New Jersey Institute of Technology
Digital Commons @ NJIT

Theses | Electronic Theses and Dissertations

Summer 8-31-2005

Quantification of long-range power law correlations among healthy and pathologic subjects using detrended fluctuation analysis and multifractal detrended fluctuation analysis

Hardick Raval
New Jersey Institute of Technology

Follow this and additional works at: https://digitalcommons.njit.edu/theses

 Part of the Biomedical Engineering and Bioengineering Commons

Recommended Citation

Raval, Hardick, "Quantification of long-range power law correlations among healthy and pathologic subjects using detrended fluctuation analysis and multifractal detrended fluctuation analysis" (2005). *Theses*. 508.
https://digitalcommons.njit.edu/theses/508

This Thesis is brought to you for free and open access by the Electronic Theses and Dissertations at Digital Commons @ NJIT. It has been accepted for inclusion in Theses by an authorized administrator of Digital Commons @ NJIT. For more information, please contact digitalcommons@njit.edu.

Copyright Warning & Restrictions

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a, user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use" that user may be liable for copyright infringement,

This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve violation of copyright law.

Please Note: The author retains the copyright while the New Jersey Institute of Technology reserves the right to distribute this thesis or dissertation

Printing note: If you do not wish to print this page, then select "Pages from: first page # to: last page #" on the print dialog screen

The Van Houten library has removed some of the personal information and all signatures from the approval page and biographical sketches of theses and dissertations in order to protect the identity of NJIT graduates and faculty.

ABSTRACT

QUANTIFICATION OF LONG -RANGE POWER LAW CORRELATIONS AMONG HEALTHY AND PATHOLOGY SUBJECTS USING DETRENDED FLUCTUATION ANALYSIS AND MULTIFRACTAL DETRENDED FLUCTUATION ANALYSIS

by
Hardik Raval

The healthy heartbeat is traditionally thought to be regulated according to the classical principle of homeostasis whereby physiologic systems operate to reduce variability and achieve an equilibrium-like state. However recent studies reveal that under normal conditions, beat-to-beat fluctuation in heart rate display the kind of long-range correlations typically exhibited by the dynamical system far away from equilibrium. In contrast, heart rate time series from patients with severe congestive heart failure show a breakdown of this long-range correlation behavior. Two different non-linear dynamic methods namely Detrended Fluctuation Analysis (DFA) and Multifractal (MF) DFA are used for the quantification of this correlation property in non-stationary physiological time series and it revealed the presence of long-range power law correlation for the group of healthy subjects while breakdown in the long-range power law correlation for the group of subjects with cardiac heart failure.

Application of DFA analysis shows evidence for a crossover phenomenon associated with a change in short($\alpha 1$) and long($\alpha 2$) range scaling exponents. For healthy subjects, calculated value of $\alpha 1$ and $\alpha 2$ (mean value $\pm$ S.D.) are 1.31 ± 0.17 and 1.00 ± 0.07 respectively. For subjects with cardiac heart failure calculated value of $\alpha 1$ and $\alpha 2$ is 0.71 ± 0.20 and 1.24 ± 0.07 respectively i.e. only one scaling exponent is not sufficient to

characterize the entire heart-rate time series which resulted into MF-DFA approach. This suggested that there is more than one exponent values needed to characterize the heart rate time series.

Multifractal DFA is based on generalization of DFA and a MATLAB code is developed to implement the MF-DFA algorithm and to identify whether the given time series under analysis exhibits multifractality or not by generating more than one exponent values for multifractal signal. The value of α for $q>0$ for healthy is 1.04 ± 0.02 and for CHF is found to be 1.32 ± 0.02 and the value of α for $q<0$ for healthy subjects is 3.01 ± 0.26 and for CHF subjects is found to be 3.53 ± 0.14 (mean value $\pm$ S.D.) The student's t-test suggests that p-value is 0.00001 which is less than 0.05 thus the value of α for $q<0$ and $q>0$ among healthy subjects and CHF subjects are statistically different. Value of α for $q>0$ is less than that for $q<0$. And for $q=2$ MF-DFA retains monofractal DFA. Thus, MF-DFA is clearly able to discriminate among the healthy and CHF for $q<0$ as well for $q>0$. MF-DFA also determines which fluctuations i.e. (small or large) dominate for the given interbeat interval time series because for $q<0$ the slow fluctuations dominate whereas for $q>0$ large fluctuations dominate. DFA and MF-DFA were able to discriminate 23 Healthy subjects out of 26 Healthy subjects data sets i.e. true positive specificity is 0.89 and false negative specificity is 0.12 and 9 CHF subjects out of 11 CHF subjects data sets i.e. true positive specificity is 0.82 and false negative specificity is 0.19.

These methods may be of use in distinguishing healthy from pathologic data sets based on the difference in the scaling properties.

QUANTIFICATION OF LONG -RANGE POWER LAW CORRELATIONS AMONG HEALTHY AND PATHOLOGIC SUBJECTS USING DETRENDED FLUCTUATION ANALYSIS AND MULTIFRACTAL DETRENDED FLUCTUATION ANALYSIS

BY
HARDIK RAVAL

A Thesis
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Biomedical Engineering

Department of Biomedical Engineering

August 2005

APPROVAL PAGE

QUANTIFICATION OF LONG -RANGE POWER LAW CORRELATIONS AMONG HEALTHY AND PATHOLOGY SUBJECTS USING DETRENDED FLUCTUATION ANALYSIS AND MULTIFRACTAL DETRENDED FLUCTUATION ANALYSIS

Hardik Raval

Dr. Stanley Reisman, Thesis Advisor — Date
Professor of Biomedical Engineering, NJIT

Dr. Ronald Rockland, Committee Member — Date
Associate Professor of Engineering Technology, NJIT

Dr. Tara Alvarez, Committee Member — Date
Assistant Professor of Biomedical Engineering, NJIT

BIOGRAPHICAL SKETCH

Author: Hardik M. Raval

Degree: Master of Science

Date: August 2005

Graduate and Undergraduate Education:

- Master of Science in Biomedical Engineering

  New Jersey Institute of Technology, Newark, New Jersey, 2005

- Bachelor of Science in Instrumentation and Control Engineering

  Dharmsinh Desai Institute of Technology, Nadiad, Gujarat India, 2003

Major: Biomedical Engineering

"Ideals are like stars; you will not succeed in touching them with your hands, but like the seafaring man on the desert of waters, you choose them as your guides, and following them, you reach your destiny"

Carl Schurz

US Senator from Wisconsin

This thesis would not have been possible without the blessings of my Late Grandfather Kanaiyalal L. Raval and my parents. I dedicate this thesis to them.

ACKNOWLEDGEMENT

My first and foremost gratitude is to Dr. Stanley Reisman, my Graduate advisor and Thesis advisor, who has completely guided me in all technical matters during the research work. He was an overall inspiration to me throughout my study at NJIT, and will be for the years to come.

I am thankful to Dr. Ronald H. Rockland and Dr. Tara Alvarez for sparing crucial time from their schedule to be a part of the committee. Also I am thankful to Prof. Plamen Ivanov, Boston University for extending a helping hand from his side and providing me with necessary guidance whenever needed.

Special thanks to my friend Mr. Sasidhar Tadanki, Phd, Electrical Engineering, Worcester Polytechnic Institute, Worcester, MA for being a constant support and motivation during my thesis. I could not possibly miss expressing my appreciation for Mr. Nirvish Shah, who has been with me since day one of my thesis work and has helped me understand many of technical fundamentals during the research. I would like to thank all of my friends namely, Dhaval Parikh, Harin Dani, Nibu Koshy, Sudhesh Nair, Kapil Anand, Sunayana Parakh, Neeraj Vyas, Tushar Bhandari, Mitesh Patel, Amit Goswami, Malav Shah, Karthik Krishna, Rahul Jain, and Saif Pathan, for being an inspiration all throughout my thesis work.

Finally, my gratitude to God, my grandparents Kanaiyalal L. Raval, Shrimati Indumati K. Raval, my parents Mukesh K. Raval and Chhaya M. Raval, my brother Sapan M. Raval, would be wife Ashta Goenka, who have encouraged me throughout. With Your Grace, the impossible's been true till date.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

Chapter **Page**

TABLE OF CONTENTS
(Continued)

Chapter **Page**

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES
(Continued)

LIST OF FIGURES
(Continued)

Figure **Page**

CHAPTER 1

INTRODUCTION

1.1 Scope of Research

This study was conducted to develop and implement programs that revealed the presence of long-range power law correlation for a group of healthy subjects while breakdown in the long-range power law correlation for a group of subjects with cardiac heart failure. The programs will be utilized in future research projects as part of DFA and MF-DFA analysis methods to test their efficacy on various pathologic conditions other than CHF (cardiac heart failure) and determine how well the methods are able to discriminate among healthy and pathologic subjects with various other pathologies. The programs implemented were then tested and validated using test data files. The scaling exponent (α) generated by the program is used to characterize the interbeat interval time series for healthy and CHF subjects. Also all generated results fell into normal values for healthy and CHF subjects as found in the literature.

The programs were then used to perform DFA and MF-DFA analysis on a set of data from healthy and CHF subjects. The study was conducted on 27 healthy (men and women, aged 28.5 to 76, mean 52.3) and 11 subjects (men and women, aged 34 to 79, mean 56.5) with congestive heart failure. The data for the normal control group were obtained from 24-hour Holter monitor recordings of 27 healthy subjects with ECG data sampled at 128 Hz. The data for the CHF group were obtained from 24-hour Holter recordings of 11 patients with recordings sampled at 250 Hz. The outcome of the study and the results of DFA analysis on this data set were published [17] and they showed a

significant difference in the value of scaling exponent between both the groups of subjects. The study also showed that there was no effect of activity on the DFA analysis.

The DFA analysis conducted in this study reproduced these results and showed that data from normal interbeat interval series showed a presence of long-range power law correlations while data from CHF subjects showed a breakdown in the long-range power law correlations. Crossover effect was observed for DFA analysis and it showed the presence of two different scaling exponents (slow and fast) for each group of subjects (healthy and CHF). This suggested that a single exponent might not be enough to characterize the interbeat interval time series. Thus more than one exponent values is needed to characterize the interbeat interval time series which can be obtained using multifractal DFA. Thus, MF-DFA analysis was conducted on the same data set of healthy and CHF subjects. This resulted into more than one exponent values and the results revealed the significant difference between healthy and pathologic subjects. The end product of this study will assist if any research involving DFA and MF-DFA analysis shall be done to test, how well they perform for pathologies other than CHF.

Heart rate variability (HRV) is defined as the fluctuation of the heart rate from one beat to another. Non-linear methods related to chaos tend to deal with autonomous systems, i.e. systems where there is no input or where the input has a very simple form. It has been suggested that a healthy heart rhythm is chaotic and shows a fractal form that may be broken down by a disease [8]. The parameters that have been used to measure non-linear properties of HRV includes Phase Plane Plots, Return Maps, Poincare Sections, Lyapunov Exponents, the fractal dimension, 1/f scaling of power spectra, Detrended Fluctuation Analysis, Kolmogorov Entropy, Approximate Entropy, and

Multifractal analysis. These methods have detected abnormal HRV in various cardiovascular conditions such as coronary artery disease with or without previous myocardial infarction

1.2 Goals of the Study

This study was conducted with a goal to

1. Determine the presence of long-range power law correlation for healthy subjects and breakdown in the long-range power law correlation for CHF subjects using DFA and MF-DFA analysis and to reproduce the findings of Peng et al [17]. This was done using programs implemented in MATLAB 7.0. The program for DFA analysis is available on www.physionet.org, and the program for MF-DFA analysis was developed in MATLAB by the author of this thesis.
2. Determine how well DFA and MF-DFA analysis are able to discriminate among healthy subjects and subjects with cardiac heart failure using statistical tools.
3. Test the effect of length, activity, non-stationarity, trends on DFA and the effect of length on MF-DFA.

CHAPTER 2

PHYSIOLOGICAL AND ENGINEERING BACKGROUND

2.1 Cardiovascular System

The cardiovascular system is sometimes called the blood-vascular or simply the circulatory system. It consists of the heart, which is a muscular pumping device, and a closed system of vessels called arteries, veins, and capillaries. As the name implies, blood contained in the circulatory system is pumped by the heart around a closed circle or circuit of vessels as it passes again and again through the various "circulations" of the body.

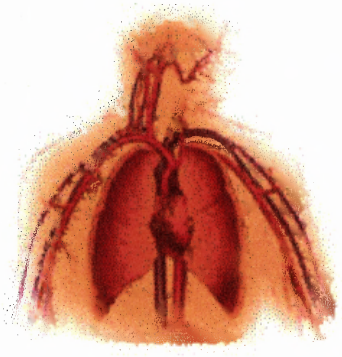

Figure 2.1 Cardiovascular system [1].

The vital role of the cardiovascular system in maintaining homeostasis [2] depends on the continuous and controlled movement of blood through the thousands of miles of capillaries that permeate every tissue and reach every cell in the body. The principle of homeostasis in heart rate fluctuations was developed by Dr. Cannon of Harvard Medical School in the 1920's, and was generally accepted for more than half a century later. This states that physiological systems normally operate to reduce variability and to maintain a

constancy of internal function. According to his theory, any physiological variable including heart rate should return to its 'normal' steady state after it has been perturbed. The principle of homeostasis suggests that variations of the heart rate are merely transient responses to a fluctuating environment [17].

2.1.1 Anatomy of Heart

The main organ of cardiovascular system is the heart. It is responsible for the circulation of blood as required by the body. It is a hollow muscular organ lying in the center of the chest (thorax) and contains four chambers: right atrium, right ventricle, left atrium and left ventricle. Each of the four chambers of the heart is different from the others because of its function. Its beating action maintains the flow of blood throughout the human body. From the heart numerous blood vessels branch out and spread to every corner of the body. These vessels form the vasculature of the cardiovascular system.

There are three main types of blood vessels, arteries which carry oxygenated blood from the heart to the various parts of the body, the veins that carry deoxygenated blood from the body back to the heart, and capillaries which join arteries and veins and are the site for gaseous and chemical exchange between various cells and blood [2].

2.1.2 Human Heart

Figure 2.2 shows the heart and its chambers. The heart is divided into right and left halves and has four chambers: an upper and a lower chamber within each half.

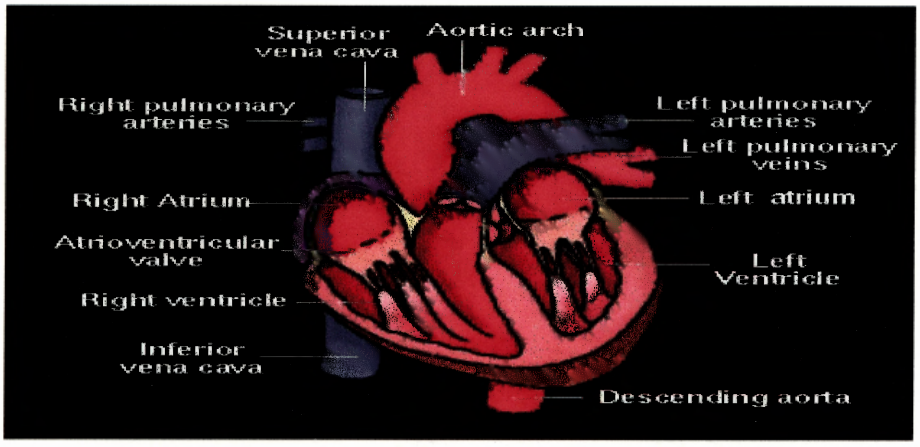

Figure 2.2 Anatomy of the Heart [3].

The upper chamber, atria, receives blood returning to the heart and transfers it to the lower chambers, the ventricles, which pump blood from the heart. The two halves of the heart are separated by a septum. This separation is important because the right half of the heart is receiving and pumping oxygen poor blood while the left side of the heart receives and pumps oxygen rich blood.

2.1.3 Nervous System

Heart rate variability depends on the rate of change of heart. The beating of the heart and its regulation is controlled of by the nervous system. The brain and spinal cord make up the central nervous system. The autonomic nervous system is responsible for the body functions, which are not under conscious control – such as the heartbeat or the digestive

system. The smooth operation of the peripheral nervous system is achieved by dividing it into sympathetic and parasympathetic systems. These are opposing actions and check on each other to provide a balance.

The nervous system is divided into a number of sub-systems. The sub-system responsible for the heart action and the circulatory system is called the autonomic nervous system. It regulates all the vital processes of the body, which are performed without consciousness. The autonomic system is further divided into sympathetic and para-sympathetic nervous system.

2.2 Heart Rate Variability

The measurement of RR interval variability is called heart rate variability (HRV). Generally lower heart rate predicts greater mortality. Wolf [4] was the first to associate the higher risk of post infarction mortality with reduced HRV in 1977. There are numerous methods namely time domain, frequency domain and now non-linear methods that are being used to quantitatively evaluate beat-to-beat cardiovascular control.

2.2.1 Electrophysiology

Excitation of the heart does not proceed directly from the central nervous system but is initiated in the sinoatrial (SA) node, the so-called pacemaker of the heart. The SA node generates an impulse of excitation that spreads across the left and right atrium of the heart and thereby causes the atria to contract. After a short time, it stimulates the atrioventricular (AV) node and thus initiating impulse in the ventricles. The impulse then proceeds down to bundle of His and then continues through special conducting fibers called the Purkinje Fibers on either side of the ventricle thereby causing simultaneous

contraction of the ventricles. Thus the frequency of the SA node mainly controls the heart rate. The following figure shows the electrical activity of the heart

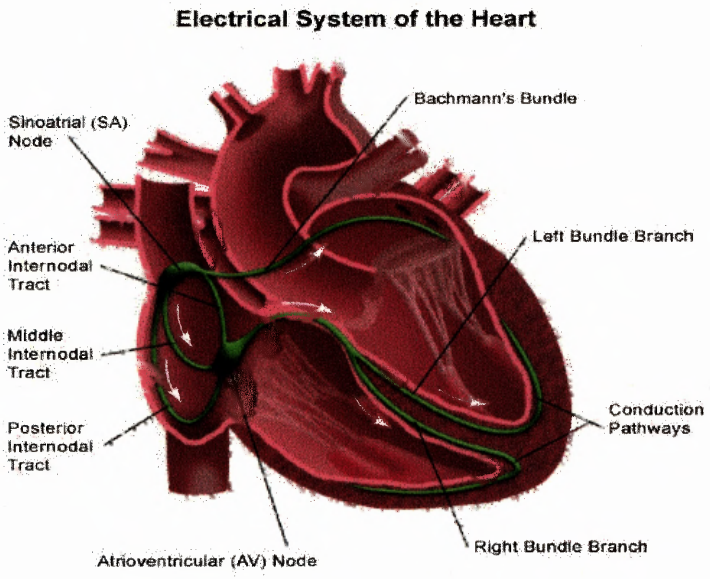

Figure 2.3 The Electrical pathway of impulse through the Heart [6].

In Figure 2.3 an electrical impulse generated by the SA node flows over the right and left atria and thereby causing them to contract. When the electrical signal reaches the AV node it is delayed slightly. The impulse then travels down the bundle of His, which then divides into right bundle branch for the right ventricle and the left bundle branch for the left ventricle. The impulse thus causes the ventricles to contract.

An ECG is a measure of the electrical activity within the heart. The ECG is used clinically to diagnose various diseases and to assess cardiac health.

2.2.2 The Concept of a Time Series

To appreciate the general clinical relevance of dynamics, consider the following common problem. What is the best way to compare a sequence of measurements obtained from two subjects, or from one individual or experimental procedure under different conditions? Conventionally, clinicians and investigators rely primarily on a comparison of means using appropriate statistical tests. However, the limitations of such traditional analyses become apparent when evaluating the data in Figure 2.4, which shows that two signals have the same means but different dynamics. Recording the instantaneous signal from any system over a continuous observation period generates a time series.

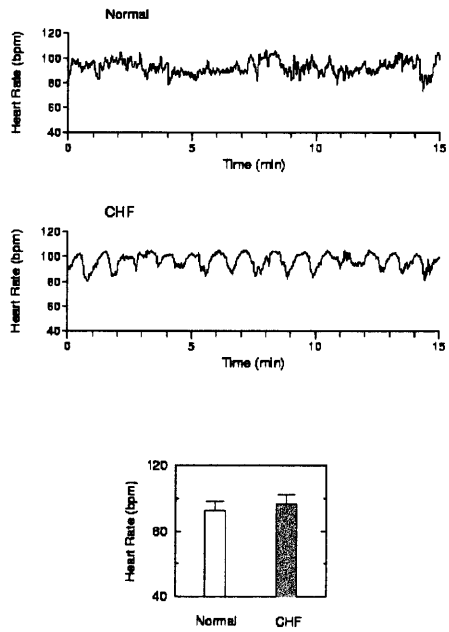

Figure 2.4 Two heart rate time series, one from a healthy subject (top) and the other from a patient with severe congestive heart failure (CHF) (middle) have nearly identical means and variances (bottom), yet very different dynamics. Note the complex, erratic pattern of the data set from the normal subject compared with the slow, periodic oscillations in heart rate with congestive failure.

2.2.3 Linear versus Nonlinear Systems

A well-known example of linear system can be given by Ohm's law: V=IR where the voltage (V) in a circuit will vary linearly with current (I), provided the resistance (R) is held constant. Two central features of linear systems are proportionality and superposition. Proportionality means that the output bears a straight-line relationship to the input. Superposition refers to the fact that the behavior of linear systems composed of multiple components can be fully understood and predicted by dissecting out these components and figuring out their individual input-output relationships. The overall output will simply be a summation of these constituent parts. The components of a linear system summate; there are no surprises or anomalous behaviors.

In contrast, even simple nonlinear systems violate the principles of proportionality and superposition. An example of a deceptively complex nonlinear equation is $y = ax(1-x)$, referred to as the logistic equation [13]. The nonlinearity of this equation, which describes a parabola, arises from the quadratic (x^2) term. Changes in the output as a function of sequential time steps can be readily plotted by a feedback procedure in which the current value of the output becomes the next value of the input, and so on. Iteration of the simple-in-form logistic equation reveals dynamics that are extraordinarily complex; depending on the value of the single parameter, **a**, the same equation can generate steady states, regular oscillations, or highly erratic behavior. Thus, for nonlinear systems, proportionality does not hold: small changes can have dramatic and unanticipated effects [14]. An added complication is that nonlinear systems composed of multiple subunits cannot be understood by analyzing these components individually. The components of a nonlinear network interact, i.e., they are coupled.

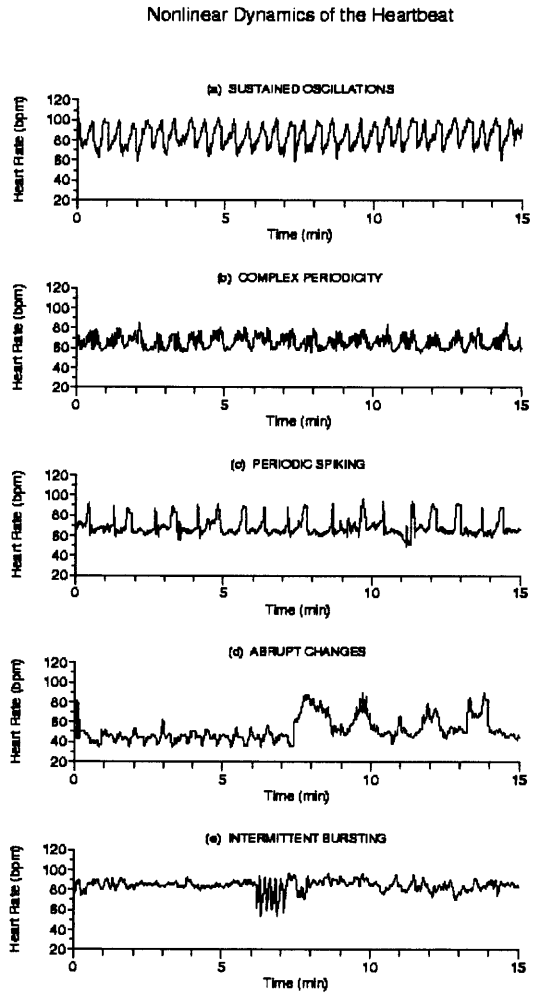

Figure 2.5 Panels (a-c) are subjects with obstructive sleep apnea syndrome, panels (d-e) are from healthy subjects at high altitude (15,000 ft) [14].

Their nonlinear coupling generates behaviors that defy explanation using traditional (linear) models (Figure 2.5). As a result, they may exhibit behavior that is characteristic of nonlinear systems, such as self-sustained, periodic waves (e.g., ventricular tachycardia); Ventricular tachycardia (VT) is a heart rhythm disorder that originates in the ventricles. It is characterized by a rapid heart rhythm during which patients may feel faint or dizzy, or even pass out. During VT, the heart does not pump blood as efficiently as it does during a normal rhythm, and rapid contractions prevent it

from filling adequately between beats, abrupt changes (e.g., sudden onset of a seizure) and, possibly, chaos.

2.3 Engineering Background

2.3.1 Fractal and Chaos

Fractals and chaos bring new sets of ideas and thus new ways of looking at nature. It is true that fractals are really everywhere [9]. The essential characteristic of fractals is that as finer details are revealed at higher magnifications the form of the details is similar to the whole: there is self-similarity which briefly means that the parts that look like the whole. Fractal structures in the human body arise from the slow dynamics of embryonic development and evolution and that this evolutionary advantage accounts for their omnipresence in biomedical phenomena as suggested by Professor Goldberger [9].

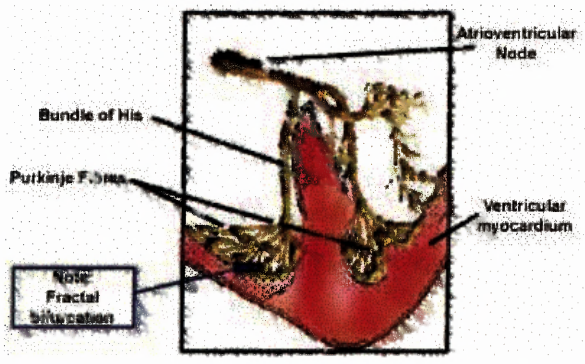

Figure 2.6 Fractals Bifurcation in the His-Purkinje Network [12].

Fractal structures, partly by virtue of their redundancy and irregularity, are robust and resistant to injury. The heart, for example, may continue to pump with little mechanical dysfunction despite extensive damage to the His-Purkinje system, which conducts cardiac electrical impulses. Chaos is best understood by comparing it to two other behaviors

which are randomness and periodicity. Random behavior never repeats itself, and is inherently unpredictable and disorganized except in a very special way e.g. the average behavior of a collection of gas molecules can be predicted with absolute precision, but the individual behavior of a single molecule cannot be predicted. Periodic behavior, on the other hand, is highly predictable because it always repeats itself over some finite time interval. A mathematical sine wave and electrocardiographic normal sinus rhythm are typical examples. Chaos is distinct from periodicity and randomness, but has characteristics of both [15]. Although chaotic behavior looks disorganized (like random behavior), it is really deterministic (like periodic behavior). Chaotic behavior exhibits a number of characteristics that distinguish it from periodic and random behavior namely chaos is more deterministic and aperiodic e.g. If one knows the equations and the initial conditions one can predict the system's behavior accurately and precisely, no matter how complex it appears. However chaotic behavior never repeats itself exactly.

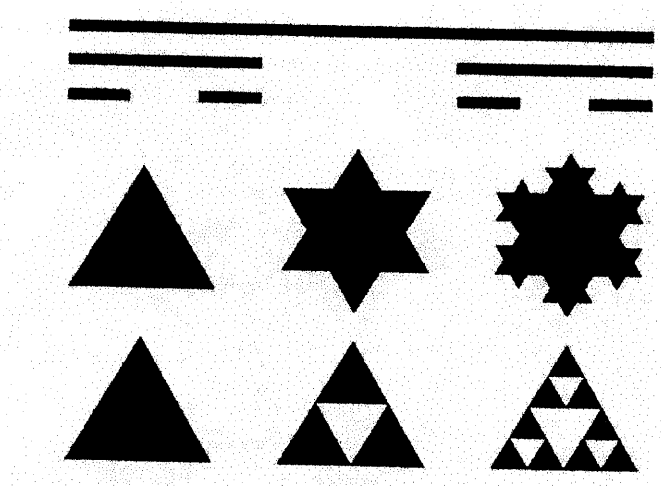

Figure 2.7 Geometrically self-similar objects.

2.3.1.1 Geometric Self-Similarity. Figure 2.7 shows geometrically self-similar objects. In Figure 2.7, the original objects are shown on the left. Then each object is shown after one iteration, and then after two iterations. For example, at the top of Figure 2.7, consider a line segment, and remove the middle third of the line segment, and then repeatedly remove the middle third of each remaining piece. Middle portion of Figure 2.7: The iterative algorithm to generate the Koch curve is to repeatedly add to each edge an equilateral triangle whose sides are one third the length of each edge. Bottom portion of Figure 2.7: The iterative algorithm to generate the Sierpinski triangle is to repeatedly remove triangles that are one quarter the area of each remaining triangle.

2.3.1.2 Statistical Self-Similarity. The pieces of biological objects are rarely exact reduced copies of the whole object [9]. Rather than being geometrically self-similar, they are statistically self-similar i.e. the statistical properties of the pieces are proportional to the statistical properties of the whole. For example, the average rate at which new vessels

branch off from their parent vessels in physiological structure can be the same for large and small vessels. This is illustrated in Figure 2.8 for the arteries in the lung.

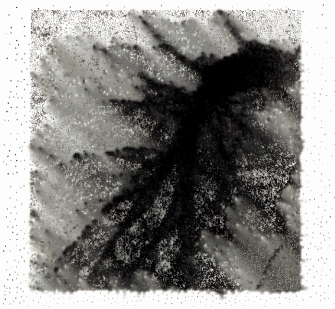

Figure 2.8 Statistical Self-Similarity in space. In the lung arterial tree, the branching patterns are similar for vessels of different sizes.

Measurements recorded from physiological systems over time can also be fractal. As shown in the Figure 2.9, the current recorded through an ion-channel by Gillis, Falke, and Misler[9] show that there are statistically self-similar bursts within bursts of the opening and closing of these ion channels.

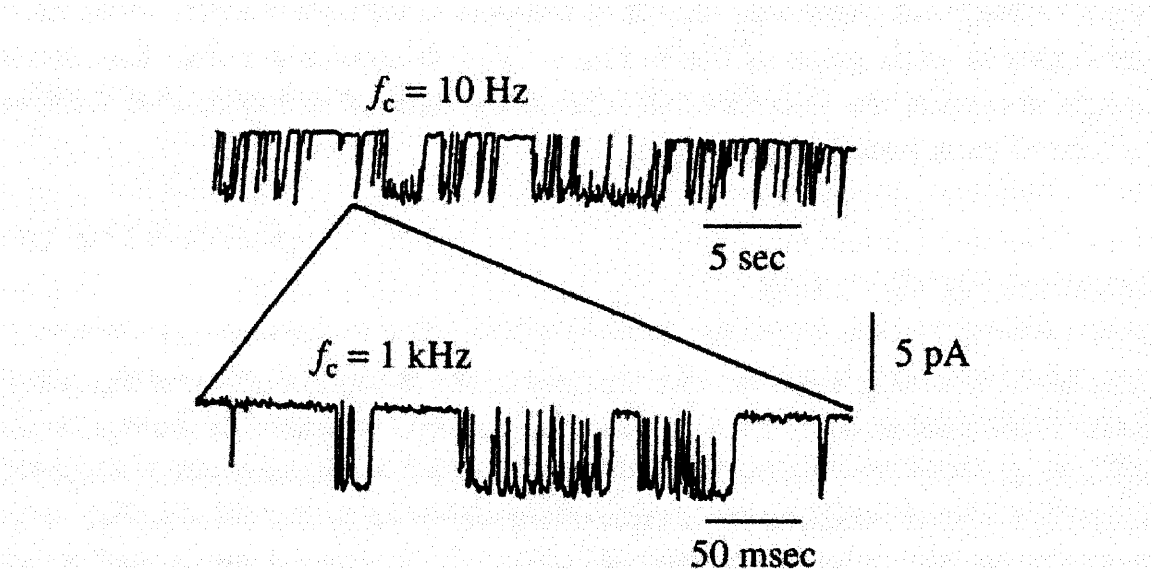

Figure 2.9 Stastical self-similarities in time.

Examples of Self-Similarity

Many physiological objects and processes are statistically self-similar. Some examples include: Systems where the branching pattern is similar at different spatial scales. These can be found in the dendrites in neurons, the airways in the lung, the ducts in the liver, the blood vessels in the circulatory system, and the distribution of flow through them.

2.3.1.3 Mathematical Description of Self-Similarity. Statistical self-similarity means that a property measured on a piece of an object at high resolution is proportional to the same property measured over the entire object at coarser resolution. Hence the value of a property L(r) when it is measured at resolution r, is compared to the value L(ar) when it is measured at a finer resolution ar, where a < 1. Statistical self-similarity means that L(r) is proportional to L(ar), namely,

$$L(ar) = k\,L(r) \qquad (2.1)$$

Where k is a constant of proportionality that may depend on a.

2.3.2 Scaling: The Measure Depends on the Resolution

The value measured for any property of an object depends on the characteristics of the object. When these characteristics depend on the measurement resolution, then the value measured depends on the measurement resolution. There is no one "true" value for a measurement. How the value depends on the measurement resolution is called the scaling relationship. Self-similarity specifies how the characteristics of an object depend on the resolution and hence it determines how the value measured for a property depends on the resolution. Thus, the self-similarity determines the scaling relationship.

2.3.2.1 Self-similarity can lead to a Power Law Scaling. The self-similarity relationship of equation 2.1 implies that there is a scaling relationship that describes how the measured value of a property L(r) depends on the scale r at which it is measured. The simplest scaling relationship determined by self-similarity has the power law form

$$L(r) = Ar^{\alpha} \tag{2.2}$$

Where A and α are constant for any particular fractal object or process.

Taking the logarithms of both sides of equation 2.2 yields

$$\log L(r) = \alpha \log(r) + b, \text{ where } b = \log A. \tag{2.3}$$

Thus, power law scaling is revealed as straight lines when the logarithm of the measurement is plotted against the logarithm of the scale at which it is measured. The

rule for self-similarity is that, there is for some measure, a constant ratio of the measure at scale r compared to that at scale ar:

$$L(r) / L(ar) = k \qquad \text{for } a < 1 \tag{2.4}$$

Suppose that for a Koch curve, there is a power law, such that

$$L(r) = Ar^{\alpha} \tag{2.5}$$

Then by substitution,

$$kL(ar) = kA(ar)^{\alpha} \tag{2.6}$$

and Equation 2.4 can be rewritten as $Ar^{\alpha} = kA(ar)^{\alpha}$, (2.7)

$$k = r^{\alpha}/(ar)^{\alpha} = 1/a^{\alpha} = a^{-\alpha} \tag{2.8}$$

$$L(ar)/L(r) = 1/k = a^{\alpha} \tag{2.9}$$

From Equation 2.6

$$L(ar) = A(ar)^{\alpha}$$

$$= Aa^{\alpha}r^{\alpha} \tag{2.10}$$

Because $Ar^{\alpha} = L(r)$,

$$L(ar) = L(r)a^{\alpha} \tag{2.11}$$

And defining $\alpha = 1\text{-}D$, where D is the fractal dimension,

$$L(ar) = L(r)a^{1-D} \tag{2.12}$$

2.3.3 Fractal Dimension: A Quantitative measure of Self-Similarity and Scaling

The properties of self-similarity and scaling can be assessed in a quantitative way by using the fractal dimension. There are many different definitions of "fractional or fractal dimension," so called because it has noninteger values [9]. When a geometrically self-similar object is examined at finer resolution, additional small replicas of the whole object are resolved.

2.3.3.1 Self-Similarity Dimension. The self-similarity dimension describes how new pieces geometrically similar to the whole object are observed as the resolution is made finer. If the scales is changed by a factor F, and if there are N pieces found that are similar to the original, then the self-similarity dimension $D_{self\text{-}similarity}$ is given by

$$N = F^{Dself\text{-}similarity}. \tag{2.13}$$

By taking logarithm on both the sides,

$$D_{self\text{-}similarity} = \log N/\log F \tag{2.14}$$

Figure 2.10 shows why the self-similarity dimension is called a "dimension."

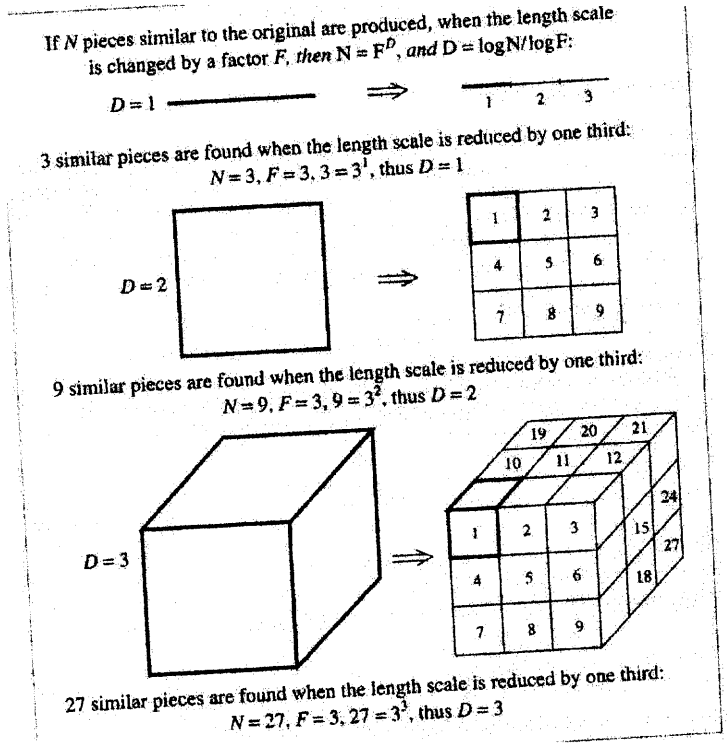

Figure 2.10 The self-similarity dimension tells us how many new pieces similar to the whole object are observed as the resolution is made finer [9]. Consider objects that are r long on each side. The length of a one-dimensional line segment is equal to r. If we reduce the scale by a factor F, then the little line segments formed are each 1/F the length of the original. Hence, F^1 of such pieces are needed to occupy the length of the original line segment, and $D_{self\text{-}similarity} = 1$. The area of a 2-D square is equal to r^2. If we reduce the scale by a factor F, then the little squares formed are each $1/F^2$ the area of the original. Hence, F^2 of such little squares are needed to occupy the area of the original square, and $D_{self\text{-}similarity} = 2$. The volume of a 3-D cube is equal to r^3. If we reduce the scale by a factor F, then the little cubes formed are each $1/F^3$ the volume of the original cube. Hence, F^3 of such pieces are needed to occupy the volume of the original cube, and $D_{self\text{-}similarity} = 3$.

2.3.3.2 Topological Dimension. The dimension introduced above describes the space-filling properties of the fractal set. The topological dimension describes the connectedness between the points in the fractal set. The topological dimension is always an integer. For curves, surfaces, or solids, the topological dimension $D_T = 1, 2,$ or 3 [9].

2.3.3.3 The Definition of a "fractal". Mandelbrot defines a "fractal" as a set of for which

$$D > D_T \tag{2.15}$$

When the fractal dimension is greater then the topological dimension then many new pieces starts appearing as they are looked at finer details. Very loosely, the topological dimension tells about the type of object the fractal is, and the fractional dimension tells how wiggly it is. For example, a line segment that has topological dimension $D_T = 1$ could be so long and wiggly that it nearly fills a two-dimensional area, and thus its fractal dimension $D \sim 2$. Since $2 > 1$, it is a "fractal."

The power-law scaling is a result of self-similarity. The fractal dimension is based on self-similarity as discussed above. Thus, the power-law scaling can be used to determine the fractal dimension. The power-law scaling describes how a property L(r) of the system depends on the scale r at which it is measured equation (2.5), and the fractal dimension describes how the number of pieces of a system depends on the scale r, namely

$$N(r) = Br^{-D} \tag{2.16}$$

where B is a constant.

2.3.3.4 Relationship Between Fractal Dimension D and Scaling Exponent α. Derive the function of dimension f(D), such that the property measured is proportional to $r^{f(D)}$. The experimentally determined scaling of the measured property is proportional to r^{α}. These powers of the scale r are equated, i.e $f(D) = \alpha$, and then solve for dimension D.

For example, assume the total length measured for a line is proportional to r^{α}, where r is the resolution used to make the measurement. Then the length of the line segment is measured at scale r, by breaking it up into pieces each of which has length r. Eq. 2.16 tells us that number of pieces is proportional to $r\, r^{-D}$. The total length measured is number of pieces time the length of each piece. Thus the total length is proportional to r^{-D} multiplied by r, i.e r^{1-D}. Since length is proportional to r^{α}, $1 - D = \alpha$, and so $D = 1 - \alpha$ for lengths and $D = 2 - \alpha$ for areas.

2.3.3.5 $1/f^{\beta}$ Power Spectra: A Characteristic of Self-Similarity. Masonori et al. investigated the statistical behaviors of heartbeat period fluctuation and estimated its power spectral density. They computed the power spectral densities of fluctuations in the heartbeat period of 15 subjects and found that the power spectral densities are inversely proportional to frequency [10]. This type of power spectral density is often called the "1/f spectrum" and such fluctuations are called the "1/f fluctuations." The 1/f fluctuations were first found in 1925 [11] in an electric current passing in a vacuum tube, and after this a variety of phenomena namely, fluctuation of cell membrane potential, frequency fluctuation of alpha brain wave, highway traffic current fluctuations, and so forth are known to be having the 1/f fluctuations.

Examples of such $1/f^{\beta}$ power spectra is shown in the Figure 2.11

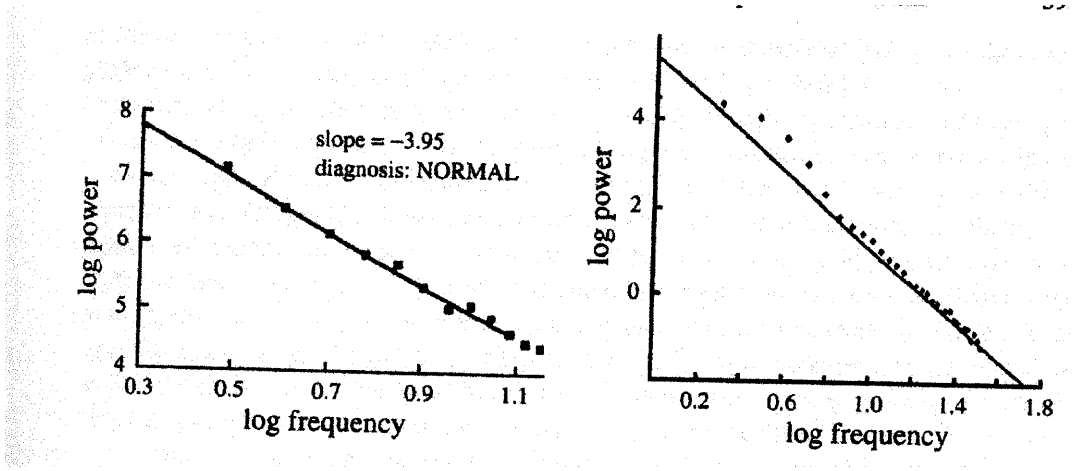

Figure 2.11 The power spectrum describes the amplitude of the fluctuations at different spatial or temporal frequencies. It has an inverse power law form for fractals in space and time. Left: Power spectrum of the fluctuations in the spatial distribution of radioisotope in the liver measured by Cargill et al. [9]. Right: Power spectrum of the fluctuations in time of the ventricular depolarization in an electrocardiogram as measured by Goldberger et al [9].

Consider for example any time series say, 2 hours of data. If we take the Fourier transform of this data this will give us the Fourier coefficients, each characterizing the different frequency present in the data and square of these coefficients will give us the power spectrum for the same data. If the plot of log of power spectrum ($S_i(f)$) v/s log of frequency (f) gives a straight line.

$$S_i(f) \sim f^{\beta} \tag{2.17}$$

Thus, different frequencies are organized in such a way that all frequencies follow the same rule. If the scale of observation is changed and the process remains the same, there is scale invariance in the process. Scale invariance means no characteristic time scale and hence exhibits a power-law relation as discussed in the previous sections.

Fractals have invariance and when people find such 1/f behavior, they start thinking for some fractal underlying this mechanism. "1/f" is a general notation to indicate the process that doesn't have a characteristic frequency. This β is related to the exponent α as $\beta = 1 - 2\alpha$. Furthermore, β can be used as an indicator of the presence and type of correlations and thus knowing the value of α using DFA one can determine the value of β and vice versa. (i) If $\beta = 0$, there is no correlation in the time series i.e. ("white noise"). White noise means there is no correlation in the time series, there is no dependence of a second point on the previous point in the give time series. (ii) If $-1 < \beta < 0$, then the time series is correlated such that positive values of the time series are likely to be close to each other, and the same is true for negative values. (iii) If $0 < \beta < 1$, then the series is correlated; however, the values of the time series are organized such that positive and negative values are more likely to alternate ("anticorrelation"). The different value of the slope β and hence α indicates the type of correlations present in the signal.

2.4 Different Measures of Heart Rate Variability (HRV) Analysis

2.4.1 Time Domain Measures of HRV Analysis

The easiest and oldest measure would be evaluation of heart rate with respect to its variation in time. This is called the time domain analysis of HRV. In time domain methods the instantaneous heart rate or the inter-beat intervals are determined and subsequently its variation over time is studied. Commonly used parameters in the time domain include 1) Mean RR interval and 2) Mean Heart Rate. Other parameters have been studied [5] and are divided as follows

The below mentioned techniques represent direct measures of RR interval:

a) SDNN: Standard Deviation of all RR intervals

b) rMSSD: the square root of the mean of the sum of the squares of differences between adjacent RR intervals

Measures derived from difference in RR intervals

a) SDNN index: Mean of the standard deviation of difference of all NN intervals for all 5-minute segments of the entire recording.

b) SDANN: Standard deviation of the average of RR intervals in all 5-min segments of the entire recording.

These methods allow the comparison of HRV during various activities like paced breathing, tilting, rest, sleep and so on. Generally the total variance of HRV increases with the length of analyzed recording [7].

2.4.2 Frequency Domain Measures of Heart Rate Variability (HRV) Analysis

Various frequency domain methods are also useful for analysis of cardiac signals. One such method is to construct the interbeat interval (IBI) signal and power spectral calculations. Variations in the frequency components of the interbeat interval (IBI) signals can be determined using these domain measures. Power spectral density analysis provides the basic information of how power distributes as a function of frequency. Normally spectral components are derived from either 5 minute or 24 hour recordings. The main advantage of spectral analysis of signals is that it allows the study of frequency-specific oscillation. Results are displayed in a graph with magnitude of variability as a function of frequency. Frequency domain measurements indicate the autonomic nervous system. The Table 2.1 below gives a comparison between time domain and frequency domain methods

Table 2.1 Comparison Between Time Domain and Frequency Domain Measurement Methods (Andrew JE Seely and Peter T Macklem Complex systems and the technology of variability analysis, Critical Care 2004, 8: R367-R384)

Variability analysis	Description	Advantages	Disadvantages	Output variables
Time Domain	Statistical calculations of consecutive intervals	Simple, easy to calculate; proven clinically useful; gross distinction of high and low frequency variations	Sensitive to artifact; requires stationarity; fails to discriminate distinct signals	SD, RMSDD Specific to HRV: SDANN
Frequency domain	Frequency distribution (plot number of observations falling in selected ranges or bins)	Visual representation of data; can fit to normal or log-normal distribution	Lacks widespread clinical application; arbitrary number of bins	Skewness (measures symmetry): positive (right tail) versus negative (left) Kurtosis (measures peakedness): flatter top (<0) versus peaked (>0)
	Frequency spectrum representation (spectral analysis)	Visual and quantitative representation of frequency contribution to waveform; useful to evaluate relationship to mechanisms; widespread HRV evaluation	Requires stationarity and periodicity for validity; sensitive to artifact; altered by posture, sleep, activity	Total power (area under curve) Specific to HRV: ULF (<0.003 Hz), VLF (0.003–0.04 Hz), LF (0.04–0.15 Hz), HF (0.15–0.4 Hz) Time spectrum analysis

2.5 Analytic Techniques to detect Non-linear Dynamic Behavior

Non-linear methods related to chaos tend to deal with autonomous systems, i.e. systems where there is no input or where the input has a very simple form. It has been suggested that a healthy heart rhythm is chaotic and shows a fractal form that may be broken down by a disease [8]. The parameters that have been used to measure non-linear properties of HRV includes, phase plane plots, return maps, Poincare sections, Lyapunov exponents, the fractal dimension, 1/f scaling of power spectra, detrended fluctuation analysis, Kolmogorov entropy, approximate entropy, and multifractal analysis. These methods have detected abnormal HRV in various cardiovascular conditions such as coronary artery disease with or without previous myocardial infarction.

2.5.1 Discrete Data

Many nonlinear systems (such as the logistic map) pass through a series of intermediate stages prior to the chaotic behavior. Often, these stages are easily recognized as oscillations between two, four, eight or more states. If a biological system is observed to behave in this manner, then the underlying rules for the behavior might be based on a nonlinear system. The next step in the analysis is the construction of a mathematical model that could reproduce the behaviors noted. This is the most difficult aspect of the problem. After developing a general theoretical model, appropriate parameters must be selected that reproduce the behavior. The final step in the process is proving that the proposed mathematical model accounts for most, if not all, of the behavior described in the biological model.

2.5.2 Continuous Data

The modeling process described above is a general framework that can be applied to many biologic behaviors; however continuous signals provide a more difficult problem. Often subtle changes in behaviors are not visible in the time series; sometimes the time series appears random. In these cases, more sophisticated techniques are needed to detect the presence of underlying structure in the behavior.

2.5.3 Phase Plane Plots

The phase plane plot is a representation of the behavior of a dynamic system in state space (the abstract mathematical area in which a behavior occurs). It typically takes the form of a graph of a position of the signal (its amplitude) on the x axis, versus the velocity of the signal (its first derivative) on the y axis. Each cycle called a trajectory or orbit, represents the behavior of the system over a period of time [15].

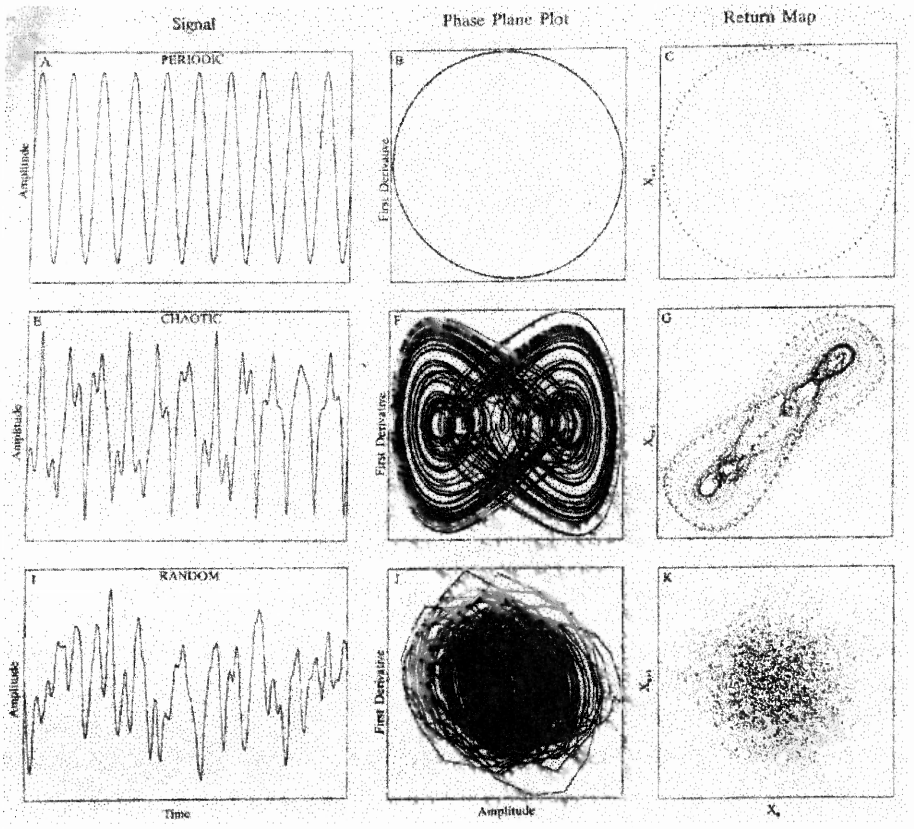

Figure 2.12 illustrates the phase plane plot of a typical chaotic system [15].

The two dimensional phase plots is the most common representation of state space, but many others are possible. Phase plane plots of periodic signals have trajectories that overlap each other precisely [Figure 2.12B], while those of random signals exhibits no definite pattern [Figure 2.12J]. In contrast although phase plane plots of chaotic signals do not have periodic trajectories, they do exhibit a definite pattern [Figure 2.12F]. A major disadvantage of the phase plane plots is its sensitivity to noise. As little as 1% noise can severely disrupt the structure of a plot. Thus the recording system and hence the data must be as noise-free as possible.

2.5.4 Return Maps

The return map is similar to the phase plane plot, but the analyzed data must be discrete (digital). Typically, the return map represents the relation between a given point in a time series plotted on the x axis, and the next point in the time series plotted on the y axis (a next amplitude plot). The temporal difference between the two points is called lag. The lag acts to smooth away some noise in the system, making return maps less sensitive to the noise compared to phase plane plots.

2.5.5 Poincare Sections

If a phase plot does not have clearly discernible pattern, this ancillary graphical technique can help reveal one. A two-dimensional phase plot is cut by a line perpendicular to the trajectories [Figure 2.13A], points on that line represents where each trajectory crossed the line [Figure 2.13B]. A graph is constructed from these points, representing the relation between adjacent trajectories [Figure 2.13C]. Sometimes this graph reveals structure that is not apparent in the phase plot itself. This method is also sensitive to noise.

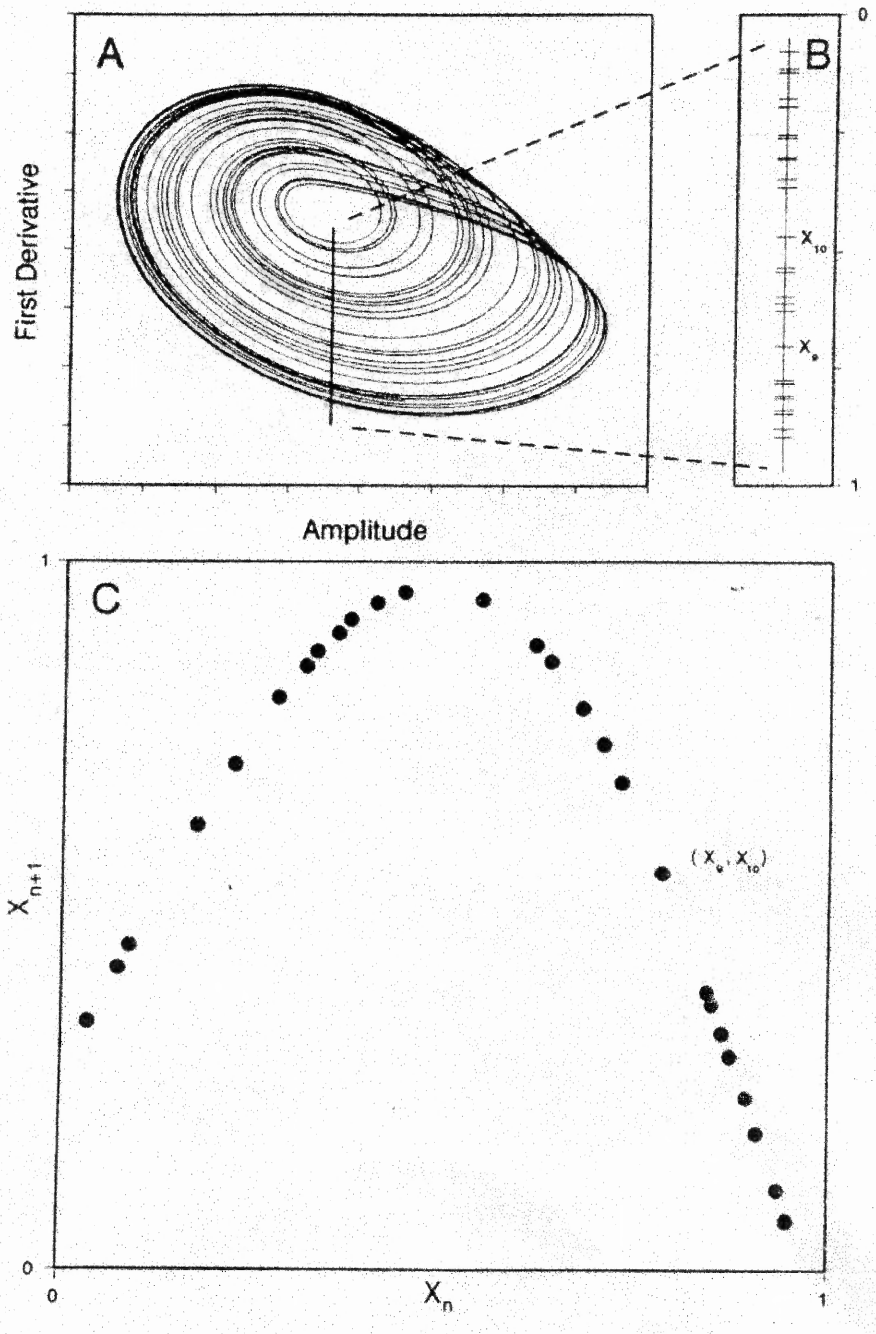

Figure 2.13 Poincare section [15]. Panel A is a phase plane plot of a typical chaotic system. The vertical line intersects a selected portion of the phase plane plot. The distribution of the points along an expansion of this line (panel B) reveals no apparent structure, but a return map constructed from these points shows a nonlinear pattern suggestive of a parabola (panel C). In panel B, X9 and X10 are the ninth and tenth trajectories-they are not adjacent. The relation between X9 and X10 is shown in graph C.

2.5.6 Lyapunov Exponents

As noted earlier, chaotic systems characteristically exhibit sensitive dependence on initial conditions. In state space, sensitive dependence manifests itself graphically as adjacent trajectories that diverge widely from their initial close positions. The Lyapunov exponent is a quantitative measure of this rate of separation [15]. The magnitude of this exponent is related to how chaotic system is; the larger the exponent, the more chaotic the system. A random signal will have an exponent of zero, because over a long period of time adjacent trajectories will converge and diverge equally. A positive Lyapunov exponent, on the other hand, indicates sensitive dependence on initial conditions and is almost without exception diagnostic of chaos. The major limitation in their calculation is that the algorithms to implement them require large amounts of data (on the order of 1,000 to 10,000 cycles). As a result the computing time itself can be limiting. However, having sensitive dependence on the initial conditions as requirement in most of the methods described above, the following section gives an introduction to a new concept called Detrended Fluctuation Analysis, which does takes into account the non-stationarity present in the data and has proven one of the most useful non-linear dynamic tools in characterizing HRV.

CHAPTER 3

DETRENDED FLUCTUATION ANALYSIS

3.1 Theory

The detrended fluctuation analysis technique is a measurement, which quantifies the presence or absence of fractal correlation properties and was introduced by Peng and his coworkers [17]. The healthy heartbeat is traditionally thought to be regulated according to the classical principle of homeostasis whereby physiologic systems operate to reduce variability and achieve an equilibrium-like state. However recent studies reveal that under normal conditions, beat-to-beat fluctuations in heart rate display the kind of long-range correlations typically exhibited by dynamical systems far from equilibrium [17].

If the RR interval time series is scale invariant, it exhibits a power-law relationship. Consider for example a time series; say 2 hours of data around 8000 points, within this boundary, if the time series exhibits a power-law it means that the series has long-range correlation and long-range refers to about 8000 points as quoted by Peng et al., which is the minimum number of data points needed for DFA analysis. The long-range correlations serve as an organizing principle for highly complex, nonlinear processes that generate fluctuations on a wide range of time scales [29]. Long-range power-law correlation means a large interval is more likely to be followed by a large interval and vice versa. A scaling exponent called α quantifies this long-range correlation in the time series. The value of α was determined by implementing DFA on two long interbeat interval time series of 24 hours which is available on www.physionet.org, and it revealed that value of α.varies between 0 – 1.5 [17]. For $0.5 < \alpha. < 1$, it refers to

long-range power-law correlation and for 0< α<0.5 refers to short-range correlation i.e. a large interval is more likely to be followed by a small interval and vice versa.

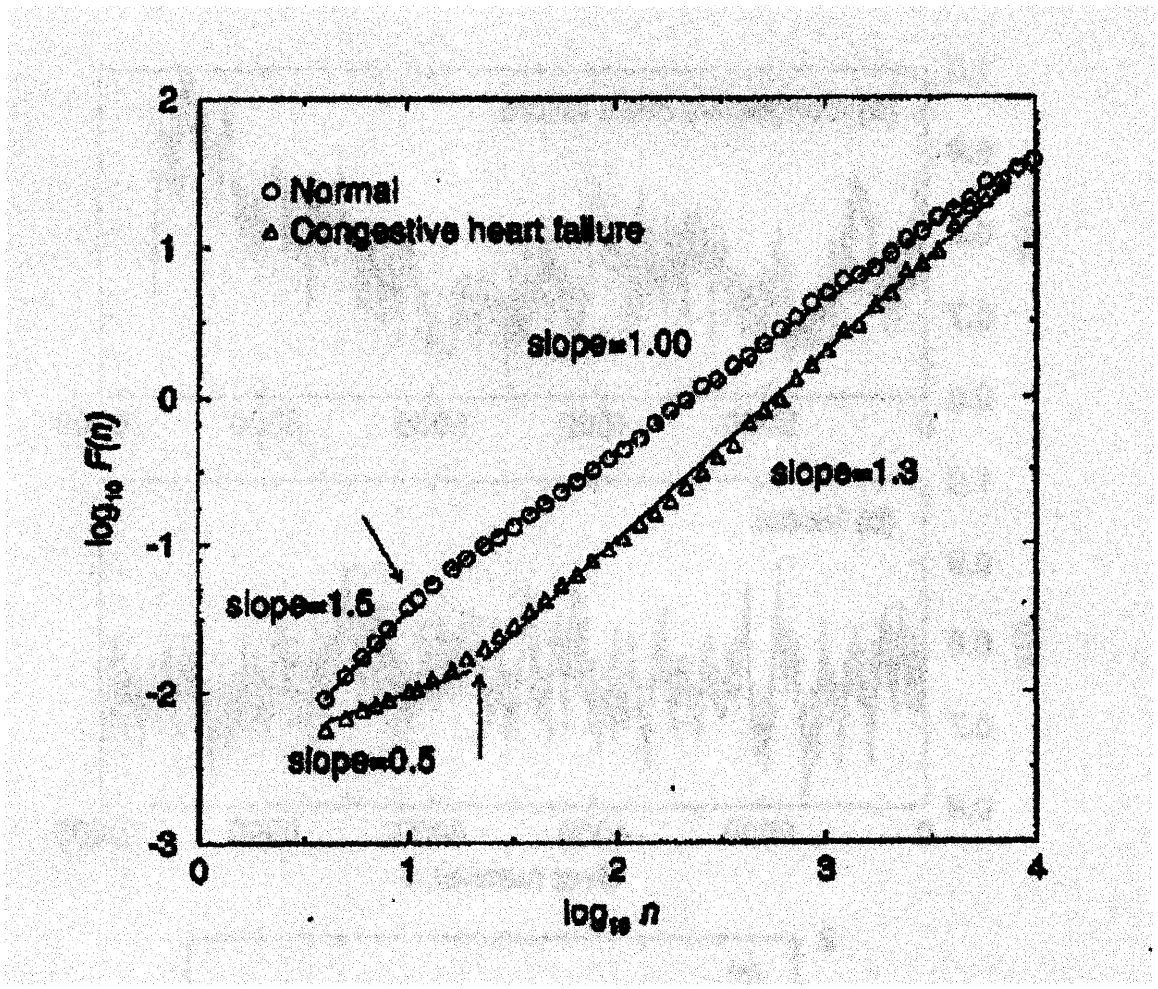

Figure 3.1 Plot of log F(n) v/s. log n for two very long interbeat interval time series (~24 hours). The circles are from a healthy subject and triangles are from a subject with congestive heart failure. Arrows indicate "crossover" phenomenon [17].

Figure (3.1) reveals the presence of two different α exponents: α1 for short range correlation and α2 for long range correlation for healthy subjects and in contrast, the heart rate time series from patients with severe congestive heart failure reveals the breakdown of this long - range correlation behavior which is characterized by the value of α1 and α2 and they are different from the one for healthy subjects. DFA thereby helps to determine the correlation in the time series and also indicates how the standard deviation of the time

series grows with the time scale i.e. how far the value deviates from the normal value which according to Peng et al. is $\alpha. = 1 \pm 0.11$ for healthy subjects and $\alpha. = 1.24 \pm 0.22$ for pathologic subjects.

Conventional methods like time domain measures and frequency domain requires that the data needed for analysis have to be stationary and DFA was introduced to address this issue. The advantage of DFA over conventional methods is that, it permits the detection of long-range correlations embedded in nonstationary time series and avoids spurious detection of apparent long-range correlations that are artifacts of nonstationarity behavior [17, 18]. Variations that arise because of extrinsic stimuli are presumed to cause a local effect, whereas variations due to the intrinsic dynamics of the system are presumed to exhibit long-range correlation. Therefore, any invariant scaling characteristics in the heart rate fluctuations obtained by these means can mostly be attributed to the intrinsic mechanism of neuroautonomic control [19].

The α exponent of the DFA calculation involves the subtraction of local trends (more likely related to the external stimuli) in order to address the correlations that are caused by nonstationarity. The DFA has also been applied to detect long-range correlations in other time series like heterogeneous DNA sequences [18, 20] and the stride interval fluctuations obtained from unconstrained human gait dynamics [21].

3.2 Algorithm

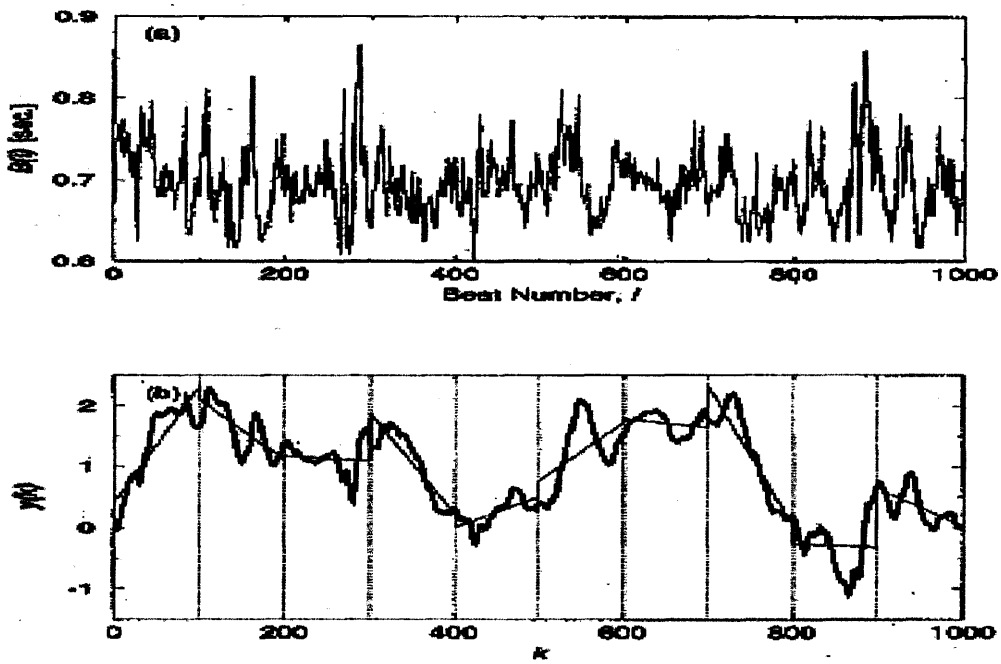

Figure 3.2 is used an example to illustrate the DFA algorithm [17].

1) The analysis is performed on a time series, for example, the interval between consecutive heart beats, with the total number of beats equal to N. First the interbeat interval time series of total length N is integrated $y(k) = \sum [B(i) - Bave]$, where B(i) is the ith beat interval and Bave is the average interbeat interval. This new series of values represent an evaluation of trends; for example, if the difference between individual RR intervals and the average RR intervals remains positive i.e. the interval between heartbeats is longer than the average interbeat interval, then the heart beat is slower than the mean, and the integrated series will increase. This trend series of data displays fractal, or scaling behavior, and the following calculation is further done to quantify this behavior.
2) The integrated time series is divided into boxes of equal length n as shown in [Figure 3.2(b)]

3) In each box of length n, a least-squares line is fit to the data representing the trend in that box [Figure 3.2(b)]. The y coordinate of each straight line segments is represented by $y_n(k)$.

4) Next the integrated time series y(k) is detrended by subtracting the local trend $y_n(k)$ in each box.

5) The root-mean square fluctuation of this integrated and detrended time series is calculated by

$$F(n) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}[y(k) - y_n(k)]^2} \tag{3.1}$$

This computation is repeated over all box sizes to provide a relationship between F(n), the average fluctuation function as a function of box size, and the box size n. Typically F(n) will increase with box size n [17]. Finally it is possible to graph the relationship between F(n) and n. Scaling or fractal correlations are present if the data is linear on a double log graph of F(n) v/s n. The slope of the graph has been termed as α, a scaling exponent, defined as $F(n) \sim n^{\alpha}$.

3.3 Interpretation

Consider for example a time series in which there is no dependence of the second point on the first point, which can be achieved either by shuffling the data points, so called surrogate data or by generating random numbers. Surrogate data refers to artificial data with known correlation values and is generated manually with an aim to test the model that we would like to implement i.e. we would know before hand what we might be expecting the method to generate, in this case DFA. Therefore, if the results generated

returns the same value that we want it to generate we can say that the method works well. Then we can implement the method on the original heart rate time series and compare the results of that with the one generated running the surrogate data set and thereby validate the use of that particular method for future use by other people.

For this type of uncorrelated data, the integrated value, y(k), corresponds to white noise, and therefore $\alpha = 0.5$. The value of α greater than 0.5 and less than or equal to 1 indicates long-range power law correlations such that a large interbeat interval (compared to the mean value) is more likely to be followed by a large interval and vice versa [17]. In contrast $0<\alpha<0.5$ exhibits a different kind of power-law correlation such that a large interval is more likely to be followed by a small interval and vice versa, so called anti-correlations. The value of $\alpha=1$ follows 1/f noise [16] and $\alpha=1.5$ indicates brown noise, which is integration of white noise [17]. α exponent can also be viewed as an indicator that describes the "roughness" of the time series: the larger the value of α, the smoother is the time series, reflecting a more periodic behavior commonly seen in pathologic condition.

There are many ways to characterize different noise sources one is to consider the spectral density, the mean square fluctuation at any particular frequency and how that varies with frequency. This would generate a noise whose spectral density varies as powers of inverse frequency, more precisely, the power spectra P(f) is proportional to $1 / f^{beta}$ for beta >= 0. When beta is 0 the noise is referred to as white noise, when it is -2, it is referred to as Brownian noise, and when it is -1, it normally referred to simply as 1/f noise which occurs very often in processes found in nature and since $\beta = 1-2\ \alpha$, the value

of α can be determined for each value of β. Thus β=0, i.e. α=0.5 refers to white noise. Figure (3.3) and Figure (3.4) shows the white noise and Brownian noise respectively.

White noise, beta = 0

Figure 3.3 White noise [30].

Brownian noise, beta = - 2

Figure 3.4 Brownian noise [30].

For Brownian noise the change, or increment, from one moment to the next is random and normally distributed. Thus, Brownian noise is an integration of white noise.

Figure 3.5 shows sample of Brownian noise, its power spectrum, and plots of power and amplitude against frequency.

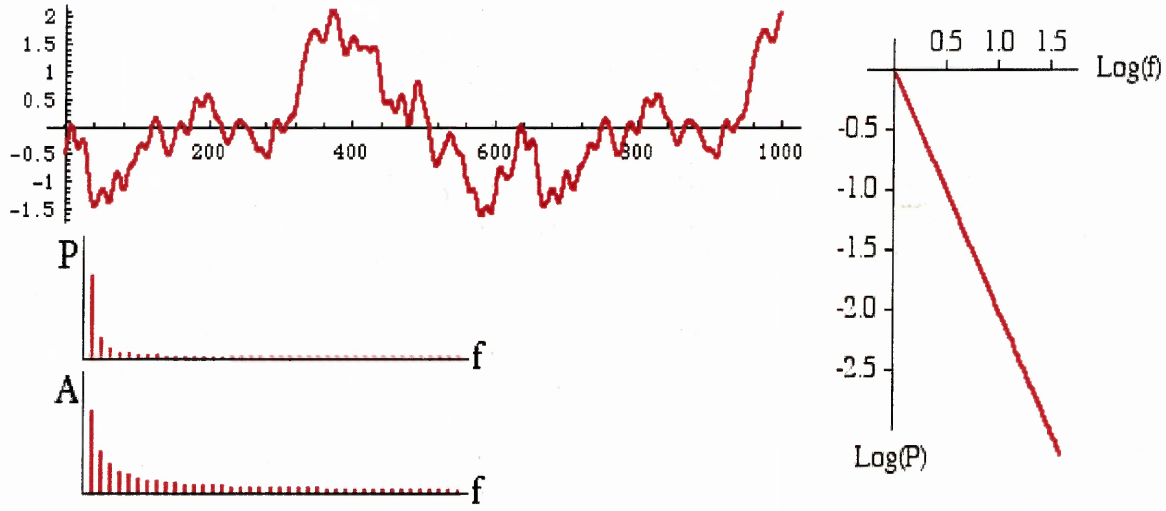

Figure 3.5 Brownian noise, its power spectrum and plots of power and amplitude against frequency [31].

If one plots the log(power) vs. log(frequency) for noise with a particular beta value, the slope gives the value of beta. This leads to the most obvious way to generate a noise signal with a particular beta. Theoretically, the scaling exponent will vary from 0.5 (white noise) to 1.5 (brown noise), but physiological signals yield scaling exponents close to 1.

3.4 Advantages and Disadvantages

DFA has advantages over other conventional fractal methods because it permits the detection of long-range correlations embedded in a raw non-stationary time series.

Importantly, it avoids the spurious detection of long-range correlations resulting from non-stationary conditions. Thus, the statistical invariant scaling characteristics in HRV obtained by these means can be mostly attributed to the intrinsic mechanism of neuroautonomic control [24].

Therefore, there is no need to rigorously control physical activity or provide an external stimulus and, this technique promotes the possibility of analyzing ambulatory

long-term recordings [25]. Although this exponent may serve as a useful indicator for selected diagnostic purposes, a major drawback is that the data requirements are greater than with other techniques and have been suggested to include at least 2 hours of data i.e. 8000 data points [17].

For practical purposes, clinical investigators are usually interested in the possibility of using substantially shorter time series. In this regard, Peng et al. noted that for short time scales, there was an apparent crossover exhibited for the scaling behavior of both data sets i.e. for healthy and pathology subjects as revealed in Figure (3.1). For the healthy subject, the α.exponent estimated from very small n (<10 beats) is larger than calculated from large n (>10 beats). This is probably due to the fact that on very short time scales, the physiologic interbeat interval fluctuation is dominated by the relatively smooth heartbeat oscillation associated with respiration, thus giving rise to a larger α. value. For longer scales, the interbeat fluctuation reflects the intrinsic dynamics of the system, approaching the standard 1/f behavior as previously noted. In contrast, the pathologic data set shows a very different crossover pattern Figure (3.1). This apparent crossover pattern motivated Peng et al. [17] to extract two different parameters from each data set over two different time scales: one short, and the other long and for each data set, α1 was calculated by making a least square fit of log F(n) v/s log (n) for $4 \leq n \leq 16$ and similarly an exponent α2 was calculated from $16 \leq n \leq 64$.

However appealing in order to simplify clinical comparison, the calculation of two scaling exponents [α1 for small ($4 \leq n \leq 16$) and α2 for large ($16 \leq n \leq 64$)] represents a somewhat arbitrary manipulation. Here α1 indicates short-range correlation i.e. $n \leq 16$ and α2 indicates long-range correlation i.e. $n \geq 16$. The value of α1 is greater than α2, because

on very short time scales (few beats to 10 beats), a physiological interbeat interval fluctuation function is dominated by relatively smooth heart beat oscillations associated with respiration, thus giving rise to a larger α value [17].

The assumption that the same scaling pattern is present throughout the signal remains unclear, and therefore techniques without this assumption are being developed and are referred to as multifractal DFA.

3.5 Multifractal Detrended Fluctuation Analysis

3.5.1 Introduction

Consider a 'population' consisting of 'members' distributed over a volume of linear size L. The population could, in fact, be the human population distributed over the surface of the earth. The population could also be considered to be meterological observation posts, which are unevenly distributed over the globe. Many variables fluctuate widely in space. Gold, for instance, is found in high concentration at many places, and in very low concentrations almost everywhere. The point is that this description holds whatever the linear scale is – to be global, on the scale of meters, or on the microscopic scale. Multifractal measures are related to the study of a distribution of physical or other quantities on a geometric support. The support may be an ordinary plane, the surface of a sphere or a volume, or it could itself be a fractal.

The concept underlying the recent development of what are now called multifractals were originally introduced by Mandelbrot in the discussion of turbulence and expanded by Mandelbrot to many other contexts. The idea that a fractal measure may be represented in terms of intertwined fractal subsets having different scaling exponents opens a new realm for the applications of fractal geometry to physical systems [26].

3.5.2 Theory

Monofractal signals are homogeneous in the sense that they have the same scaling properties, characterized locally by a single exponent α, throughout the entire signal. Therefore, monofractal signals can be indexed by a single global exponent which suggests that they are stationary from the viewpoint of their local scaling properties. Many records do not exhibit a simple monofractal behavior, which can be accounted for by a single exponent [26]. In some cases, there exist crossovers separating regimes with different scaling exponents. In other cases, the scaling behavior is more complicated, and different scaling exponents are required for different parts of the series. This occurs, e.g., when the scaling behavior in the first half of the series differs from the scaling behavior in the second half.

Multifractality occurs due to different long-range (time-) correlations of the small and large fluctuations [26]. Multifractal signals can be decomposed into many subsets—possibly infinitely many—characterized by different α exponents, which quantify the local behavior and thus relate to the local scaling of the time series. Thus multifractal signals require many exponents to fully characterize their scaling properties, and are intrinsically more complex, and inhomogeneous, than monofractals [32]. The statistical properties of the different subsets characterized by these different exponents α can be quantified by the function $F_q(s)$, where $F_q(s)$ is the fluctuation function of the subset of the time series characterized by the local exponents α. Thus, the multifractal approach for signals has the potential to describe a wide class of signals that are more complex then those characterized by a single fractal dimension such as classical $1/f$ noise.

3.5.3 Algorithm

The generalized multifractal DFA method consists of following 5 steps

1) Determine the profile

$$Y(i) = \sum_{k=1}^{i} [x_k - \langle x \rangle] \quad i=1,\ldots,N \tag{3.2}$$

Subtraction of the mean <x> is not compulsory, because it would be later eliminated by detrending in the third step

2) Divide the profile Y(i) into Ns = int(N/s) non-overlapping segments of equal length s. Since the length N of the series is often not a multiple of the considered time scale s, a short part at the end of the profile may remain. In order to disregard this part of the series, the same procedure is repeated starting from the opposite end. Thereby, 2Ns segments are obtained together.

3) Calculate the local trend for each of the 2Ns segments by a least-square fit of the series. Then determine the variance

$$F^2(v, s) = \sum_{i=1}^{s} \{Y(v-1)s + i] - y_v(i)\}^2 \tag{3.3}$$

for each segment v, v = 1,2,3....., Ns and

$$F^2(v, s) = \sum_{i=1}^{s} \{Y[N - (v-N_s)s + i] - y_v(i)\}^2 \tag{3.4}$$

For v = Ns + 1, ... 2Ns. Here $y_v(i)$ is the fitting polynomial in segment v. Linear, quadratic, cubic, or higher order polynomials can be used in the fitting procedure. Since the detrending of the time series is done by the subtraction of the polynomial fits from the profile, different order DFA differ in their capability of eliminating trends in the

series. In MF-DFA m [m^{th} order DFA] trends of order m in the profile are eliminated. Thus a comparison of results for different orders of DFA allows one to estimate the type of the polynomial trend in the time series.

4) Average over all segments to obtain the qth order fluctuation function

$$F_q(s) = \left\{ 1/2*N_s \sum_{v=1}^{2*Ns} [F^2(v, s)]^{q/2} \right\}^{1/q} \quad (3.5)$$

where in general the index variable q can take any real value. For q =2 the standard DFA procedure is retrieved. Our interest is to see how the generalized q dependent function Fq(s) depends on the time scale s for different values of q.

5) Determine the scaling behavior of the fluctuation functions by analyzing log-log plots Fq(s) versus s for each value of q. If the series xi are long-range power-law correlated, Fq(s) increases, for large value of s, as a power law

$$F_q(s) \sim s^{h(q)} \quad (3.6)$$

3.5.4 Interpretation

For a monofractal time series with compact support, h(q) (slow) is independent of q, since the scaling behavior of the variance $F^2(v,s)$ is identical for all segments v, and the averaging procedure in equation (3.5) will give just this identical scaling behavior for all values of q [26]. A function has compact support if it is zero outside of a compact set. Let S be a subset of a metric space. Then the set S is compact if, from any sequence of elements X1, X2 , ...of S, a subsequence can always be extracted which tends to some limit element X of S. Compact sets are closed and bounded, and these conditions

characterize them in finite-dimensional space. A function with compact support is only interesting in a bounded domain.

Only if small and large fluctuations scale differently there might be a dependence of h(q) on q. i.e. It would be possible to distinguish whether the given series is monofractal or multifractal because the value of h(q) remains the same despite any change in the value of q. However, if we get different value of h(q) for varying q the given series is multifractal.

For the positive values of q, segments with larger value of $F^2(v, s)$ will dominate the average value of $F_q(s)$ and thus h(q) describes the scaling behavior of the segments with large fluctuations. On the contrary, for the negative values of q, the segments v with small variance $F^2(v, s)$ will dominate the average $F_q(s)$.

For the maximum scale s=N the fluctuation function is independent of q, since the sum in equation (3.5) runs over two identical segments. For small segments $s<<N$, the sum will run over several segments, and the average value $F_q(s)$ will be dominated by $F^2(v, s)$ from the segments with small or large fluctuation if $q<0$ or $q>0$.

Hence, if we follow equation (3.6), the slope h(q) for log plot of $F_q(s)$ v/s log plot of s for $q<0$ must be greater than slope h(q) for $q>0$. Usually, the large fluctuations are characterized by a smaller scaling exponent h(q) for multifractal series than the small fluctuations [26].

3.5.5 Advantages and Disadvantages

The main advantage of multifractal DFA is, it makes possible to distinguish whether the given time series exhibits multifractality or not by generating different scaling exponents for different values of power q where q varies from -20 to 20 and thereby avoids assuming that the scaling characteristic of the entire time series remains the same as is the case with DFA.

It thereby detects the long-range correlation in the given data and henceforth is able to distinguish between the healthy and pathologic subjects, thus, it has the benefit of both DFA and itself.

A major drawback is the data requirements are greater than with other techniques and needs 24 hours of data because the result becomes statistically inconsistent for short data sets [26]. Thus, 24 hours long RR interval time series is needed for MF-DFA Analysis. Also MF-DFA only generates positive exponents which could be a possible drawback because for q=2, MF-DFA retains conventional DFA i.e MF-DFA generates the same exponent value as does DFA.

DFA results show that for $0< \alpha < 0.5$ the given signal is anticorrelated, which means that the method is not able to generate the exponent value beyond 0, which might mean greater anticorrelation. Thus, both have a common drawback of only generating positive value of the scaling exponent. Also, $F_q(s)$ becomes inaccurate for strongly anticorrelated signal i.e. h(q) is close to zero [26].

CHAPTER 4

DATA ACQUISTION AND DATA ANALYSIS

4.1 Data Acquisition

As mentioned in Section 1.2, the main goal behind this research is to i) Determine the presence of long-range power law correlation for healthy subjects and breakdown in the long-range power law correlation for CHF subjects using DFA and MF-DFA analysis and to reproduce the findings of Peng et al. (ii) Determine how well DFA and MF-DFA analysis are able to discriminate among healthy subjects and subjects with cardiac heart failure using statistical tools. (iii) Test the effect of length, activity, non-stationarity, trends on DFA and the effect of length on MF-DFA. The data used in the experiments were acquired from http://www.physionet.org/physiobank/database/nsr2db/.

This database includes beat annotation files for 26 long-term ECG recordings of subjects in normal sinus rhythm (26 men and women, aged 28.5 to 76, mean 52.3) and 11 long-term ECG recordings of subjects (11 men and women, aged 34 to 79, mean 56.5) with congestive heart failure. The data for the normal control group were obtained from 24-hour Holter monitor recordings of 26 healthy subjects with ECG data sampled at 128 Hz. The data for the CHF group were obtained from 24-hour Holter recordings of 11 patients with recordings sampled at 250 Hz.

4.2 Data Aberrancies and Data Correction

The data used in this experiment, contained some entries that were too high or too low to be considered for this data set. Such entries were removed using "deglitching" available

on PhysioNet. The code is written in MATLAB. The code identifies the outliers by comparing the data set with an Auto Regressive prediction model of order 3. If the data points lie within 25% and 75% of the predicted value, than the point is labeled valid, otherwise invalid. Only valid points are used to calculate DFA and MF-DFA exponents.

4.3 Data Analysis

Figure 4.1 compares the DFA analysis of representative eight hour interbeat interval time series of a healthy and a patient with congestive heart failure (CHF). Notice that for large time scales, the healthy interbeat interval time series shows almost perfect power-law scaling over two decades ($11 \leq n \leq 10000$) with $\alpha=1.04$, while for the pathologic data set $\alpha=1.26$.

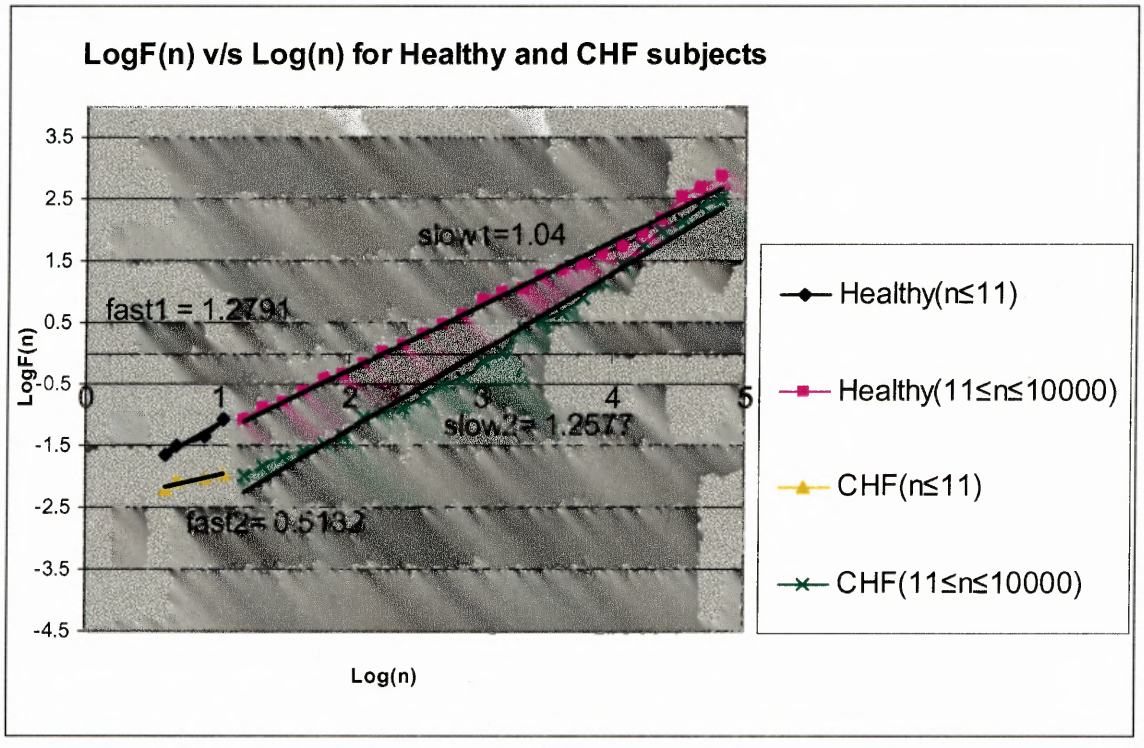

Figure 4.1 Plot of LogF(n) v/s Log(n) for two interbeat interval time series (~24hours) for healthy and subject with cardiac heart failure (CHF).

4.3.1 Crossover Phenomenon

Figure 4.1 shows the plot of LogF(n) v/s Log(n) for two interbeat interval time series for healthy subjects and subjects with cardiac heart failure. Analyzing the figure 4.1, it is noted that, there is an apparent crossover exhibited for the scaling behavior of both data sets indicated by slow1 and fast1 for healthy, and slow2 and fast2 for CHF. For the healthy subjects, α exponent estimated from very small n ($n \leq 11$ beats) is larger than that calculated from n (>11 beats) and this refers to crossover phenomenon. This is probably due to the fact that on very short time scales (a few beats to ten beats), the physiologic interbeat fluctuation is dominated by relatively smooth heart beat oscillation associated with respiration thus, giving rise to a larger α value. For longer scales, the interbeat fluctuation, reflecting the intrinsic dynamics of a complex system, approaches that of 1/f behavior as previously noted. In contrast, pathologic data set shows a very different crossover pattern (Figure 4.1). For very short time scales, the fluctuation is quite random (close to white noise, $\alpha \sim 0.5$). As time scale becomes larger, the fluctuation becomes smoother, asymptotically approaching Brownian noise, $\alpha \sim 1.5$).

Figure 4.1 shows the crossover effect resulting in two different exponents namely, slow and fast, each for healthy and subjects with cardiac heart failure. Therefore, for each data set an exponent $\alpha 1$ is calculated by making a least square fit of logF(n) v/s log(n) for $4 \leq n \leq 16$. Similarly, an exponent $\alpha 2$ is obtained from $16 \leq n \leq 64$. The two exponent's $\alpha 1$ and $\alpha 2$ were calculated for 4k, 8k, 10k, 12k, 16k, 18k, 20k, 22k, 28k, 30k, 40k, 60k, 80k, and 100k data points. For healthy subjects, calculated value of $\alpha 1$ and $\alpha 2$ (mean value $\pm$ S.D.) is 1.31 ± 0.17 and 1.003 ± 0.07 respectively. For subjects with cardiac heart failure, calculated value of $\alpha 1$ and $\alpha 2$ is 0.71 ± 0.20 and 1.24 ± 0.07, respectively.

Table 4.1 Comparison between results generated by Peng et al. and results generated using DFA analysis for Healthy and CHF subjects.

	α1(Peng et al.)	α2(Peng et al.)	α1(generated)	α2(generated)
Healthy	1.201 ± 0.178	0.998 ± 0.124	1.31 ± 0.17	1.003 ± 0.07
CHF	0.803 ± 0.259	1.12 ± 0.216	0.71 ± 0.20	1.24 ± 0.07

Table 4.2 and Table 4.3 listed below shows the effect of length on DFA for healthy and CHF subjects. The value of slow exponent for 4k, 8k, 10k, 12k, 16k, 18k, 20k, 22k, 28k, 30k, 40k, 60k and 100k among healthy subjects is 1.04 ± 0.02 and the value of slow exponent among CHF subjects is 1.32 ± 0.02 (mean value ± S.D.). Thus, 4000 data points are sufficient for DFA analysis.

Table 4.2 Effect of data length on DFA for healthy subjects

	3:4000		3:8000		3:10000		3:12000	
	slow	fast	slow	Fast	slow	fast	slow	fast
nsr001	1.1083	1.5094	1.0732	1.5551	1.0304	1.5254	1.055	1.518
nsr002	1.0244	1.3214	1.0469	1.3433	1.0221	1.3728	1.0264	1.3795
nsr003	1.1279	1.2817	0.9854	1.3331	1.0254	1.3452	1.0148	1.3354
nsr004	1.0213	1.3764	1.0198	1.3678	1.0089	1.3622	1.0402	1.3813
nsr005	1.0389	1.4385	0.9913	1.3922	0.9442	1.4354	0.9484	1.4168
nsr006	1.1373	1.4224	1.0964	1.4208	1.0737	1.4261	1.1017	1.4062
nsr007	1.0406	1.2402	0.9999	1.2631	0.997	1.2557	0.9985	1.2425
nsr008	0.882	1.2669	1.0024	1.2428	0.9851	1.3612	1.011	1.1925
nsr009	1.0759	1.1048	1.0101	1.1645	0.924	1.2728	0.9312	1.2753
nsr010	1.0282	1.3091	0.9699	1.3219	0.9845	1.3426	0.9784	1.3491
nsr011	0.9365	0.9885	0.9326	1.1226	0.9515	1.1149	0.9613	1.1071
nsr012	1.2146	1.3691	1.0661	1.4029	1.0455	1.3769	0.9624	1.3667
nsr013	0.8963	1.1836	0.837	1.1521	0.8311	1.1518	0.8409	1.1848
nsr014	1.1976	1.3239	1.1276	1.3676	1.132	1.4164	1.1477	1.3942
nsr015	0.9219	1.0699	0.945	1.223	0.9212	1.2445	0.9516	1.2169
nsr016	0.9845	1.1205	1.0284	0.994	1.0576	1.0063	1.0504	1.0262
nsr017	1.0067	1.516	0.9981	1.4887	1.0022	1.4659	1.0123	1.4381
nsr018	0.9476	1.0951	0.9744	1.1519	0.9863	1.1352	1.0358	1.1613
nsr019	1.0929	1.4371	1.1282	1.3822	1.188	1.4305	1.1997	1.3805
nsr020	1.1736	0.6183	1.0748	0.7573	1.0494	0.7952	1.064	0.8062
nsr021	0.9887	1.502	1.032	1.4259	1.0597	1.4417	1.0544	1.451
nsr022	0.8809	1.4719	0.9372	1.4882	0.9022	1.5217	0.9138	1.5084
nsr023	0.7812	1.3644	0.8741	1.3729	0.8722	1.4076	0.9225	1.364
nsr024	1.0671	1.5338	1.1073	1.4894	1.0541	1.5669	1.0449	1.5289
nsr025	1.0303	1.3292	0.9717	1.3181	0.9766	1.3142	0.9782	1.325
nsr026	0.8529	1.4961	0.8636	1.4873	0.8549	1.4485	0.8581	1.4232
AVER AGE	1.0176	1.2957	1.0035	1.3087	0.9953	1.3283	1.0039	1.3145

Table 4.2 Effect of data length on DFA for healthy subjects (continued)

3:16000		3:18000		3: 20000		3: 22000	
slow	fast	slow	Fast	Slow	fast	Slow	Fast
1.0466	1.4966	1.0486	1.4911	1.0519	1.4878	1.0508	1.4851
1.0218	1.3718	1.0154	1.3780	1.0189	1.3701	1.0053	1.3756
1.0220	1.2672	1.0323	1.2546	1.0272	1.2345	1.0287	1.2510
1.0357	1.4059	1.0405	1.4254	1.0463	1.4046	1.0524	1.4035
0.9685	1.3821	0.9492	1.3877	0.9667	1.3833	0.9710	1.3747
1.0361	1.4711	1.0025	1.4721	1.0057	1.4870	1.0017	1.4809
0.9845	1.2209	0.9987	1.1748	1.0014	1.1492	1.0034	1.1242
0.9700	1.1120	0.9331	1.0843	0.9625	1.0822	0.9925	1.0331
0.9924	1.1593	1.0117	1.1735	1.0143	1.1749	1.0192	1.1667
0.9934	1.3508	0.9852	1.3396	0.993	1.3372	0.9963	1.3098
0.9855	1.1922	0.9975	1.2463	0.9974	1.2726	1.0082	1.3104
0.9969	1.3798	0.9917	1.4018	0.9829	1.4225	0.9835	1.4175
0.8452	1.1976	0.8388	1.1912	0.8439	1.1763	0.8495	1.1542
1.1218	1.4529	1.1261	1.4567	1.0959	1.4628	1.1049	1.4641
0.9802	1.2044	0.9837	1.2094	0.9984	1.2339	1.0052	1.2442
1.1062	1.0145	1.1137	1.0173	1.1436	0.9906	1.1586	0.9609
0.9758	1.4907	0.9356	1.5053	0.9351	1.5038	0.9204	1.5033
1.0720	1.1519	1.0375	1.1424	1.0551	1.1667	1.0568	1.1604
1.1736	1.3978	1.1625	1.3450	1.1504	1.3100	1.1516	1.3133
1.0860	0.7793	1.0678	0.7154	1.0522	0.7190	1.0860	0.7341
1.0138	1.4525	1.0073	1.4489	1.0010	1.4499	1.0193	1.4491
0.8817	1.5214	0.8925	1.5243	0.8893	1.5294	0.8944	1.5260
0.9222	1.4234	0.9270	1.4445	0.9448	1.4165	0.9450	1.3980
1.0312	1.5436	1.0096	1.5409	1.0086	1.5401	1.0197	1.5365
0.9592	1.3502	0.9640	1.3423	0.9757	1.3603	0.9766	1.3525
0.8601	1.4161	0.8776	1.4106	0.8745	1.4270	0.889	1.4229
AVERAGE							
1.0031	1.3156	0.9980	1.3124	1.0014	1.3112	1.0073	1.3058

Table 4.2 Effect of data length on DFA for healthy subjects (continued)

3:28000		3:30000		3:40000		3:60000	
slow	fast	slow	Fast	slow	Fast	slow	fast
1.0252	1.493	1.0355	1.5	1.0212	1.4125	1.0203	1.4656
0.9931	1.3578	0.9898	1.3391	0.99	1.3201	0.992	1.3121
1.0826	1.2403	1.0804	1.2611	1.0418	1.3075	1.0616	1.3331
1.0536	1.3643	1.0561	1.3226	0.966	1.3732	0.9569	1.323
0.9688	1.3592	0.956	1.3726	0.966	1.3732	0.9735	1.3616
1.0055	1.4614	1.0083	1.4882	0.9257	1.4435	0.9406	1.4621
1.0094	1.129	1.0024	1.1305	0.9867	1.1523	1.0661	0.9939
0.9514	1.0116	0.9619	1.0225	0.9696	0.9086	1.0006	0.8529
1.0648	1.1669	1.0624	1.1823	1.0691	1.2098	1.0862	1.2018
1.007	1.3213	0.9967	1.3294	0.9455	1.3699	0.9379	1.3663
0.9742	1.3692	1.0162	1.3687	0.9915	1.364	1.0323	1.402
0.9845	1.4177	0.9742	1.4195	0.98	1.4235	0.9866	1.3161
0.8629	1.2089	0.8683	1.1922	0.805	1.2771	0.8844	1.2483
1.088	1.449	1.1045	1.4674	1.1288	1.4562	1.114	1.4631
1.0112	1.2562	1.0146	1.2489	1.0287	1.2209	1.0171	1.1548
1.1524	0.989	1.1254	0.9975	1.1281	1.0363	1.0767	0.9838
0.9454	1.4543	0.9852	1.4683	0.9386	1.4597	0.909	1.3221
1.0469	1.1388	1.0519	1.1295	1.0911	1.1222	1.0944	1.1425
1.1433	1.3431	1.1417	1.3435	1.1197	1.3402	1.0108	1.2739
1.1082	0.7787	1.1049	0.7881	1.1162	0.8089	1.1435	0.8594
0.9929	1.4454	0.9959	1.4472	0.8514	1.5177	0.9798	1.4292
0.8773	1.5238	0.8756	1.5231	0.8514	1.5177	0.8677	1.5017
0.9645	1.3904	0.9723	1.3922	0.9856	1.3961	0.969	1.4251
1.0352	1.5331	1.0346	1.5337	1.0663	1.5172	1.0794	1.5506
0.956	1.3587	0.9607	1.3515	0.9668	1.3453	1.0029	1.3341
0.8888	1.4171	0.8858	1.4204	0.8973	1.3859	0.9005	1.3923
AVERAGE							
1.0074	1.3060	1.010	1.3092	0.9933	1.3099	1.0039	1.2873

Table 4.2 Effect of data length on DFA for healthy subjects (continued)

3:80000		3:100000	
slow	Fast	Slow	fast
1.0312	1.4342	1.0123	1.3912
0.9982	1.2998	1.0669	1.2660
1.0433	1.2889	1.0669	1.2660
0.9566	1.3221	1.0206	1.3243
0.9786	1.3015	0.9147	1.2568
0.9434	1.4424	0.9800	1.4602
1.0561	0.9912	1.0845	1.0191
1.0002	0.8512	1.0093	0.8599
1.0982	1.1989	1.1253	1.3039
0.9456	1.3452	1.0016	1.3641
1.0592	1.4010	1.0497	1.4030
1.0091	1.3582	0.9994	1.3744
0.9012	1.2662	0.886	1.2764
1.1271	1.4540	1.1281	1.4503
0.9862	1.1460	1.0059	1.0059
1.0276	0.8455	1.0426	0.9230
0.8937	1.3452	0.8682	1.3763
1.0664	1.1445	1.0680	1.1095
1.0119	1.2880	1.0510	1.3026
1.1845	0.9461	1.1946	0.9708
0.9825	1.4216	1.0352	1.3700
0.9600	1.4869	0.9833	1.4890
0.9753	1.4262	1.0036	1.3853
1.0884	1.4261	1.1075	1.3470
1.0087	1.2814	1.0285	1.2743
0.9126	1.3723	0.9308	1.3398
AVERAGE			
1.0094	1.2724	1.0255	1.2657

Table 4.3 Effect of Data Length on DFA for CHF subjects

	3:4000		3:8000		3:10000		3:12000	
	slow	fast	slow	Fast	Slow	fast	slow	fast
Chf201	1.1389	0.696	1.1882	0.7743	1.1342	0.7206	1.1511	0.7683
Chf202	1.4038	0.5686	1.2843	0.4805	1.3666	0.5151	1.4585	0.5637
Chf203	1.2525	0.6764	1.1989	0.8355	1.2858	0.8843	1.4099	0.9129
Chf204	1.1831	0.3915	1.0983	0.5306	1.2866	0.9536	1.2233	0.7158
Chf205	1.2043	0.6327	1.3066	0.586	1.224	0.5803	1.2129	0.6577
Chf206	1.1817	0.9109	1.2131	0.8802	1.174	0.9012	1.1854	0.8938
Chf207	1.2745	1.3108	1.243	1.287	1.2218	1.3151	1.2086	1.3113
Chf208	1.2622	0.6063	1.2318	0.9589	1.1919	1.0036	1.2161	1.0622
Chf209	1.4311	1.1421	1.364	0.9185	1.2438	0.9628	1.2029	0.8912
Chf210	1.2784	0.9038	1.2701	0.8352	1.175	0.805	0.9734	0.5439
Chf211	1.2324	0.4337	1.1933	0.4124	1.1489	0.3862	1.1215	0.394
AVER-AGE	1.2584	0.7520	1.2356	0.7726	1.2229	0.8207	1.2148	0.7922

Table 4.3 Effect of Data Length on DFA for CHF subjects (continued)

3:16000		3:18000		3: 20000		3: 22000	
Slow	fast	slow	Fast	Slow	fast	Slow	fast
1.1289	0.6953	1.1361	0.6642	1.1556	0.625	1.1352	0.6101
1.4809	0.5925	1.4033	0.58	1.3757	0.5027	1.4285	0.6142
1.4488	0.9344	1.4174	0.8724	1.4134	0.8539	1.4076	0.8431
1.2089	0.9059	1.3946	0.864	1.3224	0.8288	1.1731	0.7978
1.1637	0.6512	1.1914	0.6638	1.1764	0.6761	1.1812	0.6762
1.2018	0.9062	1.195	0.8733	1.2013	0.9033	1.2058	0.9206
1.1934	1.3195	1.1941	1.3238	1.1984	1.3014	1.2226	1.2862
1.1738	0.9138	1.1743	0.888	1.1547	0.9309	1.1509	0.9261
1.0978	0.883	1.061	0.8732	1.0477	0.8693	1.0683	0.9207
0.7923	0.4754	0.7774	0.4509	0.8196	0.4345	0.8424	0.424
1.1766	0.4063	1.1874	0.4297	1.185	0.4327	1.1787	0.4152
AVERAGE							
1.1879	0.7894	1.1938	0.7712	1.1863	0.7598	1.1813	0.7667

Table 4.3 shows the effect of data length on DFA for subjects with cardiac heart failure.

Table 4.3 Effect of data length on DFA for CHF subjects (continued)

3:28000		3:30000		3:40000		3:60000	
slow	fast	slow	Fast	Slow	fast	Slow	fast
1.1416	0.6067	1.1385	0.5949	1.1439	0.5833	1.1552	0.706
1.3784	0.5451	1.397	0.6388	1.3436	0.5419	1.4136	0.596
1.384	0.8645	1.379	0.8621	1.3561	0.8458	1.3208	0.786
1.3012	0.6678	1.3491	0.7519	1.2972	0.7128	1.4005	0.6185
1.185	0.6912	1.1575	0.6782	1.1598	0.6572	1.1798	0.6341
1.2011	0.9145	1.1911	0.9194	1.1878	0.9144	1.2238	0.8432
1.2218	1.292	1.2171	1.2926	1.2446	1.2513	1.2329	1.2309
1.1565	0.9087	1.1423	0.9027	1.0834	0.9171	1.1367	0.8631
1.181	0.9292	1.2195	0.938	1.2302	0.9529	1.2521	0.9623
0.8859	0.4434	0.9489	0.4453	0.9904	0.4747	0.9386	0.4904
1.2065	0.4299	1.1891	0.4352	1.2402	0.4601	1.1753	0.4343
AVERAGE							
1.2039	0.7539	1.2117	0.7699	1.2070	0.7555	1.2208	0.7422

Table 4.3 Effect of data length on DFA for CHF subjects (continued)

3:80000		3:100000	
slow	Fast	Slow	fast
1.1549	0.8026	1.1467	0.8317
1.4832	0.6899	1.4729	0.6786
1.3251	0.8227	1.3639	0.8407
1.3008	0.6814	1.3119	0.6140
1.1137	0.6010	1.1578	0.5909
1.2403	0.8575	0.8724	0.8471
1.2605	1.2246	1.2542	1.1621
1.1309	0.8300	1.1077	0.8238
1.3429	0.9911	1.3588	0.9553
1.0902	0.5762	0.8837	0.3806
1.1682	0.4782	1.1761	0.5030
AVERAGE			
1.2373	0.7777	1.1914	0.7479

4.3.2 ANOVA Analysis

ANOVA (Analysis of variance) is used for the statistical analysis of the data sets. ANOVA is used to compare the effect of length on $\alpha 1$(fast) and $\alpha 2$(slow) exponent of DFA for healthy and CHF subjects. The null hypothesis for the experiment is that, if the generated p-value using ANOVA analysis is less then 0.05, reject the null hypothesis, and if the generated p-value is greater then 0.05, do not reject the hypothesis.

SUMMARY	Count	Sum	Average	Variance
Row 1	26	26.458	1.0176	0.012
Row 2	26	26.093	1.0036	0.0059
Row 3	26	25.88	0.9954	0.0067
Row 4	26	26.104	1.004	0.0065
Row 5	26	26.082	1.0032	0.0058
Row 6	26	25.95	0.9981	0.0055
Row 7	26	26.037	1.0014	0.0051
Row 8	26	26.19	1.0073	0.0053
Row 9	26	26.193	1.0074	0.0055
Row 10	26	26.261	1.0101	0.0051
Row 11	26	25.828	0.9934	0.0076
Row 12	26	26.104	1.004	0.0053
Row 13	26	26.246	1.0095	0.0048
Row 14	26	26.665	1.0256	0.0055
Column 1	14	14.611	1.0436	0.0006
Column 2	14	14.211	1.0151	0.0005
Column 3	14	14.64	1.0457	0.0013
Column 4	14	14.275	1.0196	0.0013
Column 5	14	13.536	0.9668	0.0008
Column 6	14	14.259	1.0185	0.0041
Column 7	14	14.229	1.0164	0.001
Column 8	14	13.632	0.9737	0.0012
Column 9	14	14.485	1.0346	0.0035
Column 10	14	13.763	0.9831	0.0007
Column 11	14	13.894	0.9924	0.0015
Column 12	14	14.177	1.0127	0.0041
Column 13	14	11.991	0.8565	0.0008
Column 14	14	15.744	1.1246	0.0007
Column 15	14	13.771	0.9836	0.0013
Column 16	14	15.196	1.0854	0.003
Column 17	14	13.326	0.9519	0.0021
Column 18	14	14.584	1.0417	0.0019
Column 19	14	15.725	1.1232	0.0037
Column 20	14	15.506	1.1076	0.0024
Column 21	14	14.014	1.001	0.0025
Column 22	14	12.607	0.9005	0.0013
Column 23	14	13.059	0.9328	0.0034
Column 24	14	14.754	1.0539	0.0011
Column 25	14	13.756	0.9826	0.0006
Column 26	14	12.347	0.8819	0.0005

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Rows	0.0241	13	0.0019	1.0525	0.4006	1.7504
Columns	1.5948	25	0.0638	36.244	9E-79	1.54
Error	0.572	325	0.0018			
Total	2.1909	363				

Figure 4.2 ANOVA for effect of length on DFA for healthy subjects.

SUMMARY	Count	Sum	Average	Variance
Row 1	11	13.843	1.2584	0.0081
Row 2	11	13.592	1.2356	0.005
Row 3	11	13.453	1.223	0.0048
Row 4	11	13.364	1.2149	0.017
Row 5	11	13.067	1.1879	0.0324
Row 6	11	13.132	1.1938	0.0329
Row 7	11	13.05	1.1864	0.0262
Row 8	11	12.994	1.1813	0.0245
Row 9	11	13.243	1.2039	0.0181
Row 10	11	13.329	1.2117	0.0164
Row 11	11	13.277	1.207	0.0121
Row 12	11	13.429	1.2208	0.0175
Row 13	11	13.611	1.2373	0.0144
Row 14	11	13.106	1.1915	0.0362
Column 1	14	16.049	1.1464	0.0002
Column 2	14	19.69	1.4065	0.0032
Column 3	14	18.963	1.3545	0.005
Column 4	14	17.851	1.2751	0.0076
Column 5	14	16.614	1.1867	0.0019
Column 6	14	16.475	1.1768	0.008
Column 7	14	17.188	1.2277	0.0006
Column 8	14	16.313	1.1652	0.0023
Column 9	14	17.101	1.2215	0.0151
Column 10	14	13.666	0.9762	0.0281
Column 11	14	16.579	1.1842	0.0009
ANOVA				

Source of Variation	SS	df	MS	F	P-value	F crit
Rows	0.0723	13	0.0056	0.8255	0.6325	1.7961
Columns	1.7807	10	0.1781	26.441	8E-27	1.9042
Error	0.8755	130	0.0067			
Total	2.7284	153				

Figure 4.3 ANOVA for effect of length on DFA for CHF.

SUMMARY	Count	Sum	Average	Variance
Row 1	11	11.3318	1.030164	0.003259
Row 2	11	13.1061	1.191464	0.036229
Column 1	2	2.159	1.0795	0.009032
Column 2	2	2.5398	1.2699	0.082418
Column 3	2	2.4308	1.2154	0.044105
Column 4	2	2.3325	1.16625	0.042428
Column 5	2	2.0725	1.03625	0.029549
Column 6	2	1.8524	0.9262	0.005789
Column 7	2	2.3387	1.16935	0.014399
Column 8	2	2.117	1.0585	0.004841
Column 9	2	2.4841	1.24205	0.027261
Column 10	2	1.8853	0.94265	0.00695
Column 11	2	2.2258	1.1129	0.007988
ANOVA				

Source of Variation	SS	df	MS	F	P-value	F crit
Rows	0.143097	1	0.143097	10.86849	0.008058	4.964603
Columns	0.263224	10	0.026322	1.999232	0.144977	2.978237
Error	0.131663	10	0.013166			
Total	0.537984	21				

Figure 4.4 ANOVA of Healthy v/s CHF

Figure 4.2 and Figure 4.3 shows the ANOVA analysis for the effect of length on α1(fast) and α2(slow) exponent of DFA for healthy and CHF subjects. Figure 4.4 shows the ANOVA of healthy v/s CHF. As shown in the table p-value for healthy subject is 0.4006 and for CHF subject is 0.6325, which is greater then 0.05. This suggests that there is no statistical change in the values of slow exponents among healthy subjects and there is no statistical change in the value of slow exponents among CHF subjects. And the p-value for healthy versus CHF is 0.008 which is less then 0.05, thus reject the null hypothesis, which means both the data sets are statistically different. Figure 4.5 shows the plot of α2(slow) versus number of data points for healthy and CHF subjects and it is apparent that there is no effect of length on α2(slow). However, DFA analysis is done

using atleast 4000 data points as indicated in table 4.2 because the results become inconsistent for data points less than 4000, and the value of slow exponent for healthy subjects is 1.004 ± 0.07 (mean ± S.D.). The pathology group shows a significant deviation of the long range correlation exponent from the normal. For the group of heart failure the value of α exponent is 1.21 ± 0.13 (mean ± S.D.).

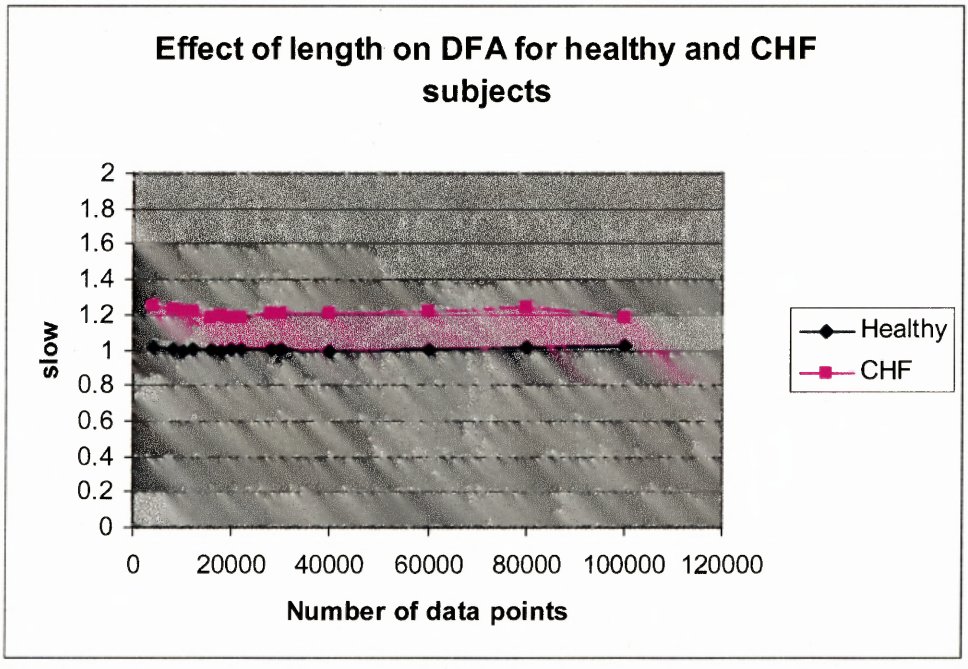

Figure 4.5 Effect of length on DFA for healthy and CHF subjects.

Tables 4.4 and 4.5 shows the effect of activity on DFA for healthy and subjects with cardiac heart failure.

Table 4.4 Effect of activity on DFA for healthy subjects

	3:33000		33001:60000		60001:100000	
	slow	Fast	Slow	fast	slow	fast
nsr001	1.0547	1.4743	1.0205	1.497	1.0674	1.5143
nsr002	1.0087	1.316	1.0324	1.096	1.0289	1.3426
nsr003	1.0721	1.2754	1.0502	1.4133	1.0745	1.2448
nsr004	1.032	1.3223	0.9413	1.2881	1.0174	1.364
nsr005	0.9594	1.3538	0.9917	1.3672	0.8463	1.2299
nsr006	1.0024	1.468	0.9998	1.4799	0.9622	1.4474
nsr007	1.0084	1.1208	1.1431	0.8702	1.0873	1.1114
nsr008	0.982	0.994	1.0047	0.8604	1.0456	0.8849
nsr009	1.048	1.406	1.0589	1.4615	1.0307	1.4332
nsr010	0.9829	1.3171	0.9747	1.3896	1.0417	1.3415
nsr011	1.0178	1.4689	1.0383	1.4998	1.0296	1.4815
nsr012	1.0414	1.272	1.0707	1.2483	1.0082	1.3924
nsr013	1.0219	1.3596	1.0121	1.4265	1.0356	1.3801
nsr014	0.9637	1.4223	0.9828	1.4669	1.002	1.4558
nsr015	1.0296	1.1968	1.0018	1.2615	1.0014	1.2566
nsr016	0.8246	1.1961	0.9137	1.4021	0.9028	1.3309
nsr017	1.1617	1.4047	1.0379	1.2334	0.9507	1.2075
nsr018	1.0016	1.1233	1.0971	1.0622	1.0308	0.9304
nsr019	0.9711	1.4812	0.891	1.3273	0.8291	1.4761
nsr020	1.0388	1.3402	1.0237	1.2451	1.01311	1.3303
nsr021	1.0569	1.1862	1.0091	1.2795	1.0116	1.0569
nsr022	0.991	1.4334	0.9878	1.4188	1.074	1.3173
nsr023	0.7808	1.4857	0.8496	1.3131	0.8992	0.9978
nsr024	0.9538	1.3975	0.9291	1.4589	1.0602	1.3445
nsr025	1.0642	1.1917	1.025	1.2064	1.1	1.1034
nsr026	1.0377	1.5322	1.005	1.5448	1.0631	1.2444
nsr027	0.967	1.3536	1.0037	1.3185	1.0429	1.2203
AVERAGE	1.0027	1.3293	1.0035	1.3124	1.0094	1.2755

Table 4.5 Effect of activity on DFA for CHF subjects

	3:30000		30001:60000		60001:90000	
	slow	fast	Slow	fast	slow	fast
chf201	1.1385	0.5949	1.1805	0.8460	1.0786	1.0105
chf202	1.3970	0.6388	1.5071	0.5908	1.5265	0.7907
chf203	1.3790	0.8621	1.2174	0.7466	1.4165	0.9467
chf204	1.3491	0.7519	1.4612	0.5275	1.2977	0.5824
chf205	1.1598	0.6782	1.1934	0.6161	1.2852	0.5708
chf206	1.2011	0.9194	1.1544	0.7649	1.3017	0.8610
chf207	1.1423	0.9027	1.1440	0.8059	1.0779	0.6782
chf208	1.0244	0.4475	0.9099	0.5770	1.2745	0.7356
chf209	1.2195	0.9380	1.3579	0.9496	1.4884	1.2132
chf210	1.1775	0.6739	1.2098	0.7524	1.2936	1.1966
chf211	1.1891	0.8545	1.1408	0.4600	1.1829	0.4167
AVERAGE	1.2161	0.7510	1.2251	0.6942	1.2930	0.8184

SUMMARY	*Count*	*Sum*	*Average*	*Variance*
Row 1	27	27.0742	1.002748	0.005301
Row 2	27	27.0957	1.003544	0.003769
Row 3	27	27.25631	1.009493	0.004841
Column 1	3	3.1426	1.047533	0.000588
Column 2	3	3.07	1.023333	0.000164
Column 3	3	3.1968	1.0656	0.000179
Column 4	3	2.9907	0.9969	0.002372
Column 5	3	2.7974	0.932467	0.005829
Column 6	3	2.9644	0.988133	0.000506
Column 7	3	3.2388	1.0796	0.00458
Column 8	3	3.0323	1.010767	0.001039
Column 9	3	3.1376	1.045867	0.000202
Column 10	3	2.9993	0.999767	0.001336
Column 11	3	3.0857	1.028567	0.000106
Column 12	3	3.1203	1.0401	0.000978
Column 13	3	3.0696	1.0232	0.000139
Column 14	3	2.9485	0.982833	0.000367
Column 15	3	3.0328	1.010933	0.000261
Column 16	3	2.6411	0.880367	0.002362
Column 17	3	3.1503	1.0501	0.011242
Column 18	3	3.1295	1.043167	0.002395
Column 19	3	2.6912	0.897067	0.005069
Column 20	3	3.07561	1.025203	0.000167
Column 21	3	3.0776	1.025867	0.000724
Column 22	3	3.0528	1.0176	0.002388
Column 23	3	2.5296	0.8432	0.003535
Column 24	3	2.9431	0.981033	0.004853
Column 25	3	3.1892	1.063067	0.001407
Column 26	3	3.1058	1.035267	0.000848
Column 27	3	3.0136	1.004533	0.001441
ANOVA				

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Rows	0.000734	2	0.000367	0.174313	0.840523	3.175141
Columns	0.252257	26	0.009702	4.610721	1.4E-06	1.70962
Error	0.109422	52	0.002104			
Total	0.362413	80				

Figure 4.6 ANOVA for effect of activity on DFA for healthy subjects.

SUMMARY	Count	Sum	Average	Variance
Row 1	11	13.3773	1.216118	0.013079
Row 2	11	13.4764	1.225127	0.027422
Column 1	2	2.319	1.1595	0.000882
Column 2	2	2.9041	1.45205	0.006061
Column 3	2	2.5964	1.2982	0.013057
Column 4	2	2.8103	1.40515	0.006283
Column 5	2	2.3532	1.1766	0.000564
Column 6	2	2.3555	1.17775	0.00109
Column 7	2	2.2863	1.14315	1.44E-06
Column 8	2	1.9343	0.96715	0.006555
Column 9	2	2.5774	1.2887	0.009577
Column 10	2	2.3873	1.19365	0.000522
Column 11	2	2.3299	1.16495	0.001166
ANOVA				

Source of Variation	SS	df	MS	F	P-value	F crit
Rows	0.000446	1	0.000446	0.098513	0.760071	4.964603
Columns	0.359702	10	0.03597	7.938002	0.001496	2.978237
Error	0.045314	10	0.004531			
Total	0.405463	21				

Figure 4.7 ANOVA for effect of activity on DFA for CHF subjects.

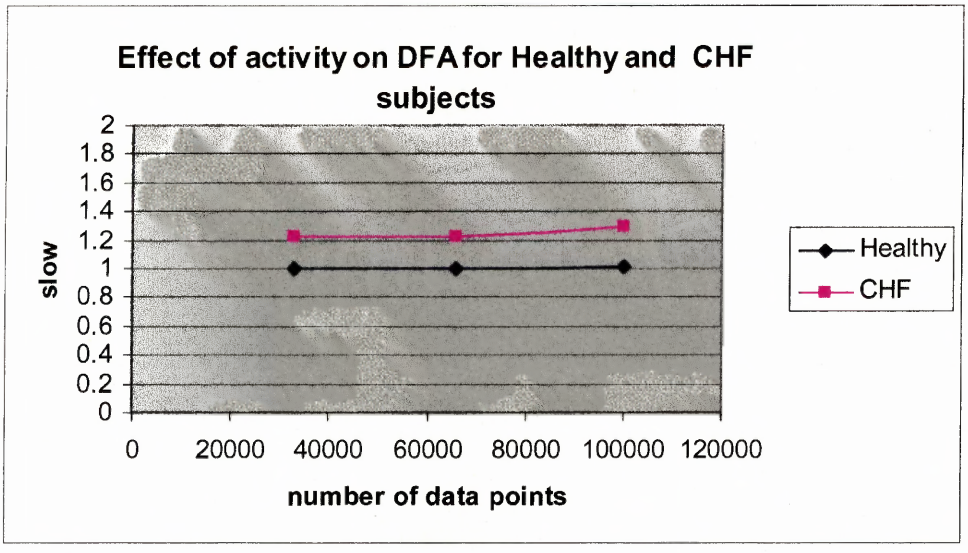

Figure 4.8 Effect of activity on DFA for healthy and CHF subjects.

Figure 4.6 and Figure 4.7 shows the ANOVA analysis for the effect of activity on DFA for healthy and CHF subjects and the p-values for the healthy group is 0.8405 and for CHF group is 0.7600 which is again greater then 0.05 and thus do not reject the null hypothesis which according to the p-value suggests, there is no effect of activity on the DFA scaling exponent for healthy as well as CHF. This can also be seen in figure 4.8.

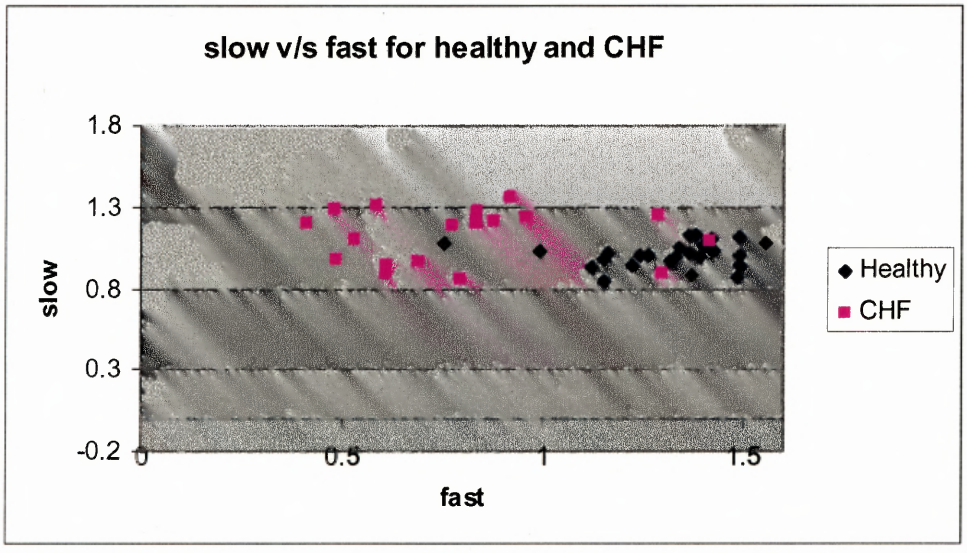

Figure 4.9 Scatter plot of scaling exponent's slow(α2) v/s fast(α1) for the healthy subjects and subjects with cardiac heart failure. Note: Good separation between healthy and heart disease subjects, with clustering of points in two distinct "clouds."

Figure 4.9 shows that good discrimination between these two groups can be achieved by using these two scaling exponents and statistically the p-value of fast among healthy subjects and subjects with cardiac failure is ~0 which is < 0.05, and thus the value of fast is significantly different among both the group of subjects.

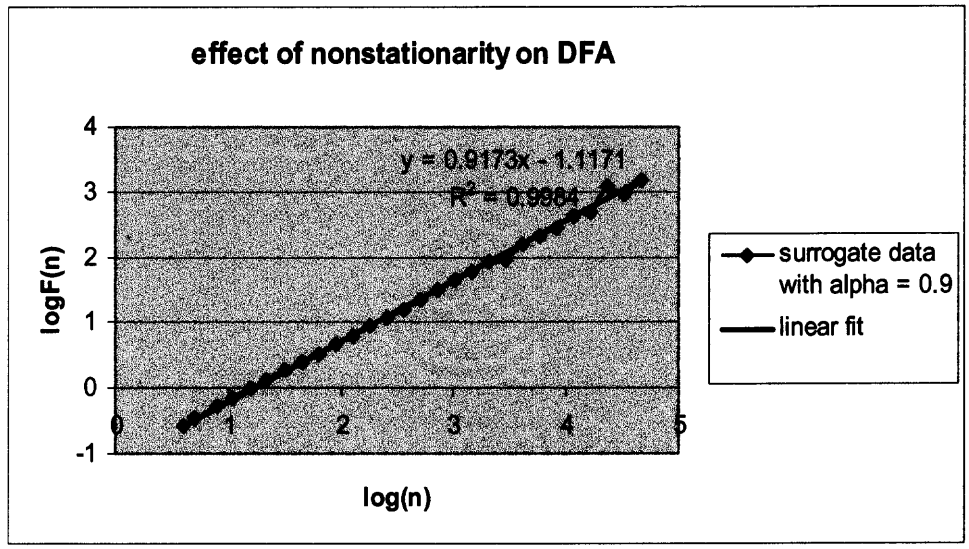

Figure 4.10 Effect of nonstationarity on DFA.

Figure 4.10 shows the plot of LogF(n) v/s Log(n) for effect of non-stationarity on the DFA. A surrogate signal with non-stationarity is taken from www.physionet.org. An artificially generated signal with known correlation α =0.9 is used. Non-stationarity is introduced by stitching together segments of data obtained from discontinuous experimental recordings, or removing some noisy and unreliable parts from continuous recordings and stitching together the remaining parts--a "cutting" procedure commonly used in preparing data prior to signal analysis. The value of α exponent obtained using DFA analysis of the signal with an artificially introduced non-stationarity remained the same as the signal without non-stationarity as revealed in fig. 4.10. Thus it can be concluded that DFA is not affected by non-stationarity present in the data.

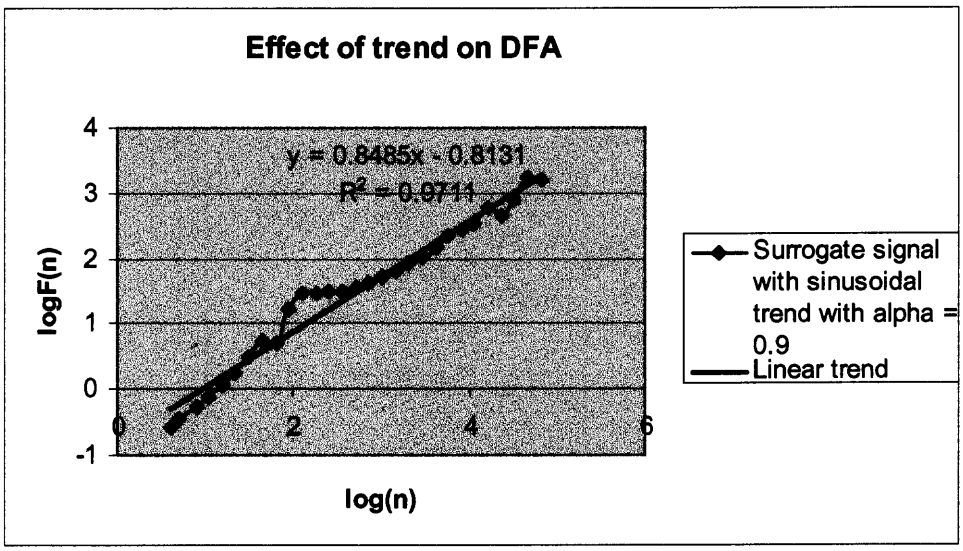

Figure 4.11 Effect of trend on DFA.

Figure 4.11 shows the effect of a sinusoidal trend on DFA. A surrogate signal with known correlation available on www.physionet.org is used to check the effect of trends on DFA.

A signal with known correlation and a trend with known scaling properties are used. The trend is superposed on the signal and the DFA analysis of this signal with superposed trend gives nearly the same α value as for the signal without any trend as shown in Figure 4.11. Thus, it can be concluded that there is no effect of trends on DFA because as discussed in the algorithm, DFA removes the effect of trend. The mathematics used to superpose the trend on the signal with known correlation is described by Ivanov et al. [33].

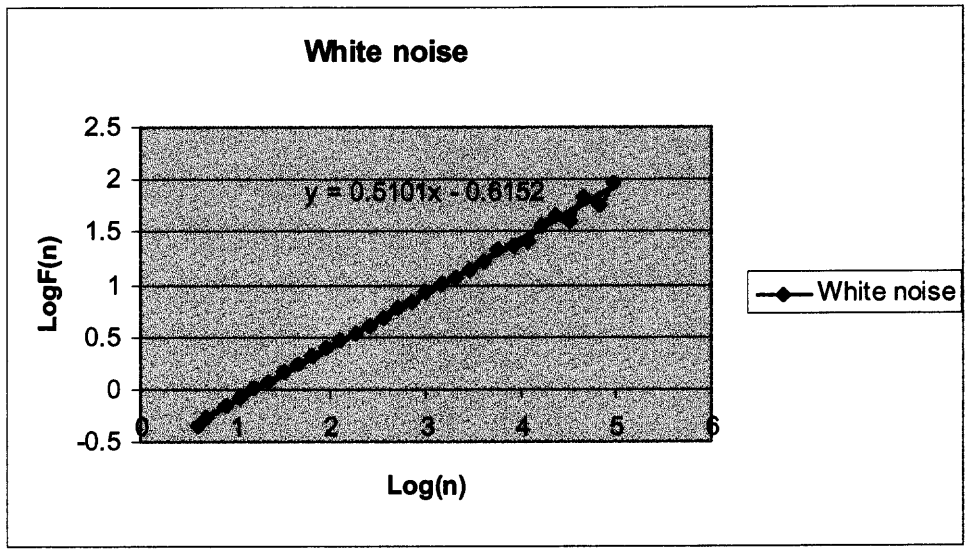

Figure 4.12 White noise.

Figure 4.12 shows the results of DFA analysis for white noise. An artificially generated noise signal available on www.physionet.org is used for the analysis. The value of α exponent using the DFA becomes α=0.5 which is white noise. Thus it can be concluded that for white noise, $\alpha = 0.5$.

Tables 4.6 and 4.7 shows the Multifractal (MF)-DFA results for Healthy and CHF subjects

Table 4.6 Multifractal DFA results for Healthy Subjects for q varying from -20 to 20

q	nsr001	nsr002	nsr003	nsr004	nsr005	nsr006	nsr007	nsr008
-20.0	3.1063	3.1034	3.1063	3.1130	3.1090	3.1063	3.1124	3.1263
-19.0	3.1110	3.1010	3.1116	3.1216	3.1116	3.1360	3.1160	3.1245
-18.0	3.1174	3.1174	3.1174	3.1174	3.1174	3.1174	3.1174	3.1174
-17.0	3.1380	3.1239	3.1090	3.1103	3.1239	3.1167	3.1129	3.1090
-16.0	3.1300	3.1233	3.1383	3.1413	3.1113	3.1210	3.1313	3.1203
-15.0	3.1289	3.1266	3.1396	3.1210	3.1396	3.1460	3.1400	3.1210
-14.0	3.1491	3.1491	3.1391	3.1491	3.1491	3.1491	3.1491	3.1491
-13.0	3.1600	3.1300	3.1239	3.1400	3.1600	3.1122	3.1600	3.1600
-12.0	3.1640	3.1780	3.1528	3.1728	3.1580	3.1438	3.1420	3.1741
-11.0	3.1880	3.1880	3.1880	3.1880	3.1880	3.1880	3.1880	3.1880
-10.0	3.1230	3.1121	3.1122	3.1406	3.1021	3.1121	3.1201	3.1121
-9.0	3.1231	3.1228	3.1228	3.1228	3.1228	3.1228	3.1228	3.1228
-8.0	3.1426	3.1320	3.1220	3.1460	3.1366	3.1426	3.1310	3.1326
-7.0	3.1292	3.1292	3.1292	3.1292	3.1292	3.1317	3.1292	3.1292
-6.0	3.0339	3.0339	3.0339	3.0339	3.0339	3.0339	3.0339	3.0339
-5.0	3.0141	3.0314	3.0450	3.0258	3.0150	3.0340	3.1141	3.1058
-4.0	2.6456	2.6505	2.6106	2.6506	2.5505	2.6206	2.6230	2.6506
-3.0	2.6720	2.6720	2.6720	2.6720	2.6720	2.6720	2.6720	2.6720
-2.0	2.1900	2.1800	2.2005	2.1960	2.2005	2.1800	2.2005	2.2109
-1.0	2.0030	2.1030	2.0430	2.0300	2.0100	2.0030	2.1201	2.0230
1.0	1.2153	1.2190	1.2253	1.2353	1.2283	1.2253	1.2300	1.2105
2.0	1.0181	0.9918	1.0281	1.0081	0.9902	1.0171	1.0097	0.9910
3.0	1.0647	1.0647	1.0647	1.0647	1.0647	1.0647	1.0647	1.0647
4.0	1.0857	1.0857	1.0857	1.0857	1.0857	1.0857	1.0857	1.0857
5.0	1.0936	1.0936	1.0936	1.0936	1.0936	1.0936	1.0936	1.0936
6.0	1.0972	1.0972	1.0972	1.0972	1.0972	1.0972	1.0972	1.0972
7.0	1.0890	1.0990	1.0760	1.0990	1.0690	1.0869	1.0990	1.0990
8.0	1.0321	1.0312	1.0321	1.0221	1.0321	1.0321	1.0321	1.0321
9.0	1.0501	1.0501	1.0501	1.0501	1.0501	1.0501	1.0501	1.0501
10.0	1.0301	1.0501	1.0501	1.0201	1.0601	1.0501	1.0501	1.0301
11.0	1.0501	1.0401	1.0501	1.0701	1.0501	1.0501	1.0440	1.0601
12.0	1.0510	1.0510	1.0510	1.0210	1.0100	1.0510	1.0510	1.0310
13.0	1.0510	1.0401	1.0501	1.0421	1.0301	1.0320	1.0401	1.0450
14.0	1.0350	1.0450	1.0250	1.0350	1.0350	1.0150	1.0350	1.0350
15.0	1.0250	1.0501	1.0250	1.0550	1.0250	1.0350	1.0250	1.0250
16.0	1.0350	1.0350	1.0350	1.0350	1.0350	1.0350	1.0350	1.0350
17.0	1.0235	1.0235	1.0335	1.0235	1.0424	1.0235	1.0235	1.0235
18.0	1.0145	1.0450	1.0145	1.0545	1.0345	1.0145	1.0345	1.0650
19.0	1.0250	1.0250	1.0250	1.0150	1.0250	1.0250	1.0250	1.0250
20.0	1.0350	1.0250	1.0150	1.0450	1.0350	1.0150	1.0350	1.0321

Table 4.6 Multifractal DFA results for Healthy Subjects for q varying from -20 to 20 (continued)

q	nsr009	nsr010	nsr011	nsr012	nsr013	nsr014	nsr015	nsr016
-20.0	3.1630	3.1263	3.1280	3.1530	3.1224	3.1187	3.1230	3.1263
-19.0	3.1116	3.1116	3.1116	3.1116	3.1160	3.1116	3.1116	3.1116
-18.0	3.1174	3.1174	3.1239	3.1174	3.1174	3.1174	3.1174	3.1174
-17.0	3.1265	3.1389	3.1413	3.1239	3.1239	3.1130	3.1210	3.1360
-16.0	3.1313	3.1383	3.1396	3.1313	3.1313	3.1313	3.1313	3.1313
-15.0	3.1396	3.1396	3.1491	3.1360	3.1656	3.1396	3.1430	3.1396
-14.0	3.1491	3.1391	3.1491	3.1491	3.1491	3.1491	3.1491	3.1491
-13.0	3.1600	3.1239	3.1600	3.1600	3.1600	3.1600	3.1600	3.1600
-12.0	3.1728	3.1528	3.1239	3.1600	3.1420	3.1728	3.1438	3.1728
-11.0	3.1228	3.1880	3.1528	3.1741	3.1880	3.1880	3.1880	3.1880
-10.0	3.1460	3.1122	3.1880	3.1880	3.1201	3.2061	3.1121	3.1810
-9.0	3.1292	3.1228	3.1122	3.1121	3.1228	3.1728	3.1228	3.1528
-8.0	3.0339	3.1220	3.1228	3.1228	3.1310	3.1880	3.1426	3.1880
-7.0	3.0258	3.1292	3.1220	3.1326	3.1292	3.1406	3.1317	3.1122
-6.0	2.6506	2.8034	3.1292	3.1292	3.0339	3.1228	3.0339	3.1228
-5.0	2.6720	2.7450	3.0339	3.0339	3.1141	3.1460	3.0340	3.1220
-4.0	2.1960	2.1911	3.0450	2.9120	2.6230	3.1292	2.6206	3.1292
-3.0	2.0300	2.6720	2.6106	2.6506	2.6720	3.0339	2.6720	3.0339
-2.0	2.2470	2.2005	2.2720	2.2372	2.2005	3.0258	2.1800	3.0450
-1.0	2.0030	2.0430	2.2005	2.1209	2.1201	2.0103	2.0030	2.1340
1.0	0.0000	0.0000	2.0430	2.0230	0.0000	0.0000	0.0000	0.0000
2.0	1.2190	1.2253	1.2454	1.2283	1.2253	1.2190	1.2300	1.2283
3.0	0.9918	1.0281	0.9062	0.9902	1.0097	0.9918	1.0097	0.9902
4.0	1.0647	1.0647	0.9108	1.0647	1.0647	1.0647	1.0647	1.0647
5.0	1.0857	1.0857	0.9041	1.0857	1.0857	1.0857	1.0857	1.0857
6.0	1.0936	1.0936	0.8945	1.0936	1.0936	1.0936	1.0936	1.0936
7.0	1.0972	1.0972	0.9918	1.0972	1.0972	1.0972	1.0972	1.0972
8.0	1.0990	1.0760	0.9902	1.0690	1.0990	1.0990	1.0990	1.0690
9.0	1.0312	1.0321	1.0647	1.0321	1.0321	1.0312	1.0321	1.0321
10.0	1.0501	1.0501	1.0857	1.0501	1.0501	1.0501	1.0501	1.0501
11.0	1.0501	1.0501	1.0936	1.0601	1.0501	1.0501	1.0501	1.0601
12.0	1.0401	1.0501	1.0972	1.0501	1.0440	1.0401	1.0440	1.0501
13.0	1.0510	1.0510	1.0690	1.0100	1.0510	1.0510	1.0510	0.9019
14.0	1.0401	1.0501	1.0321	1.0301	1.0401	1.0401	1.0401	0.9019
15.0	1.0450	1.0250	1.0501	1.0350	1.0350	0.9610	1.0350	0.9018
16.0	1.0501	1.0250	1.0601	1.0250	1.0250	0.9610	1.0250	0.9018
17.0	1.0350	1.0350	1.0501	1.0350	1.0350	0.9610	1.0350	0.9018
18.0	1.0235	1.0335	1.0350	1.0424	1.0235	0.9610	1.0235	0.9018
19.0	1.0450	1.0145	1.0235	1.0345	1.0345	0.9610	1.0345	0.9018
20.0	1.0250	1.0250	1.0345	1.0250	1.0250	0.9610	1.0250	0.9018

Table 4.6 Multifractal DFA results for Healthy Subjects for q varying from -20 to 20 (continued)

q	nsr017	nsr018	nsr019
-20.0	3.1063	3.1063	3.1063
-19.0	3.1116	3.1413	3.1174
-18.0	3.1174	3.1396	3.1389
-17.0	3.1413	3.1210	3.1383
-16.0	3.1396	3.1313	3.1313
-15.0	3.1396	3.1396	3.1396
-14.0	3.1491	3.1491	3.1239
-13.0	3.1600	3.1600	3.1528
-12.0	3.1600	3.1728	3.1880
-11.0	3.1741	3.1880	3.1880
-10.0	3.1880	3.1610	3.2061
-9.0	3.1121	3.1683	3.1880
-8.0	3.1228	3.1560	3.1810
-7.0	3.1326	3.1917	3.1528
-6.0	3.1392	3.1276	3.1880
-5.0	3.1458	3.1058	3.1122
-4.0	3.1260	3.1256	3.1228
-3.0	3.0120	3.1162	3.1220
-2.0	3.0047	3.0200	3.0067
-1.0	2.1203	2.0760	2.0873
0.0	0.0000	0.0000	0.0000
1.0	1.2106	1.2108	1.2457
2.0	1.0234	0.9553	0.9743
3.0	0.9062	1.0647	1.1621
4.0	0.9108	1.0857	1.1447
5.0	0.9041	1.0936	1.1093
6.0	0.8945	1.0972	1.0846
7.0	0.8849	1.0690	1.0688
8.0	0.8761	1.0321	1.0583
9.0	0.8685	1.0501	1.0510
10.0	0.8620	1.0735	1.0457
11.0	0.8566	1.0575	1.0418
12.0	0.8521	1.0250	1.0388
13.0	0.8484	1.0350	1.0365
14.0	0.8454	1.0235	1.0346
15.0	0.8429	1.0345	1.0331
16.0	0.8410	1.0250	1.0319
17.0	0.8394	1.2793	1.0308
18.0	0.8381	1.2795	1.0300
19.0	0.8370	1.2797	1.0293
20.0	0.8362	1.2799	1.0287

Table 4.7 MF-DFA for CHF subjects for q varying from -20 to 20

q	chf201	chf202	chf203	chf204	chf205	chf206	chf207	chf208
-20.0	3.6906	3.7083	3.6924	3.6109	3.7510	3.7415	3.7135	3.6924
-19.0	3.6913	3.7135	3.7080	3.6161	3.7620	3.7510	3.7193	3.7080
-18.0	3.6921	3.7193	3.7349	3.6220	3.7748	3.7620	3.7259	3.7237
-17.0	3.6924	3.7259	3.7369	3.6285	3.7899	3.7748	3.7260	3.6821
-16.0	3.7080	3.7332	3.6731	3.6359	3.7080	3.7490	3.2009	3.6811
-15.0	3.7349	3.7415	3.7237	3.6243	3.7490	3.7239	3.7510	3.7339
-14.0	3.7369	3.7510	3.6821	3.6538	3.6542	3.7241	3.7620	3.7329
-13.0	3.6731	3.7620	3.6852	3.6648	3.6542	3.7230	3.7330	3.7330
-12.0	3.6762	3.7748	3.6724	3.6832	3.6542	3.7222	3.7322	3.7322
-11.0	3.3573	3.7899	3.1873	3.1873	3.1873	3.1873	3.1873	3.1873
-10.0	3.2053	3.7080	3.2053	3.2053	3.2053	3.2053	3.2053	3.2053
-9.0	3.2274	3.7302	3.2274	3.2274	3.2274	3.2274	3.2274	3.2274
-8.0	3.2549	3.6579	3.2549	3.2549	3.2549	3.2549	3.2549	3.2549
-7.0	3.2904	3.6936	3.2904	3.2904	3.2904	3.2904	3.2904	3.2904
-6.0	3.3376	3.6411	3.3376	3.3376	3.3376	3.3376	3.3376	3.3376
-5.0	3.4038	3.5076	3.4038	3.4038	3.4038	3.4038	3.4038	3.4038
-4.0	3.5030	3.5074	3.5030	3.5030	3.5030	3.5030	3.5030	3.5030
-3.0	3.3244	3.3737	3.3168	3.3083	3.3543	3.3637	3.3668	3.3168
-2.0	3.2113	3.2063	3.2109	3.1956	3.1879	3.1913	3.1793	3.2113
-1.0	3.1139	3.0041	3.0341	3.0231	3.0614	3.0314	3.0461	3.0121
0.0	0.0000	1.4061	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1.0	0.1317	1.3402	0.1873	0.5795	0.0068	0.0393	0.6907	0.5934
2.0	1.2564	1.3216	1.4287	1.4611	1.4616	1.3789	1.4493	1.3546
3.0	1.4073	1.3717	1.4937	1.4212	1.4715	1.4540	1.4525	1.3573
4.0	1.5601	1.3528	1.4189	1.4128	1.4324	1.4260	1.4101	1.4069
5.0	1.5393	1.3052	1.4197	1.4153	1.4319	1.4295	1.4127	1.4025
6.0	1.5334	1.3415	1.4258	1.4158	1.4379	1.4259	1.4157	1.4059
7.0	1.5120	1.3677	1.4199	1.4259	1.4386	1.4290	1.4191	1.4031
8.0	1.5220	1.2873	1.4120	1.4260	1.4057	1.4141	1.4125	1.4070
9.0	1.5229	1.3023	1.4231	1.4260	1.4244	1.4185	1.4055	1.4090
10.0	1.5262	1.3140	1.4373	1.4236	1.4099	1.4233	1.4180	1.3800
11.0	1.5116	1.3232	1.4406	1.4360	1.4152	1.4287	1.3800	1.3961
12.0	1.5160	1.3070	1.4443	1.4360	1.4211	1.4246	1.3817	1.3808
13.0	1.5121	1.3168	1.4435	1.4360	1.4214	1.4210	1.4030	1.3808
14.0	1.5100	1.3218	1.4444	1.4360	1.4372	1.4278	1.4040	1.3808
15.0	1.5162	1.3160	1.4445	1.4360	1.4079	1.3550	1.3848	1.3608
16.0	1.5137	1.3095	1.4445	1.4360	1.4302	1.4025	1.3854	1.3608
17.0	1.5116	1.3125	1.4445	1.4360	1.4172	1.3502	1.3859	1.3607
18.0	1.5599	1.3151	1.4445	1.4360	1.4298	1.3483	1.3863	1.3607
19.0	1.5583	1.3273	1.4445	1.4360	1.4197	1.3465	1.3866	1.3606
20.0	1.5356	1.4621	1.4445	1.4360	1.4070	1.3449	1.3869	1.3806

Tables 4.8, 4.9, 4.10, 4.11 and 4.12 listed below show the effect of length on MF-DFA for healthy and CHF subjects for q varying from -20 to 20

Table 4.8 Effect of Length on MF-DFA for 4000 Data Points

q	Healthy	CHF
-20.0	3.1114	3.1811
-19.0	3.1167	3.1817
-18.0	3.1226	3.1823
-17.0	3.1292	3.1829
-16.0	3.1367	3.1367
-15.0	3.1451	3.1451
-14.0	3.1547	3.1547
-13.0	3.1658	3.1658
-12.0	3.1788	3.1788
-11.0	3.1941	3.1941
-10.0	3.1812	3.2125
-9.0	3.1835	3.2350
-8.0	3.1863	3.2630
-7.0	3.1899	3.2991
-6.0	3.3473	3.2473
-5.0	3.3147	3.2215
-4.0	3.2158	3.2116
-3.0	3.2484	3.2184
-2.0	3.0212	3.0212
-1.0	2.7032	2.8032
1.0	1.2020	1.2395
2.0	1.1594	1.1504
3.0	1.2070	1.2570
4.0	1.2145	1.1715
5.0	1.2160	1.1726
6.0	1.2128	1.1720
7.0	1.2508	1.1714
8.0	1.2516	1.1710
9.0	1.2521	1.1708
10.0	1.2524	1.1707
11.0	1.2526	1.1706
12.0	1.2527	1.1705
13.0	1.2528	1.1705
14.0	1.2529	1.1705
15.0	1.2529	1.1705
16.0	1.2530	1.1705
17.0	1.2530	1.1705
18.0	1.2530	1.1705
19.0	1.2530	1.1705
20.0	1.2530	1.1705

Table 4.9 Effect of Length on MF-DFA for 8000 Data Points

q	Healthy	CHF
-20.0	3.1109	3.6109
-19.0	3.1161	3.6161
-18.0	3.1220	3.6220
-17.0	3.1285	3.6285
-16.0	3.1359	3.6359
-15.0	3.1443	3.6443
-14.0	3.1538	3.6538
-13.0	3.1648	3.6648
-12.0	3.1777	3.6777
-11.0	3.1941	3.6941
-10.0	3.1111	3.5111
-9.0	3.1334	3.5334
-8.0	3.1612	3.5120
-7.0	3.1970	3.5070
-6.0	3.0447	3.3447
-5.0	3.1115	3.4115
-4.0	2.6512	3.5118
-3.0	2.4788	3.3788
-2.0	2.2013	3.0129
-1.0	2.0152	3.0152
1.0	1.2122	1.5800
2.0	1.0184	1.3346
3.0	1.0709	1.3254
4.0	1.0470	1.3593
5.0	1.0609	1.3269
6.0	1.0632	1.3059
7.0	1.0645	1.2930
8.0	1.0652	1.2849
9.0	1.0657	1.2799
10.0	1.0660	1.2767
11.0	1.0662	1.2747
12.0	1.0664	1.2733
13.0	1.0665	1.2725
14.0	1.0666	1.2719
15.0	1.0666	1.2716
16.0	1.0667	1.2713
17.0	1.0667	1.2712
18.0	1.0667	1.2710
19.0	1.0668	1.2710
20.0	1.0668	1.2709

Table 4.10 Effect of Length on MF-DFA for 16000 Data Points

q	Healthy	CHF
-20.0	3.1083	3.7083
-19.0	3.1135	3.7135
-18.0	3.1193	3.7193
-17.0	3.1259	3.7259
-16.0	3.1332	3.7332
-15.0	3.1415	3.7415
-14.0	3.1510	3.7510
-13.0	3.1620	3.7620
-12.0	3.1748	3.7748
-11.0	3.1899	3.7899
-10.0	3.1080	3.7080
-9.0	3.1302	3.7302
-8.0	3.1579	3.6579
-7.0	3.0294	3.6936
-6.0	3.0411	3.6411
-5.0	3.0076	3.5076
-4.0	2.8740	3.5074
-3.0	2.4637	3.3737
-2.0	2.2006	3.2063
-1.0	2.1900	3.0041
1.0	1.1854	1.4061
2.0	0.9904	1.3402
3.0	1.0409	1.3216
4.0	1.0272	1.3717
5.0	1.0728	1.3528
6.0	1.0757	1.3052
7.0	1.0785	1.3415
8.0	1.0612	1.3677
9.0	1.0638	1.2873
10.0	1.0660	1.3023
11.0	1.0680	1.3140
12.0	1.0697	1.3232
13.0	1.0411	1.3070
14.0	1.0423	1.3168
15.0	1.0433	1.3218
16.0	1.0442	1.3160
17.0	1.0449	1.3095
18.0	1.0454	1.3125
19.0	1.0459	1.3151
20.0	1.0463	1.3273

Tables 4.9, 4.10, 4.11, and 4.12 reveal that, the values of exponents among healthy and CHF subjects vary considerably.

Table 4.11 Effect of Length on MF-DFA for 32000 Data Points

q	Healthy	CHF
-20.0	3.1065	3.7065
-19.0	3.1118	3.7118
-18.0	3.1176	3.7176
-17.0	3.1241	3.7241
-16.0	3.1314	3.7314
-15.0	3.1396	3.7396
-14.0	3.1491	3.7491
-13.0	3.1600	3.7600
-12.0	3.1728	3.7728
-11.0	3.1878	3.5878
-10.0	3.1206	3.5059
-9.0	3.1228	3.4228
-8.0	3.1256	3.4126
-7.0	3.0910	3.4091
-6.0	3.0338	3.3383
-5.0	3.0405	3.4045
-4.0	2.7504	3.5039
-3.0	2.6694	3.3694
-2.0	2.5005	3.1400
-1.0	2.2994	3.0994
1.0	1.2068	1.4403
2.0	1.0248	1.3451
3.0	1.0455	1.3681
4.0	1.0471	1.3450
5.0	1.0449	1.3543
6.0	1.0427	1.3537
7.0	1.0411	1.3519
8.0	1.0400	1.3504
9.0	1.0292	1.3494
10.0	1.0286	1.3487
11.0	1.0282	1.3483
12.0	1.0279	1.3481
13.0	1.0277	1.3479
14.0	1.0275	1.3478
15.0	1.0273	1.3477
16.0	1.0272	1.3477
17.0	1.0271	1.3076
18.0	1.0271	1.3076
19.0	1.0270	1.3076
20.0	1.0270	1.3076
20	1.0263	1.3075

Table 4.12 Effect of Length on MF-DFA for 64000 Data Points

q	Healthy	CHF
-20.0	3.1065	3.7065
-19.0	3.1118	3.6118
-18.0	3.1176	3.6176
-17.0	3.1241	3.6241
-16.0	3.1314	3.6415
-15.0	3.1396	3.6496
-14.0	3.1291	3.6491
-13.0	3.1000	3.6600
-12.0	3.1728	3.6728
-11.0	3.1878	3.5878
-10.0	3.1587	3.5587
-9.0	3.1279	3.5279
-8.0	3.1256	3.5255
-7.0	3.1291	3.5010
-6.0	3.0338	3.4448
-5.0	3.0454	3.4405
-4.0	3.0504	3.5049
-3.0	2.7694	3.4694
-2.0	2.4001	3.2001
-1.0	2.1994	3.1939
1.0	1.1268	1.4003
2.0	1.0476	1.3451
3.0	1.0555	1.3287
4.0	1.0571	1.3250
5.0	1.0549	1.3243
6.0	1.0527	1.3237
7.0	1.0511	1.3219
8.0	1.0500	1.3204
9.0	1.0292	1.3094
10.0	1.0286	1.3087
11.0	1.0282	1.3083
12.0	1.0279	1.3081
13.0	1.0277	1.3079
14.0	1.0275	1.3078
15.0	1.0223	1.3077
16.0	1.0222	1.3077
17.0	1.0211	1.3076
18.0	1.0201	1.3076
19.0	1.0203	1.3076
20.0	1.0899	1.3076

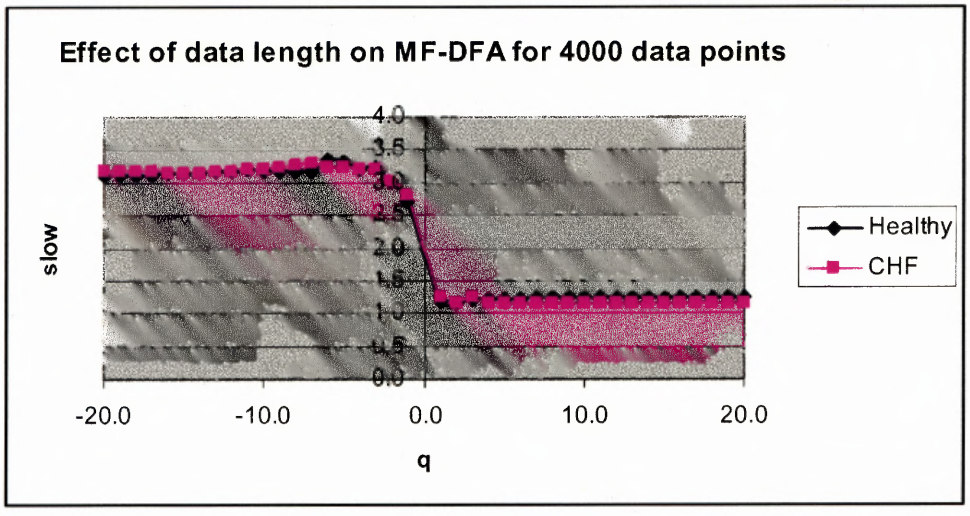

Figure 4.13 Effect of Length on MF-DFA for 4000 Data Points Note: It is not possible to distinguish between Healthy and CHF at 4000 Data Points.

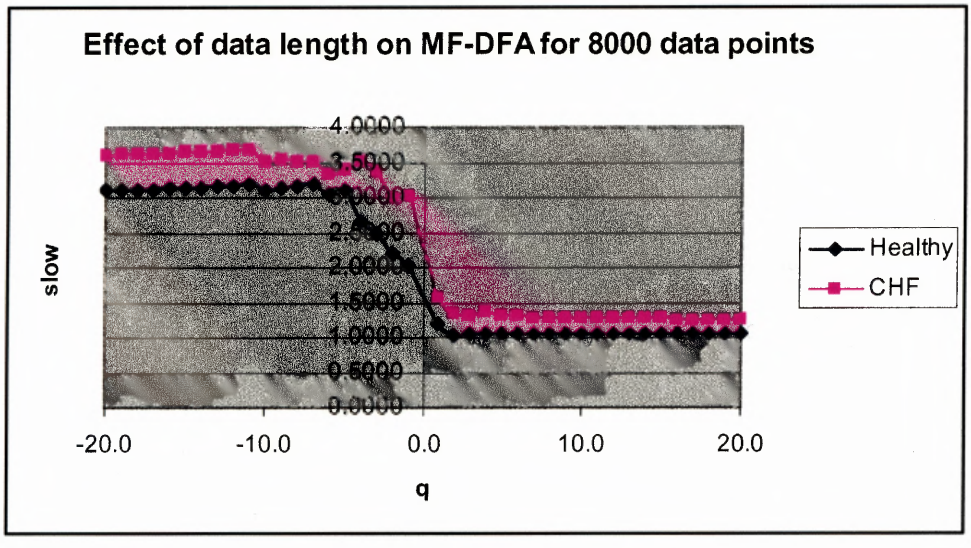

Figure 4.14 Effect of Length on MF-DFA for 8000 Data Points Note: It is possible to distinguish between Healthy and CHF at 8000 Data Points.

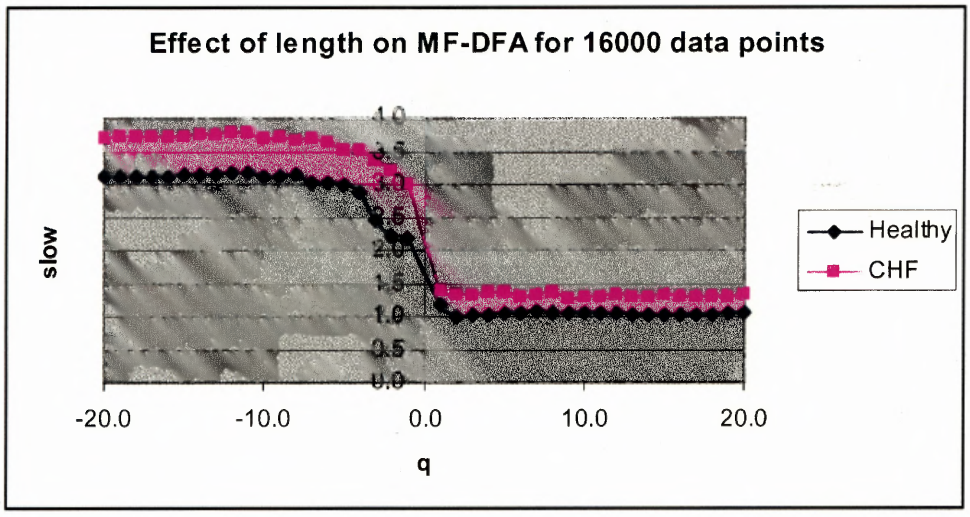

Figure 4.15 Effect of Length on MF-DFA for 16000 Data Points Note: It is possible to distinguish between Healthy and CHF at 16000 Data Points.

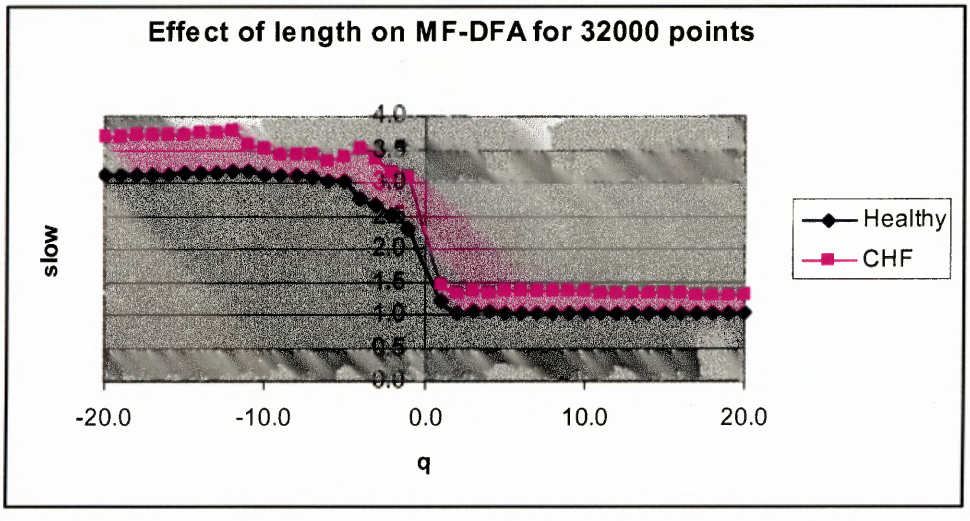

Figure 4.16 Effect of Length on MF-DFA for 32000 Data Points Note: It is possible to distinguish between Healthy and CHF at 32000 Data Points.

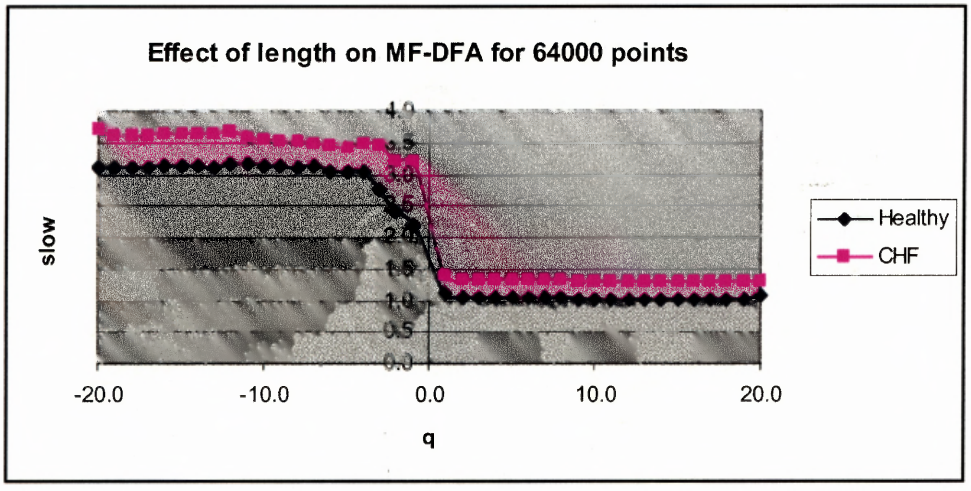

Figure 4.17 Effect of Length on MF-DFA for 64000 Data Points Note: It is also possible to distinguish between Healthy and CHF at 64000 Data Points.

Figures 4.13-4.17 show the effect of length on MF-DFA for different data points. Apparently the two groups of subjects (healthy and CHF) can be discriminated for 8000, 16000, 32000, and 64000 data points as seen in Figures 4.14-4.17. Thus it can be concluded that 8000 data points are sufficient enough to discriminate healthy subjects from subjects with cardiac heart failure.

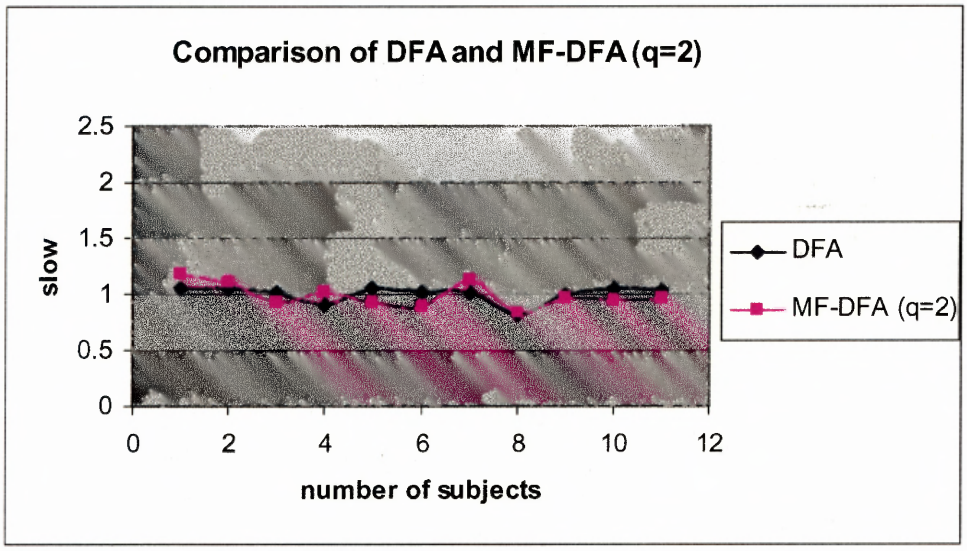

Figure 4.18 shows the Comparison of DFA and MF-DFA for q=2. Note that DFA and MF-DFA have nearly the same values of the scaling exponent for q=2.

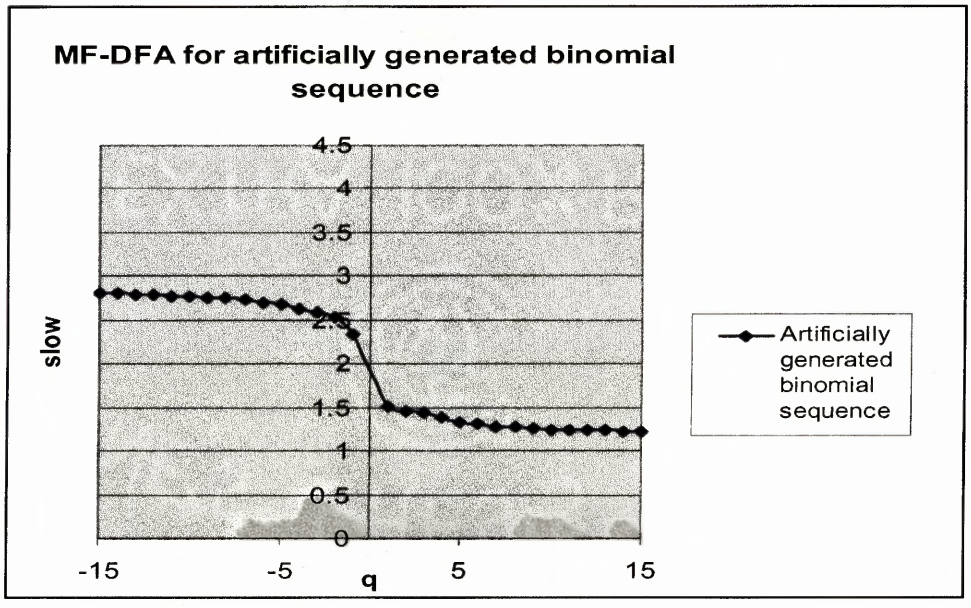

Figure 4.19 shows the plot of slow versus q for an artificially generated binomial multifractal sequence.

Binomial sequence is used because it is an artificially generated series that exhibits multifractality and thus is used as a reference to test the matlab code for MF-

DFA. The binomial sequence is artificially generated using the mathematics discussed in [27].

CHAPTER 5

CONCLUSION AND FUTURE WORK

In summary, two non-linear methods (DFA and MF-DFA) were used to analyze two different groups of subjects (healthy and subjects with cardiac heart failure). This study revealed the presence of long-range power law correlation for the group of healthy subjects while a breakdown in the long-range power law correlation for the subjects with cardiac heart failure. An α greater than 0.5 and less than or equal to 1.0 indicates long-range power law correlation, which means large interbeat intervals are more likely to be followed by large and vice versa. In contrast, $0<\alpha<0.5$ indicates a different type of power-law such that large and small values of time series are more likely to alternate. A special case of $\alpha = 1$ corresponds to 1/f noise. For $\alpha \geq 1$, correlations exist but cease to be of power law form; $\alpha = 1.5$ indicates Brownian noise, which is an integration of white noise.

The 'α' exponent can thus be viewed as an indicator that describes the "roughness" of the original time series. The larger the value of α, the smoother is the time series. Thus value of $\alpha=1.21$ for CHF, which is greater than 1.003 for healthy subjects indicates a smoother time series and thus, a pathologic condition because the larger is the heart rate variability, the healthier is the person. If the plot of LogF(n) versus Log(n) gives a straight line i.e. single value for α exponent is able to characterize the entire heart rate time series, than the given time series exhibits a monofractal behavior.

However, a crossover effect was observed as illustrated in Figure 4.1, which shows the presence of two different scaling exponents (slow and fast) for each group of subjects

(Healthy and CHF). This suggested that there is more than one exponent values needed to characterize the heart rate time series, which can be obtained using multifractal DFA.

A scatter plot of slow v/s fast for healthy and CHF subjects in DFA reveals that data from normal interbeat interval series are tightly clustered suggesting that there may exists a "universal" scaling behavior for physiological time series. In contrast, the pathologic data show more variation, which may be related to different clinical conditions and varying severity of the pathologic states.

Figure 4.17 shows the plot of q v/s slow for MF-DFA for 64000 data points. The value of slow for $q>0$ for healthy is 1.04 ± 0.02 and for CHF is found to be 1.32 ± 0.02. The value of slow for $q<0$ for healthy is 3.01 ± 0.26 and for CHF is found to be 3.53 ± 0.14 (mean value $\pm$ S.D.). Value of slow for $q>0$ is less than that for $q<0$. And for $q=2$ MF-DFA retains monofractal DFA as illustrated in Figure 4.18.

Thus, MF-DFA is clearly able to discriminate among the healthy and CHF for $q<0$ as well as for $q>0$. Both the methods were tested for the effect of length and the results suggests that, 4000 data points are needed for DFA analysis because results become inconsistent for data points less than 4000. MF-DFA suggests 8000 data points (~2 hours) are sufficient for the analysis. Therefore these methods cannot be applied to relatively short time series. There is no effect of activity, effect of non-stationarity, and effect of trend found on DFA. DFA and MF-DFA were able to discriminate 23 Healthy subjects out of 26 Healthy subjects data set i.e. true positive specificity is 0.89 and false negative specificity is 0.12 and 9 CHF subjects out of 11 CHF subjects data set i.e. true positive specificity is 0.82 and false negative specificity is 0.19.

As a further extension to this research, DFA and MF-DFA analysis methods should be used on various pathologic conditions other than CHF and determine how well the methods are able to discriminate among healthy and pathologic subjects with various other pathologies like atrial fibrillation, myocardial infarction and many more. I would like to test how well these methods are able to detect the effect of drugs in different pathologic conditions. The data sets that used for this experiment includes healthy subjects aging between (28.5 to 76 years) and CHF subjects aging between (34 to 79 years). Thus, a study should be conducted that includes data set from young subjects aging between (18 to 24 years) and check how well DFA and MF-DFA discriminate among healthy and CHF subjects. Also comparison among male and female group of subjects should be done to check the underlying non-linear dynamics prevailing among these group of subjects. These methods should be used as a predictor of pathologic conditions i.e. in this experiment it was known that the data obtained was from subjects with cardiac heart failure. However, future study should be designed such that it examines the ability of these methods to predict the onset of diseased conditions.

APPENDIX

Matlab Code for DFA and MF-DFA Analysis

This appendix contain the code used to generate the scaling exponents for DFA analysis

Dfavalue.m

```
load filename.txt;
vals=filename(3:4000);
newdata=deglitching(vals);
[slow,fast]=dfapeng(newdata);
```

The below mentioned files supports the DFA code

Dfamain.m

```
function [num, resid] = dfa(data, smallestbox, largestbox, step)
%       [num, resid] = dfa(data, smallestbox, largestbox, step)
%       dfa finds the detrended fluctuations
%     data       ... input data
%     smallestbox ... smallest window size to use
%     largestbox  ... largest window size to use
%     step       ... stepsize
dcount = 0
N = 0;
num = 0;

if nargin < 2
  smallestbox = 4;
  largestbox = 0;
  step = 1.4142136;
end
N = length(data)
idata = dfahelpr(data);
if largestbox == 0
  largestbox = N;
elseif largestbox < 0
  largestbox = N/4;
end
i = smallestbox;
j = 1;
while floor(i) <= largestbox
%  resid(j) = log10(Ardf(floor(i),idata));
%  num(j) = log10(floor(i));
 resid(j) = (ardf(floor(i),idata));
 num(j) = (floor(i));
  i = i*step;
  j = j+1;
end
```

DFApeng.m

```
function [slow,fast] = dfapeng(vals,slowlen,midlen,fastlen)
% [slow,fast] = dfapeng(vals,slowlen,midlen,fastlen)
% <vals> -- the data, typically RR intervals
% <slowlen> -- a slow time scale, 64 beats by default
% <midlen>  -- a medium time scale, 16 beats by default
% <fast>    -- a fast time scale, 4 beats by default
% Returned values:
% <slow>    -- the power-law slope on the slow to mid time scales
% <fast>    -- the power-law slope on the mid to fast time scales
if nargin < 4
  fastlen = 4;
end
if nargin < 3
  midlen = 16;
end
if nargin < 2
  slowlen = 64;
end
[num,resid] = dfamain( vals, fastlen, slowlen, sqrt(2) );
% fit the power-law of the faster time scales
goods = find( num >= fastlen & num <= midlen);
p = polyfit( log(num(goods)), log(resid(goods)), 1 );
fast = p(1);
% fit the power-law of the slower time scales
goods = find( num >= midlen & num <= slowlen);
p = polyfit( log(num(goods)), log(resid(goods)), 1 );
slow = p(1);
```

DFA helpr

```
function intdata = int_data(data)
%         intdata = int_data(data)
%         int_data integrates
%       data    ... input data
%       intdata ... integrated data
sum = 0;
data = data - mean(data);
for i = 1:length(data)
   sum = sum+data(i);
   intdata(i) = sum;
end
```

Deglitching: This function was used to clean certain data points that were too high or too low to be considered for the code.

Deglitching.m

```
function datum=deglitching(vals)
[a,b]=size(vals);
times(1)=vals(1);
for i=2:a
  times(i)=times(i-1)+vals(i);
```

```
end
labels=ones(a,b);
[labs,resids] = ardeglch(vals, labels);
c=1;
for i=1:a
  if labs(i)==1
    datum(c,1)=vals(i);
   c=c+1;
 end
end
```

ardeglch.m

```
function [labs,resids] = ardeglch(data, inlabels, modelorder, iqrcrit)
% labs = arresid(data, inlabels, modelorder, iqrcrit)
% identifies outliers in a time series by looking at the residuals of a forward and backward AR fit.
% Excludes from the fit beats labeled 0 in <inlabels>
% iqrcrit gives the criteria in inter-quartile range units for an outlier.
% Returns a vector of the same length as the data which contains a 0 for any beat marked as bad
either in inlabels or from the AR fit.
% [labs,resids] = ardeglch...
% gives the actual values of the residuals as an optional second argument
if nargin < 3
  modelorder = 3;
end


if nargin < 4
  iqrcrit = 3.5;
end
% fit the forward model
labforward = arresid(data, inlabels, modelorder);
% fit the backward model
labbackward = arresid( data((length(data)):-1:1), inlabels((length(data)):-1:1), modelorder);
% take the smaller of the two residuals, remembering
% to put labbackward back in forward order.
labels = min(labforward, labbackward(length(data):-1:1) );
resids = labels;
% Compute the interquartile range and limits for outliers.
lims = prctile(labels,[25 50 75]);
iqrange = lims(3) - lims(1);
bottom = lims(1) - iqrange*iqrcrit;
top = lims(3) + iqrange*iqrcrit;
% bogus points are marked as 666666 in <labels> or as 0 in <inlabels>
labs = (labels > bottom & labels < top & labels ~= 666666 & inlabels ~= 0 );
```

The following file is used to implement multifractal DFA algorithm

mfdfa.m

```
vals = nsr(3:length_size);
%   q=-2;
 for q = -20:20;
 if q ~= 0
```

```
% q=-2;
[slow,fast]=dfapengmul(vals,q);
  slow1(q+21) = slow;
  fast1(q+21) = fast;
  end
 end
 slow1 = fliplr(slow1);
 fast1 = fliplr(fast1);
```

The below mentioned files supports the code for multifractal DFA.

```
function [slow,fast,resid,num] = dfapengmul(vals,q,slowlen,midlen,fastlen)
% function [slow,fast] = dfapengmul(vals,slowlen,midlen,fastlen)
% [slow,fast] = dfapeng(vals,slowlen,midlen,fastlen)
% <vals> -- the data, typically RR intervals
% <slowlen> -- a slow time scale, 64 beats by default
% <midlen>  -- a medium time scale, 16 beats by default
% <fast>    -- a fast time scale, 4 beats by default
% Returned values:
% <slow>    -- the power-law slope on the slow to mid time scales
% <fast>    -- the power-law slope on the mid to fast time scales
if nargin < 5
 fastlen = 4;
end
if nargin < 4
 midlen = 16;
end
if nargin < 3
 slowlen = 64;
end
[num,resid] = muldfa( vals, fastlen, slowlen, sqrt(2),q );
% fit the power-law of the faster time scales
goods = find( num >= fastlen & num <= midlen);
p = polyfit( log(num(goods)), log(resid(goods)), 1 );
fast = p(1);
% fit the power-law of the slower time scales
goods = find( num >= midlen & num <= slowlen);
p = polyfit( log(num(goods)), log(resid(goods)), 1 );
slow = p(1);
muldfa.m
function [num, resid] = muldfa(data, smallestbox, largestbox, step, q)

%       [num, resid] = dfa(data, smallestbox, largestbox, step)
%       dfa finds the detrended fluctuations
%       data       ... input data
%       smallestbox ... smallest window size to use
%       largestbox  ... largest window size to use
%       step       ... stepsize
dcount = 0;
N = 0;
num = 0;
N = length(data);
q
if nargin < 3
```

```
 smallestbox = 3;
 largestbox = floor(N/4);
 step = 1.414;
end
N = length(data);
idata = dfahelpr(data);
if largestbox == 0
 largestbox = (floor(N/4));
elseif largestbox < 0
 largestbox = (floor(N/4));
end
i = smallestbox;
j = 1;
%  q=-2;
while floor(i) <= largestbox
resid(j) = log10(multifractaldfa(floor(i),idata,q));
  num(j) = log10(floor(i));
%  resid(j) = (multifractaldfa(floor(i),idata,q));
% resid(j) = dfa(floor(i),idata);
 num(j) = (floor(i));
 i = i*step;
 j = j+1;
end
%plot(num,resid)
%axis([0 4 -4 0]);
```

Multifractaldfa.m

```
%data=xlsread('nsdata05.xls',1,'a1:a10000');
%idata = dfahelpr(data);
%q=2;
function resid = multifractaldfa(boxlength,intdata,q)
%load multifractaldata.txt;
%vals=multifractaldata(3:8000);
%intdata=dfahelpr(vals);
%         resid = Ardf(boxlength,intdata)
%         Ardf calculates the average rms of the detrended fluctuation
%       intdata ... integrated input data
%       resid   ... result
%s=64;
%boxlength=s;
%q=-10;
s=boxlength;
N = length(intdata);
%N = length(idata);
Ns=floor(N/s);

bb = 0;
for k = 0:boxlength:(Ns-1)
   datausedcount = k+boxlength;
   ypoint = int_data((k+1):datausedcount); % divides the time series into segments
                                %ypoint = idata(k+1:k+boxlength);
   ypoint = detrend(ypoint);            % detrend: subtracts
   vari=(sum(ypoint.^2))/s;
                                 %    vari=(sum(ypoint.^2))
                                %     bb = bb+(1/((vari)^(q*0.5)));
```

```
    bb_temp(k+1) = (1/((vari)^(q*0.5)));
    clear ypoint;
end
temp1 = max(bb_temp);
temp2 = min(bb_temp);
scale_val = (temp1+temp2);
bb_temp = bb_temp./scale_val;
bb1 = bb;
bb = sum(bb_temp);
cc=0;
for l=0:boxlength:(Ns-1)
    datausecountrev=l+boxlength;
                                    %ypoint=idata(N-l-boxlength+1:N-l);
    ypoint=int_data((N-datausecountrev+1):(N-l));
    ypoint=detrend(ypoint);
    varir=(sum(ypoint.^2))/s;
                                    %    varir=(sum(ypoint.^2))
    cc_temp(l+1)=(1/((varir)^(q*0.5)));
end
cc_temp = cc_temp./scale_val;
cc1 = cc;
cc = sum(cc_temp);
                        % resid = 1/(((cc+bb)/(2*Ns))^(1/(q)));
resid = (1/(cc+bb)^(1/(q)));
resid = resid/(2*Ns);
resid = resid./(scale_val^(1/(q)));
%disp(resid);
```

Loop_fun.m

This file when run generates multitude of exponents for q varying from -20 to 20

```
data=['filename.txt'];
celldata = cellstr(data);
strings = char(celldata);
for i = 1:1
    nsr = feval(@load,strings(i,:));
    length_size = 33000;
    j=1;
        while length_size <= 98000
            newmfdfa1;
            slow_result(j,:) = slow1;
            fast_result(j,:) = fast1;
            length_size
            i
            length_size = length_size+length_size
            j = j+1
        end
            result = [slow_result; fast_result]';
%               xlswrite('muldfa_healthy',result,i+3);
            xlswrite('resultforplotmfdfa002',result,i);
        clear result
end
```

REFERENCES

1. http://training.seer.cancer.gov/module_anatomy/unit7_1_cardvasc_intro.html Retrieved on 10 May 2005.

2. Sherwood Lauralee, "Human Physiology from Cells to Systems," 2001.

3. www.sciencebob.com/ lab/bodyzone/heart.html Retrieved on 23 May 2005.

4. M. M. Wolf, G. A. Varigos, D Hunt, J. G. Sloman, "Sinus arrhythmia in acute myocardial infarction,", Med J Aust., 1978; 52-53.

5. "Heart Rate Variability: Standard of Measurement, Physiological Interpretation and Clinical Use," European Heart Journal vol. 17, pp. 354-381, 1996.

6. http://www.stjohnsmercy.org/healthinfo/adult/cardiac/electro.asp Retrieved on 10 June 2005

7. M. Malik, "Heart rate variability: Standards of Measurement, Physiological Interpretation and clinical use – Special Report," Circulation, 1996; 1043-1065.

8. A. L. Goldberger, D. R. Ridney, B. J. West, "Chaos and fractals in human physiology," Science American Journal, 1990; 262(2): 42-49

9. J. Bassinthwaighte, L. Liebovitch, and B. J. West, "Fractal Physiology," Oxford University Press, 1994

10. M. Kobayashi and T. Musha, "1/f fluctuation of Heartbeat Period," IEEE Transaction of Biomedical Engineering, 1982; 29:6-12.

11. J. B. Johnson, "The Schottky effect in low frequency circuits," Physical Review, 1925; 26: 71-85.

12. http://www.dchaos.com/portfolio/dchaos1/new_nonlinear_man_article.html Retrieved on 30 May 2005.

13. R.M. May, "Simple mathematical models with very complicated dynamical behavior," Nature 1976; 261:459-467.

14. http://reylab.bidmc.harvard.edu/DynaDx/tutorial/cop/lecture.html Retrieved on 23 June 2005.

15. Denton et al. "Fascinating rhythm: A primer on chaos theory and its application in cardiology," American Heart Journal; 1990:1419-1440.

16. B. Per, T. Chao, and W. Kurt "Self organized criticality: An explanation of 1/f noise," Physical Review Letter; 1987: 59: 381-384.

17. C.-K. Peng, S. Havlin, H. E. Stanley, and A. L. Goldberger, "Quantification in scaling exponents and cross over phenomenon in nonstationary heartbeat time series," Chaos 5; 1995:82-87.

18. C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, "Mosaic Organization of DNA nucleotides," Physical Review E; 1994: 49:1685.

19. P. Ivanov, A. Bunde, L. A. N. Amaral, S. Havlin, J. Fritsch-Yelle, R. M. Baevsky, H. E. Stanley, and A. L. Goldberger, "Macroscopic Finite Size Effects in Relaxational Processes," Europhys. Letter; 1999:48:594.

20 S. Buldyrev, N. V. Dokholyan, A. L. Goldberger, S. Havlin, C.-K. Peng, H. E. Stanley, and G. M. Viswanathan, "Analysis of DNA sequences using methods of statistical physics," Physica A; 1998:249:430.

21 J. Walleczek, "Self-organized Biological Dynamics and Nonlinear Control," Cambridge University Press, 2000.

22 E. W. Montroll and M. F. Shlesinger, "From Stochastics to Hydrodynamics," North-Holland, Amsterdam, 1984.

23 S. Havlin, R. B. Selinger, M. Schwartz, H. E. Stanley, and A. Bunde, "Random Multiplicative Processes and Transport in Structures with Correlated Spatial Disorder," Physical Review Letter; 1998:61:1438-1441.

24. S. V. Buldyrev, A. L. Goldberger, S. Havlin, C-K. Peng, H. E. Stanley, and M. Simons, "Fractal Landscapes and Molecular Evolution: Modeling the Myosin Heavy Chain Gene Family," Biophysics Journal; 1993:65:2675-2681.

25. J C Echeverr, B R Hayes-Gill, J A Crowe, M. S. Woolfson and G. D.Croaker, "Detrended fluctuation analysis: a suitable method for studying fetal heart rate variability," Physiol. Meas.; 2004:25:763–774.

26. J. Kantelhardt et al., "Multifractal detrended fluctuation analysis of nonstationary time series," Physica A; 2002: 316: 87-114.

27. J. Feder, "Fractals," Plenum Press, New York, 1988.

28. H.O. Peitgen, H Jurgens, D. Saupe, "Chaos and Fractals," Springer, New York, 1992

29. C.K.Peng et al., "Long-range Anticorrelations and Non-Gaussian Behavior of the Heartbeat," Physical Review Letter; 1993:70:1343-1346.

30. http://astronomy.swin.edu.au/~pbourke/fractals/noise/ Retrieved on 29 June 2005

31. http://classes.yale.edu/fractals/CA/OneOverF/BrownianNoise.html Retrieved on 29 June 2005

32. Plamen et al., "From 1/f noise to multifractal cascades in heartbeat dynamics," Chaos; 2001:11:3.

33. K Hu, PCh. Ivanov, Z Chen, P Carpena, H E. Stanley, "Effects of trends on Detrended Fluctuation Analysis," Physical Review E; 2001:64.