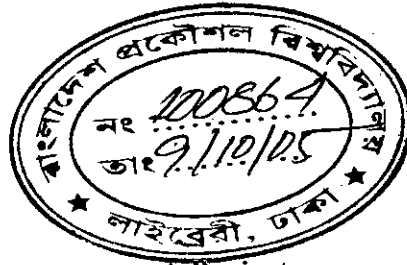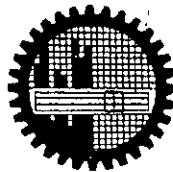# Fractal Modeling of Electrocardiogram

by

Saleh Md. Abdullah Khan Mojahedy

A Project
Submitted to the Department of
Electrical and Electronic Engineering, BUET,
in Partial Fulfillment of the Requirements for the Degree of

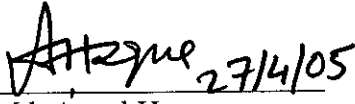## Master of Engineering in Electrical and Electronic Engineering

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
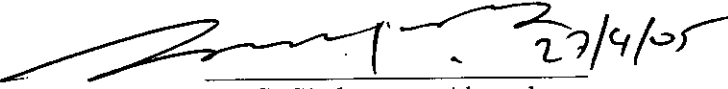BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY
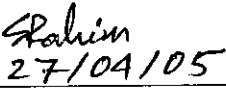
APRIL 2005

The project titled **Fractal Modeling of Electrocardiogram** Submitted by **Saleh Md. Abdullah Khan Mojahedy** Roll No: **9506123 P** Session **1994-95-96** has been accepted as satisfactory in partial fulfillment of the requirements for the degree of **Master of Engineering in Electrical and Electronic Engineering** on **27 April 2005.**

BOARD OF EXAMINERS
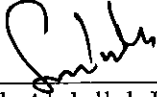
1.   Dr. Md. Aynal Haque   Chairman
Associate Professor
Dept. of EEE, BUET, Dhaka

2.   Dr. S. Shahnawaz Ahmed   Member
Professor
Dept. of EEE, BUET, Dhaka

3.   Dr. Newaz Muhammad Syfur Rahim   Member
Associate Professor
Dept. of EEE, BUET, Dhaka

CANDIDATE'S DECLARATION

It is hereby declared that this project or any part of it has not been submitted elsewhere for the award of any degree or diploma.

_____
Saleh Md. Abdullah Khan Mojahedy

ACKNOWLEDGEMENTS

CONTENTS

ABSTRACT

Modeling of data plays an important role in understanding the underlying physical mechanism generating and controlling the data. Fractal modeling is a nonlinear nonparametric model used to understand the characteristics of a signal. Fractal theory calculates the dimension of a signal, termed as fractal dimension (FD), considering the signal as an object. FD may be non-integer.

This project describes the application of fractal modeling to electrocardiogram (ECG) with an ultimate aim of detecting cardiac abnormality. The widely used four techniques such as box counting (BC), rescaled range (RS), relative dispersion (RD) and Fourier methods are applied to calculate the FD of data. In order to determine the dependency of a method on data length and other factors, the FD of a signal of known FD is calculated first by each method. It is seen that a data length of around 6000-7000 provides best result. The rescaled range method provides best result even for a lower data length.

The FD is calculated for instantaneous heart rate (IHR) calculated from measured ECG of 5 healthy adult subjects. It is observed that the rescaled range method gives consistent FD for all data sets. In order to find the applicability of fractal theory to detect cardiac condition the normal data is corrupted with random noise and thus noisy IHR time series is obtained. The FD of noisy IHR is statistically different from that of normal IHR. It indicates that fractal modeling can be applied to ECG to detect different types of pathology present in cardiac system.

List of Figures

List of Tables

Fractal Modeling of Electrocardiogram

Saleh Md. Abdullah Khan Mojahedy

Master of Engineering in Electrical and Electronic Engineering
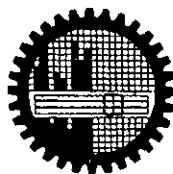
DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY

APRIL 2005

Chapter 1
INTRODUCTION
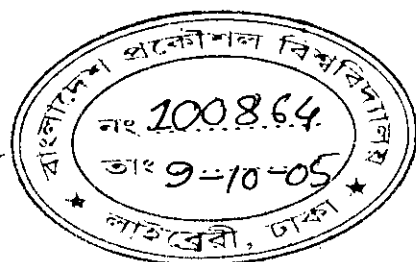
1.1 Problem Overview

The electrocardiogram (ECG) represents the electrical activity of heart. ECG is often used as a diagnostic tool to detect cardiac and other associated abnormality. Both short-term and long-term ECG is used for diagnosis. In long-term monitoring, physicians often use Holter devices attached with a subject for 24 hours to continuously monitor ECG. The ECG is normally printed in a long paper and the physicians look for some abnormal beats present in the ECG. This is a tedious job for a physician as he has to look for more than 100,000 cycles of ECG. There may be error due to fatigue and human over-looking in looking for such a huge number of cycles. So, there is a need for automatic software based recognition of ECG characteristic points to help the physicians in diagnosing the ECG.

The heart rate (HR) is calculated from the ECG and is also used as a quick diagnostic tool. It is already proved that ECG is a non-stationary signal while the HR is stationary and nonlinear [1]. So, far different methods have been developed for finding the characteristic points of ECG and the researchers are continuously trying to develop a quick method for the diagnosis of cardiac abnormality. The ultimate aim of the researchers is to develop a method that can differentiate abnormal ECG from a normal one. Fractal theory is a new concept which may be applied to detect cardiac abnormality. Fractal theory states that any signal can be viewed as an object and its calculated dimension, termed as fractal dimension (FD), can be used to understand the internal mechanism of a system generating and controlling the signal. There are different methods to calculate FD of a signal. Although FD is continuously being applied to many fields, it is yet to test its applicability in detecting cardiac conditions. Hence, the problem that should be addressed include the following:

- Out of many methods, which one is best to calculate FD of ECG or HR?

- Is ECG or HR is fractal in nature?
- How FD can be used as a diagnostic tool of cardiac abnormalities?

The above problems have been addressed in this work.

1.2 Historical Background

In recent years, a particular attention has arisen in noninvasive medical diagnostic procedures. Because biosignals recorded on the body surface reflect the internal behavior and status of the organism or its parts, they are ideally suited to provide essential information of these organs to the clinicians without any invasive measures. Here some questions arise. How are the recorded time courses of the signals to be interpreted with regard to a diagnostic decision? What are the essential features and in what code is the information hidden in the signals? These questions can be answered by the recognition of detailed signal patterns. Since signal processing is concerned with the mathematical representation of a signal and the algorithmic operation carried out to extract the information present in the signal, the systems generating and controlling the signal must be understood first. A system is any physical device that performs an operation on a signal.

Modeling is the mathematical description of a physical system. Once a model of a system is available, one can analyze its overall behavior and predict the response of its output to different inputs. The difficulty, however, lies in determining such a mathematical description. Indeed, although one can derive a model for some simple systems, such as electric motors, there are many systems whose outright complexity makes such a task impossible. In such cases, one has to resort to the numerical techniques of system identification [2]. In attempting to determine empirically the mathematical descriptions associated with systems, modeling techniques are used. The classical theory of time series analysis has been well developed over the past two decades, and excellent accounts of this theory are available [3], [4]. An important assumption that is made in the classical theory is that the structure of the series can be described by a linear model.

Two broad categories of methods are used in the modeling of time series data, namely, parametric and nonparametric methods. Nonparametric methods make no assumption about how the data are generated, and mainly use Fourier based approach and periodogram. Nonparametric methods are relatively simple, well understood and easy to compute using the FFT algorithm. However, these methods require the availability of long data records in order to obtain necessary frequency resolution required in many applications. Furthermore, these methods suffer from the leakage effects due to windowing that are inherent in finite-length data records. Often, the leakage masks weak signals that are present in the data.

On the other hand, parametric methods of data modeling eliminate the need of window functions. As a consequence, parametric methods avoid the problem of leakage and provide better resolution than do the FFT based nonparametric methods. This is specially true in applications where short data records are available due to time-variant of transient phenomena.

In all areas of medical research there is a common physiological theme. Complexity is the salient feature shared by all such systems; a feature that is attracting more and more attention in physical systems as well. Until recently, scientists have assumed that understanding such systems in different contexts, or even understanding various physiological systems in the same organism, would require completely different models. One of the most exciting prospects for the new scaling addressed herein is that it may well provide a unifying theme to many investigations, which up until now have been considered unrelated. The swirling spiral of the cochlea and the finely branched structure of the bronchial tree suggest the complex interrelations among biological development, form, and function. Relationships that depend on scale can have profound implication for physiology. Consider, for example, one of the standard problems in scaling: mass increases as the cube of a typical length, but surface area increases only as the square. Accordingly, if one species is twice the size of another likely to be eight times heavier, but have only four times the surface area. Thus the larger plants and animals must compensate for their bulk; respiration depends on surface area for the exchange of gases, as does cooling by evaporation

from the skin, and nutrition by absorption through membranes. One way to add surface to a given volume is to make the exterior more irregular, as with branches and leaves; another way is to hollow out the interior. The human lung, with about 300 million alveoli, approaches the more favorable ratio of surface to volume enjoyed by our evolutional ancestors, the single-celled microbes.

The classical scaling concepts in biology, while of great importance, are not capable of accounting for the irregular surfaces and structures seen in hearts, lungs, intestines and brains. Experiments suggest that biological processes are not continuous, homogeneous, and regular, but rather are discontinuous, inhomogeneous, and irregular [5]. Thus, a new way of analyzing such processes is required. This new perspective is that of fractal geometry and fractal statistics.

A basic theme in science is that of the invariance or symmetry of laws with respect to various types of transformations that may be performed in ordinary space and time, or even in certain abstract mathematical spaces. In dynamics there are the familiar symmetries in space and time which lead to conservation laws. For example, translational invariance implies conservation of linear momentum, whereas invariance with respect to rotation implies conservation of angular momentum. Using the known symmetries of a given system provides a framework to which the details of structure and motion must conform. In some cases, it is not known initially just what symmetry a system possesses because of lack of fundamental information about interactions among the elements of the system. However, a time-honored strategy for constructing the equations of motion for a complex system i.e., to assume a symmetry property, and then determine the forces required to maintain this symmetry over time. This approach is the variational principle, and it has been applied to such quantities as the system's entropy, mass, and energy. This strategy for determining the dynamic properties of a system has been very useful in bioengineering, where concepts such as efficiency have been introduced. A biosystem can be assumed to operate so as to maximize its efficiency. The equations that produce this effect are sought through a variational principle. In other areas, where it might at first seem that symmetry is of little use, it turns out that there is

often an unnoticed symmetry framework present that determines system behavior. As an example, the motions of individual molecules in an equilibrium gas in the lungs are random, and symmetry seems to be the last thing one would think about to describe the gas exchange process. Considered another way, however, equivalent randomness occurs everywhere in the gas, so there is a statistical symmetry with respect to spatial translation. The entire gas can be generated by taking a small element of volume, one that contains many particles with an equilibrium distributed of velocities characterized by a constant temperature, and copying it in all directions throughout the given volume. In this way, a disorganized system may be self-similar under scale changes and can be thought of as being formed by repetitious translations of a generating element. The idea of forming a large set of points from a smaller generating set is used below to provide one definition of dimension for sets of points.

In the 1960s, Benoit Mandelbrot began discussing a new geometry of nature [6], one that embraces the irregular shapes of objects such as coastlines, lightning bolts, cloud surfaces, and molecular trajectories. It was soon realized that these geometrical ideas could be applied in other areas, including non-traditional ones [7], from outside the physical sciences. A common feature of these objects, which Mandelbrot called fractals, is that their boundaries are so irregular that it is not easy to understand how to apply simple metrical ideas and operations to them. Such fundamental concepts a dimension and length measurement must be generalized. Therefore, let us consider some of the metric peculiarities of a few unusual mathematical objects, which we subsequently use to describe some biomedical systems.

Fractal modeling is an alternative way to describe some of the self similar or self affine signals and is the nonparametric modeling of the signal. First coined by Mandelbrot the fractal analysis involves the dimensional study of the objects. A fractal is an object whose Housdorff dimension is greater than its Euclidean dimension [6]. Demonstrating the fractal nature of a signal is valuable, since it provokes the development of new ways for discovering how the system works. It

has been proven that fractal geometry can play an important role in the analysis of natural phenomenon. Fractal modeling technique has been applied to a variety of fields ranging from one-dimensional signal study [8] to image processing [9], from sea-scattered radar signal [10] to medical image analysis [11]. Fractal dimension is also extensively used in understanding of the characteristics of strange attractors, the case of a chaotic signal. Some works on the modeling of biosignals using fractal model have recently been suggested [12], [13].

The topics of fractals are central concepts in a relatively new branch of science called nonlinear dynamics. Fractals are highly irregular objects and, as a result, have noninteger or fractional dimensions. The internal look alike property of fractals is referred to as self-similarity. The more closely one inspects a fractal, the more detail one sees, and the small scale structure is similar to, though not necessarily identical to, the larger scale form. As a consequence, fractal objects do not have a well-defined length. The measured length of a fractal line will vary depending on the size of the ruler used.

Fractal geometry is widespread in nature: coastlines, clouds, lightning flashes, and winding rivers, to name but a few. Example of fractal-like anatomies include the vascular system, the His-Purkinje network, the tracheobronchial tree, as well as the folds of the small bowel and brain. Fractal was adopted in the analysis of cardiac electrophysiology due to the self-similarity of cardiac electrical conduction system [14].

1.3 Aim of the Project

The ECG reflects the electrical activity of heart. Heart rate (HR), calculated from ECG, is a major indicator of cardiac conditions. Moreover, it also reflects the conditions of other functional physiological systems, such as autonomic nervous and respiratory systems. Biological data are often modeled with an ultimate aim of differentiating the data generated by a normal organ with that by a pathologic one. The objective of this research is to apply fractal dimension analysis for the modeling of HR time series data. To test the applicability of methods to calculate fractal

dimension, the methods are first applied to a simulated time series data of known fractal dimension. The methods are also applied to calculate the fractal dimension of normal and pathologic HR data with an ultimate aim to differentiate abnormal HR from a normal one.

1.4 Organization of the Dissertation

Four methods to calculate fractal dimension of HR are considered. Chapter 2 represents ECG and its wave shapes in different leads. Chapter 3 describes briefly the methods used to calculate fractal dimension (FD). FD calculation of simulated data by various methods is presented in chapter 4. Chapter 5 describes the calculation of FD of normal and abnormal HR. Chapter 6 represents the conclusions of the findings.

Chapter 2

ELECTROCARDIOGRAM

2.1 Cardiac Activity

The heart is a muscular organ located in the chest (thoracic) cavity and covered by a fibrous sac, the pericardium. It acts as the pump used to force the blood through the cardiovascular system. Its walls are composed primarily of cardiac muscle (myocardium). The inner surface of the myocardium, i.e., the surface in contact with the blood within the heart chambers is lined by a thin layer of epithelial cells (endothelium).

The human heart is divided longitudinally into right and left halves, each consisting of two chambers, an atrium and a ventricle. The cavities of the atrium and ventricle on each side of the heart communicate with each other but the right chambers do not communicate directly with those on the left. Thus the right and left ventricles are distinct.

The heart is contained in the pericardium, a membranous sac consisting of an external layer of dense fibrous tissue and an inner serous layer that surrounds the heart directly .The base of the pericardium is attached to the central tendon of the diaphragm and its cavity contains a thin serous liquid. The two sides of the heart are separated by the septum or dividing wall of tissue .The septum also includes the atrioventricular node (AV node), which plays a role in the electrical conduction through the cardiac muscles.

Function of the four chambers of the heart is different. The right atrium is elongated and lies between the inferior (lower) and superior (upper) vena cava. Its interior is complex; the anterior (front) wall being very rough, whereas the posterior (rear) wall (which form a part of the septum) and the remaining walls are smooth. At the junction of the right atrium and superior vena cava is situated the sinoatrial node

(SA node), which is the pacemaker or initiator of the electrical impulses that excites the heart. The right ventricle is situated below and to the left of the right atrium.

Communication between the atrium and ventricle is accomplished only via the AV node. Since the ventricle has to perform a more powerful pumping action, its walls are thicker than those of the atrium and its surfaces are ridged. Between the anterior wall of the ventricle and the septum is muscular ridge that is a part of the heart's electrical conduction system, known as the Bundle of His. At the junction of the right and left atrium and the right ventricle of the septum, there is another node, the atrioventricular node. The Bundle of His is attached to this node. Figure 2.1 shows the anatomy of human heart.
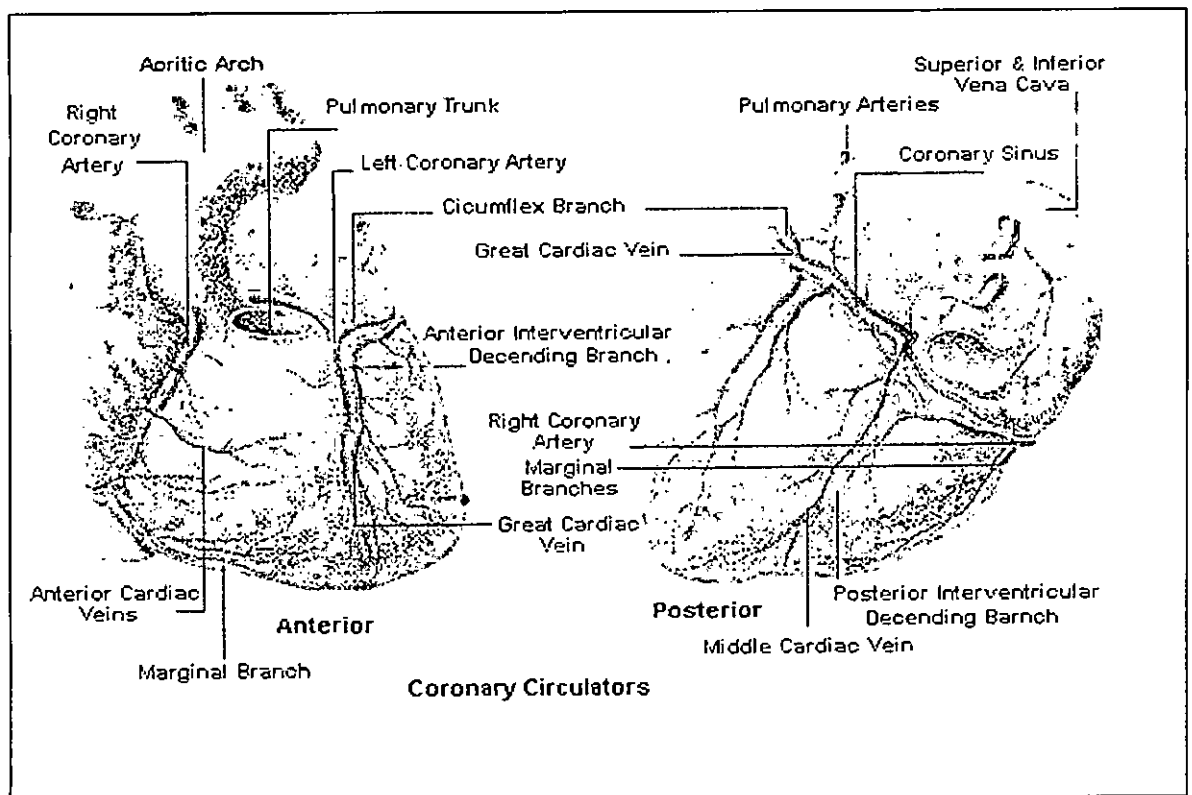
Fig. 2.1 Anatomy of human heart

The right atrium and left ventricle are joined by a fibrous tissue known as the atrioventricular ring, to which are attached the three cups of the tricuspid valve, which is the connecting valve between the two chambers.

The left atrium is smaller than the right atrium. Entry to it is through four pulmonary veins. The walls of the chambers are fairly smooth. It is joined to the left ventricle through the mitral valve, sometimes called the bicuspid valve since it consists of two cups.

The left ventricle is considered the most important chamber, for this is the power pump for all the systemic circulation. Its walls are approximately three times as thick as the walls of the right ventricle because of this function. Conduction to the left ventricle is through the left Bundle branch, which is the ventricular muscle on the septum side.

A discussion about the anatomy of the heart and of the electrical excitation system necessary to produce and control the muscular contraction should help to round out the background materials needed for understanding of cardiac dynamics. The ECG, which is a record of bio-potential events, can be used to show the relationship that exists between the electrical and mechanical events of the heart.

Figure 2.2 depicts the blood circulation within the human heart. Blood enters the heart on the right side through two main veins: the superior vena cava, which leads from the body's upper extremities and the inferior vena cava leading from the body's organs extremities below the heart. The right atrium pumps it into right ventricle. The right ventricle pumps the blood through the pulmonary artery to the lungs where it is oxygenated. The oxygen-enriched blood then enters the left atrium through the pulmonary veins, from which it is pumped into the left ventricle. The output from the ventricle is then pumped the aorta into the arteries to circulate through the body.

2.2   The Electrocardiogram

The electrocardiogram (ECG) is the graphical recording or display of the time variant biopotentials produced by the myocardium during the cardiac cycle. The

term electrocardiograph means the instrument used for this type of recording and electrocardiography means the technique used for such measurement [1].
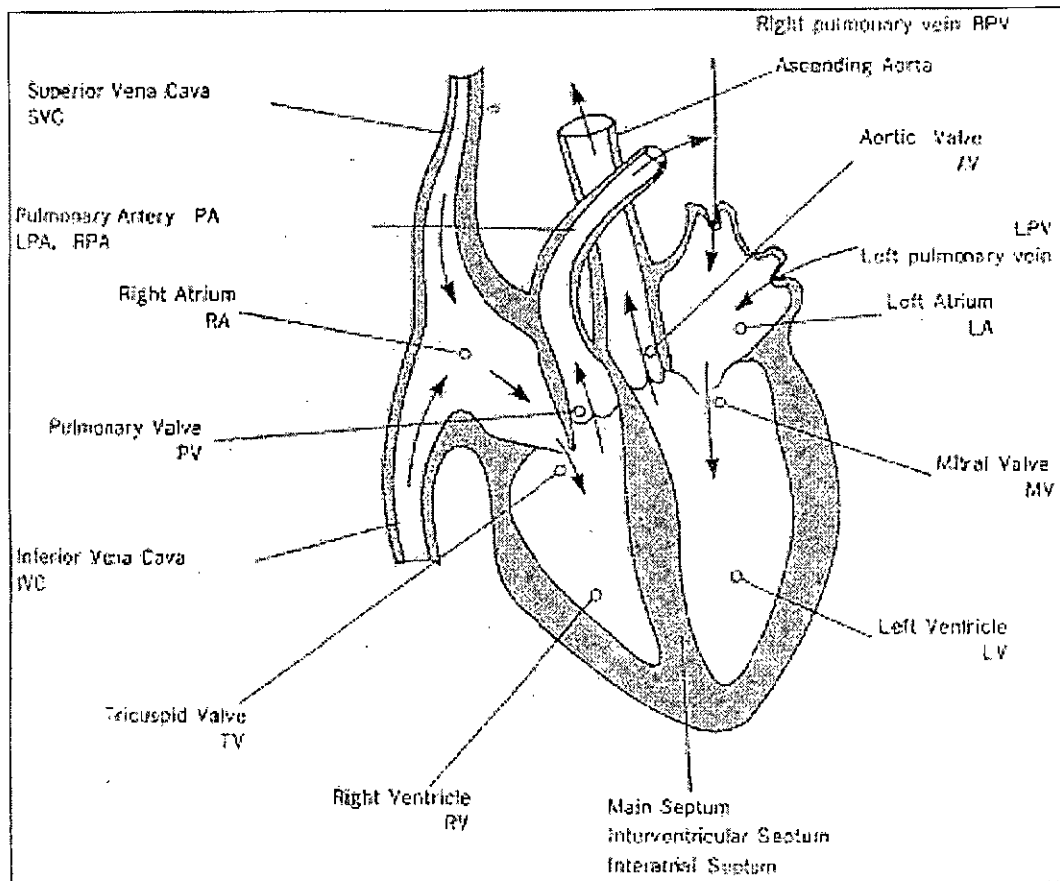
Fig. 2.2 Blood circulation in human heart

Each action potential in the heart originates near the top of the right atrium at a point called the pacemaker or SA node. The pacemaker is a group of specialized cells that spontaneously generate action potentials at a regular rate, although the rate is controlled by innervation. To initiate the heartbeat, the action potentials generated by the pacemaker propagate in all direction along the surface of both atria. The wave of activation travels parallel to the surface of the atria toward the junction of the atria and the ventricles. The wave terminates at a point near the center of the heart, called the AV node. At this point, some special fibers act as a "delay line" to provide proper timing between the action of the atria and the ventricles. Once the electrical excitation has passed through the delay line, it is rapidly spread to all parts of both ventricles by the bundle of His .The fibers in this bundle, called Purkinje

fibers, divide into two branches to initiate action potentials simultaneously in the powerful musculature of the two ventricles. The wave front in the ventricles does not follow along the surface but is perpendicular to it and moves from the inside to outside of the ventricular wall, terminating at the tip or apex of the heart. As indicated earlier, a wave of repolarization follows the depolarization wave by about 0.2 to 0.4 second. This repolarization, however, is not initiated from neighboring muscle cells but occurs as each cell returns to its resting potential independently.
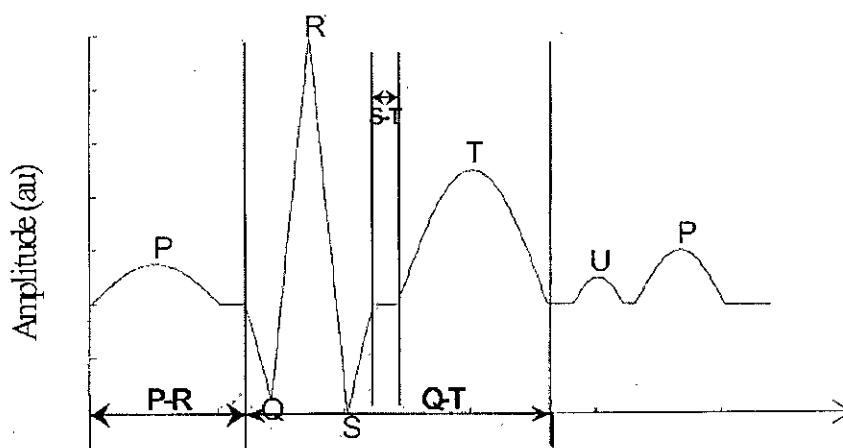
Fig. 2.3 A typical ECG wave

A typical pattern of ECG is provided in Fig. 2.3. Alphabetic designations have been given to each of the prominent features. These can be identified with events related to the action potential propagation pattern. To facilitate analysis, the horizontal segment of this waveform preceding the P wave is designated as the baseline or the isopotential line. The P wave represents depolarization of the atrial musculature. The QRS complex is the combined result of the repolarization of the atria and the depolarization of the ventricles, which occur almost simultaneously. The T wave is the wave of ventricular repolarization. The U wave, if present, is generally believed to be the result of after-potentials in the ventricular muscle. The P-Q interval represents the time during which the excitation wave is delayed in the fibers near the AV node. The shape and polarity of each of these features vary with the location of the measuring electrodes with respect to the heart, and a cardiologist normally bases his diagnosis on readings take from several electrode locations.

The P-R Interval is taken from the start of the P wave to the start of the QRS complex. It is the time taken for depolarization to pass from the SA node via the atria, AV node and His-Purkinje system to the ventricles. The QRS represents the time taken for depolarization to pass through the His-Purkinje system and the ventricular muscles. It is prolonged with disease of the His-Purkinje system. The Q-T interval is taken from the start of the QRS complex to the end of the T wave. This represents the time taken to depolarize and repolarize the ventricles. The S-T segment is the period between the end of the QRS complex and the start of the T wave. All cells are normally depolarized during this phase. The ST segment is changed by pathology such as myocardial ischemia or pericarditis. The normal heart rate can vary from 60 to 120 beats per minute (BPM). For a heart rate of 120 BPM, the normal values of waves are as follows [1]:

Amplitudes:	P wave	0.25 mV
	R wave	1.60 mV
	Q wave	25% of R wave
	T wave	0.1 to 0.5 mV
Durations:	P-R interval	0.12 to 0.20 sec
	Q-T interval	0.35 to 0.44 sec
	S-T interval	0.05 to 0.15 sec
	P wave interval	0.11 sec
	QRS interval	0.09 sec

2.3 ECG Lead Systems

The term lead is used to indicate an ECG collected by a particular group of electrodes. Because ECG signal is measured from electrodes applied to the surface of the body, the waveform of this signal is varied dependent on the placement of the electrodes. Einthoven, who developed the first clinically usable ECG in 1903, selected anatomical sites for placing electrodes. He chose the two arms Right arm (RA), Left arm (LA) and left leg (LL), designating the three different recording schemes as: lead I (RA – LA), lead II (RA – LL), lead III (LA –LL). The formation of these 3 leads is shown in Fig. 2.4.
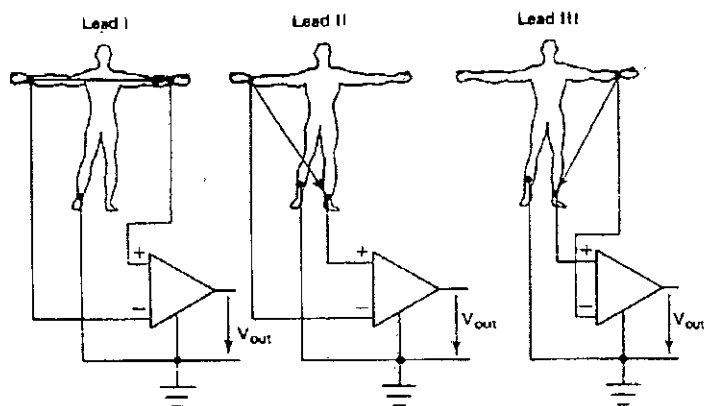
Fig. 2.4 Bipolar limb leads

Lead I, lead II and lead III are called bipolar limb leads. Although the electrodes are placed on the members, the latter act as electrolytic conductors and, in reality, the electrodes are electrically at the shoulders and left abdomen. Einthoven postulated that at any given instant of the cardiac cycle, the frontal plane representation of the electrical axis of the heart is a two dimensional vector. Further the ECG measured from any one of the three basic limb lists is a time-variant single-dimensional component of that vector. Einthoven also made that assumption that the heart (the origin of the vector) is near the center of an equilateral triangle, the apexes of which are the right and left shoulder and the crotch. By assuming that the ECG potentials at the shoulder is essentially same as the wrists and the potential at the crotch differ little from those at either ankle, he let the points of this triangle represents the electrode positions for the three limb leads. This triangle is known as the Einthoven triangle, is shown in Fig. 2.5. The sides of the triangle represent the lines along which the three projections of the ECG vector are measured. Based on this, Einthoven showed that the instantaneous voltage measured from any one of the three limb lead positions is approximately equal to the algebraic sum of the other two or that the vector sum of the projections on all three lines is equal to zero. For these to actually hold true, the polarity of the lead II measurement must be reversed.
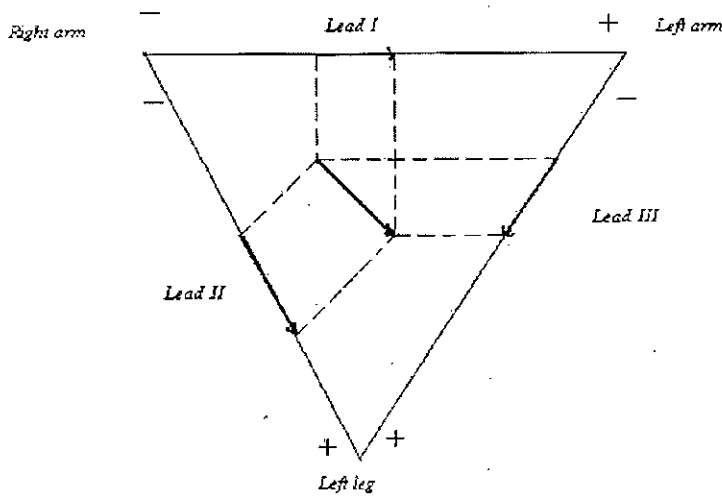
Fig. 2.5 The Einthoven triangle

V₁ Fourth intercostal space, at right sternal margin.

V₂ Fourth intercostal space, at left sternal margin.

V₃ Midway between V₂ and V₄.

V₄ Fifth intercostal space, at mid-clavicular line.

V₅ Same level as V₄, on anterior axillary line.

V₆ Same level as V₄, on mid-axillary line.
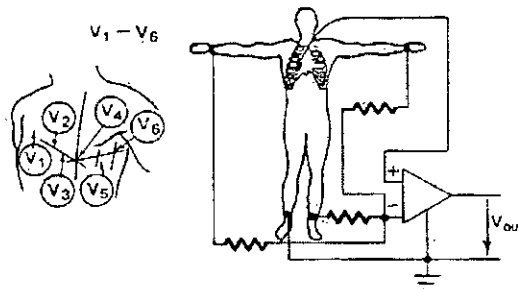
Fig. 2.6 Unipolar chest leads
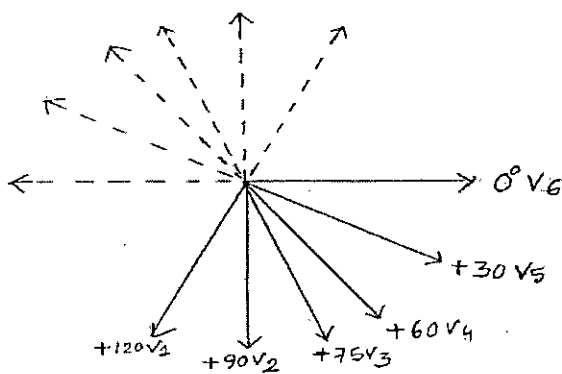
Fig. 2.7 Directions of the V leads of Wilson

It was soon recognized that the limb electrodes are quite distant from the heart and in order to examine its electrical activity more accurately, chest leads were

introduced. Since physiologists like to be able to examine events under a single electrode, Wilson (1934) introduced the V terminal, which constitutes an "indifferent" electrode. The V lead is formed by connecting resistors (5–100 kΩ) from a common point to the right arm, left arm, and left leg electrodes (Fig. 2.6). The "active" exploring electrode is placed on the chest. There are six standard unipolar V leads. Occasionally, additional sites on the right chest are used. The positioning directions of V leads are depicted in Fig. 2.7.

Goldberger (1942) introduced the augmented V leads (aV) which are also unipolar in nature. They are aVR, aVL and aVF as shown in Fig. 2.8. With the three standard, three augmented, and six V leads, the spatial direction of excitation and recovery of the heart chambers can be identified.
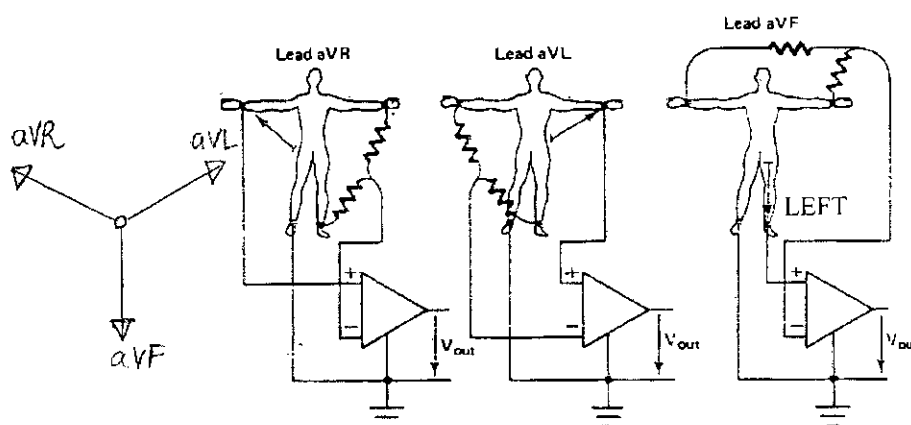
Fig. 2.8 Unipolar limb leads

At this point it may be through that there is an excessive number of leads for electrocardiography. From a mathematical viewpoint this is certainly true, since only three leads obtained from mutually perpendicular (orthogonal) axis are required to locate a vector in space. However, each ECG leads "sees" a different part of the heart. For example, the chest (V) leads provide a highly localized examination of the ventricular myocardium, and the activity that is recorded is the projection in the horizontal plane. The limb leads record the frontal-plane projection and are less highly localized. By recording a large number of leads it is possible to examine the ECG and estimate the axis of excitation and recovery by inspection, rater than by

having to plot amplitudes on the reference frames and graphically determine the vectors. Figure 2.9 shows the ECG as observed in different leads.
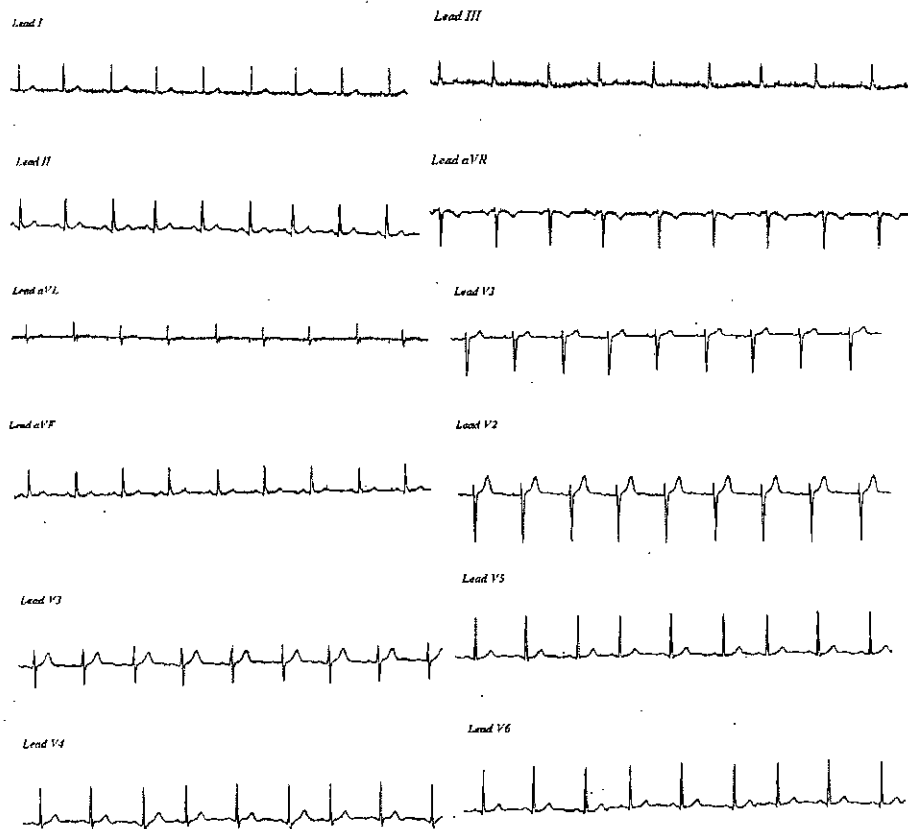
Fig. 2.9 ECG wave shapes in different leads

Beyond the lead systems already discussed, there are certain additional lead modifications that are of considerable use in the coronary care unit. The most widely used modification for ongoing ECG monitoring is the modification chest lead I (MCL I) also called the Marriott lead, named after its inventor. This lead system simulates the VI position with electrode placement as follows: positive electrode, fourth intercostal space, right sternal border, negative electrode just below the outer portion of the left clavicle, with ground just above anywhere but usually below the right clavicle. The monitor is set on lead I for bipolar tracing. Recordings obtained in this way are very useful in differentiating left ventricular ectopic rhythms from aberrant right ventricular or supraventricular rhythms. The former

situation usually necessitates prompt therapeutic action; the later is off less clinical significance.

Out of various leads, lead II is used for quick diagnosis. The typical ECG pattern as shown in Fig. 2.3 is of lead II. Heart rate is calculated from this lead.

Chapter 3

FRACTAL MODELING TECHNIQUES

3.1   Introduction

Fractal modeling is used to analyze self-similar data. When signals are self similar (self affine) or scale invariance the best way to model such signals is to use a fractal model. Many natural shapes such as coastlines, mountains and clouds are easily described by fractal models. If the graph representing the data is characterized by a continuous derivative then it is not difficult to understand that the graph has a dimension 1. The same is true if the graph is of bound variation. However it is possible for a continuous function to be sufficiently irregular to have a graph dimension strictly greater than 1. With a notion that complex objects are better described by fractal geometry than Euclidean geometry fractal concepts are finding application not only in specialized fields of science and engineering but also in many familiar fields. Fractal theory has achieved greatest success in medical imagery, texture of images of all kinds and speech signals. In this chapter the techniques to determine the fractal dimension are described.

There is no unanimously accepted method of generating fractal signal as well as calculating the fractal dimension of a time series data. Four methods namely Box counting, rescaled range, relative dispersion and Fourier are widely used to find the fractal dimension (FD) of time series data [15], [16]. The methods are briefly described below:

3.2   Box Counting Method

The idea of box counting method crop up in mind from the perception that as the fractal dimension of a graph increases the graph tends to fill the plane on which it is drawn. For one-dimensional data it is easy to say that FD will lie between 1 and 2. One would naturally assign a low value of D to the lower stochastic fractal, since it

would look like a straight line. In Fig. 3.1 a large part of the two-dimensional surface is filled by the curve, indicating a value of D closer to 2.
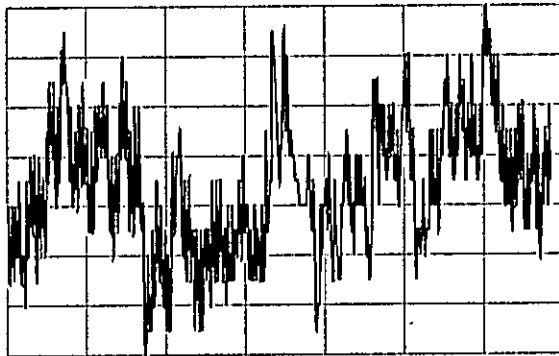
Fig. 3.1 Example of a fractal signal
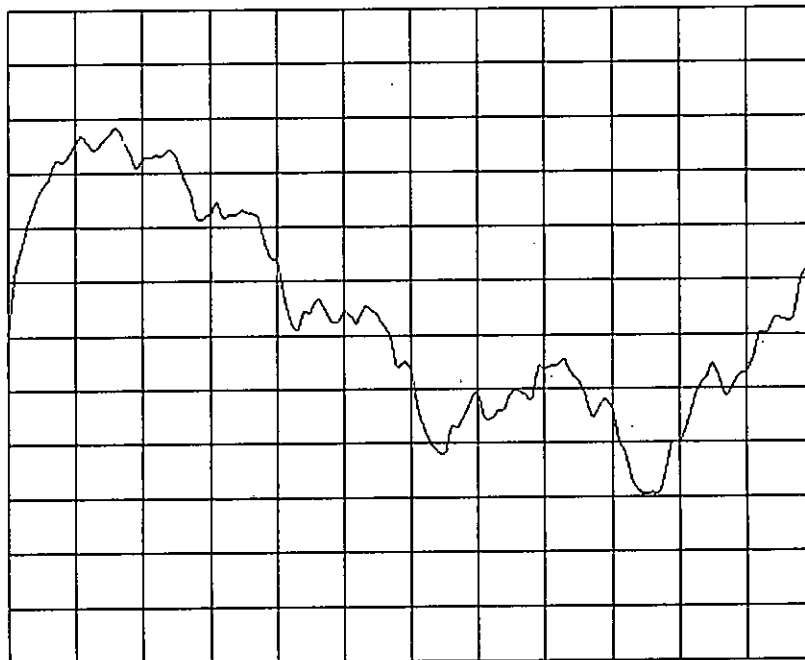
Fig. 3.2 Example of box counting

The Box dimension of a curve is widely used and easy to calculate. Box dimension is defined by–

$$D = \lim_{\delta \to 0} \frac{\log N_\delta}{-\log \delta} \qquad\qquad 3.1$$

Where N_δ is the smallest number of boxes that are needed to include the total graph and δ represents the box size. Figure 3.2 represents an example of box counting. Here, 27 boxes are required for a box size of 1/12. To calculate the box dimension the following steps are followed:

- Normalize the data so that it fit into a one by one square.

- Choose a set of box sizes so that δ tends to zero slowly. A sufficient condition is that the ratio of logarithm of two successive box sizes tends to unity. Theoretically, the FD is independent of resolution (sampling rate) of the series because of the scale-invariant property of fractals i.e. δ can be infinitely small.

- Cover the unit square with boxes and count the number of boxes required to include all the data.

- Plot log N_δ against $-\log \delta$ for various values of δ. If the time series data is fractal then the curve is a straight line. The estimate of FD is given by the slope of the straight line.

The procedural steps to calculate FD by box method are provided in appendix A1.

3.3   Rescaled Range Method

Edwin Hurst, a hydrology engineer, first used this method to solve his statistical problem of designing the Aswan dam on the Nile River. He has the problem to estimate the height of the dam capable of containing hold all the incoming water flow under a constant outflow condition. When the inflow of dam is a stationary random variable, the water level of the lake is the integral of the differences of the inflow and the constant outflow. That is each data contributes to the integral whether lowering or augmenting the result.

In our analysis we dealt with discrete data which is beyond the scope of prediction. Here a data has no effect on the successive data (as we assume when generating data

and which is a condition of being fractal signal). To get the FD by this method first we segmented the total data at our hand into some convenient equal parts of length Δt and mark starting point of each segment as t_0, t_1, t_2 etc. R/S is calculated for a set of predefined data span, u for each data segment starting at t_0, t_1, t_2 etc. Range, R(u) is defined as the maximum value minus minimum value of the integral of the differences of each data from the arithmetic mean of a span of data averaged at different starting point.

$$R(u)=\text{Max } \Sigma[X_i - <X>]-\text{Min } \Sigma[X_i - <X>] \qquad 3.2$$

$$i=t_0 \text{ to } t_0+u \text{ or}$$

$$i=t_1 \text{ to } t_1+u \text{ etc.}$$

When range is divided by the standard deviation, S(u) of the corresponding data span then we get R/S. For each u we have a number (equal to the number of segments) of R/S and for our ease to draw graph we took the average of those data and found the FD from the following relationship.

$$\text{Log}[R(u)/S(u)]=c+(2-D)\log(u) \qquad 3.3$$

The flow chart to calculate FD by rescaled range method is provided in appendix A2.

3.4 Relative Dispersion Method

In this analysis we make use of the fact that variance of a variable changes as the measurement resolution changes. Relative dispersion (RD) is defined as the standard deviation divided by the mean. In time series analysis at highest level of resolution variance is maximum. As the resolution may be decreased by averaging the signal over pairs of consecutive values the variance is decreased by a factor $1/2^{0.5}$. This happens because as a consequence to averaging mean remains same but standard deviation is reduced by the same factor. When we average two consecutive values repeatedly we actually increase the bin size geometrically starting from 2 to some desired value (bin sizes are 2, 4, 8 etc). By calculating RD for different bin sizes, n, and plotting in a log-log paper FD can be easily estimated from the following relationship.

$$\log(RD) = \log(RD_0) + (1-D)\log(n/n_0) \qquad\qquad 3.4$$

where RD_0 is RD for some reference bin size n_0. The above is an equation of straight line whose slope is directly related to Fractal Dimension of that time series. We use least square method to have the best-fit straight line. Appendix A3 provides the flow chart to calculate FD by RD method.

3.5   Fourier Method

There is a relationship between the FD and power law index (α) for time series. Results based on spectral analyses are the most common means of qualifying irregular time series. When the power spectral density (PSD) has appreciably more power at lower frequencies than higher frequencies ($\alpha > 0$) and it has no eminent peak then there is the natural tendency to approximate it by using a power law relationship $P(f) = f^{\alpha}$. A typical power spectrum of a fractal signal is given in Fig. 3.3 which exhibits a behavior which consistent with the power law relationship. Power index is given by the slope of the spectral density measured in terms of logarithmic scale. The power index is one of the measure of the irregularity of time series data. The power law index is related to FD of a time series by

$$D = (5-\alpha)/2 \qquad \text{for } 1 < D < 2 \qquad\qquad 3.5$$

In other words FD can be determined directly from its power spectrum. The slope of the power spectrum is determined by fitting a straight line using least square method. Since we are concerned with discrete data we used FFT and estimated the Power Spectral Density of discrete-time signal vector using Welch's averaged, modified periodogram method. PSD calculation flow chart is provided in appendix A4.
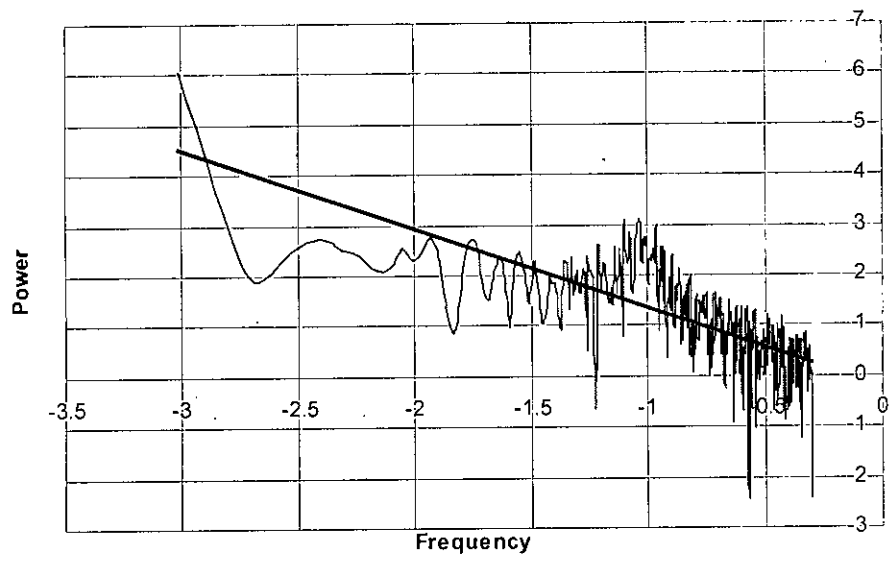
Fig. 3.3 Typical PSD of a fractal signal

Chapter 4

FRACTAL DIMENSION OF SIMULATED DATA

4.1   Introduction

Since fractal is a relatively new concept in the analysis of time series data, the validity of each method to determine fractal dimension is of great concern. Any method to calculate fractal dimension must be applied first to a signal of known fractal dimension. This chapter describes the generation of fractal signal and application of each of the methods described in the previous chapter (Chapter 3) to see the effect of method, data length, etc. on the fractal dimension.

4.2   Description of Data

There is no single unanimously accepted method to generate a signal of known fractal dimension. Two algorithms are normally used to generate a fractional Brownian motion, fBm, with a parameter H, that defines the fractal dimension of the signal. The first algorithm uses the well-known ordinary Brownian motion, such as a random walk of diffusion particles, is a fractional Brownian motion with H = 0.5. The successive differences between points of an fBm are called fractional Brownian noise, fBn. Because determining H or D from fBm is difficult, the fBn is needed to estimate fractal dimension. Fractional Brownian noise signals with positive correlations have H>0.5 and that with negative correlations have H<0.5. The other algorithm which are applicable to generate a signal with fractal dimension H in the range of 1 to 2 uses the familiar Weierstrass's equation [17]. Since our objective is to calculate fractal dimension of heart rate and it is assumed to have a value between 1 and 2, the second algorithm is used here to generate fractal signal. The Weierstrass's equation is reproduced below:

$$f(t) = \sum_{k=1}^{\infty} \lambda^{(H-2)k} \sin(\lambda^k t) \qquad\qquad 4.1$$

where $1<H<2$ and $\lambda>1$. This function is nowhere differentiable and has a fractal dimension H.
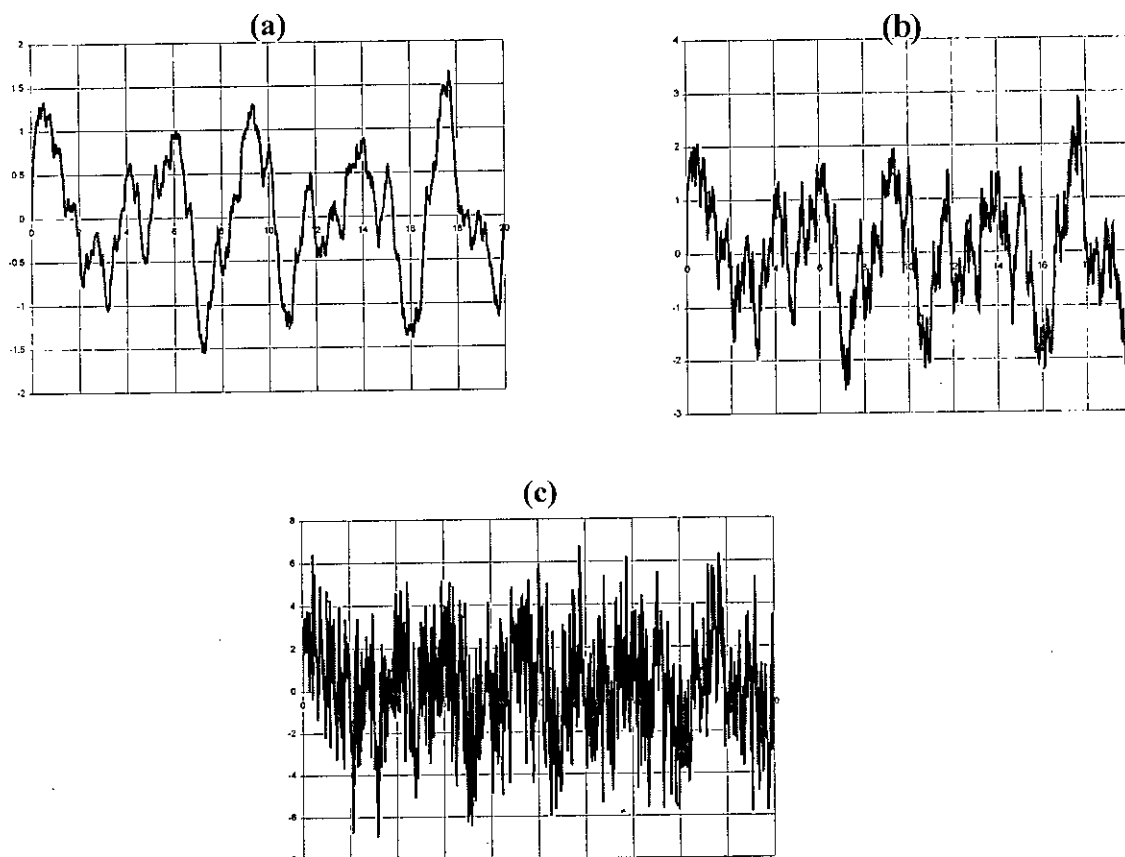
Fig. 4.1 Fractal signal generated using Weierstrass's equation for $\lambda=1.5$ and fractal dimension H of (a) 1.1 (b) 1.5 (c) 1.9

Figure 4.1 shows the graphical representation of the signals for 3 different values of fractal dimensions.

4.3   Estimation of Fractal Dimension

We applied four methods as previously described to estimate FD on generated data with s=1.5 & $\lambda=1.5$. Since the calculation of PSD using FFT requires a data length which must be a power of 2, we have considered same data length type for all methods. The dimensions were estimated for different data length by each method to observe the convergence, i.e., to estimate the data length to have an FD of 1.5. The results are depicted in Table 4.1.

Table 4.1 FD of simulated data by different methods

Data length	BC method	RS method	RD method	Fourier method
1024	1.132	1.781	1.540	1.251
2048	1.251	1.702	1.529	1.350
4096	1.412	1.581	1.538	1.435
8192	1.601	1.610	1.510	1.518

4.4   Discussion

From the obtained results we see that in case of box counting method there is a tendency of increasing FD with the increase in data length and the best result is found when data length is 6000. It is interesting to notice that rescaled range method also gives best result when data length is 6000 but there in no monotonous variation in estimated FD in this method. The relative dispersion method shows a tendency of estimating lower FD as data length becomes higher. In this both 4096 and 8192 data lengths give acceptable results. Fourier method most unexpectedly does not provide us with an appreciable result. Seeing the trend of the result we may expect that a data length of 7000 or some thing more will give a usable result. But due to limitation of the software we used we discarded the idea of testing Fourier method with a huge quantity of data. Figure 4.2 supports our expectation. In this figure, the abscissa indicates the data length and the ordinate indicates the estimated FD.
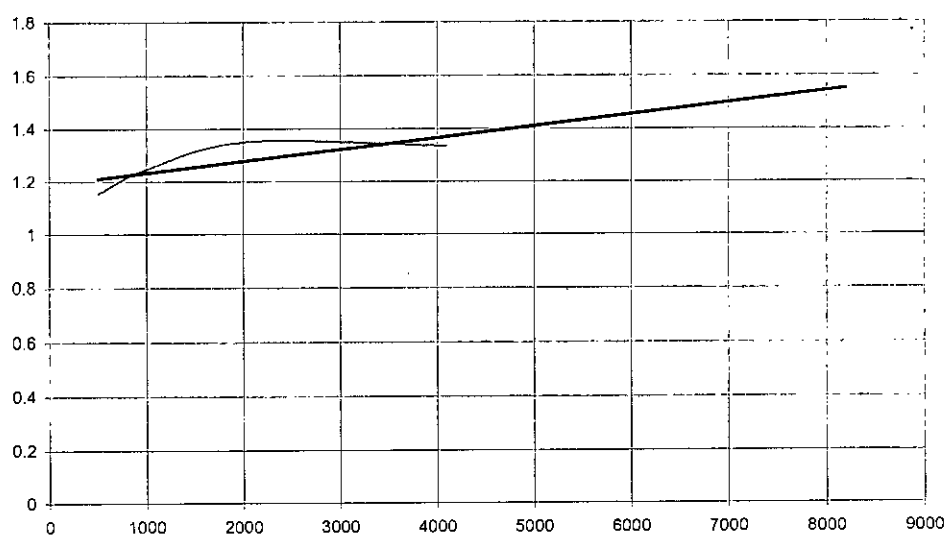
Fig. 4.2 Extrapolation of FD by Fourier method for different data length

Chapter 5

FRACTAL DIMENSION OF HEART RATE

5.1   Introduction

Cardiac activities can be assessed based on Mean Heart Rate (MHR) and Instantaneous Heart Rate (IHR). The heart can vary widely even in short period of time. The MHR can't keep track of these momentary heart rate fluctuations. IHR throws more light on this sudden change in heart rate and we in our thesis dealt with IHR. This chapter describes the calculation of IHR and the application of different methods to calculate the fractal dimension of normal and pathologic IHR time series data.

5.2   Data Description

The IHR was calculated from ECGs. The ECGs were stored in a microcomputer through a 12-bit A/D converter. The digitized signals were temporarily saved on the hard disk driver and the transferred to a file server. The frequency at which the signals are to be sampled may cause errors in the calculation of IHR. Due to the limited capacity of the computer memory, the signals were sampled at a lower frequency of 50 Hz. On the other hand, in order to calculate the IHR with accuracy of less than one beat/min (bpm), the sampling frequency (f) which satisfies the following relation is needed [18],

$$f \geq 2(60 - \Delta t)/(\Delta t)^2 \qquad\qquad 5.1$$

Where $\Delta t = 60/IHR$. For example, for a maximum IHR value of 120 bpm, the minimum sampling frequency is 476 Hz. The signals sampled at 50 Hz were restored by the sinc function,

$$x(t) = \sum_{n=-\infty}^{\infty} x(nt) \frac{\sin[2\pi f_M (t - nT)]}{2\pi f_M (t - nT)} \qquad\qquad 5.2$$

Where T is the sampling interval (i.e. 20 msec) and F_M is Nyquist reflection frequency (25 Hz). In order to minimize time of calculation, the signal restoration was made only for the period of 40 msec when three points sampled were found to

constitute a peak of ECG corresponding to each heartbeat. After restoration of the peak wave, the maximum point was looked for in restored wave, the time interval between two successive maximum points was determined to be 401(i.e. n=200) in a preliminary experiment. In the preliminary experiment, the ECGs were sampled at a frequency of 500 Hz and the IHR was calculated from the time intervals of two successive peaks (referred to as HR_{500}). Then signal values of every 80 sampling points (which correspond to the signals sampled at 50 Hz) were extracted and the ECG waves were restored by equation 5.2 changing n from 100 with a step of 5. The IHR was calculated for each wave restored with different values of n. It was found that the IHR so calculated was consistent with HR_{500} when n was more than 200 in all data sets.

The IHR was calculated from the ECG taken from 5 adult normal healthy subjects. The IHR is presented in Fig. 5.1 where the data is presented for 30 minutes. In the Figure, the alphabet in each panel indicates the person's identification and the number indicates the age of the person. The main characteristics of IHR calculated are provided in Table 5.1 where the mean and standard deviation (SD) of IHR along with total number of total beats for 30 minute IHR are provided.

Table 5.1 Mean, SD and N of calculated IHR

Subject ID	Mean (bpm)	SD (bpm)	N
H	70.1	3.2	2401
J	80.6	6.4	2362
K	75.2	6.2	2112
S	86.7	6.1	2482
T	82.4	4.6	2301

The IHR calculated is not evenly spaced in time. It does not make any sense to calculate FD of unevenly spaced heart rate data as long as we want to manipulate this dimension in taking decision about cardiac condition. So we extrapolated the IHR data to get data with equal spacing.
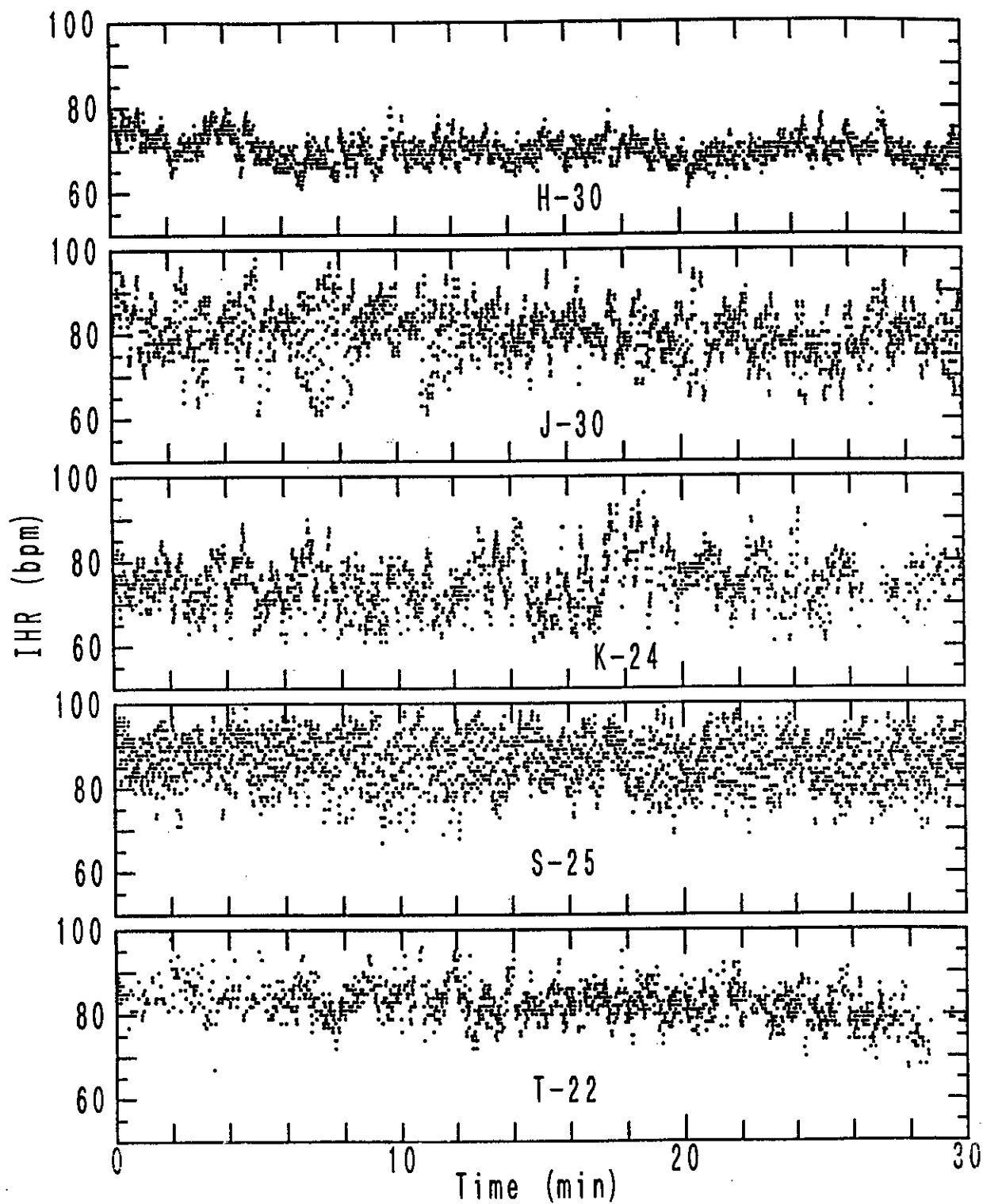
Fig. 5.1 Example of IHR calculated from the ECG of normal adult human beings

5.3   Fractal Dimension of Normal Heart Rate

After applying the four methods we got the following results with data length 2048.

Table 5.2 Fractal dimension of normal IHR with different methods

Data set	BC method	RS method	RD method	Fourier method
H	1.25	1.24	1.11	1.26
J	1.20	1.23	1.27	1.46
K	1.24	1.23	1.18	1.50
S	1.21	1.23	1.35	1.70
T	1.30	1.25	1.17	1.71

Fro the above table we notice that the rescaled range method gives almost constant FD for all data sets. This indicates that the small variation of SD of data set does not change FD. In the previous chapter it was found that this method provides best result for a data length of around 6000. Here we find consistent results for a much lower data length. On the other hand, the other 3 methods provide variable FD for different data set although all IHR are of normal subjects. The variation of FD by Fourier method is highest. Although f^{α} like characteristic is evident in HR time series [5], the poor performance of Fourier method may be due to the straight-line approximation of PSD. In the approximation process, most of the PSD in low and high frequencies are discarded which highly affects the results. Since we had IHR data length of around 2300 for all data sets, we have to confine our test with a data length of 2048. Had we been able to take a data length of around 6000, we could have better and consistent result by Fourier method as shown in the previous chapter (Chapter 4).

5.3   Fractal Dimension of Abnormal Heart Rate

Since the aim of applying fractal dimensional analysis is to detect cardiac abnormality, it is hoped that the FD of an abnormal HR time series will be different from that of normal one. There may be many types of abnormality. To find the FD

of an abnormal HR time series, the aforesaid 5 data sets were made "abnormal" by adding random white noise in each data set. The noise was generated using the 'random' function as built in the C programming language, with mean taken as 0 and variance as that of the original IHR. This noise incorporated data is termed as noisy IHR. The FD was calculated by rescaled range method. The results are presented in Table 5.3 together with the FD of normal IHR reproduced from Table 5.2.

Table 5.3 FD of normal and noisy IHR with rescaled range method

Data set	FD of normal IHR	FD of noisy IHR
H	1.24	1.08
J	1.23	1.08
K	1.23	1.06
S	1.23	1.01
T	1.25	1.06

The FD of normal IHR has mean of 1.236 and standard deviation of 0.009. In contrast, the FD of noisy IHR is 1.058 ± 0.029. To observe the difference between normal and noisy FD, analysis of variance (ANOVA) technique [19] was used. In the ANOVA technique, the f-value is calculated. It was seen that the normal FD is significantly difference from noisy FD, the f-value was found to be 74.55 and the value of $f_{critical}$ is 5.32 at 95% confidence level. This indicates that FD can differentiate abnormal IHR from normal one.

5.4 Discussion

The calculation of FD of normal and noisy HR time series data is presented in this chapter. The results show that the rescaled range method gives consistent FD for all 5 data sets of normal IHR. This finding is expected since fractal concept is based on the scale-invariance property of a time series data. It also indicates that the IHR is of fractal nature. Hence fractal theory can be used to model cardiac activity. This method also can detect abnormality of cardiac rhythm.

It is expected that different types of cardiac malfunctioning will give different FD if proper data length is selected. In this work, only one type of induced abnormality is considered. The incorporation of only random noise may not actually exhibit true abnormality of cardiac activity. This is a limitation of this work. To apply fractal theory to detect actual abnormality, we need to calculate FD of IHR of different types of pathology.

Chapter 6
CONCLUSIONS

6.1 Discussions

This works describes the application of four methods namely box counting, rescaled range, relative dispersion and Fourier methods to calculate the fractal dimension (FD) of IHR. The methods were first applied to calculate FD of a simulated data of known FD. The results show that each method is highly dependent of data length, more or less. The methods were also applied to calculate FD of normal IHR of. IHR was calculated from the ECG of 5 healthy adults and 5 data sets were constructed, each of 30 minutes duration. The results indicate that consistent FD was obtained with the rescaled range method. In this method, we tried to segment the total data into convenient numbers (usually 3 or 4). When span length approaches segment length we see that variation in RS is not marked and if we neglect that portion we get better result. The poor performance in terms of FD by other methods can be explained as follows:

- In our analysis for box counting method we used 1/100 as the lowest box dimension and 1/10 as the highest box dimension. These two are two extremities, any dimension beyond these range does not produce acceptable result. Although box counting method gives a straight line, the space-filling properties may not be truly representative as of scale-invariance for IHR. The FD by this method has a little variation, higher than rescaled range method but much less than other two methods.

- For relative dispersion (RD) method, we notice that in case of small bin size there is no appreciable variation in RD. But in high resolution, there is a sluggish variation of the slope and we calculated the slope of the curve neglecting the variable portion. This approximation may be one cause of variable FD.

- In case of Fourier method, we neglected very low frequency and very low power to have a straight-line curve of PSD. This may affect the result. Also, truly smooth PSD is not attainable, there is spurious peaks in frequencies other than the harmonic ones.

6.2   Future Perspectives

In this work, we have tried to apply fractal theory to model HR dynamics. We have shown that HR time series data can be modeled with FD. Also, FD can distinguish between normal and abnormal cardiac rhythm. But, there is more to do in this field. In this work, we have calculated only noise-corrupted IHR in an attempt to test the applicability of fractal theory to detect cardiac abnormality. Since calculation of FD needs a long data length of IHR and it is not available at the present moment, we are unable to calculate FD of pathologic IHR. IHR of different types of cardiac patients should be subjected to the calculation of FD and then we can certainly say whether fractal can be used as a diagnostic tool. Moreover since the range of the dimension is from 1 to 2, it may not be possible to accommodate numerous cardiac conditions within this narrow range. Calculation of FD of cardio-electrical images may provide better results in this direction.

APPENDIX

A1: Flow chart for box counting method

```
                    ┌─────────────┐
                   /  Read Data   /
                  └─────────────┘
                         │
                         ▼
                  ┌──────────────┐
                  │ Find maximum │
                  │     and      │
                  │ minimum value│
                  └──────────────┘
                         │
                         ▼
                  ┌──────────────┐
                  │ Normalize data│           ┌──────────────┐
                  │               │           │ Display output│
                  └──────────────┘           └──────────────┘
                         │                          ▲
                         ▼                          │ No
                  ┌──────────────┐                  ◇
                  │ Partition the │               Is any
                  │ unit square int│   Yes─────  box left?
                  │ small boxes of │              
                  └──────────────┘                  ▲
                         │                          │
                         ▼                          │
                  ┌──────────────┐                  │
                  │   Get a box  │◄─────            │
                  └──────────────┘           ┌──────────────┐
                         │                    │ Increase count│
                         ▼                    └──────────────┘
                  ◇                              ▲
            Find whether any data   Yes────────┘
               falls within the
                selected box
                         │
                         └──────── No ──────────
```
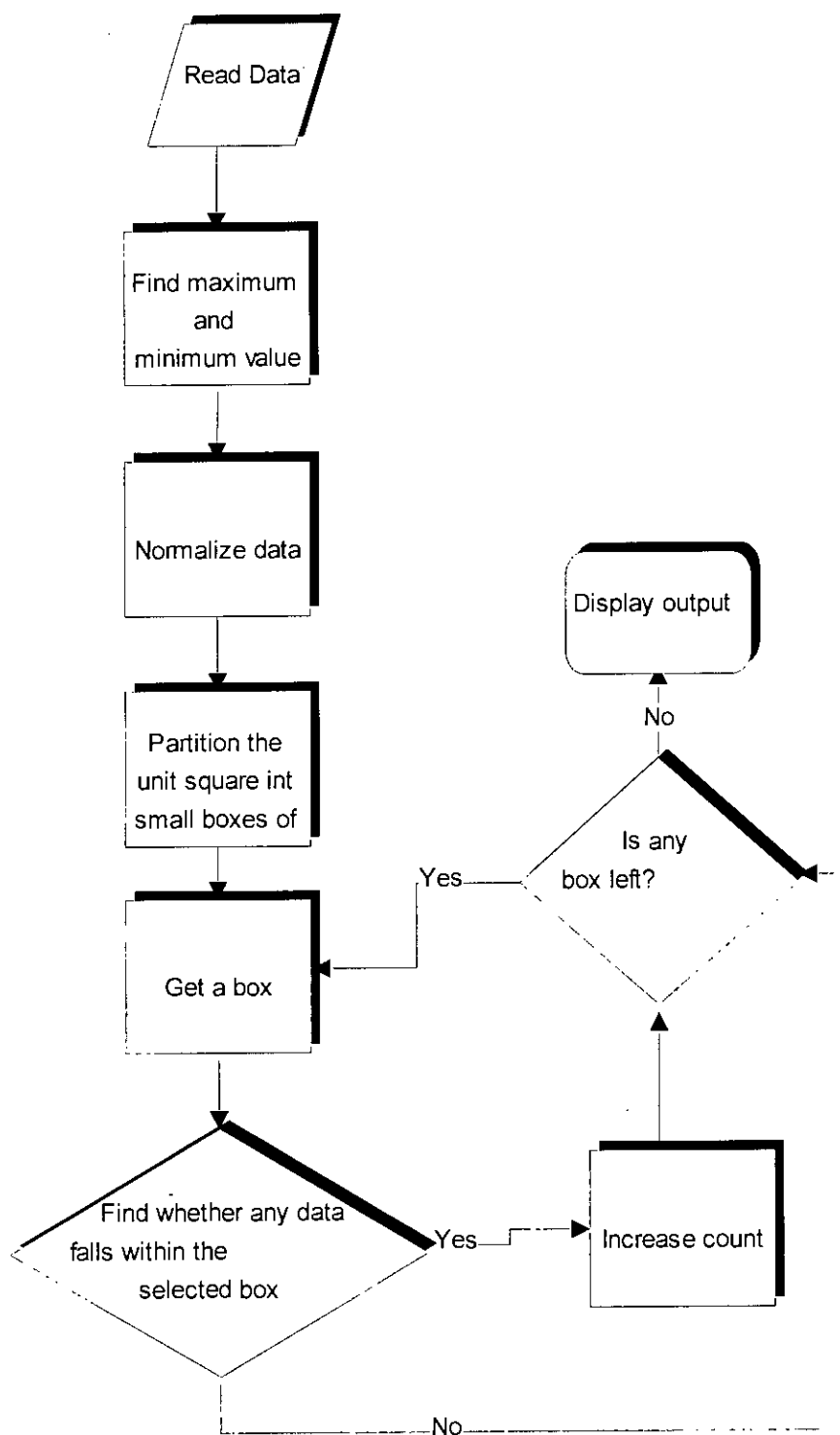
A2: Flow chart for rescaled range method
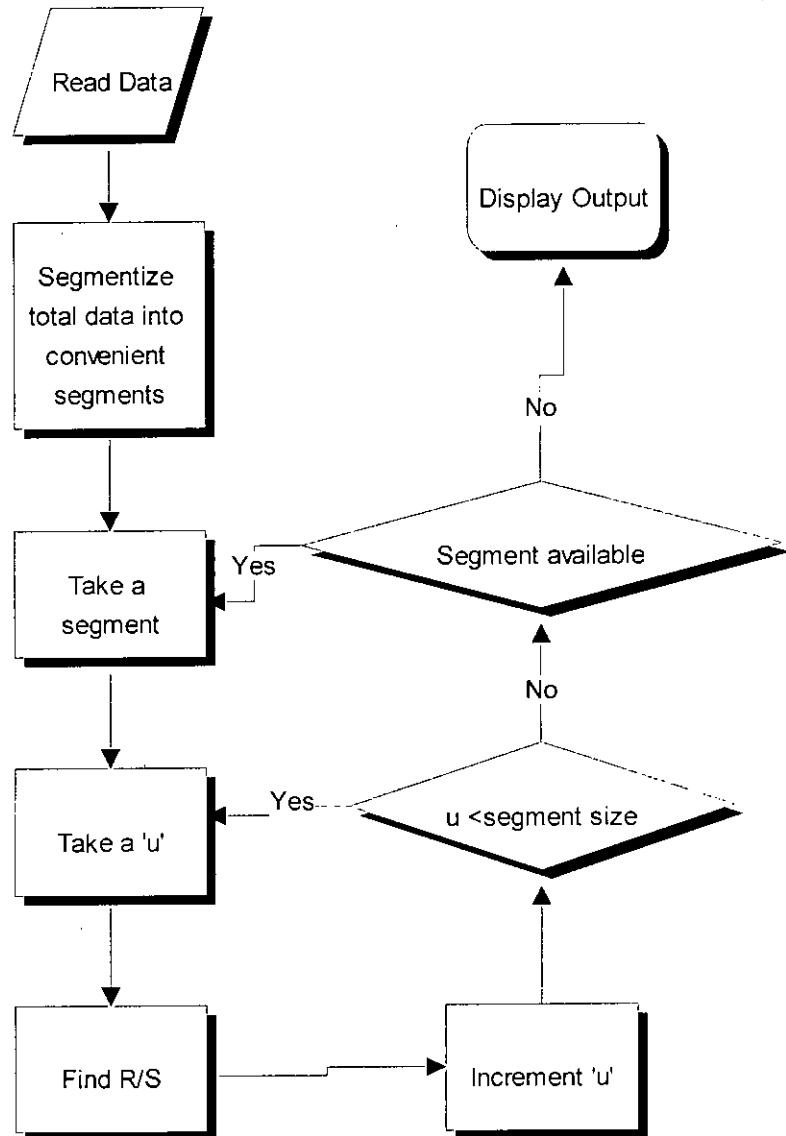
A3: Flow chart for relative dispersion method
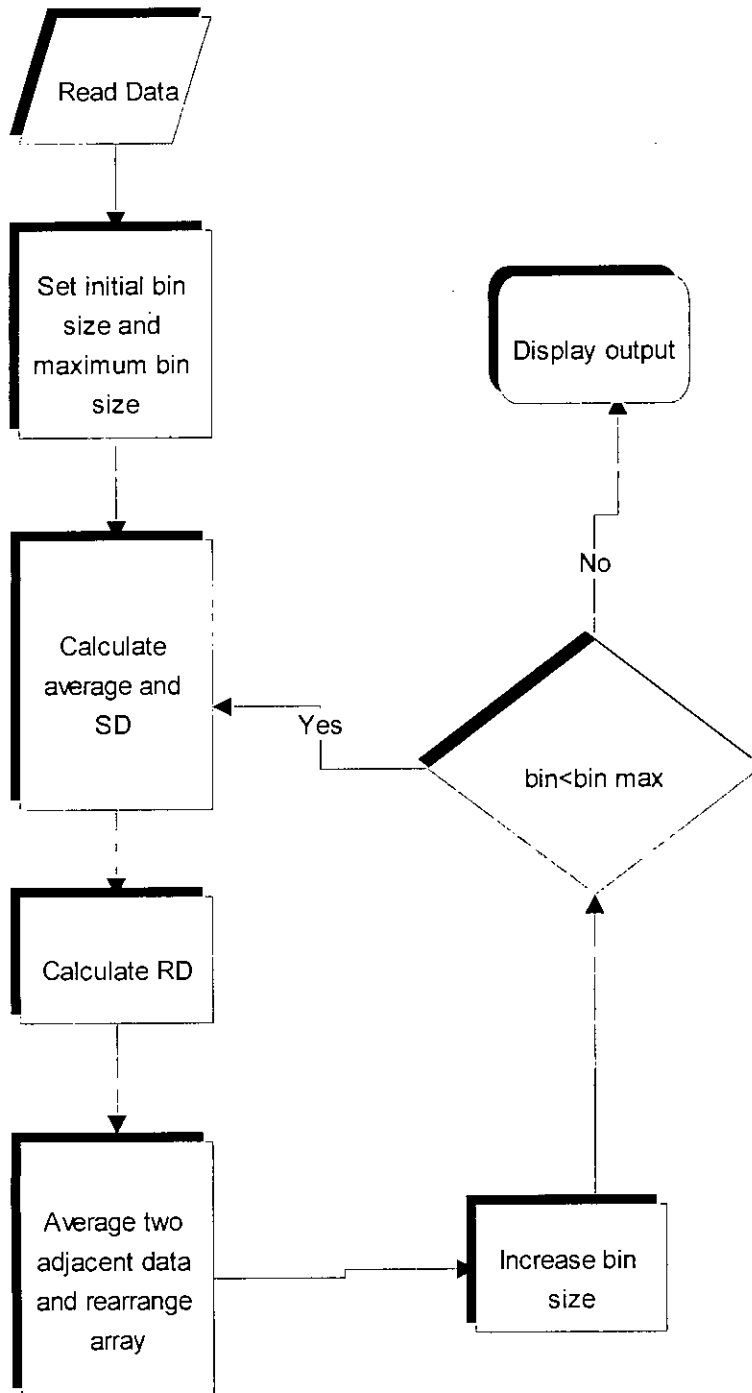
A4: Flow chart for Fourier method

```
        ┌──────────────┐
        │              │
        │  Read Data   │
        │              │
        └──────┬───────┘
               │
               ▼
        ┌──────────────┐
        │    FFT by    │
        │   MATLAB     │
        └──────┬───────┘
               │
               ▼
        ┌──────────────┐
        │    PSD by    │
        │   MATLAB     │
        └──────┬───────┘
               │
               ▼
        ┌──────────────┐
        │ Plot the graph│
        │    in XL     │
        └──────┬───────┘
               │
               ▼
        ┌──────────────┐
        │              │
        │ Find the Slope│
        │              │
        └──────┬───────┘
               │
               ▼
        ┌──────────────┐
        │              │
        │  Calculate D │
        │              │
        └──────────────┘
```
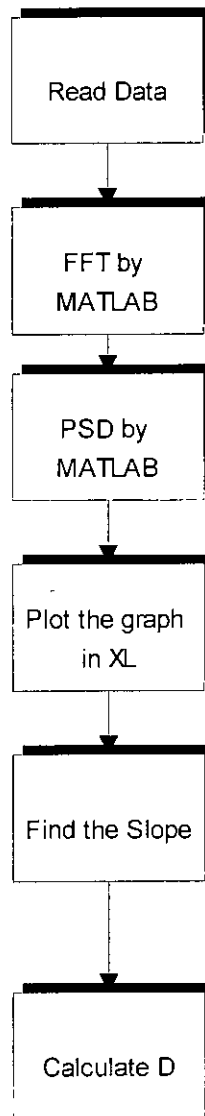
REFERENCES

[1] Leo Schamroth, *An Introduction to Electrocardiography*, Blackwell Scientific, 1990.

[2] G. E. P. Box and G. M. Jenkins, *Time Series, Forecasting and Control*, San Francisco: Holden-day, 1970.

[3] E. J. Hannan, *Multiple Time Series*, London: Wiley, 1970.

[4] T. S. Rao and M. M. Gabr, "A test of linearity of stationary time series," *Journal of Time Series Analysis*, vol. 1, no. 2, 1981.

[5] W. Deering and B. J. West, "Fractal physiology," *IEEE Engineering in Medicine and Biology Magazine*, vol. 12, no. 6, pp. 40-46, 1992.

[6] B. B. Mandelbrot, *The Fractal Geometry of Nature*, New York: W. H. Freeman, 1983.

[7] A. L. Goldberger, D. R. Rigney and B. J. West, "Chaos and fractals in human physiology," *Scientific American*, vol. 262, no. 2, pp. 42-49, 1990.

[8] X. Zhu, B. Cheng and D. M. Titterington, "Fractal model of one-dimensional discrete signal and its implementation," *IEE Proc.-Vis Image Signal Process.*, vol. 141, no. 5, pp. 318-324, 1994.

[9] T. J. Dennis and N. G. Dessipris, " Fractal modeling in image texture analysis," *IEE Pro. F*, vol. 136, no. 5, pp. 227-235, 1989.

[10] T. lo, H. Leung, J. Litva and S. Haykin, "Fractal characterization of sea-scattered signals and detection of sea-surface targets," *IEE Pro. F.*, vol. 140, no. 4, pp. 243250, 1993.

[11] C. C. Chien, J. S. Dapnte and D. Fox, "Fractal feature analysis and classification in medical imaging," *IEEE Trans. Med. Imaging*, vol. 8, no. 2, pp. 133-142, 1989.

[12] A. Accardo, M. Affinito, M. Carrozzi and F. Bouquet, "Use of the fractal dimension for the analysis of electroencephalographic time series," *Biological Cybernetics*, vol. 77, pp. 339-350, 1997.

[13] A. L. Goldberger and B. J. West, "Fractals in physiology and medicine," *Yale Journal of Biology and Medicine*, vol. 60, pp. 421-435, 1987.

[14] A. L. Goldberger, "Is the normal heartbeat chaotic or homeostatic?"" *News in Physiological Sciences*, vol 6, pp. 87-91, 1991.

[15] K. Falconer, *Fractal Geometry, Mathematical Foundations and Applications*, John Wiley & Sons: 1990.

[16] H. E. Schepers, J. H. G. M. van Beek and J. B. Bassingthwaighte, "Four methods to estimate the fractal dimension from self-affine signals," *IEEE Engineering in Medicine and Biology Magazine*, vol. 12, no. 6, pp. 57-71, 1992.

[17] C. Tricot, *Curves and Fractal Dimension*, New York: Springer-Verlag, 1995.

[18] M. A. Haque, M. K. Hasan and H. Tazawa, "Investigation of the nonlinearity in the heart rate dynamics," *Medical Engineering & Physics*, vol. 23, no. 2, pp. 111-115, 2001.

[19] R. E. Walpole, *Introduction to Statistics*, Collier Macmillan, 2nd Edition.