<td>5</td>
<td>MAXIMUM ENTROPY SPECTRAL ANALYSIS OF THE THIRD HEART SOUND</td>
<td>62</td>
</tr>
<tr>
<td>5.1</td>
<td>Numerical Experiments with Synthetic Data.</td>
<td>63</td>
</tr>
</tbody>
</table>
5.1.1 Model Order Estimation. ........................................ 63
5.1.2 Experiments with Synthetic Data. ............................. 65
5.2 Comparison of MESA and FFT Spectral Analysis of S3. .......... 68
  5.2.1 Signal Preprocessing. ........................................... 68
  5.2.2 Fourier Transform Method. ..................................... 74
  5.2.3 Maximum Entropy Method. ...................................... 74
  5.2.4 Comparison of the FFT MESA Techniques. .................. 78
5.3 Correlation of MESA and Echocardiographic Data. ............. 78
5.4 Discussion. .......................................................... 78

Chapter 6 ARMA MODELLING:
RECOMMENDATIONS FOR FURTHER WORK .......................... 81
  6.1 Auto-regressive Moving Average (ARMA) Modelling. ........ 81
  6.2 ARMA Model From The Maximum Entropy (AR) Model. ....... 82
     6.2.1 Homomorphic Systems. ...................................... 82
     6.2.2 The Complex Cepstrum. ..................................... 85
     6.2.3 Maximum Entropy ARMA Spectral Estimate. ............... 86
     6.2.4 ARMA Spectrum from AR Coefficients. .................... 88
  6.3 Deconvolution of the Chest Wall Transfer Function. .......... 90
  6.4 Conclusions and Recommendations for Further Work. .......... 91

A The Entropy of a Gaussian Sequence. ................................ 93

B Theorem Due to Szëgo. .................................................. 95

C Spectral Factorization. ................................................ 97

BIBLIOGRAPHY ............................................................... 99
List of Figures

1.1 Functional diagram of the circulatory system. ............................ 3
1.2 Electrical activity of the Heart. .............................................. 5
1.3 Relationship of the Heart Sounds to other cardiac events. ............... 9

2.1 PCG Data Acquisition System .................................................. 13
2.2 Lead ii ECG ................................................................. 14
2.3 Digitized ECG signal ........................................................... 16
2.4 Digitized PCG signal ........................................................... 16
2.5 Ultrasound Wave Propogation ............................................... 18
2.6 Propogation, Reflection and Reflection of Ultrasound. ................. 18
2.7 M-mode Echocardiogram ....................................................... 20
2.8 2-D Echocardiogram ............................................................ 20
2.9 Beam Steering ................................................................. 22
2.10 Long Axis View ............................................................... 23
2.11 Short axis view ............................................................... 23
2.12 Apical two-chamber view ................................................... 24
2.13 Apical four-chamber view ................................................... 24

3.1 Process in obtaining FFT spectra ............................................. 29
3.2 A typical FFT spectrum ....................................................... 30
3.3 Histogram of S3 energy spectral distribution ............................ 34
List of Tables

3.1 Details of the subjects. .............................................. 33
3.2 Position of window for S3 in cardiac cycle. .................... 33
3.3 % energy per frequency band. .................................... 35
3.4 Echo & Phono Derived Parameters ................................. 35
3.5 Correlation coefficients for parameters vs freq bands .......... 40
3.6 Correlation coefficients for age vs some parameters. .......... 40
5.1 Freq peaks in the FFT spectra of S3 .............................. 75
5.2 Details of subjects. ................................................. 75
5.3 MESA peaks for all subjects. ..................................... 77
5.4 Correlation coefficients. ......................................... 77
Declaration

I declare that this thesis is a record of original work and that it contains no material which has been accepted for the award of any other degree or diploma in any University.

To the best of my knowledge and belief, this thesis contains no material previously published or written by any other person, except where due reference is given in the text of the thesis.

I consent to this thesis being made available for photocopying or loan.

Gary J. Ewing  
1988
Acknowledgements

I wish to express thanks to my supervisor, Dr. J. Mazumdar who has lent the much needed encouragement.

I also wish to thank my friend and colleague Dr. N. Fazzalari (Nick) for his collaboration with respect to 2-D echocardiographic data collection, and our many discussions. My thanks are also extended to Dr. E. Goldblatt and the Department of Cardiology of the Adelaide Children's Hospital for the extensive co-operation and assistance offered.

Last but not least my thanks go to Dr. E. Van vollenhoven of Leiden University, (The Netherlands), for his many useful comments and his presentation at European conferences of the results of this work.

The work of part of this thesis was supported by a grant from the Channel 10 Children’s Medical Research Foundation of South Australia and an URG grant from the University of Adelaide.
Summary

There has been in the past and still is controversy over the genesis of the third heart sound (S3). Recent studies, strongly suggest that S3 is a manifestation of a sudden intrinsic limitation in the expansion of the left ventricle. The thesis has aimed to explore that hypothesis further using combined echocardiographic and spectral analysis techniques. Spectral analysis was carried out via conventional fast fourier transform methods and the maximum entropy method. The efficacy of these techniques, in relation to clinical and scientific application, was explored further. Briefly discussed was the application of autoregressive-moving average (ARMA) modelling for spectral analysis of S3, in relation to further work. Following is a brief synopsis of the thesis:

CHAPTER 1 This gives an historical and general introduction to heart sound analysis. Discusses briefly is the physiology of the heart and heart sounds and the diagnostic implications of S3 analysis.

CHAPTER 2 Here is discussed the instrumentation system used and phonocardiographic and echocardiographic data acquisition. Data preprocessing and storage is also covered.

CHAPTER 3 In this chapter the application of a FFT method and correlation of resultant spectral parameters with echocardiographic parameters is reported.

CHAPTER 4 The theoretical development of the maximum entropy technique (based on published papers and expanded) is discussed here. Numerical experiments with the method and associated problems are also discussed.

CHAPTER 5 The MEM is applied to the spectral analysis of S3 and compared with the FFT method. Correlation analysis of MEM derived spectral parameters with echocardiographic data is performed.

CHAPTER 6 Here ARMA modelling and application to further work is discussed. An ARMA model from the maximum entropy coefficients is derived. The application of this model to the deconvolution of the chest wall transfer function is discussed as an approach for further work.
Publications and Abstracts

During the course of my study the following papers and abstracts have been published, or have been presented at learned societies.

(i) An oral presentation “A Non-invasive Study of the Third Heart sound in Children by Phono and Echocardiography”

(ii) “A Spectral Analysis of the Third Heart Sound in Children”
Authors: G. Ewing, J. Mazumdar, N. Fazzalari, E. Goldblatt

(iii) An invited paper of the same title and authors as in (i) was published in Acta Cardiologica vol.39(4):241-254, 1984

(iv) “Use of the Maximum Entropy Method for Spectral Analysis of the Third Heart Sound in Children”

Australasian Physical & Engineering Sciences in Medicine (1986) vol. 9 No. 3

Publications in Preparation

“From AR to ARMA Spectra for Heart Sound Analysis.”
Authors: G. Ewing and J. Mazundar
To be presented at the IEEE 1989 Australian Symposium on Signal Processing and Applications. 17-19 April, Adelaide, Australia.
Note: In this paper some of the techniques discussed in Chapter 6 will be applied to heart sound data.
Chapter 1

GENERAL INTRODUCTION

1.1 Basic Introduction to the Heart and the Heart Sounds.

The primary role of the heart is to serve as a muscular pump propelling blood around the circulatory system (see figure 1.1). The arteries, which receive this blood at high pressure and velocity conduct it throughout the body, are thickly walled with elastic fibrous tissue and and a wrapping of muscle cells. The arterial tree terminates in short, narrow, vessels called arterioles, from which blood enters simple tubes known as capillaries. From the capillaries, the blood, now depleted of oxygen and nutrients and burdened with waste products, moving more slowly and under low pressure, enters small vessels called venules, which converge to form veins, ultimately guiding the blood back to the heart.

The heart is a four-chambered, hollow, muscular organ lying between the lungs in the middle mediastinum (space between the plura of the lungs). The heart is shaped like an inverted cone with it's apex pointed downwards.

The auscultation of the heart may reveal different phenomena called heart sounds and murmurs. The heart sounds are prolonged series of vibrations of both high and low frequency; owing to their complex nature and the existence of recognizable components, they should be called sound complexes [Luisada 1948].
The murmurs are longer series of vibrations, which may be mostly of either high or low frequency. However, as occasionally a murmur can be shorter than a sound complex, the distinction between the two is based largely on tradition and is usually decided on individual cases, largely on the basis of timing.

The heart sounds result from the interplay of the dynamic events associated with the heart beat and blood flow. During each phase of the cardiac cycle, the main directional mass movement of blood through the cavities of the heart and great vessels is determined and maintained as a function of both the muscular and the valvular apparatus. The interdependent work of these two systems is influenced to a great extent by the properties of the heart wall.

The activity of the haemodynamic system results in periodic vibrations emanating from the pulsating heart and vessels. The energy produced is transmitted to the chest wall, where these vibrations are detected by auscultation and phonocardiography, as heart sounds.

1.2 Heart Sounds and The Cardiac Cycle.

Clinicians and research workers usually base their considerations on the well known scheme of the cardiac cycle. This concept was developed early this century [Wiggers 1915]. Since then many investigations have been made into the quantitative and qualitative aspects of the cardiac cycle. Earlier measurements however, were hampered by crude technology and ignorance in the application of this. For instance, early cardiac pressure measurements would have been performed using manometers with poor high frequency response. Even in the early fifties, strain gauges were used with inadequate high frequency responses. In general the response of the measuring systems, including catheter systems, and their effects on the measurements were ignored.

The following is a basic description of the events occurring during the cardiac cycle. Let the description start when the heart is in diastole (relaxing). At the beginning of the diastolic interval, all dimensions of the ventricular chamber increase rapidly. This
Figure 1.1: Functional diagram of the circulatory system.
phase of rapid diastolic filling is very brief and merges abruptly or gradually into the phase of slow diastolic filling which persists until atrial contraction ensues. When the ventricles are maximally distended, the dimensions reach a plateau at the end of the rapid filling phase and change minimally during the remainder of diastole. During this period of constant ventricular volume the ventricles are in a state of diastasis. The diastolic interval ends with the onset of atrial systole (contraction).

Atrial systole is usually initiated by a wave of electrical activity (depolarisation of myocardium), emanating from the sino-atrial (S-A) node (see figure 1.2). As the wave of contraction spreads over the atrium, atrial systole occurs. Contraction of the atrial musculature reduces the capacity of the atrial chambers and displaces blood forward into the ventricles or backward into the great veins, depending on which course offers the least resistance.

As the wave of excitation extends rapidly along the Purkinje system (see figure 1.2), and spreads over the endocardial surface of the ventricles, muscular contraction occurs and the atrio-ventricular valves close. Until the pressures becomes high enough to open the semilunar valves (the valves feeding into the main arteries), all four valves are shut. Hence the contracting muscles elevate the pressure inside the ventricles without changing their volumes. This is the period of isovolumic contraction. At the onset of systole, the length of the ventricles is abruptly shortened as the atrioventricular diaphragm rapidly descends. The diameter, circumference and external length of the ventricle simultaneously expand. As the full thickness of the ventricular walls becomes excited, pressure in each ventricle exceeds the corresponding arterial pressure and blood is very rapidly ejected from the ventricles. The associated reduction in the ventricular volumes is rapid during early systole and slowing during the last part of systole.

In man (and all mammals) two heart sounds are consistently heard by auscultation. Phonocardiographic tracings often reveal up to four sounds. The sounds heard during auscultation are called the first (S1) and the second (S2) heart sounds respectively, with respect to their temporal relationship, and are systolic sounds. Phonocardiography
Figure 1.2: Electrical activity of the Heart.
often yields third (S3) and fourth (S4) heart sounds, especially in children and in cases of heart disease. These are diastolic sounds.

Many hypotheses have been put forward to explain the origin of the heart sounds, some being controversial at the time [Potain, 1900; Luisada, 1948; Smith et al, 1950; Sabbah and Stein, 1976; Stein and Sabbah, 1978; Ionescu and Stonescu, 1980]. Much was contributed to these controversies by the inadequacies of the instrumentation of the past, including the introduction of delays in the monitoring of intracardiac pressures. However with the advent of echocardiography (ultrasound imaging of the heart, which will be discussed later), the movement of intracardiac structures could be monitored with virtually no time delay. Concerning the S1, S2 and probably S4 these controversies have largely been resolved; however there still exists some controversy with respect to S3. The genesis of these sounds will now be discussed in more detail.

1.2.1 First Heart Sound.

S1 occurs during early ventricular systole, at which time blood is accelerated in the ventricles, surging towards the atrioventricular (AV) valves. Laniado (1973) demonstrated that the AV valves close after the crossover point when the ventricular pressure exceeds the atrial pressure. The closure of the mitral and tricuspid valves coincides with two major components of the first sound. It has also been shown recently [Luisada, 1983] that the mitral valve and its associated chordae contribute small fractions of energy to S1; and a good correlation between ventricular wall tension and S1 energy has been shown.

1.2.2 Second Heart Sound.

This sound is apparently associated with the vibrations of aortic and pulmonary valves just after closure and probably the ascending aorta. It has been recently pointed out by Stein and Sabbah, in their works on the second heart sound production mechanisms (1978), that even though coaptation of the leaflet is silent the rapid vibrations of
the closed leaflets, that begin immediately after coaptation, create the sound. This hypothesis has been supported by other research workers [Anastassiades et al, 1976; Kotler et al, 1978].

1.2.3 Third Heart Sound.

This is a low frequency sound occurring during early diastole, during the rapid filling phase of the ventricle. It occurs from 0.13 to 0.20 seconds after S2. Due to its low intensity and low pitch, S3 is not commonly detected via auscultation. There has been controversy over the genesis of the third heart sound since the beginning of this century [Potain, 1900]. Consequently various hypotheses have been put forward in an effort to explain the mechanisms generating S3. These include:

1. S3 originates in vibrations of the mitral valve and associated structures [Dock et al, 1955; Fleming, 1969];

2. S3 results from the impact of the heart against the chest wall [Reddy et al, 1981].

3. S3 is a consequence of the termination of the rapid filling at the moment the expanding ventricular wall reaches the limit of its passive distensibility. [Potain, 1900; Kuo et al, 1957; Craige et al, 1983].

The hypothesis suggesting that S3 is caused by vibrations of the mitral valve and associated apparatus, seems to have been discounted on the basis of various meticulous studies including that of El Gamal and Smith (1970). According to them S3 was observed in subjects whose mitral cusps and chordae had been removed. Reddy's suggestion that S3 is due to the impact of the ventricle on the chest wall has also been discounted by the recent work of Ozawa [Ozawa et al., 1983; Ozawa et al., 1983] who observed the presence of S3 in open chested dog studies.
1.2.4 Fourth Heart Sound.

S4 occurs during late diastole and is and like S3 is a low frequency low intensity sound. This sound is also known as an atrial presystolic sound or atrial gallop and occurs concurrently with atrial systole just prior to ventricular systole. Observe figure 1.3 which depicts the heart sounds and their temporal relationships with other cardiac events.

1.3 Diagnostic Implications of the Third Heart Sound.

The third heart sound occurs in a high proportion of normal children but when present in adults it is usually a manifestation of some physiopathologic disease. In the latter case the S3 is essentially a clinical sign, which can be monitored in an objective non-invasive fashion, namely the phonocardiogram. The pathogenic S3 is usually associated with left ventricular failure or excessive blood flow through the mitral valve in early diastole; e.g. mitral regurgitation, ventricular septal defect etc. Increased cardiac output; through conditions such as hyperthyroidism, anaemia, fever and pregnancy; may give rise to an S3. There have not been any differences reported between the physiological and the pathological S3, except for the case of constrictive pericarditus, where the S3 has a high pitched quality known as "pericardial knock". This suggests that the genesis of the S3 is the same for both normal physiological and pathological production of the sound. Despite the lack of a full understanding of the mechanisms of production of S3, clinicians have regarded the presence or absence of an S3 as having significant clinical value in diagnosis and prognosis. An example of this is the use of the occurrence or not of a protodiastolic gallop in the Killip classification for patients after acute myocardial infarction [Killip & Kindall, 1967], where it is used as one of the discriminating clinical signs. Further support to the clinical value of S3 has been given recently by its use as an catheterization indicator in consideration for surgery for aortic regurgitation [Abdulla et al, 1981]. It’s use as a pointer to patients who may
Figure 1.3: Relationship of the Heart Sounds to other cardiac events.
benefit from the administration of digitalis has also been reported [Lee et al, 1982].

An S3 can occur in either the right or left ventricle, however the majority emanate from the latter. Common in textbooks on phoncardiography is the description of a right ventricular S3 in subjects with right heart disease.

1.4 Scope of the Research.

In this modern age of high technology medicine with its major advances in cardiological diagnostic procedures, heart disease still presents with the challenge of solving etiological problems. An early form of cardiological diagnostic technology was of course the stethoscope. This facilitated the clinical practice of auscultation, an important differential diagnostic tool. Thus the clinician could investigate the complex acoustical phenomena arising from the effects of vibrating cardiac structures (and blood) and the propagation through tissues to the chest wall. This approach is limited however by the bounds of human perception. Man has applied more advanced technology to overcome these limitations. In relation to heart sound analysis, one such application of technology has been spectral analysis. Fast fourier transform (FFT) methods have been used successfully to perform spectral analyses of the first and second heart sounds. This has greatly improved the understanding of their generation and the resulting phonocardiogram with respect to other cardiodynamic events. However, especially when applying it to the short duration third heart sound, the FFT suffers from a fundamental limitation in frequency resolution determined by the window size.

The aim of the studies described in this thesis, is to investigate methodologies to extract the frequency contents of ausculatory sounds, the third heart sound in particular, such that a part of the way towards the final goal of discerning pathologies of myocardial tissues is reached by correlating heart sound spectral energy with echocardiographically derived parameters of cardiac structures.
Chapter 2

THE INSTRUMENTATION AND DATA COLLECTION

Introduction

2.1 Phonocardiography an Extension of Auscultation.

The normal range of human hearing lies within the range 20 Hz to 20000 Hz, with maximum sensitivity of hearing lying in the speech range; about 1000 Hz to 3000 Hz. In order to be heard low frequency sounds, such as the third heart sound, must attain energy levels thousands of times greater than those needed by vibrations within the speech range. The audible frequencies of heart sounds range from less than 20 Hz to greater than 300 Hz. As mentioned above the lower frequencies require much more energy to be heard. This is a disadvantage when the clinician is listening for low frequency sounds such as the third heart sound.

A phonocardiogram (PCG) is a graphical recording of the heart sounds; usually sound amplitude plotted against time. This is obtained by placing a specially designed transducer against the chest wall. Those signals are then amplified and displayed on an oscilloscope or as a chart recorder tracing. Compared with ordinary auscultation, phonocardiography has the advantages of providing a quantified ausculatory record and it makes interpretation of the record independent of auditory acuity. Further, phonocardiography facilitates electronic and computer processing of the ausculatory
record, to extract greater information. e.g. spectral phonocardiography.

## 2.2 Phonocardiogram Data Acquisition.

There have been attempts to standardise the practice of phonocardiography, such as a recent work by Van Vollenhoven [Van Vollenhoven et al, 1979].

Heart sound recordings were carried out in a soundproofed room in the department of cardiology at the Adelaide Children’s Hospital. During the same session 2-D echocardiography was also performed.

A schematic diagram of the data collection process is shown in figure 2.1. The lead ii electrocardiogram (ECG) (see figure 2.2) and PCG at the site of the apex were recorded simultaneously into a Hewlett-Packard H.P. 3964A four channel frequency modulated tape recorder with a tape speed of $\frac{3}{4}$ inches per second.

In the PCG acquisition system, the heart sounds were detected by a microphone (H.P. 21050A) which is a piezo-electric crystal transducer, with flat frequency response from DC to 2 kHz. The output of the microphone was then fed into a PCG/ECG preamplifier, which was locally constructed, and had a flat frequency response from 1 Hz to 2 kHz.

For computer processing of the heart sounds, the analogue PCG and ECG signals were required to be digitized. i.e. The analogue signal from the tape recorder was sampled at regular intervals (this sampling frequency must meet the Nyquist criterion, which will be briefly discussed in chapter 3). The value of the sampled waveform was represented as a number in computer memory.

This process was accomplished by means of a microprocessor based system. The ECG and PCG signals were digitized by means of a two channel, 8-bit analogue to digital converter controlled by an Intel 8085 microprocessor based system (sdk85) with 8k memory.

Approximately two seconds of data could be digitized at a time, at a rate of 2042
Figure 2.1: Block diagram of the microprocessor based heart sound data collection system.
Figure 2.2: Diagram for connection of the Lead ii ECG.
samples/sec. The data from each channel (ECG & PCG) occupied 4096 bytes of memory space. Each sampled datum was represented as an unsigned hexadecimal number. The digitized ECG and PCG data could be graphically displayed, to detect any artifacts, by displaying the data held in local memory, on an oscilloscope through a two-channel digital to analogue converter, built into the system.

2.3 Data Storage.

The data stored in local microprocessor memory was transferred to a VAX 750 computer through a serial interface within the system, at a rate of 1200 baud, for storage and further analysis. The PCG and ECG data transferred from the microprocessor memory to the VAX computer was in hexadecimal format. Thus the data was then converted to decimal format by means of a computer program. The data files were demeaned (i.e. any dc offset in the data was removed), then normalized with respect to the root mean square value of the data file. Initially 4096 bytes, each of PCG and ECG data were sent to the VAX.

A computer program was then run which detected the first “R” wave peak in the ECG data and then deposited the next 2048 ECG and PCG samples in new files. Since the ECG and PCG were recorded simultaneously, both ECG and PCG files represent the same time, in the cardiac cycle, starting at the “R” wave.

Figures 2.2 and 2.3 show respectively, plots of digitized ECG and PCG data stored on computer disk files.
Figure 2.3: Digitized ECG signal.

Figure 2.4: Digitized PCG signal.
2.4 Introduction to Echocardiography.

Echocardiography is one of the major non-invasive techniques currently used to evaluate cardiac function. This modality uses ultrasound signals reflected from cardiac structures to provide data on cardiac anatomy and function. Ultrasound is high frequency sound, that is sound with frequency above the upper limit for human audition (about 20,000 Hz). However the frequencies used in medical ultrasonography occupy the range of about 1 to 10 megahertz (1 x 10⁶ to 10 x 10⁶ Hz). Ultrasound propagates in the form of longitudinal waves, where the flow of energy is along the axis of wave propagation (see fig 2.4). The velocity of propagation depends on the mechanical properties of the medium and can be expressed mathematically as:

\[ v = f\lambda \]  \hspace{1cm} (2.1)

where \( v \) is the propagation velocity, \( f \) is the frequency and \( \lambda \) is the wavelength. In most materials \( v \) is not dependent upon frequency and in human soft tissue is considered to be constant at 1540 m/sec. This implies that \( \lambda \) must change with \( v \). The different media which propagate ultrasound have a property called acoustic impedance. This can be expressed as

\[ z = \rho v \]  \hspace{1cm} (2.2)

where \( z \) is the acoustic impedance, \( \rho \) is the material density and \( v \) is the propagation velocity. Ultrasound imaging is based on the fact that different tissue structures have a difference in acoustic impedance and the interfaces of these structures cause ultrasound reflection (and refraction, see fig. 2.5). The ultrasound pulse takes finite time to travel through the media to the reflecting interface and back to the transducer (which is now in receiving mode). This time is related to the depth in the tissue of the reflecting surface, the relation being:

\[ \text{depth} = \frac{1}{2} \nu \tau \]  \hspace{1cm} (2.3)
Figure 2.5: Comparison of longitudinal and transverse waves. (Dot density show motion.)

Figure 2.6: Propagation, Reflection and Refraction of Ultrasound.
where \( t \) = time taken from transmission to reception of the pulse.

Several sub-modalities have been developed over the years, including M-mode echocardiography, 2-D echocardiography and doppler echocardiography. Figures 2.6 & 2.7 compare diagramatically the information obtained from M-mode and 2-D echocardiography. Doppler ultrasound techniques will not be discussed here. The M-mode technique produces an ultrasound pulse, from a single element transducer, directed along a single beam. This resultant pencil beam image of the heart is displayed against time for viewing. Thus in this (moving) mode, dynamic cardiac events are able to be studied. The M-mode technique is well understood and conventions for recording and measuring data have been accepted. M-mode echocardiography is capable of extremely high temporal resolution and is thus useful for the study of fast dynamic cardiac events, e.g. valve leaflet motion. If regional abnormalities exist, this technique may give a misleading picture of global cardiac performance. It is therefore important to use M-mode in conjunction with 2-D echocardiography, which portrays the global picture of events.

2.5 2-Dimensional Echocardiography.

The technology for 2-D echocardiography requires a transducer with multiple elements which produces either a fan of ultrasound beams over a sector, or a series of parallel beams over the heart area. This series of beams scan a two dimensional area of the heart. These 2-D images are converted to a dynamic real time scan by redisplaying the images at a rate of 30 per second.

Several types of transducer have found application in 2-D echocardiography today. These include the mechanical scan, the linear array and the phased array transducers. The most sophisticated of these transducers is the phased array. This transducer is designed to physically fit and thus scan, through the intercostal spaces (between the ribs). It is called an array as it consists of multiple elements, typically 32, which are activated with an appropriate delay such that the resultant wave front is electronically
Figure 2.7: M-mode Echocardiogram.

Figure 2.8: 2-D Echocardiographic cross section with beam directions numbered. CW: chest wall, RV: right ventricle, LV: left ventricle, AM: anterior mitral PM: posterior mitral, AO: aorta, LA: left atrium
"steered" (fig. 2.8). This steering action is accomplished in a dynamic fashion to achieve a scanning action which is rapid (high temporal resolution) and scans over a wide angled sector (approaching 80°).

For accurate ultrasonic interrogation of the heart, various 2-D views have to be taken. Thus the transducer has to be placed in appropriate orientations with respect to the heart, to obtain various cross-sectional views. The orientation of the transducer with the corresponding cross-sectional views have become standardised (based on the work of the American Society of Echocardiography in 1979). Shown in figures 2.10 to 2.13 are some of the more frequently used transducer orientations. The parasternal long axis view is adopted to facilitate the study of the aorta, mitral valve, the left ventricle and the left atrium. To obtain this view, the transducer is placed in the left parasternal region, in the third or fourth intercostal space. The plane containing the ultrasound beams is parallel to the line joining the apex to the aorta, which is the long axis of the left ventricle. Another standard view is the parasternal short axis view. Here the transducer is located as in the previous view and then rotated 90° clockwise such that the ultrasound beam is perpendicular to the long axis of the left ventricle. The above mentioned views are shown in fig. 2.10 and fig. 2.11 The apical two and four chamber views are other standard views, these are shown in fig. 2.12 and fig. 2.13.

These views were also used in this study. In the four chamber view the transducer is placed over the apex at the point of maximum impulse. The ultrasonic beam is parallel to the line from the right scapula to the left flank, transecting the heart from the apex to the base. This view displays all chambers of the heart along their long axes.

To record the two chamber view, the transducer is placed as in the four chamber view and then rotated 90° clockwise. Now the beam is perpendicular to the interventricular septum and only two chambers, the left ventricle and atrium, are displayed. In both the two and four chamber views the mitral orifice is visible, but these views are mutually perpendicular.
Figure 2.9: Beam steering.
The steer angle depends on the relative phase (delay) of the pulses.
Figure 2.10: Long axis view.
IVS: interventricular septum, MO: mitral orifice
AML: anterior mitral leaflet

Figure 2.11: Short axis view.
Figure 2.12: Apical two-chamber view.

Figure 2.13: Apical four-chamber view.
2.6 Echocardiographic Measurements.

Volume estimates of the left ventricle were determined using 2-D echocardiography. The most common method for determining ventricular volume is based on the assumption that the shape of the left ventricle approximates that of a prolate ellipsoid. A prolate ellipsoid has a volume given by the formula:

\[ V = \frac{4}{3} \pi a b_0 b_1 \]  

(2.4)

where \( a \) is the semi-major axis and \( b_0, b_1 \) are independent semi-minor axes. Two-dimensional echocardiography was used to measure the minor axis of the left ventricle from parasternal long axis views. Although this is not a semi-minor axis, simple mathematical transformation can change this by dividing by two. The volume formula can now be expressed as:

\[ V = \frac{4}{3} \pi \left( \frac{A}{2} \right) \left( \frac{B_0}{2} \right) \left( \frac{B_1}{2} \right) \]

(2.5)

where \( A, B_0 \) and \( B_1 \) now represent the major and minor axes respectively. At this point, two major assumptions have to be made:

(i) the minor axes \( B_0 = B_1 = B \)

(ii) the major axis \( A = 2B \)

Thus the expression for the volume 2.4 becomes

\[ V = \frac{4}{3} \pi B \left( \frac{B}{2} \right) \left( \frac{B}{2} \right) \]

\[ = \frac{\pi}{3} B^3 \]

(2.6)

Computation can be further simplified, by using \( \pi \approx 3 \). Therefore the formula 2.6 becomes simply:

\[ V = B^3 \]

(2.7)

However, as the ventricle dilates, its form becomes more spherical and the values of the long and short axes approximate each other. Teicholz's formula,
\[ V = \frac{7B^3}{2.4 + B} \]  

(2.8)

seems to be the most satisfactory one for compensating for this distortion. This approach makes the following assumptions:

1. The left ventricle approximates an ellipsoid of revolution;
2. the minor axes of the left ventricle decrease proportionately in all directions during contraction;
3. that 2-D echocardiography can be used to accurately measure the minor axis; and
4. the left ventricle remains geometrically similar for normal and pathological states.

Some of these assumptions may not be strictly applicable to this study.

The determination of a mitral valve dimension was carried out using 2-D apical views. Standard two and four chamber views were used, together with intermediate views clockwise and anticlockwise from the four chamber view. These views were obtained by careful rotation of the transducer to about 45° on either side of the four chamber view. Hence four diameter measurements were obtained for mitral valve dimension. In this study, both the average and the maximum dimension were used in regression analysis involving spectral energies of the third heart sound. The average and maximum values obtained for M.V. diameters are shown in table form in the next chapter, together with LVED volumes obtained using Teicholz's formula. Ejection Fraction (EF) and Stroke Volume (SV) were calculated from the echocardiographically determined LV volumes.
Chapter 3

ANALYSIS OF THE THIRD HEART SOUND BY THE FAST FOURIER TRANSFORM METHOD.

Introduction

In this chapter a study involving FFT analysis and echocardiographic studies is reported. The FFT technique was used to determine the spectral distribution of the third heart sound (S3) in 14 subjects between the ages of 2 and 19 years (see table 3.1). Spectral energies in 15 Hz frequency bandwidths were correlated with various echocardiographically derived parameters.

3.1 The Fast Fourier Transform.

All the heart sound signals, which exist as analogue voltage signals in the real world, have been converted to digital signals (as described in chapter 2) to facilitate computer processing. Thus, if the heart sound signal is represented by \( x(t) \) say, it is sampled with a constant interval, \( T \). This produces from the continuous function \( x(t) \), a discrete time series \( x(nT) \), where \( n \) is the sample number. The main interest is the frequency composition of \( x(t) \) and in this thesis the concern is with estimating the spectrum of a semi-random process \( x(t) \) by analysing the discrete time series obtained by sampling
a realisation of a sample function. When dealing with a continuous function, say \( x(t) \), the spectral properties of this function are obtained via the Fourier transform. viz.

\[
X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} \, dt
\]  

(3.1)

### 3.2 The Fast Fourier Transform in Heart Sound Analysis.

Until the advent of the FFT algorithm [Cooley & Tukey, 1965], frequency analysis of heart sounds was carried out using band pass filters. Heart sounds were passed through banks of narrow band pass filters with different centre frequencies. The outputs of these filter systems were recorded, the result being the spectral plot of the heart sounds. Some research workers used filters with tunable centre frequencies to scan through the spectrum.

It has only been in recent times that the FFT has been used extensively in the spectral analysis of heart sounds, [Yoganathan, et al., 1976; Hearn et al., 1979; Longhini et al., 1979; Sarkady et al, 1980], (more than 10 years after the introduction of the FFT algorithm).

The PCG and ECG data were sampled at a rate of 2042 Hz (this gave a Nyquist frequency of about 1000 Hz which was more than adequate), and preprocessed as described earlier. These files were then simultaneously plotted by means of a graphics terminal (Visual 500). In doing this, the ECG was used as a time reference for the PCG plot which aided in obtaining the starting and end points for the third heart sounds for each file (5 for each subject). Only subjects with good quality PCGs for which good quality S3s could easily be identified were used. Each PCG file was then multiplied by a file containing a hamming window \((0.54 + 0.46\cos \theta)\) co-positioned with the S3, but zero everywhere else (see table 3.2). This had the effect of extracting the S3 from the PCG file and multiplying it by the window function. A conventional FFT was then applied to these files to produce the S3 spectra (figure 3.1 shows the process).
Figure 3.1: Process in obtaining FFT spectra.
Figure 3.2: A typical FFT spectrum.
For each subject an average of five spectra were obtained to produce a spectrum for that subject. A typical spectral plot is shown in figure 3.2.

### 3.3 Derived Spectral Parameters.

The frequency spectrum thus obtained for each subject was divided into 15 Hz bands from 0 - 90 Hz. The energy for frequency bands > 90 Hz was negligible. Then for each of these 15 Hz bands, the square of the amplitude was numerically integrated, using a trapezoidal technique. i.e.

$$I = \int_a^b f(x) \, dx \approx h\left[\frac{1}{2}f(a) + f(x_1) + \ldots + f(x_{n-1}) + \frac{1}{2}f(b)\right]$$

(3.2)

where the interval of integration is subdivided into n equal subintervals \((h = \frac{b-a}{n})\), and \(f(x)\) is approximated in each interval by a linear function. On having \(f(x)\) with a continuous second derivative, the error(\(\varepsilon\)) has bounds \(kM_s \leq \varepsilon \leq kM_l\) where \(k = \frac{(b-a)^3}{12h^2}\), \(M_s\) and \(M_l\) are the smallest and the largest values respectively, of the second derivative of \(f(x)\) in the interval of integration. This error was insignificant when compared with the experimental and digitising errors.

A parameter called the energy distribution coefficient \((EDC_f)\) was determined in the frequency domain, for each subject. As the window length represented only a small fraction (typically 5\%), the spectrum was considered time-invariant and the EDC was calculated using the following formula:

$$EDC_f = \frac{\sum_{n=1}^N nx^2(n)}{\sum_{n=1}^N x^2(n)}$$

(3.3)

where \(n = \) data point number, \(x(n) = \) amplitude of S3 spectrum at \(n\). This parameter indicates the "centre of mass" of the spectrum under consideration.

The energy distribution of S3 is shown in figure 3.3 in the form of a histogram with the ordinate representing the relative energy in 15 Hz bands from 0 to 90 Hz. Table 3.3 gives the individual energy distribution for each subject as well as the average.
distribution. It was found that the energy was predominately distributed in the lower frequency bands, approximately 50% (47 + 16%) existing in the 0-15 Hz band.

As an index of energy distribution, the energy distribution coefficient (EDC) was calculated as described above. The higher the value of EDC the more the energy is distributed towards the higher end of the spectrum. Perusal of table 3.4 indicates that the EDC's vary considerably (c.v. \( \approx 50\% \)).

R-R intervals for each subject were obtained and the average R-R interval over 5 records of ECG data were found. The R-R intervals are stated in data point numbers to show up variability more accurately as each data point represents \( \frac{n}{2042} \) seconds in real time.

### 3.4 Correlation of Spectral and Echocardiographic Data.

The previously discussed parameters were correlated against the relative energy in the six 15 Hz energy of the S3 spectra which were considered. The Pearson’s correlation coefficient \( r \) was used to test for significant correlations. The correlation coefficient \( \rho \) is given, for the entire population, by

\[
\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y}
\]  

(3.4)

For the sample population we have

\[
r = \frac{s_{xy}}{s_x s_y}
\]  

(3.5)

where \( r = \) Pearsons correlation coefficient and

\[
s_x = \left[ \frac{(\sum_i x_i - \bar{x})^2}{n-1} \right]^{\frac{1}{2}}, \quad \bar{x} = \frac{1}{n} \sum_{i=1}^{N} x_i
\]  

(3.6)

\[
s_y = \left[ \frac{(\sum_i y_i - \bar{y})^2}{n-1} \right]^{\frac{1}{2}}, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^{N} y_i
\]  

(3.7)

\[
s_{xy} = \frac{\sum_{i} (x_i - \bar{x})(y_i - \bar{y})}{n-1}
\]  

(3.8)

Thus equation 3.8 becomes
Table 3.1: Details of the subjects.

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>SEX</th>
<th>AGE</th>
<th>PATHOLOGY</th>
<th>HT (cm)</th>
<th>WT (kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F</td>
<td>8</td>
<td>Fallot's Tetrology</td>
<td>108</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>2.5</td>
<td>Marpan's syndrome and mild M.V. prolapse</td>
<td>97</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>11</td>
<td>Normal</td>
<td>144</td>
<td>27</td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>7</td>
<td>VSD post-op</td>
<td>105</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>12.5</td>
<td>Epstein's anomaly</td>
<td>144</td>
<td>35</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>7</td>
<td>Fallot's tet. post-op</td>
<td>117</td>
<td>20</td>
</tr>
<tr>
<td>7</td>
<td>M</td>
<td>4</td>
<td>ASD post-op</td>
<td>102</td>
<td>18</td>
</tr>
<tr>
<td>8</td>
<td>M</td>
<td>4</td>
<td>ASD</td>
<td>104</td>
<td>18</td>
</tr>
<tr>
<td>9</td>
<td>M</td>
<td>15.5</td>
<td>Hypertension</td>
<td>138</td>
<td>45</td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>10.5</td>
<td>Septal aneurism</td>
<td>139</td>
<td>29</td>
</tr>
<tr>
<td>11</td>
<td>M</td>
<td>5</td>
<td>Small VSD infundibular</td>
<td>103</td>
<td>18</td>
</tr>
<tr>
<td>12</td>
<td>M</td>
<td>19</td>
<td>Cytotoxic drugs</td>
<td>185</td>
<td>80</td>
</tr>
<tr>
<td>13</td>
<td>M</td>
<td>13</td>
<td>Minimal L.V. outflow</td>
<td>145</td>
<td>35</td>
</tr>
<tr>
<td>14</td>
<td>F</td>
<td>3.5</td>
<td>Endocardial fibroelastosis</td>
<td>90</td>
<td>13</td>
</tr>
</tbody>
</table>

Table 3.2: Position of window for S3 in cardiac cycle.

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>AVE. WINDOW CENTRE (DATA POINT NO.)</th>
<th>AVE WINDOW CENTRE / AVE R-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>810 ± 40</td>
<td>0.72 ± 0.04</td>
</tr>
<tr>
<td>2</td>
<td>911 ± 39</td>
<td>0.78 ± 0.03</td>
</tr>
<tr>
<td>3</td>
<td>842 ± 17</td>
<td>0.47 ± 0.01</td>
</tr>
<tr>
<td>4</td>
<td>1067 ± 122</td>
<td>0.79 ± 0.09</td>
</tr>
<tr>
<td>5</td>
<td>805 ± 6</td>
<td>0.57 ± 0.001</td>
</tr>
<tr>
<td>6</td>
<td>823 ± 23</td>
<td>0.68 ± 0.03</td>
</tr>
<tr>
<td>7</td>
<td>799 ± 28</td>
<td>0.65 ± 0.02</td>
</tr>
<tr>
<td>8</td>
<td>866 ± 8</td>
<td>0.69 ± 0.01</td>
</tr>
<tr>
<td>9</td>
<td>776 ± 1</td>
<td>0.42 ± 0.001</td>
</tr>
<tr>
<td>10</td>
<td>1124 ± 78</td>
<td>0.58 ± 0.04</td>
</tr>
<tr>
<td>11</td>
<td>913 ± 79</td>
<td>0.68 ± 0.06</td>
</tr>
<tr>
<td>12</td>
<td>1319 ± 166</td>
<td>0.53 ± 0.07</td>
</tr>
<tr>
<td>13</td>
<td>1064 ± 93</td>
<td>0.52 ± 0.05</td>
</tr>
<tr>
<td>14</td>
<td>836 ± 20</td>
<td>0.85 ± 0.02</td>
</tr>
</tbody>
</table>
Figure 3.3: Histogram of S3 energy spectral distribution.
Table 3.3: % energy per frequency band.

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>0-15</th>
<th>15-30</th>
<th>45-60</th>
<th>60-75</th>
<th>75-90</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>56.9</td>
<td>41.4</td>
<td>1.1</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>2</td>
<td>19.5</td>
<td>20.9</td>
<td>20.0</td>
<td>17.4</td>
<td>13.3</td>
</tr>
<tr>
<td>3</td>
<td>21.2</td>
<td>22.2</td>
<td>21.2</td>
<td>18.2</td>
<td>11.7</td>
</tr>
<tr>
<td>4</td>
<td>52.5</td>
<td>28.3</td>
<td>9.9</td>
<td>4.3</td>
<td>2.7</td>
</tr>
<tr>
<td>5</td>
<td>44.6</td>
<td>31.7</td>
<td>14.2</td>
<td>5.2</td>
<td>2.5</td>
</tr>
<tr>
<td>6</td>
<td>59.3</td>
<td>18.7</td>
<td>10.2</td>
<td>5.5</td>
<td>3.4</td>
</tr>
<tr>
<td>7</td>
<td>72.0</td>
<td>14.3</td>
<td>3.6</td>
<td>3.4</td>
<td>3.4</td>
</tr>
<tr>
<td>8</td>
<td>48.6</td>
<td>32.0</td>
<td>11.8</td>
<td>3.5</td>
<td>2.2</td>
</tr>
<tr>
<td>9</td>
<td>22.8</td>
<td>23.2</td>
<td>20.5</td>
<td>16.2</td>
<td>11.0</td>
</tr>
<tr>
<td>10</td>
<td>69.7</td>
<td>16.9</td>
<td>3.7</td>
<td>3.3</td>
<td>3.3</td>
</tr>
<tr>
<td>11</td>
<td>51.4</td>
<td>29.7</td>
<td>10.0</td>
<td>4.0</td>
<td>2.7</td>
</tr>
<tr>
<td>12</td>
<td>50.3</td>
<td>32.0</td>
<td>9.5</td>
<td>3.6</td>
<td>2.5</td>
</tr>
<tr>
<td>13</td>
<td>49.4</td>
<td>32.4</td>
<td>11.2</td>
<td>3.0</td>
<td>2.0</td>
</tr>
<tr>
<td>14</td>
<td>40.7</td>
<td>32.2</td>
<td>16.1</td>
<td>6.3</td>
<td>2.8</td>
</tr>
</tbody>
</table>

AVE: 47± 16  29± 7  12± 6  7± 5  5± 4  3± 2

Table 3.4: Echo & Phono Derived Parameters. (dimensional unit is the cm)

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>EF</th>
<th>TEICHOLZ</th>
<th>EDV</th>
<th>ESV</th>
<th>EDCf</th>
<th>MAX</th>
<th>AVE</th>
<th>R-R AVE</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.53</td>
<td>30.9</td>
<td>14.5</td>
<td></td>
<td>14.9</td>
<td>2.5</td>
<td>2.1</td>
<td>1118</td>
<td>37</td>
</tr>
<tr>
<td>2</td>
<td>0.40</td>
<td>50.9</td>
<td>30.6</td>
<td></td>
<td>38.6</td>
<td>2.4</td>
<td>2.0</td>
<td>1161</td>
<td>64</td>
</tr>
<tr>
<td>3</td>
<td>0.48</td>
<td>47.4</td>
<td>24.6</td>
<td></td>
<td>35.9</td>
<td>2.4</td>
<td>2.2</td>
<td>1797</td>
<td>176</td>
</tr>
<tr>
<td>4</td>
<td>0.29</td>
<td>28.3</td>
<td>20.1</td>
<td></td>
<td>15.5</td>
<td>1.9</td>
<td>1.7</td>
<td>1344</td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1401</td>
<td>28</td>
</tr>
<tr>
<td>6</td>
<td>0.74</td>
<td>16.1</td>
<td>4.2</td>
<td></td>
<td>13.9</td>
<td>1.8</td>
<td>1.4</td>
<td>1211</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>0.79</td>
<td>46.0</td>
<td>9.7</td>
<td></td>
<td>8.5</td>
<td>2.0</td>
<td>1.4</td>
<td>1232</td>
<td>70</td>
</tr>
<tr>
<td>8</td>
<td>0.57</td>
<td>29.6</td>
<td>12.7</td>
<td></td>
<td>16.3</td>
<td>1.9</td>
<td>1.4</td>
<td>1257</td>
<td>38</td>
</tr>
<tr>
<td>9</td>
<td>0.59</td>
<td>78.6</td>
<td>32.2</td>
<td></td>
<td>34.9</td>
<td>2.7</td>
<td>2.1</td>
<td>1855</td>
<td>145</td>
</tr>
<tr>
<td>10</td>
<td>0.51</td>
<td>29.6</td>
<td>14.5</td>
<td></td>
<td>9.2</td>
<td>1.9</td>
<td>1.6</td>
<td>1941</td>
<td>67</td>
</tr>
<tr>
<td>11</td>
<td>0.46</td>
<td>41.0</td>
<td>22.3</td>
<td></td>
<td>15.9</td>
<td>1.8</td>
<td>1.5</td>
<td>1337</td>
<td>92</td>
</tr>
<tr>
<td>12</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>15.9</td>
<td>-</td>
<td>2.6</td>
<td>2468</td>
<td>435</td>
</tr>
<tr>
<td>13</td>
<td>0.32</td>
<td>70.0</td>
<td>47.4</td>
<td></td>
<td>15.8</td>
<td>2.2</td>
<td>2.0</td>
<td>2037</td>
<td>51</td>
</tr>
<tr>
<td>14</td>
<td>0.36</td>
<td>129.5</td>
<td>83.1</td>
<td></td>
<td>20.3</td>
<td>3.3</td>
<td>3.0</td>
<td>938</td>
<td>18</td>
</tr>
</tbody>
</table>
Here, $x_i$ is the spectral energy of the $i_{th}$ 15 hz frequency band for n subjects and $y_i$ is the parameter being tested for correlation.

Table 3.5 presents the correlation coefficients calculated for various cardiac parameters against the relative energy in the various frequency bands of S3.

The mitral valve orifice diameter (both average and maximum) correlated significantly ($r=-0.53, p < 0.05$) with energy in the 0 to 15 Hz band. The negative correlation coefficients imply an inverse relationship between mitral valve orifice size and energy in the 0 to 15 Hz band. Thus, as the mitral orifice increases in size the energy in the frequency band decreases.

Table 3.6 yields the values of correlation coefficients for sets of parameters for the subjects involved in the study. A most significant result is that of the high negative correlation between age and average window centre/average R-R interval ( $r=-0.76$, $p < 0.005$). The latter term indicates the relative part of the cycle in which S3 occurs. The negative correlation implies that as children become older, S3 occurs earlier in the cardiac cycle.

Further perusal of Table 3.6 shows that the average length of R-R interval is highly correlated with age ($r=0.88, p < 0.002$), as expected. Also significant is the correlation of EDC$_f$ vs age ($r=0.52, p < 0.05$). This suggests that as a child subject becomes older, the energy spectral distribution of S3 tends toward the higher frequencies.

### 3.5 Discussion.

The genesis of S3 has been clearly associated with the rapid filling phase of diastole [Kuo et al., 1957; Craigie et al., 1983; Adolph et al., 1970; Crevasse et al., 1962]. In the study described in this chapter, mitral valve orifice size has been shown to exhibit negative correlation with energy in the 0 to 15 Hz band of S3 which is the dominant frequency band. This implies that as M.V. orifice size is decreased, more energy is
involved in the production of S3.

Assuming an ideal fluid (blood is not ideal or even newtonian) and laminar flow, the equation of continuity states that

\[ A_a V_a = A_m V_m \]  \hspace{1cm} (3.10)

where \( A_a, V_a \) are the cross-sectional area and the velocity of flow respectively, for the atrium and \( A_m, V_m \) are the cross-sectional area and velocity of flow at the mitral orifice. Equation 3.10 can be rearranged as,

\[ V_m = \left[ \frac{A_a}{A_m} \right] V_a \]  \hspace{1cm} (3.11)

Consequently, as M.V. orifice size decreases the velocity of the in-rushing blood is increased. Since kinetic energy is proportional to \( V^2 \), there is an increase in energy imparted to the blood filling the left ventricle. Therefore the driving energy associated with the L.V. expansion during the rapid filling phase is increased. As has been discussed in the introductory chapter, the mechanisms that have been put forward to explain the genesis of S3 are:

. valvular;
. tapping theory and
. intrinsic limit to L.V. expansion.

The present work has shown that S3 occurs earlier in the cardiac cycle with any increase in age of child subjects. This supports the hypothesis that S3 is due to the L.V. reaching it's elastic limit during diastole; since it is well known that the compliance of the L.V. decreases with increase in age, the L.V. expansion would be checked earlier in diastole with older subjects as the elastic limit of the L.V. is reached earlier. This notion is supported further by the finding of a significant correlation of EDC\(_f\) with increase in age of the subjects. That is, the spectral energy of S3 is distributed more towards the the higher frequency end of the spectrum as the subject becomes older. This again is consistent with an increase in stiffness of the L.V. with age. The resonant frequencies of the L.V. increase with stiffness. Also it is well known
that higher frequencies are more attenuated by passage through body tissues than lower frequencies. Therefore as the frequency distribution of S3 is shifted to higher frequencies as the child becomes older, it would be expected that the energy in S3 would decrease with age. This in fact is what is observed in clinical practice, with the third heart sound usually disappearing by adulthood, but may re-occur with cardiac pathology or other factors such as exercise.

It is interesting to note here that S3 occurs earlier in patients suffering from constrictive pericarditis [Ozawa et al., 1983; Crevasse et al., 1962] and diastolic overload [Ozawa et al., 1983]. In the case of constrictive pericarditis, the compliance of the ventricle is greatly reduced. Following in the same line of thought as earlier, the limit of distensibility of the ventricle is reached earlier, resulting in earlier production of S3.

With diastolic overload, the ventricle is "biased" further along its force-length curve. This results in reduced time, from the start of ventricular expansion in diastole, before the ventricular expansion is checked, resulting in an earlier S3. Further, the above results support the hypothesis that S3 is the result of an intrinsic limitation to the expansion of the L.V. due to an early diastolic pressure rise caused by an increased visco-elasticity of the myocardium [Van de Werf et al., 1984; Van de Werf et al., 1984].

3.6 Limitations.

The data collection effort depended upon the availability of child subjects at Adelaide Children's Hospital. Data collection was performed on suitable patients as they presented. These tended to be pathological as they were hospital patients, but cases whose pathologies were known not to directly affect S3 were chosen. Some did not produce an acceptable S3, either because it was physiologically not apparent, or sometimes due to technical difficulties. In some cases echocardiographic data was unacceptable. These cases were excluded from the study.

Since most of the subjects studied were pathological, scatter must have been introduced into the correlation analysis. With this in mind, the fact that statistically
significant correlations were obtained suggests that real trends were observed. To check
the validity of the above-mentioned statements, a study using a statistically significant
number of normal subjects needs to be performed.

Another limiting effect was the relatively poor resolution of the FFT method for
spectral analysis. The time duration of S3 is relatively short (order of 50 ms). This
short observation time, combined with the spectral blurring effects of the window func-
tion accounts for the poor resolution of the FFT method. This problem is somewhat
reduced by the application of the maximum entropy method for spectral analysis. A
study using this method is discussed in chapter five.
Table 3.5: Correlation coefficients for parameters vs freq bands. (r critical = 0.50, p ≤ 0.05)

<table>
<thead>
<tr>
<th>FREQ BAND cline 4-5</th>
<th>LVED</th>
<th>LVES</th>
<th>M.V. diam. AVE</th>
<th>MAX</th>
<th>R-R</th>
<th>E.F.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-15</td>
<td>-0.41</td>
<td>-0.40</td>
<td>-0.53</td>
<td>-0.53</td>
<td>-0.11</td>
<td>0.36</td>
</tr>
<tr>
<td>15-30</td>
<td>0.2</td>
<td>0.34</td>
<td>0.38</td>
<td>0.31</td>
<td>-0.24</td>
<td>-0.49</td>
</tr>
<tr>
<td>30-45</td>
<td>0.45</td>
<td>0.39</td>
<td>0.46</td>
<td>0.46</td>
<td>0.14</td>
<td>-0.29</td>
</tr>
<tr>
<td>45-60</td>
<td>0.25</td>
<td>0.17</td>
<td>0.34</td>
<td>0.37</td>
<td>0.22</td>
<td>-0.07</td>
</tr>
<tr>
<td>60-75</td>
<td>0.16</td>
<td>0.07</td>
<td>0.23</td>
<td>0.28</td>
<td>0.23</td>
<td>-0.03</td>
</tr>
<tr>
<td>75-90</td>
<td>0.13</td>
<td>0.05</td>
<td>0.13</td>
<td>0.15</td>
<td>0.22</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Table 3.6: Correlation coefficients for age vs some parameters.

<table>
<thead>
<tr>
<th>PARAMETER</th>
<th>AGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDC&lt;sub&gt;t&lt;/sub&gt; ave win centre</td>
<td>r = -0.76, (p &lt; 0.005)</td>
</tr>
<tr>
<td>EDC&lt;sub&gt;t&lt;/sub&gt; ave R-R</td>
<td>r = 0.88, (p &lt; 0.005)</td>
</tr>
<tr>
<td>AGE</td>
<td>r = 0.52, (p &lt; 0.05)</td>
</tr>
</tbody>
</table>
Chapter 4

MAXIMUM ENTROPY SPECTRAL ANALYSIS

Introduction

Classical spectral analysis requires the assumptions, about the signal under analysis, of long samples of data and of stationarity. However, in real applications of biomedical spectral analysis both these assumptions are usually violated. In the case of the spectral analysis of the third heart sound, the time duration is short enough to consider it stationary; but the assumption of a long signal history is obviously erroneous. It was therefore necessary to invoke a method suitable for short data length signals.

Estimation of the spectral density (SD) of signals has been traditionally performed by the FFT since the introduction of the algorithm by Cooley and Tukey [1965]. The first use of the Fourier Transform for spectral analysis however is attributed to Schuster [1898] who coined the word “periodogram” now in common use. Schuster took more of an empirical approach but later Weiner [1930] introduced the theoretical framework for the FT analysis of stochastic signals, based on the autocorrelation approach. It was then Blackman and Tukey who implemented Weiner’s approach in a practical way [1959] which became the most popular method of spectral estimation until the advent of the FFT which has dominated the scene until relatively recently.

As discussed in chapter 3, there are some limitations associated with the FFT method of spectral analysis. The major limitation of the FFT approach to spectral
analysis is that of frequency resolution; i.e. the capability of distinguishing between closely spaced spectral peaks. For the FFT, the frequency resolution is approximately equal to \( \frac{1}{T} \), where \( T \) is the available data time. Hence, when dealing with short data lengths, the resolution is severely restricted. Another problem inherent in the FFT method is the effect of spectral "leakage". In FFT analysis a real signal represents a truncated function (i.e. it is not infinitely long), which is equivalent to multiplying it by a "window" function. The resultant FFT spectrum contains energy due to both the signal itself and the window function; i.e. window function energy leaks into the signal spectrum. In fact the result is the spectrum of the convolution of the signal and window functions. This leakage can be reduced by appropriate design of the window function, but this always results in reduced frequency resolution.

Due to these limitations in the FFT method there has in the last decade or so, been intensive research into alternative SD estimators. These methods tend to be parametric spectral analysis techniques. That is they present a parametric model for the signal generator and the technique tries to estimate the parameter values which yield the true spectrum. The most discussed model in the literature is the AR (all pole) model, in particular in relation to speech analysis. This is probably due to the fact that it is an appropriate model for speech analysis (which generates a large input into the literature) and the AR model is more tractable than the more general ARMA model.

The two most popular approaches to AR spectral analysis are the Maximum Entropy Method [Burg, 1967] and Linear Prediction [Markel & Gray, 1976]. Philosophically the approaches are quite different but give, in theory, identical spectral estimates. Recently however, it has been shown that the autocorrelation method of Linear Prediction propagates errors to a greater extent than does the Burg method [Alexander & Rhee, 1987].

The following is a development of the Maximum Entropy Method along the lines of Smylie et al [1973] and Ulrych & Bishop [1975], but is extended to include more detail and further interpretation.
4.1 Entropy

Consider a discrete random variable $X$, with sample space $s = \{x_i\}$, $x_i$ with probability of occurring $p_i$. Also consider $x_i$ as a discrete "message signal"; define self information of any element $x_i$ being received by:

$$I(x_i) = - \log(p_i)$$

Entropy is defined as the average self information i.e.

$$H = - \sum_{i=1}^{n} P_i \log P_i$$  \hspace{1cm} (4.1)

note:

1. $h > 0$

2. If for some $i$, $P_i = 1$ while for the other $p_i = 0$, then $H = 0$.

3. Entropy reaches a maximum when all $P_i$ are equal.

If $X$ is continuous

$$H = - \int_{-\infty}^{\infty} p_x(x) \log p_x(x) \, dx$$ \hspace{1cm} (4.2)

if it exists, where $p_x(x)$ is a probability density.

4.2 Linear Digital Filtering

Consider a time sequence of length $N$, $\{x_1, x_2, \ldots, x_N\}$. Apply this to a linear filter with impulse response $\{h_1, h_2, \ldots, h_M\}$, having output $\{y_1, y_2, \ldots, y_{N+M-1}\}$. $y_n$ is given by a convolution sum. i.e.

$$y_1 = \sum_{r=1}^{M} h_r x_{2-r} = h_1$$

$$y_2 = \sum_{r=1}^{M} h_r x_{3-r} = h_2$$
\[ y_3 = \sum_{r=1}^{M} h_r x_{4-r} = h_3 \] (4.3)

\[ \vdots \]

\[ y_M = \sum_{r=1}^{M} h_r x_{M+1-r} = h_M \]

i.e. \( h_r \) is the "impulse response". It can be shown that

\[ \sum_{n=1}^{N} x_n z^{-n} \cdot \sum_{m=1}^{M} h_m z^{-m} = \sum_{r=1}^{N+M-1} (\sum_{m=1}^{M} h_m x_{n+1-m}) z^{-r} \] (4.4)

which is the z-transform form of the convolution theorem. If we define

\[ X(z) = \sum_{n=1}^{M} x_n z^{-n} \]

and \( H(z) = \sum_{m=1}^{M} h_m z^{-m} \) (4.5)

and \( Y(z) = \sum_{r=1}^{N+M-1} y_r z^{-r} \)

here is defined respectively, the z-transforms of \( \{x_n\}, \{h_m\} \) and \( \{y_r\} \) and this results in

\[ Y(z) = H(z) \cdot X(z) \] (4.6)

Suppose a system to be considered has bounded inputs and bounded outputs. i.e.

\[ \sum |h_r| < \infty \] (4.7)

\[ \Rightarrow \] \( H(z) \) has all it's poles inside \( |z| = 1 \) i.e. the unit circle. Suppose the input is \( x_n = e^{j2\pi fn} \) for \(-\infty < n < \infty \) then

\[ y_n = \sum_{r=-\infty}^{\infty} h_r e^{-j2\pi f(n+1-r)} \] (4.8)

\[ = x_{n+1} \sum_{r=-\infty}^{\infty} e^{-j2\pi fr} \] (4.9)

\[ = x_{n+1} H(e^{-j2\pi f}) \] (4.10)

where \( H(e^{-j2\pi f}) \) is the frequency response of the system. Define \( R_{yy}(k) = E[y_n y_{n+k}] \) as the autocovariance function (acvf) of the output of the filter, and the spectral density
Linear Digital Filter

\[ x_n \rightarrow \text{Linear Digital Filter} \rightarrow y_n \]

Figure 4.1:

\[ S_{yy}(f) = T_s \sum_{k=-\infty}^{\infty} R_{yy}(k)e^{-j2\pi f k T_s} \]
where \( T_s \) is the sampling interval.

Now

\[ R_{yy}(k) = E\left[ \left( \sum_{r=1}^{M} h_r x_{n+1-r} \right) \left( \sum_{s=1}^{M} h_s^* x_{n+1-s} \right) \right] \]  \hspace{1cm} (4.11)

\[ = \sum_{r=1}^{M} \sum_{s=1}^{M} h_r h_s^* E[x_{n+1-r} x_{n+1-s}^*] \]  \hspace{1cm} (4.12)

\[ = \sum_{r=1}^{M} \sum_{s=1}^{M} h_r h_s^* R_{xx}(k + r - s) \]  \hspace{1cm} (4.13)

Then

\[ S_{yy}(f) = \sum_{r=1}^{M} \sum_{s=1}^{M} h_r h_s^* \sum_{k=-\infty}^{\infty} R_{xx}(k + r - s) e^{-j2\pi f k T_s} \]  \hspace{1cm} (4.14)

\[ = \sum_{r=1}^{M} h_r e^{-j2\pi f r T_s} \sum_{s=1}^{M} h_s^* e^{j2\pi f s T_s} \]  \hspace{1cm} (4.15)

\[ = H(z) H^*(z) S_{zz}(f) \]  \hspace{1cm} (4.16)

\[ = S_{zz}(f) |H(z)|^2 \]  \hspace{1cm} (4.17)

4.3 Weiner Filtering

Suppose there is known a sequence of observations \( \{x_1, x_2, \ldots, x_n\} \) on some signal \( \{s_1, s_2, \ldots, s_{M+N-1}\} \) which is the "desired signal". It is required to find linear combinations \( \{y_1, y_2, \ldots, y_{M+N-1}\} \) of the \( \{x_n\} \) which approximate the desired signal.

Suppose the linear digital filter has impulse response \( \{h_r\} \) \( (r = 1, 2, \ldots, M) \), assume \( E[x_n] = 0 \) then, \( y_n = \sum_{r=1}^{M} h_r x_{n+1-r} \) for \( n = 1, 2, \ldots, N + M - 1 \). It is required to find
the best \( \{h_r\} \) in some sense. The minimum mean square (MMS) error criterion will be used to choose the \( \{h_r\} \).

### 4.3.1 Mean Square Error.

Let \( P = E[|e_n|^2] \) where \( e_n = s_n - y_n \) is the filter error. i.e.

\[
P = E[(s_n - y_n)(s_n - y_n)]
\]

\[
P = E[s_n s_n^*] - E[s_n y_n^*] - E[y_n s_n^*] + E[y_n y_n^*]
\]

\[
P = E[|s_n|^2] - \sum_{r=1}^{M} h_r^* E[s_n x_{n+1-r}^*] - \sum_{r=1}^{M} h_r E[s_n^* x_{n+1-r}]
\]

\[
+ \sum_{r=1}^{M} \sum_{s=1}^{M} h_r h_s^* E[x_{n+1-r} x_{n+1-s}^*]
\]

\[
P = R_{xx}(0) - h^H R_{xx}^H - R_{xx}^H h + h^H R_{xx} h
\]

\[\text{note.}^1\text{ where:}
\]

\[
h = \begin{bmatrix}
h_1 \\
h_2 \\
\vdots \\
h_M \\
\end{bmatrix}
\]

\[
R_{xx} = \begin{bmatrix}
R_{xx}(0) & R_{xx}(-1) & \ldots & R_{xx}(1-M) \\
R_{xx}(1) & \ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots & \ddots \\
R_{xx}(M-1) & R_{xx}(M-2) & \ldots & R_{xx}(0)
\end{bmatrix}
\]

\[
R_{zz} = \begin{bmatrix}
R_{zz}(0) \\
\ddots \\
R_{zz}(M-1)
\end{bmatrix}
\]

where \( R_{xz}(k) = E[x_{n+k}^* s_n] \) and \( R_{zz}(k) = E[x_{n-k}^* s_n^*] \).

\( R_{xx}(0) \) is the power of the desired signal. The matrix \( R_{xx} \) is bidiagonal and an example of the Toeplitz matrix. Note that in this case, since \( R_{xz}^*(r) = R_{zx}(-r) \) then \( R_{xz}^H = R_{zx} \).

It can be shown the eigenvalues of \( R \) are all non-negative. Hence it can be shown \( |R| \geq 0 \). Let \( \lambda \) be an eigenvalue of \( R \), and \( z \) the corresponding eigenvector. Hence

\[1\text{vectors are shown in bold face. The superscript H denotes the hermitian transpose.}\]
There are in general $M$ eigenvalues and corresponding eigenvectors for a square matrix of order $M$. The eigenvalues (can be chosen to) form an orthogonal set. It is now shown that the eigenvalues of $R$ are all non-negative. Premultiply equation 4.22 by $z^H$ i.e.

$$z^H R z = \lambda z^H z = |z|^2 \geq 0$$

$$= \sum_{r=0}^{M-1} \sum_{s=0}^{M-1} z_r^* R_{xs} (s-r) z_s \geq 0$$

$$= E \left[ \sum_{r=0}^{M-1} z_r^* x_r \sum_{s=0}^{M-1} x_s z_s \right] \geq 0$$

$$= E \left[ \sum_{r=0}^{M-1} z_r^2 x_r^2 \right] \geq 0$$

But $|z|^2 = 0 \implies \lambda \geq 0$

Now form the matrix $Z$ with all the eigenvectors as columns; and the matrix $\Lambda$ as a diagonal matrix with eigenvalues in the diagonal. viz.

$$\Lambda = \begin{bmatrix}
\lambda_1 & \lambda_2 & \cdots & 0 \\
0 & \lambda_2 & \cdots & 0 \\
& & \ddots & \ddots \\
0 & \cdots & 0 & \lambda_M
\end{bmatrix}$$

Then the various eigenvector relations:

$$Rz_1 = \lambda_1 z_1$$
$$Rz_2 = \lambda_2 z_2$$
$$\vdots$$
$$Rz_M = \lambda_M z_1$$

can be written $RZ = \Lambda Z$. Then $|R||Z| = |\Lambda||Z|$, since $|Z| \neq 0$, results in $|R| = |\Lambda| = \lambda_1, \lambda_2, \ldots, \lambda_M \geq 0$.

4.3.1.1 Minimum Mean Square Error.

As shown earlier in equation ( ) $P = R_{zz}(0) - h^H R_{xx} - R_{xx}^H h + h^H R_{zz} h$ The $h$ required is the one which minimises $P$. i.e. set
\[ \frac{\partial P}{\partial h} = 0 \] (4.25)

i.e.

\[ -2R_{ax}^H + 2h^H R_{zz} = 0 \] (4.26)

or

\[ R_{xx}^H h_0 = R_{ax} \] (4.27)

this is known as the \textit{WEINER-HOPF} equation. At this minimum, the resultant filter error power

\[ P_{\text{min}} = R_{ss}(0) - h_0^H R_{ax} - R_{ex}^H h_0 + h_0^H R_{zz} h_0 \]

\[ = R_{ss}(0) - h^H R_{ax} \]

4.3.2 The Prediction Error Filter.

Propose that the desired signal is the next value of \( \{x_n\} \), i.e. \( s_n = x_{n+1} \). Then

\[ R_{ax} = E[x_n^* s_n] \] (4.30)

\[ = E[x_n^* x_{n+m+1}] \] (4.31)

\[ = R_{xx}(m+1), m = 0, 1, \ldots M - 1 \] (4.32)

Thus the Weiner-Hopf equation can be written:

\[
\begin{bmatrix}
R_{xx}(0) & R_{xx}(-1) & \ldots & R_{xx}(1-M) \\
R_{xx}(1) & R_{xx}(0) & \ldots & R_{xx}(2-M) \\
\vdots & \vdots & & \vdots \\
R_{xx}(M-1) & R_{xx}(M-2) & \ldots & R_{xx}(0)
\end{bmatrix}
\begin{bmatrix}
h_{0,1} \\
h_{0,2} \\
\vdots \\
h_{0,M}
\end{bmatrix}
= 
\begin{bmatrix}
R_{xx}(1) \\
R_{xx}(2) \\
\vdots \\
R_{xx}(M)
\end{bmatrix}
\]

\[ \sum_{r=1}^{M} h_{0,r} R_{xx}(m-r) = R_{xx}(m), m = 1, 2, \ldots M \] (4.34)
The error

\[ e_n = \delta_n - y_n \]

\[ = x_{n+1} - \hat{x}_{n+1} \]

\[ = x_{n+1} - \sum_{m=1}^{M} h_{0,m} x_{n+1-m} \]

\[ = \sum_{m=0}^{M} a_m x_{n+1-m} \]  \hspace{1cm} (4.35)

where \( a_0 = 1 \) and \( a_r = -h_{0,r}, 1 < r \leq M \)

The vector

\[
\begin{bmatrix}
  a_0 \\
  a_1 \\
  \vdots \\
  a_M
\end{bmatrix}
\]

is the impulse response of the prediction error filter. \( \{a_m\} \) constitutes the prediction error filter. Now the Weiner-Hopf equations become:

\[
\sum_{k=1}^{M} h_{0,k} R_{zz}(m-k) = R_{zz}(m) \] \hspace{1cm} (4.36)

i.e.

\[
0 = R_{zz}(m) - \sum_{k=1}^{M} h_{0,k} R_{zz}(m-k) \] \hspace{1cm} (4.37)

i.e.

\[
0 = \sum_{k=0}^{M} a_k R_{zz}(m-k), m = 1, 2 \ldots M \] \hspace{1cm} (4.38)

This demonstrates that the Weiner-Hopf equations can be written in terms of the p.e. filter coefficients.

49
Figure 4.2: Prediction Error Filter

\[ y_n = \sum_{k=1}^{M} h_k x_{n-k+1} \]

WHERE \( h_k \) IS IMPULSE RESPONSE OF THE PREDICTIVE FILTER
Let $P_M$ denote the output power of the p.e. filter, i.e.

$$P_M = R_{ss}(0) - \mathbf{R}_{sx}^H \mathbf{R}_{xx}^{-1} \mathbf{R}_{sx}$$

$$= R_{ss}(0) - \mathbf{R}_{sx}^H h_0$$

$$= P_{\text{min}}$$

(4.39)

(4.40)

where the errors are:

$$e_n = \sum_{m=0}^{M} a_m x_{n+1-m}$$

Now if $s_n = x_{n+1}$, that is the desired signal is the next observation, then $R_{xx}(0) = R_{xx}(0)$. Then

$$P_M = R_{xx}(0) - \sum_{k=1}^{M} h_{0,k} R_{xx}(-k)$$

$$= \sum_{k=0}^{M} a_k R_{xx}(-k)$$

(4.41)

(4.42)

Combining this equation with the Weiner-Hopf equations 4.38, the result is:

$$\begin{bmatrix} R_{xx}(0) & R_{xx}(-1) & \cdots & R_{xx}(-M) \\ R_{xx}(1) & R_{xx}(0) & \cdots & R_{xx}(1-M) \\ \vdots & \vdots & \ddots & \vdots \\ R_{xx}(M) & R_{xx}(M-1) & \cdots & R_{xx}(0) \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_M \end{bmatrix} = \begin{bmatrix} P_M \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

(4.43)

4.4 Maximum Entropy Spectral Analysis.

A spectral estimate is desired which is a function of the acvf of the observed time sequence $\{x_n\}$, and which maximises the entropy (or randomness) associated with the sequence, i.e. Of all possible extensions to the acvf, the one which maximises the entropy is chosen. In order to proceed, two results are needed:

1. The entropy of a gaussian sequence is $H = \frac{1}{2} \log |R_{xx}|$ (see appendix A). Where $|R_{xx}|$ is the determinant of $R_{xx}$.
2. \( \lim_{m \to \infty} |R_{xx}|^{\frac{1}{m+1}} = 2f_B \exp \Delta t \int_{-f_B}^{f_B} \log |S_x(F)| \, df \)

Where \( \Delta t \) is the sampling interval and \( f_B = \frac{1}{2\Delta t} \). The result of item (2) (see appendix B) follows from the theorem due to Szego [1920].

Now the entropy \( H \) is unbounded, so the entropy rate per sample value or “entropy rate” \( (h) \), is maximised.

\[
h = \lim_{m \to \infty} \frac{H}{m+1} \tag{4.44}
\]

\( \hat{S}_x(f) \) is required s.t. it maximises \( h \), where

\[
h = \lim_{m \to \infty} \left( \frac{1}{m+1} \right) \frac{1}{2} \log |R_{xx}|
\]

\[
= \lim_{m \to \infty} \frac{1}{2} \log |R_{xx}|^{\frac{1}{m+1}} \tag{4.46}
\]

Now \( s_x(f) \) is the power spectral density of the whole sequence and is the Fourier Transform (FT) of the acvf of the whole sequence. i.e.

\[
S_x(F) = \Delta t \sum_m R_{xx}(m) e^{-j2\pi mf \Delta t} \tag{4.47}
\]

Substituting equation item (2) into equation 4.46 we have:

\[
h = \frac{1}{2} \log(2f_B) + \frac{1}{4f_B} \int_{-f_B}^{f_B} \log |S_x(f)| \, df \tag{4.48}
\]

Substituting equation 4.48 into 4.49 we have:

\[
h = \frac{1}{4f_B} \int_{-f_B}^{f_B} \log \left[ \sum_{m=-\infty}^{\infty} e^{-j2\pi mf \Delta t} \right] \, df + \text{const} \tag{4.49}
\]

It is required that an \( \hat{S}_x(f) \) is chosen s.t. it corresponds to the FT of the \( 2M+1 \) known values of the acvf, but which adds no information to the acvf lags for \( m \geq M + 1 \); i.e. the extension of the known acvf. That is

\[
\frac{\partial h}{\partial R_{xx}(m)} = 0 \text{ for } -M \leq m \leq M + 1 \tag{4.50}
\]

and

\[
R_{xx}(m) = \Delta t \int_{-f_B}^{f_B} \hat{S}_x(f) e^{-j2\pi mf \Delta t} \, df \text{ for } 0 \leq m \leq M + 1 \tag{4.51}
\]
Carrying out this differentiation equation 4.50 becomes:

\[
\int_{-\frac{T}{2}}^{\frac{T}{2}} \frac{e^{-j2\pi m f \Delta t}}{\hat{S}_x(f)} \, df = 0 \text{ for } |m| \geq M + 1 \tag{4.52}
\]

Equation 4.52 suggests that \( \frac{1}{\hat{S}_x(f)} \) may be expressed as a truncated Fourier series, since the integral = 0 for \( |m| \geq M + 1 \). i.e.

\[
\frac{1}{\hat{S}_x(f)} = \sum_{n=-M}^{M} c_n e^{-j2\pi n f \Delta t} \tag{4.53}
\]

[where \( c_n^* = c_{-n} \) since \( \frac{1}{\hat{S}_x(f)} \) is real].

Now \( \hat{S}_x(f) \) can be written in terms of the z-transform. Put \( z = e^{j2\pi f \Delta t}, f = \frac{\log z}{2\pi \Delta t} \), then \( df = \frac{1}{2\pi j \Delta t} \, dz \). Write \( \hat{S}_x(f) \equiv \hat{S}_x \left( \frac{\log z}{2\pi \Delta t} \right) \) simply as \( \hat{S}_x(z) \). Then as \( f \to -\frac{1}{2\Delta t}, z \to e^{-j\pi} = -1 \) That is as \( f \) travels from \(-\frac{1}{2\Delta t}\) to \( \frac{1}{2\Delta t} \), \( z \) travels around the unit circle in an anticlockwise direction. Therefore equation 4.52 becomes:

\[
\int \frac{z^{-m-1}}{\hat{S}_x(z)} \, dz = 0, \text{ } |m| \geq M + 1 \tag{4.54}
\]

and

\[
R_{xx}(m) = \frac{1}{2\pi j \Delta t} \int \hat{S}_x(f) z^{m-1} \, dz, \text{ } 0 \leq |m| \geq M \tag{4.55}
\]

Recall that \( \frac{1}{\hat{S}_x(z)} \) can be represented as a finite Fourier series. Thus

\[
\frac{1}{\hat{S}_x(z)} = G(z)G^*(z^{-1})^2 \text{ (see appendix C: Spectral factorization)}
\]

Where

\[
G_m(z) = \sum_{k=0}^{M} g_k z^{-k}
\]

Now equation 4.55 becomes:

\[
R_{xx}(m) = \frac{1}{2\pi j \Delta t} \int \frac{z^{m-1}}{G_m(z)G_m^*(z^{-1})} \, dz \tag{4.56}
\]
Convolve $R_{zz}(m)$ with $\{g_m\}$, that is:

\[
\sum_{k=0}^{M} g_k R_{zz}(m - k) = -\frac{1}{2\pi j \Delta t} \int \sum_{k=0}^{M} \left[ \frac{z^{m-k-1}}{G_M(z) G_M(z^{-1})} \right] \, dz \\
= \frac{1}{2\pi j \Delta t} \int \frac{z^{m-1} \sum g_k z^{-k}}{G_M(z) G_M(z^{-1})} \, dz \\
= \frac{1}{2\pi j \Delta t} \int \frac{z^{m-1}}{G_M(z^{-1})} \, dz
\]  
(4.57)

If $G_M(z)$ is chosen to be minimum phase, then

\[
\frac{z^{m-1}}{G_M(z^{-1})}
\]

is analytic for all $m > 0$. For $m = 0$

\[
\sum_{k=0}^{M} g_k R_{zz}(m - k) = \frac{1}{2\pi j \Delta t} \int \frac{z^{-1}}{(g_0 + g_1 z + g_2 z^2 + \ldots)} \, dz \\
= \frac{1}{2\pi j \Delta t g_0} \int \frac{z}{(z + a_2 z^2 + a_3 z^3 + \ldots)} \, dz
\]
(4.60)

Where $a_0 = 0$, $a_1 = 1$, $a_i = g_0^* g_{i-1}^*$.

\[
= \frac{1}{2\pi \Delta t g_0} \left[ \log e^{j\phi} \right]_{\phi=0}^{\phi=2\pi} \\
= \frac{1}{\Delta t g_0^*}
\]
(4.62)

Thus

\[
\sum_{k=0}^{M} g_k R_{zz}(m - k) = \begin{cases} \frac{1}{\Delta t g_0^*} & m = 0 \\ 0 & 1 \leq m \leq M \end{cases}
\]
(4.64)

But recall

\[
\sum_{k=0}^{M} a_k R_{zz}(m - k) = \begin{cases} P_M & m = 0 \\ 0 & 1 \leq m \leq M \end{cases}
\]
(4.65)

where $P_M$, $\{a_k\}$ are the minimum output power and the impulse response respectively, of the prediction error filter for input $\{x_n\}$. In particular

\[
\Delta t g_0^* g_k R_{zz}(m - k) \equiv \sum_{k=0}^{M} a_k R_{zz}(m - k) = \begin{cases} 1 & m = 0 \\ 0 & 1 \leq m \leq M \end{cases}
\]
(4.66)

i.e.
\[ g_k = a_k \frac{1}{g_0^* \Delta t P_M}, \quad k = 0, 1, \ldots M \]  

Recall \( a_0 = 1 \Rightarrow g_0 = \frac{1}{g_0^* \Delta t P_M} \) i.e.

\[ |g_0|^2 = \frac{1}{\Delta t P_M} \]

Form the z-transform transfer function of the p.e. filter.

\[ A(z) = \sum_{k=0}^{M} a_k z^{-k} = 1 + \sum_{k=1}^{M} a_k z^{-k} \]  

Then

\[ G_M(z) G_M^*(z^{-1}) = |g_0|^2 A_M(z) A_M^*(z^{-1}) \]  

[i.e. \( G_M(z) \) and \( A_M(z) \) differ by only a scaling constant.]

Thus if \( \frac{1}{\hat{S}_x(f)} = G_M(z) G_M^*(z^{-1}) \) then

\[ \hat{S}_X(f) = \frac{\Delta t P_M}{|A_M(z)|^2} \]  

That is

\[ \hat{S}_X(f) = \frac{\Delta t P_M}{\left| 1 + \sum_{k=1}^{M} a_k z^{-k} \right|^2} \]  

4.5 Algorithm to Solve the Prediction Error Filter Equations.

A method was needed for solving the p.e. equations for \( P_M \) and \( \{a_m\} \) for a given sequence \( \{x_n\} \), including some method for determining the "best" prediction error filter. The \( P_M \) and the associated \( \{a_m\} \) were found using the well known Levinson-Durbin recursion algorithm. This will now be briefly outlined.
The prediction error filter equations can be written as:

\[
\begin{bmatrix}
R_{zz}(0) & R_{zz}(-1) & \ldots & R_{zz}(1-M) \\
R_{zz}(1) & R_{zz}(0) & \ldots & R_{zz}(2-M) \\
\vdots & \vdots & \ddots & \vdots \\
R_{zz}(M-1) & R_{zz}(M-2) & \ldots & R_{zz}(0)
\end{bmatrix}
\begin{bmatrix}
a_{M-1}^1 \\
a_{M-1}^2 \\
\vdots \\
a_{M-1}^{M-1}
\end{bmatrix}
= 
\begin{bmatrix}
P_M \\
0 \\
\vdots \\
0
\end{bmatrix}
\tag{4.72}
\]

This is the filter of order (M-1), and the solution to this set of equations gives the p.e. filter coefficients. Now invert equations 4.72:

\[
\begin{bmatrix}
R_{zz}(0) & R_{zz}(-1) & \ldots & R_{zz}(1-M) \\
R_{zz}(1) & R_{zz}(0) & \ldots & R_{zz}(2-M) \\
\vdots & \vdots & \ddots & \vdots \\
R_{zz}(M-1) & R_{zz}(M-2) & \ldots & R_{zz}(0)
\end{bmatrix}
\begin{bmatrix}
a_{M-1}^{(M-1)*} \\
a_{M-1}^{(M-1)*} \\
\vdots \\
a_{M-1}^{(M-1)*}
\end{bmatrix}
= 
\begin{bmatrix}
P_M \\
0 \\
\vdots \\
0
\end{bmatrix}
\tag{4.73}
\]

Take the complex conjugates, since \(R_{zz}(k) = R_{zz}(-k)\) and \(P_{M-1} = P_{M-1}^*\):

\[
\begin{bmatrix}
R_{zz}(0) & R_{zz}(-1) & \ldots & R_{zz}(1-M) \\
R_{zz}(1) & R_{zz}(0) & \ldots & R_{zz}(2-M) \\
\vdots & \vdots & \ddots & \vdots \\
R_{zz}(M-1) & R_{zz}(M-2) & \ldots & R_{zz}(0)
\end{bmatrix}
\begin{bmatrix}
a_{M-1}^{(M-1)*} \\
a_{M-1}^1 \\
\vdots \\
a_{M-1}^{(M-1)*}
\end{bmatrix}
= 
\begin{bmatrix}
P_M \\
0 \\
\vdots \\
0
\end{bmatrix}
\tag{4.74}
\]

(These are the “backward filter” equations).

Now the algorithm takes a linear combination of these equations, viz.

\[
R^{M-1} (a_{M-1} + \rho_M a_{M-1}^{(M-1)*}) = 
\begin{bmatrix}
P_{M-1} \\
0 \\
\vdots \\
P_M
\end{bmatrix}
+ \rho_M 
\begin{bmatrix}
0 \\
0 \\
\vdots \\
P_M
\end{bmatrix}
\tag{4.75}
\]

where \(\rho_M\) is some number to be determined in optimal fashion, s.t. \(|\rho_M| < 1\). Now the order of the equations is increased by one, by adding a coefficient \(a_{M-1}^{(M-1)} = 0\) in the \(a_{M-1}^{M-1}\) column, augment \(R\) and adjust the R.H.S. accordingly.

\[
\begin{bmatrix}
R_{zz}(0) & R_{zz}(-1) & \ldots & R_{zz}(-M) \\
R_{zz}(1) & R_{zz}(0) & \ldots & R_{zz}(1-M) \\
\vdots & \vdots & \ddots & \vdots \\
R_{zz}(M) & R_{zz}(M-1) & \ldots & R_{zz}(0)
\end{bmatrix}
\begin{bmatrix}
1 \\
a_{M-1}^1 \\
\vdots \\
a_{M-1}^{(M-1)*}
\end{bmatrix}
+ \rho_M 
\begin{bmatrix}
0 \\
a_{M-1}^{(M-1)*} \\
\vdots \\
a_{M-1}^{(M-1)*}
\end{bmatrix}
\tag{4.76}
\]

\[
\begin{bmatrix}
P_{M-1} \\
0 \\
\vdots \\
P_M
\end{bmatrix}
+ \rho_M 
\begin{bmatrix}
\Delta_M^* \\
0 \\
\vdots \\
\Delta_M^{M-1}
\end{bmatrix}
\tag{4.77}
\]

56
Now compare this with the order M equations. viz.

\[
\begin{bmatrix}
R_{xx}(0) & R_{xx}(-1) & \ldots & R_{xx}(-M) \\
R_{xx}(1) & R_{xx}(0) & \ldots & R_{xx}(1-M) \\
\vdots & \vdots & \ddots & \vdots \\
R_{xx}(M) & R_{xx}(M-1) & \ldots & R_{xx}(0)
\end{bmatrix}
\begin{bmatrix}
1 \\
a_{M-1}^1 \\
\vdots \\
a_{M-1}^{M-1}
\end{bmatrix}
= 
\begin{bmatrix}
P_M \\
0 \\
\vdots \\
0
\end{bmatrix}
\tag{4.78}
\]

The two sets of equations are equivalent provided:

\[a_M^M = a_{M-1}^{M-1} + \rho_M a_{M-m}^{M-1}, \ m = 0, 1, \ldots M\]

where \(\rho_M = a_M^M\) (partial correlation coefficient)

and

\[P_M = P_{M-1} + \rho_M \Delta_M\]

\[0 = \Delta_M^* + \rho_M P_{M-1}\]

where

\[
\Delta_M = \sum_{m=0}^{M} R_{xx}(-m) a_{M-m}
\]

\[
= \left[ \sum_{m=0}^{M} R_{xx}(-m) a_{M-m}^{M-1} \right]^{*}
\]

Note that:

\[
\Delta_M = -\rho_M^* P_{M-1}
\]

so \(P_M^* = P_{M-1} - \rho_M^* \rho_M P_{M-1}\)

\[= P_{M-1}(1 - |\rho_M|^2)\]

since \(P_M \geq 0 \ \forall \ M, \Rightarrow |\rho_M| \leq 1\)

It can be shown that if \(R\) is positive semi-definite, \((x^H R x \geq 0 \ \forall \ x)\), then \(|\rho_M| \leq 1\) for all orders of \(M\). The optimal choice for \(\rho_M\) is not obtained by solving

\[
\Delta_M + \rho_M P_{M-1} = 0
\]

In fact what is required is a p.e. filter of order \(M\), whose coefficients are \(1, a_M^M, \ldots a_M^M = \rho_m\), which minimises the output power. The following method for determining the
optimum coefficients was first proposed by Burg. Note that the p.e. filter can be run in both the forward and backward directions. viz.

\[ e_{f,n}^M = \sum_{m=0}^{M} a_n^M x_{n-m+1} \quad \text{forward} \quad (4.79) \]

\[ e_{b,n}^M = \sum_{m=0}^{M} a_{M-m}^M x_{n-m+1} \quad \text{backward} \quad (4.80) \]

This gives rise to two output power functions for any given p.e. filter. \( P_{f,m} = E[|e_{f,n}^M|^2] \) and \( P_{b,m} = E[|e_{b,n}^M|^2] \). These are estimated by summing over \( n = M + 1, \ldots, N \) where \( N \) is the number of data samples. Burg’s formula is to find \( \rho_M \) such that

\[ P_M = \frac{1}{2} (P_{f,m} + P_{b,M}) \quad \text{is minimised. viz.} \]

\[ \rho_M = \frac{-2 \sum_{n=M+1}^{N} e_{f,n}^{(M-1)} e_{b,(n-1)}^{(M-1)*}} {\sum_{n=M+1}^{N} \left[ |e_{f,n}^{(M-1)}|^2 + |e_{b,(n-1)}^{(M-1)}|^2 \right]} \quad (4.81) \]

The recursion is initiated by:

\[ \rho_0 = \sigma_e^2 = \frac{1}{N} \sum_{n=1}^{N} |x_n|^2, \quad e_{f,n}^0 = e_{b,n}^0 = x_n \]

The update equations for the recursion are:

\[ e_{f,n}^M = e_{f,n}^{M-1} + \rho_M e_{b,(n-1)}^{M-1} \]
\[ e_{b,n}^M = e_{b,(n-1)}^{M-1} + \rho_M e_{f,n}^{M-1} \]

### 4.5.1 Order Selection.

The Burg algorithm is based on the maximisation of the entropy rate, conditional on the coefficients of the p.e. filter satisfying the Levison-Durbin recursion. Thus as the model order increases (in the recursion), the power output of the p.e. filter steadily (monotonically) decreases (may reach a plateau due to residual gaussian noise). Obviously the ideal filter is the one which the output power is at a minimum. This point, is for a number of reasons not easily obtainable (see section 5.1 of chapter 5). Alternative criteria have been offered to obtain the optimum model order of the p.e. filter. These include the following.
1. **Final Prediction Error (FPE) Criterion.**

This selects the order such that the one step-ahead prediction error is a minimum. It was proposed by Akaike [1969], viz.

\[ FPE(m) = P_M \left( \frac{N + M + 1}{N - M - 1} \right) \]

Here \( P_M \) is monotonically decreasing, while the term in brackets is monotonically increasing. Therefore there must be a minimum. It tends to underestimate the optimum filter order.

2. **Akaike Information Criterion (AIC).**

Again this criterion was proposed by Akaike [1972]. This is based on the minimisation of the log-likelihood of the p.e. filter variance as a function of filter model order. It is defined as:

\[ AIC(m) = \log[P_M] + \frac{2M}{N} \]

This criterion is inconsistent; i.e. larger data sets do not give improved selection of model order.

3. **Autoregressive Transfer Function Criterion (CAT).**

A third criterion has been proposed by Parzen [1976]. This is defined as:

\[ CAT(m) = \frac{1}{N} \sum_{m=1}^{M} \frac{N - m}{NP} - \frac{N - M}{NP} \]

This also underestimates the optimum order.

4.5.2 **Computational Load.**

The computational cost of performing MESA will in most cases be higher than that for the FFT. This however depends greatly on the model order for the prediction filter. Consider figure 4.3 which depicts a diagram for the MESA algorithm.

In costing the algorithm, only real and complex multiplications (and divisions) were considered. The cost of a complex operation was valued at twice the equivalent


\begin{equation}
\begin{align*}
P(0) &= \sum_{t=1}^{N} x(t) \\
b_4(1) &= x(1) \\
b_4(N-1) &= x(N) \\
b_4(t) &= b_4(t-1) = x(t) \\
\text{for } t=2 \text{ to } N-1 \\
\end{align*}
\end{equation}

\begin{equation}
\begin{align*}
\text{num} &= \text{num}+b_4(1) \cdot b_2(t) \\
\text{den} &= \text{den}+b_4^2(t)+b_2^2(t) \\
\text{for } t=1 \text{ to } N-m \\
a'(t) &= a(t) \\
\text{for } t=1 \text{ to } m-1 \\
b_4(t) &= b_4(t) - a'(m) \cdot b_2(t) \\
b_2(t) &= b_2(t+1) - a'(m) \cdot b(t+1) \\
\text{for } t=1 \text{ to } N-m \\
\end{align*}
\end{equation}

\begin{equation}
\begin{align*}
a(t) &= a'(t) - a(m) \cdot a'(m-l) \\
\text{for } t=1 \text{ to } m-1 \\
\end{align*}
\end{equation}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{mesa_algorithm}
\caption{MESA Algorithm}
\end{figure}

60
real operation.

**Calculation of $a(m)$**: For $t=1$ to $N-m \rightarrow 3(N-m) + 1$ equivalent operations. The calculation of all the $(N-m)$ values of $b_1(t)$ & $b_2(t)$ must be done for $m=1$ to $M-1$, each iteration at a total cost of $2(N-m)$ calculations disregarding $P(m)$ calculation. The calculation of $a(t)$ requires $(m-1)$ values each iteration for $(M-1)$ iterations. Hence the total cost of finding the filter coefficients is:

$$
= \sum_{m=1}^{M} [1 + 5(N - m) - 2(N - M)] + \sum_{m=2}^{M} (m - 1) 
= 3NM + 2M^2 + 1 - \sum_{m=1}^{M} m + \sum_{m=2}^{M} m
$$

(4.82)

$$
= 3NM + 2M^2 
$$

(4.83)

(4.84)

If $M$ is in the order of $\frac{N}{2}$, then the computational load $\approx 2N^2$. It is suggested in the literature that the model order never exceed $\frac{N}{2}$ as the situation becomes unstable. However, from experience, a more realistic value is for $M \approx \frac{N}{4}$, where the computational load $\approx \frac{7}{8}N^2$. Compare this to an N point FFT which requires in the order of $0.5N \log_2 N$ complex multiplication (where the number of points in an FFT operation is always $\geq$ the number of data points). Therefore the ratio of the computational load for an N point FFT to the MESA algorithm using $M = \frac{N}{4}$ is:

$$
R_{\text{mesa}}^{\text{fft}} = \frac{\frac{7}{8}N^2}{0.5N \log_2 N} = \frac{1.75N}{\log_2 N}
$$

(4.85)

(4.86)

Thus the MESA algorithm compares unfavourably with the FFT in terms of its computational loading; however for small $N$, which is where MESA is most appropriate, $R_{\text{mesa}}^{\text{fft}}$ is an acceptable tradeoff for increased resolution.
Chapter 5

MAXIMUM ENTROPY SPECTRAL ANALYSIS OF THE THIRD HEART SOUND

Introduction

In chapter 3 a FFT technique was used to determine the spectral distribution of the third heart sound in a number of child subjects. It was found, however, that the FFT method produced spectra with poor frequency resolution when using short data lengths (as is the case with S3). Thus, any sharp frequency peaks in the S3 spectrum could not be resolved. Here, another approach to spectral analysis was adopted with a view to obtaining a better resolution. The mathematical development of the MESA technique was discussed in the previous chapter. The MESA technique has been demonstrated to produce superior spectral resolution when compared with more traditional methods, especially for short data lengths [Burg, 1967; Kay 1981; Ulrych, 1975]. Another advantage of the MEM is that one can use a simple rectangular window as there is not any spectral "leakage" associated with the application of the MEM.

This chapter compares the results obtained from the FFT method and the MESA approach; first in consideration of experiments on synthetic data, then using the heart sound data of Chapter 3. Spectral peaks for heart sound data obtained from MESA were subjected to correlation analysis with various 2-D echocardiographic and other
previously described parameters.

5.1 Numerical Experiments with Synthetic Data.

5.1.1 Model Order Estimation.

It was found, as discussed on page 59 of Chapter 4, that the FPE criterion underestimated the optimal model order while the AIC criterion overestimated the optimal model order; this being true for both real and synthetic data. In obtaining the minimum, the FPE criterion yielded model orders around 10% of the window width while the AIC criterion gave model orders approaching 50% of the window width. Theoretically both the prediction error power and the AIC are monotonically decreasing functions of the order number. However, with the addition of gaussian noise, they in fact reach a plateau. It was therefore decided to include in the algorithm a test for the flatness of both the AIC and prediction error power curves. Shown in figure 5.1 is the flowchart for finding the optimal model order from the prediction error power curve.

Shown in figure 5.1 is the flowchart for testing the flatness of the FPE vs model order curve. This algorithm was performed for every iteration of the Levinson-Durbin recursion. The curve was tested to be flat to within a certain tolerance (Tol); if not, this tolerance value was increased by a small amount ($\delta_1$) and the test retried. This continued until the test succeeded. The present model order value ($m$) was then compared with the previous value obtained and if the same a counter was incremented, however if not the same the counter was reset and the tolerance value was decreased by $\delta_2$ (with $\delta_1 <\delta_2$). After storing the present value for $m$, the counter was tested to see if it had remained constant (optimal model order had not changed) for a preset number of iterations. If this was true, a flag was set, to mark that the optimal model order for this criterion was obtained. Thus this algorithm automatically adjusted, from a preset value, the tolerance on the setting of the flatness.

This algorithm was also applied to the AIC curve.
Set Flag

Counter = Test

FALSE

FALSE

TRUE

Increment counter

Reset Counter

Tol = Tol - \( \delta_2 \)

FALSE

TRUE

PINT = m

FALSE

Tol = Tol + \( \delta_1 \)

1 - \frac{P(m)}{P(m-1)} < Tol

TRUE

\( P(m) = P(m-1) (1 - A(m)^2) \)

Iteration \( m \)

\( \delta_1 \ll \delta_2 \)

Figure 5.1: Flowchart for flatness testing.
Shown in tables 5.1 and 5.2 are the model orders obtained using the FPE and the AIC criteria for:

- a 10 Hz sinewave (single sine); and
- a composite waveform of 10 Hz, 20 Hz and 40 Hz sinewaves (multi-sine).

However in table 5.1, the flatness testing algorithm described earlier, has not been applied but has been in table 5.2. It can be seen that in the former case, the FPE and the AIC gave greatly divergent results as mentioned earlier. In the case of the latter however, the results were consistent.

### 5.1.2 Experiments with Synthetic Data.

In order to verify that the maximum entropy algorithm that was employed worked satisfactorily, trials with synthetic data were carried out. A computer program was written to generate the synthetic data. This program was designed to produce data that looked similar to a third heart sound in the time domain. This was achieved by the summing of three sinusoids with frequencies of 10, 20 and 40 Hertz. They had amplitudes of 1.0, 0.5 and 0.25 respectively. This of course produced a repetitive waveform, hence the program windowed a negative half cycle of this waveform to produce a waveform something like a third heart sound (See figure 5.2 ). The sinusoid frequencies and amplitudes could also be manually set, to say, have a single sinusoid.

To simulate more realistically the actual data, a random phase jitter component was incorporated which was set as a percentage of one cycle of the waveform. Gaussian noise was also added to the synthetic data. This was generated using the following
Figure 5.2: Synthetic Data
formula [G.E. Box et al, 1958]:

\[ E(t) = -2 \log u_1 \cdot \sin(2\pi u_2) \]

where \( u_1 \) and \( u_2 \) are uniformly distributed random variables.

The variance of the noise (noise power) was set either in absolute terms or calculated by the program from a signal-to-noise ratio in decibels. Shown on the following pages in figures 5.3 to 5.12 are the results of spectral analysis of synthetic data.

Depicted in figure 5.3 and figure 5.4 is the spectral estimate, based on a window length of a half cycle, for a 10 Hz sinewave using both FFT and MESA. It is obvious that the FFT did not resolve the 10 Hz peak whereas the MESA approach produced a sharp peak at approximately 10 Hz. Now scrutiny of figures 5.5 and 5.6 shows a similar comparison of the FFT and MESA approaches, but with the data consisting of three sinusoids as discussed on page 65. Here the MESA technique produces three peaks of appropriate frequencies and amplitude. The FFT method on the other hand could not resolve the individual peaks, but produced a spectrally blurred composite.

It was quite clear from the above and similar experiments that the MESA method gave superior resolution compared with FFT approaches.

Establishing that MESA performed well with short observation lengths on simulated data, it now seemed appropriate to test the method with the introduction of gaussian noise and/or random phase jitter.

Synthetic data containing the three sinusoids described earlier was generated, but this time gaussian noise was added. The spectra of these data were estimated using the MESA algorithm. Shown in figure 5.7 is the spectrum obtained for the data where the signal to noise ratio (snr) is 80 db. It is clear, that the only effect on the spectrum due to the gaussian noise was to raise the levels, which is to be expected. Again with a snr of 70 db the shape of the spectrum was not affected. This was again true for a snr of 65 db; however with the snr at 60 db the spectrum was starting to change shape. For instance the lower end of the spectrum was shifted towards the origin, with the 10 Hz peak actually shifted off the plot. Perusal of figure 5.10, which is the spectral
plot of the synthetic data with a snr of 40 db, shows a further shift of the spectrum towards the origin. The 40 Hz peak was plotted at about 30 Hz while the other two peaks had disappeared off the plot. Of course the levels are all raised.

It was found that the length of the cardiac cycle for individual subjects varied considerably (see R-R interval values in Chapter 3). This introduces a phase jitter component into the phonocardiogram. To approximate this effect a random phase jitter capability was built into the synthetic data generating program. The phase jitter was uniformly distributed and the maximum jitter could be set as a percentage of cycle length. This phase jitter was then passed through a 3-point moving average filter to give some correlation between adjacent points.

5.2 Comparison of MEM and FFT Spectral Analysis of S3.

5.2.1 Signal Preprocessing.

The data files were pre-processed to demean and to normalize with respect to the root mean square value of the data files, which were of length 2048 samples long. The third heart sounds were identified by visual inspection of the data files via a graphics terminal display. Once located, the S3 data was extracted from the data file by application of a window function to the file. This window file was also of 2048 samples long and was zero except for the region of S3, where in the case of FFT analysis was a “Hamming” window, but in the case of MESA analysis was a rectangular window.

Only subjects with good quality PCG’s were used for which S3 could be easily identified. The average of five log spectra (geometric mean) for each subject was obtained.

The frequency spectrum obtained for each subject was divided into 50 Hz bands from 0 to 300 Hz. The energy for frequency bands > 300 Hz was negligible. Then for each of these frequency bands the square of the amplitude was numerically integrated
Figure 5.3: FFT of Synthetic Data (10 Hz sinewave)

Figure 5.4: MESA of Synthetic Data (10 Hz sinewave)
Figure 5.5: FFT of Synthetic Data (10 Hz, 20 Hz and 40 Hz sinewaves).

Figure 5.6: MESA of Synthetic Data (10 Hz, 20 Hz and 40 Hz sinewaves).
Figure 5.7: Synthetic Data with Gaussian noise (SNR 80db)

Figure 5.8: Synthetic Data with Gaussian noise (SNR 65db)
Figure 5.9: Synthetic Data with Gaussian noise (SNR 60db)

Figure 5.10: Synthetic Data with Gaussian noise (SNR 40db)
Figure 5.11: Synthetic Data with Phase Jitter (1 %).

Figure 5.12: Synthetic Data with Phase Jitter (0.5 %).
over, using a trapezoidal technique. The energy distribution coefficient was obtained as in Chapter 3.

5.2.2 Fourier Transform Method.

Frequency peaks were evident in the FFT of the third heart sound, but they were not sharply defined. The first peak especially was very broad and dominant. Tabulated in Table 5.1 are the observed peaks greater than 5 db in amplitude. The frequency spectra were obtained over the range 0 to 300 Hz. This range was divided into three subranges called, the low-frequency range (0 to 80 Hz), the medium-frequency range (80 to 200 Hz) and the high-frequency range (200 to 300 Hz); after the style of published works on the first and second heart sounds [Yoganathan, 1976; Yoganathan, 1976].

5.2.3 Maximum Entropy Method.

The frequency peaks for all subjects are categorized in the same ranges as for the FFT peaks in Table 5.1 In the case of MESA, the frequency peaks are in general sharply defined. This suggests that those peaks are consistent over several records, as an average in the frequency domain were taken.

Even though there was a variation in the number of peaks observed, the first peak was always the dominant one and the amplitude of each subsequent frequency peak decreased as frequency increased.

Only one subject (subject no. 13) had a frequency peak in the high frequency range for MESA spectra, compared to all subjects studied using FFT spectra. These high frequency peaks in the FFT spectra, were probably due to leakage effects. As S3 occurs only over a short time, the window used to extract it was narrow. This increased the sidelobe energy in the window spectrum; the total spectrum that was observed of course, was the result of the convolution of the window function and the third heart sound. A narrow window also reduced the frequency resolution of the FFT spectra, thus producing broad peaks.
Table 5.1: Freq peaks in the FFT spectra of the S3 of 10 subjects.
(numbering as for MESA spectra)

Table 5.2: Details of subjects.
Figure 5.13: FFT spectrum of the S3 of subject 11.

Figure 5.14: MESA spectrum of the S3 of subject 11.
Table 5.3: MESA peaks for all subjects.

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>0-80 Hz</th>
<th>80-200 Hz</th>
<th>200-300 Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>36</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>80</td>
<td>170</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>24,36</td>
<td>130,146</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>18</td>
<td>105,130</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>18,56</td>
<td>96</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>25,68</td>
<td>92,115,131</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>10,12</td>
<td>100</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>39</td>
<td>125,150</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>38</td>
<td>125,152</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>18,70</td>
<td>180</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>9,63</td>
<td>100</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>15</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>25</td>
<td>166</td>
<td>244</td>
</tr>
<tr>
<td>14</td>
<td>21,66</td>
<td>117</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>14,50</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>34,55</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>17</td>
<td>20</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>18</td>
<td>20,42</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>19</td>
<td>18,24</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td>20,36</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 5.4: Correlation coefficients.

<table>
<thead>
<tr>
<th>FREQ PEAK</th>
<th>TEICHOLZ</th>
<th>MAX MV SIZE</th>
<th>AVE MV SIZE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EDV</td>
<td>ESV</td>
<td>0.7 (p &lt; 0.025)</td>
</tr>
<tr>
<td>1st</td>
<td>0.4</td>
<td>0.4</td>
<td>0.7 (p &lt; 0.025)</td>
</tr>
<tr>
<td>2nd</td>
<td>0.6 (p &lt; 0.025)</td>
<td>0.5 (p &lt; 0.05)</td>
<td>0.7 (p &lt; 0.025)</td>
</tr>
<tr>
<td>3rd</td>
<td>0.7 (p &lt; 0.025)</td>
<td>0.6 (p &lt; 0.025)</td>
<td>0.9 (p &lt; 0.005)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FREQ PEAK</th>
<th>EDCf</th>
<th>AGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>0.3</td>
<td>0.4</td>
</tr>
<tr>
<td>2nd</td>
<td>0.8 (p &lt; 0.005)</td>
<td>0.6 (p &lt; 0.025)</td>
</tr>
<tr>
<td>3rd</td>
<td>0.7 (p &lt; 0.05)</td>
<td>0.7 (P &lt; 0.05)</td>
</tr>
</tbody>
</table>
5.2.4 Comparison of the FFT MESA Techniques.

These studies indicate that the Fourier Transform is incapable satisfactorily resolving the frequency peaks in the third heart sound and introduces unwanted leakage. However the maximum entropy method is capable of satisfactory resolution without leakage effects. Careful attention has to be paid however to determining the correct model order for the prediction filter (see Chapter 4) in the MESA. Shown in figure 5.15 is a histogram of S3 energy versus frequency, as averaged over all subjects. It can be seen (as expected) most of the energy lies in the lower frequencies.

5.3 Correlation of MESA and Echocardiographic Data.

Table 5.4 depicts correlation coefficients for various cardiac parameter measurements against MESA frequency peaks. Only the first, second and third peaks were used, as these were dominant and only two subjects’ spectra exhibited more than three peaks. It can be seen that all peaks correlated well with mitral valve size. The first peak had a positive but not statistically significant correlation with L.V. volumes, whereas the second and third peaks had significant correlations.

5.4 Discussion.

This study has shown that there exists a significant positive correlation between M.V. size and the dominant frequencies in S3. Also there exists a positive correlation between left ventricular volumes and S3 frequencies.

If, however, the ventricle was acting like a resonant chamber it would have been reasonably expected to find a negative correlation between ventricular volume (and probably M.V.) and frequency.

Since a high correlation coefficient does not necessarily reflect a causal relationship, the high correlation between M.V. size and frequency peaks may reflect an underlying
Figure 5.15: Histogram of S3 energy vs Frequency
relationship. For example, the M.V. size may be a good indicator of the state of the heart, which in turn reflects decreased compliance of the ventricle. If this were true it would explain the positive correlations, and emphasize that the mechanical state of the myocardium is important in the production of S3.

It has been demonstrated that S3 is associated with the ventricular wall reaching an intrinsic limit to it’s longitudinal expansion [Ozawa et al, 1983]; also it is well known that S3 occurs at the time of transition from the rapid filling wave (RFW) to the slow filling wave (SFW) of the apexcardiogram (ACG). This represents a deceleration or “negative jerk” as observed by Ozawa.

The work of Aubert [Aubert et al, 1982; 1983; 1985] indicates that the external S3 is not simply a passively filtered version of an internal S3 as he found intracardiac S3’s were unobservable, whereas external S3’s were readily observable. Ozawa also found that intracardiac S3s were not readily observable.

Ozawa dismissed Reddy’s “tapping theory” [1981] on the basis of his observation of an S3 in open chested dog studies. Aubert, however asserts that Ozawa only observed the acceleration of the ventricular wall. Ozawa, by tapping on the inside of an open chested dog preparation, observed a “positive jerk” through his accelerometer recordings. This he cited as further evidence that Reddy’s tapping theory is incorrect. However, if the S3 is associated with the transition from the RFW to the SFW of the ACG, an external negative jerk would be expected, even if there was interaction between the heart and the chest wall.

In consideration of the previously mentioned works and my own studies, it appears that the third heart sound may be generated by diastolic thrusting (associated with the RFW, and the ventricle reaching an intrinsic limit in its expansion) which is probably interacting with overlying structures.
Chapter 6

ARMA MODELLING: RECOMMENDATIONS FOR FURTHER WORK

Introduction

The methods that have been applied earlier in this thesis for spectral analysis have included the FFT method and MESA. Of these methods, the former is an example of a moving average (MA) technique or equivalently an all zero model approach [Cadzow, 1982], whereas MESA as shown in Chapter 4 is an auto-regressive (AR) or all-pole model approach. The former is appropriate for modelling spectra that are characterized by sharp notches and broad peaks, while the latter is better suited for estimating spectra with sharp peaks. Combining these two forms together results in the ARMA model (to be discussed in the next section). The heart sound signals most likely contain both poles and zeros, the poles arise from resonances while zeros arise from cancellation effects. The ARMA method would then be the most appropriate approach to heart sound spectral modelling.

6.1 Auto-regressive Moving Average (ARMA) Modelling.

The most general expression describing the behaviour of a discrete time, linear, time-invariant causal system, is the following constant coefficient difference equation:
\[ y(n) = \sum_{i=0}^{q} b_i x(n - i) - \sum_{k=1}^{p} a_k y(n - k), \quad n \geq 0 \]  \hfill (6.1)

where \( \{b_i\} \) and \( \{a_k\} \) characterize the system. The assumption of \( b_0 = 1 \) can be made without loss of generality, since the input \( x(n) \) can be scaled to allow for filter gain. Now taking the z-transform of equation 6.1 yields:

\[
Y(z) = \sum_{i=0}^{q} b_i z^{-i} X(z) - \sum_{k=1}^{p} a_k z^{-k} Y(z) \quad \hfill (6.2)
\]

\[
\Rightarrow \sum_{k=0}^{p} a_k z^{-k} Y(z) = \sum_{i=0}^{q} b_i z^{-i} X(z) \quad \hfill (6.3)
\]

\[
\Rightarrow H(z) = \frac{\sum_{i=0}^{q} b_i z^{-i}}{\sum_{k=0}^{p} a_k z^{k}} \quad \hfill (6.4)
\]

where \( H(z) = \frac{Y(z)}{X(z)} \) is the transfer function of the system.

The solution of equation 6.1 can be found by taking the inverse z-transform of equation (2), given the initial conditions: \( y(0), \ldots, y(1-p) \). Equation (4) can be written as:

\[
H(z) = b_0 \frac{\prod_{i=1}^{q} (1 - Z_i z^{-1})}{\prod_{k=1}^{p} (1 - P_k z^{-1})} \quad \hfill (6.5)
\]

where \( Z_i \) and \( P_k \) represent respectively the zeros and the poles of \( H(z) \).

### 6.2 ARMA Model From The Maximum Entropy (AR) Model.

#### 6.2.1 Homomorphic Systems.

Linear systems, which have been discussed thus far, are relatively easy to analyze and characterize. This is so because they obey the principle of superposition which leads to powerful and elegant mathematical representations. There also exists classes of nonlinear systems which obey a generalized principle of superposition and are called
homomorphic systems. These systems are represented by linear transformations between input and output vector spaces.

To discuss these systems the following general operators are defined:

1. $\diamond$: a rule for combining systems inputs (addition, multiplication, convolution, etc.).
2. $\triangle$: a rule for combining system outputs.
3. $\bullet$: a rule for combining inputs with scalars.
4. $\heartsuit$: a rule for combining outputs with scalars.

Then using $H$ to denote the system transformation, the generalized superposition equations are:

$$H[x_1(n)\triangle x_2(n)] = H[x_1(n)\triangle x_2(n)]$$  \hspace{1cm} (6.6)

and

$$H[c\bullet x_1(n)] = c\heartsuit H[x_1(n)]$$  \hspace{1cm} (6.7)

Clearly, linear systems are a special case where $\diamond$ and $\triangle$ are addition, while $\bullet$ and $\heartsuit$ are multiplication. If the theory of linear vector spaces is to be employed, then the operators must obey the postulates of vector addition and scalar multiplication; i.e.

$$x_1(n)\diamond x_2(n) = x_2(n)\diamond x_1(n)$$

$$y_1(n)\triangle y_2(n) = y_2(n)\triangle y_1(n)$$  \hspace{1cm} (6.8)

and

$$x_1(n)\diamond [x_2(n)\diamond x_3(n)] = [x_1(n)\diamond x_2(n)]\diamond x_3(n)$$

$$y_1(n)\triangle [y_2(n)\triangle y_3(n)] = [y_1(n)\triangle y_2(n)]\triangle y_3(n)$$  \hspace{1cm} (6.9)

If the operations $\diamond$ and $\triangle$ on the input side, while $\bullet$ and $\heartsuit$ on the output side, respectively represent vector addition and scalar multiplication, then all such systems can be represented as a cascade of three systems as shown in figure 6.1. This
Figure 6.1: Canonic representation of some classes of homomorphic systems.

Figure 6.2: Canonic representation of multiplicative homomorphic systems.
representation shown in figure 6.1 is known as the canonic form of the homomorphic system. The system $D_\diamond$ performs the following:

$$D_\diamond[x_1(n) \diamond x_2(n)] = D_\diamond[x_1(n)] + D_\diamond[x_2(n)]$$

$$= \tilde{x}_1(n) + \tilde{x}_2(n)$$

(6.10)

$$d_\diamond[c \triangle x_1(n)] = c D_\diamond[x_1(n)]$$

$$= c\tilde{x}_1(n)$$

(6.11)

The effect of the system $D_\diamond$ is to transform the signals $x_1(n)$ and $x_2(n)$, which are combined according to the rule $\diamond$, into a linear combination of the signals. The system $L$ is a linear system and thus obeys the principle of superposition. $\hat{y}(n)$ is a linear combination of the inputs $\tilde{x}_1(n)$ and $\tilde{x}_2(n)$. Now $D_\triangle^{-1}$ transforms from addition to $\triangle$. Since the system $D_\diamond$ is determined by the operations of $\diamond$ and $\triangle$, it is a characteristic of the class and thus known as the characteristic system for the operation of $\diamond$.

If the class of homomorphic system under consideration has the same input and output operations, then all systems in the class differ only in the linear part. This implies once the characteristic system for the class is known, only a linear filtering problem remains to be solved.

Shown in figure 6.2 is an example of a homomorphic system where the $\diamond$ operation is multiplication while the $\triangle$ operation is exponentiation. This type of system can be useful in applications such as deconvolution, where the input signal is a product of two or more other signals.

### 6.2.2 The Complex Cepstrum.

The complex cepstrum is the result of a useful transformation in application to convolutional homomorphic systems. Let $S(z)$ be the z-transform of a stable minimum phase time series (i.e. all the poles and zeros are within the unit circle). Now set $\tilde{S}(z) = \log[S(z)]$. This has a Laurent series expansion, including the unit circle in its region of convergence, of :-
\[ \tilde{S}(z) = \sum_{k=0}^{\infty} c_k z^{-k} \]  

(6.12)

Put \( z = e^{j\omega} \); i.e. on the unit circle with \( \omega = 2\pi f \). Now equation 6.12 becomes:

\[ \tilde{S}(z) = c_0 + \sum_{k=0}^{\infty} c_k e^{-jk\omega} = \log|S(\omega)| \]  

(6.13)

\( \{c_k\} \) are called cepstral coefficients.

Writing \( S(\omega) = |S(\omega)|e^{-j\theta(\omega)} \) gives real part:

\[ \log|S(\omega)| = c_0 + \sum_{k=1}^{\infty} c_k \cos(k\omega) \]  

(6.14)

and imaginary part:

\[ \theta(\omega) + 2\lambda\pi = \sum_{k=1}^{\infty} c_k \sin(k\omega), \ \lambda \text{ integer} \]  

(6.15)

Taking the derivative:

\[ \theta'(\omega) = \sum_{k=1}^{\infty} kc_k \cos(k\omega) \]  

(6.16)

where \( \theta'(\omega) > 0 \) for \( \omega \) near a pole. \( \theta'(\omega) < 0 \) for \( \omega \) near a zero. This leads to a set of cepstral pole coefficients \( \{c^+_k\} \) and also to a set of cepstral zero coefficients \( \{c_k\} \).

### 6.2.3 Maximum Entropy ARMA Spectral Estimate.

From the well known Wiener-Khinchine theorem

\[ \hat{R}_k = \int_{-\infty}^{\infty} S(f)e^{j2\pi fk} \, df, \ |k| \leq M \]  

(6.17)

where \( \hat{R}_k \) is the observed autocorrelation (normalised autocovariance) lags of a wide sense stationary stochastic process, with finite variance \( \sigma^2 \) and strictly continuous power spectrum \( P(f) \);

i.e. \( S(f) = \frac{P(f)}{\sigma^2} \) and \( \int_{-\infty}^{\infty} S(f) \, df = 1 \)

Applying the maximum entropy condition for the unknown lags \( R_k, k > |M| \):
subject to the 2M+1 constraints imposed by the known autocorrelation lags. It's solution is straight forward and often called the “maximum Entropy Principle”; viz.

\[
S(f) = \exp \left\{ \sum_{k=-M}^{M} \lambda_k e^{j2\pi f k} \right\}
\]  

where \( \lambda_k \in \mathbb{R} \). Taking real logarithms the log-spectral density is:

\[
\tilde{S}(f) = \sum_{k=-M}^{M} \tilde{R}_k e^{-j2\pi f k}
\]

where \( \tilde{R}_k, \tilde{S}(f) \geq 0 \) and \( \tilde{S}(f) \in \mathbb{R}, \mathcal{C}^\infty \). The equation 6.20 may be considered to be obtained via the fourier transform of the finite autocorrelation \( \tilde{R}_k \). Through the inverse FT of \( \tilde{R}_k \) is obtained the cepstrum. This leads to the well known Burg AR(m) spectrum. It has been shown recently [Liefhebber & Boeke, 1987] that by applying 4M+1 constraints to equation 6.18, that a more general, ARMA(M,M) spectrum can be obtained. However these extra 2M constraints do not necessarily imply more information. It can be shown that by only applying 2M+1 constraints to equation 6.19 can yield an ARMA spectrum.

Consider the spectral factorization

\[
S(f) = H(f) \cdot H(-f), \quad f \in \Omega
\]

where \( H(f) = \sum_{k=0}^{\infty} y_k e^{-j2\pi f k} \)

From equation 6.21:

\[
\log|S(f)| = \log|H(f)| + \log|H(-f)|
\]

or

\[
\tilde{S}(f) = \tilde{H}(f) + \tilde{H}(-f)
\]

( where \( H(z) \) is the transfer function of a LTI minimum phase system )

From the IFT of equation 6.23
\[ \hat{R}_k = \hat{y}_k + \hat{y}_{-k} \]  

(6.24)

where \( \{ \hat{y}_k \} \) are called the complex cepstrum. With \( \hat{y}_k \in \mathbb{R}, y_k \) min phase, \( \hat{y}_k \) causal and \( \hat{R}_k = \hat{R}_{-k} \).

The sequence \( \{ y_k \} \) can be expressed [Oppenheim & Schafer 1968] as:

\[ \hat{y}_k = U_{+k} \hat{R}_K \]  

(6.25)

where \( U_{+k} = \begin{cases} 1 & k > 0 \\ \frac{1}{2} & k = 0 \\ 0 & k < 0 \end{cases} \)

Now equation 6.19 may be written as:

\[ S(f) = \exp \left\{ \hat{y}_0 + 2 \sum_{k=1}^{M} \hat{y}_k \cos(2\pi f k) \right\} \]  

(6.26)

The ARMA spectral model can then be written as:

\[ \hat{S}_{ARMA}(f) = \exp \left\{ c_0 + 2 \sum_{k=1}^{P+Q} c_k \cos(2\pi f k) \right\} \]  

(6.27)

where \( c_k = \) cepstral coefficients

\[ c_k = c_{MA} - c_{AR} \]

6.2.4 ARMA Spectrum from AR Coefficients.

Cepstral coefficients can be obtained from prediction filter coefficients using the recursion formula [Schroeder,1981]:

\[ c_k = -a_k - \sum_{k=1}^{M-1} \frac{k}{M} c_k a_{m-k}, \quad 1 \leq n \leq M \]  

(6.28)

Then by use of equation 6.18 it is possible to obtain AR and MA coefficients; i.e.\( \{ c^+ \} \) are the pole (AR) coefficients and \( \{ c^- \} \) are the zero (MA) coefficients (see figure 6.3.

Thus the ARMA model can be constructed from these coefficients. The equation 6.19 can be written in its more usual form:
\[ R_{x}(0), R_{x}(1) \ldots R_{x}(M - 1) \]

**AR prediction coefficients**
Levinson algorithm, AIC

**Compute cepstral coefficients eqn 6.29**

**Compute group delay \( \Theta(f) \)**
(-ve derivative of the phase)

**If \( \Theta(f) \geq 0 \) IDFT\(^+ \) \( \rightarrow kc^+(k) \)**
**If \( \Theta(f) < 0 \) IDFT\(^- \) \( \rightarrow kc^-(k) \)**

**Find ARMA power spectrum from eqn 6.30**

**NOTE:** equation 6.30 is equivalent to:

\[
S_{db}^{ARMA}(f) = \frac{20P_m \Delta f}{\log 10} \sum_{k=1}^{P+q} c_k^{ARMA} \cos 2\pi f k
\]

where \( c_k^{ARMA} = c_k^{MA} - c_k^{AR} \)
\( = c_k^+ - c_k^- \)

Figure 6.3: Approach for finding ARMA coefficients.
\[ \hat{S}_{ARMA}(f) = \frac{\sigma^2 \sum_{k=0}^{d} b_k e^{-j2\pi kf}}{|1 + \sum_{k=1}^{p} a_k e^{-j2\pi kf}|^2} \] (6.29)

### 6.3 Deconvolution of the Chest Wall Transfer Function from the Third Heart Sound.

The signal that arrives at the phonocardiograph transducer is the result of the convolution of the intracardiac phono with the transfer function of the transmission path of these heart sounds.

![Diagram](image)

**Figure 6.3**

To show this a little more explicitly, let \( P_i \) and \( P_e \) represent the intracardiac phono and the external phono respectively. Then:

\[
    p_e(t) = h(t) * p_i(t) \quad \text{time domain} \tag{6.30}
\]

\[
    P_e(z) = H(z) \cdot P_i(z) \quad \text{z domain} \tag{6.31}
\]

i.e. \( H(z) = \frac{P_e}{P_i} \)
When applying an ARMA modelling technique to the external phono, a z-transform rational expression, such as shown in equation 6.4, is obtained for the signal. This expression contains a component due to the signal poles and zeros as well as the poles and zeros for the transmission path, as shown by the RHS of equation 6.32. It may be possible by use of “blind” deconvolution techniques, such as homomorphic deconvolution [Stockham & Cannon, 1975], bayesian signal deconvolution [Hunt & Trussell, 1976], a kalman filtering approach to deconvolution [Katayama & Takashi, 1988] etc. By applying one or more of these (or some hitherto unknown) methods it may be possible to remove the effects of the transmission path on the phonocardiogram. It would then be possible to obtain an estimate of the internal phono from the external phono. In the case of S3 especially, this would be invaluable in clinical evaluation of cardiac pathologies.

6.4 Conclusions and Recommendations for Further Work.

It has been demonstrated by this thesis that the spectral analysis of the third heart sound coupled with other techniques, such as echocardiography, are useful in the research context and potentially useful in the clinical context. The work in this thesis has shown that spectral analysis of S3 using the FFT method has produced some useful information, but is quite limited due to the poor spectral resolution of the method. MESA on the other hand overcomes the problems associated with the FFT method, but introduces other problems. The major problem introduced by use of MESA is
finding the correct model order. Even though there are tests available to find the optimum model order, an amount of subjective input is still required; also MESA assumes an AR model for the data which may not be appropriate. Applying the method outlined in this chapter to obtain an ARMA model from the MESA coefficients, would overcome these problems.

It could be fruitful to apply the methods discussed in this chapter using a large number of subjects to set up a data base. If in particular, the deconvolution problem was solved, then internal S3s could be obtained non-invasively. Further research should be done using simultaneously recorded internal and external phonos to develop and verify the above technique. A study should then be done using adult subjects with pathological S3’s to develop a clinically useful indicator of the effect of therapeutic measures on the pathology.
Appendix A

The Entropy of a Gaussian Sequence.

Consider the entropy of a 2-dimensional gaussian signal. Let $\Sigma = E[(x_i - \mu_x)(y_j - \mu_y)]$, which forms an outer product that produces the covariance matrix.

viz. (with zero means)

$$\Sigma = \begin{bmatrix} E[|x|^2] & E[xy^*] \\ E[x^*y] & E[|y|^2] \end{bmatrix} = \begin{bmatrix} \sigma^2_x & \rho_{xy}\sigma_x\sigma_y \\ \rho_{yx}\sigma_x\sigma_y & \sigma^2_y \end{bmatrix}$$

The determinant of the covariance matrix, $|\Sigma| = \sigma^2_x\sigma^2_y(1 - \rho^2_{xy})$. The bivariate gaussian probability density can thus be represented as:

$$p(x) = \frac{1}{\sigma_x\sigma_y2\pi\sqrt{1 - \rho^2_{xy}}} \exp \left\{ -\frac{1}{2} \begin{bmatrix} x \\ y \end{bmatrix} \Sigma^{-1} \begin{bmatrix} x \\ y \end{bmatrix} \right\}$$

(A.1)

$$p(x) = \frac{1}{\sigma_x\sigma_y2\pi\sqrt{1 - \rho^2_{xy}}} \exp \left\{ -\frac{1}{2(1 - \rho^2_{xy})} \left[ \frac{x^2}{\sigma^2_x} + \frac{y^2}{\sigma^2_y} - \frac{2\rho_{xy}}{\sigma_x\sigma_y} \right] \right\}$$

(A.2)

The entropy $H$ is

$$H = -\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x,y) \log[P(x,y)] \, dx \, dy$$

(A.3)

$$H = -\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x,y) \{ \log[\sigma_x\sigma_y2\pi\frac{1}{2}(1 - \rho^2_{xy})]$$

$$- \frac{1}{2(1 - \rho^2_{xy})} \left[ \frac{x^2}{\sigma^2_x} + \frac{y^2}{\sigma^2_y} - \frac{2\rho_{xy}}{\sigma_x\sigma_y} \right] \} \, dx \, dy$$

(A.4)

$$H = \log \left[ \frac{1}{2} \frac{1}{\sigma_x\sigma_y} \frac{1}{2\pi(1 - \rho^2_{xy})} \right]$$

$$+ \frac{1}{2(1 - \rho^2_{xy})} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x,y) \left[ \frac{x^2}{\sigma^2_x} + \frac{y^2}{\sigma^2_y} \right] \, dx \, dy - \frac{\rho^2_{xy}}{2(1 - \rho^2_{xy})}$$

(A.5)

since $\rho_{xy} = \int p(x,y) \frac{xy}{\sigma_x\sigma_y} \, dx \, dy$  

(A.6)
Now extend the analysis to the case of an N-dimensional gaussian signal. Consider
the N-variate gaussian probability density:

$$p(x) = \frac{1}{(\sqrt{\det(\Sigma)})^{\frac{N}{2}}} \exp \left\{ -\frac{1}{2} (x - \mu)^H \Sigma^{-1} (x - \mu) \right\}$$

(A.9)

Where as earlier, \( \Sigma = E[(x_k - \mu_k)(y_k - \mu_k)^T] \) is the covariance matrix and \( \mu_k = E[k] \).

The entropy \( H \) is :

$$H = -\int_{-\infty}^{\infty} p(x) \cdot \left\{ \log \left[ \frac{1}{(\sqrt{\det(\Sigma)})^{\frac{N}{2}}} \right] - \frac{1}{2} (x - \mu)^H \Sigma^{-1} (x - \mu) \right\} \, dx$$

(A.10)

$$= \log[(\sqrt{\det(\Sigma)})^N] \frac{1}{2} + \frac{1}{2} \int_{-\infty}^{\infty} p(x)(x - \mu)^H \Sigma^{-1} (x - \mu) \, dx$$

(A.11)

$$= \frac{1}{2} \log||\Sigma|| + \frac{1}{2} \log[(2\pi)^N] + \frac{1}{2}$$

(A.12)

$$= \frac{1}{2} \log||\Sigma|| + \text{const... As for 2-D case.}$$

(A.13)
Appendix B

Theorem Due to Szégo.

Let $R_{xz}$ be the $(M+1)$ by $(M+1)$ acvf matrix of a (weakly) stationary sequence \{x_1, x_2 \ldots x_n\}, whose true power spectral density is $S_z(f)$, which is assumed to be band limited; i.e. $S_z(f) \neq 0$ only if $|f| \leq f_B$. Let the eigenvalues of $R_{xz}$ be $\lambda_0, \lambda_1, \ldots \lambda_M$.

Let $g$ be a continuous function, then Szégo's theorem says:

$$
\lim_{M \to \infty} \frac{g(\lambda_0) + \ldots + g(\lambda_M)}{M + 1} = -\frac{1}{2f_B} \int_{-f_B}^{f_B} g[2f_B S_z(f)] \, df
$$

(B.1)

(This is a generalization of the result: $R_{xz}(0) = \int_{-f_B}^{f_B} S_z(f) \, df$).

Thus equation B.1 can be derived as follows:

Let $g(x) = \log|x|$, then

$$
\lim_{M \to \infty} \frac{\log[\lambda_0] + \ldots + \log[\lambda_M]}{M + 1} = \lim_{M \to \infty} \frac{\log[\lambda_0 \lambda_1 \ldots \lambda_M]}{M + 1} = \lim_{M \to \infty} \log[\lambda_0 \lambda_1 \ldots \lambda_M]^{\frac{1}{M+1}}
$$

(B.3)

$$
= -\frac{1}{2f_B} \int_{-f_B}^{f_B} \log[2f_B S_z(f)] \, df
$$

(B.4)

But it has been shown that $|R_{xz}| = \lambda_0 \lambda_1 \ldots \lambda_M \Rightarrow$

$$
\lim_{M \to \infty} \log[|R_{xz}|^{\frac{1}{M+1}}] = \frac{1}{2f_B} \int_{-f_B}^{f_B} \{ \log[2f_B] + \log[S_z(f)] \} \, df
$$

(B.5)

$$
= \log[2f_B] + \frac{1}{2f_B} \int_{-f_B}^{f_B} \log[S_z(f)] \, df
$$

(B.6)

Provided the limit exists, then:

$$
\lim_{M \to \infty} \log[|R_{xz}|^{\frac{1}{M+1}}] = 2f_B \exp\left\{ \int_{-f_B}^{f_B} \log[S_z(f)] \, df \right\}
$$

(B.7)
If bandwidth is determined by the Nyquist frequency, (i.e. $f_B = \frac{1}{2\Delta t}$), then this is the required result.
Appendix C

Spectral Factorization.

Consider a signal $y(t)$ which is a filtered version of a white noise signal $w(t)$, the filter having impulse response $h(t)$ (assume discrete time signals); i.e. $Y(z) = H(z) \cdot W(z)$.

Now the power spectral density of $y(t)$ is:

$$S_y(z) = \sum_{k=-\infty}^{\infty} E[y(t)y^*(t-k)]z^{-k}$$  \hspace{1cm} (C.1)

This is the $z$-transform case of the Weiner-Klinchine theorem.

If $y(t)$ is ergodic, this is$^1$:

$$S_y(z) = \sum_{k=-\infty}^{\infty} \sum_{t=-\infty}^{\infty} y(t)y^*(t-k)z^{-k}$$  \hspace{1cm} (C.2)

$$= \sum_{t=-\infty}^{\infty} y(t)z^{-t} \sum_{k=-\infty}^{\infty} y^*(t-k)z^{-k}$$  \hspace{1cm} (C.3)

$$= Y(z) \cdot Y^*(z^{-1})$$  \hspace{1cm} (C.4)

Similarly $S_w(z) = W(z) \cdot W^*(z^{-1})$, but

$$S_z(z) = \sum_{m=-\infty}^{\infty} R_{ww}(m)z^{-m}$$  \hspace{1cm} (C.5)

$$= \sum_{m=-\infty}^{\infty} \sigma^2_w \delta_{m0} z^{-m}$$  \hspace{1cm} (C.6)

$$= \sigma^2_w$$  \hspace{1cm} (C.7)

where $\sigma^2_w$ is the white noise variance (energy) and $\delta_{m0}$ is the kronecka delta.

Now

$$S_y = H(z) \cdot W(z) \cdot W^*(z^{-1}) \cdot H^*(z^{-1})$$

$$E[y(t)y^*(t-k)] = \sum_{t=-\infty}^{\infty} y(t)y^*(t-k)p(t) = \sum_{t=-\infty}^{\infty} y(t)\int y^*(t-k)$$

where $p(t)$ is a uniform distribution.
\[ S_v(z) = \sigma_v^2 H(z) \cdot H(z^{-1}) \]

Where \( S_v(z) \in \mathbb{R}, > 1 \).

Conversely any (well behaved) real positive function of \( z \) can be factorized into the factors \( H(z), H^*(z^{-1}) \). There are many such factorizations.

It is always possible to find \( H(z) \), for a given real positive \( S_v(z) \), such that the zero's of \( H(z) \) lie completely inside the unit circle (no deterministic component in the signal).

Then \( H(z) \) is said to be \textit{minimum phase}. \( H^*(z) \) then has the conjugate zeros and \( H^*(z^{-1}) \) has its zeros' outside the unit circle.
Bibliography


101


ERRATA

Chapter 3
On page 32 insert “bands” after “15 Hz energy” in section 3.4.

Chapter 4
Page 45 section 4.3 : The sequence “\{y_1, y_2, \ldots y_{M+N-1}\}” should read “\{y_1, y_2, \ldots y_{M+N-1}\}”
The following equations are amended as follows:
Equation 4.74 on page 56 is amended to:
\[
\begin{bmatrix}
R_{xx}(0) & R_{xx}(-1) & \ldots & R_{xx}(1-M) \\
R_{xx}(1) & R_{xx}(0) & \ldots & R_{xx}(2-M) \\
\vdots & \vdots & \ddots & \vdots \\
R_{xx}(M-1) & R_{xx}(M-2) & \ldots & R_{xx}(0)
\end{bmatrix}
\begin{bmatrix}
a_{M-1}^{(M-1)} \\
a_{M-1}^{(M-1)} \\
\vdots \\
a_{M-1}^{(M-1)} \\
1
\end{bmatrix}
= 
\begin{bmatrix}
0 \\
0 \\
\vdots \\
1
\end{bmatrix} \quad (4.74)
\]
Equation 4.76 on page 56 is amended to:
\[
\begin{bmatrix}
R_{xx}(0) & R_{xx}(-1) & \ldots & R_{xx}(-M) \\
R_{xx}(1) & R_{xx}(0) & \ldots & R_{xx}(1-M) \\
\vdots & \vdots & \ddots & \vdots \\
R_{xx}(M) & R_{xx}(M-1) & \ldots & R_{xx}(0)
\end{bmatrix}
\begin{bmatrix}
1 \\
a_{1}^{M-1} \\
\vdots \\
a_{1}^{M-1} \\
0
\end{bmatrix}
+ \rho M
\begin{bmatrix}
0 \\
0 \\
\vdots \\
1 \\
1
\end{bmatrix}
\]
\[
= 
\begin{bmatrix}
P_{M-1} \\
0 \\
\vdots \\
P_{M-1}
\end{bmatrix}
+ \rho M
\begin{bmatrix}
\Delta_{M}^{*} \\
0 \\
\vdots \\
0
\end{bmatrix} \quad (4.76)
\]
Equation 4.78 on page 57 is amended to:
\[
\begin{bmatrix}
R_{xx}(0) & R_{xx}(-1) & \ldots & R_{xx}(-M) \\
R_{xx}(1) & R_{xx}(0) & \ldots & R_{xx}(1-M) \\
\vdots & \vdots & \ddots & \vdots \\
R_{xx}(M) & R_{xx}(M-1) & \ldots & R_{xx}(0)
\end{bmatrix}
\begin{bmatrix}
1 \\
a_{M}^{1} \\
\vdots \\
a_{M}^{1}
\end{bmatrix}
= 
\begin{bmatrix}
P_{M} \\
0 \\
\vdots \\
0
\end{bmatrix} \quad (4.78)
\]

Chapter 6
Below equation 6.16 on page 86, replace “pole” with “complex pole” and “zero” with “complex zero”.
Below equation 6.27 should read:
“where P = AR order, Q = MA order
c_k = cepstral coefficients
with c_k = c_{AR} - c_{MA}”
The following equations are amended as follows:
In the diagram on page 89:
“eqn 6.29” should read “eqn 6.28”
and “eqn 6.30” should read “eqn 6.29”