Durham E-Theses

Essays on the nonlinear and nonstochastic nature of stock market data

Vorlow, Constantine Euripides

How to cite:

Vorlow, Constantine Euripides (2002) *Essays on the nonlinear and nonstochastic nature of stock market data*, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/4154/

Use policy

THE UNIVERSITY OF DURHAM

SCHOOL OF ECONOMICS, FINANCE AND BUSINESS

Essays on the Nonlinear and Nonstochastic Nature of Stock Market Data

Constantine Euripides Vorlow

Submitted for the Qualification of Ph.D. in Finance

September 2002

1 8 JUN 2003

Supervisor: Prof. Antonios Antoniou

External Examiners: Prof. Lucio Sarno and Dr. Turalay Kenc

CONTENTS

LIST OF TABLES

LIST OF FIGURES

x

ACKNOWLEDGMENT

ABSTRACT

The nature and structure of stock-market price dynamics is an area of ongoing and rigourous scientific debate. For almost three decades, most emphasis has been given on upholding the concepts of Market Efficiency and rational investment behaviour. Such an approach has favoured the development of numerous linear and nonlinear models mainly of stochastic foundations. Advances in mathematics have shown that nonlinear deterministic processes i.e. "chaos" can produce sequences that appear random to linear statistical techniques. Till recently, investment finance has been a science based on linearity and stochasticity. Hence it is important that studies of Market Efficiency include investigations of chaotic determinism and power laws. As far as chaos is concerned, there are rather mixed or inconclusive research results, prone with controversy. This inconclusiveness is attributed to two things: the nature of stock market time series, which are highly volatile and contaminated with a substantial amount of noise of largely unknown structure, and the lack of appropriate robust statistical testing procedures. In order to overcome such difficulties, within this thesis it is shown empirically and for the first time how one can combine novel techniques from recent chaotic and signal analysis literature, under a univariate time series analysis framework. Three basic methodologies are investigated: Recurrence analysis, Surrogate Data and Wavelet transforms. Recurrence Analysis is used to reveal qualitative and quantitative evidence of nonlinearity and nonstochasticity for a number of stock markets. It is then demonstrated how Surrogate Data, under a statistical hypothesis testing framework, can be simulated to provide similar evidence. Finally, it is shown how wavelet transforms can be applied in order to reveal various salient features of the market data and provide a platform for nonparametric regression and denoising. The results indicate that without the invocation of any parametric model-based assumptions, one can easily deduce that there is more to linearity and stochastic randomness in the data. Moreover, substantial evidence of recurrent patterns and aperiodicities is discovered which can be attributed to chaotic dynamics. These results are therefore very consistent with existing research

indicating some types of nonlinear dependence in financial data. Concluding, the value of this thesis lies in its contribution to the overall evidence on Market Efficiency and chaotic determinism in financial markets. The main implication here is that the theory of equilibrium pricing in financial markets may need reconsideration in order to accommodate for the structures revealed.

CHAPTER 1

Introduction

1.1 Chaos and Markets

The term "*Chaos*" comes from the Greek word "$\chi \acute{a}o\varsigma$" and in its original context is used to signify a complete lack of order. In mathematics it has a deeper meaning. Chaos is used to characterise *nonlinear deterministic dynamical systems*. The mathematical theory of *Chaos*[1] deals with systems that exhibit nonlinearities, determinism, non-periodic cycles and strong imbalances. In such systems, violently oscillating behaviour can occur even for very small perturbations in their initial states. As a paradigm, it is very popular in physical sciences as the phenomena there reveal a plethora of characteristics consistent with chaos. Moreover, chaos in the sense of *nonlinear determinism* can be detected easily in such sciences as the processes examined are usually known or controlled and the outside factors, natural or artificial, can be easily determined most of the times. Thus any external processes that may be generating noise contaminating experiment measurements, can normally be determined a priori. This makes research work more easy as the nonsystematic noise components can be identified and filtered. Usually, the nonstochastic nature of many physical or physiological phenomena, such as EEG signals and earthquake measurements, can be revealed successfully even with large signal to noise ratios in the original measurements.

Social sciences and especially Economics or Finance, are a different case. The structures of the phenomena observed there are "assumed". The assumptions are based on the notion of a well defined equilibrium relationship and logical premises which should ensure stability. Balance is ensured within a "rational" context and usually under the protection of an all-important "*ceteris paribus*" clause. Major contributing factors here are the decision rules of individual agents. Their attitude is

---

[1]Chaos was developed by many different independent efforts. See Gleick's book (1987) where all the issues regarding this theory are covered in a non-technical way. The book by Prigogine and Stengers (1984) is also an excellent reading and has become one of the recent science bestsellers in this area.

assumed to abide to certain expected utility axioms. For example, financial market models were designed to reside upon a convenient framework based on market clearing conditions owing themselves to the *Rational Expectations Hypothesis*. Adam Smith's thinking that all agents in an economy will try individually to improve their position and their collective efforts would lead to a kind of overall optimum, seems to have influenced greatly economic thought in this case. In our example of the stock market, all those who participate, say the traders, are assumed to be "rational". According to the *Efficient Market Hypothesis* (EMH), their beliefs about the market and about each other as well, will be reflected in the market prices as all available information for a homogeneous commodity, is common knowledge. As a result, all deviations from rational expectations would be reflected in market prices and immediately counterbalanced by the trading community. After almost three decades of continuous debate on the EMH issue and a range of often utterly unanticipated market shocks, the EMH remains still a dubious, not to say a controversial hypothesis. Especially after the market crash of 1987, many research papers have focused on the issue of whether the weak and semi-strong form of the EMH should hold or not.

The establishing of the EMH has necessarily been based on "linear thinking". During the 60s and 70s, most of the models devised to explain stock market dynamics were based on linear premises. The tests used on price data, in order to explain generating processes, were themselves linear as well. Most of these were evolving around explorations of linearity based on first and second moments and were mainly tests of independence and randomness.[2] Successive non-random price changes would allow for trading rules leading to above-normal earnings and hence the EMH would not hold. The weak form of EMH was put under scrutiny by tests of randomness and the semi-strong form by market anomaly studies.

The EMH, although very attractive as a theory, did not provide any explanations on what exactly is the significance for the stock market as a system to have all

---

[2]See earlier works of Working (1934, 1960a, 1960b, 1962, 1963, 1947a and 1974b also in Working and Peck 1977), Samuelson (1965) and Fama (1965) and the excellent book by Granger and Morgenstern (1970). See also Cowles (1933,1937 and 1960). Most of the earlier and seminal works can be found in the famous collection by Cootner (1964).

relevant information reflected on the prices of the securities. The answer was given by the development of the *Asset Pricing Theory*. Its foundations were the *modern portfolio theory* and *the Capital Asset Pricing Model* also known as "CAPM". Although we will discussing these more extensively in chapter 2, we should mention here that their theoretical premises were to be put to the test alongside the EMH. Moreover, given that most risk strategy policies have their foundations within these theories, it is inevitable that any proof against the EMH, would have serious implications on how agents perceive and account for financial risk.

1.2 Information and noise

Information, even when freely disseminated, is not always accurate and free from irrelevant to the issue news. In most sciences, these are termed as *"noise"*. According to Black (1986), *"Noise makes financial markets possible, but also makes them imperfect"*. *Informational efficiency* describes a market situation in which all security related information is reflected on the price of the security. There are no implications though about the context and the quality of the information itself. In this sense, the information that is contained in security market prices may or may not be related to their fundamental values. The notion then of *fundamental efficiency* is a more precise proposition as this certifies that prices are indeed indicators of the intrinsic value of a security. The fundamental empirical research question is then posed: *How may one distinguish between noise and information in such a way that information on the fundamental value is obtained and used?* Another way of putting this question is *what is the nature and the quality of noise?* Unfortunately, in social sciences one may only speculate on these issues as noise and information is not a physically observable quantity. In physics or engineering for example, many experiments deal with matter or energy which can be measured accurately even when their processes are not always free from measurement errors and random, unanticipated factors-events. Knowledge of physical laws and careful calibration of the instruments may enable the researcher to approximate extremely well the noise processes and "filter them out".

In Finance particularly, there are two major problems. Firstly, the measu-

rement functions themselves are not well defined. We can only assume that the prices reflect all socioeconomic activity that generates them. Secondly we have no real or exact knowledge of the laws behind the dynamics of the processes observed and termed as "financial markets". In other words, we anticipate that all system dynamics are captured by the pricing mechanisms and reflected fully in the observable quantities such as closing prices, volumes etc. etc. There is no way of a-priori establishing the informational quality of these observations or the nature and causes of the noise, if this indeed exists within the measurements. Above these issues, there is often no clear-cut indication of how natural, liquidity or regulatory factors restrain the dynamics and affect the equilibrium positions of markets. For these issues we can only speculate. Frequently research will hide behind the protective cape of stochasticity and randomness and according the law of large numbers, the great assumption will be that eventually, all non-systemic factors should fade away on average, revealing the true dynamics of the markets.

Due to necessity, various models have been devised through the years to account for most the matters we discussed so far. Based on *noise-theory* and empirical evidence, economists have tried to explain if and how or why markets are infected by unrelated information to fundamental asset values. Models such as the ones referring to *"noise trading"* and ideas such as *"irrational exuberance"*[3], attempt to show how contaminating information due to ill-informed or non-rationally behaved investors may cause feedback and frictions to the system that can not be arbitraged away in the EMH fashion. In his early "Noise" paper, Black (1986) provides an excellent, very insightful discussion of these issues. As Black analyses his ideas, one discerns in these something very close to the notion of chaos although they were never identified as such.

1.3 Complexity and Chaos

Financial markets as socioeconomic phenomena or systems, are undoubtedly very "complex" (Anderson and Arrow, 1988). There is no globally agreed definition

---

[3]In Shiller (2000): *"Was Federal Reserve Chairman Alan Greenspan right when he referred to current investor behaviour as "irrational exuberance" in a 1996 speech?"*.

of complexity. According to Peters (1999a), they exhibit "*global structure and local randomness*". Structure leads to strength, randomness leads to "innovation and resilience". Chaos and complexity are not exactly synonyms. Day (1994) provides a broader definition. According to his, complex dynamics will not tend to a limit cycle, fixed point or explode. their behaviour may be discontinuous and may be sufficiently described by a set of nonlinear differential or difference equations that may also contain stochastic terms. Both Peter's and Day's definitions allow for chaos and its predecessor, *catastrophe theory*, developed by the French mathematician René Thom in 1975.

So why one would accept complexity or chaos as a valid paradigm for financial dynamics? An important point here is that complex-chaotic systems react more efficiently to outside stimulus and have the ability to self organise. This is what makes the search for chaos in finance an interesting one. The complexity is reflected in their observable output and financial time series are such an output. Empirical investigation has revealed a wealth of levels of complexity, structure and randomness which have not been adequately or convincingly explained until today.

Due to the significant amount of feedback, the dynamics of financial market systems are assumed to be influenced by a multitude of interacting agents in a highly nonlinear fashion. This alone is a sufficient reason for complexity. And chaotic, unlike stochastic processes, allow for such complexity. Unfortunately, standard linear methodologies will not pick up these processes or will misinterpret them for randomness. Not surprisingly, early research on prices and returns has reported small autocorrelations (linear feedbacks) and weak nonstochastic structures. This led to the initial support of the market efficiency notion. Original sequences were often revealed to be stationary and random, individually identically distributed. Since then, the continuous exploitation of arbitrage opportunities led to increased scaling in measurements (high frequency tick by tick data), and availability of information on various levels. The new data revealed novel structures. Today the research focuses on *extreme events, scaling* or *power-laws, heavy-tailed* distributions, *chaos* and *fractals, long memory, switching regimes* and *asymptotic stationarity* such as in the ARCH-GARCH type of models. During the last ten years or so, tools from the

area of statistical physics have been employed for the investigation of the dynamics of financial time series. The combination of ideas from economics and physics has lead to the development of the new term "econophysics". Conferences are held with this context, books are written and physics departments hold tutorials and workshops on "phynance."[4] Moreover, numerous papers have appeared in *Physica A* and other interdisciplinary journals identifying the need to provide plausible explanations to the complexity of financial phenomena. These facts show the concerns of the academic community about the issues discussed here as well as the necessity to break free from conventional modelling approaches and investigate financial dynamics with every available tool. Especially when the observables from the financial markets are claimed to resemble the complexity and dynamics of some phenomena in the physical sciences.

1.4 Mandelbrot, chaos and finance

The EMH implicitly assumes investor rationality. However, many markets, especially the emerging ones, are characterised by nonrational in the sense of *non risk averse* investment behaviour. This kind of behaviour may be the cause of market imperfections that lead to nonlinearities and feedback in the price generating processes. Benoit Mandelbrot, a student of Paul Lévy, initially wrote a number of original papers between 1962 and 1972 (see collections Mandelbrot, 1999b and 1997a) which have been by and large ignored and he eventually was treated as an outsider. He was the first to identify that prices did not exhibit any continuity and that there was substantial volatility clustering. Patterns seemed to change continuously and the root mean squared deviations did not seem to stabilise asymptotically. This was a clear indication that stochastic Brownian motion was not an adequate representation for the dynamics of such sequences. Using elements of the "stable Paretian" theory, he deduced that price changes were not Gaussian, but *"fat-tailed"*, with numerous outliers. Following these findings he advanced to search for cyclical and long memory structures that could explain variants of Brownian motions for

---

[4]See for example conferences: http://felix.physics.sunysb.edu/UG/phy68004.html, or http://www.unifr.ch/econophysics/principal/events/dublin_99.html or the book by Thomas Baas (2000) "*The Predictors: Adventures in High Phynance*".

price-fluctuation dynamics. Most of this history is beautifully described in Mirow-ski's 1990 paper (see also Mirowski, 1989). Since then, the advances in information technology, have enabled researchers to employ tools that could find what Mandel-brot saw in the data. The results were not conclusive as we shall see in the next chapter.

Mandelbrot (1999c) made a comeback in 1999 with his Physica A article "*Re-normalisation and fixed points in finance, since 1962*". A modified version appeared in Scientific American. There he discusses that financial time series may have a kind of a "fractal" nature. He has demonstrated that series can be simulated that appear as the original financial returns sequences and exhibit the same stylised facts, but are not exactly stochastic. Mandelbrot discusses his idea about a "*multifractal*" structure. The skeptical would point out that he has used "reverse engineering" without providing a theoretically sound framework that allows for this kind of stru-cture. His point though, well taken, is that he proposes this as a means of estimating the "probability of what the market might do" rather than forecasting a financial shock.

On the opposite side of the issue is Granger (1994) and Malliaris and Stein (1999). They do not readily accept chaos and fractality as a plausible explanation for market dynamics. Granger's skepticism is based on the fact that long-term forecasting is not possible under the deterministic chaos hypothesis. Malliaris and Stein show through Bayesian statistical analysis that a model can be formulated that allows for types of noise consistent with a GARCH generating processes and random price fluctuations but not necessarily deterministic chaos.

Concluding, it seems that the question of deterministic-chaotic financial dy-namics, as posed by Mandlebrot's work and his followers, is an open issue that has generated considerable debate and led to no conclusive result so far. The nature of this question is also a philosophical one that may as well lead us back to rethinking or reevaluating at least fundamental premises in Adam Smith's and other pioneers' works. Until an adequate analytical dynamical equilibrium model can be devised that generates output explained clearly and globally by market observations, this question will remain open and create a lot of controversy and scientific debate.

1.5 About this thesis

This thesis belongs to the domain of nonlinear time series analysis and evolves around three empirical essays. These essays are focused on revealing hidden features of stock market data that are consistent with the hypothesis of a deterministic rather than stochastic underlying data generating process. We concentrate on the empirical investigation of nonlinear deterministic structures (i.e., chaos) within financial time series and more precisely on daily frequency closing prices of stock-market indices and their corresponding logarithmic returns. Chapters 5, 6 and 8 are the empirical essays. Chapter 2 contains the literature review and chapters 3, 4 and 7 are dedicated to background theoretical information on the techniques and methodologies used in the empirical part. The mainstream approach is that of **a univariate discrete time series analysis framework**.

Problem Statement: *The problem tackled in this thesis is how an array of novel qualitative and quantitative time series tools can be utilised to reveal hidden patterns in the data, not explained by the premises of the Efficient Market Hypothesis. We examine the methodologies of Recurrence Analysis, Surrogate Data Analysis and Wavelet Transforms.*

In this thesis we use mainly two different approaches. We employ methodologies developed for the examination of chaotic deterministic time series and Wavelets. These methodologies have been applied successfully in physical sciences and produced often conclusive results. As far as we know, only a few applications of these exist in the financial literature, under different contexts. Based on research from chaos theory and dynamical systems we endorse here *Recurrence* and *Surrogate Data analysis* on financial time series. We rely on a basic topological theorem by Takens (1981) which dictates that all available dynamics are captured within the output of any system. Using this output, and under certain general assumptions, we can regenerate these dynamics in a kind of a *pseudo phase-space* through a technique called *"Embedding"*. Provided that this embedding is successful, we can then apply statistical techniques to examine the existence or absence of stochastic randomness or to reveal qualitative information on the structure of the dynamics. For the latter,

we employ *Recurrence plots* originated by Eckmann *et al.*, (1987) and "*Recurrence Quantification Analysis*", as developed by Webber and Zbilut (1994). Chapters 3-5 deal with this subjects. Following the discovery of structures that reveal substantial absence of stochasticity, we turn to "*Surrogate Data Analysis*" in chapter 6, in order to provide a statistical hypothesis testing framework to refute stochastic linearity for our series. This is a methodology based on permutation testing, very similar to Efron's idea of Bootstrapping but with significant differences. The results there indicate that there is statistically significant absence of stochastic linearity in stock market returns. For both empirical essays 5 and 6, The data used are in their raw form. They have not been pre-whitened or pre-filtered in any way as preprocessing may affect the validity of tests and results and hinder our efforts to reveal deterministic structures.

Given that Takens (1987) theorem is based on an infinite amount of noise-free observations and following the results of previous research which indicates the existence of a significant amount of noise in finite financial time series, we propose a nonparametric model-free filtering approach based on *Wavelet theory*. Wavelets are functions which enable us to model time series in time and frequency domain. Moreover, wavelets are able to capture localised events and denoise highly irregular sequences. Their application ensures that any sequence is examined as a whole as well as in parts. In this way, even the smallest discontinuity is identified and presented. Needless to say that wavelet filtering has already been applied successfully on chaotic time series and revealed a wealth of information while identifying correctly nonlinear deterministic structures. In econometrics, they have been used so far mainly for long-memory empirical investigations on financial time series. Using the continuous wavelet transform we show that there is significant self-similar structure within the returns sequences of the FTSE ALL SHARE market index. Moreover, we use the discrete version of the transform to conduct wavelet based denoising. According to the properties of wavelet functions, such filtering should leave any deterministic signatures in the returns untouched. The pure noise obtained shows that the filtering has worked remarkably well, without any model-based parametric assumption about the data generating process. We then examine the denoised sequence

with the methodologies already covered in chapters 5 and 6, in order to establish whether there are indeed deterministic chaotic structures. The results indicate that there are substantial nonlinear aperiodicities and deterministic components. The dynamics seem to be interrupted only during severe financial shocks. The overall attitude of the index is inconsistent with the assumptions of the EMH and stochasticity. Concluding, the results suggest that indeed, there is no homogeneous structure based on stochastic premises, but nonlinear determinism is evident within the dynamics. The implication here is that the EMH should probably be modified in order to account for these kind of structures.

1.5.1 Software and data

For the empirical parts of this thesis we have utilised widely available software. We concentrated on using mainly General Public License software (GPL) such as R[5] for some of the statistics and GNUPLOT[6] for some of the diagrams. We experimented with the same algorithms and data under both Microsoft Windows 2000[7] and Linux operational systems (Linux Mandrake[8] versions 7 to 8.2). More precisely, we used S-Plus 2000[9], R versions 1.4.0 and 1.5.1 and Matlab[10] (releases 11 and 12) for most of the numerical simulations and estimations. R is an object-oriented GPL S-Plus clone, regarded by most as far superior to its copyrighted counterpart.

In order to obtain some particular measurements and graphical plots which are not part of the standard econometric and signal analysis libraries, such as *recurrence plots* (chapters 5 and 8), we used 3rd party applications, freely available for research over the web. For recurrence plots and recurrence quantification analysis, we used VRA (Visual Recurrence Analysis) version 4.2[11] under Microsoft Windows 2000. We are grateful to E. Kononov, the programmer of VRA and financial analyst for his useful comments and guidance. The VRA program is extremely versatile and produces

---

[5]Software's homepage: http//www.r-project.org
[6]Related webpage: http://www.gnuplot.info
[7]Microsoft©: http://www.microsoft.com/windows2000/
[8]Operational System's webpage: http://www.mandrakelinux.com/en/ also available from ftp://ftp.mirror.ac.uk/sites/sunsite.uio.no/pub/unix/Linux/Mandrake/Mandrake/8.2/i586/
[9]http://www.insightful.com
[10]Mathworks website: http://www.mathworks.com
[11]Available on-line from: http://pw1.netcom.com/~eugenek/download.html

very good results with datasets up to a million observations in size. For obtaining most of the *chaotic invariant statistics*, we used **TISEAN** 2.01[12] under both **Microsoft Windows 2000** and **Linux Mandrake** 8.2. Finally, we should mention that this thesis' typesetting was conducted entirely on LaTeX 2_ε.[13]

1.6 Concluding remarks

So far we have clarified that this thesis is based on the empirical investigation of **nonlinear deterministic structures** in stock market data. As such an investigation, it will be concentrating on a specific set of methodological tools proven to work in other disciplines. It is evident from the examination of the "nonlinearities" empirical finance literature, that there is an extensive set of methodologies which are not utilised in this work. For example, we avoid the area of nonlinear econometrics such as *Switching Regime* models and *Time-Varying Smooth Transition Autoregressive models* (STAR). We have also not considered *Artificial Neural Networks* approach for the standardised criticism that applies to these models regarding their poor out-of-sample performance and overfitting of data. Most of the models mentioned above are treated in texts such as Franses and Dijk (2000), Tong (1990, 1993) and Granger and Teräsvirta (1993) among others. Our approach differs to theirs in the sense that we adopt a more "systemic" view of the dynamics. Without parameterising the phase-space of the market system, using techniques from topology and chaos theory, we simulate and examine the dynamics in a pseudo phase-space. Our main interest is to discover limit attitudes consistent with chaotic determinism. A possible direction for future research would be to see how well our approach would compete with the earlier mentioned methodologies.

---

[12]Software's homepage: `http://www.mpipks-dresden.mpg.de/~tisean/TISEAN_2.1/index.html`
[13]The MikTex version, which is available online from: `http://www.miktex.org`

CHAPTER 2
Literature review

2.1 Introduction

A thesis that deals with the structure of time series of stock market data, will have sooner or later to provide a discussion on the issue of *Market Efficiency* (ME). The Efficient Market Hypothesis is explained in more detail in section 2.2 and has been the most influential premise in early financial literature. It has shaped the methodological approach of many an empirical or theoretical problem. Numerous pieces of research have provided very strong support to this idea whereas other results have shown even stronger evidence against it. Till present, views are mixed end ME remains a controversial scientific area.

In this thesis we demonstrate the use of an alternative univariate time series analysis framework, in exploring and if possible explaining the dynamics of stock market prices. In this sense we expand the toolkit used so far to support or refute market efficiency. Consequently, a relevant discussion to efficiency can not be absent from the literature review. In this chapter we present early as well as recent works on the field. We also refer to research that has provided new views on the underlying dynamics of financial time series, such as *overreaction* and *contrarian* investment. We also provide a brief outline of the new approach that suggests the existence of deterministic rather than stochastic structures for stock market prices. We will be examining this in chapters 4,5 and 6 where we explore the applicability of methodologies derived from chaos theory in financial time series analysis. Finally, we provide a brief discussion on wavelets as a new method of time series analysis and decomposition. These will be more rigorously treated in chapters 7 and 8.

2.2 The basics

A large number of studies over the past century have suggested that certain asset prices, or at least some aspects of the factors of their generating mechanisms, can be predicted. It is still debatable though whether the academic community

has endorsed a fully acceptable theory about what these generating mechanisms may be. Systematic examination of stock market prices structure and variability has intensified during the last half of the past century. The last thirty years have been invested in the examination of price movements in the capital markets under the concept of ME. The core problem being that of the evaluation of the Efficient Market Hypothesis (EMH). The EMH (not to be confused with the notion of "Pareto efficiency"), as presented by Fama (1970), states (in its *"strong"* form) that all available information is fully reflected in the assets' prices. The origins of this idea have been originally attributed to Louis Bachelier (1900). His dissertation (reprinted in Cootner, 1964), based on the random walk and martingale models, was to influence financial thought profoundly. Bachelier introduced the "Brownian motion" (or "Wiener Process") as a model for interpreting the Paris stock market price fluctuations.[14]

It was Fama (1970) that provided what was publicly accepted as early evidence in support of the EMH. His research identified the three classifications of financial markets, according to their informational context: *weak form, semi-strong form* and *strong form* efficient. Weak form efficiency implies that all relevant information is included in the past history of stock prices or returns. Semi-strong form, encompassing weak form efficiency, suggests the incorporation of other publicly available information such as new issues, changes in dividend yields and stock splits to name but a few. Finally, strong form efficiency contains the information set of both previous forms and focuses on privileged-private information i.e., on the monopolistic access to specific information by some group of investors only.

Since the 60s, a large number of empirical studies has supported or attacked the EMH. So strong were the early research findings in favour of the EMH that Jensen *et al.*, (1978) have called it the best established empirical fact in economics. Recent results tend though to stress the fact that the EMH fails to provide an adequate

---

[14]The discovery of Brownian motion is usually attributed to Einstein due to his famous 1905 paper but was anticipated by Bachelier five years earlier. Bachelier's models were studied again, some 60 years later by the Nobel laureate Samuelson (1965) and during the 60s by Mandelbrot (see in Mandelbrot's collections, 1997a and 1999b). Prices under the Brownian motion framework are allowed negative values. Samuelson remedied this by introducing the geometric Brownian motion, defined as the pathwise solution $X_t = X_0 \exp(\sigma W_t + mt)$ of the linear stochastic differential equation $dX_t = \sigma X_t dW_t + \mu X_t dt$ (where W_t denotes a standard Brownian motion and $m = \mu - \frac{1}{2}\sigma^2$).

representation of market reality. On the other hand, no competing theory proposed so far, has been accepted by both the academic community and the markets as a candidate for the replacement of the EMH.

The tests of the EMH can be clusterised in 3 main categories according to Fama (1991). Weak form efficiency tests are now being termed as tests of *"return predictability"*. Volatility tests, autocorrelation in returns tests, methods of forecasting returns with dividend yields and term structure variables, tests for market anomalies and chaotic dynamics belong to this wide category. Semi-strong form tests are termed *"event studies"*. These focus on the economic impact of firm-specific or economy-wide events on asset prices. As"events" may be characterised investment and financing decisions which include changes in management and corporate control, mergers and acquisitions, earnings or macroeconomic indicators announcements, changes in the regulatory and political environment and other. Finally and intuitively, strong-form tests are renamed *"tests of private information"*. These study the impact of insider trading, the performance of money managers, investment consulting services such as the well known case of the "Value Line enigma" (Copeland and Mayers, 1982) and other similar activities.

2.2.1 The early literature

The idea of an "intrinsic" or "fundamental" value around which the security price is expected to fluctuate, appeared not to be consistent with reality. Cowles (1933) suggested that forecasting on the base of fundamental analysis was a waste of money and brokerages operating in this manner were not outperforming the market. Working (1934) realised that cumulated series of individually, identically distributed (i.i.d.) shocks had the same appearance as those patterns exhibited by stock prices. Again, Working (1960) provided evidence in support of the randomness of stock prices by demonstrating the existence of spurious correlation between random walk data averaged over time. Kendall (1953) found that stock prices did indeed follow a random walk as earlier suggested by Working. Granger and Morgenstern (1963) and Godfrey, Granger and Morgenstern (1964), using spectral analysis, supported the same result. However, all these results have been remarkably accounted for in

Bachelier's original thesis in 1900 who has anticipated many mathematical results.

2.2.1.1 The EMH in developed and developing markets

Kendall (1963) using serial correlation coefficients for the first difference of weekly London Stock Exchange price indices, has examined the random character of the price fluctuations. This method was applied to weekly data from the New York Stock Exchange by Cootner (1962) and Moore (1964). The result was some evidence of correlation. Using the set of the 30 Dow Jones industrial index US companies, Fama (1965) produced evidence in support of the dependence between the price changes. Kemp and Reid (1971), by applying non-parametric methods on British data, have also found evidence of the same kind of dependence. Conrad and Juttner (1973), utilising parametric and non-parametric methodologies, have rejected the random walk hypothesis for the German Stock market. Serial correlation tests were employed by Dryden (1970), Jennergen and Korsvold (1974 and 1975) and Roux and Gilbertson (1978) in their studies of testing the weak form efficiency. Australian share prices were examined by Praetz (1969, 1973, 1979). Wong and Kwong (1984), using daily closing prices of 28 Hong Kong stocks, concluded that weak form efficiency cannot be supported for the Hong Kong market. Ang and Pohlman (1978) studied the Australian market as well as that of Malaysia, Hong-Kong, Japan and Singapore. Solnik (1973) collected data from 234 stocks from Belgian, British, Dutch, French, Italian, German, Swiss and Swedish stock markets. Cooper (1982), using spectral analysis, runs tests and correlation analysis, has reviewed weekly and daily data from 36 countries and found that only for the USA and UK markets the random walk hypothesis is clearly supported.

A number of papers, utilising conventional statistical methodologies, examined the applicability of the EMH in developing economies. Hong (1978) has tested for weak form efficiency on the Singapore Stock Exchange and did not reject the random walk hypothesis. Ang and Pohlman (1978) supported the weak efficiency for the same stock exchange. Barnes (1986) finds the Kuala Lumpur stock market to be inefficient. The inefficiency of the Kuwaiti stock market has been examined by Gandhi et al., 1980. Butler and Malaikah (1992) report that the Kuwaiti market

is not inefficient and they provide evidence for the lack of efficiency of the Saudi Arabian market. Panas (1990) has investigated the efficiency of the Athens stock market using monthly returns and an array of nonparametric tests. Recently, Los (1999) examined the ten year weekly data history (1986-1996) of six east Asian markets. He ranked them according to their level of efficiency: Singapore, Thailand, Indonesia, Malaysia, Hong Kong and Taiwan (the latter being the most inefficient ones). Antoniou *et al.*, 1997, examining the evolution of the Istanbul Stock Exchange suggest that efficiency comes with high trading volume, reliable information and institutional reforms.

2.3 Bubbles, fads, anomalies and overreaction

In the last two decades, a number of financial phenomena have been exhibited for which no undisputable probabilistic models have yet been developed. For example, Smith *et al.*, (2000) during a set of experiments, had all uncertainty regarding the formation of price expectations[15] eliminated by making known to all participants the distribution of future dividends. Thus one would expect a deterministic price process. In fact, more than 50% of the set of experiments were exhibiting price bubbles followed by crashes with respect to intrinsic dividend value. This is a clear contradiction to what should occur under the EM hypothesis.

There is also a growing collection of literature focusing on tests of weak and strong form efficiency, providing evidence against the EMH. Evidence of predictability of stock and bond returns, has come under the guise of autocorrelation patterns within their structure. Fama and French (1988) claimed that returns could not be predicted by using their history as stock prices follow a random walk. Conrad and Kaul (1988 and 1989) provide evidence that short-horizon returns are positively correlated. Lo and Mackinley (1988 and 1990) support the same finding. Poterba and Summers (1988) and Fama and French (1988) also report negative autocorrelation for long-horizon stock returns. Moreover, fads, mean reversion, market overreaction and time-varying expected returns, are all concepts emerging from the attempt to model market inefficiency and provide a suitable explanation for correlation in re-

---

[15]Based solely on dividend forecasts.

turns. The question remains, as to which of these three concepts is the correlation clearly attributed to. Until now studies have not led to a widely accepted distinction between them. The fads literature suggests that psychological factors, noise and feedback trading can influence investors in their investment decisions (Shiller, 1984 and 1989 and Black, 1986).

DeBondt and Thaler (1985 and 1987) introduce the overreaction model (see also Shefrin and Statman, 1985). Under the EMH, investors' reactions to market signals at any time, are specific to the available information set at that time. This information set contains past experience and expertise (asset price history, fundamentals, heuristics, scientific theory). New information is filtered and classified according to a set of investors' learning rules which have not been identified yet. In theory, the most popular available learning rule is that of Bayes but this has not been established as an apt characterisation of how investors react to new information. There has been substantial evidence that individuals tend to overweight recent information and disregard prior input (as "*a-priori beliefs*"). Although Bayes rule implies that investors should "weigh" new information by some consistent rule and refine this through a trial and error process, they seem to attribute disproportionate importance to short-run economic developments. Examining this short memory attitude, De Bondt and Thaler (1985) find that "losers" portfolios outperform "winners" by 25% even though the latter exhibit a higher degree of riskiness. This anomaly is being exploited by the so-called "Contrarians" in order to achieve abnormal returns. The main criticism to the De Bondt and Thaler conclusions is that systematic risk and the size effect are factors that can explain sufficiently this market anomaly. "Fundamentally riskier" assets should be awarded with higher returns and contrarian investment is more risky by definition.

Chan (1986) suggests that time-varying expected returns are responsible for the "overreaction" of certain investors. These are caused by a positive covariance between contrarian portfolios' betas and the risk premium. De Bondt and Thaler (1987) re-examine the issue and conclude that the overreaction effect still exists after controlling for differences in size and beta. The beta of their contrarian portfolio is positive in up-markets and negative in down-markets which suggests that time var-

ying betas are not suitable for risk-adjustment of the returns of this type of portfolio. Ball and Kothari (1989) are in accordance with Chan's approach but do not detect a positive covariance between beta and the risk premium. Zarowin (1989) attributes overreaction to the "size" phenomenon. Lehmann[16] (1990) provides evidence against market efficiency. Lo and MacKinlay (1990) show that when positive cross-covariances across securities exist (even if returns are serially dependent) then there is space for profitable contrarian investment strategies. Chopra *et al.*, (1992) detect stock price overreaction for small firms. They also reveal the distinctiveness between the January and overreaction effects. Jegadesh and Titman (1993) report that there is no sufficient evidence to prove that systematic risk is an important factor. They conclude that a more sophisticated model is needed to interpret the patterns in expected returns. The buying winners and selling losers strategy could be attributing to the temporary deviation of prices from their long-run (equilibrium) prices and thus creating the overreaction effect. Moreover, information about long-term and short-term prospects of companies is filtered diversely by individual investors[17] and this fact may generate a climate suitable to contrarian strategies. Conrad and Kaul (1993) report a methodological obstacle that could be leading to the exaggeration of the overreaction effect. Lakonishok *et al.*, (1994) note that *out-of-favour* (value) stocks, outperform *glamour* stocks in the long run simply because investors are not informed about them (they are prone to "judgmental errors"). Another plausible explanation is that they have merely identified an ex post pattern in the data.

Between the general overreaction hypothesis and the concept of arbitrage are some significant theoretical links. These are mainly based on the hypothesis of existence of rational and quasi-rational investors (those who are unable to revise their beliefs according to the Bayes rule). The generated rational equilibrium framework has been proved to be unstable and the presence of rational agents is not sufficient to guarantee any stability for the economy under the assumption of heterogeneous information. Basu (1977) has addressed the overreaction anomaly via the P/E (price-earnings ratio) and distinguished between undervalued and overvalued

---

[16]Through examination of the profits of zero net investment portfolios that earn only if investors overreact.

[17]Depending mainly and most likely on their risk tolerance.

companies describing the reactions to their assets' price adjustments. It is obvious that the whole spectrum of the overreaction issue is closely related to the efficient-market hypothesis. In fact it implies *weak-form inefficiency* and easily expands this to semi-strong and strong forms. Moreover, it violates the rationality principle of investors, since the present value of expected future dividends is no longer the basis of rational investment decisions.

Summers (1986) and Poterba and Summers (1988) expose the mean reverting movement of asset prices. They suggest that stock prices exhibit temporary deviations from the path that their fundamental (intrinsic) values were expected to follow, only to return gradually to this rationally determined equilibrium trajectory (Engel and Morris (1991) review this subject extensively). Poterba and Summers report that the mean reversion is stronger for small firms. They calculate variance ratios for horizons of two to eight years on monthly excess NYSE returns from 1926 to 1985. They report a less than one variance ratio for all investment horizons greater than two years, although the eight year return variance would be expected to be eight times the variance of the one-year returns (to claim that the EMH holds). Fama and French (1988a) extend this methodology to multi-year returns and test for mean reversion for investment horizons of one up to ten years. They use inflation adjusted NYSE monthly data for various industry groups for the same period as Poterba and Summers (1988). They too conclude that stock prices are mean reverting and mean reversion is stronger for stocks of small firms. When dividends are used to measure fundamental values, Fama and French (1988b) again find that stock prices are mean reverting. If prices were mean reverting, one would expect returns to be negatively correlated to the difference between prices and dividends. This was demonstrated in their article and for one more time, mean reversion was more intensive for small firms. Campbell and Shiller (1988) included earnings in their estimation of fundamental values and using S&P500 inflation adjusted returns and excess returns from 1971 to 1978. They found that prices were mean reverting for one-year, three-year and ten-year horizons.

Evidence of mean reversion has caused a lot of researchers to conclude that the stock market is inefficient. However there are studies that have posed some questions

against this conclusion. Some argue that mean reversion could be the facet of a more sophisticated expression of the EMH. It is also argued that observations are too few to provide powerful estimates when testing with long horizons. Mankiw *et al.*, (1991), Nelson and Kim (1990), Kim *et al.*, (1991) effectively argue that statistical tests lack the power to suggest evidence of mean reversion and consequently the collapse of the EMH. Secondly, there is a suggestion that if stock prices are mean reverting, markets might still be efficient. The constancy of real interest rates may be assumed incorrectly and this is a hypothesis that may have to be relaxed. If real interest rates are mean reverting then so too will be the stock prices. Variability of real interest rates could be due to the change of risk tolerance of individual investors or riskiness of stocks (see Fama and French, 1988b, Black, 1990 and Cecchetti *et al.*, 1990). Poterba and Summers (1988) oppose this view as this change would have to be unrealistically large to cause mean reversion of the observed magnitude. Moreover, as in Lo and MacKinlay (1988), evidence of feedback trading in the short run is not in accordance to with changes in the riskiness of stocks or the risk aversion of investors. Miller et. al (1994) attribute mean reversion to infrequent trading when drawing results from stock market indexes. Bodoukh *et al.*, (1994) argue that the autocorrelation patterns of short-horizon returns have been overstated and that they are caused most likely by institutional factors.

2.3.1 Asset-pricing models and "anomalies"

As part of the EMH debate, we can consider the literature about "calendar anomalies" in stock returns. Returns seem to be higher on average before holidays and at the end of the month. Inversely, Monday returns are on average lower than returns on other days as Cross (1973) and French (1980) point out. There is also a "January" effect as especially small stocks have on average higher returns during that month.[18] Lakonishock and Levi (1982) argue that calendar effects should not contradict the EMH if adjustments for interest gains on certain days during the trading period are made. Intraday patterns are also anomalous. Harris (1986) finds

---

[18]Keim (1983) and Roll (1983a and 1983b) show that the higher January return on small stocks is realised the last trading December day and the first five January days. This seems to be occurring because investors drop underperforming stock in December for tax reasons.

that most of the average daily return is localised in time around the beginning and the end of the day. Ariel (1987) shows that most of the monthly average return comes in the first half.

There are also suggestions that asset-pricing models fail to reflect efficiency. These models should predict that the market betas are describing in a effective manner, cross-sectional expected returns. Basu (1977 and 1983) reports positive correlation between expected returns and P/E ratios. Banz (1981) reports negative relation between risk-adjusted returns and size. Bhandari (1988) reveals correlation between returns and leverage and demonstrates that small firms are highly levered. Fama and French (1991) and Chan *et al.*, (1991) have found correlation with book-to-market ratios. Fama and French (1991) also detect that leverage is highly correlated with book-to-market ratios. Reinganum (1981) and Basu (1983) also report size and P/E relationship.

2.4 A new approach: chaotic determinism

Since the late 80s, a number of papers have explored the possibility of chaotic evidence in financial markets[19]. The statement of chaos is a very strong one as it implies a deterministic, often very complex DGP process whereas the previous stochastic paradigm was allowing for a purely random behaviour. Such stochasticity could often be approximated by some linear function. Moreover, chaotic processes may appear to behave randomly and can be predictable in a short term basis. Most linear methodologies for detecting stochasticity usually fail to produce a correct answer when applied to non-stochastic chaotic data. Finding chaotic patterns in financial time series involves the application of non-standard methodologies and the search for self-similar and fractal structures within the data. Chaotic systems are also highly dependent on initial states. Small shocks in initial conditions will lead to large deviations from the initial state of the system. Usually, this is another way we can explore the existence of chaos in dynamical systems.

It becomes apparent that under the chaotic or deterministic non-linear dy-

---

[19]We dedicate chapter 4 to the basics chaos and dynamical systems theory and the metric tests used to determine whether a sequence is chaotic-deterministic or stochastic.

namical systems view, the "linear" view of stock market efficiency is abandoned. Sudden clustering of volatile sequences and bursts of instability may appear random but could just as well be chaotic. Standard techniques may be misleading as they might report a chaotic process as a purely random one. Unfortunately, there is no well established and robust statistical test for chaos as a null or an alternative hypothesis.

One of the most applied tool so far has been the *BDS* test named after Brock, Dechert and Scheinkman (1987) based on the Grassberger and Procaccia (1983) correlation dimension concept (also used as a proof of existence of chaotic dynamics). The drawback here is that the BDS test detects whether the data is identically independently distributed (i.i.d.) and the rejection of the null does not necessarily imply the existence of chaos. Unfortunately, early efforts have applied the BDS test as a direct test for chaos and their results have been biased if not erroneous. Recently the BDS test has been criticised in more expanded frameworks such as the one in Brooks and Henry (2000a and 200b). They showed that for a particular class of heteroscedastic models, the test fails remarkably to detect common misspecification in the sense that it confuses quite different types of nonlinear structures such as the *Threshold Autoregressive* (STAR) and a range of GARCH-type models.

Other statistics that can provide support for the existence of chaotic deterministic dynamics are the *Lyapunov exponents* which should also be computed with caution (Sauer *et al.*, 1998). These show whether there is an exponential divergence from initial conditions, something that could imply the presence of chaos. More recently, the *"Surrogate Data Methodology"* has been developed as a means of distinguishing weakly deterministic and possibly chaotic patterns from linear stochastic ones. Other tools for distinguishing chaotic dynamics are visual. Diagrams such as scatter plots of lagged values or phase plots and recurrence plots can provide evidence of chaotic or deterministic dynamics. We explore these issues in the following chapters and especially in chapters 5 and 8.

Revealing empirical evidence for the existence of chaos in financial data has been met with mixed feelings and various degrees of success. Attempts to design a theoretical framework that would explain why determinism should characterise

such series, seems to have enjoyed less publicity and fame. Among early attempts to provide an analytical framework for market structure that allows for chaotic dynamics were that of Savit (1988) and Shaffer (1991). The first, combining theory based on logistic maps and the cobweb, attempted to explain the problem of price movements in financial markets. Shaffer (1991), driven by the 1987 stock market crash, demonstrated that some simple and widely accepted fundamental factors were enough to generate chaotic paths for profits and stock returns. Slight perturbation of initial parameters could lead to large increase of the volatility and dividend pay-outs could induce chaos. The conclusion was that even regulatory procedures on trading strategies could not prohibit the existence of chaos in the data. More recently Malliaris and Stein (1999) have presented their version on this issues. While they do not conclude on a chaotic underlying process, they suggest that maybe market dynamics are expressions of a high dimensional stochastic mechanism. They point that due to the absence of information on all the finer parameters defining the market system, certain deterministic aspects can be masked as randomness. They do not agree with the hypothesis of chaos as this excludes long term forecasting in markets which is apparently not true.

Hinich and Patterson (1985), Scheinkman and LeBaron (1989), Brock *et al.*, (1991), Hsieh (1991), Liu *et al.*, (1992), Lane *et al.*, (1996), Gilmore (1996) and Yadav *et al.*, (1996 and 1999) among others, have all examined whether stock returns exhibit non-linear dependence and dynamics consistent with chaos theory. Ashley and Patterson (1989) detect nonlinearities consistent with chaotic dynamics but fail prove the existence of chaos. Among early attempts for the discovery of chaos in non stock-market data should be mentioned the papers of Brock and Sayers (1988) who found nonlinearities in U.S. labour data and Barnett and Chen (1988) who detected similar artifacts in U.S. monetary measures. Frank *et al.*, (1988) also detected nonlinearities in Japan's GNP. Most of this research focuses around bilinear-bispectral techniques (see Granger and Andersen 1978) or the estimation of the invariant statistical measures that can characterise chaotic systems such us the correlation dimension.

Ramsey *et al.*, (1990) find that after correcting for nonstationarity and small-

sample bias, no evidence of low dimensional dynamics can be revealed. Most of these researches though have been focused on stock index returns and provided weak evidence for or against chaotic behaviour. Yadav *et al.*, (1999) examined the UK stock market stocks over a period of 20 years. They concluded that nonlinear dependence exists but it seems to be the result of an EGARCH element and no strong evidence for the existence of low dimensional deterministic dynamics is obtained.

Approaches such as Brock *et al.*, (1991), De Grauwe *et al.*, (1993), Hsieh (1991), Mayfield and Mizrach (1992), Peters (1991 and 1994), Scheinkman and Le-Baron (1989), LeBaron (1992, 1993, 1994) and Bollerslev (1990) have highlighted the issue regarding the origins of the financial volatility. Regarding "chaos", most findings reported no convincing evidence of chaotic dynamics but often indicated strong nonlinear dependencies among the data. The issue raised was the possibility of forecasting financial time series that might contain chaotic components or nonlinear structures. This fact is of great importance to investors who are strongly interested in market timing dynamic trading and similar investment strategies, which are clearly depending on the short-term forecastability of these time series (LeBaron, 1994). Similar research has been undertaken with futures data. Yang and Borsen (1993) conclude that the can not provide strong support for or against chaos in futures data but also show that the GARCH specification is not totally adequate in characterising the dynamics. Investigating nonlinearities in emerging markets, Antoniou *et al.*, (1997) show that non-linear terms can be found insignificant and then their model collapses to a random walk. They argue that non-linearities may not necessarily imply market inefficiency. The conclusion from most of all the researches outlined here is that further analysis is required on the underlying structure and dimensionality of financial time series.

One of the most early and appreciated researchers of chaos in finance is LeBaron. His work involves research not only on the structures of asset-price and returns time series but also on the volatilities of these. As LeBaron (1994) states, *"One of the largest deviations from pure randomness in financial series is volatility persistence"*. Forecasting return movements may be a tedious and unsuccessful process but *magnitudes of the movements are predictable*. LeBaron demonstrated that the DGP

for a substantial set of financial time series appears to follow the pattern:

$$r_t = log(p_t) - log(p_{t-1})$$

$$r_t = f(\sigma_t^2)r_{t-1} + \epsilon_t$$

$$\sigma_t^2 = \sum_{i=1}^{N} r_{t-i}^2$$

for t=1,2,...,T, where $f(\cdot)$ is a decreasing function of conditional variance *indicating higher local predictability during periods of lower volatility*. Relating these findings to investment strategies we should note that risk is relative to market volatility. Volatile periods reward risk-lovers with excess returns whereas passive strategies are fruitful during relatively stable markets.

In another of his working papers LeBaron (1993) provides evidence of *local instabilities in the volatility-volume process*. "A 10% shock to the volume process may be amplified several times in volume over several days in the future". This power-law feature cannot be captured when utilising linear techniques. Although this may not imply the existence of deterministic chaos in volume and volatility, it does strongly state that *local instability* prevails.

Closing this brief review of chaotic literature in finance, one feels obliged to present a counterpoint. Granger has seriously criticised since very early the whole idea of chaos in economics. This is evident from numerous replies to Mandelbrot in the past. More recently, Granger (1994) discussing a collection of articles on chaos: *"If the purpose of this research effort is to show that it is possible to produce economic theory that has chaos as an outcome then this certainty has been achieved, but it is unclear why more than one paper is required that shows this. The recent papers produce more general models but are still not realistic and are not in a form that is empirically testable in any specific fashion"*. Evidence is thus required that the chaos paradigm *can reproduce many features of the actual economy and not just a few and also predict features that have not been previously tested or explained*. It remains to be seen if a robust technique could be constructed that can pinpoint nonlinearities in financial data and withstand specification testing at the same time. Surrogate data analysis (which we will be exploring in chapter 6), seems to provide a promising framework towards this direction. One recently designed analysis framework, based

on earlier mathematical research is wavelet analysis of time series. Due to the intermittent and volatile nature of financial time series, it seems a very suitable alternative to other parametric approaches. We discuss this briefly below but we will be returning to this issue in chapters 7 and 8 were we review the theory more in-depth and provide an exploratory analysis of financial time series.

2.5 Wavelets in finance

Non stochastic structure, cycles and even low dimensional chaos can also be detected by implementing *Wavelet* functions. The potential of this approach, although obvious from the examination of its mathematical foundations and properties, has not been fully exploited in Economics or Finance yet. Wavelets can also be used to detect (a)cyclical components and high frequency localised in time bursts that "conventional" techniques would fail to capture. Following is a brief analysis of why wavelets are more useful than other standard Fourier-transform methodologies. Other tools have also been developed and used for detecting various kinds of cyclicity within the structure of any univariate or multivariate time series model, such as spectral decomposition or seasonal filtering techniques (e.g. X11 filter or seasonal ARIMA models), trigonometric regression or the somewhat "brute-force" method of seasonal dummies to name but a few. These existing techniques suffer from the fact that even when implemented successfully, they are either *time* or *frequency* localised, thus enabling us to view only one dimension of the cyclical pattern. Furthermore, these techniques, tend to under-perform when the time series signals are contaminated with noise or contain non-stochastic components, such as chaotic cycles and low dimensional chaotic noise. Although wavelets will be more extensively covered in chapter 7, we can provide a brief introduction here for the sake of clarity. In most disciplines, the most traditional technique for the cyclical decomposition of any signal, is that of Fourier (or Spectral) analysis (FA). The spectrum of any sequence of numbers is defined as the variance decomposition of this sequence. Any spectral peak signifies the existence of a cycle at that frequency. It is this cycle that contributes substantially to the variance of the sequence. Thus a series of observations along time can be "translated" (i.e., decomposed) into high

and low frequency components, depending on the existence of respectively high and low frequency cycles within the original sequence.

The Fourier Transform (FT) decomposes a function into sinusoids of different frequencies. The original function is described in terms of orthogonal basis functions of sines and cosines, of infinite duration. These sinusoids when summed up, reconstruct the original function. The FT of a function $f(x)$ is defined as:

$$F(s) = \int_{-\infty}^{+\infty} f(x)e^{-i2\pi xs}dx \qquad (2.1)$$

Suppose we have a sample time series $\{x_0, ..., x_{t-1}\}$ with mean $\bar{x}$ of zero (or the mean has been already removed). *Parseval's* theorem states that the sample variance of $\{x_t\}$ is:

$$s^2 = 1/n\sum_{t=0}^{n-1} x_t^2 = 1/n\sum_{j=0}^{n-1} \mid A(2\pi j/n) \mid^2, \qquad (2.2)$$

where:

$$A(\omega) = \sum_{t=0}^{n-1} x_t e^{-i\omega t} \qquad (2.3)$$

is the discrete Fourier transform (DFT) of $\{x_t\}$. The DFT analyses discretely sampled time series. Equation (2.2) shows that the variance of the series can be decomposed into a set of frequencies. The expression $\mid A(2\pi j/n) \mid^2$ is the spectrum of the series. Unfortunately FA is not without its drawbacks. Firstly, FA requires the use of the entire range of the sequence for an evaluation of the spectrum at any specific frequency point. For instance, to evaluate $\hat{f}$ at point ω_0

$$\hat{F}(\omega_0) = \int_{-\infty}^{\infty} f(t)e^{-i\omega_0 t}dt \qquad (2.4)$$

one must use $f(t)$ along its entire range (FA assumes an infinite time horizon). Secondly, the method demands a great deal of information for sufficiently capturing any local fluctuation. The Fourier representation of time localised events, requires many terms of $Ae^{in\omega}$. This signifies in itself that a large set of cycles of different frequencies must be measured (sampled) within the time series in order for the method to be able to successfully detect any kind of shock or irregularity in the

variance. Moreover, the DFT or FT has difficulty with functions with transient components (i.e., components localised in time such as any sharp oscillations, spikes-shocks or violent regime shifts within the time series). Any local oscillations may contribute to the calculated FT but the locality of such oscillations will be lost. There is no way of establishing whether the value of $\hat{F}(\omega)$ at a particular point in time t derives from frequencies during a few selected periods or the entire range of the signal.

Financial time series may be either non-stationary Gaussian, stationary non-Gaussian, chaotic of some deterministic/fractal nature or combination of these (stochastic contaminated with chaotic components or chaotic with embedded stochastic noise). One possible reason for this may be that in financial time series, one deals with a DGP which is not well defined. Rational financial decisions that will eventually generate or formulate financial market indicators, are based upon the observation of fundamental economic variables as well as unobserved emotional factors. The wavelet transform does not depend upon any presumptions on the stationarity of the time series or the distribution properties of the underlying DGP. What the wavelet transform does, is to generate in a linear manner different levels of scales of coefficients preserving thus the time element of the data.

The literature on wavelets is vast and focuses greatly on image analysis and compression, signal analysis, data transfer rates analysis and sound and vibration modelling to name but a few important areas. In the area of Economics and Finance, there are a number of recent publications, mostly in the form of working papers for the past decade. Ramsey et al., (1995) analyse U.S. stock prices using wavelet methodologies. They report evidence of non-randomness in the data and limited evidence of quasi-periodicity of large amplitude shocks. A demonstration of wavelet techniques on Spanish stock market data can be found in Arino and Vidakovic (1995). Early attempts to provide an insight are Cody (1992 and 1994 not so technical papers) and Klimasauskas (1992). Strang's (1989) SIAM review paper is an excellent starting point. More recently Ramsay and Zhang (1995), use waveform dictionaries[20] to analyse foreign exchange data. Ramsay and Zhang (1994) apply the

---

[20]Waveform dictionaries are function libraries based on wavelets.

same methodology to stock market index data demonstrating again their complexity and implying that random walks might not be a plausible explanation for such series.

2.6 Conclusions

There is no doubt that financial time series exhibit nonlinearities and highly volatile and intermittent structures. So far a multitude of traditional econometric methods has not been able to provide adequate and robust explanations for such structures and dynamics. Chaos has been an attractive alternative to stochasticity for financial time series but there has been no conclusive evidence for or against its existence in these sequences. With the advent of cheap computational power, previously impossible to apply techniques have been used and new results have come to light. New mathematical methodologies have also been proposed for the analysis of such data, namely the wavelets and neural networks.[21] An entire new area has slowly been forming called "*econophysics*". This simply is the combination of ideas, models and techniques from the physical sciences, applied in solving economic and financial problems.

Financial markets are extremely complex systems and constantly bombarded by external noise. If we are to adopt this "complex systems" approach when examining these markets, we should have a good idea about the underlying structure of the system producing the market dynamics. External factors such as noise, regulatory frameworks and information impose shocks and restrictions on how a system functions and more importantly on how it reaches an equilibrium state. The instabilities caused by these factors can lead to large deviations from an "equilibrium" point, thus implying a chaotic and complex overall market structure or at least a chaotic mechanism that may be masking the true dynamics of the overall market. Chaos is also an attractive alternative to the simple stochastic-probabilistic framework of such dynamics. Although a chaotic system may tend to deviate from a "steady state", depending on initial conditions at time of externally caused shock, it is also capable of "self organising" and returning to a stable path before being

---

[21]We will not be addressing the applicability of neural networks in this thesis. This is left for future research.

thrown off-balance again. This is not possible with most of the linear or stochastic systems. In general, chaotic structures can withstand shocks better than stochastic ones. This may be why results are dubious when modelling a complex-chaotic phenomenon with "stochastic" tools that are incapable of reporting correctly on nonlinear deterministic processes.

Chaos also represents a radical change of perspective on business cycles. Business cycles receive an endogenous explanation and are traced back to the strong nonlinear deterministic structure that can pervade the economic system. This is different from the (currently dominant) exogenous approach to economic fluctuations, based on the assumption that economic equilibria are determinate and intrinsically stable, so that in the absence of continuing exogenous shocks the economy tends towards a steady state, but because of stochastic shocks a stationary pattern of fluctuations is observed (Barnett *et al.*, 1997).

In this thesis we will be examining from a new perspective, the applicability of certain methodologies derived from chaos theory and signal analysis, in order to reveal structures that exist but have not been reported so far in financial time series. More precisely we will be addressing the issue of whether the stock market returns are characterised by stochastic randomness and whether there is evidence of a linear structure in the data generating processes. We will be examining this using both quantitative and qualitative information drawn from the time series themselves.

In the following chapter we will be discussing the most popular statistical processes that have been said to provide an adequate representation of the dynamics of financial time series. Following that chapter we will be providing the basic theoretical framework for dynamical models and chaos. We next explore financial time series prices and returns for deterministic dynamics using *recurrence* diagrams. These can discern high dimensional nonlinear dynamics and the possible existence of chaos. Following this chapter, we show how a hypothesis testing framework can be arranged and applied to rule out the possibility of the existence of linear or transformed linear dynamics in these series. The following two chapters are based on a more relaxed approach driven from the successful application of wavelets in other disciplines. We provide the basic framework of *multiresolution* analysis and *wavelet*

transforms in chapter 7 and following this we explore asset returns and prices in various ways and provide a new point of view of their dynamics and structure. In this chapter we also combine techniques from econometrics and chaos theory and suggest that the underlying dynamics may indeed be deterministic but appear to be stochastic, due to the high level of signal corruption caused by noise and volatility.

CHAPTER 3

Statistical processes for asset prices

3.1 Introduction

Since Bachelier's thesis in 1900 (in Cootner, 1963), the prices of financial assets have been considered as a result of various generating mechanisms. Until today, no robust evidence has been revealed to maintain a specific hypothesis for asset price data generating processes (DGP). Unlike in the physical sciences, the evolution mechanisms and theory governing the dynamics of the complex socio-economic and financial phenomena, are largely unknown. Moreover there is no possibility of a controlled experiment which would provide evidence for or against a specific DGP structure. One can only speculate on a parametrisation of the DGP provided that some theoretic assumptions are met within a market-clearing equilibrium framework. Parametric approaches, although very attractive have failed to provide until today a clear-cut answer regarding the nature of the asset price dynamics.

A more recent approach has suggested to "let the data speak for themselves" and proposed a nonparametric examination of the information from the stock markets. Such model-free approaches can only provide qualitative information about the data involved. To obtain a more precise view of the whole phenomenon, one must often resort to semi-parametric approaches without though providing a complete specification of the price processes. Again here, although there is a parameter estimated on a specific property of the DGP, we cannot obtain a fully parametric expression for the complete dynamics.

Originally, financial innovations were demonstrated to be following stochastic processes, either discrete or continuous. With the wake of Chaos theory and the research on deterministic nonlinear dynamical systems, implications for or evidence of chaotic processes in financial time series has been reported by many researchers such as Mandelbrot (1963), Peters (1991), Evertsz (1995a), Evertsz and Berkner (1995b), Hsieh (1991) and LeBaron (1988). Their views have been largely criticised and the academic community is reluctant to adopt the new paradigm of Chaos

theory as no strong supporting evidence has been discovered yet. Secondly, it is hard, even for academics, to interpret irregular and volatile -apparently random- financial phenomena by merely adopting the view that everything is predetermined but too complex to forecast and anticipate, at least in the long run.

The examination of nonlinear determinism in financial time series, requires the use of non standard methodologies and this might prove disturbing. Secondly, the academic community is still in the dark regarding what economists are trying to prove. There is no mathematical equilibrium model yet that could act as a point of origin and proven to be robust. Moreover, some of the most promising methodologies are still "under development". For these reasons, economists have turned to other disciplines such as mathematics, physics, engineering or even medicine where turbulent and highly irregular phenomena exhibit similar characteristics. Technical analysis methods could also favour the existence of chaos in financial phenomena but these are not entirely accepted as scientific and fail-proof.

Since this thesis is focused on novel methodologies that explore the possibilities of various non-standard DGP structures for asset prices, it would be helpful to include a brief outline and some basic mathematical properties that have been the object of interest so far. We begin from the widely adopted martingale and Brownian processes. We then present the Levỳ, fractional Brownian and $1/f$ processes as in Mandelbrot (1997 and 1998) and Wornel (1993) and finally present the fractal and multifractal processes. In the following chapters we show how new methodologies such as *wavelets, recurrence plots* and *surrogate data analysis* can be used to shed new light into the problem of the nature of the asset prices' DGPs.

3.2   Statistical properties of asset returns

Asset prices are the output of the market dynamics. This is the real observable quantity one can obtain from the financial news services around the world. As one is really interested in dynamics, the first differences of asset prices are the ones that show how the latter fluctuate around a mean value. In this thesis we will be mainly focusing on closing prices and the corresponding logarithmic returns. These can be obtained by taking the 1st differences of the natural logarithms of the closing prices

and are also called *"continuously compounded returns"*. The logarithmic function is a positive monotonic (i.e., invertible) transformation and provides sequences that are independent of the corresponding levels of (the closing) prices.

Among the early attempts to provide a stylised approach to economic or financial time series structure were those of Granger (1966) and Fama (1965). Here we follow Cont's (2001) more recent work and provide a list of *stylised empirical facts* of asset returns and volatilities:

- Asset prices up to daily frequencies appear to be a unit root process I(1) whereas the returns usually exhibit stationarity.

- Linear correlations of asset returns are often insignificant in low-medium frequencies. In high frequencies such as tick-by-tick data, they are more prevalent due to microstructure effects.

- The unconditional distribution of asset returns appears to behave in a Pareto-like fashion, displaying fat tails which exclude normality and infinite variance stable laws.

- The conditional distribution of the residuals after correcting returns for the heavy tails (as in the previous point), still exhibit heavy tails although these are weaker than those of the unconditional distribution of returns..

- Gain-loss asymmetry is more obvious in asset returns than in exchange rates. Large drops are more prevalent then equally large gains.

- As one increases the time scale, returns distributions approximate the normal i.e, the distribution of returns varies with the time scale.

- Irregularly distributed bursts within the history of asset returns and non-homogeneous patterns of volatility, give an intermittent character to these series. Intermittence is a very common characteristic of physiological time series and all methodologies that we examine in the empirical chapters of this thesis have been applied on such data.

- Asset returns usually exhibit long memory and long-range dependence.

There are also a number of stylised facts about the volatility of asset returns:

- Volatility clustering is another characteristic of asset returns. High-volatility events seem to cluster in time.

- Most measures of asset return volatility are negatively correlated with the original returns sequences. This has been termed as the "leverage effect".

- Trading volume is usually correlated with all measures of volatility.

- Fine-scale volatility can be better predicted from coarse-gained measures of volatility than the other way around.

3.3 Martingales and Random Walks

According to the *"martingale model"*, tomorrow's price is expected to be the same as today's. A stochastic process x_t follows a martingale with respect to a sequence of information sets Ω_t if

$$E_t(x_{t+1}|\Omega_t) = x_t \tag{3.1}$$

where Ω_t is the information set at time t i.e., $\Omega_t = \{x_1, x_2, ..., x_t\}$ and x_t being the best forecast for x_{t+1}. This holds true for any information set Ω_t. What the martingale model states is that $(x_{t+1} - x_t)$ is a *"fair game"*. A stochastic process y_t is a fair game if it has the property

$$E_t(y_{t+1}|\Omega_t) = 0 \tag{3.2}$$

Furthermore, y_t is a *"submartingale"* if

$$E_t(y_{t+1}|\Omega_t) > y_t \tag{3.3}$$

This itself implies that $E_t[(y_{t+1} - y_t)|\Omega_t] \geq 0$. The analogous concept here is a that of a "subfair game".

It is obvious that a process x_t is a martingale *iff*[22] $(x_{t+1} - x_t)$ is a fair game.

---

[22] *iff*: if and only if.

A fair game can also be called a *"martingale difference"*. Under the martingale framework, increments in prices (adjusted for dividends) are unpredictable given *only* the existing information set Ω_t. This is fully reflected in prices i.e., we can obtain no improvement on the predictions of the rates of return by exclusively utilising information in Ω_t. This premise is the core of the *"Efficient Market Hypothesis"* (EMH) celebrated by Fama's 1970 paper. Depending on the relevance of the information set Ω_t, markets are divided in *weak-form, semi-strong form* and *strong-form* efficient. As already defined in the previous chapters, weak-form means that Ω_t includes past prices and returns alone, semi-strong implies that Ω_t includes all publicly available information and finally, strong-form suggests that Ω_t includes also private information. Strong-form implies semi-strong form and semi-strong implies respectively weak-form efficiency. Clearly the reverse does not hold.

Denoting ϵ_t as the martingale difference (or fair game) process, we can rewrite (1) as:

$$x_{t+1} = x_t + \epsilon_t \tag{3.4}$$

The random walk model states that the price at any time t is equal to the past price at time $t - 1$ plus a residual at time t. Using P_t and ϵ_t to denote respectively the price and the residual observation at time t, we can then write the random walk model as:

$$P_t = P_{t-1} + \epsilon_t \tag{3.5}$$

where $E[\epsilon_t] = 0$ and $\text{cov}[e_t e_{t-s}] = 0$, $\forall\ s \neq 0$. In its general form (3.5) states simply that the best possible "linear" prediction for tomorrow's price is today's price. By relaxing the assumptions about the residual component e_t, one can obtain less general variations of the random walk model. If the sequence of the residuals is uncorrelated, then (3.5) is a second order martingale. If the residuals are independently distributed, then (3.5) is a random walk in a *strict-sense*. Finally, if the e_t is an *i.i.d.*[23] Normal process then (3.5) is a Wiener process.

The martingale model in its form is less restrictive than that of the random-walk model. The latter requires independence of higher conditional moments of the

---

[23]*i.i.d.*: Independent, Identically Distributed

probability distribution of price changes whereas the former demands only independence of the conditional expectation of the price changes from Ω_t. More precisely, the martingale rules out any kind of dependence between the conditional moments of the conditional expectation of the martingale differences $x_{t+1} - x_t$ and the information Ω available at time t, as in (3.2). That enables the distinction between correlation and independence to be realised within a time series of price changes analysis framework. As detected by Mandelbrot (1963) and Fama (1965), price changes though uncorrelated, have a tendency to be dependent over time and exhibit volatile periods interchanged with relatively calm ones. This implies dependence on higher order conditional moments (i.e., variance, skewness and kurtosis).

At this point, it would be insightful to note that Granger and Morgenstern (1970) caution about the relevance of the random walk hypothesis. They point out that the hypothesis does not state clearly (as misconceived by many researchers) that the price changes are unpredictable. On the contrary, it is stated that price changes, in **absolute terms** are not predictable on the basis of any **linear** combination of the past history of these changes. It does not rule out any underlying nonlinear relationships nor contradicts the possibility of predicting the **relative** price change of one stock compared to another. This is a very strong statement as it does not rule out the possibility of the professional ability of some financial analysts to outperform the market (for some period and not on average). The random walk model concentrates on the absolute price changes and makes no statement about the state of the market and its competitiveness.

Equilibrium in financial markets is defined under the framework that martingales and fair games provide. Rates of return are a fair game if and only if the corresponding prices series plus cumulated dividends discounted back to the present, is a martingale. Defining the rate of return as the sum of dividend yield plus capital gain minus one, denoting the dividends as D and defining ρ a constant, the stock price P_t at time t can be written as

$$P_t = (1 + \rho)^{-1} E(P_{t+1} + D_{t+1} | \Omega_t) \tag{3.6}$$

Under (3.5) the present value of stock price equals the sum of expected future price

plus dividends, discounted at rate ρ.[24]

The EMH ensures that all relevant information is incorporated into the prices of stocks and therefore errors in the forecasted prices are zero and uncorrelated with the set of all prior information Ω_t. This is known as the *"orthogonality property"*.

$$E_t \epsilon_t \equiv E_t(P_{t+1} - E_t P_{t+1}) \equiv E_t P_{t+1} - E[E_t P_{t+1}] \equiv 0 \qquad (3.7)$$

A necessary but not sufficient condition for both the random walk and martingale hypothesis to hold is that the price series or their logarithms are I(1) (integrated of order one i.e., they have a "unit root"). There is a vast literature on unit root testing.

The tests of the random walk hypothesis concentrate mainly on tests of the linear relationships within the historical sequences of price changes. Types of nonlinear dependence have not been captured by the random walk literature in its early stages.

3.4 Self-similarity, Affinity, Scale Invariance and Fractals

Before presenting the multifractal processes, it would be useful to outline the basics of the underlying theory. A *Self-similar* process is invariant in distribution under different scaling of space and time. Samorodnitsky and Taqqu (1994) define as "self-similar" a process $X = \{X(t), t \in \mathbb{R}\}$ if, for any $\alpha > 0$, the finite dimensional distributions of $\{X(at), t \in \mathbb{R}\}$ identical to the finite dimensional distributions of $\{a^H X(t), t \in \mathbb{R}\}$. As H we denote a non-negative scaling coefficient which is the index of self-similarity. More analytically, the increments of a random process $\{X(t), t \in \mathbb{R}\}$ satisfy the self-similarity property with parameter $H > 0$ *iff* $\forall t \in \mathbb{R}$

---

[24] As Leroy (1989) argues, (3.5) implies that using the prices P_t themselves without the dividends, generally leads to a formulation that is in conflict with the martingale definition (3.1). Under the fair game concept, the variability of the dividend-price ratio (due to the fluctuations in current dividends related to the variables that predict the future dividends) implies that the variation of the conditionally expected rate of capital gains must have an offsetting effect. This is needed to maintain the nonrandomness of the conditionally expected rate of return. In the general EM literature, the prices are supposed to be following a martingale and are regarded to be including reinvested dividends.

$$X(t_0 + t) - X(t_0) \stackrel{\mathrm{p}}{=} h^{-H}\left[X(t_0 + ht) - X(t_0)\right] \tag{3.8}$$

where $\stackrel{\mathrm{p}}{=}$ denotes equality in probability. The increments of this process $X(t)$ can be stationary or not. For instance, Brownian motion is a self-similar process with index $H = 1/2$ and stationary increments. Other processes can exhibit though *long-range dependence* or *memory* such as fractionally integrated processes.

A random process $X(t)$, $-\infty < t < \infty$, is *statistically self-similar* if its statistics are invariant to dilations (expansions) or compressions of the waveform in time (Wornell, 1995). More analytically, a random process X(t) is statistically self-similar with parameter H if for any $a > 0$, $a \in \mathbb{R}$,

$$X(t) \stackrel{\mathrm{p}}{=} \alpha^{-H} X(\alpha t) \tag{3.9}$$

If this holds for all finite-dimensional joint probability distributions, then self-similarity is defined in a *strict* sense. *Wide-sence* statistical self similarity is interpreted with second order statistics i.e., means and covariances. Thus (3.9) can be redefined as

$$M_x(t) \triangleq E[x(t)] = \alpha^{-H} M_x(\alpha t) \tag{3.10}$$

and

$$R_x(t, s) \triangleq E[x(t), x(s)] = \alpha^{-2H} R_x(\alpha t, \alpha s) \tag{3.11}$$

Geometrically, *affinity* is defined as a transformation which preserves parallel lines. Algebraically, any invertible linear transformation is an affine one. Falconer[25] (1990) defines as *Affine transformation* a mapping S of $\mathbb{R}^n$ on $\mathbb{R}^n$ of the form:

$$S(x) = T(x) + b \tag{3.12}$$

where T is a linear transformation of $\mathbb{R}^n$ and b is a vector in $\mathbb{R}^n$. Mandelbrot (1977) defines self-affinity for a time to vector function $X(t)$ with respect to exponent $\alpha > 0$ and initial time point t_0: $X(t)$ is self affine if there exists an exponent

---

[25] "An affine transformation S is a combination of a translation, rotation, dilation and perhaps a reflection".

$log(r_e)/log(r) = \alpha > 0$ such that $\forall h > 0$, the function

$$h^{-\alpha} X[h(t - t_0)] \tag{3.13}$$

is independent of h.

Self-Similarity and self-affinity[26] are two terms not to be confused. The latter expresses invariance under some linear reductions and dilations whereas the first one corresponds to the well known isotropic (homothetic) family of transformations (reductions).

3.5 Multifractal and $1/f$ processes

Fractional Brownian motion was introduced by Mandelbrot and Van Ness (1968) but together with fractional Gaussian noise models were first proposed by Kolmogorov (in Wornell, 1995). Multifractal analysis focuses on the local singular behaviour of measures of functions in a geometrical and statistical function (Riedi, 1995). Intuitively, multifractals could be of great use in finance as they provide a useful tool for modelling erratic or irregular (intermittent) data. The Multifractal Market Hypothesis states that speculative markets follow a piecewise fractional Brownian motion process. This model emerged in financial literature because of its mathematical properties that appear to fit observed financial phenomena. Intensive scientific research has revealed multifractal patterns in various areas other than finance and exchange rate economics such as in the study of network data traffic, geophysics and turbulence.

Before presenting the multifractal model, an outline of the ordinary and fractional Brownian motion is provided. A gentle and intuitive introduction is found in Mandelbrot (1982). For a more elaborate analysis of these areas, we recommend Falconer (1990), Devaney (1989) and Peitgen et al., (1992). Readings especially focused in finance are Peters (1991, 1994, 199a, 1999b) and more recently Mandelbrot (1997 and 1998) as well as Barnett and Serletis (2000).

Both "ordinary" Brownian (or Wiener) and fractional Brownian motions are

---

[26]As introduced by Mandelbrot (1977).

random processes X_{t_i} of Gaussian increments with mean $E(X_{t_2} - X_{t_1}) = 0$ and variance:

$$var(X_{t_2} - X_{t_1}) \propto |t_2 - t_1|^{2H} \tag{3.14}$$

where the H exponent is allowed to vary between (0,1). If $H = \frac{1}{2}$ then the Brownian motion generalises to a fractional Brownian motion. Increments of X are *statistically self-similar with parameter* H while $(X_t - X_{t_0})$ and $(X_{(rt)} - X_{(t_0)})/r^H$ have the same finite dimensional joint distribution functions $\forall$ t_0 and $r > 0$. Fractional Brownian motion is defined as the stochastic integral, for $t > 0$

$$X_H(t) = \frac{1}{\Gamma(H + 1/2)} \left\{ \int_{-\infty}^{0} [(t - s)^{H-1/2} - (-s)^{H-1/2}]dW(s) + \int_{0}^{t} (t - s)^{H-1/2}dW(s) \right\} \tag{3.15}$$

where $\{W(s), -\infty < s < \infty\}$ denotes a Wiener process extended to the real line. We can classify fractional Brownian motion to three categories according to the value of H: $H < \frac{1}{2}$, $H = \frac{1}{2}$ and $H > \frac{1}{2}$. The price of H is also directly related to the correlation structure of the fractional Brownian motion. If $H = \frac{1}{2}$ then the process is an *ordinary Brownian* motion i.e., exhibits probabilistically independent increments with 0 linear correlation. For $H > \frac{1}{2}$, increments are positively correlated and for $H < \frac{1}{2}$ they are negatively correlated.[27]

The autocorrelation between times t and say $2t$, given the value $L_H(0)$ at the beginning of the subinterval, is defined as:

$$E\{(L_H(t) - L_H(0) - E[L_H(t) - L_H(0)]) \ (L_H(2t) - L_H(t) - E[L_H(2t) - L_H(t)])\} \tag{3.16}$$

More analytically, the logarithm $L_H(t)$ of the asset prices, within each time line subinterval Δ_t, follows a fractional Brownian motion. The sequence of its increments have a zero average and variance $\sigma^2 \Delta_t^{2H}$:

$$L_H(t) = L_H(0) + \frac{1}{\Gamma(H + 1/2)} \int_{0}^{t} (t - s)^{H-1/2}dW(s) \tag{3.17}$$

---

[27]Exhibit more erratic oscillations.

with $0 < s < t$, for a given $L_H(0)$ and $dW(s)$ the differential of a standard Wiener process. The gamma function is introduced in $\Gamma(H + 1/2)$ so that the fractional integral in (3.15) becomes an ordinary one when $H - 1/2$ is an integer. Weighting each independent increment $dW(s)$ by $(t - s)^{H-1/2}$ enables us to calculate at any time t within a subinterval, the price of $L_H(t)$ by summing up all the increments up to time t (given the initial value $L_H(0)$ at the beginning of the subinterval).

An important class of statistically scale-invariant or self-similar random processes is the $1/f$ processes. The family of $1/f$ statistically self-similar random processes are defined as having power spectra that follow a power law of the form

$$S_x(w) \sim \frac{{\sigma_x}^2}{|w|^\gamma} \tag{3.18}$$

where $\gamma = 2H + 1$ some spectral parameter. Mandelbrot (1982) shows that the sample paths of $1/f$ processes are fractals. Graphs of sample paths of fractal processes exhibit high irregularity. Their *topological dimension*[28] is more than unity in the plane. For these processes there is a direct relationship between the self-similarity parameter H and the fractal dimension D. An increase in H causes a decrease in the fractal dimension D. The spectral parameter γ increases with H and this leads to a redistribution of power from high to low frequencies (e.g. trends or long cycles) leading to a smoother sample path graph. For processes like financial time series such as stock index price indices, usually $\gamma \approx 1$ and generally it is expected to vary between 0 and 2. Moreover, $1/f$ processes exhibit statistical dependence. Their autocorrelation function (ACF) decays slowly and this renders modelling methods such autoregressive moving average models (ARMA) inappropriate for identifying their long-range dependence.

For fractional Brownian motions, $1 < \gamma < 3$ and for classical Brownian motions, which is a special case, $\gamma = 2$. Fractional Gaussian processes which correspond to derivatives of fractional Brownian motions, have $-1 < \gamma < 1$ and stationary *white* Gaussian noise is a special case with $\gamma = 0$. For γ 's -1 and 3 we derive degenerate models.

---

[28]The topological dimension of graphs of pure random processes is unity i.e., they are one-dimensional curves in the plane.

3.6 Levy processes

A random function $X(t)$, defined on $[0, \infty)$, is called a *Levy stable motion* if $X(0) = 0$, the increments of the process are independent and stationary and for each $t \geq 0$ and a parameter $\lambda \in \mathbb{R}$, the characteristic function Φ_t is defined by

$$\Phi_t(\lambda) = E\{e^{i\lambda X_t}\} = e^{-t\psi(\lambda)} \tag{3.19}$$

where:

$$\psi(\lambda) = c|\lambda|^\alpha \{1 + i\beta \operatorname{sign}(\lambda) \mathrm{w}(\lambda, \alpha)\} \tag{3.20}$$

with $w(\lambda, \alpha) = \tan(\pi\alpha/2)$ for $\alpha \neq 1$ and $w(\lambda, 1) = -(2/\pi) \log|\lambda|$. In (3.20), $\alpha \in [0, 2]$ is the characteristic exponent, $\beta \in [1, 1]$ is the symmetry parameter and $c > 0$ is the scale parameter.

3.7 Long memory processes

A generalisation of the standard ARIMA(p,d,q) models (Box and Jenkins, 1976) are the *fractional ARIMA* models introduced by Granger and Joyeux (1997) and Hosking (1981). Defining X_t a stationary process, then we can write the dth order backward difference as

$$(1 - B)^d X_t \equiv \sum_{k=0}^{\infty} \binom{d}{k} (-1)^k X_{t-k} = \epsilon_t \tag{3.21}$$

where $d \in \mathbb{R}$ and B is the backshift lag operator. From elementary combinatorics, the number of k size combinations of d objects is given from the following formula:

$$\binom{d}{k} \equiv \frac{d!}{k!(d-k)!} = \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)} \tag{3.22}$$

The simplest form of a fractional ARIMA process is the fractional ARIMA(0,d,0). In this case, the residual term $\{\epsilon_t\}$ is a white noise process with mean 0 and variance σ_e^2. In figure (3.1) we depict the autocorrelation function of a FARIMA sequence which exhibits smooth and slow decay, an indicative characteristic of such a process.
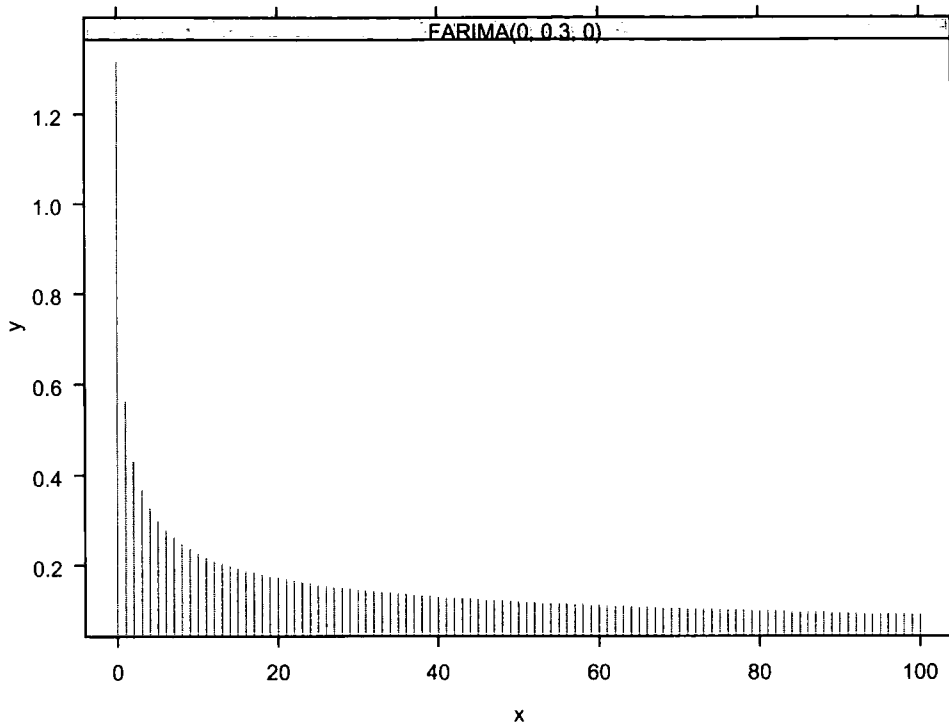
Figure 3.1: The autocorrelation function (ACF) of a fractional ARIMA process, ARFIMA(0,0.3,0).

It can be shown that a $\{X_t\}$ zero mean fractional ARIMA$(0,d,0)$ process with $-\frac{1}{2} < d < \frac{1}{2}$ is stationary and invertible (Hosking, 1981). The variance of such a process is given by

$$\mathrm{Var}\{X_t\} = \frac{\sigma_\epsilon^2 \Gamma(1-2d)}{[\Gamma(1-d)]^2} = s_0 \qquad (3.23)$$

and the autocovariance of $\{X_t\}$ is

$$s_k \equiv E\{X_t, X_{t-k}\} = \frac{\sigma_\epsilon^2 (-1)^k \Gamma(1-2d)}{\Gamma(1+k-d)\Gamma(1-k-d)} \qquad (3.24)$$

3.8 Conclusions

In this thesis we will be dealing with financial time series, mainly closing prices and returns of indices and stocks from various stock markets, and will be exploring

the structure of their data generating processes. For this reason, we concentrated here on the stylised facts about these sequences. We also described the most basic generating processes according to the existing literature, relevant to our subject. Our intention was to provide the necessary theoretical background and scientific jargon for the following discussions on the dynamical structures of the time series that we will examine.

In the following chapters, we will be utilising techniques developed specifically to detect such processes as the ones described here. Since the main theme is "*non-linearity*", in the continuation of the thesis we will be dealing with a more refined issue, and more specifically the existence of *nonlinear determinism*. In the next chapter we introduce the theory behind chaotic deterministic dynamics and show how specific *invariant* measures can be used to produce metrics which enable us to characterise a process as nonlinear deterministic or purely stochastic.

CHAPTER 4
Dynamical systems

4.1 Introduction

In the previous chapters we examined the research so far regarding the dyna-
mics of financial time series as well as their basic structures and characteristics. We
have seen so far that the basic hypothesis of Market Efficiency and linear stochastic
dynamics, does not seem to provide an adequate explanation for the finer details
revealed in stock market data. If one selects to test for stochastic randomness in
these sequences, the alternative hypothesis intuitively is determinism and possibly
chaos. That is why, in this chapter we provide a brief introduction to the theoretical
framework on dynamical systems, chaos and time series embedding. The scope is to
produce a platform necessary to comprehend and critically evaluate the empirical
methodologies, their application and their outcome on real financial data.

We begin by introducing the notion of dynamical systems and their properties.
Following the introductory section, we provide a collection of ideas, theories and
criteria used to classify dynamical systems as nonlinear-deterministic and distinguish
them from other types of systems such as purely stochastic ones. Finally we provide
the description of a set of the most popular statistical-mathematical tools used to
examine dynamical systems through their outputs.

The study of dynamical systems is mainly relevant to dynamical macroeco-
nomic literature. Economic systems are represented there in their analytical form
and one seeks usually the closed form solution that provides information on the
equilibrium conditions or the steady states of such systems. Under such a framew-
ork, economic theory has justified -in some cases- nonlinearity and chaoticity in the
economy as a driving force that leads to stability or deviations from it. The most
famous early model in macroeconomics was the so called "cobweb". That shows the
path to market equilibrium when market forces are incited by some usually external
factor.

Setting up an abstract and complicated mathematical chaotic market model is

not the scope of this thesis. Our aim rather is to investigate empirically and possibly capture any existing nonlinearities or non-stochasticities. These may be apparent in the data generating processes realised within the dynamics of the financial markets. In doing so we utilise some novel techniques which are mostly based on the theory outlined in this chapter. Such methodologies have been used successfully in the physical sciences were the deterministic dynamics of the phenomena are more easily detected. Our intention is to provide an additional empirical analysis framework that could shed new light in the investigation of the nature of financial time series data generating processes. The presentation of the basic theory here is not exhaustive. This is outside the general scope of this thesis. There are many books that cover the issues discussed in this chapter. Our discussions here are guided mainly by Williams (1997), Kantz and Schreiber (1997) and Kaplan and Glass (1995). One of the earlier and prestigious writings on chaotic dynamics is the monograph by David Ruelle (1989). We should also mention the book by Urbach (1999) who seems to be presenting mostly the theory while providing little or no empirical evidence.

4.2   Dynamical systems and Chaos

A *system* can simply be defined as *dynamical* when it evolves from an initial state, according to a set of rules dictated by system of differential equations or discrete difference equations. In the first case, when the solution is continuous in time, it is called a *flow*. In the second case, it is referred to as a *map*.

In the simplest cases, we can characterise the state of the system by a set of real quantities $x_1, x_2, ..., x_d$. Two different instances of the quantities $x_1, x_2, ..., x_d$ and $x'_1, x'_2, ..., x'_d$ represent two different states. The "closeness" of any two states is dependent on the degree of closeness of the x_i and x'_i values of for all i's. If the respective states of a system are termed as being close then we can represent the evolution of that system as a set of ordinary differential equations:

$$\dot{x}_i = f_i(x_1, ..., x_d), \text{ for } i = 1, ..., d. \tag{4.1}$$

Dynamical systems have a finite number of degrees of freedom and the evo-

lution of any system can be represented in its *phase* or *state space*. A phase space is simply a coordinate system formed by all the variables that are used in the mathematical formulation of that system. In an d-dimensional phase space, a set of d-dimensional "embedding vectors" defines a point i.e., a state of the system at a certain instant of its life. The sequence of such instances over time defines the "trajectory" of the system. Using the definition (4.1), we can consider the quantities $x_1, x_2, ..., x_d$ as representing a point x in d-dimensional space. In such a case, we can represent the corresponding state of the system by this point x. This is also known as simply the *phase* point while the entire space is the phase space (or state space). Sequential states of a system can be represented as a motion of a phase point along a curve. This is called the *phase trajectory* or simply *trajectory*.

We can define vectors in phase space combining the $x_1, x_2, ..., x_d$ points to form $f(x)$ vector components such as

$$(f_1(x_1, x_2, ..., x_d), ...f_d(x_1, x_2, ..., x_d)). \tag{4.2}$$

In this way we can abbreviate the system of differential equations (4.1) as

$$\dot{x} = f(x). \tag{4.3}$$

A more useful notation of the system of equations in (4.3) is:

$$\dot{x}(t) = f(\mathbf{x}(t)), \text{ for } \mathbf{x} = (x_1, x_2, ..., x_d) \tag{4.4}$$

This enables us to write the solution to system (4.4) for the continuous case as the flow **F**:

$$\mathbf{x}(t) = \mathbf{F}^t(\mathbf{x_0}) \tag{4.5}$$

where $\mathbf{x_0}$ is the *vector of initial conditions* or the system's "start" and $\mathbf{F}^t$ is referred to as the flow or "flow map".

In the study of dynamical or chaotic systems, it is of primary importance to detect and measure the trajectory of the orbit (or simply the orbit) $\mathbf{x}(t)$ in 4.5. What one usually tries to investigate, as we will see in the following sections, is the

"shape" of this orbit and any deviations from it. The way this orbit behaves will classify the system as chaotic, nonlinear, deterministic, linear or stochastic. Small deviations from the orbit $\mathbf{x}(t)$ are denoted as $\partial\mathbf{x}$. That enables us to define the time evolution of the system in (4.4) as:

$$\dot{\partial}\mathbf{x} = \frac{\partial}{\partial\mathbf{x}}(\mathbf{F}^t(\mathbf{x}_0))\partial\mathbf{x} \tag{4.6}$$

with

$$\mathbf{J} = \frac{\partial\mathbf{F}^t}{\partial\mathbf{x}} \tag{4.7}$$

for $\mathbf{J}$ being the *Jacobian* matrix of f after t steps (or iterations), t being the continuous time variable.

For the discrete case, the map of the dynamical system can be defined as:

$$\mathbf{x}_k = \mathbf{F}^k(\mathbf{x}_0) \tag{4.8}$$

where $\mathbf{x}$ is a d-dimensional state vector, $\mathbf{x}_0$ the initial state vector, t the discrete time variable and $\mathbf{F}^k : \mathbb{R}^d \to \mathbb{R}^d$ the system function. In the discrete case we can show that the Jacobian matrix for the map after k-iterations is as in the continuous case (4.7) :

$$\mathbf{J} = \frac{\partial\mathbf{F}^k}{\partial\mathbf{x}} \tag{4.9}$$

There are two main classifications of dynamical systems according to their "behaviour" in phase space. A "*Hamiltonian*" or *conservative* system retains volume in phase space. Such a *frictionless* system exhibits no loss of energy and continues its motion forever. A "*dissipative*" or "*nonconservative* system is a system that contracts volume in phase space. Generally, if the energy of the system fluctuates with time, then this is called a dissipative system. In such a system, the trajectories converge with time to a bounded subset of the phase space. For this case we can represent the Jacobian as $\mathbf{J} = e^{\mathbf{A}}$ where the eigenvalues of the matrix $\mathbf{A}$ are λ_i for $i=1,2,...,d$. Following this representation, the "trace" of matrix $\mathbf{A}$ will be:

$$\text{Trace}(\mathbf{A}) = \lambda_1 + \lambda_2 + ... + \lambda_d \tag{4.10}$$

For a dissipative system, although the eigenvalues of $\mathbf{A}$ can be both negative and positive, the Trace($\mathbf{A}$) will be a negative value. This implies that the system will be locally unstable.

Trajectories in phase space that lie sufficiently close together (are "concentrated") are said to be lying in the *"basin of attraction"* and the bounded region to which all trajectories converge is called an *"attractor"*. Basically, any limit set which collects these trajectories is an attractor. In the study of dynamical systems, one is interested in a well bounded sub-region (or sub-space) of the phase space that "collects" the trajectories of the system's orbits. That is where the system's dynamics are said to be concentrated and that has important implications for the characterisation, modelling and study of the system.

Attractors are classified according to their topology. There are three basic types of attractors: *points, limit cycles* and *torus*. When all trajectories converge to a point in phase space, that is termed a *"fixed point"* attractor. In the case of the attractor being a closed curve due to a periodic attitude of the system, this is called a *"limit cycle"*. If the system exhibits a quasiperiodic motion, the attractor is some kind of a closed surface, representing a kind of a *"torus"*. A limit cycle attractor is one dimensional whereas a torus is a two dimensional one. Finally, when the attractor is of a non-integer dimension, this is termed as a "strange attractor" and its dimension is called a *"fractal"* dimension. The main property of such attractors is *their sensitive dependence on initial conditions*. Small perturbations in the initial states of the system can lead to entirely different final states for that system. In such a case, trajectories that begin closely in phase space, diverge exponentially after some iterations. There is a kind of a confusion in the literature about the "dimensionality" of dynamical systems, especially in sciences which are not purely mathematical such as the social or medical sciences. In these, one can come across a number of papers that endeavour to discover nonlinear determinism or dynamics on areas that were believed to be governed by stochastically driven systems. In such instances, the term *"dimensions"* may be used to express the number of degrees of freedom, the number of variables or the number of differential equations needed to explain the dynamics of the system. What is of importance here is to note

and emphasise that when the issue of *"attractor dimensionality"* will be addressed throughout this thesis, we will be implying the *dimensions of the sub-region of the phase space where the attractor lies*. This effectively describes the topology of the sub-manifold where the dynamics are concentrated.

Another useful measure for the classification of the behaviour of dynamical systems is the set of *Lyapunov exponents* for that system. The eigenvalues λ_i for $i = 1, 2, ..., d$ are related to the logarithms of the eigenvalues of the linearised dynamics across the attractor which are also termed as the Lyapunov exponents. If the sum of the Lyapunov exponents is positive, the dynamics of the system will follow an explosive path end there will be no attractor. In such a case, the system is non-chaotic. In the opposite case, the system is termed *chaotic* and the steady state trajectories form the attractor of such a system in the phase space. As we mentioned earlier, this attractor has a non-integer dimension and it is popularly defined as a *"strange attractor"*. On such attractors, the dynamics are characterised by *stretching* and *folding* of the phase space. Stretching indicates the divergence of nearby trajectories and folding indicates the concentration of the dynamics in a finite region of the phase space. The contraction of an attractor in phase space i.e., the strength of the concentration of trajectories to the attractor is indicated by the magnitude of a negative Lyapunov exponent.

4.2.0.1 Definition of Chaos and stylised facts

The meaning of the Greek word[29] "Chaos" is easily misunderstood as "disorder" and "unpredictability". In the Encyclopedia of Mathematics (Hazewinkel, 1997), the following definition applies:

"Chaos describes a situation where typical solutions (or orbits) of a differential equation (or typical solutions of some other model describing deterministic evolution) do not converge to a stationary periodic function (of time) but continue to exhibit a seemingly unpredictable behaviour."

---

[29]The word "chaos" apparently has also lead to the making of the word "gas" which was created to show the nature of disorder in the movements of the gas molecules.

It is obvious that the term refers to dynamical deterministic systems and may thus be used as an alternative paradigm to stochastic systems in various other sciences. There are some general and basic properties or stylised facts about chaotic systems[30]:

1. Only deterministic processes can produce chaotic results.

2. Only nonlinear systems have been proved (so far) to produce such processes that lead to chaos.

3. Feedback in systems such as recurrent structures or lagged effects, can cause dynamics that may lead to deterministic chaos.

4. Chaotic systems may be very simple and exhibit sensitive dependence on initial conditions. It is not possible though to determine a chaotic system's prior history.

5. Chaotic systems allow relatively accurate short term predictability.

6. The dimension of their attractor is non-integer (or fractal). This means that the phase space trajectory where the dynamics are concentrated, lies on a hyperplane (or manifold) with non-integer dimensions..

7. Chaos is a self generated process and not necessarily caused by external factors.

8. Chaos is broad-band. This implies that there is a broad Fourier spectrum such as for an uncorrelated noise process, with some peaks due to some periodicity.

9. Chaotic processes can be controlled.

The first property implies that small perturbations on initial conditions may lead to entirely different paths of evolution for a chaotic system although such a system is deterministic. The second fact refers to what has been mentioned earlier about the local instability and the positive Lyapunov exponents of such systems. The last property points out the "strangeness" of the attracting set of the dynamics of chaotic systems. There is extensive literature which demonstrates those key

---

[30]Williams (1997) provides a rather exhaustive list. Here we include the most interesting points.

properties (see Devaney, 1989 and Falconer, 1990). The most famous chaotic systems are the *Lorenz, Hènon* and *Rössler* systems.

4.2.0.2 Chaotic attractors

We defined the notion of chaotic attractors earlier in this section. Here we outline the basic characteristics in more detail. A chaotic attractor is defined in Grebogi *et al.*, (1992) as "*a complex phase space surface to which the trajectory is asymptotic in time and on which it wanders chaotically*". Eckmann and Ruelle (1985) prefer a different definition, that of extreme sensitivity to initial conditions. Generally, chaotic attractors are characterised by their distinctive features:

- A cycle-trajectory in a chaotic regime exhibits complexity. A trajectory may never repeat it self (overlap). This is called "aperiodicity".

- The trajectories of chaotic attractors never intersect themselves.

- Trajectories that are close in time may diverge exponentially in the future and lead to entirely different states for the system. This is due to the sensitivity to initial conditions.

- The dimensions of a chaotic attractors are usually noninteger quantities.

In figure (4.1) we show an example of an attractor, the projection of the Hènon attractor[31] in 2 dimensions. In figure (4.2) we show the Lorenz attractor, projected in 3 dimensions.

As explained earlier, chaotic attractors will exhibit fractal dimensions. In chaos theory, the word "fractal" was introduced by Benoit Mandelbrot (1977 and 1982) in order to signify this fractional nature of the dimensions. Fractals are usually complicated graphs which are derived from the complex solutions of the set of equations representing the dynamics of chaotic systems. They exhibit all the important characteristics of chaotic systems, namely complexity and self-similarity. The graphs of fractal are usually very elaborate but not informative or useful when

---

[31]The Hènon map is defined as $(X, Y) \rightarrow (1.4 - X^2 + 0.3Y, X)$ (see Nusse and Yorke, 1994).
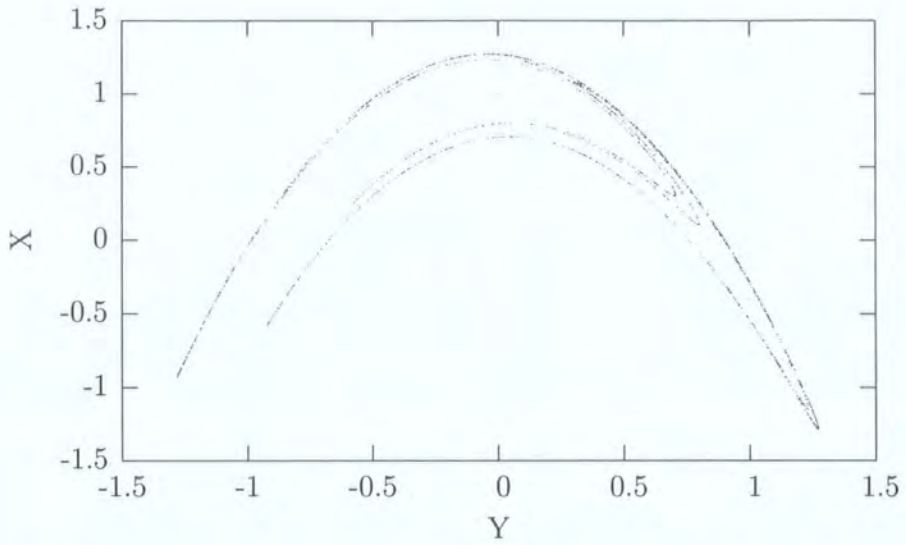
Figure 4.1: The Hènon chaotic attractor. An example of how orbits of trajectories concentrate in a small subspace of the phase space without overlapping or crossing each other.
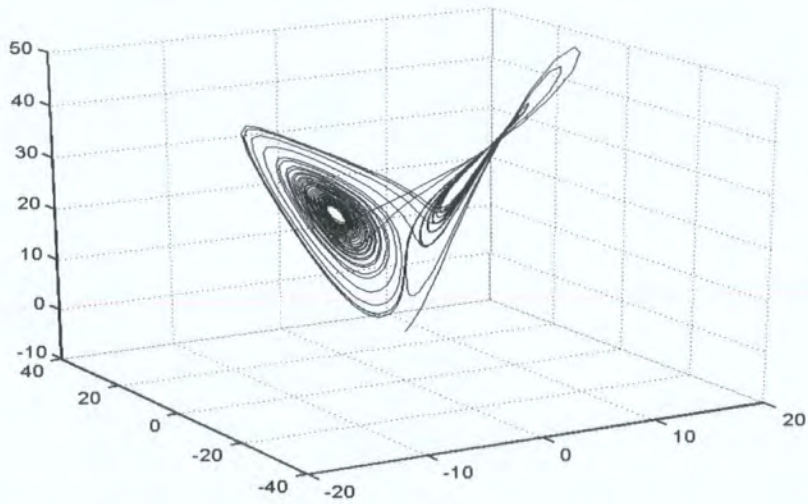
Figure 4.2: The Lorenz attractor represented in three-dimensional phase-space (constructed with Matlab).

Figure 4.3: The Mandelbrot fractal, also known as the "Gingerbread man".

it comes to financial time series analysis. Fractals are derived as graphical representations of solutions and are not really structures that can be fitted to time series so as to reveal chaotic data generating processes. In order to reveal determinism or chaos we really need to follow other paths, such as the reconstruction of dynamics based on pseudo phase-spaces derived from observed quantities. This is a complex methodology based on topology and embedding theorems by Takens and Whitney and will be discussed in the following pages.

Concluding our brief discussion on attractors, we present a set of fractals which are parts of the original fractal graph based on the basic Mandelbrot (1982) fractal in figure (4.3). Figures (4.4) and (4.5) show close-up sections ("zoomings") on the original fractal. It is easily understood that we can obtain even more detail by magnifying the original fractal and reveal patterns that exhibit great similarity with the original. All graphs were generated with Winfract version 18.21 for windows, a freeware package accompanying Wegner *et al.*, (1992).

4.2.1 Ergodic Theory

Ergodic theory is very important for the classification of phenomena as chaotic. The theory is based on probability statements concerning the state of a dynamical
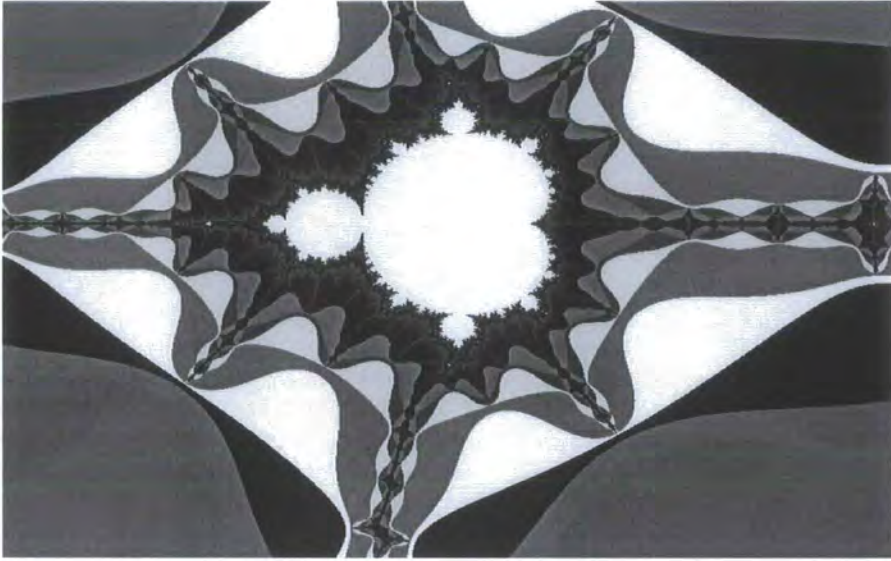
Figure 4.4: A magnification of an area from figure (4.3). This is a part from the tail on the left half of the fractal in figure (4.3), the area contained in a white box.

Figure 4.5: Another closeup of an area from figure (4.4) this time. Notice the self-similar structures repeating themselves along the edges of the fractal.

systems in space and time. This is achieved by the equation -when feasible- of space and time averages derived from the system's activity. Ergodicity lies behind the equation between the probability that a dynamical system is at some state A (i.e., some set of possible values describing this state) and the limiting proportion of time spent by the process in A. The key is to find an *invariant* probability distribution P, with respect to the dynamic function f, that generates the observed dynamics. If X is a random variable with distribution P, then (the mapping) $f(X)$ will also be of the same distribution. Under these conditions, the function f is said to be *measure-preserving* with respect to P.

What is of interest in dynamical system analysis is to be able to define quantities that are invariant under smooth coordinate transformations in phase space. This implies that whatever is our starting point on the attractor, we should always derive the same value for these quantities. The invariance of these particular measurements is what exactly ergodicity allows. In plain terms, it allows for the "rotation" of the phase space and helps us thus obtain different views of the dynamics of the system without altering their nature. It follows that it is easier to observe ergodicity in dissipative systems where dynamics are concentrated in a small portion of the phase space which we call attractor or basin of attraction. A measure can thus be defined that will calculate the time spent by a trajectory in a specific area of the phase space. Ergodicity can be established when this measure is independent of the initial conditions. *In other words, ergodicity ensures that time averages taken over a typical trajectory of the system will equal phase space averages.*

In a similar way of defining the mean of any sample sequence of a random variable, we can define the *time-average* of a random process over an interval $[-T, T]$ as:

$$\frac{1}{2T} \int_{-T}^{T} x(t)dt \tag{4.11}$$

As T tends to infinity, for many random processes, the time-average will reach the *ensemble*-average[32] $\mu_x = \frac{1}{n} \sum_{i=1}^{n} x_i$. Such processes are called *ergodic in mean* random processes. The time-average is then defined as an average over an infinite

---

[32]Or phase-space average.

interval

$$M_x = \langle X(t) \rangle = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t) dt \qquad (4.12)$$

The time-average for the second moment is similarly defined as

$$\langle X^2(t) \rangle = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x^2(t) dt \qquad (4.13)$$

An ergodic process is wide-sense stationary i.e., the expected value is independent of the time origin $(E[x_i] = E[x_{j+i}], \forall\ i, j)$ and the correlation and covariance between any two subsequences must depend only on the time difference between them. The inverse does not necessarily hold i.e., not all wide-sense stationary processes are ergodic. These are important notions for time series analysis as any inference is based on a set of sample statistics. This implies that the information we retrieve is from an incomplete realisation of an unknown data generating process. Without the assumption of ergodicity and thus wide-sense stationarity, we would not ever be able to estimate any parameters of the random process.

Ergodicity is not a testable property of time-series. It is usually assumed. For a more rigourous treatment of the subject one may refer to the books by Tong (1990), Granger and Newbold (1986) or the more recent Medio and Lines (2001).

4.3 Phase Space Reconstruction

A scalar measurement is a projection of the unobserved internal variables of a system onto an interval on the real axis. It is preferable to reconstruct a new phase-space so as the attractor in that space is equivalent to the original one. In order to guarantee that the quantities computed for the reconstructed attractor are identical to those in the original state space, the structure of the tangent space i.e., the liberalisation of the dynamics at any point in that state space, must be preserved by the reconstruction process.

To explain how phase space can be reconstructed in practice, we need to introduce a number of measures that help us determine the general shape and size of this space and the dynamics that govern the system. These are the various dimension calculation measures, the Lyapunov exponents and a number of entropy-

based criteria. We discuss these in the following sections.

4.3.1 Invariant measures

The central idea behind the chaotic analysis of dynamical systems is that their seemingly stochastic i.e., random behaviour can be modelled by a system with a small "number of modes", "degrees of freedom" or "dimensions". The key element is that although such systems may be fairly complicated and complex in structure, with relatively calm periods succeeding erratic and oscillatory behaviour in an aperiodic fashion, near to the transition to this "unstable" behaviour, one could describe adequately the state and dynamics of the system using a "low-dimensional" model. In the case of artificially constructed dynamical systems with a-priori known structure, the dimension can be easily established. In the opposite case, dimensionality must be determined by observation of the system's output. For example, if the system exhibits a stable oscillatory behaviour, its dynamics can be represented by some kind of a limit cycle attractor in the appropriate phase space, its topology being expected to be a one dimensional closed curve. One of the most popular approaches, as we will see in the progress of this thesis, is to examine the dimensionality of the attractor in a reconstructed phase space from the observable output time series of the system. A kind of a "pseudo phase space" that retains the basic characteristics of the dynamical structure and attitude of the system under examination.

4.3.1.1 Euclidian Dimension d and Similarity dimension D_S

In order to comprehend the dimensionality of a dynamical system's attractor, we have to perceive it as an object with a compact "volume" of a certain size that we wish to measure. The attractor will be an orbit, a trajectory of a set of N equiprobable and distinguishable states (Shaw, 1984), localised in the phase space within a region of radius E. If each single state can be contained in a "sphere" of radius e, the dimension d of the region containing all N states is related to the total size:

$$N = \left(\frac{E}{e}\right)^d \tag{4.14}$$

Taking the logarithms on both sides and solving for d:

$$\log(N) = d \log\left(\frac{E}{e}\right) \Leftrightarrow d = \frac{\log N}{\log(E/e)} \tag{4.15}$$

More intuitively, let us first recall that a line, a plane surface and a cube have dimensions 1, 2, and 3 respectively. This integer dimension is the *Empirical* or *Euclidian dimension* and it coincides with the *degrees of freedom*. If we now imagine a line segment, a square and a cube divided into similar objects (forms) of half their spatial ("Euclidian") size, their numbers will be 2^1, 2^2 and 2^3 respectively. We denote these exponents which pertain to the Euclidian dimension, d. Thus, if an object consists of n^d smaller objects of size $1/n$ of the same form, d is the dimension. Consequently, when an object consists of b similar objects of size $1/a$ then the *similarity dimension* D_s is:

$$D_s = \frac{\log b}{\log a} \tag{4.16}$$

The similarity dimension can also take non-integer values and is helpful in determining the dimensionality of *fractal* shapes. As a measure, it is cumbersome and difficult to compute but is the basis of all other dimension definitions.

4.3.1.2   Hausdorff Dimension D_H

This measure is based on the topology concept of *covering* sets. For $D > 0$ and $\epsilon > 0$ real numbers, we construct a covering of a set S by k *spheres*[33] with diameters smaller than e and radii r_1, r_2, r_3,...,r_k. The D-dimensional Hausdorff measure is:

$$M_D(S) = \lim_{\epsilon \to \infty} \inf_{r_k < \epsilon} \sum_1^k r_i{}^D \tag{4.17}$$

The Hausdorff dimension D_H is the value of D for which this measure varies from infinity to zero. It is a generalisation of the Euclidian and Similarity (d, D_S) dimensions and can be applied to any set of points through the method of covering. It is obvious that for D_H, not only the mathematical determination but also the

---

[33] D-dimensional closed surfaces.

computational estimation is difficult.

4.3.1.3 Capacity (Kolmogorov) Dimension D_0 and Box-counting dimension

Like the Hausdorff dimension, the *Capacity dimension* is based on the covering of a set by spheres. It has been introduced by Kolmogorov (1959) and is also called *Kolmogorov dimension*. It is defined as:

$$D_0 = \lim_{\epsilon \to 0} \frac{\log N(S, e)}{\log 1/\epsilon} \tag{4.18}$$

where $N(S, e)$ is the minimum number of spheres of size $1/\epsilon$ needed to cover the set S. The previous definition is equivalent to:

$$N(S, \epsilon) \propto (1/\epsilon)^{D_0(S)} \tag{4.19}$$

This expression demonstrates the *power law* relation between the number of spheres and their size and exists in all definitions of fractal dimensions.[34] When all sizes ϵ of spheres are constant, the Hausdorff dimension D_H is equal to the Capacity dimension. The following relation holds:

$$D_0 \leq D_H \leq d \tag{4.20}$$

The Capacity dimension is a geometric and not a probabilistic measure. In this sense it does not take into account the frequency (clustering) of points in the covering spheres and does not capture efficiently the finer structure of the fractal shape of the attractor.

4.3.1.4 Information Dimension D_1

An extension of the definition of the Capacity dimension is the *Information dimension* which is based on entropic measures from *Information theory*. As a

---

[34] *Scale invariance* for fractal shapes implies the existence of such a power law as the only scale invariant function is power function (*def*: a function $f(x)$ is scale invariant if $f(x) \propto f(\lambda x)$ for all λ.

measure itself, it is suitable for the analysis of stochastic and uncorrelated data. We recall from Information theory that the amount of information associated with the occurrence of an event q with probability $P(q)$ is given by

$$I(q) = -\log P(q) = \log 1/P(q) \tag{4.21}$$

Under certainty of the occurrence of q, $P(q) = 1$ and $I(q) = 0$. Under uncertainty, as $P(q) \to 0$, information $I(q) \to \infty$. Moreover, for two independent events p and q, $I(p + q) = I(p) + I(q)$.

Covering a set S with the minimum number $N(S, \epsilon)$ of spheres of size ϵ, we can calculate the probability $P_i(S, \epsilon)$ of a random point of the attractor to reside in the ith sphere. The associated information for that event is $I_i(S, \epsilon) = -\log P_i(S, \epsilon)$. The *information entropy* i.e., average information for every point to be in a specific sphere is

$$I(S, \epsilon) = \sum_{i=1}^{N(S,\epsilon)} -P_i(S, \epsilon) \log P_i(S, \epsilon) \tag{4.22}$$

The Information entropy is a measure of predictability. In other words, it measures (see Weaver and Shannon, 1949 and Shannon, 1948) the information needed to observe the system at a certain state i i.e., our degree of ignorance about the system's state. It is also called the *"Shannon information criterion"*. In the case of uniformly distributed stochastic (random) data in S, $I(S, \epsilon)$ takes a maximum value which implies minimum predictability. In the case of the data clustering in some area of the phase space, for example a certain sphere, then the average information is zero. The information dimension is defined as:

$$D_1 = \lim_{\epsilon \to 0} \frac{I(S, \epsilon)}{\log 1/\epsilon} \tag{4.23}$$

Again the information dimension is a generalisation of the capacity dimension D_0. In the case that all probabilities of points residing in spheres are equal,[35] $D_0 = D_1$. Moreover, if this distribution is in d-dimensional space, then $D_1 \equiv d$ which means that the size of the spheres is directly related to the probability of points residing in

---

[35]Uniform distribution of points.

them i.e., $P_i(S, \epsilon) \propto \epsilon^d$. In the opposite case of non-uniformly distributed points in the phase space, the information dimension is smaller that the capacity dimension, $D_1 \leq D_0$.

4.3.1.5 Correlation dimension D_2

The most widely used attractor descriptive parameter is the "correlation dimension" (D_2). This complexity measure is related to the system's *topological dimension*, or more intuitively, with its degrees of freedom. It is useful in determining whether the system under examination is stochastic or deterministic. For the latter kind of systems, D_2 converges to a finite value whereas for randomly behaving systems, it does not converge to any number. That has established D_2 as a commonly used statistic for determining the noisy or deterministic nature of any dynamical system (for more discussion see works by Theiler, 1987, 1988a and 1988b, Schuster, 1984, Sauer and Yorke, 1993, Ruelle, 1989, Abarbanel, 1995 and Ding *et al.*, 1993).

The Grassberger and Procaccia (1983) method for determining D_2 has been the most popular in chaotic literature. The first step is the *phase space reconstruction* via *time delay coordinate embedding* which will be discussed in more detail in section 4.3.2.1. Briefly we can describe that as in Takens (1981), from a single observable output series of the system, we construct n-dimensional vectors:

$$\vec{x} = \{x(t), x(t+\tau), ..., x(t + (n-1) \cdot \tau)\} \tag{4.24}$$

where τ a fixed increment called the *time delay* and n is the *embedding dimension*. The *correlation integral* $C(e)$ is then defined as a function of the distances between these "reconstructed" vectors as:

$$C(e) = \lim_{N \to \infty} \frac{1}{N^2} \sum_{i \neq j} \Theta(e - |\vec{x_i} - \vec{x_j}|) \tag{4.25}$$

where N is the sample size and Θ is the Heaviside function. $C(e)$ is the probability measure that two arbitrary points $\vec{x_i}$, $\vec{x_j}$ will close by a distance less than e. The correction to this expressed by Theiler (1986) is to use vectors that are distanced by at least W data points ($|i - j| > W$), where W is a measure of the temporal

correlation of the signal. In that way we may avoid temporal spurious correlations. If an attractor exists and it is a curve in the phase space, $C(e)$ will be proportional to e^1. In the case of the attractor being a two dimensional surface, the correlation integral will be proportional to e^2 and if it is a fixed point then $C(e) \sim e^0$. In general:

$$C(e) \sim e^{D_2} \tag{4.26}$$

where D_2 is as we defined earlier the correlation dimension. This means that for small e, the correlation integral can be approximated by a power of the distance e. For sufficiently large number of observations and the embedding dimension, we obtain:

$$D_2 = \lim_{e \to 0} \frac{\log C(e)}{\log e} \tag{4.27}$$

The common practice is to plot $\log C(e)$ against $\log e$ and determine D_2 from the slope of the curve. For an attractor with unknown topology it is necessary to calculate an array of $C(e)$ for several different embedding dimensions (see comments by Sauer and Yorke, 1993). If the signal is deterministic, the correlation dimension should converge towards a value as the embedding dimension increases.

4.3.1.6   Kolmogorov Entropy

The most important measure that characterises chaotic motion is the *Kolmogorov Entropy* or K-Entropy (*KE*). Classical information entropy from thermodynamics measures the lack of order in a physical system, or our degree of surprise of finding the system at a certain state and was defined as in equation (4.22). To calculate *KE* we need the vector $\vec{x}(t)$ which comprises of the points $X_i(t)$ for $i = 1, ..., d$ that define the trajectory of a dynamical system's attractor in d-dimensional space. The phase space is partitioned into hypercubes of size e^d. We define probabilities $P_{i_0,...,i_n}$ that a trajectory resides in hypercube i_0 at time $t = 0$, i_1 at $t = \tau$, i_2 at $t = 2\tau$, etc. etc. According to Shannon, we can define a quantity

$$K_n = -\sum_{i_0,...,i_n} P_{i_0,...,i_n} \ln P_{i_0,...,i_n} \tag{4.28}$$

which is proportional to the information needed to locate the system on a specific trajectory $i_0, ..., i_n$ with precision e^d (Schuster, 1984). The quantity $K_{n+1} - K_n$ is the information needed to predict which hypercube will contain the trajectory at $(n + 1)\tau$ given the trajectories up to time $n\tau$. This is the loss of information from state n to state $n + 1$. The average loss of information is given by K-entropy which is defined as:

$$K \equiv \lim_{\tau \to 0^+} \lim_{e \to 0} \lim_{N \to \infty} \frac{1}{N\tau} \sum_{n=0}^{N-1} (K_{n+1} - K_n) \qquad (4.29)$$

As $e \to 0$, k becomes independent of the specific box/hypercube and for maps with discrete time steps we can omit τ. K for random systems where initially adjacent points are distributed with equal probability is infinite. For deterministic chaotic systems where the divergence is exponential, it is larger than zero and constant.

4.3.1.7 Lyapunov exponents

As we have described earlier in this chapter, *Lyapunov exponents* (Wolf *et al.*, 1985) are useful for the characterisation and classification of attractors. They provide us with a quantitative measure of the existence and the level of chaos in a dynamical system. The magnitude of the Lyapunov exponents enables us to quantify the dynamics of a systems attractor in information theoretic terms. That is, the measure the rate at which a system's evolution creates or destroys information. Moreover, Lyapunov exponents measure the *mean exponential divergence* of initially close orbits (trajectories) in phase space with time. The more rapid this divergence is within a certain period of time, the more chaotic the system is - this being a clear indication of sensitive dependence on initial conditions.[36]

There is one Lyapunov exponent for every dimension in phase space. A positive Lyapunov exponent implies divergence i.e., a form of "stretching" in the phase space. In this way one can establish how rapidly nearby points diverge from one another. Inversely, a negative Lyapunov exponent implies "contraction" and indicates the speed with which a system returns to balance after it has been perturbed. To illustrate, in the case of three-dimensional dynamics with a point attractor, three

---

[36]This is a prerequisite for the existence chaos in a system by definition (see section 4.2.0.1 and page 52).
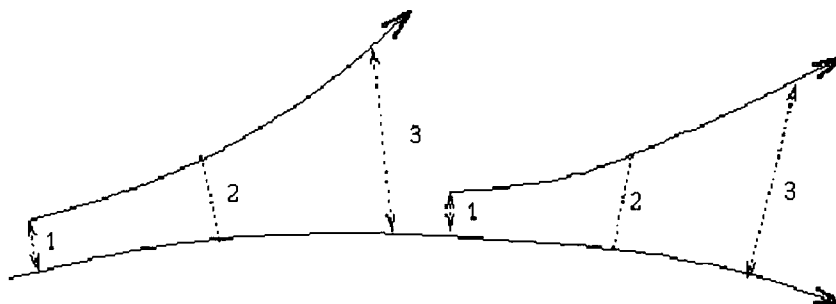
Figure 4.6: Sensitivity to initial conditions makes close trajectories diverge exponentially with time: Divergence(at time t)=Divergence(at time 0)$\times e^{\lambda t}$ where λ is the Lyapunov exponent.

negative Lyapunov exponents exist, as all three dimensions will "collapse" to a fixed point in the phase space. For the same dynamics with a limit cycle attractor, two negative and one zero exponents exist as convergence is only in two dimensions. In the case of a strange attractor, the exponents are one negative, one positive and a zero one. A positive (+) exponent indicates a local average rate of expansion, a negative exponent (-) indicates contraction whereas a zero exponent (0) indicates the direction of the flow itself. The above information is summarised in the following list which shows the signs of the spectrum of Lyapunov exponents for a three-dimensional case:

1. Fixed point attractor: (-,-,-)

2. Periodic limit cycle: (0,-,-)

3. Quasiperiodic attractor (or torus): (0,0,-)

4. Chaotic attractor: (+,0,-)

As described earlier, the negative exponent causes diverging points to remain within the range of the attractor whereas the positive allows sensitive dependence on initial conditions and therefore a *steady state* is defined by how far points will diverge before they obtain again a stable distance. In figure (4.6) we show graphically this exponential divergence.

Lyapunov exponents are defined in terms of the long time average rate of evolution of a perturbation to a system's orbit trajectory. We can deduce from the discussion above that there exists a relationship between Lyapunov exponents and the eigenvalues of the linearised dynamics averaged across the attractor (Kugiumtzis *et al.*, 1993 and 1994, see also 1997a and 1997b for further discussions). For a continuous system, the eigenvalues of (4.7) are:

$$l_i(t) = e^{\lambda_i t} \tag{4.30}$$

where λ_i is the system's dynamics rate of contraction-expansion in phase space. Following (4.30) one can define the Lyapunov exponents for that system as:

$$\lambda_i = \lim_{t \to \infty} \frac{1}{t} \ln |l_i(t)|, \text{ for } i = 1, 2, ..., d_{em} \tag{4.31}$$

where d_{em} is the embedding dimension which is equal to the number of Lyapunov exponents when the limit exists. Similarly, for the discrete case (4.9), the Lyapunov exponents are defined as:

$$\lambda_i = \lim_{k \to \infty} \frac{1}{k} \ln |l_i(k)|, \text{ for } i = 1, 2, ..., d_{em} \tag{4.32}$$

For the simplest case of a 1-dimensional map, the Lyapunov exponent can be defined as:

$$
\begin{aligned}
\lambda(x_0) &= \lim_{n \to \infty} \frac{1}{n} \ln \left| \frac{\mathrm{d} f^n(x_0)}{\mathrm{d} x_0} \right| \\
&= \lim_{n \to \infty} \frac{1}{n} \ln \left| \prod_{i=0}^{n-1} f'(x_i) \right| \\
&= \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln |f'(x_i)|.
\end{aligned}
\tag{4.33}
$$

The third expression of equation (4.33) explains that the Lyapunov exponent is the mean growth rate of the infinitesimal distance between trajectories along a reference trajectory through x_0 (Diks, 1999). When the systems are ergodic, the Lyapunov exponent has the same value for almost all choices of the initial conditions for x_0. For

this reason the latter is usually dropped as an argument of λ. For larger dimensions than unity, the k Lyapunov exponents are defined as the mean logarithms of the moduli of the eigenvalues of the Jacobian of the map along the reference trajectory. When we refer to a spectrum of Lyapunov exponents, usually we order these by decreasing order, i.e., $\lambda_1 > \lambda_2 > \lambda_3 > ... > \lambda_k$. In a similar fashion we define the spectrum of Lyapunov exponents for continuous-time dynamical systems. In this last case, there will always be a zero exponent as any perturbation along the direction of the flow will exhibit a mean growth rate of zero.[37]

In theory, when the evolution equations of the dynamical system are given, the obtaining of estimates for Lyapunov exponents and/or their spectrum is usually straightforward. In cases such as in this thesis, where Lyapunov exponents must be obtained from dynamics which are reconstructed from empirical time series, such a task is extremely difficult (Diks, 1999). Since the seminal work of Wolf *et al.*, (1985), several methods have been proposed for the estimation of the largest Lyapunov exponents from a time series. These methods are very sensitive to the presence of noise and require huge amounts of data, a fact that has been repeatedly ignored in the relevant financial and economic literature.

An interesting relationship described by Pesin (1977) states that the information entropy K_1 is the sum of the positive Lyapunov exponents in such a way that the largest Lyapunov exponent coincides with a positive value of K_1. That implies that for 1-dimensional maps, the Lyapunov exponent as defined in equation (4.33), will be equal to the information entropy.

4.3.2 How to reconstruct the phase space

As we already discussed, embedding theorems make possible the reconstruction of the phase space dynamics. Given a single output sequence of scalar numbers from a dynamical system, we can construct a "pseudo-phase space" that would encompass all dynamics. This requires and infinite amount of noise-free observations that are generated from a system's operation. In practice, obtaining noise free data in vast amounts is very difficult if not impossible. This applies most to social sciences.

---

[37]As long as the attractor is not a fixed point.

When one has in hand a time ordered sequence of scalars, what actually is observed is not a complete picture of the system that generated them but just a part of its output. The problem is to try to "guess" the dynamics from this sequence which is the projection of these dynamics onto $\mathbb{R}$. Due to this projection, a large part of the information about the dynamical system has been "distorted" i.e., transformed, masked or even lost. There is though a way to identify a specific flow of information from the unobserved dynamics or state variables of the system to the projections, if our measurements on the projected sequence are unaffected from this "distortion".

Topological measures such as dimensions, Lyapunov exponents and entropies are only invariant under smooth non-singular transformations. In this sense, dynamics can be reconstructed adequately either through obtaining a large set of very good quality data or by denoising these very carefully so as not to destroy the fine details of the attractors while reconstructing the dynamics. This implies that for transforming-denoising the original data, we should only use functions which are smooth (continuous) and invertible i.e., *monotonic* everywhere. Under this conditions, the pseudo-phase space we reconstructed from the filtered observations will mimic most faithfully the true dynamics of the system. In practice this is easier said than done. Various techniques have been proposed. Most of these are based on the *embedding* of time series obtained from the dynamics os the systems under examination. We already briefly discussed time delay embedding earlier. This is the most popular method for time series embedding. Below, we will be discussing the issue to more extent. Embedding is defined as follows:

Definition: An *embedding* to a compact smooth manifold A into $\mathbb{R}^m$ is a map F which is a one-to-one *immersion* on A i.e., a one-to-one C^1 map with a Jacobian $DF(x)$ which has full rank everywhere.

The crucial question here is under which conditions the projection due to the scalar measurements and the subsequent reconstruction by time delays, forms an embedding.

Whitney (1936) proved that every D-dimensional smooth manifold can be

embedded in the $\mathbb{R}^{2D+1}$ for integer D, and that the set of maps forming an embedding is a dense and open set in the space of C^1 maps. Sauer *et al.*, (1991) extended this to fractal sets A with box counting dimension $D_F < d$. They show that almost every C^1 map from A to the $\mathbb{R}^m$ with $m > 2D_F$ forms an embedding. These theorems and their extensions are of great importance. They simply enable us to recreate the dynamics of a very wide range of dynamical systems in some carefully constructed "pseudo" phase-space, from a single observable sequence.

4.3.2.1 Time-delay embedding

Time-delay embedding, also known as *delay coordinate embedding* or *method of time delays*, is the most popular phase-space reconstruction method. It was introduced by Packard *et al.*, (1980) but apparently developed independently by Ruelle who did not publish. In this section we briefly present the mathematical framework that supports this reconstruction methodology.

The Whitney (1936) embedding theorem and its generalisations refer to sets and spaces. In time series analysis we deal with sequences of numbers, which are projections of dynamics from higher dimensions i.e., of the true trajectories of the dynamical orbits in the original phase-space. We term these true orbits $\mathbf{z}(t)$. What we actually observe is time series $x(t)$[38] which is a mapping of $z(t)$ to $\mathbb{R}$ via a function $h(\cdot)$:

$$x(t) = h(\mathbf{z}(t)) \qquad (4.34)$$

We need to unfold the original dynamics solely from process $x(t)$. This requires a generation of a matrix which will contain as columns, vectors which are *lagged subsamples* of the original sequence $x(t)$. The length of those subsamples is determined by a lag or *delay* parameter, usually denoted τ. This parameter is straightly related to the *embedding window* . We construct n of these lagged subsamples preserving the time order of the observations in $x(t)$ and construct vector (4.24) given here in

---

[38] We assume for simplicity that $x(t)$ is homogeneously sampled in time, i.e., that we have no missing observations. The approach described here can be extended to cover the case of non-uniformly sampled data but we find this outside the general scope of this thesis.

slightly different notation:

$$\mathbf{x}(t) = \{x(t), x(t+\tau), x(t+2\tau), ..., x(t+(n-1)\tau)\}' \tag{4.35}$$

The *embedding window* is thus:

$$\tau_n = (n-1)\tau \tag{4.36}$$

In this way we obtain n observations from $x(t)$, skipping every τth. Vectors $\mathbf{x}(t)$ are filled with repeated observations from $x(t)$ in the sense that two vectors $\mathbf{x}(t)$ and $\mathbf{x}(t+\tau)$ will have almost the same elements. For example, embedding an ordered sequence of numbers from 1 to 10 (i.e., $x(t) = \{1, 2, 3, ..., 10\}$) for time delay $\tau = 1$ and $n = 3$, would produce a 4×3 matrix say $\mathbf{V}$:

$$\mathbf{V} = \begin{pmatrix} 3 & 2 & 1 \\ 4 & 3 & 2 \\ 5 & 4 & 3 \\ 6 & 5 & 4 \end{pmatrix} \tag{4.37}$$

The parameter n is called *embedding dimension*. In order to unfold the dynamics from $x(t)$ effectively, we need to choose very carefully the value of parameters τ and n. For that reason a number of procedures exist such as the *Average* or *Auto Mutual Information* criterion and the autocorrelation function for the time delay and the *False Nearest Neighbours* criterion for the embedding dimension.

4.3.3 Phase-space reconstruction: some practical issues

The embedding theorems ensure that a phase-space can be reconstructed to emulate the original dynamics of any dynamical system. The practical issue arising is how this is achieved given a single time series as an observable output from a dynamical system. Based on these theorems, procedures have been developed that enable us to determine the embedding parameters: the time delay τ and the embedding dimension. We present below the most popular which will be also utilised in the empirical work that follows. The most common way to determine τ is by

the calculation of the *Average Mutual Information*. Embedding dimensions can be determined via an iterative technique called the *False Nearest Neighbours*.

4.3.3.1 Average Mutual Information

The Average Mutual Information criterion (AMI) has been introduced by Fraser and Swinney (1986) and is the average amount of common information contained in a set of variables. More precisely, it provides a measurement of the nonlinear correlation between different sub-samples of a given series. If the series exhibit non-linearities the mutual information will be significantly greater than zero for small lag values. It can be regarded as a statistic analogous to the linear correlation coefficient, r, but sensitive to any kind of relationship and not just some linear dependence.

For a pair of discrete random variables X and Y, *mutual information* is defined as:

$$I(X;Y) = H(X) + H(Y) - H(X,Y) \tag{4.38}$$

where $H(X)$ and $H(Y)$ denote *entropies* and $H(X,Y)$ denotes *joint entropy*. From (4.38) we can easily see that for two discrete sequences[39] a_i and b_j, the amount of information we can gain for a_i by b_j measured in bits is:

$$\log_2 \left[\frac{P_{AB}(a_i, b_j)}{P_A(a_i)P_B(b_j)} \right] \tag{4.39}$$

where $P_A(a_i)$ and $P_B(b_j)$ is the normalised histogram of the distribution of the a_i and b_j sequences respectively. $P_{AB}(a_i, b_j)$ is the joint distribution for both sequences. This is another way of expressing mutual information. If a_i and b_j are independent, then

$$P_{AB}(a_i, b_j) = P_A(a_i)P_B(b_j) \tag{4.40}$$

and mutual information is zero. If we average the mutual information over all sequences, we derive the AMI given by:

---

[39]These can be different histories of the same time series.

$$I_{AB} = \sum_{a_i, b_j} P_{AB}(a_i, b_j) \log_2 \left[\frac{P_{AB}(a_i, b_j)}{P_A(a_i) P_B(b_j)} \right] \tag{4.41}$$

For a signal $s(n)$, we choose a_i to be the measurement of the sequence at time n and b_j to be the measurement of the sequence at time $n + T$ thus the AMI is defined as a function of the time lag T and is usually denoted $I(T)$. For $T = 0$ the AMI is the entropy. In order to sufficiently define the time delay τ for embedding, we choose it to be equal to the lag where the AMI reaches it's first minimum value.

4.3.3.2 False Nearest Neighbours

A set of points being close to any chosen point in embedded (i.e., reconstructed) space, are called *nearest neighbours*. The issue is discerning between *false* and *true* neighbouring points. It follows that some points will be close to others in the true dynamics, and some other points will appear to be close in embedded space but in reality, they are very far apart in the true phase space. They only appear to be close due to the reconstruction.

The FNN procedure works as follows: For a given delay[40] τ, we choose arbitrarily a central point and start increasing the number of embedding dimensions. We observe the neighbouring points. If their distance to the central point remains constant, these are termed true neighbours as opposed to false neighbours whose distance increases. We choose as embedded dimension the one that decreases the percentage of false nearest neighbours to zero. More analytically, from the delay vector $\mathbf{x}(t)$ in (4.35), reconstructed for some dimension d, we are interested in locating the nearest neighbour for each point in the reconstructed space. We denote the neighbour of $\mathbf{x}(t)$ as $\mathbf{x}_{NN}(t)$. We add one dimension to d and reexamine the distance between $\mathbf{x}(t)$ and $\mathbf{x}_{NN}(t)$ in $d + 1$ space. From the transition from d to $d + 1$:

$$\begin{aligned} \mathbf{x}(t) &\longrightarrow [\mathbf{x}(t), x(t + d\tau)] \\ \mathbf{x}_{NN}(t) &\longrightarrow [\mathbf{x}_{NN}(t), x_{NN}(t + d\tau)] \end{aligned} \tag{4.42}$$

---

[40]This implies that τ has to be a priori estimated and the accuracy of the FNN technique relys heavily apon this parameter.

If distance $|x(t + d\tau) - x_{NN}(t + d\tau)|$ is small, then $\mathbf{x}(t)$ and $\mathbf{x}_{NN}(t)$ are also near neighbours in $d+1$ dimension. The dimension in which the unfolding of the attractor is complete, is identified when the percentage of false nearest neighbours drops to zero. Usually we accept a threshold value of 1%. This dimension is called the *global integer embedding dimension* and is denoted d_e.

4.3.3.3 Other approaches

The AMI and FNN criteria are not the only ones available for the reconstruction of the phase space dynamics. The autocorrelation function has been used in earlier work in order to determine the time delay. As we will also demonstrate in the next chapter, it does not always provide clear results as the autocorrelation function is based in the first 2 moments of the data and is essentially a linear function. The underlying structure in the time series may be a result of nonlinearity and determinism. The data may also be contaminated with unknown noise (Sauer, 1992 and So *et al.*, 1996). This hinders the accurate determination of τ and the reconstruction may thus be "blurred". In this sense we may not be able to determine accurately the embedding dimension. Secondly, false neighbours may appear to be true due to the inaccurate unfolding of the dynamics from the observable sequence.

As with AMI, the FNN criterion is not the only way to determine the value of the dimension of embedding. Kennel and Isabelle (1992) have proposed a short-term nonlinear predictability criterion based on surrogate tests.[41] The optimum embedding dimension and time delay should produce the best predictability when comparing this with surrogate data simulations.

There are more techniques which we will not cover here. We found for our purposes that AMI and FNN produce good results with relatively economical computations and algorithmic complexity. Moreover, most of the existing phase-space reconstruction empirical literature is incorporating these, which makes our results comparable. In practice it is preferred to combine more than one techniques in calculating the embedding parameters in order to double-check their validity. For

---

[41]We dedicated a separate chapter on surrogate data analysis (SDA). At this point it suffices to say that SDA is a simulation technique similar to bootstrapping where one is interested in rejecting the null of linearity. More on this sector is covered in chapter 6.

instance, correlation dimension can be used to support the FNN results. Secondly, data allowing, one may choose to estimate a range of values and check whether the reconstructions based on this range provide similar answers.

4.4   Conclusions

In this chapter we presented the basic theory behind chaotic systems and outlined a set of tools used to produce metrics that would enable us to characterise any process as chaotic. The methodology is based on examining a single observable quantity which is a result of a system's functioning i.e., a single time series generated by the dynamics of the system. These metrics will remain stable for certain transformations of the phase-space dynamics, hence the name "chaotic invariants". This allows us to manipulate the dynamics obtained or "guessed" from the series in such a way so as to detect and possibly visualise their chaotic structure. In this approach, all the techniques discussed in the current chapter are purely univariate. These measurements though can be misleading and their calculation is prone to serious deviations. For this reason, a number of methodologies have been developed in order to validate results based on measurements for the detection of chaos and to provide a clearer view of the dynamics.

In the following chapter we examine "recurrence plots" based on a kind of nonlinear, high-dimensional autocorrelation matrix. This approach is rather visual and qualitative, but enables us to view the dynamics in even very high dimensions. It can also provide the base for a quantitative framework and can be used to cross-check measurements obtained by chaotic invariants calculations. As we shall see, it can provide a powerful tool for establishing whether time series are results of deterministic or stochastic data generating processes or even whether these two processes interchange within the dynamics.

In the chapter after the next, we examine how one can combine statistical measurements and a hypothesis testing framework to obtain an answer on whether the sequences examined are linear-stochastic or deterministic. This approach is called "Surrogate Data Analysis" and is based on a similar idea to bootstrapping and requires a number of simulations. These simulations are designed to produce

data with the same autocorrelation and spectral density structure. In this sense, any nonlinear structure and determinism in the original sequences can be clearly contrasted to the simulated "linear" counterparts. This enables us to refute linearity with a good amount of statistical certainty. Summarising, we can say here that there is a wide array of statistical tests and computations one can apply on data to determine (with various degrees of success) whether they have been produced by a deterministic or a stochastic system. In this thesis we focus on a specific approach which concentrates on a subset of these and we suggest possible strategies for analysing nonlinear and possibly chaotic financial data.

CHAPTER 5
Recurrence Analysis

5.1 Introduction

In most financial time series analysis applications there is no clear evidence of the nature of the noise "contaminating" the data. Moreover, the academic community is still arguing[42] about the nature of the *Data Generating Process* (DGP)(e.g., see Mandelbrot, 1999, 2001a-2001d, Marchesi and Lux, 1999, Malliaris and Stein, 1999). Any assumptions about the statistical properties of noise should be carefully validated by thoroughly examining the consistency of our modelling results. Furthermore, noise may cover the dynamics or the fine details of the underlying DGP. One technique that allows us to view these dynamics and assess on the presence and level of noise is presented in this thesis.

Our aim[43] is to show the usefulness of "*recurrence analysis*" as a univariate time series analysis tool. Recurrence analysis consists of "*recurrence plots*" and "*recurrence quantification analysis*". Here we focus on re-introducing the *recurrence Plots* (RP) for financial time series analysis. In their seminal paper Eckmann *et al.* (1987) have presented the RP as a new visual method for the qualitative assessment of time series. It is a tool for the detection of hidden patterns, similarities, temporal correlations, intermittence and structural changes in the data. Until now, it has been used mainly for the study of physiological time series and generally in phenomena which usually exhibit nonlinear determinism. Similar work in economics was undertaken by Gilmore (1990, 1992, 1996 and 2001) who introduced the *Close Returns Plot* (CRP, see also latest application by Belaire-Franch *et al.*, 2002). Both above RP and CRP techniques "suffer" from the fact that the interpretation of their graphical output is not as straightforward as in the "conventional" linear autocor-

---

[42]For interesting discussions and reviews on the topic one may refer to Chiarella (1992), Creedy and Martin (1994) or the more recent Opong *et al.*, (1999).

[43]Part of this chapter has been presented in Belgium at the Parallel Applications in Statistics and Economics, 7th International Workshop (PASE, April 2000) and appears in the proceedings, a special issue of the Neural Network World journal (vol 10, No 12, p 131-145).

relation (ACF) or partial autocorrelation function (PACF) plots that *frequentists* have been accustomed to. Moreover, an understanding of the notion of *embedding* of time-series plus a background in chaos and information theory is deemed necessary. With RP though we can visualise successfully the underlying dynamics in many different dimensions and enhance the information provided by the ACF and PACF plots. An other important point is that the time series examined may not be stationary. On top of these, RP are computationally very cheap. In the following sections of this chapter we provide a brief outline of the theoretical foundations and practical issues required to generate successfully an RP, so as to display any kind of structure within the examined time series (elements of these explained in more detail in the previous chapter). In section 5.2 and 5.3 we provide guidelines for the construction and interpretation of RPs by using illustrative examples with simulated time series. Following these sections, we generate RPs of real financial time series and derive the Recurrence Quantification Analysis results from these plots. Below we briefly discuss the Close Returns Plots, an idea similar to RPs, introduced by Gilmore (1990).

5.1.1 Close Returns plots

Close returns plots have been developed by Gilmore (1990) and revised by Midlin *et al.*, (1990), Midlin and Gilmore (1991) and Gilmore (1992). They are based on the same premises as the recurrence plots i.e., on the idea that different segments of time series may display very similar behaviour. The starting point for these diagrams is the observable time sequence x_t though *without an embedding* as in Eckmann *et al.*, (1987). If one of the observations of x_i occurs near an unstable periodic orbit, the subsequent observations will lie near this orbit for a period of time, before being repelled away from it (Gilmore, 1992). If these evolve near his orbit for a sufficiently long time, they will return to the x_i neighbourhood after some time T. This parameter T also indicates the length of the orbit. The conclusion is that we can retrieve topological information for the attractor when distances $|x_i - x_{i+T}|$ are small. If x_{i+1} is near x_{i+1+T}, then x_{i+2} will be near x_{i+2+T}, and so on and so forth. Close returns plots indicate a series of consecutive observations for

which $|x_i - x_{i+T}|$ is small. Effectively a close return plot can be approximated by a recurrence plot when the latter is constructed without embedding at all.

Gilmore (1992) has used close returns plots on exchange rate data. His improved scheme involve the plotting of distances $|x_t - x_{t+T}|$ as a function of t in the horizontal axis against T on the vertical axis, where t and T range from 1 to N, where N is the sample size. A threshold value ϵ is determined as a percentage of the maximum absolute value of differences between observations across the sample. This is usually between 1% and 5%. If $|x_t - x_{t+T}| < \epsilon$, point (t, T) is coded black (and in the opposite case white). Chaotic data produce plots with almost horizontal line segments whereas stochastic sequences will produce plots without any pattern. Gilmore (1992) checks via simulations the robustness of such plots at the presence of noise and concludes that his diagrams can provide clear and conclusive results even under such circumstances. He also uses extensively exchange rate time series and stock market data to put his plots to the test. He concludes that there is evidence of nonlinear structure but he cannot obtain a clear answer on whether this nonlinearity is due to chaotic determinism or some other reason. He reveals clear aperiodic structures which could lead us to conclude on some deterministic component for the series but more research is needed for this.

Gilmore's work was not conclusive about chaos but did not exclude it as well. The main difference is that close returns plots do not require time series embedding. In this chapter we will be discussing the use of recurrence plots on stock market data and will be revealing qualitative and quantitative information that could support the existence of chaotic attitude in their dynamics. More recent works influenced by Gilmore's research are Choo (1997) and McKenzie (2000).

5.2 RPs and phase-space reconstruction

The philosophy behind RP's is based on the concept that the dynamics of a multi-dimensional system can be recreated and predicted from a single history of anyone of its observable output variables. Using the *time-delayed* copies of this observable we can reconstruct, under certain circumstances[44], the phase-space of

---

[44]Infinite amount of noise-free data, as Takens (1981) describes.

the dynamical system.[45] Two factors are needed for this purpose: the value of the *embedding dimension* d_e of the time-series and the *time-delay* τ. The *method of delays* or *delay coordinate embedding* is regarded as the most popular phase-space reconstruction technique (Abarbanel, 1995). When the embedding is correct, the relevant theorems described in Takens (1981) and Sauer *et al.*, (1991) guarantee (see also the preprint by Sauer and Yorke, 1992) that certain properties of the original system i.e., its *dynamic invariants* are preserved in the reconstruction, enabling us to identify its dynamics. Recently Iwanski and Bradley (1998) suggested that higher dimensional space embedding is not needed and that for $d_e=1$ the RP capture the same information as do those that are generated using higher embedding dimensions.

The correct selection of the *time delay* parameter τ is crucial for the proper phase space reconstruction of the underlying dynamics. The careful selection of the embedding dimension parameter allows for the system's dynamics to be "unfolded".[46]

The criterion used for the determination of the delay parameter τ is the *average mutual information* (AMI). AMI has been introduced by Fraser and Swinney (1986) and is the average amount of common information that allows for a time separation of the dynamics. In order to sufficiently define the time delay τ we choose it to be equal to the lag where the AMI reaches it's first minimum value. To demonstrate this, we simulated an ARCH(2) process, series x, of size 5000, from normal random data. The series x is depicted in figure 5.1 (a). The AMI plot in figure 5.1 (b) clearly indicates that we should set the time delay parameter equal to 3 where the AMI reaches its first minimum.

In this empirical chapter, in order to determine the embedding dimension d_e for the vectors $y(n)$, we employed an improvement of the *global false nearest neighbours* (FNN) statistic (Kennel *et al.*, 1992) as it is implemented in the TISEAN package (Kantz and Schreiber, 1999).[47] For clean chaotic signals, the percentage of false

---

[45]See section 4.3.3, page 71 onwards.

[46]There is a large bibliography covering this topic. One may refer to Abarbanel (1995) or Kantz and Schreiber (1997) for a discussion.

[47]In this research we employed a variant of the FNN algorithm given in the TISEAN (Kantz and Schreiber, 1999) library. It determines false neighbours when the ratio of the distance of the iteration and that of the nearest neighbour exceeds a given threshold the point. In their algorithm, the authors (Hegger and Kantz, 1999) have implemented a new second stricter criterion for the

(a) ARCH(2) process.

(b) AMI
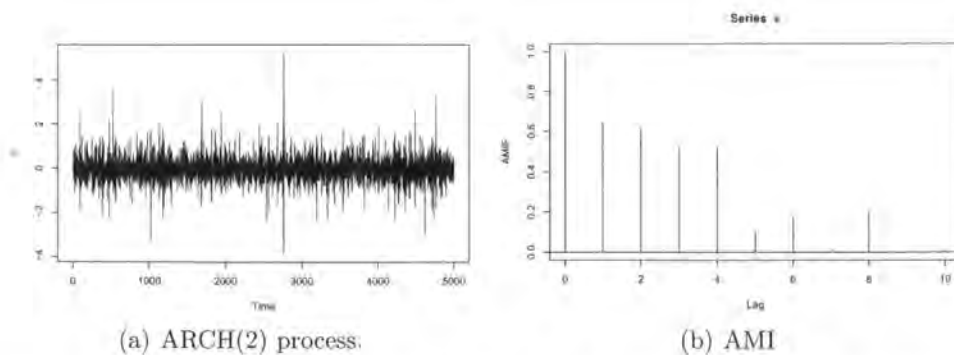
Figure 5.1: An ARCH(2) process and the Average Mutual Information plot.
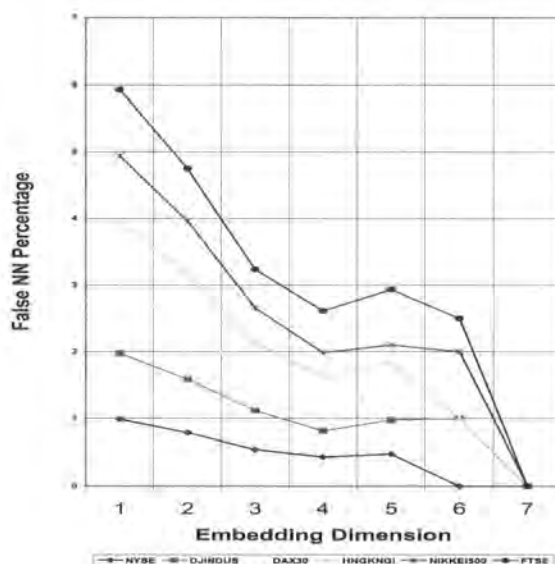
Figure 5.2: False Nearest Neighbour Percentages.

nearest neighbours will converge to zero when dimension d_E is reached and will stay to that level for successive calculations. This is a good indication for the existence of noise and its level in our data. In figure 5.2 we have plotted the graphs of the FNN percentages for embedding dimensions 0 to 7, for the six stock market index returns in table 1.[48]

---

closeness of points. If the distance to the nearest neighbour becomes smaller than the standard deviation of the data divided by the threshold, that point is omitted. For a critique on FNN one may refer to (Abarbanel, 1995). For a very insightful critique of delay coordinate embedding one should refer to Kugiumtzis (1996).

[48]We have not filtered any of the 6 series of returns so we expect a level of noise to be present.

5.2.1 RPs and correlation dimension

Both close returns and recurrence plots provide us with significant advantage over the calculation of certain invariant statistics such as the correlation dimension. This becomes more obvious when examining real life financial time series data which are always not noise-free and limited in length. As we have already discussed, the methods concentrated on the calculation of certain statistical quantities that would reveal chaos, require theoretically an infinite amount of noise-free observations. Such samples were not available in finance until recently with the generation of high-frequency tick-by-tick databases. With historical data, one is limited to samples of a few thousands observations when dealing even with daily data. That is why graphical approaches such as the one proposed in this chapter are deemed as a more appropriate tool for the detection of determinism. Brock (1986) among others, has drawn our attention to the fact that correlation dimension can not be calculated accurately on noisy data (see also Castro and Saure, 1997). Secondly, the correlation dimension estimation procedure will not provide us with any information about the underlying DGP. Finally, all these chaotic invariant statistics should not be readily accepted unless they were put to the test within a *Surrogate Data Analysis* framework which we will be examining in the next chapter.

Recurrence plots do not seem to suffer from what was just discussed. Firstly we demonstrate here that when noise contaminates the original DGP, these plots will be able to recover the "larger picture" of the dynamics effectively. Secondly, their results are not disturbed by small size samples. Rather, we should be careful when analysing large data-sets to try to reveal the finer patterns by magnifying parts of the plots or by careful examination of the recurrence matrix itself. Finally, recurrence plots are able to reveal information which could support or contradict the hypothesis of a chaotic DGP or the presence of it for some parts of the history of the series. In the following section we provide a demonstration of these plots using both simulated and real data.

5.3 Recurrence Plots

There are two types of RP, the thresholded and the unthresholded. The thresholded RP (also known as *recurrence matrix*) is generated by the comparison of all embedded vectors y_k with each other and by drawing a point when the distance between the vectors is below some threshold. A point is drawn at the coordinates (i,j) when the i'th and the j'th embedded vectors are less than some arbitrary distance r apart:

$$\|y_i - y_j\| < r \tag{5.1}$$

with i corresponding to the horizontal axis and j to the vertical one. The unthresholded RP is a coloured version of the thresholded with the points being coloured according to the $i - j$ vectors distance. The usual colour coding requires dark colours for long distances and light colours for short ones. In this chapter we concentrate on the analysis of thresholded RPs. The RPs are symmetric around the main diagonal (45° axis), as the distance between the i'th embedded vector to the j'th embedded vector is the same as that of the j'th to the i'th.

As Eckmann *et al.*, (1987) point out, the i, j are in fact times. This effectively implies that through an RP we visualise a *natural time correlation* structure of the series. With RPs we can easily locate a wealth of recurring patterns. Moreover we can visualise and detect trends, abrupt changes or drifting dynamics. Recurring patterns appear in the RP as diagonal line segments, parallel to the main diagonal. If the time series is a random process, any patterns should be absent. The length of these lines is related to the inverse of the largest positive *Lyapunov exponent*. Due to the construction of the RP, the correlation integral $C(r)$ is given by the number of darkened points in the RP divided by the total number of points. Another salient feature of the RP is that the distribution of the line segment's length, is related to the entropy of the time series. This distribution is exponential and the exponent is equal to the entropy of the series.

When visually inspecting an RP one must examine the way the distribution of the dots varies for different distances from the main diagonal. For stationary series we should encounter a distribution of a number of dots around the diagonal that is independent from their distance form it. In other words, changes in the density

(a) Thresholded (b) Unthresholded
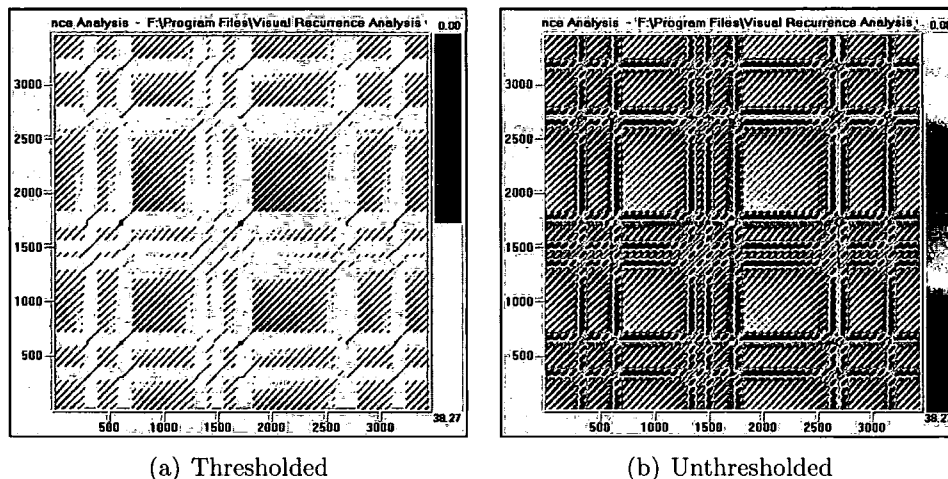
Figure 5.3: Examples of recurrence plots for the Lorenz system.

of the points around and close to the diagonal is an indication of non-stationarity. This is clearly demonstrated in the sine wave in figure 5.5 (a), where we can see due to the periodic nature of the series, a specific periodic recurrent pattern of dark (vectors that are closer) and light areas (vectors that are further apart) repeated throughout the surface of the RP.

In figure 5.3 we have generated the thresholded and unthresholded RPs of the variable x from the Lorenz chaotic system of equations.[49] It is fairly straightforward that both graphs support the chaotic nature of this deterministic system.

In Figure 5.5 (b) we produce an RP of a white-noise process. It is evident that there is no periodic component or structure. There is also a extremely large number of "stray" points uniformly distributed in the graph in the sense that none of these belong to a straight line segment parallel to the main diagonal. This is indicative of the "stochastic" nature of the signal.

One may be able though to discover the nature the high dimensional pseudo-random number generator that has been used to generate these series and capture

---

[49]The three Lorenz differential equations define system that exhibits chaotic dynamics:

$$
\begin{aligned}
\dot{x} &= -\sigma x + \sigma y \\
\dot{y} &= -xz + rx - y \\
\dot{z} &= xy - bz
\end{aligned}
\tag{5.2}
$$

where $x, y, z \in \mathbb{R}$ and $\sigma, r, b > 0$. We chose arbitrarily the σ, r and b to be positive parameters.

the periodicity in the algorithm provided that the series are of sufficient length. For the same signal in Figure 5.5 (b) we generated another RP in (d) using a higher embedding dimension d_e in order to capture the high dimensionality of the unknown pseudo-random number generator. As seen in figure 5.5 (d) it is evident from the cross like formation in the middle of the RP that there is some underlying pattern in the random number generating procedure. Also line segments parallel to the main diagonal can be discerned in various parts of the RP. Figure 5.5 (c) is an RP of a Brownian motion without any delay coordinate embedding carried out.

For an embedding with embedding dimension d_E=35 and time delay τ=19, the RP becomes the straight diagonal (the 45° axis) line. This implies that there is no *deterministic* temporal relation between the points x_t and x_{t-i} for any i[50] and hence no recurrence can be viewed in the RP.

For the ARCH(2) process described in section 3, for τ=3 and $d_e = 1$ we have created the RP in figure 5.5 (e) where some kind of structure can be discerned. This RP is different than the one in figure 5.5 (b). Although the ARCH(2) process is based on a computer generated pseudo-random process, we can see that there is a collection of horizontal and vertical lines. These lines correspond to the volatility clusters in various parts of the ARCH(2) signal.

To illustrate the importance of the correct selection of the time delay parameter τ, we generated a sine wave signal of 5000 observations length and contaminated it with a fractional noise (an ARFIMA(0,.4,0) signal). We use the AMI criterion and calculate the FNN percentage and for time delay τ=1 and embedding dimension d_e=7 we produce the RP in figure 5.6 (a). The periodical attitude of the signal is evident and very close to figure 5.5 (a). The careful reader will be able to discern the "mosaic" like pattern of white and dark areas which are more fuzzy now due to the presence of fractional noise. This noise was not able though to conceal the deterministic nature of the underlying sine wave whose main characteristics we managed to retrieve by delay coordinate embedding.

Had we chosen to use the ACF to determine the time delay, we would have calculated a τ=4 and a d_e=6. These parameters lead to the RP in figure 5.6 (b)

---

[50]As the increments of a Brownian motion are i.i.d. Gaussian noise.
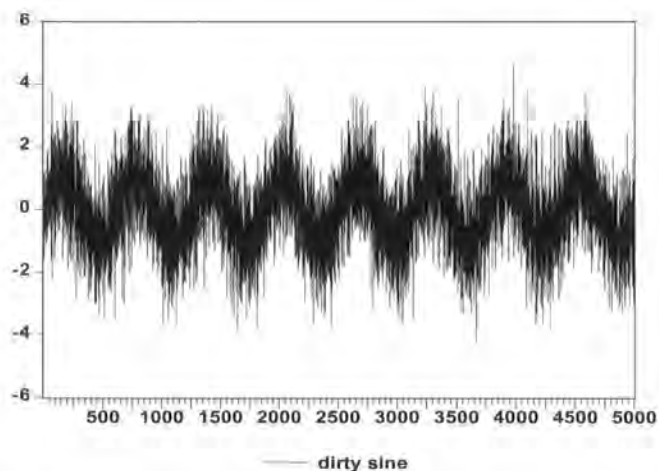
Figure 5.4: Sine with ARFIMA(0,0.4,0) noise.

which is of a much poorer quality.

In order to overcome the obvious difficulties of interpreting the RP and providing a framework for quantitative results, Webber and Zbilut (1994, see also Zbilut and Webber, 1992) have suggested a parametrisation according to:

1. the percentage of recurrence (darkened points in the recurrence matrix);

2. the percentage of determinism (darkened points included in diagonal segments);

3. the Shannon entropy of the distribution of these lines and

4. the inverse of the longest diagonal line.

They have termed this *Recurrence Quantification Analysis* (RQA) and produce various applications on physiological time series (see Zbilut *et al.*, 1998b, 2000 and 2002). Zbilut *et al.*, (1998a) have also introduced the *Cross Recurrence Quantification* (CRQ) analysis.[51]

---

[51]They propose this as a filter able to recognise and extract signals contaminated with large amounts of noise. They demonstrate that CRQ can extract information from a sequence with very low signal-to-noise ratio (large noise variances) and detect successfully hidden periodicities.
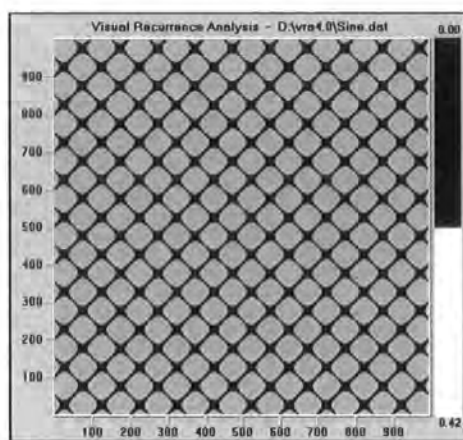
(a) Sine wave
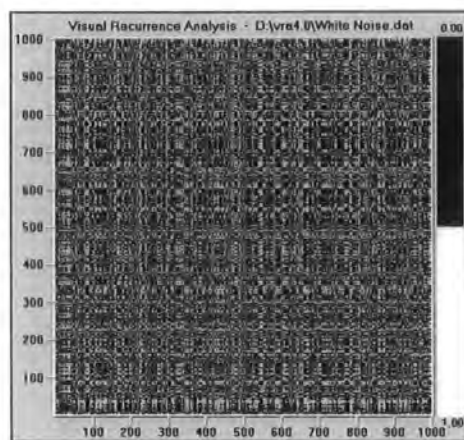
(b) White noise
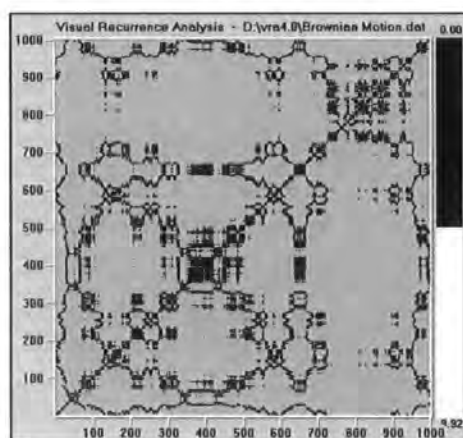
(c) Brownian motion
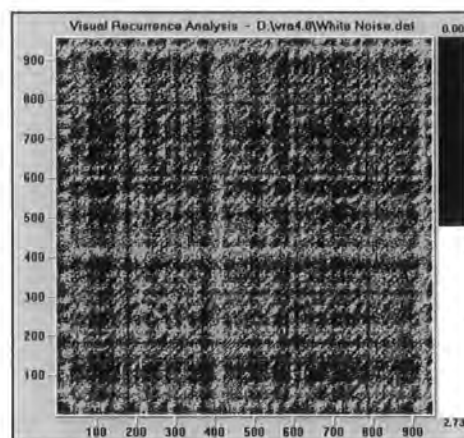
(d) White noise with high embedded dimension
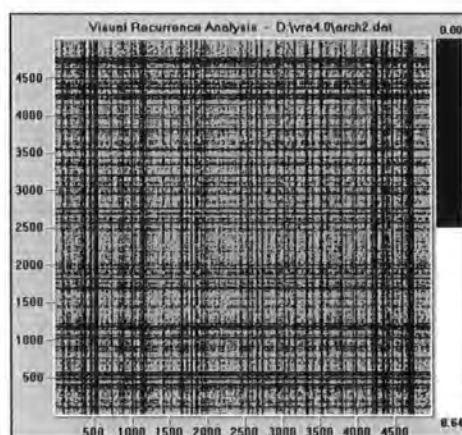
(e) ARCH(2)

Figure 5.5: Thresholded recurrence plots of various simulations.

(a) Embedding via AMI

(b) Embedding via ACF
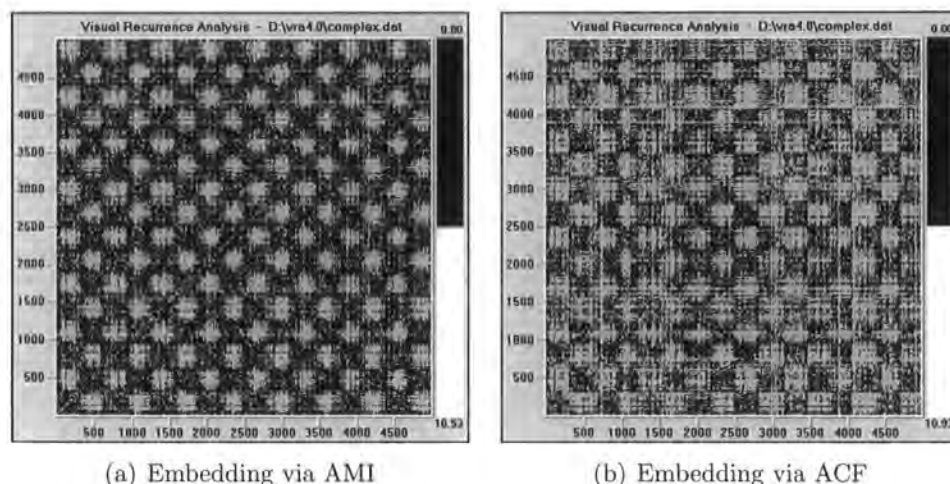
Figure 5.6: Examples of the power of the AMI criterion over the ACF.

5.4 Recurrence plots and Market Efficiency

An efficient market is one for which no pattern can be identified in the activity of the investors. In an efficient market, with a frictionless flow of information, a "buy-and-hold" is the best trading strategy. In other words, there are *no pure arbitrage opportunities*. The absence of predictability in the asset or stock index returns series guarantees that the *Efficient Market Hypothesis* (EMH) holds. Any deterministic pattern in the histories of returns, could then be evidence against the EMH. The most popular approaches have been mainly the tests for the decay of predictability in the histories of returns and the BDS test (Chiarella, 1992, Opong *et al.*, 1999, Creedy and Martin, 1994, Sheinkman and LeBaron, 1989) as signs of chaos.

Chaos describes a notion of disorder, irregularity and unpredictability. In the physical sciences context, "chaos" means that a physical phenomenon may appear to be behaving irregularly but upon closer examination one may find it to possess considerable regularity. Low order chaos i.e., "low dimensional order of nonlinear determinism" is a form of non-stochasticity. In the EMH context, any test that concentrates on whether the returns exhibit low dimensional chaotic patterns or not, could therefore be regarded as a *weak form efficiency* or *returns predictability* test (Fama, 1991). Chaos implies determinism of a nonlinear structure and allows for short term forecasting under certain conditions. In our case, the real problem lies in the detection of low order determinism. In financial time series analysis

one deals with unknown DGPs, usually contaminated with stochastic noise which could easily be mistaken for high dimensional chaos as certain "linear tools" as the autocorrelation or partial autocorrelation functions or even spectral analysis, fail to detect the nonlinear structure within this high frequency information (Granger and Lin, 1994). Another big issue is the frequency the observations used. Determinism could be more apparent in weekly or monthly aggregates than in high-frequency intra-daily data and vice-versa. Secondly, caution is needed when dealing with the nonstationarity of the observed series as de-trending, differencing or using price deflated information could introduce, alter or even destroy any deterministic pattern in the data (see Abarbanel, 1995 and Kantz and Schreiber, 1997). In this chapter we are using returns which we constructed as the first differences of the logarithms of the actual closing prices series. By taking first differences we have introduced a high-pass filter (Abarbanel, 1995) in the series and this can hinder correct determination of the underlying dynamics.

It would be astonishing and highly debatable to assume that financial time series are governed mainly or entirely by chaotic mechanisms. Until today there has been a relatively small number of cases where chaos has been detected in economic time series. This could be due to problematic and limited data and the use of techniques that are unable to provide robust evidence in favour of chaos. Given the limitations and problems of the techniques and the available data, one would be safer to examine whether financial time series are governed by *irregular stochasticity* with some degree of determinism. This premise implies that the DGP is consisting of stochastic and deterministic components which come in the form of coloured or white noise, irregular cycles, trends, volatility clustering, long memory etc. etc.

A certain classification of chaotic or non-chaotic systems can be made using the embedding theory of nonlinear dynamical systems as defined by Takens (1981) and this is what we follow in this chapter. What we propose is the use of RPs as a qualitative tool to compare the structure of the dynamics of the generating system of financial time series with that of purely random time series. Any deviation from the form of a purely stochastic signal's RP i.e., structure of any kind around the diagonal instead of uniformity, should be treated as evidence against the EMH. One may even

Price INDEX	Description	τ	d_e
DJINDUS	DOW JONES INDUSTRIAL	1	6
FTSE	FTSE ALL SHARE	3	6
NYSE	NYSE COMPOSITE	2	5
NIKKEI500	NIKKEI 500	3	6
HNGKNGI	HANG SENG	3	5
DAX30	DAX 30 PERFORMANCE	2	7

Table 5.1: Time Delays and Embedding Dimensions for the 6 time series

be able to detect periods where the series are "more" stochastic and periods where they exhibit a strong recurrent structure. Secondly, by comparing the RPs of the same series in two or more different histories, we could detect any change of the dynamics for that series. Thirdly, by comparing the RPs for different frequencies of the same series we could deduce whether there is a certain underlying dynamical cycle that is not apparent in high-frequency but is evident in lower frequencies. By using cross-recurrence analysis we could also deduce whether the series include a cycle of some periodicity and try to extract this information.

We collected from Datastream, daily data for 6 major stock market indices covering the period from January 1991 to October 1999. The original series are depicted in figure 5.8. We calculated the relevant time delays τ and embedding dimensions d_e with the AMI criterion and the FNN method respectively. The results are tabulated in table 5.1.

It is obvious from our discussion so far that these series clearly do not indicate a random walk (see figure 5.7). The formations in the RPs resemble more that of the ARCH(2) process in figure 5.5 (e). We can notice the relatively blank areas and the dark horizontal and vertical lines. These correspond to periods of crises (such as the Asian), shocks and in general of volatility clustering in the original series (mainly during the period after 1996-1997 for the European and the US indices). These events are so strong that dominate the whole phase space. The lighter areas are related to periods of relative calmness. The RPs (in subfigures a-f in 5.7) for the European and US indices are very close and exhibit horizontal and vertical thick line formations in the same approximately time point. The Japanese and Hong-Kong
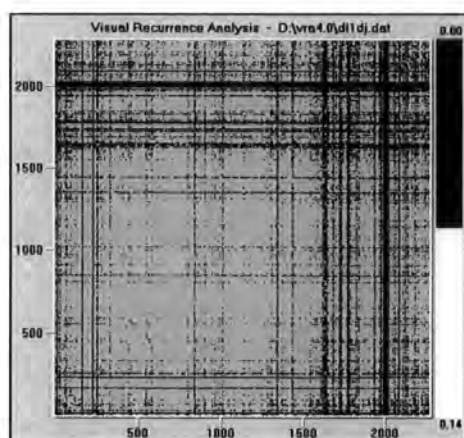
(a) Dow Jones Industrial
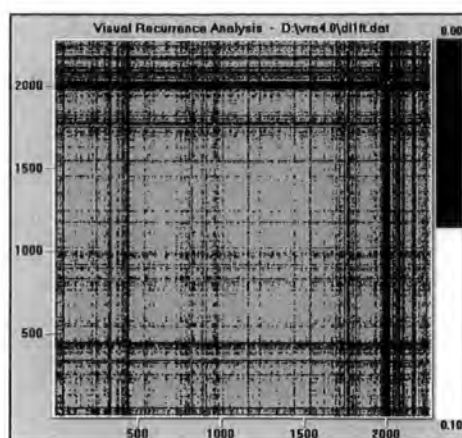
(b) FTSE ALL SHARE
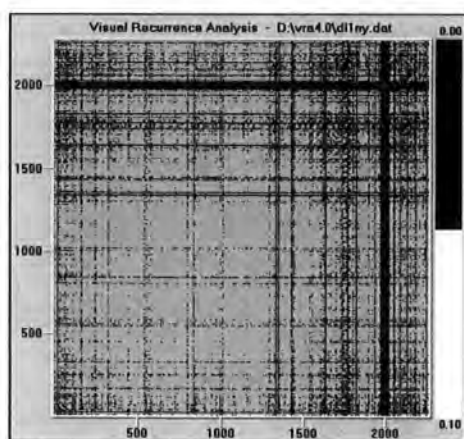
(c) NYSE
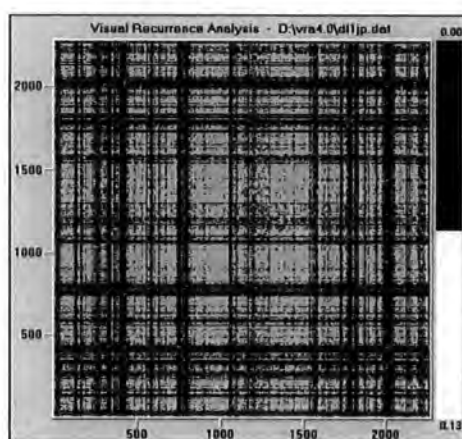
(d) NIKKEI 500
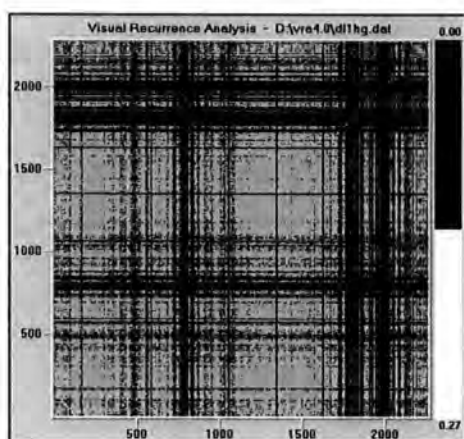
(e) HANG-SENG
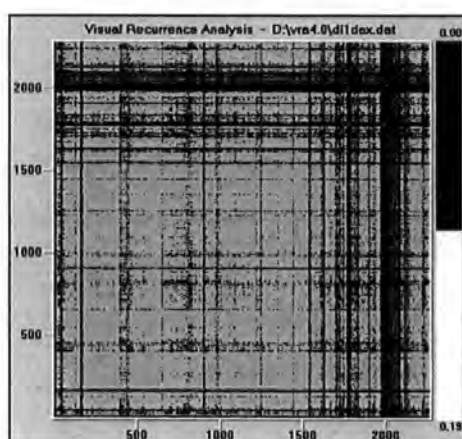
(f) DAX 30

Figure 5.7: Thresholded recurrence plots of the indices in table 5.1.
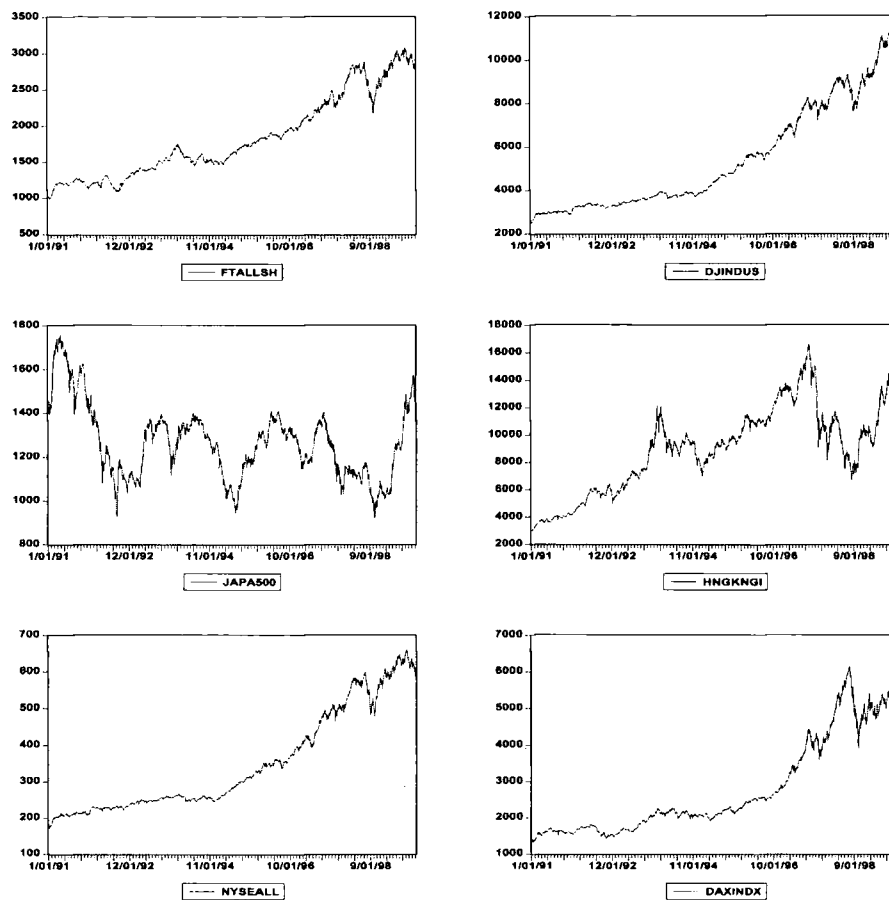
Figure 5.8: The stock market indices time series in Table 5.1.

indices exhibit some differences in their RPs as the original series are not following closely the movement and trends of those of the European and the US ones. We can still though discern the same kind of formations where intense volatility makes its presence. Another feature is that there seems to be for some RPs (Hang-Seng and NIKKEI 500 especially) a fade repetition of some recurrence structure throughout the plot. The careful eye will discern recurrent patterns that repeat themselves in different times and scales. These come in the form of a collection of horizontal and vertical lines or entire sub-areas (boxes) of the RPs. Of course this could be due to the volatility or variance inhomogeneity of these series, incorrect embedding or the general nature of the constructions of RPs. It remains to be seen if these are indications of some kind of self-similar processes.

One immediate conclusion is that these series are far from a random walk.

(a) Daily data   (b) Hourly data
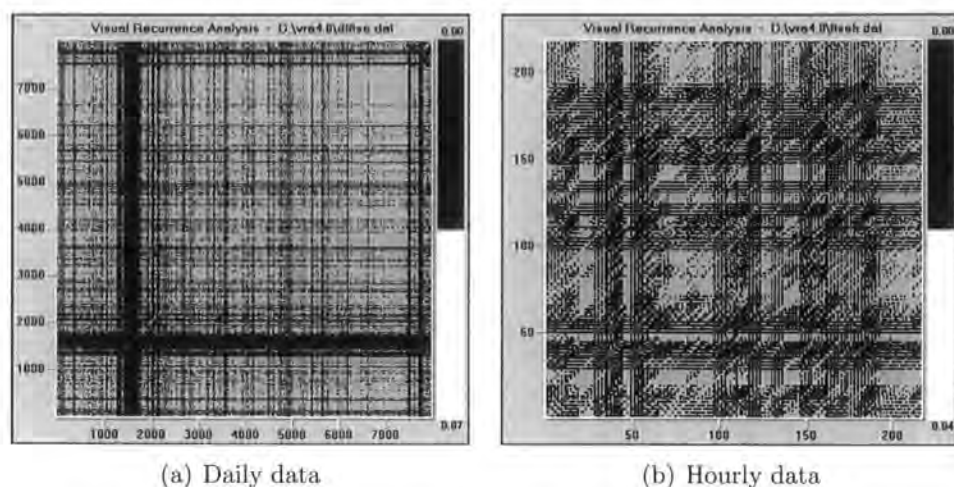
Figure 5.9: Daily and hourly data frequency recurrence plots.

We cannot though detect clearly any formation of line segments, parallel to the RP's diagonal which would be indicative of nonlinear determinism. This may be mainly due to the presence of noise in the returns, which covers the deterministic nature of a possible chaotic attractor in the series by leading us to false delay coordinate embedding. Secondly, as mentioned earlier, these returns are a product of differencing and this could intensify the presence of noise and alter the underlying deterministic patterns in the original series (Abarbanel, 1995). It remains to be seen if one can denoise the returns without "tinkering" with the determinism or stochasticity of the series. For this reason we experiment with wavelet based filtering and denoising in chapter 8. A very interesting discussion on the subject is by Theiler and Eubanks (1993).[52]

In figure 5.9 (a) we have generated the RP for a very long FTSE ALL SHARE history (daily data since 1969). If we compare this to the same series RP in figure 5.7 (b) for the last decade, we can clearly see that the recurrences in the last decade are present in both graphs (strong vertical lines on the right hand side of the RPs appearing thinner in 5.9 (a) due to the scale of the plot). The figure 5.7 (b) is the upper right corner of 5.9 (a). In the latter we can also see the turbulent period of the oil crisis as a thick black line between the 1000th and 2000th observation. An

---

[52]Originally a SFI Working Paper under the No. 93-05-026. This is available online from: http://www.santafe.edu/sfi/publications/wpabstract/199305026

interesting finding comes from the hourly FTSE ALL SHARE data in figure 5.9 (b). In this figure we have a clear indication of determinism as there are parts all over the RP that contain line segments parallel to the main diagonal.

5.4.1 "Zooming in"

Most of the RPs we discuss here exhibit significant patterns in the sense that they differ substantially from the ones that were derived from pseudo-random data (the closest thing to purely stochastic processes). We indicated that the ARCH process sequence and the original financial time series generate similar artifacts in their corresponding RPs. The volatility clustering effect seems to develop nonlinearities which are detected from the RPs and spread throughout the phase-space dynamics as thick horizontal and vertical segments.

In the very limited previous research that has been focused around RPs, there has been no significant discovery with the exception of the work by Gilmore which focused on returns plots. The main problem being here the sheer size of a recurrence matrix. For low embeddings, a RP of a million points will be approximately a matrix of a million by million pixels or points. Graphical representation of such plots with computers capable of depicting only 1024 by 768 points at a time,[53] requires a considerable amount of "normalisation". Only in that way the actual RP will fit on a screen or even an A4 page. Even under these circumstances though, patterns can be revealed. For the sake of accuracy, we produce here magnifications of segments of RPs on the FTSE ALL SHARE index closing prices and the corresponding logarithmic returns. We use the same data set as in chapter 8 which spans 3 decades (1970-2001) in daily frequency, totaling 8192 observations. The AMI criterion returned a delay of 19 for the levels and 3 for the returns series. The FNN criterion returned an embedding dimension of 9 for the levels and 7 for the returns.

It is fairly obvious that the RP in figure 5.10 exhibit patterns consistent with absence or stochastic randomness. This may also be more apparent from the examination of the unthresholded plots. It is not though possible to discern the finer details of the recurrence matrices from most of these RPs. Examination could be-

---

[53] One of the "large" graphic cards VGA definitions for IBM PC compatible computers.

(a) Closing Prices, d_e=9 and τ=19.

(b) Returns, d_e=7 and τ=3.

(c) Closing Prices, d_e=9 and τ=19.

(d) Returns, d_e=7 and τ=3.
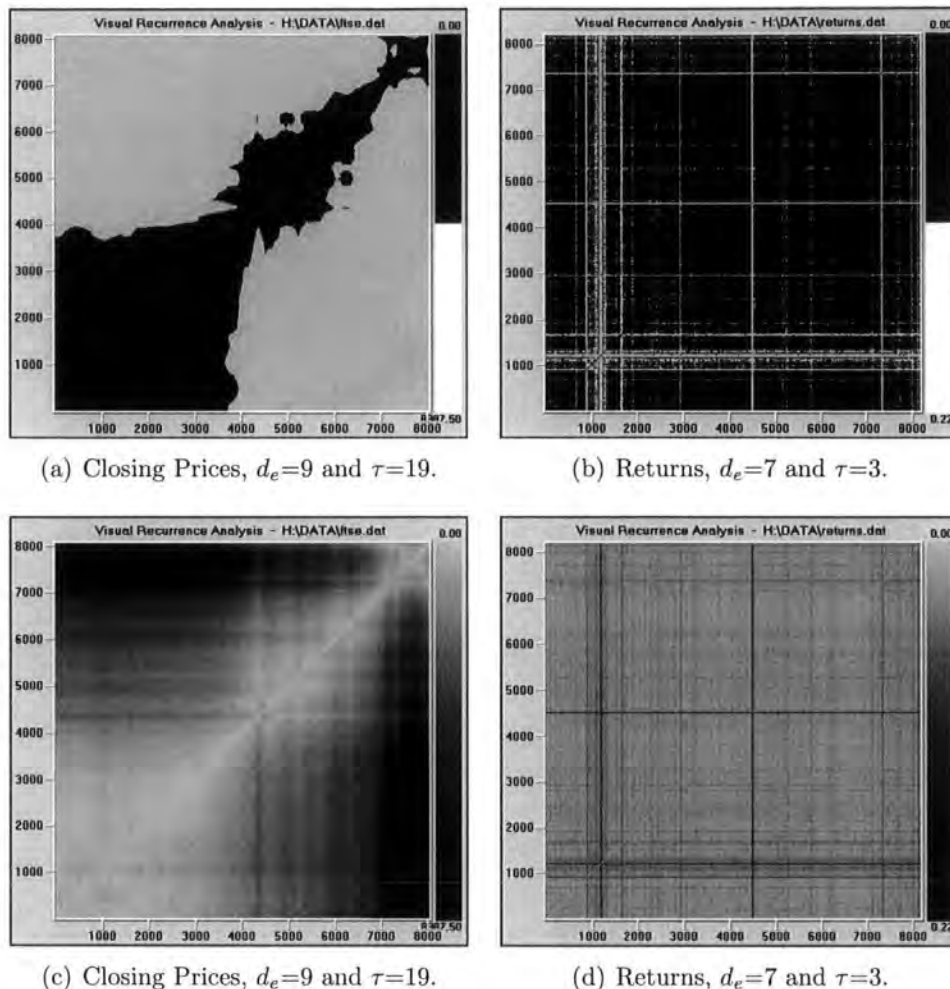
Figure 5.10: Thresholded and unthresholded recurrence plots of closing prices and returns of the **FTSE ALL SHARE, 1970-2001**, daily frequency.

nefit greatly from the magnification of sub-regions of the RPs. That would allow the identification of segments parallel to the main diagonal, indicating the presence or nonlinear deterministic recurrences.

A closer inspection of the RPs in figure 5.11 reveals their interesting characteristics. Firstly, we can easily discern in (a), (b) and (c) that the plots are replete with small parallel segments to the main diagonal. This implies the presence of deterministic recurrences. The plot in figure 5.11 (c) refers to the period before and after the oil crises of the 70s. We can clearly see the recurrent patterns and compa-

(a) Magnification 1

(b) Magnification 2

(c) The 73-74 oil crisis

(d) The market crash of 1987
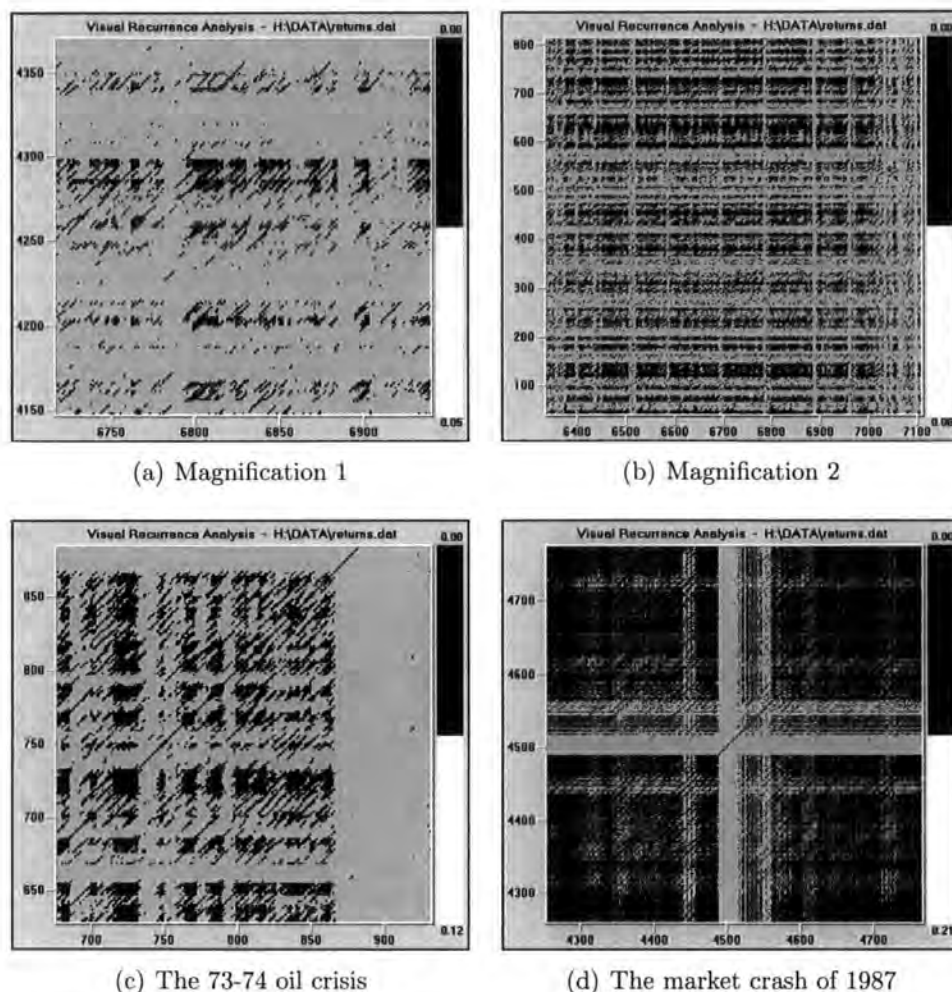
Figure 5.11: Discovering details of the RPs in figure 5.10 (b).

ring them to figure 5.3 we can also detect some similarities. Another interesting structure is the break of the recurrences during the crisis period at the right edge of RP in (c). This is also very evident in plot (d) where we have focused on the 1987 crash. The dynamics break there as well. The extreme volatility shocks during these periods create observations that are much more far apart than the rest points in phase space. The time-delay embedding can not under these circumstances help us to detect any recurrences for these periods. What is interesting though is that these "gaps" in the dynamics coincide mostly with localised shocks in the history of the time series. The more longer lasting, violent or turbulent the shocks the wider these "gaps" are.

One can support more strongly the presence of chaos after computing the range of invariant measures needed to characterise a process as chaotic. Moreover a "Surrogate Data Analysis" (see following chapter) would provide further evidence of the presence of deterministic structures. In Chapter 8, we filter the FTSE returns series using wavelets. This process may leave untouched the structure of the attractor enabling us to view its dynamics more clearly as we will demonstrate.

5.4.2 Recurrence Quantification Analysis

Following Webber and Zbilut (1994) we conduct Recurrence Quantification Analysis (RQA). Earlier, while calculating the AMI, the Shannon entropy for the FTSE data was found to be 1.151465 for the returns sequences and 2.360013 for the levels. The "spatio-temporal" entropy[54] was found to be 0 for the entire recurrence plot in figure 5.10 (b). For the magnified portions of this plot in figures 5.11 (c) and (d), the spatio-temporal entropy was found to be 72% and 80% respectively (for a sample of 1000 observations around the time of the financial shock). The results here indicate that indeed there is some form of periodicity and non-randomness in the data. The RQA results are presented in table 5.2. The most interesting statistic here is the "Maximum Line" length which corresponds to the largest segment of continuous points parallel to the main diagonal of the RP. Eckmann *et al.*, (1987) show that this is inversely proportional to the largest positive Lyapunov exponent. Cao and Cai (2000) identified this recently as a lower bound for the largest positive Lyapunov exponent. For the FTSE returns this was calculated to be 0.0411 which is positive and indicative of low dimensional deterministic dynamics. The RQA measurement gave a lower bound of 0.004 approximately which is consistent with a

---

[54]"Spatio-Temporal Entropy (STE) measures the image *structureness* in both space and time domains. Essentially, it compares the global distribution of colours over the entire recurrence plot with the distribution of colours over each diagonal line of the recurrence plot. The higher the combined differences between the global distribution and the distributions over the individual diagonal lines, the more structured the image is. In physical terms, this quantity compares the distribution of distances between all pairs of vectors in the reconstructed state space with that of distances between different orbits evolving in time. The result is normalised and presented as a percentage of "maximum" entropy (randomness). That is, 100% entropy means the absence of any structure whatsoever (uniform distribution of colours, pure randomness), while 0% entropy implies "perfect" structure (distinct colour patterns, perfect "structureness" and predictability)." Kononov (1999).

RQA Statistic	Value
Mean	0.000
Standard Deviation	0.010
Mean Distance	89.149
% Recurrence	34.339
% Determinism	27.678
Ratio	0.806
Entropy	5.009
Maximum Line	263
Trend	-1.683

Table 5.2: FTSE returns RQA results from RP in 5.10 (b).

lower bound measurement. Though, in our case, a large maximum line length may indicate less instability than a smaller one. This could imply the presence of a limit point attractor.

In order to follow the fluctuations of the maximum line length across time, a "windowed" version of RQA (Zbilut *et al.*, 2001 and 2002) is possible. This "windowed" RQA was applied on the returns sequence which was 8192 observations length. For $d_e=7$ and $\tau=3$, we generated 779 rolling subsamples of 260 observations each.[55] The choice of window refers to the number of observations per year for daily data and the shift corresponds to a week's observations. In figure 5.12 we summarise the results of the RQA on the FTSE returns. Diagram (a) refers to the mean of each selected subset of input points. Diagram (b) is the standard deviation of each selected subset. Diagram (c) refers to the mean of the rescaled distances in the upper triangular area of the RP. Diagram (d) refers to the actual recurrence as a percentage from the number of recurrences in the upper triangular area of the RP. Diagram (e) refers to the percentage of the determinism as the number of recurrent points forming parallel line segments to the RP's main diagonal by the number of recurrences as in (d). If there are no recurrent points, the number is set to -1. The ratio depicted in (f) is simply the percentage of determinism divided by the percentage of recurrent points. In diagram (g) we have the Shannon information entropy of the distribution of parallel segments to the main diagonal of the RP

---

[55]We had a window of 260 observations which was rolled on the returns sequence with a step (shift) of 5 observations each time. This generated 779 samples or "epochs".

(a) Mean of input points

(b) Std Dev of input points

(c) Mean of rescaled distances

(d) Recurrent points

(e) Deterministic points

(f) Ratio

(g) Entropy of line distribution

(h) Longest diagonal line segment
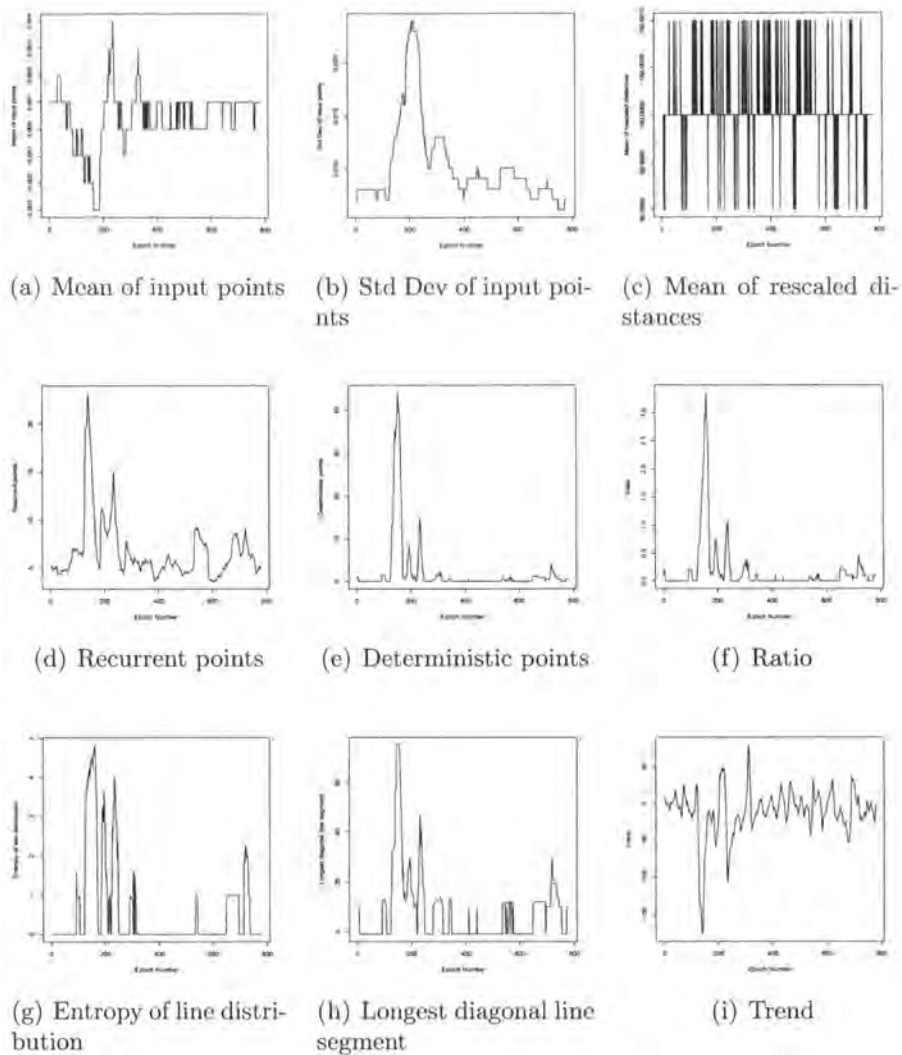
(i) Trend

Figure 5.12: RQA results on the FTSE returns.

(measured in bits). Diagram (h) is the longest diagonal line segment plot. This is inversely proportional to the largest positive Lyapunov exponent. Finally, the trend diagram (i) is the slope of the regression of the percentage of recurrence in (d) on the displacement from the main diagonal, expressed in units of percentage of local recurrence per 1000 points.

RQA is not always very straightforward. Usually we should be able to locate upper and lower bounds of the (a-i) sequences in figure 5.12. We can clearly see that there is a large peak around epoch 200 which coincides with the recurrences

(a) Mean of input points

(b) Std Dev of input points

(c) Mean of rescaled distances

(d) Recurrent points

(e) Deterministic points

(f) Ratio

(g) Entropy of line distribution

(h) Longest diagonal line segment

(i) Trend
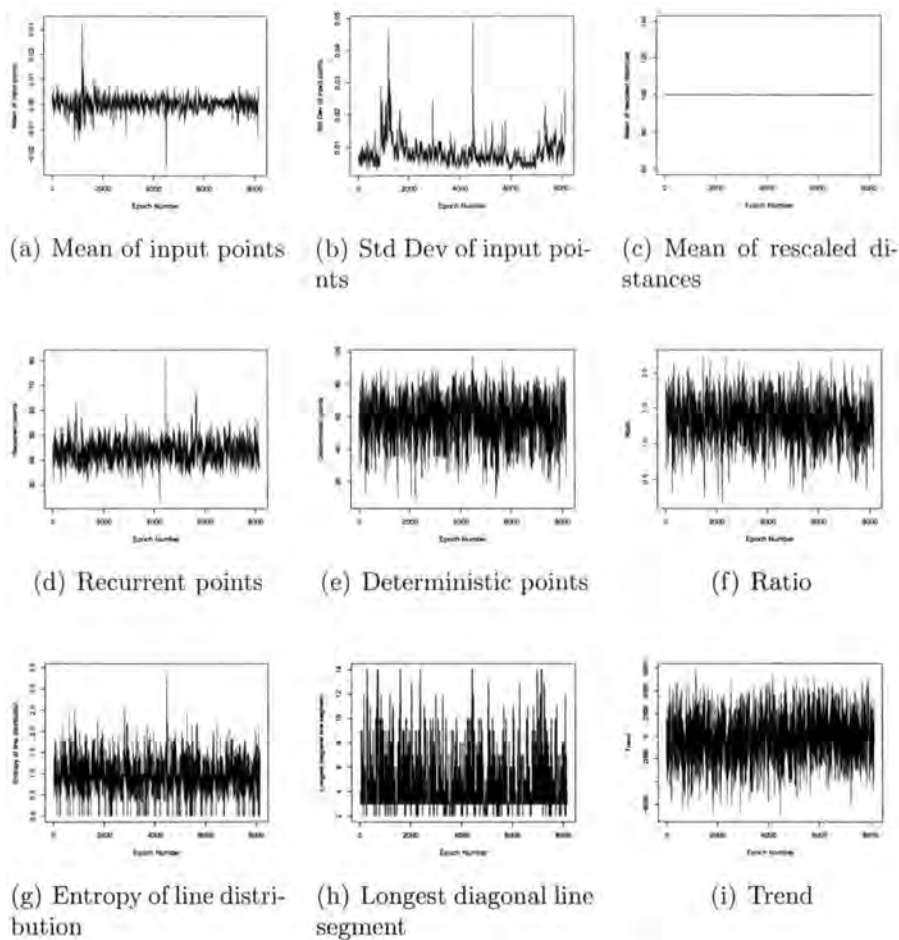
Figure 5.13: RQA results on the FTSE returns.

of the oil crisis. It seems that the volatility in that period has a very significant impact on the phase space dynamics. Overall the structures in these diagrams are consistent with some nonlinear deterministic nature in the returns sequences and not some purely stochastic structure. The maximum line length in (h) fluctuates as well with a maximum value during the oil crisis period. Apparently, the strongest determinism is located around that time. In figures 5.13 (a) to (i) we have generated the RQA measurements for the maximum possible epochs (8170 epochs, for 5 days windows with 1 day shift at a time) in order to correlate the RQA measurements to the original returns sequence topology. From plots (a) and (d) we see that the highest peaks are located close to the observations related to the oil crisis and the

crash of 1987. The maximum recurrence and entropy is located on the crash of 1987. The maximum line length fluctuates between 2 and 14 for the whole period which implies the presence of a positive largest Lyapunov exponent. It is straightforward from are analysis here that there is a strong deterministic characteristic for the FTSE returns, whether one looks at the series with a weekly window or a daily one. Similar characteristics were obtained for other frequencies as well which implies that there is more than stochastic randomness at hand. All results are consistent with the corresponding RPs.

An interesting discussion can be generated around the structures of the plots of RQA measurements especially in figure 5.12. In these we examine the recurrences on a weekly rolling window and we can see clearly that they exhibit high spikes for the epochs around 200. We identified this as the effect of the oil crisis of the 70s. An interesting research objective would be to relate these changes in dynamics with a "transition to chaos". If there are elements of chaotic determinism in financial time series, we could detect via RQA the point where stochastic dynamics change to low dimensional deterministic dynamics. From figure 5.12 (e) we see the % of deterministic points is close to 0 but during crisis periods this number shoots up to very high levels. This could be an indication that in times of "financial chaos", there is indeed a specific structure within the turbulence and the volatility of the markets which could be detected according to our time window. To analyse further, is seems that when using a weekly rolling window, one can see in 5.12 (e) that there are periods of "stochastic" calmness interrupted by periods of deterministic structure in the dynamics (during the various financial crisis). Looking at 5.13 (e), we can see that using a daily window, one can use information to forecast the high-frequency dynamics more accurately as determinism is quite high. Comparing 5.13 (e) with 5.12 (e) one can start understanding how RQA could be used to determine time-windows for risky periods.

5.5 Conclusions

In this chapter we followed a geometric-topological approach in analysing returns of stock market indices in daily frequencies. We used the AMI and FNN

criteria to establish an embedding in order to reconstruct the phase space dynamics. These dynamics we visualised with thresholded and unthresholded RPs. It is shown that these indices are not following a random walk which could be perceived as a contradiction to the Efficient Market Hypothesis. There are also visual implications of the existence of self-similarity in the recurring patterns and the dynamics of the series, as these are viewed through the RPs. It is fairly intuitive that one may make "careful" use RPs to model time-series after they have been properly filtered. This may reveal more deterministic features. Nevertheless, recurrence quantification analysis on the index returns has revealed interesting characteristics. We indicated that it is possible to discover through this technique when series become "more chaotic" or whether indeed the DGP is characterised by stochastic randomness or determinism. We believe that it is fairly self explanatory that recurrence plots and the corresponding quantification analysis can provide very powerful analytical tools.

In the following chapter we focus on a statistical hypothesis framework called "Surrogate Data Analysis". We will be showing how ideas discussed so far, can be used to set up a hypothesis of linearity which when refuted would provide a plausible explanation for the underlying dynamics. We will be using various simulations and real-life data to demonstrate the power of such an approach.

CHAPTER 6
Surrogate Data Analysis

6.1 Introduction

It is a widely accepted fact that there has been no well defined statistical hypothesis testing procedure publicised so far that would determine whether a time series is chaotic or not. There is no widely established and powerful hypothesis testing framework of a nonlinear deterministic null against the alternative of stochastic data generating process. Many empirical problems require detecting whether the system under examination is characterised by deterministic or stochastic dynamics. Moreover, it is often necessary to establish not only the nature of the dynamics but also the dimensionality of the system. This becomes increasingly difficult when one takes into account that many phenomena that may in fact be nonlinear-deterministic, may posses dynamics of higher dimensions which could render them indistinguishable from stochastic processes. This creates the need for a hypothesis testing framework that would also detect successfully weak determinism.

In this chapter we suggest the use of the *Surrogate Data Analysis* framework in order to detect nonlinearity and determinism. Surrogate analysis enables us essentially to test whether the dynamics are consistent with linearly filtered noise or a nonlinear dynamical process. We describe the various null hypothesis considered under this framework, the most commonly used discriminating (test) statistics, the most popular surrogate data generating procedures, their properties and the pitfalls of the whole approach. We finally apply a number of tests to financial time series in order to dismiss or accept stochasticity in their data generating process. A direct implication is that the surrogate data analysis framework can produce more evidence for or against the Efficient Market Hypothesis (EMH). Moreover it can provide a suitable "data-mining" platform or an improved description of the dynamics of the data, predetermining thus the appropriateness of a parametric or nonparametric analysis approach.

6.2 Previous research

During the last decades, a number of methodologies were designed to detect nonlinearities in time series. McLeod and Li (1983), Keenan (1985) and Hinich (1982) among others were some of the most popular earlier attempts to define tests for the detection of nonlinearity in time series. Bilinear models (Granger and Anderson, 1978), threshold autoregressive (TAR) models (Tong, 1978) or Volterra expansions (Nisio, 1960) have also enjoyed some success in this area. At the same time, applications focused on determining the independence of time series by means of functionals of first and second moments and the autocorrelation and partial autocorrelation functions (ACF and PACF respectively). Ljung and Box (1978) and later McLeod and Li (1983) (both in Harvey, 1994) proposed statistics for detecting model misspecification under the null hypothesis of an autoregressive moving average (ARMA) underlying process. Since the eighties, another modelling approach has enjoyed wide acceptance from the scientific community. The autoregressive conditional heteroscedasticity (ARCH) model (Engle, 1982) and later the generalised (GARCH) version of it (Bollerslev, 1986) have been used in many applications in order to explain inherent nonlinearities in economic-financial data and the peculiar structure of their volatilities. These models came as an answer to Mandelbrot's (1963) observation that financial returns series exhibited interchanging periods of volatility clustering and relative smoothness of dynamics ("tranquillity": see also other works of Mandelbrot, 2001a, 2001b, 2001c, 2001d, 1999a, 1999d, 1997a, 1997b, 1997c, 1997d and 1997e).

In this research we focus on determining between stochastic and deterministic dynamics for financial time series. In this respect, establishing nonlinearity for our series is not enough if this nonlinearity does not necessarily imply or provide evidence for the existence deterministic dynamics. Tests have been suggested during the past in this respect. Liu *et al.*, (1992) have proposed the use of the correlation exponent as means to decide whether and economics series is chaotic or not. Earlier Brock (1986), Brock and Baek (1991) and Brock and Sayers (1988) have entertained similar ideas. Brock *et al.*, (1987 and 1991) have proposed the BDS test as a means to determine the independence or *i.i.d.'ness* of a time series. Erroneously this has been used

frequently by others as a test for chaos although it is a test of independence. It was a first bold attempt though towards the designing of a test for chaos and was based on the correlation integral which is a chaotic invariant (Grassberger and Procaccia, 1983). Others followed with suggestions on the original ideas and applications such as the late Hiemstra[56] and Kelejian (1992), LeBaron et al., (1988) and Hsieh (1991). Ramsey and Yuan (1989 and 1990) have concentrated on the statistical validity of correlation dimension calculations with small data sets and explored the potential of such computations with economic data. An early approach similar to the one we suggest in this chapter was that of Sugihara and May (1990) who proposed the use of a nonlinear forecasting statistic as a means of distinguishing between deterministic and stochastic processes.

In order to bypass the shortcomings from the absence of a well-defined powerful statistical test for chaos, a new methodology has been proposed during the last decade. This framework is called "Surrogate Data Analysis" (SDA) (see in Theiler et al., 1992a, 1992b, 1992c, 1993). The general SDA procedure has been also described in Theiler (1995), Theiler and Prichard (1996) and Theiler and Rapp (1996). Among early workings on this area we have to mention the paper by Takens (1993).

The SDA is regarded as a logical and consistent statistical framework for testing hypothesis about time series. It has been applied extensively on physical signals and physiological time series but has seen very little or no application in the social sciences. This approach is based on an idea similar to bootstrapping (Efron, 1979, 1982 and Efron and Tibshirani, 1986). In fact it involves a technique such as bootstrapping, a kind of "permutation testing" (see Good, 1994 and Moore, 1999) as it will be explained. The aim of this technique is to simulate-produce a set of "surrogate" copies of the original time series that exhibit usually the same autocorrelation and spectral structure. A null hypothesis is formulated and the appropriate statistic is computed for the original and each of the surrogate series. If the original series statistic is outside the range of the surrogates one, we can safely conclude that the original series is inconsistent with the null hypothesis. Although the basic manipulations of the Bootstrap and SDA appear to be the same, the two

---

[56]One of the people who encouraged me to undertake this Ph.D.

methodologies are fundamentally different (Smith, 2001). The former makes a variety of assumptions about the data generating process (DGP) and then establishes the uncertainty of a discriminating statistic by assuming these assumptions hold. Through bootstrapping, an estimation of this uncertainty is computed and a distribution is constructed which is consistent with the true uncertainty of the observed results. The SDA on the other hand, identifies a process inconsistent with the one believed to be the DGP and attempts to establish that the value of the statistic obtained from the data is unlikely under the null hypothesis. A distribution is simulated from a well defined known "surrogate" process in order to show whether the observed result is inconsistent with this distribution or not, identifying thus the insignificance or significance of the "real" result. We should note here that surrogate data are created without replacement whereas in bootstrapping, resampling is done with replacement in order to generate an empirical distribution of a statistic. This is clearly why Moore (1999) argues[57] that the SDA approach is based rather on a permutation testing framework than a bootstrapping one. Moreover in SDA the samples are generated to be consistent with a null hypothesis which is one more strongly diversifying characteristic.

A problem that arises with either bootstrap or permutation tests is the complications that temporal dependencies (i.e., recurrences, autocorrelations, aperiodic or periodic cycles in general) can produce interesting artifacts that may mask though the dynamics of the sequences themselves and distort our view of the noise that contaminates the processes. It becomes increasingly difficult to discern between data and noise, especially when the DGPs of the noise or the data are unknown. Techniques that require the individual realisations to be independent, suffer most from these artifacts. Textbook approach suggests preprocessing the sequences via some kind of linear filters such as ARMA models which can remove linear dependencies, an approach that many researchers do not recommend. Abarbanel (1992, 1993 and 1995) and Theiler and Eubank (1993) among others demonstrate that nonlinearity in chaotic data is harder to detect when the original series are preprocessed. On the other hand, many techniques designed for the detection of nonlinear recurrences,

---

[57]Kaplan and Glass (1995) refer to the SDA method as "bootstrapping".

such as RQA,[58] lead to inaccurate conclusions when nonstationarity[59] is present. In general, there is a thin line between choosing to analyse between raw observations and transformations of those. In Finance for instance, the analysis of logarithmic returns may be of more empirical interest than examining the levels of the relevant closing prices. And differencing the series to obtain returns, is essentially a "high-pass" filter which may magnify the noise signature.

In this chapter we follow mainly the formulations and discussions by Schreiber and Schmitz (2000), Kaplan and Glass (1995) and Kantz and Schreiber (1997). It seems that SDA has been developed independently by a number of researchers in the area. Generally, the paper by Theiler *et al.*, (1992) is regarded as the starting point of all discussions.[60] Very good insight is provided also by the famous paper by Theiler *et al.*, (1992) in "Nonlinear Modelling and Forecasting" by the Santa Fe Institute which specialises in areas of complexity and chaos theory research. Theiler *et al.*, (1993) focus on tests for nonlinearity with a null of linearly correlated Gaussian noise. This is not a investigation for chaos *per se* but for nonlinearity, searching for a significant statistic that would discriminate the original series from its linear surrogates. Surrogate data analysis is also discussed more recently in Schreiber (1998) and Urbach (1999) and mentioned in Williams (1997). SDA is a relatively new idea and has not found its way yet into the economic-social time series analysis curriculum or textbooks (although bootstrapping has). A forthcoming paper by Kugiumtzis (2000) included in the book "Nonlinear Deterministic Modelling and Forecasting of Economic and Financial Time Series"[61] sets the spot in our discipline.

6.3   Why use Surrogate Data?

Complex systems can exhibit chaos in the sense that their dynamics could be represented with a set of nonlinear (dynamical) differential equations. Analytically,

---

[58]RQA: Recurrence Quantification Analysis, see previous chapter.

[59]Schmitz and Schreiber (1999) note that a common approach is the segmentation of a series in parts that can be considered nearly stationary and the performance of the tests there. This though may not be feasible in small data sets or in the presence of slowly varying stationarity.

[60]In the classic Physica D volume 58 issue which is dedicated to nonlinear time series analysis and cited in hundreds of applications as a "seminal paper" collection volume.

[61]by A. Soofi and L.Cao editors, still unpublished while this thesis was being written.

this can be proven for a number of abstract, physical or theoretical systems. In empirical sciences, when one tries to determine this from an observable quantity generated by the system, approximating the dynamics with a set of equations may prove difficult. Moreover, determining whether the system exhibits stochastic or deterministic dynamics may be an even more difficult task. A starting point would be a set of system properties that define chaotic structure or behaviour. In the textbook approach, chaos or nonlinear determinism is characterised mainly by:

- determinism,

- non-cyclical periodicity (aperiodicity),

- "boundedness" of dynamics and

- sensitive dependence on initial conditions.

Although one could rush to regard these points as "stylised facts", one should also note that it has frequently been noted in the chaos literature that the above characteristics or properties of chaotic systems may not always clarify whether a system is deterministic or stochastic. While chaos can not develop from linear systems, there is a wide variety of such systems that would produce measurements which could easily be mistaken to exhibit chaotic structure. As discussed in Kaplan and Glass (1995), all four characteristics mentioned above, may not help us conclude whether a sequence is truly random or chaotic. Moreover, parametric and nonparametric tests (such as the RUNS test, the BDS, the Augmented Dickey-Fuller or the portmanteau Q-test for instance) may detect successfully independence or stationarity but do not provide an alternative hypothesis that would support non-randomness in the sense of "chaos".[62] Most tests operate in a "linear" function space which deprives them of the ability to discern deterministic nonlinearities and these are frequently misinterpreted as stochastic randomness. An added problem is the nature and length of the data and the noise contaminating them.

---

[62]Specifically, independence as a null is not a very "interesting" hypothesis for most data as argued in Schreiber (1998): "*It becomes relevant when the residual errors of a time series model is evaluated. For example in the BDS test (Brock et al., 1988 and 1991), an ARMA model is fitted to the data. If the data are linear, then the residuals are expected to be independent*". The importance of this remark becomes even stronger when BDS is used erroneously to detect chaos instead of independence.

In social sciences the data come from unknown data generating processes and are "contaminated" by noise of unknown nature. As we have already discussed, the limited length and precision of the data may lead to miscalculated nonlinearity statistics. On the other hand, one can only suggest a suitable "stochastic" hypothesis testing framework when backed by sound theoretical reasoning and robust supportive empirical findings. This can sometimes be fairly arbitrary and misleading. Any kind of manipulation of deterministic data may destroy the fine structures of the dynamics and provide false evidence of stochasticity (see Abarbanel, 1995 and Theiler and Eubank, 1993). We should also note her that so far we have concentrated on the description of the dynamics of the DGP but not commented on the noise that may be masking them. In Economics or Finance, one of the true problems is deciding on the nature of the noise. This is quite difficult in social sciences as one does not usually know the cause of it (for an excellent discussion see Black, 1990). Usually, one determines the distribution properties of the noise according to sound theoretical assumptions and empirical evidence. In social sciences in general, it is extremely difficult or even impossible to clearly identify the structure of the noise contaminating the processes examined. Only assumptions can be made and these should be verified by some appropriate statistical hypothesis framework, often dictated by the theoretical approach and the surrounding scientific literature. This difficulty is augmented by the fact that most series in Finance or Economics should be regarded as strictly stochastic. This would allow specific market-clearing or equilibrium conditions to be met under the relative utility constrained or unconstrained optimisation frameworks.

In this chapter we suggest the alternative of SDA as it provides a way round the problems discussed above. As an added plus, it is based on a statistical hypothesis testing procedure. Under this, there are a number of different nulls that can be considered. Usually the null hypothesis formulated under SDA is the one that postulates that the dynamics are linear with Gaussian white noise random inputs (Kaplan and Glass, 1995). This is called the *linear-dynamics null hypothesis*. If the time series in question is chaotic, "in principle" this null should be rejected.

Concluding this discussion it would be appropriate to mention here what is also pointed out in Kantz and Schreiber (1999) and Kaplan and Glass (1995). Quite a few pieces of past research in chaos are satisfied with simply the computation of single chaotic invariants. In that sense, the mere determination of a positive Lyapunov exponent, a finite fractal dimension or a suitable BDS test is evidence enough to refute stochastic randomness. In reality this should not be regarded as an accurate conclusion. As pointed out earlier, linearly structured dynamics can yield outputs which can "mask" themselves as deterministic. The only safeguard against this is the determination of scaling regions i.e., confidence intervals which when containing the computed discriminating statistics, would lead us to accept their validity. In other words, one should check what would the values of the statistic be if the data were indeed generated by a linear model. If this is not consistent with the one obtained by the original series, this could constitute evidence of nonlinearity.

6.4 The general SDA hypothesis testing framework

The objective of SDA is to determine the nature of the data generating process behind the series. As a second objective, and according to which kind of "discriminating" statistic (as it is called) we use to test the null hypothesis, SDA will determine the significance of this statistical measure. The methodology focuses on the comparison of the results with those of linear-stochastic sequences. Moreover, an insight of the different states of the system in examination may be provided as a by-product of the SDA approach. The discriminating statistic is compared to the distribution of statistic values consistent with the given hypothesis. Wherever there is a difference we reject the null. Again, rushing to conclude under SDA about randomness or determinism should be avoided. As Theiler et al., (1996) point out, "The null hypothesis corresponds to an answer of 'no', and is the default conclusion in the lack of contrary evidence. One does not positively prove (or disprove) the null hypothesis; instead one attempts to reject the null hypothesis by showing that the data are unlikely to have resulted from it".

6.4.1 SDA Hypothesis and Statistics

Theiler and Prichard (1996) provide two classifications of discriminating statistics: "*pivotal*" and "*non-pivotal*". A test statistic t is called *pivotal* if its probability distribution (given the data generating process) is the same for all processes F consistent with the hypothesis. If this does not hold, the statistic is called *non-pivotal*. Let ϕ be a specific hypothesis and $\mathcal{F}_\phi$ the set of all processes (systems) consistent with that hypothesis. It is well known from basic statistical inference theory that a hypothesis is *simple* if all DGPs consistent with $\mathcal{F}_\phi$ constitute a singleton set. Otherwise the hypothesis is termed as *composite*. It follows that under a composite hypothesis, the problem is not only to generate the surrogate data consistent with the process F but also to estimate $F \in \mathcal{F}_\phi$ (Small *et al.*, 2001). The type of the hypothesis will dictate the type of the statistic. Theiler and Prichard (1996) propose the use of a pivotal test statistic in the case of a composite null. Unless the discriminating statistic is pivotal, one has to specify precisely the process F. In the case a non-pivotal statistic must be used with a composite hypothesis, they suggest that a type of "*constrained-realisation*"[63] process should be utilised for the generation of the surrogates. This is also discussed in detail in Small and Judd (1998). It effectively implies that as well as generating surrogate data that are typical realisations of a model of the original observations, it must be ensured that these surrogates are also realisations of a process that provides identical estimate values with the estimates of the parameters from the original data. This is what clearly discriminates SDA from bootstrapping as we already discussed above.

The SDA hypothesis testing procedure is not entirely as straightforward as one would like. Not all statistics perform equally well and not all hypotheses are clear-cut and well defined to suit the purpose (Small *et al.*, 2001). The Correlation Dimension D_2 as we will discuss, presents a favourable choice for a discriminating statistic. Small *et al.*, (2001) demonstrate a method where the hypothesis tested is not known beforehand, but is determined by the "pivotalness" of the test statistic. They show that nonlinear models can be used to generate surrogates as well as test

---

[63]Assume a time series X_n and $\hat{F} \in \mathcal{F}_\phi$ the process estimated from X_n. Let z_i be the surrogate data set generated from $F_i \in \mathcal{F}_\phi$. Let $\hat{F}_i \in \mathcal{F}_\phi$ be the process estimated by z_i. We call surrogate z_i a "*constrained realisation*" if $\hat{F}_i = \hat{F}$ and "*unconstrained*" if $\hat{F}_i \neq \hat{F}$.

nonlinear hypothesis. Because D_2 is pivotal, they avoid the necessity to employ a constrained realisation surrogate generator method. They suggest the generation of many noise driven simulations of the original data, based on nonlinear models, and the comparison of the distributions of a test statistic for each model (or groups of models) with that of the original observations. This comparison will determine the acceptance or rejection of the hypothesis that the original data sequence was generated by a process with the same structure as one of the models.

6.4.1.1 Types of null Hypothesis

The different memberships to specific classes of dynamical systems are reflected in different types of null hypothesis. Theiler *et al.*, (1992) describe three basic types of the null:

- H_1: white noise;

- H_2: linearly filtered white noise;

- H_3: monotonic nonlinear transformation of linearly filtered noise.

Assume t_{orig} is the value of the statistic calculated on the original data and t_{surr} the value of the same measure calculated for each surrogate sequence. As we already mentioned, the null is rejected when the test statistic from the original series lies outside the range of prices of the test statistics for the surrogate data or when t_{orig} and $E[t_{\text{surr}}]$ are sufficiently different. Rapp *et al.*, (1994 and 2001) suggest three criteria for accepting or rejecting the null:

- Criterion 1:

  The null is rejected if $t_{\text{orig}} > t_{\text{surr}}$ or $t_{\text{orig}} < t_{\text{surr}}$ for all t_{surr}. This is termed the *"nonparametric criterion"* (Hope, 1968). A single larger surrogate statistic value can lead under this criterion to a rejection of the null.

- Criterion 2:

  This is based on a Monte-Carlo probability estimation of the surrogate null (p-value) P:

$$P_{\text{crit}_2} = \frac{\text{number of cases that } t_{\text{surr}} \preceq t_{\text{orig}}}{\text{number of surrogate samples} + 1} \tag{6.1}$$

The numerator of (6.1) may also contain the $\succeq$ instead of the $\preceq$ inequality operator. This is determined by the type of the discriminating statistic used on the data. This criterion has good power when the number of surrogate samples is small.

- Criterion 3:

  In the case the values of t_{surr} are distributed normally, we can define the following criterion:

  $$Z = \frac{|t_{\text{orig}} - E[t_{\text{surr}}]|}{\sigma_{\text{surr}}} \tag{6.2}$$

  where σ_{surr} is the standard deviation of $E[t_{\text{surr}}]$. The p-value then is given by

  $$P_{\text{crit}_3} = \frac{1}{2}(1 - \text{erf}(\frac{Z}{\sqrt{2}})) \tag{6.3}$$

  where "erf" is the error function:

  $$\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-s^2} ds \tag{6.4}$$

  For one-tailed test, the null is rejected at $\alpha = 5\%$ statistical significance level if $Z \preceq 1.65$ and for $\alpha = 1\%$ if $Z \preceq 2.33$.

A legitimate question that arises is how many samples of surrogate data should be simulated. This number in general depends on the level of significance α. For a given α and a one-tailed test, $1/\alpha - 1$ surrogates are required on a minimum basis. For a two-tailed test, one should acquire at least $2/\alpha - 1$ surrogate samples. For example, for $\alpha = 5\%$, one would normally require 19 surrogate data sets for a one tailed test and 39 for a two-tailed one.

The three hypothesis H_1, H_2 and H_3 are rejected for the simple shuffled surrogates, FT surrogates and AAFT surrogates (and variants) respectively, for data generated by a nonlinear system. However, we reiterate that rejecting these hypotheses does not necessarily imply that the cause is a nonlinear deterministic DGP. For example, from the rejection of hypothesis H_3, it is only safe to conclude that it is unlikely that the original observations sprang from a *monotonic nonlinear transformation of linearly filtered noise*.

6.4.1.2 The SDA discriminating statistics

Usually, the type of the null hypothesis will dictate the nature of the discriminating statistic. One may use the ACF in order to refute linearity and Gaussianity although this may not be very interesting given the nature of our problem (Kantz and Schreiber, 1999). Schreiber and Schmitz (1997) examine the power of an array of statistics of nonlinearity and complexity measures and find that it is preferable to use an algorithm with good overall performance such as higher order autocorrelations or nonlinear predictor errors. They evaluate different nonlinear observables which require the embedding of the time series in phase space. It is known that the embedding parameters should be determined accurately in this case and that may prove to be a weakness of the use of these proposed statistics. Given that we can successfully determine the time delay τ and the embedding dimension d_e parameters for the original series, by fixing these quantities for all surrogate data sequences we can conduct the relevant tests. It is incorrect to recalculate these observables for every surrogate sequence (Kantz and Schreiber, 1999). Popular statistics to use within the SDA framework are:

- The Grassberger and Procaccia (1983) measure of correlation dimension D_2. This is based on the correlation sum $C(\epsilon)$ which at scale ϵ is defined as:

$$C(e) = \text{constant} \times \sum_{|i-j|>t_{min}} \Theta(\|\vec{x_i} - \vec{x_j}\| - \epsilon) \qquad (6.5)$$

for a fixed minimum distance t_{min} and Θ the Heaviside function. The linear processes consistent with the H_1, H_2 and H_3 hypothesis stated above, are effectively all forms of filtered noise and therefore infinitely dimensional. It follows that a measure such as the correlation dimension D_2 will be also infinite. Hence D_2 can be regarded as a pivotal statistic (for a proof see Small *et al.*, 2001). Schreiber and Schmitz (1997) propose the use of the *Taken's estimator* (Theiler, 1988) for the correlation dimension D_2, which is the maximum likelihood estimator of the Grassberger-Procaccia measure:

$$t^{ML}(m, \tau, \epsilon) = \frac{C_m(\epsilon)}{\int_0^\epsilon \frac{C_m(\epsilon')}{\epsilon'} d\epsilon'} \qquad (6.6)$$

They use the Takens (1993) estimator and the BDS (follows) statistic to provide single values for the correlation sum within the SDA hypothesis testing framework.

- The Brock *et al.*, (1988) statistic. Schreiber and Schmitz (1997) propose a simpler version:

$$t^{\text{BDS}}(m, \tau, \epsilon) = \frac{C_m(\epsilon)}{C_1(\epsilon)^m} \qquad (6.7)$$

in order to obtain an asymptotic form of the probability distribution.

- The False Nearest Neighbour statistic (FNN) by Kennel *et al.*, (1992), described earlier in this thesis. This is used to calculate the minimum embedding dimension d_e in the state reconstruction. If this criterion does not saturate for a very low percentage of false neighbours (usually 1%), then there is evidence that our series is either stochastic or severely contaminated with stochastic noise.

- The Mutual information (MI) criterion. Again, this statistic is used as the FNN criterion in delay coordinate embedding (see previous chapters). MI is based on measures of general correlation, both linear and nonlinear and can show on average how much of information past subsequences can offer on future subsequences. For detailed discussions see Granger and Lin (1994) and Fraser and Swinney (1986).

- A version of the nonlinear prediction error defined as

$$t^{\text{NLPE}}(m, \tau, \epsilon) = (\sum [x_{n+1} - F(x_n)]^2)^{\frac{1}{2}} \qquad (6.8)$$

where F is a locally constant predictor and the prediction over one time step is performed by averaging the future values of all neighbouring delay vectors in m dimensions which are closer than a small distance ϵ.

- Higher order autocovariances (or cumulants).

- A *time reversibility* statistic[64] which is based on the concept of "directionality",

under which series differ when viewed in reverse time ordering:

$$t^{\text{REV}}(\tau) = \frac{\langle (x_n - x_{n-\tau})^3 \rangle}{\langle (x_n - x_{n-\tau})^2 \rangle} \tag{6.9}$$

This statistic is known as "time reversal asymmetry". The time reversibility statistics are explained in Diks et al., (1995) and Diks (1999). He proposes an approach that does not rely on SDA but can be combined with a Monte-Carlo like test procedure if needed. In general, these kind of statistics[65] belong to a wider class of simple nonlinear averages statistics and are related to the bispectral analysis and the bilinearity principle (Rao and Gabr, 1984). A similar to time-reversibility measure is the "crinkle" statistic introduced by Theiler based on the 4th moment of the series X_n:

$$\frac{\langle (X_{t-1} - 2X_t + X_{t+1})^4 \rangle}{\langle (X_t^2) \rangle^2} \tag{6.10}$$

for $t = 1, 2, 3, ..., n$. Note that $X_{t-1} - 2X_t + X_{t+1}$ refers to the 2nd difference of the series.

- Volterra polynomials are also considered for the generation of correlation based statistics (Kugiumtzis, 1999).

- The largest Lyapunov exponents (LLE) is another popular statistic in SDA as an indication of the complexity of the DGP.

6.5 Types of Surrogates

Over the last ten years, various procedures for generating surrogate data have been proposed. Most of them are based in permutations of the original data set and

---

[64]A stationary time series X_n is said to be *reversible* of its probabilistic properties are invariant under time reversal i.e., $(X_n, X_{n+1}, X_{n+3}, ..., X_{n+k})$ and $(X_{n+k}, X_{n+k-1}, X_{n+k-2}, ..., X_1)$ have the same joint probability distribution for all k and n (Weiss, 1975 and Lawrence, 1991). This is a very important property as any presence of "directionality" excludes linear Gaussian processes. Lawrence (1991) argues: "The view taken here is that directionality is an aspect of time series analysis which deserves wider recognition; for instance, it does not make sense to forecast with a time series model which is reversible, when past data are definitely irreversible. In simulating inputs to a system based on directional historical data, directional simulated data should be used. Such obvious requirements are not met by the use of Gaussian ARMA models. "

[65]Time-reversibility statistics will be covered more analytically in the last part of this chapter.

are relatively computationally cheap. As Theiler and Prichard (1996) and Small *et al.*, (2001) point out, the type of null hypothesis dictates the type of the statistic to use and this in turn dictates the type of procedure to be used in generating the surrogate data. In this section we will concentrate our discussion on the most popular surrogate data generators and examine their properties.

6.5.1 Simple shuffled surrogate data

This is the simplest type of surrogates and should be generally avoided unless we are certain of the Gaussianity of the underlying DGP process. The surrogates are generated by shuffling the order of the original observations (via Monte-Carlo). A Gaussian independent identically distributed (i.i.d.) data set is simulated via a pseudo-random number generator and the original sequence is reordered so that it has the same rank distribution as the simulated set. The hypothesis tested here is essentially that the given series is IDD according to H_1. These surrogates are constructed to exhibit the same mean and variance as the original sequence. This approach as well as the type of the hypothesis tested is not regarded as very interesting and at the same time does not provide any valuable insight in the alternative of a nonlinear deterministic DGP hypothesis.

6.5.2 Fourier Transform based surrogate data

Under the SDA framework, a suitable null is that the series under examination are linear and more precisely that the DGP is linear stochastic with Gaussian innovations. This is effectively what hypothesis H_2 pertains to. It is imperative that the surrogate series are "properly" generated (Schreiber, 1998) as any differences with the original series may be misinterpreted as evidence of nonlinearity. Moreover, the correct size of the test will be directly related to the way surrogate data are produced. The surrogate data are constructed to be consistent with the null hypothesis. The most common surrogate data generation algorithm is the one that utilises a discrete Fourier transform (DFT) of the original observations. The phases for each frequency are replaced with random numbers from the interval $[0, 2\pi)$ while the magnitude at each frequency i.e., the power spectrum is kept as is. The last

stage incorporates an inverse Fourier transform (IFT) which produces the surrogate sequences. The process can be summarised in the following 3 steps:

1. First compute the Fourier transform of the original sequence, i.e., an amplitude $A(\omega)$ and a phase $\phi(\omega)$ at each frequency ω.

2. Replace the phases $\phi(\omega)$ with random numbers between 0 and 2π.

3. Compute the inverse Fourier transform of the amplitude and the randomised phase and obtain the surrogate data.

This approach is an improvement over the previous "simply shuffling" algorithm and results to surrogate data which have the same amplitude and power spectrum as the original sequence. Intuitively, the autocorrelation function which is the Fourier transform of the power spectrum, is the same as well. In this sense, the original and surrogate sequences can not be distinguished by means of autocorrelation functions. The rationale behind this process is explained clearly in Kaplan and Glass (1995). For a linear Gaussian process, all dynamics and properties are based on the first and second moments and the autocorrelation function. In order for the structure to be preserved in the surrogate data, the mean, the variance and the autocorrelation function should be identical to that of the original sequence. This approach results to surrogate data which have the same amplitude and power spectrum as the original sequence. Intuitively, the autocorrelation function which is the Fourier transform of the power spectrum, is the same as well. In this sense, the original and surrogate sequences can not be distinguished by means of autocorrelation functions under the null. The process described above can be compared to that of a nonparametric bootstrap (Theiler *et al.*, 1993). The surrogate data will have exactly the same power spectrum of the original series and by the Wiener-Kintchine theorem they will also exhibit the same autocorrelation function. The reason why we engage in such a practice to design FT surrogates, is also analysed in Kaplan and Glass (1995). The FT surrogates are also called *"phase randomised"* surrogates.

6.5.3 Amplitude Adjusted Fourier transformed surrogate data

The FT (or phase randomised) surrogates, may produce spurious results if the original time series consists of linearly correlated noise that has been transformed by a static, monotone nonlinearity (Rapp *et al.*, 1994). Following Theiler *et al.*, (1992), a consistent surrogate data generator procedure with hypothesis H_3 is a modification of the previous algorithm of FT surrogates:

1. Starting with the original sequence X_n, generate an i.i.d. Gaussian data set Y.

2. Reorder Y so that it has the same rank distribution as X_n.

3. Use the FT surrogate algorithm to generate a surrogate Y_i of the Y i.i.d. sequence.

4. Reorder the original data X_n to generate a surrogate X_{ni} which will have the same rank distribution as Y_i.

These surrogates are referred to as "*Amplitude adjusted Fourier Transformed*" surrogates or AAFT for sort. This has been the algorithm of choice so far for many applications. Again, it fails to perform properly in the presence of linear correlations (Palus, 1995). Kugiumtzis (1999, 2000 and 2001) discusses that in AAFT surrogates we assume inherently that the static transform in the null is monotonic. This is not as safe assumption for real life data. Both non-monotonicity of transformations and nonlinear dynamics may lead to linear correlations that would tend towards the rejection of the null when we use statistics sensitive to these correlations. Schreiber and Schmitz (1996) propose an improved procedure for the generation of AAFT surrogates ("Iteratively" refined surrogates: IAFFT). This involves a two step modification which processes the original linear correlations in the first step and the original data cumulative density function as a second step. IAFFT surrogates are shown to approximate reasonably well the original linear correlations. A small linear correlation bias and variance is introduced though in the surrogate data, which when combined may be significant. Schreiber (1998) proposes a simulated annealing algorithm to overcome these problems. This comes though with increased algorithmic complexity

and computational cost. A correction of the bias just discussed, leads to yet another modification of the AAFT surrogates, the corrected AAFT surrogates or CAAFT.

6.5.4 Corrected AAFT surrogates

Corrected Amplitude Adjusted Fourier Transform surrogates (CAAFT) were introduced by Kugiumtzis (2000) as a correction to AAFT surrogates. The design mimics exactly the original data cumulative density function and linear correlations (on average). The CAAFT surrogates are statically transformed realisations of a Gaussian autoregressive process, which have the same amplitude distribution and autocorrelation as the original data X_n. The generated surrogate data are like AAFT surrogates, but corrected to match the autocorrelation function. The correction requires a linear interpolation for the graph of the relation between the Gaussian and the transformed autocorrelation, for lags up to a given lag $\tau_{\max}$. This interpolation function is then used to estimate the ACF of the series given the ACF of the original sequence X_n. Based on this autocorrelation, the coefficients of the corresponding AR(p) model are estimated and an autoregressive time-series is generated. This series is transformed to match the amplitude distribution of the original sequence X_n. The same process is then repeated a set number of times to obtain a statistic for the candidate AR models. Finally the most proper is selected, so as the autocorrelation of the generated surrogate sequence approximates best the autocorrelation of X_n (comparison done via Euclidean norm). Based on this AR model, a number of realisations are generated and transformed to match the amplitude distribution of X_n. This algorithm is provided by Kugiumtzis (2000).

6.5.5 ARMA based surrogates

An alternative approach is to generate the surrogates by fitting directly the original series to a finite order ARMA(p,q) model:

$$y_t = y_0 + \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{i=0}^{q} b_i e_{t-i} \tag{6.11}$$

This involves choosing the correct order values for the AR() and MA() parameters p and q (usually using some informational criterion such as the Akaike (1974) (AIC) or the Schwarz-Bayesian (1978) (BIC) -the latter being more powerful for small data sets- and by examining the stationarity of the residual processes). Over-fitting should be avoided and a parsimonious model should be obtained. Stability can be checked through the Yule-Walker equations. When an appropriate ARMA model has been fitted on the data, one can generate the surrogate sequences by introducing Gaussian i.i.d. random numbers into the error terms e_t and iterate the equation (6.11). This will produce surrogates with slightly different statistics than those of the FT surrogate procedure (Theiler *et al.*, 1993). An improvement would be bootstrapping the residual sequences e_t, reshuffling them randomly and re-inserting them into equation (6.11). By this way the Gaussian assumption is relaxed, we can obtain a wider class of time series but understandably the null is slightly different in such a case. Again here we reiterate the fact that since we are focusing in linear processes with non-Gaussian innovations, the behaviour we obtain may not be linear. References for this can be found in Theiler *et al.*, (1993).

Since FT and ARMA surrogates depend on the autocorrelation function of the original data set through their the Fourier and AR coefficients, both kinds of surrogates may behave quite similarly. The basic difference is that the FT based surrogates, will match exactly the first and second moments of the original sequence plus the autocorrelation function. ARMA surrogate statistics will usually be approximate to the original i.e., not exactly the same. Moreover, while FT surrogates are produced by a procedure that attempts to replicate the data (spectral and autocorrelation structure), ARMA ones are obtained by a procedure that attempts to generate a model that fits the original series.

Another difference is that the FT-based surrogates will be series generated all at once, of the same length of the original data. The ARMA surrogate procedure, works iteratively on the data and generates new points, one at every iteration and can generate any length of series. That could be potentially useful for various applications or may pose a problem as small errors can be sequentially propagated to large future deviations when modelling long term phenomena.

Essentially, the FT approach is nonparametric as it does not directly attempt to fit any kind of linear or nonlinear model to the data. One can argue that the large number of Fourier coefficients may be compared to the parsimony of a few parameter ARMA model and that should lead us to accept the latter as a more desirable solution (provided that the parameters are carefully chosen). In practice, the application dictates the modelling approach. Theiler *et al.*, (1993) argue that the FT surrogates will be more suitable for testing hypothesis whereas ARMA surrogates may be more appropriate for determining confidence intervals.

6.5.6 Pseudo-periodic surrogates

An interesting use of the SDA framework is reported in Theiler (1992, 1994 and 1995) and Theiler and Rapp (1996). They address the problem of detecting the presence of temporal correlations between cycles. They suggest that the surrogates there should also exhibit periodicity and propose the decomposition of the signal to its cycles and then the shuffling of the individual cycles. The purpose here is to reject the null of an absence of dynamical correlation between the cycles.[66]

Based on the idea discussed above, a modification of the surrogate data generation techniques was proposed by Small and Judd (1998a and 1998b), Small *et al.*, (2001a, 1001b and 2002a) and Small and Tse (2002). They term these surrogates "*pseudo-periodic*" (PPS). This kind of surrogates allows for the detection of aperiodic determinism, a fine cyclicity in the signals that may be destroyed by any phase randomisation. The hypothesis they test is that of a series being generated by noise driven by some periodic mechanism against that of non-periodic deterministic dynamics. Pseudo-periodic surrogates must exhibit two properties:

1. The surrogate data must exhibit the same periodic structure as the original sequence.

2. There must be no other deterministic structure present in the data.

In order to use the PPS framework, delay-coordinate embedding is applied to the data as a first step. For more details on this kind of surrogates one should refer to

---

[66]This approach may prove an interesting test for business-cycle applications in macroeconomic time-series analysis.

the articles by Small.

6.6 SDA applications

In this part we have generated a number of simulations of data that are characterised by deterministic nonlinearity or stochasticity and conducted SDA tests, in order to examine the performance of a subset of algorithms we have presented earlier. Continuing, we applied SDA on real life data. We choose 9 financial markets and execute the SDA on their main stock market indices. More precisely, we select USA (Dow Jones industrial average and NYSE New-York stock exchange indices), Canada (Toronto stock exchange TSE 300 index), Germany (DAX 30 Performance index), Hong-Kong (Hang-Seng index), Japan (Nikkei 225 stock average index), Mexico (IPC Bolsa index), Greece (ASE main market index), and UK (FTSE ALL SHARE index). In that way we cover 3 series from North America, 3 series from Europe and 3 series worldwide (including Latin America). Data are daily closing prices and the time span ranges from 1/1/1980 till 31/5/2002 i.e., 5849 daily observations. The Greek and Mexican time series are slightly smaller in length. Greek observations start from 30/9/1988 (while Mexico starts from 4/1/1988). The source is Datastream International. The statistics we apply under an SDA framework are the Mutual Information based, the "crinkle" and the time-reversibility test. We avoid using any fractal dimension based estimation statistic. We do this for various reasons. The most basic ones are that proper embedding of the series is required and this may be tricky as already described in the previous chapter. Secondly, for financial data, dimension calculations do not saturate at specific levels due to the existence of heteroscedastic noise in the sequences. According to some preliminary tests[67] the correlation integral based dimension calculations did not offer any valid or strong insight in the nature of the DGP as they did not converge to low dimensions for a number of embedded lags.

Among all data, only the UK market (figure 6.1) seemed initially to saturate at a level between 20 and 25 dimensions. All other markets' estimations exhibited highly oscillatory behaviour. This may provide some evidence for higher dimension

---

[67]Results available from the author upon request.
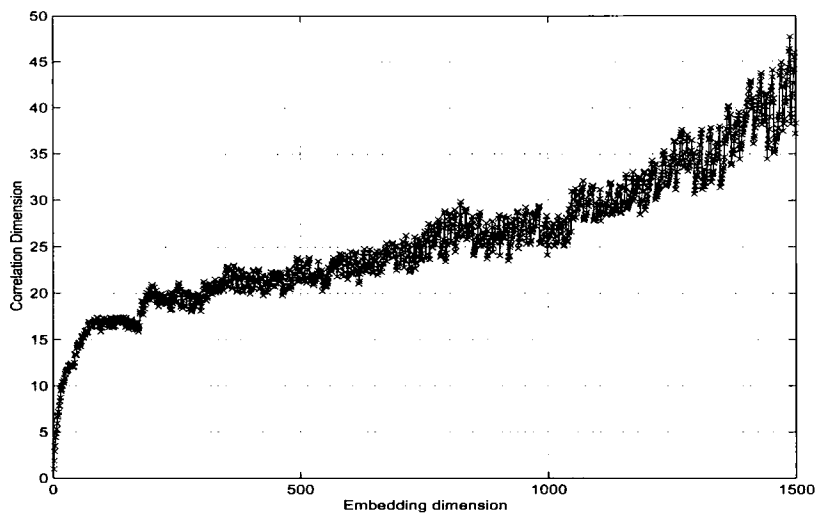
Figure 6.1: The correlation dimension plot for the UK market index returns. The results depicted here are not to be trusted as we do not have a real saturation for low embedding dimensions and the diagram shoots off for large ones. The saturation appears to be smooth for low embedding dimensions but becomes very erratic for larger ones as one would expect, given the nature of the data.

dynamics behind the DGP of this specific market. These results though are highly questionable as saturation should really be achieved for low dimension levels. The presence of noise in our data is expected to create severe problems in dimension estimation calculations.

6.6.1 Simulations

In figures (6.2) and (6.3) we show three types of surrogate data realisations for 3500 observations of the Lorenz "x" variable series. The three Lorenz differential equations define system that exhibits chaotic dynamics:

$$
\begin{aligned}
\dot{x} &= -\sigma x + \sigma y \\
\dot{y} &= -xz + rx - y \\
\dot{z} &= xy - bz
\end{aligned}
\tag{6.12}
$$

where $x, y, z \in \mathbb{R}$ and $\sigma, r, b > 0$.[68]
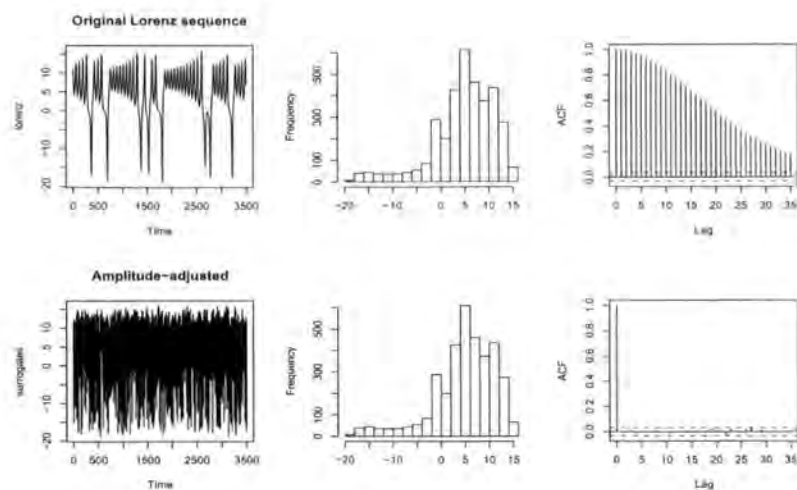
Figure 6.2: Original Lorenz sequence and amplitude adjusted surrogates. Note that although the histograms show the same distribution, the ACF functions differ.
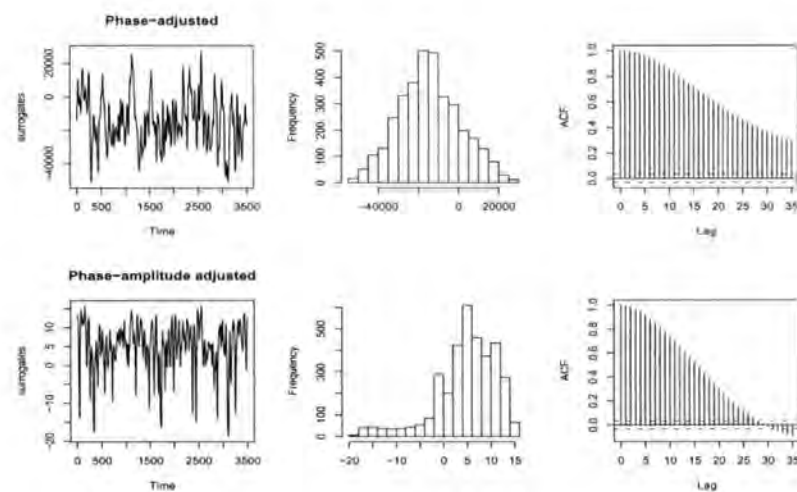
Figure 6.3: Phase and phase-amplitude adjusted surrogates of the Lorenz sequence in figure (6.2). Note the relationship between the ACF functions of these realisations and the original data. The phase-adjusted data are different as seen in their histogram.
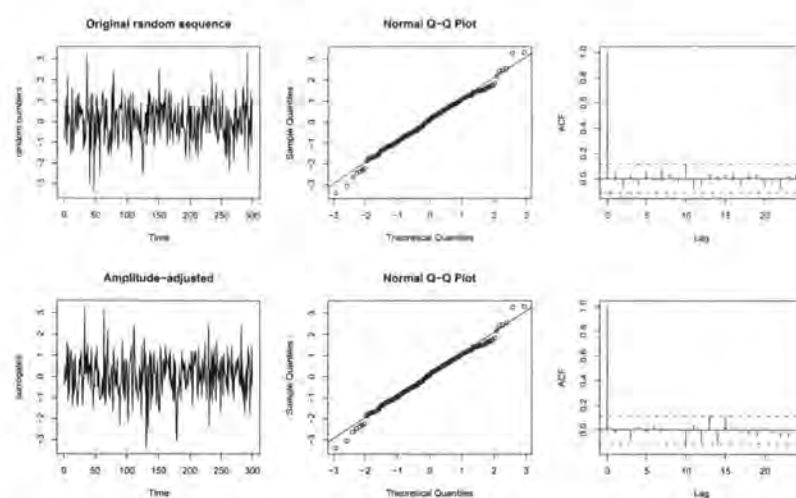
Figure 6.4: 300 normal random observations and their amplitude adjusted surrogates. Note the approximate equality of distributions and ACF functions supporting the null hypothesis.

In figure (6.2) we show the original Lorenz sequence with the amplitude adjusted surrogate sequence, histograms and ACF. In figure (6.3), we show the phase (FT) and phase-amplitude adjusted (AAFT) surrogates for comparison. These graphs were generated with the original SDA algorithms as described in the appendix of Theiler *et al.*, (1992) via the TSERIES version 0.9-0 package (Trapletti, 2002) provided in "R" statistical programming language by Ihaka and Gentleman (1996).[69] We also generated 100 normally distributed random numbers based on the algorithm AS 241 by Wichura (1988), also implemented in "R". The power of SDA is obvious from the inspection of figures (6.2 to 6.5). The Lorenz sequence, clearly chaotic, leads the procedure to reject the null whereas in the second case of the Gaussian random numbers, one can surely fail to reject it.[70] In table (6.1) we have included the first 20 coefficients of the ACF of a simulated AR(1) autoregressive process. Note the very small values of the difference of the ACF coefficients for the original sequences from that of the surrogate data and the very small standard error, indicating that

---

[68]We chose arbitrarily the σ, r and b positive parameters.

[69]Both library and language are available from http://www.r-project.org under general public license (GPL), free for scientific research (see also introduction).

[70]Note that these figures are generated by single realisations of the surrogate data sets for each algorithm. In practice, a number of those sequences must be generated in order to actually refute the null.
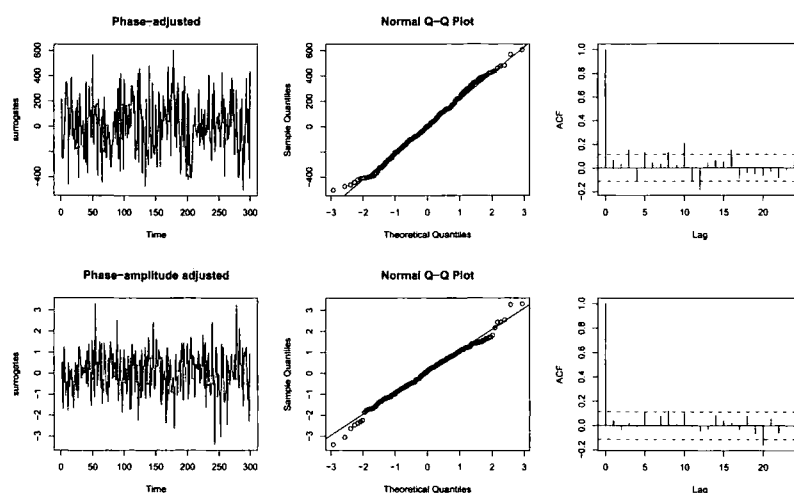
Figure 6.5: Phase and phase-amplitude adjusted surrogates of the random sequence in top figure (6.4). Again the same conclusion is reached supporting the null.

the null should be rejected.

In order to show how SDA can be used with a discriminating statistic, we generated the same length (3500) normally distributed random numbers. We used the Mutual Information criterion to test the three null hypotheses with simple random surrogates, phase randomised FT surrogates and AAFT surrogates. We test both the Lorenz and random sequences.

6.6.2 SDA with financial time series

Continuing our demonstration of the SDA framework, we apply the Mutual Information criterion test on the 9 stock market indices. The original closing prices time series are plotted in figure (6.8). We first conduct a test for i.i.d. vs. any dependence (H_1) and plot the results for the Mutual Information criterion in figure (6.9). We then conduct tests with hypothesis H_2 and H_3 with FT phase randomised and AAFF surrogates respectively in figures (6.10-6.11).

It is evident from figure (6.9) that H_1 should be clearly rejected at $\alpha = 5\%$ for all markets. Diagrams for Hong-Kong and Japan are less clear, as the confidence bounds are close to the statistic values. In fact, all tests for all hypotheses were rejected. The difficulty that generally arises in obtaining clearer or more powerful
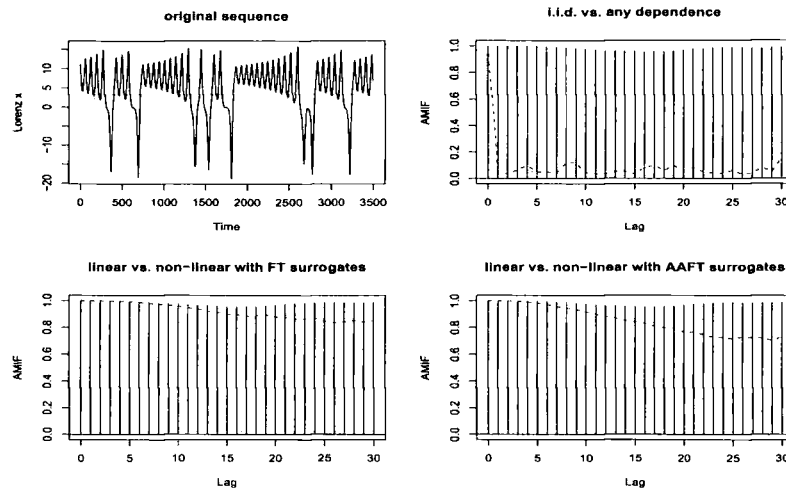
Figure 6.6: SDA with the Lorenz series. It is straightforward that the null is rejected in all three cases. Level of significance is $\alpha = 5\%$. 30 surrogate data samples were generated and the Mutual Information criterion used. Dashed lines indicate the confidence bounds.
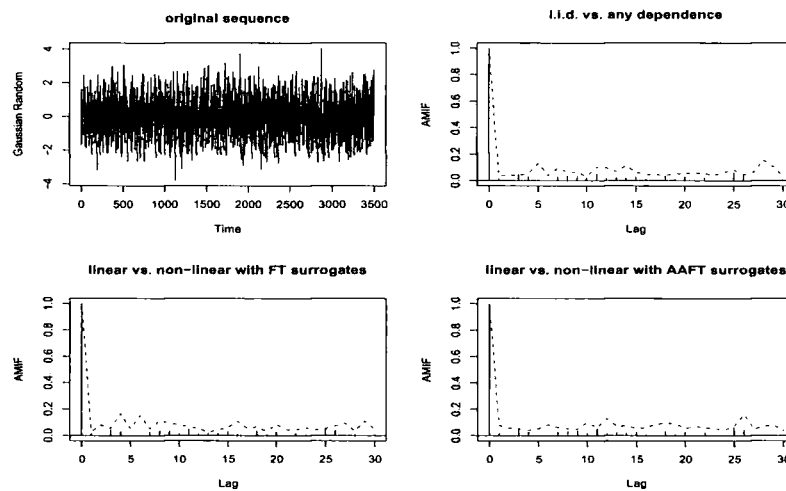
Figure 6.7: SDA with the normal random series. It is straightforward that the null is accepted in all three cases. Level of significance is $\alpha = 5\%$. 30 surrogate data samples were generated and the Mutual Information criterion used. Dashed lines indicate the confidence bounds.

Lag	Original	bias	std. error
τ_1	0.364570	-0.007051	0.04593
τ_2	0.078227	-0.003578	0.05120
τ_3	0.038625	0.002531	0.05483
τ_4	0.003403	-0.002702	0.05494
τ_5	0.032144	0.003304	0.06458
τ_6	-0.040358	0.012629	0.06560
τ_7	-0.051926	-0.001310	0.07939
τ_8	-0.030304	0.003288	0.08890
τ_9	-0.030866	0.018067	0.07058
τ_{10}	-0.051793	-0.002556	0.06319
τ_{11}	-0.103726	0.007442	0.07271
τ_{12}	-0.041744	0.005070	0.07315
τ_{13}	0.067786	0.019490	0.05620
τ_{14}	0.080881	0.016454	0.05907
τ_{15}	0.124059	-0.015057	0.05862
τ_{16}	0.096451	-0.018595	0.06360
τ_{17}	0.059631	-0.016266	0.05869
τ_{18}	-0.047576	-0.006706	0.05894
τ_{19}	-0.012457	0.005777	0.06999
τ_{20}	0.017235	0.011584	0.08179

Table 6.1: The first 20 lags of the autocorrelation function of a simulated AR(1) process and the bias of the surrogate data ACF statistic. 50 phase-amplitude adjusted FT surrogate sequences were generated.
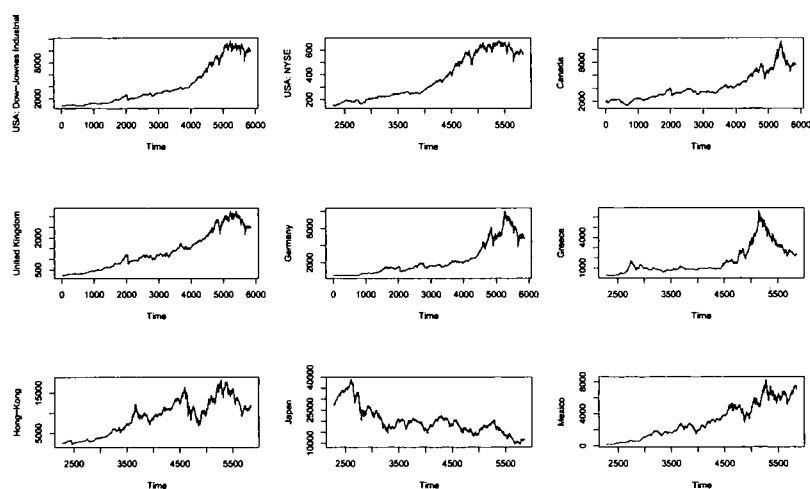
Figure 6.8: Daily closing prices for the 9 stock market indices. See text for further details.
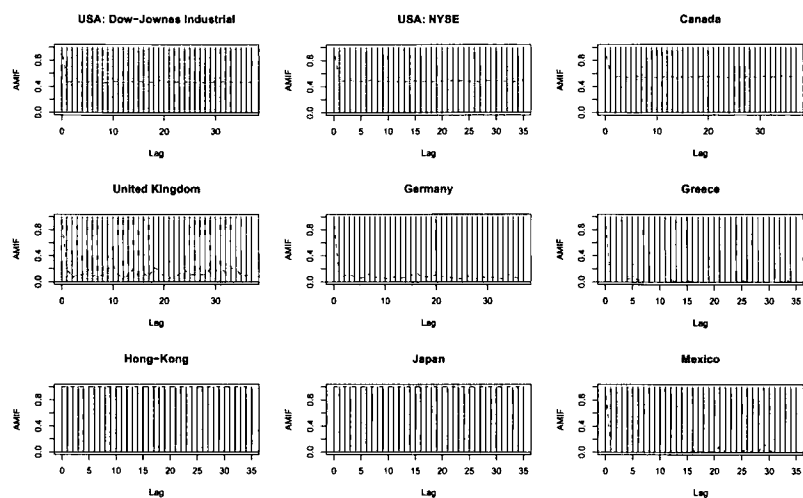
Figure 6.9: Testing for i.i.d. vs. any dependence for 9 markets. SDA was conducted with 50 surrogate samples, $\alpha = 5\%$ and the Mutual Information criterion. Simple shuffling was used here to generate surrogates consistent with H_1.
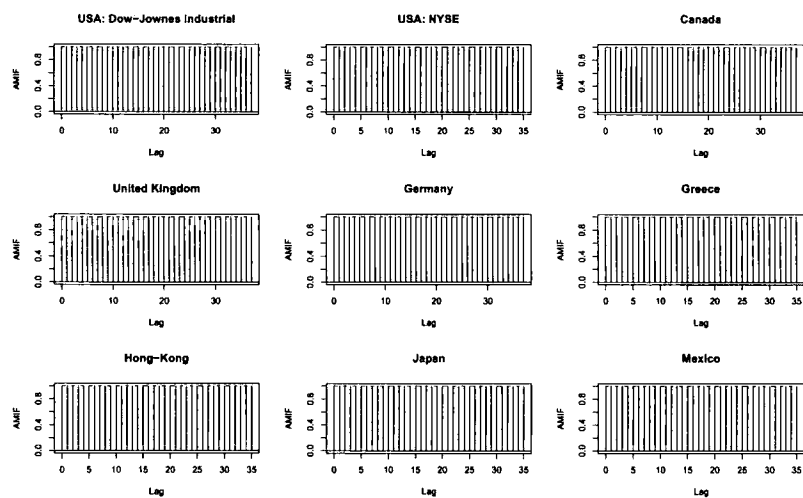
Figure 6.10: Testing for H_2 i.i.d. vs. any dependence for 9 markets. SDA was conducted with 50 surrogate samples, $\alpha = 5\%$ and the Mutual Information criterion. Phase randomised (FT) surrogates were used in this case.
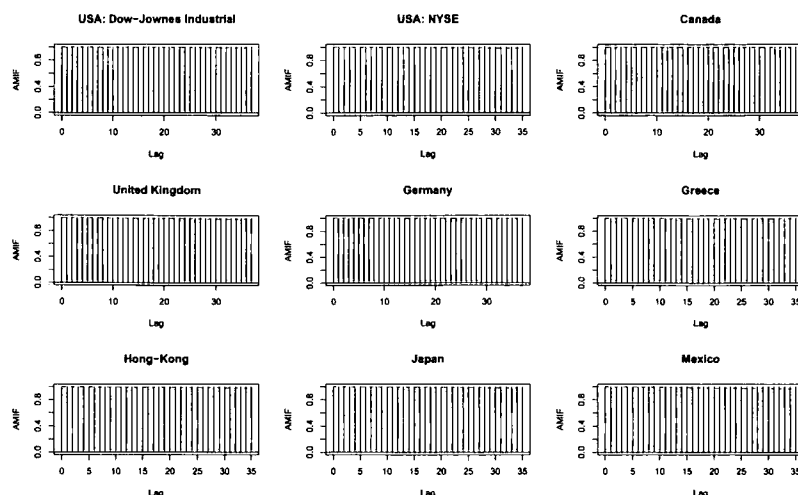
Figure 6.11: Testing the linearity hypothesis H_3 for the closing prices in figure (6.8). Phase randomised amplitude-adjusted (AAFT) surrogates were in this case (FT).

results for H_2 and H_3 testing is usually attributed to the nonstationarity of the levels data. For this reason we repeated the process for the returns sequences.

We applied the same SDA framework on the returns sequences generated by the original financial market indices for every country Y_t. We generated the continuously compounded returns as $\ln(P_t) - \ln(P_{t-1})$ where P_t is closing price at time t. Then we simulated 50 simple Gaussian shuffled, phase-randomised and phase-randomised amplitude-adjusted surrogates in order to test respectively for H_1, H_2 and H_3. The results are displayed in figures (6.12) to (6.14). It is evident that for the returns sequences, all three hypothesis for linearity should be rejected as the calculated statistics for the original series are exceeding the confidence boundaries.

6.6.2.1 Theiler's "crinkle" statistic

This test is based on the 4th moment and was developed by Theiler *et al.*, (1992) in order to demonstrate problems arising with surrogate data analysis. We use the statistic as in equation (6.10) on the 9 indices and the corresponding returns series to test hypothesis H_3.[71] The results are summarised in table (6.2). 50 phase randomised (FT) amplitude adjusted surrogates ("polished") where generated for

---

[71]This test assumes de-meaned or stationary data.
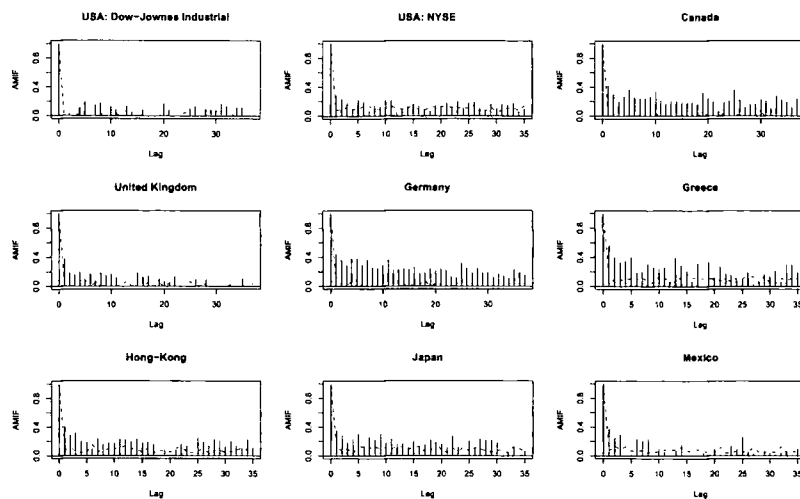
Figure 6.12: Testing for H_1 with SDA on the returns sequences.
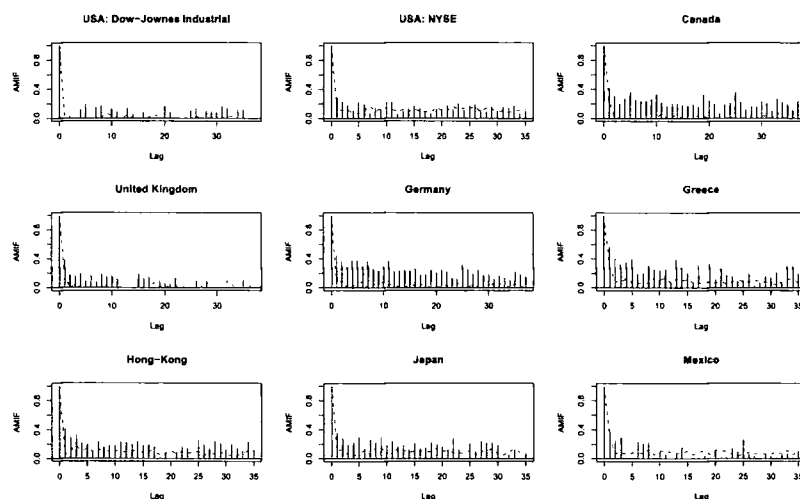
Figure 6.13: Testing for H_2 with SDA on the returns sequences.

each test. The test is one-sided (right) t-student based which means that for a 5% significance level, 19 surrogates would suffice. We can reject the null for all variables at $\alpha = 2\%$ except for US (NYSE at 6% and Dow-Jones at 10%) and Japan (at 4%). Therefore, evidence of nonlinearity is extremely strong in almost all of the returns sequences.
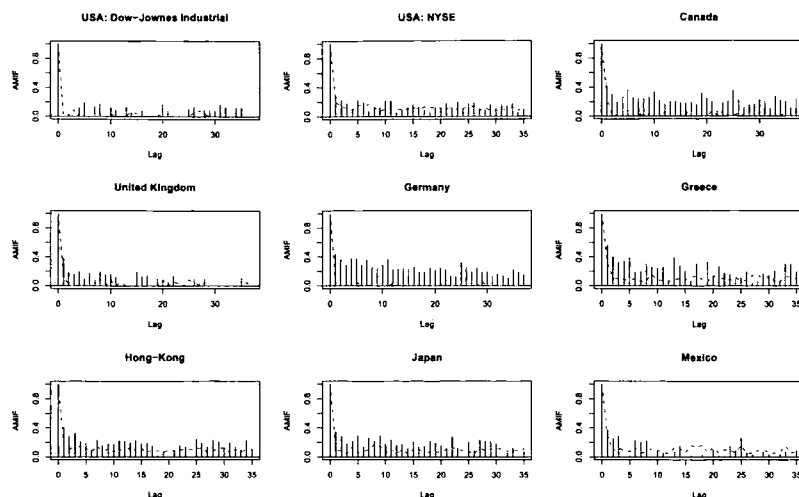
Figure 6.14: Testing for H_3 with SDA on the returns sequences.

Markets	t-statistic Returns	p-value Returns
Canada	7.310	0.02
Dow-Jones (US)	1.219	0.10
NYSE (US)	1.800	0.06
Germany	8.910	0.02
Greece	10.887	0.02
UK	5.105	0.02
Hong-Kong	5.133	0.02
Mexico	10.733	0.02
Japan	3.818	0.04

Table 6.2: Theiler's "Crinkle" statistic SDA results on returns sequences. 50 phase randomised (FT) amplitude adjusted surrogates ("polished") where generated for each test. The test is one (right) sided t-student based which means that for a 5% significance level, 19 surrogates would suffice. We can reject the null for all variables at $\alpha = 2\%$ except for US (NYSE at 6% and Dow-Jones at 10%) and Japan (at 4%). Therefore, evidence of nonlinearity is extremely strong in almost all of the returns sequences.

6.6.2.2 Time reversibility

Before we provide the results of our analysis, we should look into the time-asymmetry issue a bit deeper. Time reversibility is not a new concept. It is based on the asymmetry of data across time. Early attempts to detect this asymmetry were based mainly on observation of graphical displays. Namely the time series plots and delay maps (scatter diagrams of autocorrelations of various orders). Any kind of asymmetry in these plots was treated as an indication of time-irreversibility and therefore nonlinearity. For example, in Tong (1990) reverse data plots are suggested i.e., time series plots with the time axis reversed.

As we already discussed, a sequence X is time-reversible if any n-length sub-sample has the same joint probability distribution when ordered reversely in time as in its proper (original or natural) sequencing. This effectively means that a reversible time series should not look "different" when viewed in forward (original or natural) or reverse time. Any linear Gaussian random process or static transformation of it, is inherently time reversible. This property makes the time-reversibility statistic an essential tool in detecting nonlinearities and determinism.

It has been shown that time reversibility is essentially restricted to ARMA linear Gaussian processes (Weiss, 1975). Later research has indicated that, under certain regularity conditions for the existence of all-order moments, general linear processes are time-reversible (Hallin *et al.*, 1988). It was conceived that higher order cumulants (moments around the mean) may reflect temporal order. Second order was not enough to reveal irreversible time-structure. Recently Cheng (1999) has provided a sufficient condition which does not require moments of order higher than two. These findings essentially indicate that we can detect reversibility through comparative measures of distance such as the Euclidean norm or other kind of metrics.

Phillip Rothman's doctoral dissertation (1991) was dedicated to irreversibility in economic time series and more precisely business cycles. He introduced a time-reversibility statistic that was shown to posses good power when testing against bilinear and threshold autoregressive processes (TAR). He also showed that this test statistic, converged more quickly than the BDS and the Hinich (1982)

bispectrum tests.[72] In order to provide clarifications on some issues regarding business cycles, Ramsey and Rothman (1996) demonstrated the application of the "symmetric-bicovariance" function as a statistic for time-asymmetry (also known as the "Ramsey-Rothman" time reversibility (TR) test):

$$\hat{\gamma}_{2,1}(k) = \hat{B}_{2,1}(k) - \hat{B}_{1,2}(k) \tag{6.13}$$

where k is the order of the lag and $\hat{B}_{2,1}(k)$ and $\hat{B}_{1,2}(k)$ are moment estimators of the bicovariances $E[X_t^2 \cdot X_{t-k}]$ and $E[X_t \cdot X_{t-k}^2]$ respectively. The Fortran based algorithms for the test were provided and described thoroughly in Rothman (1997).

Smitchz and Shreiber (1998) use surrogate data analysis and a time-reversibility statistic to examine unevenly sampled time series. They propose the use of the following measure:

$$\gamma = \frac{1}{(\sigma^2)^{\frac{3}{2}}(N-1)} \sum_{n=2}^{N} (\frac{X_n - X_{n-1}}{t_n - t_{n-1}})^3 \tag{6.14}$$

where X_n and X_{n-1} is the sequence sample at times t_n and t_{n-1} respectively. This is a good indicator for nonlinearity but not very informative about what the source of this nonlinearity may be. If the DGP is a linear process, then $\gamma \approx 0$ for both the original sequence and the surrogates. Nonlinearities in the data will cause $\gamma \neq 0$ which implies that the statistical test here should be two tailed.

In table (6.3) we present the result of a test for H_3 as a null. We used 50 "polished" surrogate samples (which is more than enough for a 5% significance level). We can reject the null for the US market (Dow-Jones at $\alpha = 3\%$ and NYSE at $\alpha = 7\%$). We can also reject the null for the Greek market at $\alpha = 4\%$. The results are not that surprising for this version of the time-reversibility criterion. Diks 1999 reports the same results for his reversibility criterion (which does not involve SDA). Scmitz and Schreiber (2000) have concluded similarly for the same test. They used 1500 daily returns of the BUND Future but generated 19 surrogate samples based on the simulated annealing algorithm. They too could not refute the null. This does not constitute of course evidence of linearity, merely we can not reject the null

---

[72]See also the articles by Hinich (1982), Hinich and Patterson (1985, 1989 and 1993) and Hinich and Rothman (1998).

Markets	t-statistic Returns	p-value Returns
Canada	0.110	0.44
Dow-Jones (US)	2.934	0.03
NYSE (US)	2.075	0.07
Germany	-1.156	0.93
Greece	1.910	0.04
UK	0.414	0.31
Hong-Kong	0.227	0.38
Mexico	-0.830	0.81
Japan	-0.377	0.66

Table 6.3: **Time reversibility SDA results on the levels of the each index and the corresponding returns sequences. 50 phase randomised (FT) amplitude adjusted surrogates ("polished") where generated for each test. The results here do not support strongly the rejection of the null as in table (6.2). Only the US markets and the Greek can support the rejection of the null for a low significance level.**

of linearity as it is expressed by H_3 with these data. All of the above results while not supportive to the existence of nonlinearity, are consistent with the random walk hypothesis for the logarithms of the index level prices.

One explanation for our results may be that the time-asymmetry statistic should be modified to capture higher moments if this is important in describing the data. Secondly, these series contain heteroscedasticity which can be captured by an (G)ARCH model. If we assume that such a process is at hand here, the surrogate data should be modified in order to exhibit such a structure. This is something that has not been tried yet. GARCH processes are not monotonic transformations and in such a capacity, there is no surrogate data testing framework yet to consider such processes. Following Diks (1999) the question posed here is: can a GARCH DGP give rise to a reversible logarithmic return sequence? If the answer is yes, then reversibility as a null hypothesis would exclude GARCH processes as well. This is an interesting area for future research.

6.7 Conclusions

In this chapter we presented the general framework of SDA as a test procedure for weak determinism. We described the types of linearity hypothesis that can be used with the original observations, the surrogate data generation procedures and how one should interpret the results. SDA is a recent development and has seen very little application in our discipline. It has an advantage over the bootstrap in the sense that we do not have to assume a specific parametric model for the DGP. We merely need to simulate surrogate samples which exhibit the same spectral and distribution structure as the original sequence. By using some appropriate test criterion we can then compare its values when applied to the original observations and the surrogates. If the values differ substantially, we may refute the null of the existence of linearity in the DGP. That does not necessarily imply that our data are nonlinear deterministic. What we may safely assume is that there is no linear dependence among the data according to the rejected null. This can be evidence of nonlinearity. One should be careful not to accept such a result as a clear indication of chaoticity in the series. Further testing and the use of measures suitable for characterising chaotic dynamics may be required.

By applying this framework to financial data, we concluded that for our series, some statistics such as the "crinkle" and the Mutual Information seem to provide results consistent with the existence of nonlinear structure in the data. Time-reversibility statistics were less clear in the same data sets but our results were consistent with previous work in this area.

Summarising, we should carefully advocate the use of SDA as a supplementary testing procedure to those already applied in mainstream econometric analysis. SDA is quite different to bootstrapping and provides useful insight into the possible nature of the DGP. One should combine this with other techniques when trying to establish a parametric model for the data. Moreover, it could be used as a "second opinion" when testing the validity of parametric approaches. For instance one could examine regression residuals and test them for nonlinearity using Mutual Information or time-reversibility based criteria. SDA is not without flaws. Some problems may appear due to random number generation routines, or the surrogate data generation

procedures in general. Research is constant in the field and we should soon see SDA related procedures that may be even more useful or appropriate for the analysis of financial or macroeconomic data.

In the following chapter we deviate from our "chaotic" discussions in order to provide the basic theoretical background needed in wavelet analysis. Wavelets are a counterpart to Fourier analysis and can cope very well with nonlinear and volatile time sequences such as financial time series. They can also provide a very good tool for denoising these and can allow us to visualise their structure in various resolutions. In the next chapters we will demonstrate how wavelets can be used to provide qualitative and quantitative information on financial time series. We will also show how they can be used in conjunction with what was discussed so far to reveal very interesting facts about the structure of these series.

CHAPTER 7
Wavelet Theory

7.1 Introduction

As we have seen in chapters 5 and 6, the results of the nonlinear methodologies rely heavily on the quality of the data. The noisy and volatile nature of stock market series, creates the need of pre-filtering in order to clearly reveal interesting patterns. As we have discussed and is also stressed in the literature (Abarbanel, 1995) not all data preprocessing methodologies are appropriate when searching for determinism and nonlinearities. For this reason we explore a novel approach, based on *wavelet transforms*, so as to provide a new means of financial time series denoising. Our aim is to reveal any details that could be amplified with the use of nonlinear techniques already discussed and thus provide evidence for or against Market Efficiency and linear stochastic dynamics.

In the previous chapters we examined ways of determining whether a time series is characterised by deterministic or stochastic dynamics. We suggested a range of metrics that can be calculated, a number of invariant statistics that can reveal the amount of determinism that governs the data generating processes. It is important to determine in finance what is the structure of time sequences of observations as this has direct implications for the structure of the markets and the psychology of the participating agents. Especially, in a theoretical point of view, notions such as Market Efficiency can be refuted or accepted on the basis of the structure of these sequences. This has a profound impact on the whole framework of finance as most of the intertemporal equilibrium relationships and market clearing conditions are based on such notions. Moreover, risk estimation techniques are strongly dependent on the fact that financial time series are stochastic processes.

As we discussed so far it is not always possible to establish determinism in financial data. This is for various reasons. Until recently, large data sets were not available for markets and this has an effect on the quality of results we obtain while calculating chaotic invariant measures. Secondly, financial time series are extremely

volatile and noisy with very unclear dynamics. Any attempt to parameterise on these dynamics, will greatly affect the nature and shape of any existing attractor. This could result in determining that these series are stochastic when indeed they are deterministic or vice versa. One should carefully preprocess data with smooth transformations which do not affect the original dynamics and make sure at the same time that such transformations can be reversed so the original information can be retrieved with no loss. One of these transformations is the *wavelet transform*, which has been used so far with great success in the physical sciences, often analysing chaotic multidimensional dynamics.

In this chapter[73] we provide an introduction into wavelet theory in order to explain the mathematical framework which enables us to filter financial time series without any parametric assumptions on their underlying distributions. We will demonstrate the ability of wavelets to deal with finite and highly irregular samples in continuous and discrete time. We will also show the advantages of wavelets over Fourier analysis and how these can be used within a regression framework to provide smoothed sequences. In the following chapters we will be applying this theory to denoise UK stock market time series. Furthermore, we will be incorporating techniques already discussed in the previous chapters in order to establish whether the data are governed by stochastic randomness or chaotic determinism or both.

7.1.1 The basic framework

The limitations of the applicability of Fourier analysis on various experimental and empirical data has led scientists to develop the *hierarchical representations* of functions. The basic concept behind these methods (also called *multiresolution methods*) is to represent signals with (i.e., transform them to) a set of coefficients. Each of these coefficients provides limited information about the *frequency* and *the position* of the signal. The inverse transformation can then be used to retrieve the original data from the coefficients with no loss of information. Instead of carrying

---

[73]We focus on the purely theoretical framework of wavelet analysis, we will be citing scientific work that is relevant to our presentation. In the next chapter we deal with empirical applications of the theory and we will be citing and discussing there the relevant applied literature. For more details refer to a plethora of books on the subject such as: Chui (1992a, 1992b, 1997), Burrus *et al.*, (1997), or Strang and Nguyen (1996).

any analysis in real (original) time and space, one may choose to work in the functional space of the transformation which usually exhibits a wealth of useful properties.

Although the roots of wavelet theory can be found a hundred or more years back, it is only during the past decade that it has enjoyed such popularity and growth. Wavelets have become by now a necessary tool in many signal analysis related investigations. They are most suitable for the examination of nonstationary signals that exhibit sharp irregularities and brief but violent structural breaks. This is mainly due to the ability of wavelets to establish not only the frequency patters in a signal but also the time stamp of any non periodic shock or singularity. They are especially useful in the analysis of signals of various different nonstationarity classes, of non-homogeneous in time (or space) signals[74] and fractal or long memory time series.

Wavelets have many important characteristics which enable them to be a versatile tool in signal analysis. Although it is not our purpose to provide an extensive analytical framework of the wavelet theory in this thesis, we shall cover the most important aspects. Introductory it is important to mention that their most distinctive characteristic is their ability to be constructed through *dilations* i.e., changes in scale and *translations* i.e., changes in position. The scale actually refers to the width of the wavelet function utilised. Changes in scale result in changes in the wavelet width. It can also be defined as the distance between the oscillations of the wavelet. This feature allows them to be very useful for data compression and analysis of irregularly spaced or nonstationary series. Wavelet functions can also filter information very effectively even when the nature of the noise contaminating the data is unknown.

To demonstrate here a result of wavelet compression, we provide figure 7.1. The left image of Einstein is the original whereas the right image is a compressed version of the original using a Haar wavelet. It is obvious that there has been very little optical information lost due to compression. An image can be regarded as a two-dimensional signal. A financial time sequence is nothing else than one-dimensional signal as we already noted. In the same fashion one would apply wavelet analysis on

---

[74]Non-homogeneously sampled in time such as series with missing observations.
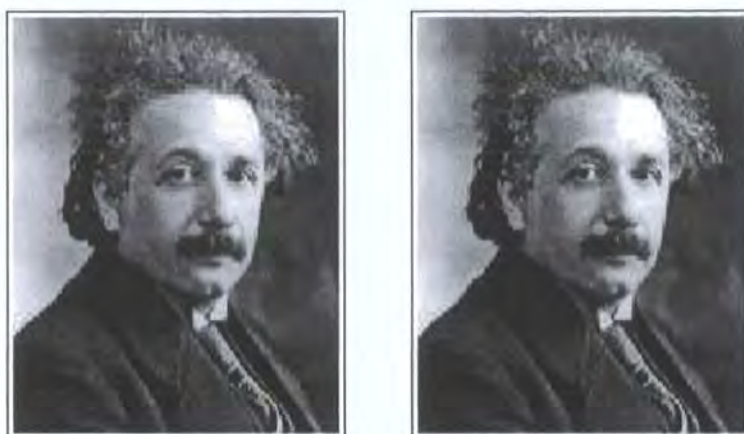
Figure 7.1: Wavelet compression of Einstein's photo. Left original scan-
ned image, right 30% compressed image. It is obvious from an
eye inspection that no visible loss of information has occurred.

a picture, one can decompose, compress and denoise a single one-dimensional time
series.

Another example of the power of wavelet filtering is in figure 7.2 featuring a
picture of Ingrid Daubechies.[75] Wavelets were used here to denoise a detail of the
image.[76]

So what exactly are wavelets and how do they work? Wavelets are orthogonal
functions that have *compact support* i.e., they exist only within a certain range and
not outside of it.[77] A crude but effective explanation of the functionality of wavelets
is to imagine them as functions that we can expand or contract and move along
the horizontal axis in order to "imitate" a specific event in the history of a signal.
Something like a mathematical magnifying glass that enables us to concentrate on
a very specific event in the history of a signal. The dilations enable wavelets to

---

[75]Ingrid Daubechies is a very active researcher in the field of wavelet analysis and author of
the classical wavelet book "Ten Lectures on Wavelets," (1992). She is the inventor of smooth
orthonormal wavelets of compact support.

[76]From a demonstration of the Wavelab (1995) wavelet library for MATLAB©available on the
WWW from: http://www-stat.stanford.edu/~wavelab/

[77]The *support* of a function refers to the region of the parameter domain over which the function
has a nonzero price. Functions that are supported over a bounded interval are said to have *compact
support*. *Box* functions are an example of compactly supported function.
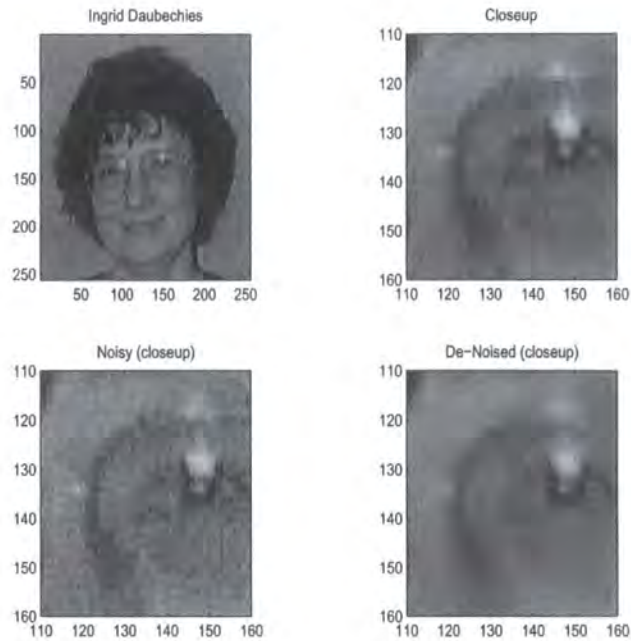
Figure 7.2: Demonstration of wavelet denoising of two-dimensional signals.

scale themselves to distinguish local characteristics of time series in various different scales and the translations allow them to cover the whole length of the series. Most of the wavelet families of functions form a complete orthonormal system with finite support (compactness) which allows them to exhibit such functionality and at the same time enable the inverse wavelet transformation to exist, similarly to the Fourier transform. Their substantial advantage over the Fourier transform is that the latter provides us only with the global frequencies of a specific event as it works on an infinite interval.

In spectral analysis, each Fourier coefficient contains complete information about the structure of the time series at one frequency but no information at the other frequencies. It should also be noted that Fourier analysis assumes an infinite and periodic signal whereas most empirical work is carried on finite and aperiodic series. Moreover, Fourier analysis uses sine, cosine and imaginary exponential functions whereas wavelet analysis has a much wider class of functions available. Discrete and continuous wavelet functions exist with different usefulness in signal analysis. Wavelet theory has also delivered us the *Wavelet packet transform* which
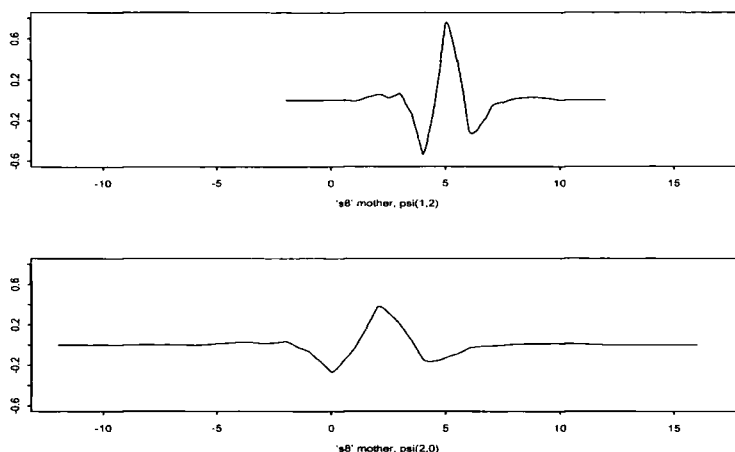
Figure 7.3: An example of how translations and dilations affect the position and shape of the s8 wavelet.

enables the use of *libraries* (sets) of wavelet functions to be applied on a single signal analysis at the same time. Among the basic properties of wavelets we should also include

1. *Linear-time complexity*: Wavelet transforms and inverse wavelet transforms can accomplished with very fast algorithms (faster then Fourier transforms) which work in linear time.

2. *Sparsity*: in practice, many of the coefficients of a wavelet decomposition are zero or negligible. This allows compression of information and acceleration of iterative procedures.

3. *Adaptability*: Wavelets are extremely flexible functions and can be adapted to represent a wide variety of other functions such as uni and multidimensional signals of about any kind of structure.

Wavelet functions provide what is known as a *time-scale* view of the signal. The signal is decomposed into simple functions as in spectral analysis. Instead of using sine and cosine functions, a *scaled* and *shifted* elementary function is used, known as the *mother wavelet* and usually represented as $\psi(t)$. The mother wavelet is scaled (or *dilated*) by a quantity a and time-shifted (or *translated*) by a quantity
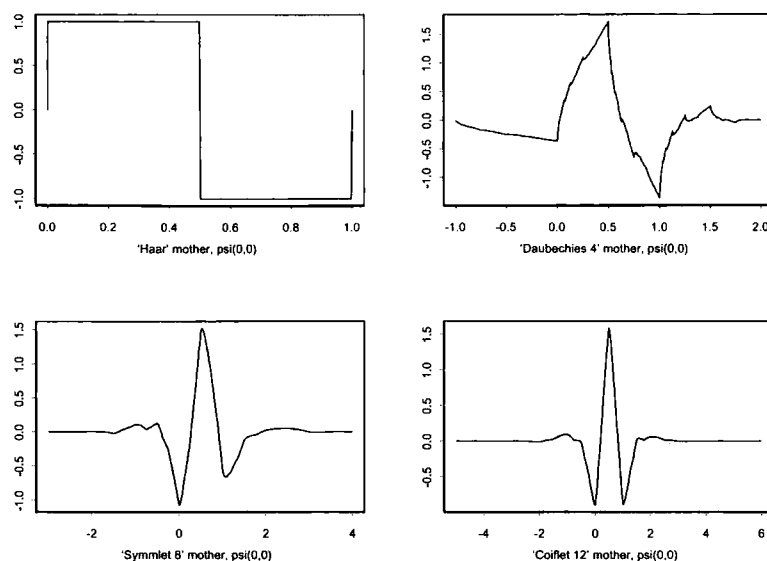
Figure 7.4: 4 types of mother wavelet functions from left to right, from top to bottom: The Haar, Daubechies 4, Symmlet 8 and Coiflet 12 wavelet functions.

b. These dilations and translations generate the wavelet basis function:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi(\frac{t-b}{a}), \qquad a,b \in \mathbb{R}, \qquad a > 0 \tag{7.1}$$

The representation that is obtained by decomposing a signal using 7.1 is called a *time-scale representation* also known as *wavelet representation*. The normalisation factor $1/\sqrt{a}$ allows the equality of the Euclidean norms (lengths) of the mother wavelet and the basis functions $\|\psi\| = \|\psi_{a,b}\|$.

Wavelet functions are localised or *compactly supported*. This means that they have finite length unlike the sinusoids used in spectral analysis which are of infinite length. They also exhibit irregularity, asymmetry and lack of smoothness which enables them to provide a better approximation of signals with irregularities.

The dilations and translations of the mother wavelet constitute the *Wavelet Transform* (WT) which is defined in continuous or discrete functional space. Reconstructing the original sequence from the decomposition by applying a reverse process (much alike as the inverse Fourier transform), requires the mother wavelet
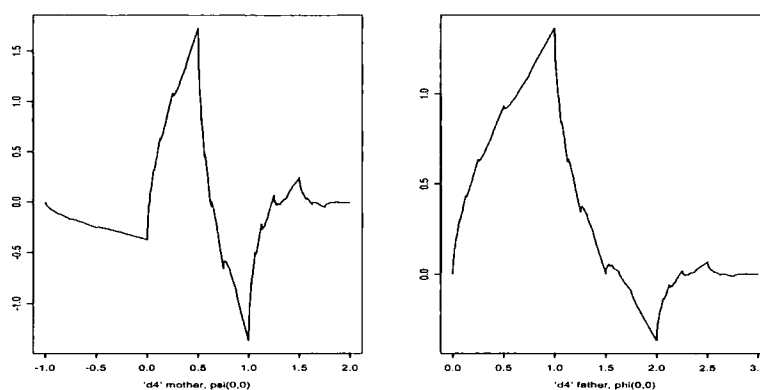
Figure 7.5: The Daubechies 4 mother wavelet and its scaling function (or father wavelet).

$\psi(t)$ to satisfy the following *admissibility condition*:

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty \tag{7.2}$$

where $\hat{\psi}(\omega)$ is the Fourier transform of the mother wavelet $\psi(t)$. By (7.2) is implied that $\int_{-\infty}^{\infty} \psi(t)dt = \hat{\psi}(\omega) = 0$ as C_ψ should be a finite quantity. This explains why mother wavelet basis functions are designed to have finite energy and are compacted (i.e., locally concentrated) as opposed to the Fourier basis functions. The constant C_ψ varies according to the choice of the wavelet. In the following sections we will provide a more rigourous treatment of the basic concepts of wavelet theory.

7.2 The Continuous and Discrete Wavelet Transforms

There are two approaches to time series analysis when using wavelets, the continuous and the discrete. In the continuous case, what is simply carried out is a sliding of a wavelet function across the signal in such a way that every wavelet overlaps the next to it. This has a result the generation of a lot of redundant information. A way round this is to use the discrete version of the wavelet transform which consists of sliding a wavelet function across the series but skipping observations such that there is no overlap between successive wavelet positions.

The continuous Wavelet Transform (CWT) is defined as the sum over all time

of the signal or function $f(t)$ times the scaled (dilated i.e., stretched or compressed) shifted (translated) versions of the mother wavelet function ψ:

$$W(scale, position) = \int_{-\infty}^{\infty} f(t)\psi(\text{scale, position}, t)dt \qquad (7.3)$$

More precisely, in the case of the analysis of a function of continuous variables, the CWT is defined as:

$$W_\psi(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t)\psi(\frac{t-b}{a})dt \qquad (7.4)$$

and the inverse transform is defined as:

$$f(t) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{a^2}[W_\psi(a, b)]\frac{1}{\sqrt{a}}\psi(\frac{t-b}{a})dadb \qquad (7.5)$$

where $\psi(t)$ is the mother wavelet and $a, b \in \mathbb{R}$ continuous variables with $a \neq 0$. The parameter a is called the "scale" or "dilation" parameter which determines the amount of stretching (expansion) or compression of the wavelet. The parameter "b" is the shift or "translation" parameter which dispositions the wavelet from its original position on the time axis along this time axis itself. When the scale is small ($|a| \ll 1$), the wavelet function is highly concentrated (shrunken-compressed) with frequency contents mostly in higher frequency bands. In the opposite case ($|a| \gg 1$), the wavelet is stretched and contains mostly low frequencies. For small scales we obtain thus a more detailed view of the signal (or "higher resolution") whereas for a larger scale we obtain a more global view of the signal. This implies that the wavelets are the mathematical analogue of a "magnifying glass", enabling us to "zoom-in" on very short-lived phenomena, i.e., singularities such as sharp shocks-breaks in time series with very short duration. Changing the b translation parameter, we can localise in time i.e., "move" the wavelets across the history of the signal. The CWT operates in all possible scales and ranges of the signal in a continuous manner. The resulting wavelet coefficients are functions of these two parameters. Thus, the initial function-signal $f(t)$ is projected onto a particular wavelet and the retrieved transformation belongs now not to the original functional (time) space but to a time-

frequency domain. By repeating this process we break down the original signal into its components. This is called "Wavelet Decomposition". The CWT function (7.4) ensures that when the wavelet resembles the signal, the retrieved wavelet coefficients will be large. The inverse transformation (7.5) ensures a perfect reconstruction of the original signal without any loss of information.

From our discussion above, it is obvious that the CWT provides us with a large amount of extra information on top of that of the original signal. To reduce this we subsample the information of the CWT at a discrete smaller number of resolution levels, by selecting a specific finite set of wavelet coefficients. We *critically* sample s and u translations and dilations respectively:

$$s = 2^{-j} \qquad (7.6)$$

and

$$u = k2^{-j} \qquad (7.7)$$

obtaining the Discrete Wavelet Transform (DWT) resolution in time and frequency. The term "critical sampling" implies that we select the minimum number of coefficients necessary to ensure that no loss of information occurs with respect to the original signal or function. In figure 7.6 one can see the differences between DWT and CWT for the step function in figure 7.11.

7.3 Definitions

Before we start our discussion on wavelet theory,[78] it is essential that we review necessary terminology from functional analysis to provide the basic jargon used in the discussions that will follow.

A *vector space* is a set i.e., a collection of *"things"* for which scalar addition and multiplication are defined. A *function space* is a *linear* vector space of finite or infinite dimensions where functions can be defined as its vectors, and scalar

---

[78]Although we are showing a number of wavelet applications in this thesis, It is not our intention to provide a comprehensive or rigourous treatment of the wavelet theory and its background. Bibliography is already vast. An initial starting point for introductory, intermediate or advanced concepts should be the **Lucent Technologies** website at: *http://www.wavelets.org.*
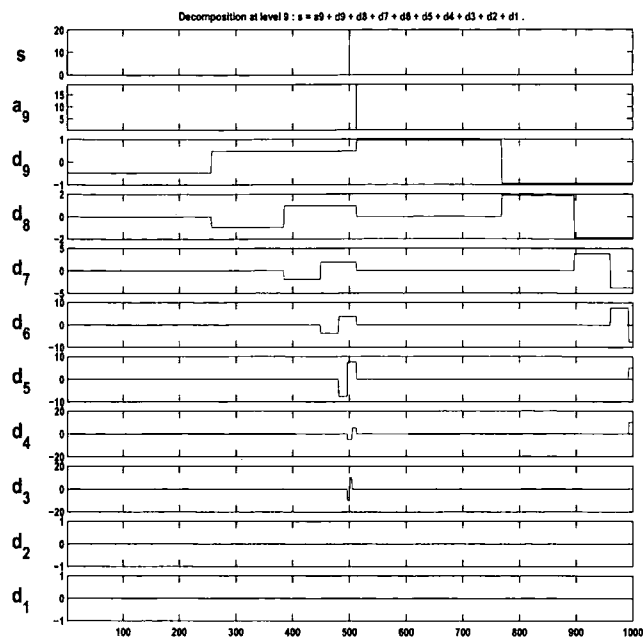
Figure 7.6: DWT of the step function in figure 7.11 with a level 9 Haar wavelet.

multiplication and addition are defined. The *inner product* of two vectors $f(t)$ and $g(t)$ is a scalar a and is obtained by the integral:

$$a = \langle f(t), g(t) \rangle = \int f(t)g(t)dt \qquad (7.8)$$

The inner product on any vector space V is a mapping from $V \times V$ to $\mathbb{R}$ that is symmetric, bilinear and positive definite. The *norm* or length of a vector f is defined by

$$\|f\| = \sqrt{|\langle f, f \rangle|} \qquad (7.9)$$

The definition (7.9) refers to the 2-norm is defined in finite-dimensional vector space but this can also be expanded to n-norms where $n > 2$. Two vectors with non-zero norms are called *orthogonal* if their inner product is zero. The space of all functions $f(t)$ with a well defined integral of the square of their modulus is called $L^2(\mathbb{R})$. The "L" stands for "Lebesque" integral and the "2" denotes the square of the modulus of the function. The "$(\mathbb{R})$" signifies that the independent variable of integration $t\epsilon\mathbb{R}$ (i.e., t is a real number).

A *basis* for a vector space consists of the minimum set of vectors from which all other vectors in that vector space can be generated through linear combinations. For a vector space V, a collection of vectors $u_1, u_2, ...\epsilon V$ are said to be *linearly independent* if

$$c_1 u_1 + c_2 u_2 + ... = 0 \text{ iff } c_1 = c_2 = ... = 0 \tag{7.10}$$

A collection of linearly independent vectors $u_1, u_2, ...\epsilon V$ forms a *basis* for V if every vector $v \epsilon V$ can be written as a linear combination of the u_i vectors:

$$v = \sum_i c_i u_i \tag{7.11}$$

for real numbers $c_i \epsilon \mathbb{R}$. The vectors in a basis for V are said to *span* V. When a basis is formed by a finite number of elements $u_1, ..., u_m$, then V is *finite-dimensional* with dimension m. Otherwise V is said to be *infinite-dimensional*.

Assuming that we represent two real-valued time series $\mathbf{X}$ and $\mathbf{Y}$ as column vectors:

$$\mathbf{X} = \left[\begin{array}{cccc} X_0, & X_1, & ..., & X_{N-1} \end{array} \right]^T$$

and

$$\mathbf{Y} = \left[\begin{array}{cccc} Y_0, & Y_1, & ..., & Y_{N-1} \end{array} \right]^T$$

where 'T' denotes transpose, then the inner product is:

$$\langle \mathbf{X}, \mathbf{Y} \rangle \equiv \mathbf{X}^T \mathbf{Y} = \sum_{t=0}^{N-1} X_t Y_t \tag{7.12}$$

and the squared norm of any of the two vectors, say $\mathbf{X}$:

$$\|\mathbf{X}\|^2 \equiv \langle \mathbf{X}, \mathbf{X} \rangle = \sum_{t=0}^{N-1} X_t^2 \equiv \mathcal{E}_\mathbf{X}; \tag{7.13}$$

defines the quantity $\mathcal{E}_\mathbf{X}$ which is called 'energy' of the $\mathbf{X}$ vector.

A very important definition is that of "orthonormality". An $N \times N$ real-valued

matrix $\mathcal{A}$ called "*orthonormal*" if

$$\mathcal{A}^T \mathcal{A} = I_N \equiv \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

i.e., when the transposed matrix is also the inverse.

7.3.1 A brief history of wavelets

Wavelets can be regarded as a relatively new approach to signal analysis. During the last decade they have become enormously popular due to their versatility and economy. Applications range from fingerprint recognition, image or voice analysis and compression, military applications and medicine to finance. The roots of the theory lie a century ago in the work of Weierstrass (1895) who described a family of functions that bear his name. These functions were constructed by superimposing scale copies of a given base function and were what today we term as "fractals" i.e., everywhere continuous but non-differentiable. Fifteen years later Alfred Haar (1910) constructed the first orthonormal system of *compactly supported* functions[79] which today we call the *Haar basis* functions. The "compact support" property of Haar's wavelets ensured that these functions would exist only within a finite interval and vanish outside it. This development was to form the foundation of modern wavelet theory. The only drawback was that Haar's wavelets were not continuously differentiable which limited their applicability. Some decades later Dennis Gabor (1946) described a nonorthogonal basis of wavelets based on Gaussian functionals. He thus described the "short term" Fourier transform which today is also known as "Gabor" or "windowed" transform. Until wavelet theory was developed, Gabor analysis was used as a modified spectral analysis which was able to cope better with localised events.

The term *wavelet* was introduced in seismology. Ricker (1940) used it to describe disturbances that follow sharp seismic shocks. Morlet *et al.*, (1982) demon-

---

[79] Actually the first mention of wavelets appears in an appendix of Alfred Haar's thesis

strated that these small waves could be modelled with Gabor's functions.

7.4 Wavelets vs Spectral analysis

The classic tool for frequency analysis of time series has been Fourier analysis also known as spectral analysis. This technique focuses on the frequency localisation of events. Fourier analysis enables us to determine the frequency of occurrence of a specific event that affects the prices of the time series. It answers to "how often" but not to "when". The main difficulty being that each Fourier coefficient contains complete information on the behaviour of the series at a single frequency but not information on the rest of the frequencies. Secondly, one basic requirement of Fourier analysis is that the examined series are of infinite length and more or less periodic functions. In real life however, data and especially financial time series are almost always aperiodic and of finite length. Wavelets were designed as an answer to the need to localise also in time. They enable us to determine not only the frequency of an event but also its time stamp. In order to explain this ability further, we generated two figures. Figure 7.7 shows the Fourier coefficients of a signal of 1000 pseudo-random numbers which are normally distributed. As we can see all frequencies contribute the same to the dynamics of the sequence. This is exactly the "flat spectrum" as discussed by Granger (1968). We can thus determine the randomness of the data from this figure. Figure 7.8 shows the scalogram of the same sequence of random numbers. Although we have not discussed yet the scalograms in detail, we see that we can obtain information on a number of scales (here limited to 40) as well as in time. We can thus see the values of the coefficients in lower scales (higher frequencies) and in larger scales (lower frequencies). This is not possible in figure 7.7 as we only obtain the magnitude of coefficients by frequency. In the following section we briefly describe how Fourier transforms work in order to contrast them to the wavelet transforms and see clearly their advantages and limitations.
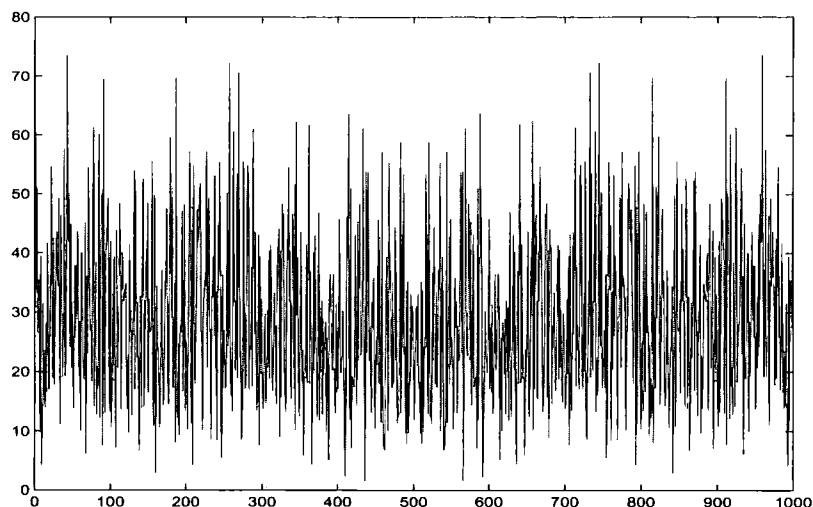
Figure 7.7: Fourier coefficients from spectral decomposition of a signal of 1000 random normal observations.
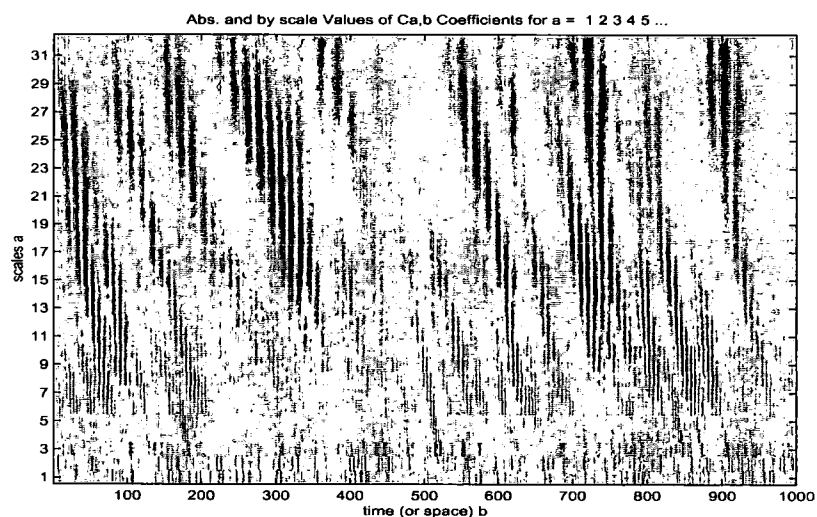
Figure 7.8: Wavelet coefficients of the signal in Figure 7.7.

7.4.1 Fourier Series and transforms

Joseph Fourier in 1807, the father of frequency analysis, asserted that any 2π-periodic function can be explained as the sum of sines and cosines. That is known today as Fourier or spectral analysis. To explain the requirements for Fourier analysis, we consider the set of square integrable functions $L^2(0, 2\pi)$ or the space of

all squared 2π periodic integrable functions which have a finite energy form:

$$\int_0^{2\pi} |f(t)|^2 dt < \infty, \qquad t \in (0, 2\pi) \tag{7.14}$$

This is how we define any piecewise continuous function $f(t)$ which can be defined on the entire real number axis $\mathbb{R}(-\infty, \infty)$ as $f(t) = f(t - 2\pi)$ for $t \in \mathbb{R}$. Any function belonging to the $L^2(0, 2\pi)$ functional space can be expressed in a series of sinusoidal functions (sines and cosines) i.e., expanded into a Fourier series:

$$f(t) = \sum_{-\infty}^{\infty} c_n e^{int} \tag{7.15}$$

where $i = \sqrt{-1}$ imaginary number. The Fourier coefficients c_n are then defined as:

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-int} dt \tag{7.16}$$

and (7.15) converges uniformly to $f(t)$:

$$\lim_{M,N \to \infty} \int_0^{2\pi} \left| f(t) - \sum_{-M}^{N} c_n e^{int} \right|^2 dt = 0 \tag{7.17}$$

An important point here is Parseval's theorem which states for Fourier analysis that the Fourier series are constrained by the following identity:

$$\frac{1}{2\pi} \int_0^{2\pi} |f(t)|^2 dt = \sum_{-\infty}^{\infty} |c_n|^2 \tag{7.18}$$

In Fourier analysis, we can thus define an orthonormal basis in $L^2(0, 2\pi)$ by dilating a single exponential function e^{it} so that $e^{it} = e^{int}$. Any 2π-periodic function that is square integrable can be represented by superimposing scaled transformations of the basis function e^{it}. We should recall from basic trigonometry that:

$$e^{it} = \cos(t) + i\sin(t) \tag{7.19}$$

This effectively implies that we can deconstruct-reconstruct signals by using sinusoidal wave functions of different frequencies. We can use Fourier transforms to analyse a signal in the time domain in order to examine its frequency content. Using Fourier theory we can transfer any function from the time-domain to the frequency-domain (i.e., from the original functional space to Fourier functional space). Then we can analyse the signal for its frequency content as the Fourier coefficients (7.16) of the transformation represent the contribution of each sine and cosine functional at each frequency (i.e., the contribution of each frequency to the energy of the signal). The original signal can always be retrieved by an inverse Fourier transform which brings back the signal from Fourier functional space to the original time-domain. Much alike wavelet transforms, Fourier transforms are defined in a discrete or continuous space. The discrete Fourier transform (DFT) estimates the Fourier transform of a function or signal for a finite (discrete) number of sampled points which provide a typical representation of the original continuous function or signal.

Fourier analysis assumes that the signal of function $f(t)$ under analysis is periodic and infinite. If $f(t)$ is aperiodic (such as most financial or economic time series), the superimposition of sine and cosine functions provides a poor representation. An improvement on the original Fourier idea was the windowed Fourier transform (WFT, or short term Fourier transform (STFT) known also as Gabor analysis). The WFT basis provides a representation in time as well as in frequency. With WFT, $f(t)$ is segmented into sections according to a window which defines the size of these sections (hence the term "windowed") which are analysed individually.

For uniformly sampled in time signals, in order to approximate using DFT the Fourier integral (7.16) to compute the Fourier coefficients, we use the Fast Fourier Transformation (FFT). The FFT is a technique that has been applied widely to the analysis in the frequency domain of many time series. The FFT requires $n \log n$ arithmetic operations which are as we will see more than the Discrete Wavelet Transform (DWT) but the same as the Maximum Overlap Wavelet Transform described later in this chapter. Both FFT and DWT require series of 2^n length (where n integer). The algoritmic complexity though of the DWT is much simpler than that of the FFT. Due to the orthonormality of the basis, both inverse FFT and DWT

Standard Basis

Gabor Basis

Fourier Basis

Wavelet Basis
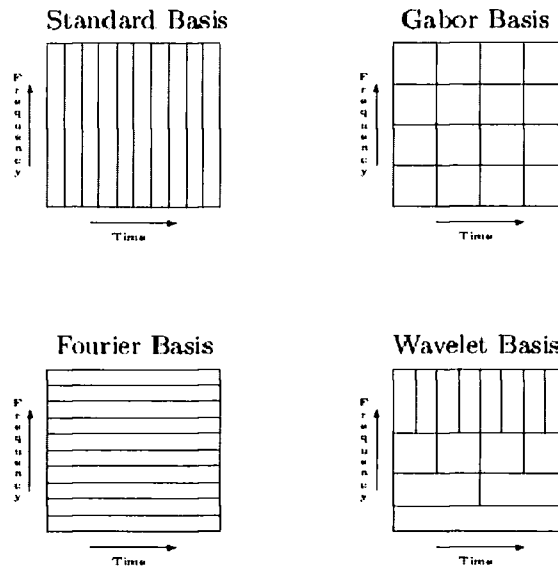
Figure 7.9: The effect of different bases in time-frequency analysis

are the inverse transform matrix of the original transformation which implies that both DWT and FFT can be conceptualised as rotation of the functional space to a different domain. Whereas the basis functions differ between FFT and DWT, both are localised in frequency. In Fourier analysis we obtain the *power spectrum* which shows how much power or energy is contributed by every frequency. The equivalent of a power spectrum in wavelet theory is the *wavelet scalogram* which shows how this energy is contributed in time but also in scale. An example of a wavelet scalogram for a series of 1000 normally distributed random numbers is provided in figure 7.8. Another scalogram, which shows how wavelets capture the periodic nature of a sine wave contaminated with white noise, is shown in figure 7.10. The power of wavelet transforms over Fourier transforms to localise in time, is clearly demonstrated in figure 7.13 which contains the scalogram of the step function in figure 7.11. Notice that the break in the level of the function is clearly identified by the scalogram (figure 7.13) whereas such an information is not obtained by the inspection of the Fourier coefficients of the same function given in figure 7.12.

Scalograms provide us with a visualisation of the time frequency localisation of the wavelet transform. Each detail coefficient is plotted as a rectangle (cell) filled with a colour which represents the magnitude of the coefficients. The exact
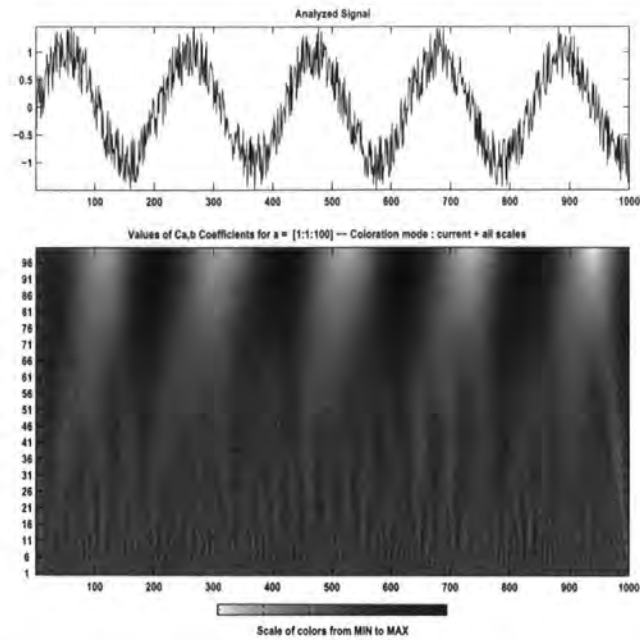
Figure 7.10: Noisy sine wave scalogram. The periodic character of the underlying signal is evident from the scalogram's structure as well as the contribution to the signal's frequency of the smaller and larger scales.
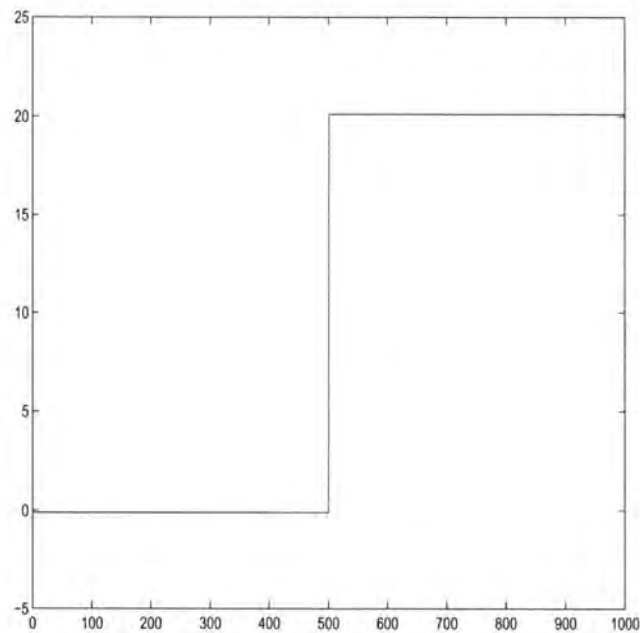
Figure 7.11: A step function with a break at observation 500.
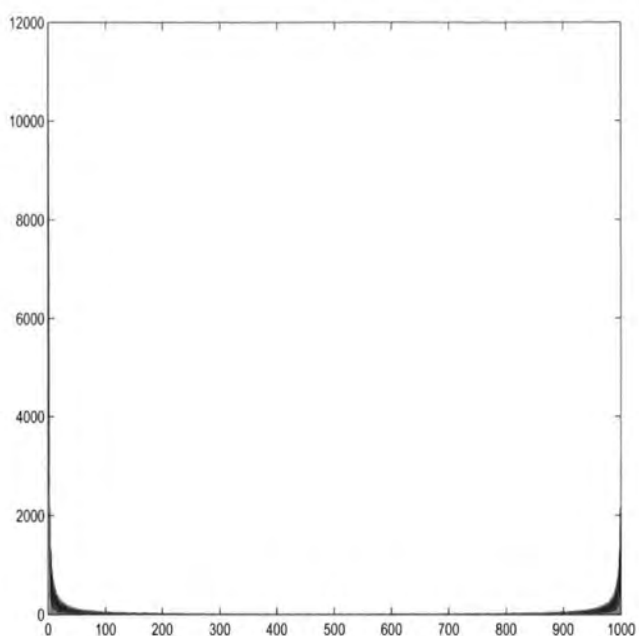
Figure 7.12: The Fourier coefficients of the step function in figure 7.11.
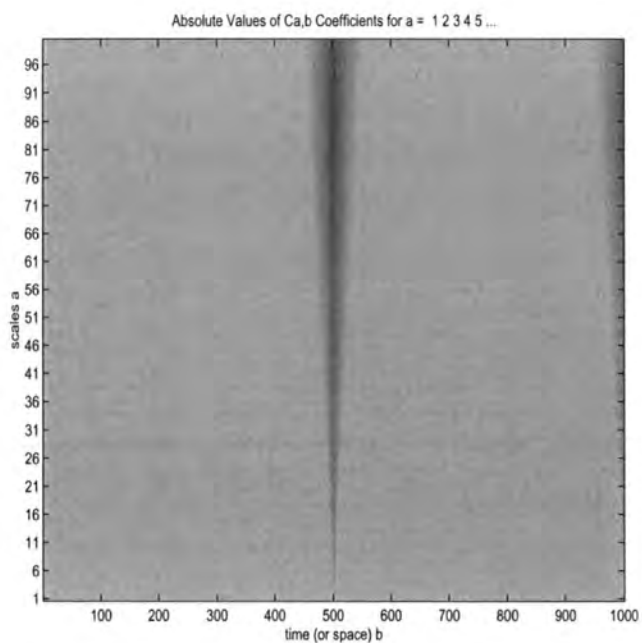
Figure 7.13: The scalogram (using a Haar wavelet transform) of the step function in figure 7.11. The break at 500 is very clearly indicated by the wavelet coefficients.

location and size of each cell refers to the time interval and the frequency range of the coefficient. After careful inspection of a scalogram, it becomes apparent that low level coefficients are wider and shorter, implying a wide localisation in time but narrower localisation in frequency. The higher the level of a coefficient the thinner and taller the cell as we have a better localisation in frequencies (allowing for larger frequency ranges) but this localisation is provided in a smaller time range than the previous levels. The heights of the cells grow at a power of 2 as the levels increase.

An interesting comparison between the FFT periodogram and the wavelet scalogram is displayed in figures (7.14) and (7.15). Although one can easily see the relation between the white noise process and the Brownian motion[80] in the spectrograms, the visual inspection of the respective periodograms does not allow us to conclude the same. We can clearly see that the Brownian motion is nonstationary in the periodogram of figure (7.15) with the low frequencies indicating a drift or a trend. In the case of the scalograms in figure (7.14) we can also see that the high frequency structure of the two series is almost identical which is something the periodogram does not reveal.

Apart from the different nature and usefulness of spectrums and scalograms, there are other differences between wavelet and Fourier transforms which make the former more elegant for time series analysis of signals of a specific nature. Wavelet functions are localised not only in frequency but also in time. Fourier functionals provide only frequency localisation. In order to demonstrate this dissimilarity, figure 7.9 is provided where it is shown that wavelet basis varies allowing for a multitude of different windows which capture lower, medium and higher frequencies more effectively than other methods. In this way, sharp discontinuities, shifts or breaks can be captured by short length basis whereas long basis functions provide the time identity of these events. Wavelet transforms are characterised by an infinity of basis functions which are translations and dilations of the initial "mother" wavelet. This is exactly what enables one to capture "structure" more efficiently choosing wavelet transforms over any other transform. What we actually achieve through wavelet functions is a series of "correlations" of these functions with every single part of the

---

[80]Generated as a cumulative sum of the white noise.
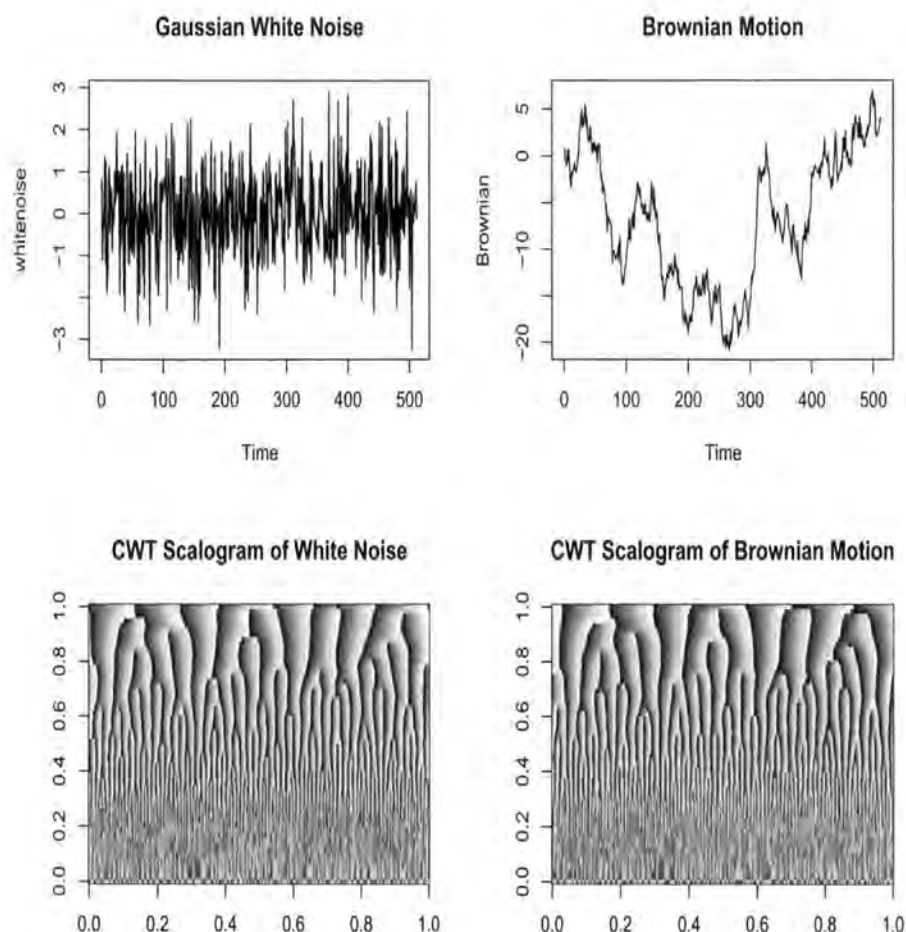
Figure 7.14: The scalogram of a White noise process compared to that of the related Brownian motion. Although the low frequency wavelet coefficients seem to be different as the Brownian motion is exhibiting trends, it is easy to see at the top of the scalograms that they both have almost identical high frequency content.

time series. By shrinking or expanding the wavelet and rolling it on the series we can obtain high correlations where the wavelet resembles the structure of the series and low where this structure is different. Moreover, these correlations are obtained over all possible scales. This is clearly depicted in the scalogram which shows areas of high and low "correlations" i.e., areas where the series dynamics at a certain scale can be approximated by the wavelet function.

A careful selection of the mother wavelet is necessary in order to obtain accurate results. As we shall see, one chooses wavelets that "look" like the series investi-

Gaussian White noise periodogram

Brownian motion periodogram
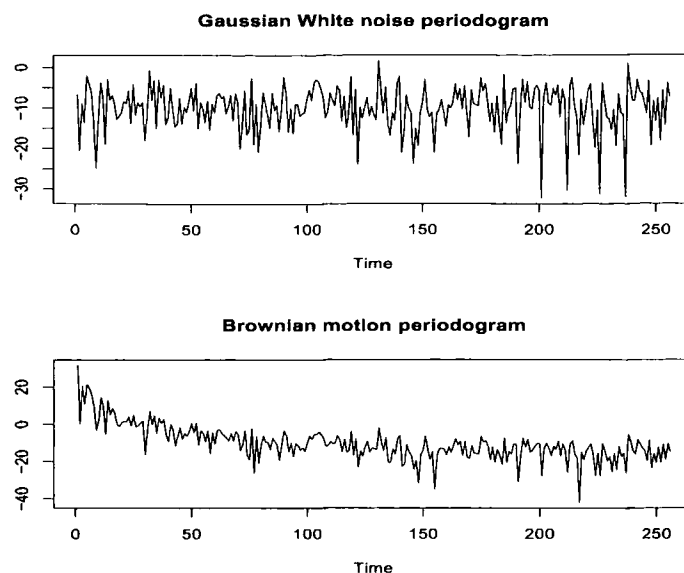
Figure 7.15: The periodograms of the White noise and Brownian seque-
nces in figure (7.14.

gated. Smooth wavelets can capture smoother cycles while highly irregular wavelets
can detect localised sharp shocks in the sequences. In the next chapter we will be
showing how continuous wavelet transforms can provide us with very informative
scalograms. We will be using these scalograms not only to identify specific historic
events in the history of the data but also self similar structures in various scales that
could imply the presence of deterministic dynamics.

7.5 Criteria and Properties of Wavelet Functions

Wavelets exhibit a number of useful properties as discussed below. As already
discussed, selecting a particular wavelet function over another implies a tradeoff
between these different properties.

1. *Linearity*

   A wavelet transform $W[f(t)]$ of any function $f(t)$ for any constant a satisfies
   the linearity condition:

   $$W[af(t)] = aW[f(t)] \tag{7.20}$$

And this can be generalised for more than one functions:

$$W[af(t) + bg(t)] = aW[f(t)] + bW[f(t)] \tag{7.21}$$

2. *Smoothness.*

   Although some particular wavelets are not smooth functions (such as the Haar) they do generally exhibit smoothness which enables them to efficiently capture the characteristics of signal structures. Smoothness is measured by the number of derivatives that exist for each wavelet. For example, the discontinuity of the Haar wavelet function implies that it is not differentiable. There are also some wavelets that although continuous, are not differentiable everywhere.

3. *Invariance*

   Invariance to translations is what allows the commutativity of differentiation of any wavelet transform:

   $$W[f(t - b_0)] = W(a, b - b_0) \tag{7.22}$$

   Invariance to dilations allows to determine whether the function f exhibits singularities:

   $$W[f(\frac{t}{a_0})] = \frac{1}{a_0}W(\frac{a}{a_0}, \frac{b}{b_0}) \tag{7.23}$$

4. *Temporal or Spatial Localisation.*

   As already mentioned, wavelets can localise in time-space and as well as in frequency enabling the detection of both the time and frequency signatures of singularities within sequences. The compact support is one reason for this characteristic. We obtain very compact wavelets that can produce very good localisation in time-time. This property is in general inversely related to the previous one. Smoother wavelets enable wider support widths and improved localisation.

5. *Frequency Localisation*

Again, in this case as in the time-space localisation property, smoothness is a determining factor. Smoother wavelets allow better frequency localisation.

6. *Zero mean and Vanishing Moments*

The zero mean of wavelets is implied by

$$\int_{-\infty}^{\infty} \psi(t)dt = 0 \tag{7.24}$$

In most applications it is usually necessary to have in addition the first M moments zero. Higher vanishing moments enable wavelets to improve their approximation of higher degree polynomial structures. Smoothness is also closely related to this property. The following condition is satisfied by every mother wavelet ψ with M vanishing moments:

$$\int t^{m}\psi(t)dt = 0 \qquad m = 0, 1, ..., M - 1 \tag{7.25}$$

Such wavelets are referred to as mth-order wavelets. The conclusion here is that wavelets possessing many zero moments enable one to ignore most regular polynomial components of signals and actually capture more efficiently small-scale fluctuations or higher-order features.

7. *Symmetry*

Symmetry follows form the definition that the ϕ and ψ functions are symmetric. This property also allows wavelet coefficients to be stable and not to "drift" relative to the signal analysed avoiding thus *phase-shifts*. The family of orthogonal wavelet functions which have compact support (with the exception of the Haar family) are not symmetric. Biorthogonal wavelets can be either symmetric or anti-symmetric.

8. *Orthogonality*

Although non-orthogonal wavelet functions (e.g. the biorthogonal) have been constructed, the orthogonality principle is a central characteristic for most of the wavelets applied in practice.

9. *Compactness*

Wavelet functions are defined within a closed numerical set and their value is zero for any number outside that set.

10. *Boundedness*

Boundedness is shown by

$$\int |\psi(t)|^2 dt < \infty \tag{7.26}$$

and can be written also as

$$|\psi(t)| < (1 + |t|^n)^{-1} \quad \text{or} \quad |\hat{\psi}(\omega)| < (1 + |k - \omega_0|^n)^{-1} \tag{7.27}$$

where ω_0 is the dominant wavelet frequency and n is an as large as possible integer.

11. *Basis self-similarity*

All wavelets of a given family $\psi_{a,b}(t)$ have the same number of oscillations as the basis mother wavelet $\psi(t)$ as they are dilated and translated (i.e., transformed) versions of it. This is why wavelets are exceptionally useful for signals of self-similar or fractal structure.

7.6 Multiresolution Analysis

The *multiresolution analysis* (MRA) is an important concept in wavelet theory which makes wavelet functions extremely attractive for signal analysis. It is regarded as the core of wavelet analysis. When a series is decomposed MRA allows us to obtain information on various levels of frequency and time detail. Through MRA the signal's (or function's) resolution is adapted to a specific level of detail analysis, allowing the examination to "zoom" into the particular characteristics of its structure. The basic idea behind MRA is the decomposition of the whole functional space $L^2(\mathbb{R})$ into *orthogonal projections* i.e., subspaces so as each of those subspaces to contain a part of the original signal (or function). The subspaces provide us with the finer details of the original signal (or function) at 2^j (for any integer

j) total number of scales $\Delta t' = 1, 1/2, ..., (1/2)^j$ which range from lower to higher frequencies.

The original idea belongs to Mallat (1989) which had in mind a "scale-invariant" representation of the signal (referring actually to images from image analysis i.e., 2-dimensional signals). MRA works like a magnifying glass, providing information on finer details of the signal but at the same time allowing for a "zooming out" to provide a (more) complete view of the overall structure. To understand MRA even better, one should regard wavelet transforms as a kind of a band-pass filter where larger scales imply lower frequencies and small bandwidth. While computing the wavelet transform of a signal progressing our computations from lower scales to larger scales, we can stop at each stage and compute the inverse wavelet transform. As we do that we are actually using the remaining coefficients for that stage while setting the smaller scale coefficients to zero. We thus obtain and "interim" version of the series i.e., a subset of the characteristics of the signal, as we build a sequence of smooth (low-pass), detailed (band-pass) or rough (high-pass) versions of the original data. Thus, by starting from the lowest resolution, we can add details to create the higher resolution versions of the original signal ending with a complete synthesis of the signal at the highest resolution.

The formal definition of MRA requires the definition of a function space (sequence of closed subspaces) $V_j \subset L^2(\mathbb{R})$, $j \in \mathbb{Z}$, such that:

$$V_j = \{f \in L^2(\mathbb{R}) : f \quad \text{is piecewise constant on } [k2^{-j}, (k+1)2^{-j}], k \in \mathbb{Z}\} \quad (7.28)$$

This sequence (7.28) of space represents an array of subspaces on increasing resolution as j increases. Each subscpace V_j defined in (7.28), consists of piecewise constant functions over intervals of $\mathbb{R}$ which are precisely two times the length of the V_{j-1}. This sequence as defined above, exhibits the following properties:

1. $\cdots \subset V_3 \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset \cdots$

   $\longleftarrow$ coarser $\qquad\qquad$ finer $\longrightarrow$

   As j increases, the space V_j becomes more and more like the original function space.
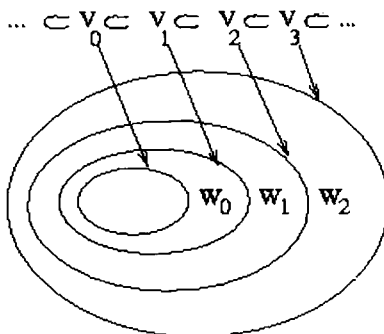
Figure 7.16: Nested subspaces.

2. $x(\cdot) \in V_0$ if and only if $x_{0,k}(\cdot) \in V_0$ also, where $x_{0,k}(t) \equiv x(t - k)$ for $k \in \mathbb{Z}$

3. $x(\cdot) \in V_0$ if and only if $x_{j,0}(\cdot) \in V_j$ also, where $x_{j,0}(t) = x(\frac{t}{2^j})/\sqrt{2^j}$ for $j \in \mathbb{Z}$

4. The following two technical conditions hold:

$$\overline{\bigcup_{j \in \mathbb{Z}} V_j} = L^2(\mathbb{R}) \quad \text{and} \quad \bigcap_{j \in \mathbb{Z}} V_j = \{0\}$$

(here '0' refers to function that is 0 for all t) there exists a scaling function $\phi(\cdot) \in V_0$ such that $\{\phi_{0,k}(\cdot) : k \in \mathbb{Z}\}$ forms an orthonormal basis for V_0, where $\phi_{0,k}(t) \equiv \phi(t - k)$ and V_j called approximation space for scale $\lambda_j = 2^j$. Property (4) is also known as "*completeness*".

7.7 Wavelet Regression

Assume the following regression model:

$$Y_i = f(x_i) + \sigma \epsilon_i \tag{7.29}$$

where $\epsilon_i \sim N(0, 1)$, $n = 2^j$ and $x_i = i/n$. In this section we will show the steps for estimating f using wavelet functions. We represent $f(x)$ as:

$$f(x) = \sum_{k=0}^{2^{j_0}-1} a_{j_0,k} \phi_{j_0,k}(x) + \sum_{j=j_0}^{\infty} \sum_{k=0}^{2^j-1} \beta_{j,k} \psi_{j,k}(x) \tag{7.30}$$

where

$$a_{j_0,k} = \int f(x)\phi_{j_0,k}(x)dx \quad \text{and} \quad b_{j,k} = \int f(x)\psi_{j,k}(x)dx \qquad (7.31)$$

We define the *empirical scaling coefficients* (ESC) and the *empirical wavelet coefficients* (EWC) S_k and $D_{j,k}$ respectively as:

$$S_k = \frac{1}{\sqrt{n}}\sum_i \phi_{j_0,k}(x_i)Y_i \quad \text{and} \quad D_{j,k} = \frac{1}{\sqrt{n}}\sum_i \psi_{j,k}(x)dx \qquad (7.32)$$

Intuitively, the following conditions hold for ESC and EWC:

$$
\begin{aligned}
E(S_k) &= \frac{1}{\sqrt{n}}\sum_i \phi_{j_0,k}(x_i)E(Y_i) = \\
&= \frac{1}{\sqrt{n}}\sum_i \phi_{j_0,k}f(x_i) \\
&= \frac{\sqrt{n}}{n}\sum_i \phi_{j_0,k}f(x_i) \approx \sqrt{n}\int \phi_{j_0,k}f(x)dx \\
&= \sqrt{n}\,a_{j_0,k} \qquad\qquad\qquad\qquad\qquad\qquad (7.33)
\end{aligned}
$$

In a similar way we can show that $E(D_{j,k}) \approx \sqrt{n}\beta_{j,k}$. We construct vector $\mathbf{Z}$ which contains all the empirical coefficients, where $Z_1 = S_1$, $Z_2 = S_2$,...,$Z_n = D_{J-1,2^{j-1}-1}$. Defining $\mathbf{W}$ and $n \times n$ matrix such that the first row is the first father wavelet (scaling function) evaluated at $x_1, ..., x_n$, the second row is the second father wavelet evaluated at the same points etc. etc. adding all the basis elements one row at a time. Then:

$$\frac{1}{\sqrt{n}} \cdot \begin{bmatrix} \phi_{j_0,0}(x_1) & \cdots & \phi_{j_0,0}(x_n) \\ \vdots & \vdots & \vdots \\ \phi_{j_0,2^{j_0}-1}(x_1) & \cdots & \phi_{j_0,2^{j_0}-1}(x_n) \\ \psi_{j_0,0}(x_1) & \cdots & \psi_{j_0,0}(x_n) \\ \vdots & \vdots & \vdots \\ \psi_{J-1,0}(x_1) & \cdots & \psi_{J-1,0}(x_n) \\ \vdots & \vdots & \vdots \\ \psi_{J-1,2^{J-1}-1}(x_1) & \cdots & \psi_{J-1,2^{J-1}-1}(x_n) \end{bmatrix} \qquad (7.34)$$

and $\mathbf{Z}$ can be written as:

$$\mathbf{Z} = \mathbf{WY} \qquad (7.35)$$

This matrix of the empirical wavelet coefficients $\mathbf{Z}$ need not be calculated as the product in (7.35). In practice, we approximate the highest level of scaling coefficients $a_{J-1,k}$ by Y_k since,

$$E(Y_k) = f(k/n) \approx \int f(x)\phi_{J-1,k}(x)dx = a_{J-1,k} \qquad (7.36)$$

A cascade algorithm is applied to obtain the rest of the coefficients and then we multiply this by $1/\sqrt{n}$ to obtain $\mathbf{Z}$. We estimate the scaling coefficients as:

$$\hat{a}_{j0,k} = \frac{1}{\sqrt{n}} Z_{j0,k} \qquad (7.37)$$

For the detail coefficients (or mother wavelets) we obtain $\hat{\beta}_{j,k}$ as a "shrunken" version of the $Z_{j,k}$. A methodology of "non-linear shrinkage" is used here called "*thresholding*". A non-linear approach is preferred in this case as it allows for a better detection of structural changes in the function or sequence. Thresholding generates a sparse vector of wavelet coefficients where most of them are 0 except in the case of the ones that correspond to the jumps or breaks in the continuity of the function or the sequence.[81] We estimate σ as:

$$\hat{\sigma} = \frac{\text{median}(|Z_{J-1,k}| : k = 0, ..., 2^{J-1} - 1}{0.6745} \qquad (7.38)$$

We then insert the computed estimates in formula (7.30) and obtain the estimation $\hat{f}$ where:

$$\hat{f} = \mathbf{W}^T \hat{\mathbf{Z}} \qquad (7.39)$$

and

$$\hat{\mathbf{Z}}^T = \sqrt{n}(\hat{a}_{j0,0}, ..., \hat{\beta}_{J-1,2^{J-1}-1}) \qquad (7.40)$$

---

[81] More analysis is provided on this in the section of "thresholding" in this chapter

7.8 Wavelet Shrinkage-Thresholding

The popularity of wavelets in nonparametric estimation of functions is due to the principle of *wavelet shrinkage*.[82] This refers to removing the denoising of signals by "shrinking" (reducing) the below some specific "threshold" wavelet coefficients to zero. In order to demonstrate how this works, we assume an additive i.i.d. noise process ϵ_i which contaminates a discrete sequence f_i producing the series y_i which is the data under examination:

$$y_i = f_i + \sigma \epsilon_i \tag{7.41}$$

where $\epsilon_i \sim N(0, 1)$. Instead of assuming a certain parametric structure of the data generating process, we follow a "nonparametric regression" approach:

1. The wavelet transform is applied to y_i. Function f must be approximated (estimated) as accurately as possible which means that the estimation $\hat{f}$ should be determined with a small mean square error which implies that the L_2 risk function

$$R(\hat{f}, f) = \frac{1}{n} \sum_{i=1}^{n} E(\hat{f}_i - f_i)^2 \tag{7.42}$$

should be small.

2. Some wavelet coefficients are reduced towards zero (shrink).

3. The inverse wavelet transform produces then a "smoothed" version of y_i.

This idea was refined by Donoho and Johnstone (1994, 1995) and Donoho (1995) and augmented by Nason (1995). Wavelet shrinkage as a smoothing technique is becoming extremely popular in sciences such as seismology, medicine, physics and statistics among others. The ability of wavelets to represent a signal in wavelet space with a few coefficients and successfully capture localised events is what allowed the "shrinkage" concept to be developed. The most popular shrinkage method introduced by Donoho and Johnstone, is called "*Waveshrink*" and focuses around the

---

[82]In this section, in order to describe the methodology with accuracy, we follow closely Bruce and Gao (1996).

estimation of the *shrinkage coefficients*. This estimation is based on three principles (Bruce and Gao, 1996) which also define the steps of the estimation procedure in the same order:

1. The most important features of any signal can be represented by just a few coefficients.

2. Noise affects all wavelet coefficients (at all scales). As it has been already shown, using matrix notation, a wavelet decomposition comprises of the multiplication (projection) of the signal vector $\mathbf{y}$ in (7.41) with the wavelet matrix $\mathbf{w}$ producing the transformation $\mathbf{W} = \mathbf{W}\mathbf{y}$ which can also be denoted as

$$\mathbf{w} = \mathbf{W}\mathbf{f} + \mathbf{W}\sigma\epsilon \qquad (7.43)$$

due to the orthonormality of the transformation in (7.43). Hence $\mathbf{W}\sigma\epsilon$ is i.i.d. normal noise with variance σ^2 assigned to every component of $\mathbf{W}$ i.e., noise is spread over all the wavelet coefficients.

3. By "shrinking" specific wavelet coefficients to zero (usually setting the smaller ones to zero), noise can be eliminated while the important structure is preserved. This principle may be understood better by the explanation in Hubbard (1994). What actually is achieved by wavelet shrinkage it to project the function or signal onto a functional (wavelet) space where noise and true signal are disentangled, thus retrieving through the inverse wavelet transform a "clean" sequence. What is of importance is to locate the subset of the wavelet coefficients that capture the largest energy portion of the original "unclean" signal.

Under the Waveshrink approach Donoho and Johnstone propose to wavelet-decompose f, threshold the coefficients and then reconstruct the function from the thresholded coefficients, producing the estimation $\hat{f}$. There are a number of different thresholding rules and functions available which lead to different amounts of smoothing according to the wavelet functions used in the decompositions and the signal's overall structure.

7.8.1 Thresholding functions and rules

In this section we present, following very closely Bruce and Gao (1996), the most popular thresholding functions or "policies" available in wavelet nonparametric denoising of signals:

Threshold policies

1. **Hard thresholding**

   Under the hard thresholding policy, we only keep the coefficients with absolute values above a fixed threshold level $\lambda > 0$.

   The hard thresholding function is defined as:

   $$d_{j,k}^{\text{hard}} = \begin{cases} 0, & \text{if } |d_{j,k}| \leq \lambda \\ d_{j,k}, & \text{otherwise} \end{cases} \qquad (7.44)$$

   Bruce and Gao (1996)[83] have shown that hard shrinking leads to larger error variance because of the discontinuity of the shrinkage function.

2. **Soft thresholding**

   Soft thresholding shrinks all coefficients towards 0:

   $$d_{j,k}^{\text{soft}} = \text{sign}(d_{j,k})(|d_{j,k}| - \lambda)_+ \qquad (7.45)$$

   Bruce and Gao (1996) again showed that this policy tends to have larger bias as all large coefficients are shanked towards zero by an amount λ. These characteristics of soft and hard thresholding have lead to the development of the next policy:

3. **Semi-soft Thresholding**

   The shrinkage function for the semi-soft policy is defined as:

---

[83]See also Gao and Bruce (1997).

$$
d_{\lambda_1, \lambda_2} = \begin{cases} 0, & \text{if } |x| \leq \lambda_1 \\ \text{sign}(d_{j,k}) \frac{\lambda_2 (|d_{j,k}| - \lambda_1)}{\lambda_2 - \lambda_1}, & \text{if } \lambda_1 < |d_{j,k}| \leq \lambda_2 \\ d_{j,k} & \text{if } |d_{j,k}| > \lambda_2 \end{cases} \tag{7.46}
$$

For values of $d_{j,k}$ near the lower threshold λ_1, the semisoft policy acts like the soft policy whereas for values of $d_{j,k}$ above the upper threshold λ_2, $d_{\lambda_1, \lambda_2} = d_{j,k} = d_{j,k}^{\text{hard}}$ i.e., equivalent to the hard policy. From the above definition (7.46) it follows that we can define hard and soft policies as limiting cases of the semi-soft function for $\lambda_1 = \lambda_2$ and $\lambda_2 = \infty$ respectively.

4. **Quantile thresholding**

This policy focuses on shrinking only a fixed percentage of the smallest wavelet coefficients, say keeping only 10% or the top 100 coefficients on a signal of 1000 observations:

$$
d_{j,k}^{\text{quantile}} = \begin{cases} 0, & \text{if } d_{j,k} \leq q \\ d_{j,k}, & \text{otherwise} \end{cases} \tag{7.47}
$$

where q is the q-th quantile, arbitrarily set.

5. **Universal thresholding**

As proposed by Donoho and Johnstone (1994), where threshold

$$
\lambda = \frac{\sigma \sqrt{2 \log n}}{\sqrt{n}} \tag{7.48}
$$

is imposed on a transformed series of y_i/n, where n is the length of the series and σ is the scale of the noise on a standard deviation scale. We can accommodate both hard or soft thresholding approach with this universal threshold λ.

Research in wavelet analysis has provided a number of different rules which can be combined with the threshold functions-policies in order to obtain various degrees of smoothing. The most basic ones that we also apply in this thesis are the following:
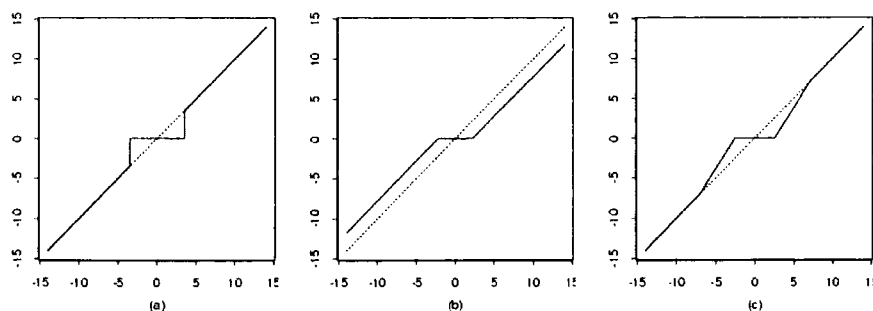
Figure 7.17: Threshold policies from left to right: hard (a), soft (b) and the semi-soft (c) shrinkage.

Threshold Rules:

Researchers in wavelets are constantly exploring the possibilities of new threshold policies and rules. Here we present the most popular, namely the *universal, minimax, top* and *adapt.*

1. **The Universal rule** The universal threshold is defined as $\lambda_j = \sqrt{2 \log(n)}$, where n is the sample size. This method yields the largest thresholds. Intuitively, applying the universal scheme will result in a high degree of smoothness.

2. **The Minimax rule** The threshold λ_j is determined via the minimax rule such as to minimise a theoretical upper bound on the asymptotic risk. This will always be smaller than the universal and thus provide less smoothing. Donoho and Johnstone (1994) and Bruce and Gao (1996) have computed the minimax threshold values for a range of sample sizes both for the soft and hard thresholding policies.

3. **The Top rule** Instead of using statistical theory to determine the level of the threshold, using the top rule one decides for the number of coefficients that will be reduced to zero. This is equivalent to hard shrinkage with threshold $\lambda = |d_j|$ where j corresponds to the $\frac{1-\text{top}}{n}$th quantile of the wavelet coefficients.

4. **The Adapt rule** Donoho and Johnstone (1995) provide a threshold rule which adapts on the signal at every multiresolution level. The basis of this approach is the minimisation of the *Stein's Unbiased Risk Estimator*, popularly know as "SURE" at each resolution level.

7.9   A simple wavelet: The Haar

The simplest way to demonstrate how wavelet functions work would be to use the Haar wavelet. The Haar is regarded as the most basic and simple type of wavelet. With no loss of generality, we can concentrate on the discrete case where the Haar roots lie in the projection operation called the *Haar transform*. This transform may be regarded as a prototype for all other wavelet transforms. Haar (1910) showed that certain square-like wave functions can be translated and scaled to create a basis that spans the L^2 set. It was only years later that this has historically identified as a wavelet system.

The Haar transform is an easy way to demonstrate how wavelets work. One prerequisite for most wavelet analysis approach is that the length of the one-dimensional signal analysed is of an integer power of 2 i.e., 2^j for some integer j.[84] Let us represent all one-dimensional signals of length 2^j as vectors of the same length. These vectors will belong to a vector space say V^j. A basis can be defined for this space. The basis functions for the V^j vector spaces are called *scaling functions* and are usually represented by the letter ϕ in wavelet literature. A simple basis for the V^j vector space would be a set of scaled and translated compactly supported *box* functions:

$$\phi_i^j(x) = \phi(2^j x - i) \qquad i = 0, ..., 2^j - 1 \qquad (7.49)$$

where

$$\phi(x) = \begin{cases} 1 & \text{for } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases} \qquad (7.50)$$

The function ϕ 7.50 is called the *Haar father wavelet* or *Haar scaling function*.

To demonstrate the functionality of the Haar transform better, let us consider two numbers a and b. We use a simple linear transform that replaces these numbers by their average and their difference:

$$\mu = \frac{a+b}{2} \qquad (7.51)$$

---

[84]Modifications of this rule exists as we shall see, which allow sequences of different lengths of those of power of 2 to be regarded for analysis.

and

$$d = b - a \tag{7.52}$$

There is no loss of information in the above transforms as both a and b numbers can be immediately recovered as:

$$a = \mu - \frac{d}{2} \tag{7.53}$$

and

$$b = \mu + \frac{d}{2} \tag{7.54}$$

What has been demonstrated here is the basic idea behind the Haar wavelet transform. Assume a series x_n of 2^n observations $x_{n,k}$:

$$x_n = \{x_{n,k} \mid 0 \le k < 2^n\} \tag{7.55}$$

To each pair of observations $a = x_{2k}$ and $b = x_{2k+1}$ the average and difference transformations (7.51) and (7.52) above are applied. There will be 2^{n-1} a and b numbers pairs for $k = 0...2^{n-1}$ from the two transformations:

$$\mu_{n-1,k} = \frac{x_{n,2k} + x_{n,2k+1}}{2} \tag{7.56}$$

and

$$d_{n-1,k} = x_{n,2k+1} - x_{n,2k} \tag{7.57}$$

In that way, the x_n original series are split into two parts, the μ_{n-1} (the average transformations $\mu_{n-1,k}$) and d_{n-1} (the difference transformations $d_{n-1,k}$) of 2^{n-1} observations each. Using the averages $\mu_{n-1,k}$ and differences $d_{n-1,k}$, one can retrieve the original signal x_n. The $\mu_{n-1,k}$ averages are a more "coarse" representation of the x_n data i.e., one can say that the transformed series is of "less detail" than the original one. The amount of detail lost due to the transformation depends highly on the nature of the original series. If these are fairly smooth, the transformation resembles closely the original information and the finer details of the series can be represented effectively. The $d_{n-1,k}$ differences are the information needed to retrieve

the original signal.

The same transformation steps (7.56) and (7.57) above can be applied again on the transformed series x_{n-1} generating two more sequences of averages μ_{n-2} and differences d_{n-2} and obtaining an even more coarse series x_{n-2} of length 2^{n-2}. This can be repeated n times in total. At the end we will have retrieved n vectors of details d_i with $0 \leq i < n$. Each of these vectors will constitute of 2^i coefficients. Finally we will have retrieved a vector x_0 which will be of the most coarse scale and will only contain one value, $x_{0,0}$ which will be the average of all the samples of the original series. Again, by using the inverse transform we can regenerate the original series, starting from x_0 and working back using the d_i elements. The total number of wavelet coefficients is:

$$1 + \sum_{i=0}^{n-1} 2^j = 2^n \tag{7.58}$$

which is exactly the length of the series x_n. The Haar transform is equivalent to the application of a $N \times N$ square matrix of power of two dimension ($N = 2^n$) to the series x_n. The computational cost is proportional to the length of the series but much less than that of a Fourier transformation. The operations required for the latter are of the order $n \log n$ whereas the wavelet transform only requires n^2 transformations due to the hierarchical structure of the transform. This fact produces remarkable economy in computational time and is one of the features that make wavelets so attractive solutions to filtering signals.

7.10 The Maximal Overlap Discrete Wavelet Transform

The *maximal overlap discrete wavelet transform* (MODWT) is also known as the *undecimated DWT, shift invariant DWT* or *non-decimated DWT*. MODWT is regarded as a modification on the DWT and has five important properties which distinguish it from the latter:

1. The Partial DWT requires a power of two sample size i.e., for any j integer, the sample size must be j^2. MODWT can be applied to signals of any size. The drawback is that it requires more than $O(N)$ computations of the DWT

and specifically $O(n \log_2 n)$ multiplications which is as many as the fast Fourier transform (FFT).

2. The MODWT can be utilised for multiresolution analysis. Shifting circularly the signal by any amount results into shifting the detail and smoothness coefficients by the same amount.

3. The MODWT makes it easier to align the structural features of the multiresolution analysis with the original signal.

4. The MODWT wavelet variance estimator is asymptotically more efficient than the one that is based on the DWT.

5. the MODWT-based spectra are invariant to circular shifts of the signal.

We provide here a brief introduction of the mathematical background for the MODWT following the framework of Percival and Walden (1999). For a thorough discussion of the process one may refer to Percival and Mofjeld (1997), Percival and Walden (1999) or Gençay *et al.*, (2002). We define $\widetilde{\mathbf{W}} = \widetilde{\mathcal{W}}\mathbf{X}$ the Jth order partial MODWT which comprises of $J + 1$ N-length vectors of wavelet coefficients $\widetilde{\mathbf{W}}_1, ..., \widetilde{\mathbf{W}}_j, \widetilde{\mathbf{V}}_j$:

$$\widetilde{\mathbf{W}} \equiv [\widetilde{\mathbf{W}}_1 \widetilde{\mathbf{W}}_2 ... \widetilde{\mathbf{W}}_j \widetilde{\mathbf{V}}_j]^T \tag{7.59}$$

Every vector $\widetilde{\mathbf{W}}_j$ is related to $\lambda_j = 2^{j-1}$ size changes of length and the vector $\widetilde{\mathbf{V}}_j$ is associated with averages of length λ_j and higher. DWT can still be obtained from MODWT for sequences of $N = 2^j$ length with subsampling and rescaling. In the same fashion of the DWT matrix $\mathcal{W}$, the MODWT matrix $\widetilde{\mathcal{W}}$ is constructed from a total of $J + 1$ $N \times N$ submatrices and is expressed as

$$\widetilde{\mathcal{W}} = [\widetilde{\mathcal{W}}_1, \widetilde{\mathcal{W}}_2, ..., \widetilde{\mathcal{W}}_J, \widetilde{\mathcal{V}}_j]^T \tag{7.60}$$

where $\mathcal{W}_1$ is $N \times N$ matrix:

$$\widetilde{\mathcal{W}_1} \equiv \begin{bmatrix} \tilde{h}_0 & 0 & 0 & \cdots & 0 & 0 & 0 & \tilde{h}_3 & \tilde{h}_2 & \tilde{h}_1 \\ \tilde{h}_1 & \tilde{h}_0 & 0 & \cdots & 0 & 0 & 0 & 0 & \tilde{h}_3 & \tilde{h}_2 \\ \tilde{h}_2 & \tilde{h}_1 & \tilde{h}_0 & \cdots & 0 & 0 & 0 & 0 & 0 & \tilde{h}_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \tilde{h}_3 & \tilde{h}_2 & \tilde{h}_1 & \tilde{h}_0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & \tilde{h}_3 & \tilde{h}_2 & \tilde{h}_1 & \tilde{h}_0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & \tilde{h}_3 & \tilde{h}_2 & \tilde{h}_1 & \tilde{h}_0 \end{bmatrix} \qquad (7.61)$$

The rows of $\tilde{\mathbf{h}}_1 = \mathbf{h}_1/\sqrt{2}$ are the rescaled wavelet filter coefficients which we circularly shift by $m - 1$ for $m = 1, ..., N$. Generalising this we can show that for rescaled wavelet coefficients $\tilde{\mathbf{h}}_j = \mathbf{h}_j/2^{j/2}$ and scaling filter coefficients $\tilde{\mathbf{g}}_J = \mathbf{g}_J/2^{J/2}$, we can construct the remaining sub-matrices in (7.60) in a similar way as in (7.61). That leaves matrix $\widetilde{\mathcal{V}_J}$ to be defined. This has the same structure as $\widetilde{\mathcal{W}_J}$ but circularly shifted scaling coefficients are used instead of the wavelet coefficients.

Just as in the DWT, the MODWT is an energy preserving transform and the total energy of any time series $\mathbf{X}$ can be partitioned by the MODWT scaling and wavelet coefficients:

$$\|\mathbf{X}\|^2 = \sum_{j=1}^{J} \|\widetilde{\mathbf{W}}_j\|^2 + \|\widetilde{\mathbf{V}}_J\|^2 \qquad (7.62)$$

This allows for the formulation of the MODWT wavelet variance, covariance and correlation.

7.11 Conclusions

In this chapter we presented the basic outline of the wavelet analysis framework. Understandably, one can not cover such an issue in a few pages. Wavelet theory is a subject for a Ph.D. thesis alone. The presentation here though has clearly indicated that wavelet filters can be used in conjunction with financial data, to provide interesting answers. The highly volatile and intermittent character of financial time series, their possible chaotic structure and their irregular periodicities, makes these sequences ideal candidates for wavelet analysis.

The fact that wavelet transforms have clear advantages over Fourier tran-

sforms, does not necessarily invalidate the latter. It is clear from our discussion so far that for financial time series wavelets seem to be more appropriate. These series are of limited length, they exhibit volatility clustering and violent localised shocks. Hence Fourier analysis is not very helpful as it will misinterpret singularities such as market crashes as periodic artifacts. Spectral analysis is still helpful though, especially for long periodic signals which could be obtained for example from high frequency data sets such as tick by tick financial data. Of course we could still apply wavelet transforms and obtain even more information. The conclusion is that wavelets do not render other techniques useless but can be very effective when used on financial data.

In the next chapter we will be using wavelets to produce qualitative information for financial time series and their volatilities. We will be using the continuous transforms to generate scalograms which indicate any periodic or non-periodic structure and can also reveal self similarity, chaotic or stochastic characteristics for the time series examined. Following this investigation we will be applying the discrete wavelet transform to decompose financial time series in various resolutions and recover their high, medium and low frequency structures. We will also be using the same approach to denoise these series and demonstrate how wavelet smoothing can provide residual sequences remarkably white with very little computational effort. We will be using this approach on 30 years of daily data for the FTSE ALL SHARE index closing prices and logarithmic returns. Combining the methodologies and theory discussed in chapters 3-6 we demonstrate that wavelets can amplify the nonlinear characteristics of the series by filtering out noise efficiently. The recurrence plots reveal a very strong deterministic chaotic nature for the FTSE returns and the calculation of certain invariant measures confirms these results.

CHAPTER 8
Time-frequency analysis

8.1 Introduction

In the previous chapter we have provided a brief introduction on the most basic aspects of wavelet theory. In this chapter we are using wavelet transforms exploring their applicability in describing financial time series.[85] Our intention is to demonstrate that a wavelet semiparametric (and boldly atheoretic) approach can provide more insight on the structure and behaviour of financial time series. Our main interest is to show that by using wavelets we do not only capture salient features and provide better approximations for financial time series but we also reveal details of nonlinear structure and recurrent patterns in various scales.

The first part of this chapter is dedicated to the demonstration of continuous wavelet transforms (CWT) on the daily closing prices of the FTSE ALL SHARE (FTSE) index for the last 30 years. We also apply the CWT on the continuously compounded returns (or logarithmic) returns of the series and the realised volatility of the FTSE. We run the CWT on the above data sets using different wavelet functions and cross-examine the results. Our main purpose is to provide a cartography through wavelet transforms of the dynamics of the series under examination for various different time scales and investigate any localised events such as stock market crashes etc. etc. We also investigate seasonal or self similar recurrent structures that may appear in the wavelet scalograms as a result of the time-scale analysis.

As a second application of wavelet theory in the analysis of financial time series we examine the same daily data as above using the discrete wavelet transform (DWT). We provide a multiresolution analysis (MRA) approach and explore the dynamics of the series in various scales-frequencies via the DWT. We then use

---

[85]Our aim is not to provide an exhaustive demonstration of the applicability of wavelet theory on the analysis of financial time series. Such an endeavour would require time and space that would take us beyond the limits of this thesis. It is fairly self-explanatory that the wavelet approach potential in signal analysis is enormous. Here we demonstrate some key applications which are in accordance with previous research and provide a sane basis for critical evaluation of past findings and incentive for future advances.

as a second step the Waveshrink algorithm, as described in the previous chapter, to denoise the series. We use wavelets to reduce the amount of noise in a non-parametric fashion. We compare the fitted values and residuals with the original sequences and check for normality and stationarity. We furthermore examine the fits and residuals for traces of deterministic recurrences and structures using the methodologies described in chapters 4 to 6. We investigate mainly by inspecting and comparing the recurrence plots (RP) of the sequences and by calculating chaotic invariant statistics.

8.1.1 Previous research

Wavelets are becoming more popular day by day as the academic community appreciates their ability to detect localised events as well as periodic structures. Schleicher (2002) provides an interesting introduction to the general subject. One of the first applications was by Greenblatt (1994) who used wavelets for outlier detection. Jensen (1994) uses wavelets to estimate fractionally integrated processes. He shows an alternative way to estimate the fractional differencing parameter and shows the advantages of wavelets over the existing method of Geweke and Porter-Hudak (1983). Jensen has developed a consistent OLS estimate of a fractionally integrated processes' differencing parameter, using continuous wavelet theory as constructed from smoothing kernels. He derived the asymptotic biasness and variance of the OLS estimate and tested the consistency of the estimate with a number of Monte Carlo experiments. Jensen (1999b) continues his long memory related research applying compactly supported wavelets to the ARFIMA(p, d, q) long-memory process to develop an alternative maximum likelihood estimator of the differencing parameter d. He shows that this is invariant to the unknown mean of the process and the model specification as well as the noise contaminating the data. Again, he finds the wavelet based maximum likelihood estimator to be robust to model specification and as such he proposes it as an attractive alternative semiparametric estimator to the Geweke and Porter-Hudak (1983) one. Olmeda and Fernandez (2000), provide a counter-balance by drawing our attention to the pitfalls of using wavelet filtering for denoising and forecasting purposes. Capobianco has also a series of papers

on the field (see Capobianco 1997,a 1997b, 1999a, 1999b and 1999c). Capobianco (2001) uses wavelets for describing financial returns processes. He studies the Nikkei stock index in high frequency and shows results about modelling with GARCH when the data have been preprocessed by wavelet transforms. He demonstrates that one can obtain better volatility prediction power for one step ahead forecasts implying that latent volatility features can be detected more efficiently. Capobianco (2002) uses multiresolution analysis to approximate volatility processes. He focuses on intra-day dynamics and again shows how wavelet transforms can improve our view of volatility dynamics provided by a GARCH specification or wavelet pre-processed sequences. These findings are consistent with his earlier work (see Capobianco 1997a and 1997b).

In this chapter we follow a slightly different approach. We do not concentrate on long memory specifications and drift parameters for high and ultra-high frequency data. We rather use a combined approach of wavelet filtering and chaos theory in order to establish whether our series are governed by deterministic dynamics or if these dynamics are partially deterministic and contaminated by stochastic randomness. We use wavelets in a continuous and discrete multiresolution framework to show their overall functionality. We then use discrete transforms to denoise financial returns. We establish the whiteness of the residuals and continue by applying an array of tests for nonlinearity and determinism on the filtered series. To this extent, our approach is very novel.

8.1.2 Datasets and their descriptive statistics

The dataset utilised in this chapter comprises of 8192^{86} daily observations of the FTSE index closing prices. We decompose this set via both discrete and continuous wavelet transforms. In the same fashion we decompose the continuously compounded (logarithmic) returns of the FTSE and the realised volatility (that being calculated as the squared returns sequences). All daily sequences start form the 10th of July 1970 and end the 30th of November 2001. All daily closing prices

---

[86]Although the Maximum Overlap DWT (see MODWT in previous chapter) approach allows us to examine time series of any length and not just of a power of 2, we have selected this sample size in order to compare MODWT with other discrete wavelet transforms that require sequences of length $n = 2^j$.
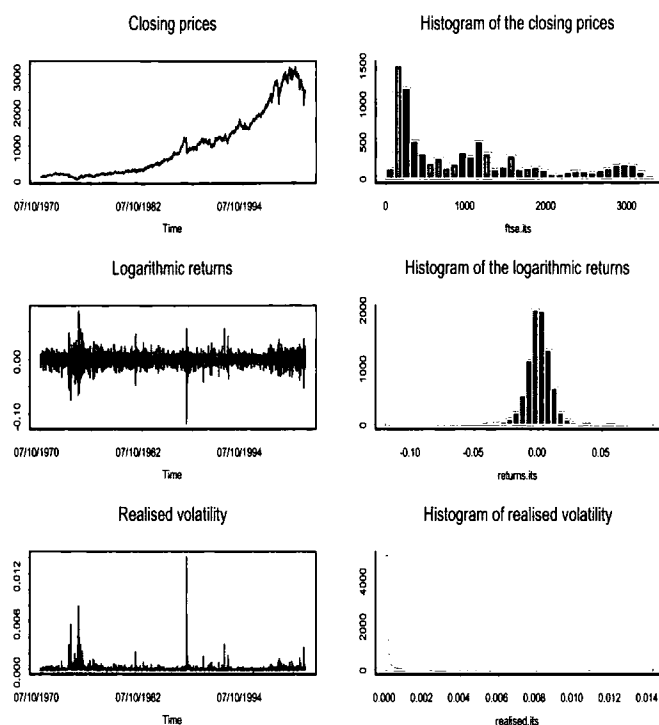
Figure 8.1: Time series plot of the data and their corresponding distributions

of the FTSE have been downloaded from Datastream.

In table 8.1 we display the descriptive statistics for the FTSE index, its logarithmic returns and the realised volatility. From this table and the inspection of the relevant distribution histograms in figure 8.1 we can deduce that the distribution of the closing prices of the FTSE index is positively skewed whereas the corresponding returns are leptokurtic and the realised volatility positively skewed as well. These results are according to the stylised facts for this kind of data (see Cont, 2001). The Jarque-Bera (see Bera and Jarque, 1981) test for normality clearly rejects the null for all sequences, as expected.

As a second stage examination we generate the autocorrelation (ACF) and partial autocorrelation (PACF) function plots of the three series in figure 8.2. The actual values of the ACF and PACF coefficients are listed in tables 8.2 and 8.3 together with the corresponding Q-statistic values and their χ^2 probability values. These can help as distinguish any ARIMA structure in the sequences.

Statistic	index	returns	realised volatility
minimum	61.92	-0.1191000	0.000e+000
Q1	221.60	-0.0048350	5.287e-006
median	768.90	0.0003343	2.859e-005
mean	1001.00	0.0003669	9.995e-005
Q3	1514.00	0.0058370	9.368e-005
maximum	3266.00	0.0894300	1.419e-002
st.deviation	899.0817	0.009992	0.000335
skewness	0.941687	-0.332639	19.48220
kurtosis	2.741635	12.32305	608.4498
Jarque-Berra	1233.378 (0.0)	29815.85 (0.0)	1.26e+08 (0.0)

Table 8.1: Descriptive statistics. Jarque-Bera p-values within parenthesis.
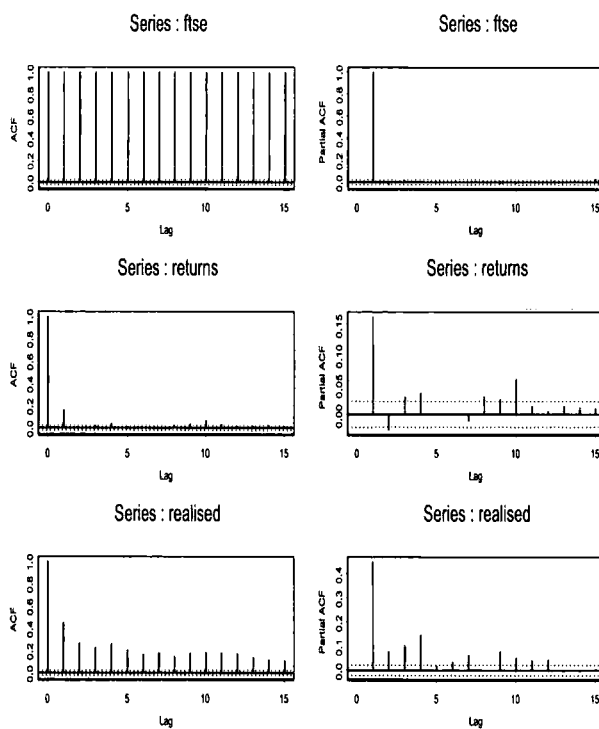
Figure 8.2: The ACF (left column) and PACF (right column) functions of the series.

lags	index	Q-Stat	returns	Q-stat	volatility	Q-stat	prob.
1	1.0000	8188.6	1.0000	217.78	1.0000	1667.80	0.00
2	0.9997	16373	0.1630	217.78	0.4512	2243.80	0.00
3	0.9993	24552	0.0000	222.33	0.2651	2665.40	0.00
4	0.9990	32727	0.0236	237.81	0.2268	3234.60	0.00
5	0.9987	40897	0.0434	239.22	0.2635	3575.90	0.00
6	0.9983	49062	0.0131	239.33	0.2041	3813.50	0.00
7	0.9980	57223	0.0038	239.98	0.1702	4079.80	0.00
8	0.9976	65379	-0.0089	246.09	0.1802	4259.70	0.00
9	0.9973	73531	0.0273	256.13	0.1481	4512.20	0.00
10	0.9969	81678	0.0350	291.90	0.1755	4783.90	0.00
11	0.9966	89820	0.0660	300.82	0.1820	5051.10	0.00
12	0.9962	97958	0.0330	302.47	0.1805	5297.90	0.00
13	0.9959	106091	0.0142	305.77	0.1734	5463.00	0.00
14	0.9955	114219	0.0200	309.83	0.1418	5579.30	0.00
15	0.9952	122342	0.0223	312.16	0.1190	5682.20	0.00

Table 8.2: Autocorrelation function coefficients with their corresponding Q-statistics and their probability values.

lags	index	Q-Stat	returns	Q-stat	volatility	Q-stat	prob.
1	0.9997	8188.6	0.1630	217.7800	0.4512	1667.80	0.00
2	-0.0255	16373	-0.0273	217.78	0.0773	2243.80	0.00
3	0.0168	24552	0.0288	222.33	0.1030	2665.40	0.00
4	-0.0047	32727	0.0358	237.81	0.1472	3234.60	0.00
5	-0.0092	40897	0.0008	239.22	0.0209	3575.90	0.00
6	0.0039	49062	0.0023	239.33	0.0356	3813.50	0.00
7	0.0017	57223	-0.0116	239.98	0.0643	4079.80	0.00
8	0.0066	65379	0.0298	246.09	-0.0011	4259.70	0.00
9	-0.0149	73531	0.0254	256.13	0.0792	4512.20	0.00
10	-0.0125	81678	0.0590	291.90	0.0531	4783.90	0.00
11	0.0052	89820	0.0138	300.82	0.0416	5051.10	0.00
12	0.0003	97958	0.0057	302.47	0.0444	5297.90	0.00
13	-0.0058	106091	0.0133	305.77	-0.0070	5463.00	0.00
14	-0.0023	114219	0.0114	309.83	-0.0069	5579.30	0.00
15	0.0291	122342	0.0102	312.16	0.0069	5682.20	0.00

Table 8.3: Partial autocorrelation function coefficients with their corresponding Q-statistics and probability values.

8.2 A time-frequency approach: the CWT scalograms

In the CWT and DWT analysis that follows, we utilise two wavelet functions. The Haar and the symmlet 8 (or s8 for short) wavelet. These two "mother" functions are depicted in figure 8.3 together with their corresponding "father" or scaling functions. These wavelets were chosen after careful experimentation. Given that a wavelet should resemble the time series analysed, the Haar is not really an appropriate choice. It is though a very simple function and as we shall demon-
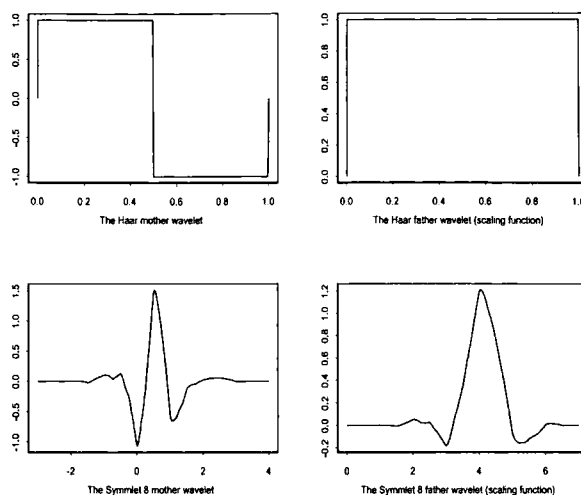
Figure 8.3: From top to bottom: the Haar and s8 mother wavelet funtions and their corresponding scaling functions.

strate, the information we obtain from the CWT qualitatively similar to that of the s8 wavelet. The s8 mother function seems a more likely candidate and produces slightly more detailed scalograms. The overall structure of those though does not differ significantly from the Haar versions. In figures 8.4 - 8.6 we have produced the CWT scalograms for the first 8000 observations of the FTSE ALL SHARE index closing prices, the corresponding logarithmic returns and the realised volatility. We have produced 2 scalograms for each case, one for every wavelet function utilised. From an initial inspection of the various scalograms, we can see that the Haar ones provide a finer decomposition ("thinner") than the symmlet 8 (s8) for the returns (figure 8.5) and the realised volatility (figure 8.6), due the "blocky" discontinuous structure of the Haar mother wavelet depicted in figure (8.3). Because of the smoothness of the symmlet 8 mother wavelet, the CWT scalograms for that wavelet are more "elegant", sometimes even more informative, with smoother transitions between low and high valued wavelet coefficients and clearer bifurcations. An interesting point derived from the comparison of the Haar and s8 scalograms is that for both the levels and the transformations of the FTSE series, the story they deliver is the same.

By careful examination of the index closing prices scalograms in figure 8.4, we can see that both the Haar and sym8 CWT coefficients change patterns after the 3000th (mainly 4000th i.e., roughly the 1st half of the history of the series) observation onwards and especially from the short "booming" period before the 1987 crash. Until that point, the prices of the wavelet coefficients are lower, indicative of the relative smoothness or lack of excessive volatility and of very weak positive trend. Both Haar and s8 versions of the story are quite similar with a small difference. For the larger scales of between 150 and 250 days, the Haar CWT scalogram reports wider and fewer periods of smaller coefficients than that of the s8 wavelet. We attribute this difference to the structure of the wavelet function itself. Both wavelets though are able to recognise that for the lower scales there is a relative absence of trend (i.e., of a low frequency component) whereas for larger scales and larger time "windows", a very weak trend is more obvious for the first half of the index series. One may recognise those as the darker areas (*"tree trunk"* like formations) i.e., collections of high coefficients on the top of both scalograms in figure 8.4. It is obvious form the two scalograms that the wavelet coefficients are able to capture the change of the pattern after the 4000th observation. They also capture the increase in volatility and trend of the series. The difference between the Haar and the s8 scalogram is more evident on the right bottom half of the graphs where for the s8 wavelet, one can identify more easily the bifurcations formed by the coefficients. This mainly shows how small and large sequences of coefficients interchange and may imply a multifractal structure or some kind of aperiodic cyclicity.

The patterns discussed so far, clearly change for the last half of the series. It is obvious from the time series plot on top of figure 8.4, that there is a increase of the steepness and the variance of the index sequence. This follows up historically the occurrence of the 1987 stock market crash. The crash occurs in the vicinity on the 4500th x-axis coordinate, where both Haar and s8 CWT scalograms show a concentration of high coefficients on all scales (shown as an inverted dark peak). We see that regardless of the choice of the mother wavelet, the actual timing of the crash of 1987 is detected successfully. Following that point, the volatility of the series seems to increase considerably with finer bifurcations of wavelet coefficients

occurring in low, medium and large scales. It is obvious that the frequency and the intensity of the aperiodic cycle structures has changed for the last half of the series. The keen eye can also identify the rest of the famous crises as they occur after observation 7000 such as the Asian crisis, the NASDAQ and others. These and the effect of the incident of the 11th of September 2001 can be seen in figure 8.7 where we have produced the s8 CWT scalogram of the whole 8192 index observations. In our analysis so far we have chosen to limit our scalograms to the first 8000 observations in order to exclude the intensive fluctuations of the last part of the history of the series. Although our discrete and continuous analysis has included all 8192 observations, we choose to truncate the sample in order to avoid depicting the large valued coefficients at the end of the scalograms by excluding 192 points. We do this as we are mainly interested in the 1987 crash which seems more isolated and clear to interpret. We examine though the scalogram of this last cluster of observations at the end of this section. We can thus concentrate on the oil crisis, the 1987 and Asian markets crashes and avoid "blurring" of the results at the edge of the series because of the increased concentration of high valued coefficients due to the clustering of well known recent events. It would be interesting to see in a couple of years how these scalograms would have evolved with the inclusion of the recent and future history of the series.

In figure 8.7, we can clearly see after the vertical line the change in the scalogram's pattern. We can also locate the intense oscillations following the Asian crisis, the NASDAQ crash and the September the 11th events at the darker regions of the right edge of the scalogram. An interesting point is that the oil crisis of the 70s is not that evident from the levels of the index as in the scalograms of the returns and the realised volatility. This is more clearly shown in figure 8.8.

For further analysis we selected three sub-samples from equal distinct sub-periods to examine. We utilised only the s8 CWT scalograms. The first period covers 500 sample points, starting from the 1000th one and ending on the 1500th one. It covers the daily observations ranging from 09/05/1974 to 08/04/1976. The second and the third have both length of 1000 observations. The second starts on the 4000th one and ends on the 5000th observation. It covers the timespan

07/11/1985 - 07/09/1989. The third and last subsample refers to the period between the 7000th and 8000th observation i.e., the dates 08/05/1997 and 07/03/2001. The analysis of the first, second and third sub-samples is displayed in figures 8.9, 8.10 and 8.11 respectively. In table 8.5 we have listed the 19 largest shocks or oscillations encountered in the history of the whole sample by date of occurrence and position in the sample for reference reasons. In all the above mentioned figures, we choose to display on top the corresponding realised volatility sequences which provides an adequate representation of the magnitude of the oscillations, as these are captured by the wavelet coefficients.

Subsample	Dates	Range	Size
1	09/05/1974-08/04/1976	1000-1500	500
2	07/11/1985-07/09/1989	4000-5000	1500
3	08/05/1997-07/03/2001	7000-8000	1500

Table 8.4: The 3 subsamples used in figures 8.9-8.11.

Dates	Observation
06/12/1973	890
14/12/1973	896
01/03/1974	951
02/01/1975	1170
24/01/1975	1186
27/01/1975	1187
29/01/1975	1189
30/01/1975	1190
07/02/1975	1196
10/02/1975	1197
11/03/1975	1218
17/04/1975	1245
19/10/1987	4507
20/10/1987	4508
21/10/1987	4509
22/10/1987	4510
26/10/1987	4512
10/04/1992	5676
11/09/2001	8134

Table 8.5: Dates and positions of th 19th largest oscillations in the FTSE series as these are identified by the 19 largest DWT wavelet coefficients.
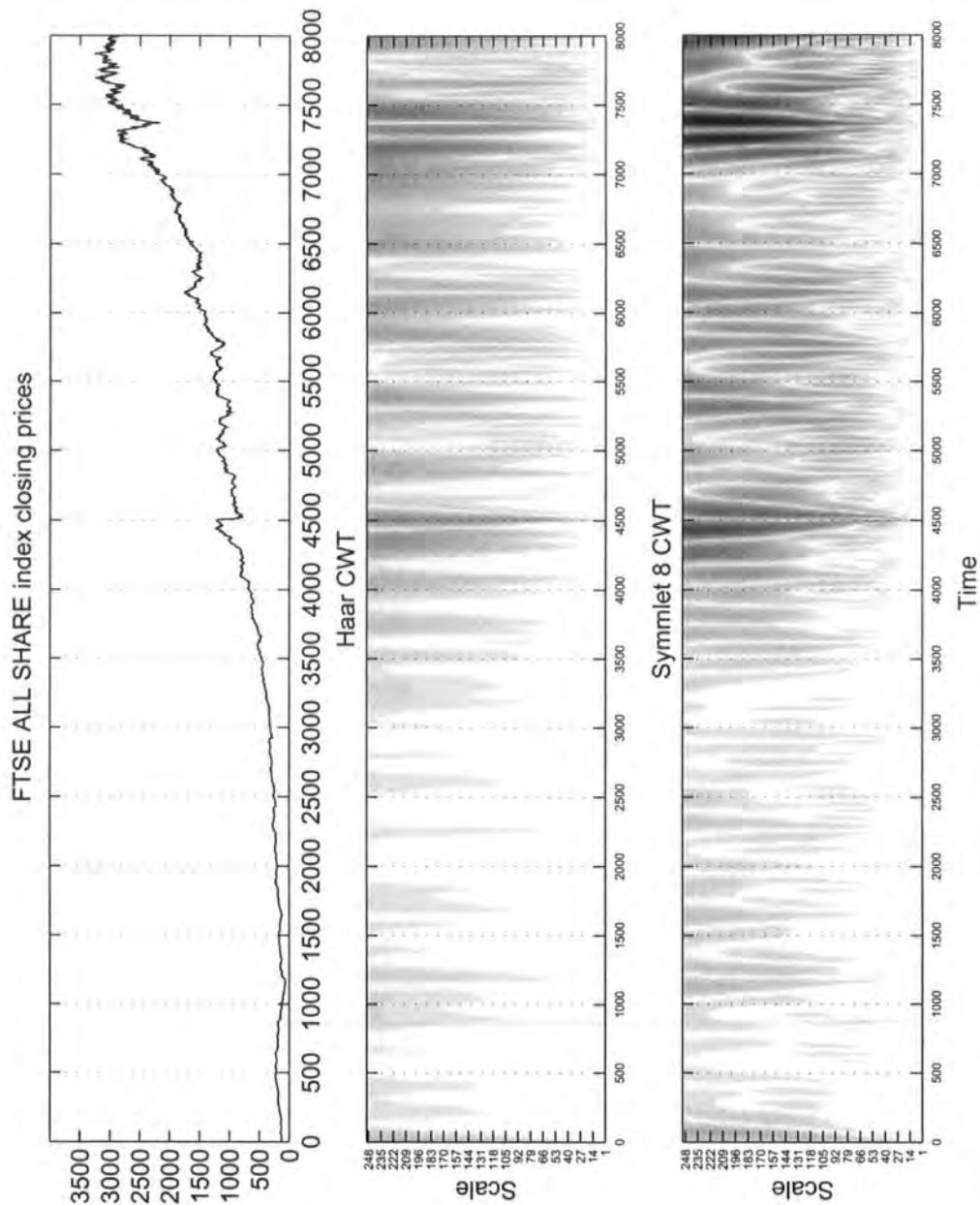
Figure 8.4: CWT of the FTSE ALL SHARE index closing prices. From top to bottom: actual time series, CWT scalograms using a Haar wavelet and a symmlet 8. The darker the colouring of the coefficients, the larger their value.
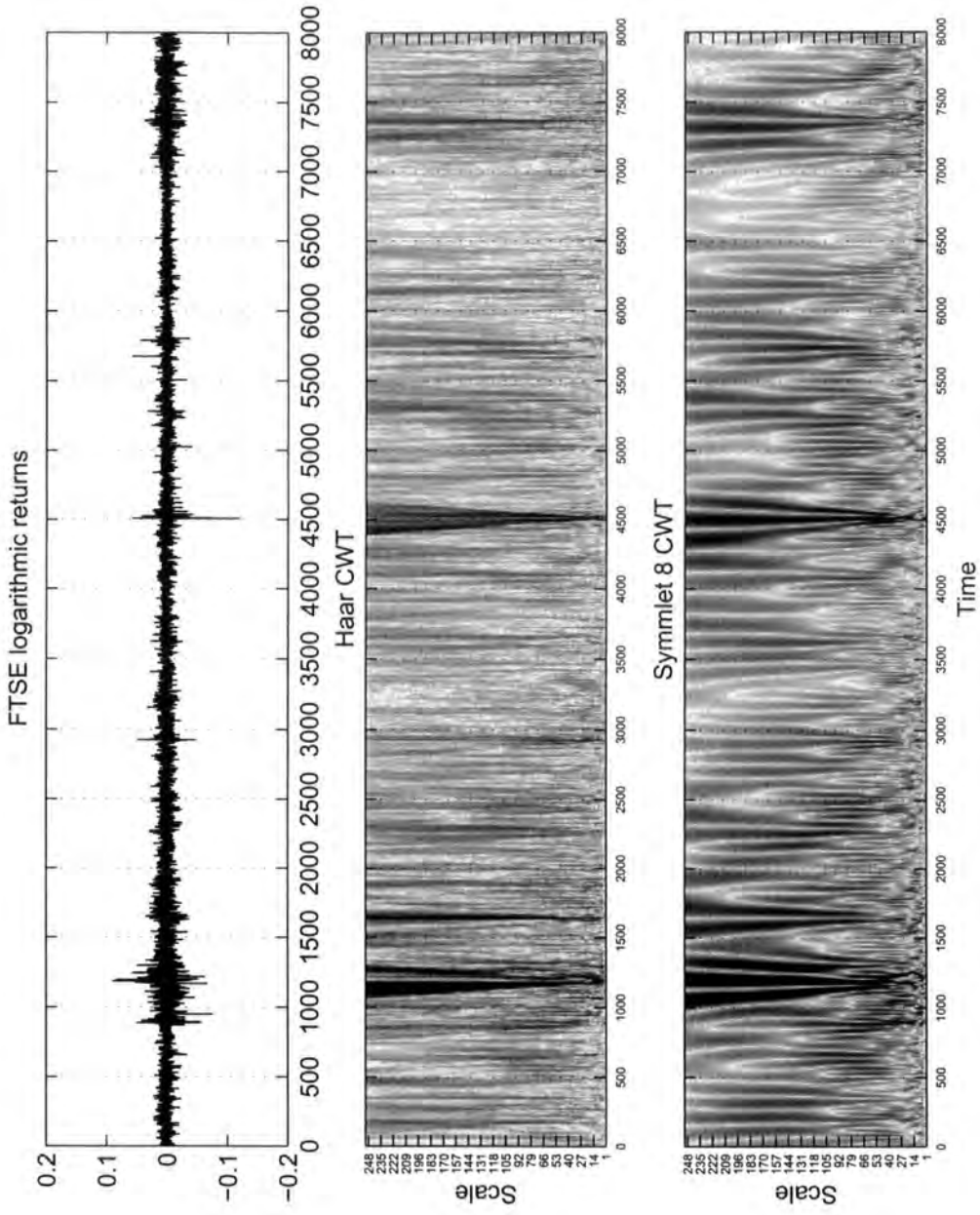
Figure 8.5: CWT of the FTSE ALL SHARE index returns. From top to bottom: actual time series, CWT scalograms using a Haar wavelet and a symmlet 8. The darker the colouring of the coefficients, the larger their value.
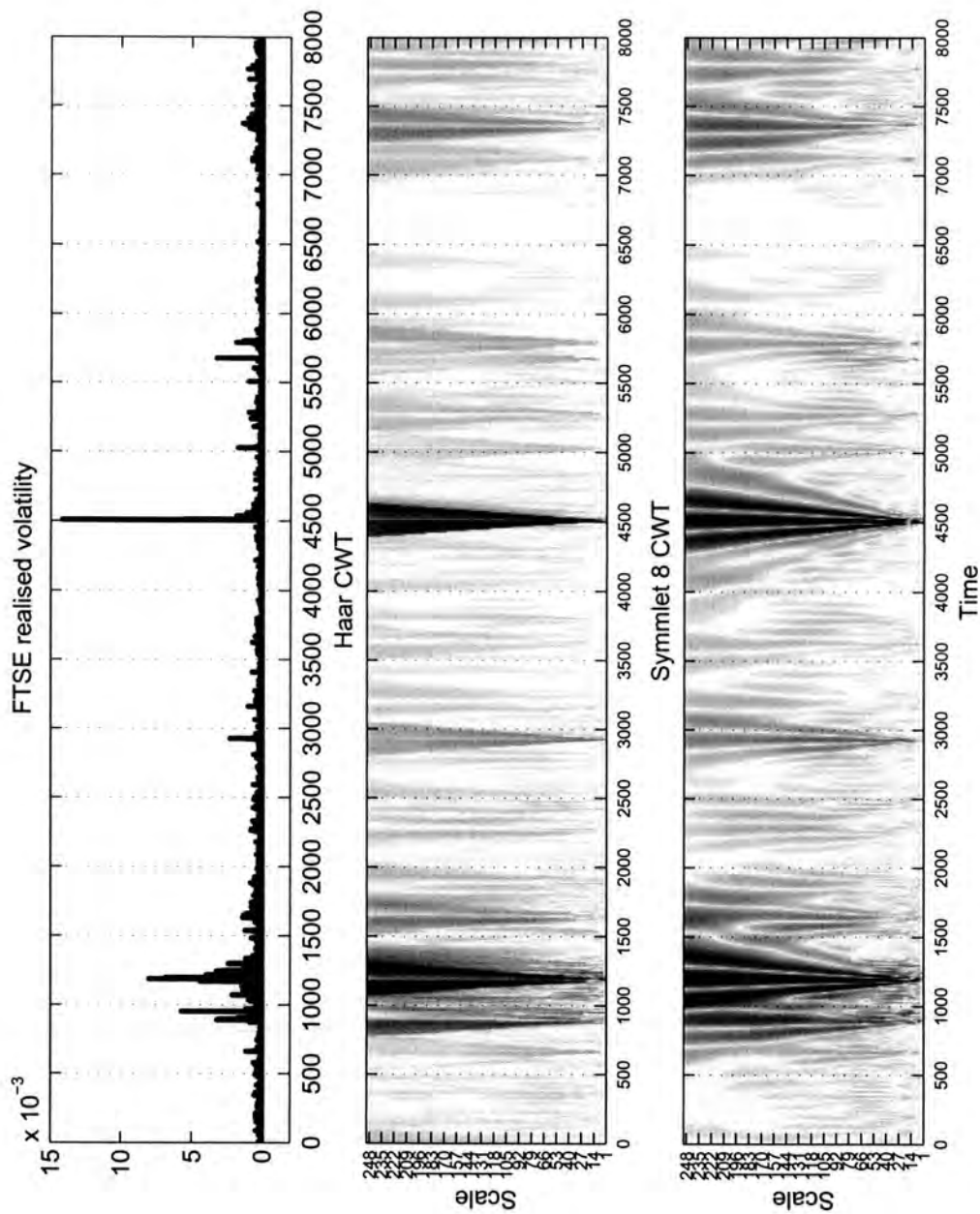
Figure 8.6: CWT of the FTSE ALL SHARE index realised volatility. From top to bottom: actual time series, CWT scalograms using a Haar wavelet and a symmlet 8. The darker the colouring of the coefficients, the larger their value.
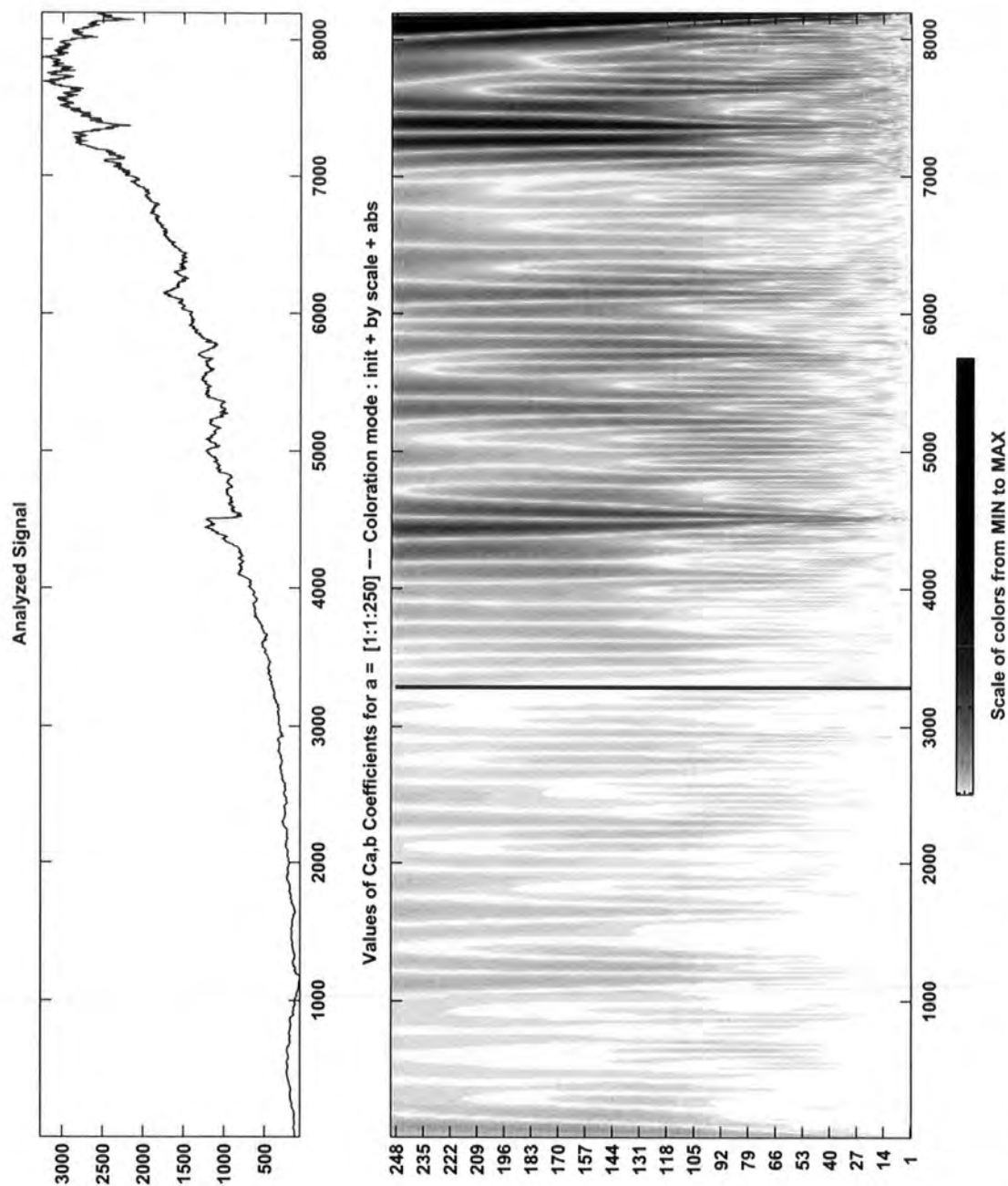
Figure 8.7: CWT of the FTSE ALL SHARE index closing prices. From top to bottom: actual time series and CWT scalogram using a Haar wavelet. The darker the colouring of the coefficients, the larger their value.
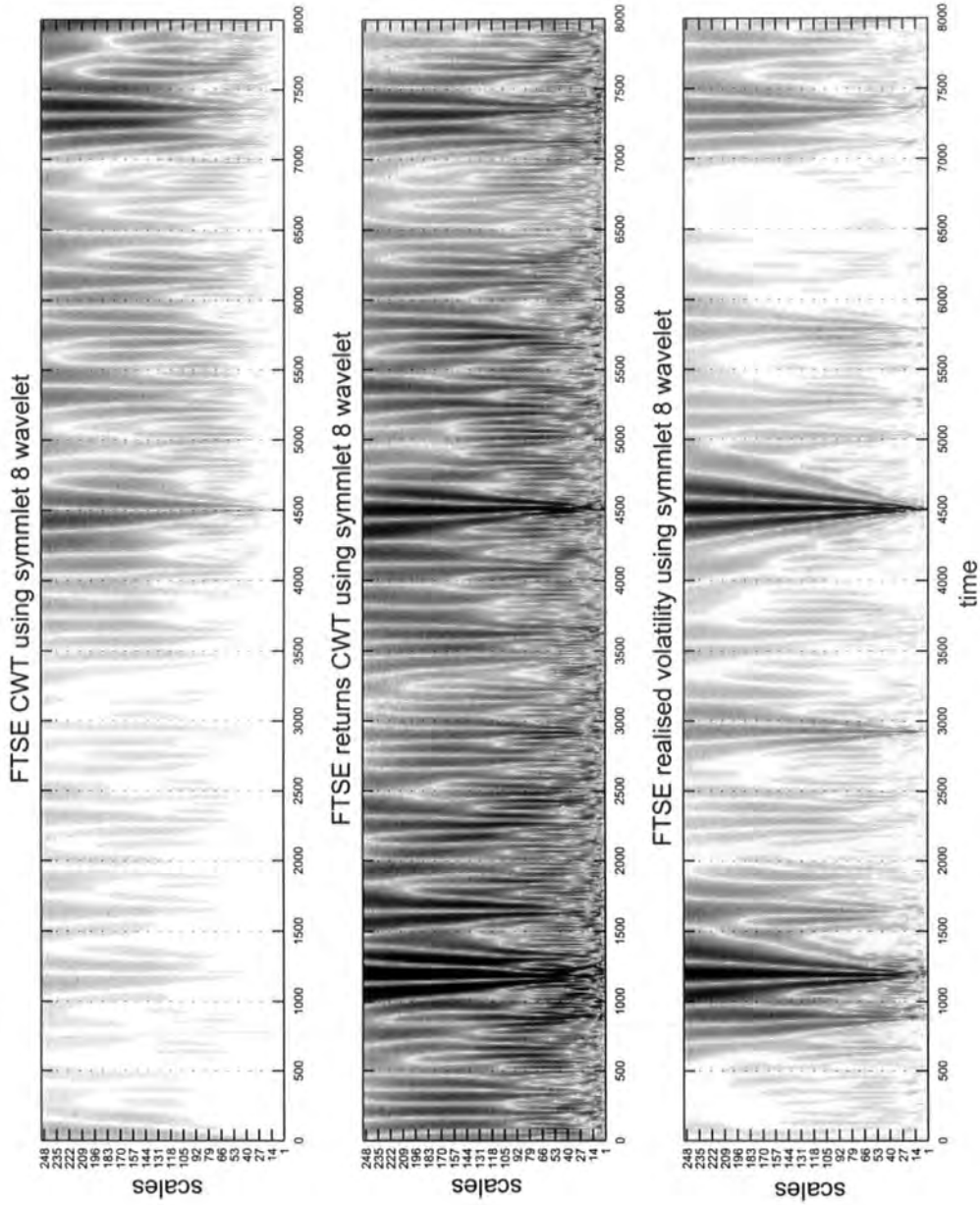
Figure 8.8: Comparison of the s8 CWT scalograms of figures 8.4, 8.5 and 8.6. It is obvious that whether we examine the actual series, returns or realised volatility scalograms, we can easily detect the locality of financial shocks and changes in trend or variation.
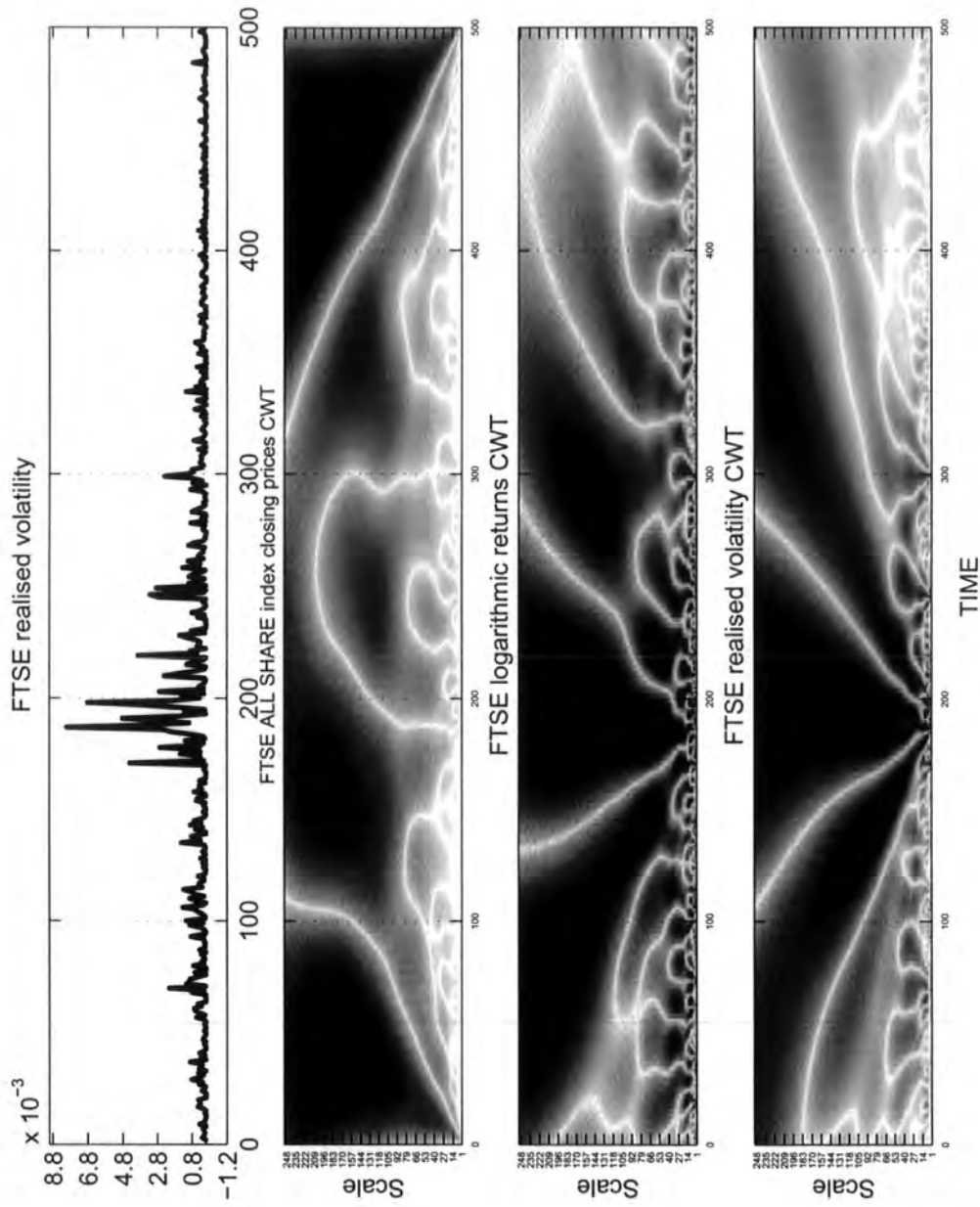
Figure 8.9: s8 CWT scalograms of the period 09/05/1974 till 08/04/1976. 500 observations starting on the 1000th.
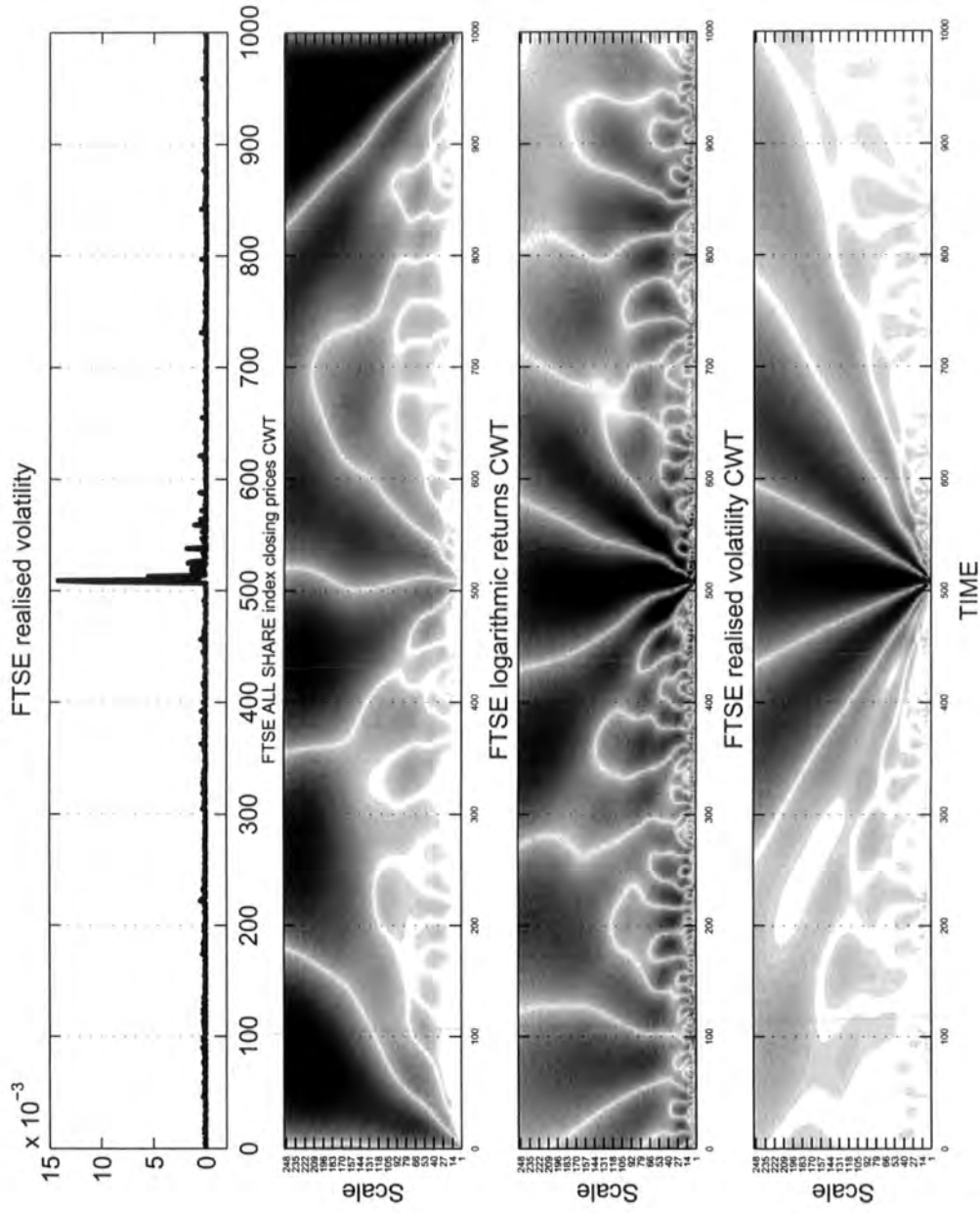
Figure 8.10: s8 CWT scalograms of the period 07/11/1985 till 07/09/1989. 1000 observations starting in th 4000th.
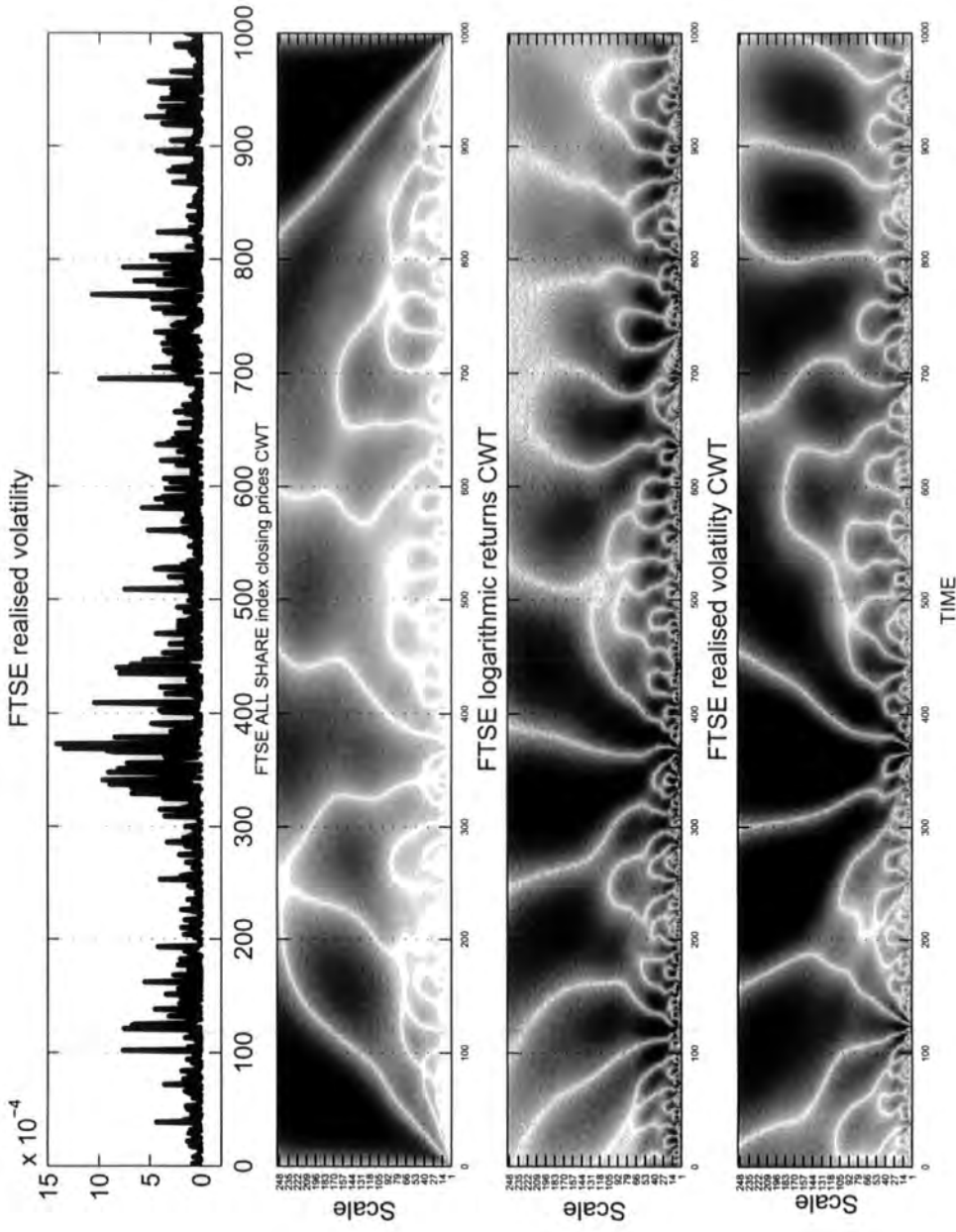
Figure 8.11: s8 CWT scalograms of the period 08/05/1997 till 07/03/2001. 1000 observations starting in the 7000th.

8.3 The DWT of the FTSE

In the previous section we used CWT coefficients in scalograms to establish facts about financial time series, their fluctuation and volatility patterns. In this section we concentrate on the application of the discrete wavelet transform (DWT) for the analysis and denoising of the same sequence used in the preceding CWT experimentation. Specifically we concentrate on the analysis of the FTSE ALL SHARE index and its logarithmic returns. Our main purpose here being to show how a multiresolution analysis (MRA) approach can be used to describe and smooth the series. As a second step we utilise the WAVESHRINK algorithm, as described in the previous chapter, in order to denoise the levels and returns sequences.[87] As a last step we use the techniques demonstrated in chapters 2 and 3 to detect deterministic patterns and recurrences, using both the information from the CWT scalograms and the recurrence plots of the smoothed and residual series. We find evidence of deterministic structures in parts of the smoothed series.

In the analysis that follows, we have applied the Maximum Overlap Discrete Wavelet Transform (MODWT) as defined in the previous chapter for the reasons stated there. We examined the series using the Haar, daubechies 6 and 20 (d6, d20), symmlet 20 (s20) and coiflet 30 (c30) wavelets MODWT as in figure 8.12. We found that the results do not differ remarkably for the d20, s20 and c30 wavelets so we concentrate in this chapter on the report of the Haar and d20 DWT and MRA results. For the Waveshrink algorithm, we experimented with the above wavelets for various levels and smooth factors and we found that the best results were not that much due to the selection of an "appropriate" wavelet function as for the correct selection of the algorithm parameters. We concluded that a d6 wavelet was producing adequate smoothing. The data we use are again the 8192 observations of the FTSE ALL SHARE index closing prices and their continuously compounded returns.

---

[87]We compared the residuals from this smoothing with the ones obtained from a GARCH(1,1) estimation (assuming both Gaussian and Student-t distributions) and we deduced that the WAVESHRINK approach provides much "cleaner" normal residuals than that of the GARCH(1,1).
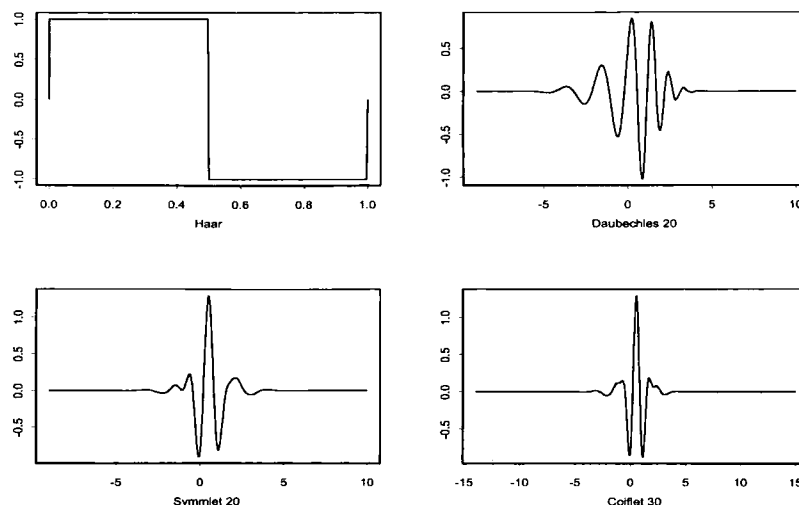
Figure 8.12: From left to right, top to bottom: the Haar, Daubechies 20 (d20), Symmlet 20 (s20) and Coiflet 30 (c30) mother wavelet functions.

8.3.1 Multiresolution Analysis

In figures 8.13 and 8.14 we display the multiresolution approximations and decompositions of the FTSE ALL SHARE index closing prices and its returns respectively. The MRA in these figures was conducted with a Haar wavelet. We used 8 levels of approximation which we found to provide adequate analysis. For each of these figures, the right part contains the detail coefficients sequences which when added to the smooth series s8, generate the reconstruction sequences for each of the 8 levels. For example, in order to obtain the smoothed sequence S6 of the FTSE index in figure 8.13, one just needs to ad to S8 the detail coefficient sequences D8, and D7 i.e., S6 is simply:

$$S6 = S8 + D8 + D7 \qquad (8.1)$$

In figure 8.15 we show how the FTSE ALL SHARE index closing prices are approximated (smoothed) by the S6 smooth level. In figure 8.16 we show the original series overlapped by the S6 sequence for the period 1985-1988 and the level 6 residuals ϵ_6 which are computed as:

$$\epsilon_i = \text{DATA} - \sum_{i=1}^{j} Di \qquad (8.2)$$
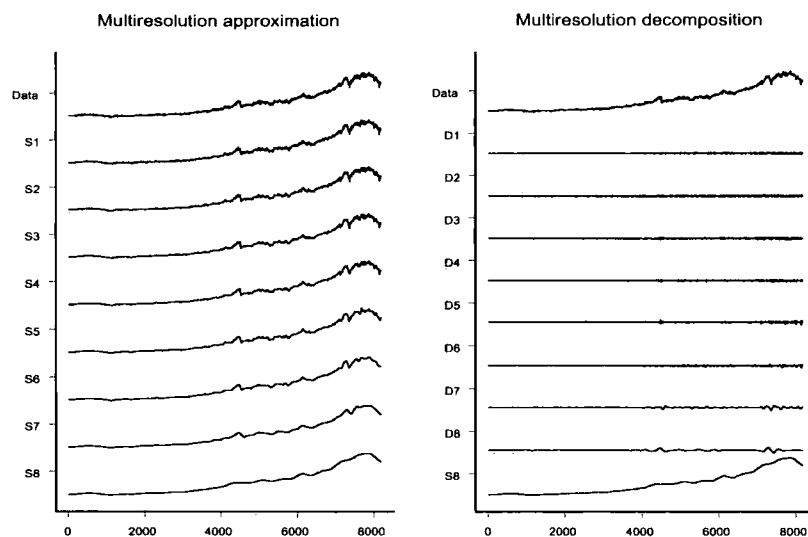
Figure 8.13: FTSE ALL SHARE index closing prices MRA.
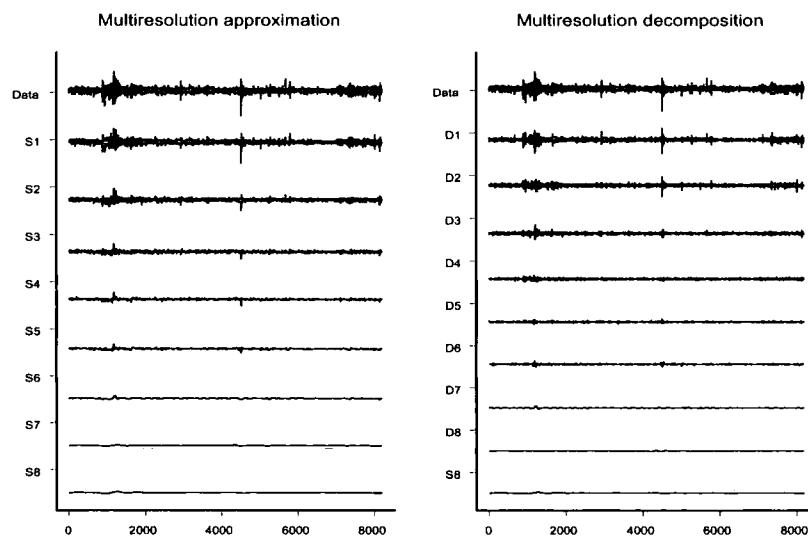
Figure 8.14: FTSE ALL SHARE index logarithmic returns MRA.

FTSE index closing prices

S6 approximation with a Haar wavelet
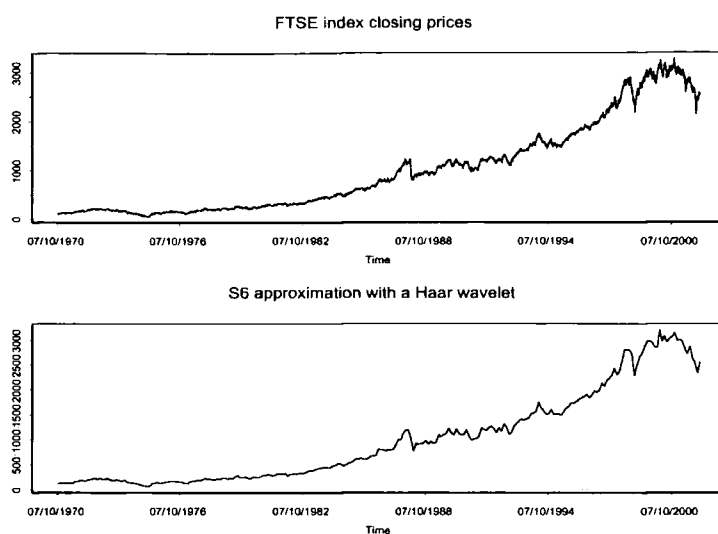
Figure 8.15: The actual FTSE ALL SHARE closing prices and the S6 level smoothed series.

for $j=6$. As we can see, the arbitrarily chosen MRA S6 level of smoothing follows the FTSE very close, especially during the 1987 crash.

The Haar DWT transform coefficients for the first 6 levels, the S6 smooth and the inverse discrete wavelet transform of the FTSE series is depicted in figure 8.17.

The DWT of the FTSE logarithmic returns sequence, computed for 13 levels, is represented in figure 8.18. Again we utilised here a Haar wavelet. In this attempt we have accounted for all the possible discrete wavelet decomposition levels to demonstrate the 2^j order of the DWT algorithm. We should recall here that $8192 = 2^{13}$ which allows for 13 levels of discrete decomposition. In practice, one may choose to concentrate on the first 5 to 7 levels as higher decompositions provide no further information on the variability of the series. In the same figure we may see how the oil crisis and the crash of 1987 have been picked up by the $d1$ to $d4$ coefficients series, seen as negative and positive spikes on those levels.
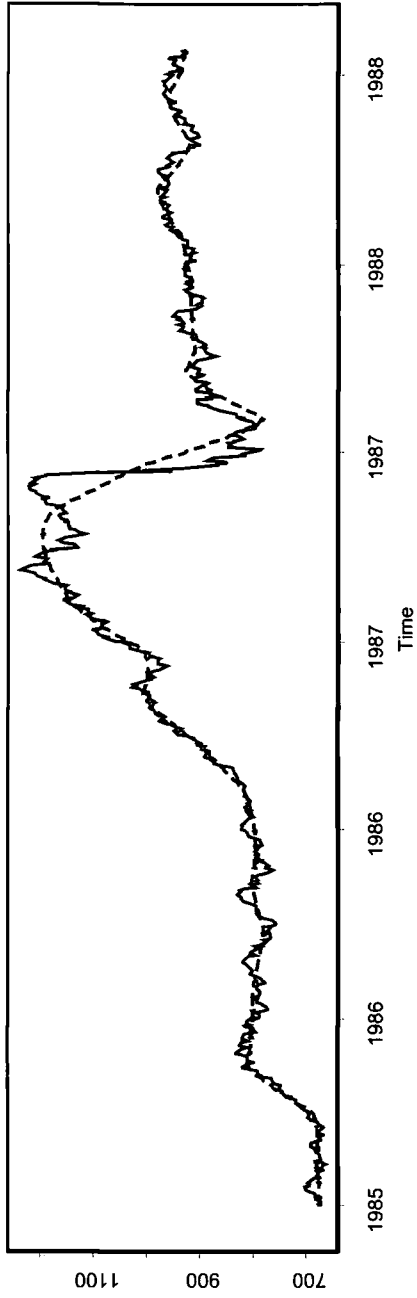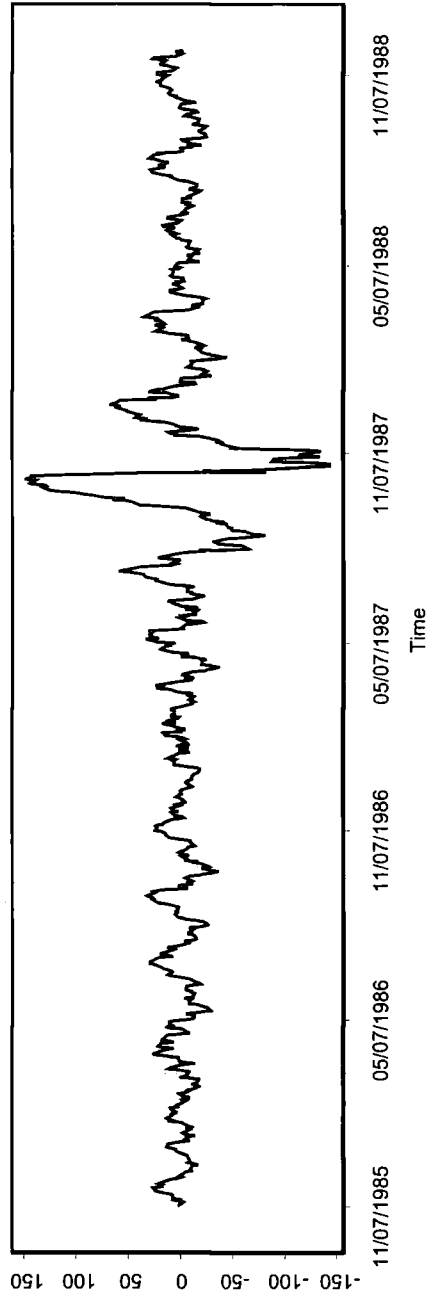
Figure 8.16: A detail of figure 8.15. The 1985-1988 timeline with the crash of '87. Top graph: the original series and the S6 smooth. Bottom graph: the residuals as defined in equation (8.2).
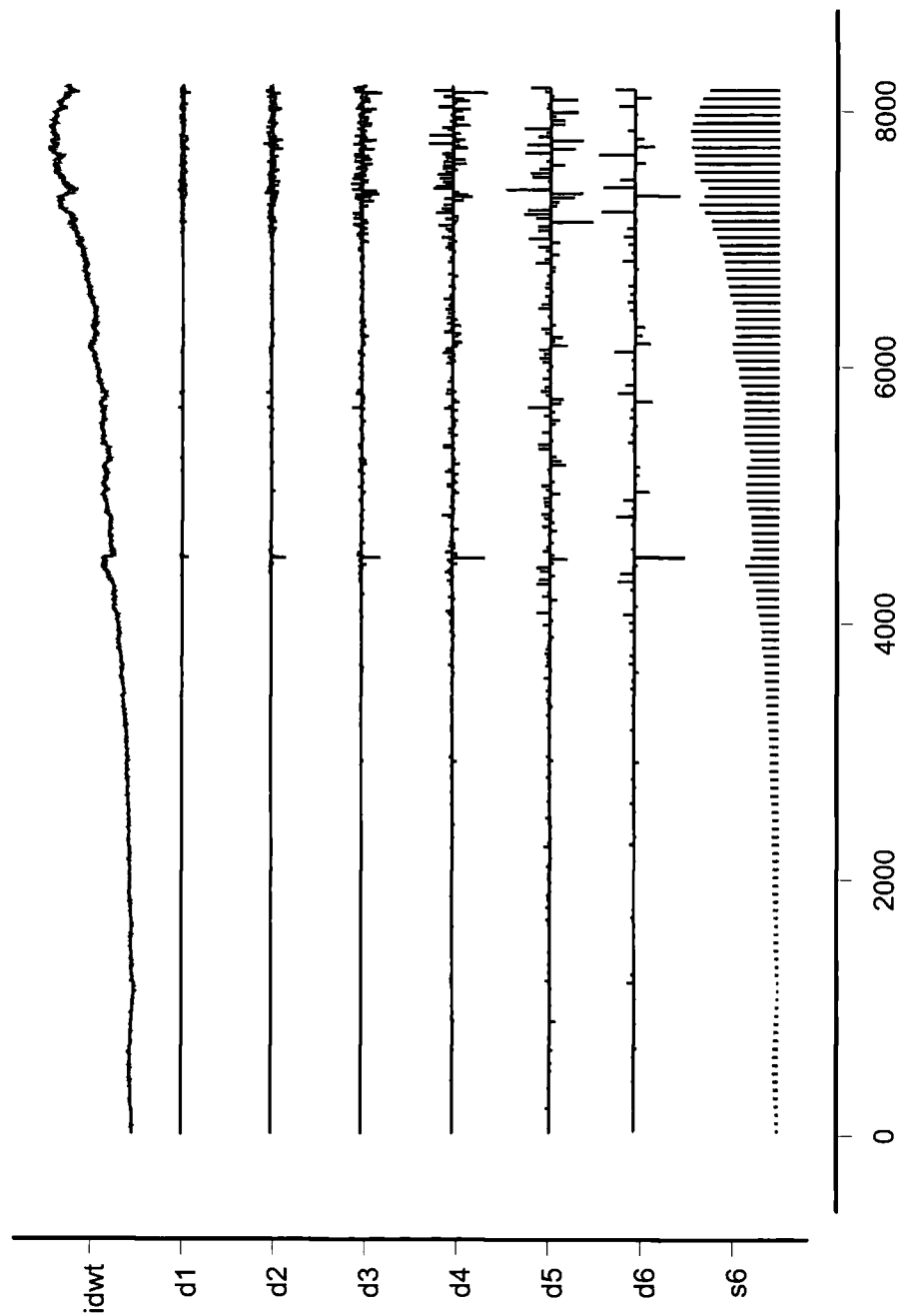
Figure 8.17: The FTSE ALL SHARE Haar inverse discrete wavelet transform. DWT Haar coefficients for 6 levels: $D_1, D_2, ..., D_6$.

FTSE logarithmic returns

Haar DW decomposition on 13 levels
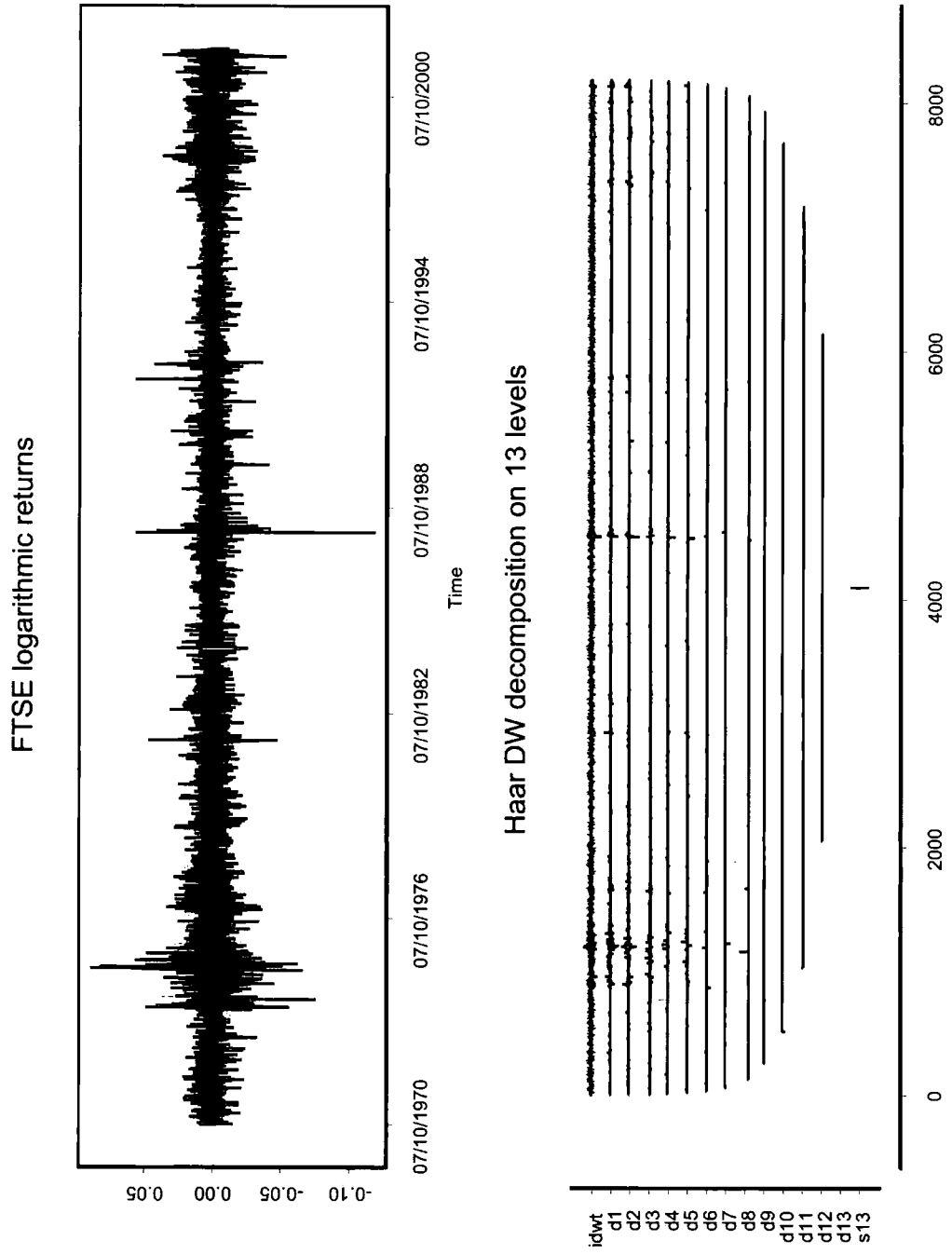
Figure 8.18: The DWT of the FTSE logarithmic returns.

8.3.2 Waveshrinking the FTSE

In this section we use the tools described earlier in chapter 4, to smooth the FTSE ALL SHARE index returns. More precisely, we utilised the discrete wavelet transforms to obtain the wavelet coefficients for each level (scale) of the decompositions. We then used the Waveshrink methodology to retain the most important coefficients. Finally by conducting the inverse wavelet transforms with the retained coefficients, we obtained the smoothed versions of the returns sequences. This analysis can be outlined in the 3 steps:

1. Choosing a suitable mother wavelet function.

2. Denoising the returns using the Waveshrink approach as described in the previous chapter.

3. Obtaining the fitted values and residuals and examining their structure.

The first step refers to the selection of a wavelet function that approximates the series adequately. Suitable candidates should be functions that are short enough to localise on singularities of the returns sequence (such a the 1987 crash). As returns do not exhibit any trend, it is not necessary to use functions that are much "stretched" in time i.e., the horizontal axis. For our analysis purposes we experimented with Daubechies, Haar, Coiflet and Symmlet wavelets. We also examined the possibility of using biorthogonal wavelets. The results were similar. High and low order wavelets were used. In all cases we used the maximum overlap or nondecimated discrete wavelet transform (MODWT). Some wavelets performed better than others. After tedious experimentation and calibration of the Waveshrink denoising approach, the conclusion derived was that the choice of the wavelet function was less important than the Waveshrink functions and rules utilised during the denoising process.

On every step, the normality of the residual noise was tested using the Jarque-Bera and Kolmogorov-Smirnov statistics. The QQplots were also generated and the empirical quantiles of the Waveshrink residuals were checked with those of the standard normal distribution. A straight line would indicate clearly the normality of the residuals. Comparing our results to the ones obtained by a GARCH(1,1)

Variable	ADF test statistic	τ Critical Values
FTSE returns	-35.56652	$\alpha_{1\%} =$ -2.5661
Waveshrink fit	-25.09999	$\alpha_{5\%} =$ -1.9394
Waveshrink residuals	-48.87310	$\alpha_{10\%} =$ -1.6156

Table 8.6: Augmented Dickey-Fuller unit root test results with MacKinnon critical values for rejection of hypothesis of a unit root. It is obvious that even for 1% statistical significance level the Waveshrink residuals are a stationary process.

Parameter	Value
Mother wavelet	Daubechies 6
Shrink rule	top 1900 largest coefficients
Shrink function	hard thresholding
Scale rule	all coefficients considered
Scale function	mean absolute deviation

Table 8.7: The parameters of the Waveshrink denoising process on the FTSE returns using the d6 wavelet.

specification of the series (which is a common approach in the literature) made obvious that the residuals of the GARCH processes do not follow closely the assumed underlying distributions.

The results we demonstrate in figures 8.19-8.22 are the ones that were obtained from the application of the Waveshrink denoising process via a Daubechies 6 (d6) mother wavelet. The elements of this approach are summarised in table 8.7. The best results were obtained by using the "hard" thresholding function and by keeping the top 1900 wavelet coefficients in size. These were capturing the energy of the FTSE returns signal sufficiently. The actual, fitted and residual processes are displayed in figures 8.19-8.20. In figure 8.20 we can see better the result of the denoising process on the FTSE returns sequence.

Figure (8.21) compares the distributions of the actual returns and the fitted (denoised) series. It is obvious that the "fat-tailed" structure has remained as a key element of the underlying process, though the variance of the series has been reduced considerably due to the omission of the process captured by the residuals. In table 8.8 we display the descriptive statistics for the fitted and residual sequences.

Statistic	fitted	residuals
Mean	0.000367001895299	2.21030297151e-08
Median	0.000592908206597	-1.7054051905e-05
Maximum	0.0804649099356	0.027347065248
Minimum	-0.119011327871	-0.0269594711128
Standard dev.	0.0057592733608	0.00735614022242
Skewness	-1.35125715114	-0.0247104586328
Kurtosis	78.3633836757	3.10564694301

Table 8.8: The descriptive statistics of the Waveshrink obtained proces-
ses.
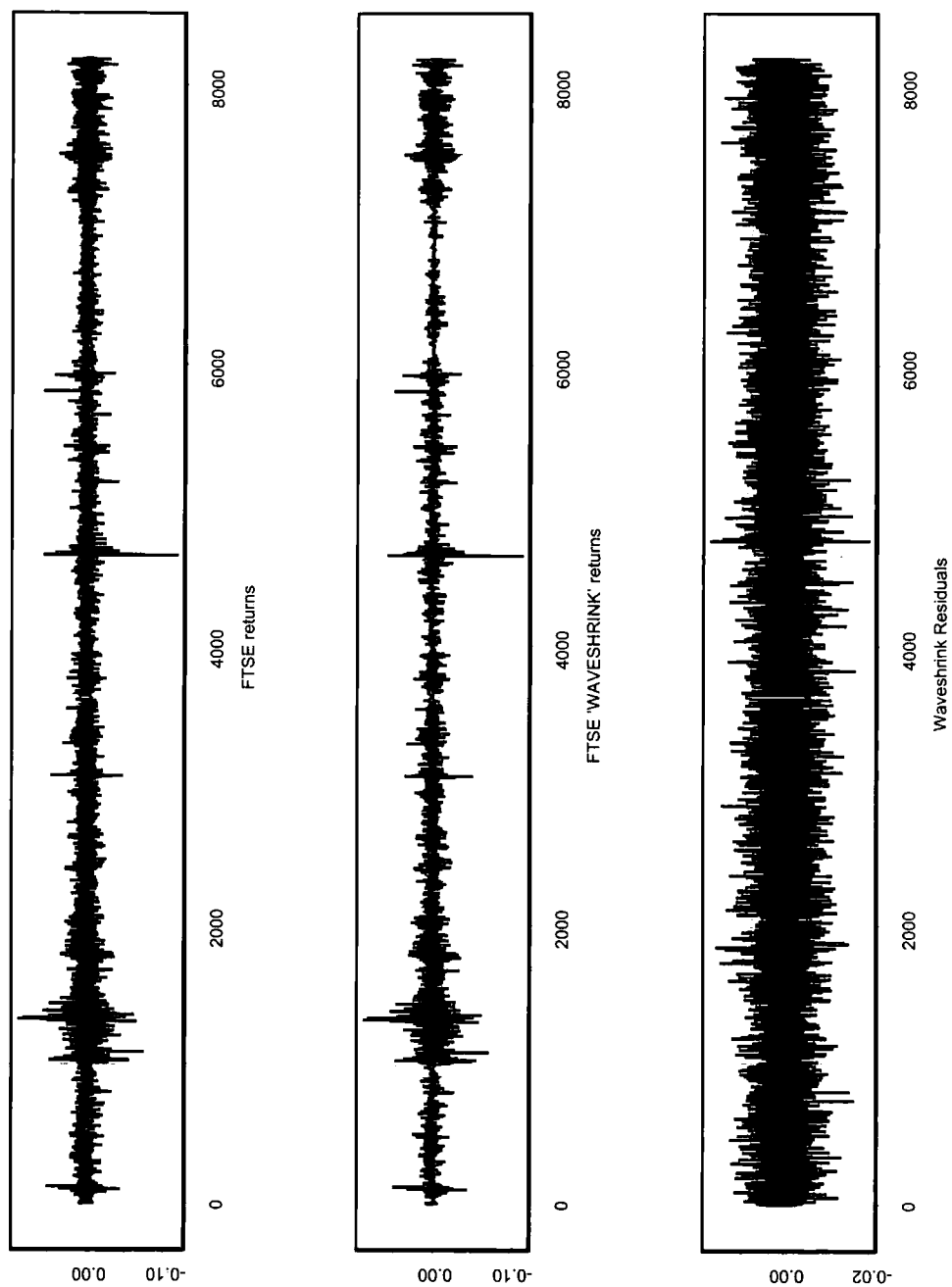
Figure 8.19: The results of Waveshrink denoising algorithm. The FTSE returns sequences, the Waveshrink fitted values and the Waveshrink residuals.
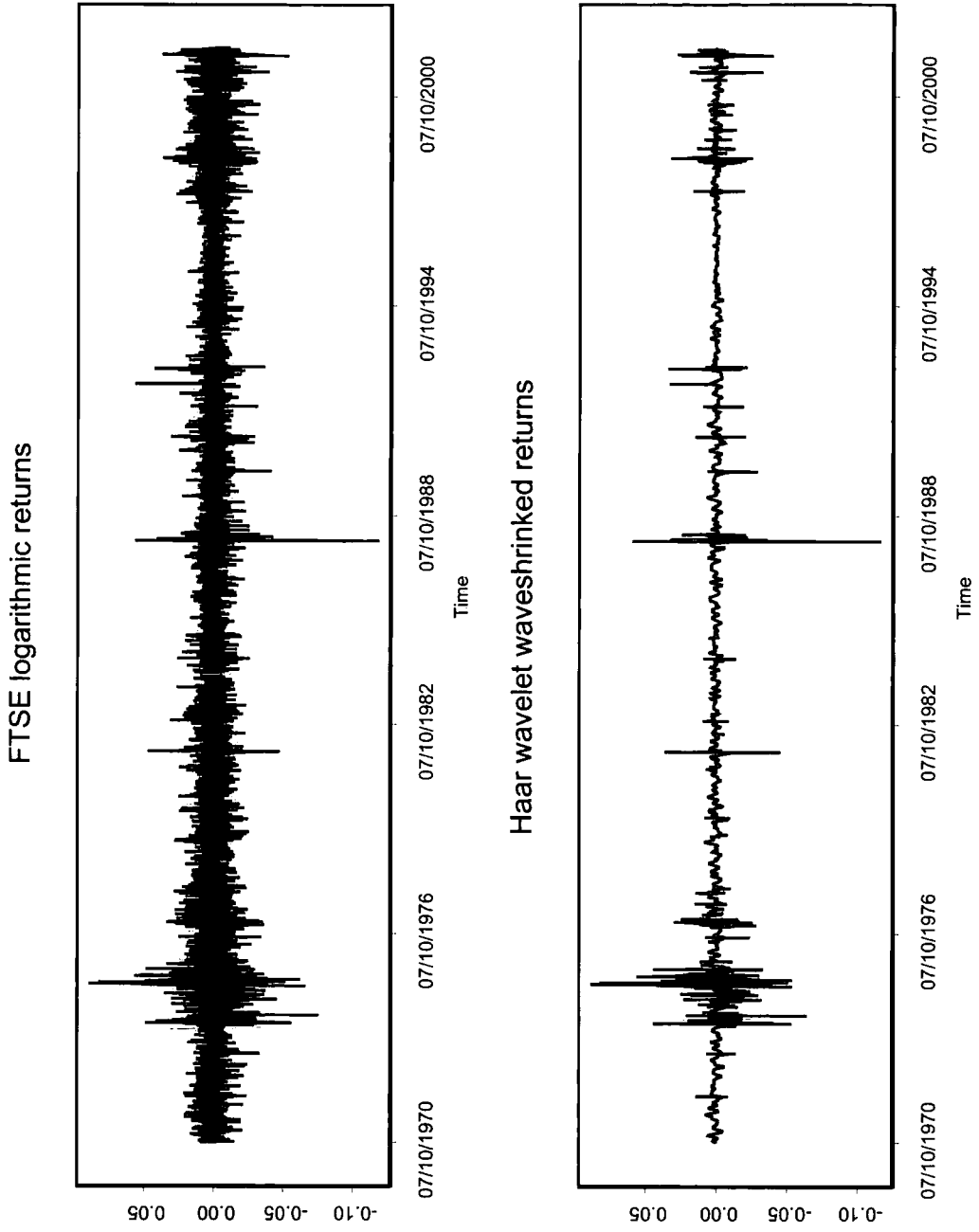
Figure 8.20: A closeup of the series in figure (8.19) showing better the structure of the Waveshrink fit.

FTSE logarithmic returns

Haar wavelet waveshrinked returns

FTSE returns histogram

Waveshrinked returns histogram
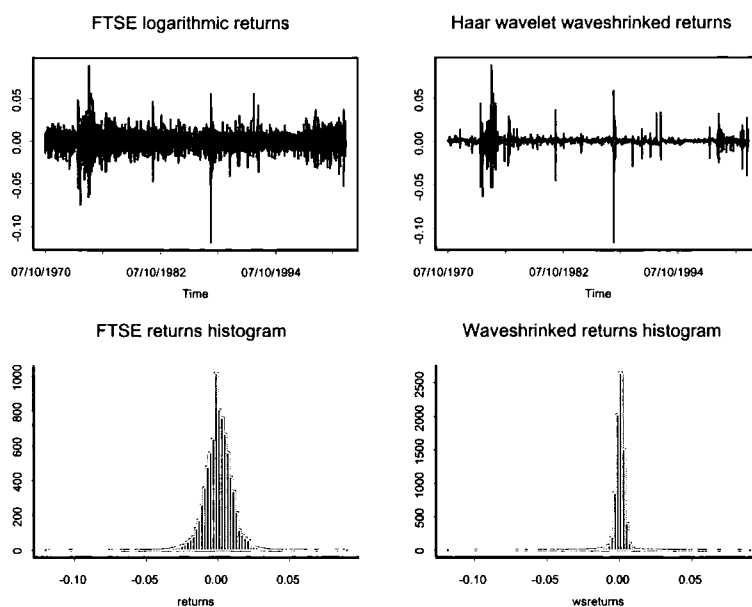
Figure 8.21: A comparison of the distributions of the FTSE returns and the Waveshrink fit.

In table 8.6 we report the results of stationarity test via the Augmented Dickey-Fuller (Dickey and Fuller 1979) procedure for various levels of statistical significance. It is straightforward the processes are I(0) or stationary from the ADF statistic values obtained even at a 1% level of significance. Moreover we have conducted the Jarque-Bera (Bera and Jarque, 1981) and Kolmogorov-Smirnov one sample (Conover, 1971), tests to establish the normality of the Waveshrink obtained residuals. These results are listed in tables 8.11 and 8.10 respectively. The Jarque-Bera statistic is distributed as a χ^2 with 2 degrees of freedom and the critical value for $\alpha=5\%$ level of statistical significance is 5.991. That leads us to accept normality for the residuals.[88] The very low p-value leads us to accept again normality for the Waveshrink residuals. Conducting a RUNS[89] test for independence for the Waveshrink residuals returned a value for the statistic of -0.6661 and a p-value of 0.50. Because the statistic is less than the critical value 1.96 for $\alpha=5\%$, the null hypothesis of independence is accepted. The opposite was shown for the fitted values with a RUNS

---

[88]We are testing the null hypothesis that the distribution of the residuals is the normal.

[89]The RUNS statistic for independence is approximately distributed as a N(0,1).
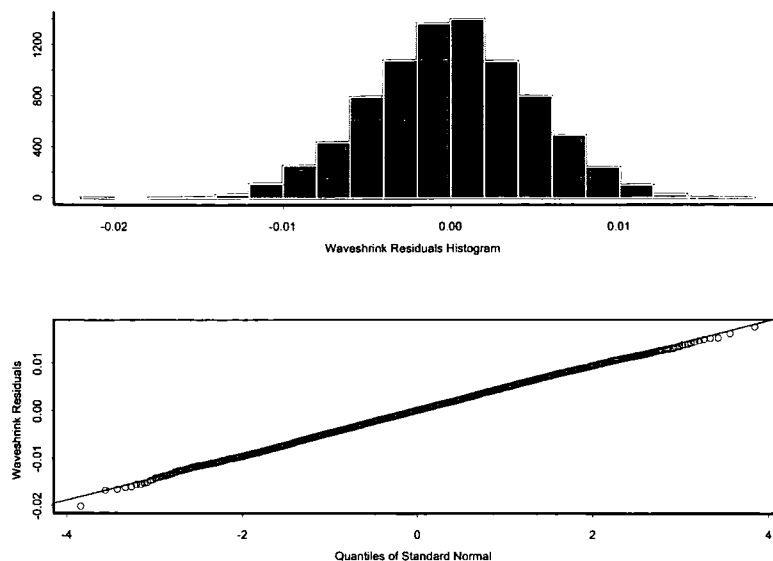
Figure 8.22: Assessing the normality of the Waveshrink residuals.

statistic of -5.9457.

In table 8.9 we have tabulated the first 20 values of the autocorrelation (ACF) and partial autocorrelation(PACF) functions for the Waveshrink fitted values and residuals with the corresponding Q-statistic values. As we can see, the values of the coefficients are *extremely small* and thus can be regarded as negligible quantities.[90] These functions are depicted in figure 8.23.

8.4 Nonlinear Determinism

Following the denoising of the returns series, we applied the theory and methodologies described in chapters 4 to 6 in order to establish whether the sequence exhibited nonlinear determinism or not. Our main rationale here being that wavelets would have removed the nonsystematic components of the variability of the series, leaving a clear data generating process structure. If indeed the FTSE returns series was contaminated by a normally distributed noise process that masked the dynamic structure of the underlying generating process, the inverse discrete wavelet

---

[90]From a personal communication with Prof. Chris Chatfield.

lag	Waveshrink Residuals			Waveshrink fitted values		
	ACF	PACF	Q-STAT	ACF	PACF	Q-STAT
1	0.008	0.008	0.438	0.299	0.299	659.520
2	-0.095	-0.095	66.272	0.152	0.069	830.210
3	-0.096	-0.095	133.990	0.248	0.204	1282.500
4	-0.083	-0.093	184.400	0.241	0.130	1710.100
5	-0.066	-0.087	216.220	0.157	0.038	1892.300
6	-0.070	-0.102	252.070	0.135	0.030	2026.200
7	-0.097	-0.140	321.040	0.186	0.084	2282.100
8	-0.027	-0.083	326.310	0.116	-0.014	2380.500
9	-0.012	-0.084	327.300	0.124	0.047	2493.500
10	0.013	-0.062	328.630	0.176	0.081	2722.800
11	-0.004	-0.079	328.750	0.150	0.036	2887.800
12	-0.024	-0.097	332.890	0.092	-0.010	2950.100
13	-0.009	-0.085	333.430	0.094	0.000	3015.700
14	-0.028	-0.115	339.300	0.115	0.013	3112.500
15	0.017	-0.070	341.470	0.035	-0.062	3121.300
16	0.004	-0.089	341.570	0.021	-0.031	3124.500
17	0.002	-0.089	341.590	-0.031	-0.104	3131.500
18	0.003	-0.092	341.670	0.001	-0.009	3131.600
19	0.013	-0.086	342.850	0.022	0.012	3135.100
20	0.035	-0.060	352.120	0.040	0.040	3146.700

Table 8.9: The first 20 values of the autocorrelation (ACF) and partial autocorrelation(PACF) functions for the Waveshrink fitted values and residuals with the corresponding Q-statistic values.

Series	K-S statistic	p-value
Waveshrink residuals	0.0121	0.0091
Waveshrink fit	0.1939	0.0000

Table 8.10: Kolmogorov-Smirnov (K-S) test of composite normality.

Series	J-B statistic	p-value
Waveshrink residuals	4.6702	0.0968
Waveshrink fit	1941626	0.0000

Table 8.11: Jarque-Bera (J-B) test of normality. The test is distributed as a χ^2 distribution with 2 degrees of freedom test. The critical value for $\alpha = 5\%$ is 5.991.

ACF of Waveshrink residuals

PACF of Waveshrink residuals
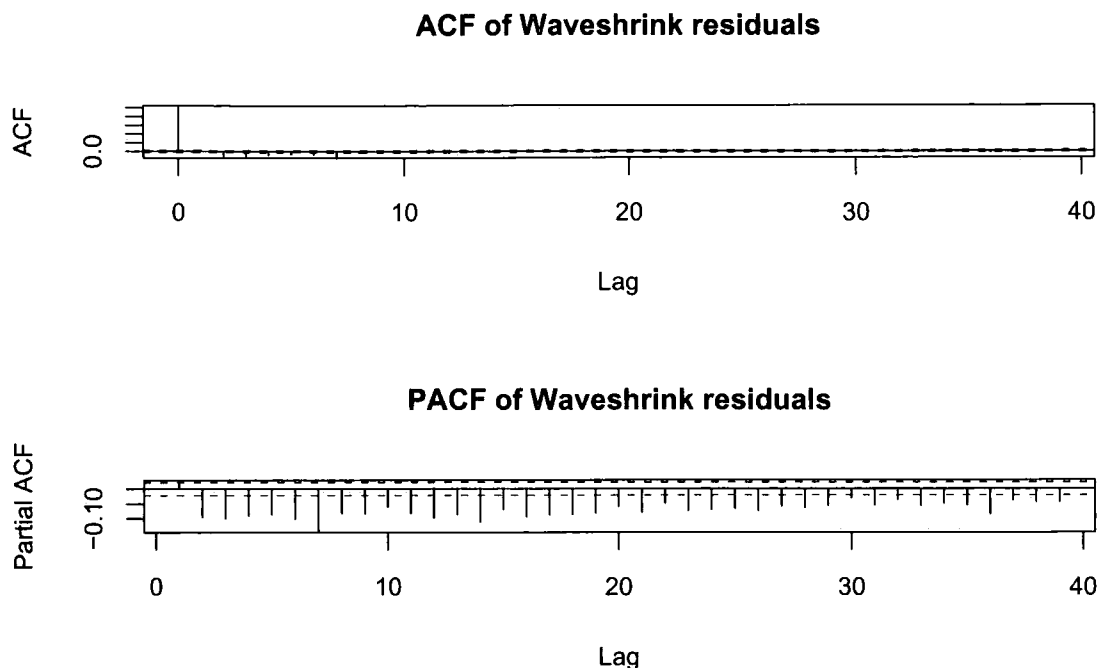
Figure 8.23: ACF and PACF functions of the Waveshrink residuals.

transform would provide a "cleaner" representation of this structure.

As we have already discussed in previous chapters, we have to be extremely careful when preprocessing time series prior to attractor reconstruction calculations. This is called "bleaching" (Theiler and Eubank, 1993). Abarbanel (1995) suggests the use of FIR (Finite Impulse Response filters) for this kind of manipulations. In general, one should transform the series via a smooth monotonic function i.e., a function that is continuous (smooth dynamics) and invertible. That will ensure that certain topological conditions are met for the dynamics (Takens, 1981) i.e., the attractor structure has not been distorted from the transformation. Wavelet transforms qualify for this only in their discrete case (DWT).[91] What we follow here is a complete investigation of the dynamics of the Waveshrink filtered series.

We reconstructed the dynamics of the phase space via delay coordinate em-

---

[91]From a personal communication with Prof. James B. Ramsey: "The discrete wavelet transform is a one to one onto mapping, but not the continuous (CWT) or the various oversampled versions (MODWT). You can not really talk about a monotonic function, except in exceptional circumstances; and the one-to-one'ness (sic) holds only for the full set of coefficients, not any approximation that may be involved. Tests for nonlinearity should be easily incorporated in the wavelet framework; a more precise answer depends on the wavelet transformation used and the nature of the nonlinearity to be detected."

LAG	RETURNS	FIT	RESIDUALS
Entropy	1.15146500	0.35874360	2.19788100
2	0.03936236	0.14599210	**0.01608972**
3	0.03003536	0.10321960	0.02195078
4	**0.02377835**	0.07943605	0.02113038
5	0.02662920	0.06850084	0.01860383
6	0.02332719	0.06125449	0.01745880
7	0.02598398	0.05747043	0.01962191
8	0.02235857	0.05159686	0.01714247
9	0.02376081	0.04784449	0.01734900
10	0.02548636	0.04632784	0.01672281
11	0.02215267	0.04012481	0.01860825
12	0.02033222	**0.03733414**	0.01474768
13	0.02081499	0.03750026	0.01481370
14	0.01942279	0.03720081	0.01346067
15	0.01808479	0.03443327	0.01407659

Table 8.12: **First 15 values of the Average Mutual Information (AMI) functions (see figure 8.24) of the actual FTSE returns, Waveshrink fits and Waveshrink residual sequences. Bold numbers indicate the 1st local minimums. An interesting finding is that both the returns and residual sequences exhibit a local minimum as well where the 1st minimum of the Waveshrink fit is ($\tau = 12$).**

bedding. in table 8.12 we display the results of the computation of the Average Mutual Information criterion (AMI). The lag order for the time delay τ (i.e., the first minimum of the AMI function) for the returns, the Waveshrink fitted returns and the Waveshrink residuals were respectively 4, 12 and 2 as indicated by the bold numbers in table 8.12. In figure 8.24 we display the three AMI functions with arrows that indicate the locality of the 1st minimum.

Using the time delays τ in table 8.12, we calculated the embedding dimensions for the 3 sequences via the False Nearest Neighbors (FNN) methodology, as this is described in chapters 4 and 5. The results for the FTSE returns, Waveshrink fitted values and residuals are tabulated in tables 8.14, 8.15 and 8.13 respectively.

An interesting finding is the graceful convergence of both the AMI and FNN criteria functions for the Waveshrink fitted values (table 8.15) and figure 8.25. In the following sections we construct diagrams in order to visualise more clearly the
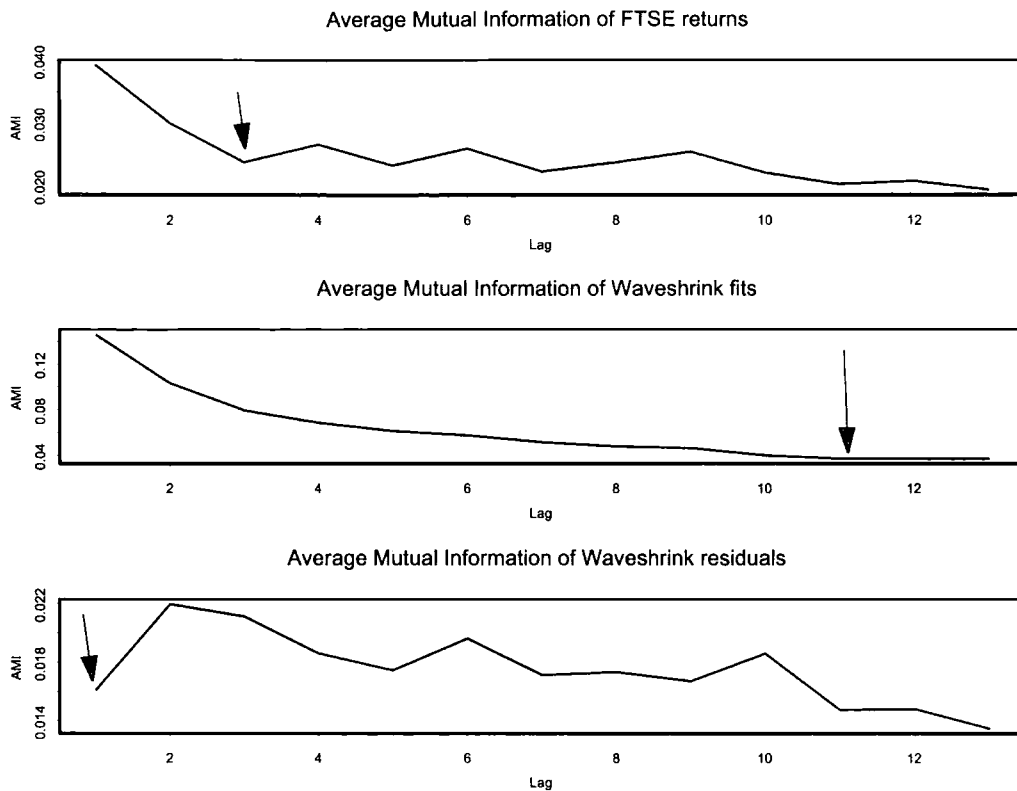
Figure 8.24: The Average Mutual Information functions of the three sequences in table (8.12), 2nd-15th lag. Arrows indicate where first minimum is located.

Embedding Dimension	Fraction of FNN	Average size of neighborhood	Average squared size of neighborhood
1	9.976801e-01	5.772006e-05	6.860344e-08
2	8.809821e-01	2.872275e-03	1.300435e-05
3	6.441538e-01	8.508392e-03	8.085705e-05
4	5.319444e-01	1.059754e-02	1.174292e-04
5	4.814815e-01	1.118932e-02	1.277096e-04

Table 8.13: FNN algorithm results for the Waveshrink residuals.

Embedding Dimension	Fraction of FNN	Average size of neighborhood	Average squared size of neighborhood
1	9.974302e-01	2.857821e-05	2.667477e-08
2	8.738098e-01	8.947465e-04	1.396515e-06
3	5.888966e-01	2.701687e-03	8.370320e-06
4	4.402939e-01	3.669201e-03	1.416172e-05
5	4.242424e-01	4.066201e-03	1.697205e-05
6	4.000000e-01	4.021737e-03	1.678563e-05

Table 8.14: FNN algorithm results for the FTSE returns.
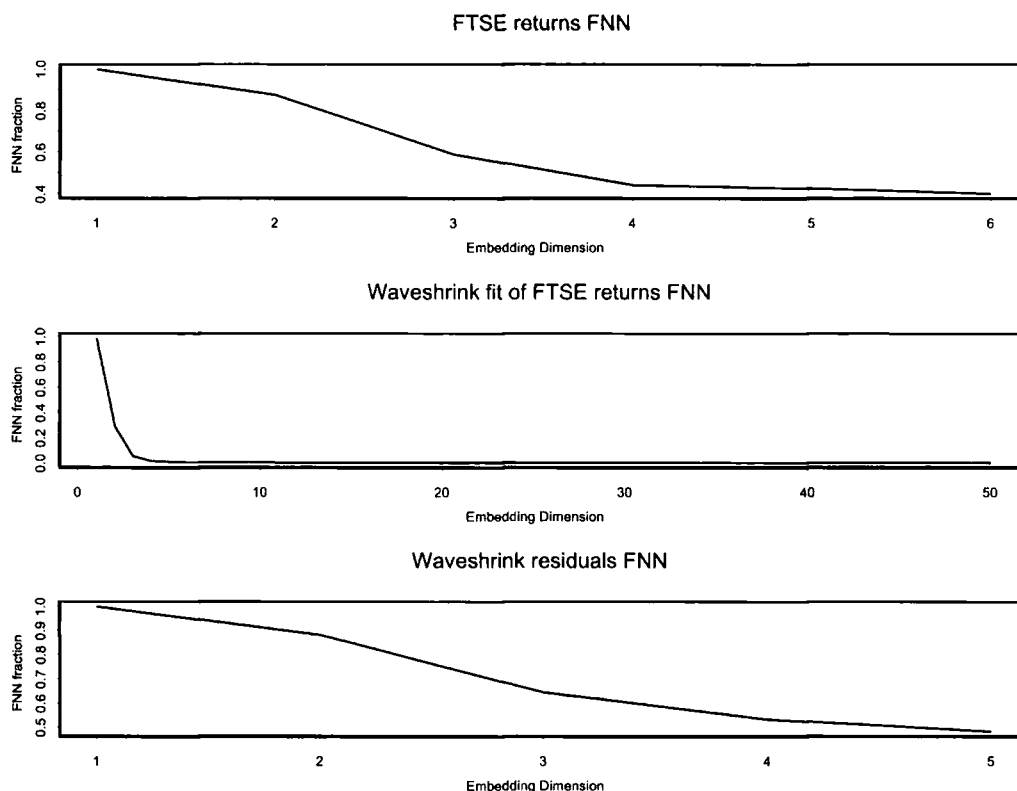
Figure 8.25: Embedding dimensions according to the FNN criterion as in tables (8.14), (8.13) and (8.15). It is obvious that the embedding dimension for the returns and residuals sequences is low whereas the Waveshrink obtained fitted values have a much larger embedding dimension.

Embedding Dimension	Fraction of FNN	Average size of neighborhood	Average squared size of neighborhood
1	9.850582e-01	2.042813e-05	1.500015e-08
2	2.896907e-01	3.252177e-04	2.764600e-07
3	5.093644e-02	9.384883e-04	1.181869e-06
4	1.486000e-02	1.396898e-03	2.285555e-06
5	6.946857e-03	1.542079e-03	2.702321e-06
6	4.324121e-03	1.607818e-03	2.894843e-06
7	4.419446e-03	1.658367e-03	3.047319e-06
8	4.040834e-03	1.694972e-03	3.163020e-06
9	4.270623e-03	1.728534e-03	3.271531e-06
10	3.789673e-03	1.757896e-03	3.369958e-06
11	2.486325e-03	1.784847e-03	3.461092e-06
12	2.610966e-03	1.806778e-03	3.533881e-06
13	2.464403e-03	1.826278e-03	3.599295e-06
14	1.438021e-03	1.841930e-03	3.654511e-06
15	1.511031e-03	1.854205e-03	3.699481e-06
16	1.586798e-03	1.863491e-03	3.734467e-06
17	1.667222e-03	1.873007e-03	3.771303e-06
18	1.751313e-03	1.881180e-03	3.803668e-06
19	1.832845e-03	1.890046e-03	3.838984e-06
20	1.149425e-03	1.896096e-03	3.861257e-06
21	1.197605e-03	1.901297e-03	3.879315e-06
22	1.247920e-03	1.907320e-03	3.900388e-06
23	1.303215e-03	1.912803e-03	3.919247e-06
24	1.358081e-03	1.917058e-03	3.933527e-06
25	1.415762e-03	1.918389e-03	3.937148e-06
26	1.477105e-03	1.919160e-03	3.938654e-06
27	1.540041e-03	1.917816e-03	3.931240e-06
28	1.605996e-03	1.914844e-03	3.916603e-06
29	1.675042e-03	1.912064e-03	3.902694e-06
30	1.745201e-03	1.910670e-03	3.894237e-06
31	1.821494e-03	1.907147e-03	3.877712e-06
32	1.904762e-03	1.901691e-03	3.852412e-06
33	1.989390e-03	1.896082e-03	3.827146e-06
34	2.073255e-03	1.891692e-03	3.806227e-06
35	2.869440e-03	1.889659e-03	3.795823e-06
36	1.484781e-03	1.889323e-03	3.791188e-06
37	1.536098e-03	1.889931e-03	3.790570e-06
38	1.592357e-03	1.889186e-03	3.783675e-06
39	1.643385e-03	1.891993e-03	3.791525e-06
40	0.000000e+00	1.894156e-03	3.797516e-06

Table 8.15: FNN algorithm results for the Waveshrink fitted values.

dynamics of the series we produced via the Waveshrink algorithm. We will be using delay, phase and recurrence plots.

8.4.1 Exploring dynamics via lag and phase plots.

A "lag plot" or "time-delay plot" is simply a scatter-plot of the series against its lags. These are very common plots in physical sciences and show the time evolution of certain phenomena revealing periodic behaviours. What actually we obtain from these is a 2-dimensional diagram of the orbits of any existing attractors. A delay plot can be extended to 3 dimensions if this provides meaningful information. A phase-plot is a scatter diagram of the series against its derivatives. In our case, a phase-plot would be the scatter-plot of the FTSE against its returns or the fitted values.

We used the time delays τ and the embedding dimensions to generate the recurrence plots for the three sequences, returns, fitted values and residuals. We also generated the phase-plots of the fitted returns sequences. In figure 8.26 (a) we have the delay plot the first 100 observations of the Waveshrink fitted returns. The lag order for this scatter diagram is the same as the time-delay we determined: $\tau = 12$.

In figure 8.26 (b) the first 100 observations of the FTSE residuals, the Waveshrink fit and residual sequences are displayed for the better comprehension of figure 8.26 (a). with a lag order of 12. The fitted sequence is replete with such patterns indicating clearly that an attractor or cycle of some certain structure exists. The patterns indicated here are distorted during periods of intense market volatility. This fact is confirmed by the recurrence plots in figure 8.41. In order to visualise the dynamics around a period of strong fluctuations, we focused on the 1987 market crash.

In figures 8.27 and 8.28 we have included the delay plots of the periods a year before the 1987 crash and exactly during the crash respectively. It is obvious that the financial shocks of 1987 have caused a serious perturbations of the market dynamics which are evident in the structure of the orbits that the FTSE returns were following. The whole 8192 observations sequence of the Waveshrink fitted

(a) Delay plot
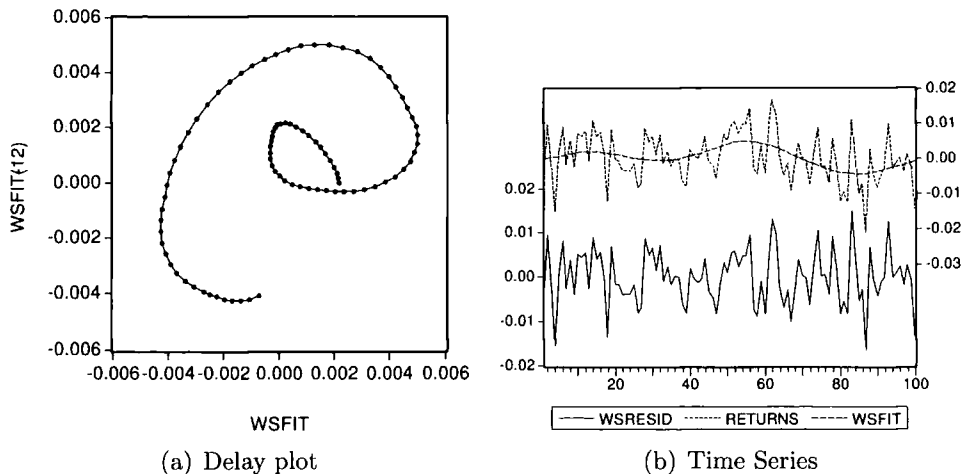
(b) Time Series

Figure 8.26: Subfigure (a): Phase diagram of the first 100 days of the Waveshrink fitted values. Subfigure (b): First 100 observations, time series plots. Top: FTSE returns sequence with the Waveshrink fitted values. Bottom: Waveshrink residuals.

FTSE returns is replete with such patterns that interrupt when there is a significant shock. We found that during the period from 1/1/1988 to 1/8/1988, the dynamics show evidence of recovering to a smoother structures as before the shock (in figure 8.29).

In figure 8.30, we chose to project a history of the orbit of the fitted returns. We collected a daily sequence of the Waveshrink fitted observations, spanning 2 years before the market crash. The dynamics, as these are projected in 3-dimensional space, exhibit the smoothness and periodicity that we detected earlier from the 2-dimensional delay plots. As we collect more and more observations, the dynamics become less clear, as subsequent volatile periods generate outliers that mask the smoothness of the dynamics. This is evident from the examination of the delay plot in figure 8.31. These periods can also be discerned in the corresponding recurrence plots in figure 8.41 as wide blank gaps between the recurrences. An interesting research question here would be to examine the dynamic structure of these identified periods where the "smooth" functioning of the markets breaks down.

In order to examine how the orbits unfold for different delay windows, we generated a sequence of delay plots in figures 8.33 and 8.34, for the first 100 and
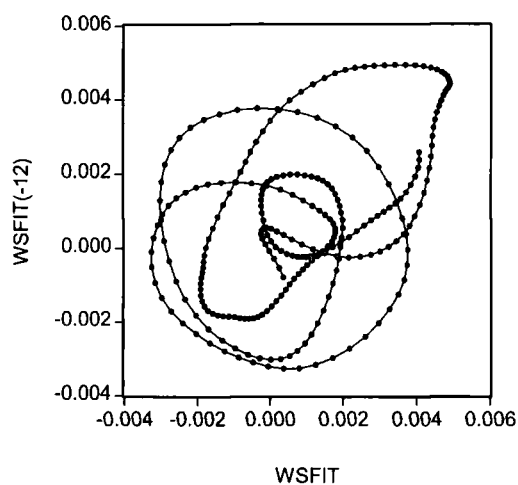
Figure 8.27: Delay plot showing the cycle the FTSE index returns were following just a year before the crash of 1987.
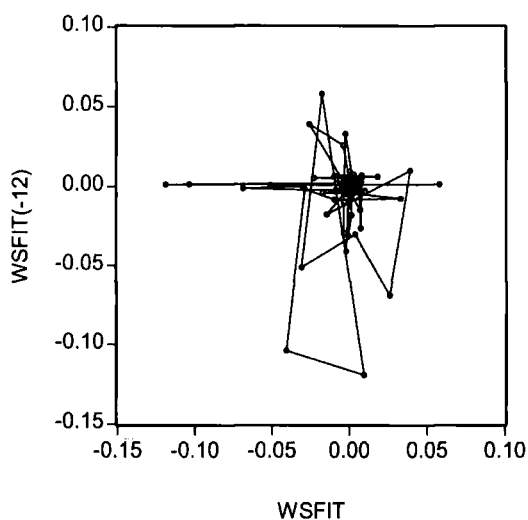
Figure 8.28: Delay plot showing the effect of the crash of 1987 on the FTSE returns dynamics. It is obvious that the presence of outliers have concealed the smoothness of the dynamics that we could discern in figure (8.27). It was found that removing the outliers could not recover the smoothness of the dynamics.
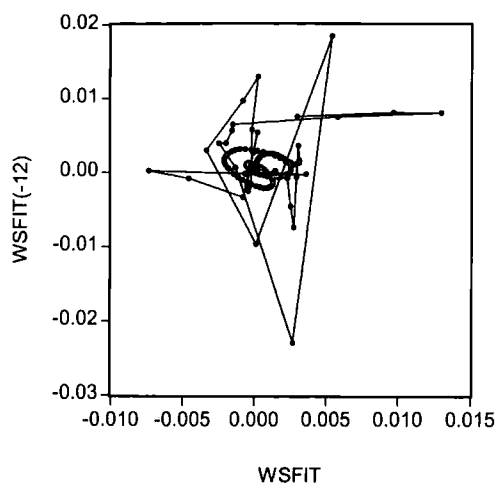
Figure 8.29: Evidence of recovering of the cycle in the period 01/01/1988 01/08/1988 (in the centre).
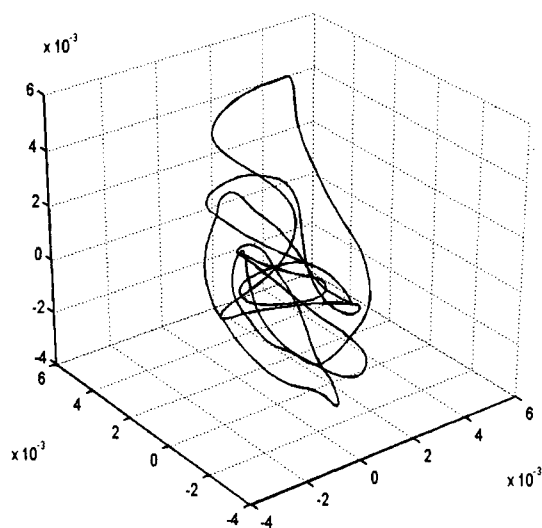
Figure 8.30: 3D delay plot showing the cycle the FTSE index returns were following 2 years before the crash of 1987.
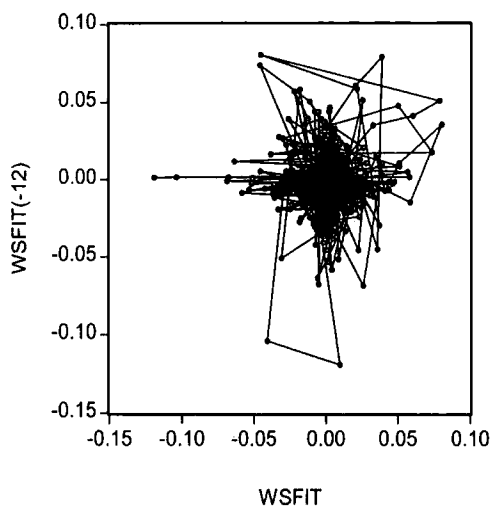
Figure 8.31: The delay plot of the FTSE Waveshrink fit returns for the whole sample period 10/7/1970 to 30/11/2001.

200 days respectively. These sets of "delay-cartoons" can show us how the orbits develop as well as the effect of the selection of the correct delay time τ for examining the dynamics. For $\tau=1$, the delay plot is almost like a straight line which indicates 1st order serial correlation. As the delay parameter becomes larger, the delay plots acquire a cyclical pattern which becomes even more distorted as the τ value becomes larger. The orbits of the attractor unfold more gracefully for values closely around the correct time delay. In figure 8.34, we can see how the smoothness in figure 8.33 is being distorted with the inclusion of more observations which introduce volatility. Interesting patterns can be revealed for the original sequence of the FTSE ALL SHARE index closing prices. We calculated via AMI a very large time-delay factor of a value of 87. We generated the delay plot for the whole sequence in figure 8.35. There we can see clearly a certain time-dependence in the closing prices which becomes more and more diffused for larger closing prices of the index i.e., smaller values seem to have stronger time-dependence than larger ones.

The important message from the delay plots in figures 8.26-8.34 is that given we have faithfully reconstructed the dynamics by careful delay coordinate embedding, we can use simple scatter diagrams to view these and discover certain a-periodic characteristics. The very interesting point here is that we have to be very careful

(a) Delay plot   (b) Delay plot
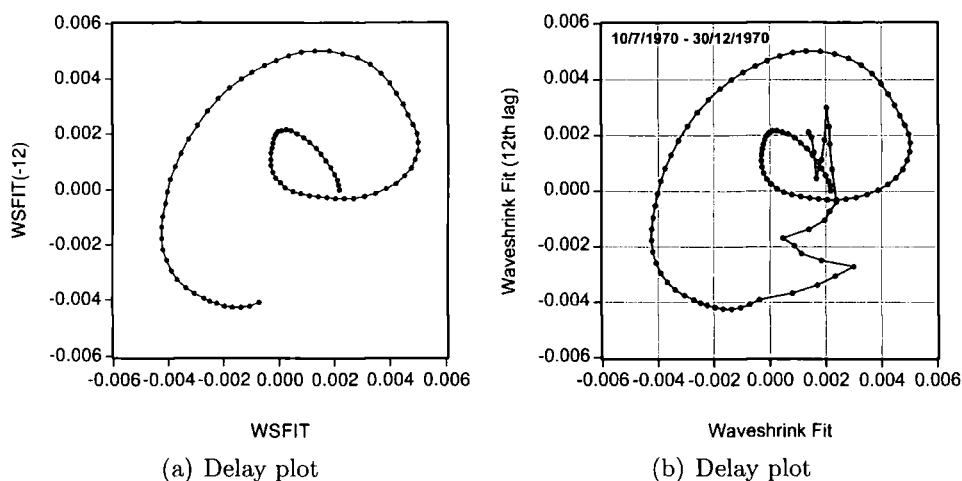
Figure 8.32: How the cycle is distorted.

when examining the finer structure of these dynamics. If we just generate a delay plot for the complete sequence, this will probably be very fuzzy or incomprehensive. The scatter diagrams will be crammed with swarms of points which reveal no sensible pattern (depicted clearly in figure 8.31). In figure 8.32 we can clearly see how this can become a problem. Subfigure (a) is the delay plot corresponding to the first hundred observations of the fitted returns. Subfigure (b) is the same as (a) with the inclusion of a few more observations (spanning precisely the period between 10/7/1970 and 30/12/1970). We can see clearly how the orbit becomes more jerky for the new information.

Past research in the field has failed to reveal interesting patterns because of this fact. If instead one "zooms" into sub-periods, one can discover very interesting details. The conclusion would be that while increasing our sample size, there seems to be a tendency of the attractor to shift, expand, contract and change periodicity as new information comes in. This may well be the effect of the change of structure introduced by increasing financial volatility since the 70s.

Delay plots are not the only tool to visualise dynamics. We can also generate phase plots. In figure 8.36 we generate phase plots for the same period between 10/7/1970 and 30/12/1970, for the FTSE returns (a) and the Waveshrinked fitted values (b) respectively. We can clearly see that the phase plot (b) of the fitted values is much more informative and that there is again evidence of cyclicity. We can also

Waveshrink fit, lagged scatterplots for first 100 days
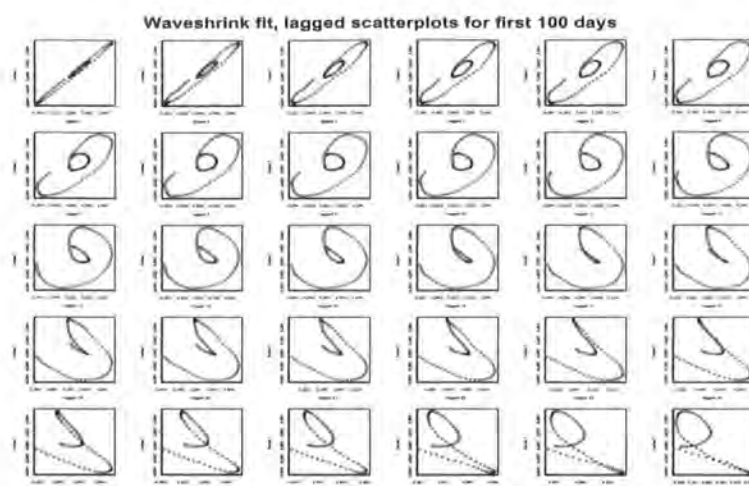
Figure 8.33: Delay plots of the first 100 days. Delay τ=1 to 40.

Waveshrink fit, lagged scatterplots, first 200 days
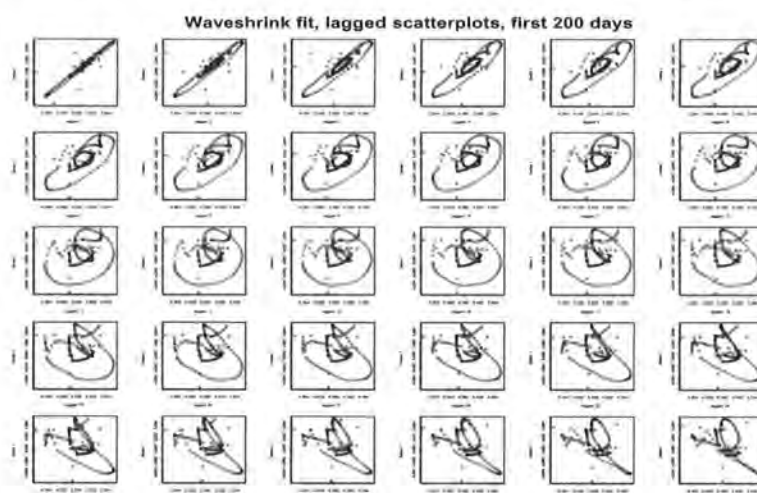
Figure 8.34: Delay plots of the first 200 days. Delay τ=1 to 40.

see how this is distorted exactly as in figure 8.32.

Finally we can retrieve more interesting visual information from the simple scatter plots of the FTSE closing prices against the returns and the Waveshrink fit, as in figures 8.37 and 8.38 respectively. Here we attempt to generate a phase plot with a difference. The scatter diagram in figure 8.37 is indeed a phase plot as defined earlier and covers the whole period. Instead of plotting the returns against their first derivative (1st difference), we choose to plot the actual closing prices of the
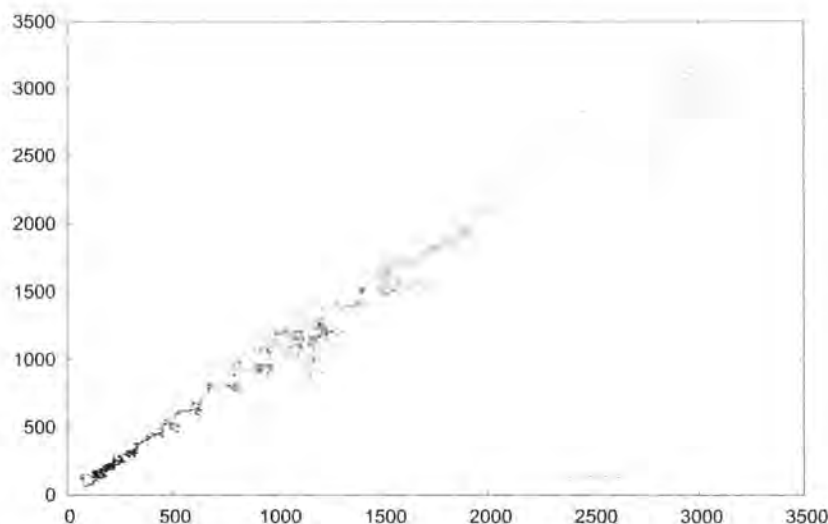
Figure 8.35: Delay plot: FTSE ALL SHARE INDEX with delay $\tau=87$.

(a) Returns
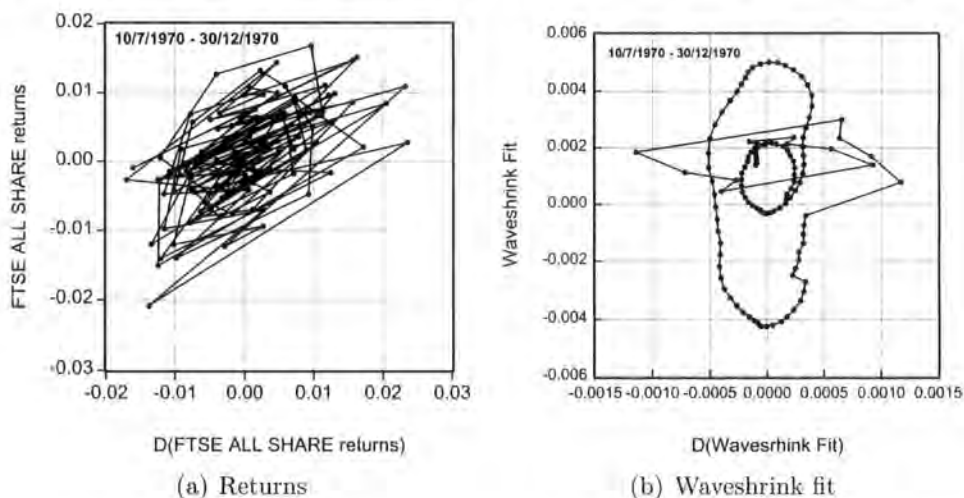
(b) Waveshrink fit

Figure 8.36: 2-Dimensional phase-plot of the FTSE returns and the Waveshrink fit.

index against the 1st derivative which is simply the returns themselves. We can see no clear structure there. If one though replaces the returns with the fitted values, as these have been "cleaned" by Waveshrink, we obtain a much more informative phase plot in figure 8.38. We can clearly see a pattern revealing itself as some kind of a *vortex* or *spiral orbit*. Some parts of the dynamics are though distorted. It remains to be seen if one can improve that image by better filtering of noise in the original returns sequences. What we clearly see though is that there exists some

Figure 8.37: Scatter (phase) plot: **FTSE ALLSHARE INDEX** against **FTSE logarithmic returns.**

structure in the dynamics of the FTSE closing prices (if we accept that Waveshrink has correctly identified the nonsystematic component of the returns sequence).

8.4.2 Searching for nonlinear dependence

Following our visual inspection of the dynamics, we conduct here a sequence of tests in order to obtain a more definite answer on the nature of the structure of the returns sequences. We concentrate on analysing the Waveshrink fitted values and the Waveshrink residuals via RQA (see chapter 5) and by computing the largest Lyapunov exponent and the BDS test (Brock *et al.*, 1986).

In table 8.16 we have tabulated the results of the calculations for the Lyapunov exponent, using up to 9 embedded dimensions. The corresponding graph is in figure 8.39. The maximal Lyapunov exponent was calculated to be positive and close to 0 (as it can be also deduced from the slopes of the curves in the above figure) and could imply the presence of a limit point attractor. There is though a strong indication of instability that could be explained by deterministic dynamics. The information from table 8.18 seems also to be supporting this scenario.

Continuing our calculations of chaotic invariance metrics, we estimated the correlation and capacity dimension values. The correlation dimension was estimated to be 1.86 which shows that we can not exclude the presence of low dimensional

time	d2	d3	d4	d5	d6	d7	d8	d9
0	-7.271145	-7.202302	-7.160007	-7.118117	-7.07683	-7.011904	-6.937624	-6.91934
1	-5.509712	-6.357552	-6.910519	-7.03832	-7.003449	-6.992186	-6.912909	-6.910013
2	-5.033595	-5.956734	-6.777051	-6.98232	-6.970116	-6.943216	-6.851219	-6.866161
3	-4.812322	-5.70629	-6.706622	-6.94426	-6.920337	-6.883374	-6.775203	-6.784495
4	-4.642391	-5.503484	-6.625691	-6.893774	-6.856007	-6.819489	-6.705294	-6.707279
5	-4.536468	-5.387661	-6.545305	-6.829669	-6.767481	-6.761345	-6.646933	-6.652775
6	-4.468019	-5.28041	-6.491634	-6.764757	-6.698928	-6.718224	-6.610766	-6.618373
7	-4.392264	-5.223751	-6.441175	-6.746184	-6.69653	-6.698041	-6.599961	-6.609992
8	-4.329657	-5.165006	-6.417873	-6.751848	-6.704675	-6.702241	-6.609735	-6.63878
9	-4.312444	-5.150391	-6.428015	-6.772029	-6.765853	-6.731491	-6.634256	-6.696717
10	-4.297177	-5.158731	-6.454546	-6.789794	-6.770654	-6.772241	-6.673363	-6.766084
11	-4.30532	-5.161997	-6.481984	-6.818569	-6.814396	-6.811364	-6.712728	-6.829305
12	-4.33772	-5.21149	-6.516485	-6.852583	-6.858897	-6.841146	-6.740695	-6.849782
13	-4.203048	-5.08653	-6.463403	-6.842752	-6.862809	-6.857106	-6.754526	-6.835975
14	-4.127285	-5.011987	-6.410891	-6.813765	-6.849366	-6.852617	-6.751421	-6.802695
15	-4.069028	-4.939197	-6.383165	-6.798482	-6.84625	-6.828882	-6.729829	-6.751636
16	-4.025461	-4.862711	-6.346866	-6.753655	-6.820865	-6.790686	-6.699028	-6.708716
17	-3.98885	-4.814983	-6.30588	-6.707423	-6.772429	-6.747605	-6.663863	-6.676416
18	-3.960939	-4.751951	-6.246264	-6.663817	-6.731538	-6.706903	-6.63094	-6.648751
19	-3.916913	-4.706795	-6.205829	-6.623326	-6.702769	-6.67754	-6.60506	-6.63889
20	-3.875599	-4.674112	-6.168017	-6.587522	-6.670793	-6.662794	-6.597697	-6.665944
21	-3.86031	-4.654021	-6.183111	-6.576747	-6.663955	-6.664534	-6.603242	-6.716312
22	-3.835916	-4.641725	-6.195877	-6.581382	-6.67805	-6.679715	-6.625319	-6.769025
23	-3.822345	-4.626398	-6.197597	-6.588406	-6.6989	-6.700061	-6.6554	-6.816522
24	-3.800894	-4.61773	-6.214719	-6.62037	-6.719893	-6.718291	-6.681332	-6.834134
25	-3.761239	-4.583639	-6.193282	-6.610883	-6.717088	-6.729272	-6.69149	-6.824392
26	-3.750595	-4.55701	-6.156524	-6.595566	-6.704003	-6.730083	-6.691945	-6.796684
27	-3.737304	-4.53082	-6.134899	-6.582288	-6.698049	-6.716586	-6.67235	-6.75121
28	-3.720834	-4.486848	-6.088976	-6.538124	-6.673751	-6.691099	-6.640063	-6.708924
29	-3.707755	-4.455796	-6.038359	-6.489456	-6.634183	-6.660532	-6.609539	-6.67686
30	-3.701383	-4.411802	-5.991457	-6.440786	-6.600937	-6.630686	-6.584441	-6.651943
31	-3.692149	-4.37649	-5.941794	-6.39598	-6.576809	-6.608794	-6.563097	-6.646561
32	-3.689308	-4.347115	-5.897257	-6.366598	-6.546945	-6.596072	-6.556804	-6.676534
33	-3.690044	-4.329093	-5.896254	-6.333486	-6.538167	-6.594756	-6.562357	-6.72524
34	-3.68868	-4.315132	-5.90719	-6.35425	-6.552216	-6.604233	-6.579687	-6.769411
35	-3.684383	-4.305417	-5.923646	-6.399619	-6.575561	-6.618019	-6.603718	-6.802923
36	-3.680952	-4.294647	-5.951081	-6.425445	-6.597122	-6.630951	-6.626674	-6.809292
37	-3.666941	-4.29027	-5.942626	-6.429194	-6.592251	-6.636692	-6.632902	-6.796246
38	-3.665805	-4.274849	-5.927875	-6.413548	-6.583328	-6.639424	-6.630283	-6.775563
39	-3.676363	-4.271757	-5.913153	-6.411211	-6.587121	-6.63811	-6.61713	-6.749789
40	-3.665615	-4.2572	-5.869218	-6.3514	-6.563483	-6.628139	-6.596423	-6.730632
41	-3.667731	-4.254617	-5.822498	-6.290857	-6.526304	-6.612282	-6.577275	-6.706169
42	-3.679349	-4.235997	-5.799621	-6.241456	-6.462198	-6.59389	-6.558796	-6.662457
43	-3.666012	-4.221624	-5.761033	-6.178312	-6.425435	-6.575848	-6.536909	-6.619841
44	-3.665184	-4.221215	-5.710672	-6.158996	-6.407885	-6.562404	-6.528244	-6.610788
45	-3.663635	-4.222998	-5.694237	-6.13243	-6.396353	-6.558447	-6.53218	-6.641962
46	-3.66059	-4.224461	-5.687519	-6.17204	-6.411993	-6.565279	-6.548501	-6.696498
47	-3.650914	-4.225269	-5.705311	-6.213243	-6.431165	-6.578519	-6.572169	-6.756991
48	-3.647624	-4.234907	-5.743078	-6.236089	-6.47321	-6.591046	-6.591307	-6.789013
49	-3.654933	-4.243547	-5.752128	-6.246065	-6.488149	-6.594681	-6.591661	-6.763805
50	-3.645627	-4.231486	-5.750458	-6.243817	-6.514893	-6.59389	-6.580468	-6.705096

Table 8.16: Maximal Lyapunov Exponent
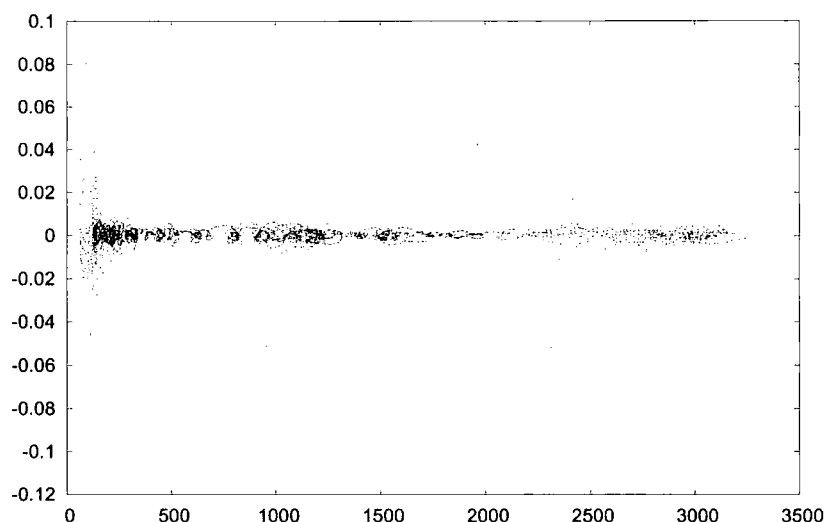
Figure 8.38: Scatter (phase) plot: FTSE ALLSHARE INDEX against Waveshrink fit.

dynamics. Capacity dimension was estimated to be 1.25.

In tables 8.19-8.21 we have included the results of the BDS test calculations following Cromwell *et al.*, (1994). This was conducted on the Waveshrink denoised returns, the actual returns and the residuals. We calculated the statistic for embedding dimensions according to the corresponding AMI values for time-delay τ and for various sizes of neighbourhoods of close points **e**. The null hypothesis H_0 is that the sequences are series of i.i.d. random variables (independently identically distributed) and the test statistic is asymptotically normal.

For α=5% statistical significance level, the critical value is 1.96. It is straightforward that for all sequences, we have to reject the null of independence. As we argued in chapter 5, it is preferable to use such tests within a surrogate data analysis framework. The only difference here is that by preprocessing the original FTSE returns sequences with wavelet filters, we may have altered the significance levels for the whole statistical hypothesis testing procedure. Our BDS test results should be treated with caution as they may be misleading, at least for the Waveshrink denoised sequences. Further research is needed to establish the power of such tests when wavelet preprocessing is applied to the test data.
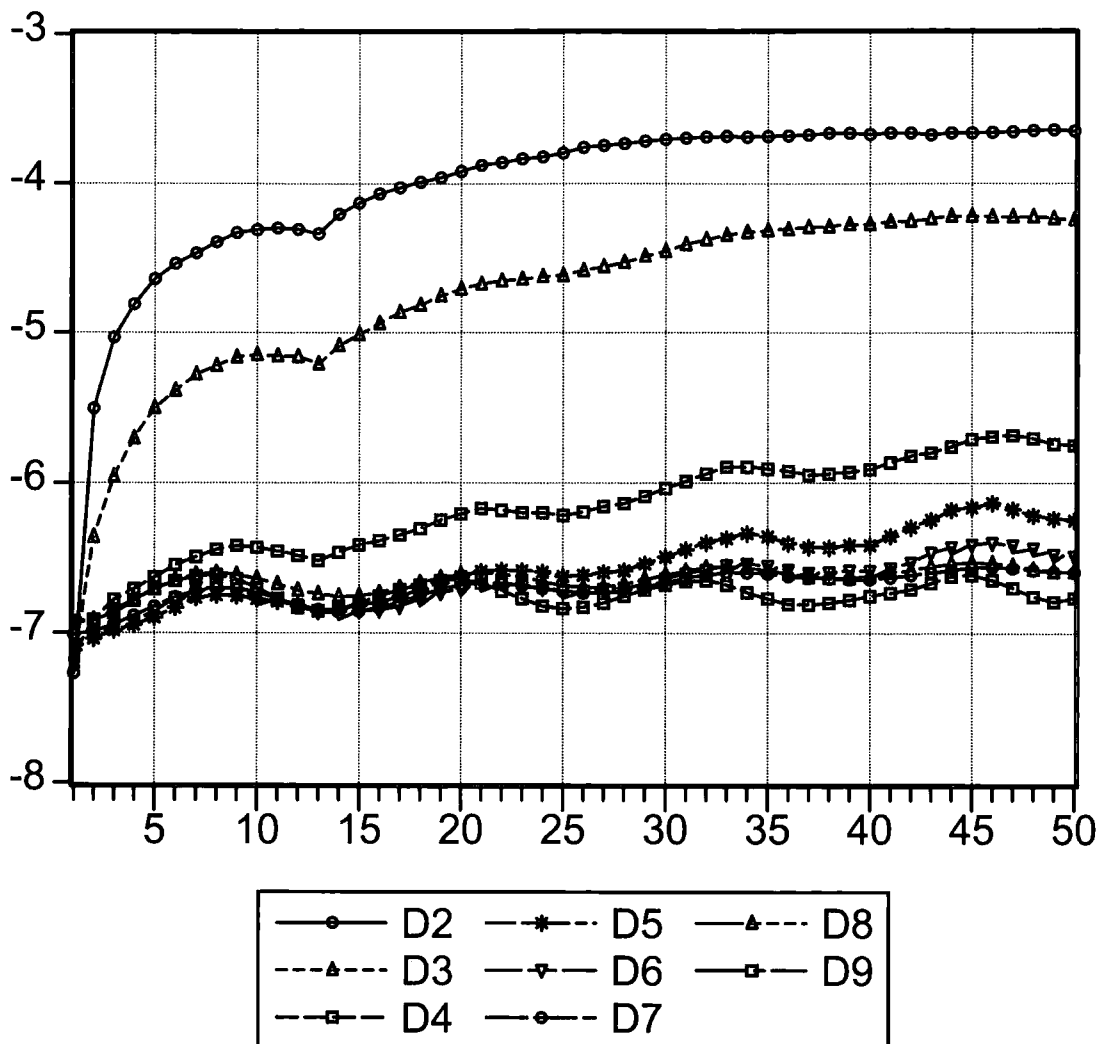
Figure 8.39: Maximal Lyapunov Exponent

8.4.3 Recurrence Quantification Analysis

The visual information from the delay and phase plots as well as the calculated metrics suggest that there may be a low dimensional deterministic dynamics within the generating process for the FTSE Waveshrink-denoised returns time series. The next logical step was to investigate these using recurrence quantification analysis (RQA) (Webber and Zbilut, 1994) as described in chapter 5.

In figure 8.41 we generated the recurrence plots (RPs) for the Waveshrink residuals and fitted values, using the embedding parameters as determined earlier. We generated both thresholded and unthresholded versions of the RPs. Subfigures

D_e	C_d
2	0.122
3	0.172
4	0.218
5	0.261
6	0.313
7	0.379
8	0.455
9	0.545
10	0.647
11	0.765
12	0.902
13	1.06
14	1.23
15	1.41
16	1.64
17	1.75
18	1.79
19	1.83
20	1.85

Table 8.17: Correlation dimension C_d for Waveshrink fit. Proposed value for dimension is C_d=1.86

Type	Value
Capacity	1.25
Information	1.46
Correlation	1.86

Table 8.18: Dimension estimation for Waveshrink fitted values.

(a-b) in 8.41 refer to the residuals. These RPs are devoid of any significant structure as expected. Magnifications of various sections of these RPs did not reveal any deterministic recurrences. Subfigures (c) and (d) refer to the fitted returns. As we can clearly see there, the dynamics are very strong and more intense than those reported earlier in chapter 5 for the original returns sequences. Various magnifications of the thresholded RP of subfigure (d) revealed remarkable structures. Subfigures (e) and (f) show clearly that there are very strong deterministic recurrences. By comparing these RPs with the ones demonstrated in chapter 5, we can clearly see that the fitted returns' dynamics resemble more the ones in the Lorenz RP (figure 5.3 on page 84) than the actual FTSE returns one (figures 5.10 and 5.11 on page 95).

D_e	e=0.0029	e=0.0058	e=0.0086	e=0.0115
2	1.851752e+02	104.7695	76.3662	65.3040
3	2.217260e+02	104.7264	74.4398	63.1530
4	2.679491e+02	103.3663	71.1855	59.9907
5	3.340442e+02	103.4542	68.8693	57.6563
6	4.287900e+02	104.8513	67.3118	55.7507
7	5.646849e+02	107.3747	66.2819	54.3269
8	7.597906e+02	110.9035	65.6517	53.2347
9	1.041572e+03	115.3997	65.3556	52.3979
10	1.451400e+03	120.8850	65.3504	51.7829
11	2.052455e+03	127.4431	65.6091	51.3263
12	2.939294e+03	135.1792	66.2306	51.1331
13	4.256397e+03	144.0559	67.0498	51.0377
14	6.224739e+03	154.1624	68.0639	51.0613
15	9.178386e+03	165.5993	69.2377	51.1704
16	1.363079e+04	178.5201	70.5883	51.3863
17	2.036672e+04	193.0740	72.1009	51.6785
18	3.060562e+04	209.5349	73.7857	52.0357
19	4.620032e+04	228.1187	75.6384	52.4591
20	7.000361e+04	249.0734	77.6774	52.9596
21	1.064126e+05	272.7013	79.8978	53.5170
22	1.622306e+05	299.2581	82.3255	54.1639
23	2.480387e+05	329.1081	84.9535	54.8897
24	3.801729e+05	362.6921	87.7794	55.6870
25	5.840141e+05	400.5221	90.8073	56.5437
26	8.989943e+05	443.1813	94.0402	57.4555
27	1.386315e+06	491.2678	97.4904	58.4241
28	2.141330e+06	545.4998	101.1644	59.4490
29	3.312328e+06	606.7626	105.0859	60.5310
30	5.130501e+06	675.9171	109.2816	61.6949
31	7.956396e+06	754.1236	113.7692	62.9312
32	1.235319e+07	842.5252	118.5421	64.2314
33	1.919837e+07	942.5995	123.6209	65.5949
34	2.986618e+07	1055.8158	129.0329	67.0316
35	4.650643e+07	1183.9533	134.7818	68.5379
36	7.248517e+07	1329.1457	140.8855	70.1144
37	1.130651e+08	1493.6304	147.3633	71.7610
38	1.765086e+08	1680.0780	154.2483	73.4803
39	2.757708e+08	1891.5935	161.5566	75.2742
40	4.311253e+08	2131.6122	169.3129	77.1453

Table 8.19: BDS test statistics for the Waveshrink fitted values. Embedding dimension D_e ranges from 2 to 40.

D_e	e=0.005	e=0.01	e=0.015	e=0.02
2	19.5158	22.3919	25.7312	28.5046
3	25.4327	27.7493	30.8304	33.6993
4	29.7449	31.3975	33.6811	36.1074
5	34.9725	35.1123	36.0299	37.7345
6	40.4283	38.6889	38.0289	38.8859
7	48.2491	43.1466	40.3045	39.9644
8	58.0616	47.9569	42.5383	40.9241
9	69.8953	53.8184	45.1180	41.9534
10	84.1499	60.8204	47.9608	43.0562

Table 8.20: BDS test statistics for the FTSE ALL SHARE returns sequence. Embedding dimension D_e ranges from 2 to 10.

D_e	e=0.0037	e=0.0074	e=0.011	e=0.0147
2	5.4147	5.2439	4.9389	4.0153
3	10.0185	9.2285	8.2396	6.7585
4	14.1093	12.8060	11.3598	9.4550
5	18.4796	16.3753	14.1669	11.7171
6	22.1520	19.3503	16.3747	13.4423
7	27.1757	23.2274	19.0033	15.4200
8	32.2873	27.0065	21.3365	17.0626
9	41.2059	31.2639	23.7854	18.6433
10	53.1275	36.1510	26.2818	20.1204

Table 8.21: BDS test statistics for the Waveshrink residuals sequence. Embedding dimension D_e ranges from 2 to 10.

Variable	Std Normal	p-value
FTSE returns	-0.7721	0.4401
Waveshrink fit	-5.9457	2.753e-09
Waveshrink residuals	-0.6661	0.5053

Table 8.22: Runs test results for the three sequences.

Variable	LLE
FTSE index	0.0056543
FTSE returns	0.032205
Waveshrink fit	0.008199
Waveshrink residuals	0.135448

Table 8.23: Largest Lyapunov Exponents (LLE).
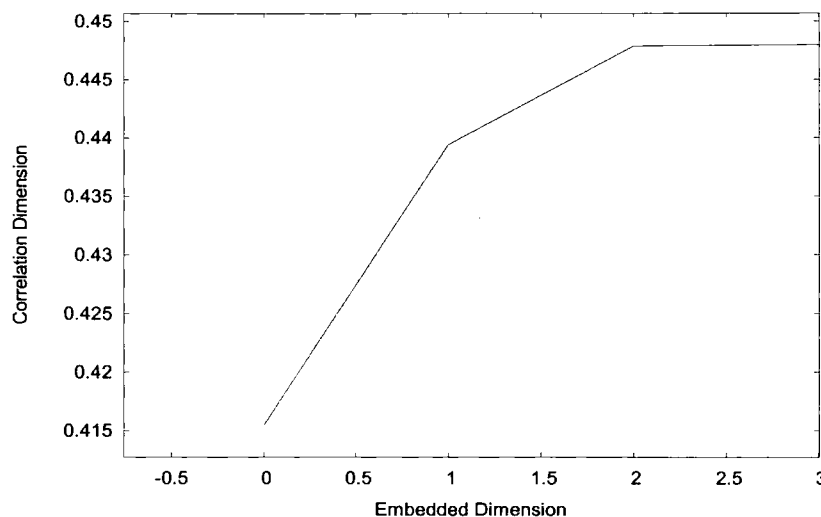
Figure 8.40: Waveshrink fit correlation dimension calculations.

RQA Statistic	Value
Mean	0.000
Standard Deviation	0.006
Mean Distance	93.306
% Recurrence	29.256
% Determinism	45.26
Ratio	1.547
Entropy	1.658
Maximum Line	906
Trend	-3.214

Table 8.24: Waveshriked FTSE returns RQA results from RP in 8.41 (d).

In table 8.24 we present the results of RQA on the wavelet pre-filtered returns. Comparing these results to the ones in chapter 5 (table 5.2, page 98), we can clearly see that the filtering has improved the metrics that indicate determinism. The standard deviation has diminished whereas the mean distance has become larger. The percentage of determinism has risen by 17.58 % whereas the entropy has fallen. This indicates that wavelets have managed to pick up and filter out a substantial stochastic component of the series. The maximum line length has also increased indicating that we have picked up even longer recurrences. All these findings here lead to the conclusion that wavelet pre-filtering and denoising may be the way ahead in establishing whether financial time series are governed partially or fully by

(a) Residuals

(b) Residuals

(c) WS fit

(d) WS fit

(e) Zoom 1

(f) Zoom 2
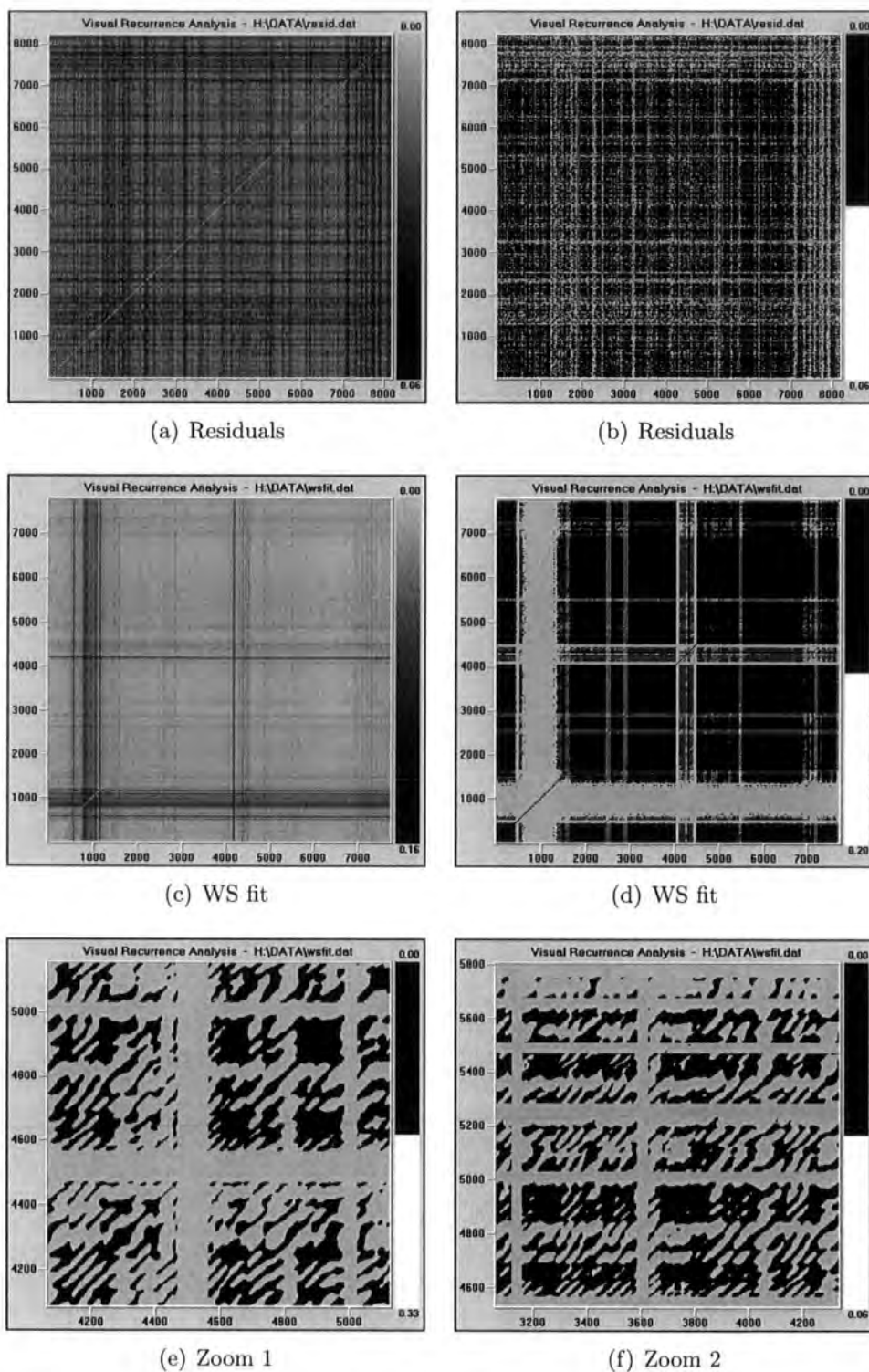
Figure 8.41: Recurrence plots for Waveshrink fit and residual sequences.

nonlinear deterministic processes.

8.5 Conclusions

Wavelet transforms can be used on data in order obtain information on various frequencies as well as in time. This is a clear advantage over spectral analysis which can only focus on frequencies. Financial time series exhibit volatility and sharp localised fluctuations. This makes them ideal candidates for wavelet analysis.

In this chapter we showed how continuous and discrete wavelet transforms can be applied on financial time series. We also showed how the results from these transforms can be used to gain an insight on the dynamics and the structure of these time series. Initially we obtained from continuous transforms the wavelet scalograms. We revealed a wealth of structures in various scales and showed how these graphs can identify significant events that altered the structure of the sequences. More precisely, using the FTSE ALL SHARE daily time series, we were able to identify the timing of shocks such as the oil crisis of the 70s or the 1987 stock market crash. We were also able to find how these events translate in various scales.

As a second application we used discrete wavelet transforms to obtain views on the structure of the FTSE closing prices and returns at various scales. Following that, we used the Waveshrink algorithm to denoise the returns sequences and showed that the obtained residuals (noise) were stationary, i.i.d. normal random variables. All results seem to concur that Waveshrink was able to filter out nonsystematic components. The fitted values were then passed through an array of qualitative and quantitative tests in order to establish what kind of dynamics were governing their data generating process. Careful examination of their structure revealed that we can not rule out deterministic nonlinearities and chaos. There is a strong indication that the dynamics are not stochastic or random. We supported this finding using recurrence quantification analysis, by calculating measures such as Lyapunov exponents and correlation dimensions and finally by conducting statistical tests for independence (BDS and Q-statistics) and nonlinearity (RQA).

The final conclusion of this chapter is that wavelets can be used in conjunction with other linear and nonlinear techniques in order to provide a clearer view of the

data generating processes of financial time series. Wavelet filtering can provide an enormous wealth of information and detail for the dynamics of the time series. Moreover, it can be used to obtain cleaner processes. We experimented with a single time series of the FTSE logarithmic returns and revealed very clear and interesting dynamics. Future research may be focused on refining this information and combining the tools used here possibly under a multivariate framework.

CHAPTER 9
Conclusions

9.1 What can we deduce from our results

As we discussed in the introduction, the emergence of chaos theory in finance implies a substantial paradigm shift. Complexity as a descendent of chaos theory, while contributing to the dynamical systems overall theoretical framework, stops at the border of regular behaviour and unpredictable chaotic irregularity. In this sense, tools from nonlinear deterministic system analysis are required to characterise chaotic structures or attitudes. The nonlinear methodologies utilised in this thesis can deduce whether a system's dynamics abide to nonlinear determinism or stochastic randomness.

Until recently we were not well equipped to detect any significant shifts from randomness. Most commonly the BDS (Borck *et al.*, 1987) test for independence was used, which is though inappropriate for application to small, noisy data sets. Secondly, this test can reveal wether the data generating process is an individually identically distributed sequence. The alternative to this is not necessarily a chaotic process. The introduction of qualitative tools such as the *close returns* (Gilmore, 1993) and *recurrence* plots (by Eckmann *et al.*, 1987), gave us the opportunity to visualise dynamics, with a few non-critical and non-parametric assumptions and very little or no manipulation of the data. Moreover these tools enable us to deduce in a clearer fashion whether the dynamics involved are deterministic or stochastic. Wavelets (Daubechies, 1992) on the other hand have provided a new filtering approach that enables global and localised view of dynamics and structure for any sequence whether being stochastic, deterministic or a result of mixed processes. Furthermore, they can be applied to finite and highly discontinuous sequences, even to series which are non-homogeneously distributed in time. Wavelets do not suffer from the shortcomings of Fourier analysis and have been used successfully in many sciences with multidimensional data. Wavelet filtering has also been extremely successful in the exploratory analysis of sequences that exhibit chaotic dynamics (see Wornell

237

1993 and 1995). Combining the so far mentioned techniques we revealed that the dynamics behind the stock market series are characterised by absence of stochasticity. Moreover we found strong presence of nonlinear deterministic structures. We achieved this with a minimal set of parametric or theoretical model assumptions and the results appear to be reproducible. Another important aspect of this research is that we established nonlinearities in the data by avoiding using invariant statistical measures as a basic approach as these have been proven to be prone to miscalculations and to often lead to inconsistencies, due to the quality of the financial data used.

The results from all three empirical essays point to the same direction. We believe that we have compelling evidence from both qualitative and quantitative sources that stock market returns are not characterised entirely by stochastic data generating processes. Visual inspection of the dynamics as these are revealed in chapters 5 and 8, implies that the structure of these sequences does not agree with the expected theoretical facts. In chapter 5 (figure 5.11, page 96) the magnified areas of the recurrence plots indicate that the returns have considerable deterministic recurrent structure. This is also identified by the quantification procedure in page 98 (see table 5.2). These findings were the initial point for the path followed in the rest of the thesis.

In econometric theory the important issue is to put specific hypothesis to the test. Parametric model fitting is assumed preferable to any atheoretic approach. Even then though, the dynamics of any system defined by a set of structural relationships should be put to the test via a statistical hypothesis testing procedure. *Surrogate Data Analysis* (SDA, Theiler *et al.*, 1992) as examined in chapter 6, provides a link between such an approach and the suggested strategy in this thesis. In section 6.3 (page 107) we discuss our rationale. It is important to stress again that the SDA scheme is a prerequisite for most approaches that involve the calculation of chaotic invariants for the determination of chaos in time series. Although this methodology exists since 1992, it seems that it has been largely ignored by the empirical finance literature. This fact puts previous research results under serious scrutiny. Our experimentation with various types of tests in chapter 6 has provided a clear

view on how this methodology can be utilised to furnish us with a direct answer. The main conclusion is that the set of linear hypotheses referred to in section 6.4.1.1 (page 112) can be strongly refuted. The alternatives to these hypotheses allow directly for nonlinear and nonstochastic structures. The implications are that we can not rule out with certainty deterministic and chaotic structures for the financial returns sequences. This provides further support to our investigation objectives.

So far we had achieved to provide strong implications of deterministic structures and nonlinearities in index closing prices and return sequences. The main problem, always identified in approaches like ours, is the lack of clean "observations". This was clarified from our introduction. Intuitively the logical course of action would be to consider transformations of the data that could ensure substantial noise filtering without tampering with the underlying dynamics of the sequences. Unfortunately not all preprocessing approaches are useful when it comes to revealing unknown dynamics. If indeed the case was that nonlinear determinism is at hand, a number of pre-filtering techniques should be ruled out. Abarbanel (1995) and Theiler and Eubank (1993) among others have shown that there are a few techniques that would capture nonsystematic noise components efficiently and provide us with series that contain "untouched" structural information. In the case of chaotic determinism, the approach should be very sensitive to the presence of any attractors and allow them to exist in the dynamics of the "clean" data. Using a slightly different approach we employed wavelet transforms and more precisely the discrete transform version and the denoising strategy called "Waveshhrink" (Donoho and Johnstone 1994 and 1995, also in Bruce and Gao, 1996) based on wavelet regression. In section 8.3.2 (page 205) we demonstrated how this can be applied to returns sequences such as the FTSE index, to provide a denoised version of these. The noise obtained was remarkably white. Furthermore, the transform was monotonic-invertible which implied that the loss of original information was minimal. This ensured that chaotic information had been preserved to a very large degree (if not entirely). The examination of the preprocessed sequences with the methodologies outlined in the previous chapters revealed qualitative information on the existence of cycles consistent with chaos and determinism. Moreover, quantitative measurements provided

strong implications for the existence of determinism. Visual information from sca-
lograms (in section 8.2, page 185) suggested also the existence of some type of self
affinity and self similarity, characteristic indications of fractal structures. These re-
sults are also reproducible. The evidence here suggests that the processes are purely
aperiodic and nonstochastic. Yet after all this evidence it would be a very strong
and somewhat rushed premise to suggest that chaos is at hand. A plausible answer
would be to entertain the idea of mixed processes, chaotic and stochastic. For such a
position though there is yet no well defined and structured methodology that allows
for robust testing. This could be an interesting area for future research. There is
a new though scientific concept that could explain the structures revealed in this
thesis, that of the *self-organised critical phenomena* (Bak, 1994) and we discuss this
in subsection 9.2.1 that follows.

Similar to this thesis' context research has failed to provide in the past, strong
evidence of low dimensional dynamics in stock market indices. For example, Abh-
yankar *et al.*, (1997) study the S&P 500 index and although they report nonlinear
structure they do not support low dimensional deterministic chaos. Peters (1994),
using a very small set of 496 monthly observations, provides evidence of low dimen-
sional deterministic dynamics for the same index. His data though are preprocessed
with the use the of Consumer Price Index. Andreadis (2000) uses the same index
and transformations of it via a *self-criticality* (Bak *et al.*, 1989) approach to demon-
strate the existence of fractional Brownian dynamics for the series and absence of
chaos, supporting the indications of Hsieh (1991). Willey (1992) investigates the
NASDAQ 100 and S&P 500 daily index but fails to reach conclusive results that
support the existence of chaos.

We could refer to a plethora of similar research results that have been obtai-
ned over the last ten years. The conclusion is that noise in the stock index time
series, affects the performance of the methodologies utilised and leads to inconclu-
sive results. To adopt the philosophy of our approach in chapter 6, any research
that is based on statistical hypothesis testing is prone to misinterpretations due to
inconsistent use of testing procedures or miscalculation of the power of the tests.
Secondly, parametric approaches and preprocessing set limits to the upper bounds

of statistical tests, an issue which is largely ignored in the literature. This implies that one may be conducting a test at a 5% level of significance though the real level may be much larger, leading thus to acceptance of linearity, log-normality and stochasticity when in fact we should have been rejecting these. And last but not least, the refutal the presence of chaos, via statistical hypothesis testing, implies clearly that on the basis of the data set, one can not support its existence. It does not necessarily mean that nonlinear determinism is not there or in other similar data (or even a larger sample of the same data).

In this thesis, with the exception of chapter 6, where tests can be conducted within a "safeguarded" statistical environment, we avoid the pitfalls discussed so far. We adopt a purely *topological time series analysis* approach and search for deterministic structure, avoiding methodologies and calculation of statistics that can produce misleading results due to the nature of our data. We thus provide compelling qualitative evidence as well as quantitative indications against linearity and stochastic randomness for our data set. The performance of the tools used can be improved by future research and this may lead to clearer and stronger support for the hypothesis of chaotic dynamics.

We believe that the usefulness of our approach is self-explanatory. It is fairly straightforward that recurrence analysis can be used as a *weak from efficiency test*. Very effectively, recurrence plots and recurrence quantitative analysis (RQA) can provide evidence for or against stochastic randomness for stock return sequences, the way we indicated in chapter 5. Lack of structure in recurrence plots can be easily quantified by RQA which can provide a clear-cut indication for the existence market efficiency. The only prerequisite here is careful *delay coordinate embedding* and preprocessing of the data, if needed. In the case of suspicions of large noise contents, one may choose to filter the sequences via finite impulse response filters (following Abarbanel, 1995) or choose the methodology suggested here, i.e., the *discrete wavelet transforms*. Moreover the results can be supported by an SDA statistical hypothesis testing framework, as we already demonstrated.

9.2 Chaos, Self-organised Criticality and the future

We already discussed that our empirical approach aims to provide further evidence in support for nonlinearity and nonstochasticity for stock returns, in the context of novel weak form market efficiency tests. There is a wealth of previous research, mostly during the last 10 years that deals with this issue in an analytical or empirical framework. One can refer to collections of articles by Trippi (1995), Creedy and Lance (1994), Creedy and Martin (1994 and 1997) or Barnet *et al.*, (1996) for a concentrated view of the research so far. There are many articles which deal with the issue of market dynamics that would allow for chaotic prices or returns and the ones closer to this thesis were discussed in chapter 2.

Our research here concentrated in the examination of daily frequency time series of stock indices, mainly the FTSE ALL SHARE. One may argue that there is a substantial part of the dynamics that would evolve intradaily, at higher frequencies. Aggregation to the closing prices of every day would not always allow the "true" high frequency dynamic components to survive. An interesting thus path for future research would be to identify if the methodologies utilised here would reveal the same results for higher or lower frequencies. Following Abarbanel (1995), Peters (1994) and Kantz and Schreiber (1997), we can note that collecting data for a large period of time i.e., over 30 years at daily frequencies, would ensure under certain general assumptions that we have most of the necessary information to properly unfold the dynamics of the phase space. Of course, in every empirical approach, there is always the question of how much data should we use. Intuitively, a comparative approach of using a long time series of stock market returns in various frequencies would provide a clearer view into the true dynamics of the market. Unfortunately, high-frequency data are available for a sort period of time relatively to daily and lower frequency observations. For a very detailed view on the issue of high-frequency data, one may refer to the work of Gwilym and Sutcliffe (1999) as well as Goodhart and O'Hara (2000). Most of the research in the high-frequency field (see for instance the research from Olsen and Associates[92]) seems to be pointing towards a specific direction: time dependent structures with highly asymmetric dynamics. Time-varying frequencies

---

[92]On the WWW under: http://www.olsen.ch/research/working_papers.html

and irregular spacing of events generate structures that do not usually leave a stamp of their dynamics when these series are sequenced in lower frequencies.

Experimentation with higher frequencies than daily has provided some interesting results: scaling laws and long memory components. Müller *et al.*, (1993) suggest a scaling law for absolute price changes for FX data and suggest fractional noise structures and long memories for their volatility structures. Their findings are in accordance to the frameworks laid by Mandelbrot (1999a and 1999c) where he demonstrates clearly how a fractal dynamics in time series can be explained as long memories and fractional noise by existing empirical "nonlinear" methodologies. An interesting point here is that ARCH and GARCH type approaches do not sufficiently capture the long memory structures due to the fractality or multifractality of the dynamics. Our contribution to this area is that by using wavelet transforms we were able to reveal via scalograms (see chapter 8) the evolution of the dynamics at different scales for the index and returns time series. We were also able to identify the structural and volatility breaks caused by sever financial crises with accuracy and to discern how these shocks affected the whole series at various time scales. The same gaps in the phase space can be also revealed in the recurrence plots in chapter 5, where the vertical and horizontal line segments identify the time and severity of the shocks.

Finding a plausible explanation for the structures revealed is not an easy task. This is still a liquid issue and an area of ongoing research. A quick answer would be that external events may be affecting (contaminating) the dynamics of the sequences under examination. That is probably why recurrent patterns seem to be interrupted and the selfsimilar bifurcations in the wavelet scalograms appear not to be sharing the same structure during the periods of strong fluctuations. An interesting area of future research would be to identify the change of phase-space dynamics during crisis and attempt to model the structures there independently. What though theoretical approach may one adopt in order to justify our results and suggested research strategy? We suggest that the answer lies in some newly developed notions and theories.

The assumption of complexity in market dynamics inevitably brings us clo-

ser to the issues of self organisation. If indeed nonlinear determinism is a major power behind the movements of prices, the market as a complex system should exhibit a certain degree of stability, instability and self-organisation. The epicentre of the discussion hereon will be the work and ideas expressed mostly by Bak (1996) and Holland (1995). These refer to the area of physical science regarding *critical phenomena* from complex systems and *hidden order*.

9.2.1 Self organised criticality

Self-organised criticality (SOC) (Bak, 1987) provides a new way of viewing physical and social phenomena. The basic assumption is that part of nature is perpetually out of balance though organised in some critical state where anything can occur within a framework of well-defined statistical laws. SOC can explain some types of complex patters which can be discerned during catastrophic events who are characterised by fractal dynamics. According to Bak (1995) "*Complexity originates from the tendency of large dynamical systems to organize themselves into a critical state, with avalanches or "punctuations" of all sizes. In the critical state, events which would otherwise be uncoupled become correlated. The apparent, historical contingency in many sciences, including geology, biology, and economics, finds a natural interpretation as a self-organized critical phenomenon.*".

The basic example in Bak's work is the *sandpile* model. In its critical state it mimics a range of phenomena associated with complexity. A sandpile can be created by sand grains let to fall down. In the beginning the pile is flat and individual grains stay close to where they have originally landed. Up to this point, classical physics can provide an adequate explanation for the dynamic attitude of every sand grain, using basic Newtonian laws. As the process of dropping sand on the pile continues though, the pile gets larger and steeper and eventually landslides will occur. At some point, large landslides as "avalanches" will span most of the surface of the sandpile. At that point the system is out of balance and the behaviour of individual grains can not provide an explanation of the overall system dynamics. This is the *critical state* of the sandpile. Any slight disturbance can trigger an avalanche of its own dynamics. The nature of the sand slides is fractal. Their size distribution

(number of sand grains per slide) follows a power-law such as the fractal dynamics introduced by Mandelbrot and other "chaoticians". At the end of the day, Bak suggests that significant events do not happen gradually or smoothly but abruptly, in the form of "avalanches". The basic characteristics of a self-organised system according to Bak are:

- The system is open and dissipative.

- The system organises itself in a critical state with regular aperiodic avalanches of variable size due to its dissipative structure.

- The system's dynamics are embedded in a single spatiotemporal fractal structure.

- Catastrophic instability for a self-organised system may occur when it is forced to certain optimal states which can take it out of its self-organised state.

To use Bak's own words: "*self-organisation can be described by an inescapable divergence of the size of avalanches*". Avalanches of size say 1, 2, 3, 4,...,n would occur with probabilities of 1, 1/2, 1/3, 1/4,...,1/n respectively. Thus, the probability of occurrence is inversely proportional to the size of the slide. This law is in agreement with the existence of fractal patterns. The avalanche size according to the number of avalanches follows a $1/f$ noise or *flicker noise* law. This law characterises many fractal and long-memory phenomena and is increasingly discovered in finance.

Bak's views allow for the presence of SOC in financial phenomena. This has not been bypassed in the literature (see Pis'mak (2001), Strozzi *et al.*, (2002), Sornette (2003), Brock (1991) and Middleton (1996) among others). A common feature of SOC models is the presence of separation of time scales i.e., the system evolves at a slow rate until one of its elements reaches a certain threshold. This triggers the shock or avalanche, in a form of increased activity which will occur within a very short time window. When the shock is over, the system evolves again, following the same patterns until the next avalanche. Since it is an open dissipative system, there is provision for external driven factors or forces. For example, in the case of a financial market, SOC allows for interaction of agents from various sectors as well

as activity that is triggered by some cognitive process depending on the filtering of relevant but outside the market information (such as political events and other non quantitative observations). Institutional, economical, social, natural phenomena and physical laws, can all be accounted for within a SOC framework. Moreover, the *dissipative* nature of the system under SOC, allows for a wealth of other features that make the theory appealing to many disciplines.

In this thesis we provided evidence that financial time series indicate a stock market dynamical structure with aperiodic, nonstochastic and probably dissipative chaotic structure. Any constraints to the agent's expected utility functions can cause according to Bak such dissipation. This also occurs when due to their *bounded rationality* (see Rubinstein, 1998), agents misinterpret or fail to interpret aspects or signals of their environment. As their choices are discrete, dissipation will also occur discretely.

Concluding, SOC can be a very interesting new paradigm that may provide an adequate explanation for the types of dynamics observed in stock market systems, not accounted for in the classic financial theory literature.

9.2.2 Investigating the causes of chaos and the future of research

It seems that as empirical analysis tools are improved, scientific research will be offered more means and opportunities to discover deterministic and chaotic patterns in asset prices. The fundamental question though will prevail: "*What causes chaos in the market dynamics?*". An answer can not be easily found. So far we have been partially successful on mainly detecting nonlinear-chaotic dynamics. As already discussed, financial volatility and turbulence may be explained under a SOC framework. This is an area of ongoing research. No convincing or widely accepted analytical model has been presented yet that explains the existence of an economic or even market-psychology mechanism that leads to the realisation of critical phenomena and emergent market behaviour.

In his forthcoming book Los (2003) discusses chaos and financial turbulence. Regarding the causes of chaotic crashes, he provides a sensible idea (page xivii) which we quote below:

"Now some financial crises are more dangerous than others. For example it may not be dangerous to speed up the trading and price formation activity in a financial market and encounter a crisis, because the financial market may move through a so-called safe financial crisis. Whereas slowing down trading and price formation may lead to an unsafe financial "blues sky catastrophe". It may cause a financial crisis in which the pricing system close to an attractor suddenly heads for the attractor at infinity: the market pricing process breaks down and can't recover."

According to Los (2003), financial turbulence could be perceived as an "efficiency enhancing phenomenon" and distinguished from catastrophes or crises. Any discontinuity in persistent financial time series could be regarded and measured as a crisis and should only be detected in low liquidity markets (especially currency markets). Los implies that chaos can be detected by a market's transition from smooth to turbulent dynamics. In this case, our findings (especially in chapter 8) could provide an invaluable view of the structure of the market shocks. By using a combination of wavelets and topological time series analysis tools, we could time and examine closely the break down of smooth market functioning, examine the frequency and locality of the crises and also analyse the structural changes of the dynamics during the shock's history. It is evident from our preliminary results (in chapter 8) that there are is substantial knowledge to be gained from such an approach.

Recently, Alvarez-Ramirez *et al.* (2002) attempted to provide a model that could produce deterministic prices which resemble actual stock market data. Attempting to capture the complexity of stock market prices' behaviour, they proposed a framework where fundamentalist and chartist patterns of trading behaviour induce transient and oscillatory price dynamics. They suggest that deterministic modelling approaches should be used cautiously and be regarded as a complement to stochastic methodologies, agreeing on the issue of difficulty in short-term forecasting as in Malliaris and Stein (1999).

An interesting area of current and future research is the "Compass Rose" structures discovered in phase portraits of asset returns. These formations, can increase asset price forecastability, help to calibrate volatility models and are attributed to

price discreteness and market microstructure. Introduced through the Journal of Finance by Crack and Ledoit (1996), it has found its way to random walk tests (Fang, 2002) and chaos detection (Kramer and Runde, 1997 and Gleason et al., 2000). Our approach has revealed evidence of nonperiodic cycles and chaos. A future research path will take us to investigate how asset price dynamics, as revealed in chapters 5, 6 and 8, can generate through the presence of noise and price discreteness, patterns like the compass rose.

Allen and Phang (1994) adopt a more radical strategy. They choose to explain complexity in financial markets through an *evolutionary economics* approach. Using various simulations they attempt to examine chaotic market dynamics through a self adaptive trading model. Assuming a fixed chaotic attractor, using an evolutionary and learning process, their model developed successful trading strategies. Even in difficult to forecast situations, the model managed to make profits and acquire knowledge of long term dynamics. Their conclusion was that market efficiency related investment strategies were abandoned for the survival of the trading system. They did not though investigate time-varying chaotic dynamics and such an exercise could be a very interesting area for future research. The experiment by Allen an Phang shows an interesting case where investment behaviour renders market equilibrium conditions that lead to efficiency irrelevant, when the dynamics of the price mechanism exhibit chaotic determinism. On the same lines, Farmer and Lo (1999) discuss that evolutionary and ecological models of financial markets can provide further insight on why and how agents compete and adapt in complex dynamical markets though on a *suboptimal level*. They discuss how evidence of chaotic and deterministic market dynamics contradicts the EMH and at the same time can be explained very realistically by agent-based models. Hommes (2002) combines an evolutionary approach with bounded rationality. Using nonlinear adaptive systems he finds that certain stylised facts (fat tails, volatility clusters and long memory) can be explained by evolutionary models. He suggests that experimental testing can provide important insight on whether asset prices are driven by only news on fundamentals or "market psychology".

In general, computer simulations, artificial intelligence and complexity theory,

combined with evolutionary economics and agent-based methodologies are regarded as the cutting-edge of *experimental economics*. The relevant literature usually points to the general direction of chaotic determinism and nonlinearity in the markets and economies investigated. We can safely assume that in the near future there is going to be a strong interest to put theories such as the *EMH* and *Rational Expectaions* to the test. In this respect, we believe that the methodological approach of this thesis can be easily combined with experimental and agent-based economic models to produce really interesting, if not ground-breaking results.

Concluding, it would be interesting to note the implication of chaos in asset and derivative pricing. Savit (1988 and 1988) uses purely chaotic models to illustrate the expected effects of nonlinearities on options calculations. Both chaotic and probability based option pricing evaluations appear to be entirely different although the random and chaotic price processes are indistinguishable. They appear to have the same unconditional distributions. Savit (1989) suggests that the knowledge of the true conditional distribution can be aided by chaotic-determinism related theories and methodologies. In such a case, short-term predictability may improve and, with the use of options, one may achieve long-term improvements in investment strategies. Empirical research on this field seems also to be increasing. Adrangi and Chartrath (2001 and 2003), Adrangi *et al.* (2001) and Chartrath *et al.* (2001 and 2002) investigate the possibility of chaotic price formations in futures and especially commodity futures markets. They do not in general discover strong evidence for chaos and the nonlinearities are explained mostly by GARCH structures. We believe that our approach could significantly improve the results of their research and provide more clear answers on the issue of chaotic determinism in the price processes investigated.

9.3 Conclusion

By now it should be more apparent that this thesis follows a path between *inductive reasoning* and *emergence*, while borrowing some elements from *reductionism*[93] in the sense that every definable process is assumed to be computable. We have used a combination of empirical approaches to analyse data rather in a "data-

mining" fashion than a parametric model one. Our intention was to "let the data speak" to borrow a popular expression. We investigated patterns searching for correlations and recurrent structures that could agree with certain financial realities, as these are reflected within some well defined models of Economics and Finance. We also believe that our research is somewhere between relativism and reductionism since our results may support a process that leads to better understanding of financial dynamics and allows for improved model formulation and prediction. As a final point our results are reproducible (see Antoniou and Vorlow, 2002).

In empirical economics and finance most of the formulations follow the *Occam's Razor* rule, a reductionistic approach which ensures compact congruent representations. Between competing models and ideas the simplest is the most preferable. Within the context of this thesis, it remains truly a philosophical question of whether financial phenomena follow stochastic or deterministic paths. We provided the arguments against determinism in the introduction. The research path we followed so far has provided strong indications for the absence of linearity and simple stochastic rules that lead to "convenient" stabilities. In a true "wavelet" fashion, we attempted to see the *"forest and the trees"* (see[94] Grapps 1995). According to this statement, our approach was a study of emergence. Moreover, by involving minimal assumptions about the systems examined, we allowed for relationships between the market systems and the overall financial environment. These relationships though did not dictate the methodologies and course of analysis applied.

With respect to the hypothesis of "Market Efficiency" (EMH), our answer should be conditional to what is the true notion of an efficient market. If we follow Granger and Morgestern (1970) (see also page 37 in this thesis), the random walk may imply that:

"*...price changes, in* **absolute terms** *are not predictable on the basis*

---

[93] "*The logical positivist version of reductionism also implies the unity of science insofar as the definability of the theoretical entities of the various sciences in terms of the observable would constitute the common basis of all scientific laws. Although this version of reductionism is no longer widely accepted, primarily because of the difficulty of giving a satisfactory characterisation of the distinction between theoretical and observational statements in science, the question of the reducibility of one science to another remains controversial.*" by James Schombert, see also online: `http://zebu.uoregon.edu/~js/glossary/reductionism.html`

[94] Also online under `http://www.amara.com/current/wavelet.html`

*of any **linear** combination of the past history of these changes. It does not rule out any underlying nonlinear relationships nor contradicts the possibility of predicting the **relative** price change of one stock compared to another..."*

According to the above statement, our answer should be that nonlinear determinism and chaos, do not contradict the random walk model and hence the EMH. We should also consider that the attitude of chaotic deterministic systems may be forecasted only in the short run as long run forecasting is inhibited by the exponential divergence and instability of their dynamics. This itself may suggest that techniques and technology and since with chaotic models we can not determine (accurately or at all) the long run equilibrium or disequilibrium states of any nonlinear deterministic system, having stock markets follow such laws would render them de facto unpredictable[95] and hence "efficient" in the broad sense. Essentially, our answer here would involve references to recent financial literature such as Peters (1999a and 1999b), Bass (1999), Taleb (2001) and May (1999). Quoting the last:

"...it is impossible to have nonlinearity in finance and equilibrium in economics. They are mutually exclusive. Put another way, once we destroy linearity in finance, the concept of equilibrium in economics must also fall, and vice versa."

Under this rationale, admitting the presence of "chaos" in market efficiency definitions, would destabilise the theoretical framework of any market equilibrium model, such as the ones based on discounted expected utilities. We agree in that sense with Bethlehem (1977) who points out that the biggest problem the EMH poses is that it is concerned with efficiency in an *ex-ante* sense. Efficiency though can only be judged *ex-post*.

The final conclusion is necessarily that further research, surely interdisciplinary, is needed to provide a further incite on the dynamics we described in this thesis. The answer should lie somewhere between chaos theory, volatility and long memory models. Research should also focus on the concept of bounded rationality

---

[95]We thank the external examiners for their useful comments and discussion on this issue.

and how the EMH could be modified to allow for irrational investment behaviour that could explain nonlinear determinism in asset price dynamics. We believe that the combination of chaotic models and wavelets is very promising. Current scientific endeavours and faculties of thought appear to be in favour of our research strategy to this point.

BIBLIOGRAPHY

Abarbanel, H. D. I., 1995. Analysis of Observed chaotic Data. Springer-Verlag, New York.

Abarbanel, H. D. I., Brown, R., Kennel, M. B., 1992. Local Lyapunov exponents from observed data. J. Nonlinear Sci. 2, 343–365.

Abarbanel, H. D. I., Kennel, M. B., 1993. Local false neighbors and dynamical dimensions from observed chaotic data. Phys. Rev. E 47, 3057–3068.

Abhyankar, A., Copeland, L., Wong, W., 1997. Uncovering nonlinear structure in real-time stock-market indexes: The S&P 500 index, the DAX, the Nikkei 225 and the FTSE-100. Journal of Business and Economic Statistics 15, 1–14.

Adler, R. J., Feldman, R. E., Taqqu, M. S., 1998. A practical guide to heavy tails: statistical techniques and applications. Birkhuser, Boston.

Adrangi, B., Chatrath, A., 2001. Nonlinear dependence in palladium and platinum futures: Some tests for chaos. Finance India 15 (3), 865–83.

Adrangi, B., Chatrath, A., 2003. Non-linear dynamics in futures prices: Evidence from the coffee, sugar and cocoa exchange. Applied Financial Economics 13 (4), 245–56.

Adrangi, B., Chatrath, A., Raffiee, K., 2001. The demand for US air transport service: A chaos and nonlinearity investigation. Transportation Research: Part E: Logistics and Transportation Review 37 (5), 337–53.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Automatic Control AC-19, 716–723.

Allen, P. M., Phang, H. K., 1994. Managing Uncertainty in Complex Systems: Financial Markets. In: Leydesdorff, L., Van den Besselaar, P. (Eds.), Evolutionary economics and chaos theory: New directions in technology studies, 1st Edition. Printer Publishers Ltd., London, UK, pp. 125–135.

Alligood, K. T., Sauer, T. D., Yorke, J. A., 1997. Chaos; An Introduction to dynamical systems. Springer Verlag, New York, Berlin, Heidelberg.

Allingham, D., West, M., Mees, A. I., 1998. Wavelet reconstruction of nonlinear dynamics. International Journal of Bifurcation and Chaos 8 (11), 2191–2201.

Alvarez-Ramirez, J., Ibarra-Valdez, C., Fernandez-Anaya, G., 2002. Complex dynamics in a simple stock market model. International journal of bifurcation and chaos in applied sciences and engineering 12 (7), 1565–1577.

Anderson, P. W., Arrow, K. J., Pines, D., 1988. The economy as an evolving complex system: The proceedings of the Evolutionary Paths of the Global Economy Workshop, (1987), Santa Fe, New Mexico. Addison-Wesley Pub. Co.

Andreadis, I., 2000. Self-criticality and stochasticity of an S&P 500 index time series. Chaos, Solitons & Fractals 11 (7), 1047–1059.

Ang, J. S., Pohlman, R. A., 1978. A note on the price behavior of far eastern stocks. Journal of International Business Studies 9 (1), 103–07.

Antoniou, A., Ergul, N., Holmes, P., 1997. Market efficiency, thin trading and nonlinear behaviour: Evidence from an emerging market. European Financial Management 3 (2), 175–90.

Antoniou, A., Vorlow, C. E., 2000. Recurrence plots and financial time series analysis. Neural Network World 10 (1-2), 131–146.

Antoniou, A., Vorlow, C. E., 2002. Recurrence Quantification Analysis of Wavelet pre-filtered Index Returns. Discussion paper, University of Durham, Dept. Economics and Finance.

Ariel, R. A., 1987. A monthly effect in stock returns. Journal of Financial Economics 18 (1), 161–74.

Arino, M., 1996. Forecasting time series via discrete wavelet transform. Discussion paper, IESE, Universidad de Navarra.

Arino, M., Morettin, P., Vidakovic, B., 1995. On wavelet scalograms and their applications in economic time series. Preprint, ISDS working papers, Duke University.

Ashley, R., Patterson, D., 1996. The importance of being nonlinear: A frequency-domain approach to nonlinear model identification and estimation. In: Barnett, W. A., Kirman, A. P., Salmon, M. (Eds.), Nonlinear dynamics and economics: Proceedings of the Tenth International Symposium in Economic Theory and Econometrics. Cambridge University Press, International Symposia in Economic Theory and Econometrics series, pp. 297–306.

Ashley, R. A., Hinich, M. J., Patterson, D. M., 1990. Nonlinear serial dependence in industrial stock returns. In: Lawrence, K. D., Guerard, John B., J., Reeves, G. R. (Eds.), Advances in mathematical programming and financial planning. Vol. 2. JAI Press, pp. 163–81.

Ashley, R. A., Patterson, D. M., 1989. Linear versus nonlinear macroeconomies: A statistical test. International Economic Review 30 (3), 685–704.

Baas, T. A., 1999. The predictors, 1st Edition. H. Holt and Co., New York.

Baillie, R. T., Bollerslev, T., Mikkelsen, H. O., 1996. Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 74 (1), 3–30.

Baillie, R. T., Bollerslev, T. P., 1987. A multivariate generalized ARCH approach to modeling risk premia in forward foreign exchange rate markets. Michigan State Econometrics and Economic Theory Workshop Paper 8708.

Bak, P., 1996. How nature works: the science of self-organized criticality. Copernicus, New York, USA.

Bak, P., Paczuski, M., Shubik, M., 1996. Price variations in a stock market with many agents. Yale Cowles Foundation Discussion Paper 1132.

Bak, P., Tang, C., 1990. Self-organized criticality. Physics Today , 27–28.

Bak, P., Tang, C., Wiesenfeld, K., 1987. Self-organized criticality: an explanation of 1/f noise. Physical Review Letters 59, 381–384.

Bak, P., Tang, C., Wiesenfeld, K., 1988. Self-organized criticality. Physical Review A 38, 364–374.

Ball, R., Kothari, S. P., 1989. Nonstationary expected returns: Implications for tests of market efficiency and serial correlation in returns. Journal of Financial Economics 25 (1), 51–74.

Banz, R. W., 1981. The relationship between return and market value of common stocks. Journal of Financial Economics 9 (1), 3–18.

Barnes, P., 1986. Thin trading and stock market efficiency: The case study of the Kuala Lumpur stock exchange. Journal of Business Finance & Accounting 13 (4), 289–307.

Barnett, W. A., Chen, P., 1988. The aggregation-theoretic monetary aggregates are chaotic and have strange attractors: An econometric application of mathematical chaos. In: Barnett, W. A., Berndt, E. R., White, H. (Eds.), Dynamic econometric modeling: Proceedings of the Third International Symposium in Economic Theory and Econometrics. Cambridge University Press, International Symposia in Economic Theory and Econometrics series, pp. 199–245.

Barnett, W. A., Gallant, A. R., Hinich, M. J., Jungeilges, J. A., Kaplan, D. T., Jensen, M. J., 1998. A single-blind controlled competition among tests for nonlinearity and chaos. Journal of Econometrics 82 (1), 157–193.

Barnett, W. A., Medio, A., Serletis, A., 1997. Nonlinear and complex dynamics in economics. Wustl economics working papers.

Barnett, W. A., Serletis, A., 2000. Martingales, nonlinearity, and chaos. Journal of Economic Dynamics and Control 24 (5-7), 703–724.

Basu, S., 1977. Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. Journal of Finance 32 (3), 663–82.

Basu, S., 1983. The relationship between earnings' yield, market value and return for NYSE common stocks: Further evidence. Journal of Financial Economics 12 (1), 129–56.

Belaire-Franch, J., Contreras, D., Tordera-Lledo, L., 2002. Assessing nonlinear structures in real exchange rates using recurrence plot strategies. Physica D: Nonlinear Phenomena 171 (4), 249–264.

Bera, A., Jarque, C., 1981a. Efficient tests for normality, heteroskedasticity and serial independence of regression residuals: Monte-Carlo evidence. Economics Letters 7, 313–318.

Bera, A., Jarque, C., 1981b. Efficient Tests for Normality, Heteroskedasticity and Serial Independence of Regression Residuals: Monte-Carlo evidence. Economics Letters 7, 313–318.

Bethlehem, R. W., 1997. Fundamental analysis: Its limitations and use in portfolio selection. Unpublished master's thesis, University of South Africa.

Bhandari, L. C., 1988. Debt/equity ratio and expected common stock returns: Empirical evidence. Journal of Finance 43 (2), 507–28.

Black, F., 1986. Noise. Journal of Finance 41 (3), 529–43.

Black, F., 1990. Mean reversion and consumption smoothing. Review of Financial Studies 3 (1), 107–14.

Bollerslev, T., 1986. Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 31 (3), 307–27.

Bollerslev, T., 1987. A conditionally heteroskedastic time series model for speculative prices and rates of return. Review of Economics and Statistics 69 (3), 542–47.

Bollerslev, T., Chou, R. Y., Kroner, K. F., 1992. ARCH modeling in finance: A review of the theory and empirical evidence. Journal of Econometrics 52 (1-2), 5–59.

Bollerslev, T., Engle, R. F., Nelson, D. B., 1994. ARCH models. In: Engle, R. F., McFadden, D. L. (Eds.), Handbook of econometrics. Vol. 4. Elsevier North-Holland, Handbooks in Economics, vol. 2. Amsterdam; London and New York, pp. 2959–3038.

Bollerslev, T., Hodrick, R. J., 1992. Financial market efficiency tests. National Bureau of Economic Research Working Paper 4108.

Boudoukh, J., Richardson, M. P., Whitelaw, R. F., 1994. A tale of three schools: Insights on autocorrelations of short-horizon stock returns. Review of Financial Studies 7 (3), 539–73.

Box, G. E. P., Jenkins, G. M., 1976. Time series analysis: forecasting and control, rev Edition. Holden-Day, San Francisco.

Brock, W., Lakonishok, J., LeBaron, B., 1992. Simple technical trading rules and the stochastic properties of stock returns. Journal of Finance 47 (5), 1731–64.

Brock, W. A., 1986. Distinguishing random and deterministic systems: Abridged version. Journal of Economic Theory 40 (1), 168–95.

Brock, W. A., 1988. Nonlinearity and complex dynamics in economics and finance. In: Anderson, P. W., Arrow, K. J., Pines, D. (Eds.), The economy as an evolving complex system: The proceedings of the Evolutionary Paths of the Global Economy Workshop, (1987), Santa Fe, New Mexico. Addison-Wesley, Santa Fe Institute Studies in the Sciences of Complexity, vol. 5, pp. 77–97.

Brock, W. A., 1990. Chaos and complexity in economic and financial science. In: von Furstenberg, G. M. (Ed.), Acting under uncertainty: Multidisciplinary conceptions. Kluwer Academic, Theory and Decision Library. Series A, pp. 423–50.

Brock, W. A., 1991. Understanding macroeconomic time series using complex systems theory. Structural Change and Economic Dynamics 2 (1), 119–141.

Brock, W. A., 1993. Pathways to Randomness in the Economy: Emergent Nonlinearity and Chaos in Economics and Finance. Estudios Economicos 8 (1), 3–55.

Brock, W. A., 2000. Whither nonlinear? Journal of Economic Dynamics and Control 24 (5-7), 663–78.

Brock, W. A., Baek, E. G., 1991. Some theory of statistical inference for nonlinear science. Review of Economic Studies 58 (4), 697–716.

Brock, W. A., de Lima, P. J. F., 1996. Nonlinear time series, complexity theory, and finance. In: Maddala, G. S., Rao, C. R. (Eds.), Statistical methods of finance. Elsevier North-Holland, Handbook of Statistics series, vol. 14. Amsterdam, pp. 317–61.

Brock, W. A., Dechert, W. D., 1991. Non-linear dynamical systems: Instability and chaos in economics. In: Hildenbrand, W., Sonnenschein, H. (Eds.), Handbook of mathematical economics. Vol. 4. North-Holland, Handbooks in Economics, no. 1, Amsterdam, pp. 2209–35.

Brock, W. A., Dechert, W. D., Scheinkman, J. A., 1987. A test for independence based on the correlation dimension. Discussion paper, Social Systems Research Institute, University of Wisconsin.

Brock, W. A., Hommes, C. H., 1997. A rational route to randomness. Econometrica 65 (5), 1059–95.

Brock, W. A., Hommes, C. H., 1998. Heterogeneous beliefs and routes to chaos in a simple asset pricing model. Journal of Economic Dynamics and Control 22 (8-9), 1235–1274.

Brock, W. A., Hsieh, D. A., LeBaron, B., 1991. Nonlinear dynamics, chaos, and instability: Statistical theory and economic evidence. MIT Press.

Brock, W. A., Magill, M. J. P., 1979. Dynamics under uncertainty. Econometrica 47 (4), 843–68.

Brock, W. A., Sayers, C. L., 1988. Is the business cycle characterized by deterministic chaos? Journal of Monetary Economics 22 (1), 71–90.

Brockett, P., Hinich, M. J., 1983. Probability bounds on downtimes. Graduate School of Business, University of Texas at Austin, working paper.

Brockett, P. L., Hinich, M. J., Patterson, D., 1988. Bispectral-based tests for the detection of Gaussianity and linearity in time series. Journal of the American Statistical Association 83 (403), 657–64.

Brooks, C., Henry, O. T., 2000a. Can portmanteau nonlinearity tests serve as general mis-specification tests? evidence from symmetric and asymmetric GARCH models. Economics Letters 67 (3), 245–51.

Brooks, C., Henry, O. T., 2000b. Linear and non-linear transmission of equity return volatility: Evidence from the Us, Japan and Australia. Economic Modelling 17 (4), 497–513.

Bruce, A., Donoho, D., Gao, H.-Y., 1996. Wavelet analysis (for signal processing). IEEE Spectrum 33 (10), 26–35.

Bruce, A., Gao, H.-Y., 1996a. Applied Wavelet Analysis with S-PLUS. Springer, New York.
URL http://www.springer-ny.com/catalog/np/jul96np/DATA/0-387-94714-0.html

Bruce, A., Gao, H.-Y., 1996b. Understanding WaveShrink: Variance and bias estimation. Biometrika 83 (4).
URL ftp://ftp.statsci.com/pub/gao/varbias.ps.Z

Buckheit, J., Chen, S., Donoho, D., Johnstone, I., Scargle, J., 1995. Wavelab reference manual.
URL citeseer.nj.nec.com/article/buckheit95wavelab.html

Burrus, C. S., Gopinath, R. A., Guo, H., 1997. An Introduction to Wavelets and Wavelet Transforms: A Primer, prentice hall Edition. Prentice Hall, Paramus.

Campbell, J. Y., Lo, A. W., MacKinlay, A. C., 1997. The econometrics of financial markets. Princeton University Press, Princeton, N.J.

Campbell, J. Y., Shiller, R. J., 1988. Stock prices, earnings, and expected dividends. Journal of Finance 43 (3), 661–76.

Capobianco, E., 1988. Statistical analysis of financial volatility by wavelet shrinkage. Working paper.

Capobianco, E., 1997a. On the multiresolution analysis of financial time series models. In: 6th International PASE Workshop, Computers in Finance and Economics. Marianske Lazne, Czech Republic.

Capobianco, E., 1997b. Wavelet-denoised financial time series. In: International Workshop on Stochastic Model Building and Variable Selection.

Capobianco, E., 1999a. High frequency stock returns volatility: Feature extraction and pattern recognition via wavelet transforms. Working paper.

Capobianco, E., 1999b. Statistical analysis of financial volatility by wavelet shrinkage. Methodology And Computing In Applied Probability 1 (4), 423–443.

Capobianco, E., 1999c. Wavelets for high frequency financial time series. Working paper.

Capobianco, E., 2001. Wavelet transforms for the statistical analysis of returns generating stochastic processes. International Journal of Theoretical and Applied Finance 4 (3), 511–34.

Capobianco, E., 2002. Multiresolution approximation for volatility process. Quantitative Finance 2 (2), 91–110.

Carmona, R., Hwang, W.-L., Torr Esani, B., 1998. Practical Time-Frequency Analysis: Gabor and Wavelet Transforms with an Implementation in Wavelet Analysis and Its Applications Vol. 9. Academic Press, San Diego.

Casdagli, M., DesJardins, D., Eubank, S., Farmer, J. D., Gibson, J., Theiler, J., Hunter, N., 1992. Nonlinear modeling of chaotic time series: Theory and application. In: Kim, J. H., Stringer, J. (Eds.), Applied chaos. Wiley, New York, pp. 335–380.

Castro, R., Sauer, T., 1997. Correlation dimension of attractors through interspike intervals. Phys. Rev. E 55 (1), 287–290.

Cecchetti, S. G., Lam, P.-S., Nelson, C. M., 1990. Mean reversion in equilibrium asset prices. American Economic Review 80 (3), 398–418.

Chan, K. C., 1988. On the contrarian investment strategy. Journal of Business 61 (2), 147–63.

Chan, L. K. C., Hamao, Y., Lakonishok, J., 1991. Fundamentals and stock returns in japan. Journal of Finance 46 (5), 1739–64.

Chatrath, A., Adrangi, B., Dhanda, K. K., 2002. Are commodity prices chaotic? Agricultural Economics 27 (2), 123–37.

Chatrath, A., Adrangi, B., Shank, T., 2001. Nonlinear dependence in gold and silver futures: Is it chaos? American Economist 45 (2), 25–32.

Chen, S.-H., Lux, T., Marchesi, M., 2001. Testing for non-linear structure in an artificial financial market. Journal of Economic Behavior & Organization 46 (3), 327–342.

Cheng, Q., 1999. Miscellanea. on time-reversibility of linear processes. Biometrika 86 (2), 483–486.

Chiarella, C., 1990. The elements of a nonlinear theory of economic dynamics. Lecture Notes in Economics and Mathematical Systems, vol. 343. (Springer).

Chiarella, C., 1992. Developments in nonlinear economic dynamics: Past, present and future. Working Paper 14, School of Finance and Economics, University of Technology, Sydney.

Choo, E. S., 1997. A topological test for chaos. In: Creedy, J., Martin, V. L. (Eds.), Nonlinear Economic Models: cross-sectional, time series and neural network applications. Elgar, Aldershot, U.K., pp. 177–190.

Chopra, N., Lakonishok, J., Ritter, J. R., 1992. Measuring abnormal performance: Do stocks overreact? Journal of Financial Economics 31 (2), 235–68.

Chui, C. K., 1992a. An Introduction to Wavelets. Academic Press, San Diego.

Chui, C. K., 1997. Wavelets: A Mathematical Tool for Signal Processing. Society for Industrial & Applied Mathematics, Philadelphia.

Chui, C. K. E., 1992b. Wavelets: A Tutorial in Theory and Applications. Academic Press, San Diego.

Cody, M. A., 1993. A wavelet analyzer. Dr Dobb's journal 18 (4), 44.

Cody, M. A., 1994. The wavelet packet transform. Dr Dobb's journal 20 (4), 44.

Coifman, R. R., Donoho, D., 1995. Time-invariant wavelet de-noising. In: Antoniadis, A., Oppenheim, G. (Eds.), Wavelets and Statistics. Vol. 103 of Lecture Notes in Statistics. Springer-Verlag, New York, pp. 125–150.
URL ftp://playfair.stanford.edu/pub/donoho/TIDeNoise.ps.Z

Conover, W. J., 1971. Practical nonparametric statistics. Wiley, New York.

Conrad, J., Kaul, G., 1988. Time-variation in expected returns. Journal of Business 61 (4), 409–25.

Conrad, J., Kaul, G., 1989. Mean reversion in short-horizon expected returns. Review of Financial Studies 2 (2), 225–40.

Conrad, J., Kaul, G., 1993. Long-term market overreaction or biases in computed returns? Journal of Finance 48 (1), 39–63.

Conrad, K., Juttner, D. J., 1973. Recent behaviour of stock market prices in Germany and the random walk hypothesis. Kyklos 26 (3), 576–99.

Cont, R., 2001. Empirical properties of asset returns: Stylized facts and statistical issues. Quantitative Finance 1 (2), 223–36.

Cooper, J. C. B., 1982. World stock markets: Some random walk tests. Applied Economics 14 (5), 515–31.

Cootner, P. H., 1962. Stock prices: Random vs. systematic changes. Industrial Management Review III (Spring 1962), 25–45.

Cootner, P. H., 1964. The random character of stock market prices. Cambridge, Mass., M.I.T. Press.

Copeland, T. E., Mayers, D., 1982. The value line enigma (1965-1978): A case study of performance evaluation issues. Journal of Financial Economics 10 (3), 289–321.

Cowles, A. I., 1933. Can stock market forecasters forecast? Econometrica 1 (3), 309–324.

Cowles, A. I., 1937. Some a posteriori probabilities in stock market action. Econometrica 5 (3), 280–294.

Cowles, A. I., 1960. A revision of previous conclusions regarding stock price behavior. Econometrica 28 (4), 909–915.

Crack, T. F., Ledoit, O., 1996. Robust structure without predictability: The "compass rose" pattern of the stock market. Journal of Finance 51 (2), 751–762.

Creedy, J., Lye, J., Martin, V. L., 1994. Non-linearities and the long-run real exchange rate distribution. In: Creedy, J., Martin, V. L. (Eds.), Chaos and non-linear models in economics: Theory and applications. Elgar, Aldershot, U.K., pp. 196–212.

Creedy, J., Martin, V. L., 1994a. A model of the distribution of prices. In: Creedy, J., Martin, V. L. (Eds.), Chaos and non-linear models in economics: Theory and applications. Elgar, Aldershot, U.K., pp. 100–110.

Creedy, J., Martin, V. L., 1994b. The strange attraction of chaos in economics. In: Creedy, J., Martin, V. L. (Eds.), Chaos and non-linear models in economics: Theory and applications. Elgar, Aldershot, U.K., pp. 7–29.

Creedy, J., Martin, V. L. e., 1994c. Chaos and non-linear models in economics: Theory and applications. Elgar, Aldershot, U.K.

Cribari-Neto, F., Jensen, M. J., 1997. Matlab as an econometric programming environment. Journal of Applied Econometrics 12 (6), 735–44.

Cromwell, J. B., Labys, W. C., Terraza, M., 1994. Univariate tests for time series models. Sage Publications, Thousand Oaks, Calif.

Cross, F., 1973. The behavior of stock prices on Fridays and Mondays. Financial Analysts Journal Nov.-Dec., 67–69.

Cutler, D. M., Poterba, J. M., Summers, L. H., 1989. What moves stock prices? Journal of Portfolio Management 15 (3), 4–12.

Cutler, D. M., Poterba, J. M., Summers, L. H., 1990. Speculative dynamics and the role of feedback traders. American Economic Review 80 (2), 63–68.

Cutler, D. M., Poterba, J. M., Summers, L. H., 1991. Speculative dynamics. Review of Economic Studies 58 (3), 529–46.

Daubechies, I., 1992. Ten Lectures on Wavelets. Society for Industrial & Applied Mathematics, Philadelphia.

Day, R. H., 1994. Complex economic dynamics: An introduction to dynamical systems and market mechanisms. Vol. 1. MIT Press.

De Bondt, W. F. M., Thaler, R. H., 1985. Does the stock market overreact? Journal of Finance 40 (3), 793–805.

De Bondt, W. F. M., Thaler, R. H., 1987. Further evidence on investor overreaction and stock market seasonality. Journal of Finance 42 (3), 557–81.

De Bondt, W. F. M., Thaler, R. H., 1989. A mean-reverting walk down Wall Street. Journal of Economic Perspectives 3 (1), 189–202.

De Grauwe, P., Dewachter, H., 1992. Chaos in the Dornbusch model of the exchange rate. Kredit und Kapital 25 (1), 26–54.

De Grauwe, P., Dewachter, H., Embrechts, M., 1993. Exchange rate theory: chaotic models of foreign exchange markets. Blackwell.

De Grauwe, P., Vansanten, K., 1990. Deterministic chaos in the foreign exchange market. Centre for Economic Policy Research Discussion Paper 370.

Devaney, R. L., 1989. An introduction to chaotic dynamical systems, 2nd Edition. Addison-Wesley, Redwood City, Calif.

Dickey, D. A., Fuller, W. A., 1979. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association 74 (366), 427–431.

Diks, C., 1996. Estimating invariants of noisy attractors. Phys. Rev. E 53 (5), 4263–4266.

Diks, C., 1999. Nonlinear time series analysis: methods and applications. Vol. 4 of World Scientific Series: Nonlinear time series and chaos. River Edge NJ, Singapore.

Diks, C., van Houwelingen, J. C., Takens, F., DeGoede, J., 1995. Reversibility as a criterion for discriminating time series. Phys. Lett. A 210 (2-3), 221–228.

Ding, M., Grebogi, C., Ott, E., Sauer, T., Yorke, J. A., 1993. Plateau onset of correlation dimension: When does it occur? Phys. Rev. Lett. 70, 3872–3875.

Dixon, D. D., Zak, M., Zbilut, J. P., 1994. A method for detecting possible non-determinism in a time series. XXX Preprint: `chao-dyn/9402004` .

Donoho, D. L., 1995. De-noising by soft-thresholding. IEEE Transactions on Information Theory 41 (3), 613–627.
URL `http://playfair.Stanford.EDU/reports/donoho/denoiserelease3.ps.Z`

Donoho, D. L., Johnstone, I. M., 1994. Ideal spatial adaptation by wavelet shrinkage. Biometrika 81 (3), 425–455.
URL `http://playfair.Stanford.EDU/reports/donoho/ausws.ps.Z`

Donoho, D. L., Johnstone, I. M., 1995. Adapting to unknown smoothness by wavelet shrinkage. Journal of the American Statistical Association 90, 1200–1224.
URL `http://playfair.Stanford.EDU/reports/donoho/isaws.ps.Z`

Dryden, M. M., 1970a. Filter tests of UK share prices. Applied Economics 1 (4), 261–75.

Dryden, M. M., 1970b. A statistical study of UK share prices. Scottish Journal of Political Economy 17 (3), 369–89.

Eckmann, J., Kamphorst, S. O., Ruelle, D., 1987. Recurrence plots of dynamical systems. Europhys. Lett. 4, 973–977.

Eckmann, J.-P., Ruelle, D., 1985. Ergodic theory of chaos and strange attractors. Rev. Mod. Phys. 57, 617–656.

Efron, B., 1979. Bootstrap methods: another look at the jackknife. Ann. Statist. 7, 1–26.

Efron, B., 1982. The Jackknife, the Bootstrap and Other Resampling Plans. SIAM.

Efron, B., Tibshirani, R., 1993. An introduction to the bootstrap. Chapman & Hall, New York, monographs on statistics and applied probability, vol. 57.

Engle, R. F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50 (4), 987–1007.

Engle, R. F., Granger, C. W. J., 1997. Co-integration and error correction: Representation, estimation, and testing. In: Hamouda, O. F., Rowley, J. C. R. (Eds.), Discrete and continuous systems, cointegration and chaos. Elgar, Elgar Reference Collection. Foundations of Probability, Econometrics and Economic Games, vol. 10, pp. 295–320.

Evertsz, C. J. G., 1995. Fractal geometry of financial time series. In: Evertsz, C. J. G., Peitgen, H. O., Voss, R. F., Mandelbrot, B. (Eds.), The Mandelbrot Festschrift. World Scientific.

Evertsz, C. J. G., Berkner, K., 1995. Large deviation and self-similarity analysis of graphs: Dax stock prices. Chaos Solitons and Fractals 6 (1), 121–130.

Falconer, K. J., 1990. Fractal geometry: mathematical foundations and applications. New York, Chichester.

Fama, E. F., 1965. The behaviour of stock market prices. Journal of Business 38, 34–105.

Fama, E. F., 1970. Efficient capital markets: A review of theory and empirical work. Journal of Finance 25 (2), 383–417.

Fama, E. F., 1971. Risk, return, and equilibrium. Journal of Political Economy 79 (1), 30–55.

Fama, E. F., 1991. Efficient capital markets: II. Journal of Finance 46 (5), 1575–617.

Fama, E. F., 1997. Mandelbrot on price variation (Mandelbrot and the Stable Paretian Hypothesis). In: Mandelbrot, B. B. (Ed.), Fractals and scaling in finance: Discontinuity, concentration, risk. Springer, pp. 444–56.

Fama, E. F., 1998. Market efficiency, long-term returns, and behavioral finance. Journal of Financial Economics 49 (3), 283–306.

Fama, E. F., et al., 1969. The adjustment of stock prices to new information. International Economic Review 10 (1), 1–21.

Fama, E. F., French, K. R., 1988a. Dividend yields and expected stock returns. Journal of Financial Economics 22 (1), 3–25.

Fama, E. F., French, K. R., 1988b. Permanent and temporary components of stock prices. Journal of Political Economy 96 (2), 246–73.

Fama, E. F., French, K. R., 1992. The cross-section of expected stock returns. Journal of Finance 47 (2), 427–65.

Fama, E. F., French, K. R., 1996. The CAPM is wanted, dead or alive. Journal of Finance 51 (5), 1947–58.

Fama, E. F., Jensen, M. C., 1984. Organizational forms and investment decisions. Working paper, Boston, Mass.

Fama, E. F., MacBeth, J. D., 1973. Risk, return, and equilibrium: Empirical tests. Journal of Political Economy 81 (3), 607–36.

Fang, Y., 2002. The compass rose and random walk tests. Computational Statistics & Data Analysis 39 (3), 299–310.

Farmer, D. J., Lo, A. W., 1999. From the Academy Frontiers of Science Symposium, Frontiers of finance: Evolution and efficient markets. Proceedings of the National Academy of Sciences of the United States of America 96 (18), 9991–9992.

Flandrin, P., 1998. Time-Frequency/Time-Scale Analysis. Academic Press, San Diego, Wavelet Analysis and Its Applications Series, Vol. 10.

Frank, M., Stengos, T., 1988a. The stability of Canadian macroeconomic data as measured by the largest lyapunov exponent. Economics Letters 27 (1), 11–14.

Frank, M. K., Gencay, R., Stengos, T., 1988. International chaos? European Economic Review 32 (8), 1569–84.

Frank, M. Z., Stengos, T., 1988b. Chaotic dynamics in economic time-series. Journal of Economic Surveys 2 (2), 103–33.

Frank, M. Z., Stengos, T., 1988c. Some evidence concerning macroeconomic chaos. Journal of Monetary Economics 22 (3), 423–38.

Franses, P. H., Dijk, D. v., 2000. Nonlinear time series models in empirical finance. New York, cambridge University Press.

Fraser, A. M., Swinney, H. L., 1986. Independent coordinates for strange attractors from mutual information. Phys. Rev. A 33 (2), 1134–1140.

Frazier, M. W., Axler, S., Gehring, F. W., Ribet, K. A., 1999. An Introduction to Wavelets Through Linear Algebra. Springer-Verlag, New York, undergraduate Texts in Mathematics.

Fredrikson, E. B., 1971. Frontiers of investment analysis. Intext Educational Publishers, Scranton, The Intext series in finance.

French, K. R., 1980. Stock returns and the weekend effect. Journal of Financial Economics 8 (1), 55–69.

Gabor, D., 1946. Theory of communication. Journal of the Institute for Electrical Engineers 93, 429–439.

Gandhi, D. K., Saunders, A., Woodward, R. S., 1980. Thin capital markets: A case study of the Kuwaiti stock market. Applied Economics 12 (3), 341–49.

Gao, H.-Y., Bruce, A., 1997. WaveShrink with firm shrinkage. Statistica Sinica 7 (4), 855–874.
URL ftp://ftp.statsci.com/pub/gao/firm.ps.Z

Gencay, R., Selcuk, F., Whitcher, B., 2001. An Introduction to Wavelets and Other Filtering Methods in Finance and Economics. Academic Press, San Diego.

Geweke, J., Porter-Hudak, S., 1983. The estimation and application of long memory time series models. Journal of Time Series Analysis 4, 221–238.

Ghose, D., Kroner, K. F., 1993. The relationship between GARCH and stable processes: Finding the source of fat tails in financial data. Discussion Paper 93-1.

Gilles, C., LeRoy, S. F., 1997. Econometric aspects of the variance-bounds tests: A survey. In: Lo, A. W. (Ed.), Market efficiency: Stock market behaviour in theory and practice. Vol. 2. Elgar Reference Collection. International Library of Critical Writings in Financial Economics, vol. 3., UK, pp. 65–103.

Gilmore, C. G., 1993. A new test for chaos. Journal of Economic Behavior & Organization 22 (2), 209–237.

Gilmore, C. G., 1996. Detecting linear and nonlinear dependence in stock returns: New methods derived from chaos theory. Journal of Business Finance & Accounting 23 (9-10), 1357–1377.

Gilmore, C. G., 2001. An examination of nonlinear dependence in exchange rates, using recent methods from chaos theory. Global Finance Journal 12 (1), 139–151.

Gleason, K. C., Lee, C. I., Mathur, I., 2000. An explanation for the compass rose pattern. Economics Letters 68 (2), 127–133.

Gleick, J., 1987. Chaos: making a new science. Viking, New York, N.Y., U.S.A.

Good, P. I., 1994. Permutation tests: a practical guide to resampling methods for testing hypotheses. Springer-Verlag,.

Goodhart, C. A. E., O'Hara, M., 2000. High-frequency data in financial markets: Issues and applications. In: Goodhart, C. A. E., Payne, R. (Eds.), The foreign exchange market: Empirical studies with high-frequency data. St. Martin's Press; London: Macmillan Press, New York, pp. 503–545.

Granger, C. W. J., 1964. Spectral analysis of economic time series. Princeton University Press, Princeton Studies in Mathematical Economics, no. 1.

Granger, C. W. J., 1980. Forecasting in business and economics. Academic Press, New York.

Granger, C. W. J., 1990. Modelling economic series: readings in econometric methodology. Oxford University Press.

Granger, C. W. J., 1994. Is chaotic economic theory relevant for economics? Journal of International and Comparative Economics 3, 139–145.

Granger, C. W. J., 1997. The typical spectral shape of an economic variable. In: Hamouda, O. F., Rowley, J. C. R. (Eds.), Time series models, causality and exogeneity. Elgar, Elgar Reference Collection. Foundations of Probability, Econometrics and Economic Games, vol. 8., UK, pp. 407–18.

Granger, C. W. J., 1999. Empirical modeling in economics: specification and evaluation. Cambridge University Press.

Granger, C. W. J., Andersen, A. P., 1978. An introduction to bilinear time series models. Vandenhoeck und Ruprecht, Gttingen.

Granger, C. W. J., Godfrey, M., Morgestern, O., 1964. The random-walk hypothesis of stock market behavior. Kyklos 17, 1–30.

Granger, C. W. J., Joyeux, R., 1997. An introduction to long-memory time series models and fractional differencing. In: Hamouda, O. F., Rowley, J. C. R. (Eds.), Discrete and continuous systems, cointegration and chaos. Elgar, Elgar Reference Collection. Foundations of Probability, Econometrics and Economic Games, vol. 10., UK, pp. 439–53.

Granger, C. W. J., Morgenstern, O., 1970. Predictability of stock market prices. Lexington, Mass., Heath Lexington Books.

Granger, C. W. J., Morgestern, O., 1963. Spectral analysis of new york stock market prices. Kyklos 16, 1–27.

Granger, C. W. J., Tersvirta, T., 1993. Modelling nonlinear economic relationships. Oxford University Press, Advanced texts in econometrics.

Graps, A., 1995. An introduction to wavelets. IEEE Computational Sciences and Engineering 2 (2), 50–61.

Grassberger, P., 1983a. Generalized dimensions of strange attractors. Phys. Lett. A 97, 227–230.

Grassberger, P., 1983b. On the fractal dimension of the Hénon attractor. Phys. Lett. A 97, 224–226.

Grassberger, P., Procaccia, I., 1983a. Estimation of the Kolmogorov entropy from a chaotic signal. Phys. Rev. A 28, 2591–2593.

Grassberger, P., Procaccia, I., 1983b. Measuring the strangeness of strange attractors. Physica D 9, 189–208.

Grassberger, P., Procaccia, I., 1983c. On the characterization of strange attractors. Phys. Rev. Lett. 50, 346–349.

Grebogi, C., Ott, E., Yorke, J. A., 1982. Chaotic attractors in crisis. Phys. Rev. Lett. 48, 1507–1510.

Greenblatt, S. A., 1996. Wavelets in econometrics: An application to outlier testing. In: Gilli, M. (Ed.), Computational Economic Systems: Models, Methods & Econometrics. Vol. 5 of Advances in Computational Economics. Kluwer Academic Publishers, Dordrecht, pp. 139–160.

Gwilym, O., Sutcliffe, C., 2001. Problems encountered when using high frequency financial market data: Suggested solutions. Journal of Financial Management and Analysis 14 (1), 38–51.

Haar, A., 1910. Zur theorie der orthogonalen funktionensysteme. Mathematische Annalen 69, 331–371.

Hallin, M., Lefevre, C., Puri, M., 1988. On time-reversibility and the uniqueness of moving average representations for non-Gaussian stationary time-series. Biometrika (1), 170–171.

Harris, L., 1986. A transaction data study of weekly and intradaily patterns in stock returns. Journal of Financial Economics 16 (1), 99–117.

Hazewinkel, M., 1997. Encyclopaedia of Mathematics. Kluwer Academic Publishers, Norwell.

Hegger, R., Kantz, H., 1997. Embedding of sequences of time intervals. Europhys. Lett. 38 (4), 267–272.

Hegger, R., Kantz, H., 1999. Improved false nearest neighbor method to detect determinism in time series data. Phys. Rev. E 60 (4), 4970–4973.

Hegger, R., Kantz, H., Schreiber, T., 1999. Practical implementation of nonlinear time series methods: The TISEAN package. Chaos 9 (2), 413–435.

Hiemstra, C., Kelejian, H. H., 1992. A small sample issue relating to tests for normality. Working Paper 92-5, University of Maryland, Department of Economics.

Hinich, M. J., 1982. Testing for Gaussianity and linearity of a stationary time series. Journal of Time Series Analysis 3, 169–176.

Hinich, M. J., Patterson, D. M., 1985. Evidence of nonlinearity in daily stock returns. Journal of Business and Economic Statistics 3 (1), 69–77.

Hinich, M. J., Patterson, D. M., 1989. Evidence of nonlinearity in the trade-by-trade stock market return generating process. In: Barnett, W. A., Geweke, J., Shell, K. (Eds.), Economic complexity: Chaos, sunspots, bubbles, and nonlinearity. Proceedings of the Fourth International Symposium in Economic Theory and Econometrics. Cambridge University Press, International Symposia in Economic Theory and Econometrics series, pp. 383–409.

Hinich, M. J., Patterson, D. M., 1993. Intraday nonlinear behavior of stock prices. In: Day, R. H., Chen, P. (Eds.), Nonlinear dynamics and evolutionary economics. Oxford University Press, Oxford, pp. 201–14.

Hinich, M. J., Rothman, P., 1998. Frequency-domain test of time reversibility. Macroeconomic Dynamics 2 (1), 72–88.

Holland, J. H., 1995. Hidden order: how adaptation builds complexity. Addison-Wesley, Reading, Mass.

Hommes, C. H., 2002. Papers from the Arthur M. Sackler Colloquium of the National Academy of Sciences: Economic Agents and Markets as Emergent Phenomena, Modeling the stylized facts in finance through simple nonlinear adaptive systems. Proceedings of the National Academy of Sciences of the United States of America 99 (3), 7221–7229.

Hong, H., 1978. Predictability of price trends on stock exchanges: A study of some far eastern countries. Review of Economics and Statistics 6 (4), 619–21.

Hosking, J. R. M., 1981. Fractional differencing. Biometrika 68, 165–176.

Hsieh, D. A., 1991. Chaos and nonlinear dynamics: Application to financial markets. Journal of Finance 46 (5), 1839–77.

Hubbard, B. B., 1998. The World According to Wavelets: The Story of a Mathematical Technique in the Making, 2nd Edition. A. K. Peters Limited Publisher Record, Natick.

Ihaka, R., Gentleman, R., 1996. R: A language for data analysis and graphics. Journal of Computational and Graphical Statistics 5 (3), 299–314.

Ikegawa, S., Shinohara, M., Fukunaga, T., Zbilut, J. P., Webber, C. L., 2000. Nonlinear time-course of lumbar muscle fatigue using recurrence quantifications. Biological cybernetics 82 (5), 373–383.

Iwanski, J. S., Bradley, E., 1998. Recurrence plots of experimental data: To embed or not to embed? Chaos 8 (4), 861–871.

Jansen, M., 2001. Noise Reduction by Wavelet Thresholding. Springer-Verlag, New York, lecture Notes in Statistics; Vol. 161.

Jawerth, 1994. Practical Guide to Wavelets. CRC Press, Boca Raton.

Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. Journal of Finance 48 (1), 65-91.

Jennergren, L. P., Korsvold, P. E., 1974. Price formation in the Norwegian and Swedish stock markets-some random walk tests. Swedish Journal of Economics 76 (2), 171-85.

Jensen, M. C., 1969. Risk, the pricing of capital assets, and the evaluation of investment portfolios. Journal of Business 42 (2), 167-247.

Jensen, M. C., 1972. Studies in the theory of capital markets. Praeger, New York.

Jensen, M. C., 1975. Tests of capital market theory and implications of the evidence. Financial Analysts Research Foundation.

Jensen, M. C., 1978. Some anomalous evidence regarding market efficiency. Journal of Financial Economics 6 (2/3), 95-101.

Jensen, M. C., Bennington, G. A., 1970. Random walks and technical theories: Some additional evidence. Journal of Finance 25 (2), 469-82.

Jensen, M. C., Smith, C. W., 1984. The Modern theory of corporate finance. McGraw-Hill, New York.

Jensen, M. J., 1994. Wavelet analysis of fractionally integrated processes. Working paper, University of Missouri.

Jensen, M. J., 1997. Making wavelets in finance. Financial Engineering News (1), 1-10.

Jensen, M. J., 1999a. An approximate wavelet mle of short- and long-memory parameters. Studies in Nonlinear Dynamics and Econometrics 3 (4), 239-53.

Jensen, M. J., 1999b. Using wavelets to obtain a consistent ordinary least squares estimator of the fractional differencing parameter. Journal of Forecasting 18, 17–32.

Jensen, M. J., 2000. An alternative maximum likelihood estimator of long-memory processes using compactly supported wavelets. Journal of Economic Dynamics and Control 24 (3), 361–87.

Jensen, M. J., Whitcher, B., 2000. Wavelet-based estimation procedures for seasonal long-memory models. In: Computing in Economics and Finance 2000.

Judd, K., Mees, A., 1998. Embedding as a modeling problem. Physica D 120 (3-4), 273–286.

Judd, K., Small, M., 2000. Towards long-term prediction. Physica D 136 (1-2), 31–44.

Kaiser, G., 1999. A Friendly Guide to Wavelets. Birkhauser Boston Secaucus Springer-Verlag, New York, Cambridge.

Kantz, H., Schreiber, T., 1997. Nonlinear time series analysis. Cambridge University Press, Cambridge nonlinear science series, vol. 7.

Kaplan, D. T., Glass, L., January 27 1992. Direct test for determinism in a time series. Physical review letters 68 (4), 427–430.

Kaplan, D. T., Glass, L., 1995. Understanding nonlinear dynamics. Springer-Verlag, New York, textbooks in mathematical sciences.

Keenan, D. M., 1985. A tukey nonadditivity-type test for time series nonlinearity. Biometrika 72, 39–44.

Keim, D. B., 1983. Size-related anomalies and stock return seasonality: Further empirical evidence. Journal of Financial Economics 12 (1), 13–32.

Kemp, A. G., Reid, G. C., 1971. The random walk hypothesis and the recent behaviour of equity prices in Britain. Economica, N.S. 38 (149), 28–51.

Kendall, M. G., Bradford Hill, A., 1953. The Analysis of Economic Time-Series Part I: Prices. Journal of the Royal Statistical Society. Series A 116 (1), 11–34.

Kennel, M. B., 1997. Statistical test for dynamical nonstationarity in observed time-series data. Phys. Rev. E 56 (1), 316–321.

Kennel, M. B., Abarbanel, H. D. I., 1994. False neighbors and false strands: A reliable minimum embedding dimension algorithm. INLS Preprint .

Kennel, M. B., Brown, R., Abarbanel, H. D. I., 1992. Determining embedding dimension for phase-space reconstruction using a geometrical construction. Phys. Rev. A 45, 3403–3411.

Kennel, M. B., Isabelle, S., 1992. Method to distinguish chaos from colored noise and to determine embedding parameters. Phys. Rev. A 46, 3111–3118.

Kim, M. J., Nelson, C. R., Startz, R., 1991. Mean reversion in stock prices? a reappraisal of the empirical evidence. Review of Economic Studies 58 (3), 515–28.

Klimasauskas, K., 1992. Wavelets. Advanced Technology for Developers 1 (3).

Kolmogorov, A., 1959. Entropy per unit time as a metric invariant of automorphisms. Dokl. Akad. Nauk., SSSR 124, 124:754.

Kononov, E., November 15, 1999. Visual recurrence analysis (VRA). Version 4.2.
URL http://pw1.netcom.com/~eugenek

Kramer, W., Runde, R., 1997. Chaos and the compass rose. Economics Letters 54 (2), 113–118.

Kugiumtzis, D., 1996. State space reconstruction parameters in the analysis of chaotic time series - the role of the time window length. Physica D 95, 13–28.

Kugiumtzis, D., 1997a. Assessing different norms in nonlinear analysis of noisy time series. Physica D 105, 62–78.

Kugiumtzis, D., 1997b. Correction of the correlation dimension for noisy time series. International Journal of Bifurcation and Chaos 7 (6), 1283–1294.

Kugiumtzis, D., 1999. Test your surrogate data before you test for nonlinearity. Phys. Rev. E 60 (3), 2808–2816.

Kugiumtzis, D., 2000. Surrogate data test for nonlinearity including nonmonotonic transforms. Phys. Rev. E 62 (1), R25–R28.

Kugiumtzis, D., 2001. On the reliability of the surrogate data test for nonlinearity in the analysis of noisy time series. International Journal of Bifurcation and Chaos 11 (7), 1881–1896.

Kugiumtzis, D., Lillekjendlie, B., Christophersen, N., 1993. Chaotic time series, part I. estimation. Modeling Identification & Control 15 (4), 205–224.

Kugiumtzis, D., Lillekjendlie, B., Christophersen, N., 1994. Chaotic time series, part II: System identification and prediction. Modeling Identification & Control 15 (4), 225–243.

Kugiumtzis, D., Lingjaerde, O. C., Christophersen, N., 1998. Regularized local linear prediction of chaotic time series. Physica D 112 (3-4), 344–360.

Lakonishok, J., Levi, M., 1982. Weekend effects on stock returns: A note. Journal of Finance 37 (3), 883–89.

Lakonishok, J., Shleifer, A., Vishny, R. W., 1994. Contrarian investment, extrapolation, and risk. Journal of Finance 49 (5), 1541–78.

Lane, J. A., Peel, D. A., Raeburn, E. J., 1996. Some empirical evidence on the time-series properties of four UK asset prices. Economica 63 (251), 405–26.

Lawrence, A., 1991. Directionality and reversibility in time series. International Statistical Review 59, 67–79.

LeBaron, B., 1988. Stock return nonlinearities: Some initial tests and findings. Ph.d., University of Chicago.

LeBaron, B., 1992. Some relations between volatility and serial correlations in stock market returns. Journal of Business 65 (2), 199–219.

LeBaron, B., 1994. Chaos and nonlinear forecastability in economics and finance. Philosophical transactions. Physical sciences and engineering (1688), 397.

LeBaron, B., 1999. Technical trading rule profitability and foreign exchange intervention. Journal of International Economics 49 (1), 125–143.

LeBaron, B., 2000. Agent-based computational finance: Suggested readings and early research. Journal of Economic Dynamics and Control 24 (5-7), 679–702.

LeBaron, B., Arthur, W. B., Palmer, R., 1999. Time series properties of an artificial stock market. Journal of Economic Dynamics and Control 23 (9-10), 1487–1516.

Lehmann, B. N., 1990. Fads, martingales, and market efficiency. Quarterly Journal of Economics 105 (1), 1–28.

LeRoy, S. F., 1989. Efficient capital markets and martingales. Journal of Economic Literature 27 (4), 1583–1621.

LeRoy, S. F., 1996. Stock price volatility. In: Maddala, G. S., Rao, C. R. (Eds.), Statistical methods of finance. Elsevier North-Holland, Handbook of Statistics series, vol. 14. Amsterdam, pp. 193–208.

LeRoy, S. F., 1997. Risk aversion and the martingale property of stock prices. In: Lo, A. W. (Ed.), Market efficiency: Stock market behaviour in theory and practice. Vol. 1. Elgar, International Library of Critical Writings in Financial Economics, vol. 3, pp. 82–92.

Liu, T., Granger, C. W. J., Heller, W. P., 1992. Using the correlation exponent to decide if an economic series is chaotic. Working Paper, University of California 91-21R.

Ljung, G. M., Box, G. E. P., 1994. On a measure of lack of fit in time series models. In: Harvey, A. (Ed.), Time series. Vol. 1. Elgar; distributed in the U.S. by Ashgate Brookfield Vt., Elgar Reference Collection. International Library of Critical Writings in Econometrics, vol. 5. Aldershot, U.K., pp. 82–88.

Lo, A. W., 1989. Long-term memory in stock market prices. Working Papers in Economics (AWPE) Database 2984.

Lo, A. W., MacKinlay, A. C., 1987a. A simple specification test of the random walk hypothesis. Working Papers in Economics (AWPE) Database 8612.

Lo, A. W., MacKinlay, A. C., 1987b. Stock market prices do not follow random walks: evidence from a simple specification test. NBER working paper series 2168, National Bureau of Economic Research, Cambridge, MA.

Lo, A. W., MacKinlay, A. C., 1988. The size and power of the variance ratio test in finite samples: a monte-Carlo investigation. NBER technical working paper series 66, National Bureau of Economic Research, Cambridge, MA.

Lo, A. W., MacKinlay, A. C., 1989. An econometric analysis of nonsynchronous-trading. Working Papers in Economics (AWPE) Database 2960.

Lo, A. W., MacKinlay, A. C., 1990a. Data-snooping biases in tests of financial asset pricing models. Review of Financial Studies 3 (3), 431–467.

Lo, A. W., MacKinlay, A. C., 1990b. When are contrarian profits due to stock market overreaction? Review of Financial Studies 3 (2), 175–205.

Lo, A. W., MacKinlay, A. C., 1997. Maximizing predictability in the stock and bond markets. Macroeconomic Dynamics 1 (1), 102–34.

Lo, A. W., MacKinlay, A. C., 1999. A non-random walk down Wall Street. Princeton University Press.

Los, C. A., 1999. Nonparametric testing of the high-frequency efficiency of the 1997 asian foreign exchange markets. Journal of Multinational Financial Management 9 (3-4), 265–89.

Los, C. A., 2000. Nonparametric efficiency testing of asian stock markets using weekly data. In: Fomby, T. B., Hill, R. C. (Eds.), Applying kernel and nonparametric estimation to economic topics. JAI Press, pp. 329–63, advances in Econometrics, vol. 14.

Los, C. A., July 2003. Financial Market Risk: Measurement & Analysis. Routledge International Studies in Money and Banking. Taylor & Francis Books Ltd, London, UK.

Mallat, S., 1989. A theory for multiscale signal decomposition: The wavelet representation. IEEE Trans. on Pattern and Machine Intelligence 11 (7), 674–693.

Mallat, S. G., 1998. A Wavelet Tour of Signal Processing. Morgan Kaufmann Publishers, San Francisco.

Malliaris, A. G., Stein, J. L., 1999a. Financial modelling: From stochastics to chaotics and back to stochastics. In: Malliaris, A. G. (Ed.), Foundations of futures markets: Selected essays of A. G. Malliaris. Elgar, pp. 225–40.

Malliaris, A. G., Stein, J. L., 1999b. Methodological issues in asset pricing: Random walk or chaotic dynamics. Journal of Banking and Finance 23 (11), 1605–35.

Mandelbrot, B. B., 1971. When can price be arbitraged efficiently? a limit to the validity of the random walk and martingale models. Review of Economics and Statistics 53 (3), 225–36.

Mandelbrot, B. B., 1972a. Correction of an error in "the variation of certain speculative prices" (1963). Journal of Business 45 (4), 542–43.

Mandelbrot, B. B., 1972b. Statistical methodology for nonperiodic cycles: From the covariance to R/S analysis. Annals of Economic and Social Measurement 1 (3), 259–90.

Mandelbrot, B. B., 1977. Fractals: form, chance, and dimension. W.H. Freeman, San Francisco.

Mandelbrot, B. B., 1982. The fractal geometry of nature. W.H. Freeman, San Francisco.

Mandelbrot, B. B., 1997a. Fractals and scaling in finance: Discontinuity, Concentration, Risk: selecta volume E. Springer, New York.

Mandelbrot, B. B., 1997b. Limitations of efficiency and of martingale models. In: Mandelbrot, B. B. (Ed.), Fractals and scaling in finance: Discontinuity, concentration, risk. Springer, Selecta, vol. E., pp. 492–512.

Mandelbrot, B. B., 1997c. Nonlinear forecasts, rational bubbles, and martingales. In: Mandelbrot, B. B. (Ed.), Fractals and scaling in finance: Discontinuity, concentration, risk. Springer, Selecta, vol. E., pp. 471–91.

Mandelbrot, B. B., 1997d. Self-affine variation in fractal time. In: Mandelbrot, B. B. (Ed.), Fractals and scaling in finance: Discontinuity, concentration, risk. Springer, Selecta, vol. E., pp. 513–25.

Mandelbrot, B. B., 1997e. Three fractal models in finance: Discontinuity, concentration, risk. Economic Notes 26 (2), 171–211.

Mandelbrot, B. B., 1997f. The variation of certain speculative prices. In: Mandelbrot, B. B. (Ed.), Fractals and scaling in finance: Discontinuity, concentration, risk. Springer, Selecta, vol. E., pp. 371–418.

Mandelbrot, B. B., 1997g. The variation of the prices of cotton, wheat, and railroad stocks, and of some financial rates. In: Mandelbrot, B. B. (Ed.), Fractals and scaling in finance: Discontinuity, concentration, risk. Springer, Selecta, vol. E., pp. 419–43.

Mandelbrot, B. B., 1999a. A multifractal walk down Wall Street. Scientific American 280 (2), 70–74.

Mandelbrot, B. B., 1999b. Multifractals and 1/f noise: wild self-affinity in physics (1963-1976): selecta volume N. Springer, New York.

Mandelbrot, B. B., 1999c. Renormalization and fixed points in finance, since 1962. Physica A 263 (1), 477–491.

Mandelbrot, B. B., 1999d. Survey of multifractality in finance. Yale Cowles Foundation Discussion Paper 1238.

Mandelbrot, B. B., 2001a. Scaling in financial prices I: Tails and dependence. Quantitative Finance 1 (1), 113–23.

Mandelbrot, B. B., 2001b. Scaling in financial prices II: Multifractals and the star equation. Quantitative Finance 1 (1), 124–30.

Mandelbrot, B. B., 2001c. Scaling in financial prices III: Cartoon Brownian motions in multifractal time. Quantitative Finance 1 (4), 427–40.

Mandelbrot, B. B., 2001d. Scaling in financial prices IV: Multifractal concentration. Quantitative Finance 1 (6), 641–49.

Mandelbrot, B. B., 2001e. Stochastic volatility, power laws and long memory. Quantitative Finance 1 (6), 558–59.

Mandelbrot, B. B., 2002. Gaussian self-affinity and fractals: Globality, the Earth, 1/f noise, and R/S. Springer.

Mandelbrot, B. B., Ness, J. W. v., 1968. Fractional Brownian motion, fractional noises and application. SIAM Review 10, 422–437.

Manetti, C., Giuliani, A., Ceruso, M.-A., Webber, C., Zbilut, J. P., 2001. Recurrence analysis of hydration effects on nonlinear protein dynamics: multiplicative scaling and additive processes. Physics letters: A 281 (5), 317–324.

Marchesi, M., Lux, T., 1999. Scaling and criticality in a stochastic multi-agent model of financial mar. Nature 397, 498 – 500.

Martinerie, J. M., Albano, A. M., Mees, A. I., Rapp, P. E., 1992. Mutual information, strange attractors, and the optimal estimation of dimension. Phys. Rev. A 45 (10), 7058–7064.

May, C. T., 1999. Nonlinear pricing: theory & applications. J. Wiley, New York.

Mayfield, E. S., Mizrach, B., 1992. On determining the dimension of real-time stock-price data. Journal of Business and Economic Statistics 10 (3), 367–74.

McKenzie, M. D., 2001. Chaotic behavior in national stock market indices: New evidence from the close returns test. Global Finance Journal 12 (1), 35–53.

Mcleod, A. I., Li, W. K., 1983. Diagnostic checking ARMA time series models using squared-residual autocorrelations. J. Time Series Anal. 4, 269–273.

McLeod, A. I., Li, W. K., 1994. Diagnostic checking arma time series models using squared-residual autocorrelations. In: Harvey, A. (Ed.), Time series. Vol. 2. Elgar, Elgar Reference Collection. International Library of Critical Writings in Econometrics, vol. 5, pp. 301–05.

Medio, A., Lines, M., 2001. Non-Linear Dynamics. Cambridge University Press, New York.

Meyer, Y., 1993. Wavelets: Algorithms and Applications. Society for Industrial & Applied Mathematics, Philadelphia.

Middleton, E., 1996. Adaptation level and 'animal spirits'. Journal of Economic Psychology 17 (4), 479–498.

Miller, M. H., Muthuswamy, J., Whaley, R. E., 1994. Mean Reversion of Standard & Poor's 500 Index Basis Changes: Arbitrage-Induced or Statistical Illusion? Journal of Finance 49 (2), 479–513.

Mirowski, P., 1989. 'Tis a Pity Econometrics Isn't an Empirical Endeavor: Mandelbrot, Chaos, and the Noah and Joseph Effects. Ricerche Economiche 43 (1-2), 76–99.

Mirowski, P., 1990. From Mandelbrot to Chaos in Economic Theory. Southern Economic Journal 57 (2), 289–307.

Moore, J., 1999. Bootstrapping, permutation testing and the method of surrogate data. Physics in Medicine and Biology 44, L11–L12.

Morlet, J., Arens, G., Fourgeau, E., Giard, D., 1982. Wave propagation and sampling theory – part 1: Complex signal and scattering in multilayered media. Geophysics 47 (2), 203–221.

Muller, U., Dacorogna, M., Dave, R., Pictet, G., Olsen, R., Ward, R., 1993. Fractals and intrinsic time – a challenge to econometricians. The 39th International Conference of the Applied Econometrics Association on Real Time Econometrics http://citeseer.nj.nec.com/muller93fractals.html.

Muller, U. A., Dacorogna, M. M., Pictet, O. V., 1998. Heavy tails in high-frequency financial data. In: Adler, R. J., Feldman, R. E., Taqqu, M. S. (Eds.), A practical guide to heavy tails: Statistical techniques and applications. Birkhauser, Boston; Basel and Berlin, pp. 55–77.

Nason, G. P., 1995. Choice of the threshold parameter in wavelet function estimation. In: Antoniadis, A., Oppenheim, G. (Eds.), Wavelets and Statistics. Vol. 103 of Lecture Notes in Statistics. Springer-Verlag, New York, pp. 261–280.

Nelson, C. R., Kim, M. J., 1990. Predictable stock returns: Reality or statistical illusion? Working Paper 3297.

Newland, D. E., 1994. An Introduction to Random Vibrations, Spectral and Wavelet Analysis. Halsted Press Publisher Record, New York.

Nusse, H. E., Yorke, J. A., 1994. Dynamics: numerical explorations. Springer-Verlag, New York, applied mathematical sciences, v. 101.

Ogden, R. T., 1996. Essential Wavelets for Statistical Applications and Data Analysis. Birkhauser Boston, Cambridge.

Olmeda, I., Fernndez, E., 2000. Filtering with wavelets may be worst than you think, computing in Economics and Finance, Society for Computational Economics.

Opong, K. K., et al., 1999. The behaviour of some UK equity indices: An application of Hurst and BDS tests. Journal of Empirical Finance 6 (3), 267–82.

Ott, E., Sauer, T., Yorke, J. A. (Eds.), 1994. Coping with chaos; Analysis of chaotic data and the exploitation of chaotic systems. Wiley & Sons Inc., New York.

Packard, N. H., Crutchfield, J. P., Farmer, J. D., Shaw, R. S., 1980. Geometry from a time series. Phys. Rev. Lett. 45, 712–716.

Panas, E. E., 1990. The behaviour of Athens stock prices. Applied Economics 22 (12), 1715–27.

Peitgen, H.-O., Jrgens, H., Saupe, D., 1992. Chaos and fractals: new frontiers of science. Springer-Verlag, New York.

Percival, D. B., Walden, A. T., 2000. Wavelet Methods for Time Series Analysis. Cambridge University Press, New York, series in Statistical and Probabilistic Mathematics, No. 4.

Pesaran, M. H., Potter, S. M. e., 1993. Nonlinear dynamics, chaos and econometrics. Journal of Applied Econometrics, Volume 7, Supplement 1992. New York; (Wiley).

Pesin, Y. B., 1977. Lyapunov characteristic exponents and smooth ergodic theory. Usp. Mat. Nauk. 32, 55.

Peters, E. E., 1991. Chaos and order in the capital markets: a new view of cycles, prices, and market volatility. Wiley, New York.

Peters, E. E., 1994. Fractal market analysis: applying chaos theory to investment and economics. New York (J. Wiley & Sons).

Peters, E. E., 1999a. Complexity, risk and financial markets. Chichester Wiley, New York.

Peters, E. E., 1999b. Patterns in the dark: understanding risk and financial crisis with complexity theory. New York (Wiley).

Pis'mak, Y. M., 2001. Self-organization in a model of economic system with scale invariant interactions. Physica A: Statistical Mechanics and its Applications 299 (1-2), 311–318.

Poterba, J. M., Summers, L. H., 1986. The persistence of volatility and stock market fluctuations. American Economic Review 76 (5), 1142–51.

Poterba, J. M., Summers, L. H., 1988. Mean reversion in stock prices. Journal of Financial Economics 22 (1), 27–59.

Praetz, P., Wilson, E. J. G., 1978. The distribution of stock market returns: 1958-1973. Australian Journal of Management 3 (1), 79–90.

Praetz, P. D., 1972. The distribution of share price changes. Journal of Business 45 (1), 49–55.

Praetz, P. D., 1973. A spectral analysis of Australian share prices. Australian Economic Papers 12 (20), 70–78.

Praetz, P. D., 1975. Testing the Efficient Markets Theory on the Sydney wool Futures Exchange. Australian Economic Papers 14 (25), 240–49.

Praetz, P. D., 1977. A comparison of the stable and student distributions as statistical models for stock prices: Comment. Journal of Business 50 (1), 76–77.

Praetz, P. D., 1979. Testing for a flat spectrum on efficient market price data. Journal of Finance 34 (3), 645–58.

Prichard, D., Theiler, J., 1994. Generating surrogate data for time series with several simultaneously measured variables. Phys. Rev. Lett. 73, 951–954.

Prichard, D., Theiler, J., 1995. Generalized redundancies for time series analysis. Physica D 84, 476–493.

Prigogine, I., Stengers, I., 1984. Order Out of Chaos: Man's New Dialogue with Nature. Bantam Books, New York.

Ramsey, J. B., 1988. Economic and financial data as nonlinear processes. New York University Economic Research Reports 88-30.

Ramsey, J. B., 1996. If nonlinear models cannot forecast, what use are they? Studies in Nonlinear Dynamics and Econometrics 1 (2), 65–86.

Ramsey, J. B., Lampart, C., 1998. The decomposition of economic relationships by time scale using wavelets: Expenditure and income. Studies in Nonlinear Dynamics and Econometrics 3 (1), 23–42.

Ramsey, J. B., Rothman, P., 1996. Time irreversibility and business cycle asymmetry. Journal of Money, Credit, and Banking 28 (1), 1–21.

Ramsey, J. B., Sayers, C. L., Rothman, P., 1990. The statistical properties of dimension calculations using small data sets: Some economic applications. International Economic Review 31 (4), 991–1020.

Ramsey, J. B., Zhang, Z., 1994. The application of wave form dictionaries to stock market index data. New York University Economic Research Reports 94-05.

Ramsey, J. B., Zhang, Z., 1995. The analysis of foreign exchange data using waveform dictionaries. New York University Economic Research Reports 95-03.

Rao, T. S., Gabr, M., 1984. An Introduction to Bispectral Analysis and Bilinear Time Series Models. Vol. 24 of Lecture Notes in Statistics. Springer-Verlag, Berlin.

Reinganum, M. R., 1981. Misspecification of capital asset pricing: Empirical anomalies based on earnings' yields and market values. Journal of Financial Economics 9 (1), 19–46.

Ricker, N., 1940. The form and nature of seismic waves and the structure of seismograms. Geophysics 5, 348–366.

Riedi, R., 1995. An improved multifractal formalism and self-similar measures. Journal of Mathematical Analysis and Applications 189, no 2, 462–490.

Roll, R., 1983a. On computing mean returns and the small firm premium. Journal of Financial Economics 12 (3), 371–86.

Roll, R., 1983b. Vas ist das? Journal of Portfolio Management 9 (2), 18–28.

Rose, P. S., 1993. Readings on financial institutions and markets, 5th Edition. Irwin, Homewood, IL.

Rothman, P., 1990. Characterization of the time irreversibility of economic time series. Ph.d. thesis, New York University.

Rothman, P., 1994. The time reversibility test with application to financial data. In: Semmler, W. (Ed.), Business cycles: Theory and empirical methods. Kluwer Academic, Recent Economic Thought Series. Dordrecht and Boston, pp. 389–403.

Rothman, P., 1996. International evidence of business-cycle nonlinearity. In: Barnett, W. A., Kirman, A. P., Salmon, M. (Eds.), Nonlinear dynamics and economics: Proceedings of the Tenth International Symposium in Economic Theory and Econometrics. Cambridge University Press, International Symposia in Economic Theory and Econometrics series. Cambridge, pp. 333–41.

Rothman, P., 1997. FORTRAN programs for running the tr test: A guide and examples. Studies in Nonlinear Dynamics and Econometrics 1 (4), 203–08.

Roux, F., Gilbertson, B., 1978. Some further comments on the Johannesburg Stock Exchange as an efficient market. The Investment Aanalysts Journal 11 (4).

Rubinstein, A., 1998. Modeling bounded rationality. MIT Press, Cambridge, Mass.

Ruelle, D., 1989. Chaotic Evolution and Strange Attractors. Cambridge University Press, New York, lezioni Lincee Lectures.

Ruelle, D., 1993. Chance and chaos. Princeton, NJ (Princeton University Press).

Ruelle, D., 1995. Turbulence, strange attractors, and chaos. Singapore ; River Edge, NJ (World Scientific).

Samorodnitsky, G., Taqqu, M. S., 1994. Stable non-Gaussian random processes: stochastic models with infinite variance. Chapman & Hall, New York.

Samuelson, P., 1965. Proof that properly anticipated prices fluctuate randomly. Industrial Management Review , 1–10.

Sauer, T., 1992. A noise reduction method for signals from nonlinear systems. Physica D 58, 193–201.

Sauer, T., 1993. Time series prediction by using delay coordinate embedding. In: Weigend, A. S., Gershenfeld, N. A. (Eds.), Time series prediction: Forecasting the future and understanding the past. Addison Wesley, Reading, MA, pp. 175–193.

Sauer, T., Yorke, J. A., 1992. Almost every map preserves dimension. Preprint.

Sauer, T., Yorke, J. A., 1993. How many delay coordinates do you need? International Journal of Bifurcation and Chaos 3 (3), 737–744.

Sauer, T., Yorke, J. A., Casdagli, M., 1991. Embedology. J. Stat. Phys. 65, 579–616.

Sauer, T. D., Temkin, J. A., Yorke, J. A., 1998. Spurious Lyapunov exponents in attractor reconstruction. Phys. Rev. Lett. 81 (20), 4341–4344.

Savit, R., 1988. When random is not random: An introduction to chaos in market prices. Journal of Futures Markets 8 (3), 271–90.

Savit, R., 1989. Nonlinearities and chaotic effects in options prices. Journal of Futures Markets 9 (6), 507–18.

Scheinkman, J. A., LeBaron, B., 1989a. Nonlinear dynamics and GNP data. In: Barnett, W. A., Geweke, J., Shell, K. (Eds.), Economic complexity: Chaos, sunspots, bubbles, and nonlinearity. Proceedings of the Fourth International Symposium in Economic Theory and Econometrics. Cambridge University Press, International Symposia in Economic Theory and Econometrics series, pp. 213–240.

Scheinkman, J. A., LeBaron, B., 1989b. Nonlinear dynamics and stock returns. Journal of Business 62 (3), 311–37.

Scheinkman, J. A., Woodford, M., 1994. Self-organized criticality and economic fluctuations. American Economic Review 84 (2), 417–21.

Schleicher, C., 2002. An introduction to wavelets for economists.

Schmitz, A., Schreiber, T., 1998. Testing for nonlinearity in unevenly sampled time series. http://xxx.lanl.gov/abs/chao-dyn/9804042.

Schmitz, A., Schreiber, T., 1999. Testing for nonlinearity in unevenly sampled time series. Phys. Rev. E 59 (4), 4044–4047.

Schneider, P., Grassberger, P., 1997. Studying attractor symmetries by means of cross-correlation sums. Nonlinearity 10, 749–762.

Schreiber, T., 1997. Detecting and analysing nonstationarity in a time series using nonlinear cross predictions. Phys. Rev. Lett. 78 (5), 843–847.

Schreiber, T., Schmitz, A., 1996. Improved surrogate data for nonlinearity tests. Phys. Rev. Lett. 77 (4), 635–638.

Schreiber, T., Schmitz, A., 1997a. Classification of time series data with nonlinear similarity measures. Phys. Rev. Lett. 79 (8), 1475–1478.

Schreiber, T., Schmitz, A., 1997b. Discrimination power of measures for nonlinearity in a time series. Phys. Rev. E 55 (5), 5443–5447.

Schreiber, T., Schmitz, A., 2000. Surrogate time series. Physica D 142 (3-4), 346–382.

Schuster, H. G., 1984. Deterministic chaos : an introduction. Physik-Verlag.

Schwartz, G., 1978. Estimating the dimension of a model. Ann. Statist. 6 (461-464).

Shaffer, S., 1991. Structural shifts and the volatility of chaotic markets. Journal of Economic Behavior and Organization 15 (2), 201–14.

Shannon, C. E., July and October 1948. A mathematical theory of communication. Bell System Technical Journal 27, 379–423 and 623–656.

Shannon, C. E., Weaver, W., 1949. The Mathematical Theory of Communication. University of Illinois Press, Urbana.

Shaw, R. S., 1984. Modelling chaotic systems. Springer Ser. in Synerg. 11, 218.

Shefrin, H., Statman, M., 1985. The disposition to sell winners too early and ride losers too long: Theory and evidence. Journal of Finance 40 (3), 777–90.

Shiller, R. J., 1989. Market volatility. Cambridge, Mass. and London (MIT Press), xii, 464.

Shiller, R. J., 2000. Irrational Exhuberance, 1st Edition. Princeton University Press, Princeton, New Jersey.

Shiller, R. J., Perron, P., 1984. Testing the random walk hypothesis: Power versus frequency of observation. Discussion Paper 732.

Silverman, B. W., Vassilicos, J. C., 2000. Wavelets: The Key to Intermittent Information. Oxford University Press, New York.

Small, M., Judd, K., 1998a. Correlation dimension: A pivotal statistic for non-constrained realizations of composite hypotheses in surrogate data analysis. Physica D 120 (3-4), 386–400.

Small, M., Judd, K., 1998b. Detecting nonlinearity in experimental data. International Journal of Bifurcation and Chaos 8 (6), 1231–1244.

Small, M., Judd, K., 1999. Detecting periodicity in experimental data using linear modeling techniques. Phys. Rev. E 59 (2), 1379–1385.

Small, M., Judd, K., Mees, A., 2001a. Testing time series for nonlinearity. Statistics and Computing 11, 257–268.

Small, M., Judd, K., Mees, A., 2002a. Modeling continuous processes from data. Phys. Rev. E 6504 (4), 6704–6704.

Small, M., Tse, C. K., 2002. Applying the method of surrogate data to cyclic time series. Physica D 164 (3-4), 187–201.

Small, M., Yu, D., Harrison, R., 2001b. A surrogate test for pseudo-periodic time series data. Physical Review Letters 87, 188–201.

Small, M., Yu, D. J., Simonotto, J., Harrison, R. G., Grubb, N., Fox, K. A. A., 2002b. Uncovering non-linear structure in human ECG recordings. Chaos Solitons Fractals 13 (8), 1755–1762.

Smith, L., October 2001. The Maintenance of Uncertainty. Mathematical Institute, University of Oxford, Oxford, OX1 3LB, U.K.

Smith, V. L., Suchanek, G. L., Williams, A. W., 1991. Bubbles, crashes, and endogenous expectations in experimental spot asset markets. In: Smith, V. L. (Ed.),

Papers in experimental economics. Cambridge University Press, Cambridge, pp. 339–71.

Smith, V. L., van Boening, M., Wellford, C. P., 2000. Dividend timing and behavior in laboratory asset markets. Economic Theory 16 (3), 567–83.

So, P., Ott, E., Schiff, S. J., Kaplan, D. T., Sauer, T., Grebogi, C., 1996. Detecting unstable periodic orbits in chaotic experimental data. Phys. Rev. Lett. 76 (25), 4705–4708.

Solnik, B. H., 1973. Note on the validity of the random walk for European stock prices. Journal of Finance 28 (5), 1151–59.

Soofi, A. S., Cao, L., 1920. Modelling and Forecasting Financial Data: Techniques of Nonlinear Dynamics. Kluwer Academic Publishers, Norwell.

Sornette, D., 2003. Critical market crashes. Physics Reports 378 (1), 1–98.

Strang, G., 1989. Wavelets and dilation equations: a brief introduction. SIAM Review 31, 614–627.

Strang, G., Nguyen, T., 1996. Wavelets and Filter Banks. Wellesley-Cambridge Press, Wellesley.

Summers, L. H., 1986a. Do we really know that financial markets are efficient? In: Edwards, J., et al. (Eds.), Recent developments in corporate finance. Cambridge University Press, Cambridge, pp. 13–24.

Summers, L. H., 1986b. Does the stock market rationally reflect fundamental values? Journal of Finance 41 (3), 591–601.

Takens, F., 1980. Detecting strange attractors in turbulence. In: Rand, D. A., Young, L. S. (Eds.), Dynamical Systems and Turbulence (Warwick 1980). Vol. 898 of Lecture Notes in Mathematics. Springer-Verlag, Berlin, pp. 366–381.

Takens, F., 1993. Detecting nonlinearities in stationary time series. International Journal of Bifurcation and Chaos 3 (2), 241–256.

Taleb, N., 2001. Fooled by randomness: the hidden role of chance in the markets and in life. Texere, New York.

Teolis, A., 1998. Computational Signal Processing with Wavelets. Birkhauser Boston, Cambridge.

Terasvirta, T., Tjostheim, D., Granger, C. W. J., 1994. Aspects of modelling nonlinear time series. In: Engle, R. F., McFadden, D. L. (Eds.), Handbook of econometrics. Vol. 4. Elsevier North-Holland, Handbooks in Economics, vol. 2, pp. Analysis, Statistical Information Theory, and Other Special.

Theiler, J., 1986. Spurious dimensions from correlation algorithms applied to limited time-series data. Phys. Rev. A 34, 2427–2432.

Theiler, J., 1987. Efficient algorithm for estimating the correlation dimension from a set of discrete points. Phys. Rev. A 36, 4456–4473.

Theiler, J., 1988a. Lacunarity in a best estimator of fractal dimension. Phys. Lett. A 133, 195–200.

Theiler, J., 1988b. Quantifying chaos: Practical estimation of the correlation dimension. Ph.D. thesis, Caltech.

Theiler, J., 1992. Adventures in bootstrapping: How (and why) to generate surrogate data. CNLS/T12, MS-B213, Los Alamos National Laboratory, Los Alamos, N.M.

Theiler, J., 1994. Two tools to test time series data for evidence of chaos and/or nonlinearity. Integr. Physiol. Behav. Sci. 29 (3), 211–216.

Theiler, J., 1995. On the evidence for low-dimensional chaos in an epileptic electro-encephalogram. Phys. Lett. A 196, 335–341.

Theiler, J., Eubank, S., 1993. Don't bleach chaotic data. Chaos 3, 771–782.

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., Farmer, J. D., 1992a. Testing for nonlinearity in time series: the method of surrogate data. Physica D 58, 77–94.

Theiler, J., Galdrikian, B., Longtin, A., Eubank, S., Farmer, J. D., 1992b. Detecting nonlinear structure in time series. In: Vohra, S., Spano, M., Schlesinger, M., Pecora, L., Ditto, W. (Eds.), Proc. of the 1st Experimental Chaos Conference. World Scientific, Singapore, pp. 47–53.

Theiler, J., Galdrikian, B., Longtin, A., Eubank, S., Farmer, J. D., 1992c. Using surrogate data to detect nonlinearity in time series. In: Casdagli, M., Eubank, S. (Eds.), Nonlinear modeling and forecasting. Proceedings of the workshop held September, 1990, in Santa Fe, New Mexico. Addison Wesley, pp. 163–188.

Theiler, J., Linsay, P. S., Rubin, D. M., 1993. Detecting nonlinearity in data with long coherence times. In: Weigend, A. S., Gershenfeld, N. A. (Eds.), Time series prediction: Forecasting the future and understanding the past. Addison Wesley, pp. 429–455.

Theiler, J., Prichard, D., 1996. Constrained-realization monte-Carlo method for hypothesis testing. Physica D 94, 221.

Theiler, J., Prichard, D., 1997. Using "surrogate surrogate data" to calibrate the actual rate of false positives in tests for nonlinearity in time series. In: Cutler, C. D., Kaplan, D. T. (Eds.), Nonlinear Dynamics and Time Series. Vol. 11 of Fields Inst. Communications. American Math. Soc., pp. 99–113.

Theiler., J., Rapp, P. E., 1996. Re-examination of the evidence for low-dimensional, nonlinear structure in the human electroencephalogram. Electroencephalogr. Clin. Neurophysiol. 98, 213.

Tong, H., 1978. On a threshold model. In: Noordhoff, S. (Ed.), Pattern Recognition and Signal Processing. Amsterdam.

Tong, H., 1983. Threshold models in non-linear time series analysis. Vol. 21. Springer-Verlag, New York.

Tong, H., 1990. Non-Linear Time Series: A Dynamical System Approach. Oxford University Press, New York.

Tong, H., 1993. Dimension estimation and models. World Scientific, Singapore.

Trulla, L. L., Giuliani, A., Zbilut, J. P., Webber, C. L., 1996. Recurrence quantification analysis of the logistic equation with transients. Physics letters: A 223 (4), 255–261.

Urbach, R., 2000. Footprints of Chaos in the Markets: Analyzing Non-Linear Time Series in Financial Markets and Other Real Systems. Financial Times - Prentice Hall Philadelphia, London.

Vidakovic, B., 1999. Statistical Modeling by Wavelets. Wiley-Interscience, Hoboken, wiley Series in Probability and Statistics; Vol. 345.

Webber, C. L., Zbilut, J. P., 1994a. Dynamical assessment of physiological systems and states using recurrence plot strategies. Journal of Applied Physiology 76 (2), 965–73.

Webber, C. L., Zbilut, J. P., 1994b. Dynamical assessment of physiological systems and states using recurrence plot strategies. modeling in physiology. Journal of applied physiology 76 (2), 965.

Wegner, T., Peterson, M., Tyler, B., Branderhorst, P., 1992. Fractals for Windows. SAMS Publisher, Indianapolis.

Weierstrass, K., 1895. Mathematische werke. Vol. 2. Mayer and Muller, Berlin.

Weiss, G., 1975. Time-reversibility of linear stochasitic process. Journal of Applied Probability 12, 831–836.

Whitney, H., 1936. Differential manifolds. Annals of Mathematics 37 (3), 645–680.

Willey, T., 1992. Testing for nonlinear dependence in daily stock indices. Journal of Economics and Business 44 (1), 63–76.

Williams, G. P., 1997. Chaos theory tamed. Joseph Henry Press, Washington, D.C.

Wolf, A., Swift, J. B., Swinney, H. L., Vastano, J. A., 1985. Determining Lyapunov exponents from a time series. Physica D 16, 285–317.

Wong, K. A., Kwong, K. S., 1984. The behaviour of Hong Kong stock prices. Applied Economics 16 (6), 905–17.

Working, H., 1934. A random-difference series for use in the analysis of time series. Journal of the American Statistical Association 29 (185), 11–24.

Working, H., 1960a. Note on the correlation of first differences of averages in a random chain. Econometrica 28 (4), 916–918.

Working, H., 1960b. Price effects of futures trading. Calif. Stanford University Food Research Institute, Stanford.

Working, H., 1962. New concepts concerning futures markets and prices. The American Economic Review 52 (3), 431–459.

Working, H., 1963. Futures markets under renewed attack. Working paper, Food Research Institute.

Working, H., 1974a. Measurement of cycles in speculative prices. Working paper, Stanford University.

Working, H., 1974b. Measurement of cycles in speculative prices. Working paper, Stanford University.

Working, H., Peck, A. E., 1977. Selected writings of Holbrook Working, Professor Emeritus, Food Research Institute, Stanford University. Chicago Board of Trade, Chicago.

Wornell, G., 1993. Wavelet-based representation for the 1/f family of fractal processes. Proc. IEEE (Sept. 1993).

Wornell, G., 1995. Signal Processing with Fractals: A Wavelet Based approach. Prentice Hall.

Yadav, P., Paudyal, K., Pope, P., 1996. Non-linear dependence in daily stock returns: Evidence from Pacific Basin markets. Advances in Pacific Basin Financial Markets 2 (B), 349–377.

Yadav, P., Paudyal, K., Pope, P., 1999. Non-linear dependence in individual stock returns: Does trading frequency matter. Journal of Business Finance and Accounting 26, 651–679.

Yang, S.-R., Brorsen, B. W., 1993. Nonlinear dynamics of daily futures prices: Conditional heteroskedasticity or chaos? Journal of Futures Markets 13 (2), 175–91.

Young, R. K., 1992. Wavelet Theory and Its Applications. Kluwer Academic Publishers, Norwell.

Zarowin, P., 1989a. Does the stock market overreact to corporate earnings information? Journal of Finance 44 (5), 1385–99.

Zarowin, P., 1989b. Short-run market overreaction: Size and seasonality effects. Journal of Portfolio Management 15 (3), 26–29.

Zbilut, J. P., Giuliani, A., Webber, C. L., 1998a. Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification. Phys. Lett. A 246 (1-2), 122–128.

Zbilut, J. P., Giuliani, A., Webber, C. L., 1998b. Recurrence quantification analysis and principal components in the detection of short complex signals. Physics letters: A 237 (3), 131–136.

Zbilut, J. P., Giuliani, A., Webber, C. L., 2000. Recurrence quantification analysis as an empirical test to distinguish relatively short deterministic versus random number series. Physics letters: A 267 (2), 174–179.

Zbilut, J. P., Thomasson, N., Hoeppner, T. J., Webber, C. L., 2001. Recurrence quantification in epileptic EEGs. Physics letters: A 279 (1), 94–102.

Zbilut, J. P., Webber, C. L., 1992. Embeddings and delays as derived from quantification of recurrence plots. Physics letters: A 171 (3-4), 199.

Zbilut, J. P., Zaldivar-Comenges, J. M., Strozzi, F., 2002. Recurrence quantification based Liapunov exponents for monitoring divergence in experimental data. Phys. Lett. A 297 (3-4), 173–181.