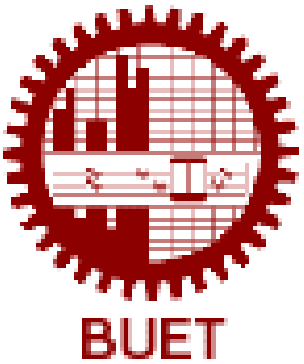
DETECTION OF CARDIAC ABNORMALITY USING FRACTAL ANALYSIS

by

Md. Meganur Rhaman

A Project Report
Submitted to the Department of
Electrical and Electronic Engineering, BUET,
in partial fulfilment 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

March 2010

The project titled **Detection of Cardiac Abnormality Using Fractal Analysis** Submitted by **Md. Meganur Rhaman,** Roll No: **040506223 F**, Session **April 2005** has been accepted as satisfactory in partial fulfilment of the requirements for the degree of **Master of Engineering in Electrical and Electronic Engineering** on **14 March 2010**.

BOARD OF EXAMINERS

1.	_____ Dr. Md. Aynal Haque Professor Dept. of EEE, BUET, Dhaka	Chairman
2.	_____ Dr. S. M. Mahbubur Rahman Assistant Professor Dept. of EEE, BUET, Dhaka	Member
3.	_____ Dr. Celia Shahnaz Assistant Professor Dept. of EEE, BUET, Dhaka	Member

DECLARATION

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

_____

Md. Meganur Rhaman

ACKNOWLEDGEMENTS

I would like to express my gratitude to my erudite supervisor Dr. Md. Aynal Haque, Professor Department of Electrical and Electronic Engineering, BUET for his kind supervision to initiate the idea of this paper. His continuous assistance inexorable enthusiasm encouragement, valuable suggestions were outstanding which made this task a reality. His innovative idea, modern outlook inspired the authors to perform this task. It is noteworthy as well as ineffable that his generosity and helping attitude was the key to make the job a wholesome one.

I wish to express my thanks and regards to the Head of the Department of Electrical and Electronic Engineering, BUET, for his support and cooperation.

Sincere thanks to my parents, wife, colleagues, friends and well wishers for their encouragement, advice and support to complete my project.

Table of Contents

ABSTRACT

Fractal is an object whose Housdorff dimension is greater than its Euclidean dimension. This report presents the application of fractal theory to the analysis of Electrocardiogram (ECG). Three methods for calculating fractal dimension (FD) namely, relative dispersion (RD), rescaled range (RS) and power spectral density (PSD) methods are applied here. At first, the three methods are applied to simulated data with known FD. The FD by all the methods depends on data length which is taken as 1024, 2048 and 4096. The RS and PSD methods are less biased to data length and give more consistent result than RD method. The best result is obtained for a data length of 4096.

The same three methods are applied to calculate FD of instantaneous heart rate (IHR) derived from ECG of both normal and abnormal records of MIT-BIH data base. The abnormalities considered here are premature ventricular contraction (PVC), bundle branch block (BBB) and atrial premature beat (APB). The number of data sets for whose ECG are taken is 3 in each case except for BBB where 4 data sets are taken. The FD of IHR of normal ECG is by RD, RS and PSD methods are found as 1.46±0.071, 1.66±0.006 and 1.80±0.025, respectively, where FD is provided as mean ± standard deviation. The respective values of FD for records containing PVC are 1.43±0.035, 1.51±0.029 and 1.71±0.010 while that for left BBB are 1.28±0.015, 1.74±0.017 and 1.70± 0.043. The FD values for APB data are 1.35±0.015, 1.56±0.015 and 1.74±0.006. It is seen that the FD of ECG containing PVC, BBB and APB type abnormalities is different from that of normal ECG. Hence, it is can be said that fractal analysis can distinguish these types of abnormalities from the normal ones. However, due to small number of data sets studied here, absolute conclusion regarding the applicability of FD analysis to detect cardiac abnormality can not be made. This study shows the potential use of fractal analysis in the estimation of cardiac condition.

LIST OF FIGURES

LIST OF TABLES

Chapter 1
INTRODUCTION

1.1 Introduction to Fractals

A fractal is an object whose Housdorff dimension is greater than its Euclidean dimension. The explanation of fractal should be observed on the same way as the biologists observe the explanation of life. There is no solid and swift definition. It gives the impression best to regard 'fractal' as a set, which has the properties of well arrangement, enough abnormalities to be described in conventional geometrical language and fractal dimension (FD) greater than its topological dimension. So we can say fractal is a mathematical investigation for characterizing complex, replicating geometrical patterns at different scale lengths. Signifying the fractal behaviour of a signal is inconsistent, since it incites the growth of new way for determining how the arrangement works. It has been established that fractal geometry can play significant role in the analysis of natural observable fact. Fractal dimension is also expansively used in understanding of the characteristic of extraordinary attractors.

Fractals:

- Fractals are of rough or fragmented geometric form that can be subdivided in parts, each of which is (at least approximately) a reduced copy of the whole.
- They are puckered objects that defy conventional measures, such as length and are most often characterized by their fractal dimension.
- They are mathematical sets with a high degree of geometrical complexity that can model many natural phenomena. Almost all natural objects can be observed as fractals (coastlines, trees, mountains, and clouds).
- Their fractal dimension strictly exceeds topological dimension.
- It is a quantity, very often non-integer, often it is the only one measure of fractals.
- It measures the degree of fractal frontier disintegration or abnormality over multiple scales.
- It determines how fractal differs from Euclidean objects (point, line, plane, circle etc.)

Mono-Fractals / Multi-Fractals:

- Just a small group of fractals have one certain fractal dimension, which is scale invariant. These fractals are mono-Fractals.
- The most of natural fractals have different fractal dimensions depending on the scale. They are composed of many fractals with the different fractal dimension. They are called “multi-Fractals“.
- To characterize set of multi-Fractals (e.g. set of the different coastlines) we do not have to establish all their fractal dimensions, it is enough to estimate their fractal dimension at the same scale.

Self-similarity/ Semi-self similarity:

- Fractal is strictly self-similar if it can be expressed as a union of sets, each of which is an exactly reduced copy (is geometrically similar to) of the full set (Sierpinski triangle, Koch flake). The most fractal looking in nature does not display this precise form.
- Natural objects are not union of exact reduced copies of whole. A magnified view of one part will not precisely reproduce the whole object, but it will have the same qualitative appearance. This property is called statistical self-similarity or semi-self-similarity.

Fractal analysis:

- A compilation of mathematical procedures used to establish fractal dimension (or any other fractal characteristic) or set of fractal dimensions (in the case of multi-fractals) with the smallest error.
- Now a day very often used to illustrate properties if natural objects. Methods are under continuous scientific improvement.

1.2 Fractal Dimension

The term "fractals" is derived from the Latin word ‘**Fractus**’, the adjectival form of **‘Franger’**, or "to break" [1]. Unlike conventional geometry, which deals with lines, triangles, circles, spheres and cones, fractal geometry is concerned with broken or "fractured" shapes as so commonly found in nature. Such shapes simply do not vary in degree from conventional geometry (for example, clouds that are not spheres, trees that are

not cones and rivers that do not run straight). Fractals have been used to explain objects and geometrical formations [2]. Many structures demonstrate an underlying geometric regularity, known as scale invariance or self-similarity that is the tendency of natural forms to repeat them. If these objects are observed at different size scales, there is the same fundamental pattern that is encountered. This repetitive pattern defines the fractional, or fractal dimension of the object structure.

Fractal patterns exhibits, in many cases, chaotic behaviour. Chaos may be defined as the pattern that lies between the determinism and randomness of a system and it refers to a constrained type of unpredicted turbulent dynamics. It is emphasized that all fractal dimensional systems are chaotic and that the data they generate will be a periodic, complex, and apparently unpredictable. The analysis of the dynamics of human biomedical or biological signals is an important area of investigation to help control and to be able to predict the onset of pathological conditions. Fractal behaviour is exhibited by the heart in electrocardiogram signals and by the brain in electroencephalogram (EEG) signals [3], [4]. It has been emphasized that the demonstration of chaotic behaviour in humans opens out the possibility of rapid diagnosis and effective therapeutic control of conditions ranging from epilepsy to cardiac arrest [5]-[7]. What is required is a practical method or procedure that utilizes or translates chaos and/or fractal theory concept into a simple and straightforward manner to help distinguish between normal and pathological behaviour [8]

Fractals are characterized by a scaling law that relates two variables: the scale factor and the object being measured. This scaling relationship is described by a power law, which in turn describes the inherent physical attributes of the object being analyzed. The exponent of the power law refers to the dimension [9]. From fractal linguistics perspective, fractal images are strictly limited visualizations of recursive equations that cannot be represented to the limit as an image on a computer. The Koch curve shown in Figure 1 is not a fractal, as the iteration level is not depicted at the level of infinity and has been referred to as pre-fractal [2]

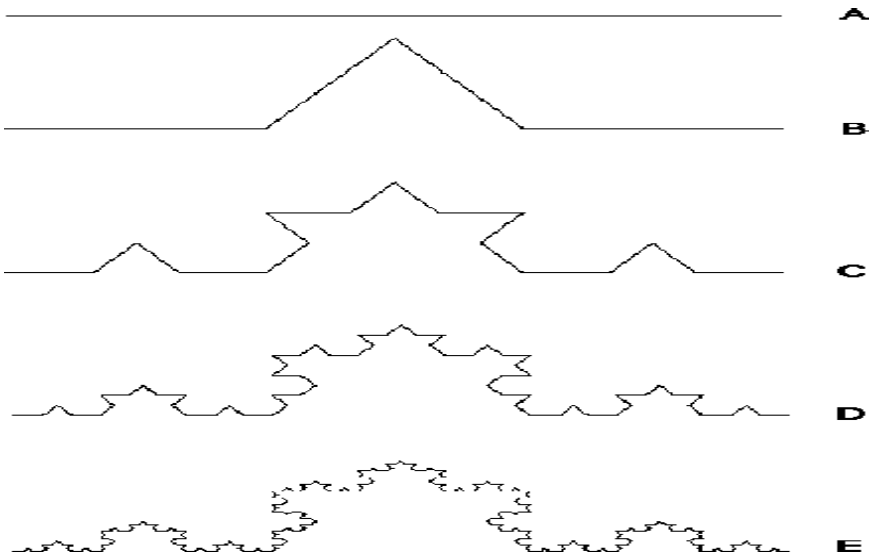


Fig 1.1: The Koch curve displaying the iteration process over several generations

A fractal object is the result of repeated transformations of a geometrical figure that leads to self-similar patterns [1]. This self-similarity is a function of the scale invariance observed within the pattern between successive transformations. Viewing a fractal would therefore reveal identical patterns at different observation scales. In fact an important definition of fractals is that this self-similarity continues indefinitely. To explain the concept of fractal dimension, it is necessary to explain the dimension in the first place. A line is one dimensional and a plane is two dimensional. We can break a line into N self-similar pieces, each with magnification factor N. A square can be decomposed into N^2 self-similar copies of itself, each of which must be magnified by a factor of N to yield the original figure as shown in figure 1.2.

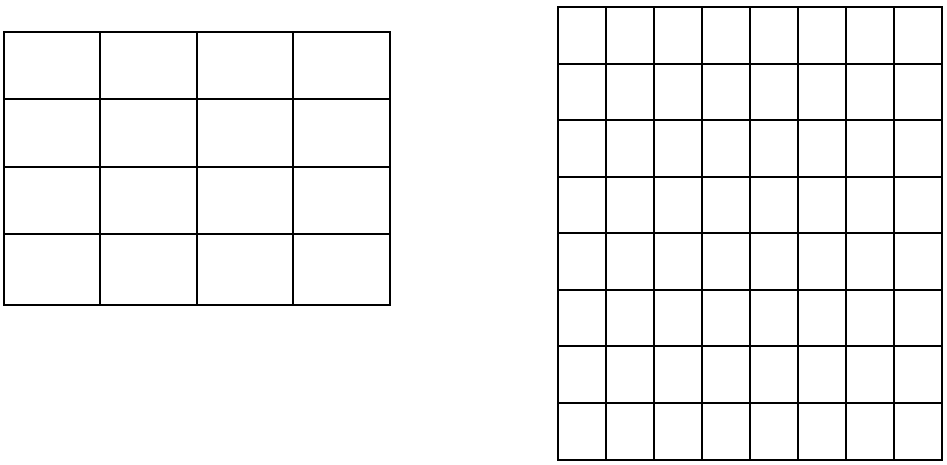


Fig 1.2: Breaking of a square into N^2 self similar pieces, each with magnification factor N

Now we see an alternative way to specify the dimension of self similar objects. The dimension is simply the exponent of the number of self similar pieces with magnification factor N into which the figure can be broken.

$$Dimension = \frac{\log(Number\ \ of\ \ \ \ self-similar\ \ pieces)}{\log(magnificaiton\ \ \ \ factor)}$$

$$= \frac{\log N^2}{\log N}$$

$$= 2$$

So, the dimension of the square is 2. This we take as the definition of fractal dimension of a self similar object:

$$Fractal\ Dimension = \frac{\log(Number\ \ of\ \ \ \ self-similar\ \ pieces)}{\log(magnificaiton\ \ \ \ factor)} \qquad (1.1)$$

Now we can compute the dimension of S for the Sierpinski triangle consisting of three self similar pieces, each with magnification factor 2. So the fractal dimension is

$$Fractal\ Dimension = \frac{\log(Number\ \ of\ \ \ \ self-similar\ \ pieces)}{\log(magnificaiton\ \ \ \ factor)}$$

$$= \frac{\log 3}{\log 2}$$

$$= 1.58$$

Fig1.3: Construction of Sierpinski gasket

So, the dimension of S is somewhere between 1 and 2. Unfortunately, similarity dimension is meaningful only for a small class of self-similar sets. Very roughly, fractal dimension is a measure of how "complicated" a self similar figure is. It provides a description of how much space a set fills. It is a measure of the prominence of the irregularities of a set when viewed at very small series. Fractal system, both in time and space, has neighbour to neighbour correlation at all levels of scale. In spatial statistics, the variance of the local measure depends on the resolution of measurement. For instance, the

variance of regional blood flow in the heart increases as the resolution of measurement is increased. In an experiment where the heart is divided into many sections for the measurement of the concentration of a flow proportional marker in each of the sections, the relation between the resolution mass of observed tissue element, m, and the variance, Var, of the regional blood flows, is well described by a power law function:

$$\text{Var}(m) = \text{Var}(m_0)\left(\frac{m}{m_0}\right)^{2H-2} \tag{1.2}$$

Where H is the Hurst coefficient, H=2-D and D is the fractal dimension lying between 1.0 and 2.0. The spatial distribution is considered to have fractal properties. The consequence of D is that it describes very compactly the relation between the variance of the signal and the time scale. No assumptions on the underlying system for which the signal is observed are made and therefore it is a purely statistical tool. The fractal dimension estimated these techniques form a time series in between 1 to 2 i.e. for one dimensional function y(t), the fractal D lies between the Euclidean dimension E(=1) and E+1.

1.3 Historical Background

There is long history of using image analysis to determine the morphology of a system associated with an image. Recently, techniques based on concepts from the field of artificial life have been used for image analysis [3]. The measures of complexity they use involve fractal dimension and percolation. The concept of fractal dimension to describe structures which look the same at all length scale, was first proposed by Mandelbrot [1]. Although in strict terms, this is a purely mathematical concept, there are many examples in nature that closely approximate a fractal object, though only over particular range of scale. From the historical analysis fractals initiate to take form in the seventeenth century when mathematician and philosopher Leibniz measured recursive self-similarity (although he made the oversight of thinking that only the straight line was self-similar in this sense). It captured until 1872 before a function appeared whose graph would today be considered fractal, when Karl Weierstrass [10] gave an example of a function with the non-intuitive possessions of being everywhere continuous but nowhere differentiable. In 1904, Helge von Koch, dissatisfied with Weierstrass's very conceptual and analytic definition, gave a more geometric definition of a similar function, which is now called the Koch curve [2]. In 1915, Waclaw Sierpinski constructed his triangle and, one year later, his carpet. Originally these geometric fractals were described as curves rather than the 2D shapes that they are

known as in their modern constructions. In 1918, Bertrand Russell recognized a "supreme beauty" within the emerging mathematics of fractals. The idea of self-similar curves was taken further by Paul Pierre Lévy, who, in his 1938 paper Plane or Space Curves and Surfaces Consisting of Parts Similar to the Whole described a new fractal curve, the Lévy C curve [10]. Georg Cantor also gave examples of subsets of the real line with unusual properties—these Cantor sets are also now recognized as fractals. Iterated functions in the complex plane were investigated in the late 19th and early 20th centuries by Henri Poincaré, Felix Klein, Pierre Fatou and Gaston Julia [10]. However, without the aid of modern computer graphics, they lacked the means to visualize the beauty of many of the objects that they had discovered. In the 1960s, Benoît Mandelbrot started investigating self-similarity in papers. Finally, in 1975 Mandelbrot created the word "fractal" to denote an object whose Hausdorff-Besicovitch dimension is greater than its topological dimension. He illustrated this mathematical definition with striking computer-constructed visualizations. These images captured the popular imagination; many of them were based on recursion, leading to the popular meaning of the term "fractal". The fractal analysis can also be applied further in image analysis, fluid dynamics, investing natural phenomenon and so on. By determining the Fractal Dimension of an ECG signal, an estimation of heart condition can be made, which can be helpful in medical science [11].

1.4 Objective of the Project

The aim in this study to apply frequently a time varying signal is required to be examined to pull out characteristics of awareness. For example, medical diagnosis often requires the analysis of time varying cardiac, respiratory, or brain signals in order to detect cardiac, pulmonary, or mental problems. In industrial control, electrical sensors produce signals in response to sensed parameters which occur over time in a manufacturing process and a control system responds to characteristics detected in the resultant signals. In general, most temporal processes are analyzed using Fourier Transform technique (frequency domain); chaos dynamics (position-velocity phase plane) and other complex mathematical techniques have been applied to signal analysis. A common drawback of these methods is that they are often complex, not easily amenable to analysis, and require some data pre-processing procedures, such as filtering, etc. Thus there remains a need for simple and practical methods for analyzing such time varying electrical signals. Images and shapes within images which cannot be represented by Euclidean geometry have been analyzed by

Fractal geometry. Unlike conventional geometry, which deals with lines, triangles, circles, spheres and cones, fractal geometry is concerned with broken or "fractured" shapes as so commonly found in nature.

1.5 Organization of the Report

The rest of the report is organized as the following:

- Chapter 2 contains the method of calculation of fractal dimension (FD) of time series data. The methods considered are relative dispersion (RD), power spectral density (PSD) and rescaled range (RS) methods.
- Chapter 3 contains the calculation of FD of simulated data. Data of known FD is simulated and FD is calculated to compare the applicability of different methods.
- Chapter 4 describes the ECG and its different patterns for normal and abnormal cardiac conditions.
- Chapter 5 contains the calculation of FD of heart rate (HR). FD based analysis of HR time series data of healthy persons as well as pathologic patients are performed using three different methods. The abnormalities considered are premature ventricular contraction (PVC), bundle branch block (BBB), atrial premature contraction (APC) and so on.
- Chapter 6 discusses the findings of the project and future perspectives.

Chapter 2

CALCULATION OF FRACTAL DIMENSION

2.1 Introduction

Fractal analysis is a mathematical investigation for characterizing complex, replicating geometrical patterns at different scale lengths. We shall inspect three methods, relative dispersion (RD), power spectral density (PSD) and rescaled range (RS) methods, [2], [12], [13], for analyzing one-dimensional time series data. All of these methods give the value of fractal dimension D. When adjacent elements in a time series or a spatial distribution are positively correlated, the measured variance will drop less rapidly as the resolution is decreased. Based on this property, the fractal dimension (FD) can also be estimated directly from a RD analysis. The statistical method, Hurst's rescaled range (RS) analysis from which the fractal dimension can be calculated depends on the scaling properties of self-affine signal [2]. For analysis of flows in rivers the range is of the integral of deviation from the mean of the interval and it is normalized by dividing by the standard deviation of the differences from the mean. This method can also be used to estimate the fractal dimension from a fractional Brownian noise. The modification of RS was performed by using the differences from the mean. The signals analyzed by a specific power spectrum of the form [9]:

$$S(f) \approx f^{B}$$

Where, f is the frequency. Thus Fourier analysis is another procedure used to estimate fractal dimension.

2.2 Relative Dispersion (RD) Method

Originally, the RD method was introduced in 1988 for evaluating the fractal characteristics of regional flow distributions in the heart distribution analysis [2]. This technique, which we will discuss in detail below, is a one-dimensional approach that can be applied to an isotropic signal of any dimension. Making estimates of the variance of the signal at each of several different levels of resolution form the basis of the technique; for fractal signals a plot of the log of the standard deviation versus the log of the measuring element size (the measure of resolution) gives a straight line with a slope of 1 - D, where D is the fractal dimension. For a specified value of H, the Hurst coefficient, we can determine the FD as D

= 2 – H with 1 < D < 2. H is a measure of irregularity; the irregularity or anticorrelation in the signal is maximal at H near zero. White noise with zero correlation has H = 0.5. Smoother correlated signals have H near to 1.0. It is reported that there is a strong relationship between the measure of variation (the coefficient of variation) and the resolution of measurement. This is a simple one dimensional spatial analysis which we level as RD analysis, where RD is derived by dividing the standard deviation by the mean. Mathematically it can be expressed as:

$$RD = \frac{SD}{Mean}$$

$$Standard Deviation = \frac{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2}{N} \tag{2.1}$$

where, x_i = Random variable

$\bar{x}$ = Mean of the variables

N = Number of Samples

The dependence of the variance of a variable also found by the same fractal analysis of time series, for example, the velocity of blood cells passing through a small artery as a function of time. If a signal y(t) is divided into intervals of length Δt and the mean is calculated for each interval, one can calculate the variance of these means; there will be a lower variance when the interval length is increased. The relationship between the variance and the sample time may fit a power law, and if so, the fractal dimension of the time series can be calculated by same RD analysis as was used for a spatial measure. In time series analysis, one stars by measuring at the highest of resolution, the variance of the signal over its whole time course. The next step is to average the signal over pairs of consecutive values, find the variance and repeat for longer strings of values of y(t). String length increasing geometrically, grouping by 2, 4, 8 etc, is efficient because pairs of average are combined for a next iteration. When the signal is uncorrelated (white noise), with H = 0.5, one expects that the standard deviation, when two consecutive values are averaged, is reduced by a factor $\frac{1}{2^{0.5}}$, or when n consecutive values are averaged, by $\frac{1}{n^{0.5}}$. The mean remains the same. Now one will find that with a correlation function fBn, and H not equal to 0.5, the SD will be proportional to n^{H-1}, n being the bin size or time resolution interval. By calculating the RD (RD = SD / Mean) for different bin sizes, n and fitting the square law function:

$$RD = RD_0\left(\frac{n^{H-1}}{n_0}\right) \quad (2.2)$$

where, RD_0 is the RD for some reference bin size n_0.

The whole data set is used for each calculation of RD (n) at each level of resolution or number of pieces, n. The exponent can be best estimated by a log-log transformation.

$$\log (RD)=\log(RD_0) +(H-1)\log(n/n_0) \quad (2.3)$$

Here, H-1 = slope (2.4)

For fractal dimension, D=2-H (2.5)

D=2-(1+slope)

So, D =1-slope (2.6)

So, we can easily estimate the fractal dimension from this equation.

The fit is little improved when longer signals are used. The RD analysis yields essentially the same result only the first three or four points on the graph (i.e. the smallest bin sizes 2, 4, 8 and 16) are used. With very large bin sizes, the number of bins is of course small and the measure of the variance is less accurate. This relationship is an important advantage of RD analysis, because especially in spatial statistics, signals may not be very long due to limited resolution and domain size.

In the years since its introduction in 1988, RD analysis has been fairly extensively applied, for example to regional flow distributions in the heart, the lung and more recently, the kidney. While it is intuitively obvious that greater heterogeneity will be observed when the spatial resolution of the observations is more refined, when the spatial distributions are fractal, the heterogeneity observed at each of many different levels of resolution is describable in terms of two numbers, a fractal dimension, D, and the variance at a chosen level of resolution. Such two-number descriptors allow the results from various laboratories around the world to be compared even when the raw observations have not been made on pieces of tissue of the same size at all these laboratories. Dispersion analysis is also applicable to signals that are functions of time (i.e., true one-dimensional signals such as might be obtained for blood pressure, heart rate, or local concentration of a solute).

2.3 Rescaled Range (RS) Method

The basis of the rescaled range analysis was laid by Hurst [2]. Mandelbrot and Wallis examined and further elaborated the method. Feder [10] gives an excellent overview of the history, theory and applications, and adds some more statistical experiments. Rescaled range analysis is a statistical method to analyze long records of natural phenomena. There are two factors used in this analysis: firstly the range R, this is the difference between the minimum and maximum 'accumulated' values or cumulative sum of X(t,τ) of the natural phenomenon at discrete integer-valued time t over a time span τ, and secondly the standard deviation S, estimated from the observed values X(t,τ). Hurst found that the ratio R/S is very well described for a large number of natural phenomena by the following empirical relation [2]:

$$\frac{R(\tau)}{S(\tau)} \propto \tau^{H}$$

Where τ is the time span, and H the Hurst exponent. The coefficient C was taken equal to 0.5 by Hurst. R and S are defined as

$$R(\tau) = \max_{1 \le t \le \tau} X(t,\tau) - \min_{1 \le t \le \tau} X(t,\tau)$$

and, $$S(\tau) = \left(\frac{1}{\tau} \sum_{t=1}^{\tau} \left[\xi(t) - \langle \xi \rangle_{\tau} \right]^2 \right)^{\frac{1}{2}}$$

where $$\langle \xi \rangle_{\tau} = \frac{1}{\tau} \sum_{t=1}^{\tau} \xi(t)$$

and, $$X(t,\tau) = \sum_{u=1}^{t} \left[\xi(u) - \langle \xi \rangle_{\tau} \right]$$

The relation between the Hurst exponent and the fractal dimension is simply

D=2-H.

We calculate the individual calculations for each interval length. A straight line is fitted in the log-log plot:

Log[R($\mathcal{T}$)/S($\mathcal{T}$)] = C + H log($\mathcal{T}$)

Where H = slope.

So, Fractal dimension, D= 2-H

With the help of this equation we can easily evaluate fractal dimension in rescaled range analysis, this method handles observations in time. The graphical representation uses time in the abscissa, and the observed value in the ordinate. The estimation from the R/S

analysis depends on the way of way of subsets from the signal is taken to calculate the rescaled range. With longer signal, larger lags can be used along with more points per lag. Also, the length of the shortest lag can be increased, which proves the fit further.

2.4 Power Spectral Density (PSD) Method

The power spectrum (the square of the amplitude from the Fourier transform) of a unpolluted fractional Brownian motion is known to be described by a power law function [2]:

$$|A|^2 = 1/ f^{\beta}$$

Where |A| is the magnitude of the spectral density at frequency f, with an exponent equal to $\beta = 2H + 1$.

In general, fractal signals always have such a very broad spectrum. When the derivative is taken from a fractal signal, β is reduced by two. Thus, for fractional Brownian noise, fBn, β is expected to be: $\beta=2H+1$. Here again, a straight line is fitted from a log-log plot, and H is calculated from the slope β. Power spectrum method applies the power law variation of time series. A strong relationship exists between fractal dimension and power law index of time series. In the frequency domain, fractal time series exhibit power law properties:

$$P(f) \sim f^{-\alpha}$$

Where P(f) is the power spectral density f, and the exponent α is the so called power-spectral index. For the values region between D or FD=1 and FD = 2 the following relationship between FD and α is valid [2], [12]:

$$FD = (5- \alpha)/2, \text{ for } 1<FD<2$$

In other words, the fractal dimension of a time series can be calculated directly from its power spectrum. The influence for longer signals in the Fourier analysis is not clear from the individual plots, but the estimates are quite accurate even for shorter signals.

Chapter 3

FRACTAL DIMENSION OF SIMULATED DATA

3.1 Introduction

In calculating the fractal dimension (FD), we face some typical problems. The first problem is how to distinguish between the effects from different signal generating algorithm and the estimation methods. Another problem is, the production of a truly fractal signal is itself a non-trivial exercise and may be as difficult as producing random noise. So, we must concern about capability of generating algorithm as well as the analysis algorithms. Some judgment may require a very long signal, whereas other provides good result at shorter signal length. It has already been established that FD depends on data length, the method of calculation and obviously, on the variation characteristics of the signal [12]. To find out fractal dimension and its dependency on different parameters, we have simulated a time series data with known fractal dimension (D=1.1 to 1.9). This chapter describes the estimation of FD of the simulated data with three methods described in Chapter 2. Based on the analysis of simulated data in the specified three methods, our aim is to find out the best method that gives fractal dimension closer to that dimension with which the data is simulated. All the possible parameters such as size, time interval between data are varied to reach at the desired solution of the FD. Then we will use these methods to estimate the fractal dimension of instantaneous heart rate (IHR) data of healthy persons and define a range of FD of healthy subjects for each method. Then we will use the methods to determine fractal dimension of IHR of subjects with different abnormalities and determine the range of variation of FD for each abnormality type.

3.2 Data Simulation

There is no unique method to generate a signal of known fractal dimension. Since our ultimate objective is to calculate FD of heart rate and we know that it lies between 1 and 2, we use the familiar Weierstrass function [10] to generate a time series signal. The function is reproduced below:

$$f(t) = \sum_{k=1}^{\infty} \lambda^{(H-2)k} \sin(\lambda^k t) \qquad (3.1)$$

where 1 <H < 2 and λ represents FD which is greater than 1. The data was simulated for signal H = 1.5 and λ = 1.5. The detailed representation of simulated data is presented in [14].

3.3 FD with Relative Dispersion Analysis

The simulated time series data is loaded from the 'time_value.mat' file in the first place. The data series is then segmented into intervals and individual mean of each segment is calculated and saved in 'mean.mat' file. The means, equal to the number of segmented intervals, are loaded from the 'mean.mat' file and the final mean is calculated and saved in 'mean_final.mat' file. The final mean of the data series is then used to determine the standard deviation of each individual data point from the final mean. Finally, the standard deviation is normalized by dividing with the total number of data point in the data series, which yields the relative dispersion RD of the data series. The results are then plotted against $\log_2(n)$ versus $\log_2(RD)$, where n is the bin size or the time resolution interval. The MATLAB function 'polyfit()' is used to fit a least square straight line over the $\log_2(n)$ versus $\log_2(RD)$ curve. The slope of the straight line is then used to calculate the fractal dimension. The important equations are listed below:

$$RD = \frac{SD}{Mean}$$

$$Standard\,Deviation = \frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{N}$$

$\text{Log}(RD)=\log(RD_0) +(H-1)\log(n/n_0)$

Fractal dimension, D=2-H

Fractal dimension, D=1-slope

The typical curve of $\log_2(n)$ versus $\log_2(RD)$ plot along with the least square approximated line for a typical data length (N =4096) is provided in Fig. 3.1. The different parameters with different time intervals are shown in the Tables 3.1.

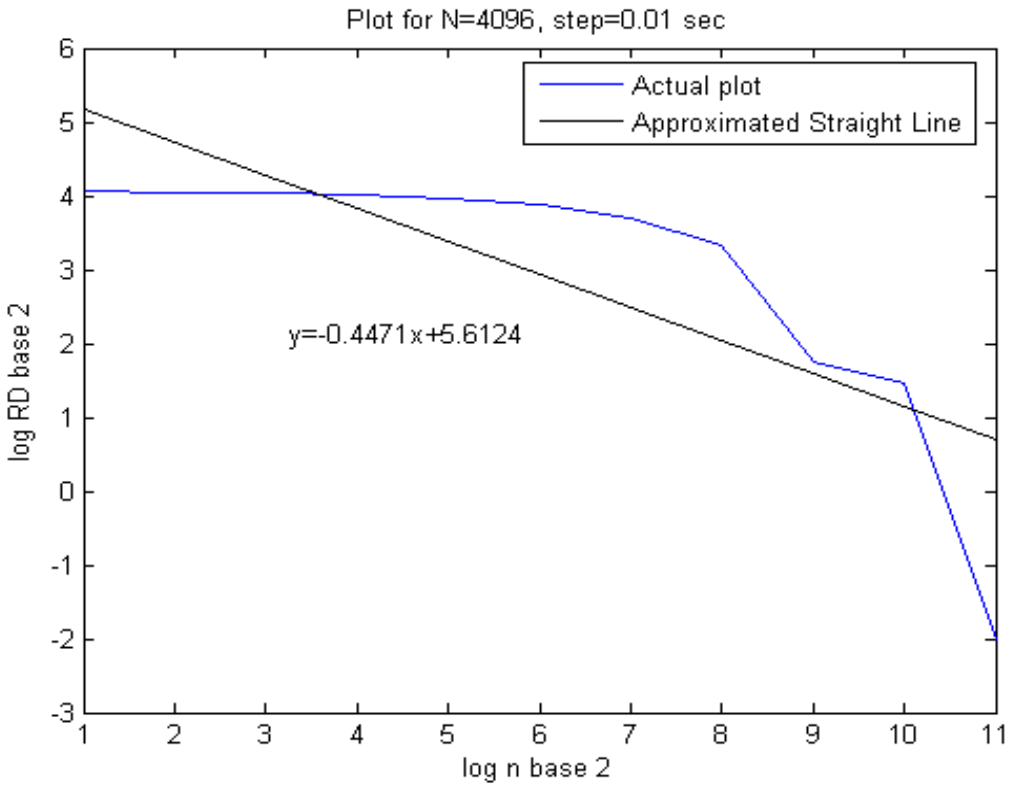


Fig 3.1: Actual and approximated straight line for relative dispersion method using 4096 simulated data.

Table3.1: Relative dispersion method using 4096 simulated data

Bin size n	**Mean**	**SD**	**RD**	**$\log_2 n$**	**$\text{Log}_2(\text{RD})$**	**Hurst coefficient**	**Fractal dimension**
2048	0.0744	0.0183	0.2466	11	-2.019	0.553	1.447
1024	0.0744	0.2047	2.7532	10	1.461		
512	0.0744	0.2510	3.3732	9	1.755		
256	0.0744	0.7474	10.0495	8	3.3291		
128	0.0744	0.9704	13.0498	7	3.706		
64	0.0744	1.0984	14.7707	6	3.885		
32	0.0744	1.1577	15.5683	5	3.961		
16	0.0744	1.1992	16.1257	4	4.011		
8	0.0744	1.2199	16.4046	3	4.036		
4	0.0744	1.2305	16.5468	2	4.048		
2	0.0744	1.2362	16.6240	1	4.055		

The results obtained with other 2 values of N (1024 and 2048) are also calculated. The results are provided in Table 3.2.

Table 3.2: FD of simulated data by RD method

Data size, N	Fractal Dimension, D
4096	1.4471
2048	1.2182
1024	1.2892

The best result is achieved with N=4096. But use of longer and closely spaced data do not provide any clear indication about the performance of RD method as we tried to reach the known fractal dimension by analyzing various data series. In general, with the increase of data length, calculated fractal dimension in general, gets closer to the actual result D=1.5.

3.4 FD with Rescaled Range Analysis

After loading the simulated data from 'time_value.mat' the data is read from the file in different lag size, i.e. in the different bin size or resolution time interval. Mean of the data series is then calculated and from the mean the difference of individual data point of all lags are determined. The integral of the difference for each data point over the lag time interval is then saved in 'data_slot.mat' file. Maximum and minimum of integral for each lag interval is then calculated to find out the range of each interval and saved in 'slot_range.mat' file. Standard deviation of the data points for each lag interval is determined to find the ratio of rescaled range to standard deviation for each lag interval and saved in 'slot_avg_r_sd.mat' file. Finally, the R/S ratio is calculated from the average of each slot R/S ratio. The results are then plotted against $\log_2(T)$ versus $\log_2(R/S)$, where T is the length of the time spans as stated by Mandelbrot and Wallis. As previous, The Matlab function 'polyfit ()' is used to fit a least-square straight line. The slope of the straight line is then used to calculate the fractal dimension. The typical results of plotting and FD calculation for a data length of 1024 are provided in Fig. 3.2 and Table 3.2, respectively.

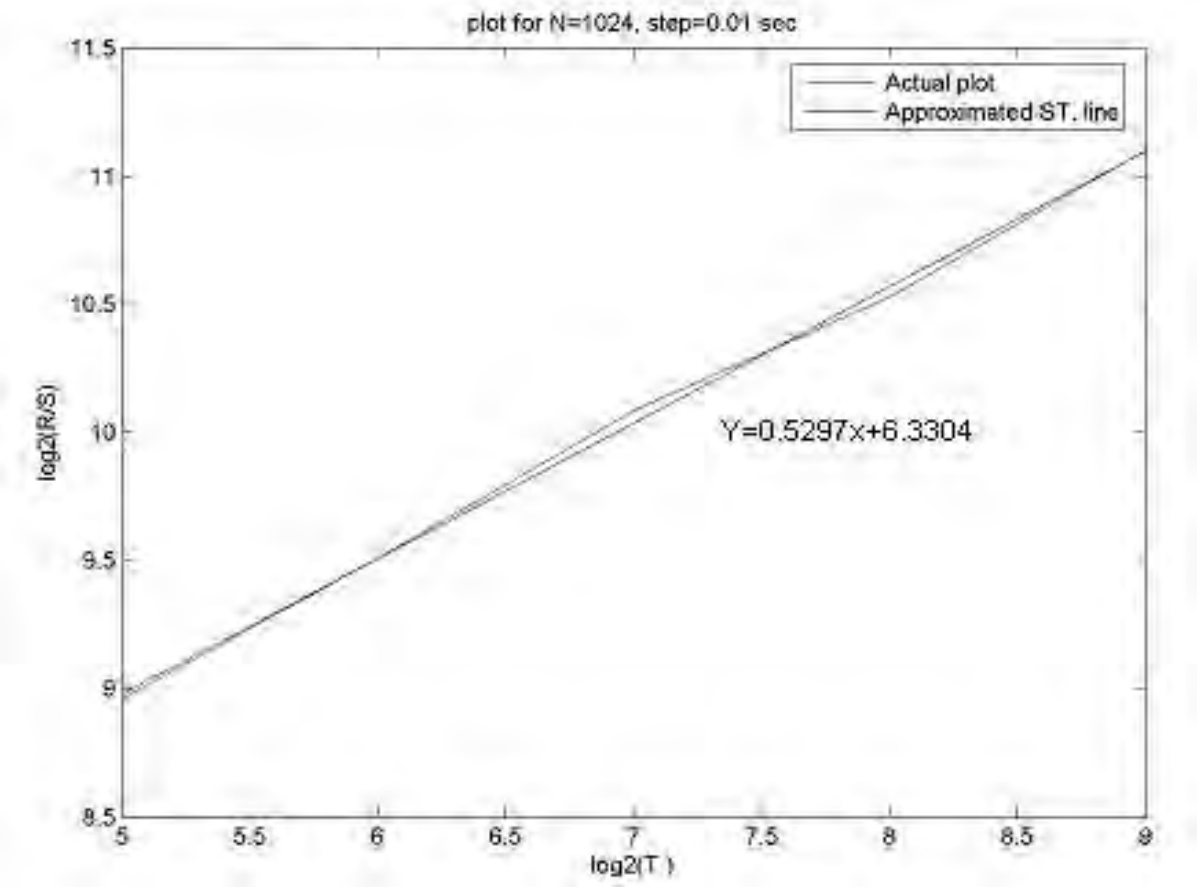


Fig: 3.2: Rescaled range analysis using 1024 simulated data.

Table 3.3: Rescaled Range analysis using 1024 simulated data.

Slot size, T	**R/S**	**$\log_2$(T)**	**$\log_2$(R/S)**
512	2198.600	9	11.1024
256	1478.300	8	10.5297
128	1087.100	7	10.0862
64	727.8666	6	9.50750
32	499.7584	5	8.96510

The results of FD with different data length are provided in Table 3.4.

Table 3.4: FD by Rescaled Range analysis

Data size, N	**Fractal Dimension , D**
1024	1.4703
2048	1.4617
4096	1.4744

From the graph, it is seen that in rescaled range analysis there is no linear variation of fractal dimension with data length or closely spaced data series. The best result is obtained with data size, N=4096. But one may find a quick hint from the fact that, in general, data series is suited better by RS analysis rather than RD analysis.

3.5 FD with Power Spectral Density Analysis

Theoretically, power spectral analysis yields the least biased results .The power spectrum is the Fourier transform of the autocorrelation function. The autocorrelation function is closely connected with the power spectrum of f, defined by

$$S(\omega) = \lim_{T\to\infty} \frac{1}{2T}\left|\int_{-T}^{T} f(t)e^{it\omega}dt\right|$$

The power spectrum reflects the strength of frequency ω in harmonic decomposition of f.

At first simulated data is loaded from the 'time_value.mat' file. Power spectral density is estimated by using MATLAB function 'PSD'. Function Format:

[Pxx,F] = psd(x,nfft,fs,window,noverlap)

It returns a vector of frequencies, the same size as Pxx at which the PSD is estimated. The PSD is plotted on log scale by using 'log log' command. Least-square straight line is then fitted over the PSD plotted. The slope of the estimated straight line is used in calculating the fractal dimension of the simulated data. The plots of the power spectral densities against the normalized frequency in the log scale are shown in Figure 3.3.

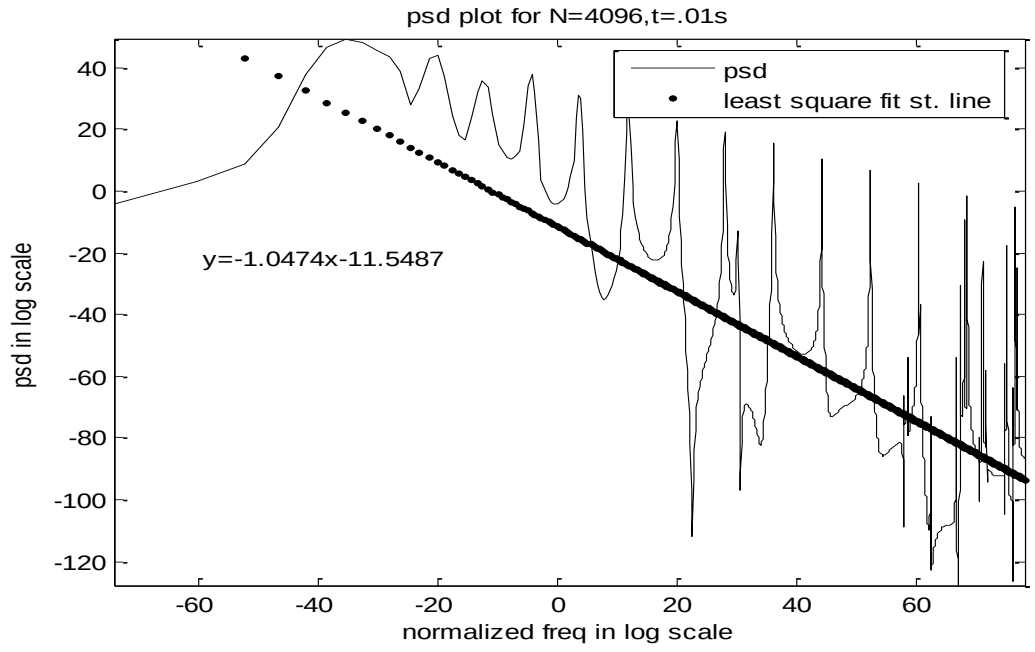


Fig 3.3: PSD and approximated straight line using 4096 simulated data.

Measuring the slope of the fitted straight line, we have calculated the FD for different data length (1024, 2048 and 4096) using the following equation:

D= (5-Slope)/2

The found results are provided in Table 3.5.

Table: 3.5: FD by PSD method

Data size, N	Fractal Dimension, D
1024	1.4788
2048	1.4542
4096	1.4763

3.6 Discussion

As in the case of RD method, the best result is achieved with longer data series. From the analysis it is obvious that, PSD and RS methods show least biased results and the FD is close to the actual dimension of 1.5. The results for the RS analysis depend on the number of subintervals, the minimum and maximum lags. It is not clear how important the number of subintervals is. In the analysis, we choose not to use overlapping sub intervals so that extra dependency between these sub intervals occurs in calculating the FD.

Since the simulated data, generated from a nonlinear equation, represents a parametric model may have different characteristics from the actual natural data. So, no single method can be established as the best one at this stage to relate the nature of data series and the calculated FD. Hence, we need to apply all the methods to calculate FD of heart rate data. The study is presented in the Chapter 5.

Chapter 4
ELECTROCARDIOGRAM

4.1 Introduction

An electrocardiogram (ECG or EKG) is a recording of the electrical activity of the heart over time produced by an electrocardiograph, usually in a noninvasive recording via skin electrodes. Its name is made of different parts: electro, because it is related to electrical activity, cardio, Greek for heart, gram, a Greek root meaning "to write".

Electrical impulses in the heart originate in the sinoatrial node and travel through the heart muscle where they impart electrical initiation of systole or contraction of the heart. The electrical waves can be measured at selectively placed electrodes (electrical contacts) on the skin. Electrodes on different sides of the heart measure the activity of different parts of the heart muscle. An ECG displays the voltage between pairs of these electrodes, and the muscle activity that they measure, from different directions, also understood as vectors. This display indicates the overall rhythm of the heart and weaknesses in different parts of the heart muscle. It is the best way to measure and diagnose abnormal rhythms of the heart, particularly abnormal rhythms caused by damage to the conductive tissue that carries electrical signals, or abnormal rhythms caused by levels of dissolved salts (electrolytes), such as potassium, that are too high or low. In myocardial infarction (MI), the ECG can identify damaged heart muscle. But it can only identify damage to muscle in certain areas, so it can't rule out damage in other areas. The ECG cannot reliably measure the pumping ability of the heart; for which ultrasound-based (echocardiography) or nuclear medicine tests are used.

4.2 ECG Leads

The term lead is used to indicate a particular group of electrode. Because ECG signal is measured from electrodes placed on the surface of the body, the waveform of the signal is varies depending on the placement of the electrodes.

4.2.1 Limb Leads

Leads I, II and III are the so-called limb leads because at one time, the subjects of electrocardiography had to literally place their arms and legs in buckets of salt water in order to obtain signals for Einthoven's string galvanometer. They form the basis of what is known as Einthoven's triangle. Eventually, electrodes were invented that could be placed directly on the patient's skin. Even though the buckets of salt water are no longer necessary, the electrodes are still placed on the patient's arms and legs to approximate the signals obtained with the buckets of salt water. They remain the first three leads of the modern 12 lead ECG.

Lead I is a dipole with the negative (white) electrode on the right arm and the positive (black) electrode on the left arm.

Lead II is a dipole with the negative (white) electrode on the right arm and the positive (red) electrode on the left leg.

Lead III is a dipole with the negative (black) electrode on the left arm and the positive (red) electrode on the left leg.

4.2.2 Augmented V Leads

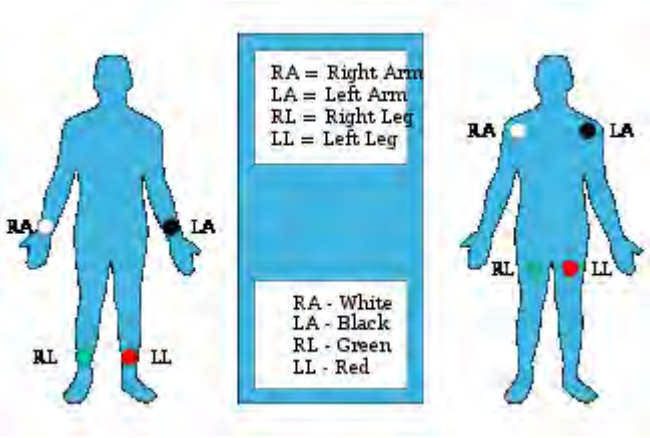


Fig 4.1 Augmented Limb Leads

Leads aVR, aVL, and aVF are '**augmented limb leads'**. They are derived from the same three electrodes as leads I, II, and III. However, they view the heart from different angles (or vectors) because the negative electrode for these leads is a modification of 'Wilson's central terminal', which is derived by adding leads I, II, and III together and plugging them into the negative terminal of the ECG machine.

- Lead aVR or "augmented vector right" has the positive electrode (white) on the right arm. The negative electrode is a combination of the left arm (black) electrode and the left leg (red) electrode, which "augments" the signal strength of the positive electrode on the right arm.
- Lead aVL or "augmented vector left" has the positive (black) electrode on the left arm. The negative electrode is a combination of the right arm (white) electrode and the left leg (red) electrode, which "augments" the signal strength of the positive electrode on the left arm.
- Lead aVF or "augmented vector foot" has the positive (red) electrode on the left leg. The negative electrode is a combination of the right arm (white) electrode and the left arm (black) electrode, which "augments" the signal of the positive electrode on the left leg.

The augmented limb leads aVR, aVL, and aVF are amplified in this way because the signal is too small to be useful when the negative electrode is Wilson's central terminal. Together with leads I, II, and III, augmented limb leads aVR, aVL, and aVF form the basis of the hexaxial reference system, which is used to calculate the heart's electrical axis in the frontal plane.

4.2.3 V Leads

The V leads V1, V2, V3, V4, V5, and V6 are placed directly on the chest. Because of their close proximity to the heart, they do not require augmentation. Wilson's central terminal is used for the negative electrode, and these leads are considered to be unipolar. The precordial leads view the heart's electrical activity in the so-called horizontal plane. The heart's electrical axis in the horizontal plane is referred to as the Z axis. Leads V1, V2, and V3 are referred to as the right precordial leads and V4, V5, and V6 are referred to as the left precordial leads. The QRS complex should be negative in lead V1 and positive in lead V6. The QRS complex should show a gradual transition from negative to positive between

leads V2 and V4. The equiphasic lead is referred to as the transition lead. When the transition occurs earlier than lead V3, it is referred to as an early transition. When it occurs later than lead V3, it is referred to as a late transition. There should also be a gradual increase in the amplitude of the R wave between leads V1 and V4. This is known as R wave progression. Poor R wave progression is a nonspecific finding. It can be caused by conduction abnormalities, myocardial infarction, cardiomyopathy, and other pathological conditions.

Lead V1 is placed in the fourth intercostal space to the right of the sternum.

Lead V2 is placed in the fourth intercostal space to the left of the sternum.

Lead V3 is placed directly between leads V2 and V4.

Lead V4 is placed in the fifth intercostal space in the midclavicular line (even if the apex beat is displaced).

Lead V5 is placed horizontally with V4 in the anterior axillary line

Lead V6 is placed horizontally with V4 and V5 in the midaxillary line.

4.2.4 Waves and intervals

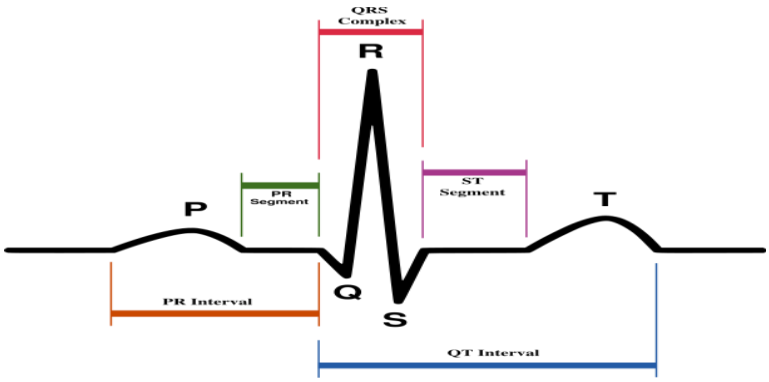


Fig 4.2: Schematic representation of normal ECG

A typical signal from one ECG lead is quasi-periodic with periodicity of one second. Within one period, called a cardiac cycle, the ECG is comprised of 4 to 5 bumps of different characteristics. This pattern of bumps, known as waves, represents the electrical activity coordinating a series of muscle contraction that makes one heart beat. The electrical current generated by atrial depolarization is recorded as the P wave and that generated by ventricular depolarization is recorded as the Q, R, and S waves: the QRS complex. Atrial repolarization is recorded as atrial T wave (Ta) and ventricular repolarization as ventricular T wave or simply T wave. Because atrial repolarization

normally occurs during ventricular depolarization, the atrial T wave (Ta) is buried or hidden in the QRS complex. In a normal cardiac cycle, the P wave occurs first, followed by the QRS complex and the T wave as shown in figure 4.2. The section of ECG between the waves and complexes are called segments and intervals: the PR segment, the ST segment, the TP segment, the PR interval, the QT interval and the RR interval. Intervals include waves and complexes, whereas segments do not. When electrical activity of heart is not being detected, the ECG is a straight line, flat line- the isoelectric line or base line.

4.3 Typical Values of Normal ECG Parameters

Heart rate	120 beats per minute (BPM)
Amplitudes:	
P wave	0.25mV
R wave	1.60mV
Q wave	25% of R wave
T wave	0.1 to 0.5mV
Duration:	
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	0.11sec
QRS interval	0.09 sec

4.4 Loading of Data

The source of the ECGs included in the MIT-BIH Arrhythmia Database is a set of over 4000 long-term Holter recordings that were obtained by the Beth Israel Hospital Arrhythmia Laboratory between 1975 and 1979. Approximately 60% of these recordings were obtained from inpatients. The database contains 23 records (numbered from 100 to 124 inclusive with some numbers missing) chosen at random from this set, and 25 records (numbered from 200 to 234 inclusive, again with some numbers missing) selected from the same set to include a variety of rare but clinically important phenomena that would not be well-represented by a small random sample of Holter recordings. Each of the 48 records is slightly over 30 minutes long. An initial set of beat labels was produced by a simple slope-sensitive QRS detector, which marked each detected event as a normal beat.

Two identical 150-foot chart recordings were printed for each 30-minute record, with these initial beat labels in the margin. For each record, the two charts were given to two cardiologists, who worked on them independently. The cardiologists added additional beat labels where the detector missed beats, deleted false detections as necessary, and changed the labels for all abnormal beats. They also added rhythm labels, signal quality labels, and comments. The annotations were transcribed from the paper chart recordings. Once both sets of cardiologists' annotations for a given record had been transcribed and verified, they were automatically compared beat-by-beat, and another chart recording was printed. This chart showed the cardiologists' annotations in the margin, with all discrepancies highlighted. Each discrepancy was reviewed and resolved by consensus. The corrections were transcribed, and the annotations were then analyzed by an auditing program, which checked them for consistency and which located the ten longest and shortest R-R intervals in each record (to identify possible missing or falsely detected beats).

The analog outputs of the playback unit were filtered to limit analog-to-digital converter (ADC) saturation and for anti-aliasing, using a passband from 0.1 to 100 Hz relative to real time, well beyond the lowest and highest frequencies recoverable from the recordings. The bandpass-filtered signals were digitized at 360 Hz per signal relative to real time using hardware constructed at the MIT Biomedical Engineering Center and at the BIH Biomedical Engineering Laboratory. The sampling frequency was chosen to facilitate implementations of 60 Hz (mains frequency) digital notch filters in arrhythmia detectors. Since the recorders were battery-powered, most of the 60 Hz noise present in the database arose during playback. In those records that were digitized at twice real time, this noise appears at 30 Hz (and multiples of 30 Hz) relative to real time. The ECG’s are taken for 30 minutes and sampled at 360Hz. The database includes ECG (Holter recordings) of 50 persons, including both normal and abnormal heart condition.

The database contains approximately 109,000 beat labels. Sixteen were corrected in the first seven years after the database was released in 1980 (in records 104, 108, 114, 203, 207, 217, and 222); in addition, all of the left bundle branch block beats in record 214 were originally labelled as normal beats. The rhythm labels have been more substantially revised and now include notations for paced rhythm, bigeminy, and trigeminy, which were missing in early copies.

Table 4.1: The beat types in MIT-BIH data base (entire records)

Record	N							V	F	O	N	E	P	F	O	Q	Record
	.	L	R	A	a	J	S	V	F	!		.	L	R	A	a	J
100	2239	-	-	33	-	-	-	1	-	-	-	-	-	-	-	-	-
101	1860	-	-	3	-	-	-	-	-	-	-	-	-	-	-	-	2
102	99	-	-	-	-	-	-	4	-	-	-	-	-	2028	56	-	-
103	2082	-	-	2	-	-	-	-	-	-	-	-	-	-	-	-	-
104	163	-	-	-	-	-	-	2	-	-	-	-	-	1380	666	-	18
105	2526	-	-	-	-	-	-	41	-	-	-	-	-	-	-	-	5
106	1507	-	-	-	-	-	-	520	-	-	-	-	-	-	-	-	-
107	-	-	-	-	-	-	-	59	-	-	-	-	-	2078	-	-	-
108	1739	-	-	4	-	-	-	17	2	-	-	1	-	-	-	11	-
109	-	2492	-	-	-	-	-	38	2	-	-	-	-	-	-	-	-
111	-	2123	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-
112	2537	-	-	2	-	-	-	-	-	-	-	-	-	-	-	-	-
113	1789	-	-	-	6	-	-	-	-	-	-	-	-	-	-	-	-
114	1820	-	-	10	-	2	-	43	4	-	-	-	-	-	-	-	-
115	1953	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
116	2302	-	-	1	-	-	-	109	-	-	-	-	-	-	-	-	-
117	1534	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-
118	-	-	2166	96	-	-	-	16	-	-	-	-	-	-	-	10	-
119	1543	-	-	-	-	-	-	444	-	-	-	-	-	-	-	-	-
121	1861	-	-	1	-	-	-	1	-	-	-	-	-	-	-	-	-
122	2476	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
123	1515	-	-	-	-	-	-	3	-	-	-	-	-	-	-	-	-
124	-	-	1531	2	-	29	-	47	5	-	-	5	-	-	-	-	-
200	1743	-	-	30	-	-	-	826	2	-	-	-	-	-	-	-	-
201	1625	-	-	30	97	1	-	198	2	-	-	10	-	-	-	37	-
202	2061	-	-	36	19	-	-	19	1	-	-	-	-	-	-	-	-
203	2529	-	-	-	2	-	-	444	1	-	-	-	-	-	-	-	4
205	2571	-	-	3	-	-	-	71	11	-	-	-	-	-	-	-	-
207	-	1457	86	107	-	-	-	105	-	472	-	-	105	-	-	-	-
208	1586	-	-	-	-	-	2	992	373	-	-	-	-	-	-	-	2

209	2621	-	-	383	-	-	-	1	-	-	-	-	-	-	-	-	-
210	2423	-	-	-	22	-	-	194	10	-	-	-	1	-	-	-	-
212	923	-	1825	-	-	-	-	-	-	-	-	-	-	-	-	-	-
213	2641	-	-	25	3	-	-	220	362	-	-	-	-	-	-	-	-
214	-	2003	-	-	-	-	-	256	1	-	-	-	-	-	-	-	2
215	3195	-	-	3	-	-	-	164	1	-	-	-	-	-	-	-	-
217	244	-	-	-	-	-	-	162	-	-	-	-	-	1542	260	-	-
219	2082	-	-	7	-	-	-	64	1	-	-	-	-	-	-	133	-
220	1954	-	-	94	-	-	-	-	-	-	-	-	-	-	-	-	-
221	2031	-	-	-	-	-	-	396	-	-	-	-	-	-	-	-	-
222	2062	-	-	208	-	1	-	-	-	-	-	212	-	-	-	-	-
223	2029	-	-	72	1	-	-	473	14	-	16	-	-	-	-	-	-
228	1688	-	-	3	-	-	-	362	-	-	-	-	-	-	-	-	-
230	2255	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-
231	314	-	1254	1	-	-	-	2	-	-	-	-	-	-	-	2	-
232	-	-	397	1382	-	-	-	-	-	-	-	1	-	-	-	-	-
233	2230	-	-	7	-	-	-	831	11	-	-	-	-	-	-	-	-
234	2700	-	-	-	-	50	-	3	-	-	-	-	-	-	-	-	-

Symbols used in plots

Symbol	Meaning
· or N	Normal beat
L	Left bundle branch block beat
R	Right bundle branch block beat
A	Atrial premature beat
a	Aberrated atrial premature beat
J	Nodal (junctional) premature beat
S	Supraventricular premature beat
V	Premature ventricular contraction
F	Fusion of ventricular and normal beat
[	Start of ventricular flutter/fibrillation
!	Ventricular flutter wave
]	End of ventricular flutter/fibrillation
e	Atrial escape beat
j	Nodal (junctional) escape beat

E	Ventricular escape beat
/	Paced beat
f	Fusion of paced and normal beat
x	Non-conducted P-wave (blocked APB)
Q	Unclassifiable beat
|	Isolated QRS-like artifact
Rhythm annotations appear below the level used for beat annotations:	
AB	Atrial bigeminy
AFIB	Atrial fibrillation
AFL	Atrial flutter
B	Ventricular bigeminy
BII	2° heart block
IVR	Idioventricular rhythm
N	Normal sinus rhythm
NOD	Nodal (A-V junctional) rhythm
P	Paced rhythm
PREX	Pre-excitation (WPW)
SBR	Sinus bradycardia
SVTA	Supraventricular tachyarrhythmia
T	Ventricular trigeminy
VFL	Ventricular flutter
VT	Ventricular tachycardia
Signal quality and comment annotations appear above the level used for beat annotations:	
qq	Signal quality change: the first character (`c' or `n') indicates the quality of the upper signal (clean or noisy), and the second character indicates the quality of the lower signal
U	Extreme noise or signal loss in both signals: ECG is unreadable
M or MISSB	Missed beat
P or PSE	Pause
T or TS	Tape slippage

Chapter 5

FRACTAL ANALYSIS OF HEART RATE

5.1 Introduction

The ECG represents the cardiac electrical activities of the heart and it provides information about the cardiac conditions. The mean and instantaneous heart rates, derived from ECG, can provide vital information about the functioning of heart. The heart rate can vary widely even for a short period of time, instantaneous heart rate (IHR) can provide more information. There have been different methods developed for the analysis of the IHR. Due to the large number of patients in intensive care units and the need for continuous observation of such conditions, several techniques for automated abnormality detection have been developed. Classical techniques based on time and frequency domain analysis have long been used to detect cardiac abnormality. However, since ECG as well as IHR is nonlinear in nature, these techniques seem to provide only a limited amount of information as they normally ignore the underlying nonlinear dynamics. So, there has been an increasing interest in applying nonlinear techniques in analyzing electrocardiogram to detect cardiac abnormalities. A technique of nonlinear analysis, the fractal analysis, is recently having its popularity to many researchers working on nonlinear data for which most mathematical models produce intractable solutions. This chapter provides the calculation of fractal dimension (FD) of IHR derived from normal and abnormal ECG to estimate the cardiac conditions. All the ECG records have been taken from MIT-BIH data base. We have used data records 115, 117 and 122 as the representatives of normal ECG. Three types of abnormalities have been considered. They are premature ventricular contraction (PVC), bundle branch block (BBB) and atrial premature beat (APB). The data sets taken for PVC are 116, 119 and 221, for BBB are 109, 111, 207 and 214, while that for APB are 209, 220 and 232.

5.2 Calculation of IHR

The IHR was calculated from ECG of MIT-BIH data base. The calculation is performed after detecting R waves of ECG. For this, the QRS complex of ECG was detected. The detailed description of detecting QRS complex is provided in [14] and reproduced in Appendix. Figure 5.1 represents the ECG of 3 records of healthy subjects.

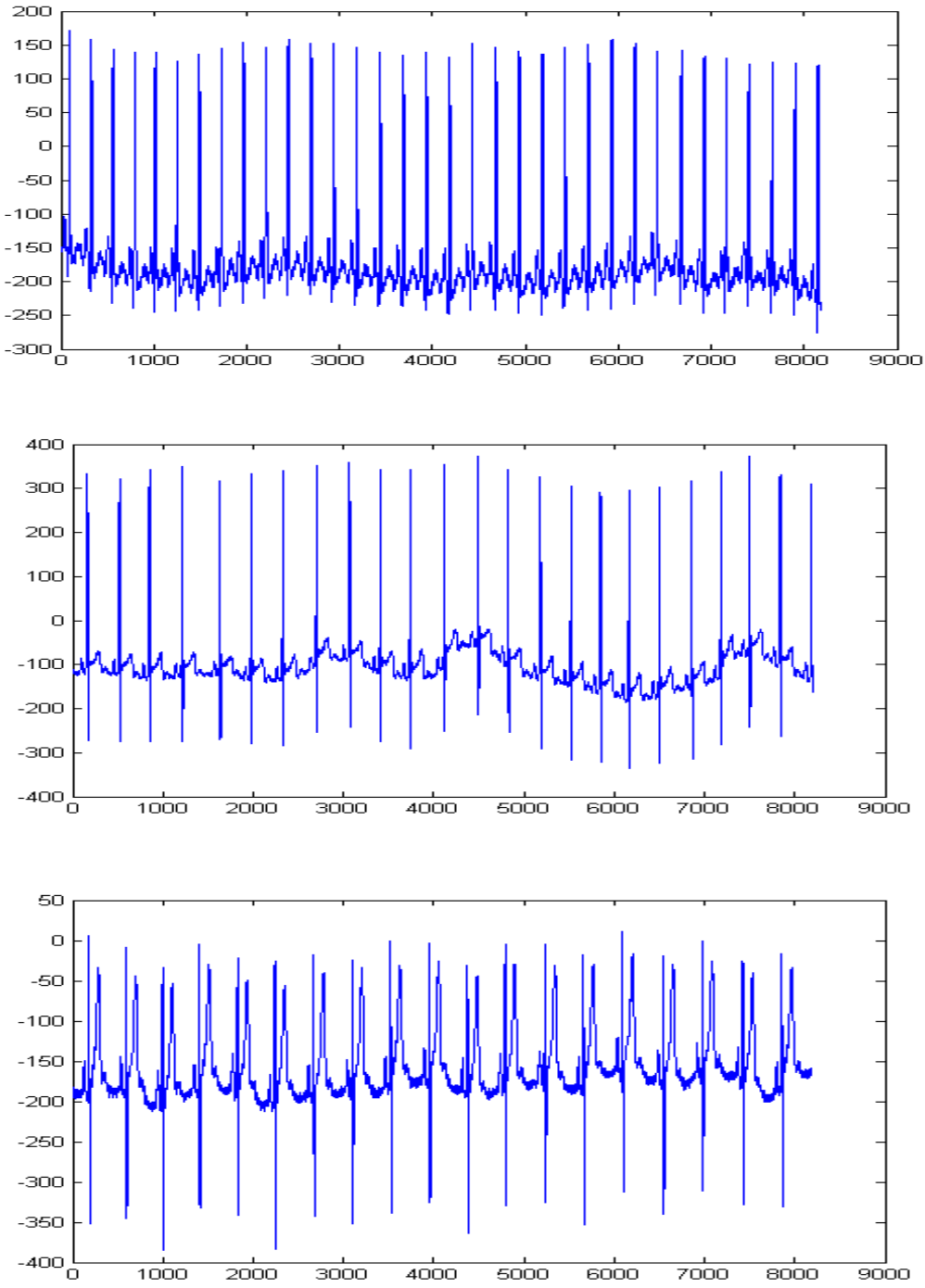


Fig 5.1: ECG of healthy subjects. From top panel: Record 115, 117 and 122

5.3 FD of IHR of Healthy Subjects

The IHR data of healthy persons are loaded using MATLAB code 'mitdb'. The three methods namely relative dispersion (RD), rescaled range (RS) and power spectral density (PSD) methods are applied to calculate the FD of each data set. The procedures for

calculating FD by each method are described in Chapter 3. In Chapter 3, it was found that FD depends on number of data points taken for calculation and we have found consistent and best results by all methods for a data length of 4096. So, in calculating FD by each method, 4096 data points are taken. The results found by the three methods are provided in Table 5.1.

Table 5.1: Fractal Dimension of IHR of Healthy Persons by three methods

Data set	FD by RD method	FD by RS method	FD by PSD method
115	1.40	1.66	1.80
117	1.54	1.66	1.82
122	1.45	1.65	1.77

5.4 FD of IHR of Subjects with PVC Beats

Premature ventricular contraction (PVC), also known as ventricular premature beat (VPB) or extra systole is a relatively common condition where the heartbeat is initiated by the heart ventricles rather than the sinoatrial node, the normal heartbeat initiator. This may be perceived as a "skipped beat" or felt as palpitations in the chest. In a normal heartbeat, the ventricles contract after the atria, forcing blood both to the body and to the lungs. In a PVC, the ventricles contract first which means that circulation is inefficient. However, single beat PVC arrhythmias do not usually pose a danger and can be asymptomatic in healthy individuals**.** Premature ventricular contraction can occur in a healthy person of any age, but becomes more frequent in the elderly, and is more commonly found in men.

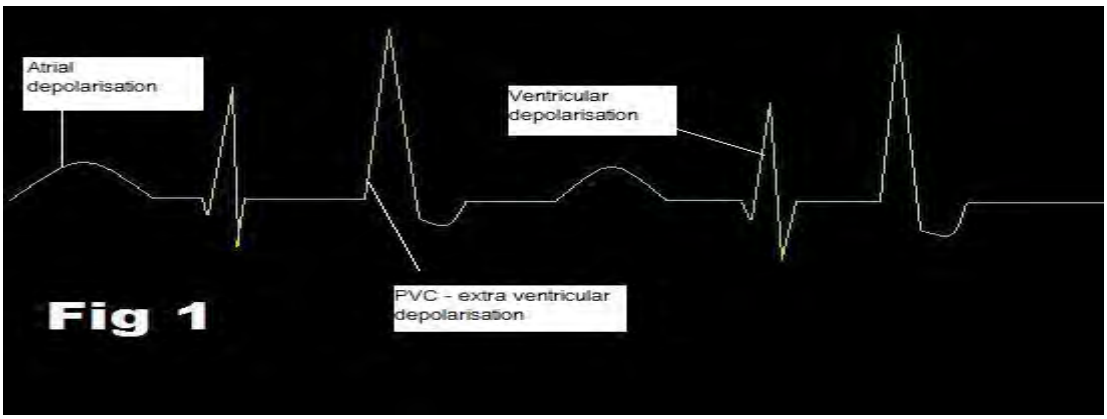


Fig 5.2: A typical ECG of a patient with PVC

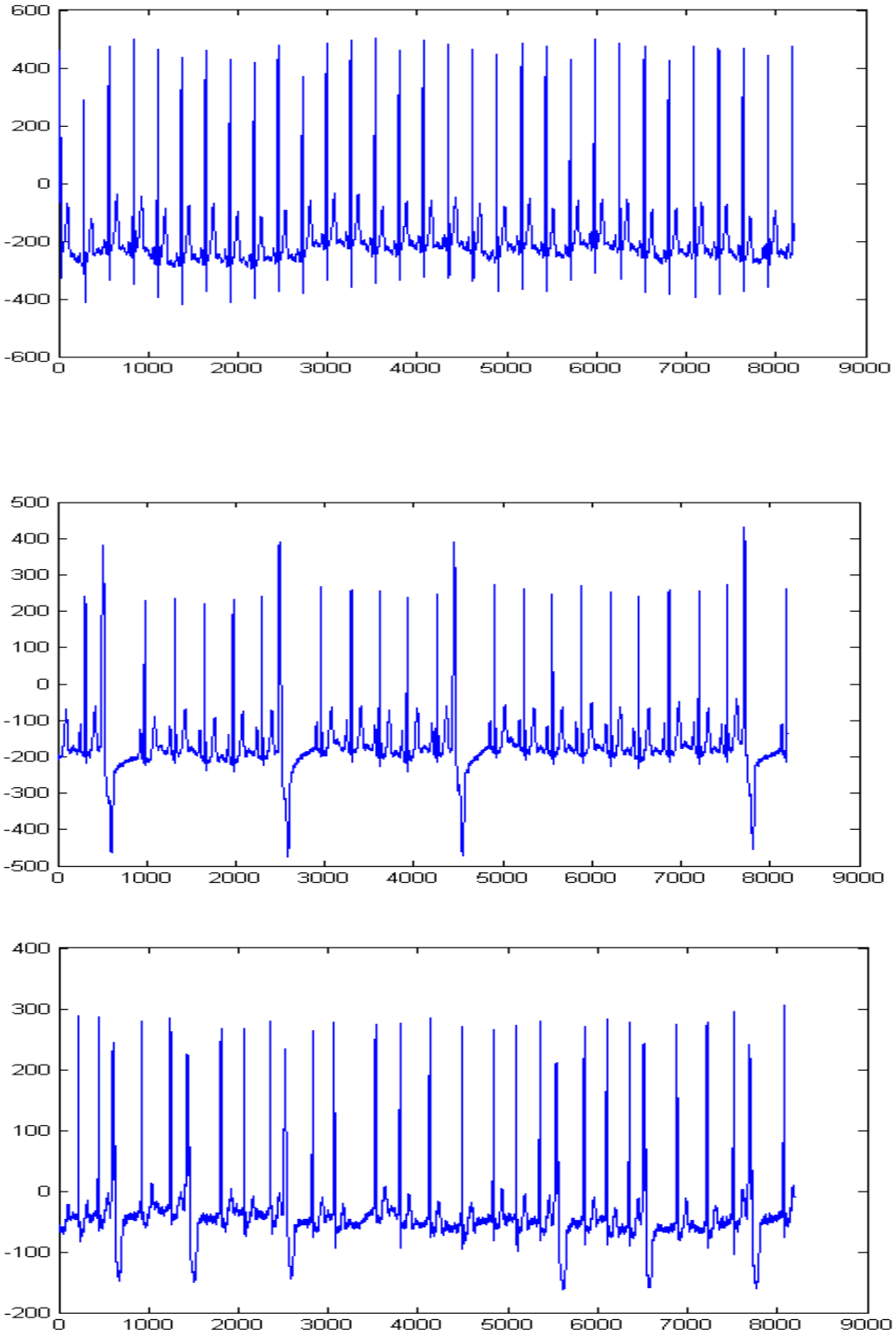


Fig 5.3: ECG of PVC subjects. From top panel: Record 116, 119 and 221

The abnormalities present in the records are provided below:

Record	*N*							*V*	*F*	*O*	*N*	*E*	*P*	*F*	*O*	*Q*	*V*
	.	*L*	*R*	*A*	*a*	*J*	*S*	*V*	*F*	*!*	*e*	*j*	*E*	*P*	*f*	*p*	*Q*
116	2302	-	-	1	-	-	-	109	-	-	-	-	-	-	-	-	-
119	1543	-	-	-	-	-	-	444	-	-	-	-	-	-	-	-	-
221	2031	-	-	-	-	-	-	396	-	-	-	-	-	-	-	-	-

In the above chart 'V' denotes 'Premature Ventricular Contraction'. These records contain only PVC and normal beats. Fractal dimensions obtained from these data sets are provided in Table 5.2.

Table 5.2: Fractal Dimension of IHR of PVC beats by three methods

Data set	FD by RD method	FD by RS method	FD by PSD method
116	1.47	1.53	1.71
119	1.43	1.48	1.70
221	1.40	1.53	1.72

5.5 FD of IHR of Subjects with BBB

Excitation of the ventricular myocardium develops in an orderly manner because of the conduction system of the A-V node, bundle of His (with left and right branches) and Purkinje fibers. This conduction system has a propagation velocity that is higher than in ventricular muscle. The net result is that both ventricles contracts almost simultaneously and the beat that results develops the maximum force. BBB is of two types: left bundle branch block (LBBB) and right bundle branch block (RBBB). A block of excitation in one of the branches will cause the effected side to be excited, that is, depolarize late and this fact will be revealed in a prolongation of the duration of the QRS. So, in BBB, the ventricles will be excited serially, the effected side being excited last. Excitation first appears on the surface of the thin walled right ventricles owing to the nature of the conduction system and differing thickness of ventricular Myocardium. In ECG signal, the right sided V leads show a large, broad, downward wave. Left-sided V leads show a large, broad, upward wave. QRS is prolonged and downward in V_5-V_6. Little evidence delayed

conduction is shown in V_{1-3}. We have considered data records 109, 111, 207 and 214 that contain LBBB only. The types of beats in these records are provided below:

	N							*V*	*F*	*O*	*N*	*E*	*P*	*F*	*O*	*Q*	*V*
Record	.	*L*	*R*	*A*	*a*	*J*	*S*	*V*	*F*	*!*	*e*	*j*	*E*	*P*	*f*	*p*	*Q*
109	-	*2492*	-	-	-	-	-	*38*	*2*	-	-	-	-	-	-	-	-
111	-	*2123*	-	-	-	-	-	*1*	-	-	-	-	-	-	-	-	-
207	-	*1457*	*86*	*107*	-	-	-	*105*	-	*472*	-	-	*105*	-	-	-	-
214	-	*2003*	-	-	-	-	-	*256*	*1*	-	-	-	-	-	-	-	*2*

Although the records contain other types of abnormalities as well, LBBB are prominent. The ECG waves of these records are provided in Figure 5.4.

Table 5.3 provides the FD of IHR from LBBB patients.

Table 5.3: Fractal Dimension of IHR of LBBB beats by three methods

Data set	FD by RD method	FD by RS method	FD by PSD method
109	1.27	1.73	1.68
111	1.26	1.76	1.66
207	1.29	1.72	1.76
214	1.29	1.74	1.70

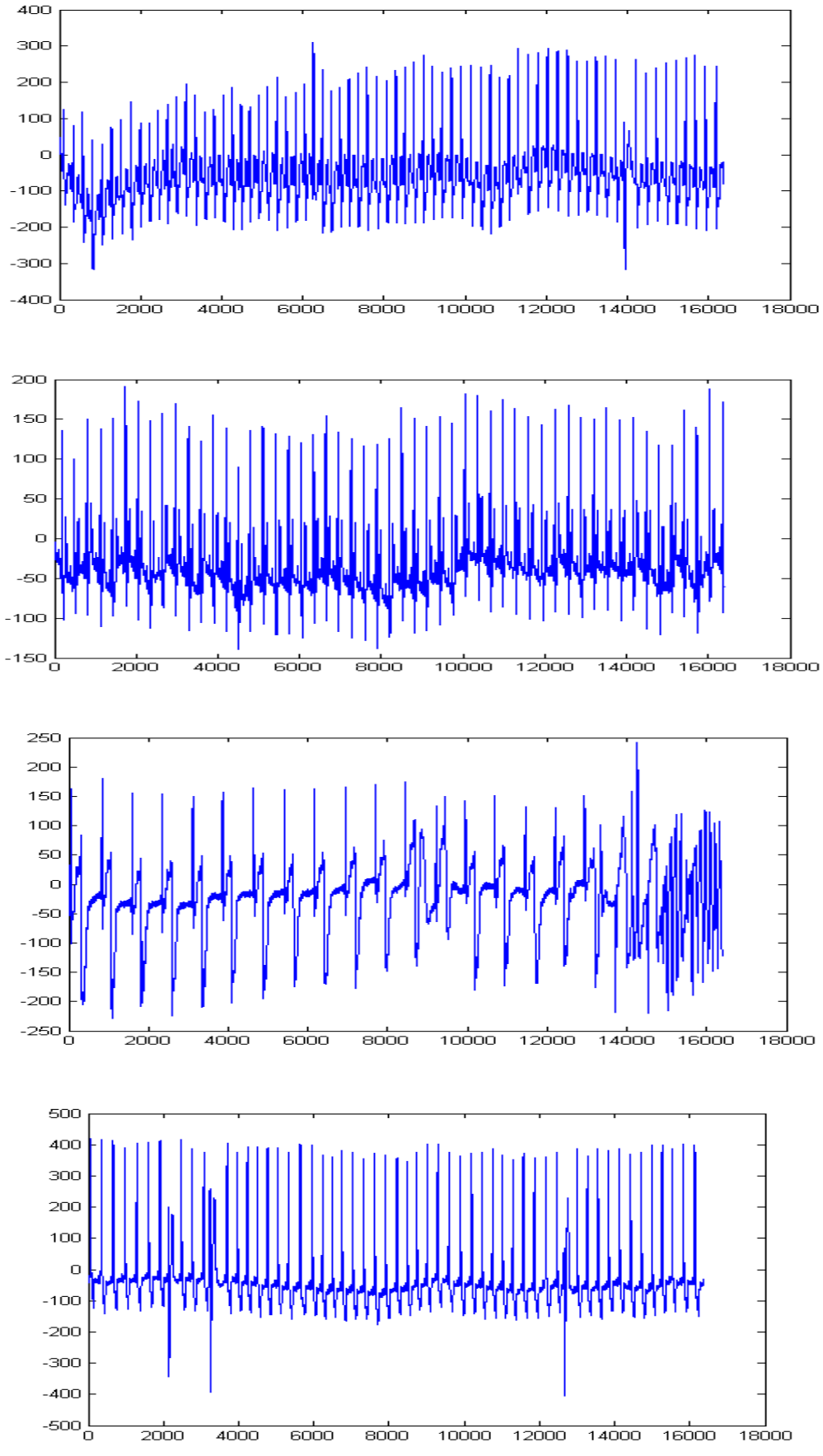


Fig 5.4: ECG of LBBB subjects. From top panel: Record 109, 111, 207 and 214

5.6 FD of IHR of Subjects with APB

The term atrial premature beat describes beats arising from the atrium and occurring before the expected sinus beats. Premature beats can occur randomly or in a pattern. P-wave morphology differs from sinus beats and varies, depending on the origin of the impulse in the atria. In general, beats arising in the high atria have upright P waves in leads II, III and aVF, and those originating in the low atria have inverted P waves in those leads. Depending on the prematurity of the atrial impulse and refractoriness of the AV node and His-Purkinje system, the P wave may conduct with normal or prolonged PR interval with narrow or aberrant QRS complex or may block and not be followed by a QRS complex. Premature atrial beats generally depolarize and reset the sinus node so that the next sinus beat occurs sooner than the expected sinus P wave. Expressed differently, the interval from sinus to atrial premature beat plus the next sinus beat measures less than two sinus cycles. This phenomenon, incomplete compensatory pause, is fairly common with atrial premature beats. Atrial premature beats (APBs) are slightly less frequent than ventricular premature beats (VPBs) but much more frequent than supraventricular premature beats arising in the atrio-ventricular (AV) junction. Often, APBs are not associated with heart disease. In up to 64% of healthy young individuals, some APBs are found in an ambulatory electrocardiogram (Holter ECG), and in most cases without symptoms.

It has been empirically noted that alcohol in excess, fatigue, cigarettes, fever, anxiety and infectious diseases of all sorts may result in atrial premature beats, and elimination of these factors may correct the disorder. Atrial premature beats also occur in some patients with CHF and in some with myocardial ischemia. They may also be noted in patients with underlying myocardial disease, in those with pulmonary disease and systemic hypoxia. Premature beats are usually perceived by patients as "palpitations," and sometimes are of concern to patients who notice "skipped" beats or "fluttering" that are frightening. Often atrial premature beats cause no symptoms and are not recognized by the patient. Atrial premature beats require no specific therapy. If they are a source of aggravation or concern to the patient and one of the above-mentioned factors can be identified and easily corrected, this should be done. Some ECG signals of data with APB are shown Figure 5.5.

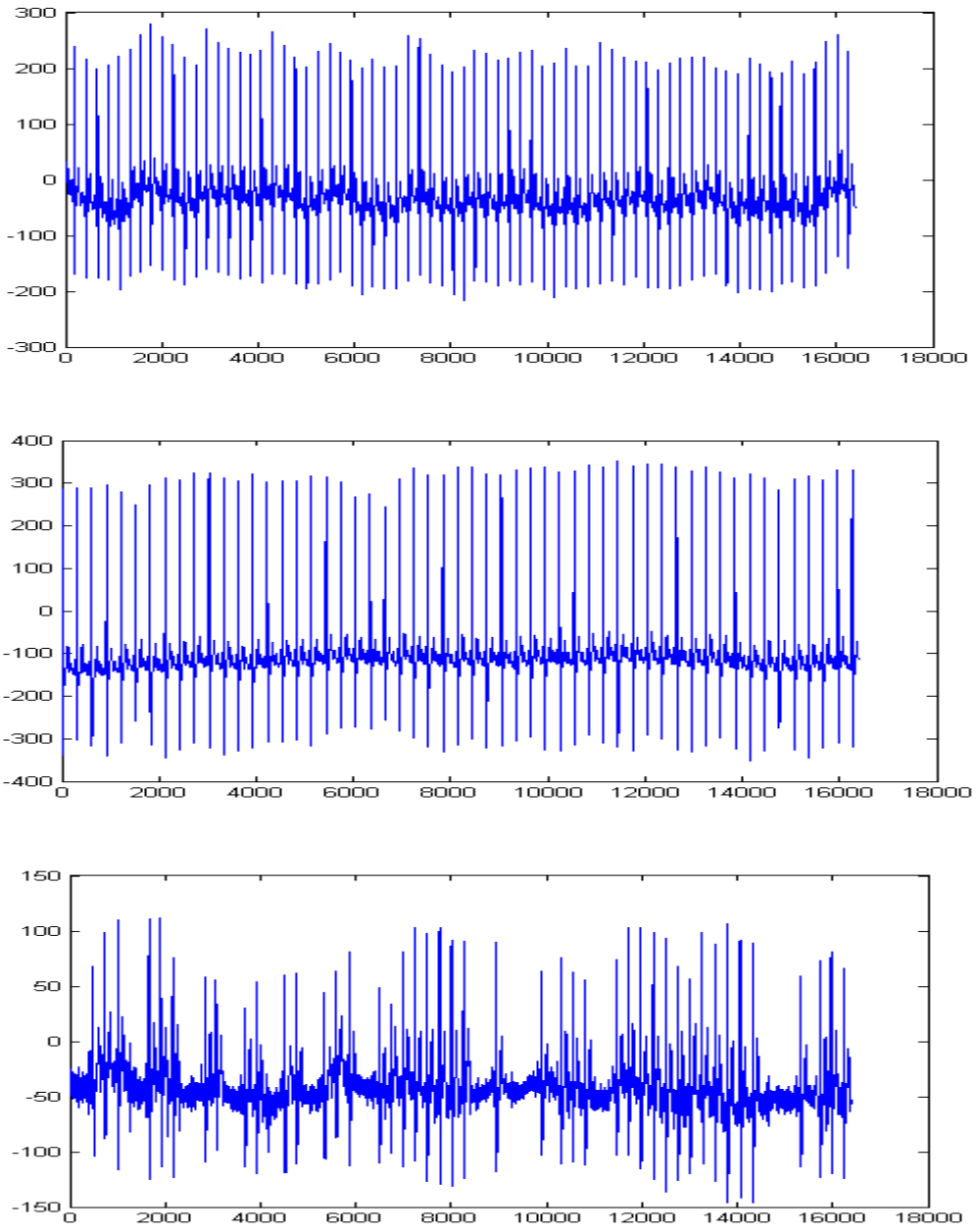


Fig 5.5: ECG of APB subjects. From top panel: Record 209, 220 and 232

The types of beats present in these records are provided below where A denotes APB:

	N							*V*	*F*	*O*	*N*	*E*	*P*	*F*	*O*	*Q*	*V*
Record	.	*L*	*R*	*A*	*a*	*J*	*S*	*V*	*F*	*!*	*e*	*j*	*E*	*P*	*f*	*p*	*Q*
209	2621	-	-	383	-	-	-	1	-	-	-	-	-	-	-	-	-
220	1954	-	-	94	-	-	-	-	-	-	-	-	-	-	-	-	-
232	-	-	397	1382	-	-	-	-	-	-	-	1	-	-	-	-	-

Fractal dimensions obtained from these data are provided in Table 5.4.

Table 5.4: Fractal Dimension of IHR of APB by three methods

Data set	FD by RD method	FD by RS method	FD by PSD method
209	1.35	1.58	1.75
220	1.33	1.55	1.74
232	1.36	1.56	1.74

5.7 Discussion

As in the simulated data, RS and PSD methods provide consistent results. The RD method was found to be very much biased with data length. The summary of the FD, (mean ± standard deviation), in different beat types are provided in Table 5.5.

Table 5.5: FD of IHR for different beat types

Beat type	RD method	RS method	PSD method
Normal (3 data sets)	1.46 ± 0.071	1.66 ± 0.006	1.80 ± 0.025
PVC (3 data sets)	1.43 ± 0.035	1.51 ± 0.029	1.71 ± 0.010
LBBB (4 data sets)	1.28 ± 0.015	1.74 ± 0.017	1.70 ± 0.043
APB (3 data sets)	1.35 ± 0.015	1.56 ± 0.015	1.74 ± 0.006

From the above Table, it is seen that all the three methods can distinguish abnormal beats (PVC, LBBB and APB) from the normal beat type. It should be noted that we have calculated FD of only a few data sets in each case, normal and abnormal. So, we, at this stage, are unable to make absolute comments about the applicability of fractal analysis to estimate cardiac abnormality. However, the analysis shows that there is a great potential of fractal analysis to analyze ECG.

Chapter 6
CONCLUSION

6.1 Discussions

The fractal dimension provides a measure of the heterogeneity of the system and takes the correlation structure into account. In the simulation study where we have taken D as 1.5 implies that we are not dealing with a random distribution, but that there is correlation between adjacent regions. We have compared three numerical methods to estimate the fractal dimension. Relative dispersion analysis is well suited for long signals. We obtained best result for Weierstrass function using RS analysis. This report represents a brief application of fractal theory to model analyze heart rate dynamics. The fractal dimension for normal hearts as well as hearts with various abnormalities is calculated here using three methods. We have found that the three methods applied here can distinguish some typical abnormalities (PVC, LBBB and APB) form the normal ECG. However, due to a very small number of data sets analyzed, we can not make absolute and generalized comments regarding the applicability of the methods.

6.2 Future Perspectives

To make this method applicable to medical science, especially to a nonlinear and nonstationary signal like ECG, much works are needed to be done. A large number of data sets, both of normal and abnormal ECG, should be subjected to the analysis. More rigorous methods of calculating fractal dimension should also be investigated. By determining the fractal dimension of an ECG signal, an estimation of heart condition can be made, which can be helpful in medical science. The same methods can be used for the one-dimensional analysis of other special signals, for the analysis of thicknesses of tree growth rings, and mud layers. For the analysis of two- and three-dimensional signals, the methods can be extended to account for anisotropy, making the analysis more complicated but adhering to the same basic theory. The fractal analysis can also be applied further in image analysis, fluid dynamics, investing natural phenomenon and so on.

Appendix-A: QRS Detection Algorithm

An ECG beat is defined as the signal sample from one R-wave to the next. Figure A.1 shows the block diagram of the QRS detection algorithm.

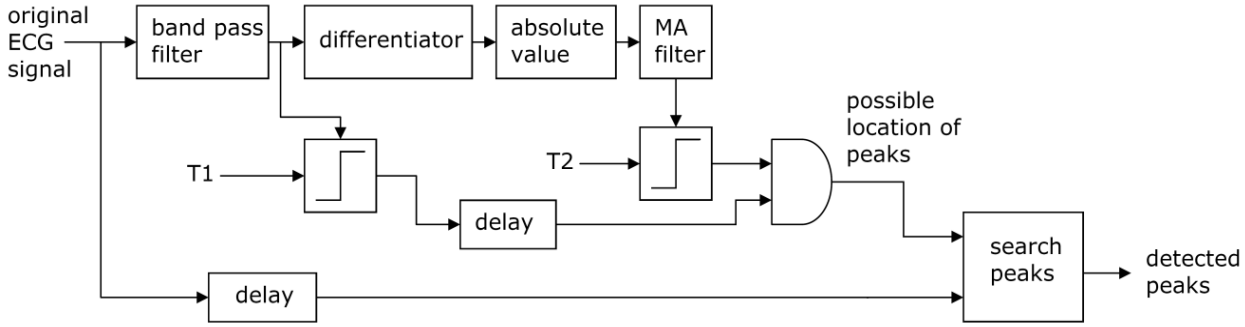


Fig A.1: Block diagram of QRS detection algorithm

At first the ECG signal is passed through a linear phase bandpass filter (4 hz to 40 hz) for smoothing operation and reducing base line shifting. The impulse response, magnitude response and the phase response of the bandpass filter is shown in the figure A.2 and A.3. The group delay of the filter is 150.

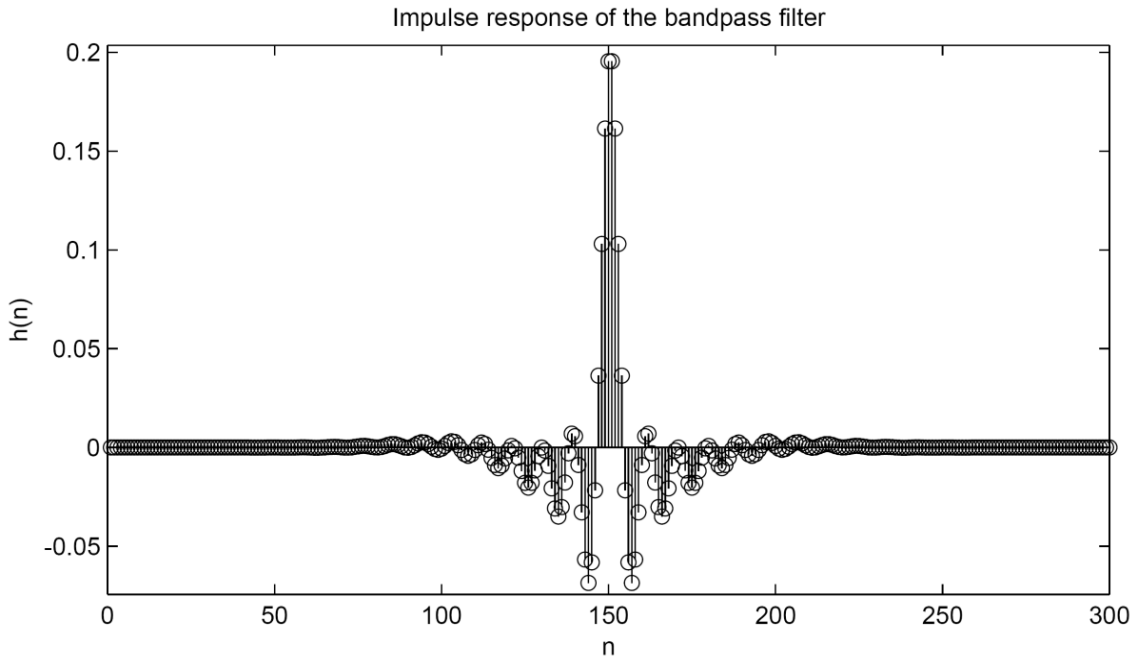


Fig A.2: Impulse response of the bandpass filter

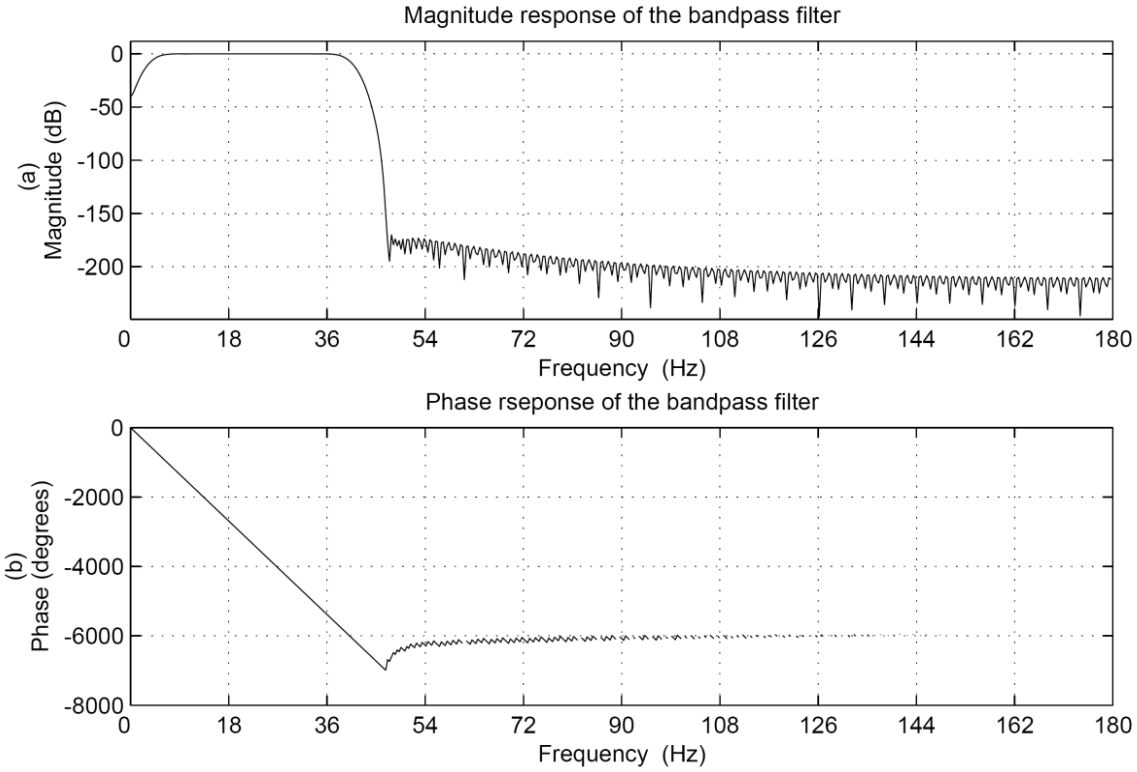


Fig A.3: (a) The magnitude and (b) phase response of the bandpass filter

Differentiation of the filtered signal provides the slope information of QRS complexes. Since there are quick rise and fall times of the QRS complex in the ECG signals, the derivative makes it easier to detect the time of occurrence of the QRS complexes. The transfer function of the five-point differentiation equation is given by

$$H(z) = \left(\frac{1}{8}\right)(-z^{-2} - 2z^{-1} + 2z^{1} + z^{2}) \tag{A.1}$$

The absolute value of the output of derivative filter can be found by the following operation

$$y(n) = \sqrt{x(n)^2} \tag{A.2}$$

This absolute valued signal is then passed through a moving average filter which produces high value at the region of QRS complex. The window size has to be taken properly, neither so wide that merges the QRS complex and T wave together, nor so narrow that produces several peaks in the integration waveform. It is calculated from the equation below:

$$y(n) = \left(\frac{1}{N}\right)[x(n-(N-1)) + x(n-(N-2)) + ... + x(n)] \tag{A.3}$$

where N is the width of the integration window. This work takes N as 18

As seen in figure A.1, the algorithm sets two thresholds T1 and T2 to make decisionsT1 is set 40 for the filtered ECG, and T2 is set 7 for the signals produced by the moving window integration. The thresholded filtered signal is then delayed by 10 samples and a logical 'and' operation is performed with the thresholded moving squared average signal. As a result, possible location of peaks of the original ECG signal is found. These locations are delayed by160 samples from the original ECG. Searching in these regions, we get the peaks. If there are more than one peak in the vicinity of 50 samples, the highest peak is considered from those.

The sequences of the QRS detection algorithm is shown in figure A.4 where the ECG signal is taken from the record 100/MLII of MIT-BIH arrhythmia database.

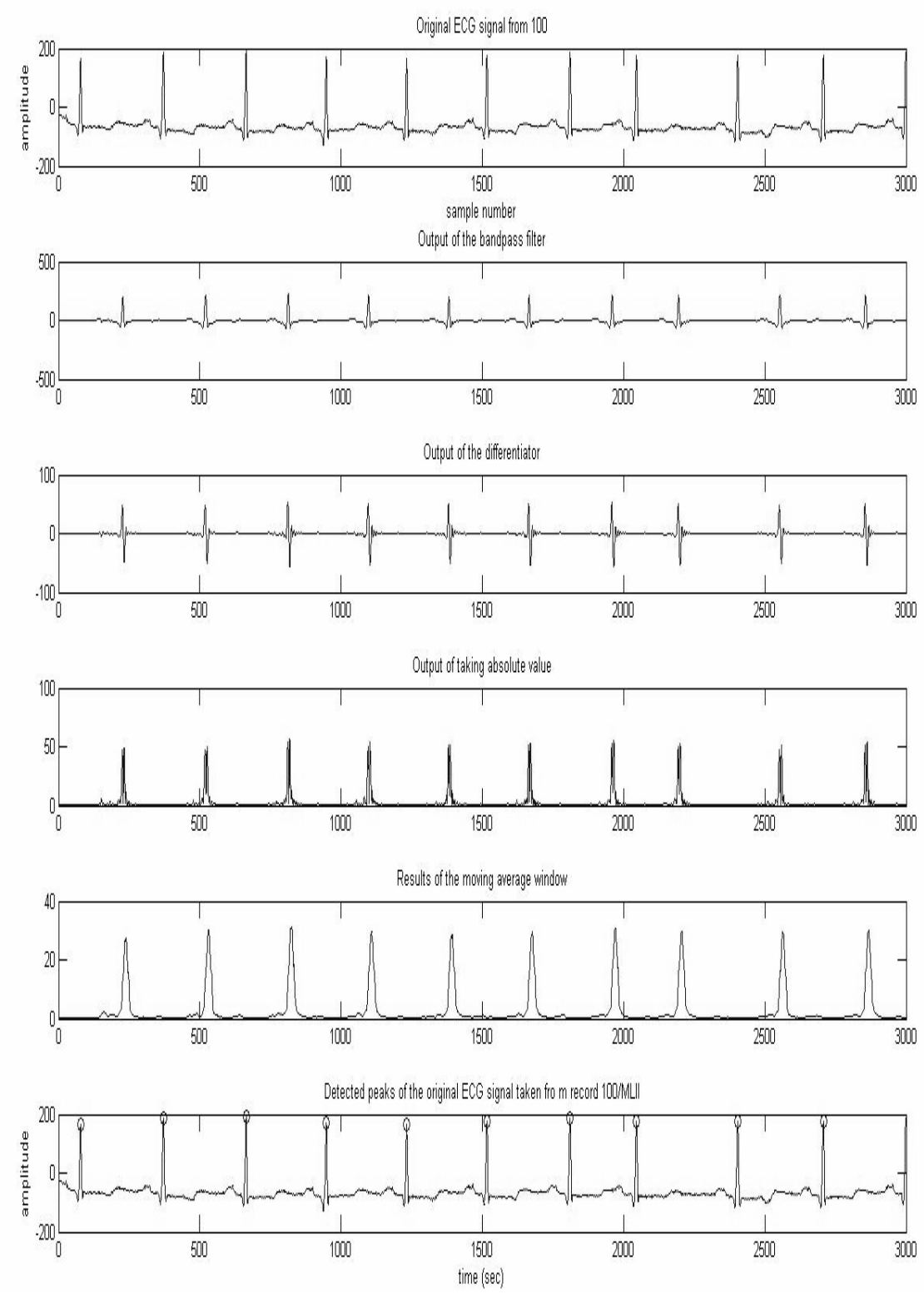


Fig A.4: The sequences of peak detection of the ECG signal taken from the record 100 (MLII) of MIT-BIH arrhythmia database.

REFERENCES

[1] B. B. Mandelbrot, *The Fractal Geometry of Nature*, Freeman, New York, 1982.

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

[3] W. Deering and B. J. West, "Fractal Physiology," *IEEE EMB Mag.*, vol. 12, pp. 40-46, 1992.

[4] A. L. Goldberger, D. R. Rigney and B. J. West, "Chaos and fractals in human physiology," *Scientific American*, vol. 262, pp. 42-49, 1990.

[5] A. Al-Hazimi, N. Al-Ama, A. Syarnic, R. Qosti and K. Abdel-Galil, "Time domain analysis of heart rate variability in diabetic patients with and without autonomic neuropathy," *Annals of Saudi Medicine*, vol. 22, pp. 2002.

[6] T. Koshino, Y. Kimura, Y. Kameyama, T. Takahashi, T. Yasui, H. Chisaka, J. Sugawara and K. Okamura, "Fractal and periodic heart rate dynamics in fetal sheep: comparison of conventional and new measures based on fractal analysis," *Am J Heart Circ Physiol.*, vol. 284, pp.1858-1864, 2003.

[7] A. L. Goldberger and D. R. Rigney, "Defending against sudden death: Fractal mechanisms of cardiac stability," *Proc. 9th Ann. IEEE/EMBS Conference*, pp. 1313-1317, 1987.

[8] B. J. West, "Physiology in fractal dimensions: error tolerance," *Ann Biomed Eng.* vol.18, pp.135-149, 1990.

[9] C. K. Peng, S. Havlin, H. E. Stanley and A. L. Goldberger, "Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series," *Chaos*, vol. 5, pp. 82-87, 1995.

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

[11] D. P. Francis, K. Wilson, P. Georgiadou, R. Wensel, L. C. Davies, A. Coats and M. Piepoli, "Physiological basis of fractal complexity properties of heart rate variability in man," *J Physiol.*, vol. 542, pp. 619-629, 2002.

[12] 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 EMB Mag.*, vol. 12, pp. 57-71, 1992.

[13] P. Bak and K. Chen, "Self-organized criticality," *Sci Amer*, vol. 246, pp. 46-53, 1991.

[14] S. M. A. K. Mojahedy, *Fractal Modeling of Electrocardiogram*, M. Eng. Project, EEE, BUET, 2005.