2010-01-01

The Fractal Market Hypothesis: Applications to Financial Forecasting

Jonathan Blackledge
Technological University Dublin, jonathan.blackledge@tudublin.ie

The Fractal Market Hypothesis:

Applications to Financial Forecasting

Lecture Notes 4

Prof. Dr. Jonathan Blackledge
Stokes Professor
School of Electrical Engineering Systems
College of Engineering and Built Environment
Dublin Institute of Technology

Distinguished Professor
Centre for Advanced Studies
Institute of Mathematics
Polish Academy of Sciences
Warsaw University of Technology

http://eleceng.dit.ie/blackledge
jonathan.blackledge@dit.ie

ISBN: 978-83-61993-01-83
Pinted in Poland

Preface

Most financial modelling systems rely on an underlying hypothesis known as the Efficient Market Hypothesis (EMH) including the famous Black-Scholes formula for placing an option. However, the EMH has a fundamental flaw; it is based on the assumption that economic processes are normally distributed and it has long been known that this is not the case. This fundamental assumption leads to a number of shortcomings associated with using the EMH to analyse financial data which includes failure to predict the future volatility of a market share value.

This book introduces a new financial risk assessment model based on Lévy statistics and considers a financial forecasting system that uses a solution to a non-stationary fractional diffusion equation characterized by the Lévy index. Variations in the Lévy index are considered in order to assess the future volatility of financial data together with the likelihood of the markets becoming bear or bull dominant, thereby providing a solution to securing an investment portfolio. The key hypothesis associated with this approach is that a change in the Lévy index precedes a change in the financial signal from which the index is computed and can therefore be used as a risk management index. It is shown that there is a quantitative relationship between Lévy's characteristic function and a random scaling fractal signal obtained through a Green's function solution to the fractional diffusion equation. In this sense, the model considered is based on the Fractal Market Hypothesis and two case studies are considered in this respect: (i) application of the hypothesis by predicting the volatility associated with foreign exchange markets; (ii) application to the ABX index.

Acknowledgments

The author is supported by the Science Foundation Ireland Stokes Professorship Programme. Case Study I was undertaken at the request of GAM https://www.gam.com. The ABX data used in Case Study II was provided by the Systemic Risk Analysis Division, Bank of England.

Contents

1 Introduction

In 1900, Louis Bachelier concluded that the price of a commodity today is the best estimate of its price in the future. The random behaviour of commodity prices was again noted by Working in 1934 in an analysis of time series data. In the 1950s, Kendall attempted to find periodic cycles in indices of security and commodity prices but did not find any. Prices appeared to be yesterday's price plus some random change and he suggested that price changes were independent and that prices apparently followed random walks. The majority of financial research seemed to agree; asset price changes appeared to be random and independent and so prices were taken to follow random walks. Thus, the first model of price variation was based on the sum of independent random variations often referred to as Brownian motion.

Some time later, it was noticed that the size of price movements depend on the size of the price itself. The Brownian motion model was therefore revised to include this proportionality effect and a new model developed which stated that the log price changes should be Gaussian distributed. This model is the basis for the equation

$$\frac{1}{S}\frac{d}{dt}S(t) = \sigma dX + \mu dt$$

where S is the price at time t, μ is a drift term which reflects the average rate of growth of an asset, σ is the volatility and dX is a sample from a normal distribution. In other words, the relative price change of an asset is equal to some random element plus some underlying trend component. This model is an example of a log-normal random walk and has the following important properties: (i) Statistical stationarity of price increments in which samples of data taken over equal time increments can be superimposed onto each other in a statistical sense; (ii) scaling of price where samples of data corresponding to different time increments can be suitably re-scaled such that they too, can be superimposed onto each other in a statistical sense; (iii) independence of price changes.

It is often stated that asset prices should follow Gaussian random walks because of the Efficient Market Hypothesis (EMH) [1], [2] and [3]. The EMH states that the current price of an asset fully reflects all available information relevant to it and that new information is immediately incorporated into the price. Thus, in an efficient market, models for asset pricing are concerned with the arrival of new information which is taken to be independent and random.

The EMH implies independent price increments but why should they be Gaussian distributed? A Gaussian Probability Density Function (PDF) is chosen because price movements are presumed to be an aggregation of smaller ones and sums of independent random contributions have a Gaussian PDF because due to the Central Limit Theorem [4], [5]. This is equivalent to arguing that all financial time series used to construct an 'averaged signal' such as the Dow Jones Industrial Average

are statistically independent. However, this argument is not fully justified because it assumes that the reaction of investors to one particular stock market is independent of investors in other stock markets which, in general, will not be the case as each investor may have a common reaction to economic issues that transcend any particular stock. In other words asset management throughout the markets relies on a high degree of connectivity and the arrival of new information sends 'shocks' through the market as people react to it and then to each other's reactions. The EMH assumes that there is a rational and unique way to use available information, that all agents possess this knowledge and that any chain reaction produced by a 'shock' happens instantaneously. This is clearly not physically possible.

In this work, we present an approach to analysing financial signals that is based on a non-stationary fractional diffusion equation derived under the assumption that the data are Lévy distributed. We consider methods of solving this equation and provide an algorithm for computing the non-stationary Lévy index using a standard moving window principle. We also consider case studies in which the method is use to assess the ABX index and predict market volatility focusing on foreign currency exchange [18], [19].

In common with other applications of signal analysis, in order to understand the nature of a financial signal, it is necessary to be clear about what assumptions are being made in order to develop a suitable model. It is therefore necessary to introduce some of the issues associated with financial engineering as given in the following section [6], [7] and [8].

2 The Black-Scholes Model

For many years, investment advisers focused on returns with the occasional caveat 'subject to risk'. Modern Portfolio Theory (MPT) is concerned with a trade-off between risk and return. Nearly all MPT assumes the existence of a risk-free investment, e.g. the return from depositing money in a sound financial institute or investing in equities. In order to gain more profit, the investor must accept greater risk. Why should this be so? Suppose the opportunity exists to make a guaranteed return greater than that from a conventional bank deposit say; then, no (rational) investor would invest any money with the bank. Furthermore, if he/she could also borrow money at less than the return on the alternative investment, then the investor would borrow as much money as possible to invest in the higher yielding opportunity. In response to the pressure of supply and demand, the banks would raise their interest rates. This would attract money for investment with the bank and reduce the profit made by investors who have money borrowed from the bank. (Of course, if such opportunities did arise, the banks would probably be the first to invest our savings in them.) There is elasticity in the argument because of various 'friction factors' such as transaction costs, differences in borrowing and lending rates, liquidity laws etc., but on the whole, the principle is sound because the market is saturated with arbitrageurs whose purpose is to seek out and exploit irregularities or miss-pricing.

The concept of successful arbitraging is of great importance in finance. Often loosely stated as, 'there's no such thing as a free lunch', it means that one cannot ever make an instantaneously risk-free profit. More precisely, such opportunities cannot exist for a significant length of time before prices move to eliminate them.

2.1 Financial Derivatives

As markets have grown and evolved, new trading contracts have emerged which use various tricks to manipulate risk. Derivatives are deals, the value of which is derived from (although not the same as) some underlying asset or interest rate. There are many kinds of derivatives traded on the markets today. These special deals increase the number of moves that players of the economy have available to ensure that the better players have more chance of winning. To illustrate some of the implications of the introduction of derivatives to the financial markets we consider the most simple and common derivative, namely, the option.

2.1.1 Options

An option is the right (but not the obligation) to buy (call) or sell (put) a financial instrument (such as a stock or currency, known as the 'underlying') at an agreed date in the future and at an agreed price, called the strike price. For example, consider an investor who 'speculates' that the value of an asset at price S will rise.

The investor could buy shares at S, and if appropriate, sell them later at a higher price. Alternatively, the investor might buy a call option, the right to buy a share at a later date. If the asset is worth more than the strike price on expiry, the holder will be content to exercise the option, immediately sell the stock at the higher price and generate an automatic profit from the difference. The catch is that if the price is less, the holder must accept the loss of the premium paid for the option (which must be paid for at the opening of the contract). If C denotes the value of a call option and E is the strike price, the option is worth $C(S,t) = \max(S - E, 0)$.

Conversely, suppose the investor speculates that an asset is going to fall, then the investor can sell shares or buy puts. If the investor speculates by selling shares that he/she does not own (which in certain circumstances is perfectly legal in many markets), then he/she is selling 'short' and will profit from a fall in the asset. (The opposite of a short position is a 'long' position.) The principal question is how much should one pay for an option? Clearly, if the value of the asset rises, then so does the value of a call option and vice versa for put options. But how do we quantify exactly how much this gamble is worth? In previous times (prior to the Black-Scholes model which is discussed later) options were bought and sold for the value that individual traders thought they ought to have. The strike prices of these options were usually the 'forward price', which is just the current price adjusted for interest-rate effects. The value of options rises in active or volatile markets because options are more likely to pay out large amounts of money when they expire if market moves have been large, i.e. potential gains are higher, but loss is always limited to the cost of the premium. This gain through successful 'speculation' is not the only role that options play. Another role is Hedging.

2.1.2 Hedging

Suppose an investor already owns shares as a long-term investment, then he/she may wish to insure against a temporary fall in the share price by buying puts as well. Clearly, the investor would not want to liquidate holdings only to buy them back again later, possibly at a higher price if the estimate of the share price is wrong, and certainly having incurred some transaction costs on the deals. If a temporary fall occurs, the investor has the right to sell his/her holdings for a higher than market price. The investor can then immediately buy them back for less, in this way generating a profit and long-term investment then resumes. If the investor is wrong and a temporary fall does not occur, then the premium is lost for the option but at least the stock is retained, which has continued to rise in value. Since the value of a put option rises when the underlying asset value falls, what happens to a portfolio containing both assets and puts? The answer depends on the ratio. There must exist a ratio at which a small unpredictable movement in the asset does not result in any unpredictable movement in the portfolio. This ratio is instantaneously risk free. The reduction of risk by taking advantage of correlations between the

option price and the underlying price is called 'hedging'. If a market maker can sell an option and hedge away all the risk for the rest of the options life, then a risk free profit is guaranteed.

Why write options? Options are usually sold by banks to companies to protect themselves against adverse movements in the underlying price, in the same way as holders do. In fact, writers of options are no different to holders; they expect to make a profit by taking a view of the market. The writers of calls are effectively taking a short position in the underlying behaviour of the markets. Known as 'bears', these agents believe the price will fall and are therefore also potential customers for puts. The agents taking the opposite view are called 'bulls'. There is a near balance of bears and bulls because if everyone expected the value of a particular asset to do the same thing, then its market price would stabilise (if a reasonable price were agreed on) or diverge (if everyone thought it would rise). Clearly, the psychology and dynamics (which must go hand in hand) of the bear/bull cycle play an important role in financial analysis.

The risk associated with individual securities can be hedged through diversification or 'spread betting' and/or various other ways of taking advantage of correlations between different derivatives of the same underlying asset. However, not all risk can be removed by diversification. To some extent, the fortunes of all companies move with the economy. Changes in the money supply, interest rates, exchange rates, taxation, commodity prices, government spending and overseas economies tend to affect all companies in one way or another. This remaining risk is generally referred to as market risk.

2.2 Black-Scholes Analysis

The value of an option can be thought of as a function of the underlying asset price S (a Gaussian random variable) and time t denoted by $V(S,t)$. Here, V can denote a call or a put; indeed, V can be the value of a whole portfolio or different options although for simplicity we can think of it as a simple call or put. Any derivative security whose value depends only on the current value S at time t and which is paid for up front, is taken to satisfy the Black-Scholes equation given by[9]

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0$$

where σ is the volatility and r is the risk. As with other partial differential equations, an equation of this form may have many solutions. The value of an option should be unique; otherwise, again, arbitrage possibilities would arise. Therefore, to identify the appropriate solution, certain initial, final and boundary conditions need to be imposed. Take for example, a call; here the final condition comes from the arbitrage argument. At $t = T$

$$C(S,t) = \max(S - E, 0)$$

The spatial or asset-price boundary conditions, applied at $S = 0$ and $S \to \infty$ come from the following reasoning: If S is ever zero then dS is zero and will therefore never change. Thus, we have

$$C(0, t) = 0$$

As the asset price increases it becomes more and more likely that the option will be exercised, thus we have

$$C(S, t) \propto S, \quad S \to \infty$$

Observe, that the Black-Sholes equation has a similarity to the diffusion equation but with additional terms. An appropriate way to solve this equation is to transform it into the diffusion equation for which the solution is well known and with appropriate transformations gives the Black-Scholes formula [9]

$$C(S, t) = SN(d_1) - Ee^{r(T-t)}N(d_2)$$

where

$$d_1 = \frac{\log(S/E) + (r + \frac{1}{2}\sigma^2)(T - t)}{\sigma\sqrt{T - t}},$$

$$d_2 = \frac{\log(S/E) + (r - \frac{1}{2}\sigma^2)(T - t)}{\sigma\sqrt{T - t}}$$

and N is the cumulative normal distribution defined by

$$N(d_1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_1} e^{\frac{1}{2}s^2} ds.$$

The conceptual leap of the Black-Scholes model is to say that traders are not estimating the future price, but are guessing about how volatile the market may be in the future. The model therefore allows banks to define a fair value of an option, because it assumes that the forward price is the mean of the distribution of future market prices. However, this requires a good estimate of the future volatility σ.

The relatively simple and robust way of valuing options using Black-Scholes analysis has rapidly gained in popularity and has universal applications. Black-Scholes analysis for pricing an option is now so closely linked into the markets that the price of an option is usually quoted in option volatilities or 'vols'. However, Black-Scholes analysis is ultimately based on random walk models that assume independent and Gaussian distributed price changes and is thus, based on the EMH.

The theory of modern portfolio management is only valuable if we can be sure that it truly reflects reality for which tests are required. One of the principal issues with regard to this relates to the issue of assuming that the markets are Gaussian distributed. However, it has long been known that financial time series do not adhere to Gaussian statistics. This is the most important of the shortcomings relating

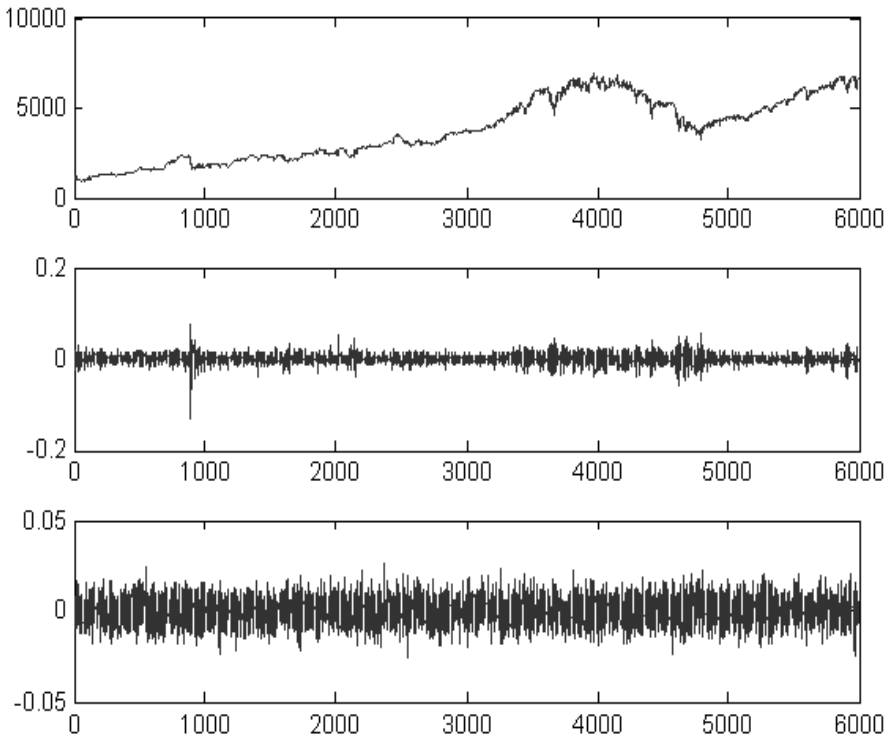

Figure 1: Financial time series for the FTSE value (close-of-day) from 02-04-1984 to 12-12-2007 (top), the log derivative of the same time series (centre) and a zero-mean Gaussian distributed random signal (bottom).

to the EMH model (i.e. the failure of the independence and Gaussian distribution of increments assumption) and is fundamental to the inability for EMH-based analysis such as the Black-Scholes equation to explain characteristics of a financial signal such as clustering, flights and failure to explain events such as 'crashes leading to recession. The limitations associted with the EMH are illustrated in Figure 1 which shows a (discrete) financial signal $u(t)$, the log derivative of this signal $d \log u(t)/dt$ and a synthesised (zero-mean) Gaussian distributed random signal. The log derivative is considered in order to: (i) eliminate the characteristic long term exponential growth of the signal; (ii) obtain a signal on the daily price differences[1]. Clearly, there is a marked difference in the characteristics of a real financial signal and a random Gaussian signal. This simple comparison indicates a failure of the statistical independence assumption which underpins the EMH.

These limitations have prompted a new class of methods for investigating time

[1]The gradient is computed using forward differencing.

series obtained from a range of disciplines. For example, Re-scaled Range Analysis (RSRA), e.g. [10], [11], which is essentially based on computing and analysing the Hurst exponent [12], is a useful tool for revealing some well disguised properties of stochastic time series such as persistence (and anti-persistence) characterized by non-periodic cycles. Non-periodic cycles correspond to trends that persist for irregular periods but with a degree of statistical regularity often associated with non-linear dynamical systems. RSRA is particularly valuable because of its robustness in the presence of noise. The principal assumption associated with RSRA is concerned with the self-affine or fractal nature of the statistical character of a time-series rather than the statistical 'signature' itself. Ralph Elliott first reported on the fractal properties of financial data in 1938. He was the first to observe that segments of financial time series data of different sizes could be scaled in such a way that they were statistically the same producing so called Elliot waves. Since then, many different self-affine models for price variation have been developed, often based on (dynamical) Iterated Function Systems (IFS). These models can capture many properties of a financial time series but are not based on any underlying causal theory of the type attempted in this work.

A good stochastic financial model should ideally consider all the observable behaviour of the financial system it is attempting to model. It should therefore be able to provide some predictions on the immediate future behaviour of the system within an appropriate confidence level. Predicting the markets has become (for obvious reasons) one of the most important problems in financial engineering. Although, at least in principle, it might be possible to model the behaviour of each individual agent operating in a financial market, one can never be sure of obtaining all the necessary information required on the agents themselves and their modus operandi. This principle plays an increasingly important role as the scale of the financial system, for which a model is required, increases. Thus, while quasi-deterministic models can be of value in the understanding of micro-economic systems (with known 'operational conditions'), in an ever increasing global economy (in which the operational conditions associated with the fiscal policies of a given nation state are increasingly open), we can take advantage of the scale of the system to describe its behaviour in terms of functions of random variables.

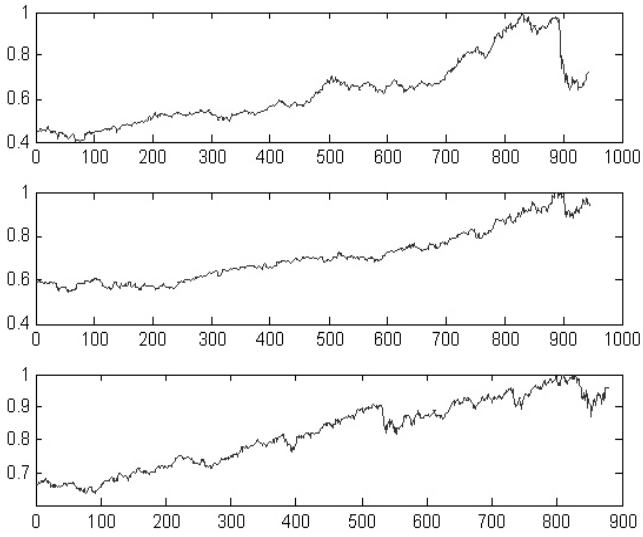

Figure 2: Evolution of the 1987, 1997 and 2007 financial crashes. Normalised plots (i.e. where the data has been rescaled to values between 0 and 1 inclusively) of the daily FTSE value (close-of-day) for 02-04-1984 to 24-12-1987 (top), 05-04-1994 to 24-12-1997 (centre) and 02-04-2004 to 24-09-2007 (bottom)

3 Market Analysis

The stochastic nature of financial time series is well known from the values of the stock market major indices such as the FTSE (Financial Times Stock Exchange) in the UK, the Dow Jones in the US which are frequently quoted. A principal aim of investors is to attempt to obtain information that can provide some confidence in the immediate future of the stock markets often based on patterns of the past, patterns that are ultimately based on the interplay between greed and fear. One of the principle components of this aim is based on the observation that there are 'waves within waves' that appear to permeate financial signals when studied with sufficient detail and imagination. It is these repeating patterns that occupy both the financial investor and the systems modeller alike and it is clear that although economies have undergone many changes in the last one hundred years, the dynamics of market data do not appear to change significantly (ignoring scale). For example, Figure 2 shows the build up to three different 'crashes', the one of 1987 and that of 1997 (both after approximately 900 days) and what may turn out to be a crash of 2007. The similarity in behaviour of these signals is remarkable and is indicative of the quest to understand economic signals in terms of some universal phenomenon

from which appropriate (macro) economic models can be generated. In an efficient market, only the revelation of some dramatic information can cause a crash, yet post-mortem analysis of crashes typically fail to (convincingly) tell us what this information must have been.

In modern economies, the distribution of stock returns and anomalies like market crashes emerge as a result of considerable complex interaction. In the analysis of financial time series, it is inevitable that assumptions need to be made to make the derivation of a model possible. This is the most vulnerable stage of the process. Over simplifying assumptions lead to unrealistic models. There are two main approaches to financial modelling: The first approach is to look at the statistics of market data and to derive a model based on an educated guess of the 'mechanics' of the market. The model can then be tested using real data. The idea is that this process of trial and error helps to develop the right theory of market dynamics. The alternative is to 'reduce' the problem and try to formulate a microscopic model such that the desired behaviour 'emerges', again, by guessing agents' strategic rules. This offers a natural framework for interpretation; the problem is that this knowledge may not help to make statements about the future unless some methods for describing the behaviour can be derived from it. Although individual elements of a system cannot be modelled with any certainty, global behaviour can sometimes be modelled in a statistical sense provided the system is complex enough in terms of its network of interconnection and interacting components.

In complex systems, the elements adapt to the aggregate pattern they co-create. As the components react, the aggregate changes, as the aggregate changes the components react anew. Barring the reaching of some asymptotic state or equilibrium, complex systems keep evolving, producing seemingly stochastic or chaotic behaviour. Such systems arise naturally in the economy. Economic agents, be they banks, firms, or investors, continually adjust their market strategies to the macroscopic economy which their collective market strategies create. It is important to appreciate that there is an added layer of complexity within the economic community: Unlike many physical systems, economic elements (human agents) react with strategy and foresight by considering the implications of their actions (some of the time!). Although we can not be certain whether this fact changes the resulting behaviour, we can be sure that it introduces feedback which is the very essence of both complex systems and chaotic dynamical systems that produce fractal structures.

Complex systems can be split into two categories: equilibrium and non-equilibrium. Equilibrium complex systems, undergoing a phase transition, can lead to 'critical states' that often exhibit random fractal structures in which the statistics of the 'field' are scale invariant. For example, when ferromagnets are heated, as the temperature rises, the spins of the electrons which contribute to the magnetic field gain energy and begin to change in direction. At some critical temperature, the spins form a random vector field with a zero mean and a phase transition occurs in which the magnetic field averages to zero. But the field is not just random, it is a self-affine

random field whose statistical distribution is the same at different scales, irrespective of the characteristics of the distribution. Non-equilibrium complex systems or 'driven' systems give rise to 'self organised critical states', an example is the growing of sand piles. If sand is continuously supplied from above, the sand starts to pile up. Eventually, little avalanches will occur as the sand pile inevitably spreads outwards under the force of gravity. The temporal and spatial statistics of these avalanches are scale invariant.

Financial markets can be considered to be non-equilibrium systems because they are constantly driven by transactions that occur as the result of new fundamental information about firms and businesses. They are complex systems because the market also responds to itself, often in a highly non-linear fashion, and would carry on doing so (at least for some time) in the absence of new information. The 'price change field' is highly non-linear and very sensitive to exogenous shocks and it is probable that all shocks have a long term effect. Market transactions generally occur globally at the rate of hundreds of thousands per second. It is the frequency and nature of these transactions that dictate stock market indices, just as it is the frequency and nature of the sand particles that dictates the statistics of the avalanches in a sand pile. These are all examples of random scaling fractals [13].

3.1 Autocorrelation Analysis and Market Memory

When faced with a complex process of unknown origin, it is usual to select an independent process such as Brownian motion as a working hypothesis where the statistics and probabilities can be estimated with great accuracy. However, using traditional statistics to model the markets assumes that they are games of chance. For this reason, investment in securities is often equated with gambling. In most games of chance, many degrees of freedom are employed to ensure that outcomes are random. In the case of a simple dice, a coin or roulette wheel, for example, no matter how hard you may try, it is physically impossible to master your roll or throw such that you can control outcomes. There are too many non-repeatable elements (speeds, angles and so on) and non-linearly compounding errors involved. Although these systems have a limited number of degrees of freedom, each outcome is independent of the previous one. However, there are some games of chance that involve memory. In Blackjack, for example, two cards are dealt to each player and the object is to get as close as possible to 21 by twisting (taking another card) or sticking. In a 'bust' (over 21), the player loses; the winner is the player that stays closest to 21. Here, memory is introduced because the cards are not replaced once they are taken. By keeping track of the cards used, one can assess the shifting probabilities as play progresses. This game illustrates that not all gambling is governed by Gaussian statistics. There are processes that have long-term memory, even though they are probabilistic in the short term. This leads directly to the question, does the economy have memory? A system has memory if what happens today will affect what happens in the future.

Memory can be tested by observing correlations in the data. If the system today has no affect on the system at any future time, then the data produced by the system will be independently distributed and there will be no correlations. A function that characterises the expected correlations between different time periods of a financial signal $u(t)$ is the Auto-Correlation Function (ACF) defined by

$$A(t) = u(t) \odot u(t) = \int_{-\infty}^{\infty} u(\tau)u(\tau - t)d\tau.$$

where $\odot$ denotes that the correlation operation. This function can be computed either directly (evaluation of the above integral) or via application of the power spectrum using the correlation theorem

$$u(t) \odot u(t) \Longleftrightarrow \mid U(\omega) \mid^2$$

where $\Longleftrightarrow$ denotes transformation from real space t to Fourier space ω (the angular frequency), i.e.

$$U(\omega) = \mathcal{F}[u(t)] = \int_{-\infty}^{\infty} u(t) \exp(-i\omega t)dt$$

where $\mathcal{F}$ denotes the Fourier transform operator. The power spectrum $\mid U(\omega) \mid^2$ characterises the amplitude distribution of the correlation function from which we can estimate the time span of memory effects. This also offers a convenient way to calculate the correlation function (by taking the inverse Fourier transform of $\mid U(\omega) \mid^2$). If the power spectrum has more power at low frequencies, then there are long time correlations and therefore long-term memory effects. Inversely, if there is greater power at the high frequency end of the spectrum, then there are short-term time correlations and evidence of short-term memory. White noise, which characterises a time series with no correlations over any scale, has a uniformly distributed power spectrum.

Since prices movements themselves are a non-stationary process, there is no ACF as such. However, if we calculate the ACF of the price increments du/dt, then we can observe how much of what happens today is correlated with what happens in the future. According to the EMH, the economy has no memory and there will therefore be no correlations, except for today with itself. We should therefore expect the power spectrum to be effectively constant and the ACF to be a delta function. The power spectra and the ACFs of log price changes $d \log u/dt$ and their absolute value $\mid d \log u/dt \mid$ for the FTSE 100 index (daily close) from 02-04-1984 to 24-09-2007 is given in Figure 3. The power spectra of the data is not constant with rogue spikes (or groups of spikes) at the intermediate and high frequency portions of the spectrum. For the absolute log price increments, there is evidence of a power law at the low frequency end, indicating that there is additional correlation in the signs of the data.

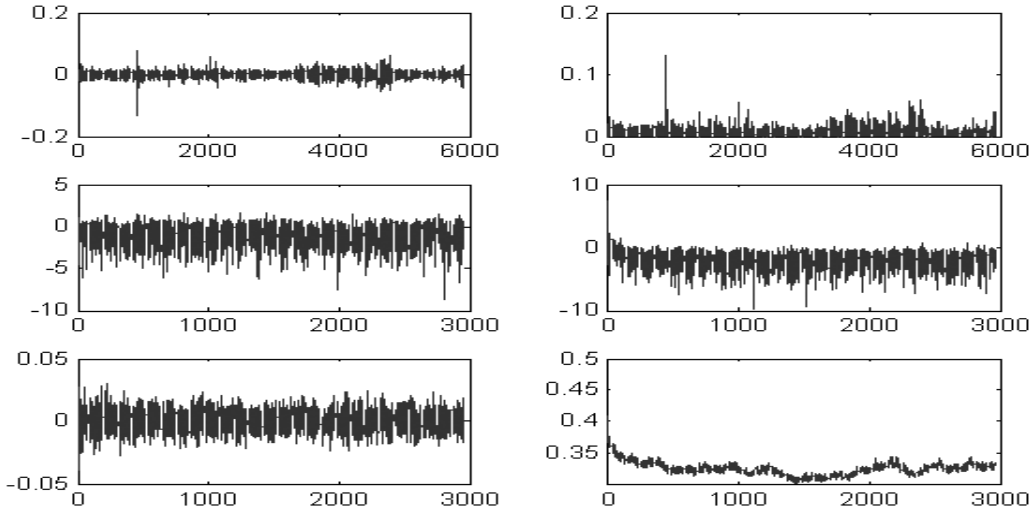

Figure 3: Log-power spectra and ACFs of log price changes and absolute log price changes for FTSE 100 index (daily close) from 02-04-1984 to 24-09-2007. Top-left: log price changes; top-right: absolute value of log price changes; middle: log power spectra; bottom: ACFs.

The ACF of the log price changes is relatively featureless, indicating that the excess of low frequency power within the signal has a fairly subtle effect on the correlation function. However, the ACF of the absolute log price changes contains a number of interesting features. It shows that there are a large number of short range correlations followed by an irregular decline up to approximately 1500 days after which the correlations start to develop again, peaking at about 2225 days. The system governing the magnitudes of the log price movements clearly has a better long-term memory than it should. The data used in this analysis contains 5932 daily price movements and it is therefore improbable that these results are coincidental and correlations of this, or any similar type, what ever the time scale, effectively invalidates the independence assumption of the EMH.

3.2 Stochastic Modelling and Analysis

Developing mathematical models to simulate stochastic processes has an important role in financial analysis and information systems in general where it should be noted that information systems are now one of the most important aspects in terms of regulating financial systems, e.g. [14], [15]. A good stochastic model is one that accurately predicts the statistics we observe in reality, and one that is based upon some well defined rationale. Thus, the model should not only describe the data, but

also help to explain and understand the system.

There are two principal criteria used to define the characteristics of a stochastic field: (i) The PDF or the Characteristic Function (i.e. the Fourier transform of the PDF); the Power Spectral Density Function (PSDF). The PSDF is the function that describes the envelope or shape of the power spectrum of a signal. In this sense, the PSDF is a measure of the field correlations. The PDF and the PSDF are two of the most fundamental properties of any stochastic field and various terms are used to convey these properties. For example, the term 'zero-mean white Gaussian noise' refers to a stochastic field characterized by a PSDF that is effectively constant over all frequencies (hence the term 'white' as in 'white light') and has a PDF with a Gaussian profile whose mean is zero.

Stochastic fields can of course be characterized using transforms other than the Fourier transform (from which the PSDF is obtained) but the conventional PDF-PSDF approach serves many purposes in stochastic systems theory. However, in general, there is no general connectivity between the PSDF and the PDF either in terms of theoretical prediction and/or experimental determination. It is not generally possible to compute the PSDF of a stochastic field from knowledge of the PDF or the PDF from the PSDF. Hence, in general, the PDF and PSDF are fundamental but non-related properties of a stochastic field. However, for some specific statistical processes, relationships between the PDF and PSDF can be found, for example, between Gaussian and non-Gaussian fractal processes [16] and for differentiable Gaussian processes [17].

There are two conventional approaches to simulating a stochastic field. The first of these is based on predicting the PDF (or the Characteristic Function) theoretically (if possible). A pseudo random number generator is then designed whose output provides a discrete stochastic field that is characteristic of the predicted PDF. The second approach is based on considering the PSDF of a field which, like the PDF, is ideally derived theoretically. The stochastic field is then typically simulated by filtering white noise. A 'good' stochastic model is one that accurately predicts both the PDF and the PSDF of the data. It should take into account the fact that, in general, stochastic processes are non-stationary. In addition, it should, if appropriate, model rare but extreme events in which significant deviations from the norm occur.

New market phenomenon result from either a strong theoretical reasoning or from compelling experimental evidence or both. In econometrics, the processes that create time series such as the FTSE have many component parts and the interaction of those components is so complex that a deterministic description is simply not possible. As in all complex systems theory, we are usually required to restrict the problem to modelling the statistics of the data rather than the data itself, i.e. to develop stochastic models. When creating models of complex systems, there is a trade-off between simplifying and deriving the statistics we want to compare with reality and simulating the behaviour through an emergent statistical behaviour. Stochastic sim-

ulation allows us to investigate the effect of various traders' behavioural rules on the global statistics of the market, an approach that provides for a natural interpretation and an understanding of how the amalgamation of certain concepts leads to these statistics.

One cause of correlations in market price changes (and volatility) is mimetic behaviour, known as herding. In general, market crashes happen when large numbers of agents place sell orders simultaneously creating an imbalance to the extent that market makers are unable to absorb the other side without lowering prices substantially. Most of these agents do not communicate with each other, nor do they take orders from a leader. In fact, most of the time they are in disagreement, and submit roughly the same amount of buy and sell orders. This is a healthy non-crash situation; it is a diffusive (random-walk) process which underlies the EMH and financial portfolio rationalization.

One explanation for crashes involves a replacement for the EMH by the Fractal Market Hypothesis (FMH) which is the basis of the model considered in this work. The FMH proposes the following: (i) The market is stable when it consists of investors covering a large number of investment horizons which ensures that there is ample liquidity for traders; (ii) information is more related to market sentiment and technical factors in the short term than in the long term - as investment horizons increase and longer term fundamental information dominates; (iii) if an event occurs that puts the validity of fundamental information in question, long-term investors either withdraw completely or invest on shorter terms (i.e. when the overall investment horizon of the market shrinks to a uniform level, the market becomes unstable); (iv) prices reflect a combination of short-term technical and long-term fundamental valuation and thus, short-term price movements are likely to be more volatile than long-term trades - they are more likely to be the result of crowd behaviour; (v) if a security has no tie to the economic cycle, then there will be no long-term trend and short-term technical information will dominate. Unlike the EMH, the FMH states that information is valued according to the investment horizon of the investor. Because the different investment horizons value information differently, the diffusion of information will also be uneven. Unlike most complex physical systems, the agents of the economy, and perhaps to some extent the economy itself, have an extra ingredient, an extra degree of complexity. This ingredient is consciousness.

3.3 Fractal Time Series and Rescaled Range Analysis

A time series is fractal if the data exhibits statistical self-affinity and has no characteristic scale. The data has no characteristic scale if it has a PDF with an infinite second moment. The data may have an infinite first moment as well; in this case, the data would have no stable mean either. One way to test the financial data for the existence of these moments is to plot them sequentially over increasing time periods to see if they converge. Figure 4 shows that the first moment, the mean, is stable,

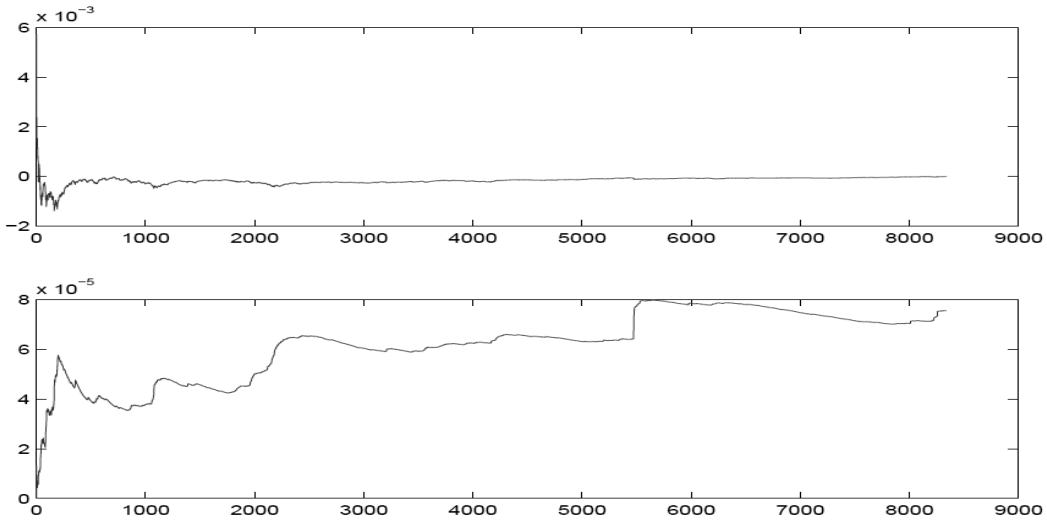

Figure 4: The first and second moments (top and bottom) of the Dow Jones Industrial Average plotted sequentially.

but that the second moment, the mean square, is not settled. It converges and then suddenly jumps and it is observed that although the variance is not stable, the jumps occur with some statistical regularity. Time series of this type are example of Hurst processes; time series that scale according to the power law,

$$\left\langle \frac{d}{dt} \log u(t) \right\rangle_t \propto t^H$$

where H is the Hurst exponent.

H. E. Hurst (1900-1978) was an English civil engineer who built dams and worked on the Nile river dam project. He studied the Nile so extensively that some Egyptians reportedly nicknamed him 'the father of the Nile.' The Nile river posed an interesting problem for Hurst as a hydrologist. When designing a dam, hydrologists need to estimate the necessary storage capacity of the resulting reservoir. An influx of water occurs through various natural sources (rainfall, river overflows etc.) and a regulated amount needed to be released for primarily agricultural purposes. The storage capacity of a reservoir is based on the net water flow. Hydrologists usually begin by assuming that the water influx is random, a perfectly reasonable assumption when dealing with a complex ecosystem. Hurst, however, had studied the 847-year record that the Egyptians had kept of the Nile river overflows, from 622 to 1469. Hurst noticed that large overflows tended to be followed by large overflows until abruptly, the system would then change to low overflows, which also tended to be followed by low overflows. There seemed to be cycles, but with no predictable

period. Standard statistical analysis revealed no significant correlations between observations, so Hurst developed his own methodology. Hurst was aware of Einstein's (1905) work on Brownian motion (the erratic path followed by a particle suspended in a fluid) who observed that the distance the particle covers increased with the square root of time, i.e.

$$R = \sqrt{t}$$

where R is the range covered, and t is time. This is the same scaling property as discussed earlier in the context of volatility. It results, again, from the fact that increments are identically and independently distributed random variables. Hurst's idea was to use this property to test the Nile River's overflows for randomness. In short, his method was as follows: Begin with a time series x_i (with $i = 1, 2, ..., n$) which in Hurst's case was annual discharges of the Nile River. (For markets it might be the daily changes in the price of a stock index.) Next, create the adjusted series, $y_i = x_i - \bar{x}$ (where $\bar{x}$ is the mean of x_i). Cumulate this time series to give

$$Y_i = \sum_{j=1}^{i} y_j$$

such that the start and end of the series are both zero and there is some curve in between. (The final value, Y_n has to be zero because the mean is zero.) Then, define the range to be the maximum minus the minimum value of this time series,

$$R_n = \max(Y) - \min(Y).$$

This adjusted range, R_n is the distance the systems travels for the time index n, i.e. the distance covered by a random walker if the data set y_i were the set of steps. If we set $n = t$ we can apply Einstein's equation provided that the time series x_i is independent for increasing values of n. However, Einstein's equation only applies to series that are in Brownian motion. Hurst's contribution was to generalize this equation to

$$(R/S)_n = cn^H$$

where S is the standard deviation for the same n observations and c is a constant. We define a Hurst process to be a process with a (fairly) constant H value and the R/S is referred to as the 'rescaled range' because it has zero mean and is expressed in terms of local standard deviations. In general, the R/S value increases according to a power law value equal to H known as the Hurst exponent. This scaling law behaviour is the first connection between Hurst processes and fractal geometry.

Rescaling the adjusted range was a major innovation. Hurst originally performed this operation to enable him to compare diverse phenomenon. Rescaling, fortunately, also allows us to compare time periods many years apart in financial time series. As discussed previously, it is the relative price change and not the change itself

that is of interest. Due to inflationary growth, prices themselves are a significantly higher today than in the past, and although relative price changes may be similar, actual price changes and therefore volatility (standard deviation of returns) are significantly higher. Measuring in standard deviations (units of volatility) allows us to minimize this problem. Rescaled range analysis can also describe time series that have no characteristic scale, another characteristic of fractals. By considering the logarithmic version of Hurst's equation, i.e.

$$\log(R/S)_n = \log(c) + H\log(n)$$

it is clear that the Hurst exponent can be estimated by plotting $\log(R/S)$ against the $\log(n)$ and solving for the gradient with a least squares fit. If the system were independently distributed, then $H = 0.5$. Hurst found that the exponent for the Nile River was $H = 0.91$, i.e. the rescaled range increases at a faster rate than the square root of time. This meant that the system was covering more distance than a random process would, and therefore the annual discharges of the Nile had to be correlated.

It is important to appreciate that this method makes no prior assumptions about any underlying distributions, it simply tells us how the system is scaling with respect to time. So how do we interpret the Hurst exponent? We know that $H = 0.5$ is consistent with an independently distributed system. The range $0.5 < H \leq 1$, implies a persistent time series, and a persistent time series is characterized by positive correlations. Theoretically, what happens today will ultimately have a lasting effect on the future. The range $0 < H \leq 0.5$ indicates anti-persistence which means that the time series covers less ground than a random process. In other words, there are negative correlations. For a system to cover less distance, it must reverse itself more often than a random process.

3.4 The Joker Effect

After this discovery, Hurst analysed all the data he could including rainfall, sunspots, mud sediments, tree rings and others. In all cases, Hurst found H to be greater than 0.5. He was intrigued that H often took a value of about 0.7 and Hurst suspected that some universal phenomenon was taking place. He carried out some experiments using numbered cards. The values of the cards were chosen to simulate a PDF with finite moments, i.e. $0, \pm 1, \pm 3, \pm 5, \pm 7 \text{and} \pm 9$. He first verified that the time series generated by summing the shuffled cards gave $H = 0.5$. To simulate a bias random walk, he carried out the following steps.

1. Shuffle the deck and cut it once, noting the number, say n.
2. Replace the card and re-shuffle the deck.
3. Deal out 2 hands of 26 cards, A and B.

4. Replace the lowest n cards of deck B with the highest n cards of deck A, thus biasing deck B to the level n.

5. Place a joker in deck B and shuffle.

6. Use deck B as a time series generator until the joker is cut, then create a new biased hand.

Hurst did 1000 trials of 100 hands and calculated $H = 0.72$. We can think of the process as follows: we first bias each hand, which is determined by a random cut of the pack; then, we generate the time series itself, which is another series of random cuts; then, the joker appears, which again occurs at random. Despite all of these random events $H = 0.72$ would always appear. This is called the 'joker effect'. The joker effect, as illustrated above, describes a tendency for data of a certain magnitude to be followed by more data of approximately the same magnitude, but only for a fixed and random length of time. A natural example of this phenomenon is in weather systems. Good weather and bad weather tend to come in waves or cycles (as in a heat wave for example). This does not mean that weather is periodic, which it is clearly not. We use the term 'non-periodic cycle' to describe cycles of this kind (with no fixed period). Thus, if markets are Hurst processes, they exhibit trends that persist until the economic equivalent of the joker comes along to change that bias in magnitude and/or direction. In other words rescaled range analysis can, along with the PDF and PSDF, help to describe a stochastic time series that contains within it, many different short-lived trends or biases (both in size and direction). The process continues in this way giving a constant Hurst exponent, sometimes with flat episodes that correspond to the average periods of the non-periodic cycles, depending on the distribution of actual periods.

The following is a step by step methodology for applying R/S analysis to stock market data. Note that the AR(1) notation used below stands for auto regressive process with single daily linear dependence. Thus, taking AR(1) residuals of a signal is equivalent to plotting the signals one day out of phase and taking the day to day linear dependence out of the data.

1. Prepare the data P_t. Take AR(1) residuals of log ratios. The log ratios account for the fact that price changes are relative, i.e. depend on price. The AR(1) residuals remove any linear dependence, serial correlation, or short-term memory which can bias the analysis.

$$V_t = \log(P_t/P_{t-1})$$

$$X_t = V_t - (c + mV_{t-1})$$

The AR(1) residuals are taken to eliminate any linear dependency.

2. Divide this time series (of length N) up into A sub-periods, such that the first and last value of time series are included i.e. $A \times n = N$. Label each sub-period I_a with $a = 1, 2, 3, ..., A$. Label each element in I_a with $X_{k,a}$ where $k = 1, 2, 3, ..., n$. For each I of length n, calculate the mean

$$e_a = (1/n) \sum_{i=1}^{k} N_{k,a}$$

3. Calculate the time series of accumulated departures from the mean for each sub interval.

$$Y_{k,a} = \sum_{i=1}^{k} (N_{i,a} - e_a)$$

4. Define the range as

$$R_{I_a} = \max(Y_{k,a}) - \min(Y_{k,a})$$

where $1 \leq k \leq n$.

5. Define the sample standard deviation for each sub-period as

$$S_{I_a} = \sqrt{\frac{1}{n} \sum_{k=1}^{n} (N_{k,a} - e_a)^2}$$

6. Each range, R_{I_a} is now normalized by dividing by its corresponding S_{I_a}. Therefore the re-scaled range for each I_a is equal to R_{I_a}/S_{I_a}. From step 2 above, we have A contiguous sub-periods of length n. Therefore the average R/S value for each length n is defined as

$$(R/S)_n = \frac{1}{A} \sum_{a=1}^{A} (R_{I_a}/S_{I_a})$$

7. The length n is then increased until there are only two sub-periods, i.e. $n = N/2$. We then perform a least squares regression on $\log(n)$ as the independent variable and $\log(R/S)$ as the dependent variable. The slope of the equation is the estimate of the Hurst exponent, H.

The R/S analysis results for the NYA (1960-1998) for daily, 5-daily, 10-daily and 20-daily returns are shown in Figure 5. The Hurst exponent is $0.54 \leq H \leq 0.59$, from daily to 20-daily returns indicating that the data set is persistent, at least up to 1000 trading days. At this point the scaling behaviour appears to slow down. The $(R/S)_n$ values show a systematic deviation from the line of best fit which is plotted in the Figures. From the daily return results this appears to be at about 1000 days.

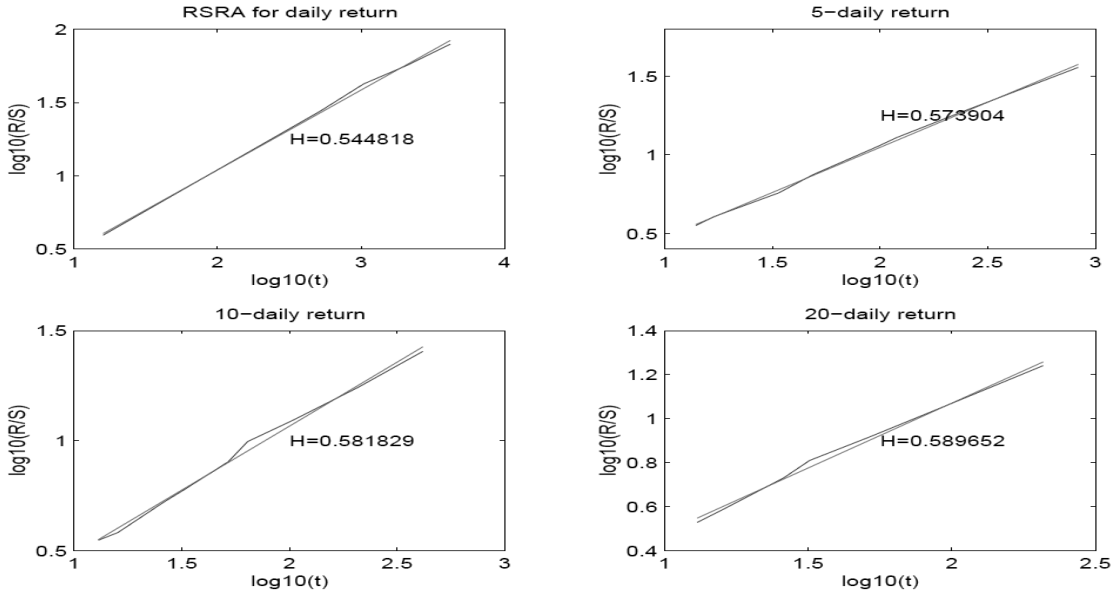

Figure 5: Rescaled Range Analysis results for the Dow Jones Industrial Average 1960-89

The 5-daily, 10-day and 20-day return results appear to agree a value of about 630 days. This is also where the correlation function starts to increase. This deviation is more pronounced the lower the frequency with which the data is sampled. The results show that there are certain non-periodic cycles that persist for up to 1000 days which contribute to the persistence in the data, and after these are used up, the data (the walker) slows down. These observations can be summarized as follows: The market reacts to information, and the way it reacts is not very different from the way it reacts previously, even though the information is different. Therefore the underlying dynamics and the statistics of the market have not changed. This is especially true of fractal statistics. (The 'fractal statistics' referred to are the fractal dimension and the Hurst exponent.) The results clearly imply that there is an inconsistency between the behaviour of real financial data and the EMH lognormal random walk model which is compounded in the following points:

1. The PSDF of log price changes is not constant. Therefore price changes are not independent.

2. The PDF of log price changes are not Gaussian, they have a sharper peak at the mean and longer tails.

In addition, the following properties are evident:

1. Asset price movements are self-affine, at least up to 1000 days.
2. The first moment of the PDF is finite but the second moment is infinite (or at least very slow to converge).
3. If stock market data is viewed as a random walk then the walk scales faster than the square root of the time up until approximately 1000 days and then slows down.
4. Large price movements tend to be followed by large movements and vice versa, although signs of increments are uncorrelated. Thus volatility comes in bursts. These cycles are referred to as non-periodic as they have randomly distributed periods.

Hurst devised a method for analysing data which does not require a Gaussian assumption about the underlying distribution and in particular, does not require the distribution to have finite moments. This method can reveal subtle scaling properties including non-periodic cycles in the data that spectral analysis alone cannot find.

4 The Classical and Fractional Diffusion Equations

For diffusivity $D = \sigma^{-1}$, the homogeneous diffusion equation is given by

$$\left(\nabla^2 - \sigma\frac{\partial}{\partial t}\right) u(\mathbf{r}, t) = 0$$

where $\mathbf{r} \equiv (x, y, z)$ is the three-dimensional space vector and ∇^2 is the Laplacian operator given by

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

The field $u(\mathbf{r}, t)$ represents a measurable quantity whose space-time dependence is determined by the random walk of a large ensemble of particles, a strongly scattered wavefield or information flowing through a complex network. We consider an initial value for this field denoted by $u_0 \equiv u(\mathbf{r}, 0) = u(\mathbf{r}, t)$ at $t = 0$. The diffusion equation can be derived using a random walk model for particles undergoing inelastic collisions. It is assumed that the movements of the particles are independent of the movements of all other particles and that the motion of a single particle at some interval of time is independent of its motion at all other times. This derivation (usually attributed to Einstein [20]) is given in the following section for the one-dimensional case. For completeness, the three-dimensional case is considered in Appendix A.

4.1 Derivation of the Classical Diffusion Equation

Let τ be a small interval of time in which a particle moves some distance between λ and $\lambda + d\lambda$ with a probability $p(\lambda)$ where τ is long enough to assume that the movements of the particle in two separate periods of τ are independent. If n is the total number of particles and we assume that $p(\lambda)$ is constant between λ and $\lambda + d\lambda$, then the number of particles that travel a distance between λ and $\lambda + d\lambda$ in τ is given by

$$dn = np(\lambda)d\lambda$$

If $u(x, t)$ is the concentration (number of particles per unit length) then the concentration at time $t + \tau$ is described by the integral of the concentration of particles which have been displaced by λ in time τ, as described by the equation above, over all possible λ, i.e.

$$u(x, t + \tau) = \int_{-\infty}^{\infty} u(x + \lambda, t)p(\lambda)d\lambda \tag{1}$$

Since, τ is assumed to be small, we can approximate $u(x, t + \tau)$ using the Taylor series and write

$$u(x, t + \tau) \simeq u(x, t) + \tau\frac{\partial}{\partial t}u(x, t)$$

Similarly, using a Taylor series expansion of $u(x+\lambda,t)$, we have

$$u(x+\lambda,t) \simeq u(x,t) + \lambda\frac{\partial}{\partial x}u(x,t) + \frac{\lambda^2}{2!}\frac{\partial^2}{\partial x^2}u(x,t)$$

where the higher order terms are neglected under the assumption that if τ is small, then the distance travelled, λ, must also be small. We can then write

$$u(x,t) + \tau\frac{\partial}{\partial t}u(x,t) = u(x,t)\int_{-\infty}^{\infty} p(\lambda)d\lambda$$

$$+\frac{\partial}{\partial x}u(x,t)\int_{-\infty}^{\infty}\lambda p(\lambda)d\lambda + \frac{1}{2}\frac{\partial^2}{\partial x^2}u(x,t)\int_{-\infty}^{\infty}\lambda^2 p(\lambda)d\lambda$$

For isotropic diffusion, $p(\lambda) = p(-\lambda)$ and so p is an even function with usual normalization condition

$$\int_{-\infty}^{\infty} p(\lambda)d\lambda = 1$$

As λ is an odd function, the product $\lambda p(\lambda)$ is also an odd function which, if integrated over all values of λ, equates to zero. Thus we can write

$$u(x,t) + \tau\frac{\partial}{\partial t}u(x,t) = u(x,t) + \frac{1}{2}\frac{\partial^2}{\partial x^2}u(x,t)\int_{-\infty}^{\infty}\lambda^2 p(\lambda)d\lambda$$

so that

$$\frac{\partial}{\partial t}u(x,t) = \frac{\partial^2}{\partial x^2}u(x,t)\int_{-\infty}^{\infty}\frac{\lambda^2}{2\tau}p(\lambda)d\lambda$$

Finally, defining the diffusivity as

$$D = \int_{-\infty}^{\infty}\frac{\lambda^2}{2\tau}p(\lambda)d\lambda$$

we obtain the diffusion equation

$$\left(\frac{\partial^2}{\partial x^2} - \sigma\frac{\partial}{\partial t}\right)u(x,t) = 0$$

where $\sigma = 1/D$. Note that this derivation does not depend explicitly on $p(\lambda)$. However, there is another approach to deriving this result that is informative with regard to the discussion given in the following section and is determined by $p(\lambda)$. Under

the condition that $p(\lambda)$ is a symmetric function, equation (1) - a correlation integral - is equivalent to a convolution integral. Thus, using the convolution theorem, in Fourier space, equation (1) becomes

$$U(k, t+\tau) = U(k,t)P(k)$$

where U and P are the Fourier transforms of u and p given by

$$U(k, t+\tau) = \int_{-\infty}^{\infty} u(x, t+\tau) \exp(-ikx) dx$$

and

$$P(k) = \int_{-\infty}^{\infty} p(x) \exp(-ikx) dx$$

respectively. Suppose we consider a Probability Density Function (PDF) $p(x)$ that is Gaussian distributed. Then the Characteristic Function $P(k)$ is also Gaussian given by say (ignoring scaling constants)

$$P(k) = \exp(-a \mid k \mid^2) = 1 - a \mid k \mid^2 + ...$$

Let

$$P(k) = 1 - a \mid k \mid^2, \quad a \to 0$$

We can then write

$$\frac{U(k, t+\tau) - U(k,t)}{\tau} = -\frac{a}{\tau} \mid k \mid^2 U(k,t)$$

so that as $\tau \to 0$ we obtain the equation

$$\sigma \frac{\partial}{\partial t} u(x,t) = \frac{\partial^2}{\partial x^2} u(x,t)$$

where $\sigma = \tau / a$ and we have used the result

$$\frac{\partial^2}{\partial x^2} u(x,t) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} k^2 U(k,t) \exp(ikx) dk$$

This approach to deriving the diffusion equation relies on specifying the characteristic function $P(k)$ and upon the conditions that both a and τ approach zero, thereby allowing $\sigma = \tau / a$ to be of arbitrary value. This is the basis for the approach considered in the following section with regard to a derivation of the anomalous or fractional diffusion equation.

4.2 Derivation of the Fractional Diffusion Equation for a Lévy Distributed Process

Lévy processes are random walks whose distribution has infinite moments. The statistics of (conventional) physical systems are usually concerned with stochastic fields that have PDFs where (at least) the first two moments (the mean and variance) are well defined and finite. Lévy statistics is concerned with statistical systems where all the moments (starting with the mean) are infinite. Many distributions exist where the mean and variance are finite but are not representative of the process, e.g. the tail of the distribution is significant, where rare but extreme events occur. These distributions include Lévy distributions [21],[22]. Lévy's original approach to deriving such distributions is based on the following question: Under what circumstances does the distribution associated with a random walk of a few steps look the same as the distribution after many steps (except for scaling)? This question is effectively the same as asking under what circumstances do we obtain a random walk that is statistically self-affine. The characteristic function $P(k)$ of such a distribution $p(x)$ was first shown by Lévy to be given by (for symmetric distributions only)

$$P(k) = \exp(-a \mid k \mid^{\gamma}), \quad 0 < \gamma \leq 2 \tag{2}$$

where a is a constant and γ is the Lévy index. For $\gamma \geq 2$, the second moment of the Lévy distribution exists and the sums of large numbers of independent trials are Gaussian distributed. For example, if the result were a random walk with a step length distribution governed by $p(x)$, $\gamma \geq 2$, then the result would be normal (Gaussian) diffusion, i.e. a Brownian random walk process. For $\gamma < 2$ the second moment of this PDF (the mean square), diverges and the characteristic scale of the walk is lost. For values of γ between 0 and 2, Lévy's characteristic function corresponds to a PDF of the form (see Appendix B)

$$p(x) \sim \frac{1}{x^{1+\gamma}}, \quad x \to \infty$$

Lévy processes are consistent with a fractional diffusion equation as we shall now show [23]. Consider the evolution equation for a random walk process to be given by equation (1) which, in Fourier space, is

$$U(k, t+\tau) = U(k,t)P(k)$$

From equation (2),

$$P(k) = 1 - a \mid k \mid^{\gamma}, \quad a \to 0$$

so that we can write

$$\frac{U(k,t+\tau) - U(k,t)}{\tau} \simeq -\frac{a}{\tau} \mid k \mid^{\gamma} U(k,t)$$

which for $\tau \to 0$ gives the fractional diffusion equation

$$\sigma \frac{\partial}{\partial t} u(x,t) = \frac{\partial^\gamma}{\partial x^\gamma} u(x,t), \quad \gamma \in (0,2] \tag{3}$$

where $\sigma = \tau/a$ and we have used the result

$$\frac{\partial^\gamma}{\partial x^\gamma} u(x,t) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \mid k \mid^\gamma U(k,t) \exp(ikx) dk \tag{4}$$

The solution to this equation with the singular initial condition $u(x,0) = \delta(x)$ is given by

$$u(x,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(ikx - t \mid k \mid^\gamma / \sigma) dk$$

which is itself Lévy distributed. This derivation of the fractional diffusion equation reveals its physical origin in terms of Lévy statistics.

4.3 Generalisation

The approach used to derived the fractional diffusion equation given in the previous section can be generalised further for arbitrary PDFs. Applying the correlation theorem to equation (1) we note that

$$U(k, t+\tau) = U(k,t) P^*(k)$$

where the characteristic function $P(k)$ may be asymmetric. Then

$$U(k, t+\tau) - U(k,t) = U(k,t)[P^*(k) - 1]$$

so that as $\tau \to 0$ we obtained a generalised anomalous diffusion equation given by

$$\frac{\partial}{\partial t} u(x,t) = \frac{1}{\tau} \left(\int_{-\infty}^{\infty} u(x+y,t) p(y) dy - u(x,t) \right)$$

4.4 Green's Function for the Fractional Diffusion Equation

Let

$$u(x,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} U(x,\omega) \exp(i\omega t) d\omega$$

where ω is the angular frequency so that equation (3) can be written as

$$\left(\frac{\partial^\gamma}{\partial x^\gamma} + \Omega^2 \right) U(x,\omega) = 0$$

where $\Omega^2 = -i\omega\sigma$ and we choose the positive root $\Omega = i(i\omega\sigma)^{\frac{1}{2}}$. For an ideal impulse located at x_0, the Green's function g is then defined in terms of the solution of [24]

$$\left(\frac{\partial^\gamma}{\partial x^\gamma} + \Omega^2\right) g(x \mid x_0, \omega) = -\delta(x - x_0)$$

Fourier transforming this equation with regard to x and using equation (4), we obtain an expression for g given by (with $X \equiv \mid x - x_0 \mid$)

$$g(X, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\exp(ikX)dk}{(k^{\frac{\gamma}{2}} - \Omega)(k^{\frac{\gamma}{2}} + \Omega)}$$

The integral has two roots at $k = \pm\Omega^{\frac{2}{\gamma}}$ and for $k = \Omega^{\frac{2}{\gamma}}$, the Green's function is given by

$$g(x \mid x_0, \omega) = \frac{i}{2\Omega_\gamma} \exp(i\Omega_\gamma \mid x - x_0 \mid)$$

where

$$\Omega_\gamma = i^{\frac{2}{\gamma}} (i\omega\sigma)^{1/\gamma}$$

Suppose we consider the fractional diffusion equation

$$\left(\frac{\partial^2}{\partial x^2} - \sigma^q \frac{\partial^q}{\partial t^q}\right) u(x, t)$$

where we call q the 'Fourier Dimension'. Using the result

$$\frac{\partial^q}{\partial t^q} u(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} U(x, \omega)(i\omega)^q \exp(i\omega t) d\omega$$

the Green's function is defined as the solution to

$$\left(\frac{\partial^2}{\partial x^2} + \Omega_q^2\right) g(x \mid x_0, \omega) = \delta(x - x_0)$$

where (ignoring the negative root)

$$\Omega_q = i(i\omega\sigma)^{\frac{q}{2}}$$

In this case, the Green's function is given by

$$g(x \mid x_0, \omega) = \frac{i}{2\Omega_q} \exp(i\Omega_q \mid x - x_0 \mid)$$

This analysis provides a relationship between the Lévy index and the Fourier Dimension given by

$$\frac{1}{\gamma} = \frac{q}{2}$$

Gaussian processes associated with the classical diffusion equation are thus recovered when $\gamma = 2$ and $q = 1$ and $\gamma \in (0, 2] \equiv q \in (\infty, 1]$

4.4.1 Green's Function for $q = 2$

When $q = 2$, we recover the Green's function for the wave equation For $\Omega_2 = -\omega\sigma$, we have

$$g(X, \omega) = \frac{1}{2i\omega\sigma} \exp(-i\omega\sigma X)$$

Fourier inverting, using the convolution theorem and noting that

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-i\omega\sigma X) \exp(i\omega t) d\omega = \delta(t - \sigma X)$$

and

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{2i\omega} \exp(i\omega t) d\omega = \frac{1}{4} \mathrm{sgn}(t)$$

where

$$\mathrm{sgn}(t) = \begin{cases} +1, & t < 0 \\ -1, & t > 0 \end{cases}$$

we obtain an expression for the time-dependent Green's function given by

$$G(x \mid x_0, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} g(x \mid x_0, \omega) \exp(i\omega t) d\omega$$

$$= \frac{1}{4\sigma} \mathrm{sgn}(t - \sigma \mid x - x_0 \mid)$$

which describes the propagation of a wave travelling at velocity $1/\sigma$ with a wavefront that occurs at $t = \sigma \mid x - x_0 \mid$.

4.4.2 Green's Function for $q = 1$

For $q = 1$, $\Omega_1 = i\sqrt{i\omega\sigma}$ and

$$G(X, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\exp(-\sqrt{i\omega\sigma} X)}{2\sqrt{i\omega\sigma}} \exp(i\omega t) d\omega$$

Using the result

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{\exp(-a\sqrt{p})}{2\sqrt{p}} \exp(pt) dp = \frac{1}{\sqrt{\pi t}} \exp[-a^2/(4t)]$$

where a is a constant, then, with $p = i\omega$ we obtain

$$G(x \mid x_0, t) = \frac{1}{\sqrt{\pi\sigma t}} \exp[-\sigma(x - x_0)^2/(4t)], \quad t > 0$$

which is the Green's function for the classical diffusion equation, i.e. a Gaussian function.

5 Green's Function Solution

We consider a Green's function solution to the equation

$$\left(\frac{\partial^2}{\partial x^2} - \sigma^q \frac{\partial^q}{\partial t^q}\right) u(x,t) = -F(x,t)$$

when $F(x,t) = f(x)n(t)$ where $f(x)$ and $n(t)$ are stochastic functions described by PDFs $\Pr[f(x)]$ and $\Pr[n(t)]$ respectively. Although a Green's function solution does not require the source function to be separable, utilising a separable function in this way allows a solution to be generated in which the terms affecting the temporal behaviour of $u(x,t)$ are clearly identifiable. Although we consider the fractional diffusion equation for values of $q \in (1,\infty)$, for generality, we consider a solution for $q \in (-\infty,\infty)$. Thus, we require a general solution to the equation

$$\left(\frac{\partial^2}{\partial x^2} - \sigma^q \frac{\partial^q}{\partial t^q}\right) u(x,t) = -f(x)n(t), \quad -\infty \leq q \leq \infty$$

Let

$$u(x,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} U(x,\omega) \exp(i\omega t) d\omega$$

and

$$n(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} N(\omega) \exp(i\omega t) d\omega$$

Then, using the result

$$\frac{\partial^q}{\partial t^q} u(x,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} U(x,\omega)(i\omega)^q \exp(i\omega t) d\omega$$

we can transform the fractional diffusion equation to the form

$$\left(\frac{\partial^2}{\partial x^2} + \Omega_q^2\right) U(x,\omega) = -f(x)N(\omega)$$

The Green's function solution is then given by

$$U(x_0,\omega) = N(\omega) \int_{-\infty}^{\infty} g(x \mid x_0,\omega) f(x) dx \tag{5}$$

under the assumption that u and $\partial u/\partial x \to 0$ as $x \to \pm\infty$.

5.1 General Series Solution

The evaluation of $u(x_0, t)$ via direct Fourier inversion for arbitrary values of q is not possible because of the irrational nature of the Green's function with respect to ω. To obtain a general solution, we use the series representation of the exponential function and write

$$U(x_0, \omega) = \frac{iM_0 N(\omega)}{2\Omega_q} \left[1 + \sum_{m=1}^{\infty} \frac{(i\Omega_q)^m}{m!} \frac{M_m(x_0)}{M_0} \right] \tag{6}$$

where

$$M_m(x_0) = \int_{-\infty}^{\infty} f(x) \mid x - x_0 \mid^m dx$$

We can now Fourier invert term by term to develop a series solution. Given that we consider $-\infty < q < \infty$, this requires us to consider three distinct cases.

5.1.1 Solution for $q = 0$

Evaluation of $u(x_0, t)$ in this case is trivial since, from equation (5)

$$U(x_0, \omega) = \frac{M(x_0)}{2} N(\omega) \quad \text{or} \quad u(x_0, t) = \frac{M(x_0)}{2} n(t)$$

where

$$M(x_0) = \int_{-\infty}^{\infty} \exp(- \mid x - x_0 \mid) f(x) dx$$

5.1.2 Solution for $q > 0$

Fourier inverting, the first term in equation (6) becomes

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{iN(\omega)M_0}{2\Omega_q} \exp(i\omega t) d\omega =$$

$$\frac{M_0}{2\sigma^{\frac{q}{2}}} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{N(\omega)}{(i\omega)^{\frac{q}{2}}} \exp(i\omega t) d\omega$$

$$= \frac{M_0}{2\sigma^{\frac{q}{2}}} \frac{1}{(2i)^q \sqrt{\pi}} \frac{\Gamma\left(\frac{1-q}{2}\right)}{\Gamma\left(\frac{q}{2}\right)} \int_{-\infty}^{\infty} \frac{n(\tau)}{(t-\tau)^{1-(q/2)}} d\tau$$

The second term is

$$-\frac{M_1}{2}\frac{1}{2\pi}\int_{-\infty}^{\infty} N(\omega)\exp(i\omega t)d\omega = -\frac{M_1}{2}n(t)$$

The third term is

$$-\frac{iM_2}{2.2!}\frac{1}{2\pi}\int_{-\infty}^{\infty} N(\omega)i(i\omega\sigma)^{\frac{q}{2}}\exp(i\omega t)d\omega = \frac{M_2\sigma^{\frac{q}{2}}}{2.2!}\frac{d^{\frac{q}{2}}}{dt^{\frac{q}{2}}}n(t)$$

and the fourth and fifth terms become

$$\frac{M_3}{2.3!}\frac{1}{2\pi}\int_{-\infty}^{\infty} N(\omega)i^2(i\omega\sigma)^{q}\exp(i\omega t)d\omega = -\frac{M_3\sigma^{q}}{2.3!}\frac{d^{q}}{dt^{q}}n(t)$$

and

$$i\frac{M_4}{2.4!}\frac{1}{2\pi}\int_{-\infty}^{\infty} N(\omega)i^3(i\omega\sigma)^{\frac{3q}{2}}\exp(i\omega t)d\omega = \frac{M_4\sigma^{\frac{3q}{2}}}{2.4!}\frac{d^{\frac{3q}{2}}}{dt^{\frac{3q}{2}}}n(t)$$

respectively with similar results for all other terms. Thus, through induction, we can write $u(x_0,t)$ as a series of the form

$$u(x_0,t) =$$

$$\frac{M_0(x_0)}{2\sigma^{q/2}}\frac{1}{(2i)^q\sqrt{\pi}}\frac{\Gamma\left(\frac{1-q}{2}\right)}{\Gamma\left(\frac{q}{2}\right)}\int_{-\infty}^{\infty}\frac{n(\tau)}{(t-\tau)^{1-(q/2)}}d\tau$$

$$-\frac{M_1(x_0)}{2}n(t)+\frac{1}{2}\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{(k+1)!}M_{k+1}(x_0)\sigma^{kq/2}\frac{d^{kq/2}}{dt^{kq/2}}n(t)$$

Observe that the first term involves a fractional integral (the Riemann-Liouville integral), the second term is composed of the source function $n(t)$ alone (apart from scaling) and the third term is an infinite series composed of fractional differentials of increasing order $kq/2$. Also note that the first term is scaled by a factor involving $\sigma^{-q/2}$ whereas the third term is scaled by a factor that includes $\sigma^{kq/2}$ so that

$$\lim_{\sigma\to 0}u(x_0,t) = -\frac{M_1(x_0)}{2}n(t)$$

$$+\frac{M_0(x_0)}{2\sigma^{q/2}}\frac{1}{(2i)^q\sqrt{\pi}}\frac{\Gamma\left(\frac{1-q}{2}\right)}{\Gamma\left(\frac{q}{2}\right)}\int_{-\infty}^{\infty}\frac{n(\tau)}{(t-\tau)^{1-(q/2)}}d\tau$$

5.1.3 Solution for $q < 0$

In this case, the first term becomes

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{iN(\omega)M_0}{2\Omega_q}\exp(i\omega t)d\omega$$

$$=\frac{M_0}{2}\sigma^{\frac{q}{2}}\frac{1}{2\pi}\int_{-\infty}^{\infty}N(\omega)(i\omega)^{\frac{q}{2}}\exp(i\omega t)d\omega=\frac{M_0}{2}\sigma^{\frac{q}{2}}\frac{d^{\frac{q}{2}}}{dt^{\frac{q}{2}}}n(t)$$

The second term is the same is in the previous case (for $q > 0$) and the third term is

$$-\frac{iM_2}{2.2!}\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{N(\omega)i}{(i\omega\sigma)^{\frac{q}{2}}}\exp(i\omega t)d\omega$$

$$=\frac{M_2}{2.2!}\frac{1}{\sigma^{q/2}}\frac{1}{(2i)^q\sqrt{\pi}}\frac{\Gamma\left(\frac{1-q}{2}\right)}{\Gamma\left(\frac{q}{2}\right)}\int_{-\infty}^{\infty}\frac{n(\tau)}{(t-\tau)^{1-(q/2)}}d\tau$$

Evaluating the other terms, by induction we obtain

$$u(x_0,t)=\frac{M_0(x_0)\sigma^{q/2}}{2}\frac{d^{q/2}}{dt^{q/2}}n(t)-\frac{M_1(x_0)}{2}n(t)$$

$$+\frac{1}{2}\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{(k+1)!}\frac{M_{k+1}(x_0)}{\sigma^{kq/2}}\frac{1}{(2i)^{kq}\sqrt{\pi}}\frac{\Gamma\left(\frac{1-kq}{2}\right)}{\Gamma\left(\frac{kq}{2}\right)}$$

$$\times\int_{-\infty}^{\infty}\frac{n(\tau)}{(t-\tau)^{1-(kq/2)}}d\tau$$

Here, the solution is composed of three terms: a fractional differential, the source term and an infinite series of fractional integrals of order $kq/2$. Thus, the roles of fractional differentiation and fractional integration are reversed as q changes from being greater than to less than zero.

5.2 Fractional Differentials

Fractional differentials $\hat{D}^q$ of any order q need to be considered in terms of the definition for a fractional differential given by

$$\hat{D}^q f(t)=\frac{d^n}{dt^n}[\hat{I}^{n-q}f(t)],\quad n-q>0$$

where n is an integer and $\hat{I}$ is the fractional integral operator (the Riemann-Liouville transform),

$$\hat{I}^p f(t) = \frac{1}{\Gamma(p)} \int_{-\infty}^{t} \frac{f(\tau)}{(t-\tau)^{1-p}} d\tau, \quad p > 0$$

The reason for this is that direct fractional differentiation can lead to divergent integrals. However, there is a deeper interpretation of this result that has a synergy with the issue over whether a fractionally diffusive system has 'memory' and is based on observing that the evaluation of a fractional differential operator depends on the history of the function in question. Thus, unlike an integer differential operator of order n, a fractional differential operator of order q has 'memory' because the value of $\hat{I}^{q-n} f(t)$ at a time t depends on the behaviour of $f(t)$ from $-\infty$ to t via the convolution with $t^{(n-q)-1}/\Gamma(n-q)$. The convolution process is of course dependent on the history of a function $f(t)$ for a given kernel and thus, in this context, we can consider a fractional derivative defined via the result above to have memory. In this sense, the operator

$$\frac{\partial^2}{\partial x^2} - \sigma^q \frac{\partial^q}{\partial t^q}$$

describes a process that has a memory association with the temporal characteristics of the system it is attempting to model. This is not an intrinsic characteristic of systems that are purely diffusive $q = 1$ or propagative $q = 2$. Appendix C provides a short overview of the fractional calculus to accompany the solution methods used here with regard to the fractional diffusion equation.

5.3 Asymptotic Solutions for an Impulse

We consider a special case in which the source function $f(x)$ is an impulse so that

$$M_m(x_0) = \int_{-\infty}^{\infty} \delta(x) \mid x - x_0 \mid^m dx = \mid x_0 \mid^m$$

This result immediately suggests a study of the asymptotic solution

$$u(t) = \lim_{x_0 \to 0} u(x_0, t)$$

$$= \begin{cases} \frac{1}{2\sigma^{q/2}} \frac{1}{(2i)^q \sqrt{\pi}} \frac{\Gamma\left(\frac{1-q}{2}\right)}{\Gamma\left(\frac{q}{2}\right)} \int_{-\infty}^{\infty} \frac{n(\tau)}{(t-\tau)^{1-(q/2)}} d\tau, & q > 0; \\ \frac{n(t)}{2}, & q = 0; \\ \frac{\sigma^{q/2}}{2} \frac{d^{q/2}}{dt^{q/2}} n(t), & q < 0. \end{cases}$$

The solution for the time variations of the stochastic field u for $q > 0$ are then given by a fractional integral alone and for $q < 0$ by a fractional differential alone. In

particular, for $q > 0$, we see that the solution is based on the convolution integral (ignoring scaling)

$$u(t) = \frac{1}{t^{1-q/2}} \otimes n(t), \quad q > 0 \tag{7}$$

where $\otimes$ denotes convolution integral over t which in ω-space is given by (ignoring scaling)

$$U(\omega) = \frac{N(\omega)}{(i\omega)^{q/2}}$$

This result is the conventional random fractal noise model for Fourier Dimension q [25], [26] and [27]. Table 1 quantifies the results for different values of q with conventional name associations. Note that Brown noise conventionally refers to the integration of white noise but that Brownian motion is a form of pink noise because it classifies diffusive processes identified by the case when $q = 1$. The field u has the following fundamental property for $q > 0$:

$$\lambda^{q/2}\Pr[u(t)] = \Pr[u(\lambda t)]$$

This property describes the statistical self-affinity of u. Thus, the asymptotic solution considered here, yields a result that describes a random scaling fractal field characterized by a Power Spectral Density Function (PSDF) of the form $1/ \mid \omega \mid^q$ which is a measure of the time correlations in the signal.

q-value	t-space	ω-space (PSDF)	Name
$q = 0$	$\frac{1}{t} \otimes n(t)$	1	White noise
$q = 1$	$\frac{1}{\sqrt{t}} \otimes n(t)$	$\frac{1}{\lvert\omega\rvert}$	Pink noise
$q = 2$	$\int^t n(t)dt$	$\frac{1}{\omega^2}$	Brown noise
$q > 2$	$t^{(q/2)-1} \otimes n(t)$	$\frac{1}{\lvert\omega\rvert^q}$	Black noise

Table 1: Noise characteristics for different values of q. Note that the results given above ignore scaling factors.

Note that $q = 0$ defines the Hilbert transform of $n(t)$ whose spectral properties

in the positive half space are identical to $n(t)$ because

$$\frac{1}{t} \otimes n(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} [-i\pi \text{sign}(\omega)] N(\omega) \exp(i\omega t) d\omega$$

where

$$\text{sgn}(\omega) = \begin{cases} 1, & \omega > 0; \\ -1, & \omega < 0. \end{cases}$$

The statistical properties of the Hilbert transform of $n(t)$ are therefore the same as $n(t)$ so that

$$\Pr[t^{-1} \otimes n(t)] = \Pr[n(t)]$$

Hence, as $q \to 0$, the statistical properties of $u(t)$ 'reflect' those of n, i.e.

$$\Pr\left[\frac{1}{t^{1-q/2}} \otimes n(t)\right] = \Pr[n(t)], \quad q \to 0$$

However, as $q \to 2$ we can expect the statistical properties of $u(t)$ to be such that the width of the PDF of $u(t)$ is reduced. This reflects the greater level of coherence (persistence in time) associated with the stochastic field $u(t)$ for $q \to 2$.

The asymptotic solution compounded in equation (7) is equivalent to considering a model for u based on the equation

$$\left(\frac{\partial^2}{\partial x^2} - \frac{\partial^q}{\partial t^q}\right) u(x,t) = \delta(x) n(t), \quad q > 0, \quad x \to 0 \tag{8}$$

where $n(t)$ is 'white noise'. However, we note that for $\mid x \mid > 0$ the function $\exp(i\Omega_q \mid x \mid)$ does not affect the ω^{-q} scaling law of the power spectrum, i.e. $\forall x$,

$$\mid U(x,\omega) \mid^2 = \frac{\mid N(\omega) \mid^2}{4\sigma^q \omega^q}, \quad \omega > 0$$

Thus for a 'white noise' spectrum $N(\omega)$ the Power Spectrum Density Function of U is determined by ω^{-q} and any algorithm for computing an estimate of q given $\mid U(\omega) \mid^2$ therefore applies $\forall x$ and not just for the case when $x \to 0$. Since we can write

$$U(x,\omega) = N(\omega) \frac{i}{2\Omega_q} \exp(i\Omega_q \mid x \mid) = N(\omega) \frac{1}{2(i\omega\sigma)^{q/2}}$$

$$\times \left(1 + i(i\omega\sigma)^{q/2} \mid x \mid - \frac{1}{2!}(i\omega\sigma)^q \mid x \mid^2 + \ldots\right)$$

unconditionally, by inverse Fourier transforming, we obtain the following expression for $u(x,t)$ (ignoring scaling factors including σ):

$$u(x,t) = n(t) \otimes \frac{1}{t^{1-q/2}} + i \mid x \mid n(t)$$

$$+\sum_{k=1}^{\infty} \frac{i^{k+1}}{(k+1)!} \mid x \mid^{2k} \frac{d^{kq/2}}{dt^{kq/2}} n(t)$$

Here, the solution is composed of three terms composed of (i) a fractional integral, (ii) the source term $n(t)$; (iii) an infinite series of fractional differentials of order $kq/2$.

5.4 Rationale for the Model

A Hurst process describes fractional Brownian motion based on the generalization of Brownian motion quantified by a generalisation of the $\sqrt{t}$ to t^H, $H \in (0, 1]$ scaling law The latter scaling law makes no prior assumptions about any underlying distributions. It simply tells us how the system is scaling with respect to time. Processes of this type appear to exhibit cycles, but with no predictable period. The interpretation of such processes in terms of the Hurst exponent H is as follows: We know that $H = 0.5$ is consistent with an independently distributed system. The range $0.5 < H \leq 1$, implies a persistent time series, and a persistent time series is characterized by positive correlations. Theoretically, what happens today will ultimately have a lasting effect on the future. The range $0 < H \leq 0.5$ indicates anti-persistence which means that the time series covers less ground than a random process. In other words, there are negative correlations. For a system to cover less distance, it must reverse itself more often than a random process.

Given that random walks with $H = 0.5$ describe processes whose macroscopic behaviour is characterised by the diffusion equation, then, by induction, Hurst processes should be characterised by generalizing the diffusion operator

$$\frac{\partial^2}{\partial x^2} - \sigma \frac{\partial}{\partial t}$$

to the fractional form

$$\frac{\partial^2}{\partial x^2} - \sigma^q \frac{\partial^q}{\partial t^q}$$

where $q \in (0, 2]$. Fractional diffusive processes can therefore be interpreted as intermediate between classical diffusive (random phase walks with $H = 0.5$; diffusive processes with $q = 1$) and 'propagative process' (coherent phase walks for $H = 1$; propagative processes with $q = 2$), e.g. [43] and [44]. The relationship between the Hurst exponent H, the Fourier dimension q and the Fractal dimension D_F is given by (see Appendix D)

$$D_F = D_T + 1 - H = 1 - \frac{q}{2} + \frac{3}{2} D_T$$

where D_T is the topological dimension. Thus, a Brownian process, where $H = 1/2$, has a fractal dimension of 1.5.

Fractional diffusion processes are based on random walks which exhibit a bias with regard to the distribution of angles used to change the direction. By induction, it can be expected that as the distribution of angles reduces, the corresponding walk becomes more and more coherent, exhibiting longer and longer time correlations until the process conforms to a fully coherent walk. A simulation of such an effect is given in Figure 6 which shows a random walk in the (real) plane as the (uniform) distribution of angles decreases. The walk becomes less and less random as the width of the distribution is reduced. Each position of the walk $(x_j, y_j), \quad j = 1, 2, 3, ..., N$ is computed using

$$x_j = \sum_{i=1}^{j} \cos(\theta_i), \quad y_j = \sum_{i=1}^{j} \sin(\theta_i)$$

where

$$\theta_i = \alpha\pi \frac{n_i}{\|\mathbf{n}\|_\infty}$$

and n_i are random numbers computed using the linear congruential pseudo random number generator

$$n_{i+1} = an_i \mathrm{mod} P, \quad i = 1, 2, ..., N, \quad a = 7^7, \quad P = 2^{31} - 1$$

The parameter $0 \leq \alpha \leq 2\pi$ defines the width of the distribution of angles such that as $\alpha \to 0$, the walk becomes increasingly coherent or 'propagative'.

In considering a t^H scaling law with Hurst exponent $H \in (0, 1]$, Hurst paved the way for an appreciation that most natural stochastic phenomena which, at first site, appear random, have certain trends that can be identified over a given period of time. In other words, many natural random patterns have a bias to them that leads to time correlations in their stochastic behaviour, a behaviour that is not an inherent characteristic of a random walk model and fully diffusive processes in general. This aspect of stochastic field theory is the basis for Lévy processes [21] and it is through application of a Lévy distribution that we have derive the fractional diffusion equation as shown in Section 4.2.

5.5 Non-stationary Model

Real financial signals are non-stationary. An example of this is illustrated in Figure 7 which shows a non-stationary walk in the complex plane obtained by taking the Hilbert transform of the corresponding signal, i.e. computing the analytic signal

$$s(t) = u(t) + \frac{i}{\pi t} \otimes u(t)$$

and plotting the real and imaginary component of this signal in the complex plane.

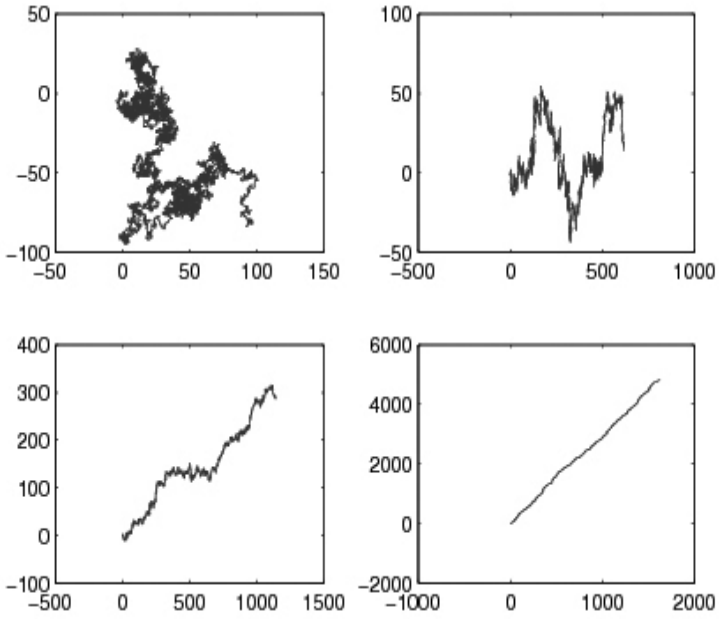

Figure 6: Random phase walks in the plane for a uniform distribution of angles $\theta_i \in [0, 2\pi]$ (top left), $\theta_i \in [0, 1.9\pi]$ (top right), $\theta_i \in [0, 1.8\pi]$ (bottom left) and $\theta_i \in [0, 1.2\pi]$ (bottom right).

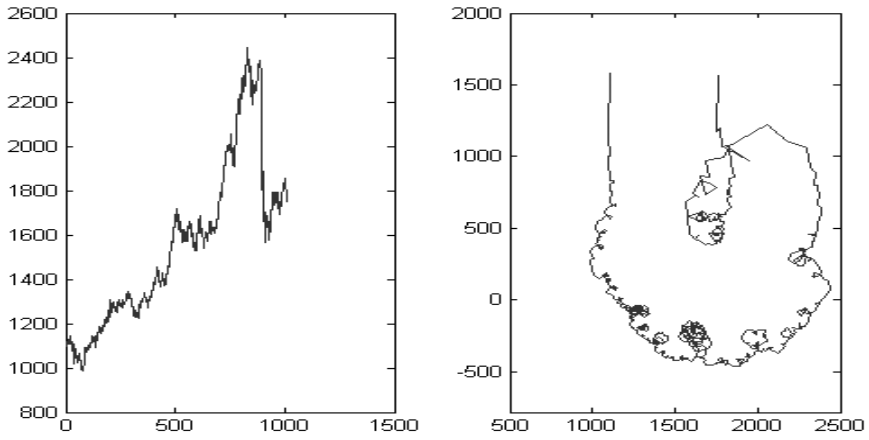

Figure 7: Non-stationary behaviour in the complex plane (right) obtained by computig the Hilbert transform of the signal (left) - FTSE Close-of-Day from 02-04-1984 to 24-12-1987.

The fractional diffusion operator in equation (8) is appropriate for modelling fractional diffusive processes that are stationary. Since financial signals are highly non-stationary, the model must therefore be extended to the non-stationary case if it is to be of value for financial signal analysis. In developing a non-stationary fractional diffusive model, we could consider the case where the diffusivity is time variant as defined by the function $\sigma(t)$. However, a more interesting case arises when the characteristics of the diffusion process itself change over time becoming more or less diffusive as $q \to 1$ and $q \to 2$ respectively. In thise sense we consider q to be a function of time τ and introduce a non-stationary fractional diffusion operator given by

$$\frac{\partial^2}{\partial x^2} - \sigma^{q(\tau)} \frac{\partial^{q(\tau)}}{\partial t^{q(\tau)}}$$

This operator is the theoretical basis for the results presented in this work. In terms of using this model to develop an asset management metric that is predictive of future behaviour we consider the following hypothesis: a change in $q(\tau)$ precedes an associated change in a financial signal. The goal of this model/hypothesis is to be able to compute a function - namely $q(\tau)$ - which is a measure of the non-stationary behaviour of an economic signal with regard to its future behaviour. This is because, in principle, the value of $q(\tau)$ should reflect the early stages of a change in the behaviour of $u(t)$ provided an accurate enough numerical algorithm for computing $q(\tau)$ is available. This is discussed in the following section.

6 Financial Signal Analysis

If we consider the case where $q(\tau)$ is a relatively slowly varying function of time so that

$$\left|\frac{\partial q}{\partial \tau}\right| << \left|\frac{\partial u}{\partial t}\right|$$

then we can consider $q(\tau)$ to be constant over a 'window of time' τ. Ignoring scaling, for a quasi-stationary segment of a financial signal,

$$u(t,\tau) = \frac{1}{t^{1-q(\tau)/2}} \otimes n(t), \quad q > 0$$

which has characteristic spectrum

$$U(\omega,\tau) = \frac{N(\omega)}{(i\omega)^{q(\tau)/2}}$$

The PSDF is characterised by $\omega^{-q(\tau)}$, $\omega > 0$ and our problem is thus, to compute $q(\tau)$ from the data

$$P(\omega,\tau) =\mid U(\omega,\tau) \mid^2 = \frac{\mid N(\omega) \mid^2}{\omega^{q(\tau)}}, \quad > 0$$

. Consider the PSDF

$$\hat{P}(\omega,\tau) = \frac{c}{\omega^{q(\tau)}}$$

with logarithmic transformation

$$\ln \hat{P}(\omega,\tau) = C + q(\tau) \ln \omega$$

where $C = \ln c$. The problem is then reduced to implementing an appropriate method to compute q (and C) by finding a best fit of the line $\ln \hat{P}(\omega)$ to the data $\ln P(\omega)$. Application of the least squares method for computing q for a given τ, which is based on minimizing the error

$$e(q,C) = \| \ln P(\omega) - \ln \hat{P}(\omega,q,C) \|_2^2$$

with regard to q and C, leads to errors in the estimates for q. The reason for this is that relative errors at the start and end of the data $\ln P$ may vary significantly especially because any errors inherent in the data P will be 'amplified' through application of the logarithmic transform required to linearise the problem. In general, application of a least squares approach is very sensitive to statistical heterogeneity [28] and may provide values of q that are not compatible with the rationale associated with the model (i.e. values of $1 < q < 2$ that are intermediate between diffusive and propagative processes). For this reason, an alternative approach is considered which, in this work, is based on Orthogonal Linear Regression (OLR) [29].

Applying a standard moving window, $q(\tau)$ is computed by repeated application of OLR based on the m-code available from [30]. This provides a numerical estimate of the function $q(\tau)$ whose behaviour 'reflects' the 'state' of the financial signal. Since q is, in effect, a statistic, its computation is only as good as the quantity (and quality) of data that is available for its computation. For this reason, a relatively large window is required whose length is compatible with the number of samples available.

6.1 Numerical Algorithm

The principal numerical algorithm associated with the application of the model as follows:

Step 1: Read data (financial time series) from file into operating array $a[i], i = 1, 2, ..., N$.

Step 2: Set length $L < N$ of moving window w to be used.

Step 3: For $j = 1$ assign $L + j - 1$ elements of $a[i]$ to array $w[i], i = 1, 2, ..., L$.

Step 4: Compute the power spectrum $P[i]$ of $w[i]$ using a Discrete Fourier Transform (DFT).

Step 5: Compute the logarithm of the spectrum excluding the DC, i.e. compute $\log(P[i])\forall i \in [2, L/2]$.

Step 6: Compute $q[j]$ using the OLR algorithm.

Step 7: For $j = j + 1$ repeat Step 3 - Step 5 stopping when $j = N - L$.

Step 8: Write the signal $q[j]$ to file for further analysis and post processing.

The following points should be noted:

(i) The DFT is taken to generate an output in standard form where the zero frequency component of the power spectrum is taken to be $P[1]$.

(ii) With $L = 2^m$ for integer m, a Fast Fourier Transform can be used

(iii) The minimum window size that should be used in order provide statistically significant values of $q[j]$ is $L = 64$ when q can be computed accurate to 2 decimal places.

The m-code for this algorithm is given in Appendix F and an example of the output generated by this algorithm for a 1024 element 'look-back' window is given in Figure 8 using Dow Jones Industrial Average Close-of-Day data obtained from [31]. Table 2 provides some basic statistical information with regard to $q(\tau)$ for this data whose mean value is 1.27. Application of the Null Hypothesis test with

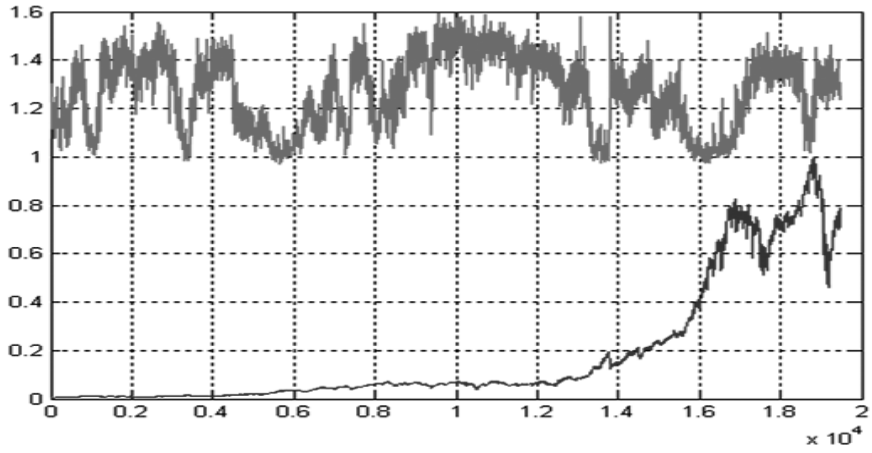

Figure 8: Computation of $q(\tau) \in [1, 2]$ (above) using a 1024 element window obtained from a time series composed of Dow Jones Industrial Average Close-of-Day data $u(t)$ from from 02-11-1932 to 24-05-2010 after normalisation (below).

regard to whether or not $q(\tau)$ is Gaussian distributed is negative, i.e. the 'Composite Normality' is of type 'Reject'. Thus, $q(\tau)$ is not normally distributed.

A closer inspection of data such as that given in Figure 8 reveals a qualitative relationship between trends in the financial signal $u(t)$ and $q(\tau)$ in accordance with the theoretical model considered. In particular, over periods of time in which $q(\tau)$ increases in value, the amplitude of the financial signal $u(t)$ decreases. Moreover, and more importantly, an upward trend in $q(\tau)$ appears to be a precursor to a downward trend in $u(t)$, a correlation that is compatible with the idea that a rise in the value of $q(\tau)$ relates to the 'system' becoming more propagative, which in stock market terms, indicates the likelihood for the markets becoming 'bear' dominant in the future.

Results of the type given in Figure 8 not only provides for a general appraisal of different macroeconomic financial time series, but, with regard to the size of the selected window used, an analysis of data at any point in time. The output can be interpreted in terms of 'persistence' and 'anti-persistence' and in terms of the existence or absence of after-effects (macroeconomic memory effects). For those periods in time when $q(\tau)$ is relatively constant, the existing market tendencies usually remain. Changes in the existing trends tend to occur just after relatively sharp changes in $q(\tau)$ have developed. This behaviour indicates the possibility of using the time series $q(\tau)$ for identifying the behaviour of a macroeconomic financial system in terms of both inter-market and between-market analysis. These results support the possibility of using $q(\tau)$ as an independent volatility predictor to give a risk assessment associated with the likely future behaviour of different time series.

Further, because this analysis is based on a self-affine model, the interpretation of a financial signal via $q(\tau)$ should, in principle, be scale invariant.

Statistical Parameter	Value for $q(\tau)$
Minimum Value	0.9702
Maximum value	1.5947
Range	0.6245
Mean	1.2662
Median	1.2936
Standard Deviation	0.1444
Variance	0.0209
Skew	-0.2422
Kertosis	1.9043
Composite Normality	Reject

Table 2: Statistical parameters associated with $q(\tau)$ computed for the Dow Jone Industrial Average Close-of-Day data from 02-11-1928 to 24-05-2010 given in Figures 8

6.2 Macrotrend Analysis

In order to develop a macrotrend signal that has optimal properties with regard the assessment of risk in terms of the likely long-term future behaviour of a financial signal, it is important that the filter used is: (i) consistent with the properties of a Variation Diminishing Smoothing Kernel (VDSK); (ii) that the last few values of the trend signal are 'data consistent'. VDSKs are convolution kernels with properties that guarantee smoothness around points of discontinuity of a given signal where the smoothed function is composed of a similar succession of concave or convex arcs equal in number to those of the signal. VDSKs also have 'geometric properties' that preserve the 'shape' of the signal. There are a range of VDSKs including the Gaussian function and, for completeness, Appendix E provides an overview of the principal analytical properties, including some basic Theorem's and Proof's relating to VDSKs.

In practice, the computation of the smoothing process using a VDSK must be performed in such a way that the initial and final elements of the output data are entirely data consistent with the input array in the locality of any element. Since a VDSK is a non-localised filter which tends to zero at infinity (see Appendix E), in order to optimise the numerical efficiency of the smoothing process, filtering is undertaken in Fourier space. However, in order to produce a data consistent macrotrend signal using a Discrete Fourier Transform, wrapping effects must be eliminated. The solution is to apply an 'end point extension' scheme which involves

padding the input vector with elements equal to the first and last values of the vector. The length of the 'padding vectors' is taken to be at least half the size of the input vector. The output vector is obtained by deleting the filtered padding vectors.

Figure 9 is an example of a macrotrend analysis obtained using the VDSK filter $\exp(-\alpha\omega^2)$ which include the normalised gradients computed using a 'forward differencing scheme' illustrating 'phase shifts' associated with the two signals.

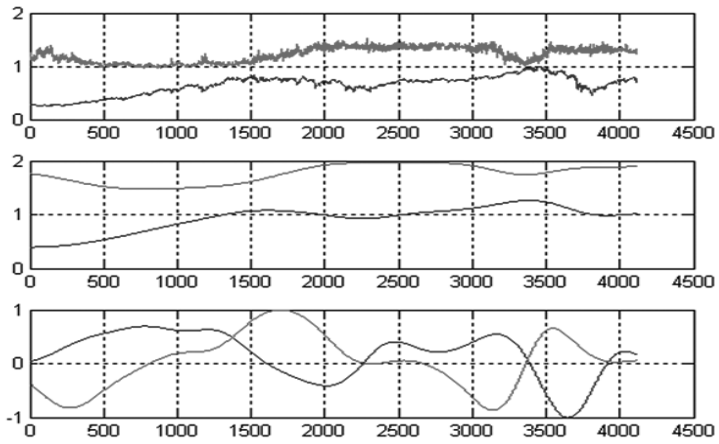

Figure 9: Analysis of Dow Jones Industrial Average Close-of-Day data from 18-01-1994 to 25-05-2010. Top: Data $u(t) \in [0,1]$ and $q(\tau) \in [1,2]$ computed using a window of 1024; Centre: Associated macrotrends obtained for $\alpha = 0.01$; Bottom: Normalised gradients of macrotrends computed using a forward differencing scheme.

7 Case Study I: Market Volatility

Since the publication of the preference-free option by Black and Sholes in 1973 [9], option pricing theory has developed into a standard tool for designing, pricing and hedging derivative securities of all types. For an ideal market, six inputs are required: the current stock price, the strike price, the time to expiry, the risk-free interest rate, the dividend and the volatility. The first three inputs are known from the outset but the last three must be estimated. The problem with this is that the correct values for these parameters are only known when the option expires which means that the future values of these quantities need to be known or accurately estimated if an option is to be priced correctly. The most important of these parameters is the volatility. This is because a change in volatility has the biggest impact on the price of an option. Volatility measures variability or dispersion about a central tendency and is a measure of the degree of price movement in a stock, a futures contract or any other market.

Fundamental to the Black-Scholes model is that financial asset prices are random variables that are log-normally distributed. Therefore, the relative price changes are usually measured by taking the differences between the logarithmic prices which are taken to be normally distributed. This is consistent with the log-normal model discussed in Section 1 where the volatility is proportional to $d(\ln S)/dt$. Thus, for a stock price signal $(S_1, S_2, ...S_n)$, using a forward differencing scheme, we let

$$v_i = \ln\left(\frac{S_{i+1}}{S_i}\right)$$

and define the volatility as

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(v_i - \bar{v})^2}$$

where $\bar{v}$ is the mean given by

$$\bar{v} = \frac{1}{n}\sum_{i=1}^{n} v_i$$

This result defines the volatility in terms of the standard deviation of the sample signal v_i which is a measure of the dispersion of a stock price signal whose variations in time are computed using a moving window.

In order to compare volatilities for different interval lengths, it is common to express volatility in annual terms by scaling the estimate with an annualisation factor (normalising constant) h which is the number of intervals per annum such that $\sigma_{\text{annual}} = \sigma\sqrt{h}$ where $h = 252$ for daily data (252 being approximately, the number of trading days per annum), $h = 52$ for weekly data and $h = 12$ for monthly data. Defining volatility in terms of variations in the standard deviation of an assets

returns from the mean implies that large values of volatility fluctuate over a wide range leading to high risk. This means that if we assume that the mean of the log-relative returns is zero, then, a 10% volatility represents the following (according to a normal distribution): in one year, returns will be within [-10%, +10%] with a probability of 68.3% (1 standard deviation from the mean); within [-20% , +20%] with a probability of 95.4% (2 standard deviations), and within [-30% , +30%] with a probability of 99.7% (3 standard deviations). For this reason volatility is usually presented in terms of a percentage.

In this case study we use the value of $q(\tau)$ to predict the volatility associated with foreign exchange markets. This is based on observing the positions in $q(\tau)$ where the derivative is zero (i.e. when $q'(\tau) \equiv dq(\tau)/d\tau = 0$) after generating a macrotrend (as discussed in Section 6.2) and applying a threshold. Above this threshold, the volatility is predicted to reduce and below the threshold, it is predicted to increase. The rationale for this approach relates to the idea that as q approaches 1 the 'economy' becomes diffusive and as q approaches 2, it becomes propagative. This is equivalent to the Lévy index approaching 2 and 1, respectively. The problem is to determine optimum values for the length of the moving window L, the smoothing parameter α and the threshold T which is achieved by undertaking a range of tests using historical data. Figure 10 shows an example of this approach based on the US to Australian dollar exchange rate from 03/08/2003 to 16/04/2010, close-of-day. The positions at which $q'(\tau) = 0$ coupled with the applied threshold, define the 'critical points' at which a prediction is made on the future volatility that, in this example, is computed using a 30 day 'look-back' window. This depends upon whether the value of q at which $q'(\tau) = 0$ is above (volatility predicted to decrease) or below (volatility predicted to increase) the threshold.

Table 3 provides a quantification of the method. Here, the date at which $q'(t) = 0$ is provide together with a prediction on whether the volatility will go up or down, the volatility at the date given (i.e. the date at which the prediction is made) and the volatility 30 days after the date of the prediction. Also provided is a *True/False* criterion associated with each prediction that, for this example, yields a prediction accuracy of 69% .

Figure 11 shows the same analysis applied to the Japanese Yen to US Dollar exchange from 03/08/2003 to 18/04/2010 with an assessment of the predictions being given in Table 4 that, in this example, provides 75% accuracy. As a final example, Table 5 shows the results obtained for the Euro/US Dollar exchange rate from 28/10/2002 to 09/04/2010.

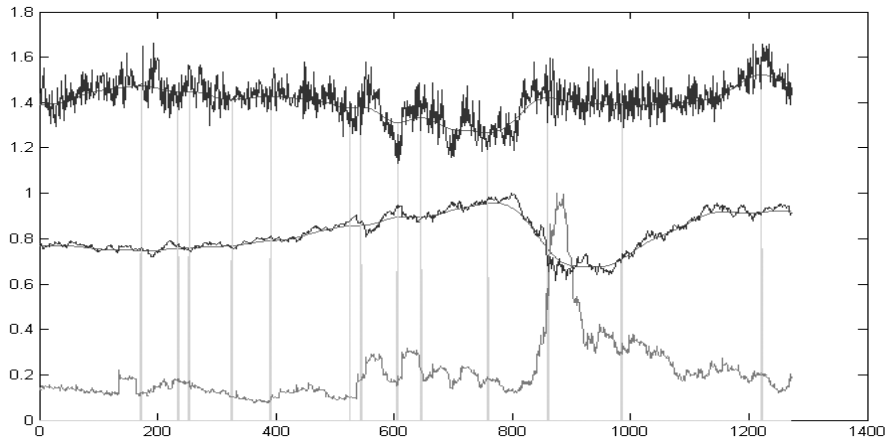

Figure 10: US/Australian dollar exchange rate from 03/08/2003 to 16/04/2010 for $L = 512$, $\alpha = 0.002$ and $T = 1.4$. The signals are as follows: On the scale of 1 to 2 are $q(\tau)$ and the associated macrotrend; on the scale of 0 to 1 are the normalised exchange, associated macrotrend and the normalised volatility (grey). The positions in time at which $q' = 0$ is indicated by the spikes.

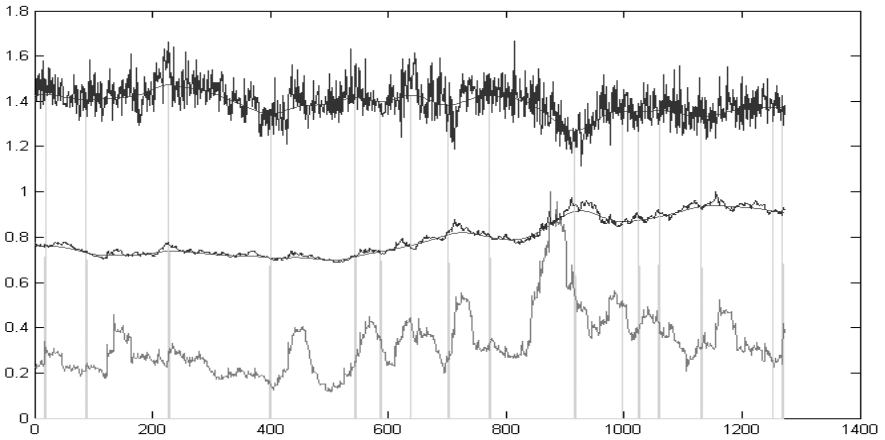

Figure 11: Japanese Yen to US dollar exchange rate from 03/08/2003 to 18/04/2010 for $L = 512$, $\alpha = 0.002$ and $T = 1.4$. The signals are as follows: On the scale of 1 to 2 are $q(\tau)$ and the associated macrotrend; on the scale of 0 to 1 are the normalised exchange data, associated macrotrend and the normalised volatility (grey). The positions in time when $q' = 0$ is indicated by the spikes.

Date at which $q'(t) = 0$	Prediction: Up $\Uparrow$ or Down $\Downarrow$	30 day volatility at date	30 day volatility after date	True/False $\surd/\times$
2 Mar 2006	$\Downarrow$	7.34%	9.58%	$\times$
29 May 2006	$\Downarrow$	11.58%	9.95%	$\surd$
23 Jun 2006	$\Downarrow$	10.94%	7.93%	$\surd$
4 Oct 2006	$\Uparrow$	6.59%	5.95%	$\times$
3 Jan 2007	$\Uparrow$	5.58%	8.52%	$\surd$
11 Jul 2007	$\Uparrow$	6.51%	16.71%	$\surd$
7 Aug 2007	$\Uparrow$	11.42%	18.26%	$\surd$
1 Nov 2008	$\Uparrow$	10.00%	18.94%	$\surd$
26 Feb 2008	$\Uparrow$	15.58%	14.52%	$\times$
20 May 2008	$\Uparrow$	11.57%	8.07%	$\times$
9 Oct 2009	$\Uparrow$	34.55%	59.36%	$\surd$
2 Apr 2010	$\Uparrow$	18.69%	21.54%	$\surd$
1 Mar 2010	$\Downarrow$	13.23%	8.27%	$\surd$

Table 3: Assessment of volatility predictions obtained for the US dollar to Australian dollar exchange rate from 03/08/2003 to 16/04/2010.

Date at which $q'(t) = 0$	Prediction: Up $\Uparrow$ or Down $\Downarrow$	30 day volatility at date	30 day volatility after date	True/False $\surd/\times$
25 Jul 2005	$\Downarrow$	9.53%	7.07%	$\surd$
27 Oct 2005	$\Uparrow$	7.59%	10.29%	$\surd$
19 May 2006	$\Downarrow$	9.55%	8.33%	$\surd$
16 Jan 2007	$\Uparrow$	5.08%	10.27%	$\surd$
7 Aug 2007	$\Uparrow$	8.03%	11.24%	$\surd$
5 Oct 2007	$\Uparrow$	11.49%	11.44%	$\times$
17 Dec 2007	$\Downarrow$	14.23%	10.48%	$\surd$
5 Mar 2008	$\Uparrow$	8.66%	14.13%	$\surd$
9 Jun 2008	$\Downarrow$	10.90%	9.30%	$\surd$
26 Dec 2008	$\Uparrow$	17.56%	15.17%	$\times$
20 Apr 2009	$\Uparrow$	16.57%	13.66%	$\times$
28 May 2009	$\Uparrow$	12.77%	13.24%	$\surd$
16 Jul 2009	$\Uparrow$	13.80%	10.33%	$\times$
26 Oct 2009	$\Uparrow$	9.94%	13.41%	$\surd$
13 Apr 2010	$\Uparrow$	8.29%	14.20%	$\surd$
6 May 2010	$\Uparrow$	11.49%	14.42%	$\surd$

Table 4: Assessment of volatility predictions associated with Japanese Yen to US dollar exchange rate from 03/08/2003 to 18/04/2010.

Date at which $q'(t) = 0$	Prediction: Up $\Uparrow$ or Down $\Downarrow$	30 day volatility at date	60 day volatility after date	True/False $\surd/\times$
17 Dec 2003	$\Uparrow$	9.82%	11.09%	$\surd$
17 Jun 2004	$\Downarrow$	12.17%	9.09%	$\surd$
3 Nov 2004	$\Uparrow$	7.70%	8.56%	$\surd$
20 Jul 2005	$\Downarrow$	10.43%	9.03%	$\surd$
27 Dec 2005	$\Uparrow$	8.62%	8.90%	$\surd$
26 Jun 2006	$\Downarrow$	9.42%	7.97%	$\surd$
7 Feb 2007	$\Uparrow$	5.65%	5.42%	$\times$
28 May 2007	$\Uparrow$	4.56%	5.26	$\surd$
12 Mar 2008	$\Uparrow$	7.64%	8.95%	$\surd$
19 Nov 2008	$\Downarrow$	22.16%	20.64%	$\surd$
12 Mar 2009	$\Uparrow$	14.21%	20.64%	$\surd$
29 Jun 2009	$\Downarrow$	18.98%	10.92%	$\surd$
30 Sep 2009	$\Uparrow$	7.93%	8.66%	$\surd$

Table 5: Assessment of volatility predictions associated with Euro to US dollar exchange rate from 28/10/2002 to 09/04/2010

8 Case Study II: Analysis of ABX Indices

ABX indices serve as a benchmark of the market for securities backed by home loans issued to borrowers with weak credit. The index is administered by the London-based Markit Group which specialises in credit derivative pricing [45].

8.1 What is an ABX index?

The index is based on a basket of Credit Default Swap (CDS) contracts for the subprime housing equity sector. Credit Default Swaps operarte as a type of insurance policy for banks or other holders of bad mortgages. If the mortgage goes bad, then the seller of the CDS must pay the bank for the lost mortgage payments. Alternatively, if the mortgage stays good then the seller makes a lot of money. The riskier the bundle of mortgages the lower the rating.

The original goal of the index was to create visibility and transparency but it was not clear at the time of its inception that the index would be so closely followed. As subprime securities have become increasingly uncertain, the ABX index has become a key point of reference for investors navigating risky mortgage debt on an international basis. Hence, in light of the current financial crisis (i.e. from 2008-date), and given that most economist agree that the subprime mortgage was a primary catalyst for the crisis, analysis of the ABX index has become a key point of reference for investors navigating the world of risky mortgage debt.

On asset-backed securities such as home equity loans the CDS provides an insurance against the default of a specific security. The index enables users to trade in a security without being limited to the physical outstanding amount of that security thereby given investors liquid access to the most frequently traded home equity tranches in a basket form. The ABX uses five indices that range from triple-A to triple-B minus. Each index chooses deals from 20 of the largest sub-prime home equity shelves by issuance amount from the previous six months. The minimum deal size is $500 million and each tranche referenced must have an average life of between four and six years, except for the triple-A tranche, which must have a weighted average life greater than five years. Each of the indices is referenced to by different rated tranches, i.e. AAA, AA, A, BBB and BBB-. They are selected through identification of the most recently issued deals that meet the specific size and diversity criteria. The principal 'market-makers' in the index were/are: Bank of America, Bear Stearns, Citigroup, Credit Suisse, Deutsche Bank, Goldman Sachs, J P Morgan, Lehman Brothers, Merrill Lynch (now Bank of America), Morgan Stanley, Nomura International, RBS Greenwich Capital, UBS and Wachovia. However, during the financial crisis that developed in 2008, a number of changes have taken place. For example, on September 15, 2008, Lehman Brothers filed for bankruptcy protection following a massive exodus of most of its clients, drastic losses in its stock, and devaluation of its assets by credit rating agencies and in 2008 Merrill Lynch was acquired by Bank of America at which point Bank of America merged its global banking and wealth management division with the newly acquired firm. The Bear Stearns Companies, Inc. was a global investment bank and securities trading and brokerage, until its collapse and fire sale to J P Morgan Chase in 2008.

ABX contracts are commonly used by investors to speculate on or to hedge against the risk that the underling mortgage securities are not repaid as expected. The ABX swaps offer protection if the securities are not repaid as expected, in return for regular insurance-like premiums. A decline in the ABX index signifies investor sentiment that subprime mortgage holders will suffer increased financial losses from those investments. Likewise, an increase in the ABX index signifies investor sentiment looking for subprime mortgage holdings to perform better as investments.

8.2 ABX and the Sub-prime Market

Prime loans are often packaged into securities and sold to investors to help lenders reduce risk. More than $500B of such securities were issued in the US in 2006. The problem for investors who bought 2006's crop of high-risk mortgage originations, was that as the US housing market slowed as did mortgage applications. To prop up the market, mortgage lenders relaxed their underwriting standards lending to ever-riskier borrowers at ever more favourable terms.

In the last few weeks of 2006, the poor credit quality of the 2006 vintage subprime

mortgage origination started to become apparant. Delinquencies and foreclosures among high-risk borrowers increased at a dramatic rate, weakening the performance of the mortgage pools. In one security backed by subprime mortgages issued in March 2006, foreclosure rates were already 6.09% by December that year, while 5.52% of borrowers were late on their payments by more than 30 days. Lenders also began shutting their doors, sending shock waves through the high-risk mortgage markets throughout 2007. The problem kept new investor money at bay, and dramatically weakened a key derivative index tied to the performance of 2006 high-risk mortgages, i.e. the ABX index. As a result the ABX suffered a major plummet of the index starting in December 2006 when BBB- fell below 100 for the first time. The most heavily traded subindex, representing loans rated BBB-, fell as hedge funds flocked to bet on the downturn and pushed up the cost of insuring against default. This led to a knock-on effect as lenders withdrew from the ABX market

In early 2007 the issues were seen as: (i) Which investors were bearing the losses from having bought sub-prime mortgage backed securities? (ii) How large and concentrated were these losses? (iii) Had this sub-prime securitization distributed their risk among many players in the financial system or were the positions and losses concentrated among a few players? (iv) What were the potential systemic risk effects of these losses? We now know that the systemic risk had a devastating affect on the global economy and became known as the 'Credit Crunch'. One of the catalysts for the problem was a US bill allowing bankruptcy judges to alter loan balances which nobody dealing in CDS had considered. The second key factor was the speed of deterioration of the ABX Indices in 2007 which shocked investors and left them waiting to see the bottom of the market before getting back in - they are still waiting. The third key factor was the failure of the US Treasury to provide foreclosure relief for distressed home owners which congress had approved. The following series of reactions (denoted by $\rightarrow$) were triggered as a result: The treasury said it won't take steps to prevent home foreclosures, so that prices of mortgage securities collapsed $\rightarrow$ bank equity was wiped out $\rightarrow$ banks, with shrunken equity capital, were forced to cut back on all types of credit $\rightarrow$ financing for anything, especially residential mortgage loans, dried up $\rightarrow$ market values of homes declined further $\rightarrow$ mortgage securities declined further, and the downward spiral becomes self perpetuating.

8.3 Effect of ABX on Bank Equities

At the end of February 2007 a price of 92.5 meant that a protection buyer will need to pay the protection seller 7.5% upfront and then 0.64% per year. At the time, this kind of mortgage yield was about 6.5%, so the upfront charge was more than the yield per year. By April 2009 the A grade index had fallen to 8 meaning that the protection seller would want 92% upfront which meant that the sub-prime market 'died'. In July 2007 AAA mortgage securities started trading at prices materially

below par, or below 100. Until then, many banks had bulked up mortgage securities that were rated AAA at the time of issue. This was because they believed that AAA bonds could always be traded at prices close to par, and consequently the bonds' value would have a very small impact on the earnings and equity capital. The mystique about AAA ratings dated back more than 80 years. From 1920 onward, the default experience on AAA rated bonds, even during the Great Depression, was nominal.

The way the securities are structured is that different classes of creditors, or different tranches, all hold ownership interests in the same pool of mortgages. However, the tranches with the lower ratings - BBB, A, AA - take the first credit losses and they are supposed to be destroyed before the AAA bondholders lose anything. Typically, AAA bondholders represent about 75-80% of the entire mortgage pool. During the Great Depression (1929-1933), national average home prices held their value far better than they have since 2007. The assumptions that a highly liquid trading market and gradual price declines, have proved to be wrong. Beginning in the last half of 2007, the price declines of AAA bonds was steep, and the trading market suddenly became very illiquid. Under standard accounting rules, those securities must be marked to market every fiscal quarter, and the banks' equity capital shrank beyond all expectations. Hundreds of billions of dollars have been lost as a result. However, the losses in mortgage securities, and from financial institutions such Lehman that were undone by mortgage securities, dwarf everything else. Before the end of each fiscal quarter, bank managements must also budget for losses associated with mortgage securities. But since they cannot control market prices at a future date, they compensate by adjusting what they can control, which is all discretionary extensions of credit. Banks cannot legally lend beyond a certain multiple of their capital.

8.4 Credit Default Swap Index

This index is used to hedge credit risk or to take a position on a basket of credit entities. Unlike a credit default swap, a credit default swap index is a completely standardised credit security and may therefore be more liquid and trade at a smaller bid-offer spread. This means that it can be cheaper to hedge a portfolio of credit default swaps or bonds with a CDS index than to buy many CDS to achieve a similar effect. Credit-default swap indexes are benchmarks for protecting investors owning bonds against default, and traders use them to speculate on changes in credit quality. There are currently two main families of CDS indices: CDX and iTraxx. CDX indices contain North American and Emerging Market companies and are administered by CDS Index Company and marketed by Markit Group Limited, and iTraxx contain companies from the rest of the world and are managed by the International Index Company (IIC). A new series of CDS indices is issued every six months by Markit Group and IIC. Running up to the announcement of each series,

a group of investment banks is polled to determine the credit entities that will form the constituents of the new issue. This process is intended to ensure that the index does not become 'cluttered with instruments that no longer exist, or which trade illiquidly. On the day of issue a fixed coupon is decided for the whole index based on the credit spread of the entities in the index. Once this has been decided the index constituents and the fixed coupon is published and the indices can be actively traded.

8.5 Analysis of Sub-Prime CDS Market ABX Indices using the FMH

The US Sub-Prime Housing Market is widely viewed as the source of the current economic crisis. The reason that it has had such a devastating effect on the global economy is that investment grade bonds were purchased by many substantial international financial institutions but in reality the method used to designate the relatively low risk required for investment grade securities was seriously flawed. This resulted in the investment grade bonds becoming virtually worthless very quickly when systemic risks that wrongly had been ignored undermined the entire market. About 80% of the market was designated investment grade (AAA - highest, AA and A - lowest) with protection provided by a high risk grades (BBB- and BBB). The flawed risk model was based on an assumption that the investment grades would always be protected by the higher risk grades that would take all of the first 20% of defaults. Once defaults exceeded 20% the 'house of cards' was demolished. It is therefore of interest to see if a FMH based analysis of the ABX indices could have been used a predictive tool in order to develop a superior risk model.

Figure 12 shows the ABX index for each grade using data supplied by the Systemic Risk Assessment Division of the Bank of England. During the second week of December 2006 the BBB- index slipped to 99.76 for a couple of days but then recovered. In March 2007 the index for BBB- slipped just below 90 and seemed to be recovering and by mid-May was above 90 again. In June 2007 the BBB- really began to slide and this time it never recovered and was closely followed by the collapse of the BBB index after which there was no further protection for the investment grades. The default swaps work like an insurance so that if the cost of insuring against risk becomes greater than the annual return from the loan then the market is effectively dead. By February 2008 the AAA grade was below this viable level.

The results of applying the FMH based on the algorithms discussed in Section 6 is given in Figures 13-17. Table 6 provides a list of the statistical variables associated with $q(t)$ for each case. In each case, $q(t)$ initially has values > 2 but this falls rapidly prior to a change of the index. Also, in each case, the turning point of the normalised gradient of the Gaussian filtered signal (i.e. point in time of the minimum value) is an accurate reflection of the point in time prior to when the index

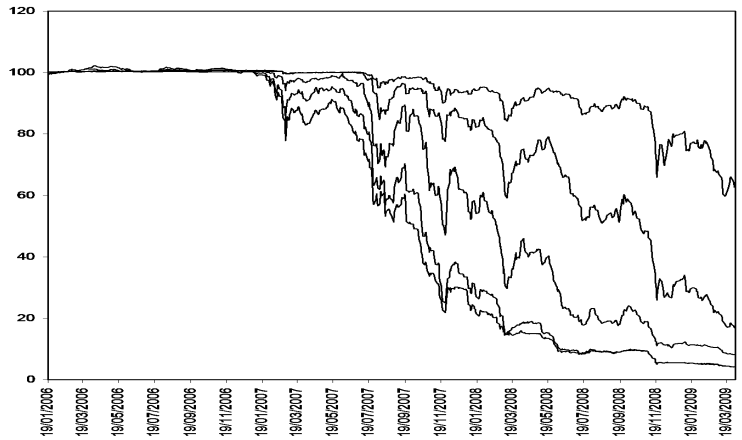

Figure 12: Grades for the ABX Indices from 19 January 2006 to 2 April 2009 based on Close-of-Day prices. The indices are differentiated by their increasing rate of decay with regard to the following order: AAA, AA, A, BBB, BBB-

falls rapidly relatively to the prior data. This turning point occurs before the equivalent characteristic associated with the smoothed index. The model consistently 'signals' the coming meltdown with sufficient notice for orderly withdrawal from the market. For example, the data used for Figure 13 reflects the highest Investment Grade and would be regarded as particularly safe. The normalised gradient of the output data provides a very early signal of a change in trend, in this case, at around approximately 180 days from the start of the run, which is equivalent to early April 2007 at which point the index was just above 100. In fact the AAA index appears to be viable as an investment right up to early November 2008 after which is falls dramatically. In Figure 15, a trend change is again observed in the normalised gradient at approximately 190 days which is equivalent mid April 2007. It is not until the second week of July 2007 that this index begins to fall rapidly. In Figure 17 the normalised gradient signals a trend change at around 170 for the highest risk grade. This is equivalent to the third week of March 2007. At this stage the index was only just below 90 and appeared to be recovering.

9 Discussion

In terms of the non-stationary fractional diffusion model considered in this work, the time varying Fourier Dimension $q(\tau)$ can be interpreted in terms of a 'gauge' on the characteristics of dynamical system characterised by fractional diffusive processes. In a statistical sense, $q(\tau)$ is just another measure that may, or otherwise, be of value

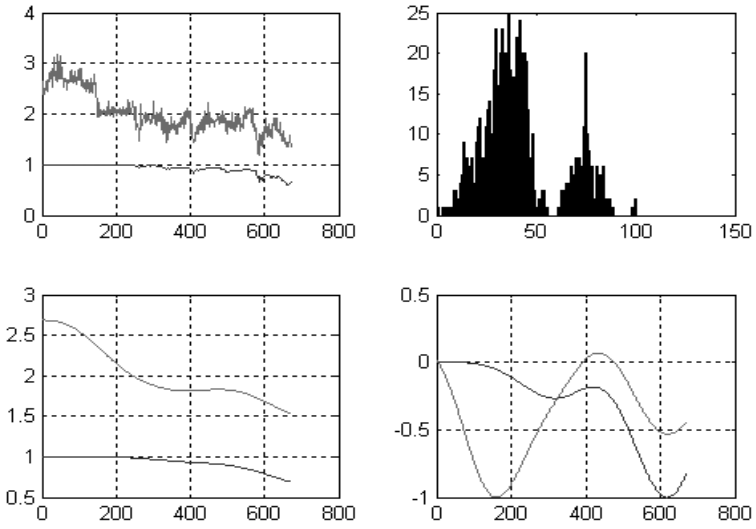

Figure 13: Analysis of AAA ABX.HE indices (2006 H1 vintage) by rating (closing prices) from 24-07-2006 to 02-04-2009. Top-left: AAA (normalised) data and $q(t)$; Top-right: 100 bin histogram; Bottom-left: Macotrends for $\beta = 0.1$; Bottom-right: Normalised gradients of macrotrends.

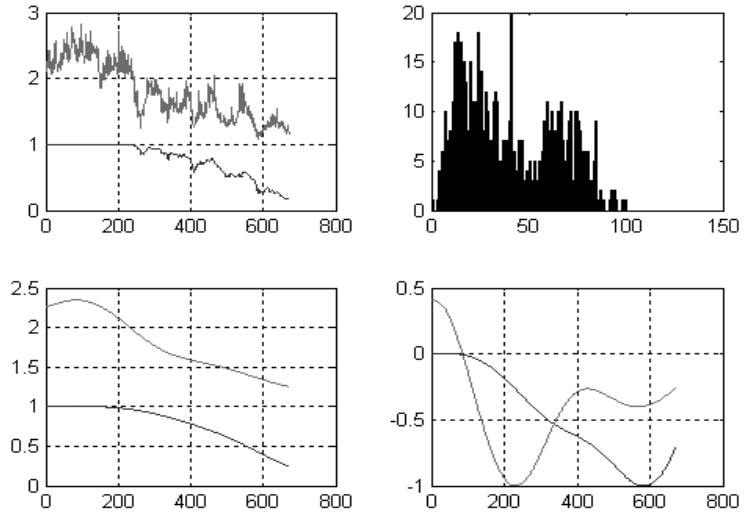

Figure 14: Analysis of AA ABX.HE indices (2006 H1 vintage) by rating (closing prices) from 24-07-2006 to 02-04-2009 for a 128 moving window. Top-left: AA data (normalised) and $q(t)$; Top-right: 100 bin histogram; Bottom-left: Macotrends for $\beta = 0.1$; Bottom-right: Normalised gradients of macrotrends.

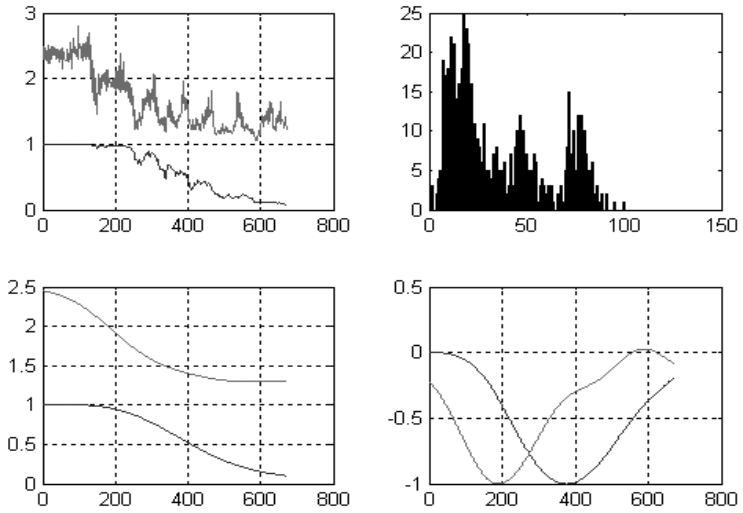

Figure 15: Analysis of A ABX.HE indices (2006 H1 vintage) by rating (closing prices) from 24-07-2006 to 02-04-2009 for a 128 size moving window. Top-left: AA data (normalised) and $q(t)$; Top-right: 100 bin histogram; Bottom-left: Macotrends for $\beta = 0.1$; Bottom-right: Normalised gradients of macrotrends.

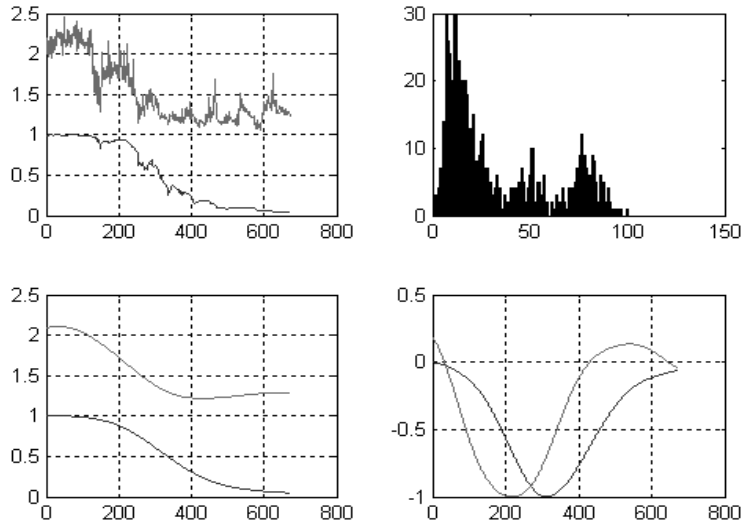

Figure 16: Analysis of BBB ABX.HE indices (2006 H1 vintage) by rating (closing prices) from 24-07-2006 to 02-04-2009 for a moving window with 128 elements. Top-left: AA data (normalised) and $q(t)$; Top-right: 100 bin histogram; Bottom-left: Macotrends for $\beta = 0.1$; Bottom-right: Normalised gradients of macrotrends.

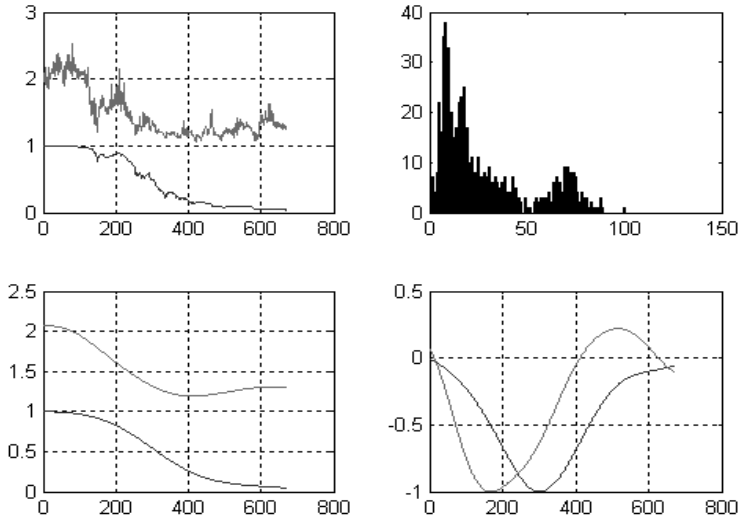

Figure 17: Analysis of BBB- ABX.HE indices (2006 H1 vintage) by rating (closing prices) from 24-07-2006 to 02-04-2009 for a moving window of size 128 element. Top-left:AA data (normalised) and $q(t)$; Top-right: 100 bin histogram; Bottom-left: Macotrends for $\beta = 0.1$; Bottom-right: Normalised gradients of macrotrends.

Statistical Parameter	AAA	AA	A	BBB	BBB-
Min.	1.1834	1.0752	1.0522	1.0610	1.0646
Max.	3.1637	2.8250	2.7941	2.4476	2.5371
Range	1.9803	1.7499	1.7420	1.3867	1.4726
Mean	2.0113	1.7869	1.6663	1.5141	1.4722
Median	1.9254	1.7001	1.4923	1.3425	1.3243
SD	0.3928	0.4244	0.4384	0.3746	0.3476
Variance	0.1543	0.1801	0.1922	0.1404	0.1208
Skew	0.7173	0.3397	0.6614	0.8359	1.0345
Kertosis	2.7117	1.8479	2.0809	2.2480	2.7467
CN	Reject	Reject	Reject	Reject	Reject

Table 6: Statistical values associated with $q(t)$ computed for ABX.HE indices (2006 H1 vintage) by rating (closing prices) from 24-07-2006 to 02-04-2009. Note that the acronyms SD and CN stand for 'Standard Deviation' and 'Composite Normality' repectively.

to market traders. In comparison with other statistical measures, this can only be assessed through its practical application in a live trading environment. However, in terms of its relationship to a stochastic model for financial time series data, $q(\tau)$ appears to provide a measure that is consistent with the physical principles associated with a random walk that includes a directional bias, i.e. fractional Brownian motion. The model considered, and the signal processing algorithm proposed, has a close association with re-scaled range analysis for computing the Hurst exponent H [12]. In this sense, the principal contribution of this work has been to consider an approach that is quantifiable in terms of a model that has been cast in terms of a specific fractional partial differential equation. As with other financial time series, their derivatives, transforms etc., a range of statistical measures can be used to characterise $q(\tau)$ and it is noted that in all cases studied to date, the composite normality of the signal $q(\tau)$ is of type 'Reject'. In other words, the statistics of $q(\tau)$ are non-Gaussian. Further, assuming that a financial time series is statistically self-affine, the computation of $q(\tau)$ can be applied over any time scale provided there is sufficient data for the computation of $q(\tau)$ to be statistically significant. Thus, the results associated with the Close-of-Day data studied in this work are, in principle, applicable to signals associated with data over a range of time scales.

As shown in Section 4.4 we obtain a simple algebraic relationship between Fourier Dimension q and the Lévy index γ given by $\gamma^{-1} = q/2$. Adhering to the numerical range of the Lévy index, we note that for $\gamma \in (0, 2]$ then $q \in (\infty, 1]$. However, q is related to other parameters including the Hurst exponent and the Fractal Dimension as shown in Appendix D. For a fractal signal where $D_F \in (1, 2)$, then $q \in (2, 1)$ and $\gamma \in (1, 2)$. Thus, the Fourier Dimension is simply related to the Fractal Dimension and Lévy index. The approach used is therefore consistent with assuming that financial signals are fractal signals and we may therefore consider equation (7) to be a classification of the 'Fractal Market Hypothesis' (FMH) based on Lévy statistics in contrast to the Efficient Market Hypothesis that is based on Gaussian statistics. In the context of the ideas presented in this work, the FMH has a number of fundamental differences with regard to the EMH which are tabulated in Table 7.

The material presented in this work has been exclusively concerned with a one-dimensional model for fractional diffusion. However, we note that for the three-dimensional case, equation (1) becomes

$$u(\mathbf{r}, t + \tau) = \int_{-\infty}^{\infty} u(\mathbf{r} + \lambda, t) p(\lambda) d\lambda$$

where $\mathbf{r}$ denotes the three-dimensional space vector $\mathbf{r} \equiv (x, y, z)$ and that in (three-dimensional) Fourier space

$$U(\mathbf{k}, t + \tau) = U(\mathbf{k}, t) P(k)$$

where $\mathbf{k}$ is the spatial frequency vector (k_x, k_y, k_z) and $k \equiv \mid \mathbf{k} \mid$. Thus, for the Characteristic Function

$$P(k) = \exp(-a \mid k \mid^{\gamma}) = 1 - a \mid k \mid^{\gamma}, \quad a \to 0,$$
$$\frac{U(\mathbf{k}, t+\tau) - U(\mathbf{k}, t)}{\tau} = -\frac{a}{\tau} U(\mathbf{k}, t) \mid k \mid^{\gamma}$$

and using the Reisz definition of a fractional n-dimensional laplacian given by

$$\nabla^{\gamma} \equiv -\frac{1}{(2\pi)^n} \int d^n \mathbf{k} \mid k \mid^{\gamma} \exp(i\mathbf{k} \cdot \mathbf{r})$$

we obtain (for $\tau \to 0$) the three dimensional homogeneous fractional diffusion equation

$$\left(\nabla^{\gamma} - \sigma \frac{\partial}{\partial t} \right) u(\mathbf{r}, t) = 0$$

However, the Green's function for the three-dimensional case is given by [24]

$$g(r, \omega) = \frac{\exp(i\Omega_{\gamma} r)}{4\pi r}$$

where $r \equiv \mid \mathbf{r} \mid$ which is not characterised by a $\omega^{-\frac{1}{\gamma}}$-type scaling law. In two-dimensions, the Green's function is [24] (ignoring scaling)

$$g(r, \omega) \sim \frac{\exp(i\Omega_{\gamma} r)}{\sqrt{\Omega_{\gamma} r}}$$

and it is clear that

$$\mid g(r, \omega) \mid \sim \frac{1}{\sqrt{r(\omega\sigma)^{\frac{1}{\gamma}}}}, \quad \sigma \to 0$$

which is characterised by a $\omega^{-\frac{1}{2\gamma}}$ scaling law. The inability for the three-dimensional fractional diffusion equation to yield $\omega^{-\frac{1}{\gamma}}$-type fractal noise is overcome if we consider a model based on the separable case [32]. We consider the equation

$$\left(\frac{\partial^{\gamma}}{\partial x^{\gamma}} + \frac{\partial^{\gamma}}{\partial y^{\gamma}} + \frac{\partial^{\gamma}}{\partial z^{\gamma}} - \sigma \frac{\partial}{\partial t} \right) u(x, y, z, t) = n(x, y, z, t)$$

and a solution where

$$n(x, y, z, t) = \delta(x) n_x(t) + \delta(y) n_y(t) + \delta(z) n_z(t)$$

and

$$u(x, y, z, t) = u_x(x, t) + u_y(y, t) + u_z(x, t)$$

The source function n is then taken to model a system characterised by a separable spatial impulse with separable white noise function (n_x, n_y, n_z). This model is used for the morphological analysis of Hyphal growth rates, for example, where the fractal dimension provides an estimate of the metabolic production of filamentous fungi [33].

10 Conclusion

The value of $q(\tau)$ characterises stochastic processes that can vary in time and are intermediate between fully diffusive and propagative or persistent behaviour. Application of Orthogonal Linear Regression to financial time series data provides an accurate and robust method to compute $q(\tau)$ when compared to other statistical estimation techniques such as the least squares method. As a result of the physical interpretation associated with the fractional diffusion equation discussed in Section 5.4, we can, in principle, use the signal $q(\tau)$ as a predictive measure in the sense that as the value of $q(\tau)$ continues to increases, there is a greater likelihood for propagative behaviour of the markets. This is reflected in the data analysis based on the examples given in which the Gaussian VDSK $\exp(-\alpha\omega^2)$ has been used to smooth both $u(t)$ and $q(\tau)$ to obtain macrotrends in which the value of α determines the level of detail in the output. From the examples provided in Figure 9 and other trials that have been undertaken (details of which lie beyond the scope of this text), it is clear that the 'turning point' (i.e. positions in time where $q'(t) = 0$) appear to 'flag' a future change in the trend of the signal $u(t)$. This feature is reflected in the 'cross-over points' of the normalised gradients for u' and $q'(t)$ illustrated in Figure 9, i.e. points in time when $u'(t), \quad \|u'\|_\infty = 1$ and $q'(\tau), \quad \|u'\|_\infty = 1$ are approximately the same and whose gradients are of opposite polarity. These characteristics are a consequence of the phase shifts that exist in the gradients of $u(t)$ and $q(\tau)$ over different frequency bands. Although the interpretation of these phase shifts requires further study, from the results obtained to date, it is clear that they provide an assessment of the risk associated with a particular investment portfolio and has been used to develop a model for forecasting volatility as discussed in Section 7. However, application of this model for predicting market behavour is dependent on optimising two parameters, namely, the length of the moving window L and the filter parameter α. Evaluation of historical data for a given time series (with a specific sampling rate) is required in order to optimise the values of these parameters.

The model used is predicated on the assumption that financial time series are Lévy distributed and one of the key results of this work has been to provide connectivity between Lévy distributed processes used to derive the fractional diffusion equation and random scaling fractal signals in terms of a Green's function solution to this equation. The value of the Lévy index $\gamma = 2/q$ is then used to gauge the likely future behavior of the signal under the FMH: a change in γ precedes a change in the signal.

EMH	FMH
Gaussian Statistics	Non-Gaussian Statistics
Stationary Process	Non-stationary Process
No memory - no historical correlations	Memory - historical correlations
No repeating patterns at any scale	Many repeating patterns at all scales - 'Elliot waves'
Continuously stable at all scales	Continuously unstable at any scale - 'Lévy Flights'

Table 7: Principal differences between the Efficient Market Hypothesis (EMH) and the Fractal Market Hypothesis (FMH).

11 Appendices

Appendix A: Einstein's Derivation of the Diffusion Equation

Let τ be a small interval of time in which a particle moves between λ and $\lambda + d\lambda$ with probability $P(\lambda)$ where $\lambda = \sqrt{\lambda_x^2 + \lambda_y^2 + \lambda_z^2}$ and τ is small enough to assume that the movements of the particle are τ-independent. If $u(\mathbf{r}, t)$ is the concentration (i.e. the number of particles per unit volume) then the concentration at time $t + \tau$ generated by a source function $F(\mathbf{r})$ is given by

$$(\mathbf{r}, t + \tau) = \int_{-\infty}^{\infty} u(\mathbf{r} + \lambda, t) P(\lambda) d\lambda + F(\mathbf{r}, t) \qquad (A.1)$$

Since $\tau << 1$, we may approximate $u(\mathbf{r}, t + \tau)$ as

$$u(\mathbf{r}, t + \tau) = u(\mathbf{r}, t) + \tau \frac{\partial}{\partial t} u(\mathbf{r}, t)$$

and write $u(\mathbf{r} + \lambda, t)$ in terms of the Taylor series

$$\begin{aligned} u(\mathbf{r} + \lambda, t) =& u + \lambda_x \frac{\partial u}{\partial x} + \lambda_y \frac{\partial u}{\partial y} + \lambda_z \frac{\partial u}{\partial z} \\ &+ \frac{\lambda_x^2}{2!} \frac{\partial^2 u}{\partial x^2} + \frac{\lambda_y^2}{2!} \frac{\partial^2 u}{\partial y^2} + \frac{\lambda_z^2}{2!} \frac{\partial^2 u}{\partial z^2} \\ &+ \lambda_x \lambda_y \frac{\partial^2 u}{\partial x \partial y} + \lambda_x \lambda_z \frac{\partial^2 u}{\partial x \partial z} + \lambda_y \lambda_z \frac{\partial^2 u}{\partial y \partial z} + ... \end{aligned}$$

However, higher order terms can be neglected since, if $\tau << 1$, then the distance travelled, λ, must also be small. Equation (A.1) may then be written as

$$\begin{aligned} u + \tau \frac{\partial u}{\partial t} = F(\mathbf{r}, t) &+ \int_{-\infty}^{\infty} u P(\lambda) d\lambda \\ &+ \int_{-\infty}^{\infty} \left(\lambda_x \frac{\partial u}{\partial x} + \lambda_y \frac{\partial u}{\partial y} + \lambda_z \frac{\partial u}{\partial z} \right) P(\lambda) d\lambda \\ &+ \int_{-\infty}^{\infty} \left(\frac{\lambda_x^2}{2!} \frac{\partial^2 u}{\partial x^2} + \frac{\lambda_y^2}{2!} \frac{\partial^2 u}{\partial y^2} + \frac{\lambda_z^2}{2!} \frac{\partial^2 u}{\partial z^2} \right) P(\lambda) d\lambda \\ &+ \int_{-\infty}^{\infty} \left(\lambda_x \lambda_y \frac{\partial^2 u}{\partial x \partial y} + \lambda_x \lambda_z \frac{\partial^2 u}{\partial x \partial z} + \lambda_y \lambda_z \frac{\partial^2 u}{\partial y \partial z} \right) P(\lambda) d\lambda \end{aligned}$$

Assuming that $P(\lambda)$ is normalized we have

$$\int_{-\infty}^{\infty} P(\lambda)d\lambda = 1$$

so that

$$\begin{aligned}
\tau\frac{\partial}{\partial t}u = &\int_{-\infty}^{\infty}\frac{\lambda_x^2}{2}\frac{\partial^2 u}{\partial x^2}P(\lambda)d\lambda + \int_{-\infty}^{\infty}\frac{\lambda_x\lambda_y}{2}\frac{\partial^2 u}{\partial x\partial y}P(\lambda)d\lambda \\
&+ \int_{-\infty}^{\infty}\frac{\lambda_x\lambda_z}{2}\frac{\partial^2 u}{\partial x\partial z}P(\lambda)d\lambda + \int_{-\infty}^{\infty}\frac{\lambda_y\lambda_x}{2}\frac{\partial^2 u}{\partial y\partial x}P(\lambda)d\lambda \\
&+ \int_{-\infty}^{\infty}\frac{\lambda_y^2}{2}\frac{\partial^2 u}{\partial y^2}P(\lambda)d\lambda + \int_{-\infty}^{\infty}\frac{\lambda_y\lambda_z}{2}\frac{\partial^2 u}{\partial y\partial z}P(\lambda)d\lambda \\
&+ \int_{-\infty}^{\infty}\frac{\lambda_z\lambda_x}{2}\frac{\partial^2 u}{\partial z\partial x}P(\lambda)d\lambda + \int_{-\infty}^{\infty}\frac{\lambda_z\lambda_y}{2}\frac{\partial^2 u}{\partial z\partial y}P(\lambda)d\lambda \\
&+ \int_{-\infty}^{\infty}\frac{\lambda_z^2}{2}\frac{\partial^2 u}{\partial z^2}P(\lambda)d\lambda + \int_{-\infty}^{\infty}\lambda_x\frac{\partial u}{\partial x}P(\lambda)d\lambda \\
&+ \int_{-\infty}^{\infty}\lambda_y\frac{\partial u}{\partial y}P(\lambda)d\lambda + \int_{-\infty}^{\infty}\lambda_z\frac{\partial u}{\partial z}P(\lambda)d\lambda + F(\mathbf{r},t)
\end{aligned}$$

which may be written as a matrix equation of the following form

$$\frac{\partial}{\partial t}u(\mathbf{r},t) = \nabla\cdot\mathbf{D}\nabla u(\mathbf{r},t) + \mathbf{V}\cdot\nabla u(\mathbf{r},t) + F(\mathbf{r},t)$$

where $\mathbf{D}$ is the diffusion tensor given by

$$\mathbf{D} = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{pmatrix}$$

where

$$D_{ij} = \int_{-\infty}^{\infty}\frac{\lambda_i\lambda_j}{2\tau}P(\lambda)d\lambda = \frac{1}{2\tau}\langle\lambda_i\lambda_j\rangle$$

and $\mathbf{V}$ is a flow vector which describes any drift velocity that the particle ensamble may have and is given by

$$\mathbf{V} = \begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix} \qquad V_i = \int_{-\infty}^{\infty}\frac{\lambda_i}{\tau}P(\lambda)d\lambda = \frac{1}{\tau}\langle\lambda_i\rangle$$

Note that as $\lambda_i\lambda_j = \lambda_j\lambda_i$, the diffusion tensor is diagonally symmetric (i.e. $D_{ij} = D_{ji}$). For isotropic diffusion where $\langle\lambda_i\lambda_j\rangle = 0$ for $i \neq j$ and $\langle\lambda_i\lambda_j\rangle = \langle\lambda^2\rangle$ for $i = j$ and with no drift velocity so that $\mathbf{V} = \mathbf{0}$, then

$$\frac{\partial}{\partial t}u(\mathbf{r},t) = \nabla \cdot \begin{pmatrix} D & 0 & 0 \\ 0 & D & 0 \\ 0 & 0 & D \end{pmatrix} \nabla u(\mathbf{r},t) + F(\mathbf{r},t)$$

$$= D\nabla^2 u(\mathbf{r},t) + F(\mathbf{r},t)$$

where

$$D = \int_{-\infty}^{\infty} \frac{\lambda^2}{2\tau} P(\lambda) d\lambda$$

Appendix B: Evaluation of the Lévy Distribution

We wish to show that the Characteristic Function $P(k) = \exp(-a \mid k \mid^{\gamma}), \quad 0 < \gamma \leq 2$ is equivalent to a Probability Density Function given by $p(x) \sim x^{-(1+\gamma)}, \quad x \to \infty.$, i.e. we wish to prove the following:

Theorem

$$\frac{1}{x^{1+\gamma}} \leftrightarrow \exp(-a \mid k \mid^{\gamma}), \quad 0 < \gamma \leq 2, \quad x \to \infty$$

where $\leftrightarrow$ denotes transformation from real to Fourier space[2].

Proof of Theorem
For $0 < \gamma < 1$, and since the characteristic function is symmetric, we have

$$p(x) = \mathrm{Re}[f(x)]$$

where

$$f(x) = \frac{1}{\pi} \int_0^{\infty} e^{ikx} e^{-k^{\gamma}} dk$$

$$= \frac{1}{\pi} \left(\left[\frac{1}{ix} e^{ikx} e^{-k^{\gamma}} \right]_{k=0}^{\infty} - \frac{1}{ix} \int_0^{\infty} e^{ikx} (-\gamma k^{\gamma-1} e^{-k^{\gamma}}) dk \right)$$

$$= \frac{\gamma}{2\pi i x} \int_{-\infty}^{\infty} dk H(k) k^{\gamma-1} e^{-k^{\gamma}} e^{ikx}, \quad x \to \infty$$

where

$$H(k) = \begin{cases} 1, & k > 0 \\ 0, & k < 0 \end{cases}$$

For $0 < \gamma < 1$, $f(x)$ is singular at $k = 0$ and the greatest contribution to this integral is the inverse Fourier transform of $H(k)k^{\gamma-1}$. Noting that

$$\mathcal{F}^{-1} \left[\frac{1}{(ik)^{\gamma}} \right] \sim \frac{1}{x^{1-\gamma}}$$

where $\mathcal{F}^{-1}$ denotes the inverse Fourier transform, and that

$$H(k) \leftrightarrow \delta(x) + \frac{i}{\pi x} \sim \delta(x), \quad x \to \infty$$

[2]The author acknowledges Dr K I Hopcraft, School of Mathematical Sciences, Nottingham University, England, for his advice in respect of this result.

then, using the convolution theorem, we have

$$f(x) \sim \frac{\gamma}{i\pi x} \frac{i^{1-\gamma}}{x^{\gamma}}$$

and thus

$$p(x) \sim \frac{1}{x^{1+\gamma}}, \quad x \to \infty$$

For $1 < \gamma < 2$, we can integrate by parts twice to obtain

$$f(x) = \frac{\gamma}{i\pi x} \int_0^{\infty} dk k^{\gamma-1} e^{-k^{\gamma}} e^{ikx}$$

$$= \frac{\gamma}{i\pi x} \left[\frac{1}{ix} k^{\gamma-1} e^{-k^{\gamma}} e^{ikx} \right]_{k=0}^{\infty}$$

$$+ \frac{\gamma}{\pi x^2} \int_0^{\infty} dk e^{ikx} [(\gamma-1)k^{\gamma-2} e^{-k^{\gamma}} - \gamma (k^{\gamma-1})^2 e^{-k^{\gamma}}]$$

$$= \frac{\gamma}{\pi x^2} \int_0^{\infty} dk e^{ikx} [(\gamma-1)k^{\gamma-2} e^{-k^{\gamma}} - \gamma (k^{\gamma-1})^2 e^{-k^{\gamma}}], \quad x \to \infty$$

The first term of this result is singular and therefore provides the greatest contribution and thus we can write,

$$f(x) \simeq \frac{\gamma(\gamma-1)}{2\pi x^2} \int_{-\infty}^{\infty} H(k) e^{ikx} (k^{\gamma-2} e^{-k^{\gamma}}) dk$$

In this case, for $1 < \gamma < 2$, the greatest contribution to this integral is the inverse Fourier transform of $k^{\gamma-2}$ and hence,

$$f(x) \sim \frac{\gamma(\gamma-1)}{\pi x^2} \frac{i^{2-\gamma}}{x^{\gamma-1}}$$

so that

$$p(x) \sim \frac{1}{x^{1+\gamma}}, \quad x \to \infty$$

which maps onto the previous asymptotic as $\gamma \to 1$ from the above.

Appendix C: A Short Overview of Fractional Calculus

In a famous letter from l'Hospital to Leibnitz written in 1695, l'Hospital asked the following question: 'Given that $d^n f/dt^n$ exists for all integer n, what if n be $\frac{1}{2}$'. The reply from Leibnitz was all the more interesting: 'It will lead to a paradox ... From this paradox, one day useful consequences will be drawn'.

Fractional calculus (e.g. [34], [35], [36] and [37]) has been studied for many years by some of the great names of mathematics since the development of (integer) calculus in the late seventeenth century. Relatively few papers and books exist on such a naturally important subject. However, a study of the works in this area of mathematics clearly show that the ideas used to define a fractional differential and a fractional integral are based on definitions which are in effect, little more than generalizations of results obtained using integer calculus. The classical fractional integral operators are the Riemann-Liouville transform [34]

$$\hat{I}^q f(t) = \frac{1}{\Gamma(q)} \int_{-\infty}^{t} \frac{f(\tau)}{(t-\tau)^{1-q}} d\tau, \quad q > 0$$

and the Weyl transform

$$\hat{I}^q f(t) = \frac{1}{\Gamma(q)} \int_{t}^{\infty} \frac{f(\tau)}{(t-\tau)^{1-q}} d\tau, \quad q > 0$$

where

$$\Gamma(q) = \int_{0}^{\infty} t^{q-1} \exp(-t) dt.$$

For integer values of q (i.e. when $q = n$ where n is a non-negative integer), the Riemann-Liouville transform reduces to the standard Riemann integral. This transform is just a (causal) convolution of the function $f(t)$ with $t^{q-1}/\Gamma(q)$. For fractional differentiation, we can perform a fractional integration of appropriate order and then differentiate to an appropriate integer order. The reason for this is that direct fractional differentiation can lead to divergent integrals. Thus, the fractional differential operator $\hat{D}^q$ for $q > 0$ is given by

$$\hat{D}^q f(t) \equiv \frac{d^q}{dt^q} f(t) = \frac{d^n}{dt^n} [\hat{I}^{n-q} f(t)].$$

Another (conventional) approach to defining a fractional differential operator is based on using the formula for n^{th} order differentiation obtained by considering the definitions for the first, second, third etc. differentials using backward and then

generalising the formula by replacing n with q. This approach provides us with the result [34]

$$\hat{D}^q f(t) = \lim_{N\to\infty}\left[\frac{(t/N)^{-q}}{\Gamma(-q)}\sum_{j=0}^{N-1}\frac{\Gamma(j-q)}{\Gamma(j+1)}f\left(t-j\frac{t}{N}\right)\right].$$

A review of this result shows that for $q = 1$, this is a point process but for other values it is not, i.e. the evaluation of a fractional differential operator depends on the history of the function in question. Thus, unlike an integer differential operator, a fractional differential operator has 'memory'. Although the memory of this process fades, it does not do so quickly enough to allow truncation of the series in order to retain acceptable accuracy. The concept of memory association can also be seen from the result

$$\hat{D}^q f(t) = \frac{d^n}{dt^n}[\hat{I}^{n-q} f(t)]$$

where

$$\hat{I}^{q-n} f(t) = \frac{1}{\Gamma(n-q)}\int_{-\infty}^{t}\frac{f(\tau)}{(t-\tau)^{1+q-n}}d\tau, \quad n-q>0$$

in which the value of $\hat{I}^{q-n} f(t)$ at a point t depends on the behaviour of $f(t)$ from $-\infty$ to t via a convolution with the kernel $t^{n-q}/\Gamma(q)$. The convolution process is of course dependent on the history of the function $f(t)$ for a given kernel and thus, in this context, we can consider a fractional derivative defined via the result above to have memory.

C.1 The Laplace Transform and the Half Integrator

It informative at this point to consider the application of the Laplace transform to identify an ideal integrator and then a half integrator. The Laplace transform is given by

$$\hat{L}[f(t)] \equiv F(p) = \int_0^\infty f(t)\exp(-pt)dt$$

and from this result we can derive the transform of a derivative given by

$$\hat{L}[f'(t)] = pF(p) - f(0)$$

and the transform of an integral given by

$$\hat{L}\left[\int_0^t f(\tau)d\tau\right] = \frac{1}{p}F(p).$$

Now, suppose we have a standard time invariant linear system whose input is $f(t)$ and whose output is given by

$$s(t) = f(t) \otimes g(t)$$

where the convolution is causal, i.e.

$$s(t) = \int_0^t f(\tau) g(t-\tau) d\tau.$$

Suppose we let

$$g(t) = H(t) = \begin{cases} 1, & t > 0; \\ 0, & t < 0. \end{cases}$$

Then, $G(p) = 1/p$ and the system becomes an ideal integrator:

$$s(t) = f(t) \otimes H(t) = \int_0^t f(t-\tau) d\tau = \int_0^t f(\tau) d\tau.$$

Now, consider the case when we have a time invariant linear system with an impulse response function by given by

$$g(t) = \frac{H(t)}{\sqrt{t}} = \begin{cases} \mid t \mid^{-1/2}, & t > 0; \\ 0, & t < 0. \end{cases}$$

The output of this system is $f \otimes g$ and the output of such a system with input $f \otimes g$ is $f \otimes g \otimes g$. Now

$$g(t) \otimes g(t) = \int_0^t \frac{d\tau}{\sqrt{\tau}\sqrt{t-\tau}} = \int_0^{\sqrt{t}} \frac{2x dx}{x\sqrt{t-x^2}}$$

$$= 2 \left[\sin^{-1} \left(\frac{x}{\sqrt{t}} \right) \right]_0^{\sqrt{t}} = \pi.$$

Hence,

$$\frac{H(t)}{\sqrt{\pi t}} \otimes \frac{H(t)}{\sqrt{\pi t}} = H(t)$$

and the system defined by the impulse response function $H(t)/\sqrt{\pi t}$ represents a 'half-integrator' with a Laplace transform given by

$$\hat{L} \left[\frac{H(t)}{\sqrt{\pi t}} \right] = \frac{1}{\sqrt{p}}.$$

This result provides an approach to working with fractional integrators and/or differentiators using the Laplace transform. Fractional differential and integral operators can be defined and used in a similar manner to those associated with conventional or integer order calculus and we now provide an overview of such operators.

C.2 Operators of Integer Order

The following operators are all well-defined, at least with respect to all test functions $u(t)$ say which are (i) infinitely differentiable and (ii) of compact support (i.e. vanish outside some finite interval).

Integral Operator:

$$\hat{I}u(t) \equiv \hat{I}^1 u(t) = \int_{-\infty}^{t} u(\tau)d\tau.$$

Differential Operator:

$$\hat{D}u(t) \equiv \hat{D}^1 u(t) = u'(t).$$

Identify Operator:

$$\hat{I}^0 u(t) = u(t) = \hat{D}^0 u(t).$$

Now,

$$\hat{I}[\hat{D}u](t) = \int_{-\infty}^{t} u'(\tau)d\tau = u(t)$$

and

$$\hat{D}[\hat{I}u](t) = \frac{d}{dt}\int_{-\infty}^{t} u(\tau)d\tau = u(t)$$

so that

$$\hat{I}^1\hat{D}^1 = \hat{D}^1\hat{I}^1 = \hat{I}^0.$$

For n (integer) order:

$$\hat{I}^n u(t) = \int_{-\infty}^{t} d\tau_{n-1} ... \int_{-\infty}^{\tau_2} d\tau_1 \int_{-\infty}^{\tau_1} u(\tau)d\tau,$$

$$\hat{D}^n u(t) = u^{(n)}(t)$$

and

$$\hat{I}^n[\hat{D}^n u](t) = u(t) = \hat{D}^n[\hat{I}^n u](t).$$

C.3 Convolution Representation

Consider the function

$$t_+^{q-1}(t) \equiv \mid t \mid^{q-1} H(t) = \begin{cases} \mid t \mid^{q-1}, & t > 0; \\ 0, & t < 0. \end{cases}$$

which, for any $q > 0$ defines a function that is locally integrable. We can then define an integral of order n in terms of a convolution as

$$\hat{I}^n u(t) = \left(u \otimes \frac{1}{(n-1)!} t_+^{n-1} \right)(t)$$

$$= \frac{1}{(n-1)!} \int_{-\infty}^{t} (t-\tau)^{n-1} u(\tau) d\tau$$

$$= \frac{1}{(n-1)!} \int_{-\infty}^{t} \tau^{n-1} u(t-\tau) d\tau$$

In particular,

$$\hat{I}^1 u(t) = (u \otimes H)(t) = \int_{-\infty}^{t} u(\tau) d\tau.$$

These are classical (absolutely convergent) integrals and the identity operator admits a formal convolution representation, using the delta function, i.e.

$$\hat{I}^0 u(t) = \int_{-\infty}^{\infty} \delta(\tau) u(t-\tau) d\tau$$

where

$$\delta(t) = \hat{D} H(t).$$

Similarly,

$$\hat{D}^n u(t) \equiv \hat{I}^{-n} u(t) = \int_{-\infty}^{\infty} \delta^{(n)}(\tau) u(t-\tau) d\tau = u^{(n)}(t).$$

On the basis of the material discussed above, we can now formally extend the integral operator to fractional order and consider the operator

$$\hat{I}^q u(t) = \frac{1}{\Gamma(q)} \int_{-\infty}^{\infty} u(\tau) t_+^{q-1}(t-\tau) d\tau$$

$$= \frac{1}{\Gamma(q)} \int_{-\infty}^{t} u(\tau) t_+^{q-1} (t-\tau) d\tau$$

where

$$\Gamma(q) = \int_0^{\infty} t^{q-1} \exp(-t) dt, \quad q > 0$$

with the fundamental property that

$$\Gamma(q+1) = q\Gamma(q).$$

Here, I^q is an operator representing a time invariant linear system with impulse response function $t_+^{q-1}(t)$ and transfer function $1/p^q$. For the cascade connection of I^{q_1} and I^{q_2} we have

$$\hat{I}^{q_1}[\hat{I}^{q_2} u(t)] = \hat{I}^{q_1+q_2} u(t).$$

This classical convolution integral representation holds for all real $q > 0$ (and formally for $q = 0$, with the delta function playing the role of an impulse function and with a transfer function equal to the constant 1).

C.4 Fractional Differentiation

For $0 < q < 1$, if we define the (Riemann-Liouville) derivative of order q as

$$\hat{D}^q u(t) \equiv \frac{d}{dt}[\hat{I}^{1-q} u](t) = \frac{1}{\Gamma(1-q)} \frac{d}{dt} \int_{-\infty}^{t} (t-\tau)^{-q} u(\tau) d\tau,$$

then,

$$\hat{D}^q u(t) = \frac{1}{\Gamma(1-q)} \int_{-\infty}^{t} (t-\tau)^{-q} u'(\tau) d\tau \equiv \hat{I}^{1-q} u'(t).$$

Hence,

$$\hat{I}^q[\hat{D}^q u] = \hat{I}^q[\hat{I}^{1-q} u'] = \hat{I}^1 u' = u$$

and $\hat{D}^q$ is the formal inverse of the operator $\hat{I}^q$. Given any $q > 0$, we can always write $\lambda = n - 1 + q$ and then define

$$\hat{D}^\lambda u(t) = \frac{1}{\Gamma(1-q)} \frac{d^n}{dt^n} \int_{-\infty}^{t} u(\tau)(t-\tau)^{-q} d\tau.$$

D^q is an operator representing a time invariant linear system consisting of a cascade combination of an ideal differentiator and a fractional integrator of order $1-q$. For D^λ we replace the single ideal differentiator by n such that

$$\hat{D}^0 u(t) = \frac{1}{\Gamma(1)} \frac{d}{dt} \int_{-\infty}^{t} u(\tau) d\tau = u(t) \equiv \int_{-\infty}^{\infty} u(\tau)\delta(t-\tau)d\tau$$

and

$$\hat{D}^n u(t) = \frac{1}{\Gamma(1)} \frac{d^{n+1}}{dt^{n+1}} \int_{-\infty}^{t} u(\tau) d\tau$$

$$= u^{(n)}(t) \equiv \int_{-\infty}^{\infty} u(\tau)\delta^{(n)}(t-\tau)d\tau.$$

In addition to the conventional and classical definitions of fractional derivatives and integrals, more general definitions are available including the Erdélyi-Kober fractional integral [38]

$$\frac{t^{-p-q+1}}{\Gamma(q)} \int_0^t \frac{\tau^{p-1}}{(t-\tau)^{1-q}} f(\tau)d\tau, \quad q > 0, \quad p > 0$$

which is a generalisation of the Riemann-Liouville fractional integral and the integral

$$\frac{t^p}{\Gamma(q)} \int_t^\infty \frac{\tau^{-q-p}}{(\tau-t)^{1-q}} f(\tau)d\tau, \quad q > 0, \quad p > 0$$

which is a generalization of the Weyl integral. Further definitions exist based on the application of hypergeometric functions and operators involving other special functions such as the Maijer G-function and the Fox H-function [39]. Moreover, all such operators leading to a fractional integral of the Riemann-Liouville type and the Weyl type to have the general forms (through induction)

$$\hat{I}^q f(t) = t^{q-1} \int_{-\infty}^{t} \Phi\left(\frac{\tau}{t}\right) \tau^{-q} f(\tau)d\tau$$

and

$$\hat{I}^q f(t) = t^{-q} \int_{t}^{\infty} \Phi\left(\frac{t}{\tau}\right) \tau^{q-1} f(\tau)d\tau$$

respectively, where the kernel Φ is an arbitrary continuous function so that the integrals above make sense in sufficiently large functional spaces. Although there are a number of approaches that can be used to define a fractional differential/integral, there is one particular definition, which in terms of its 'ease of use' and wide ranging applications, is of significant value and is based on the Fourier transform, i.e.

$$\frac{d^q}{dt^q}f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}(i\omega)^q F(\omega)\exp(i\omega t)d\omega$$

where $F(\omega)$ is the Fourier transform of $f(t)$. When $q = 1, 2, 3...$, this definition reduces to a well known result that is trivial to derive in which, for example, the 'filter' $i\omega$ (for the case when $q = 1$) is referred to as a 'differentiator'. When $q < 0$, we have a definition for the fractional integral where, in the case of $q = -1$, for example, the filter $(i\omega)^{-1}$ is an 'integrator'. When $q = 0$ we just have $f(t)$ expressed in terms of its Fourier transform $F(\omega)$. This Fourier based definition of a fractional derivative can be extended further to include a definition for a 'fractional Laplacian' ∇^q where for n dimensions

$$\nabla^q \equiv -\frac{1}{(2\pi)^n}\int d^n\mathbf{k}k^q\exp(i\mathbf{k}\cdot\mathbf{r}), \quad k = \mid \mathbf{k} \mid$$

and $\mathbf{r}$ is an n-dimensional vector. This is the fractional Riesz operator. It is designed to provide a result that is compatible with the case of $q = 2$ for $n > 1$, i.e. $\nabla^2 \Longleftrightarrow -k^2$ (which is the reason for introducing the negative sign). Another equally valid generalization is

$$\nabla^q \equiv \frac{1}{(2\pi)^n}\int d^n\mathbf{k}(ik)^q\exp(i\mathbf{k}\cdot\mathbf{r}), \quad k = \mid \mathbf{k} \mid$$

which introduces a q dependent phase factor of $\pi q/2$ into the operator.

C.5 Fractional Dynamics

Mathematical modelling using (time dependent) fractional Partial Differential Equations (PDEs) is generally known as fractional dynamics [40], [41]. A number of works have shown a close relationship between fractional diffusion equations of the type (where p is the space-time dependent PDF and σ is the generalized coefficient of diffusion)

$$\nabla^2 p - \sigma\frac{\partial^q}{\partial t^q}p = 0, \quad 0 < q \leq 1$$

and

$$\nabla^q p - \sigma\frac{\partial}{\partial t}p = 0, \quad 0 < q \leq 2$$

and continuous time random walks with either temporal or spatial scale invariance (fractal walks). Fractional diffusion equations of this type have been shown to produce a framework for the description of anomalous diffusion phenomena and Lévy-type behaviour. In addition, certain classes of fractional differential equations are known to yield Lévy-type distributions. For example, the normalized one-sided Lévy-type PDF

$$p(x) = \frac{a^q}{\Gamma(q)} \frac{\exp(-a/x)}{x^{1+q}}, \quad a > 0, \quad x > 0$$

is a solution of the fractional integral equation

$$x^{2q} p(x) = a^q \hat{I}^{-q} p(x)$$

where

$$\hat{I}^{-q} p(x) = \frac{1}{\Gamma(q)} \int_0^x \frac{p(y)}{(x-y)^{1-q}} dy, \quad q > 0.$$

Another example involves the solution to the anomalous diffusion equation

$$\nabla^q p - \tau \frac{\partial}{\partial t} p = 0, \quad 0 < q \leq 2.$$

Fourier transforming this equation and using the fractional Riesz operator defined previously, we have

$$\frac{\partial}{\partial t} P(k,t) = -\frac{1}{\tau} k^q P(k,t)$$

which has the general solution

$$P(k,t) = \exp(-t \mid k \mid^q / \tau), \quad t > 0.$$

which is the characteristic function of a Lévy distribution. This analysis can be extended further by considering a fractal based generalization of the Fokker-Planck-Kolmogorov (FPK) equation [42]

$$\frac{\partial^q}{\partial t^q} p(x,t) = \frac{\partial^\beta}{\partial x^\beta} [s(x) p(x,t)]$$

where s is an arbitrary function and $0 < q \leq 1$, $0 < \beta \leq 2$. This equation is referred to as the fractal FPK equation; the standard FPK equation is of course recovered for $q = 1$ and $\beta = 2$. The characteristic function associated with $p(x,t)$ is given by

$$P(k,t) = \exp(-a k^\beta t^q)$$

where a is a constant which again, is a characteristic of a Lévy distribution. Finally, d-dimensional fractional master equations of the type [43], [44]

$$\frac{\partial^q}{\partial t^q} p(\mathbf{r}, t) = \sum_{\mathbf{s}} w(\mathbf{r} - \mathbf{s}) p(\mathbf{s}, t), \quad 0 < q \leq 1$$

can be used to model non-equilibrium phase transitions where p denotes the probability of finding the diffusing entity at a position $\mathbf{r} \in R^d$ at time t (assuming that it was at the origin $\mathbf{r} = \mathbf{0}$ at time $t = 0$) and w are the fractional transition rates which measure the propensity for a displacement $\mathbf{r}$ in units of $1/(\text{time})^q$. These equations conform to the general theory of continuous time random walks and provide models for random walks of fractal time.

Appendix D: Dimensional Relationships

In this appendix we consider the relationships between the Hurst Exponent, the Topological Dimension and Fractal and Fourier Dimensions.

Suppose we cut up some simple one-, two- and three-dimensional Euclidean objects (a line, a square surface and a cube, for example), make exact copies of them and then keep on repeating the copying process. Let N be the number of copies that we make at each stage and let r be the length of each of the copies, i.e. the scaling ratio. Then we have

$$Nr^{D_T} = 1, \quad D_T = 1, 2, 3, ...$$

where D_T is the topological dimension. The similarity or fractal dimension is that value of D_F which is usually (but not always) a non-integer dimension 'greater' that its topological dimension (i.e. 0,1,2,3,... where 0 is the dimension of a point on a line) and is given by

$$D_F = -\frac{\log(N)}{\log(r)}$$

The fractal dimension is that value that is strictly greater than the topological dimension. In each case, as the value of the fractal dimension increases, the fractal becomes increasingly 'space-filling' in terms of the topological dimension which the fractal dimension is approaching. A fractal exhibits structures that are self-similar. A self-similar deterministic fractal is one where a change in the scale of a function $f(x)$ (which may be a multi-dimensional function) by a scaling factor λ produces a smaller version, reduced in size by λ, i.e.

$$f(\lambda x) = \lambda f(x)$$

A self-affine deterministic fractal is one where a change in the scale of a function $f(x)$ by a factor λ produces a smaller version reduced in size by a factor λ^q, $q > 0$, i.e.

$$f(\lambda x) = \lambda^q f(x)$$

For stochastic fields, the expression

$$\Pr[f(\lambda x)] = \lambda^q \Pr[f(x)]$$

describes a statistically self-affine field - a random scaling fractal. As we zoom into the fractal, the shape changes, but the distribution of lengths remains the same.

There is no unique method for computing the fractal dimension. The methods available are broadly categorized into two families: (i) Size-measure relationships, based on recursive length or area measurements of a curve or surface using different measuring scales; (ii) application of relationships based on approximating or fitting

a curve or surface to a known fractal function or statistical property, such as the variance.

Consider a simple Euclidean straight line λ of length $L(\lambda)$ over which we 'walk' a shorter 'ruler' of length δ. The number of steps taken to cover the line $N[L(\lambda), \delta]$ is then L/δ which is not always an integer for arbitrary L and δ. Since

$$N[L(\lambda), \delta] = \frac{L(\lambda)}{\delta} = L(\lambda)\delta^{-1},$$

$$\Rightarrow 1 = \frac{\ln L(\lambda) - \ln N[L(\lambda), \delta]}{\ln \delta}$$

$$= -\left(\frac{\ln N[L(\lambda), \delta] - \ln L(\lambda)}{\ln \delta}\right)$$

which expresses the topological dimension $D_T = 1$ of the line. In this case, $L(\lambda)$ is the Lebesgue measure of the line and if we normalize by setting $L(\lambda) = 1$, the latter equation can then be written as

$$1 = -\lim_{\delta \to 0}\left[\frac{\ln N(\delta)}{\ln \delta}\right]$$

since there is less error in counting $N(\delta)$ as δ becomes smaller. We also then have $N(\delta) = \delta^{-1}$. For extension to a fractal curve f, the essential point is that the fractal dimension should satisfy an equation of the form

$$N[F(f), \delta] = F(f)\delta^{-D_F}$$

where $N[F(f), \delta]$ is 'read' as the number of rulers of size δ needed to cover a fractal set f whose measure is $F(f)$ which can be any valid suitable measure of the curve. Again we may normalize, which amounts to defining a new measure F' as some constant multiplied by the old measure to get

$$D_F = -\lim_{\delta \to 0}\left[\frac{\ln N(\delta)}{\ln \delta}\right]$$

where $N(\delta)$ is taken to be $N[F'(f), \delta]$ for notational convenience. Thus a piecewise continuous field has precise fractal properties over all scales. However, for the discrete (sampled) field

$$D = -\left\langle \frac{\ln N(\delta)}{\ln \delta} \right\rangle$$

where we choose values δ_1 and δ_2 (i.e. the upper and lower bounds) satisfying $\delta_1 < \delta < \delta_2$ over which we apply an averaging processes denoted by $\langle\ \rangle$. The most common approach is to utilise a bi-logarithmic plot of $\ln N(\delta)$ against $\ln \delta$, choose values δ_1 and δ_2 over which the plot is uniform and apply an appropriate

data fitting algorithm (e.g. a least squares estimation method or, as used in this work, Orthogonal Linear Regression) within these limits.

The relationship between the Fourier dimension q and the fractal dimension D_F can be determined by considering this method for analysing a statistically self-affine field. For a fractional Brownian process (with unit step length)

$$A(t) = t^H, \quad H \in (0, 1]$$

where H is the Hurst dimension. Consider a fractal curve covering a time period $\Delta t = 1$ which is divided up into $N = 1/\Delta t$ equal intervals. The amplitude increments ΔA are then given by

$$\Delta A = \Delta t^H = \frac{1}{N^H} = N^{-H}$$

The number of lengths $\delta = N^{-1}$ required to cover each interval is

$$\Delta A \Delta t = \frac{N^{-H}}{N^{-1}} = N^{1-H}$$

so that

$$N(\delta) = N N^{1-H} = N^{2-H}$$

Now, since

$$N(\delta) = \frac{1}{\delta^{D_F}}, \quad \delta \to 0,$$

then, by inspection,

$$D_F = 2 - H$$

Thus, a Brownian process, where $H = 1/2$, has a fractal dimension of 1.5. For higher topological dimensions D_T

$$D_F = D_T + 1 - H$$

This algebraic equation provides the relationship between the fractal dimension D_F, the topological dimension D_T and the Hurst dimension H. We can now determine the relationship between the Fourier dimension q and the fractal dimension D_F.

Consider a fractal signal $f(x)$ over an infinite support with a finite sample $f_X(x)$, given by

$$f_X(x) = \begin{cases} f(x), & 0 < x < X; \\ 0, & \text{otherwise.} \end{cases}$$

A finite sample is essential as otherwise the power spectrum diverges. Moreover, if $f(x)$ is a random function then for any experiment or computer simulation we must necessarily take a finite sample. Let $F_X(k)$ be the Fourier transform of $f_X(x)$, $P_X(k)$ be the power spectrum and $P(k)$ be the power spectrum of $f(x)$. Then

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F_X(k) \exp(ikx) dk,$$

$$P_X(k) = \frac{1}{X} |F_X(k)|^2$$

and

$$P(k) = \lim_{X \to \infty} P_X(k)$$

The power spectrum gives an expression for the power of a signal for particular harmonics. $P(k)dk$ gives the power in the range k to $k + dk$. Consider a function $g(x)$, obtained from $f(x)$ by scaling the x-coordinate by some $a > 0$, the f-coordinate by $1/a^H$ and then taking a finite sample as before, i.e.

$$g_X(x) = \begin{cases} g(x) = \frac{1}{a^H} f(ax), & 0 < x < X; \\ 0, & \text{otherwise.} \end{cases}$$

Let $G_X(k)$ and $P'_X(k)$ be the Fourier transform and power spectrum of $g_X(x)$, respectively. We then obtain an expression for G_X in terms of F_X,

$$G_X(k) = \int_0^X g_X(x) \exp(-ikx) dx =$$

$$\frac{1}{a^{H+1}} \int_0^X f(s) \exp\left(-\frac{iks}{a}\right) ds$$

where $s = ax$. Hence

$$G_X(k) = \frac{1}{a^{H+1}} F_X\left(\frac{k}{a}\right)$$

and the power spectrum of $g_X(x)$ is

$$P'_X(k) = \frac{1}{a^{2H+1}} \frac{1}{aX} \left| F_X\left(\frac{k}{a}\right) \right|^2$$

and, as $X \to \infty$,

$$P'(k) = \frac{1}{a^{2H+1}} P\left(\frac{k}{a}\right)$$

Since $g(x)$ is a scaled version of $f(x)$, their power spectra are equal, and so

$$P(k) = P'(k) = \frac{1}{a^{2H+1}} P\left(\frac{k}{a}\right)$$

If we now set $k = 1$ and then replace $1/a$ by k we get

$$P(k) \propto \frac{1}{k^{2H+1}} = \frac{1}{k^q}$$

Now since $q = 2H + 1$ and $D_F = 2 - H$, we have

$$D_F = 2 - \frac{q-1}{2} = \frac{5-q}{2}$$

The fractal dimension of a fractal signal can be calculated directly from q using the above relationship. This method also generalizes to higher topological dimensions giving

$$q = 2H + D_T$$

Thus, since

$$D_F = D_T + 1 - H$$

then $q = 5 - 2D_F$ for a fractal signal and $q = 8 - 2D_F$ for a fractal surface so that, in general,

$$q = 2(D_T + 1 - D_F) + D_T = 3D_T + 2 - 2D_F$$

and

$$D_F = D_T + 1 - H = D_T + 1 - \frac{q - D_T}{2} = \frac{3D_T + 2 - q}{2}$$

Appendix E: Variation Diminishing Smoothing Kernels

Variation Diminishing Smoothing Kernels (VDSK) are convolution kernels with properties that guarantee smoothness and thereby, eliminate Gibbs' effect around points of discontinuity of a given function. Further the smoothed function can be shown to be made up of a similar succession of concave or convex arcs equal in number to those of the function. Thus, we consider the following question: let there be given a continuous or discontinuous function f whose graph is composed of a succession of alternating concave or convex arcs. Is there a smoothing kernel (or a set of them) which produces a smoothed function whose graph is also made up of a similar succession of concave or convex arcs equal in number to those of f?.

E.1 Laguerre-Pôlya Class Entire Functions

The class of kernels which relate to this question are a class of entire functions which shall be called class E originally studied earlier by E Laguerre and G Pôlya. An entire function $E(z), z \in \mathbf{C}$ belongs to the class E

$\Longleftrightarrow$

$$E(z) = \exp(bz - cz^2) \prod_{\lambda=1}^{\infty} \Big(1 - \frac{z}{a(\lambda)}\Big) \exp[z/a(\lambda)], \qquad (E.1.1)$$

where $b, c, a(\lambda) \in \mathbf{R}$, $c \geq 0$, *and*

$$\sum_{\lambda=1}^{\infty} a^{-2}(\lambda) < \infty. \qquad (E.1.2)$$

where $\Longleftrightarrow$ is taken to denote 'if and only if' - *iff*. The convergence of the series $(E.1.2)$ guarantees that the product in $(E.1.1)$ converges and represents an entire function. Laguerre proved, and Pôlya added a refinement, that a sequence of polynomials, having real roots only, which converge uniformly in every compact set of the complex plane $\mathbf{C}$, approaches a function of class E in the uniform limit of such a sequence. For example,

$$\exp(-z^2) = \lim_{\lambda \to \infty} \Big(1 - \frac{z^2}{\lambda^2}\Big)^{\lambda^2},$$

and the polynomials $(1 - z^2/\lambda^2)$ have real roots only. In this definition, it is not assumed that the $a(\lambda)$ are distinct. To include the case in which the product has a finite number of factors or reduces to 1 without additional notation, it is assumed that certain points on all the $a(\lambda)$ may be ∞. Furthermore, it is assumed, without loss of generality, that the roots $a(\lambda)$ are arranged in an order of increasing absolute values,

$$0 < |a(1)| \leq |a(2)| \leq |a(3)| \leq \ldots$$

Examples of functions belonging to class E are

$$1,\ 1-z,\ \exp(z),\ \exp(z^2),\ \cos z$$

$$\frac{\sin z}{z},\ \Gamma^{-1}(1-z),\ \Gamma^{-1}(z)$$

Note that the product of two functions of this class produce a new function of the same class.

E.2 Variation Diminishing Smoothing Kernels (VDSKs)

A function k is variation diminishing iff it is of the form

$$k(x) = (2\pi i)^{-1} \int_{-i\infty}^{i\infty} [E(z)]^{-1} \exp(zx)\, dz, \qquad (E.2.1)$$

where $E(z) \in E$ is given by

$$E(z) = \exp(bz - cz^2) \prod_{\lambda=1}^{\infty} \Big(1 - \frac{z}{a(\lambda)}\Big) \exp[z/a(\lambda)], \qquad (E.2.2)$$

with $b, c, a(\lambda) \in \mathbf{R}$, $c \geq 0$, and

$$\sum_{\lambda=1}^{\infty} a^{-2}(\lambda) < \infty$$

In other words, a frequency function k is variation diminishing iff its bilateral Laplace transform equals $[E(z)]^{-1}$:

$$[E(z)]^{-1} = \int_{-\infty}^{\infty} k(x) \exp(-zx)\, dx. \qquad (E.2.3)$$

In order to define a smoothing kernel, the function k given in $(E.2.1)$ must be an even function. For, if $k(x)$ is even, then the corresponding bilateral Laplace transform $[E(z)]^{-1}$ is also even. This fact follows readily from

$$[E(z)]^{-1}$$

$$= \int_{-\infty}^{\infty} k(x) \exp(-zx)\, dx = \int_{-\infty}^{\infty} k(-x) \exp(-zx)\, dx$$

$$= \int_{-\infty}^{\infty} k(x) \exp(zx)\, dx = [E(-z)]^{-1}$$

Conversely, if $[E(z)]^{-1}$ is even, then its inverse bilateral transform is even since a component of convergence of $(E.2.3)$ contains the imaginary axis. This follows from the fact that the component of convergence of each one of the functions which compose $E(z)$ contains completely the imaginary axis. Further, it follows that

$$[E(iu)]^{-1} = K(u), \tag{E.2.4}$$

where $K(u)$ is the FT of k. From the evenness of $[E(z)]^{(-1)}$ it follows that $K(u)$ is real, hence k is even. But $E(z)$ is even iff $b = 0$ and $a(2\lambda-1) = -a(2\lambda)$, $\lambda = 1, 2, \ldots$. Therefore $E(z)$ is taken to be

$$E(z) = \exp(-cz^2) \prod_{\lambda=1}^{\infty} \Big(1 - \frac{z^2}{a^2(\lambda)}\Big), \tag{E.2.5}$$

with $c, a(\lambda) \in \mathbf{R}$, $c \geq 0$, and

$$\sum_{\lambda=1}^{\infty} a^{-2}(\lambda) < \infty$$

Equation $(E.2.4)$ establishes the relationship between the bilateral Laplace transform and the Fourier transform of k. Thus, any analysis associated with use of the bilateral Laplace transform can be undertaken in terms of the Fourier transform.

Using equation $(E.2.4)$ the Fourier transform of $(E.2.1)$ is given by

$$k(x) \leftrightarrow K(u) = [E(iu)]^{-1} = \exp(-cu^2) \prod_{\lambda=1}^{\infty} \Big(\frac{a^2(\lambda)}{a^2(\lambda) + u^2}\Big), \tag{E.2.6}$$

where $\leftrightarrow$ denotes transformation from real to Fourier space, $c, a(\lambda) \in \mathbf{R}$, $c \geq 0$, and $\sum_{\lambda=1}^{\infty} a^{-2}(\lambda) < \infty$.

Because equation $(E.2.6)$ is a variation diminishing function by construction and $|K(0)| \leq 1$, then the following result holds.

Theorem E.2.1 (VDSKs)

k defined as in equation $(E.2.6)$

$\Longrightarrow$

1. *k is a smoothing kernel belonging to SK_1,*
2. *k is variation diminishing,*
3. *$k(x) \geq 0$, $x \in \mathbf{R}$.*

In order to make a complete study of the VDSKs, such kernels will be divided in three classes: *The Finite VDSKs*, *The Non-Finite VDSKs*, and *The Gaussian VDSK*.

E.3 The Finite VDSKs

The finite and the non-finite VDSKs are kernels which can be synthesized from the following basic function:

$$\mathrm{e}(x) = \frac{1}{2}\exp(-|x|),\ x \in \mathbf{R}. \tag{E.3.1}$$

The finite VDSKs are made up by a finite number of convolutions of functions $a(\lambda)\,\mathrm{e}[a(\lambda)x]$, $\lambda = 1, 2, \ldots$. Clearly $\mathrm{e}(x)$ is a VDSK with mean $\nu = 0$ and variance $\sigma^2 = 2$ and its Fourier transform is given by

$$\mathrm{e}(x) \leftrightarrow \frac{1}{1+u^2}. \tag{E.3.2}$$

Note that if $a > 0$, then $a\,\mathrm{e}(ax)$ is again a VDSK. Using the similarity property of the Fourier transform and equation (E.3.2), its Fourier transform is given by

$$a\,\mathrm{e}(ax) \leftrightarrow \frac{a^2}{a^2+u^2}. \tag{E.3.3}$$

Its mean ν again vanishes and its variance takes the value $\sigma^2 = 2/a^2$.

Let $a(1), a(2), \ldots, a(n) > 0$ be constants, some or all of which may be coincident. The following VDSKs are introduced

$$k_\lambda(x) = a(\lambda)\,\mathrm{e}[a(\lambda)x],\ \lambda = 1, 2, \ldots, n. \tag{E.3.4}$$

The combination of these functions by convolution gives a new VDSKs with properties quantified in the following theorem.

Theorem E.3.1 (Properties of The Finite VDSKs)

1. $a(\lambda) > 0$, $\lambda = 1, 2, \ldots,$
2. $k_\lambda(x) = a(\lambda)\,\mathrm{e}[a(\lambda)x],$
3. $k = k_1 \otimes k_2 \otimes \cdots \otimes k_n,$
4. $K(u) = \prod_{\lambda=1}^{n} \left(a^2(\lambda)/(a^2(\lambda)+u^2)\right)$

$\Longrightarrow$

A. k *is a VDSK,*
B. $k(x) \leftrightarrow K(u),$
C. k *has mean* $\nu = 0,$
D. k *has variance* $\sigma^2 = \sum_{\lambda=1}^{n} \left(2/a^2(\lambda)\right) < \infty.$

Proof. A. The assertion follows from mathematical induction.

B. It follows from Convolution Theorem and mathematical induction.

C. Let $k_\lambda(x) \leftrightarrow K_\lambda(u)$. Then because each k_λ is a VDSK, it follows that the respective mean, ν_λ, is given by

$$\nu_\lambda = iK'_\lambda(0) = 0,\ \lambda = 1, 2, \ldots, n$$

Moreover, if $n = 2$, then the mean ν of k is given by

$$\nu = iK'(0) = i(K_1K_2)'(0) = i(K_1K_2{}' + K_1{}'K_2)(0) = i(0) = 0$$

The assertion follows from this result and mathematical induction.

D. Let $k_\lambda(x) \leftrightarrow K_\lambda(u)$. Then because k_λ is a VDSK, it follows that the respective variance, σ_λ^2, is given by

$$\sigma_\lambda^2 = -K''(0) = \frac{2}{a^2\lambda},\ \lambda = 1, 2, \ldots, n$$

Furthermore, from the result given in C above, if $n = 2$, then the mean σ^2 of k is given by

$$\sigma^2 = -K''(0) = -(K_1K_2)''(0)$$

$$= (-K_1K_2{}'' - 2K_1{}'K_2{}' - K_1{}''K_2)(0) = \frac{2}{a^2(1)} + \frac{2}{a^2(2)}$$

The assertion follows from this result and mathematical induction. From the explicit expression of $K(u)$ given in Theorem E.3.1. it follows that

$$K(u) = \prod_{\lambda=1}^{n} \left(\frac{a^2(\lambda)}{a^2(\lambda) + u^2} \right)$$

$$= \prod_{\lambda=1}^{n} \left(\frac{a(\lambda)}{a(\lambda) - iu} \right) \left(\frac{-a(\lambda)}{-a(\lambda) - iu} \right)$$

$$= \prod_{\lambda=1}^{n} \left(\frac{a(\lambda)}{a(\lambda) - iu} \right) \prod_{\lambda=1}^{n} \left(\frac{-a(\lambda)}{-a(\lambda) - iu} \right)$$

$$= \prod_{\lambda=1}^{2n} \left(\frac{d(\lambda)}{d(\lambda) - iu} \right)$$

where $d(\lambda) = a(\lambda)$ for $\lambda = 1, 2, \ldots, n$ and $d(\lambda) = -a(\lambda)$ for $\lambda = n+1, n+2, \ldots, 2n$. Thus k is of degree $2n$ and the following theorem holds.

Theorem E.3.2 (Degree of Differentiability of The Finite VDSKs)

k a finite VDSK,

$\Longrightarrow$

1. $k \in C^{2n-2}(\mathbf{R}, \mathbf{R})$,
2. $k \in C^{2n-1}(\mathbf{R}, \mathbf{R})$ *except at* $x = 0$, *where*

$$k^{2n-1}(0^+),\ k^{2n-1}(0^-)$$

both exist.

The asymptotic behaviour of k and its Fourier transform, K, will be now studied.

Theorem E.3.3 (Asymptotic Behaviour of The Fourier transform of The Finite VDSKs)

1. k a finite VDSK,

2. $k(x) \leftrightarrow K(u)$

$\Longrightarrow$

$$|K(u)| = O(|u|^{-2n}),\ |u| \to \infty$$

Proof. k is made up of a finite convolution operations of functions $k_\lambda(x) = a(\lambda)\,\mathrm{e}[a(\lambda)x]$, where $a(\lambda) > 0,\ \lambda = 1, 2, \ldots, n$; and whose FT, $K_\lambda(u)$, satisfy the inequality

$$|K_\lambda(u)| = \left|\frac{a^2(\lambda)}{a^2(\lambda) + u^2}\right| \leq \frac{a^2(\lambda)}{|u|^2},\ \lambda = 1, 2, \ldots, n$$

Thus

$$|K(u)| = \left|\prod_{\lambda=1}^{n} K_\lambda(u)\right| \leq \prod_{\lambda=1}^{n}\left(\frac{a^2(\lambda)}{|u|^2}\right) = |u|^{-2n}\prod_{\lambda=1}^{n} a^2(\lambda). \qquad (E.3.5)$$

From the above theorem we construct the following corollarys.

Corollary E.3.4 (Absolute and Quadratic Integrability of The Fourier transform of The Finite VDSKs)

1. k a finite VDSK,

2. $k(x) \leftrightarrow K(u)$

$\Longrightarrow$

$K(u) \in L(\mathbf{R}, \mathbf{R}) \cap L^2(\mathbf{R}, \mathbf{R})$.

Corollary E.3.5 (Absolute and Quadratic Integrability of The Finite VDSKs)

k a finite VDSK,

$\Longrightarrow$

$k(x) \in L(\mathbf{R}, \mathbf{R}) \cap L^2(\mathbf{R}, \mathbf{R})$.

The Fourier transform $K(u)$ of the Fourier transform of k is given by

$$K(u) \leftrightarrow 2\pi k(-x)$$

Since k is a even function then

$$K(u) \leftrightarrow 2\pi k(x)$$

This result in conjunction with Corollary E.3.4. and Riemann-Lebesgue Lemma proves the following theorem.

Theorem E.3.6 (Asymptotic Behaviour of The Finite VDSKs)

k a finite VDSK

$\Longrightarrow$

$k(x) \to 0$ *as* $|x| \to \infty$.

E.4 The Non-Finite VDSKs

We now study kernels k holding the property

$$k(x) \leftrightarrow K(u) = \prod_{\lambda=1}^{\infty} \left(\frac{a^2(\lambda)}{a^2(\lambda) + u^2} \right) \tag{E.4.1}$$

which are non-finite kernels. In particular, the infinite product in equation $(E.4.1)$ may have only a finite number of factors, so that the finite VDSKs of the last section are included. Kernels holding equation $(E.4.1)$ can be synthesized from the basic kernel

$$\mathrm{e}(x) = \frac{1}{2} \exp(-|x|), \; x \in \mathbf{R}$$

The non-finite VDSKs are composed of a non-finite number of functions $a(\lambda)\,\mathrm{e}[a(\lambda)x]$, $\lambda = 1, 2, \ldots$. The properties of such kernels are given in the following theorem.

Theorem E.4.1 (Properties of The Non-Finite VDSKs)

1. $a(\lambda) > 0$, $\lambda = 1, 2, \ldots,$
2. $k_\lambda(x) = a(\lambda)\,\mathrm{e}[a(\lambda)x]$,
3. $k = k_1 \otimes k_2 \otimes \cdots \otimes k_n \ldots,$
4. $K(u) = \prod_{\lambda=1}^{\infty} \left(a^2(\lambda)/(a^2(\lambda) + u^2) \right)$

$\Longrightarrow$

A. k is a VDSK,
B. $k(x) \leftrightarrow K(u)$,
C. k has mean $\nu = 0$,
D. k has variance $\sigma^2 = \sum_{\lambda=1}^{\infty} \left(2/a^2(\lambda) \right) < \infty$.

Since k (Theorem E.4.1) is made up by a non-finite number of convolution operationw, then it is of degree infinity, which leads to the following.

Theorem E.4.2 (Degree of Differentiability of The Non-Finite VDSKs)

k a non-finite VDSK

$\Longrightarrow$

$k \in C^\infty(\mathbf{R}, \mathbf{R})$.

The asymptotic behaviour of the Fourier transform of a non-finite kernel is established in the following theorem.

Theorem E.4.3 (Asymptotic Behaviour of The Fourier transform of The Non-Finite VDSKs)

1. k a non-finite VDSK,
2. $k(x) \leftrightarrow K(u)$,
3. $R, p > 0$

$\Longrightarrow$

$|K(u)| = O(|u|^{-2p})$, $|u| \to \infty$.

Proof. Choose $N > p$ and so large that $|a(\lambda)| \geq R$ when $\lambda > N$ which is possible since $|a(\lambda)| \to \infty$ as $\lambda \to \infty$. Set

$$K_N(u) = \prod_{\lambda=N+1}^{\infty} \left(\frac{a^2(\lambda)}{a^2(\lambda) + u^2} \right)$$

By equation $(E.3.5)$, it follows that

$$|K(u)| \leq \frac{|K_N(u)|}{|u|^{2N}} \prod_{\lambda=1}^{N} a^2(\lambda)$$

Because $|K_N(u)|$ never vanishes and is continuous for all $u \in \mathbf{R}$, then it has a positive lower bound. Hence, for a suitable constant M

$$|K(u)| \leq \frac{M}{|u|^{2N}}$$

In particular, if $p = 1$ in the above theorem and because k is a variation diminishing function, the following corollary results.

Corollary E.4.4 (Absolute Integrability of The Non-Finite Kernels and Their FT)

1. k a non-finite VDSK,

2. $k(x) \leftrightarrow K(u)$

$\Longrightarrow$

$k, K \in L(\mathbf{R}, \mathbf{R})$.

Application of the symmetry property of the Fourier transform, the Riemann-Lebesgue Lemma and the above corollary proves the following theorem.

Theorem E.4.5 (Asymptotic Behaviour of The Non-Finite VDSKs)

k a non-finite VDSK

$\Longrightarrow$

$k(x) \to 0$ *as* $|x| \to \infty$.

Some examples of non-finite VDSKs are:

$$\frac{\pi}{4} \operatorname{sech}^2(\frac{\pi x}{2}) \leftrightarrow u \operatorname{csch} u$$

$$= \prod_{\lambda=1}^{\infty} \left(\frac{\lambda^2 \pi^2}{\lambda^2 \pi^2 + u^2} \right), \qquad (E.4.2)$$

$$\frac{1}{2} \operatorname{sech}(\frac{\pi x}{2}) \leftrightarrow \operatorname{sech} u = \prod_{\lambda=1}^{\infty} \left(\frac{(2\lambda - 1)^2 \pi^2}{(2\lambda - 1)^2 \pi^2 + u^2} \right). \qquad (E.4.3)$$

Note that a non-finite VDSK does not necessarily belongs to $L^2(\mathbf{R}, \mathbf{R})$, e.g. the kernel given by equation $(E.4.3)$.

E.5 The Gaussian VDSK

The Gaussian VDSK, k, is defined by the relation

$$k(x) \leftrightarrow K(u) = \exp(-cu^2),\ c > 0. \tag{E.5.1}$$

With $c \to 1/4c^2$, the Gaussian VDSK is now defined as

$$k(x) \leftrightarrow K(u) = \exp(-u^2/4c^2),\ c > 0. \tag{E.5.2}$$

The basic properties of the above kernel follow directly and are collated together in the following theorem.

Theorem E.5.1 (Basic Properties of The Gaussian VDSK)

1. $k(x) = c\,\text{gauss}(cx)$, $c > 0$,
2. $K(u) = \exp(-u^2/4c^2)$, $c > 0$,
3. $p > 0$

$\Longrightarrow$

A. k *is a VDSK,*
B. $k(x) \leftrightarrow K(u)$,
C. k *has mean* $\nu = 0$,
D. k *has variance* $\sigma^2 = 1/2c^2$,
E. $k, K \in L(\mathbf{R}, \mathbf{R}) \cap L^2(\mathbf{R}, \mathbf{R})$,
F. $k, K \in C^\infty(\mathbf{R}, \mathbf{R})$,
G. $|k(x)| = o(|x|^{-p})$,
H. $|K(u)| = o(|u|^{-p})$.

If in equation $(E.5.1)$, c is considered as a variable, say t, then after taking the inverse Fourier transform with respect to x we obtain a real valued function of two variables, i.e.

$$k(x,t) = \frac{1}{\sqrt{4\pi t}} \exp(-x^2/4t). \tag{E.5.3}$$

This new function is the familiar *source solution* of the diffusion equation

$$\left(\frac{\partial^2}{\partial x^2} - \frac{\partial}{\partial t}\right) k(x,t) = 0 \tag{E.5.4}$$

E.6 Geometric Properties of The VDSKs

We consider the general geometric properties shared by the finite, non-finite and the Gaussian VDSKs where k denotes either a finite, non-finite or Gaussian VDSK throughout.

Theorem E.6.1 (Geometric Properties of The VDSKs)

1. k *a VDSK,*
2. $f\colon \mathbf{R} \to \mathbf{R}$ *bounded and convex (concave)*

$\Longrightarrow$

A. For $a, b \in \mathbf{R}$

$$V[k(x) \otimes f(x) - a - bx] \leq V[f(x) - a - bx], \qquad (E.6.1)$$

B. $(k \otimes f)(x)$ *is convex (concave).*

Proof. A. Inequality $(E.6.1)$ follows by a direct application of the variation diminishing property of k.

B. It is well known that f is convex iff

$$\Delta_h^2 f(x) = f(x + 2h) - 2f(x + h) - f(x) \geq 0,$$

for all $x \in \mathbf{R}$, $h > 0$. Because k is a non-negative function, then

$$\Delta_h^2[(k \otimes f)(x)] = \Delta_h^2 \left[\int_{-\infty}^{\infty} k(y) f(x - y)\, dy \right]$$

$$= \int_{-\infty}^{\infty} k(y) \Delta_h^2 f(x - y)\, dy \geq 0$$

Thus the inequality follows. The case for which f is concave follows using a similar argument but $\Delta_h^2 f(x) \leq 0$, for all $x \in \mathbf{R}$, $h > 0$.

The geometric significance of inequality (E.6.1) is that the number of intersections of the straight line $y = a + bx$, $a, b \in \mathbf{R}$, with $(k \otimes f)(x)$ does not exceed the number of intersections of $y = a + bx$ with $y = f(x)$. As a special instance of such an inequality, it follows that $(k \otimes f)(x)$ is non-negative if f is non-negative.

Corollary E.6.2 (Non-Negativity of $k \otimes f$**)**

1. k *a VDSK,*

2. $f : \mathbf{R} \to \mathbf{R}$, $f \geq 0$, *and bounded*

$\Longrightarrow$

$(k \otimes f)(x) \geq 0$, $x \in \mathbf{R}$.

From the above results, it is clear that if f is composed of a succession of alternating convex or concave arcs, then $k \otimes f$ is also made up of a similar succession of convex or concave arcs equal in number to those of f. Thus, a VDSK is *shape preserving.*

Appendix F: M-Code for Computing the Fourier Dimension

```
%INPUT PARMETERS:
%w - length of look-back window - must be set to be smallar than data
%input array size (no check on this!!)
%
%INPUT DATA:
%Input data is default 'Data.txt' where 'PATH' must be set by user.
%
%OUTPUT DATA:
%Output data written to default Q.txt (q-values of analysis)
%'PATH' must be set by the user. Output is simple column vector
%with floating point values of q set to 2 decimal place accuracy.
%
%OUTPUT STATiSTICS
%
%qmin - minimum value of q
%qmax - Maximum value of q
%qrange - numerical range of computed q
%qmean - mean value of q
%qstd - standard deviation of q
%qvar - variance of q
%qmedian - median of q
%qskew - skewness of q
%qkurtosis - kurtosis of q
%normality - Bera-Jarque parametric null hypothesis test
%            of composite normality;
%            normality=0: data is likely to be Gaussian distributed
%            normality=1; data is likely to be non-Gaussian distributed
%
%GRAPHICS
%
%Figure 1 -
%Above:  input data (blue) and q signal (red)
%Below: 100-bin histogram of of q signal
%
%Clear workspace and set window size
clear
w=128;
```

```
%Read data, calculate size
%and setup working arrays
fid=fopen('C:\PATH\Filename.txt','r');
[data n]=fscanf(fid,'%g',[inf]);
fclose(fid);
s=size(data);
n=s(1);
k=1;
for i=1:n
    y(k)=data(i);
k=k+1;
end

m=n-w-1;
for i=1:n
    d(i)=data(i);
end
d=d';
k=1;
for i=1:n
    y(k)=d(i);
k=k+1;
end

for k=1:w
    xw(k)=k;
end

%Start moving window process
for k=1:m
for i=1:w
   yw(i)=y(i+k);
end
yw=fft(yw);
yw=abs(yw).^2;
for i=1:w/2
    if yw(i)>0
ydata(i)=log(yw(i));
xdata(i)=log(xw(i));
    end
end
```

```
%orthogonal linear fit
 p=linortfit(xdata,ydata);
 q(k)=-p(2)/2;
end
for i=1:m
    dw(i)=y(i+w);
    dx(i)=i;
end

%Normalize data
dw=dw./max(dw);

%Plot normalized data and q
figure(1)
subplot(2,1,1), plot(dx,dw,'b-',dx,q,'r-');
grid on;

%Compute histogram
h=hist(q,100);

%Plot histogram
subplot(2,1,2), bar(h), colormap(hot);

%Compute statistics of q
qmin=min(q)
qmax=max(q)
qrange=range(q)
qmean=mean(q)
qstd=std(q)
qvar=var(q)
qmedian=median(q)
qskew=skewness(q)
qkurtosis=kurtosis(q)

%Check for composite normality
normality=jbtest(q)

%Write q values to file
 fid=fopen('C:\PATH\Q.txt','wt');
 fprintf(fid,'%6.2f\n',q);
 fclose(fid);
```

```
function x=linortfit(xdata,ydata)
fun=inline('sum((p(1)+p(2)*xdata-ydata).^2)/(1+p(2)^2)','p','xdata','ydata');
x0=flipdim(polyfit(xdata,ydata,1),2);
options=optimset('TolX',1e-6,'TolFun',1e-6);
x=fminsearch(fun,x0,options,xdata,ydata);
```

References

[1] Fama, E., *Efficient Capital Markets: A Review of Theory and Empirical Work*, Journal of Finance, Vol. 25, 383-417, 1970.

[2] Burton, G. M., *Efficient Market Hypothesis*, The New Palgrave: A Dictionary of Economics, Vol. 2, 120-23, 1987.

[3] Copeland, T. R., Weston, J. F. and Shastri, K., *Financial Theory and Corporate Policy*, 4th Edition, Pearson Addison Wesley, 2003.

[4] Samuelson, P., *Proof That Properly Anticipated Prices Fluctuate Randomly*, Industrial Management Review, Vol. 6, 41-49, 1965.

[5] Fama, E. *The Behavior of Stock Market Prices*, Journal of Business, Vol. 38, 34-105, 1965.

[6] Watsham, T. J. and Parramore, K., *Quantitative Methods in Finance*, Thomson Business Press, 1996.

[7] Martin, J. D. Cox, S. H. McMinn R. F. and Maminn, R. D., *The Theory of Finance: Evidence and Applications*, International Thomson Publishing, 1997.

[8] R. C. Menton, R. C. *Continuous-Time Finance*, Blackwell Publishers, 1992.

[9] Black F. and Scholes M., *The Pricing of Options and Corporate Liabilities*, Journal of Political Economy, Vol. 81(3), 637-659, 1973.

[10] Hurst, H. *Long-term Storage Capacity of Reservoirs*, Trans. of American Society of Civil Engineers, Vol. 116, 770-808, 1951.

[11] Mandelbrot, B. B. and Wallis, J. R., *Robustness of the Rescaled Range R/S in the Measurement of Noncyclic Long Run Statistical Dependence*, Water Resources Research, Vol. 5(5), 967-988, 1969.

[12] Mandelbrot, B. B, *Statistical Methodology for Non-periodic Cycles: From the Covariance to R/S Analysis*, Annals of Economic and Social Measurement, Vol. 1(3), 259-290, 1972.

[13] Mandelbrot, B. B., *The Fractal Geometry of Nature*, Freeman, 1983.

[14] Lichtenberg, A. J. and Lieberman, M. A., *Regular and Stochastic Motion: Applied Mathe-matical Sciences*, Springer-Verlag, 1983.

[15] DeMark, T. R., *The New Science of Technical Analysis*, Wiley, 1994.

[16] Matthews, J. O., Hopcraft, K. I., Jakeman, E. and Siviour, G. B., *Accuracy Analysis of Measurements on a Stable Power-law Distributed Series of Events*, J. Phys. A: Math. Gen. 39, 1396713982, 2006.

[17] Lee, W. H., Hopcraft, K. I. and E. Jakeman, *Continuous and Discrete Stable Processes*, Phys. Rev. E 77, American Physical Society, 011109, 1-4. 2009.

[18] Murphy, J. J., *Intermarket Technical Analysis: Trading Strategies for the Global Stock, Bond, Commodity and Currency Market*, Wiley Finance Editions, Wiley, 1991.

[19] Murphy, J. J., *Technical Analysis of the Futures Markets: A Comprehensive Guide to Trading Methods and Applications*, New York Institute of Finance, Prentice-Hall, 1999.

[20] Einstein, A., *On the Motion of Small Particles Suspended in Liquids at Rest Required by the Molecular-Kinetic Theory of Heat*, Annalen der Physik, Vol. 17, 549-560, 1905.

[21] Shlesinger, M. F., Zaslavsky, G. M. and Frisch U. (Eds.), *Lévy Flights and Related Topics in Physics*, Springer 1994.

[22] Nonnenmacher T. F., *Fractional Integral and Differential Equations for a Class of Lévy-type Probability Densities*, J. Phys. A: Math. Gen., Vol. 23, L697S-L700S, 1990.

[23] Abea, S. and Thurnerb, S., *Anomalous Diffusion in View of Einsteins 1905 Theory of Brownian Motion*, Physica, A(356), Elsevier, 403-407, 2005.

[24] Evans, G. A., Blackledge, J. M., and Yardley P., *Analytical Solutions to Partial Differential Equations*, Springer, 1999.

[25] Feder, J., *Fractals*, Plenum Press, 1988.

[26] Falconer, K. J., *Fractal Geometry*, Wiley, 1990.

[27] Bak, P., *How Nature Works*, Oxford University Press, 1997.

[28] Lvova, I., *Application of Statistical Fractional Methods for the Analysis of Time Series of Currency Exchange Rates*, PhD Thesis, De Montfort University, 2006.

[29] Rao, C. R., *Linear Statistical Inference and its Applications*, Wiley, 1973.

[30] http://www.mathworks.com/matlabcentral/fileexch ange/loadFile.do?objectId=6716&objectType=File

[31] http://uk.finance.yahoo.com/q/hp?s=%5EDJI

[32] Eisenhart, L. P., *Enumeration of Potentials for which One-particle Schrödinger Equations are Separable*, Phys. Rev. 74, 87-89, 1948.

[33] Blackledge, J. M. and Barry, D., *Morphological Analysis based on a Fractional Dynamic Model for Hyphal Growth*, EG UK Theory and Practice of Computer Graphics 2010, To be Published, 2010

[34] Oldham, K. B. and Spanier, J., *The Fractional Calculus*, Academic Press, 1974.

[35] Dold, A. and Eckmann, B. (Eds.), 1975, *Fractional Calculus and its Applications*, Springer, 1975.

[36] Miller, K. S. and Ross B., *An Introduction to the Fractional Calculus and Fractional Differential Equations*, Wiley, 1993.

[37] Samko, S. G., Kilbas, A. and Marichev, O. I., *Fractional Integrals and Derivatives: Theory and Applications*, Gordon and Breach, 1993.

[38] Sneddon, I. N., *The use in Mathematical Physics of Erdélyi-Kober operators and of some of their Generalizations*, Lectures Notes in Mathematics (Eds. A Dold and B Eckmann), Springer, 37-79, 1975.

[39] Kiryakova, V., *Generalized Fractional Calculus and Applications*, Longman, 1994.

[40] Hilfer, R., *Foundations of Fractional Dynamics*, Fractals Vol. 3, No. 3, 549-556, 1995.

[41] Compte, A. *Stochastic Foundations of Fractional Dynamics*, Phys. Rev E, Vol. 53, No. 4, 4191-4193, 1996.

[42] Jespersen, S., Metzler R. and Fogedby, H. C., *Lévy Flights in External Force Fields: Langevin and Fractional Fokker-Planck Equations and Their Solutions*, Phys. Rev. E, Vol. 59, No. 3, 2736- 2745, 1995.

[43] Hilfer, R., *Exact Solutions for a Class of Fractal Time Random Walks*, Fractals, Vol. 3, No. 1, 211-216, 1995.

[44] Hilfer, R. and Anton, L. *Fractional Master Equations and Fractal Time Random Walks*, Phys. Rev. E, Vol. 51, No. 2, R848-R851, 1995.

[45] http://www.markit.com/en/home.page