# Option pricing using path integrals

by

**Dr. Frédéric D.R. Bonnet**

B.Sc. (Mathematical and Computer Science with Honours), 1998.
Ph.D. in Science (Theoretical and Astrophysics), 2002.
The University of Adelaide.

Thesis submitted for the degree of

**Doctor of Philosophy**

in

School of Electrical and Electronic Engineering

Faculty of Engineering, Computer and Mathematical Sciences
University of Adelaide, Australia

July, 2008

Chapter 1

# Introduction

IN this chapter we review the historical aspect of the financial markets and the various techniques used to analyse those markets.

Economists and mathematicians are known to draw models to help them predict parameters or trends in financial markets (Allison and Abbott 2000). However, in recent years more engineers and physicists are becoming involved in the analysis of economic systems and are bringing new concepts and tools to some long standing problems in the research of quantitative methods for financial markets. The standard approach used by physicists is to apply techniques used in statistical mechanics, to obtain fresh insights into the dynamics of financial markets. For example attempts have been made to study the dynamics using a maximal entropy approach (Michael and Johnson 2003) and Buchen and Kelly (1996) found that the data they used was not sufficient to uniquely determine distribution of assets. Alternatively, a modified Ising model to study stochastic resonance and model financial crashes (Krawieecki and Holyst 2003) has been used. Stochastic Differential Equations (SDE) have also been exploited in the evaluation of option pricing (Haven 2002, Faller and Petruccione 2003, Stanislavsky 2003), and have been found to be successful in developing a theory of non-Gaussian option pricing that allows closed form solutions for European options, which can be exercised exclusively on a fixed day of expiration and not before[1]. The approach of Borland (2002a) and Borland (2002b) uses stochastic processes with statistical feedback (Borland 1998a) as a model for stock prices. Such processes were developed within the Tsallis generalized thermostatistics framework (Tsallis 1988a).

Options are ancient financial tools, they are used for speculative purposes or for hedging major market transactions against unexpected changes in the market environment. These can produce large fluctuations in the prices of the assets, and options are intended to prevent the destruction of large amounts of capital. Historically, ancient Romans, Grecians and Phoenicians traded options against outgoing cargo from their local sea ports. In more recent times option pricing techniques have their roots in early work by Castelli who published in 1877 a book entitled *The Theory of Option in Stocks and Shares*. The earliest known analytical valuation for the option was offered by Louis Bachelier in his dissertation at La Sorbonne (Bachelier 1964). Louis Bachelier discovered the treatment of stochastic phenomena five years before Einstein's related but much more famous work on Brownian motion (Einstein 1905b) and twenty three years before Wiener's mathematical development (Wiener 1923), which led to assisting the discovery in 1973 by Fisher Black and Merton Scholes of what is famously called

[1] As in the case of American options that can be exercised anytime during the life of the option.

Black-Scholes model (Black and Scholes 1973). Present-day options are contracts between two parties where each party has the right, but is not obligated, to buy or sell assets. This right has a value that must be purchased at a given price. This price usually depends on the value of the asset in question, hence the name derivative security pricing and risk management of such financial instruments is a major focus of financial market research.

In the Black-Scholes option pricing model, one can assume markets where underlying assets follow a geometric Brownian motion, which is described by an Ito stochastic differential equation (SDE). The disadvantage of the Black-Scholes model is that it is based on several assumptions, namely:

1. markets are efficient,
2. the returns are normally distributed,
3. no commissions are charged,
4. interest rates remain constant and known,
5. the stock pays no dividend during the option's life dividends,
6. European exercise terms are used.

Since 1973 the original Black-Scholes option pricing model has been improved and extended considerably. Merton (1973) included the effect of dividends, three years later Ingerson (1976) relaxed the assumption of no taxes or transaction costs, and Merton removed the restriction of constant interest rates. Currently, we are now in a position of being able to determine the values of a large variety of different options. However many challenges remain, in particular, for options with non-standard payoff features[2] and multi–factor options (i.e., that depend on several underlying assets), these are the options that are commonly traded in real markets and allow anticipated exercise and/or depend on the history of the underlying asset.

Thus the challenge remains to develop more efficient computational tools that the financial analyst can apply to the ever-growing array of more complex derivative instruments being developed by financial markets. One promising approach involves

[2]such as various barrier offers, having features that are designed to reduce option cost by removing protection from risk, considered highly volatile.

the use of path integrals, which have been well developed by physicists. In fact the most important use of path integrals in financial markets is made in the determination of a fair price of financial derivatives in particular options. The relevance of path integrals to this field was first recognized in 1988 by the theoretical physicist Dash (2004) who wrote two seminal unpublished papers on the subject entitled *Path Integrals and Option Pricing I and II*. Since then, many theoretical physicists have entered the field and papers on the subject have begun to appear on the Los Alamos archive (Otto 1998, Otto 1999).

Path integrals were first introduced by Wiener (1921b), see also (Wiener 1921a), and then further developed by Feynman (1948) to perform calculations in quantum electrodynamics (QED). The method has now become a very important one as well as representing a very powerful tool for elementary particle physicists who study the theory of strong interactions in quantum chromodynamics (QCD), where it is possible to use lattice gauge theory to study the fundamental particles of nature that make up the entire universe, such as quarks and gluons (Bonnet *et al.* 2000a, Bonnet *et al.* 2001a, Bonnet *et al.* 2002a, Bonnet *et al.* 2002c, Zhang *et al.* 2004). In that theory a powerful technique, using cutting edge scientific computing (Bonnet *et al.* 2000b, Bonnet *et al.* 2001b, Bonnet *et al.* 2002b), is used to approximate the path integral and calculate both gauge dependent and gauge independent quantities and observables (Bonnet *et al.* 2004). These techniques may be of use in the future for option pricing (Ilinski 1997) where one could formulate a lattice gauge theory for path integrals once a better understanding of the path integral for option pricing is known, which ultimately leads to Monte Carlo simulation of the stock market. Engineers do use path integrals, for example, when studying radio frequency (RF) propagation (Coleman 2002). Applications of path integral techniques (Schulman 1981) to financial markets have also recently started to appear in the literature: for example (Matacz 2000, Montagna *et al.* 2002, Ingber 2000, Chiarella *et al.* 1999, Chiarella and El-Hassan 1997) and the references therein.

The advantage of path integrals when calculating an observable is that they take into consideration the associated fluctuations and give an alternative approach to the Black-Scholes-Merton (BSM) model. Moreover, in the simplest case of a so-called European option, the BSM equation can be explicitly solved to obtain an analytical formula for the price of the option, but when we consider the more complex financial derivatives discussed earlier, the BSM formula fails to give an analytical result. Appropriate

numerical procedures have been developed in the literature to price exotic financial derivatives with path-dependent features as discussed in Hull (2000b) and Paul and Baschnagel (1999). Moreover it is widely recognized that the simplicity of the popular Black-Scholes model, which relates derivative prices to current stock prices and quantifies risk through a constant volatility parameter, is no longer sufficient to capture modern market phenomena especially since the 1987 crash. It has been observed that real markets display high volatility—the modeling of volatility as a stochastic process has been motivated *a priori* by empirical studies of stock price returns in which estimated volatility is observed to exhibit random characteristics (Fouque *et al.* 2000).

The feature of the path integral technique that makes it useful for option pricing is that it provides a way of tracking the evolution of the state of the system over time. In the option pricing context, the state of the system is the vector of prices of underlying assets; usually stocks and bonds. So if we know the state of the system at a particular initial time, it is then possible to estimate the state of that system at a later time in the future. In general, one may thus calculate some transition probability that is used to calculate an expectation value of some functional of a given stochastic process. In the option pricing context this means calculating the expectation (under the risk neutral measure) of the payoff over possible outcomes for the underlying securities. For example for a European call option at the maturity time $T$, the quantity of interest would be $\max\{S(T) - K, 0\}$, where $K$ is the strike price and $S(T)$ is the price of an asset at the maturity time $T$.

An accurate method for calculating option prices, with the path integral approach, remains to be formulated. Path integrals are useful in this case because they involve a quantized theory, resulting in the fact that all of the fluctuations arising from the system are taken into consideration and therefore provides a complete description. Unlike a classical approach, each fluctuation is associated with a path, and each path is weighted according to some transition probability.

So far, various attempts have been made to calculate option prices using path integrals however, up to now all of the formulations that have been published, lead to an analytic exact solution of the path integral due to the insertion of relatively simple probability transition functions (Matacz 2000, Montagna *et al.* 2002, Ingber 2000). One of the aims of this program is to develop methods away from simple geometric Brownian motion models that are used in much of the literature to-date this will allow us to price options on assets driven by more realistic models.

To achieve this goal, we investigate numerical solutions to path integrals. Existing financial data is used to formulate a numerical model for the transition probabilities, which is then inserted as input data with a set of parameters into the path integral, hence leading to a more accurate calculation of the option price. This may be repeated iteratively. The results obtained are then compared to results obtained from the Black-Scholes equations for option pricing.

The advantages of this method are that, (1) real data is used as input, (2) an automated numerical solution arises and (3) the method provides an approach that is applicable to more general models rather than the frequently used Black-Scholes model or the binomial models presented in Rubinstein (1994) and Jackwerth and Rubinstein (1996), where various limiting assumptions are used such as (a) the underlying asset return follows a binomial process, (b) the binomial tree recombines, (c) the ending nodal values are ordered from the lowest to highest.

## 1.1 History of Financial Markets

The concept of trading is a very old concept—in fact it is the basis of virtually all civilizations. Trading some goods, or processing goods to obtain other goods to either generate profit or maintain a living has occurred since the beginning of mankind.

In the following sections we quickly summarize the history of mathematical finance and how it came to what it is today.

We start this section with background on Bachelier's work— Bachelier is believed to be the founder of modern option pricing theory.

### 1.1.1 Bachelier Theory

It has been recognized that *Louis Bachelier* (1870-1946), Fig. 1.1, is the founder of mathematical finance and the "father of modern option pricing". He was born in Le Havre in 1870, he then moved to Paris around 1872 where he worked for some time at the Paris Bourse. During that time he became familiar with the workings of financial markets. Louis Bachelier started his PhD in mathematical physics at the Sorbonne under the supervision of the famous French mathematician Poincaré.

His thesis entitled *Théorie de la Spéculation* (Bachelier 1964, Bachelier 1900) was published in 1900. The thesis dealt with the probabilistic modeling of financial markets

**Figure 1.1. Louis Bachelier.** Louis Bachelier (1870-1946) when he was young.

and triggered the beginning of two scientific theories: The theory of Brownian motion and the mathematical modeling of financial markets—five years before Einstein's famous 1905 paper on Brownian motion. Bachelier worked out, in his doctoral thesis, the distribution functions for the what is now known as the Wiener process (the stochastic process that underlies Brownian motion). The integral equation verified the distribution (later called the Chapman–Kolmogorov equation) and linked it mathematically to Fourier's heat equation. Bachelier's work was initially underestimated by the academic community.

It appears that Einstein, in 1905, ignored Bachelier's work, but it was known to Kolmogorov who brought it to the knowledge of Paul Lévy (1886-1971) many years later. Bachelier's treatment and understanding of the theory of Brownian motion is more mathematical than in Einstein's 1905 paper, which was more focused on Brownian motion of physical particles. In his thesis Bachelier also derived the distribution of the maximum of Brownian motion on an interval and uses it to study barrier options.

Later Paul Lévy and William Feller called the Brownian motion process the Bachelier–Wiener process.

In modern works on probability and mathematical finance, Bachelier's name is frequently quoted and his work is now well recognized. Bachelier's work was formulated in the language of physics and the mathematics was not rigorous (since many of the mathematical techniques necessary to formulate it had not been developed at the time) but the results he obtained were original and basically correct.

Bachelier's work on stochastic modeling of financial markets was unearthed in the 1950s by Samuelson in the United States and an English translation of his thesis subsequently appeared (Bachelier 1964). Inspired by his work, Samuelson formulated the log–normal model for stock prices, which formed the basis for Black–Scholes option pricing model. Unfortunately Bachelier died in 1946 and did not live to see the development of modern mathematical finance.

Presently, his contribution is now well recognized. Interesting material on Bachelier's life and scientific work may be found in Eberlein and Taqqu (1986) and Courtault *et al.* (2000). Moreover in 2000 a Bachelier congress was organized, in his honor, in Paris (Geman *et al.* 2000).

### 1.1.2 History of Brownian Motion

Five years after the work by Louis Bachelier on his *Théorie de la Spéculation* Albert Einstein published four famous papers. Two of the papers were on Brownian motion (Einstein 1905b), which made an attempt to explain the phenomenon observed by the Scottish botanist Brown (1828). This was the foundation of a probabilistic formulation of statistical mechanics and what has become a well established subject of physical investigation.

Einstein's first paper on Brownian motion would appear out of date these days, but nevertheless contains the cornerstone (Hänggi 2005) of the modern theory of stochastic processes. Starting out by using arguments from classical thermodynamics and the concept of osmotic pressure of suspended particles to evaluate a particle's diffusion constant by balancing a diffusion current with drift current (through Stokes' law), he obtained a relation between two transport coefficients: The particle's diffusion constant and the fluid viscosity or friction. This relation known as the Einstein relation (Einstein 1905a, Sutherland 1905) was later generalized in terms of the famous fluctuation dissipation theorem by Callen and Welton (1951) and by the linear response theory of Kubo (1957).

Einstein's 1905 first paper on Brownian motion also contains a derivation of the (over damped) diffusion equation from which he deduces that the root mean square displacement of suspended particles is proportional to the square root of time. Moreover the trajectories of Brownian particles can be regarded as memoryless and non–differentiable (Einstein 1908).

Einstein's theory was able to predict the diffusion constant more accurately and extracted an independent estimate of the Avogadro-Loschmidt number $N$. This was carried out by measuring the distance traveled rather than the velocity (the approach used in unsuccessful attempts). The earlier attempts date back to around 1865 when Josef Loschmidt first tried to measure the size of molecules (Bader and Loschmidt 2001).

In modern terms, Einstein's model assumes that the Brownian motion is a stochastic process with continuous, independent increments, and stationary Gaussian increments. In 1905, the ideas of Borel and Lebesgue measure theory were not developed yet but only came out in the first decade of the twentieth century—as a result Einstein was not able to show that the process he had constructed actually did exist mathematically.

Later in the early 1920's Wiener (1921b) combined ideas from measure theory and Fourier series to construct Brownian motion (Wiener 1921a, Wiener 1923). Wiener and others proved many properties of the paths of Brownian motion, an activity that continues up to the present day. The development of Brownian motion led to stochastic integration. There are two key properties relating stochastic integration, these are

1. The paths of the Brownian motion have non–zero finite quadratic variation such that on an interval $(s, t)$, the quadratic variation is $(t - s)$.

2. The paths of Brownian motion have infinite variation on compact time intervals.

One should note that if Einstein were to have assumed rectifiable paths, Wiener's construction would have essentially proved the impossibility of such model. In recognition of his work, his construction of Brownian motion is often referred as the Wiener process. Wiener also constructed a multiple integral, but it was not what it is known today as the *multiple Wiener integral*. In about 1951, Îto , when trying to understand Wiener's paper, refined and greatly improved Wiener ideas and developed what is known today as Îto calculus and the stochastic differential equation (see Chapter 4 for more details). Brownian motion has had a significant impact on the theory of quantum mechanics itself. But the description of Brownian motion for general quantum system still presents true challenges (Hängi and Ingold 2005, Ankerhold *et al.* 2005). For example little is known about the modeling of quantum fluctuations in stationary non–equilibrium systems or connections between the complexity upon phase–space reduction and the microscopic quantum chaos.

In quantum mechanics the formulation as a sum over paths (Feynman 1948, Feynman and Hibbs 1965, Kleinert 2004) has its roots in the diffusive nature of the trajectories of Brownian walk in continuous time.

Brownian motion has also been used in the 1950's for Gauss–Markov processes (Onsager and Machlup 1953) with linear coefficients and has also inspired mathematicians like Khintchine, Lévy, Mandelbrot, and many physicists and engineers to go beyond Einstein's standard formulation of Brownian motion.

Finally an another interesting class of problems are the non–differentiable Brownian trajectories, in modern language such trajectories are called *fractals* and these trajectories are statistically self–similar on all scales.

We now turn to stochastic calculus, which was derived from the formulation of the Brownian motion.

### 1.1.3 History of Stochastic Calculus

Stochastic calculus really began with the work of Wiener in 1923, which represents the basic foundation for the theory, but it is not until 1951 that Îto refined the theory and greatly improved Wiener ideas (Îto 1951).

Once the foundations were established the next step in the grand work of stochastic integration comes from Kolmogorov who played a fundamental role in the development of theory of stochastic integration from the non–financial perspective, which was motivated and inspired by the theory of Markov processes.

In 1931, two years before his famous book establishing a rigorous mathematical basis for probability theory, using measure theory, Kolmogorov refers to and briefly explains Bachelier's construction of Brownian motion (Kolmogorov 1992). In this paper Kolmogorov showed that continuous Markov processes (diffusion) depends essentially on only two parameters: one for the speed of the drift and the other for the size of the purely random part (the diffusive component). He was then able to relate the probability distribution of the process to the solutions of partial differential equation, which he solved and which are known as *Kolmogorov equations*.

In the late 1920's much work had already been carried out in probability, theory, especially in the two centers Moscow and Paris. A French naturalized mathematician, Vincent Döeblin, joined the probabilist and wanted to construct a stochastic process

with continuous paths that would be consistent with Kolmogorov's analytic theory of transition probabilities for Markov processes (Risken 1984). During the second world war Döeblin was drafted and volunteered to the front. Before he went out he sketched out his ideas and put his work in the safe of the National Academy of Science of France, to be opened after 100 years. Döeblin died during that time and took his notes with him in tragic circumstances. The Academy safe was opened only in May 2000 on the request of his brother (Claude Döeblin). It was only then that the far reaching vision of his work became apparent. In his notes he utilized the new concept of Martingales proposed by Ville only in 1939 and understood the importance of studying sample paths, instead of relying exclusively on the distributional properties. One idea he had was to run Brownian motion by a random clock, which is now known as time change. The change of time was related to the diffusion coefficients and in this way he was able to give a modern treatment of diffusion many years before it was conventionally developed.

As already mentioned at the beginning of this section Kiyosi Îto extended the work of Wiener by formulating a true stochastic differential to be used in the study of Markov–processes, and the formulation of Îto calculus. A good summary of his work and contributions can be found in Varadhan and Stroock (1987).

Îto calculus has led to the theory of stochastic calculus (see Sec. 4 for more details). The theory has so many far reaching applications in so many domains, including most of modern financial modeling. Many great contributions came after Îto , which has widened the field of stochastic calculus (Øksendal 2003, Karatzas and Shreve 1988, Shreve 2004, Kloeden and Platen 1992) just to cite a few.

## 1.2 The Various Techniques Used in Finance

In this Section we describe the most commonly used techniques in finance and in the theory of mathematical finance.

For more than four decades, distributions of financial asset returns have been known to be non–Gaussian—see for example (Mandlebrot 1963, Fama 1965, Bouchaud and Potters 2004). The assumption of normality is confronted against two hard facts, which are illustrated in Fig. 1.2:

NOTE:
This figure is included on page 12
of the print copy of the thesis held in
the University of Adelaide Library.

**Figure 1.2. Daily returns for the NASDAQ and its PDF.** The graph for the time series(top graph), and the distribution function for the daily returns for the NASDAQ when the time lag is 1, that is when $\delta t = 1$ for the period of $16^{th}$ of June 1980 to $14^{th}$ of June 2006. From the graph we can see the effects mentioned in the text, items 1 and 2, that the volatility is clustering and non-Gaussian structure in the time series. Here the data was obtained from *Yahoo Finance (2008)*.

1. The empirical distribution of asset returns have **thicker tails** (commonly known as *fat–tails*) than those from a normal distribution and appears to be negatively skewed. This means more extremes values. This has some very serious implications for risk management and portfolio selection.

2. Returns are time dependent. Squared returns, absolute returns, and all measures and proxies of volatility exhibit strong serial correlations. This is known as **clustering or conditional heteroskedasticity** (Engle 1982).

Financial modeling is all about capturing and exploiting patterns in the data including the phenomena just mentioned above.

There are many tools used in financial modeling, these tools are used in financial institutions that also develop their own strategies for portfolio management, risk evaluation/management and forecasting.

Below, in the next sections we briefly introduce the main ones. These methods will be used and explained in more details in the following chapters.

### 1.2.1 Binomial Tree Diagrams

Unlike the continuous methods that we find in stochastic calculus, the *binomial tree* method is formulated on a discretized evolution of the option price over time. For a more complete discussion on the subject the reader is invited to read (van der Hoek and Elliot 2006, Stampfli and Goodman 2001, Levy 2004).

The general idea is to use some simple bifurcation diagrams that we will call tree diagrams, where each leg of the diagram is associated with the price either going up or down in the price history.

Binomial trees provide a generalizable numerical method for the valuation of options. The binomial model was first proposed by Cox *et al.* (1979). Basically the method uses a *discrete time* model by varying price over time of the underlying asset. The evaluation of the option is carried out via application of the risk neutrality assumption over the life of the option as the price of the underlying asset evolves.

These models are widely used because they are able to handle a variety of conditions for which other models cannot easily be applied. This is largely because they model the underlying asset over time as opposed to a particular point. For example binomial trees are used to value American options, which can be exercised at various points. The discrete nature of the model and its relative simplicity (i.e. mathematically), make it possible for implementation in a spreadsheet. Something that can be very attractive in environment where rapid evaluation is required.

Although slower than the Black–Scholes model formulation, the binomial tree method is considered more accurate, particularly for longer dated options, and options on securities with dividend payments. As a result of this, various versions of the binomial model are widely used by practitioners in the option markets.

For options with several sources of uncertainty (real options) or for options with complicated features, lattice methods are not so practical. Monte Carlo option models are generally used in this cases. Monte Carlo simulations are generally more expansive in terms of computation and are not used when a lattice approach will suffice.

The general framework is set in discrete time so that the evolution of the option's key underlying variable can be traced, for a given time step between valuation date and option expiration. Each node in the lattice, represents a possible price of underlying, at a particular point in time. This price evolution forms a basis for the option evaluation.

The evaluation process is iterative, starting at each final node, and then working backwards through the tree to the first node (the evaluation date), where the calculated result is the value of the option.

The option evaluation using this method is made up of a three step process

1. Price tree generation.
2. Calculation of option value at each final node.
3. Progressive calculation of option value at each earlier node; the value at the first node is the value of the option.

The tree of prices is produced by working forward from valuation date to expiration at each step it is assumed that the underlying instrument will move up or down by a specific factor ($u$ or $d$) per step of the tree where by definition $u \geq 1$ and $0 < d \leq 1$. So if $S_0$ is the current price then the next period the price will either be $S_u = S\,u$ or $S_d = S\,d$ as shown in Fig. 1.3

The up and down factors are calculated using the underlying volatility, $\sigma$ and the time duration of a step $t$ measured in years. From the condition that the log of the price is $\sigma^2 t$ we have

$$u = e^{\sigma\sqrt{t}} \tag{1.1}$$

$$d = e^{-\sigma\sqrt{t}} = \frac{1}{u}. \tag{1.2}$$

The above is the original Cox *et al.* (1979) method. There are other techniques for generating the lattice as was shown by van der Hoek and Elliot (2006).

The Cox–Ross–Rubinstein method ensures that if the underlying asset moves up and then down, it would be equivalent as if it had moved down and then up, that is the two paths merge or recombine. We call such trees recombining trees. In reality or in general this property may actually not be the case. This approximation accelerates the computation of the option price, as it reduces the number of tree nodes. This property

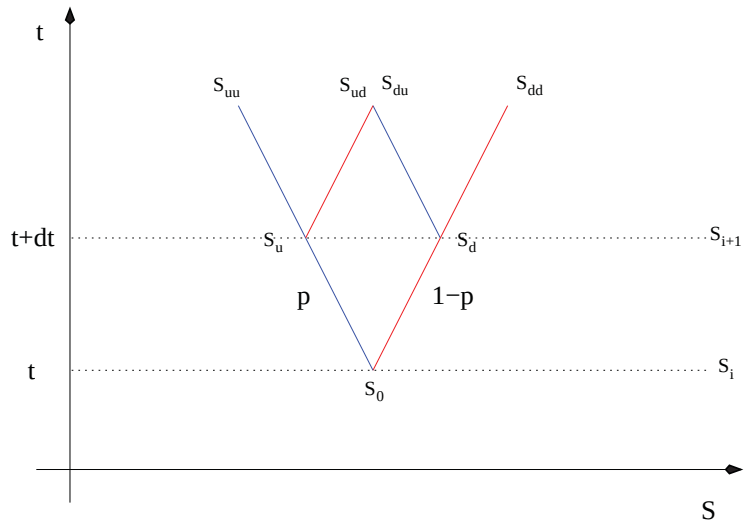


**Figure 1.3. The lattice graph for a binomial tree.** The lattice evolution graph for the first time step in the Cox *et al.* (1979) model. The diagram evolves from one time step to the next. Each asset movement is associated with a given probability $p$ and the complement of that with a probability of $1 - p$. The price of the asset with a probability of going up at the first time step is given by $S_u = S\,u$ similarly when the price of the asset goes down

also allows that the value of the underlying asset at each node is calculated directly via a formula and does not require that the tree be built first. The node value will then be

$$S_n = S_0 u^{N_u - N_d}, \tag{1.3}$$

where $N_u$ and $N_d$ are the number of up or down ticks respectively.

At each node the option value is given by its exercise value, that is for the strike price $K$ and $S$ the spot price we would have

$$\max\{(S - K), 0\}, \tag{1.4}$$

$$\max\{(K - S), 0\}, \tag{1.5}$$

for call and put options respectively.

In the above paragraph the phrase "the underlying volatility" in practice it is commonly known that the volatility is non–stationary as is simplistically assumed in some models, such as in the original Black–Scholes model. One of the major challenges in creating financial models is in modelling volatility. There are many different ways of doing so and one of the most promising exploits methods and techniques from econometrics. In the next section, we briefly introduce the general econometrics ideas and introduce a few models that will be of interest to us in later sections.

### 1.2.2 Econometrics

Econometrics techniques can be used to model time series. Time series can be observed everywhere in the world. One can see time series in any traffic situation, computer systems, demography, electronic and electrical systems, stock markets etc.

Here we are interested in time series which are generated by share prices, stock markets indices, and currency exchanges in order to model financial returns.

If we let $S_t = S(t)$ denote the price at time t for an $n$ valued series, that is for $t = 1, .., n$ with equally spaced time sampling period then the simple net returns $R_t^{(\text{Net})}$ between instant $t-1$ and instant $t$ is

$$R_t^{(\text{Net})} \equiv \frac{S_t - S_{t-1}}{S_{t-1}}, \tag{1.6}$$

and the gross return, $R_t^{(\text{Gross})}$, by

$$R_t^{(\text{Gross})} \equiv \frac{S_t}{S_{t-1}}. \tag{1.7}$$

The gross return, when compounded over $k$ periods, is defined as

$$R_t(k) \equiv \frac{S_{t+k}}{S_{t-1}} = \left(\frac{S_t}{S_{t-1}}\right)\left(\frac{S_{t+1}}{S_t}\right)\cdots\left(\frac{S_{t+k}}{S_{t+k-1}}\right). \tag{1.8}$$

The alternative, is to use the continuously compounded return, which is defined as the natural log of the gross return, Eq. (1.7), by

$$r_t \equiv \log\frac{S_t}{S_{t-1}} = \log(S_t) - \log(S_{t-1}). \tag{1.9}$$

A typical time series for the returns given by Eq. (1.9) is shown, for example, in Fig. 1.4.

As already mentioned in Section 1.2, the distribution for asset returns are both fat–tailed and display some skewness as well as clustering. The fact that the distribution is fat–tailed is due to the fact that large returns occur more frequently. When bad news occurs it is often followed by high volatility. That is negative stock market returns are usually followed by high volatility. The skewness in the probability distribution is from the fact that, in stock market data, large negative returns occur more often than large positive returns. This means that there is some asymmetry in the probability distribution. Finally the clustering comes from the fact that volatility of returns has period of high volatility separated by regions of low volatility.

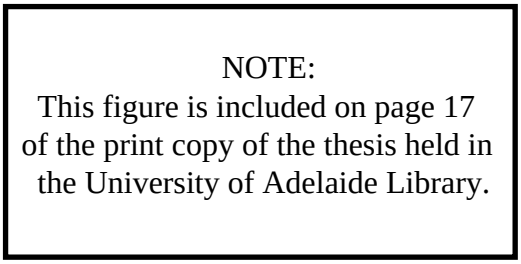


**Figure 1.4. Daily returns for the NASDAQ and S&P500.** Same as Fig. 1.2 but this time the daily returns for the NASDAQ and S&P500 when the time lag is 1, that is when $\delta t = 1$ for the period of $16^{th}$ of June 1980 to $14^{th}$ of June 2006 is graphed. From the graphs we can see that the volatility is clustering and also the non-Gaussian structure in the time series. Here the data was obtained from *Yahoo Finance (2008)*.

Econometrics focuses in describing financial returns data using regression based models of the form

$$x_t = X_t^T(\theta) + \zeta_t \quad \text{for} \quad t = 1, .., n. \tag{1.10}$$

Here $n$ is the length of the time series, $x_t$ is the return at time $t$, $X_t^T(\theta)$ is a vector of size $k$, $\theta$ is a vector of $k$ regression coefficients and $\zeta_t$ are the residuals.

The variance $\sigma_t^2$ of the residual at time $t$ is given by the expected values of the residuals, $\sigma_t^2 = \mathrm{E}[\zeta_t^2]$. In the finance literature the term *volatility* depends on the context and refers to either the variance $\sigma_t^2$ or the standard deviation. In reality, only daily or intraday data is available and for these situations the so–called $ARCH$ (Auto-Regressive Conditional Heteroskedasticity Model) (Engle 1982) and the $GARCH$ (Bollerslev 1986) its generalized version (with its variations) are useful in describing such situations. These models are useful because they describe time variation in conditional variance, which partially explains the fat–tail phenomenon present in returns. The returns also

tend to be negatively correlated returns with changes in the volatility—this can be explained using the leverage effect (Black 1976), a feature that can be adapted in $GARCH$ models.

Notably $GARCH$ models have been very successful in modeling several feature of asset prices. Many surveys have emerged on this topic, see for example (Bollerslev *et al.* 1992, Bera and Higgins 1993, Bollerslev *et al.* 1994, Palm 1996) and more recently (Li *et al.* 2002).

The structure of a *volatility model* can be described as

$$x_t = \mu_t(\theta) + \zeta_t, \tag{1.11}$$
$$\zeta_t = \sigma_t(\theta) z_t, \tag{1.12}$$

where $\mu_t(\theta)$ is the conditional expected value of $x_t$ conditioned over the filtration $\mathcal{F}_{t-1}$ and $\sigma_t(\theta)$ is the variance also conditioned over the filtration $\mathcal{F}_{t-1}$, that is

$$\mu_t(\theta) = \mathrm{E}[x_t|\mathcal{F}_{t-1}], \tag{1.13}$$
$$\sigma_t(\theta) = \mathrm{E}[(x_t - \mu_t(\theta))^2|\mathcal{F}_{t-1}]. \tag{1.14}$$

In Eq. (1.11) the return $x_t$ is decomposed into conditional mean $\mu_t(\theta)$ and a residual term $\zeta_t$. The dynamics $\mu_t(\theta)$ may be an $ARMA(p,q)$ process. The filtration $\mathcal{F}_t$ is the information set available at time $t$. It may include current, past returns, current or past residuals or any variable known at time $t$. In Eq. (1.12), $\zeta_t$ has the volatility conditional on the information available at time $t-1$ denoted by $\sigma_t$. The vector $\theta$ is unknown and needs to be fitted for. The random variable $z_t$ will be assumed to follow some distribution with mean 0 and variance 1, this term is usually called white noise.

A volatility model is a model that describes the evolution of $\sigma_t^2(\theta)$. There are two types of models for describing the dynamics of volatility:

1. In the first category, volatility is described as an exact function of a given set of variables. This category includes $GARCH$ models.

2. In the second category, volatility is described as a stochastic function. It includes stochastic volatility models. This category includes the Heston model for example.

Econometrics deals with the first category. The rest of this section will briefly go through some of the models that are currently available.

These models are of importance because as already mentioned, in order to realistically model market data one has to take into account the features observed from the empirical studies.

#### The Different Models

In this subsection we briefly summarize the different models used in econometrics. The main model used in econometrics is the $GARCH$ model with its derived models. The simplest $GARCH$ model is the linear $GARCH(p,q)$, which can be described in terms of fundamental auto-regressive model $AR(p)$ and auto-regressive with moving average model $ARMA(p,q)$ process. The problem with the $AR(p)$ and $ARCH(p)$ models is that because of the large persistence in the volatility, fitting the data with these models requires a large value $p$ (Bollerslev 1986). Something that is not very desirable because it means that it is difficult to distinguish the effect of each parameters. On the other hand because the $GARCH(p,q)$ model is made up of both $AR(p)$ and $ARMA(p,q)$ it is possible to describe real data more accurately with smaller values of $p$ and $q$, making it a more efficient model. Further details in studies with these models can be found in Box and Jenkins (1976), Hamilton (1994), or in Engle (1995).

### 1.2.3 Stochastic Calculus

To model financial data, stochastic calculus is used as an alternative to econometrics. The general idea is to extend differential equations to include random processes.

The most general form for a stochastic differential equation is written as

$$dX(t) = \alpha(t, X(t))dt + \beta(t, X(t))dW(t), \tag{1.15}$$

where $\alpha(t, X(t))$ is called the drift, and $\beta(t, X(t))$ is a diffusion term. The random process $dW(t)$ is a Wiener process. The drift and diffusion terms in Eq. (1.15) can be constant, which leads to Gaussian distributed transition probabilities. However in general, these functions are non–constant functions of $t$ and $X(t)$. In most cases, the general form of the SDE cannot be solved directly but can only be approximated numerically.

The general solution to Eq. (1.15) is given by

$$X(t) = X(0) + \int_0^t ds\, \alpha(s, X(s)) + \int_0^t \beta(s, X(s))dW(s). \tag{1.16}$$

The first integral in Eq. (1.16) is just an ordinary integral and in most cases can be evaluated without too many difficulties. The second integral is more complicated because it involves a random term. This integral cannot be regarded as a normal integral because of the non–differentiabilty of Brownian motion, so it must therefore be considered as a stochastic integral.

Stochastic differential equations are used to model many different sorts of assets, stock prices, interest models and option prices, just to name a few. In this thesis we will be considering mostly the later case.

Option pricing comes in many different forms each having different characteristics and levels of complexity associated with them. Basically an option is the right but not the obligation to buy or sell an asset at a given time in the future. Options are a form of contract between two different parties who agree on a future price. The simplest option is the European, which can only be exercised at the expiry date. In contrast to the European option is the American option, which can be exercised at any time during the life of the option. This small difference leads to completely different mathematics in both cases.

The simplest stochastic differential equation for option pricing is obtained when the drift and the diffusion terms are left constant. In that case we obtain the standard Black–Scholes model. This model is called standard geometric Brownian motion. Unfortunately, as already mentioned, these types of models do not capture all the features observed in real tick data. In Chapter 4, we explain in more detail how to carry out such calculations and stochastic calculus in general. In Chapter 5, some of the different options available on the market are explained.

### 1.2.4 Path Integrals

Another method which is still an active field of research, is the evaluation of option via the method of path integrals. Path integrals emerged from the area of quantum physics and was largely developed by Feynman (1948) for calculations in quantum mechanics and quantum field theory— in particular in quantum electrodynamics (QED) and later in quantum chromodynamics, which is a non–Abelian version of QED.

In general, path integrals in both QCD and QED cannot be solved directly without using some method of approximation. Similarly in quantum field theory, which is an extension of quantum mechanics where the particles are no longer treated as point like

particle but as fields instead. However path integrals have been mainly successful in quantum mechanics mostly.

Path integrals are complicated mathematical objects that can be used in general on only a very small set of problems. A full solution of the path integral is usually obtained without making any approximation or making use of perturbation theory.

In this work we want to investigate if it is possible to use path integrals in a financial setting as an alternative method to SDE for calculating the option price of some given asset (see Chapter 6 for more details).

In this section we present the general idea of path integrals with little mathematics, explicit calculations are carried out in Chapter 6

The general idea of path integrals is that it leads to a representation of physical quantities, usually called observables as an average with an appropriate weighting factor embedded in it.

If we consider a particle moving, as shown in Fig. 1.5, through space–time or any

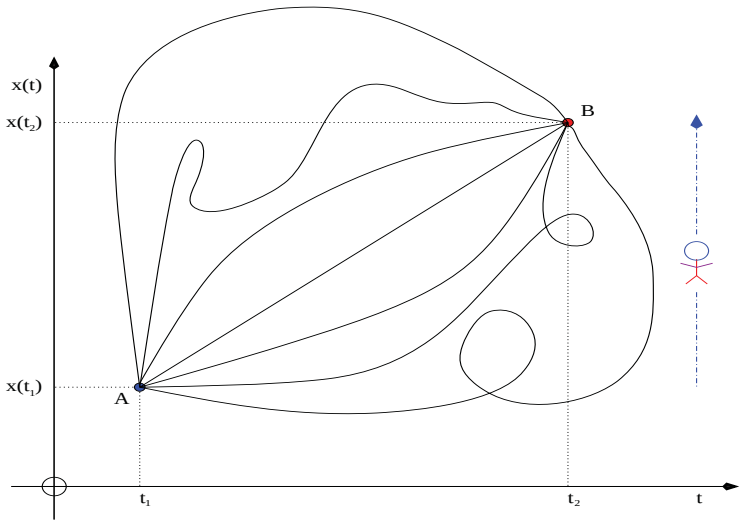


**Figure 1.5. Particle moving from $A$ to $B$.** The picture of a particle moving between two points in a given space, this is taking into consideration all the possible paths that it can take to do so.

medium between two points, $A$ and $B$, from a time $t_1$ to a time $t_2$, then the way to understand the path integral is that the quantum amplitude or transition probability for the particle to move from point $A$ to point $B$ is found by considering all possible paths that join the two points in that medium.

Each path has a probability factor associated with it— what this means is that the paths that are the least likely to happen will negligibly contribute to the path integral. This probability factor is proportional to,

$$e^{-C\,\mathcal{A}[\Omega(t,x(t))]}, \tag{1.17}$$

and is measured from what we call the action functional, $A[\Omega(t,x(t))]$, where $\Omega(t,x(t))$ is a function in space–time. In the case of quantum mechanics the constant $C = i/\hbar$. This exponential factor is the quantum analog of the Boltzmann factor $e^{E/K_B T}$ often used in statistical mechanics.

The action functional carries all of the information about the dynamics of the system, because it is given by a time integral of the Lagrangian density functional, i.e.,

$$\mathcal{A}[\Omega(t,x(t))] = \int_{t_0}^{t} d\tau \mathcal{L}[\Omega(\tau,x(\tau)),\dot{\Omega}(\tau,x(\tau))]. \tag{1.18}$$

This Lagrangian functional also carries all of the information about the system and the system it describes, that is, all the dynamics and its interactions.

From Eq. (1.17) we see that if $A[\Omega(t,x(t))]$ is very large then the exponential factor will tend to zero, and to 1 if $A[\Omega(t,x(t))] \to 0$. In that case we have the most probable path. In general we have

$$e^{-C\,\mathcal{A}[\Omega(t,x(t))]} = \begin{cases} 0 & \text{if } A[\Omega(t,x(t))] \to \infty \\ \Sigma(t,x(t)) & \text{if } 0 < A[\Omega(t,x(t))] < \infty, \\ 1 & \text{if } A[\Omega(t,x(t))] \to 0, \end{cases} \tag{1.19}$$

where $\Sigma(t,x(t))$ is some finite functional.

The path integral can then written in general by considering the integration over all possible paths joining the two points, that is

$$\mathcal{K}(x_T,T|x,t) = \int_{x(t)}^{x(T)} \mathcal{D}x[t']\, e^{-C\mathcal{A}[x(t')]}. \tag{1.20}$$

Here $\mathcal{D}x[t']$ is really the integral measure, which is given by a product of $N$ integrals. Ideally in order to obtain the correct approximation for the path integral one must take the limit as $N$ goes to infinity. In practice we normally take the value of $N$ to be finite, and evaluate the path integral that way. For more details on how path integrals are evaluated in quantum mechanics see Kleinert (2004), Zinn-Justin (2005), Zinn-Justin (2002), Rivers (1987), Feynman (1972) and Roepstorff (1994) for example.

The above discussion sets out the basic idea of the path integral. Unfortunately this path integral cannot be used on every quantum mechanical system, because of the structure of the potential $V(x)$. The potential functional is embedded inside the Lagrangian or the Hamiltonian depending of the formulation that one uses. This structure can make the Feynman path integral tractable or not, even for the simplest potential of the type $V(x) \rightarrow -\frac{1}{|x|}$ the path integral cannot be evaluated, that is because it diverges even for two time slices.

However this is a good starting point and sets out the general basic idea of path integral. In Chapter 6, we will apply these ideas and similar formulation for the path integral in a financial context for different stochastic models to model option price for underlying assets.

## 1.3 Outline of Thesis

The goal of this thesis is divided into three aims. The thesis focuses focus on two specific issues:

**Aim 1:** If we consider the evolution of a stock price through time we see a sample path that is not smooth and appears similar to the evolution of a stochastic differential equation. The aim of this project is to construct a model for the transition probabilities in a given time-window size, using real data, which will then be inserted into the path integral. But before this can be done well one must have a proper strategy to evaluate the path integral. The thesis does not eventualy achieve that goal because of the technical issues associated with the evaluation of the path integral.

**Aim 2:** Because the structure of the path integral is not of simple form, our second aim is to advance numerical methods to provide a numerical solution to the path integrals.

**Aim 3:** Is to fully automate the procedure from the raw data to an option pricing formula constructed directly from a given set of data on a particular time window.

The open questions and research challenges surrounding Aims 1 and 2 are the determination of the optimum window size, determination of the transition probability function, the boundary conditions for the problem, and the format of the data that we will use. Also a research challenge is the construction of a reliable data–driven stochastic differential model, while taking into account features such as fat–tails and clustering also known as conditional heteroskedasticity.

This ambitious approach has many challenges and in each of the Aims 1, 2, and 3, are studied as distinct projects.

The results from Aims 1 and 2 are compared to standard models from previous studies.

The overall outline of the thesis is as follows. The first chapter, Chapter 1, introduces the reader to the field of mathematical finance through a short historical overview. Chapter 2 is the first step in the construction of an array of functionals and available distributions to model financial data. One of the challenges is to model the volatility term accurately, and we explore various models from time series analysis that are available to us for constructing a volatility model from the data for a given data set. Chapter 3 reviews the fundamentals of Brownian motion and is used as a building block for subsequent chapters. Chapter 4 also reviews the fundamentals of stochastic calculus and is divided into two parts. The first part is purely analytical and states some of the fundamentals theorem of stochastic calculus that are used in the evaluation of stochastic differential equations. The second part focuses on numerical approximations for SDEs, which are to be used for verification of convergence in the construction of stochastic differential equations.

Chapters 5 discusses many of the options that are currently available, and is divided into three sections: European, exotic, and American options. Chapter 6 is on path integrals where we carry explicit calculations on given models, such as, Gaussian models, statistical feedback stochastic model and multifractal models. Each model belongs to a different class of stochastic differential equations. On a different spin, Chapter 7 describes about agent models that can be used to model the evolution of the price function using real data sets and compares it to the evolution of real data. Chapter 8 concludes this work and summarizes the findings.

## 1.4 Statement of Original Contributions

This thesis has made a number of contributions to the field of econophysics. The main innovation in this thesis is the approach to option pricing and the way of combining existing knowledge about distribution theory, time series, stochastic calculus, and path integrals, from different disciplines, combining them all in a fully automated algorithm in order to extract the option price. The idea of using scalable window size on real data as an input for model construction to then be inserted into a path integral is an original idea and to our knowledge has not been considered elsewhere.

The first step in the implementation of this idea was carried out in Chapter 2 where we build an array of distributions that describe the data set in question. From the data set we extract a prescription for the volatility, either from time series analysis or from a stochastic volatility model.

In Chapter 3 we review the fundamentals of Brownian motion and use these ideas in Chapter 4. In Chapter 4 we implement numerical approaches that estimate the solutions of stochastic differential equations. We use three different numerical schemes and use these as comparison checks to obtain insights for the solution of the stochastic differential model with the given probability distribution for that particular data set contained within that particular window size.

Chapter 5 reviews the different kind of options onto which may potentially be applied to the above method.

Chapter 6 is the main original contribution of this thesis and gives some insights on how to proceed in the evaluation of path integral when using stochastic differential equations, which spans beyond the Black–Scholes–Merton model, that is, beyond geometric Brownian motion. It also gives an indication on how to proceed for future work.

The second main contribution of this thesis in Chapter 7 is through the use of agent models where we have demonstrated an agent model for studying market bubbles. Within that framework we demonstrate the distinction between the dynamics of a specific minority game and the so called $–Game in the ability to simulate the price function.

Chapter 2

# Distributions and Time series processes

IN this chapter we describe some aspects of distribution theory. In particular we describe the various distributions that are used in econometrics and time series analysis. These distributions underpin any analysis in finance. The best way to construct a model that will capture the features observed in the markets is by understanding the correct structure of the distributions in question. We then fit some empirical data to these distributions to gain insight into which distribution to use for the construction of stochastic volatility models and time series analysis. All of these distributions are aimed for use in the path integral and stochastic differential equation model building.

In this chapter we write out some useful probability distributions. We then use these distributions to see which of these fit the empirical distribution best. These probability distributions are also used in time series analysis for conditional $GARCH$ models. For each distribution we give the probability density function, the loglikelihood, the expected value, $\mathrm{E}[x]$, and variance, and the value of the kurtosis. The kurtosis essentially measures the fatness of the tail in the distribution.

In the introduction we saw in Fig. (1.2), that the probability distribution function for the log return, Eq. (1.9), from empirical data does not follow a Gaussian distribution.

The log return is given by

$$r_{\delta t}(t) \equiv \log \frac{S(t)}{S(t-\delta t)} = \log(S(t)) - \log(S(t-\delta t)), \tag{2.1}$$

where $\delta t$ is what we call the time lag in this case—if we are looking at daily returns then the time lag $\delta t = 1$ and for weekly $\delta t = 5$ and so on. In Fig. (2.1), we show the

NOTE:
This figure is included on page 28
of the print copy of the thesis held in
the University of Adelaide Library.

**Figure 2.1. Distribution function of the NASDAQ and S&P500.** Distribution function for the log returns of the NASDAQ and S&P500 from January 1980 to June 2006 for different ticks, $\delta t = 1$ and $\delta t = 5$ corresponds to 1 and 5 day intervals. Here the data was obtained from *Yahoo Finance (2008)*.

probability distribution function for two stock indices: the NASDAQ and the S&P500, for two different values of $\delta t$, that is, daily and weekly returns. Both indices behave the same and return more or less the same looking distribution function, which is almost symmetric about 0. We do notice that as $\delta t$ increases the height of the distribution

decreases considerably but the tails become wider—this is referred to as a *fat–tailed* distribution.

In Fig. (2.2), we repeat the same procedure but this time with more $\delta t$ values, showing

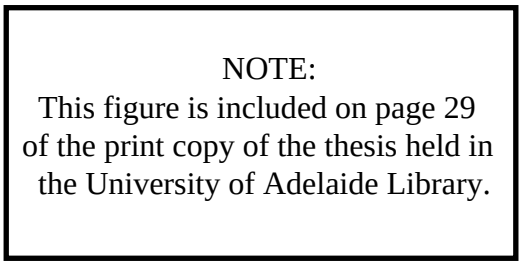


**Figure 2.2. Distribution function of the NASDAQ and S&P500.** Here we show the same graph as in Fig. 2.1 but this time the distribution function is for the log returns of the NASDAQ and S&P500 from January 1980 to June 2006 for different ticks, $\delta t = 1, 5, 20, 40$ and $\delta t = 250$ corresponding to 1,5,20,40 and 250 day intervals, here the data was obtained from *Yahoo Finance (2008)*.

how the distribution evolves as the time lag increases.

The aim is to find a distribution that will fit this data correctly, and for later use in constructing stochastic volatility models or for time series analysis.

The simplest case is the Gaussian distribution. Although it does not fit empirical data correctly, it nevertheless is a very useful distribution as we see in later chapters.

## 2.1 The Gaussian Distribution

The *Gaussian distribution* is given by

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x-\mu)^2}{2\sigma}\right) \tag{2.2}$$

where $\mu$ is the mean and $\sigma$ the variance of the distribution.

The *kurtosis* is calculated using the expected value of $\mathrm{E}[x^2]$ and $\mathrm{E}[x^4]$, i.e.

$$\aleph = \frac{\mathrm{E}[x^4]}{(\mathrm{E}[x^2])^2}. \tag{2.3}$$

Here the kurtosis can be thought of as a measure of the peakedness of the probability distribution of a real–valued random variable. The expected value of $x$ is computed from

$$\mathrm{E}[x] = \int_{\infty}^{\infty} x f(x) dx. \tag{2.4}$$

Hence using the Gaussian properties defined in Appendix A.7.2 by Eq. (A.17) and Eq. (A.18) we find the expected value for $\mathrm{E}[x^2]$ and $\mathrm{E}[x^4]$ to be

$$\mathrm{E}[x^2] = \sigma \quad \text{and} \quad \mathrm{E}[x^4] = 3\sigma^2. \tag{2.5}$$

Hence the kurtosis for the Gaussian distribution is given by

$$\aleph = 3. \tag{2.6}$$

The loglikelihood is obtained by taking the log of $f(x)$

$$\begin{aligned} L(\theta) &= \log[f(x)] \\ &= \frac{1}{2}\log[\sigma] + \frac{x^2}{2\sigma}, \end{aligned} \tag{2.7}$$

where $\theta = (w^T, b_0, b^T)$, the parameter set of the processes defined in Sec. 1.2.2 for the time series are given by $w^T = (\alpha_0, \cdots, \alpha_p, \beta_0, \cdots, \beta_q, \gamma)$ and $b^T = (b_1, \cdots, b_k)$. The loglikehood is a useful tool because it allows us to estimate unknown parameters based on known outcomes. The code for this distribution can be found in Appendix E.4.6.

## 2.2 Non-Gaussian Distributions

### 2.2.1 Student $t$–Distribution

The Student distribution[3] is a special case of the generalised hyperbolic distribution, Sec. 2.2.3. The distribution with mean, $\mu = 0$, is given by

$$f(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)(\nu-2)^{-\frac{1}{2}}\sigma^{-\frac{1}{2}}}{\sqrt{\pi}\Gamma\left(\frac{\nu}{2}\right)}\left[\frac{x^2}{\sigma(\nu-2)}+1\right]^{-\frac{\nu+1}{2}} \quad \text{with} \quad \nu > 2. \tag{2.8}$$

[3]The derivation of the t-distribution was first published in 1908 by William Sealy Gosset, while he worked at a Guinness Brewery in Dublin (Gosset 1908). He was prohibited from publishing under his own name, so the paper was written under the pseudonym 'Student'.

The kurtosis is calculated in exactly the same fashion as for the Gaussian distribution, Sec. 2.1, that is using Eq. (2.3). Using Eq. (A.20) for the calculation of $\mathrm{E}[x^2]$ one can show that the variance take the form

$$\mathrm{E}[x^2] = \sigma \quad \text{and} \quad \mathrm{E}[x^4] = 3\sigma^2 \frac{(\nu-2)}{\nu-4}, \tag{2.9}$$

leading to the kurtosis for the Student distribution of

$$\aleph = 3\left(\frac{\nu-2}{\nu-4}\right) \quad \text{with} \quad \nu > 4. \tag{2.10}$$

Furthermore the loglikelihood of the Student $t$ distribution can be calculated by simply taking the log of Eq. (2.8) that leads to the following expression,

$$\begin{aligned} L(\theta) &= \log[f(x)] \\ &= -\log\left(\Gamma\left(\frac{\nu+1}{2}\right)\right) + \log\left(\Gamma\left(\frac{\nu}{2}\right)\right) + \frac{1}{2}\log(\sigma) \\ &+ \frac{1}{2}\log(\nu-2) + \frac{\nu+1}{2}\log\left(1+\frac{x^2}{\sigma(\nu-2)}\right), \end{aligned} \tag{2.11}$$

where $\theta = (w^T, b_0, \nu, b^T)$, the parameter set of the processes defined in Sec. 1.2.2 for the time series are given by $w^T = (\alpha_0, \cdots, \alpha_p, \beta_0, \cdots, \beta_q, \gamma)$ and $b^T = (b_1, \cdots, b_k)$. The code for this distribution can be found in Appendix E.4.5

### 2.2.2 General Error Distribution

This distribution is also known as the exponential power distribution, the error distribution or the generalized error distribution. The distribution is symmetric about the mean, and the kurtosis can be varied by altering the value of the distribution shape parameters.

The general error distribution with zero mean is given by

$$f(x) = \frac{a}{\lambda 2^{(1+\frac{1}{a})}\Gamma\left(\frac{1}{a}\right)} \exp\left(-\frac{1}{2}\left|\frac{x}{\lambda}\right|^a\right), \tag{2.12}$$

where $\lambda$ is the scale factor, $a$ is the exponent (or shape parameter).

Here the variance and the fourth order expected value are calculated using the integral property Eq. (A.21) and the property of the Gamma function, Eq. (A.23), given by

$$\mathrm{E}[x^2] = \sigma = \lambda^2 2^{\frac{2}{a}} \frac{\Gamma\left(\frac{3}{a}\right)}{\Gamma\left(\frac{1}{a}\right)} \quad \text{and} \quad \mathrm{E}[x^4] = \lambda^4 2^4 \frac{\Gamma\left(\frac{5}{a}\right)}{\Gamma\left(\frac{1}{a}\right)}. \tag{2.13}$$

Eq. (2.13) implies that the scale factor, $\lambda$ is

$$\lambda = \left( \frac{2^{-\frac{2}{a}} \Gamma\left(\frac{1}{a}\right) \sigma}{\Gamma\left(\frac{3}{a}\right)} \right)^{\frac{1}{2}}. \tag{2.14}$$

Eq. (2.14) also leads to the following kurtosis for the general error distribution of

$$\aleph = \frac{\Gamma\left(\frac{5}{a}\right) \Gamma\left(\frac{1}{a}\right)}{\left(\Gamma\left(\frac{3}{a}\right)\right)^2}. \tag{2.15}$$

From Eq. (2.14), we can see how the parameter $a$ controls the shape of the distribution; for example when $a = 1$ we have the Laplacian distribution and when $a = 2$ we obtain the Gaussian distribution. When $a \to \infty$ we obtain the uniform distribution with lower and upper limits of $-(3\sigma)^{1/2}$ and $(3\sigma)^{1/2}$ respectively and a kurtosis of $9/5$. For $a < 2$ the distribution is *leptokurtic*, that is it has tails that are thicker than those for a Gaussian, on the other hand when $a > 2$ the distribution has tails that are *platykurtic*, that is, it has tails that are thinner than those of a Gaussian.

The loglikelihood defined by

$$\begin{aligned} L(\theta) &= \log[f(x)] \\ &= -\log(a) + \log(\lambda) + \left(\frac{1}{a}\right)\log(2) + \log\left(\Gamma\left(\frac{1}{a}\right)\right) + \frac{1}{2}\left|\frac{x}{\lambda}\right|^a, \end{aligned} \tag{2.16}$$

where $\theta = (w^T, a, \lambda, b_0, b^T)$, the parameter set of the processes defined in Sec. 1.2.2 for the time series are given by $w^T = (\alpha_0, \cdots, \alpha_p, \beta_0, \cdots, \beta_q, \gamma)$ and $b^T = (b_1, \cdots, b_k)$.

### 2.2.3 The Generalized Hyperbolic Lévy Motion Distribution, $\mathrm{GH}(x)$

The *generalized hyperbolic*, $\mathrm{GH}(x)$, distribution was introduced in 1977 by Barndorff-Nielsen (1977) to model the grain size distributions of wind blown sand. It can be shown (Barndorff-Nielsen and Halgreen 1977) that the generalized hyperbolic distribution generates a discontinuous Lévy process with length increments of length unity.

The one dimensional density of the generalized hyperbolic distribution is given by

$$\mathrm{GH}(x) = \mathcal{A}\left(\delta^2 + (x-\mu)^2\right)^{\frac{\lambda-\frac{1}{2}}{2}} K_{\lambda-\frac{1}{2}}\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right) \exp\left(\beta(x-\mu)\right) \tag{2.17}$$

where

$$\mathcal{A} = \left.\frac{\left(\alpha^2-\beta^2\right)^{\frac{\lambda}{2}}}{\sqrt{2\pi}\alpha^{\lambda-\frac{1}{2}}\delta^{\lambda}K_{\lambda-\frac{1}{2}}\left(\delta\sqrt{\alpha^2-\beta^2}\right)}\right|_{\alpha>0,\delta>0 \text{ and } 0\leq|\beta|<\alpha}. \tag{2.18}$$

The distribution is controlled by the embedded parameters and offer a lot more flexibility in modeling financial data than other distributions, because there are more parameters to tune the distribution with. The parameter $\alpha$ controls the shape of the distribution, $\beta$ its skewness, and $\delta$ acts as a scale factor similar to $\sigma$ in the normal distribution given above by Eq. (2.2). The location of the distribution is controlled by the parameter $\mu$ and the fatness of the tail by the parameter $\lambda$. In the limit $\delta \to \infty$ the factor $\delta/\alpha \to \sigma^2$ and we recover the normal distribution.

Here, $K_\lambda$ is the modified Bessel function of the third kind with index $\nu$ and can easily be evaluated numerically using the standard numerical recipe routines. The integral representation for $K_\nu$ is

$$K_\nu(x) = \frac{1}{2}\int_0^\infty y^{\nu-1}\exp\left(-\frac{1}{2}x\left(y+\frac{1}{y}\right)\right)dy. \tag{2.19}$$

For $\lambda = \frac{n+1}{2}$ with $n = 0, 2, \cdots$ the Bessel function $K_\lambda$ is

$$K_{n+1}(x) = \frac{\pi}{2}x^{-\frac{1}{2}}\exp(-x)\left(1+\sum_{i=1}^{n}\frac{(n+1)!}{i!(n-i)!}(2x)^{-i}\right). \tag{2.20}$$

The Bessel function given by Eq. (2.19) has the property that $K_\lambda(x) = K_{-\lambda}(x)$, which can be used when we are considering the special cases when $\lambda = 1/2$ or 1. In each case one obtains a different distribution with slightly different properties—see Secs. 2.2.4 and 2.2.5 for these cases.

The mean of the GH$(x)$ distribution is

$$\mathrm{E}[x] = \mu + \frac{\beta\delta}{\sqrt{\alpha^2-\beta 2}}\frac{K_{\lambda+1}\left(\delta\sqrt{\alpha^2-\beta^2}\right)}{K_\lambda\left(\delta\sqrt{\alpha^2-\beta^2}\right)}, \tag{2.21}$$

with variance

$$\mathrm{Var}[x] = \delta^2\left[\frac{K_{\lambda+1}(\zeta)}{\zeta K_\lambda(\zeta)} + \frac{\beta^2}{\alpha^2-\beta^2}\left\{\frac{K_{\lambda+2}(\zeta)}{K_\lambda(\zeta)} - \left(\frac{K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)}\right)^2\right\}\right]\Bigg|_{\zeta=\delta\sqrt{\alpha^2-\beta^2}}. \tag{2.22}$$

Taking the log of Eq. (2.17) we get the loglikelihood of $\mathrm{GH}(x)$ for $n$ independent observations

$$\begin{aligned} L(\theta) &= \log[\mathrm{GH}(x)] \\ &= \log(\mathcal{A}) + \left(\frac{\lambda}{2} - \frac{1}{4}\right) \sum_{i=1}^{n} \log\left(\left(\delta^2 + (x-\mu)^2\right)^{\frac{\lambda-\frac{1}{2}}{2}}\right) \\ &+ \sum_{i=1}^{n} \left\{ \log\left( K_{\lambda-\frac{1}{2}}\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right) \exp\left(\beta(x-\mu)\right)\right)\right\}, \end{aligned} \tag{2.23}$$

where $\theta = (w^T, \alpha, \beta, \delta, \mu, \lambda, b_0, b^T)$, the parameter set of the processes defined in Sec. 1.2.2 for the time series are given by $w^T = (\alpha_0, \cdots, \alpha_p, \beta_0, \cdots, \beta_q, \gamma)$ and $b^T = (b_1, \cdots, b_k)$. The generalized hyperbolic distribution allows for an almost perfect statistical match to these empirical distributions (Prause 1999, Raible 2000, Eberlein 2001). However estimating the parameters using maximum likelihood can be challenging because of the number of parameters and also because some of the parameters are hard to separate, which means that the likelihood function may have several local maxima. Alternatively by exploiting the normal variance-mean mixture structure of the $\mathrm{GH}(x)$ distribution, one may apply the EM-Algorithm (Dempster *et al.* 1977), which is a powerful algorithm for ML estimation on data containing missing values. This is particularly suitable for mixture distribution, since the mixing operation in a sense produces missing data. Moreover this algorithm is easily programmable and surely converges to the maximum, and it provides interesting insights into the model.

We now consider special cases of this distribution that is when $\lambda = 1$ and $\lambda = -1/2$. In the first case one obtains the hyperbolic distribution and in the second we get the normal inverse Gaussian distribution. The code for this distribution can be found in Appendix E.4.1

### 2.2.4 The Hyperbolic Distribution, $H(x)$

The hyperbolic distribution is obtained when $\lambda = 1$. Looking at Eq. (2.17) in Sec. 2.2.3 and setting $\lambda = 1$, the generalized hyperbolic distribution $\mathrm{GH}(x)$ simplifies to the hyperbolic distribution $H(x)$ which is defined by

$$H(x) = \frac{\sqrt{\alpha^2 - \beta^2}}{2\delta\alpha K_1\left(\delta\sqrt{\alpha^2-\beta^2}\right)} \exp\left(-\alpha\sqrt{\delta^2 + (x-\mu)^2} + \beta(x-\mu)\right)\Bigg|_{\delta\leq 0 \text{ and } |\beta|<\alpha}. \tag{2.24}$$

The code for this distribution can be found in Appendix E.4.2

### 2.2.5 The Normal Inverse Gaussian Distribution, NIG$(x)$

Another special case of the generalized hyperbolic distribution, GH$(x)$, is when $\lambda = -1/2$. The distribution simplifies then to the normal inverse Gaussian distribution NIG$(x)$ (Barndorff-Nielsen 1998) and it is defined as

$$\mathrm{NIG}(x) = \frac{\alpha\delta}{\pi} \exp\left(\delta\sqrt{\alpha^2-\beta^2} + \beta(x-\mu)\right) \frac{K_1\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)}{\sqrt{\delta^2+(x-\mu)^2}}\Bigg|_{\delta\leq 0 \text{ and } |\beta|\geq\alpha}. \tag{2.25}$$

When the mean $\mu$ and the skewness $\beta$ parameters are both zero, we have the symmetric centered NIG$(x)$ distribution, NIG$_{sc}(x)$, which has density

$$\mathrm{NIG}_{sc}(x) = \frac{\alpha\delta}{\pi} \exp(\delta\alpha) \frac{K_1\left(\alpha\sqrt{\delta^2+x^2}\right)}{\sqrt{\delta^2+x^2}}\Bigg|_{\delta\leq 0 \text{ and } |\beta|\geq\alpha}, \tag{2.26}$$

with loglikelihood $L(\theta)$ given by

$$\begin{aligned} L(\theta) &= \log[\mathrm{GH}(x)] \\ &= \log\left(\frac{\alpha\delta}{\pi}\right) + \log\left(K_1\left(\alpha\sqrt{\delta^2+x^2}\right)\right) - \log\left(\sqrt{\delta^2+x^2}\right) + \delta\alpha. \end{aligned} \tag{2.27}$$

In Fig. (2.3), we fit the Student distribution to some real data—NASDAQ and the S&P500—for the log return Eq. (2.1) for several time lags that we call $\delta t$. The idea of this fit is to see at what values of $\delta t$ the fit fails for the Student distribution. From the graph we can see that only small values of $\delta t$ gives a perfect fit. This will be useful when the process will be automated, because in the automation process one is pull out the distribution which fits from a pool of distributions without performing any manipulation on the distribution itself. The time lag corresponds to the interval in the data between each sample, so for $\delta t = 1$ we are looking at daily data and when $\delta t = 5$ it corresponds to weekly data, and so on. From the graphs we can observe that the Student distribution fits the data very well when $\delta t = 1$ but not so well when the time lag increases. For example when $deltat = 250$ we can see from the center graph the red curve does not fit the data at all, this is also shown from the $\lambda$ value. It is important to note that the value of $\lambda$ represents the goodness of fit. In general if the $\lambda$ value is greater than 1 the resulting fit is unreliable. In the Table 2.1 we can see that for large value of $\delta t$ the $\lambda$ value is of the order of 10 or higher which clearly shows that the fit is unreliable. This explains why the value of nu is less than two. The numerical results from the fit are summarized in Table 2.1.

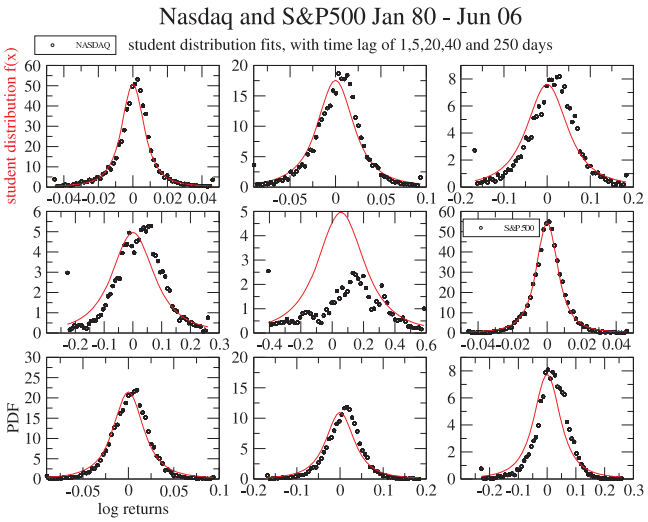


**Figure 2.3. Student distribution fit to NASDAQ and S&P500 tick data.** Fitting (red curve) of the Student distribution to the log returns distribution function for the NASDAQ and S&P500 from January 1980 to June 2006 for different ticks (black empty circles), $\delta t = 1, 5, 20, 40$ and $\delta t = 250$ for the NASDAQ and for $\delta t = 1, 5, 20, 40$ for the S&P500, corresponding to 1,5,20,40 and 250 day intervals. The top left corner graph corresponds to the fit when $\delta t = 1$, the next one to the right corresponds to $\delta t = 5$ and so on until the far right one in the second row which also corresponds to $\delta t = 1$ but this time for the S&P500 data set, obtained from *Yahoo Finance (2008)*. This is similarly the case for the bottom row far left that also corresponds to S&P500 data, but this time when $\delta t = 5$. We can see that the Student distribution produces a good fit only in this case.

So it appears that the Student distribution gives the best fit for small time lag but not for large $\delta t$ values. Hence for these value, the Student distribution cannot be used and other distribution should be used instead. The code for this fit, which uses the Student distribution (Appendix E.4.5), can be found in Appendix E.4.10. The subroutine that fits the PDF (code that gets the PDF is in Appendix E.4.8) is given in Appendix E.4.9.

| $\delta t$ | $a(1)=\nu$ | $a(2)=\sigma$ | $a(3)=\Omega$ | $\chi^2$ | $\lambda$ | |
|---|---|---|---|---|---|---|
| 1 | 2.00046206 | 0.22651291 | 1.02027035 | 175.2454376 | 0.0000100 | NASDAQ |
| 5 | 2.00161004 | 0.58023381 | 1.07071745 | 57.34449387 | 0.0010000 | |
| 20 | 2.00608969 | 0.83409035 | 1.08737576 | 26.51067924 | 0.0010000 | |
| 40 | 2.00343227 | 3.54513001 | 1.09089303 | 16.99897575 | 0.0010000 | |
| 250 | 1.99997783 | 4.41461372 | 0.26370871 | 33.04307556 | 1000.0000 | |
| 1 | 1.96208358 | -18.087507 | 0.84215897 | 6460.717285 | 0.0001000 | S&P500 |
| 5 | 2.00122166 | 0.53618139 | 1.09421635 | 48.31625366 | 0.0001000 | |
| 20 | 2.00291896 | 0.86556965 | 1.09863532 | 54.17663193 | 0.0010000 | |
| 40 | 2.00279140 | 1.78614318 | 1.08881390 | 42.76802063 | 0.0010000 | |
| 250 | 1.99738288 | 9.47337055 | 0.48427352 | 60.98373795 | 10.000000 | |

**Table 2.1. Student distribution fit result.** Result of the fits for the Student distribution, Eq. (2.8) (code Appendix E.4.5) for NASDAQ and S&P500 log returns at various time lags $\delta t$. The results show the values for each of the parameters $a(1)=\nu$, $a(2)=\sigma$ and $a(3)=\Omega$. Here, $\Omega$ can be viewed as normalization factor, which is used as a scaling factor. This factor changes for each data set and in order to obtain a more accurate fit. So $f^{\text{FIT}}(x)=\Omega f(x)$, where $f(x)$ is given by Eq. (2.8).

## 2.3 Linear Processes

### 2.3.1 The AR($p$) Model

The notation AR($p$) stands for auto-regressive time series with a weighted linear sum of size $p$. The idea is to model $n$ observations, $\sigma_t$, for $t=1,\cdots,n$ in term of white noise $z_t$ for $t=1,\cdots,n$, and a weighted linear sum of previous observations.

An auto-regressive time series model of order $p$, AR($p$) takes the following form

$$\sigma_t^2(\theta)=w+\sum_{j=1}^{p}\alpha_j\sigma_{t-j}^2(\theta)+z_t,\quad \text{for}\quad t=1,\cdots,n, \tag{2.28}$$

where $\alpha_j$, for $j=1,\cdots,p$, are the tuned auto-regressive coefficients and $z_t$ is just white noise satisfying

$$\mathrm{E}[z_t]=0\quad\text{and}\quad \mathrm{Var}[z_t]=1. \tag{2.29}$$

For such process it can be shown that $\sigma_t^2(\theta)$ is covariance stationary provided the roots of the polynomials

$$P(x)=1-\alpha_1 x-\alpha_2 x^2-\cdots-\alpha_p x^p=0 \tag{2.30}$$

all have modulus greater than 1, that is $|x_j| > 1$ for $j = 1, \cdots, p$.

If the AR($p$) process is covariance stationary then $\mathrm{E}[\sigma_t^2(\theta)] = \mu$ for all $t$ where $\mu$ is the unconditional mean of the sequence. Taking the expected value in Eq. (2.28) and using Eq. (2.29) we can show that the unconditional mean can be expressed as

$$\mathrm{E}[\sigma_t^2(\theta)] = w + \sum_{j=1}^{p} \alpha_j \mathrm{E}[\sigma_{t-j}^2(\theta)] + \mathrm{E}[z_t] \quad \longrightarrow \quad \mu = w + \mu \sum_{j=1}^{p} \alpha_j, \tag{2.31}$$

which gives the following expression for the unconditional mean

$$\mu = w \left\{ 1 - \sum_{j=1}^{p} \alpha_j \right\}^{-1}. \tag{2.32}$$

### 2.3.2 The ARMA($p, q$) Model

Auto-regressive models can be generalized into an auto-regressive model with moving average called ARMA($p, q$) by introducing an extra lagged term as follows

$$\sigma_t^2(\theta) = w + \sum_{j=1}^{p} \alpha_j \sigma_{t-j}^2(\theta) + \sum_{j=1}^{q} \beta_j z_{t-j} + z_t, \quad \text{for} \quad t = 1, \cdots, n. \tag{2.33}$$

As for the AR($p$) it can be shown that for the ARMA($p$) process, which satisfies the same conditions for the covariance as for the AR($p$) process, the unconditional mean for the ARMA($p$) is the same as for the AR($p$), Eq. (2.32).

### 2.3.3 The ARCH($p$) Model

The ARCH($p$) model by originally introduced in the early 1980s (Engle 1982) and it assumes that the conditional variance is a linear function of the past $p$ squared innovations

$$\sigma_t^2(\theta) = w + \sum_{j=1}^{p} \alpha_j \zeta_{t-j}^2, \quad \text{for} \quad t = 1, \cdots, n. \tag{2.34}$$

Eq. (2.34) says that the conditional volatility is assumed to be a moving average of squared innovations. For this model to be well defined and the conditional variance to be positive the parameters $w$ and $\alpha_j$ must satisfy the following constraints $w > 0$ and $\alpha_j \geq 0$ for $j = 1, \cdots, p$.

The unconditional variance of innovations denoted $\sigma^2$, is the unconditional expectation of $\mathrm{E}[\sigma_t^2(\theta)] = \mathrm{E}[\zeta_t^2] = \sigma^2$. Hence for the ARCH$(p)$ process the unconditional mean is given by

$$\sigma^2 = \mathrm{E}[\sigma_t^2(\theta)] = w\left\{1 - \sum_{j=1}^{p} \alpha_j\right\}^{-1}. \tag{2.35}$$

This shows that the process $\zeta_t$ is covariance stationary if and only if the sum of the auto-regressive parameters is less than 1.

The major problem with the ARCH$(p)$ process is that a large $p$ value must be taken into account in order to fit the data. A more generalized version, proposed by (Bollerslev 1986) in the late 80s, of the process called the GARCH$(p,q)$ is described in the next subsection.

### 2.3.4 The Linear GARCH$(p,q)$ Model

The generalized auto-regressive conditional heteroskedasticity model GARCH$(p,q)$ adds another set of parameters $\beta_j$ which acts as a weighted sum. The conditional variance of a GARCH$(p,q)$ is given by

$$\sigma_t^2(\theta) = w + \sum_{j=1}^{p} \alpha_j \zeta_{t-j}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2(\theta), \quad \text{for} \quad t = 1, \cdots, n, \tag{2.36}$$

$\sigma_t^2(\theta) > 0$ when the parameters $w > 0$, $\alpha_j \geq 0$ for $j = 1, \cdots, p$ and $\beta_j \geq 0$ for $j = 1, \cdots, q$.

Taking the expected value on both sides leads to the unconditional mean, which is given by

$$\sigma^2 = w\left\{1 - \sum_{j=1}^{p} \alpha_j - \sum_{j=1}^{q} \beta_j\right\}^{-1}. \tag{2.37}$$

This means that the process $\zeta_t$ is covariance stationary if and only if $\sum_{j=1}^{p} \alpha_j - \sum_{j=1}^{q} \beta_j < 1$. This is a sufficient but a necessary condition for $\zeta_t$ to be strictly stationary (Bollerslev 1986, Bougerol and Picard 1992, Nelson 1990).

In Fig. (2.4) we show the time series of the log return for the NASDAQ index from January 1980 to June 2006 in time steps of a day. From Fig. (2.4) the clustering effect becomes truly evident. We can clearly see regions of large fluctuation. As already

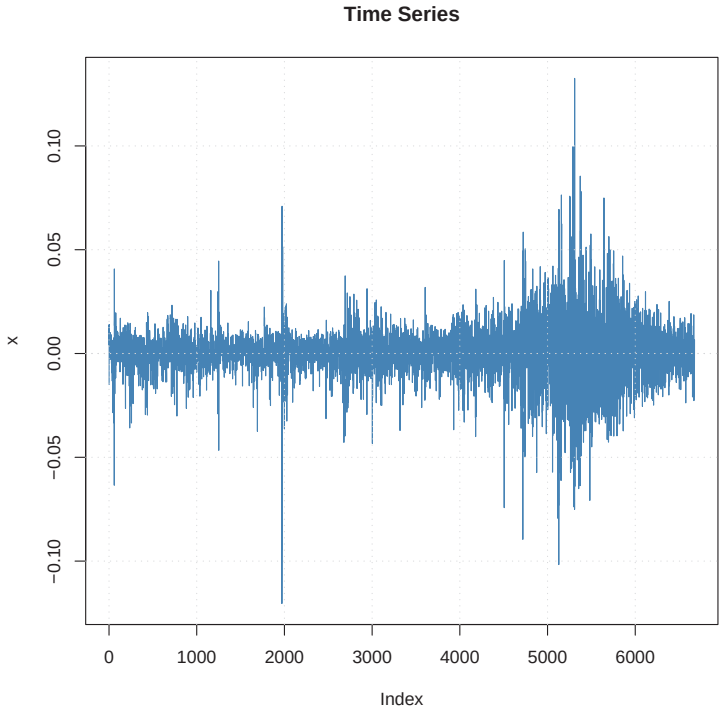


**Figure 2.4. Time series of the NASDAQ $r(t)$.** Time series for the log returns of the NASDAQ from January 1980 to June 2006 for $\delta t = 1$ corresponding to 1 day intervals. From this graph we can clearly see the different features already mentioned in the introduction, these include the periods of high volatility followed by quite periods and the clustering effect.

mentioned earlier, the probability distribution function does not follow a Gaussian distribution function. In Fig. (2.5) we show such a distribution for the time series shown in Fig. (2.4). A good measure of how the distribution deviates from the Gaussian is to use the qnorm $QQ$-Plot, shown in Fig. (2.6). In a Gaussian distribution the graph of the points should be linear, and any deviation from this indicates the non–Gaussian structure of the tails of the distribution. In Fig. (2.6), it is clear how the points deviate from the staight line and it happens almost symmetrically. This means that the distribution is not Gaussian, but fat-tailed and almost symmetric. That is, there is no skewness in the distribution.

A good starting point to time series analysis is to use a GARCH$(1,1)$ model, as it has been shown in Bollerslev (1986) that the GARCH$(1,1)$ model can successfully model exchange rates as well as stock price indices. Here we are trying to see what kind of

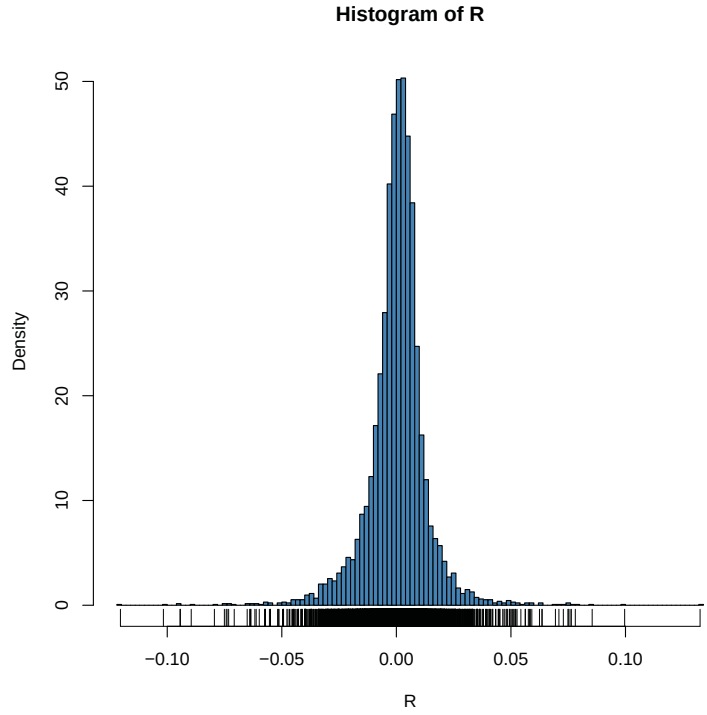


**Figure 2.5. Histogram of the NASDAQ $r(t)$.** Histogram with 100 bins for the log returns of the NASDAQ for Fig. 2.4 from January 1980 to June 2006 for $\delta t = 1$ corresponding to 1 day intervals.

results we obtain when the model is applied to the NASDAQ time series shown in Fig. (2.4).

Setting $p$ and $q$ to 1 in Eq. (2.36), and using the statistical package R (*The Comprehensive R Archive Network* April 4, 2004, by Friedrich Leisch), we obtain the results for the fitted values[4]. From the fit we see that we obtain the following values for $\alpha_1 = 0.132214$ and for $\beta_1 = 0.858471$, also given are the values for $\mu = 7.43221 \times 10^{-4}$ and $\omega = 1.93125 \times 10^{-06}$.

If we examine at the statistics in Table 2.2 we realize that the values are high and that the $p$-values are 0 for most of the Ljung-Box Test (Ljung and Box 1978, Brockwell and Davis 2002). This shows that the fit is probably not reliable. Alternatively we can combine different processes for the mean and the variance to improve the statistics and see if we obtain a much better fit. For example if we model the mean and the variance by an ARMA$(0,1)$ and GARCH$(1,1)$ respectively we can see how the statistics improve.

[4]In Appendix E.1 we give the script used to obtain the fitted values given in the Tables 2.2, 2.3, 2.4 and Table 2.5, also in Appendix. E.1 is a summary of the fit.

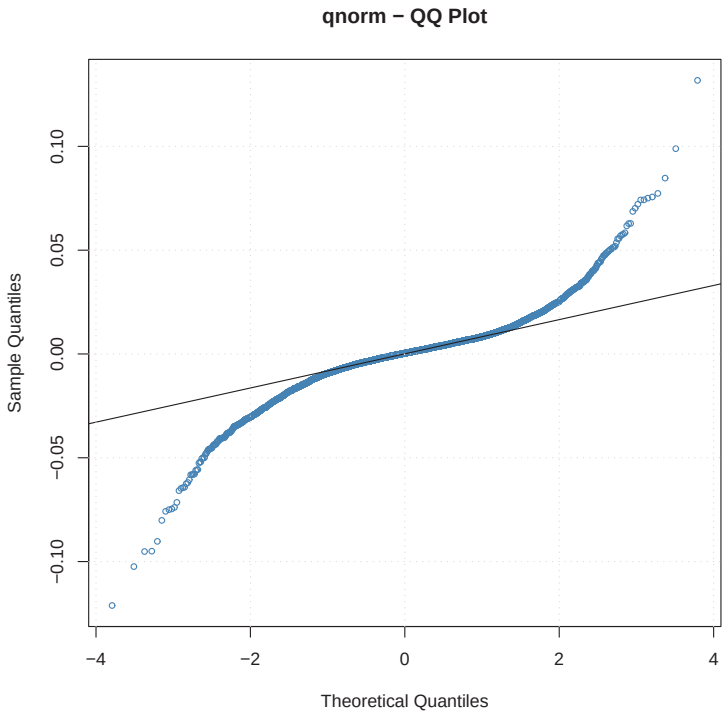


**Figure 2.6. The $qnorm - QQ$ plot of the NASDAQ for $\mathrm{GARCH}(1,1)$.** The $qnorm - QQ$ plot of the NASDAQ for $\mathrm{GARCH}(1,1)$ from the log returns of the NASDAQ from January 1980 to June 2006 for $\delta t = 1$ corresponding to 1 day intervals. From this graph we can clearly see the deviation from the normal distribution and it appears here that the distribution is symmetric.

If we repeat the fit on the same data set, but this time with the model just mentioned we observe an improvement in the statistics. In particular if we look at the resulting value of the Ljung-Box Test in Table 2.3 we clearly see an improvement in the statistics—that is, lower values that ressemble the ones obtained by Bollerslev (1986).

We may repeat the procedure by increasing the number of fitting parameters in the model. In particular in Table 2.4 we show the results from a mean and variance equation ARMA(1,1)+GARCH(1,1). Simlarly in Table 2.5 we show the results from a mean and variance equation ARMA(2,2)+GARCH(1,2). For the later one the statistics appear to be the best giving the best fit to the data.

We also try fitting using ARMA(1,2)+GARCH(2,2) and ARMA(2,2)+GARCH(1,2). This is shown in Table 2.6 and Table 2.7 respectively. From these results we observe that little improvement is gained in the fit. Furthermore the fitting procedure is rather slow.

| Coefficient(s): | Estimate | Std. Error | $t$ value | $\Pr(>\|t\|)$ | Significance |
|---|---|---|---|---|---|
| $\mu$ | 7.432e-04 | 1.006e-04 | 7.387 | 1.50e-13 | 0.001 |
| $\omega$ | 1.931e-06 | 2.581e-07 | 7.483 | 7.26e-14 | 0.001 |
| $\alpha_1$ | 1.322e-01 | 9.571e-03 | 13.815 | $<$ 2e-16 | 0.001 |
| $\beta_1$ | 8.585e-01 | 9.353e-03 | 91.783 | $<$ 2e-16 | 0.001 |
| Statistical test: | Residuals | Test statistic | Statistics | $p$–value | |
| Jarque-Bera Test: | $R$ | $\chi^2$ | 1400.505 | 0 | |
| Ljung-Box Test: | $R$ | Q(10) | 213.4334 | 0 | |
| Ljung-Box Test: | $R$ | Q(15) | 240.7022 | 0 | |
| Ljung-Box Test: | $R$ | Q(20) | 252.2096 | 0 | |
| Ljung-Box Test: | $R^2$ | Q(10) | 11.66196 | 0.3083168 | |
| Ljung-Box Test: | $R^2$ | Q(15) | 15.82622 | 0.3936941 | |
| Ljung-Box Test: | $R^2$ | Q(20) | 18.85845 | 0.5310435 | |
| LM Arch Test: | $R$ | $TR^2$ | 13.29254 | 0.3481412 | |
| | Non–Normalized | normalized | | | |
| Log Likelihood: | -21444.01 | -3.212105 | | | |

**Table 2.2. The results for a mean and variance equation** ARMA(0,0)+GARCH(1,1)**.** Results of the fit for a straight GARCH$(1,1)$ fit to the NASDAQ data of Fig. 2.4. The statistical tests are the Ljung-Box Test and the Jarque-Bera Test. Here the statistics are quite high with a return $p$–value of 0 and this indicates a lack of fit.

This suggests two things, the first one is that a mean and variance equation from a fitting model such as ARMA(1,1)+GARCH(1,1) is probably sufficient to model the volatility of the log returns for this particular time series. A different time series would result in different fitting parameters. The second one is that it is probably as good we can get it for this class of models without having to resort to higher level of sophistication. A summary of the simulation output can be found in Appendix E.1.

In the next sections we briefly review with higher level of sophistication, but we do not carry out any fitting with those models as the scope is better suited for future study.

### 2.3.5 The GARCH$(1,1)-$NIG$(x)$ Model

Although the generalized distribution provides an almost perfect match for these distribution it does not always take into account volatility clustering. Using an alternative parametrization for the NIG$_{sc}(x)$ distribution, that is by setting $q(x) = 1/\left(1+x^2\right)$,

| Coefficient(s): | Estimate | Std. Error | $t$ value | $\Pr(>\|t\|)$ | Significance |
|---|---|---|---|---|---|
| $\mu$ | 7.084e-04 | 1.140e-04 | 6.215 | 5.12e-10 | 0.001 |
| $ma_1$ | 1.803e-01 | 1.309e-02 | 13.772 | $<$ 2e-16 | 0.001 |
| $\omega$ | 1.776e-06 | 2.347e-07 | 7.570 | 3.73e-14 | 0.001 |
| $\alpha_1$ | 1.283e-01 | 9.163e-03 | 14.001 | $<$ 2e-16 | 0.001 |
| $\beta_1$ | 8.630e-01 | 8.921e-03 | 96.737 | $<$ 2e-16 | 0.001 |
| Statistical test: | Residuals | Test statistic | Statistics | $p$–value | |
| Jarque-Bera Test: | $R$ | $\chi^2$ | 1648.716 | 0 | |
| Ljung-Box Test: | $R$ | Q(10) | 20.98795 | 0.0211777 | |
| Ljung-Box Test: | $R$ | Q(15) | 38.31875 | 0.0008093 | |
| Ljung-Box Test: | $R$ | Q(20) | 49.10475 | 0.0002971 | |
| Ljung-Box Test: | $R^2$ | Q(10) | 10.49738 | 0.3979915 | |
| Ljung-Box Test: | $R^2$ | Q(15) | 15.11501 | 0.4431651 | |
| Ljung-Box Test: | $R^2$ | Q(20) | 18.52967 | 0.5525594 | |
| LM Arch Test: | $R$ | $TR^2$ | 13.14630 | 0.3584972 | |
| | Non–normalized | normalized | | | |
| Log Likelihood: | -21536.66 | -3.225983 | | | |

**Table 2.3. The results for a mean and variance equation** ARMA(0,1)+GARCH(1,1)**.** Results of the fit for a $\mathrm{ARMA}(0,1)$ for the mean and a $\mathrm{GARCH}(1,1)$ for the variance to fit the NASDAQ data of Fig. 2.4. The statistical tests are the Ljung-Box Test, LM Arch Test and the Jarque-Bera Test. Here the statistics are quite high with a return $p$–value of 0. This is an improvement when compared to Table 2.2 where a simple GARCH(1,1) was used.

$\alpha^\star = \alpha\delta$ and $\sigma^\star = \delta^{1/2}/\alpha^{1/2}$ the $\mathrm{NIG}_{sc}(x)$ distribution can be rewritten as

$$\mathrm{NIG}_{sc}(x) = \frac{(\alpha^\star)^{\frac{1}{2}}}{\pi\sigma^\star}\exp\left(\alpha^\star\right) q\left(\frac{x}{\sigma^\star\left(\alpha^\star\right)^{\frac{1}{2}}}\right)^{-1} K_1\left(\alpha^\star q\left(\frac{x}{\sigma^\star\left(\alpha^\star\right)^{\frac{1}{2}}}\right)\right). \tag{2.38}$$

Here the function $K_1(x)$ is the modified bessel function defined in Eq. (2.19). The following $\mathrm{GARCH}(1,1)-\mathrm{NIG}(x)$ model was proposed by (Forsberg and Bollerslev 2002)

$$\begin{aligned}\zeta_t|\psi_{t-1} &\sim \mathrm{NIG}((\sigma_t^\star(\theta))^2,\alpha^\star)\\ (\sigma_t^\star(\theta))^2 &= w+\alpha_1\zeta_{t-1}^2+\beta_1(\sigma_{t-1}^\star(\theta))^2\end{aligned} \tag{2.39}$$

where $\zeta_t|\psi_{t-1}$ is the conditional probability. The parameters of this process can be estimated using maximum likelihood techniques.

| Coefficient(s): | Estimate | Std. Error | $t$ value | $\Pr(>\|t\|)$ | Significance |
|---|---|---|---|---|---|
| $\mu$ | 6.340e-04 | 1.212e-04 | 5.232 | 1.67e-07 | 0.001 |
| $ar_1$ | 1.018e-01 | 8.523e-02 | 1.194 | 0.232 | 1.000 |
| $ma_1$ | 8.109e-02 | 8.544e-02 | 0.949 | 0.343 | 1.000 |
| $\omega$ | 1.777e-06 | 2.348e-07 | 7.571 | 3.71e-14 | 0.001 |
| $\alpha_1$ | 1.283e-01 | 9.170e-03 | 13.997 | $<$ 2e-16 | 0.001 |
| $\beta_1$ | 8.629e-01 | 8.934e-03 | 96.580 | $<$ 2e-16 | 0.001 |
| Statistical test: | Residuals | Test statistic | Statistics | $p$–value | |
| Jarque-Bera Test: | $R$ | $\chi^2$ | 1640.465 | 0 | |
| Ljung-Box Test: | $R$ | Q(10) | 16.81144 | 0.0786419 | |
| Ljung-Box Test: | $R$ | Q(15) | 33.36103 | 0.0041826 | |
| Ljung-Box Test: | $R$ | Q(20) | 44.47849 | 0.0012979 | |
| Ljung-Box Test: | $R^2$ | Q(10) | 10.57112 | 0.3918924 | |
| Ljung-Box Test: | $R^2$ | Q(15) | 15.28725 | 0.4309308 | |
| Ljung-Box Test: | $R^2$ | Q(20) | 18.68275 | 0.5425253 | |
| LM Arch Test: | $R$ | $TR^2$ | 13.27991 | 0.3490285 | |
| | Non–normalized | normalized | | | |
| Log Likelihood: | -21537.36 | -3.226087 | | | |

**Table 2.4. The results for a mean and variance equation** ARMA(1,1)+GARCH(1,1)**.** Results of the fit for a ARMA$(1,1)$ for the mean and a GARCH$(1,1)$ for the variance to fit the NASDAQ data of Fig. 2.4. The statistical tests are the Ljung-Box Test, LM Arch Test and the Jarque-Bera Test. Here the statistics are much better and it indicates an improvement in the fit. This is an improvement when compared to Table 2.2, but we notice little improvement when compared to the results obtained from Table 2.3.

We move to non–linear processes, these processes not only capture the fat-tail effect but also take into account the clustering effects observed in the empirical series. These processes are all variations of the GARCH$(p,q)$ processes described in Sec. 2.3.4 and Sec. 2.3.5. In general it is sufficient to consider only processes of low order in $p$ and $q$ to fit the empirical data. The following section describes these processes.

## 2.4 Non–Linear GARCH$(p,q)$ Processes

The *linear* GARCH$(p,q)$ model assumes that both positive and negative shocks of equal magnitude have an identical effect on the future volatility. However empirical

| Coefficient(s): | Estimate | Std. Error | $t$ value | $\Pr(>\|t\|)$ | Significance |
|---|---|---|---|---|---|
| $\mu$ | 3.179e-04 | 5.333e-04 | 0.596 | 0.5511 | 1.000 |
| $ar_1$ | 4.149e-01 | 9.732e-01 | 0.426 | 0.6699 | 1.000 |
| $ar_2$ | 1.294e-01 | 2.340e-01 | 0.553 | 0.5802 | 1.000 |
| $ma_1$ | -2.312e-01 | 9.737e-01 | -0.237 | 0.8123 | 1.000 |
| $ma_2$ | -1.874e-01 | 8.029e-02 | -2.333 | 0.0196 | 0.050 |
| $\omega$ | 1.878e-06 | 2.701e-07 | 6.951 | 3.62e-12 | 0.001 |
| $\alpha_1$ | 1.385e-01 | 1.321e-02 | 10.481 | $<$ 2e-16 | 0.001 |
| $\beta_1$ | 7.395e-01 | 1.057e-01 | 6.995 | 2.66e-12 | 0.001 |
| $\beta_2$ | 1.127e-01 | 9.601e-02 | 1.174 | 0.2403 | 1.000 |
| Statistical test: | Residuals | Test statistic | Statistics | $p$–value | |
| Jarque-Bera Test | $R$ | $\chi^2$ | 1621.872 | 0 | |
| Ljung-Box Test | $R$ | Q(10) | 7.391649 | 0.6880254 | |
| Ljung-Box Test | $R$ | Q(15) | 23.14657 | 0.0810771 | |
| Ljung-Box Test | $R$ | Q(20) | 34.35992 | 0.0237905 | |
| Ljung-Box Test | $R^2$ | Q(10) | 9.465821 | 0.4885379 | |
| Ljung-Box Test | $R^2$ | Q(15) | 13.76548 | 0.5433846 | |
| Ljung-Box Test | $R^2$ | Q(20) | 17.16872 | 0.6419921 | |
| LM Arch Test | $R$ | $TR^2$ | 11.9721 | 0.447923 | |
| | Non–normalized | normalized | | | |
| Log Likelihood: | -21542.28 | -3.226825 | | | |

**Table 2.5. The results for a mean and variance equation** ARMA(2,2)+GARCH(1,2)**.** Results of the fit for a ARMA$(2,2)$ for the mean and a GARCH$(1,2)$ for the variance to fit the NASDAQ data of Fig. 2.4. The statistical tests are the Ljung-Box Test, LM Arch Test and the Jarque-Bera Test. Here the statistics are much better. This is an improvement when compared to Table 2.2. Also notice a greater improvement when comparing the $p$–value, indicating a much better fit. This is an improvement on the results obtained from Table 2.4 and further improved when compared with Table 2.3.

| Coefficient(s): | Estimate | Std. Error | $t$ value | $\Pr(>\|t\|)$ | Significance |
|---|---|---|---|---|---|
| $\mu$ | 9.011e-05 | 4.220e-05 | 2.135 | 0.0328 | 0.050 |
| $ar_1$ | 8.678e-01 | 5.509e-02 | 15.753 | $<$ 2e-16 | 0.001 |
| $ma_1$ | -6.849e-01 | 5.723e-02 | -11.968 | $<$ 2e-16 | 0.001 |
| $ma_2$ | -1.412e-01 | 1.824e-02 | -7.742 | 9.77e-15 | 0.001 |
| $\omega$ | 1.881e-06 | 2.701e-07 | 6.964 | 3.32e-12 | 0.001 |
| $\alpha_1$ | 1.387e-01 | 1.323e-02 | 10.483 | $<$ 2e-16 | 0.001 |
| $\alpha_2$ | 1.000e-08 | 1.049e-06 | 0.010 | 0.9924 | 1.000 |
| $\beta_1$ | 7.401e-01 | 1.054e-01 | 7.024 | 2.16e-12 | 0.001 |
| $\beta_2$ | 1.120e-01 | 9.567e-02 | 1.171 | 0.2417 | 1.000 |
| Statistical test: | Residuals | Test statistic | Statistics | $p$–value | |
| Jarque-Bera Test | $R$ | $\chi^2$ | 1661.630 | 0 | |
| Ljung-Box Test | $R$ | Q(10) | 8.15112 | 0.6140786 | |
| Ljung-Box Test | $R$ | Q(15) | 21.33793 | 0.1263753 | |
| Ljung-Box Test | $R$ | Q(20) | 31.79398 | 0.04554363 | |
| Ljung-Box Test | $R^2$ | Q(10) | 9.316377 | 0.5023645 | |
| Ljung-Box Test | $R^2$ | Q(15) | 13.67046 | 0.550655 | |
| Ljung-Box Test | $R^2$ | Q(20) | 17.10919 | 0.6458725 | |
| LM Arch Test | $R$ | $TR^2$ | 11.87354 | 0.4558884 | |
| | Non–normalized | normalized | | | |
| Log Likelihood: | -21542.28 | -3.226825 | | | |

**Table 2.6. The results for a mean and variance equation** ARMA(1,2)+GARCH(2,2)**.** Results of the fit for a $\mathrm{ARMA}(1,2)$ for the mean and a $\mathrm{GARCH}(2,2)$ for the variance to fit the NASDAQ data of Fig. 2.4. The statistical tests are the Ljung-Box Test, LM Arch Test and the Jarque-Bera Test. Here the statistics are much better. This is an improvement when compared to Table 2.2. Also notice a greater improvement when comparing the $p$–value, indicating a much better fit. This is an improvement on the results obtained from Table 2.4 and further improved when compared with Table 2.3.

| Coefficient(s): | Estimate | Std. Error | $t$ value | $\Pr(>\|t\|)$ | Significance |
|---|---|---|---|---|---|
| $\mu$ | 3.179e-04 | 1.380e-04 | 2.304 | 0.02124 | 0.050 |
| $ar_1$ | 4.148e-01 | 1.566e-01 | 2.650 | 0.00806 | 0.010 |
| $ar_2$ | 1.294e-01 | 3.943e-02 | 3.282 | 0.00103 | 0.010 |
| $ma_1$ | -2.311e-00 | 1.580e-01 | -1.462 | 0.14367 | 1.001 |
| $ma_2$ | -1.874e-01 | 6.308e-02 | -2.970 | 0.00298 | 0.010 |
| $\omega$ | 1.878e-06 | 2.692e-07 | 6.974 | 3.09e-12 | 0.001 |
| $\alpha_1$ | 1.385e-01 | 1.315e-02 | 10.530 | $<$ 2e-16 | 0.001 |
| $\alpha_2$ | 1.000e-08 | NA | NA | NA | 1.000 |
| $\beta_1$ | 7.395e-01 | 1.051e-01 | 7.036 | 1.98e-12 | 0.001 |
| $\beta_2$ | 1.127e-01 | 9.546e-02 | 1.181 | 0.23759 | 1.000 |
| Statistical test: | Residuals | Test statistic | Statistics | $p$–value | |
| Jarque-Bera Test | $R$ | $\chi^2$ | 1621.869 | 0 | |
| Ljung-Box Test | $R$ | Q(10) | 7.391774 | 0.6880133 | |
| Ljung-Box Test | $R$ | Q(15) | 23.14673 | 0.08107385 | |
| Ljung-Box Test | $R$ | Q(20) | 34.36017 | 0.02378904 | |
| Ljung-Box Test | $R^2$ | Q(10) | 9.465918 | 0.488529 | |
| Ljung-Box Test | $R^2$ | Q(15) | 13.76557 | 0.543378 | |
| Ljung-Box Test | $R^2$ | Q(20) | 17.16884 | 0.6419846 | |
| LM Arch Test | $R$ | $TR^2$ | 11.97218 | 0.4479163 | |
| | Non–normalized | normalized | | | |
| Log Likelihood: | -21542.28 | -3.226825 | | | |

**Table 2.7. The results for a mean and variance equation** ARMA(2,2)+GARCH(2,2)**.** Results of the fit for a ARMA$(2,2)$ for the mean and a GARCH$(2,2)$ for the variance to fit the NASDAQ data of Fig. 2.4. The statistical tests are the Ljung-Box Test, LM Arch Test and the Jarque-Bera Test. Here the statistics are much better. This is an improvement when compared to Table 2.2. Also notice a greater improvement when comparing the $p$–value, indicating a much better fit. This is an improvement on the results obtained from Table 2.4 and further improved when compared with Table 2.3.

studies on stock returns have shown that they are characterized by increased volatility following negative shocks (that is bad news). This *leverage* effect was first recognized by Black (1976). The leverage effect suggests that the positive and negative shocks have an asymmetric impact on the conditional volatility of subsequent observations. It has been found that the returns for different asset classes display different leverage characteristics. The returns for equities and equity indices have negative leverage. By contrast, returns for commodities and commodity futures exhibit both positive and negative leverage effects (McKenzie *et al.* 2001). Finally exchange rate returns, where the concept of good/bad news is less defined, have no leverage effects at all. This is because the return series of one currency can be expressed in terms of another and are therefore invertible.

Since linear GARCH$(p,q)$ cannot capture these effects, various non–linear GARCH$(p,q)$ extensions have been proposed. For example for the exponential GARCH$(p,q)$ denoted as EGARCH$(p,q)$, more details may be found in Nelson (1991). For the assymmetric GARCH$(p,q)$ models, i.e. AGARCH$(p,q)$, see Engle and Ng (1993).

Another set of models are the GJR $-$ GARCH$(p,q)$, as shown in Goldsten *et al.* (1993). Here we consider asymmetric effects in AGARCH $-$ I$(p,q)$ and AGARCH $-$ II$(p,q)$ and the GJR $-$ GARCH$(p,q)$ sequences, which can be modeled by the inclusion of an extra asymmetry parameter $\gamma$.

### 2.4.1 The AGARCH $-$ I$(p,q)$ Model

The asymmetry is modeled by the extra parameter $\gamma$. For example, in the standard GARCH$(1,1)$ model when $\sigma^2_{t-1}(\theta)$ is fixed, $\sigma^2_t(\theta) = \sigma^2_t(\zeta_{t-1})$ is a parabola with a minimum at $\zeta_{t-1} = 0$. The introduction of the additional parameter shifts the parabola horizontally so that the minimum occurs at $\zeta_{t-1} = -\gamma$. The conditional variance following negative shocks can therefore be enhanced by choosing $\gamma < 0$, so that $\sigma^2_t(-\zeta_{t-1}) > \sigma^2_t(\zeta_{t-1})$ for $\zeta_{t-1} > 0$.

Using the definition of the GARCH$(p,q)$, Eq. (2.3.3), the AGARCH $-$ I$(p,q)$ model is defined as

$$\sigma^2_t(\theta) = w + \sum_{j=1}^{p} \alpha_j \left(\zeta_{t-j} + \gamma\right)^2 + \sum_{j=1}^{q} \beta_j \sigma^2_{t-j}(\theta), \quad \text{for} \quad t = 1, \cdots, n, \tag{2.40}$$

where $\sigma^2_t(\theta) > 0$ when the parameters $w > 0$, $\alpha_j \geq 0$ for $j = 1, \cdots, p$ and $\beta_j \geq 0$ for $j = 1, \cdots, q$.

Now since $\left(\zeta_{t-j}+\gamma\right)^2 = \zeta_{t-j}^2 + 2\zeta_{t-j}\gamma + \gamma^2$, Eq. (2.40), becomes

$$\sigma_t^2(\theta) = w + \sum_{j=1}^{p}\alpha_j\zeta_{t-j}^2 + \sum_{j=1}^{p}2\alpha_j\zeta_{t-j}\gamma + \sum_{j=1}^{p}\alpha_j\gamma^2 + \sum_{j=1}^{q}\beta_j\sigma_{t-j}^2(\theta). \tag{2.41}$$

Now if we add to both sides the following term $\zeta_{t-j}^2$ and $\sum_{j=1}^{q}\beta_j\zeta_{t-j}^2 - \sum_{j=1}^{q}\beta_j\zeta_{t-j}^2$ on the right hand side of Eq. (2.41) we obtain

$$\begin{aligned}\sigma_t^2(\theta) &= w + \sum_{j=1}^{p}\left(\alpha_j+\beta_j\right)\zeta_{t-j}^2 + 2\sum_{j=1}^{p}\alpha_j\zeta_{t-j}\gamma + \sum_{j=1}^{p}\alpha_j\gamma^2 \\ &+ \sum_{j=1}^{q}\beta_j\left(\sigma_{t-j}^2(\theta) - \zeta_{t-j}^2\right).\end{aligned} \tag{2.42}$$

Setting $\kappa = \max(p,q)$ with $\alpha_j = 0$ for $j > p$ and $\beta_j = 0$ for $j > q$, with $\nu_t = -\sigma_{t-j}^2(\theta) + \zeta_{t-j}^2$ we then get

$$\zeta_t^2 = w + \sum_{j=1}^{\kappa}\left(\alpha_j+\beta_j\right)\zeta_{t-j}^2 + 2\gamma\sum_{j=1}^{p}\alpha_j\zeta_{t-j} + \sum_{j=1}^{p}\alpha_j\gamma^2 - \sum_{j=1}^{q}\beta_j\nu_t + \nu_t. \tag{2.43}$$

Taking the expected value on both sides and using the fact that the expected value of white noise and of the random variable $\nu_t$ is just 0, that is $\mathrm{E}\left[\zeta_t\right] = 0 = \mathrm{E}\left[\zeta_{t-j}\right] = \mathrm{E}\left[\nu_t\right]$ we get

$$\mathrm{E}\left[\zeta_t^2\right] = w + \sum_{j=1}^{\kappa}\left(\alpha_j+\beta_j\right)\mathrm{E}\left[\zeta_{t-j}^2\right] + \sum_{j=1}^{p}\alpha_j\gamma^2. \tag{2.44}$$

This is an AR$(\kappa)$ and the condition for $\zeta_t^2$ to be covariance stationary is

$$\sum_{j=1}^{\kappa}\left(\alpha_j+\beta_j\right) < 1, \tag{2.45}$$

which is the same condition as for the standard linear GARCH$(p,q)$ process. Assuming that $\zeta_t^2$ is covariance stationary we have $\sigma^2 = \mathrm{E}\left[\zeta_t^2\right] = \mathrm{E}\left[\zeta_{t-j}^2\right]$ and so the unconditional variance for this process is

$$\sigma^2 = \frac{w + \gamma^2\sum_{j=1}^{p}\alpha_j}{1 - \sum_{j=1}^{\kappa}\left(\alpha_j+\beta_j\right)}. \tag{2.46}$$

One can then calculate the kurtosis for a particular value of $p$ and $q$. For example when $p = 1$ and $q = 1$ one finds for the AGARCH $-$ I$(1,1)$ model,

$$\aleph = \frac{\mathrm{E}[x^4]}{(\mathrm{E}[x^2])^2} = \frac{3\left(1-\alpha_1^2+\mathcal{F}\right)}{1-3\alpha_1^2} \quad \text{with} \quad \mathcal{F} = \frac{4\alpha_1^2\gamma^2\left(1-\alpha_1\right)}{w+\alpha_1\gamma^2}. \tag{2.47}$$

Furthermore one can evaluate the forecast formula for the expected volatility $\mathrm{E}\left[\sigma_t^2(\theta)|\psi_{t-1}\right]$, which is evaluated under the information set $\psi_{t-1}$ at time $t$. After some computation we find that the AGARCH $-$ I$(1,1)$ forecast is given by

$$\mathrm{E}\left[\sigma_{t+T}^2(\theta)|\psi_{t-1}\right] = \frac{w + \alpha_1\gamma^2 - (\alpha_1+\beta_1)^T}{1-(\alpha_1+\beta_1)} + (\alpha_1+\beta_1)^T \mathrm{E}\left[\sigma_t^2(\theta)|\psi_{t-1}\right]. \quad (2.48)$$

We now turn to the AGARCH $-$ II$(p,q)$ process, which is a slight variation of the AGARCH $-$ I$(p,q)$.

### 2.4.2 The AGARCH $-$ II$(p,q)$ Model

In the AGARCH $-$ II$(p,q)$ model the inclusion of $\gamma$ can also result in an enhancement of $\sigma_t^2(\theta)$ following a negative shock. The model for the process is defined by

$$\sigma_t^2(\theta) = w + \sum_{j=1}^{p} \alpha_j \left(|\zeta_{t-j}| + \zeta_{t-j}\gamma\right)^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2(\theta), \quad \text{for} \quad t = 1, \cdots, n. \quad (2.49)$$

Repeating the same procedure as in Sec. 2.4.1, which leads to Eq. (2.46), we find that the unconditional variance for this process is given by

$$\sigma^2 = \frac{w}{1 - \sum_{j=1}^{\kappa} \left(\alpha_j + \beta_j + \gamma^2 \Delta_j\right)}, \quad (2.50)$$

where $\kappa = \max(p,q)$ with $\Delta_j = \alpha_j = 0$ for $j > p$ and $\beta_j = 0$ for $j > q$ and $\Delta_j = 1$ for $j \leq 0$. The kurtosis will depend on the values that $p$ and $q$ take.

### 2.4.3 The GJR $-$ GARCH$(p,q)$ Model

The GJR $-$ GARCH$(p,q)$ model is defined as

$$\sigma_t^2(\theta) = w + \sum_{j=1}^{p} \left(\alpha_j + \gamma S_{t-j}\right) \zeta_{t-j}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2(\theta), \quad \text{for} \quad t = 1, \cdots, n, \quad (2.51)$$

where $S_t = 1$ when $\zeta_t < 0$ and $S_t = 0$ when $\zeta_t \geq 0$. Repeating the same procedure as in Sec. 2.4.1, which leads to Eq. (2.46), we find that the unconditional variance for this process is given by

$$\sigma^2 = \frac{w}{1 - \sum_{j=1}^{\kappa} \left(\alpha_j + \beta_j + \frac{\gamma}{2}\Delta_j\right)}. \quad (2.52)$$

Here, $\kappa = \max(p,q)$ with $\Delta_j = \alpha_j = 0$ for $j > p$ and $\beta_j = 0$ for $j > q$ and $\Delta_j = 1$ for $j \leq 0$. The kurtosis will depend on the values that $p$ and $q$ take.

### 2.4.4 The EGARCH$(p,q)$ Model

The EGARCH$(p,q)$ model is defined as in Eq. (2.53). For the EGARCH$(p,q)$, an asymmetric response arises from the $\sum_{j=1}^{p}\alpha_j Z_{t-j}$ term. In as EGARCH$(1,1)$, if $\alpha_1 < 0$, then a negative shock $\zeta_t$ increases the value of $\sigma_t^2(\theta)$, so that $\log\left(\sigma_t^2(-Z_{t-j})\right) > \log\left(\sigma_t^2(Z_{t-j})\right)$. The process is written as

$$\begin{aligned}\log\left(\sigma_t^2(\theta)\right) = w &+ \sum_{j=1}^{p}\alpha_j Z_{t-j} + \sum_{j=1}^{p}\phi_t\left(\left|Z_{t-j}\right| - \mathrm{E}\left[\left|Z_{t-j}\right|\right]\right)\\ &+ \sum_{j=1}^{q}\beta_j \log\left(\sigma_{t-j}^2(\theta)\right),\end{aligned} \tag{2.53}$$

for $t = 1, \cdots, n$. Here the extra random variable $Z_t$ is given by $Z_t = \zeta_t/\sqrt{\sigma_t^2(\theta)}$ and its expected value is denoted by $\mathrm{E}\left[\left|Z_{t-j}\right|\right]$.

## 2.5 The Maximum Likelihood Method

The parameters $w, \alpha_j$, and $\beta_j$ need to be estimated and the most commonly used method to do so is to use the method of maximum likelihood (Kay 1993). The method was invented by a geneticist and statistician Sir R. A. Fisher between 1912 and 1922 (Aldrich 1997, Lehmann and Casella 1998) [5]. The method has many applications in many different domains, such as linear models and generalized models, econometrics, psychometrics, data modeling in nuclear and particle physics just to mention a few. The parameter estimation is a vast field of its own and the method can be complex and computer intensive. Here we only mention the method for now.

## 2.6 Chapter Summary

In this chapter we built an array of distributions that can be used to analysise financial data and construct models.

We saw that reasonable fits can be obtained for the log returns, Eq. (2.1), using a Student distribution, Eq. (2.8), for two different data sets namely NASDAQ and S&P500.

[5] After moving from Cambridge University, in 1957, R. A. Fisher spent some time as a senior research fellow at the CSIRO in Adelaide, Australia. R. A. Fisher spent the rest of his life in Adelaide.

This was possible only when $\delta t = 1$ and we saw that as we increased $\delta t$ we lost reliability in the fits, therefore suggesting that for $\delta t > 1$ the Student distribution cannot be used but distribution like the one described in Sec 2.2.3 would probably fit the data much better because of the extra parameters which take into account the skewness of the distribution and its shape as well as scaling factors.

Nevertheless it would be possible to create relatively accurate models based on the student distribution.

Modeling the volatility of empirical data sets is not a simple task, one can either do it via stochastic differential equations or via time series analysis. In this chapter we used time series analysis to see if we were able to obtain reasonable fits to the data and we saw that it was possible to obtain a rough fit using a simple combination of ARMA and GARCH models for the mean and variance equation.

In this chapter we also build an array of models that can be used to model the volatility, these include models like assymetric GARCH models and the exponential GARCH models known as EGARCH$(p, q)$ or GJR $-$ GARCH$(p, q)$.

In the next chapter we briefly review Brownian motion as one of the fundamental building blocks of stochastic calculus. These ideas will be used in Chapter 4.

# Chapter 3

# Brownian Motion

IN this chapter we define Brownian motion and develop its basic properties. This chapter is purely written as a review chapter and is here for completeness. Since it represents the foundational basis of stochastic calculus. One of the most important properties of unbiased Brownian motion is that it is a martingale and that it accumulates quadratic variation at rate of one unit per time. It is this notion that makes stochastic calculus different from the ordinary calculus.

The construction of Brownian motion is based on the idea of a random walk. In the following subsections we explain how random walks are constructed and describe their properties. Our treatment of Brownian motion follows that of Øksendal (2003), Karatzas and Shreve (1988), and Shreve (2004).

## 3.1 Random Walks

### 3.1.1 Symmetric Random Walks

A good starting point for creating Brownian motion is with a symmetric *random walk*. A typical path of such walk is shown in Fig. (3.1). A symmetric random walk is con-

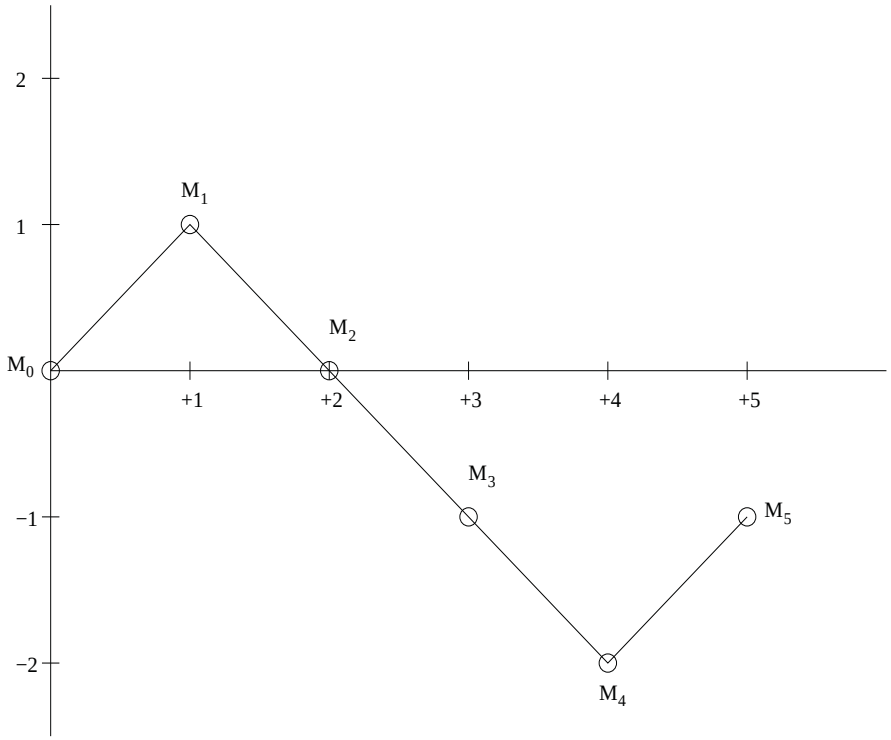


**Figure 3.1. Random walks.** Typical trajectory for a symmetric random walk. The trajectory is evolving through time.

structed by repeatedly tossing a fair coin. On each toss, each of which have equal probability, the probability of a H is $p$ and $q = 1 - p$ for a T with $p = q = 1/2$. If we denote the successive outcomes of the tosses by $w = w_1 w_2 w_3 \cdots = \prod_{i=1}^{\infty} w_i$, with $w_n$ being the outcome of the $n^{th}$ toss and we let

$$X_i = \begin{cases} +1 & \text{if } w_i = \text{H} \\ -1 & \text{if } w_i = \text{T}. \end{cases} \tag{3.1}$$

with $M_0 = 0$, then the process $M_k$

$$M_k = \sum_{i=1}^{k} X_i \quad \text{for} \quad \text{k} = 1, 2, \cdots, \tag{3.2}$$

is a symmetric random walk mean 0 and variance 1, that is $\mathrm{E}[X_i] = 0$ and $\mathrm{Var}[X_i] = \mathrm{E}[X_i^2] = 1$. Random walks have independent increments. What this means is that if we choose a $\{k\}_{i=0}^{m} \in N^+$ such that $k_i < k_{i+1}$ with $k_0$, the random variables

$$M_{k_1} = (M_{k_1} - M_{k_0}), (M_{k_2} - M_{k_1}), \cdots, (M_{k_m} - M_{k_{m-1}}) \tag{3.3}$$

are independent variables,

$$(M_{k_{i+1}} - M_{k_i}) = \sum_{j=k_i+1}^{k_{i+1}} X_j, \tag{3.4}$$

and each of these are called an increment of the random walk, having an expected value and variance

$$\mathrm{E}[M_{k_{i+1}} - M_{k_i}] = \sum \mathrm{E}[X_j] = 0 \tag{3.5}$$

$$\mathrm{Var}[M_{k_{i+1}} - M_{k_i}] = \sum_{j=k_i+1}^{k_{i+1}} \mathrm{Var}[X_j] = \sum_{j=k_i+1}^{k_{i+1}} 1 = k_{i+1} - k_i \tag{3.6}$$

respectively. Which means that the variance of the symmetric random walk over any time interval $k \in N^+$ to $l \in N^+$ is $l - k$.

The symmetric random walk is also a *martingale*[6], because if we choose $k \in N^+$ and $l \in N^+$ such that $k < l$ and compute the conditional expectation

$$\begin{aligned} \mathrm{E}[M_l|M_k] &= \mathrm{E}[(M_l - M_k) + M_k|\mathcal{F}_k] \\ &= \mathrm{E}[(M_l - M_k)|\mathcal{F}_k] + \mathrm{E}[M_k|\mathcal{F}_k] \\ &= \mathrm{E}[(M_l - M_k)] + M_k \\ &= M_k. \end{aligned} \tag{3.7}$$

as defined in Appendix A.2.1.

Finally we consider the *quadratic variation* of the symmetric random walk that is defined up to time $k$ as

$$[M, M]_k = \sum_{j=1}^{k} (M_j - M_{j-1})^2 = k, \tag{3.8}$$

which is computed along a path by taking all the one step increments $M_j - M_{j-1}$ (these are equal to $X_j$, which is either 1 or -1, depending on the path), squaring these increments, and then summing them. Since $(M_j - M_{j-1})^2 = 1$, for all $M_j - M_{j-1} = \pm 1$ the sum in Eq. (3.8) is $k$.

### 3.1.2 Scaled Symmetric Random Walks

To approximate Brownian motion, we fix a positive integer $n$ and define the *scaled symmetric random walk*

$$W^{(n)}(t) = \frac{1}{\sqrt{n}} M_{nt}, \quad \text{for} \quad nt \geq 0. \tag{3.9}$$

[6] A martingale is a stochastic process (i.e., a sequence of random variables) such that the conditional expected value of an observation at some time $t$, given all the observations up to some earlier time $s$, is equal to the observation at that earlier time $s$, as shown in Eq. (3.7), see Appendix A.2.1. Historically the name *martingale* first appeared in *Le Dictionaire de l'Académie Française* in 1762 in the fourth edition. The definition was properly defined in the sixth edition and it stated "To play a martingale, it is to play always everything that we have lost". An origin which is believed to be true (Mansuy 2005), appears to be derived from a Provençal expression "*jouga a la martegalo*", which means to play in an incomprehensible and absurd way. Having established a plausible root for the expression we can trace its origin back by studying further *Le Dictionaire Provençal* from Mistral (1979). The word *martegalo* is related as well to the people who lived in a village called Martigue to whom a particular "naivety" or "naive curiosity" was attributed. *Le Martigue* refers to the Lac of Berre that later gave birth to the city on the 21st of April 1581 by joining three villages, which were on the edges of the delta of the *Golfe de Fos*.

If $nt$ is not an integer, $W^{(n)}(t)$ is defined by linear interpolation between its values at the nearest points $s$ and $u$, at the left and right of $t$ at which $ns$ and $nu$ are integers. Brownian motion is obtained in the limit as $n \to \infty$.

Like the symmetric random walk, the scaled random walk has independent increments. That is for $\{t_i < t_{i+1} | t_0 = 0\}_{i=0}^{m} \in N^+$ such that each $nt \in N$ we have

$$W^{(n)}(t_1) - W^{(n)}(t_0), \cdots, W^{(n)}(t_m) - W^{(n)}(t_{m-1}) \tag{3.10}$$

are independent with expected value and variance

$$\begin{aligned} \mathrm{E}[W^{(n)}(t) - W^{(n)}(s)] &= 0 & (3.11)\\ \mathrm{Var}[W^{(n)}(t) - W^{(n)}(s)] &= t - s. & (3.12) \end{aligned}$$

The symmetric random walk the scaled random walk is also a martingale, because if we let $0 \leq s \leq t$ and write $(W^{(n)}(t) - W^{(n)}(s)) + W^{(n)}(s)$ and take the conditional expectation value, with respect to the filtration $\mathcal{F}_s$, we would find that

$$\mathrm{E}[W^{(n)}(t) | \mathcal{F}(s)] = W^{(n)}(s). \tag{3.13}$$

For the symmetric random walk we consider the quadratic variation of scaled random walk. For $t \geq 0$ such that $nt$ is an integer

$$\begin{aligned} [W^{(n)}(t), W^{(n)}(t)] &= \sum_{j=1}^{nt} \left[ W^{(n)}\left(\frac{j}{n}\right) - W^{(n)}\left(\frac{j-1}{n}\right) \right]^2 \\ &= \sum_{j=1}^{nt} \left[ \frac{1}{\sqrt{n}} M_{n\frac{j}{n}} - \frac{1}{\sqrt{n}} M_{n\frac{j-1}{n}} \right]^2 \\ &= \sum_{j=1}^{nt} \left[ \frac{1}{\sqrt{n}} X_j \right]^2 = \sum_{j=1}^{nt} \frac{1}{n} = t. \end{aligned} \tag{3.14}$$

This is also evaluated path by path in the limit as $n \to \infty$. The distribution of a scaled random walked $W^{(n)}(t)$ is evaluated as time $t$ converges to the normal distribution with mean 0 and variance 1.

## 3.2 Brownian Motion

We obtain Brownian motion as the limit of the scaled random walks $W^{(n)}(t)$, Eq. (3.9), as the limit $n \to \infty$. Brownian motion inherits the properties of the symmetric random walk, and typical sample paths are illustrated in Fig. 3.2. Formally Brownian motion is defined as follows

**Definition 3.2.1** *Let $(\Omega, \mathcal{F}, P)$ be a probability space. For each $w \in \Omega$ suppose there is a continuous function $W(t)$ of $t \geq 0$ that satisfies $W(0) = 0$ and that depends on $w$, then $W(t)$, for $t \geq 0$ is Brownian motion if for all $0 = t_0 < t_1 < t_2 < \cdots < t_m$ the increments*

$$W(t_1) - W(t_0), \cdots, W(t_m) - W(t_{m-1}) \tag{3.15}$$

*are independent and each of these increments are normally distributed with*

$$\mathrm{E}[W(t_{i+1}) - W(t_i)] = 0, \tag{3.16}$$

$$\mathrm{Var}[W(t_{i+1}) - W(t_i)] = t_{i+1} - t_i. \tag{3.17}$$

In Definition 3.2.1, $w$ should be thought of as the Brownian motion path—that is a random experiment is performed and its outcome is the path of the Brownian motion. Then $W(t)$ is the value of this path at time $t$ that depends on which path resulted from the random experiment.

The distribution of the Brownian motion can be summarized by the following theorem which we only state and do not prove, since the proof of this theorem may be found in Shreve (2004).

**Theorem 3.2.2 (Brownian motion)** *Let $(\Omega, \mathcal{F}, P)$ be a probability space. For each $w \in \Omega$ suppose there is a continuous function $W(t)$ of $t \geq 0$ that satisfies $W(0) = 0$ and that depends on $w$. The following three properties are equivalent*

1. *For all $0 = t_0 < t_1 < t_2 < \cdots < t_m$ the increments*

$$W(t_1) - W(t_0), \cdots, W(t_m) - W(t_{m-1}) \tag{3.18}$$

*are independent and each of these increments are normally distributed with*

$$\mathrm{E}[W(t_{i+1}) - W(t_i)] = 0, \tag{3.19}$$

$$\mathrm{Var}[W(t_{i+1}) - W(t_i)] = t_{i+1} - t_i. \tag{3.20}$$

2. *For all $0 = t_0 < t_1 < t_2 < \cdots < t_m$ the random variables $W(t_1), W(t_2) \cdots W(t_m)$ are jointly normally distributed with means equal to zero and covariance matrix*

$$\begin{pmatrix} \mathrm{E}[W^2(t_1)] & \mathrm{E}[W(t_1)W(t_2)] & \cdots & \mathrm{E}[W(t_1)W(t_m)] \\ \mathrm{E}[W(t_2)W(t_1)] & \mathrm{E}[W^2(t_2)] & \cdots & \mathrm{E}[W(t_2)W(t_m)] \\ & & & \\ \mathrm{E}[W(t_m)W(t_1)] & & \cdots & \mathrm{E}[W^2(t_m)] \end{pmatrix} = \begin{pmatrix} t_1 t_1 \cdots t_1 \\ t_1 t_2 \cdots t_2 \\ \\ t_1 t_2 \cdots t_m \end{pmatrix} \tag{3.21}$$

3. *For all* $0 = t_0 < t_1 < t_2 < \cdots < t_m$ *the random variables* $W(t_1), W(t_2) \cdots W(t_m)$ *have joint moment–generating function*

$$\begin{aligned} \varphi(u_1, u_2, \cdots, u_m) &= \mathrm{E}[\exp(u_1 W(t_1) + u_2 W(t_2) + \cdots + u_m W(t_m))] \quad (3.22) \\ &= \exp\left\{ \frac{1}{2}\left(\sum_{i=1}^{m} u_i\right)^2 t_1 + \frac{1}{2}\left(\sum_{i=2}^{m} u_i\right)^2 (t_2 - t_1) + \cdots \right. \\ &+ \left. \frac{1}{2}\left(\sum_{i=m-1}^{m} u_i\right)^2 (t_{m-1} - t_{m-2}) + \frac{1}{2} u_m^2 (t_m - t_{m-1}) \right\} \end{aligned}$$

*if any of 1., 2. or 3. holds (and hence they all hold) then* $W(t)$*, for* $t \geq 0$*, is classed as Brownian motion.*

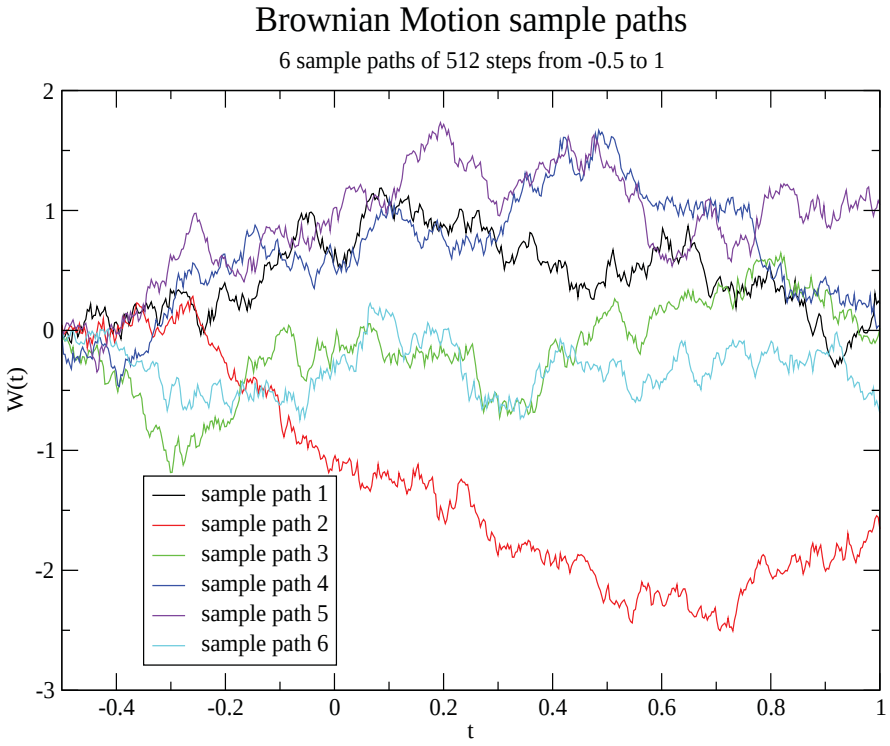


**Figure 3.2. Sample paths for Brownian motion.** Typical trajectories for the Brownian motion. Here each sample path is generated from a different random sequence.

The source code that generates these sample paths for Brownian motion can be found in Appendix E.3.5, the Gaussian random numbers were generated using the Box–Muller method (see Appendix A.8.2) and implemented in the Appendix E.3.12.

### 3.2.1 Filtration for Brownian Motion

In addition to the Brownian motion itself some notation for the amount of information available at each time is needed. This is done with a filtration and the following definition

**Definition 3.2.3 (Filtration for the Brownian motion)** *Let $(\Omega, \mathcal{F}, P)$ be a probability space on which is defined a Brownian motion $W(t)$ of $t \geq 0$. Filtration for Brownian motion is a collection of $\sigma$–algebra, $\mathcal{F}(t)$, $t \geq 0$, satisfying*

1. *(Information accumulates). For $0 \leq s \leq t$, every set in $\mathcal{F}(s)$ is also in $\mathcal{F}$. In other words there is at least as much information available at later time $\mathcal{F}(t)$ as there is at the earlier time $\mathcal{F}(s)$.*
2. *(Adaptivity). For each $t > 0$, Brownian motion $W(t)$ at time $t$ is $\mathcal{F}(t)$–measurable. In other words, the information available at time $t$ is sufficient to evaluate Brownian motion $W(t)$ at that time.*
3. *(Independence of future increments). For $0 \leq t \leq u$ the increments $W(u) - W(t)$ are independent of $\mathcal{F}(t)$. In other words, any increment of Brownian motion after time $t$ is independent of the information available at time $t$. Let $\Delta(t)$, $t \geq 0$, be a stochastic process. We say $\Delta t$ is adapted to the filtration $\mathcal{F}(t)$ if for each $t \geq 0$ the random variable $\Delta(t)$ is $\mathcal{F}(t)$–measurable.*

Properties 1 and 2 in the definition guarantee that the information available at each time $t$ is at least as much as one would learn from observing the Brownian motion up to time $t$. Property 3 says that this information is of no use for predicting future movements of Brownian motion. In asset pricing models we build, property 3 leads to the efficient market hypothesis.

Finally, one of the properties of unbiassed Brownian motion is that it is a martingale, because if we let $0 \leq s \leq t$ then applying the same reasoning as in the symmetric random walk case we see that we have the same result as in Eq. (3.13). A second property of Brownian motion is that it is also a Markov process—that is we have the following theorem,

**Theorem 3.2.4** *Let $W(t)$, for $t \geq 0$, be Brownian motion and let $\mathcal{F}(t)$ be a filtration for this Brownian motion. Then $W(t)$ is a Markov–process.*

Furthermore the transition probabilities for Brownian motion is Gaussian distributed. We will omit the proof of this theorem as it can be found in Shreve (2004).

The formalism developed above is the foundation of stochastic calculus and will be used during the rest of this thesis. Brownian motion appears in every financial model that is built using stochastic calculus. These models appear later in this thesis, so it is important to understand all of the properties of Brownian motion before trying to make any attempt in solving these models.

In the chapter on stochastic calculus, Chapter 4, we go through, in more detail, some of the concepts used in stochastic calculus.

### 3.2.2 Quadratic Variation

We would like to count the number of times a function oscillates up or down between times 0 and $T$ with the down moves adding rather than subtracting. Let us start with first order variation of a function, $f(t)$, up to time $T$ we choose on an interval $[0, T]$, some partition $\Pi = \{t_0, t_1, \ldots, t_n\}$ as in Theorem 3.2.2. If we denote the maximum step size of the partition by $||\Pi|| = \max_{j=0,\cdots,n-1}(t_{j+1} - t_j)$ then we define the first order variation as

$$V_T(f) = \lim_{||\Pi||\to 0} \sum_{j=1}^{n-1} \left| f(t_{j+1}) - f(t_j) \right|. \tag{3.23}$$

The limit here is taken so that the number $n$ goes to infinity and the length of the longest step size go to zero. Using the mean value theorem from ordinary calculus, which applies to any functions $f(t)$ which is differentiable everywhere denoted by $f'(t)$, we can obtain an expression for the derivative at a point $t_j^\star$ on a sub-interval $[t_j, t_{j+1}]$ such that

$$f'(t_j^\star) = \frac{f(t_{j+1}) - f(t_j)}{t_{j+1} - t_j} \implies f(t_{j+1}) - f(t_j) = f'(t_j^\star)(t_{j+1} - t_j). \tag{3.24}$$

This is a Riemann sum for the integral of the function $|f'(t)|$, therefore

$$V_T(f) = \lim_{||\Pi||\to 0} \sum_{j=1}^{n-1} \left| f(t_{j+1}) - f(t_j) \right| \tag{3.25}$$

$$= \lim_{||\Pi||\to 0} \sum_{j=1}^{n-1} \left| f'(t_j^\star) \right| (t_{j+1} - t_j) = \int_0^T |f'(t)| dt. \tag{3.26}$$

Equation, Eq. (3.25), defines the first order variation of a function on an interval $[0, T]$. The quadratic variation of a function can be summarized in the following definition

**Definition 3.2.5 (Quadratic variation)** *Let $f(t)$ be a function defined for $0 \leq t \leq T$. The quadratic variation of $f$ up to time $T$ is*

$$[f,f](T) = \lim_{||\Pi||\to 0} \sum_{j=1}^{n-1} \left[f(t_{j+1}) - f(t_j)\right]^2 \tag{3.27}$$

*where $\Pi = \{t_0, t_1, \ldots, t_n\}$ and $0 = t_0 < t_1 < \cdots < t_n = T$.*

The quadratic variation is 0 when $f(t)$ is a continuous function, but the Brownian path on the other hand is not differentiable anywhere with respect to time. Hence the quadratic variation is not zero.

We want to compute the quadratic variation for the Brownian path $W(t)$ that is,

$$Q_\Pi = \sum_{j=0}^{n-1} \left(W(t_{j+1}) - W(t_j)\right)^2, \tag{3.28}$$

which is summed in quadrature. To do this we employ the expected value and its variance. These are calculated as follows,

$$\mathrm{E}\left[\left(W(t_{j+1}) - W(t_j)\right)^2\right] = \mathrm{Var}[\left(W(t_{j+1}) - W(t_j)\right)] = t_{j+1} - t_j. \tag{3.29}$$

We therefore get for Eq. (3.28),

$$\begin{aligned} \mathrm{E}\left[Q_\Pi\right] &= \mathrm{E}\left[\sum_{j=0}^{n-1} \left(W(t_{j+1}) - W(t_j)\right)^2\right] \\ &= \sum_{j=0}^{n-1} \left(t_{j+1} - t_j\right) \\ &= (t_1 - t_0) + (t_2 - t_1) + \cdots + (t_n - t_{n-1}) = T. \end{aligned} \tag{3.30}$$

Moreover the variance can also be calculated, this is carried out as follows

$$\begin{aligned} \mathrm{Var}\left[\left(W(t_{j+1}) - W(t_j)\right)^2\right] &= \mathrm{E}\left[\left(\left(W(t_{j+1}) - W(t_j)\right)^2 - \mathrm{E}\left[\left(W(t_{j+1}) - W(t_j)\right)\right]\right)^2\right] \\ &= \mathrm{E}\left[\left(\left(W(t_{j+1}) - W(t_j)\right)^2 - \left(t_{j+1} - t_j\right)\right)^2\right] \\ &= \mathrm{E}\left[\left(W(t_{j+1}) - W(t_j)\right)^4\right] \\ &- 2\left(t_{j+1} - t_j\right)\mathrm{E}\left[\left(W(t_{j+1}) - W(t_j)\right)^2\right] + \left(t_{j+1} - t_j\right)^2. \end{aligned} \tag{3.31}$$

The Brownian path is normally distributed so the fourth moment, which is the kurtosis, is given by

$$\mathrm{E}\left[\left(W(t_{j+1}) - W(t_j)\right)^4\right] = 3\left(t_{j+1} - t_j\right)^2. \tag{3.32}$$

Hence the variance takes the form

$$\mathrm{Var}\left[\left(W(t_{j+1}) - W(t_j)\right)^2\right] = 2\left(t_{j+1} - t_j\right)^2. \tag{3.33}$$

We can now get an expression for the variance of $Q_\Pi$ and take the limit as the partition goes to zero, that is

$$\begin{aligned}\lim_{||\Pi||\to 0} \mathrm{Var}\left[Q_\Pi\right] = \lim_{||\Pi||\to 0} \sum_{j=0}^{n-1} 2\left(t_{j+1} - t_j\right)^2 &\leq \lim_{||\Pi||\to 0} \sum_{j=0}^{n-1} 2||\Pi||\left(t_{j+1} - t_j\right) \\ &= \lim_{||\Pi||\to 0} 2||\Pi||T = 0. \end{aligned} \tag{3.34}$$

We therefore have

$$\mathrm{E}\left[Q_\Pi\right] = T \tag{3.35}$$

$$\mathrm{Var}\left[Q_\Pi\right] = 0. \tag{3.36}$$

The above result can be inserted into the following theorem:

**Theorem 3.2.6 (Quadratic variation for the Brownian motion)** *Let $W(t)$ be a Brownian motion, then $[W, W](T) = T$ for all $T \geq 0$ almost surely.*

As a consequence of Theorem 3.2.6 we can write

$$dW(t)dW(t) = dt. \tag{3.37}$$

Furthermore we may compute the cross variation of $W(t)$ with $t$, which is

$$\lim_{||\Pi||\to 0} \sum_{j=0}^{n-1} \left(W(t_{j+1}) - W(t_j)\right)\left(t_{j+1} - t_j\right). \tag{3.38}$$

To calculate this limit we observe that

$$\begin{aligned}\lim_{||\Pi||\to 0} \left| \sum_{j=0}^{n-1} \left(W(t_{j+1}) - W(t_j)\right)\left(t_{j+1} - t_j\right) \right| &\leq \lim_{||\Pi||\to 0} \left\{ \max_{0\leq k\leq n-1} \left|W(t_{j+1}) - W(t_j)\right| T \right\} \\ &= \lim_{||\Pi||\to 0} ||\Pi|| \left|W(t_{j+1}) - W(t_j)\right| \\ &= 0. \end{aligned} \tag{3.39}$$

We therefore have for the cross variation

$$\lim_{||\Pi||\to 0} \sum_{j=0}^{n-1} \left(W(t_{j+1}) - W(t_j)\right)\left(t_{j+1} - t_j\right) = 0, \tag{3.40}$$

| $\times$ | $dW(t)$ | $dt$ |
|---|---|---|
| $dW(t)$ | $dt$ | 0 |
| $dt$ | 0 | 0 |

**Table 3.1. Brownian motion multiplication table.** The table for the Brownian motion multiplication.

which we will write as

$$dW(t)dt = 0 = dtdW(t). \tag{3.41}$$

Now if we consider the cross variation between $dt$ and $dt$, that is the limit

$$\begin{aligned} \lim_{||\Pi||\to 0} \sum_{j=0}^{n-1} \left(t_{j+1} - t_j\right)^2 &\leq \lim_{||\Pi||\to 0} \left\{ \max_{0\leq k\leq n-1} \left(t_{k+1} - t_k\right) \sum_{j=0}^{n-1} \left(t_{j+1} - t_j\right) \right\} \\ &= \lim_{||\Pi||\to 0} ||\Pi|| T = 0, \end{aligned} \tag{3.42}$$

hence here too we have

$$dt\, dt = 0. \tag{3.43}$$

These are very useful properties of the Brownian motion, these properties will simplify calculations a great deal and will be used right through the rest of this work. We will refer to these properties as the Brownian motion multiplication rule. These rules are summarized in Table 3.1

## 3.3 Chapter Summary

This chapter has summarized the main properties of Brownian motion, as well as a few of the important properties associated with it. These properties represent the building blocks of stochastic calculus and will be used for the rest of this work through all of the remaining chapters.

In the next chapter, Chapter 4, we define Îto calculus from the ground up and apply its rules and properties to examples in the context of finance. In particular we are interested in modeling volatility, such models are known as stochastic volatility models. We also explicitly write out the equations used solve these stochastic differentials. These equations were derived many years ago and usually carry the name of the people who

have discovered them. For example the Îto –Doeblin equation, the Feynman–Kac formula, and the Kolmogorov equations. The last two are used to calculate the transition probabilities of the system, which relates stochastic differential equation to partial differential equations. These are very important as they offer access to a solution, whereas in many cases stochastic differential equations cannot be explicitly calculated.

# Chapter 4

# Stochastic Calculus

IN this chapter we review some aspects of stochastic calculus by describing its fundamental properties and this is not part of the original contribution of this thesis as stochastic calculus was developed several decades ago. The content of this chapter includes Îto calculus and its integral, which then leads to a stochastic differential equation. We also describe the different numerical methods that are used to approximate these, sometime complicated, equations and the partial differential equations that are used to relate stochastic differential equations and partial differential equations.

As already discussed in Section 1.2.3, stochastic calculus is an alternative approach to binomial trees. Note that Eq. (1.16) is usually called an Îto process that is a general solution to the differential stochastic equation, Eq. (1.15), normally called an Îto stochastic differential equation. There is only a small set of these equations that are solvable explicitly, in general one has to use numerical methods to approximate such equation.

In this chapter we review the Îto process, the stochastic differential equations and then show some numerical approximations schemes for these stochastic differential equations.

## 4.1 Îto Calculus

In this section we define Îto integrals and develop their properties. These are used to model value of a portfolio that results from trading strategies of asset in continuous time. Îto calculus is used to manipulate these stochastic integrals is based on the Îto-Doeblin formula, which is different from ordinary calculus. The difference comes from the fact that Brownian motion has non–zero quadratic variation.

### 4.1.1 Îto Integral for Simple Process

In this section we try to make sense of the integral

$$\int_0^t \Delta(s)dW(s) \quad \text{for} \quad t > 0. \tag{4.1}$$

Here $W(t)$, $t \geq 0$, is a Brownian motion which is also known as a Wiener process, together with a filtration process $\mathcal{F}_t$, $t \geq 0$, for this Brownian motion. Here we assume that $\Delta(t)$ is an adapted process (see Def. A.3.1), because later we will see that $\Delta(t)$ is the value we take for an asset at time $t$, which usually depends on the price path of the asset at and up to time $t$.

In this case, because the Brownian motion path cannot be differentiated with respect to time, the integral in Eq. (4.1) cannot be treated as an ordinary Lebesgue integral. The Îto integral is defined as follows:

Let $\Pi = \{t_0, \cdots, t_n\}$ be a partition of $[0, T]$ such that $0 = t_0 < t_1 < \cdots < t_n = T$. Assume that $\Delta(t)$ is constant in $t$ on each sub-interval $[t_j, t_{j+1}]$. Such a process can be

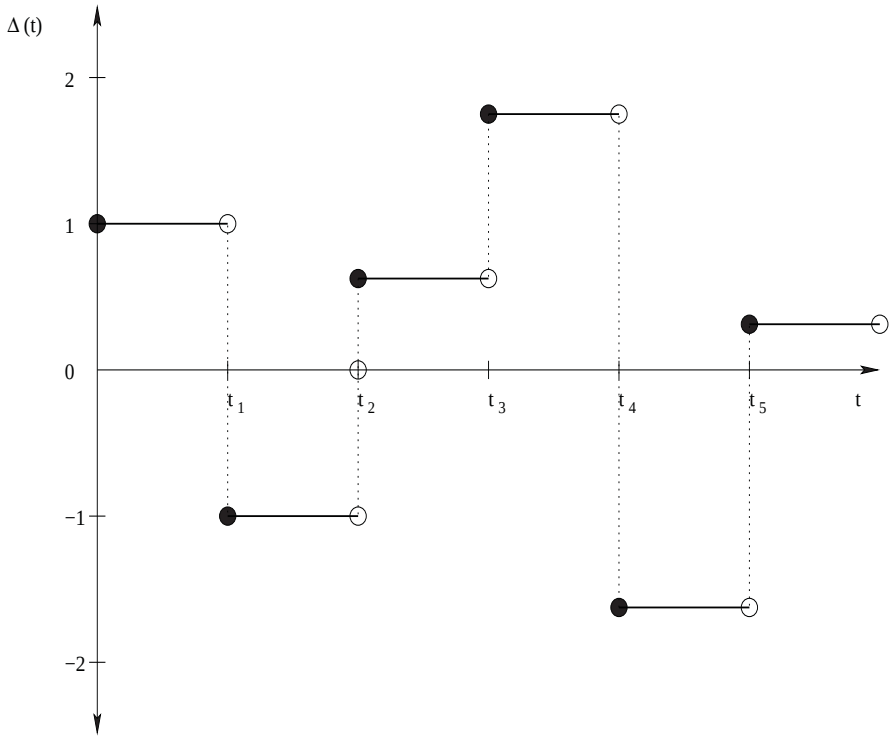


**Figure 4.1. Simple Îto process.** Simple Îto process path on a given sub-interval. We can think of the $t_i$ as the trading date for example and $\Delta(t_i)$ as the value of a given asset at the time $t_i$. The process is clearly discountinous as it evolves over true time. Here the black dots represent the points that are included in the interval while the open circles are not included in the interval.

viewed in Fig. 4.1. One way to think about the interplay between the simple process $\Delta(t)$ and $W(t)$, is to regard $W(t)$ as the price per share of an asset at time $t$. The $t_i$ as the trading dates in the asset and $\Delta(t_0), \cdots, \Delta(t_{n-1})$ as the position taken in the asset at each trading dates and held fix to the next trading date.

The gain from trading at each time $t$ is given by

$$\begin{aligned} I(t) &= \Delta(t_0)[W(t) - W(t_0)] = \Delta(0)W(t) \quad \text{for} \quad 0 \le t \le t_1 \\ I(t) &= \Delta(0)W(t) + \Delta(t_1)[W(t) - W(t_1)] \quad \text{for} \quad t_1 \le t \le t_2 \\ I(t) &= \Delta(0)W(t) + \Delta(t_1)[W(t) - W(t_1)] + \Delta(t_2)[W(t) - W(t_2)] \quad \text{for} \quad t_3 \le t \le t_3 \end{aligned}$$

and so on. In general, if $t_k \le t \le t_{k+1}$, then

$$I(t) = \sum_{j=0}^{k-1} \Delta(t_j)[W(t) + W(t_j)] + \Delta(t_k)[W(t) - W(t_k)]. \tag{4.2}$$

The process $I(t)$ in Eq. (4.2) is the Îto integral of the simple process $\Delta(t)$.

We now state some properties of the Îto integral without proof, as these proofs may found many books such as in Shreve (2004), Karatzas and Shreve (1988), Øksendal (2003), or Kloeden and Platen (1992).

**Theorem 4.1.1 (Martingale)** *The Îto integral $I(t)$, Eq. (4.2), is a martingale, that is*

$$\mathrm{E}[I(t)|\mathcal{F}_s] = I(s). \tag{4.3}$$

The Îto integral also has the isometric property,

**Theorem 4.1.2 (Îto isometry)** *The Îto integral, Eq. (4.2), satisfies*

$$\mathrm{E}[I^2(t)] = \mathrm{E}\left[\int_0^t \Delta^2(s)ds\right]. \tag{4.4}$$

The theorem on quadratic variation is another important one. The theorem is stated as follows:

**Theorem 4.1.3 (Quadratic variation)** *The quadratic variation accumulates up to time t by the Îto integral, Eq. (4.2), is*

$$[I, I](t) = \int_0^t \Delta^2(s)ds. \tag{4.5}$$

In the last two theorems we see how the quadratic variation and the variance of a process can differ. The quadratic variation is computed path by path and the result can depend on the path. The size of the quadratic variation directly depends on the size of the position. Here $\Delta(s)$ can be regarded as a risk measure.

The variance of $I(t)$ is an average over all possible paths of the quadratic variation—it is therefore usually considered more as a theoretical concept than the quadratic variation. Moreover the variance here differs from the empirical variance, which is captured directly from the data and used as an estimator. Another useful property of the Îto integral is that the square of the differential is given by

$$dI(t)dI(t) = \Delta^2(t)dW(t)dW(t) = \Delta^2(t)dt. \tag{4.6}$$

In the last equality we have used the property of the Brownian motion for quadratic variation, Eq. (3.1), that is $dW(t)dW(t) = dt$.

Having defined the Îto integral for constant integral, we now move on to the more general case where $\Delta(t)$ is allowed to vary continuously or to jump.

### 4.1.2 Îto Integral for non Simple Process

In this subsection we define the Îto integral, Eq. (4.1), for a non–simple process, that is when $\Delta(t)$ is allowed to vary continuously in time with possible jumps. Such a process can be visualized in Fig.(4.1). Here we assume that $\Delta(t)$, $t \geq 0$, is an adapted process so

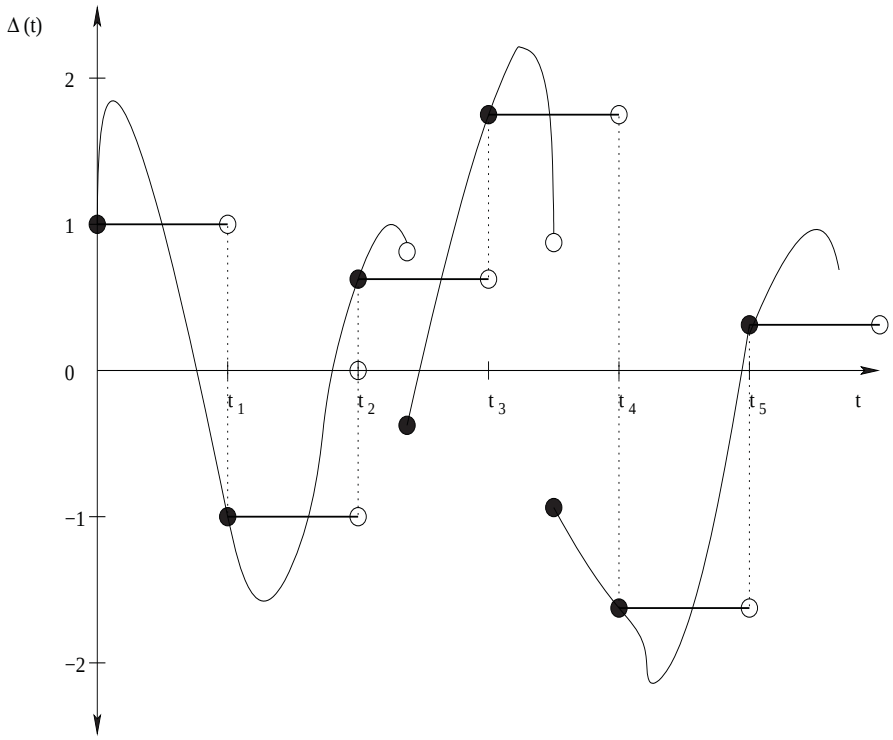


**Figure 4.2. Non–simple Îto process.** Non–simple Îto process path on a given sub-interval. This is the same as Fig. 4.1, but this time when the process is continous in time.

that the $\Delta(t)$ is square integrable. That is, the expected value of the squared integrand is finite.

Using simple processes on a given partition, typically $[0, T]$, it is possible to approximate the Îto integral for general integrands. As was carried out in the simple case, the idea is to set the approximating simple process equal to $\Delta(t_j)$ at each $t_j$ and then holding the process constant over the sub-interval $[t_j, t_{j+1}]$. In the limit of the step size approaching zero, the approximating integrand tends to the continuously varying integrand.

Generally, it is possible to choose $\Delta_n(t)$ of simple process such that as $n \to \infty$ these processes converge to the continuously varying $\Delta(t)$, i.e,

$$\lim_{n\to\infty} \mathrm{E}\left[\int_0^t |\Delta_n(t) - \Delta(t)|^2 \, dt\right] = 0. \tag{4.7}$$

Then for each $\Delta_n(t)$, the Îto integrand is defined for $0 < t < T$, Eq. (4.1). So for the general integrand, which varies continuously, the Îto integral is defined by

$$\int_0^t \Delta(s)dW(s) = \lim_{n\to\infty} \int_0^t \Delta_n(s)dW(s) \quad \text{for} \quad 0 \leq t \leq T. \tag{4.8}$$

This integral inherits all the properties of the Îto integral when the integrand is simple, as defined in Section 4.1.1. All of the properties can be summarized in the following theorem,

**Theorem 4.1.4 (Îto integral)** *Let $T$ be a positive constant and let $\Delta(t)$, $0 \geq t \geq T$, be an adapted process, that is square integrable $\mathrm{E}\left[\int_0^t \Delta^2(t)dt\right] < \infty$. Then Eq. (4.7) has the following properties*

1. **(Continuity)** *As a function of the upper limit of integration $t$, the paths of $I(t)$ are continuous.*
2. **(Adaptivity)** *For each $t$, $I(t)$ is $\mathcal{F}_t$–measurable.*
3. **(Linearity)** *If $I(t) = \int_0^t \Delta(s)dW(s)$ and $J(t) = \int_0^t \Gamma(s)dW(s)$ then*

$$I(t) \pm J(t) = \int_0^t (\Delta(s) \pm \Gamma(s))\, dW(s), \tag{4.9}$$

$$cI(t) = \int_0^t c\Delta(s)dW(s). \tag{4.10}$$

4. **(Martingale)** *$I(t)$ is a martingale.*
5. **(Îto isometry)** *The expected value of $\mathrm{E}[I^2(t)] = \mathrm{E}\left[\int_0^t \Delta^2(s)ds\right]$.*
6. **(Quadratic variation)** *$[I, I](t) = \int_0^t \Delta^2(s)ds$.*

Using the above theorem we can show that the integral of Brownian motion takes the form

$$\int_0^t W(s)dW(s) = \frac{1}{2}W^2(t) - \frac{1}{2}[W, W](t) = \frac{1}{2}W^2(t) - \frac{1}{2}t, \tag{4.11}$$

as opposed to the usual Lebesgue integral

$$\int_0^t g(s)dg(s) = \int_0^t g(s)g'(s)ds = \frac{1}{2}g^2(t) \tag{4.12}$$

that we find in ordinary calculus.

So far we have defined the Îto integral and have not said anything about how to evaluate and/or manipulate these integrals. The Îto–Doeblin formula is an essential formula for the manipulation and evaluation of these integrals.

### 4.1.3 Îto –Doeblin Formula

In almost every stochastic differential equation the drift and diffusion are composed of random and/or non–random functions and sometimes a combination of both. We therefore need a rule to differentiate expressions of the form $f(W(t))$ where $f(x)$ is a differential function and $W(t)$ is a Brownian motion.

#### For Brownian Motion

The Îto–Doeblin in differential form is given by

$$df(W(t)) = f'(W(t))dW(t) + \frac{1}{2}f''(W(t))dt. \tag{4.13}$$

Integrating this, we obtain the Îto–Doeblin formula in integral form

$$f(W(t)) - f(W(0)) = \int_0^t f'(W(s))dW(s) + \frac{1}{2}\int_0^t f''(W(s))ds, \tag{4.14}$$

which gives a precise definition for both terms appearing on the right hand side. The first term is an Îto integral and the second one is a Lebesgue integral with respect to time. We can write this into a theorem, which we state without proof, as the proof may be found in Shreve (2004), Karatzas and Shreve (1988), or Øksendal (2003).

**Theorem 4.1.5 (Îto–Doeblin for Brownian motion)** *Let $f(t,x)$ be a function for which the partial derivatives, $f_t(t,x)$, $f_x(t,x)$ and $f_{xx}(t,x)$ are defined and let $W(t)$ be a Brownian motion. Then for every $t \geq 0$*

$$\begin{aligned} f(t,W(t)) &= f(0,W(0)) \\ &+ \int_0^t f_t(s,W(s))ds + \int_0^t f_x(s,W(s))dW(s) + \frac{1}{2}\int_0^t f_{xx}(s,W(s))ds. \end{aligned} \tag{4.15}$$

In general it is better to work with the differential form. Eq. (4.15) may be rewritten in differential form as

$$\begin{aligned} df(t,W(t)) &= f_t(t,W(t))dt \\ &+ f_x(t,W(t))dW(t) + f_{tx}(t,W(t))dt\,dW(t) + \frac{1}{2}f_{xx}(t,W(t)), \end{aligned} \tag{4.16}$$

but because of the multiplication rules on the Brownian motion, Eq. (3.1),

$$dW(t)\,dW(t) = dt, \quad dt\,dW(t) = 0 = dW(t)\,dt = 0, \quad \text{and,} \quad dt\,dt = 0, \tag{4.17}$$

Eq. (4.16) takes the form of

$$df(t,W(t)) = f_t(t,W(t))dt + f_x(t,W(t))dW(t) + \frac{1}{2}f_{xx}(t,W(t)). \tag{4.18}$$

### For the Îto Process

The process for which we develop stochastic calculus is the Îto process, and almost all the processes, except those that have jumps (which fall into the category of Lévy processes) are Îto processes.

**Definition 4.1.6 (Îto processs)** *Let $W(t)$, $t \geq 0$ be a Brownian motion, and let $\mathcal{F}_t$, be an associated filtration. An Îto process is a stochastic process of the form*

$$X(t) = X(0) + \int_0^t \Delta(s)dW(s) + \int_0^t \theta(s)ds, \tag{4.19}$$

*where $X(0)$ is non–random and $\Delta(s), \theta(s)$ are adapted stochastic processes*[7].

This is most easily remembered when Eq. (4.19) is written in differential form

$$dX(t) = \theta(t)dt + \Delta(t)dW(t). \tag{4.20}$$

Using the multiplication rule for the Brownian motion, Eq. (3.1), we can compute

$$dX(t)dX(t) = \theta^2(t)dt\,dt + \theta(t)\Delta(t)dtdW(t) + \Delta(t)\theta(t)dW(t)\,dt + \Delta^2(t)dW(t)dW(t)$$

hence

$$dX(t)dX(t) = \Delta^2(t)dt. \tag{4.21}$$

This says that at each time $t$, the process $X(t)$ is accumulating quadratic variation at the rate of $\Delta^2(t)$ per unit time and hence the total quadratic variation accumulated on the time interval $[0, t]$ is

$$[X, X](t) = \int_0^t \Delta^2(s)ds. \tag{4.22}$$

We can summarize these properties into a more general theorem as we did in the case of the Brownian motion

**Theorem 4.1.7 (Îto–Doeblin formula for the Îto process)** *Let $X(t)$, $t \geq 0$, be an Îto process as defined in Eq. (4.19) and let $f(t, x)$ be a function for which the partial derivatives, $f_t(t, x)$, $f_x(t, x)$ and $f_{xx}(t, x)$ are defined and continuous. Then for every $t \geq 0$*

$$\begin{aligned} f(t, X(t)) &= f(0, X(0)) \\ &+ \int_0^t f_t(s, X(s))ds + \int_0^t f_x(s, X(s))dX(s) + \frac{1}{2}\int_0^t f_{xx}(s, X(s))d[X, X](s). \end{aligned} \tag{4.23}$$

[7] It is assumed that $\mathrm{E}\left[\int_0^t \Delta^2(s)ds\right]$ and $\int_0^t |\theta(s)|\,ds$ are finite for every, $t > 0$ so that the integrals on the right–hand side of Eq. (4.19) are defined and the Îto integral is a martingale.

Substituting Eq. (4.20) and Eq. (4.21) into Eq. (4.23) we see that Eq. (4.23) can be rewritten as

$$\begin{aligned} f(t,X(t)) &= f(0,X(0)) \\ &+ \int_0^t \left[ f_t(s,X(s)) + \theta(s) f_x(s,X(s)) + \frac{1}{2}\Delta^2(s) f_{xx}(s,X(s)) \right] ds \\ &+ \int_0^t \Delta(s) f_x(s,X(s)) dW(s). \end{aligned} \tag{4.24}$$

In Eq. (4.24), there is only one Îto integral and the others are ordinary Lebesgue integrals. However it is easier to remember Eq. (4.24) in its differential form and can be rewritten as

$$df(t,X(t)) = f_t(t,X(t)) + f_x(t,X(t))dX(t) + \frac{1}{2} f_{xx}(t,X(t))dX(t)dX(t). \tag{4.25}$$

Another important theorem is when the Îto integrand is a deterministic function, that is a non–random function.

**Theorem 4.1.8 (Îto integral of a determnistic integrand)** *Let $W(t)$, $t \geq 0$, be a Brownian motion and let $\Delta(t)$ be a non–random function of time. Define $I(t) = \int_0^t \delta(s)dW(s)$. For each $t \geq 0$, the random variable $I(t)$ is normally distributed with expected value of zero and variance $\int_0^t \Delta^2(s)ds$.*

All of these theorems and definitions are used to solve problems involving stochastic processes.

We now briefly mention multivariate stochastic calculus, since it is very useful when one is considering multi–asset models, and stochastic volatility models, which have some mean reversion incorporated in them (Fouque *et al.* 2000).

### 4.1.4 Multivariate Stochastic Calculus

#### Multiple Brownian Motion

A $d$–dimensional Brownian motion is a process

$$\widetilde{W}(t) = (W_1(t), \cdots, W_d(t)), \tag{4.26}$$

which has the properties that each $W_i(t)$ are independent and are a one dimensional Brownian motion. Associated with a $d$–dimensional Brownian motion, we have a filtration $\mathcal{F}_t$, $t \geq 0$, such that the information accumulates. That is for $0 \leq s \leq t$ every

set in $\mathcal{F}_s$ is also in $\mathcal{F}_t$. The future increments are independent, that is for $0 \leq s \leq t$, the vector of increments $\widetilde{W}(s) - \widetilde{W}(t)$ is independent of $\mathcal{F}_t$.

The quadratic variation in the $d$–dimensional case is the same as one dimensional Brownian motion, that is because each component of $W_i$ is independent hence we can write

$$[W_i, W_j](t) = \delta_{ij} t, \tag{4.27}$$

which is written informally as

$$dW_i(t) dW_j(t) = \delta_{ij} dt. \tag{4.28}$$

Here $\delta_{ij}$ is the usual Kronecker delta function, which takes the value of 1 when $i = j$ and 0 otherwise.

As in the one dimensional case we can write The Îto–Doeblin formula for the multidimensional case. Here we only consider the two dimensional case, but in general the formula can be generalized the $d$–dimensional case.

### Îto–Doeblin Formula for the Multiple Process (The $2d$–Case)

To keep the notation simple we consider the case when $d = 2$. Let $X(t)$ and $Y(t)$ be both Îto processes, which means they are processes of the form

$$X(t) = X(0) + \int_0^t \theta_1(s) ds + \int_0^t \sigma_{11}(s) dW_1(s) + \int_0^t \sigma_{12}(s) dW_2(s), \tag{4.29}$$

$$Y(t) = Y(0) + \int_0^t \theta_2(s) ds + \int_0^t \sigma_{21}(s) dW_1(s) + \int_0^t \sigma_{22}(s) dW_2(s), \tag{4.30}$$

where the integrands $\theta_i(t)$ and $\sigma_{ij}(t)$ are assumed to be adapted processes. Eq. (4.29) and Eq. (4.30) may be rewritten in their differential form as

$$dX(t) = \theta_1(t) dt + \sigma_{11}(t) dW_1(t) + \sigma_{12}(t) dW_2(t), \tag{4.31}$$

$$dY(t) = \theta_2(t) dt + \sigma_{21}(t) dW_1(t) + \sigma_{22}(t) dW_2(t), \tag{4.32}$$

or in more compact form as

$$d\widetilde{X}(t) = \widetilde{\theta}(t) dt + \Sigma(t) d\widetilde{W}(t), \tag{4.33}$$

$$\tag{4.34}$$

where

$$d\widetilde{X}(t) = \begin{pmatrix} dX(t) \\ dY(t) \end{pmatrix}, \quad \widetilde{\theta}(t) = \begin{pmatrix} \theta_1(t) \\ \theta_2(t) \end{pmatrix}, \quad \Sigma(t) = \begin{pmatrix} \sigma_{11}(t) & \sigma_{12}(t) \\ \sigma_{21}(t) & \sigma_{22}(t) \end{pmatrix} \quad \text{and}$$

$$d\widetilde{W}(t) = \begin{pmatrix} dW_1(t) \\ dW_2(t) \end{pmatrix}. \tag{4.35}$$

From Eq. (4.31) and Eq. (4.32) we can calculate the quadratic variation for this system of stochastic differential equations, i.e.,

$$\begin{aligned} d\widetilde{X}(t)[d\widetilde{X}(t)]^T &= \begin{pmatrix} dX(t) \\ dY(t) \end{pmatrix} \begin{pmatrix} dX(t) & dY(t) \end{pmatrix} \\ &= \widetilde{\theta}(t)dt\, dt \left[\widetilde{\theta}(t)\right]^T + \widetilde{\theta}(t)dt \left[d\widetilde{W}(t)\right]^T [\Sigma(t)]^T \\ &+ \Sigma(t)d\widetilde{W}(t)\, dt \left[\widetilde{\theta}(t)\right]^T \\ &+ \Sigma(t)d\widetilde{W}(t) \left[d\widetilde{W}(t)\right]^T [\Sigma(t)]^T \\ &= \Sigma(t) \begin{pmatrix} dW_1(t) \\ dW_2(t) \end{pmatrix} \begin{pmatrix} dW_1(t) & dW_2(t) \end{pmatrix} [\Sigma(t)]^T \qquad (4.36) \\ &= \Sigma(t) \begin{pmatrix} dW_1(t)dW_1(t) & dW_1(t)dW_2(t) \\ dW_2(t)dW_1(t) & dW_2(t)dW_2(t) \end{pmatrix} [\Sigma(t)]^T \\ &= \Sigma(t) [\Sigma(t)]^T dt, \quad \text{since} \quad dW_i(t)dW_j(t) = \delta_{ij}dt. \qquad (4.37) \end{aligned}$$

Hence using the multiplication rules for the Brownian motion we obtain the following

$$d\widetilde{X}(t)[d\widetilde{X}(t)]^T = \Sigma(t) [\Sigma(t)]^T dt, \tag{4.38}$$

in matrix form we write

$$\begin{pmatrix} dX(t)dX(t) & dX(t)dY(t) \\ dY(t)dX(t) & dY(t)dY(t) \end{pmatrix} = \begin{pmatrix} \sigma_{11}^2 + \sigma_{12}^2 & \sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22} \\ \sigma_{21}\sigma_{11} + \sigma_{22}\sigma_{12} & \sigma_{21}^2 + \sigma_{22}^2 \end{pmatrix} (t)dt. \tag{4.39}$$

The quadratic variation is then given by

$$[\widetilde{X}, \widetilde{X}] = \int_0^t \Sigma(s) [\Sigma(s)]^T ds. \tag{4.40}$$

The equations may be generalized into a $d$–dimensional case where $\Sigma(t)$ is a $d \times d$ matrix and $d\widetilde{X}(t), \widetilde{\theta}(t)$ and $d\widetilde{W}(t)$ are $d \times 1$ matrices defined in Eq. (4.35) above.

We now write down in a compact form the Îto–Doeblin formula for the 2–dimensional case:

**Theorem 4.1.9 (Two dimensional Îto–Doeblin formula)** *Let $f(t,x,y)$ be a function whose partial derivatives $f_t, f_x, f_{xx}, f_{xy}, f_{yx}$ and $f_{yy}$ are defined and continuous. Let $X(t)$ and $Y(t)$ be Îto processes as in Eq. (4.31) and Eq. (4.32). The two dimensional Îto–Doeblin formula in differential form is*

$$\begin{aligned} df(t,X(t),Y(t)) &= f_t(t,X(t),Y(t))dt + f_x(t,X(t),Y(t))dX(t) + f_y(t,X(t),Y(t))dY(t) \\ &+ f_{xy}(t,X(t),Y(t))dX(t)dY(t) + \\ &+ \frac{1}{2}f_{xx}(t,X(t),Y(t))dX(t)dX(t) + \frac{1}{2}f_{yy}(t,X(t),Y(t))dY(t)dY(t). \end{aligned} \tag{4.41}$$

Where we have assumed that the differential are symmetric and that the commutator of $dX(t)$ and $dY(t)$ is 0, that is $[dX(t), dY(t)] = 0$. Integrating on both sides and substituting Eq. (4.40), Eq. (4.31) Eq. (4.31) we obtain

$$\begin{aligned} f(t,X(t),Y(t)) &= \int_0^t \Big[ f_t(s,X(s),Y(s)) + \theta_1(s) f_x(s,X(s),Y(s)) + \theta_2(s) f_y(s,X(s),Y(s)) \\ &+ \frac{1}{2} f_{xx}(s,X(s),Y(s))(\sigma_{11}^2 + \sigma_{12}^2)(s) \\ &+ f_{xy}(s,X(s),Y(s))(\sigma_{21}\sigma_{11} + \sigma_{22}\sigma_{12})(s) \\ &+ \frac{1}{2} f_{yy}(s,X(s),Y(s))(\sigma_{21}^2 + \sigma_{22}^2)(s) \Big] ds \\ &+ \int_0^t \left[ f_x(s,X(s),Y(s))\sigma_{11}(s) + f_y(s,X(s),Y(s))\sigma_{21}(s) \right] dW_1(s) \\ &+ \int_0^t \left[ f_x(s,X(s),Y(s))\sigma_{12}(s) + f_y(s,X(s),Y(s))\sigma_{22}(s) \right] dW_2(s) \\ &+ f(0,X(0),Y(0)). \end{aligned} \tag{4.42}$$

From Eq. (4.42) we can easily see why it is better to work with the differential form rather than the integral forms. As a final note we point out one of the properties for the product of two Îto differentials:

**Corollary 4.1.10 (Îto product rule)** *Let $X(t)$ and $Y(t)$ be Îto processes. Then*

$$d\left(X(t)Y(t)\right) = Y(t)dX(t) + X(t)dY(t) + dY(t)dX(t). \tag{4.43}$$

This result can be seen to follow from Theorem 4.1.9, Eq. (4.41).

We now turn towards numerical schemes for stochastic differential equations. These schemes can be used to calculate the numerical solution when it is not possible to obtain analytical solutions, which is most often the case, for a given SDE.

## 4.2 Discretization Methods (SDE Numerical Approaches)

The basis of numerical methods for stochastic differential equations (Kloeden and Platen 1992, Burrage *et al.* 2000) lies on the iterated application of the Îto –Doeblin formula for the Îto process, which we shall call the Îto –Doeblin–Taylor expansion.

As we saw earlier, the Îto –Doeblin formula for the Îto process given in **Definition 4.1.6**, by Eq. (4.19),

$$X(t) = X(0) + \int_0^t a(s, X(s))ds + \int_0^t b(s, X(s))dW(s), \tag{4.44}$$

is given by **Theorem 4.1.7** with the following expression Eq. (4.23),

$$\begin{aligned} f(t, X(t)) &= f(0, X(0)) & (4.45)\\ &+ \int_0^t f_t(s, X(s))ds + \int_0^t f_x(s, X(s))dX(s) + \frac{1}{2}\int_0^t f_{xx}(s, X(s))d[X, X](s) \\ &= f(0, X(0)) + \int_0^t \mathcal{L}^0 f(s, X(s))ds + \int_0^t \mathcal{L}^1 f(s, X(s))dW(s) & (4.46) \end{aligned}$$

with the operators $\mathcal{L}^0$ and $\mathcal{L}^1$ acting on the stochastic function $f(t, X(t))$ defined as

$$\begin{aligned} \mathcal{L}^0 f(t, X(t)) &= f_t(t, X(t)) + a(t, X(t))f_x(t, X(t)) + \frac{1}{2}[b(t, X(t))]^2 f_{xx}(t, X(t)) \\ &= f_t + a\, f_x + \frac{1}{2}b^2 f_{xx}, & (4.47)\\ \mathcal{L}^1 f(t, X(t)) &= b(t, X(t))f_x(t, X(t)) = b\, f_{xx}. & (4.48) \end{aligned}$$

Where $f_x$ stands for the partial derivative with respect to $x$, that is $f_x \equiv (\partial/\partial x)f = \partial_x f$.

Let us first consider the case when $a(t, X(t)) \to a(X(t))$ and $b(t, X(t)) \to b(X(t))$. In this case, if we apply Eq. (4.45) to $f(t, X(t)) = a(X(t))$ and $f(t, X(t)) = b(X(t))$. We then obtain the following expression for the general solution for $X(t)$, Eq. (4.44)

$$\begin{aligned} X(t) &= X(0) + \int_0^t \left\{ a(X(0)) + \int_0^{s_2} \mathcal{L}^0 a(X(s_1))ds_1 + \int_0^{s_2} \mathcal{L}^1 a(X(s_1))dW(s_1) \right\} ds_2 \\ &+ \int_0^t \left\{ b(X(0)) + \int_0^{s_2} \mathcal{L}^0 b(X(s_1))ds_1 + \int_0^{s_2} \mathcal{L}^1 b(X(s_1))dW(s_1) \right\} dW(s_2), \end{aligned} \tag{4.49}$$

with the following operators

$$\begin{aligned} \mathcal{L}^0 a(X(t)) &\equiv \mathcal{L}^0 a = a(t, X(t))a_x(X(t)) + \frac{1}{2}[b(X(t))]^2 a_{xx}(X(t)), & (4.50)\\ \mathcal{L}^1 a(X(t)) &\equiv \mathcal{L}^1 a = b(X(t))a_x(X(t)), & (4.51)\\ \mathcal{L}^0 b(X(t)) &\equiv \mathcal{L}^0 b = a(t, X(t))b_x(X(t)) + \frac{1}{2}[b(X(t))]^2 b_{xx}(X(t)), & (4.52)\\ \mathcal{L}^1 b(X(t)) &\equiv \mathcal{L}^1 b = b(X(t))b_x(X(t)). & (4.53) \end{aligned}$$

This expression can be rewritten in a clearer form with the Îto integral in question and a remainder term that contains the rest of the series. This is carried out as follows,

$$X(t) = X(0) + a(X(0)) \int_0^t ds_2 + b(X(0)) \int_0^t dW(s_2) + R, \tag{4.54}$$

with the remainder term

$$\begin{aligned} R &= \int_0^t \int_0^{s_2} \mathcal{L}^0 a(X(s_1)) ds_1 ds_2 + \int_0^t \int_0^{s_2} \mathcal{L}^1 a(X(s_1)) dW(s_1) ds_2 \\ &+ \int_0^t \int_0^{s_2} \mathcal{L}^0 b(X(s_1)) ds_1 dW(s_2) + \int_0^t \int_0^{s_2} \mathcal{L}^1 b(X(s_1)) dW(s_1) dW(s_2). \end{aligned} \tag{4.55}$$

We may repeat this procedure by inserting Eq. (4.45) for a chosen $f$ and produce an expansion with higher order terms plus a reminder term. We observe that when we perform such operation the number of Îto integrals increases according to the number of iterations that have been carried out.

For example if we set $f = \mathcal{L}^1 b$ and apply Eq. (4.45) into Eq. (4.55) we obtain the following expansion,

$$\begin{aligned} R &= \int_0^t \int_0^{s_2} \mathcal{L}^0 a(X(s_1)) ds_1 ds_2 + \int_0^t \int_0^{s_2} \mathcal{L}^1 a(X(s_1)) dW(s_1) ds_2 \\ &+ \int_0^t \int_0^{s_2} \mathcal{L}^0 b(X(s_1)) ds_1 dW(s_2) + \int_0^t \int_0^{s_3} \Big\{ \mathcal{L}^1 b(X(0)) \\ &+ \int_0^{s_2} \mathcal{L}^0 \mathcal{L}^1 b(X(s_1)) ds_1 + \int_0^{s_2} \mathcal{L}^1 \mathcal{L}^1 b(X(s_1)) dW(s_1) \Big\} dW(s_2) dW(s_3). \end{aligned}$$

The solution can then be rewritten as

$$\begin{aligned} X(t) &= X(0) \\ &+ a(X(0)) \int_0^t ds_2 + b(X(0)) \int_0^t dW(s_2) + \mathcal{L}^1 b(X(0)) \int_0^t \int_0^{s_3} dW(s_2) dW(s_3) + \widetilde{R}, \end{aligned} \tag{4.56}$$

where the remainder term is given by

$$\begin{aligned} \widetilde{R} &= \int_0^t \int_0^{s_2} \mathcal{L}^0 a(X(s_1)) ds_1 ds_2 + \int_0^t \int_0^{s_2} \mathcal{L}^1 a(X(s_1)) dW(s_1) ds_2 \\ &+ \int_0^t \int_0^{s_2} \mathcal{L}^0 b(X(s_1)) ds_1 dW(s_2) + \\ &+ \int_0^t \int_0^{s_3} \left\{ \int_0^{s_2} \mathcal{L}^0 \mathcal{L}^1 b(X(s_1)) ds_1 + \int_0^{s_2} \mathcal{L}^1 \mathcal{L}^1 b(X(s_1)) dW(s_1) \right\} dW(s_2) dW(s_3). \end{aligned} \tag{4.57}$$

From the properties of the Brownian motion we saw in Sec. 4.1.2 that the quadratic variation, Eq. (4.11), is given by

$$\int_0^t W(s) dW(s) = \frac{1}{2} W^2(t) - \frac{1}{2} [W, W](t) = \frac{1}{2} W^2(t) - \frac{1}{2} t. \tag{4.58}$$

Where, $t_0$ is set to 0, but if we let $\Delta t = t - t_0$ we hence see that Eq. (4.58) becomes

$$\begin{aligned}\int_{t_0}^{t}\int_{t_0}^{s_3} dW(s_2)dW(s_3) = \int_{t_0}^{t} W(s)dW(s) &= \frac{1}{2}W^2(t-t_0) - \frac{1}{2}[W,W](t-t_0) \\ &= \frac{1}{2}\left\{(\Delta W(t))^2 - \frac{1}{2}\Delta t\right\}. \end{aligned} \tag{4.59}$$

Inserting the result obtained from Eq. (4.59) for the double Îto integral into the Îto expansion, Eq. (4.56), the solution can be recasted as,

$$\begin{aligned} X(t) &= X(t_0) + a(X(t_0))\int_{t_0}^{t} ds_2 \\ &\quad + b(X(t_0))\int_{t_0}^{t} dW(s_2) + \mathcal{L}^1 b(X(t_0))\frac{1}{2}\left\{(\Delta W(t))^2 - \frac{1}{2}\Delta t\right\} + \widetilde{R}, \end{aligned} \tag{4.60}$$

where $\widetilde{R}$ is now given by Eq. (4.61)

$$\begin{aligned} \widetilde{R} &= \int_{t_0}^{t}\int_{t_0}^{s_2} \mathcal{L}^0 a(X(s_1))ds_1 ds_2 + \int_{t_0}^{t}\int_{t_0}^{s_2} \mathcal{L}^1 a(X(s_1))dW(s_1)ds_2 \\ &+ \int_{t_0}^{t}\int_{t_0}^{s_2} \mathcal{L}^0 b(X(s_1))ds_1 dW(s_2) + \\ &+ \int_{t_0}^{t}\int_{t_0}^{s_3} \left\{\int_{0}^{s_2} \mathcal{L}^0\mathcal{L}^1 b(X(s_1))ds_1 + \int_{t_0}^{s_2} \mathcal{L}^1\mathcal{L}^1 b(X(s_1))dW(s_1)\right\} dW(s_2)dW(s_3). \end{aligned} \tag{4.61}$$

From the remainder term we see that we have multiple integrals, which are made up of $ds$ and $dW(t)$ and combinations of these. Those that are with respect to $ds$ can be treated as normal Lebesgue integrals and can be integrated in a normal manner. The others have to be treated as Îto integral and must obey the rules associated with those integrals.

All the numerical methods that are based on Taylor expansions contains a combination of these integrals plus a remainder—the number of terms depends on the level of convergence desired. The remainder term is discarded for that numerical scheme in question. The more terms included in the numerical procedure, the higher the accuracy in the approximation and the more expansive the computation will be. The scheme used will depend on the nature of the problem, if we are dealing with a *stiff* SDE[8] then one would want to consider implicit methods as opposed to explicit ones.

As an illustration, we consider in Section 4.2.1, 4.2.2 and Section 4.2.3 three different schemes and compare their performance on an applied problem. These three different schemes have order of convergence from $\gamma = 0.5$ to $\gamma = 1.5$.

[8] A stochastic differential equation is said to be *stiff* if the set of eigenvalues differ by a great amount.

In Eq. (4.60), one can notice that there are terms like $a_x$ and $b_x$ that are embedded in terms such as $\mathcal{L}^0 a$. These terms can be a numerical disadvantage, because one has to evaluate the derivative of a function at each iteration of the numerical procedure. (The source code that evaluates the derivatives numerically for a function in more that one variable is given in Appendix E.3.11 and for a function of one single variable the source code is in Appendix E.3.9.). This problem can be resolved by using procedure such as the Runge–Kutta method. In this case one is making an explicit approximation.

Implicit methods are used when we are dealing with multidimensional stochastic differential equations where the eigenvalues vary a great deal—that is stiff stochastic differential equations. In this case, one would use the implicit methods developed and reviewed in Kloeden and Platen (1992). Such methods will not be discussed here, as they are beyond the scope of this thesis.

We now come to the simplest and quickest numerical scheme available, the Euler–Marayama scheme which has a order of convergence of $\gamma = 0.5$.

### 4.2.1 The Euler–Marayama Scheme, $\gamma = 0.5$

The *Euler–Marayama* approximation is the simplest time discrete approximation of an Îto process. Given an Îto process, $X = \{X(t), t_0 \leq t \leq T\}$, satisfying Eq. (4.44) with Taylor expansion given by Eq. (4.60), we can then construct the iterative numerical scheme by just considering the first two non–trivial terms in the expansion. For a given discretization $t_0 = \tau_0 < \tau_1 < \cdots < \tau_i, \cdots < t_N = T$,

$$Y(\tau_{n+1}) = Y(\tau_n) + a(\tau_n, Y(\tau_n))\,[\tau_{n+1} - \tau_n] + b(\tau_n, Y(\tau_n))\,[W(\tau_{n+1}) - W(\tau_n)]\,, \quad (4.62)$$

for $n = 1, \cdots, N-1$ with initial value

$$Y_0 = X(0). \quad (4.63)$$

If we rewrite $\tau_{n+1} - \tau_n$ by

$$\Delta_n = \tau_{n+1} - \tau_n, \quad (4.64)$$

for the $n^{th}$ time increment and call $\delta$ the equidistant maximum time step

$$\delta \equiv \Delta_n = \frac{T - t_0}{N}, \quad (4.65)$$

for some integer $N$ we can increment the time step by

$$\tau_n = t_0 + n\delta. \tag{4.66}$$

Moreover if we define the random increments

$$\Delta W_n \equiv \Delta W(\tau_n) = W(\tau_{n+1}) - W(\tau_n), \quad \text{for} \quad n = 1, \cdots, N-1, \tag{4.67}$$

where $W(\tau_n)$ is a Brownian motion, $W = \{W(t), t \geq 0\}$ then we rewrite the Euler-Marayama scheme as

$$Y(\tau_{n+1}) = Y(\tau_n) + a(\tau_n, Y(\tau_n))\Delta_n + b(\tau_n, Y(\tau_n))\Delta W(\tau_n). \tag{4.68}$$

For the multi–dimensional case with $m$ Wiener processes and $d$ stochastic differential equations,so that we have $d$ stochastic differential equations with $m$ noise terms in each SDE. The scheme takes the form

$$Y^k(\tau_{n+1}) = Y^k(\tau_n) + a^k(\tau_n, Y^k(\tau_n))\Delta_n + \sum_{j=1}^{m} b^{kj}(\tau_n, Y^k(\tau_n))\Delta W^j(\tau_n), \tag{4.69}$$

where $b^{kj}(\tau_n, Y^k(\tau_n))$ is the $kj^{th}$ entry of an $d \times m$ matrix, similar to the one defined in Section 4.1.4 in Eq. (4.35). The source code for this numerical scheme can be found in Appendix E.3.7. The next higher order scheme is the Milstein scheme, which is explained in the next section.

### 4.2.2 The Milstein Scheme, $\gamma = 1.0$

The scheme of interest is the one developed by *Milstein*, which has an order of convergence of $\gamma = 1.0$ and is very similar to the Euler–Mirayama scheme discussed in the previous section but, with an additional term that comes from the double Îto integral in Eq. (4.59). We saw in Eq. (4.60) that a general solution to the Îto process was given by

$$\begin{aligned} X(t) &= X(t_0) + a(X(t_0)) \int_{t_0}^{t} ds_2 \\ &\quad + b(X(t_0)) \int_{t_0}^{t} dW(s_2) + \mathcal{L}^1 b(X(t_0)) \frac{1}{2} \left\{ (\Delta W(t))^2 - \frac{1}{2}\Delta t \right\} + \widetilde{R}, \end{aligned} \tag{4.70}$$

where $\widetilde{R}$ is given by Eq. (4.61). The scheme is constructed by only considering the few non–trivial terms in the series in Eq. (4.70), by considering the Euler–Marayama

scheme plus an extra non–trivial term and dropping the remainder term. Looking at Eq. (4.70), the Milstein scheme is defined as

$$\begin{aligned} Y(\tau_{n+1}) &= Y(\tau_n) + a(\tau_n, Y(\tau_n))\Delta_n + b(\tau_n, Y(\tau_n))\Delta W(\tau_n) \\ &+ \frac{1}{2} b(\tau_n, Y(\tau_n)) b_x(\tau_n, Y(\tau_n)) \left\{ (\Delta W(\tau_n))^2 - \Delta_n \right\}. \end{aligned} \tag{4.71}$$

In the multi–dimensional case the Milstein scheme has the form,

$$\begin{aligned} Y^k(\tau_{n+1}) &= Y^k(\tau_n) + a^k(\tau_n, Y^k(\tau_n))\Delta_n + \sum_{j=1}^{m} b^{kj}(\tau_n, Y^k(\tau_n))\Delta W^j(\tau_n) \\ &+ \sum_{j_1, j_2 = 1}^{m} \mathcal{L}^{j_1} b^{kj_1}(\tau_n, Y^k(\tau_n)) I_{(j_1, j_2)}(\tau_n). \end{aligned} \tag{4.72}$$

Here $I_{(j_1,j_2)}(\tau_n)$ is a double Îto integral

$$I_{(j_1,j_2)}(\tau_n) = \int_{\tau_{n+1}}^{\tau_n} \int_{\tau_n}^{s_1} dW^{j_1}(s_1) dW^{j_2}(s_2) = \begin{cases} \frac{1}{2}\left\{ (\Delta W^{j_1}(\tau_n))^2 - \Delta_n \right\} & \text{if } j_1 = j_2 \\ J^p_{(j_1,j_2)}(\tau_n) & \text{if } j_1 \neq j_2, \end{cases} \tag{4.73}$$

where $J^p_{j_1,j_2}(\tau_n)$ is an approximation for a given order $p$,

$$\begin{aligned} J^p_{(j_1,j_2)}(\tau_n) &= \Delta_n \left\{ \frac{1}{2}\xi_{j_1}\xi_{j_2} + \sqrt{\rho_p}\left(\mu_{j_1,p}\xi_{j_2} - \mu_{j_2,p}\xi_{j_1}\right) \right\} \\ &+ \frac{\Delta_n}{2\pi} \sum_{r=1}^{p} \frac{1}{r} \left[ \zeta_{j_1,r}\left(\sqrt{2}\xi_{j_2} + \eta_{j_2,r}\right) - \zeta_{j_2,r}\left(\sqrt{2}\xi_{j_1} + \eta_{j_1,r}\right) \right], \end{aligned} \tag{4.74}$$

and

$$\rho_p = \frac{1}{12} - \frac{1}{12\pi^2} \sum_{r=1}^{p} \frac{1}{r^2} \tag{4.75}$$

and $\xi_j, \mu_{j,p}, \eta_{j,r}$ and $\zeta_{j,r}$ are independent $\mathcal{N}(0,1)$ Gaussian random variables with

$$\begin{aligned} \xi_j = \frac{1}{\sqrt{\Delta_n}}\Delta W^j(\tau_n) \quad , \quad \zeta_{j_1,r} &= \sqrt{\frac{2}{\Delta_n}}\pi r a_{j,r} \quad , \quad \eta_{j,r} = \sqrt{\frac{2}{\Delta_n}}\pi r b_{j,r}, \\ \text{and} \quad \mu_{j,r} &= \frac{1}{\sqrt{\Delta_n \rho_p}} \sum_{r=p+1}^{\infty} a_{j,r}. \end{aligned} \tag{4.76}$$

The coefficients $a_{j,r}$ and $b_{j,r}$ come from the Fourier approximation of the Brownian bridge (Kloeden and Platen 1992)

$$W^j(t) - \frac{t}{\Delta_n} W^j(\Delta_n) = \frac{1}{2} a_{j,0} + \sum_{r=1}^{\infty} a_{j,r} \cos\left(\frac{2\pi r t}{\Delta_n}\right) + b_{j,r} \sin\left(\frac{2\pi r t}{\Delta_n}\right) \tag{4.77}$$

with

$$a_{j,r} = \frac{2}{\Delta_n}\int_0^{\Delta_n}\left[W^j(s) - \frac{s}{\Delta_n}W^j(\Delta_n)\right]\cos\left(\frac{2\pi rs}{\Delta_n}\right)ds \tag{4.78}$$

$$b_{j,r} = \frac{2}{\Delta_n}\int_0^{\Delta_n}\left[W^j(s) - \frac{s}{\Delta_n}W^j(\Delta_n)\right]\sin\left(\frac{2\pi rs}{\Delta_n}\right)ds. \tag{4.79}$$

This approximation depends on the size of $p$ which influences the accuracy of $J^p_{j_1,j_2}(\tau_n)$, as an approximation of $I(j_1,j_2)(\tau_n)$. The constant $p$ must be chosen such that

$$p \equiv p(\Delta_n) \geq \frac{K}{\Delta_n} \tag{4.80}$$

for some constant $K \geq 0$, which is tuned in order to obtain the strong convergence of $\gamma = 1.0$.

The souce code for this numerical scheme can be found in Appendix E.3.8.

### 4.2.3 The Order $\gamma = 1.5$ Strong Taylor Scheme

Let us now consider a $\gamma = 1.5$ order *Taylor* scheme. In the previous two sections we saw that the Taylor expansion, Eq. (4.60), with remainder term given by Eq. (4.61), could produce higher order terms by truncating the expansion series at a given order. Here we see that eliminating just a term, i.e.,

$$\int_{t_0}^{t}\int_{t_0}^{s_3}\left\{\int_0^{s_2}\mathcal{L}^0\mathcal{L}^1 b(X(s_1))ds_1\right\}dW(s_2)dW(s_3) \tag{4.81}$$

in that Taylor expansion produces a numerical scheme or order $\gamma = 1.5$. To find an expression for this numerical scheme one needs to evaluate

$$X(t) = X'(t) + \widetilde{R} - \int_{t_0}^{t}\int_{t_0}^{s_3}\left\{\int_0^{s_2}\mathcal{L}^0\mathcal{L}^1 b(X(s_1))ds_1\right\}dW(s_2)dW(s_3). \tag{4.82}$$

Now if we evaluate the operators, $\mathcal{L}^0$ and $\mathcal{L}^1$, which are acting on the functions $a(\tau_n, Y(\tau_n))$ and $b(\tau_n, Y(\tau_n))$ in Eq. (4.82) we obtain the following solution that contains multiple

Îto integrals,

$$\begin{aligned} X(t) &= X(t_0) + a(X(t_0)) \int_{t_0}^{t} ds_2 \\ &+ b(X(t_0)) \int_{t_0}^{t} dW(s_2) + \frac{1}{2} b(X(0)) b_x(X(0)) \left\{ (\Delta W(t))^2 - \frac{1}{2}\Delta t \right\} \\ &+ \int_{t_0}^{t} \int_{t_0}^{s_2} \left( a(X(s_1)) a_x(X(s_1)) + \frac{1}{2}[b(X(s_1))]^2 a_{xx}(X(s_1)) \right) ds_1 ds_2 \\ &+ \int_{t_0}^{t} \int_{t_0}^{s_2} b(X(s_1)) a_x(X(s_1)) dW(s_1) ds_2 \\ &+ \int_{t_0}^{t} \int_{t_0}^{s_2} \left( a(X(s_1)) b_x(X(s_1)) + \frac{1}{2}[b(X(s_1))]^2 b_{xx}(X(s_1)) \right) ds_1 dW(s_2) \\ &+ \int_{t_0}^{t} \int_{t_0}^{s_3} \int_{t_0}^{s_2} b(X(s_1)) \Big( b(X(s_1)) b_{xx}(X(s_1)) \\ &+ (b_x(X(s_1)))^2 \Big) dW(s_1) dW(s_2) dW(s_3). \end{aligned} \tag{4.83}$$

Here we only need to consider the Îto integrals that are given by

$$I_{(0,0)}(\tau_n) = \int_{t_0}^{t} \int_{t_0}^{s_1} ds_1 ds_2 = (\Delta_n)^2 \tag{4.84}$$

$$I_{(1,0)}(\tau_n) = \int_{t_0}^{t} \int_{t_0}^{s_1} dW(s_1) ds_2 = \frac{1}{2}[\Delta_n]^{\frac{3}{2}} \left( U_1 + \frac{1}{\sqrt{3}} U_2 \right) \Delta Z \tag{4.85}$$

$$I_{(0,1)}(\tau_n) = \int_{t_0}^{t} \int_{t_0}^{s_1} ds_2 dW(s_1) = ([\Delta W(\tau_n)]\Delta_n - \Delta Z) \tag{4.86}$$

$$\begin{aligned} I_{(1,1,1)}(\tau_n) &= \int_{t_0}^{t} \int_{t_0}^{s_2} \int_{t_0}^{s_1} dW(s_1) dW(s_2) dW(s_3) \\ &= \frac{1}{2} \left\{ \frac{1}{3}[\Delta W(\tau_n)]^2 - \Delta_n \right\} \Delta W(\tau_n). \end{aligned} \tag{4.87}$$

Inserting Eqs. (4.84-4.87) into Eq. (4.83) the order 1.5 strong Taylor expansion takes the form,

$$\begin{aligned} Y(\tau_{n+1}) &= Y(\tau_n) + a(\tau_n, Y(\tau_n))\Delta_n + b(\tau_n, Y(\tau_n))\Delta W(\tau_n) \\ &+ \frac{1}{2} b(\tau_n, Y(\tau_n)) b_x(\tau_n, Y(\tau_n)) \left\{ (\Delta W(\tau_n))^2 - \Delta_n \right\} \\ &+ \left( a(\tau_n, Y(\tau_n)) a_x(\tau_n, Y(\tau_n)) + \frac{1}{2}[b(\tau_n, Y(\tau_n))]^2 a_{xx}(\tau_n, Y(\tau_n)) \right) \Delta_n^2 \\ &+ b(\tau_n, Y(\tau_n)) a_x(\tau_n, Y(\tau_n)) \Delta Z \\ &+ \left( a(\tau_n, Y(\tau_n)) b_x(\tau_n, Y(\tau_n)) + \frac{1}{2}[b(\tau_n, Y(\tau_n))]^2 b_x(\tau_n, Y(\tau_n)) \right) \\ &\times \Big( [\Delta W(\tau_n)]\Delta_n - \Delta Z \Big) \\ &+ b(\tau_n, Y(\tau_n)) \Big( b(\tau_n, Y(\tau_n)) b_{xx}(\tau_n, Y(\tau_n)) + (b_x(\tau_n, Y(\tau_n)))^2 \Big) \\ &\times \frac{1}{2} \left\{ \frac{1}{3}[\Delta W(\tau_n)]^2 - \Delta_n \right\} \Delta W(\tau_n). \end{aligned} \tag{4.88}$$

In the general multi–dimensional case, with $d, m = 1, 2, \cdots$, the $k^{th}$ component of the order 1.5 strong Taylor schemes takes the form,

$$\begin{aligned} Y^k(\tau_{n+1}) = Y^k(\tau_n) &+ a^k(\tau_n, Y^k(\tau_n))\Delta_n + \frac{1}{2}\mathcal{L}^0 a^k(\tau_n, Y^k(\tau_n))[\Delta_n]^2 \\ &+ \sum_{j=1}^{m}\Big\{ b^{kj}(\tau_n, Y^k(\tau_n))\Delta W^j(\tau_n) + \mathcal{L}^0 b^{kj}(\tau_n, Y^k(\tau_n))I_{(0,j)}(\tau_n) \\ &+ \mathcal{L}^j a^k(\tau_n, Y^k(\tau_n))I_{(j,0)}(\tau_n)\Big\} \\ &+ \sum_{j_1,j_2=1}^{m} \mathcal{L}^{j_1} b^{kj_2}(\tau_n, Y^k(\tau_n))I_{(j_1,j_2)}(\tau_n) \\ &+ \sum_{j_1,j_2,j_3=1}^{m} \mathcal{L}^{j_1}\mathcal{L}^{j_2} b^{kj_3}(\tau_n, Y^k(\tau_n))I(j_1,j_2,j_3)(\tau_n). \end{aligned} \tag{4.89}$$

The difference here compared with the Milstein scheme, Eq. (4.71), is that we also have multiple Îto integrals with respect to different components of the Wiener process. Here the same approximation, Eq. (4.74), can be used to for these Îto integrals. We therefore use Eq. (4.76) with a few extra terms that are given in this section. This follows the notation in Kloeden and Platen (1992). Then for $j, j_1, j_2$ and $j_3 = 1, 2, \cdots, m$ and $p = 1, 2, \cdots$ we have

$$I_{(j)} = \Delta W^j(\tau_n) = \sqrt{\Delta_n}\xi_j, \tag{4.90}$$

$$I_{(j,0)} = \frac{1}{2}\Delta_n\left(\sqrt{\Delta_n}\xi_j + a_{j,0}\right), \tag{4.91}$$

with

$$a_{j,0} = -\frac{\sqrt{2\Delta_n}}{\pi}\sum_{r=1}^{p}\frac{1}{r}\zeta_{j,r} - 2\sqrt{\Delta_n\rho_p}\mu_{j,p}, \tag{4.92}$$

where $\rho_p$ is given by Eq. (4.75). The Îto integral $I_{(j_1,j_2)}(\tau_n)$ is defined by Eq. (4.73). The last Îto integral is a triple integral, which takes the form,

$$I^p_{(j_1,j_2),j_3}(\tau_n) = \begin{cases} \frac{1}{2}\left\{\frac{1}{3}(\Delta W^{j_1}(\tau_n))^2 - \Delta_n\right\}\Delta W^j(\tau_n) & \text{if } j_1 = j_2 = j_3 \\ J^p_{(j_1,j_2,j_3)}(\tau_n) & \text{otherwise,} \end{cases} \tag{4.93}$$

with

$$\phi_{j,r} = \frac{1}{\sqrt{\Delta_n\alpha_p}}\sum_{r=p+1}^{\infty}\frac{1}{r}b_{j,r} \quad \text{and} \quad \alpha_p = \frac{\pi^2}{180} - \frac{1}{2\pi^2}\sum_{r=1}^{p}\frac{1}{r^4}. \tag{4.94}$$

The other $J$ integrals, Eq. (4.74) are approximated using the following

$$J^p_{(0)} = \Delta_n \;\;,\;\; J^p_{(j)} = \sqrt{\Delta_n}\xi_j \;\text{ and }\; J^p_{(0,0)} = [\Delta_n]^2, \tag{4.95}$$

$$J^p_{(j,0)} = \frac{1}{2}\Delta_n\left(\sqrt{\Delta_n}\xi_j + a_{j,0}\right) \;\;,\;\; J^p_{(0,j)} = \frac{1}{2}\Delta_n\left(\sqrt{\Delta_n}\xi_j - a_{j,0}\right). \tag{4.96}$$

The integrals with three indices are given by,

$$\begin{aligned} J^p_{(0,0,0)} &= \frac{1}{3!}[\Delta_n]^3 \\ J^p_{(j,0,0)} &= \frac{1}{3!}[\Delta_n]^{\frac{5}{2}}\xi_j + \frac{1}{4}[\Delta_n]^2 a_{j,0} - \frac{1}{\pi}\Delta_n b_j \\ J^p_{(0,j,0)} &= \frac{1}{3!}[\Delta_n]^{\frac{5}{2}}\xi_j - \frac{1}{\pi}\Delta_n b_j \\ J^p_{(j,0,0)} &= \frac{1}{3!}[\Delta_n]^{\frac{5}{2}}\xi_j - \frac{1}{4}[\Delta_n]^2 a_{j,0} - \frac{1}{\pi}\Delta_n b_j, \end{aligned} \tag{4.97}$$

with

$$b_j = \sqrt{\frac{\Delta_n}{2}}\sum_{r=1}^{p}\frac{1}{r^2}\eta_{j,r} + \sqrt{\Delta_n \alpha_p}\phi_{j,p}, \tag{4.98}$$

$$\begin{aligned} J^p_{(j_1,0,j_2)} &= \frac{1}{3!}[\Delta_n]^2\xi_{j_1}\xi_{j_2} + \frac{1}{2}a_{j_1,0}J^p_{(0,j_2)} + \frac{1}{2\pi}[\Delta_n]^{\frac{3}{2}}\xi_{j_2}b_{j_1} \\ &- [\Delta_n]^2 B^p_{(j_1,j_2)} - \frac{1}{4}[\Delta_n]^{\frac{3}{2}}a_{j_2,0}\xi_{j_1} + \frac{1}{2\pi}[\Delta_n]^{\frac{3}{2}}\xi_{j_1}b_{j_2} \end{aligned} \tag{4.99}$$

$$\begin{aligned} J^p_{(0,j_1,j_2)} &= \frac{1}{3!}[\Delta_n]^2\xi_{j_1}\xi_{j_2} - \frac{1}{\pi}[\Delta_n]^{\frac{3}{2}}\xi_{j_2}b_{j_1} + [\Delta_n]^2 B^p_{(j_1,j_2)} \\ &- \frac{1}{4}[\Delta_n]^{\frac{3}{2}}a_{j_2,0}\xi_{j_1} + \frac{1}{2\pi}[\Delta_n]^{\frac{3}{2}}\xi_{j_1}b_{j_2} \\ &+ [\Delta_n]^2 C^p_{(j_1,j_2)} + \frac{1}{2}[\Delta_n]^2 A^p_{(j_1,j_2)}, \end{aligned} \tag{4.100}$$

$$\begin{aligned} J^p_{(j_1,j_2,0)} &= \frac{1}{2}[\Delta_n]^2\xi_{j_1}\xi_{j_2} - \frac{1}{2}[\Delta_n]^{\frac{3}{2}}\left(a_{j_2,0}\xi_{j_1} - a_{j_1,0}\xi_{j_2}\right) \\ &+ [\Delta_n]^2 A^p_{(j_1,j_2)} - J^p_{(j_1,0,j_2)} - J^p_{(0,j_1,j_2)}, \end{aligned} \tag{4.101}$$

$$\begin{aligned} J^p_{(j_1,j_2,j_3)} &= \frac{1}{\sqrt{\Delta_n}}\xi_{j_1}J^p_{(0,j_2,j_3)} + \frac{1}{2}a_{j_1,0}J^p_{(j_2,j_3)} + \frac{1}{2\pi}\Delta_n b_{j_1}\xi_{j_2}\xi_{j_3} \\ &- [\Delta_n]^{\frac{3}{2}}\xi_{j_2}B^p_{(j_1,j_3)} + [\Delta_n]^{\frac{3}{2}}\xi_{j_3}\left(\frac{1}{2}A^p_{(j_1,j_2)} - C^p_{(j_2,j_1)}\right) \\ &+ [\Delta_n]^{\frac{3}{2}}D^p_{(j_1,j_2,j_3)}, \end{aligned} \tag{4.102}$$

where

$$A^p_{(j_1,j_2)} = \frac{1}{2\pi}\sum_{r=1}^{p}\frac{1}{r}\left[\zeta_{j_1,r}\eta_{j_2,r} - \zeta_{j_2,r}\eta_{j_1,r}\right], \tag{4.103}$$

$$B^p_{(j_1,j_2)} = \frac{1}{4\pi^2}\sum_{r=1}^{p}\frac{1}{r^2}\left[\zeta_{j_1,r}\zeta_{j_2,r} - \eta_{j_1,r}\eta_{j_2,r}\right], \tag{4.104}$$

$$C^p_{(j_1,j_2)} = -\frac{1}{2\pi^2}\sum_{r,l=1,r\neq l}^{p}\frac{r}{r^2-l^2}\left[\frac{1}{l}\zeta_{j_1,r}\zeta_{j_2,r} - \frac{1}{r}\eta_{j_1,r}\eta_{j_2,r}\right], \tag{4.105}$$

$$\begin{aligned} D^p_{(j_1,j_2,j_3)} &= -\frac{1}{\pi^2 2^{\frac{5}{2}}}\sum_{r,l=1}^{p}\frac{1}{l(r+l)}\Big[\zeta_{j_2,l}\left(\zeta_{j_3,l+r}\eta_{j_1,r} - \zeta_{j_1,r}\eta_{j_1,l+r}\right) \\ &+ \eta_{j_2,l}\left(\zeta_{j_1,r}\zeta_{j_3,l+r} + \eta_{j_1,r}\eta_{j_3,l+r}\right)\Big] \\ &+ \frac{1}{\pi^2 2^{\frac{5}{2}}}\sum_{l=1}^{p}\sum_{r=1}^{l-1}\frac{1}{r(l-r)}\Big[\zeta_{j_2,l}\left(\zeta_{j_1,r}\eta_{j_3,l-r} + \zeta_{j_3,l-r}\eta_{j_1,r}\right) \\ &- \eta_{j_2,l}\left(\zeta_{j_1,r}\zeta_{j_3,l-r} - \eta_{j_1,r}\eta_{j_3,l-r}\right)\Big] \\ &+ \frac{1}{\pi^2 2^{\frac{5}{2}}}\sum_{l=1}^{p}\sum_{r=l+1}^{2p}\frac{1}{r(r-l)}\Big[\zeta_{j_2,l}\left(\zeta_{j_3,r-l}\eta_{j_1,r} - \zeta_{j_1,r}\eta_{j_3,r-l}\right) \\ &- \eta_{j_2,l}\left(\zeta_{j_1,r}\zeta_{j_3,r-l} + \eta_{j_1,r}\eta_{j_3,r-l}\right)\Big]. \end{aligned} \tag{4.106}$$

For $r > p$, the Gaussian variables $\eta_{j,r}$ and $\zeta_{j,r}$ are set to zero for $j = 1, \cdots, m$. In Fig. (4.3) and Fig. (4.4), we show the results obtained from the numerical schemes discussed above for a simple SDE for two different sample paths. The SDE takes the following form

$$dX(t) = a^2X(t)(1 + (X(t))^2)dt + a(1 + (X(t))^2)dW(t), \tag{4.107}$$

which is a reducible SDE. For such an SDE it is possible to obtain an analytic solution,

$$X(t) = \tan\left(aW(t) + \arctan(X(0))\right). \tag{4.108}$$

The analytic solution[9] can then be compared with the results obtained from the numerical schemes. Here the point of the exercise is to obtain a quick perspective on which scheme performs the best.

Without any zooming we can clearly see that the red curve (The Euler–Marayama scheme, $\gamma = 0.5$) is above all the other lines and that the blue line (the order 1.5 strong Taylor scheme, $\gamma = 1.5$) is the closest out of the three other schemes. This says that for a simulation the order 1.5 strong Taylor scheme would be the preferred scheme along

[9]The source code for this explicit solution is given by Appendix E.3.6.

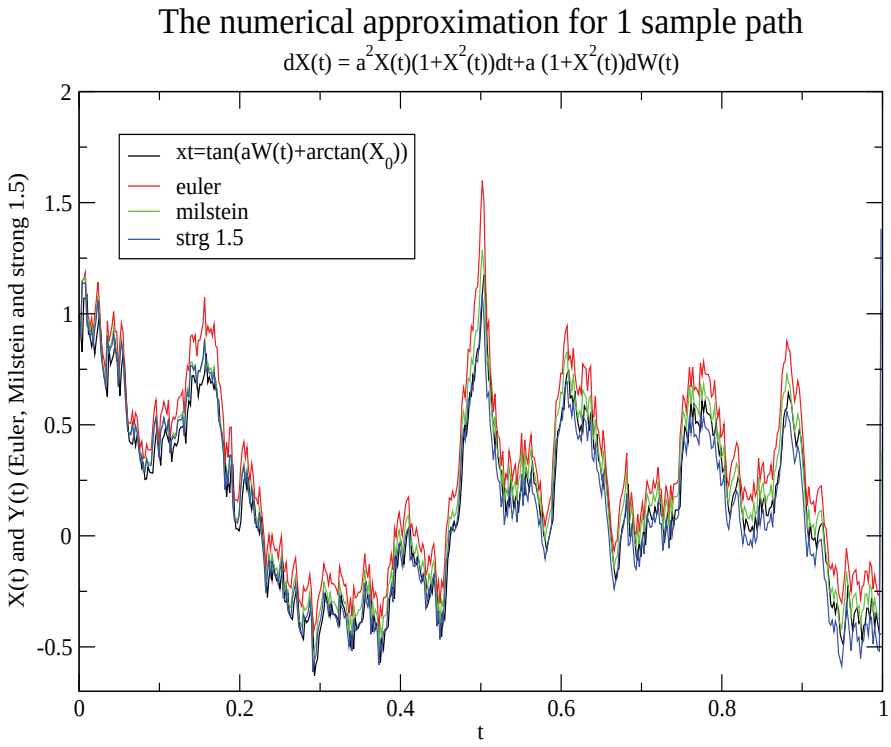


**Figure 4.3. Sample paths for the explicit solution versus the numerical schemes.** Sample paths for the exact solution, Eq. (4.108), of a stochastic differential equation, Eq. (4.107), and their numerical approximations, the Euler–Marayama method, Eq. (4.68), the Milstein method, Eq. (4.71), and the strong Taylor 1.5 method, Eq. (4.88). Here the numerical solution of the stochastic differential equation, Eq. (4.107), is compared to the explicit solution for that stochastic differential equation, Eq. (4.108). Here the noise term is just the standard Wiener process, and is Gaussian distributed.

with at least the Milstein scheme. The Euler–Marayama is very quick and simple to implement, hence may be used to obtain an overview of the solution.

The souce code for this numerical scheme can be found in Appendix E.3.10.

### 4.2.4 The Runge–Kutta Scheme a Strong Order $\gamma = 1$ Scheme

One of the main disadvantages with the Taylor expansion method is that it involves derivatives that must be evaluated at each iterations of the numerical scheme in question. On simple schemes, this may not present a problem but as the level of the complexity increases in the numerical scheme the more these derivative appear, and these are of higher order. This could pose a problem during the computation as it increases the computational expanse.

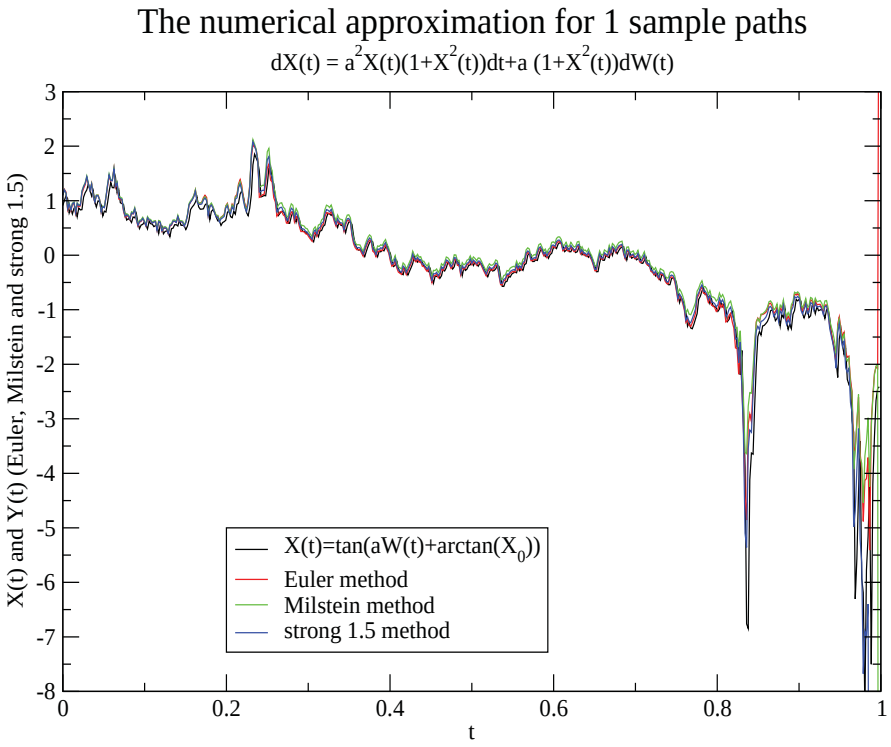


**Figure 4.4. Sample paths for the explicit solution versus the numerical schemes.** Here we show another sample path with a different random sequence for the exact solution, Eq. (4.108), of a stochastic differential equation, Eq. (4.107), and their numerical approximations, the Euler–Marayama method, Eq. (4.68), the Milstein method, Eq. (4.71), and the strong Taylor 1.5 method, Eq. (4.88). Here the same numerical procedures are applied to the same stocahstic differential equation as in Fig. 4.3 but this time for a different sample path. Here the noise term is just the standard Wiener process, and is Gaussian distributed.

The Runge–Kutta method, which can be applied to stochastic differential equation in the same manner as it is applied to normal differential equation replaces the differential term by approximations.

In this section we will write the Runge–Kutta method for the Milstein scheme. The resulting scheme is also a scheme of order 1, that is $\gamma = 1$.

In general we can always write the differential of a function as

$$b'(y) = \lim_{\Delta y \to 0} \frac{b(y + \Delta y) - b(y)}{\Delta y}. \tag{4.109}$$

Hence we can approximate via

$$b(Y(\tau_n) + \Delta Y(\tau_n)) - b(Y(\tau_n)) = b'(Y(\tau_n))\Delta Y(\tau_n) + \mathcal{O}((\Delta Y(\tau_n))^2), \tag{4.110}$$

and using the Euler–Marayama scheme, Eq. (4.68),

$$\Delta Y(\tau_n) = a(\tau_n, Y(\tau_n))\Delta_n + b(\tau_n, Y(\tau_n))\Delta W(\tau_n), \tag{4.111}$$

we deduce that

$$b(Y(\tau_n) + \Delta Y(\tau_n)) - b(Y(\tau_n)) = b'(Y(\tau_n))b(\tau_n, Y(\tau_n))\Delta W(\tau_n) + \mathcal{O}(\Delta_n). \tag{4.112}$$

Applying $\Delta W(\tau_n) = \sqrt{\Delta_n}$, with Eq. (4.111), we arrive at

$$\begin{aligned} b'(Y(\tau_n))b(\tau_n, Y(\tau_n)) &= \frac{1}{\sqrt{\Delta_n}}\Big( b(\tau_n, Y(\tau_n) + a(\tau_n, Y(\tau_n))\Delta_n + b(\tau_n, Y(\tau_n))\Delta W(\tau_n)) \\ &- b(\tau_n, Y(\tau_n))\Big). \end{aligned} \tag{4.113}$$

This expression is used in the Milstein scheme, Eq. (4.71). The resulting scheme which does not not involve any differential terms in it is given by what is called the Runge–Kutta order $\gamma = 1$ method

$$\begin{aligned} \hat{Y}(\tau_n) &= Y(\tau_n) + a(\tau_n, Y(\tau_n))\Delta_n + b(\tau_n, Y(\tau_n))\sqrt{\Delta_n} \\ Y(\tau_{n+1}) &= Y(\tau_n) + a(\tau_n, Y(\tau_n))\Delta_n + b(\tau_n, Y(\tau_n))\Delta W(\tau_n) \\ &+ \frac{1}{2\sqrt{\Delta_n}}\Big( b(\tau_n, \hat{Y}(\tau_n) - b(\tau_n, Y(\tau_n))\Big). \end{aligned} \tag{4.114}$$

Numerical methods are in most cases used as an alternative approach to analytical solution simply because analytical solutions cannot be obtained.

These methods will be used to obtained solutions in models used and built in later sections. And will be used to compare the results obtained with path integral techniques.

We next turn to another class of stochastic differential equations—that is those that involve jumps. It does not take much insight to realize that volatility in real markets causes prices to move up and down by large amounts. These movements cannot be modelled by normal stochastic differential equations and very often these models must incorporate what we call jump processes. Jump processes are built from Poisson compound process and fall in the category of Lévy processes, which, unlike the standard geometric Brownian motion, are non–Gaussian processes. What this means is that these processes capture more stylized effects, those that we observe from empirical data. The disadvantage with these models is that they are more complicated to solve and that the Lévy distribution is not fully understood. Nevertheless this class of problems is useful and is still a very active field of research. For a complete discussion on financial modeling with jump processes see Cont and Tankov (2004) and Shreve (2004).

## 4.3 Jump Process and Lévy Processes

Jump process are a very useful way of modelling financial data when the data contains large volatility in it. These processes are contructed using compound poisson process. The distribution generated for this procees are usually Lévy distribution. We are not going to discuss these distributions in this work, but leave it for further work. A complete discussion may be found in Cont and Tankov (2004) or in Shreve (2004).

## 4.4 Stochastic Volatility Models

### 4.4.1 Mean Reverting Stochastic Volatility Models

When the volatility is a Markov Ito process, it is possible to find a pricing function for the European derivative of the form $P(t, X(t), Y(t))$ from the non-arbitrage arguments, as in the Black-Scholes case (Fouque *et al.* 2000). The function $P(t, X(t), Y(t))$ satisfies a partial differential equation with *two* space dimensions ($x$ and $y$); the price of the derivative depends on the value of the process $y$, which is not directly observable.

We will derive the path integral for a mean reverting *stochastic volatility model*[10]. We first review a volatility model assuming that the volatility is a function of a mean of a mean reverting Ornstein–Uhlenbeck process (OU) defined as a solution of,

$$dY(t) = m + (y - m)e^{-\alpha t} + \beta \int_0^t \exp(-\alpha(t-s))d\hat{Z}(s), \tag{4.115}$$

so that the distribution $Y \sim \mathcal{N}(m + (y-m)e^{\alpha t}, \frac{\beta^2}{2\alpha}(1 - e^{-2\alpha t}))$. In this model, $\alpha$ is called the *rate of mean reversion* and $m$ is the long–run mean level of $Y$. The drift term pulls $Y$ towards $m$, so we would expect that $\sigma(t)$ is pulled toward the mean value of $f(Y(t))$ with respect to the long–run distribution of $Y$.

The second Brownian motion $\hat{Z}(t)$

$$\hat{Z}(t) = \rho W(t) + \sqrt{1 - \rho^2}, \tag{4.116}$$

where $Z(t)$ is also a Brownian motion that is dependent of $W(t)$. This motion is usually correlated with the Brownian motion $W(t)$ driving the asset price. The instantaneous

[10] A general discussion on mean reverting stochastic volatility model can be found in Fouque *et al.* (2000).

correlation coefficient is denoted by $\rho$ and defined by the covariation[11] of $W(t)$ and $\hat{Z}(t)$,

$$\rho\, dt = d\langle \hat{Z}, W\rangle(t) \equiv d\hat{Z}(t)dW(t). \tag{4.117}$$

In general it is found from financial data that $\rho < 0$, and there are economic arguments for a negative correlation or leverage effect between stock price and volatility stocks. It is commonly found from empirical studies that the asset prices tend to go down when the volatility goes up. Furthermore, in general the correlation may depend on time $\rho(t) \in [-1, 1]$, but we shall assume that it is independent of time for now, since in most practical situations it is taken to be such.

The stochastic process with a mean reverting (OU) function is defined as,

$$\begin{aligned} d\,X(t) &= \mu X(t)\, dt + \sigma(t)X(t)\, dW(t) \\ \sigma(t) &= f(Y(t)), \\ d\,Y(t) &= \alpha(m - Y(t))\, dt + \beta\, dZ\hat{(}t). \end{aligned} \tag{4.118}$$

We now examine the pricing function $P(t, x, y)$ by trying to construct hedged portfolio of assets that can be priced by the no–arbitrage principle. Let $f^{(1)}(t, x, y)$ be the price of a European derivative with expiration date $T_1$ and payoff function $h(X_{T_1})$ and we wish to find processes $\{a(t), b(t), c(t)\}$ such that,

$$f^{(1)}(t, x, y) = a_{T_1} X_{T_1} + b_{T_1}\beta_{T_1} + c_{T_1} f^{(2)}(t, x, y) \tag{4.119}$$

where $\beta(t) = e^{rt}$ is the price of a riskless bond under the preventing short term constant interest rate $r$, and $f^{(2)}(t, x, y)$ is the price of the European contract with the same payoff function $h$ as $f^{(1)}(t, x, y)$ but with different expiration date $T_2 > T_1 > t$. Eq. (4.119) may be interpreted as a portfolio whose payoff at time $T_1$ equals the payoff of $f^{(1)}$ moreover the portfolio is to be self–financing so that

$$df^{(1)}(t, x, y) = a(t)X(t) + b(t)re^{rt} + c(t)df^{(2)}(t, x, y). \tag{4.120}$$

If such portfolio can be found then in order for there to be no–arbitrage opportunities, it must be that

$$f^{(1)}(t, x, y) = a(t)X(t) + b(t)e^{rt} + c(t)f^{(2)}(t, x, y), \tag{4.121}$$

for $t < T_1$.

[11] correlated variation of two or more variables, not to be confused with *covariance*.

Applying the two dimensional version of Îto formula, Eq. (4.122),

$$dg(t,X(t),Y(t)) = \frac{\partial g}{\partial t}\,dt + \frac{\partial g}{\partial x}dX(t) + \frac{\partial g}{\partial y}dY(t) + \frac{1}{2}\left(\frac{\partial^2 g}{\partial x^2}\,d\langle X(t)\rangle + \frac{\partial^2 g}{\partial x \partial y}\,d\langle X,Y\rangle(t) + \frac{\partial^2 g}{\partial y^2}\,d\langle Y(t)\rangle\right), \tag{4.122}$$

where

$$d\langle X(t)\rangle = \sigma_x^2(t,X(t))\,dt \tag{4.123}$$

$$d\langle X,Y\rangle(t) = \sigma_x(t,X(t))\sigma_y(t,Y(t))\,dt \tag{4.124}$$

$$d\langle Y(t)\rangle = \sigma_y^2(t,Y(t))\,dt, \tag{4.125}$$

on both sides of Eq. (4.120) for the model

$$\begin{aligned} dX(t) &= \mu X(t)\,dt + \sigma(t)X(t)\,dW(t) \\ \sigma(t) &= f(Y(t)) \\ dY(t) &= \alpha(m - Y(t))\,dt + \beta(\rho\,dW(t) + \sqrt{1-\rho^2}\,dZ(t)), \end{aligned} \tag{4.126}$$

with Eq. (4.125)

$$d\langle X(t)\rangle = (f(y)x)^2\,dt \tag{4.127}$$

$$d\langle X,Y\rangle(t) = \beta f(y)x\,d\langle W,Z\rangle(t) = \beta f(y)x\rho\,dt \tag{4.128}$$

$$d\langle Y(t)\rangle = (\beta)^2\,dt, \tag{4.129}$$

we obtain

$$\begin{aligned} df^{(1)}(t,X(t),Y(t)) &= \frac{\partial f^{(1)}}{\partial t}\,dt + \frac{\partial f^{(1)}}{\partial x}dX(t) + \frac{\partial f^{(1)}}{\partial y}dY(t) + \\ &\qquad \frac{1}{2}\left(\frac{\partial^2 f^{(1)}}{\partial x^2}\,d\langle X(t)\rangle + \frac{\partial^2 f^{(1)}}{\partial x \partial y}\,d\langle X,Y\rangle(t) + \frac{\partial^2 f^{(1)}}{\partial y^2}\,d\langle Y(t)\rangle\right) \\ &= a(t)\,dX(t) + b(t)re^{rt}\,dt + c(t)\left[\frac{\partial f^{(2)}}{\partial t}\,dt + \frac{\partial f^{(2)}}{\partial x}dX(t)\right. \\ &+ \frac{\partial f^{(2)}}{\partial y}dY(t) + \frac{1}{2}\left(\frac{\partial^2 f^{(2)}}{\partial x^2}\,d\langle X(t)\rangle + \frac{\partial^2 f^{(2)}}{\partial x \partial y}\,d\langle X,Y\rangle(t)\right. \\ &+ \left.\left.\frac{\partial^2 f^{(2)}}{\partial y^2}\,d\langle Y(t)\rangle\right)\right] \\ &= \left(a(t) + c(t)\frac{\partial f^{(2)}}{\partial x}\right)dX(t) + c(t)\frac{\partial f^{(2)}}{\partial y}\,dY(t) \\ &+ \left[b(t)re^{rt} + c(t)\left(\frac{\partial}{\partial t} + \mathcal{M}_1\right)f^{(2)}\right]dt \end{aligned} \tag{4.130}$$

with the operator $\mathcal{M}_1$ defined as

$$\mathcal{M}_1 = \frac{1}{2}\left((f(y)x)^2\frac{\partial^2}{\partial x^2} + \beta f(y)x\rho\frac{\partial^2}{\partial x \partial y} + \beta^2\frac{\partial^2}{\partial y}\right), \tag{4.131}$$

we therefore obtain the following partial differential equation

$$\begin{aligned}
&\left(\frac{\partial}{\partial t} + \mathcal{M}_1\right) f^{(1)}\, dt + \frac{\partial f^{(1)}}{\partial x} dX(t) + \frac{\partial f^{(1)}}{\partial y} dY(t) \\
= &\left(a(t) + c(t)\frac{\partial f^{(2)}}{\partial x}\right) dX(t) + c(t)\frac{\partial f^{(2)}}{\partial y}\, dY(t) \\
+ &\left[b(t)re^{rt} + c(t)\left(\frac{\partial}{\partial t} + \mathcal{M}_1\right) f^{(2)}\right] dt.
\end{aligned} \tag{4.132}$$

Here the probability functions $f^{(1)}, f^{(2)}$ and their derivatives are evaluated at $(t, X(t), Y(t))$. If we now equate the terms to extract the coefficients $a(t)$, $b(t)$ and $c(t)$, then we see that

$$c(t)\frac{\partial f^{(2)}}{\partial y} = \frac{\partial f^{(1)}}{\partial y} \quad \longrightarrow \quad c(t) = \left(\frac{\partial f^{(2)}}{\partial y}\right)^{-1}\frac{\partial f^{(1)}}{\partial y}, \tag{4.133}$$

$$\left(a(t) + c(t)\frac{\partial f^{(2)}}{\partial x}\right) = \frac{\partial f^{(1)}}{\partial x} \quad \longrightarrow \quad a(t) = \frac{\partial f^{(1)}}{\partial x} - c(t)\frac{\partial f^{(2)}}{\partial x} \tag{4.134}$$

$$\left(\frac{\partial}{\partial t} + \mathcal{M}_1\right) f^{(1)} = b(t)re^{rt} + c(t)\left(\frac{\partial}{\partial t} + \mathcal{M}_1\right) f^{(2)}. \tag{4.135}$$

Using Eq. (4.121) to obtain an expression for the $b(t)$ coefficient leads to

$$\begin{aligned}
b(t) &= \frac{\left(f^{(1)} - a(t)X(t) - c(t)f^{(2)}\right)}{e^{rt}} \\
&= \frac{1}{e^{rt}}\left[f^{(1)} - \left(\frac{\partial f^{(1)}}{\partial x} - c(t)\frac{\partial f^{(2)}}{\partial x}\right)X(t) - c(t)f^{(2)}\right].
\end{aligned} \tag{4.136}$$

As a result Eq. (4.135) becomes

$$\begin{aligned}
\left(\frac{\partial}{\partial t} + \mathcal{M}_1\right) f^{(1)} &= r\left[f^{(1)} - \left(\frac{\partial f^{(1)}}{\partial x} - c(t)\frac{\partial f^{(2)}}{\partial x}\right)x - c(t)f^{(2)}\right] \\
&+ c(t)\left(\frac{\partial}{\partial t} + \mathcal{M}_1\right) f^{(2)} \\
&= r\left[f^{(1)} - x\frac{\partial f^{(1)}}{\partial x}\right] \\
&+ c_t\left[\left(x\frac{\partial f^{(2)}}{\partial x} - f^{(2)}\right)r + \left(\frac{\partial}{\partial t} + \mathcal{M}_1\right) f^{(2)}\right] \\
\left[\left(\frac{\partial}{\partial t} + \mathcal{M}_1\right) + rx\frac{\partial}{\partial x} - r\right] f^{(1)} &= c(t)\left[\left(\frac{\partial}{\partial t} + \mathcal{M}_1\right) + rx\frac{\partial}{\partial x} - r\right] f^{(2)}.
\end{aligned} \tag{4.137, 4.138}$$

Inserting the expression for the coefficient Eq. (4.133), we obtain the following PDE for the model defined in Eq. (4.126),

$$\left(\frac{\partial f^{(1)}}{\partial y}\right)^{-1}\left[\left(\frac{\partial}{\partial t}+\mathcal{M}_1\right)+r\left(x\frac{\partial}{\partial x}-1\right)\right]f^{(1)} = \left(\frac{\partial f^{(2)}}{\partial y}\right)^{-1} \times \left[\left(\frac{\partial}{\partial t}+\mathcal{M}_1\right)+r\left(x\frac{\partial}{\partial x}-1\right)\right]f^{(2)} \tag{4.139}$$

$$\left(\frac{\partial f^{(1)}}{\partial y}\right)^{-1}\mathcal{M}_2 f^{(1)}(t,X(t),Y(t)) = \left(\frac{\partial f^{(2)}}{\partial y}\right)^{-1} \times \mathcal{M}_2 f^{(2)}(t,X(t),Y(t)) \tag{4.140}$$

where

$$\mathcal{M}_2 = \left[\left(\frac{\partial}{\partial t}+\mathcal{M}_1\right)+r\left(x\frac{\partial}{\partial x}-1\right)\right], \tag{4.141}$$

is the standard Black–Scholes differential operator with volatility parameter $f(y)$ plus a second order term from the $Y$ diffusion process. Now the left hand side of Eq. (4.140) contains terms depending on $T_1$ and not $T_2$, similarly for the right hand side of the same equation. Thus both sides must be equal to a function that does not depend on expiration date. This function is denoted as

$$\alpha(m-y)-\beta\left(\rho\frac{(\mu-r)}{f(y)}+\gamma(t,x,y)\sqrt{1-\rho^2}\right)=\alpha(m-y)-\beta\Lambda(t,x,y). \tag{4.142}$$

Here $\gamma(t,x,y)$ is an arbitrary function and $\Lambda(t,x,y)$ is a short hand notation for the second coefficient in the LHS of Eq. (4.142).

The pricing function $P(t,X(t),Y(t))$, with the dependence on suppressed expiry date must satisfy the partial differential equation,

$$\left[\left(\frac{\partial}{\partial t}+\mathcal{M}_1\right)+r\left(x\frac{\partial}{\partial x}-1\right)+\alpha(m-y)\frac{\partial}{\partial y} -\beta\left(\rho\frac{(\mu-r)}{f(y)}+\gamma(t,x,y)\sqrt{1-\rho^2}\right)\frac{\partial}{\partial y}\right]P=0. \tag{4.143}$$

This may be written in a more compact form using equation Eq. (4.141) and Eq. (4.142) as

$$\left[\mathcal{M}_2+\alpha(m-y)\frac{\partial}{\partial y}-\beta\Lambda(t,x,y)\right]P=0. \tag{4.144}$$

The terminal condition is $P(T,x,y)=h(x)$, and since $Y(t)$ is an OU process, the $y$ domain is $(-\infty,\infty)$.

Now from Eq. (4.143) we can group the differential operator to isolate various operators. We recall that Eq. (4.143) may be written as

$$\left[\frac{\partial}{\partial t}+\frac{1}{2}\left((f(y)x)^2\frac{\partial^2}{\partial x^2}+\beta f(y)x\rho\frac{\partial^2}{\partial x\partial y}+\beta^2\frac{\partial^2}{\partial y}\right)\right.$$
$$\left.+r\left(x\frac{\partial}{\partial x}-1\right)+(\alpha(m-y)-\beta\Lambda(t,x,y))\frac{\partial}{\partial y}\right]P=0. \tag{4.145}$$

Hence we may break up the operator such that,

$$\mathcal{L}P = \left[\mathcal{L}_{\text{BS}}(f(y))+\mathcal{L}_{\text{correlation}}+\mathcal{L}_{\text{OU}}+\mathcal{L}_{\text{premium}}\right]P=0, \tag{4.146}$$
$$\mathcal{L}_{\text{BS}}(f(y)) = \frac{\partial}{\partial t}+\frac{1}{2}(f(y)x)^2\frac{\partial^2}{\partial x^2}+r\left(x\frac{\partial}{\partial x}-1\right), \tag{4.147}$$
$$\mathcal{L}_{\text{correlation}} = \beta f(y)x\rho\frac{\partial^2}{\partial x\partial y}, \tag{4.148}$$
$$\mathcal{L}_{\text{OU}} = \alpha(m-y)\frac{\partial}{\partial y}+\frac{1}{2}\beta^2\frac{\partial^2}{\partial y}, \tag{4.149}$$
$$\mathcal{L}_{\text{premium}} = -\beta\Lambda(t,x,y)\frac{\partial}{\partial y}=-\beta\left(\rho\frac{(\mu-r)}{f(y)}+\gamma(t,x,y)\sqrt{1-\rho^2}\right)\frac{\partial}{\partial y}. \tag{4.150}$$

The Lagrangian $\mathcal{L}_{\text{BS}}(f(y))$ is the Black–Scholes operator with volatility level $f(y)$, the second term, $\mathcal{L}_{\text{correlation}}$ is due to the correlation; $\mathcal{L}_{\text{OU}}$ is the infinitesimal generator of the OU process $Y(t)$; $\mathcal{L}_{\text{premium}}$ is due to the market price of volatility risk.

The function $\gamma(t,x,y)$ is the risk premium factor from the second source of randomness $Z(t)$ that drives the volatility. In the case that there is a perfect correlation we have $|\rho|=1$ and the $Z(t)$ term does not appear. Hence

$$dP(t,X(t),Y(t)) = \left[\frac{\mu-r}{f(y)}\left(xf(y)\frac{\partial P}{\partial x}+\beta\rho\frac{\partial P}{\partial y}\right)+rP+\gamma\beta\sqrt{1-\rho^2}\frac{\partial P}{\partial y}\right]dt$$
$$+ \left(xf(y)\frac{\partial P}{\partial x}+\beta\rho\frac{\partial P}{\partial y}\right)dW(t)+\beta\sqrt{1-\rho^2}\frac{\partial P}{\partial y}\,dZ(t). \tag{4.151}$$

From this expression we see that an infinitesimal fractional increase in the volatility risk $\beta$ the infinitesimal rate of returns on the option by $\gamma$ times that fraction in addition to the increase from the excess returns to risk ratio $(\mu-r/f(y))$.

### 4.4.2 Pricing With Equivalent Martingale Measure

This is an alternative derivative for model defined in Eq. (4.118). Suppose that there is an equivalent martingale measure $\mathcal{P}^\star$ under which the discounted price $\tilde{X}=e^{-rt}X(t)$ is a martingale. Then we may rewrite,

$$d\tilde{X}(t)=(\mu-r)\tilde{X}(t)\,dt+\sigma\tilde{X}(t)\,dW(t), \tag{4.152}$$

in such a way that the drift term is absorbed into the martingale term,

$$d\tilde{X}(t) = \sigma\tilde{X}(t)\left[dW(t) + \frac{(\mu - r)}{\sigma}\,dt\right], \tag{4.153}$$

we set

$$\theta = \frac{(\mu - r)}{\sigma} \tag{4.154}$$

called the market price of asset risk (Fouque *et al.* 2000) and define,

$$\tilde{W}(t) = W(t) + \int_0^t \theta\,ds = W(t) + \theta t, \tag{4.155}$$

so that

$$d\tilde{X}(t) = \sigma\tilde{X}(t)\,d\tilde{W}(t). \tag{4.156}$$

If we introduce the random variable $\xi^\theta(T)$ defined by,

$$\xi^\theta(T) = \exp\left[-\theta W(T) - \frac{1}{2}\theta^2 T\right], \tag{4.157}$$

then we see that the conditional expectation, $E$, with respect to the $\sigma$–algebra $\mathcal{F}_t$ is given by

$$\mathrm{E}\left[\xi^\theta(T)|\mathcal{F}_t\right] = \exp\left[-\theta W(T) - \frac{1}{2}\theta^2 T\right] = \xi^\theta(T) \quad \text{for} \quad 0 \le t \le T, \tag{4.158}$$

which defines a martingale denoted by $\left(\xi^\theta(T)\right)_{0\le t\le T}$. We now introduce the probability measure $\mathcal{P}^\star$ that is equivalent to $\mathcal{P}$, which means that it has the same null sets $\mathcal{P}^\star$, moreover $\mathcal{P}^\star$ has the density $\xi^\theta(T)$ with respect to $\mathcal{P}$,

$$d\mathcal{P}^\star = \xi^\theta(T)\,d\mathcal{P}. \tag{4.159}$$

In this case the expectation with respect to $\mathcal{P}^\star$ for any integrable random variable is given by

$$\mathrm{E}^\star[Z] = \mathrm{E}\left[\xi^\theta Z(T)\right], \tag{4.160}$$

it can also be shown that for any adapted and integrable process $(Z(t))$,

$$\mathrm{E}^\star[Z(t)|\mathcal{F}_s] = \frac{1}{\xi^\theta(T)}\mathrm{E}\left[\xi^\theta(T)Z(t)|\mathcal{F}_s\right], \quad \text{for any} \quad 0 \le s \le t \le T. \tag{4.161}$$

The process $\left(\xi^\theta(T)\right)_{0\le t\le T}$ is called the Radon–Nikodym process (Fouque *et al.* 2000). By the Girsanov theorem (Øksendal 2003, Karatzas and Shreve 1988, Shreve 2004), see App. A.3.1, $(\tilde{W})$ and any shift of the second independent Brownian motion of the form

$$\tilde{Z} = Z(t) + \int_0^t \gamma\,ds, \tag{4.162}$$

which will not change the drift of $\tilde{X}(t)$, are independent standard Brownian motions under a measure $\mathcal{P}^{\star(\gamma)}$ defined by

$$\begin{aligned}\frac{d\mathcal{P}^{\star(\gamma)}}{d\mathcal{P}} &= \exp\left[-\frac{1}{2}\int_t^T \left(\left(\theta_s^{(1)}\right)^2 + \left(\theta_s^{(2)}\right)^2\right) ds\right. \\ &- \left.\int_t^T \theta_s^{(1)}\, dW_s - \int_t^T \theta_s^{(2)}\, dZ(s)\right] \qquad (4.163)\end{aligned}$$

$$(4.164)$$

where

$$\theta_s^{(1)} = \frac{\mu - r}{f(Y(t))}, \qquad \text{and} \qquad \theta_s^{(2)} = \gamma(t). \qquad (4.165)$$

Here $\gamma(t)$ is any adapted process. In order to make the measure $\mathcal{P}^{\star(\gamma)}$ a well defined probability measure it may be assume that it is distributed according to $\left(\frac{\mu-r}{f(Y(t))}, \gamma(t)\right)$. Then under $\mathcal{P}^{\star(\gamma)}$, the stochastic differential equation Eq. (4.126) becomes,

$$\begin{aligned} dX(t) &= rX(t)\,dt + \sigma(t)X(t)\,d\tilde{W}(t) = (\mu - \lambda_w\sigma(t))\,X(t)\,dt + \sigma(t)X(t)\,d\tilde{W}(t) \\ \sigma(t) &= f(Y(t)), \quad \text{with} \quad \lambda_w = \frac{\mu - r}{\sigma(t)} \qquad (4.166) \\ dY(t) &= \left[\alpha(m - Y(t)) - \beta\left(\rho\frac{\mu - r}{\sigma(t)} + \gamma(t)\sqrt{1-\rho^2}\right)\right] dt \\ &+ \beta(\rho\, d\tilde{W}(t) + \sqrt{1-\rho^2}\, d\tilde{Z}(t)). \end{aligned}$$

Any allowable choice of $\gamma$ leads to an expression equivalent martingale measure $\mathcal{P}^{\star(\gamma)}$ and the possible no arbitrage derivatives prices

$$V(t) = \mathrm{E}^{\star(\gamma)}\left[e^{-r(T-t)}H|\mathcal{F}_t\right], \qquad (4.167)$$

where $H$ is the payoff.

The process $(\gamma(t))$ is called the risk premium factor or the market price of volatility risk from the second source of randomness $Z$ that drives the volatility, and parametrize the space of equivalent martingale measures $\{\mathcal{P}^{\star(\gamma)}\}$ when $\gamma = \gamma(t, X(t), Y(t))$.

With this model it is possible to hedge one derivative contract $f^{(1)}$ with stock and another derivative security $f^{(2)}$ as in Eq. (4.120). This leads to the same hedging ratios $c(t)$ and $a(t)$ as it did for the model of Eq. (4.126). Since $\gamma$ may be any arbitrary function the hedging ratios are non-unique.

## 4.5 Connection With Partial Differential Equations

There are several ways to compute a derivative security:

- The use of discretisation methods discussed in Section 4.2 and application of a Monte Carlo method to calculate $Y(\tau_n)$ for many realizations of $X(t)$, thus generating sample paths. One then computes the average $h(X(t))$, discussed below, over all these simulations to get an approximate value for the derivative price $g(t,x)$, which is the risk neutral expected discounted payoff.
- Numerically solve a partial differential equation.
- Use binomial trees, discrete models.
- Use a path integral, which is an integral based formulation, to calculate the discounted payoff (see Chapter 6).

In this section we focus on the second case, because most stochastic differential equations are solved using either Kolmogorov's equation (that is the forward and/or backward equations) or the Feynman–Kac formula. The latter is the starting point of the path integral technique used for option pricing.

In Section 4.1.3, the definition of an Îto process was given by Definition 4.1.6, Eq. (4.19), which is a special case. We now consider the most general form of a stochastic differential equation which has the form following the notation of Eq. (1.15),

$$dX(t) = \alpha(t, X(t))dt + \beta(t, X(t))dW(t), \tag{4.168}$$

which has a general solution given by Eq. (1.16) or Eq. (4.44). Here $\alpha(t, X(t))$ and $\beta(t, X(t))$ are given functions called the drift and diffusion respectively. To obtain a solution one needs to specify an initial condition of the form $X(t_0) = x$ when $t_o \geq 0$ and $x \in R$ are specified.

In general stochastic differential equations are in general difficult solve, but a one dimensional linear stochastic differential can be solved explicitly. The stochastic differential equation, which has the form

$$dX(t) = (a(t) + b(t)X(t))\, dt + (\gamma(t) + \sigma(t)X(t))\, dW(t), \tag{4.169}$$

where $a(t), b(t), \gamma(t)$ and $\sigma(t)$ are non–random functions of time, can be solved explicitly. The solution of this equation can be obtained using the definition for the processes

$$Z(t) = \exp\left\{\int_{t_0}^{t} ds\left(b(s) - \frac{1}{2}\sigma^2(s)\right) + \int_{t_0}^{t} dW(s)\sigma(s)\right\} \tag{4.170}$$

$$Y(t) = Y(0) + \int_{t_0}^{t} ds\frac{a(s) - \sigma(s)\gamma(s)}{Z(s)} + \int_{t_0}^{t} dW(s)\frac{\gamma(s)}{Z(s)}, \tag{4.171}$$

with $Z(t_0) = 1$. If we apply the one dimensional Îto–Doeblin formula for an Îto process, Theorem 4.1.7, Eq. (4.23), where $f(\xi) = e^{\xi}$, $f_{\xi}(\xi) = e^{\xi}$ and $f_{\xi\xi}(\xi) = e^{\xi}$ with

$$\xi(t) = \int_{t_0}^{t} ds\left(b(s) - \frac{1}{2}\sigma^2(s)\right) + \int_{t_0}^{t} dW(s)\sigma(s). \tag{4.172}$$

In this case we see that for the stochastic differential, $Z(t) = \exp(\xi(t))$,

$$\begin{aligned} dZ(t) &= f_{\xi}d\xi(t) + \frac{1}{2}f_{\xi\xi}d\xi(t)d\xi(t) \\ &= Z(t)\left(\sigma(t)dW(t) + \left(b(t) - \frac{1}{2}\sigma^2(t)\right)dt\right) + \frac{1}{2}Z(t)\sigma^2(t)dt \\ &= b(t)Z(t)dt + Z(t)\sigma(t)dW(t). \end{aligned} \tag{4.173}$$

Similarly the stochastic differential form of $Y(t)$, Eq. (4.171),

$$dY(t) = \left[\frac{a(t) - \sigma(t)\gamma(t)}{Z(t)}\right]dt + \frac{\gamma(t)}{Z(t)}dW(t). \tag{4.174}$$

Now using Îto product rule, Corollary 4.1.10, Eq. (4.43) for $X(t) = Y(t)Z(t)$, that is

$$d\left(X(t)Y(t)\right) = Y(t)dX(t) + X(t)dY(t) + dY(t)dX(t) \tag{4.175}$$

$$= \left(a(t) + b(t)X(t)\right)dt + \left(\gamma(t) + \sigma(t)X(t)\right)dW(t), \tag{4.176}$$

the product of these two stochastic differential equation solves the stochastic differential equation, Eq. (4.169), where $a(t), b(t), \gamma(t)$ and $\sigma(t)$ are adapted random process.

In order for the Markov property to hold the only randomness, which is allowed in Eq. (4.168), is the randomness belonging to the solution of $X(t)$ and the driving Brownian motion $W(t)$.

### The Markov Property

In the case where we have a stochastic differential equation, Eq. (4.168), if we let $0 \leq t \leq T$ being given and let $h(y)$ be a Borel–measurable function, then we can

denote the expectation of $h(X(T))$, where $X(t)$ is the solution to Eq. (4.168), with initial condition $X(t_0) = x$, by a non–random Borel–measurable function $g(t,x)$ of two dummy variables $t$ and $x$. That is

$$g(t,x) = \mathrm{E}^{t,x}\left[h(X(T)\right], \tag{4.177}$$

which translates into the following theorem:

**Theorem 4.5.1 (Markov property)** *Let $X(t)$, $t \geq 0$ be a solution to the stochastic differential equation, Eq. (4.168), with initial condition given at time 0. Then for $0 \leq t \leq T$*

$$\mathrm{E}^{t,x}\left[h(X(T)|\mathcal{F}_t\right] = g(t, X(t)). \tag{4.178}$$

When we are computing the expected value in Eq. (4.178) the only valuable piece of information is the value of $X(T)$. This means that $X(T)$ is a Markov process. This is highlighted in the following Corollary:

**Corollary 4.5.2** *Solutions to stochastic differential equations are Markov processes.*

In most cases the general solution for stochastic differential equation can only be obtained by using partial differential equation. Partial differential can be used to obtained the expected value and the transition probabilities for that particular stochastic differential equation. This is done via the Feynman–Kac and the Kolmogorov equations (both the forward and backward equations) respectively.

### 4.5.1 The Feynman-Kac Formula

The *Feynman–Kac formula* (Shreve 2004, Karatzas and Shreve 1988) is an important equation because it allows one to relate stochastic differential equation and partial differential equation. The equation gives the expected value, which is used to calculate price function. It is also used in the path integral formulation, see Chapter 6

**Theorem 4.5.3 (Feynman–Kac formula)** *Consider the stochastic differential equation,*

$$dX(t) = \alpha(t, X(t))dt + \beta(t, X(t))dW(t). \tag{4.179}$$

*Let $h(y)$ be a Borel–measurable function. Fix $T > 0$ and let $t \in [0, T]$. Define the function*

$$g(t,x) = \mathrm{E}^{t,x}\left[h(X(T))\right] = \int dX(T)h(X(T))p(X(T), T|x(t), t), \tag{4.180}$$

*then $g(t,x)$ satisfies the partial differential equation*

$$g_t(t,x) + \alpha(t,x)g_x(t,x) + \frac{1}{2}\beta^2(t,x)g_{xx}(t,x) = 0 \tag{4.181}$$

*and the terminal condition*

$$g(T,x) = h(x) \quad , \quad \forall x \in R. \tag{4.182}$$

We do not attempt to prove the Feynman–Kac formula in this work, since it is not the aim of this thesis. The proof is nevertheless very instructive and can be found in many textbooks, for example Karatzas and Shreve (1988), Shreve (2004), or Øksendal (2003). Note however that the proof does depend on the following lemma:

If we let $0 \leq s \leq t \leq T$ the Markov property, Theorem 4.5.1, implies that if we have

$$\mathrm{E}\left[h(X(T))|\mathcal{F}_s\right] = g(s,X(s)), \tag{4.183}$$

$$\mathrm{E}\left[h(X(T))|\mathcal{F}_t\right] = g(t,X(t)), \tag{4.184}$$

$$\tag{4.185}$$

then

$$\begin{aligned} \mathrm{E}\left[g(t,X(t))|\mathcal{F}_s\right] &= \mathrm{E}\left[\mathrm{E}\left[h(X(T))|\mathcal{F}_t\right]|\mathcal{F}_s\right] \\ &= \mathrm{E}\left[h(X(T))|\mathcal{F}_s\right] \\ &= g(s,X(s)). \end{aligned} \tag{4.186}$$

This proves the following lemma:

**Lemma 4.5.4** *Let $X(t)$ be a solution to the stochastic differential equation defined as in Eq. (4.168), with initial condition given at time 0. Let $h(y)$ be a Borel–measurable function, fix $T > 0$ and let $g(t,x)$ be given by Eq. (4.177), then the stochastic process*

$$g(t,X(t)) \quad \text{for} \quad 0 \leq t \leq T, \tag{4.187}$$

*is a martingale.*

This lemma is used in the proof of the Feynman–Kac theorem, Theorem 4.5.3, and also for the proof of the discounted Feynman–Kac formula.

The discounted is the equation that is used in the evaluation of option pricing and is referred to in the section on path integral technique, Sec.6.3. Here again we only state the theorem without proof.

**Theorem 4.5.5 (Discounted Feynman–Kac formula)** *Consider the stochastic differential equation, Eq. (4.168). Let $h(y)$ be a Borel–measurable function. Let $r$ be a constant. Fix $T > 0$ and let $t \in [0, T]$. Define the function*

$$\begin{aligned} f(t,x) &= \mathrm{E}^{t,x}\left[e^{-r(T-t)}h(X(T))\right] \\ &= \int dX(T)e^{-r(T-t)}h(X(T))p(X(T),T|x(t),t), \end{aligned} \tag{4.188}$$

*then $f(t,x)$ satisfies the partial differential equation*

$$f_t(t,x) + \alpha(t,x)f_x(t,x) + \frac{1}{2}\beta^2(t,x)f_{xx}(t,x) = rf(t,x) \tag{4.189}$$

*and the terminal condition*

$$f(T,x) = h(x) \quad , \quad \forall x \in R. \tag{4.190}$$

Here Eq. (4.189) is used to solve interest rate models, such as the Hull–White interest rate model and bond models, it can also be generalized to the multi–dimensional case where $\widetilde{W}(t)$ is defined as Sec. 4.1.4, Eq. (4.26).

We now move to the Kolmogorov equations which unlike the Feynman–Kac formula that returns the expected value, both the forward and backward Kolmogorov equations return the transition probabilities for given stochastic differential equations. The forward equation is also called the Fokker–Planck equation and it indicates what will happen in the future, whereas the backward Kolmogorov gives information on the initial conditions that will lead to a certain state. They both give transition probabilities for the system. In our context transition probabilities specify the probability of a process being at some point in the future or the past, given its present position.

### 4.5.2 The Forward Kolmogorov Equation (The Fokker–Planck Equation)

Our starting point here is Eq. (4.168), which is the most general of stochastic differential equation.

If we assume that a process has arbitrary initial value at time $X(t_0) = x$ and evolve forward so that at each time, $T > t$, which can be a positive number but cannot be less than or equal to 0. For $s$ and $t$ such that $0 \leq s \leq t \leq T$, let $p(y(T),T|x(t),t)$ be the transition probability of obtaining $y(T)$ at time $T$ from the state at $x(t)$ at time

$t$, with $p(y(T), T|x(t), t) = 0$ for $0 \leq s \leq t \leq T$ and $y(T) \leq 0$ (i.e., if we solve the equation with initial condition $X(t_0) = x$, then the random variable $X(T)$ has density $p(y(T), T|x(t), t)$ in the $y$ variable). Here, because we are interested in obtaining the transition probability for future times, the variables $t$ and $x$ are held constant as opposed to the backward Kolmogorov equation where $T$ and $x(T)$ are held constant.

Let $b$ be a positive constant and let $h_b(y)$ be a Borel–measurable function with continuous first and second derivative such that $h_b(x) = 0$ for all $x \leq 0$, $h_b'(x) = 0$ for all $x \geq b$ and $h_b(b) = h_b'(b) = 0$. Let $X(t)$ be the solution to the stochastic differential equation, Eq. (4.168), with initial condition $X(t_0) = x \in (0, b)$. Using Îto–Doeblin formula, Eq. (4.23), to compute $dh_b(y)$ with $y = X(t)$,

$$\begin{aligned} dh_b(y) &= h_b'(y)dX(t) + \frac{1}{2}h_b''(y)dX(t)dX(t) \\ &= h_b'(y)dX(t) + \frac{1}{2}\beta^2(t,y)h_b''(y)dt \\ &= h_b'(y)\Big(\alpha(t, X(t))dt + \beta(t, X(t))dW(t)\Big) + \frac{1}{2}\beta^2(t,y)h_b''(y)dt \\ &= \left[\alpha(t, X(t))h_b'(y) + \frac{1}{2}\beta^2(t,y)h_b''(y)\right]dt + \beta(t, X(t))h_b'(y)dW(t). \end{aligned} \tag{4.191}$$

If we let $0 \leq t \leq T$ be given and integrate Eq. (4.191) from $t$ to $T$ and take the expectation value, which is given by

$$\mathrm{E}\left[h_b(y)\right] = \int_t^T dy h_b(y)p(y, T|x, t), \tag{4.192}$$

and assuming that $X(t)$ has density of $p(y, T|x, t)$, then Eq. (4.191) looks like (with the expected value the Brownian motion given by $\mathrm{E}\left[W(t)\right] = 0$),

$$\begin{aligned} \int_0^b dy\, h_b(y)p(y, T|x, t) &= h_b(x) + \int_t^T ds \int_0^b dy\, \alpha(t,y)h_b'(y)p(y, T|x, t) \\ &+ \int_t^T ds \int_0^b dy \frac{1}{2}\beta^2(t,y)h_b''(y)p(y, T|x, t). \end{aligned} \tag{4.193}$$

Integrating the right hand side by parts once for the integral containing the $h_b'(y)$ and twice for the integral containing the $h_b''(y)$ with respect to $y$ leads to:

$$\begin{aligned} \int_0^b dy\, h_b(y)p(y, T|x, t) &= h_b(x) + \int_t^T ds \left\{ \Big[\alpha(t,y)h_b(y)p(y, T|x, t)\Big]_0^b \right. \\ &- \left. \int_0^b dy \frac{\partial}{\partial y}\Big[\alpha(t,y)p(y, T|x, t)\Big]h_b(y)\right\} \\ &+ \int_t^T ds \int_0^b dy \frac{\partial^2}{\partial y^2}\left[\frac{1}{2}\beta^2(t,y)p(y, T|x, t)\right]h_b(y). \end{aligned} \tag{4.194}$$

If we now differentiate Eq. (4.194) with respect to $T$ to obtain

$$\int_0^b dy \left[\frac{\partial}{\partial T}p(y,T|x,t) + \frac{\partial}{\partial y}\left[\alpha(t,y)p(y,T|x,t)\right] - \frac{\partial^2}{\partial y^2}\left[\frac{1}{2}\beta^2(t,y)p(y,T|x,t)\right]\right] h_b(y) = 0. \tag{4.195}$$

Hence we have

$$\frac{\partial}{\partial T}p(y,T|x,t) + \frac{\partial}{\partial y}\left[\alpha(t,y)p(y,T|x,t)\right] - \frac{\partial^2}{\partial y^2}\left[\frac{1}{2}\beta^2(t,y)p(y,T|x,t)\right] = 0. \tag{4.196}$$

This partial differential equation, Eq. (4.196), is the forward Kolmogorov equation also known as the Fokker–Planck equation. The Fokker–Planck equation has been the subject of many discussion, see for example Risken (1984), and is a very important equation in statistical physics. This equation will be used latter in this thesis for a non–Gaussian option pricing model, which is based on a Tsallis distribution (Tsallis 1988a, Curado and Tsallis 1991a).

The backward version of the Kolmogorov equation is very similar to the forward one, but instead of keeping $x(t)$ and $t$ constant, $y = X(T)$ and $T$ are kept constant.

### 4.5.3 The Backward Kolmogorov Equation

The backward Kolmogorov equation, under the same conditions as the forward one, Sec. 4.5.2, says that the transition $p(y,T|x,t)$ satisfies

$$\frac{\partial}{\partial t}p(y,T|x,t) + \alpha(t,x)\frac{\partial}{\partial x}p(y,T|x,t) + \frac{1}{2}\beta^2(t,x)\frac{\partial^2}{\partial x^2}p(y,T|x,t) = 0. \tag{4.197}$$

This equation gives the initial conditions that will lead to a certain state. The backward Kolmogorov equation is very useful on getting information about the system and will be used later in the text, especially in Chapter 6.

## 4.6 Chapter Summary

In this chapter we focused on two main aspects of the stochastic calculus the first one is its definition and how it is constructed from the Brownian motion, explained in

Chapter 3. Using these definitions and theorems we made contact with one of the most important equation, namely the Îto–Doeblin equation.

The second aspect is related to the evaluation of stochastic differential equation. In general SDE cannot be evaluated analytically directly and one needs to have some methods for approximating these equation. This is normally carried out using the numerical approaches such as the ones derived in Sec. 4.2. These methods can be the only way to obtain some insights into the SDE solutions. We implemented three methods which we have tested on a particular example where the explicit solution is known. These methods are to be used for model construct checking purposes. The strong 1.5 Taylor expansion is only of order 1.5 but there are higher order method that have been implemented by other authors. The reader is invited to see the pioneering work done by Burrage *et al.* (2000) on the numerical solutions of SDE or by Kloeden and Platen (1992).

Alternatively in order to obtain the transitions probabilities on needs to make contact with partial differential equations that are associated with these SDE. This is done using the Kolmogorov equations. Finally the Feynman–Kac formula is the equation that is used in the evaluation of the discounted option price.

In the next chapter we go into depth in describing the different types of options. Before moving to option pricing using path integrals.

Chapter 5

# Option pricing and derivatives

IN this chapter we describe some of the various options and derivatives available on the market these days. These financial instruments are widely used in today markets and the list of such financial products is still growing. Financial institutions can be very creative and usually fabricate their own products, which could involve combinations of future contracts and options. Options are attractive products because they give the right but not the obligation to exercise the option at a given time in the future. This flexibility comes at a price. Each option will exercise differently, hence the mathematics to describe these products is different in each case. This chapter explores this concept in more detail by reviewing the most popular option and most fundamental option pricing models, which was developed in the early 70s by Black-Scholes. This model is commonly known as the Black-Scholes-Merton model

## 5.1 Some Concepts

Financial traders typically deal with a special class of financial contracts called derivatives. A derivative is a financial product whose price depends upon the price of another financial product. These derivatives include, for example, *forward contracts*, *futures*, *options* and *swaps*. Here we are only going to focus on the first three of these.

A *forward contract* is when one of the parties agrees to buy, for a given amount, an asset at a specified price (called the forward price or the delivery price $K$) on a specified future date (the delivery date $T$). The other party agrees to sell the specified amount of the asset at the delivery price on the delivery date.

*Futures* are a forward contract traded on by exchange. The exchange is done through an exchange institution or a delivery house.

An *option* is a financial contract that gives the holder the right to exercise a given action (usually buying or selling) on an underlying asset of time $T$ and at a price $K$. The price $K$ is called the strike price and $T$ is called the expiration date, the exercise date or the date of maturity. There are two types of options, namely there is the *call option*, which gives the holder the right to buy the underlying asset by a certain date for a certain price. A *put option* is the reverse of a *call option*, and thus applies to selling rather than buying.

There are a few different sorts of options and each of them carry a different name, just to cite a few:

- **European options:** can only be exercised on the expiration date itself [12]
- **American options:** can be exercised at any time up to the expiration date.
- **Bermudian** and **mid–Atlantic options:** can be exercised at several moments between purchase of the option and a given expiry date.
- **Exotic** or **path dependent options:** these options have values that depend on the history of an asset price not just its value on exercise. An example would be an option to purchase an asset for the arithmetic average value of that asset over the month before expiring.

[12] Note that the terms "American" and "European" do not refer to the location of the option or the exchange. Some options trading on North American exchanges are European.

- **Barrier option:** the option can either come into existence or become worthless if the underlying asset reaches some prescribed value before expiring.
- **Asian option:** the price depends on some form of average.
- **LookBack options:** the price depends on asset price maximum or minimum.

A more complete list may be found in Appendix C.1. Options are not the only product that financial institutions trade on a daily basis.

*Swaps* involve a private agreement between two parties to exchange cash flows at a certain times in the future according to some prearranged formula.

In general the quantities that are of financial interest are those that determine the underlying price of some asset. The scale used are usually given in units (price, currency, index etc.) which are known to undergo all sorts of fluctuations as time evolves. Moreover transactions occur at random times with random intensities. It is therefore important to carefully choose the variables that will be studied.

In general the price of the asset is not constant in time, this may fluctuate for many reasons, reasons such as inflation, economic growth/recession, geopolitical situation of certain countries, interest rates, consumer confidence, investor confidence, which is itself related to the previous reasons, are only a few of the main reasons why the asset price may fluctuate over time.

Let us define $Y(t)$ as the price of a financial asset at time $t$. Since $Y(t)$ fluctuates randomly it makes it an appropriate stochastic variable to be investigated. Using $Y(t)$ one may construct other random variables for example:

1. One can study the *price changes* as a function of time, $t$,

$$Z(t) = Y(t + \Delta t) - Y(t) \tag{5.1}$$

   where $\Delta t$ is some time difference between two asset price, $\delta t$ could be as little as 1 second to several days.

2. One could also analyze *returns* defined as

$$R(t) = \frac{Y(t + \Delta t) - Y(t)}{Y(t)} = \frac{Z(t)}{Y(t)}. \tag{5.2}$$

   Returns provide a direct percentage of gain or loss in a given time period, but become a less reliable variable when examined over a long period of time.

3. Another interesting variable that may be studied is the *successive difference* of the natural logarithm of price

$$S(t) = \ln Y(t + \Delta t) - \ln Y(t). \tag{5.3}$$

We firstly refine the terms used in this context and then go through specific options. In general and so far we define and have defined $S(t)$ as the price of the underlying asset at time $t$ traded in the spot market. We denote $T$ as the expiration date of the option and denote by $\tau = T - t$ the owner of the contract. Also let $K$ denote the strike price and $\mathcal{O}[S(T)]$ denote the *pay–off* of the option.

Now depending on the type of the option—that is depending if it is an European, American or any other—the payoff of the option will be exercised differently and at different times. The pay–off is the amount of money the investor receives at time $T$ when the option expires.

The next chapter on path integrals, Chapter 6, is mostly about applying the path integral technique to option pricing— because of the nature of the problem we only consider European options, since these are the simplest ones. Path dependent options are not treated in this discourse, since the method has to be tested on simple European options first. As a result we spend more time describing the European option and give only an introductory discussion on the other types of option. Further details on exotic options may be found in Shreve (2004) for example.

## 5.2 European Options

As already mentioned, a *European option* has the property that if the price of the underlying asset at expiration $S(T)$ is smaller than $K$ it means that the owner of the option can purchase the underlying asset in the spot market for less than the strike price $K$. In such case it is not the best strategy to exercise the option. On the other hand if the price of the underlying asset at expiration is higher than $K$ then it may be optimal to exercise the option. In that case the owner will only need to pay $K$. In reality the owner does not pay anything. The owner receives what is called a cash settlement of $S(T) - K$. If the option owner wishes to actually acquire the underlying asset he/she can purchase it in the spot market. The cost will then be $S(T) - (S(T) - K) = K$. The value of the option at time $T$ is called the pay–off, that is $\mathcal{O}[S(T)]$. Mathematically a *call option* may

be expressed as

$$\mathcal{O}_c[T, S(T)] = \begin{cases} 0 & \text{if } S(T) \leq K \\ S(T) - K & \text{if } S(T) > K, \end{cases} = \max(S(T) - K, 0). \tag{5.4}$$

This is often written as $\max(S(T) - K, 0) \equiv (S(T) - K, 0)^+$.

Alternatively, using the same notation, the owner of put options will not exercise the option if $S(T)$ is larger than $K$. That is because if the owner did exercise he/she would receive only $K$ for the underlying asset whereas he/she could get could get $S(T)$ for a sale on the spot market.

Anti–symmetrically if $S(T)$ is smaller than $K$, the owner of the option should exercise it, because he/she can then sell the underlying asset for $K$ that is more than he/she would get in the spot market.

In the case the owner of the option does not own the underlying asset, he/she may have to purchase it first. This is the case for markets where actual delivery is expected. In many markets he/she would get a compensation payment in cash. Mathematically, in the case of a put option we would write the payoff as we did above for the call option, that is

$$\mathcal{O}_p[T, S(T)] = \begin{cases} K - S(T) & \text{if } S(T) < K \\ 0 & \text{if } S(T) \geq K \end{cases} = \max(0, K - S(T)). \tag{5.5}$$

Graphically we would expressed Eq. (5.4) and Eq. (5.5) as in Fig. (5.1). This figure says that in case of a put option the value of the option only takes value when $S(T) < K$ and 0 otherwise. For the call option it is the other way around, that is when $S(T) > K$.

We now turn to the most classical example of the European option evaluation, which is based on geometric Brownian motion.

### 5.2.1 The Black–Scholes–Merton Model

The Black–Scholes–Merton[13] model is based on the geometric Brownian motion which has been described in earlier chapters, and for which the general solution is known.

[13]It is a recognized fact that the original contributors to the model Black and Scholes, but at about the same time and in collaboration with Merton the three collaborated and exchanged ideas on this model. Hence we should refer to this model as the Black–Scholes–Merton and not just the Black–Scholes option pricing model, as is commonly practiced

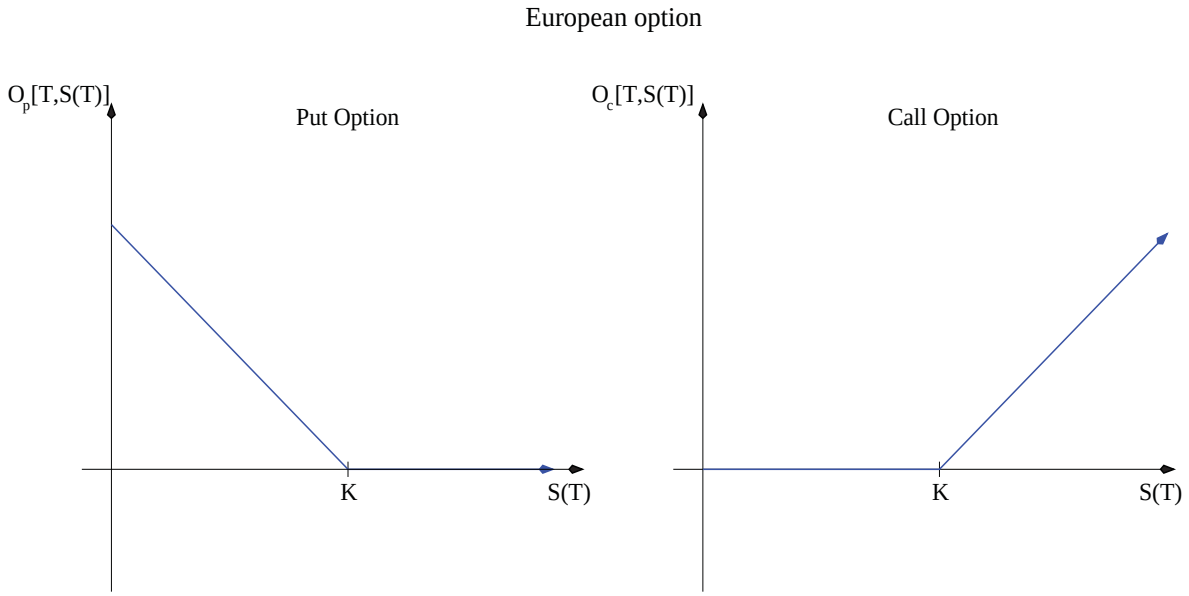


**Figure 5.1. Put and call graph of the payoff function for the European option.** The graphical description of a European option payoff function for the put (left diagram) and a call option (right diagram) The option price would typically follow the blue line from above. Here $K$ represents the strike price of the option.

The general solution of the geometric Brownian motion follows a Gaussian distribution, Eq. (2.2). However as was pointed out in the introduction, see for example Fig. (1.2) where the Gaussian distribution simply does not fit the empirical data for the daily log–returns but a non–Gaussian distribution of the type defined in Eq. (2.8) was more suitable for these fits, see Fig. (2.3). Moreover because of the skewness in the empirical data, distributions such as the one defined by Eq. (2.17) or by Eq. (2.24), which have more parameters to tune the distribution against empirical distribution such as the ones shown in Fig. (2.3) or in Fig. (1.2), are by far more realistic distributions to fit the real data.

Nevertheless the Black–Scholes–Merton model is a fundamental option pricing model and it has been applied to many different types of options, because of its Gaussian structure for which most of the integrals containing such integrands are known.

Let us consider an agent who at each time $t$ has a portfolio valued at $X(t)$. This portfolio invests in a money market account paying a constant rate of interests $r$ and a stock modeled by the geometric Brownian motion

$$dS(t) = \alpha S(t)dt + \sigma S(t)dW(t), \tag{5.6}$$

which has the solution given by Eq. (4.170); that is

$$S(t) = S(0) \exp\left\{\sigma W(t) + \left(\alpha - \frac{1}{2}\sigma^2\right) t\right\}\Big|_{t_0=0}. \tag{5.7}$$

In general, a portfolio is made up of stocks and bonds and maybe some fixed interest cash deposits. In this case suppose that at each time $t$ the investor holds $\Delta(t)$ shares of stock. The position can be random but must be adapted to the filtration associated to the Brownian motion $W(t)$, $t \geq 0$. The remainder of the portfolio is invested in the money market account. The differential $dX(t)$ for the investor portfolio value at each time $t$ is due to two factors, the capital gain $\Delta(t)S(t)$ on the stock position and the interests earnings $r(X(t) - \Delta(t)S(t))dt$ on the cash position, that is

$$\begin{aligned} dX(t) &= \Delta(t)dS(t) - r(X(t) - \Delta(t)S(t))dt \\ &= \Delta(t)\left(\alpha S(t)dt + \sigma S(t)dW(t)\right) - r(X(t) - \Delta(t)S(t))dt \\ &= \left[\Delta(t)\alpha S(t) + r(X(t) - \Delta(t)S(t))\right] + \sigma\Delta(t)S(t)dW(t) \\ &= \left[\Delta(t)\left(\alpha - r\right)S(t) + rX(t)\right]dt + \sigma\Delta(t)S(t)dW(t). \end{aligned} \tag{5.8}$$

This can be understood as follows. Firstly it is the average underlying rate of returns $r$ on the portfolio, which is reflected by the term $rX(t)dt$. Secondly, a risk premium $(\alpha - r)$ for investing in the stocks, is reflected by the term $\Delta(t)\left(\alpha - r\right)S(t)dt$. Thirdly, a volatility term proportional to the size of the stock investment is the term $\sigma\Delta(t)S(t)dW(t)$.

We can now obtain the Îto–Doeblin formula for the discounted stock price $e^{-rt}S(t)$ and the discounted portfolio value of an agent $e^{-rt}X(t)$. The differential for the discounted asset price and value of the portfolio are

$$\begin{aligned} d\left(e^{-rt}S(t)\right) &= -re^{-rt}S(t)dt + e^{-rt}dS(t) \\ &= -re^{-rt}S(t)dt + e^{-rt}\left(\alpha S(t)dt + \sigma S(t)dW(t)\right) \\ &= e^{-rt}\left(\alpha - r\right)S(t)dt + e^{-rt}\sigma S(t)dW(t), \end{aligned} \tag{5.9}$$

and

$$\begin{aligned} d\left(e^{-rt}X(t)\right) &= -re^{-rt}X(t)dt + e^{-rt}dX(t) \\ &= -re^{-rt}X(t)dt + e^{-rt}\left[\Delta(t)\left(\alpha - r\right)S(t) + rX(t)\right]dt \\ &+ e^{-rt}\sigma\Delta(t)S(t)dW(t) \\ &= \Delta(t)d\left(e^{-rt}S(t)\right), \end{aligned} \tag{5.10}$$

respectively.

Let us now consider a European call option that has a payoff as in Eq. (5.4). Black, Scholes, and Merton argued that the value of this call at any time should depend on the time to expiration and on the current value of the stock price, as well as the model parameters $r$ and $\sigma$, and the strike price $K$.

In the Black–Scholes–Merton model, only two of these are variables: that is, the time $t$ and the stock price $S(t)$.

In this context we will denote the value of the call price at time $t$ by $C(t, S(t))$. The value of the option in this case is random. So at an initial time we do not know the future stock price $S(t)$ hence $C(t, S(t))$ is also unknown. On the other hand if we replace $S(t)$ by the dummy variable $x$ that is, $x = S(t)$, $C(t, x)$ becomes a non–random function. The goal is to determine the function $C(t, x)$ so we at least have a formula for the future option values in terms of the future prices.

This is carried out by using the Îto–Doeblin formula for both, $C(t, S(t))$, and the discounted option price $e^{-rt}C(t, S(t))$ to obtain the fundamental Black–Scholes–Merton differential equation. Using the differential form, Eq. (4.25) of Eq. (4.23) in Theorem 4.1.7 on both $C(t, S(t))$ and $e^{-rt}C(t, S(t))$ we obtain

$$dC(t, S(t)) = C_t(t, S(t))dt + C_x(t, S(t))dX(t) + \frac{1}{2}C_{xx}(t, S(t))dS(t)dS(t). \tag{5.11}$$

From Eq. (5.6) we get

$$dS(t)dS(t) = \sigma^2 S^2(t)dt, \tag{5.12}$$

hence

$$\begin{aligned} dC(t, S(t)) &= C_t(t, S(t))dt + C_x(t, S(t))\left(\alpha S(t)dt + \sigma S(t)dW(t)\right) \\ &+ \frac{1}{2}C_{xx}(t, S(t))\sigma^2 S^2(t)dt \\ &= \left[C_t(t, S(t)) + \alpha S(t)C_x(t, S(t)) + \frac{1}{2}\sigma^2 S^2(t)C_{xx}(t, S(t))\right]dt \\ &+ \sigma S(t)C_x(t, S(t))dW(t). \end{aligned} \tag{5.13}$$

Similarly for the discounted option price

$$\begin{aligned} d\left(e^{-rt}C(t,S(t))\right) &= \left(-re^{-rt}C(t,S(t)) + e^{-rt}C_t(t,S(t))\right)dt \\ &+ e^{-rt}C_x(t,S(t))\left[\alpha S(t)dt + \sigma S(t)dW(t)\right] \\ &+ \frac{1}{2}e^{-rt}C_{xx}(t,S(t))\sigma^2 S^2(t)dt \\ &= e^{-rt}\left[-r + \frac{\partial}{\partial t} + \alpha S(t)\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2 S^2(t)\frac{\partial}{\partial x^2}\right]C(t,S(t))dt \\ &+ e^{-rt}\sigma S(t)C_x(t,S(t))dW(t). \end{aligned} \tag{5.14}$$

A hedging portfolio starts with some initial capital $X(0)$ and we invest in the stock and money market account so that the portfolio value $X(t)$ at each time $t \in [0,T]$ agrees with $C(t,S(t))$. This happens if and only if

$$e^{-rt}X(t) = e^{-rt}C(t,S(t)) \quad , \quad \forall t, \tag{5.15}$$

that is when

$$d\left(e^{-rt}X(t)\right) = d\left(e^{-rt}C(t,S(t))\right) \quad , \quad \forall t \in [0,T]. \tag{5.16}$$

Integrating Eq. (5.16), with $X(0) = C(0,S(0))$ we obtain

$$e^{-rt}X(t) - X(0) = e^{-rt}C(t,S(t)) - C(0,S(0)) \quad , \quad \forall t \in [0,T]. \tag{5.17}$$

But $X(0) = C(0,S(0))$, hence we retrieve Eq. (5.15). Now equating Eq. (5.10) and Eq. (5.14), we obtain

$$\begin{aligned} \Delta(t)\left(\alpha - r\right)S(t)dt &+ \sigma\Delta(t)S(t)dW(t) = \\ \left[-r + \frac{\partial}{\partial t} + \alpha S(t)\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2 S^2(t)\frac{\partial^2}{\partial x^2}\right]C(t,S(t))dt &+ \sigma S(t)C_x(t,S(t))dW(t). \end{aligned} \tag{5.18}$$

If we equate terms on both sides of the equation for the $dW(t)$ term we obtain,

$$\Delta(t) = C_x(t,S(t)) \quad , \quad \forall t \in [0,T]. \tag{5.19}$$

This is called the *delta hedging* and $C_x(t,S(t))$ is called the delta. If we now equate the $dt$ term we obtain

$$\left(\alpha - r\right)S(t)C_x(t,S(t)) = \left[-r + \frac{\partial}{\partial t} + \alpha S(t)\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2 S^2(t)\frac{\partial^2}{\partial x^2}\right]C(t,S(t)), \tag{5.20}$$

which simplifies to

$$rC(t,S(t)) = \left[\frac{\partial}{\partial t} + rS(t)\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2 S^2(t)\frac{\partial^2}{\partial x^2}\right]C(t,S(t)) \quad , \quad \forall t \in [0,T]. \tag{5.21}$$

Setting $S(t) = x$, where $x$ is a dummy variable results in the Black–Scholes–Merton partial differential equation, i.e.,

$$rC(t,x) = \left[\frac{\partial}{\partial t} + rx\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2}{\partial x^2}\right] C(t,x) \quad , \quad \forall t \in [0,T]. \tag{5.22}$$

In order to get an analytic solution for the option price, which satisfies the terminal solution

$$C(T,x) = (x-K)^+ \equiv \max\{0, x-K\}, \tag{5.23}$$

one needs to solve this partial differential equation. Here, Eq. (5.22) is a partial differential equation of the type called backward parabolic (see App. B.1.1 or Logan (2004) or Larsson and Thomée (2005) for more details). For this type of equation one needs to define the boundary conditions at $x = 0$ and $x = \infty$. Substituting $x = 0$ into Eq. (5.22) gives

$$rC(t,0) = \frac{\partial}{\partial t}C(t,0) \longrightarrow C(t,0) = C(0,0)e^{rt}. \tag{5.24}$$

Substituting $t = T$ into this equation and using the fact that $C(T,0) = (0-K)^+ = 0$, then we see that $C(0,0) = 0$. Hence,

$$C(t,0) = 0 \quad , \quad \forall t \in [0,T], \tag{5.25}$$

which is the boundary condition at $x = 0$.

As $x \to \infty$ the function grows without bound. In such a case we can impose a boundary condition at $x = \infty$ by specifying the rate of growth. One way to specify a boundary condition at $x = \infty$ for the European call is

$$\lim_{x\to\infty} = \left[C(t,x) - \left(x - e^{-r(T-t)}K\right)\right] = 0 \quad , \quad \forall t \in [0,T]. \tag{5.26}$$

In particular $C(t,x)$ grows at the same rate as $x \to \infty$. To find the solution of Black–Scholes–Merton equation, with terminal solution given by Eq. (5.23), one needs to use the risk neutral pricing formula. That is to calculate the expected value of the discounted payoff under the filtration $\mathcal{F}_t$ of the Brownian motion, i.e.,

$$C(t,S(t)) = E\left[e^{-r(T-t)}\left(S(T)-K\right)^+ | \mathcal{F}_t\right] \tag{5.27}$$

where $S(T)$ is given by Eq. (5.7), but in general is given by

$$S(T) = S(t_0)\exp\left\{\sigma(W(T) - W(t_0)) + \left(r - \frac{1}{2}\sigma^2\right)(T-t_0)\right\}. \tag{5.28}$$

Setting $\tau = T - t_0$ and taking the natural logarithm on both sides of Eq. (5.28) leads to

$$\log(S(T)) = \log(S(t_0)) + \left(r - \frac{1}{2}\sigma^2\right)\tau + \sigma(W(T) - W(t_0)). \tag{5.29}$$

But we know that $S(T)$ has a log–normal distribution,

$$\begin{aligned} \log(S(T)) &\sim \mathcal{N}(\log(S(t_0)) + \left(r - \frac{1}{2}\sigma^2\right)\tau, \sigma) \\ &= \mathcal{N}(\zeta, \sigma) \quad \text{where} \quad \zeta = \log(S(t_0)) + \left(r - \frac{1}{2}\sigma^2\right)\tau, \\ \text{Hence} \quad \frac{\log(S(T)) - \zeta}{\sigma} &\sim \mathcal{N}(0,1). \end{aligned} \tag{5.30}$$

Consequently

$$\begin{aligned} C(t, S(t)) &= E\left[e^{-r\tau}\left(S(T) - K\right)^+ | \mathcal{F}_t\right] \\ &= e^{-r\tau}\int_K^\infty \left(S(T) - K\right) \frac{1}{\sqrt{2\pi}\sigma\sqrt{\tau}S(T)} \exp\left[-\frac{1}{2}\left(\frac{\log(S(T)) - \zeta}{\sigma\sqrt{\tau}}\right)^2\right] dS(T), \end{aligned} \tag{5.31}$$

where we have replaced the volatility parameter $\sigma$ by the $\sigma\sqrt{\tau}$ the volatility over the transition period from $t$ to $T$. To evaluate the integral in Eq. (5.31) it is best to separate the integral into two distinct ones:

$$I_1 = e^{-r\tau}\int_K^\infty \frac{S(T)}{\sqrt{2\pi}\sigma\sqrt{\tau}S(T)} \exp\left[-\frac{1}{2}\left(\frac{\log(S(T)) - \zeta}{\sigma\sqrt{\tau}}\right)^2\right] dS(T), \tag{5.32}$$

$$I_2 = e^{-r\tau}\int_K^\infty \frac{-K}{\sqrt{2\pi}\sigma\sqrt{\tau}S(T)} \exp\left[-\frac{1}{2}\left(\frac{\log(S(T)) - \zeta}{\sigma\sqrt{\tau}}\right)^2\right] dS(T). \tag{5.33}$$

By making the change of variable $y = \log(S(T))$, which means that $dy = (S(T))^{-1}dS(T)$ and $e^y = S(T)$, $I_1$ becomes

$$\begin{aligned}
I_1 &= e^{-r\tau}\int_{\log(K)}^{\infty}\frac{e^y}{\sqrt{2\pi}\sigma\sqrt{\tau}}\exp\left[-\frac{1}{2}\frac{(y-\zeta)^2}{\sigma^2\tau}\right]dy,\\
&= e^{-r\tau}\int_{\log(K)}^{\infty}\frac{1}{\sqrt{2\pi}\sigma\sqrt{\tau}}\exp\left[y-\frac{1}{2}\frac{(y-\zeta)^2}{\sigma^2\tau}\right]dy,\\
&= e^{-r\tau}\int_{\log(K)}^{\infty}\frac{1}{\sqrt{2\pi}\sigma\sqrt{\tau}}\exp\left[\frac{2\sigma^2y\tau-y^2+2y\zeta-\zeta^2}{2\sigma^2\tau}\right]dy,\\
&= e^{-r\tau}\int_{\log(K)}^{\infty}\frac{1}{\sqrt{2\pi}\sigma\sqrt{\tau}}\exp\left[\frac{-\left(y-\left(\sigma^2\tau+\zeta\right)\right)^2+\left(\sigma^2\tau+\zeta\right)\sigma^2\tau}{2\sigma^2\tau}\right]dy,\\
&= e^{-r\tau}\int_{\log(K)}^{\infty}\frac{e^{\frac{1}{2}\sigma^2\tau+\zeta}}{\sqrt{2\pi}\sigma\sqrt{\tau}}\exp\left[\frac{-\left(y-\left(\sigma^2\tau+\zeta\right)\right)^2}{2\sigma^2\tau}\right]dy,\\
&= e^{-r\tau}e^{\frac{1}{2}\sigma^2\tau+\zeta}\int_{\log(K)}^{\infty}\frac{1}{\sqrt{2\pi}\sigma\sqrt{\tau}}\exp\left[\frac{-\left(y-\left(\sigma^2\tau+\zeta\right)\right)^2}{2\sigma^2\tau}\right]dy,\\
&= e^{-r\tau}e^{\frac{1}{2}\sigma^2\tau+\zeta}\mathcal{N}(d_1),
\end{aligned} \tag{5.34}$$

where $\mathcal{N}(d_1)$ is the normal distribution, Eq. (A.15)

$$d_1 = \frac{\log\left(\frac{S(T)}{K}\right)+\left(r+\frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}}. \tag{5.35}$$

The second integral, $I_2$ is simpler to evaluate. With the same change of variables we get

$$\begin{aligned}
I_2 &= -Ke^{-r\tau}\int_{\log(K)}^{\infty}\frac{1}{\sqrt{2\pi}\sigma\sqrt{\tau}}\exp\left[-\frac{1}{2}\frac{(y-\zeta)^2}{\sigma^2\tau}\right]dy,\\
&= -Ke^{-r\tau}\mathcal{N}(d_2)
\end{aligned} \tag{5.36}$$

where

$$d_2 = \frac{\log\left(\frac{S(T)}{K}\right)+\left(r-\frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}}. \tag{5.37}$$

Hence we have for the European call option

$$C(t,S(t)) = e^{-r\tau}e^{\frac{1}{2}\sigma^2\tau+\zeta}\mathcal{N}(d_1) - Ke^{-r\tau}\mathcal{N}(d_2) \tag{5.38}$$

with

$$d_1 = \frac{\log\left(\frac{S(T)}{K}\right)+\left(r+\frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}} \tag{5.39}$$

$$d_2 = \frac{\log\left(\frac{S(T)}{K}\right)+\left(r-\frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}}. \tag{5.40}$$

This is the solution of the European call option for the Black–Scholes–Merton, model for which attracted the Nobel Prize in economics in 1997.

This equation can be easily modeled and the computer code for the Black–Scholes–Merton model can be found in Appendix E.5.1 (routine name is $black\,scholes(\cdots)$).

Table. 5.1 shows the results for the Black–Scholes–Merton model for a European option when the strike price $K = 100.0$, the current price of the option is $S(0) = 0.0$, the interest rate $r = 0.1$, and the volatility $\sigma = 0.3$ for the option price $C(t, S(t))$. Also shown in the table are the Greeks[14], the $\Delta$, $\Gamma$, $\Theta$, $\nu$ and $\rho$. The Black–Scholes–Merton is a very useful model, even though it does not capture the stylized effects of the market it allows one to obtain a closed form solution for the option pricing model written above and also permits explicit computation of the Greeks, and the evaluation of the risk associated with a given option portfolio. As a result, a desirable property of a model of a financial market is that it allows for easy computation of the Greeks. The Greeks in the Black–Scholes–Merton model are very easy to calculate and this is one reason for the model's continued popularity in the market.

We now turn to exotic options, in particular the knock in/out barrier, lookback option and the Asian option.

## 5.3 Exotic Options

Sometimes the European call and put options, considered in Sec. 5.2, are called *vanilla* or even *plain vanilla* options. Their pay–off depends only on the final value of the

[14]The Greeks are used to evaluate the risk in the option, and constitute a crutial part of risk management theory in finance. Each Greek (with the exception of theta) represents a specific measure of risk associated with owning an option. An option portfolio can be adjusted accordingly (hedging) to achieve a desired exposure. For example delta, $\Delta = \frac{\partial C}{\partial S}$, measures the sensitivity to changes in the price of the underlying asset. The gamma measures the rate of change in the delta. The gamma, $\Gamma$, is the second derivative of the value function with respect to the underlying price, $\Gamma = \frac{\partial^2 C}{\partial S^2}$. Gamma is important because it indicates how a portfolio will react to relatively large shifts in price. The vega, $\nu$, which is not a Greek letter measures sensitivity to volatility. The vega is the derivative of the option value with respect to the volatility of the underlying, $\nu = \frac{\partial C}{\partial \sigma}$. The term kappa, $\kappa$, is sometimes used instead of vega, as is tau, $\tau$, though this is rare. The speed measures third order sensitivity to price. The speed is the third derivative of the value function with respect to the underlying price, $\frac{\partial^3 C}{\partial S^3}$. The theta measures sensitivity to the passage of time. The theta, $\Theta$, is the negative of the derivative of the option value with respect to the amount of time to expiry of the option, $\Theta = -\frac{\partial C}{\partial T}$. The rho, $\rho$, measures sensitivity to the applicable interest rate. The $\rho$ is the derivative of the option value with respect to the risk free rate, $\rho = \frac{\partial V}{\partial r}$.

| European Put Options | | | | | | |
|---|---|---|---|---|---|---|
| $t$ | $C(t,S(t))$ | $\Delta$ | $\Gamma$ | $\Theta$ | $\nu$ | $\rho$ |
| 0.100 | 3.558 | -0.462 | 0.042 | -16.533 | 12.490 | -4.971 |
| 0.200 | 4.879 | -0.444 | 0.029 | -10.851 | 17.487 | -9.860 |
| 0.300 | 5.824 | -0.431 | 0.024 | -8.298 | 21.204 | -14.663 |
| 0.400 | 6.571 | -0.419 | 0.020 | -6.758 | 24.241 | -19.377 |
| 0.500 | 7.191 | -0.408 | 0.018 | -5.698 | 26.832 | -24.004 |
| 0.600 | 7.720 | -0.399 | 0.016 | -4.909 | 29.100 | -28.544 |
| 0.700 | 8.179 | -0.390 | 0.015 | -4.292 | 31.118 | -32.997 |
| 0.800 | 8.582 | -0.381 | 0.014 | -3.792 | 32.935 | -37.364 |
| 0.900 | 8.940 | -0.373 | 0.013 | -3.377 | 34.585 | -41.646 |
| 1.000 | 9.260 | -0.366 | 0.012 | -3.025 | 36.093 | -45.843 |
| European Call Options | | | | | | |
| $t$ | $C(t,S(t))$ | $\Delta$ | $\Gamma$ | $\Theta$ | $\nu$ | $\rho$ |
| 0.100 | 3.955 | 0.532 | 0.042 | -20.469 | 12.490 | 4.929 |
| 0.200 | 5.667 | 0.544 | 0.029 | -14.724 | 17.487 | 9.744 |
| 0.300 | 6.996 | 0.552 | 0.024 | -12.109 | 21.204 | 14.451 |
| 0.400 | 8.121 | 0.558 | 0.020 | -10.508 | 24.241 | 19.054 |
| 0.500 | 9.113 | 0.562 | 0.018 | -9.387 | 26.832 | 23.557 |
| 0.600 | 10.007 | 0.566 | 0.016 | -8.539 | 29.100 | 27.962 |
| 0.700 | 10.826 | 0.569 | 0.015 | -7.863 | 31.118 | 32.271 |
| 0.800 | 11.584 | 0.572 | 0.014 | -7.305 | 32.935 | 36.485 |
| 0.900 | 12.290 | 0.574 | 0.013 | -6.832 | 34.585 | 40.608 |
| 1.000 | 12.952 | 0.576 | 0.012 | -6.422 | 36.093 | 44.640 |

**Table 5.1. The results for the Black–Scholes–Merton model for a European option.** Results for the Black–Scholes–Merton model for a European option when the strike price $K = 100.0$, the current price of the option is $S(0) = 0.0$, the interest rates $r = 0.1$, the volatility $\sigma = 0.3$ for the option price $S(t)$. Also shown in the table are the Greeks, the $\Delta$, $\Gamma$, $\Theta$, $\nu$ and $\rho$. This data was produced from a simulated Wiener process.

underlying asset. Options whose pay–off depends on the path of the underlying asset are called *path–dependent* or *exotic*.

In this section we discuss three different types of exotic options based on a geometric Brownian motion. This is only presented at an introductory level because, in Chapter 6, we only work with *path–independent* options for the European option case. There we aim to solve the path integral for the simplest cases first before considering more complicated cases, such as *path–dependent* options. Moreover as already mentioned, assets in real markets do not behave as per standard geometric Brownian motion—so in reality options cannot really be modeled in this way. Nevertheless these options can be approximately modeled using standard Brownian motion and therefore must be mentioned. More details about these options modeled by geometric Brownian motion can be found in (Shreve 2004).

Some of the *path–dependent* options considered here are the *barrier* options, *lookback* options and the *Asian* options. The first two options have explicit pricing formulas which are based on the reflections principle for Brownian motion. Such a formula for the Asian option is unknown. However using a change of *numéraire*[15] argument that reduces the partial differential equation to a simple form can be easily solved numerically (Logan 2004, Larsson and Thomée 2005). For more details the reader is invited to read the details in Shreve (2004), Karatzas and Shreve (1988), or in Øksendal (2003).

### 5.3.1 Knock In/Out Barrier Option

There are several types of barrier options, some "knock out" when the underlying asset price crosses a certain value called the barrier. There are four different possible scenarios in this case, two for the knock out and another two for the knock in option.

For the knock out call option we have

- *up–and–out* is when the underlying asset price begins below the barrier and crosses above. It causes to knocks out in other words it becomes worthless or expires.
- *down–and–out* is when the option has a barrier below the initial asset price and knocks out if the asset price falls below the barrier.

[15] A numéraire is the unit of account in which other assets are denominated, usually in the currency where the option has been issued. Changing the numéraire would consist of changing the currency to the one of another country. This concept is thoroughly covered in Shreve (2004), Karatzas and Shreve (1988), or in Øksendal (2003).

For knock in call option we have

- *up–and–in* option pays off zero unless they cross the barrier
- *down–and–in* option are zero unless they cross the barrier from above and goes below the barrier.

These scenarios can be represented graphically as in Fig. (5.2) In this section we treat

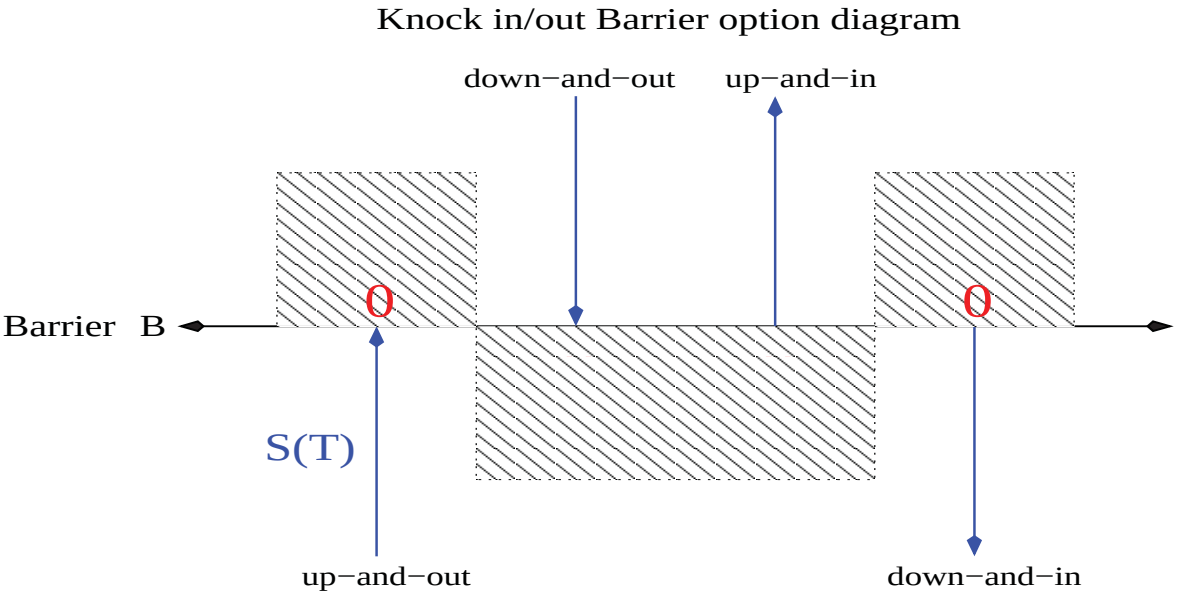


**Figure 5.2. Knock in/out Barrier option.** The graphical description of Knock in/out Barrier option. Here there are four different possible scenarios each of which are represented by the blue arrow. For example if we consider the far left arrow, then we see that the option will knock out once it passes a barrier $B$ from below, in other words when the underlying asset price begins below the barrier and crosses above, in that case it causes to knock out that is becoming worthless or expires. Similarly for the other scenarios.

only the up–and–out case on a geometric Brownian motion, but the methodology equally works well for up/down–and–in/out puts and calls.

### Up–and–Out Call

The underlying risky asset is a geometric Brownian motion for sake of the discussion[16] where $\widetilde{W}(t)$, $0 \leq t \leq T$, is a Brownian motion under the risk neutral measure $\widetilde{P}$. That is

$$dS(t) = rS(t)dt + \sigma S(t)d\widetilde{W}(t). \tag{5.41}$$

[16] the method works equally well for other underlying risky asset solution $S(t)$.

Consider a European call, expiring at time $T$, with strike price $K$ and up–and–out barrier $B$. We assume that $K < B$; otherwise the option must knock out in order to be in the money and hence could only payoff zero.

The stochastic differential equation Eq. (5.41) has the solution given by Eq. (4.170), that is

$$S(t) = S(0) \exp\left\{\sigma\widetilde{W}(t) + \left(r - \frac{1}{2}\sigma^2\right)t\right\}\Big|_{t_0=0} = S(0) \exp\left\{\sigma\widehat{W}(t)\right\}\Big|_{t_0=0} \tag{5.42}$$

where $\widehat{W}(t)$ is defined as Appendix C.2, Eq. (C.1), that is $\widehat{W}(t) = \alpha t + \widetilde{W}(t)$ and

$$\alpha = \frac{1}{\sigma}\left(r - \frac{1}{2}\sigma^2\right). \tag{5.43}$$

We defined $\widehat{M}(t) = \max_{0\le t\le T}\widehat{W}(t)$ as in Eq. (C.2), so

$$\max_{0\le t\le T} S(t) = S(0)e^{\sigma\widehat{M}(t)}. \tag{5.44}$$

The option knocks out if and only if $S(0)e^{\sigma\widehat{M}(t)} > B$; if $S(0)e^{\sigma\widehat{M}(t)} \le B$, the option pays off when

$$(S(T) - K)^+ = \left(S(0)e^{\sigma\widehat{W}(t)} - K\right)^+. \tag{5.45}$$

In other words, the pay off of the option is

$$\begin{aligned} V(T) &= \left(S(0)e^{\sigma\widehat{W}(t)} - K\right)^+ \mathcal{I}\Big|_{\left\{S(0)e^{\sigma\widehat{M}(t)}\le B\right\}} \\ &= \left(S(0)e^{\sigma\widehat{W}(t)} - K\right)\mathcal{I}\Big|_{\left\{S(0)e^{\sigma\widehat{W}(t)}\ge K; S(0)e^{\sigma\widehat{M}(t)}\le B\right\}} \\ &= \left(S(0)e^{\sigma\widehat{W}(t)} - K\right)\mathcal{I}\Big|_{\left\{\widehat{W}(t)\ge k; \widehat{M}(t)\le b\right\}}, \end{aligned} \tag{5.46}$$

where

$$k = \frac{1}{\sigma}\log\left(\frac{K}{S(0)}\right) \quad \text{and} \quad b = \frac{1}{\sigma}\log\left(\frac{B}{S(0)}\right). \tag{5.47}$$

This price of an up–and–out call satisfies a Black–Scholes–Merton equation that needs to be modified to account for the barrier. We are not going to discuss this model further in this section since it is a little out of the scope of the remaining chapters.

The next path–dependent option that we would like to discuss is the lookback option, which uses the same reflection principle (see Appendix C.2) as the knock in/out barrier option.

### 5.3.2 Lookback Option

An option whose payoff is based on the maximum that the underlying asset price attains over some interval of time prior to expiration date is called a *lookback* option.

Here we consider a *floating strike lookback* option. The payoff of this option is the difference between the maximum asset price over the time between initiation and expiration and the asset price at expiration.

Using the same geometric Brownian motion asset price as in Sec. 5.3.1, Eqs. (5.42,5.43) and Eq. (5.44) we may write the maximum of the asset price up to time $t$ as

$$Y(t) = \max_{0 \le u \le t} S(u) = S(0) e^{\sigma \widehat{M}(t)}. \tag{5.48}$$

The payoff for the lookback option considered here is,

$$V(T) = Y(T) - S(T) \ge 0 \tag{5.49}$$

at expiration time $T$. This payoff is non–negative because $Y(T) \ge S(T)$. Let $t \in [0, T]$ be given. At time $t$, the risk neutral price of the lookback option is

$$V(T) = \widetilde{E}\left[ e^{-r(T-t)} \left( Y(T) - S(T) \right) | \mathcal{F}_t \right]. \tag{5.50}$$

Because the pair of processes $(Y(T), S(T))$ has a Markov property there must exist a function $v(t, x, y)$ such that

$$V(t) = v(t, x, y). \tag{5.51}$$

This function is characterized by the option model in question. For example one can compute $v(t, x, y)$ explicitly in the case of the Black–Scholes–Merton model.

The Asian option is another path–dependent option, which takes the history of the option price into account.

### 5.3.3 Asian Option

An *Asian option* is one whose payoff includes a time average of the underlying asset price. The average is totally arbitrary and may be over the entire time period between initiation and expiration or may be over some period of time that begins later than the initiation of the option and also with the option expiration. The average may be from continuous sampling,

$$\frac{1}{T} \int_0^T S(t) dt, \tag{5.52}$$

or may be from discrete sampling

$$\frac{1}{m}\sum_{j=1}^{m} S(t_j) \quad \text{where} \quad 0 < t_1 < t_2 < \cdots < t_m = T. \tag{5.53}$$

The advantage of this is that it makes it difficult for anyone to significantly affect the payoff by manipulation of the underlying asset price.

The price of the Asian option is not known in closed form but can be evaluated by using a *change of numéraire* (see footnote 15), or by Monte Carlo simulations.

For example using the same geometric Brownian motion asset price as in Sec. 5.3.1, Eqs. (5.42,5.43) and Eq. (5.44) we can write the payoff functions for a *fixed strike Asian call* whose payoff at time $T$ is

$$V(T) = \left(\frac{1}{T}\int_0^T S(t)dt - K\right)^+. \tag{5.54}$$

Then the price at time $t$ prior to the expiration time $T$ of this call is given by the risk neutral formula

$$V(t) = \widetilde{E}\left[e^{-r(T-t)}V(T)|\mathcal{F}_t\right] \quad , \quad 0 \le t \le T. \tag{5.55}$$

The usual iterated conditioning argument shows that

$$e^{-rt}V(t) = \widetilde{E}\left[e^{-rT}V(T)|\mathcal{F}_t\right] \quad , \quad 0 \le t \le T, \tag{5.56}$$

is a martingale under the probability measure $\widetilde{P}$. This is the quantity we wish to compute. The possible avenues to this are done using methods of augmentation of state, that is $S(t)$ is augmented by defining a second process

$$Y(t) = \int_0^t S(u)du. \tag{5.57}$$

The other method consists of performing a change of numéraire and will not be discussed here any further.

## 5.4 American Option

European option contracts can only be exercised on the expiration date. An American option gives the owner the possibility to exercise the option at any time between the issue of the contract and the expiration date.

Because of this early exercise feature, such an option is at least as valuable as its European counterpart. Sometimes the difference in value is negligible or even zero and then American and European options are close or equivalent. At other times, the value of this early exercise feature, the so called *early exercise premium* can be substantial. An intermediate option between American and European is the Bermudan option (see Appendix C.1).

Because the American option can be exercised at any time prior to its expiration, it can never be worth less than the payoff associated with intermediate exercise. This is called the *intrinsic value* of the option. This is in contrast with the European option whose discounted price process is a martingale under the risk neutral measure. The discounted price process of an American option is a supermartingale (see Def. A.2.1 in Appendix A.2) under this measure. The holder of this option may fail to exercise at the optimal exercise date and in this case the discounted option price has a tendency to fall; hence the supermartingale property.

Because of the structure of the option, that is one can exercise at anytime, the option relies on the concept of stopping times (see Def. A.3.2), that is a stopping time $\tau$ has the property that the decision to stop at time $t$ must be based on the information available at time $t$.

The simplest interesting American option is the *perpetual American put option*. This is not a traded option but is a good starting point. The perpetual American put option is constructed on the same principle as Black–Scholes–Merton model, that is using the same geometric Brownian motion asset pricing formula.

The perpetual American put pays $K - S(t)$ if it is exercised at time $t$. This is its intrinsic value.

**Definition 5.4.1 (American perpetual price)** *Let $\tau$ be the set of all stopping times. The price of the perpetual American put is defined to be*

$$v_\star(x) = \max_{\tau \in T} \widetilde{\mathrm{E}} \left[ e^{-r\tau} \left( K - S(\tau) \right) \right], \tag{5.58}$$

*where $x = S(0)$ is the initial stock price. In the event that $\tau = \infty$ we interpret $e^{-r\tau}\left(K - S(\tau)\right)$ to be zero.*

The price of the option at time zero is the risk neutral expected payoff of the option, discounted from the exercise at any time, then every date is like every other date and

it is therefore reasonable to expect that the optimal exercise policy depends only on the value of $S(t)$ and not on the time variable $t$. The owner of the put option should exercise as soon as $S(t)$ falls far enough below $K$. Two questions come to mind from this. The first one is how far below the value $K$ should one exercise and how do we know it corresponds to the optimal exercise? The second question is what is the value of the put?

We are not going to attempt to answer these questions here as they are beyond the present scope and also because it is a different problem to the one we are concerned about, namely option pricing with path integral. This is is covered in Chapter 6. For a more complete discussion the reader may refer to Shreve (2004), van der Hoek and Elliot (2006), Dash (2004), and Levy (2004).

## 5.5 Chapter Summary

In this chapter we have reviewed and given specific examples of option. In particular we started by considering the European option which gives the right but not the obligation to buy or sell an asset at some time in the future. This right has a price, often called the premium. The European option is a building block of the Black–Scholes–Merton model, described in Sec. 5.2.1. Also described are other types of options such as exotic options and American options. Other options behave differently but are founded on the same principle as the European option. For example the American is really an European option with the property that it can be exercised at anytime as opposed to a fixed date in the future like in the European case. This small difference leads to a completely different problem and the mathematics is very different. This of course also applies to exotic options.

We use this background in the later chapters to evaluate the option price using path integrals. The idea behind the use of the path integral is that it provides an alternative approach to the stochastic calculus and aims at providing a more accurate method for evaluating the option price.

We begin next chapter with the simplest case where the option is path independent. In this category we have European options. We then apply the path integrals to different classes of stochastic differential equations. In each case we work with the European options.

Chapter 6

# Path integrals in finance

IN this chapter we describe the application of path integrals in a financial context. Here we will see how all of the ideas presented in Chapter 2 through Chapter 5 merge together when one is formulating the path integral in this context. We approach the discussion from a physics perspective, which is a formulation from a quantum mechanical perpective where there is a finite number of degrees of freedom. The similarities between finance and the physical theory then becomes apparent. Later in the chapter we discuss methods for evaluating the path integral, which represents the challenging part of the theory.

## 6.1 The Path Integral in Quantum Mechanics

In quantum mechanics the action is formulated in terms of the Hamiltonian instead of the Lagrangian, but the two can be related via the momentum operators. The Hamiltonian formulation is more general than the Lagrangian one.

The path integral in Hamiltonian theory is constructed using a time evolution operator, let us call it $\hat{U}(t_b, t_a)$ for now. This operator is performed on the transition probability between two space–time points.

For example if the Hamiltonian has the special form of $H \equiv H(\hat{p}, \hat{x})$, where $\hat{p}$ and $\hat{x}$ are time independent momentum and space operators, then one can get an expression for the Schröedinger equation (Schrodinger 1926b, Schrodinger 1926a, Griffiths 2004),

$$H(\hat{p}, \hat{x})|\psi(t)\rangle = i\hbar \frac{\partial}{\partial t}|\psi(t)\rangle. \tag{6.1}$$

The wave function $|\psi(t)\rangle$ is given by direct integration at any time $t_b$ from one state to any other at time $t_a$ and we obtain

$$|\psi(t)\rangle = e^{-i(t_b - t_a)\hat{H}/\hbar}|\psi(t)\rangle, \tag{6.2}$$

where the time evolution operator is given by

$$\hat{U}(t_a, t_b) = e^{-i(t_b - t_a)\hat{H}/\hbar}. \tag{6.3}$$

For that system the transition probability amplitude is given by

$$P\left(\psi(t_b), t_b|\psi(t_a), t_a\right) \equiv \langle\psi(t_b)|\psi(t_a)\rangle = \langle\psi(t_b)|\hat{U}(t_b, t_a)|\psi(t_a)\rangle. \tag{6.4}$$

These Hermitian operators ($\hat{U}^\dagger = \hat{U}^{-1}$) follow the unitarity property and have the property that

$$\hat{U}(t_b, t_a) = \hat{U}(t_b, t_c)\hat{U}(t_c, t_a) \quad \text{for} \quad t_c \in [t_a, t_b]. \tag{6.5}$$

Let us now consider any system with a wave-function $|x(t)\rangle$ for an arbitrary Hamiltonian $\hat{H} \equiv \hat{H}(\hat{p}, \hat{x})$. Then if we have a partition that is sliced into a large number, let us say $N + 1$, then for each time slice of equal width $\Delta t = t_n - t_{n-1} = \frac{t_b - t_a}{N+1} > 0$ we can write out the transition probability

$$\begin{aligned} P\left(x(t_b), t_b|x(t_a), t_a\right) &= \langle x(t_b)|\hat{U}(t_b, t_N)\hat{U}(t_N, t_{N-1}) \cdots \hat{U}(t_n, t_{n-1}) \\ &\cdots \quad \hat{U}(t_2, t_1)\hat{U}(t_1, t_a)|x(t_a)\rangle \end{aligned} \tag{6.6}$$

and after inserting a complete set of states

$$\int_{-\infty}^{\infty} dx(t_n)|x(t_n)\rangle\langle x(t_n)| = 1 \quad \text{for} \quad n = 1, \cdots, N, \tag{6.7}$$

the probability amplitudes become a product of $N$ integrals

$$\begin{aligned} P\left(x(t_b), t_b | x(t_a), t_a\right) &= \prod_{n=1}^{N}\left[\int_{-\infty}^{\infty} dx(t_n)\right]\prod_{n=1}^{N+1}\langle (x(t_n), t_n | x(t_{n-1}), t_{n-1}\rangle \\ &= \prod_{n=1}^{N}\left[\int_{-\infty}^{\infty} dx(t_n)\right]\prod_{n=1}^{N+1}\langle x(t_n)|e^{-i\Delta t\hat{H}/\hbar}|x(t_{n-1})\rangle, \end{aligned} \tag{6.8}$$

with the Hamiltonian $\hat{H} \equiv H(\hat{p}, \hat{x}, t)$. Now if we assume that the Hamiltonian is made up of two terms, that is the kinetic energy term and the potential term as

$$H(\hat{p}, \hat{x}, t) = \hat{T}(\hat{p}, t) + \hat{V}(\hat{x}, t), \tag{6.9}$$

then for sufficiently small time slices the evolution operator becomes

$$e^{-i\Delta t\hat{H}/\hbar} = e^{-i\Delta t(\hat{T}(\hat{p},t)+\hat{V}(\hat{x},t))/\hbar}. \tag{6.10}$$

This is factorisable according to the Baker–Campbell–Hausdorff formula (Sakurai 1994),

$$\hat{X} \equiv \frac{i}{2}[\hat{V}, \hat{T}] - \frac{\Delta t}{\hbar}\left(\frac{1}{6}[\hat{V}, [\hat{V}, \hat{T}]] - \frac{1}{3}[[\hat{V}, \hat{T}], \hat{V}]\right) + \mathcal{O}((\Delta t)^2). \tag{6.11}$$

If we neglect the $\mathcal{O}((\Delta t)^2)$ terms, we can calculate for the local matrix elements of $e^{-i\Delta t\hat{H}/\hbar}$ using the following simple expression,

$$\begin{aligned} \langle x(t_n)|e^{-i\Delta t\hat{H}/\hbar}|x(t_{n-1})\rangle &\approx \int_{-\infty}^{\infty} dx(t)\langle x(t_n)|e^{-i\Delta t\hat{T}(\hat{p}(t_n),t_n)/\hbar}|x(t_{n-1})\rangle \\ &\times \langle x(t_n)|e^{-i\Delta t\hat{V}(\hat{x}(t_n),t_n)/\hbar}|x(t_{n-1})\rangle \\ &= \int_{-\infty}^{\infty} dx(t)\langle x(t_n)|e^{-i\Delta t\hat{V}(\hat{x}(t_n),t_n)/\hbar}|x(t_{n-1})\rangle \\ &\times \int_{-\infty}^{\infty}\frac{dp(t_n)}{2\pi\hbar}e^{i\hat{p}(t_n)(x(t)-\hat{x}(t_{n-1}))/\hbar}e^{-i\Delta t\hat{T}(\hat{p}(t_n),t_n)/\hbar}, \end{aligned} \tag{6.12}$$

evaluating the local matrix elements

$$\langle x(t_n)|e^{-i\Delta t\hat{V}(\hat{x}(t_n),t_n)/\hbar}|x(t_{n-1})\rangle = \delta(x(t_n) - x(t))e^{-i\Delta tV(x(t_n),t_n)/\hbar} \tag{6.13}$$

this becomes

$$\begin{aligned} \langle x(t_n)|e^{-i\Delta t\hat{H}/\hbar}|x(t_{n-1})\rangle &\approx \int_{-\infty}^{\infty}\frac{dp(t_n)}{2\pi\hbar}\exp\left\{\frac{i}{\hbar}p(t_n)[x(t_n) - x(t_{n-1})]\right. \\ &- \left.\frac{i}{\hbar}\Delta t[T(\hat{p}(t_n), t_n) - V(x(t_n), t_n)]\right\}. \end{aligned} \tag{6.14}$$

Inserting Eq.(6.14) back into Eq.(6.8) we get

$$\begin{aligned} P\left(x(t_b),t_b|x(t_a),t_a\right) &= \prod_{n=1}^{N}\left[\int_{-\infty}^{\infty}dx(t_n)\right]\prod_{n=1}^{N+1}\left(x(t_n),t_n|x(t_{n-1}),t_{n-1}\right) \\ &= \prod_{n=1}^{N}\left[\int_{-\infty}^{\infty}dx(t_n)\right]\prod_{n=1}^{N+1}\left[\int_{-\infty}^{\infty}\frac{dp(t_n)}{2\pi\hbar}\right]\exp\left\{\frac{i}{\hbar}\mathcal{A}^N\right\} \end{aligned} \tag{6.15}$$

where

$$\begin{aligned} \mathcal{A}^N &= \sum_{n=1}^{N+1}\left\{p(t_n)[x(t_n)-x(t_{n-1})]-\Delta t[T(\hat{p}(t_n),t_n)-V(x(t_n),t_n)]\right\} \\ &= \sum_{n=1}^{N+1}\left\{p(t_n)[x(t_n)-x(t_{n-1})]-\Delta t H(\hat{p}(t_n),x(t_n),t_n)\right\}. \end{aligned} \tag{6.16}$$

such that in the limit as $N\to\infty$

$$\lim_{N\to\infty}\prod_{n=1}^{N}\left[\int_{-\infty}^{\infty}dx(t_n)\right]\prod_{n=1}^{N+1}\left[\int_{-\infty}^{\infty}\frac{dp(t_n)}{2\pi\hbar}\right]\equiv\int_{x(t_a)}^{x(t_b)}\mathcal{D}'x(t)\int\frac{\mathcal{D}p(t)}{2\pi\hbar}, \tag{6.17}$$

and the sum $\mathcal{A}^N$ tends to the integral

$$\begin{aligned} \mathcal{A}[p(t),x(t)] &= \int_{t_a}^{t_b}dt\left\{p(t)\dot{x}(t)-H(p(t),x(t),t)\right\} \\ &= \int_{t_a}^{t_b}dt\mathcal{L}(x(t),\dot{x}(t)) \end{aligned} \tag{6.18}$$

where $\mathcal{L}(x(t),\dot{x}(t))$ is the Lagrangian functional. Hence in the limit, with the functional integral $\mathcal{D}x(t)$, the amplitude

$$\begin{aligned} P\left(x(t_b),t_b|x(t_a),t_a\right) &\equiv \langle x(t_b)|\hat{U}(t_b,t_a)|x(t_a)\rangle \\ &= \int_{x(t_a)}^{x(t_b)}\mathcal{D}'x(t)\int\frac{\mathcal{D}p(t)}{2\pi\hbar}\exp\left\{\frac{i}{\hbar}\mathcal{A}[p(t),x(t)]\right\}, \end{aligned} \tag{6.19}$$

which has an intuitive interpretation that integrating over all paths is equivalent to summing over all histories in which a physical system can possibly evolve to. Here, Eq.(6.19) is called the *Feynman path integral* formula. For a more thorough discussion about path integrals in quantum mechanics see Kleinert (2004), Zinn-Justin (2005), Zinn-Justin (2002), Rivers (1987), Feynman (1972) and Roepstorff (1994), for example, for more details.

In the case where we have a standard kinetic energy $T=\frac{p^2}{2M}$ and a smooth potential, in the limit as $N\to\infty$, Eq.(6.15) converges to a standard quantum mechanical probability amplitude.

The above discussion sets out the basic idea of the path integral. Unfortunately this path integral cannot be used on every quantum mechanical system, because of the structure of the potential $V(x)$. This limitation certainly applies for atomic potentials, like the Coulomb potential $V(x) \to -\frac{1}{|x|}$, because the Feynman path integral diverges even for two time slices. Nevertheless we will use ideas similar to the one discussed above later in this work.

## 6.2 The Path Integral in Quantum Field Theory

The next step in the formulation of the path integral is carried out in quantum field theory. In this case we have an infinite number of degrees of freedom—in finance these theories can be used in interest rates models (Baaquie 2004).

The difference between the formulation of the path integral in quantum mechanics and quantum field theory (QFT) is that, in QFT, the particles are no longer considered as point like particles but as fields. This means that we are no longer integrating over a set of points but instead over a set of functions, where each function describes the state of the field.

In quantum field theory, that is in both Abelian and non–Abelian theories[17], we normally work directly with the Lagrangian density or functional. In these theories we are interested in calculating the expected value of fields or what we usually call correlation functions of the fields. In general one cannot evaluate the path integral directly in quantum field theory and one is obliged to resort to the use of an approximate method such as lattice gauge theory to calculate these correlations functions. Lattice gauge

[17] An Abelian theory is when any two matrices in a given group commute amongst themselves, in contrast to a non–Abelian theory where the matrices do not commute. In physics, in the first case, one may refer to quantum electrodynamics (QED), which deals with the interactions between the photons, electrons, muons, and other hadrons. In QED the photons do not interact with each other, which means that electromagnetic field strength tensor is just made up of the derivative of the fields. QED is the theory that describes electromagnetism, and has been to predict quantum phenomenon in particle physics very accurately. In the second case one may refer to quantum chromodynamics (QCD), which describes the interactions between quarks and gluons mainly. There the group representation no longer has commuting matrices in it. This means that the electromagnetic field strength tensor has an extra term in it that prevents the fields having a zero commutator. It is that extra term which says that the gluons, unlike the photons, interact with each other. Hence This theory is used to describes the interactions between the quarks and gluons inside the protons and neutrons. This difference makes a significant difference in what the two theories describe.

is a brute force method that has found itself to be very successful in describing non-perturbative effects in quantum chromodynamics (QCD).

One of the most interesting calculations in particle physics is to evaluate correlation functions for certain particles. These could be baryons or mesons. This is normally achieved by calculating $n$–point Green's functions. In quantum field theory one is talking about the propagator, when one is looking at $n$-point Green's functions. In other words these are correlation functions. These arise from Feynman diagrams, which arise in either perturbative or non–perturbative expansion from either Abelian or non-Abelian gauge theories. A propagator is a 2–point Green's function, which can be expressed in terms of a Feynman path integral where the fields arise inside the path integral. Physically what happens is that a field (i.e. a particle) is created at a point and annihilated at a later point in space–time, in other words the field is propagated in space–time. The path integral is an ensemble average where each possible path is weighted by a probability factor, which is proportional to the exponential of the action functional, which is itself composed of the Lagrangian functional. In effect, one could make the analogy that the propagator is the whole path integral and is a prescription on how the field has propagated along space–time and can be viewed as a transition probability.

In Section 6.3.1 we show how to extract the Lagrangian from an arbitrary stochastic differential equation, which is then used to evaluate the path integral hence the propagator. As an alternative approach to the path integral method one can use the drift and diffusion terms to construct the Fokker–Planck equation, which is a classical partial differential equation. The solution of this partial differential equation leads to transition probabilities for that model in question. Once one has extracted the transition probabilities there is no need to evaluate the path integral because these are essentially the same objects. This means that one either works in a partial differential framework or in a path integral framework. However once the transition probabilities have been extracted it is possible to use those to find the path, which contributes the most in the path integral. This is carried out using the Euler–Lagrange equation of motion, and is commonly known as the classical limit in quantum field theory. This is carried out in Section 6.4 and is called the instanton method. The solution of this partial differential equation can then be inserted back into the path integral and be evaluated.

## 6.3 The Path Integral in Finance

The path integral in quantum field theory has an infinite number of degrees of freedom, these can be used in interest rates models for example (Baaquie 2004).

In the introduction, Sec. 1.2.4, we showed how to arrive at the Feynman path integral from a quantum mechanical argument, that is from the Schrödinger equation.

In this section we start this discussion from the discounted Feynman–Kac formula defined in Theorem 4.5.5, that is Eq. (4.188). Theorem 4.5.5 states that for a general stochastic differential Eq. (4.168) and for a Borel–measurable function $h(y)$ the function

$$\begin{aligned} f(t,x) &= E^{t,x}\left[e^{-r(T-t)}h(X(T))\right] \\ &= \int dX(T)e^{-r(T-t)}h(X(T))p(X(T),T|x(t),t), \end{aligned} \tag{6.20}$$

satisfies the partial differential equation defined in, Eq. (4.189), where $h(X(T))$ is the terminal condition, that is the payoff function. Here the function $p(X(T),T|x(t),t)$ is the transition probability function. The transition probability is the most important part of the integrand because it describes how the system evolves in time. That is how the system transitions from one state at a given time to another state at a later or earlier time. These transition probability can be obtained from the Fokker–Planck equation, sec. 4.5.2 Eq. (4.196), or by considering the path integral. In quantum mechanics $p(X(T),T|x(t),t)$ is known as the propagator.

It can also be shown that Eq. (6.20) is the unique solution to the Cauchy problem (Shreve 2004, Karatzas and Shreve 1988). The Cauchy problem is the form that the Black–Scholes–Merton model, sec. 5.2.1, partial differential equation takes. The solution of that PDE gives the option price at time $t$ preceding expiration. Hence we will define the option price as in Eq. (5.27), that is

$$\mathcal{O}(S(t),t) = \mathrm{E}^{t,S(t)}\left[e^{-r\tau}\mathcal{F}(e^{X(T)})\right], \quad \tau = T-t. \tag{6.21}$$

Here $X(T) = \ln(S(T))$, and naturally $X(t) = \ln(S(t))$, is the solution of the stochastic differential equation, Eq. (4.168).

Now supposing that the transition probability is a continuous Markovian process, sec. A.4.1, then the transition probabilities satisfy the Chapman-Kolmogorov equation (CKE), sec. A.5, that is

$$P(X(T),T|X(t),t) = \int_{-\infty}^{\infty} P(X(T),T|x_2,t_2)P(x_2,t_2|X(t),t)dx_2, \tag{6.22}$$

then Feynman and Kac have shown that these probability distributions function have a path integral representation (Kac 1959, Feynman and Hibbs 1965, feller 1966).

### 6.3.1 The Correct Formulation of the Lagrangian Density

In this section we explicitly write down the Lagrangian density from an arbitrary stochastic differential equation. From Eq. (4.168), a general stochastic differential equation (SDE) can be written as

$$dX(t) = \alpha(t, X(t))\, dt + \beta(t, X(t))\, dW(t). \tag{6.23}$$

This equation represents a general stochastic differential equation. The random functions $\alpha(t, X(t))$ and $\beta(t, X(t))$ are called the drift term, and diffusion term respectively. The transition probability density function can be obtained using the Kolmogorov forward equation commonly known as the Fokker–Planck equation, Eq. (4.196). If we let $\mathcal{K}(y, T|x, t)$ denote the transition probability then the Fokker-Planck equation is given by

$$\begin{aligned} \frac{\partial}{\partial T}\mathcal{K}(y, T|x, t) &= \left[-\frac{\partial}{\partial y}\alpha(t, y) + \frac{1}{2}\frac{\partial^2}{\partial y^2}\beta^2(t, y)\right] \mathcal{K}(y, T|x, t) \\ &= \mathcal{G}(y, t)\mathcal{K}(y, T|x, t), \end{aligned} \tag{6.24}$$

where $\mathcal{G}(y, t)$ is the Fokker-Planck operator obtained directly from Eq. (6.23) above. A formal solution with initial value $\mathcal{K}(y, T|x, t) = \delta(y - x)$ can be derived using the Dyson series (Dyson 1949), see Appendix A.6, that is the general solution of Eq. (6.24) and may be found using the time ordered series, Eq. (A.14),

$$\begin{aligned} \mathcal{K}(y, T|x, t) &= \mathcal{T}\left[1 + \sum_1^\infty \frac{(1)^n}{n!}\int_t^t dt_1 \int_t^{t_1} dt_2 \cdots \int_t^{t_{n-1}} dt_n \mathcal{G}(y, t_1)\cdots\mathcal{G}(y, t_n)\right]\delta(y - x) \\ &= \mathcal{T}\exp\left[\int_t^T \mathcal{G}(y, \tau)d\tau\right]\delta(y - x). \end{aligned} \tag{6.25}$$

For a small time difference $\tau = T - t$, Eq. (6.25) reduces to

$$\mathcal{K}(y, T|x, t) = \left[1 + \mathcal{G}(y, t)\tau + \mathcal{O}(\tau^2)\right]\delta(y - x) = \exp\left[\mathcal{G}(y, t)\tau\right]\delta(y - x). \tag{6.26}$$

The transition probabilities are needed for the path integral solution. They are normally derived by repeatedly applying the Chapman–Kolmogorov equation, Eq. (6.27) (see Appendix A.5).

Now if we divide the time interval, $[t, T]$, into $N$ equal time steps $\Delta t$ bounded by $N-1$ equally spaced time points $t_i = t + i\Delta t$ for each $i = 0, 1, ..., N$ and the time increment is defined as $\Delta t = (T-t)/N$ then at each time step the value of $x$ is denoted by $x_i \equiv x(t_i)$ we can write down the transition probabilities as $\mathcal{K}(x_i, t_i | x_{i-1}, t_{i-1})$.

Consequently, we see how to obtain a definition of the path integral via a repeated use of the Chapman–Kolmogorov equation, that is,

$$\begin{aligned}\mathcal{K}(y, T|x, t) &= \lim_{N\longrightarrow\infty} \overbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}^{N-1} \\ &\times \mathcal{K}(y, T|x_{N-1}, t_{N-1}) \cdots \mathcal{K}(x_1, T_1|x, t)\, dx_1 \cdots dx_{N-1}, \end{aligned} \tag{6.27}$$

where $t_i = t + i\Delta t$, for each $i = 1, ..., N$. The constant $N$ being the number of $N$ equal steps.

In the limit where $N \to \infty$ we may use Eq. (6.25),

$$\begin{aligned}\mathcal{K}(y, T|x, t) &= \exp\left[\mathcal{G}(y, t)(T-t)\right] \delta(y-x) \\ &= \left[1 + \mathcal{G}(y, t)\tau + \mathcal{O}(\tau^2)\right] \delta(y-x) \\ &= \left[1 + \mathcal{G}(x, t)\tau + \mathcal{O}(\tau^2)\right] \delta(y-x) \end{aligned} \tag{6.28}$$

where for the last line in Eq. (6.28) we have used the delta function identity, $\delta(y-x)f(x) = \delta(y-x)f(y)$ and $\tau = T - t$. Here the Fokker-Plank operator is defined as

$$\mathcal{G}(y, t) = -\frac{\partial}{\partial y} D^{(1)}(y, t) + \frac{\partial^2}{\partial y^2} D^{(2)}(y, t), \tag{6.29}$$

where

$$D^{(1)}(y, t) = \alpha(t, y) \tag{6.30}$$

$$D^{(2)}(y, t) = \frac{1}{2}\beta^2(t, y). \tag{6.31}$$

If we now introduce the representation of the $\delta$ function in terms of Fourier integral

$$\begin{aligned}\mathcal{K}(y, T|x, t) &= \exp\left[-\frac{\partial}{\partial y} D^{(1)}(x, t)\tau + \frac{\partial^2}{\partial y^2} D^{(2)}(x, t)\tau\right] \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iu(y-x)} du \\ &= \left[1 - \frac{\partial}{\partial y} D^{(1)}(x, t)\tau + \frac{\partial^2}{\partial y^2} D^{(2)}(x, t)\tau + \mathcal{O}(\tau^2)\right] \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iu(y-x)} du \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left[-iuD^{(1)}(x, t)\tau - u^2 D^{(2)}(x, t)\tau + iu(y-x)\right] du \\ &= \frac{1}{\sqrt{4\pi D^{(2)}(x, t)\tau}} \exp\left[-\frac{\left[y - x - D^{(1)}(x, t)\tau\right]^2}{4\, D^{(2)}(x, t)\tau}\right]. \end{aligned} \tag{6.32}$$

Inserting this equation into the Chapman–Kolmogorov, Eq. (6.27), and denoting $x_i \equiv x(t_i)$, we obtain

$$\begin{aligned}\mathcal{K}(x_T,T|x,t) &= \lim_{N\longrightarrow\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\\ &\times\ \mathcal{K}(x_T,T|x_{N-1},t_{N-1})\cdots\mathcal{K}(x_1,T_1|x,t)\,dx_1\cdots dx_{N-1} \qquad (6.33)\\ &= \lim_{N\longrightarrow\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\prod_{i=1}^{N-1}\frac{dx_i}{\sqrt{4\pi D^{(2)}(x_i,t_i)\Delta t}}\\ &\times\ \exp\left[-\sum_{i=0}^{N-1}\frac{\left[x_{i+1}-x_i-D^{(1)}(x_i,t_i)\Delta t\right]^2}{4\,D^{(2)}(x_i,t_i)\Delta t}\right].\end{aligned}$$

Now if we discritize the derivatives so that,

$$\frac{x_{i+1}-x_i}{\Delta t}=\dot{x}(t_i)=\frac{d}{dt_i}x(t_i) \qquad (6.34)$$

that is write it as, $x_{i+1}-x_i=\dot{x}(t_i)\Delta t$, then we can rewrite the sum inside the exponential term as

$$-\sum_{i=0}^{N-1}\frac{\left[x_{i+1}-x_i-D^{(1)}(x_i,t_i)\Delta t\right]^2}{4\,D^{(2)}(x_i,t_i)\Delta t}=-\sum_{i=0}^{N-1}\frac{\left[\dot{x}(t_i)-D^{(1)}(x_i,t_i)\right]^2}{4\,D^{(2)}(x_i,t_i)}\Delta t. \qquad (6.35)$$

Furthermore by transforming the sum into an integral, the sum can be rewritten as an integral

$$-\sum_{i=0}^{N-1}\frac{\left[\dot{x}(t_i)-D^{(1)}(x_i,t_i)\right]^2}{4\,D^{(2)}(x_i,t_i)}\Delta t\rightarrow-\int_t^T\frac{\left[\dot{x}(t')-D^{(1)}(x(t'),t')\right]^2}{4\,D^{(2)}(x(t'),t')}dt', \qquad (6.36)$$

which represents the action functional,

$$\mathcal{A}[x(t)]=\int_t^T\frac{\left[\dot{x}(t')-D^{(1)}(x(t'),t')\right]^2}{4\,D^{(2)}(x(t'),t')}dt'=\int_t^T\mathcal{L}(\dot{x}(t'),x(t'),t')\,dt', \qquad (6.37)$$

where

$$\mathcal{L}(\dot{x}(t'),x(t'),t')=\frac{\left[\dot{x}(t')-D^{(1)}(x(t'),t')\right]^2}{4\,D^{(2)}(x(t'),t')} \qquad (6.38)$$

is the Lagrangian functional. Finally

$$\prod_{i=1}^{N-1}\frac{dx_i}{\sqrt{4\pi D^{(2)}(x_i,t_i)\Delta t}}\equiv\mathcal{D}x[t], \qquad (6.39)$$

is the integral measure for the path integral. In compact form the path integral may be written as

$$\mathcal{K}(x_T, T|x, t) = \int_{x(t)}^{x(T)} \mathcal{D}x[t']\, e^{-\mathcal{A}[x(t')]}. \tag{6.40}$$

In this section we have written down a general form for the path integral that can carry the information from any stochastic differential equation into the path integral. Now as already mentioned the option price is given by the discounted Feynman–Kac formula Eq. (6.20) and Eq. (6.21), hence we can write the path integral representation for the option price as

$$\begin{aligned} \mathcal{O}(S(t), t) &= \mathrm{E}^{t,S(t)}\left[e^{-r\tau}\mathcal{F}[e^{x(T)}]\right], \quad \tau = T - t \\ &= \int_{-\infty}^{\infty} dx(T) \left( \int_{x(t)}^{x(T)} \mathcal{D}x[t'] e^{-r\tau}\mathcal{F}[e^{x(T)}] e^{-\mathcal{A}[x(t')]} \right). \end{aligned} \tag{6.41}$$

This definition of the path integral is used in quantum mechanics as we saw earlier. It can be shown rigorously that this limit converges (Kac 1959, Kac 1949, Kac 1980, Glimm and Jaffe 1981, Simon 1979, MFreidlin 1985). The form of Eq. (6.32) is not unique as the differential operator did not act on Eqs.(6.30) and (6.31). Alternatively forcing the derivatives to act on the $D^{(1)}(y,t)$ and $D^{(2)}(y,t)$ factors leads to the following Fokker–Planck operator

$$\begin{aligned} \mathcal{G}(y,t) &= -\frac{\partial}{\partial y} D^{(1)}(y,t) + \frac{\partial^2}{\partial y^2} D^{(2)}(y,t) \\ &- \left[ D^{(1)}(y,t) - 2\frac{\partial}{\partial y} D^{(2)}(y,t) \right] \frac{\partial}{\partial y} + \frac{\partial^2}{\partial y^2} D^{(2)}(y,t). \end{aligned} \tag{6.42}$$

Inserting Eq. (6.42) into Eq. (6.25) and replacing $\frac{\partial}{\partial y}$ by $-\frac{\partial}{\partial x}$ we can perform the same steps as before we obtain leading for small time $\tau$ to

$$\begin{aligned} \mathcal{K}(y, T|x, t) &= \frac{1}{\sqrt{4\pi D^{(2)}(y,t)\tau}} \exp\left[ -\frac{\partial D^{(1)}(y,t)}{\partial y}\tau + \frac{\partial^2 D^{(2)}(y,t)}{\partial y^2}\tau \right. \\ &- \left. \frac{\left[ y - x - \left( D^{(1)}(y,t) - 2\frac{\partial D^{(2)}(y,t)}{\partial y} \right)\tau \right]^2}{4\, D^{(2)}(y,t)\tau} \right]. \end{aligned} \tag{6.43}$$

This leads to the following Lagrangian

$$\begin{aligned} \mathcal{L}(\dot{x}(t'), x(t'), t') &= -\frac{\partial D^{(1)}(x(t'),t')}{\partial x}\tau + \frac{\partial^2 D^{(2)}(x(t'),t')}{\partial x^2}\tau \\ &- \frac{\left[ \dot{x}(t') - \left( D^{(1)}(x(t'),t') - 2\frac{\partial D^{(2)}(x(t'),t')}{\partial x} \right)\tau \right]^2}{4\, D^{(2)}(x(t'),t')\tau}. \end{aligned} \tag{6.44}$$

These section shows how to extract the Lagrangian functional directly from the stochastic differential equation. We will use these definitions in the following sections.

### 6.3.2 Standard Gaussian Path Integrals

We now apply what we have learnt in sec. 6.3.1 to the Black–Scholes–Merton model. We saw in Eq. (5.6), with interest rates $r$, that a stock was modelled by the geometric Brownian motion

$$dS(t) = \alpha S(t)dt + \sigma S(t)dW(t), \tag{6.45}$$

for which the general solution was given by Eq. (6.46), that is

$$S(t) = S(0)\exp\left\{\sigma W(t) + \left(\alpha - \frac{1}{2}\sigma^2\right)t\right\}\Big|_{t_0=0}, \tag{6.46}$$

Here $\alpha$ is the drift rate and $\sigma$ the volatility. Let us make a change of variable by introducing a new variable $x(t) = \ln(S(t))$, hence using the Îto–Doeblin equation, Eq. (4.25), in Theorem. 4.1.7, with $f = \ln(S(t))$, we obtain

$$\begin{aligned} dx(t) &= f_t + f_S dS(t) + \frac{1}{2}f_{SS}dS(t)dS(t) \\ &= f_t + f_S\left\{\alpha S(t)dt + \sigma S(t)dW(t)\right\} + \frac{1}{2}f_{SS}\sigma^2 S^2(t)dt. \end{aligned} \tag{6.47}$$

Here $f_t = 0$, $f_S = 1/S(t)$ and $f_{SS} = -1/S^2(t)$. Using the multiplication table in Eq. (3.1), it leads to the following stochastic differential equation

$$dx(t) = \left(\alpha - \frac{1}{2}\sigma^2\right)dt + \sigma dW(t), \tag{6.48}$$

with constant drift rate $\alpha$ and volatility $\sigma$ and no dividends. Using Eq. (6.48) we can obtain the Black–Scholes–Merton PDE, Eq. (5.22). This was derived in sec. 5.2.1.

To calculate the option price we use the discounted Feynman–Kac formula defined in Eq. (6.20) in general and Eq. (6.21) for the Black–Scholes–Merton model, i.e.

$$\begin{aligned} \mathcal{O}(S(t),t) &= \mathrm{E}^{t,x}\left[e^{-r\tau}\mathcal{F}[e^{x(T)}]\right], \quad \tau = T - t. \\ &= \int_{-\infty}^{\infty} dx(T)\left(\int_{x(t)}^{x(T)} \mathcal{D}x[t']e^{-r\tau}\mathcal{F}[e^{x(T)}]e^{-\mathcal{A}[x(t')]}\right). \end{aligned} \tag{6.49}$$

Here $\mathcal{F}[e^{x(T)}]$ is the payoff function, which will depend on the type of the option.

Using the Lagrangian functional representation derived in sec. 6.3.1, Eq. (6.38), directly from Eq. (6.48). Using the definition for the SDE coeficients $D^{(1)}(x,t)$ and $D^{(2)}(x,t)$, Eqs.(6.30) and Eq. (6.31) respectively,

$$D^{(1)}(x,t) = \alpha - \frac{1}{2}\sigma^2, \tag{6.50}$$

$$D^{(2)}(x,t) = \frac{1}{2}\sigma^2. \tag{6.51}$$

Here because $D^{(1)}(x,t)$ and $D^{(2)}(x,t)$ are both constants the Lagrangian Eq. (6.38) and Eq. (6.44) would lead to the same result, which is given by

$$\begin{aligned}\mathcal{L}(\dot{x}(t'),x(t'),t') &= \frac{\left[\dot{x}(t') - D^{(1)}(x(t'),t')\right]^2}{4\,D^{(2)}(x(t'),t')} \\ &= \frac{1}{2\sigma^2}\left(\dot{x}(t') - \mu\right)^2 \quad \text{where} \quad \mu = \alpha - \frac{1}{2}\sigma^2.\end{aligned} \tag{6.52}$$

The action functional becomes

$$\begin{aligned}\mathcal{A}[x(t')] &= \int_t^T \mathcal{L}(\dot{x}(t'),x(t'),t')\,dt' \\ &= \int_t^T \frac{1}{2\sigma^2}\left(\dot{x}(t') - \mu\right)^2 dt' \\ &= \int_t^T \frac{1}{2\sigma^2}\left(\dot{x}^2(t') - 2\mu\dot{x}(t') + \mu^2\right) dt' \\ &= \frac{\mu^2}{2\sigma^2}\tau - \frac{\mu}{\sigma^2}\left(x(T) - x(t)\right) + \frac{1}{2\sigma^2}\int_t^T \left(\dot{x}(t')\right)^2 dt' \\ &= \frac{\mu^2}{2\sigma^2}\tau - \frac{\mu}{\sigma^2}\left(x(T) - x(t)\right) + \mathcal{A}_0[x(t')],\end{aligned} \tag{6.53}$$

where $\mathcal{A}_0[x(t')]$ is the action functional for a zero drift process, $dx(t) = \sigma dW(t)$, which is a martingale. So we have

$$\mathcal{A}_0[x(t')] = \frac{1}{2\sigma^2}\int_t^T \left(\dot{x}(t')\right)^2 dt'. \tag{6.54}$$

Applying the discretization method onto the time interval, $[t,T]$, by dividing it into $N$ equal time steps $\Delta t$ bounded by $N-1$ equally spaced time points $t_i = t + i\Delta t$ for each $i = 0,1,...,N$ and the time increment is defined as $\Delta t = (T-t)/N$ then at each time step the value of $x$ is denoted by $x_i \equiv x(t_i)$. Furthermore if we transform the continuous derivatives into discrete ones and similarly the integrals by discrete sums, that is substituting

$$\int_t^T \cdots dt' \longrightarrow \sum_{i=0}^{N-1} \cdots \Delta t, \quad \text{and} \tag{6.55}$$

$$\dot{x}(t') \longrightarrow \frac{x_{i+1}(t') - x_i(t')}{\Delta t}, \tag{6.56}$$

we can write down the action functional in its discrete form, that is

$$\mathcal{A}[x(t_i)] = \frac{\mu^2}{2\sigma^2}\tau - \frac{\mu}{\sigma^2}\left(x(T) - x(t)\right) + \frac{1}{2\sigma^2\Delta t}\sum_{i=0}^{N-1}\left(x_{i+1}(t') - x_i(t')\right)^2. \tag{6.57}$$

The integral measure is given by Eq. (6.39), that is

$$\prod_{i=1}^{N-1}\frac{dx_i}{\sqrt{2\pi\sigma^2\Delta t}} \equiv \mathcal{D}x[t'], \tag{6.58}$$

where $\Delta t = \tau/N$.

Now the path integral over all paths from the initial state $x(t)$ to the final state $x(T)$ is given by

$$\begin{aligned}\int_{x(t)}^{x(T)}\mathcal{D}x[t']e^{-r\tau}\mathcal{F}[e^{x(T)}]e^{-\mathcal{A}[x(t')]} \\ = \lim_{N\longrightarrow\infty}\overbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}^{N-1}\left(\prod_{i=1}^{N-1}\frac{dx_i}{\sqrt{2\pi\sigma^2\Delta t}}\right)\mathcal{F}[e^{x(T)}]e^{-\mathcal{A}[x(t_i)]},\end{aligned} \tag{6.59}$$

hence the option price, becomes

$$\begin{aligned}\mathcal{O}(S(t),t) &= \int_{-\infty}^{\infty}dx(T)\left(\int_{x(t)}^{x(T)}\mathcal{D}x[t']e^{-r\tau}\mathcal{F}[e^{x(T)}]e^{-\mathcal{A}[x(t')]}\right) \\ &= \int_{-\infty}^{\infty}dx(T)e^{-r\tau}\mathcal{F}[e^{x(T)}]\exp\left[\frac{\mu^2}{2\sigma^2}\tau - \frac{\mu}{\sigma^2}\left(x(T) - x(t)\right)\right]\mathcal{K}(x,T|x,t).\end{aligned} \tag{6.60}$$

Where the path integral is now just

$$\begin{aligned}\mathcal{K}(x,T|x,t) &= \int_{x(t)}^{x(T)}\mathcal{D}x[t']e^{-\mathcal{A}_0[x(t')]} \\ &= \lim_{N\longrightarrow\infty}\overbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}^{N-1}\exp\left[-\frac{1}{2\sigma^2\Delta t}\sum_{i=0}^{N-1}\left(x_{i+1}(t') - x_i(t')\right)^2\right]\left(\prod_{i=1}^{N-1}\frac{dx_i}{\sqrt{2\pi\sigma^2\Delta t}}\right).\end{aligned} \tag{6.61}$$

This definition has been shown to converge (Kac 1959, Kac 1949, Kac 1980, Glimm and Jaffe 1981, Simon 1979, MFreidlin 1985). The payoff $\mathcal{F}[e^{x(T)}]$ depends only on the terminal state $x(T)$ that is why we were able to take it outside the path integral in Eq. (6.60). Furthermore, Eq. (6.61) is a Gaussian integral; this means that we can apply the Gaussian identity, Eq. (A.19). This readily simplifies the evaluation of the path

integral. To see this we perform the first two integrals, that is

$$\frac{1}{2\pi\sigma^2\Delta t}\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2\sigma^2\Delta t}\left((x_2-x_1)^2+(x_1-x_0)^2\right)\right]dx_1 \tag{6.62}$$

$$= \frac{1}{2\pi\sigma^2\Delta t}\sqrt{\frac{\pi}{\frac{1}{2\sigma^2\Delta t}+\frac{1}{2\sigma^2\Delta t}}}\exp\left[-\frac{1}{2\sigma^2 2\Delta t}(x_2-x_0)^2\right]$$

$$= \frac{1}{\sqrt{2\pi\sigma^2(2\Delta t)}}\exp\left[-\frac{1}{2\sigma^2(2\Delta t)}(x_2-x_0)^2\right]. \tag{6.63}$$

Repeating this process for $dx(t_2), dx(t_3)\cdots$ we see that

$$\mathcal{K}(x,T|x,t) = \frac{1}{\sqrt{2\pi\sigma^2\tau}}\exp\left[-\frac{(x(T)-x(t))^2}{2\sigma^2\tau}\right]. \tag{6.64}$$

Having evaluated the path integral we can now insert Eq. (6.64) into Eq. (6.60), that is

$$\begin{aligned}
\mathcal{O}(S(t),t) &= \int_{-\infty}^{\infty}dx(T)\left(\int_{x(t)}^{x(T)}\mathcal{D}x[t']e^{-r\tau}\mathcal{F}[e^{x(T)}]e^{-\mathcal{A}[x(t')]}\right)\\
&= \int_{-\infty}^{\infty}dx(T)e^{-r\tau}\mathcal{F}[e^{x(T)}]\exp\left[\frac{\mu^2}{2\sigma^2}\tau-\frac{\mu}{\sigma^2}\left(x(T)-x(t)\right)\right]\\
&\times \frac{1}{\sqrt{2\pi\sigma^2\tau}}\exp\left[-\frac{(x(T)-x(t))^2}{2\sigma^2\tau}\right].\\
&= \frac{1}{\sqrt{2\pi\sigma^2\tau}}\int_{-\infty}^{\infty}dx(T)e^{-r\tau}\mathcal{F}[e^{x(T)}]\\
&\times \exp\left[-\frac{1}{2\sigma^2\tau}\left(\mu^2\tau^2-2\mu\tau\left(x(T)-x(t)\right)+(x(T)-x(t))^2\right)\right]\\
&= \frac{1}{\sqrt{2\pi\sigma^2\tau}}\int_{-\infty}^{\infty}dx(T)e^{-r\tau}\mathcal{F}[e^{x(T)}]\exp\left[-\frac{1}{2\sigma^2\tau}\left(x(T)-x(t)-\mu\tau\right)\right)^2\Bigg].
\end{aligned}$$

So we arrive at the following expression for the option price,

$$\mathcal{O}(S(t),t) = \frac{1}{\sqrt{2\pi\sigma^2\tau}}\int_{-\infty}^{\infty}dx(T)e^{-r\tau}\mathcal{F}[e^{x(T)}]\exp\left[-\frac{\left(x(T)-x(t)-\mu\tau\right)\right)^2}{2\sigma^2\tau}\Bigg]. \tag{6.65}$$

This is the same as Eq. (5.31) in the case of a European call option obtained in Sec. 5.2.1, which leads to the solution Eq. (5.38).

So we see that using path integrals we are able to arrive at the same result as when we use standard stochastic calculus. We note that when using stochastic calculus, to evaluate the option price, we are required to know the explicit solution of the stochastic differential Eq. (5.7). Here using the path integral approach we do not need to know the explicit solution of the stochastic differential equation—we are able to arrive at an explicit solution for the option price by just extracting the information directly from the

stochastic differential equation. This enables us to write down directly the Lagrangian functional and evaluate the path integral.

This is very convenient, because in many situations it is not possible to solve the stochastic differential equation directly—which means that the path integral can be used as an alternative approach. On the other hand, path integrals are themselves a very complicated mathematical objects where very often it is not possible to arrive at a solution either.

In the next section we consider a non–Gaussian model to see if it is possible to arrive at a solution using a more realistic model.

### 6.3.3 Non–Standard Gaussian Path Integrals

In this section we attempt to solve the path integral for a non-Gaussian model. The general approach here is to start by deriving the general formula for the Lagrangian (recall that the Lagrangian is the functional that defines the action functional and hence the path integral). We then try to solve the path integral by setting some of the terms, such as the drift term, to zero to obtain a simpler version of the path integral. We then see if we can successfully increase the complexity progressively until we arrive at a full solution.

It is argued that heavy non–Gaussian tails and finite hedging time make it necessary to formulate a model outside the notion of risk–free option prices (Bouchaud *et al.* 1996, Bouchaud *et al.* 2002).

As opposed to other models where the standard Black–Scholes–Merton price model is extended to account for more exotic effects, such as jump diffusion models (Merton 1976b) or Lévy noise (Hull 2000a), here we use a model developed for stock return fluctuations (Borland 2002c, Borland and Bouchaud 2004). A closed form solution for European options was successfully derived therein. This approach is based on a class of stochastic process that allows statistical feedback as a model of the underlying stock returns. In there it was also shown that the distributions of returns implied by these processes closely matched those found empirically.

In particular, they display features such as fat–tails and peaked middles that are not at all captured by the standard class of log–normal distributions. Such stochastic processes were recently introduced within a Tsallis framework (Borland 1998b). This

framework is used in statistical physics, namely within the field of Tsallis nonextensive thermostatistics (Tsallis 1988b, Curado and Tsallis 1991b).

In this setting, we assume that the log return for the stock price $S(t)$ defined as

$$Y(t) \equiv \ln\left[\frac{S(t+\delta t)}{S(\delta t)}\right], \tag{6.66}$$

where $\delta t$ is simply an increment constant value. This follows the process

$$dY = \mu\, dt + \sigma\, d\Omega, \tag{6.67}$$

across timescale $t$, where $\sigma$ represents the volatility and $\mu$ the rate of return. The driving noise is now modeled by $\Omega$, which is drawn from a non–Gaussian distribution.

To do this it is assumed that $\Omega$ follows the statistical feedback process (Borland 1998b),

$$d\Omega = P(\Omega)^{\frac{1-q}{2}}\, dW. \tag{6.68}$$

These stochastic processes can be interpreted if the driving noise follows a generalized Wiener process governed by a fat–tailed Tsallis distribution (Tsallis 1988b, Curado and Tsallis 1991b) of index $q > 1$. Hence, using Eq. (6.68), we may rewrite the stochastic process as

$$dY = \mu\, dt + \sigma\, P(\Omega)^{\frac{1-q}{2}}\, dW. \tag{6.69}$$

Here $W$ is a Gaussian distributed noise process. For $q = 1$, $\Omega$ reduces to $W$ and the standard model is recovered. It is true that in the case of $q = 1$ the distribution is a Gaussian, and that if $5/3 < q < 3$ the attractor becomes a Lévy distribution Tsallis *et al.* (2003). The probability distribution of the variable $\Omega$ evolves according to the non–linear Fokker–Planck equation (Borland 1998b)

$$\frac{\partial}{\partial t}P(\Omega, t|\Omega', t') = \frac{1}{2}\frac{\partial^2}{\partial \Omega^2}P^{2-q}(\Omega, t|\Omega', t'), \tag{6.70}$$

with $P$ given by the Tsallis distribution and defined as (Borland 2002c, Borland and Bouchaud 2004, Borland 1998b)

$$P(\Omega, t|\Omega', t') = \frac{1}{Z(t)}\left[1 - \beta(t)(1-q)(\Omega - \Omega')^2\right]^{\frac{1}{1-q}}. \tag{6.71}$$

In Fig. 6.1 we show the graphs of Eq. (6.71) at a fixed $q$, i.e. at $q = 1.43$, on a fixed interval for $\Omega \in [-1, 1]$. In Fig. 6.1 (left graph) we can see how the distribution has a slow decaying affect on the tail. These kind of effects are those that are normally

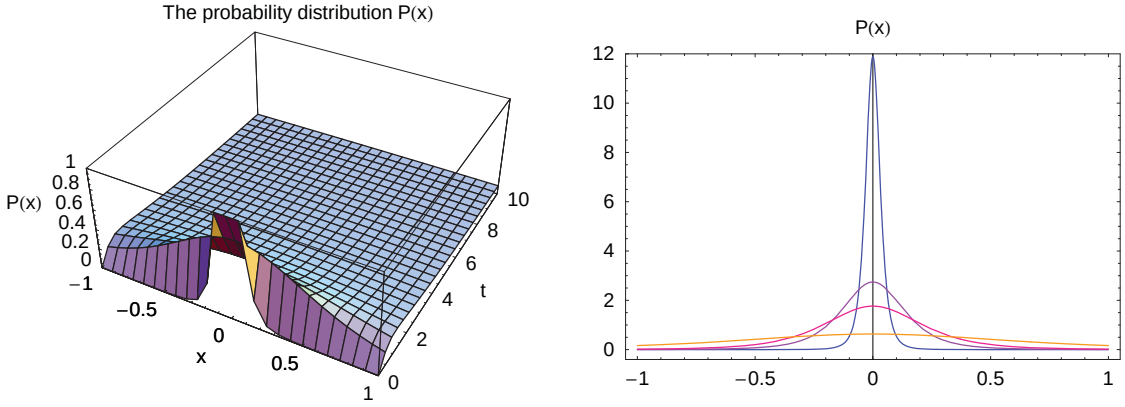


**Figure 6.1. The graph of the probability distribution function Eq. (6.71) when $q = 1.43$.** The graph of the probability distribution function Eq. (6.71) when $q = 1.43$ for the time range of $t \in [0.01, 10]$ and $-1 \leq \Omega \leq 1$ (right graph). We can see that as $t \to 0$ the distribution becomes sharply peaked which diverges to infinity. A two dimensional representation of Fig. 6.1 for the probability distribution function $P(\Omega)$ versus $\Omega$ for $\Omega \in [-1, 1]$, Eq. (6.71), when $q = 1.43$ at fixed time values. The most sharply peaked curve is when $t = 0.01$, the one in the middle is at $t = 0.05$ and the flattest one is when $t = 0.1$.

observed in real markets. On the opposite side of the scale we may observe that as $t \to 0$ the distribution becomes more and more sharply peaked, see Figure 6.1 (right graph). This seems to fit quite well the short interval intra–day empirical data from commonly know indexes, like S&P500, Dow Jones and CAC40 for example.

The time dependent $\beta(t)$ function is defined as,

$$\beta(t) = c^{\frac{1-q}{3-q}} \left[(2-q)(3-q)(t-t')\right]^{\frac{-2}{3-q}}, \tag{6.72}$$

and the normalization factor $Z(t)$ is defined as

$$Z(t) = \left[(2-q)(3-q)c(t-t')\right]^{\frac{1}{3-q}}, \tag{6.73}$$

and are both plotted in Fig. 7.9—on the left graph we show $\beta(t)$ as a function of the parameter $1.3 \leq q \leq 2.5$ and the time evolution parameter $t$ and on the right graph $Z(t)$ also as a function of the parameter $1 < q \leq 2.5$ and the time evolution parameter $t$.

In these graphs we can see that in range of $1 < q \leq 2$ the function is well behaved and that there are no singularities. The singularities arise from the Gamma function in the $c$ coefficient, given in Eq. (6.75).

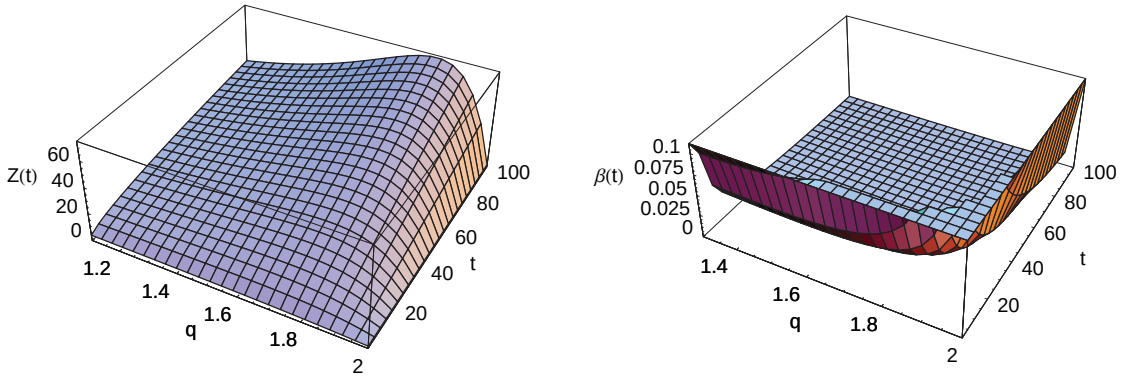


**Figure 6.2. The graph of $Z(t)$ and $\beta(t)$.** The graph of $\beta(t)$ (right graph) as a function of the parameter $1.3 \leq q \leq 2.5$ and the time evolution parameter $t$. The graph of $Z(t)$ (left graph) as a function of the parameter $1 < q \leq 2.5$ and the time evolution parameter $t$.

The $q$–dependent constant $c$ is given by

$$c = \beta(t)Z^2(t), \tag{6.74}$$

and is found to be given by the following

$$c \equiv \frac{\pi}{q-1} \frac{\Gamma^2\left(\frac{1}{q-1} - \frac{1}{2}\right)}{\Gamma^2\left(\frac{1}{q-1}\right)}, \tag{6.75}$$

where $\Gamma$ is the usual gamma function. In Fig. (6.3) we can see that for $q < 1$ the gamma function makes the $c$ function highly divergent. The coefficient $c$ is well behaved from in the range of $1 < c \leq 2.5$. For $q \geq 2.5$ the function diverges to a sharp peak.

In that process $\Omega(t)$ follows a Tsallis distribution. The transition probability function for the this process was derived by Borland (2002a) (see the references therein for further details). This is given by Eq. (6.71), that is

$$P(\Omega, t|\Omega', t') = \frac{1}{Z(t)} \left[1 - \beta(t)(1-q)(\Omega - \Omega')^2\right]^{\frac{1}{1-q}}. \tag{6.76}$$

We may insert the expression for the log return, Eq. (6.66), to obtain an expression for $\Omega(t)$ in terms of the asset price $S(t)$, in other words using Eq. (6.67) we observe that

$$\Omega(t) = \frac{1}{\sigma}\left[\ln\left(\frac{S(t+\delta t)}{S(\delta t)}\right) - \mu t\right]. \tag{6.77}$$

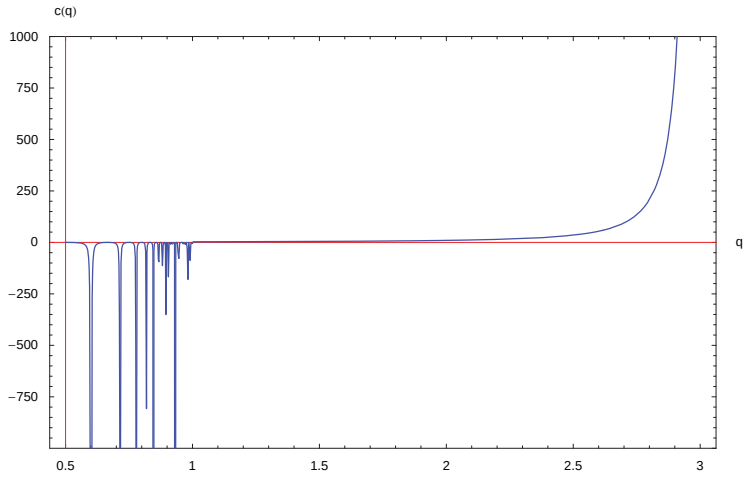


**Figure 6.3. The evolution of the coefficient $c$ as a function of the parameter $q$.** The evolution of the coefficient $c$ as a function of the parameter $q$. We can see that in the region $1 < q \leq 3$ the function is well behaved. However for $q > 2.75$ we can clearly see that the function will start to diverge.

Insering Eq. (6.77) into the transition probabilities, Eq. (6.76) with $\Omega' = 0$, we find that

$$\begin{aligned} P(\ln(S(t+\delta t))|\ln(S(t))) &= \frac{1}{Z(t)}\left[1-\beta(t)(1-q)\frac{1}{\sigma^2}\left[\ln\left(\frac{S(t+\delta t)}{S(\delta t)}\right)-\mu t\right]^2\right]^{\frac{1}{1-q}} \\ &= \frac{1}{Z(t)}\left[1-\tilde{\beta}(t)(1-q)\left[\ln\left(\frac{S(t+\delta t)}{S(\delta t)}\right)-\mu t\right]^2\right]^{\frac{1}{1-q}} \end{aligned} \tag{6.78}$$

where $\tilde{\beta}(t) = \beta(t)/\sigma^2$. This means that

$$\begin{aligned} P(\ln(S(t+\delta t))|\ln(S(t)))^{\frac{1-q}{2}} &= \frac{1}{Z^{\frac{1-q}{2}}(t)} \\ &\times \left[1-\tilde{\beta}(t)(1-q)\left[\ln\left(\frac{S(t+\delta t)}{S(\delta t)}\right)-\mu t\right]^2\right]^{\frac{1}{2}}. \end{aligned} \tag{6.79}$$

We can transform Eq. (6.69) by using $S(t+\delta t) = S(\delta t)e^Y$ with $\delta t = 0$. We obtain $S(t) = S(0)e^{Y(t)}$. Applying the Îto–Doeblin formula with $f = S(0)e^{Y(t)}$, which leads to $f_Y = S(0)e^{Y(t)}$ and $f_{YY} = S(0)e^{Y(t)}$, onto $f$ we obtain

$$\begin{aligned} dS(t) &= f_t\,dt + f_Y dY(t) + f_{YY}\,dY(t)dY(t) \\ &= S(t)\left(\mu + \frac{\sigma^2}{2}P(\Omega)^{1-q}\right)dt + \sigma S(t)P(\Omega)^{\frac{1-q}{2}}dW(t). \end{aligned} \tag{6.80}$$

Inserting Eq. (6.79) into Eq. (6.80) with $\delta t = 0$ and $S(0) = 1$ we obtain the following

$$\begin{aligned} dS(t) &= \left(\mu + \frac{\sigma^2}{2}\frac{1}{Z^{1-q}(t)}\left[1-\tilde{\beta}(t)(1-q)\left[\ln\left(S(t)\right)-\mu t\right]^2\right]\right)dt \\ &+ \frac{\sigma S(t)}{Z^{\frac{1-q}{2}}(t)}\left[1-\tilde{\beta}(t)(1-q)\left[\ln\left(S(t)\right)-\mu t\right]^2\right]^{\frac{1}{2}}dW(t). \end{aligned} \tag{6.81}$$

We can transform Eq. (6.81) a second time in the same way as Eq. (6.80) by letting the asset price $S(t)$ to transform to $x(t) = \ln(S(t))$ so that we can recast Eq. (6.81) in the same form as Eq. (6.48). In this case we have $f_S = 1/S(t)$ and $f_{SS} = -1/S(t)$, also

$$dS(t)dS(t) = \frac{\sigma^2 S^2(t)}{Z^{1-q}(t)}\left[1-\tilde{\beta}(t)(1-q)\left[\ln\left(S(t)\right)-\mu t\right]^2\right]dt. \tag{6.82}$$

Hence the stochastic differential equation for $dx(t)$

$$\begin{aligned} dx(t) &= f_t\,dt + f_S\,dS(t) + f_{SS}\,dS(t)dS(t) \\ &= \mu\,dt + \frac{\sigma}{Z^{\frac{1-q}{2}}(t)}\left[1-\tilde{\beta}(t)(1-q)\left[x(t)-\mu t\right]^2\right]^{\frac{1}{2}}. \end{aligned} \tag{6.83}$$

Using the formalism in Sect. 6.3.1 we can read off the coefficients from Eq. (6.83), these are

$$D^{(1)}(x,t) = \mu \tag{6.84}$$

$$D^{(2)}(x,t) = \frac{1}{2}\frac{\sigma^2}{Z^{1-q}(t)}\left[1-\tilde{\beta}(t)(1-q)\left[x(t)-\mu t\right]^2\right]. \tag{6.85}$$

Eq. (6.84) and Eq. (6.85) was obtained from Eq. (6.83) indeed if we use Eq. (6.30) and Eq. (6.31) with the help of Eq. (6.79) we see that we recover Eq. (6.69) as we should.

We are now in position to write down the Lagrangian functional. We will try first the Lagrangian definition which is given by Eq. (6.44), the differentials are evaluated as follows,

$$\frac{\partial}{\partial x}D^{(1)}(x,t) = 0, \tag{6.86}$$

$$\frac{\partial}{\partial x}D^{(2)}(x,t) = -\frac{\sigma^2}{Z^{1-q}(t)}(1-q)\tilde{\beta}(t)\left(x(t)-\mu t\right), \tag{6.87}$$

$$\frac{\partial^2}{\partial x^2}D^{(2)}(x,t) = -\frac{\sigma^2}{Z^{1-q}(t)}(1-q)\tilde{\beta}(t). \tag{6.88}$$

Inserting these into Eq. (6.44) the Lagrangian then becomes

$$\begin{aligned} \mathcal{L}(\dot{x}(t),x(t),t) &= -\frac{\sigma^2}{Z^{1-q}(t)}(1-q)\tilde{\beta}(t)\left(x(t)-\mu t\right) \\ &- \frac{\left[\dot{x}(t)-\left(\mu+\frac{2\sigma^2}{Z^{1-q}(t)}(1-q)\tilde{\beta}(t)\left(x(t)-\mu t\right)\right)\tau\right]^2}{\frac{2\sigma^2}{Z^{1-q}(t)}\left[1-\tilde{\beta}(t)(1-q)\left[x(t)-\mu t\right]^2\right]\tau}. \end{aligned} \tag{6.89}$$

We will also consider the Lagrangian representation given by Eq. (6.38). Consequently we may write the Lagrangian density as

$$\mathcal{L}(\dot{x}(t), x(t), t) = \frac{Z^{1-q}(t)\left[\dot{x}(t) - \mu\right]^2}{2\sigma^2\left[1 - \tilde{\beta}(t)(1-q)\left[x(t) - \mu t\right]^2\right]}, \tag{6.90}$$

which has a much simpler form than Eq. (6.89). Furthermore if we take the drift term $\mu$ to be 0, the Lagrangian takes the form of

$$\mathcal{L}(\dot{x}(t), x(t), t) = \frac{Z^{1-q}(t)\left[\dot{x}(t)\right]^2}{2\sigma^2\left[1 - \tilde{\beta}(t)(1-q)\left[x(t)\right]^2\right]}. \tag{6.91}$$

Eq.6.91 is the Lagrangian functional when the drift term is set to zero. This has the effect of simplifying the path integral, and this is acceptable as most of the physics in this model is embedded in the coloured noise.

In Sec. 6.4.1 we obtain some solutions for various cases using the method of least action principle, which consist in finding the most likely path in the path integral. This is carried out using the Euler–Lagrange equation of motion.

We now turn our attention to another class of stochastic differential equations. These models are non-Markovian and are based on the theory of fractals. The advantage of such model is that it represents a very good candidate when one wants to model volatility, because it takes memory into account. In the next section, we explore if it is possible to insert and compute the path integral using multi–fractal models.

### 6.3.4 The Multifractal Random Walk Model

It is now widely recognized that the simplicity of the popular Black-Scholes–Merton model, which relates derivative prices to current stock prices and quantifies risk through a constant volatility parameter, is no longer sufficient to capture modern market phenomena, especially since the 1987 crash. It has also been observed that real markets display high volatility as a stochastic process. The modelling of volatility as a stochastic process has been motivated *a priori* by empirical studies of stock price returns in which estimated volatility is observed to exhibit "random" characteristics (Fouque *et al.* 2000). This implies the Black–Scholes–Merton model is inadequate in describing the real market because it assumes constant volatility and also because the existence of volatility fluctuations, which are long–range correlated in time. In empirical studies (Lo 1991,

Ding and Granger 1996, Liu *et al.* 1997, Cont 2001, Muzy *et al.* 2000) it has been shown that the volatility correlation function decays very slowly in time and is well fitted by a power law. As a result, an interesting class of multifractal models, where the volatility is a Gaussian random variable with a correlation function that decays in time as a logarithm has been proposed (Muzy *et al.* 2000). In this study it was shown that the kurtosis of the process decreases only very slowly in contrast with most simple models of stochastic volatility, where the kurtosis drops exponentionaly with time (Pochart and Bouchaud 2002). This makes this model interesting for option pricing because it is consistent with smiles that flatten only very slowly with time (Bouchaud and Potters 2000a, Potters *et al.* 1998, Backus *et al.* 1997)

Multifractal models (Bacry *et al.* 2001, Muzy and Bacry 2002) have been used to account for scale invariance properties in different areas, such as energy dissipation or the velocity field in turbulent flows (Frish 1995) and in financial data (Bouchaud and Potters 2000a). The scale invariance properties of a deterministic fractal function $f(t)$ are generally characterized by the exponent, $\zeta_q$, which governs the power–law scaling of the absolute moments of its fluctuation, that is

$$m(q,l) = K_q l^{\zeta_q}. \tag{6.92}$$

Here $K_q$ is the memory kernel, with prescription $q$. Here $q$ is acting as scaling factor that governs the power law. The factor $l$ is acting as time lag in the fractal function $f(t)$. For example on can choose $m(q,l) = \sum_t |f(t+l) - f(t)|^q$. When $\zeta_q$ is linear in $q$ there is then a single scaling exponent $H$ which is involved, that is we have $\zeta_q = qH$ implying that the function $f(t)$ is *monofractal*. On the other hand if $\zeta_q$ becomes non–linear in $q$, then it implies that the function $f(t)$ is *multifractal*.

This can be extended to stochastic processes, $X(t)$ with stationary increments (Bacry *et al.* 2001, Muzy and Bacry 2002). In that case $m(q,l)$ takes the form

$$m(q,l) = \mathrm{E}\left[|X(t+l) - X(t)|^q\right], \tag{6.93}$$

where E is the expected value.

This model is the continuous time limit of a stochastic volatility model where log volatility correlations decay logarithmically. It possesses a stability property related to its scale invariance property for each timescale $\Delta t \leq T$. The returns at scale $\Delta t$ are defined as,

$$r_{\Delta t} \equiv \ln\left[\frac{\rho(t)}{\rho(t-\Delta t)}\right]. \tag{6.94}$$

This can be described by a stochastic volatility model,

$$r_{\Delta t} = \xi(t)\sigma_{\Delta t}(t) = \xi(t)e^{w_{\Delta t}(t)}, \tag{6.95}$$

where $\xi(t)$ is a standard Gaussian white noise independent of $w_{\Delta t}(t)$, which could be considered as a Gaussian process with mean and covariance defined as

$$\mu_{\Delta t} = \frac{1}{2}\ln\left(\sigma^2\Delta t\right), \tag{6.96}$$

$$C_{\Delta t} = \text{cov}\left[w_{\Delta t}(t), w_{\Delta t}(t+\tau)\right] = \lambda^2 \ln\left[\frac{T}{|\tau| + \Delta t\, e^{-3/2}}\right]. \tag{6.97}$$

Here $\sigma^2\Delta t$ is the return variance at scale $\Delta t$ and $T$ represents an integral time scale. Also $\lambda$ is a scale parameter. Typically the time parameter, $T$, and the scale parameter, $\lambda$, take values of one year and of around 0.02 respectively.

The MRW model can be expressed in a more familiar form in which the log–volatility $w_{\Delta t}(t)$ obeys an autoregressive equation whose solution is real,

$$w_{\Delta t}(t) = \mu_{\Delta t} + \int_{-\infty}^{t} d\tau\, \eta(\tau) K_{\Delta t}(t-\tau). \tag{6.98}$$

In this case the random function $\eta(\tau)$ denotes a standardized Gaussian white noise and the memory kernel $K(\cdot)$. The function $K(\cdot)$ ensures that the system evolves in a causal manner. In other words keep the system random. This function can be seen as the information flow, which implies that $w_{\Delta t}(t)$ represents the response of the market to increasing information up to date $t$.

At time $t$, the distribution of $w_{\Delta t}(t)$ is Gaussian with mean $\mu_{\Delta t}$ and variance

$$V_{\Delta t} = \int_0^{\infty} d\tau\, K_{\Delta t}^2(\tau) = \lambda^2 \ln\left[\frac{T\, e^{3/2}}{\Delta t}\right]. \tag{6.99}$$

Its autocovariance, which entirely specifies the randomness of the process is given by

$$C_{\Delta t}(\tau) = \int_0^{\infty} dt\, K_{\Delta t}(t) K_{\Delta t}(t+|\tau|). \tag{6.100}$$

Performing a Fourier transform we obtain,

$$\hat{K}_{\Delta t}^2(f) = \hat{C}_{\Delta t}(f) = 2\lambda^2 f^{-1}\left[\int_0^{Tf} \frac{\sin(t)}{t}\, dt + \mathcal{O}(f\Delta t \ln(f\Delta t)\right], \tag{6.101}$$

which shows that for $\tau$ small enough we have an expression for the kernel as a function of time,

$$K_{\Delta t}(f) \approx K_0\sqrt{\frac{\lambda^2 T}{\tau}} \quad \text{for} \quad \Delta t \ll \tau \ll T. \tag{6.102}$$

This slow power law decay of the memory kernel, Eq. (6.98), ensures the long range dependence and the multifractality of the stochastic volatility process, defined in Eq. (6.95).

### 6.3.5 The Path Integral for the Multifractal Random Walk Model

In Section 6.3.4 we reviewed the multifractal random walk model and now we derive the path integral for this model.

A multifractal is generally considered as a complicated model. It is a discrete model, which operates at all time scales, just like fractals do. Multifractal models are perfect candidates for studying volatility in asset prices. Here we are going to adopt the simplest approach to the problem, which is the direct use of the formalism developed and discussed in the earlier sections to see how far it takes us

The MRW model can be expressed in a more familiar form in which the log–volatility $w_{\delta t}(t)$ obeys an auto regressive equation whose solution is defined as in Eq. (6.98). Here $\delta t \equiv \Delta t$ by definition. Using Eq (6.102) we may rewrite the log–volatility as

$$w_{\delta t}(t) \approx \mu_{\delta t} + \int_{\infty}^{t} d\tau\, \eta(\tau) K_0 \sqrt{\frac{\lambda^2 T}{\tau}}, \quad \text{for} \quad \delta t \ll \tau \ll T. \tag{6.103}$$

The expression for the mean $\mu_{\delta t}$ is defined in Eq. (6.96) and by inserting $C_{\delta t}(0)$, using Eq. (6.97), we obtain a new expression for the mean

$$\mu_{\delta t} = \frac{1}{2} \ln\left(\sigma^2 \delta t\right) - \lambda^2 \ln\left[\frac{T}{e^{-3/2}\delta t}\right]. \tag{6.104}$$

Such that we may rewrite log–volatility as

$$w_{\delta t}(t) \approx \frac{1}{2} \ln\left(\sigma^2 \delta t\right) - \lambda^2 \ln\left[\frac{T}{e^{-3/2}\delta t}\right] + \int_{\infty}^{t} d\tau\, \eta(\tau) K_0 \sqrt{\frac{\lambda^2 T}{t-\tau}}. \tag{6.105}$$

Hence we can rewrite Eq. (6.94) as

$$\begin{aligned} r_{\Delta t} &= \xi(t)\sigma_{\Delta t}(t) = \xi(t) e^{w_{\Delta t}(t)} \\ &\approx \xi(t) \left[ e^{\ln(\sigma\sqrt{\delta t})} e^{\ln\left[\frac{T}{e^{-3/2}\delta t}\right]^{-\lambda^2}} \exp\left( \int_{\infty}^{t} d\tau\, \eta(\tau) K_0 \sqrt{\frac{\lambda^2 T}{t-\tau}} \right) \right] \qquad (6.106) \\ &= \xi(t) \left[ \sigma\sqrt{\delta t} \left[\frac{T}{e^{-3/2}\delta t}\right]^{-\lambda^2} \exp\left( \int_{\infty}^{t} d\tau\, \eta(\tau) K_0 \sqrt{\frac{\lambda^2 T}{t-\tau}} \right) \right]. \qquad (6.107) \end{aligned}$$

This is equivalent to rewriting the model in a stochastic differential equation of the form

$$\begin{aligned} dx &= e^{w_{\Delta t}(t)}\, dz + 0.\, dt, \implies \mu = 0, \\ &= \left[ \sigma\sqrt{\delta t} \left[\frac{T}{e^{-3/2}\delta t}\right]^{-\lambda^2} \exp\left( \int_{\infty}^{t} d\tau\, \eta(\tau) K_0 \sqrt{\frac{\lambda^2 T}{t-\tau}} \right) \right] dz, \qquad (6.108) \end{aligned}$$

where $\delta t\, dz$ is the standard Wiener process.

The option price for a path independent option is defined as in Eq. (6.49)

$$\mathcal{O}_{\mathcal{F}}(S,t) = e^{-r\tau} E_{(t,S)}\left[\mathcal{F}(e^{x_T})\right] = e^{-r\tau} \int_{-\infty}^{\infty} \left( \int_{x(t)=x_t}^{x(T)=x_T} \mathcal{F}(e^{x_T}) e^{-A[x(t')]} \mathcal{D}x(t') \right) dx_T,$$

where the action and Lagrangian functionals are defined as in Eq. (6.38). Here $\mu = 0$, so can rewrite the Lagrangian as

$$\mathcal{L} = \frac{1}{2\sigma^2} \left(\dot{x}(t')\right)^2, \quad \text{with} \quad \dot{x}(t') \equiv \frac{dx}{dt'}. \tag{6.109}$$

The discritization is carried out in exactly the same way as in Eq. (6.55) for the integral part and as in Eq. (6.56) for the derivative term respectively. Here $\Delta t$ is fixed at all times. However the volatility also depends on the time $t'$, we therefore get for each $i$, with a discrete log–volatility:

$$w_{\delta t}(t_i) \approx \frac{1}{2} \ln\left(\sigma^2 \delta t\right) - \lambda^2 \ln\left[\frac{T}{e^{-3/2}\delta t}\right] + \int_{\infty}^{t_i} d\tau\, \eta(\tau) K_0 \sqrt{\frac{\lambda^2 T}{t_i - \tau}}, \tag{6.110}$$

an action functional that takes the form of

$$A(x_i) = \frac{1}{2\Delta t^2} \sum_{i=0}^{N-1} \frac{(x_{i+1} - x_i)^2}{e^{2w_{\delta t}(t_i)}}. \tag{6.111}$$

The path integral is then written in the following way

$$\begin{aligned} & \int_{x(t)=x_t}^{x(T)=x_T} \mathcal{F}(e^{x_T}) e^{-A[x(t')]} \mathcal{D}x(t') \qquad (6.112) \\ = & \lim_{N\longrightarrow\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathcal{F}(e^{x_T}) e^{-A(x_i)} \frac{dx_1}{\sqrt{2\pi e^{2w_{\delta t}(t_1)}\Delta t}} \cdots \frac{dx_{N-1}}{\sqrt{2\pi e^{2w_{\delta t}(t_{N-1})}\Delta t}} \\ = & \lim_{N\longrightarrow\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathcal{F}(e^{x_T}) \exp\left[-\frac{1}{2\Delta t^2} \sum_{i=0}^{N-1} \frac{(x_{i+1} - x_i)^2}{e^{2w_{\delta t}(t_i)}}\right] \prod_{i=1}^{N-1} \frac{dx_i}{\sqrt{2\pi e^{2w_{\delta t}(t_i)}\Delta t}}. \end{aligned}$$

Let us start the integration for $x_2$ and $x_1$ and then give the general expression for the integration. We therefore have

$$\int_{-\infty}^{\infty} dx_1 \exp\left[-\frac{1}{2\Delta t^2}\left(\frac{(x_1 - x_0)^2}{e^{2w_{\delta t}(t_0)}} + \frac{(x_2 - x_1)^2}{e^{2w_{\delta t}(t_1)}}\right)\right] \frac{1}{\sqrt{2\pi e^{2w_{\delta t}(t_1)}\Delta t}}, \tag{6.113}$$

using the Gaussian equality, Eq. (A.19), defined as

$$\int_{-\infty}^{\infty} \exp\left[-a(x-z)^2 - b(z-y)^2\right] dz = \sqrt{\frac{\pi}{a+b}} \exp\left[-\frac{ab}{a+b}(x-y)^2\right], \tag{6.114}$$

we may rewrite Eq. (6.113) as

$$\begin{aligned} & \frac{1}{\sqrt{2\pi e^{2w_{\delta t}(t_1)}\Delta t}}\sqrt{\frac{\pi}{\frac{1}{2\Delta t^2 e^{2w_{\delta t}(t_1)}}+\frac{1}{2\Delta t^2 e^{2w_{\delta t}(t_0)}}}}\exp\left[-\frac{\frac{1}{2\Delta t^2 e^{2w_{\delta t}(t_1)}}\frac{1}{2\Delta t^2 e^{2w_{\delta t}(t_0)}}}{\frac{1}{2\Delta t^2 e^{2w_{\delta t}(t_1)}}+\frac{1}{2\Delta t^2 e^{2w_{\delta t}(t_0)}}}(x_2-x_0)^2\right] \\ = & \frac{\sqrt{\pi}}{\sqrt{2\pi e^{2w_{\delta t}(t_1)}\Delta t}}\sqrt{\frac{\prod_{i=0}^{1}2\Delta t^2 e^{2w_{\delta t}(t_i)}}{\sum_{i=0}^{1}2\Delta t^2 e^{2w_{\delta t}(t_i)}}}\exp\left[-\frac{1}{\sum_{i=0}^{1}2\Delta t^2 e^{2w_{\delta t}(t_i)}}(x_2-x_0)^2\right]. \end{aligned} \tag{6.115}$$

Repeating the process for the $N-1$ integrals we obtain

$$\begin{aligned} & \int_{x(t)=x_t}^{x(T)=x_T}\mathcal{F}(e^{x_T})e^{-A[x(t')]}\mathcal{D}x(t') \\ = & \lim_{N\longrightarrow\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\mathcal{F}(e^{x_T})\exp\left[-\frac{1}{2\Delta t^2}\sum_{i=0}^{N-1}\frac{(x_{i+1}-x_i)^2}{e^{2w_{\delta t}(t_i)}}\right]\prod_{i=1}^{N-1}\frac{dx_i}{\sqrt{2\pi e^{2w_{\delta t}(t_i)}\Delta t}} \\ = & \frac{(\sqrt{\pi})^{N-1}\mathcal{F}(e^{x_T})}{\sqrt{2\pi e^{2w_{\delta t}(t_{N-1})}\Delta t}}\sqrt{\frac{2\Delta t^2 e^{2w_{\delta t}(t_{N-1})}}{\sum_{i=0}^{N-1}2\Delta t^2 e^{2w_{\delta t}(t_i)}}}\exp\left[-\frac{1}{\sum_{i=0}^{N-1}2\Delta t^2 e^{2w_{\delta t}(t_i)}}(x_T-x_0)^2\right], \end{aligned} \tag{6.116}$$

where $t_{N-1}\equiv T$ and the multifractal component $w_\delta t(t_i)$ is defined as in Eq. (6.105), namely:

$$w_{\delta t}(t_i)\approx\frac{1}{2}\ln\left(\sigma^2\delta t\right)-\lambda^2\ln\left[\frac{T}{e^{-3/2}\delta t}\right]+\int_{\infty}^{t_i}d\tau'\,\eta(\tau')K_0\sqrt{\frac{\lambda^2 T}{t_i-\tau'}}, \tag{6.117}$$

with the exponential of the log–volatility expressed in Eq. (6.106) or in Eq. (6.108).

Eq. (6.117) is partially correct because if we observe Eq. (6.117), we can see that the Gaussian white noise term still remains in the equation. This suggests that using a one dimensional path integral is not sufficient and that in order to solve the integral fully one would need to consider at least a 2 dimensional path integral, that is one dimension in space and the other for the log-volatility. This is clearly beyond the scope of this thesis and would present itself as a Ph.D. thesis on its own to just tackle this problem well and accurately. This is suggested as an item for future research.

Section 6.4 describes the instanton method. The instanton method is useful in complicated, systems because it uses the principle of least action to determine the most probable path contribution to the action functional. The instanton path, in physics, gives a prescription for the way one can tunnel from one quantum state to the other.

## 6.4 The Instanton Method

As mentioned, the instanton method[18] uses the principle of least action to determine the most probable path contribution to the action functional.

The fundamental quantity of classical mechanics is the action, which is the time integral of the Lagrangian as defined in Eq. (6.38) or Eq. (6.44). The action functional has the ability to describe the entire dynamics of the system over the space in question. In other words this functional contains all of the information about the dynamics of the system.

The principle of least action states that when a system evolves from one given configuration to another between times $t_1$ and $t_2$, it does so along the path in configuration space for which the action is an extrema (normally a minimum). The minimum is the solution to the classical equation of motion, which is called the Euler-Lagrange equation. In particle physics this is known as the classical solution, which is motivated by the belief that a semi classical approach may shed some light on the underlying quantum world. In quantum physics it often happens that the ground states are degenerate, that is there is more than one vacuum state. This problem may be cured by allowing quantum tunneling between states. The prescription on the tunneling is what is called an instanton. In the financial context this concept can be used to find the most probable path that makes the greatest contribution in the path integral, this is useful because in the case of complicated models the integral may be otherwise very difficult to perform. Moreover the structure of the process (usually modeling the option price or the log return) would not be known until the problem is actually solved, however with the instanton method it is possible to obtain a solution for the equation of motion, which describes the entire system, and at the same time shed some light on the structure of the process in question.

Mathematically the principle of the least action condition can be written as

$$\delta \mathcal{A} = 0, \tag{6.118}$$

[18] Instantons have been been observed and applied in various areas in physics in particular in quantum chromodynamics (Bonnet *et al.* 2002b, Bonnet *et al.* 2000b) and the references therein. Instantons also appear in any theory that has complex topological spaces (Coleman 1985, Shifman 1994) such as in quantum mechanics, quantum field theory, QED, QCD, quantum gravity, topology, differential topology etc.

which leads to the Euler-Lagrange equation

$$\left[\frac{d}{dt}\left(\frac{\partial}{\partial \dot{x}(t)}\right) - \frac{\partial}{\partial x(t)}\right] \mathcal{L}(\dot{x}(t), x(t), t) = 0. \tag{6.119}$$

We now outline a general approach in evaluating the path integral in using the instanton solution: first we write down the Lagrangian functional; using this functional, the Euler–Lagrange equation is then evaluated using Eq. (6.119) to obtain the equation of motion for the system. One then has to find the solution of the differential equation. This solution is then inserted back into the path integral for evaluation. This path is the path that is the most likely path. The option price is then evaluated with a given payoff function[19].

### 6.4.1 The Instanton Method for the Non-Gaussian Model

#### Keeping $Z^{(1-q)}(t)$ and $\beta(t)$ Constant in Time

Taking the derivative with respect to both $x(t)$ and $\dot{x}(t)$ and supposing that $Z^{(1-q)}(t)$ and $\beta(t)$ are constant in time, namely we define $Z^{(1-q)}(t) = Z^{(1-q)}$ and $\beta(t) = \beta$ respectively, leads to the following Euler-Lagrangian equation

$$\begin{aligned} &\left[\frac{d}{dt}\left(\frac{\partial}{\partial \dot{x}(t)}\right) - \frac{\partial}{\partial x(t)}\right] \mathcal{L}(x(t), \dot{x}(t)) \\ = \; & Z^{(1-q)} \left[\frac{\ddot{x}(t)\left[1-\beta(1-q)x(t)^2\right] + \dot{x}^2(t)\beta(1-q)x(t)}{\left[1-\beta(1-q)x(t)^2\right]^2}\right] = 0. \end{aligned} \tag{6.120}$$

Now since $\left[1-\beta(1-q)x(t)^2\right] \neq 0$ similarly for $Z^{(1-q)} \neq 0$, we must have

$$\ddot{x}(t)\left[1-\beta(1-q)x(t)^2\right] + \dot{x}^2(t)\beta(1-q)x(t) = 0. \tag{6.121}$$

This equation is only true for $\beta(t)$ and $Z(t)$ constant in time, that is $\beta(t) \equiv \beta$ and $Z(t) = Z$.

In this case we may integrate and find that

$$\sinh^{-1}\left(x(t)\sqrt{\beta(1-q)}\right) = C_1 t - x(t)\sqrt{1-\beta(1-q)x(t)^2}, \tag{6.122}$$

[19] These payoffs come from the type of option taken into consideration. In Chapter 5 the payoff function for a option was explicitly written in Sec. 5.2, 5.3 and Sec. 5.4 for the European, some exotic and American options respectively

where we have to solve for $x(t)$. Alternatively we can use Maple to obtain a solution for the differential equation, Eq. (6.121). The solution obtained from the Maple software package is computed as

$$x(t) := \text{RootOf}\left(\text{hypergeom}([\frac{1}{2},\frac{1}{2}],[\frac{3}{2}],-\beta\,(-1+q)\,Z^2)\,Z\right. \tag{6.123}$$
$$\left. -\frac{1}{3}\,\frac{t\,(\text{hypergeom}([\frac{3}{2},\frac{3}{2}],[\frac{5}{2}],\sin(\frac{\%1\,S(0)}{\%2})^2)\,\%2\sin(\frac{\%1\,S(0)}{\%2})^3+3\,\%1\,S(0))}{\%2\sin(\frac{\%1\,S(0)}{\%2})\,S(0)}-\%1\,S(0)\right)$$
$$\%1 := \text{hypergeom}([\frac{1}{2},\frac{1}{2}],[\frac{3}{2}],-\beta\,(-1+q)\,S(0)^2)$$
$$\%2 := \text{RootOf}((-\beta+\beta\,q)\,Z^2+1).$$

The output in Eq. (6.123) may be simplified by evaluating the Gaussian hypergeometric functions directly and evaluating the term %2, which is just equivalent to $\sqrt{\frac{1}{\beta(1-q)}}$.

Substituting Eq. (6.123) back into the Lagrangian equation, Eq. (6.91) with $\mu = 0$, we obtain the following equation for the Lagrangian functional, using Maple

$$\mathcal{L}(\dot{x}(t),x(t),t) := \frac{1}{2}Z^{(1-q)}\,(\%3\,\%1\sin(\frac{\%2\,S(0)}{\%1})^3+3\,\%2\,S(0))^2\Big/((\tag{6.124}$$
$$-\text{hypergeom}([\frac{3}{2},\frac{3}{2}],[\frac{5}{2}],-\beta\,(-1+q)\,\%4)\,\beta\,(-1+q)\,\%4\,\%1\sin(\frac{\%2\,S(0)}{\%1})\,S(0)$$
$$+3\,\text{hypergeom}([\frac{1}{2},\frac{1}{2}],[\frac{3}{2}],-\beta\,(-1+q)\,\%4)\,\%1\sin(\frac{\%2\,S(0)}{\%1})\,S(0))^2\sigma^2$$
$$(1-(1-q)\,\beta\,\%4))$$
$$\%1 := \text{RootOf}((-\beta+\beta\,q)\,Z^2+1)$$
$$\%2 := \text{hypergeom}([\frac{1}{2},\frac{1}{2}],[\frac{3}{2}],-\beta\,(-1+q)\,S(0)^2)$$
$$\%3 := \text{hypergeom}([\frac{3}{2},\frac{3}{2}],[\frac{5}{2}],\sin(\frac{\%2\,S(0)}{\%1})^2)$$
$$\%4 := \text{RootOf}(3\,\text{hypergeom}([\frac{1}{2},\frac{1}{2}],[\frac{3}{2}],-\beta\,(-1+q)\,Z^2)\,Z\,\%1\sin(\frac{\%2\,S(0)}{\%1})\,S(0)$$
$$-t\,\%3\,\%1\sin(\frac{\%2\,S(0)}{\%1})^3-3\,t\,\%2\,S(0)-3\,\%2\,S(0)^2\,\%1\sin(\frac{\%2\,S(0)}{\%1}))^2,$$

where both Eq. (6.123) and Eq. (6.124) are written in Maple format for convenience.

This functional represents the functional for which the path integral has the highest probability, that is the most likely path. In order to calculate the action functional one

needs to integrate Eq. (6.124) with respect to time. Once the action functional has been evaluate it is then possible to calculate the path integral for this process.

### The General Case

We now return to the case when $\beta(t)$ and $Z(t)$ are not left constant but instead are allowed to vary with time, from Eq. (6.72), Eq. (6.73) and Eq. (6.75), we may take the time derivative of $\beta(t)$ and $Z(t)$ with $t' = 0$. In this case for $\beta'(t)$ we obtain the following expression,

$$\begin{aligned}\beta'(t) = \frac{d}{dt}\beta(t) &= \pi^{\frac{1-q}{3-q}}\left[(2-q)\,(3-q)\,t\right]^{-\frac{2}{3-q}}\left(\frac{\Gamma\left[\frac{1}{q-1}-\frac{1}{2}\right]^2}{(q-1)\,\Gamma\left[\frac{1}{q-1}\right]^2}\right)^{\frac{1-q}{3-q}} \\ &= -\frac{2}{(3-q)\,t}\,\beta(t). \end{aligned} \tag{6.125}$$

Similarly for the function $Z'(t)$

$$Z'(t) = \frac{d}{dt}Z(t) = \pi^{\frac{1}{3-q}}\left(\frac{(2-q)\,(3-q)\,t\,\Gamma\left[\frac{1}{q-1}-\frac{1}{2}\right]^2}{(q-1)\,\Gamma\left[\frac{1}{q-1}\right]^2}\right)^{\frac{1}{3-q}} = \frac{1}{(3-q)\,t}\,Z(t). \tag{6.126}$$

Now evaluating the Euler–Lagrange equation, we find that

$$\frac{\partial}{\partial\dot{x}(t)}\mathcal{L}(x(t),\dot{x}(t)) = \frac{Z(t)^{1-q}\,\dot{x}(t)}{1-(1-q)\,\beta(t)\,x(t)^2} \tag{6.127}$$

$$\frac{\partial}{\partial x(t)}\mathcal{L}(x(t),\dot{x}(t)) = \frac{(1-q)\,Z(t)^{1-q}\,\beta(t)\,x(t)\,\dot{x}(t)^2}{\left[1-(1-q)\,\beta(t)\,x^2(t)\right]^2}, \tag{6.128}$$

and taking the time derivative of Eq. (6.127),

$$\begin{aligned}\frac{d}{dt}\left(\frac{\partial}{\partial\dot{x}(t)}\right)\mathcal{L}(x(t),\dot{x}(t)) &= \frac{Z(t)^{-q}}{\left[1-(1-q)\,\beta(t)\,x^2(t)\right]^2} \\ &\times \Big((1-q)\,\dot{x}(t)\,\Big(\Big[1-(1-q)\,\beta(t)\,x^2(t)\Big]\,Z'(t) \\ &+ Z(t)\,x(t)\,\left(x(t)\,\beta'(t)+2\,\beta(t)\,\dot{x}(t)\right)\Big) \\ &+ Z(t)\,\Big[1-(1-q)\,\beta(t)\,x^2(t)\Big]\,\ddot{x}(t)\Big). \end{aligned} \tag{6.129}$$

In this case the definition for the time derivative of $\dot{x}(t)$ is equivalent to $\dot{x}(t)$, i.e. $\dot{x}(t) \equiv \dot{x}(t)$, similarly for $x(t)$ and $\ddot{x}(t)$.

Combining terms in the previous equation leads to the following Euler–Lagrange equation,

$$\begin{aligned} 0 &= \left[\frac{d}{dt}\left(\frac{\partial}{\partial \dot{x}(t)}\right) - \frac{\partial}{\partial x(t)}\right] \mathcal{L}(x(t), \dot{x}(t)) \\ &= \frac{Z(t)^{-q}}{\left[1-(1-q)\,\beta(t)\,x^2(t)\right]^2} \\ &\times \left((1-q)\,\dot{x}(t)\,\left[\left[1-(1-q)\,\beta(t)\,x^2(t)\right]\,Z'(t) + Z(t)\,x(t)\,\left(x(t)\,\beta'(t) + \beta(t)\,\dot{x}(t)\right)\right]\right. \\ &+ \left. Z(t)\,\left[1-(1-q)\,\beta(t)\,x^2(t)\right]\,\ddot{x}(t)\right). \end{aligned} \tag{6.130}$$

Now inserting Eqs. (6.125) and (6.126) into Eq. (6.130) we find that the Euler–Lagrange equation becomes

$$\begin{aligned} 0 &= \left[\frac{d}{dt}\left(\frac{\partial}{\partial \dot{x}(t)}\right) - \frac{\partial}{\partial x(t)}\right] \mathcal{L}(x(t), \dot{x}(t)) \\ &= \frac{Z(t)^{1-q}}{\left[1-(1-q)\,\beta(t)\,x^2(t)\right]^2} \\ &\times \left[(1-q)\,\dot{x}(t)\,\left(\frac{\left[1-(1-q)\,\beta(t)\,x^2(t)\right]}{(3-q)\,t} + x(t)\,\left(-\frac{2\,\beta(t)\,x(t)}{(3-q)\,t} + \beta(t)\,\dot{x}(t)\right)\right)\right. \\ &+ \left.\left[1-(1-q)\,\beta(t)\,x^2(t)\right]\,\ddot{x}(t)\right], \end{aligned} \tag{6.131}$$

dividing both sides of the equation by the factor sitting on the RHS we get

$$1 + t_i \frac{(3-q)}{(1-q)} \frac{\ddot{x}(t)}{\dot{x(t)}} = \frac{\beta(t)\,x(t)\,[(3-q)t\,\dot{x}(t) - 2x(t)]}{[1-(1-q)\,\beta(t)\,x^2(t)]}, \tag{6.132}$$

which is mathematically equivalent to

$$\begin{aligned} \dot{x}(t)\,[(1-q) - (3-q)\,(1-q)\,\beta(t)\,x(t)\,(x(t) - t\,\dot{x}(t))] + \\ (3-q)\,t\,\left[1-(1-q)\,\beta(t)\,x^2(t)\right]\,\ddot{x}(t) = 0. \end{aligned} \tag{6.133}$$

Eq. (6.132) is a non–trivial non–linear differential equation. Using Mathematica one obtains the following solution for $x(t)$, by isolating irrelevant functions of $q$, to simplify the preceding equation,

$$\begin{aligned} \gamma(q) &= \pi^{\frac{q}{q-3}}\,(q-3)^{1+\frac{2}{q-3}}\,(q-2)^{\frac{2}{q-3}}\,(q-1)^{1+\frac{1}{q-3}} \\ &\times \Gamma\left[\frac{1}{q-1} - \frac{1}{2}\right]^{\frac{2q}{q-3}}\,\Gamma\left[\frac{1}{q-1}\right]^{\frac{2}{q-3}} \end{aligned} \tag{6.134}$$

$$h(q) = \pi^{\frac{1}{q-3}}\,(q-3)\,(q-1)^{\frac{q}{q-3}}\,\Gamma\left[\frac{1}{q-1} - \frac{1}{2}\right]^{\frac{2}{q-3}}\,\Gamma\left[\frac{1}{q-1}\right]^{\frac{2q}{q-3}}, \tag{6.135}$$

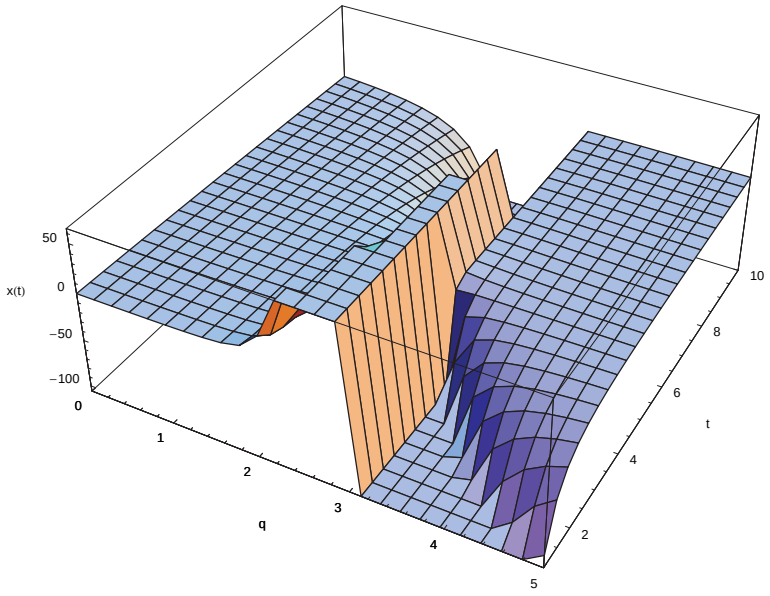


**Figure 6.4. The graph of $x(t)$ when $C_1 = C_2 = h(q) = 1$ and $\gamma(q) = 1$.** The graph of $x(t)$ when $C_1 = C_2 = h(q) = 1$ and $\gamma(q) = 1$. The discontinuities become evident at $q = 3$, we also observe divergence in the solution for $q > 2$.

we may rewrite the solution of the differential equation, Eq. (6.132), as

$$x(t) \to \frac{1}{8\,\eta(q)}\left(e^{-\frac{C_2\,\eta(q)}{q-3}}\,t^{-\frac{1+\eta(q)}{q-3}}\left(t^{\frac{2\,\eta(q)}{q-3}} - 16\,e^{2\left(C_1+\frac{C_2\,\eta(q)}{q-3}\right)}h(q)\right)\right) \tag{6.136}$$

$$x(t) \to \frac{1}{8\,\eta(q)}\left(e^{-\frac{C_2\,\eta(q)}{q-3}}\,t^{-\frac{1+\eta(q)}{q-3}}\left(e^{\frac{2\,C_2\,\eta(q)}{q-3}} - 16\,e^{2C_1}h(q)\,t^{\frac{2\,\eta(q)}{q-3}}\right)\right), \tag{6.137}$$

where the function $\eta(q)$ is given by $\eta(q) = \sqrt{1 + e^{2C_1}\gamma(q)}$.

Eqs. (6.136) and (6.137) are two equivalent solutions. We can therefore use both equations such that to extract the constant of integration one sets $x(t_0) \equiv \Omega_0$ in Eq. (6.136) and $x(t_0) \equiv \Omega_1$ in Eq. (6.137). We find that

$$C_1 = \frac{1}{2}\ln\left[\frac{16\,\Omega_1\,h(q)}{\Omega_0}\right], \tag{6.138}$$

$$C_2 = -\frac{1}{2}\ln\left(\Omega_1\right)\frac{(q-3)\sqrt{\Omega_0}}{\sqrt{\Omega_0 - \gamma(q)\Omega_1\,16\,h(q)}}. \tag{6.139}$$

In Fig. 6.4 we show graphically the structure of the solution by setting the constants to fixed values. In Fig. 6.4 we fix the $C_1 = C_2 = \gamma(q) = h(q) = 1$. If we set the constants to $C_1 = C_2 = h(q) = 1$ and $\gamma(q) = 10$ we would observe that as we increase the value of the $\gamma(q)$ by a factor of 10 the overall shape the graph remains almost identical,

however, the value $x(t)$ would scale up by a given factor, and the values of $x(t)$ for $3 \leq q \leq 5$ are pushed further back in time. In the other region, the one for $0 \leq q \leq 3$ we would notice that the non–zero values for $x(t)$ are shifted closer to the origin on the $q$ axis. This is consistent with the fat-tail effect.

One thing that these two pictures have in common is the discontinuity around $q = 3$. Now looking at Eq. (6.134) through Eq. (6.139) we see that for $q \leq 3$, $\gamma(q)$ becomes complex. This is however not consistent with the initial value of $q = 1.43$ that was determined from empirical fits by Borland (2002c), and Borland and Bouchaud (2004) on S&P 500 daily returns and NASDAQ stocks (1 minute interval). This represents a problem with the method, because if the $q$ value extracted from the empirical data cannot be used without applying some sort of renormalization, it would mean that the information extracted from the empirical fit cannot be carried into the path integral approach. This issue must be explored further to address the problem to find a way around this problem and avoid complex integration in the path integral and see if this is specific to only one data set or not. One of the possibilities to avoid complex integration is to perform what is called a Wick rotation in physics. A Wick rotation consists of mapping the time onto the imaginary plane, this way a complex time variable would become real value variable.

Using the solution of Eq. (6.136), we can also graph the probability distribution function, in Eq. (6.71). In Fig. 6.5 and Fig. 6.6 we show the graph of the probability distribution $P(x)$ for different $t$ values and different plot ranges. In Fig. 6.5 (left graph) we show the probability distribution function, $P(x)$, for $q \in [0,5]$ and $t \in [0,5]$, on this graph we can see where the function becomes complex valued and where it would be possible to integrate when this distribution is inserted into the Lagrangian functional. From this graph we clearly see that in the range of $3 \leq q \leq 5$ the distribution is well behaved and that there are no discontinuities, moreover we can see that at around $t \geq 3$ the function is starting to sharply increase to a large value. Extending into the $t$ direction to $t \leq 10$, shows that indeed there is a region where the distribution becomes singular, see Fig. 6.5 (right graph). If we now change the plotting range on both the $t$ and the $P(x)$ axis to see first how far the peak stretches and if there are other values of $t$ for which we have a singularity, as it is done in Fig. 6.6, we remark that first there does not appear to be any other singularity points and that the distribution remains smooth elsewhere. Furthermore we observe that the singularity is finite.

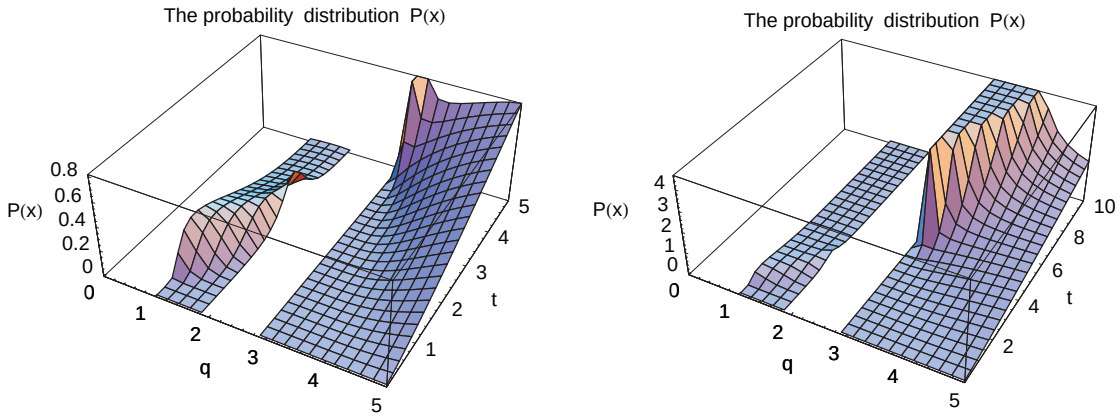


**Figure 6.5. The graph of $P(x(t))$ when $C_1 = C_2 = h(q) = 1$ and $\gamma(q) = 1$.** The graph of $P(x(t))$ when $C_1 = C_2 = h(q) = 1$ and $\gamma(q) = 1$. The discontinuities become evident. The probability distribution function will be integrable in the region of $q \in [1,2]$ and $q \in [3,5]$ for small $t$. The left shows $P(x)$ for $q \in [0,5]$ and $t \in [0.01,5]$, while the figure on the right shows the same graph but on a different $t$ range, i.e. $t \in [0.01,10]$.

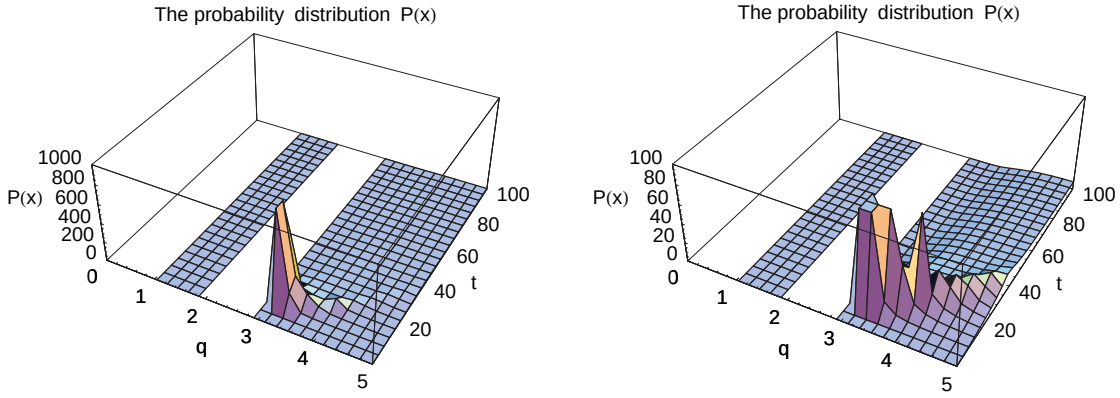


**Figure 6.6. Same graph as in Fig. 6.5 for $P(x(t))$ when $C_1 = C_2 = h(q) = 1$ and $\gamma(q) = 1$.** Same graph as in Fig. 6.5 for $P(x(t))$ when $C_1 = C_2 = h(q) = 1$ and $\gamma(q) = 1$. Here the discontinuities become more evident. This shows the structure of the peak, which is finite. The probability distribution function will be integrable in the region of $q \in [1,2]$ and $q \in [3,5]$ for small $t$. The left shows $P(x) \in [0,1000]$ for $q \in [0,5]$ and $t \in [0.01,100]$, while the figure on the right also shows $P(x(t))$, but on a different range, i.e. $P(x(t)) \in [0.01,100]$.

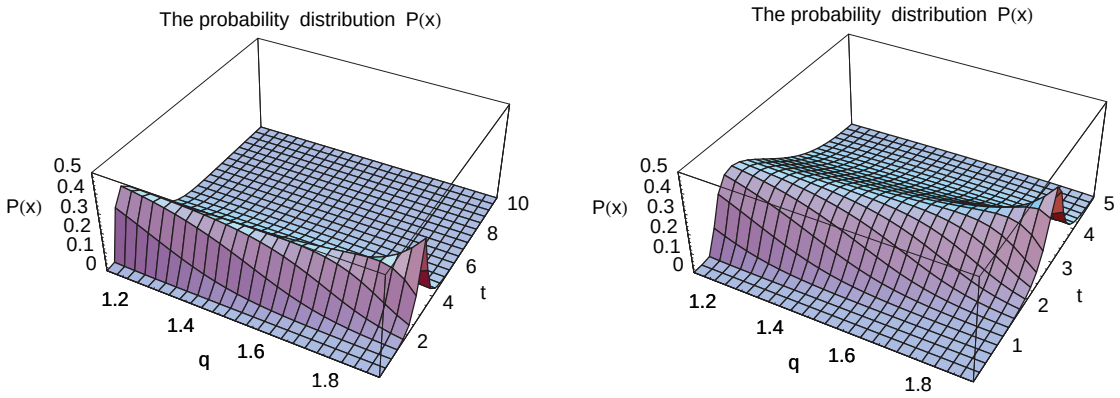


**Figure 6.7. The graph of $x(t)$.** The graph of $x(t)$ when $C_1 = C_2 = h(q) = 1$ and $\gamma(q) = 1$. Here $q$ is taken in the range of $1.1 \leq q \leq 1.8$ in both figures. On the right graph the range for $t$ is $0.01 \leq t \leq 5$ while on the left one the range is $0.01 \leq t \leq 10$.

We can also examine the distribution in the region of $q \in [1, 2]$, the region where the value of $q = 1.43$ has been extracted from empirical fits; see Borland (Borland 2002c, Borland and Bouchaud 2004) for more details. This is shown in Fig. 6.7. where can see that the distribution is well behaved, without singularities or discontinuities thus permitting integration of $x(t)$ with respect to $t$.

One of the main concerns with this solution is that now the range for the parameter $q$ has been shifted from its previous value of 1.43. We see that the solution becomes complex for $q \leq 3$, because of the function $\gamma(q)$ not been well defined. This is because in Eq. (6.134) some of the terms are negative raised to a power, which is not defined in the real plane, that is the term $(q-3)^{1+\frac{2}{q-3}}\,(q-2)^{\frac{2}{q-3}} \notin \mathcal{R}$ for $q < 3$, while the term involving the $\Gamma$ functions $\Gamma\left[\frac{1}{q-1} - \frac{1}{2}\right]^{\frac{2q}{q-3}}\,\Gamma\left[\frac{1}{q-1}\right]^{\frac{2}{q-3}}$ is defined in some of the regions for that $q$. On the other hand the first term just mentioned is perfectly defined for $q > 3$, while the term involving the $\Gamma$ functions is not. As a result, combining the two terms together makes an undefined function, i.e. Eq. (6.134), for all $q$. This represents a problem as the integration of complex functions is not always well defined, but may be overcome by taking only the absolute values of the terms $q-3$ and $q-2$. In this case, in the region of $q \in [0, 2]$, $\gamma(q)$ is well defined.

We can now examine the Lagrangian functional for the above case, as we did for the distribution $P(x)$, which is for the case when $C_1 = C_2 = h(q) = 1$ and $\gamma(q) = 1$. This is shown in Fig. (6.8).

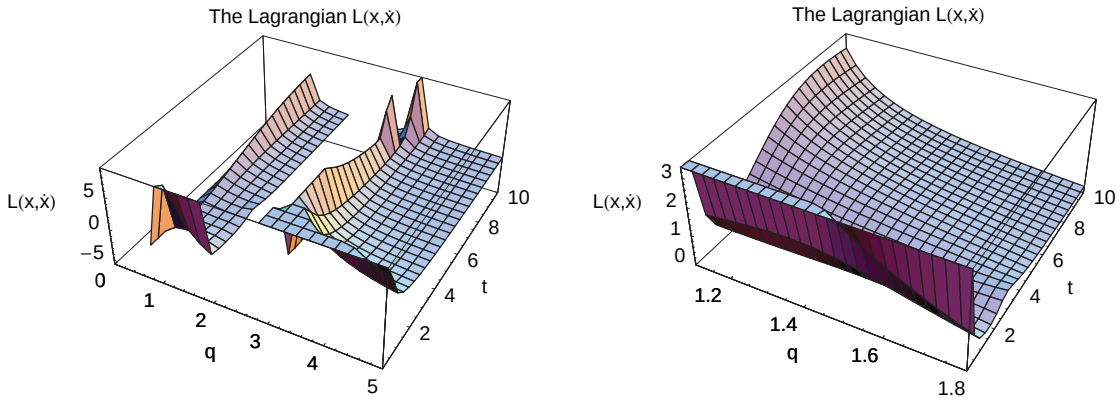


**Figure 6.8. The graph of the Lagrangian functional Eq. (6.91) for a given $x(t)$.** The graph of the Lagrangian functional Eq. (6.91) for a given $x(t)$, Eq. (6.136), when $C_1 = C_2 = h(q) = 1$ and $\gamma(q) = 1$. Here $t$ is taken in the range of $0.01 \leq t \leq 10$ in both figures. On the right graph the range for $t$ is $1.1 \leq q \leq 1.8$ while on the left one the range is $0 \leq q \leq 5$.

Similarly, we can see from the left graph in Fig. 6.8 where the discontinuities are. In this case too, when we plot $L(x, \dot{x})$ in the region $q \in [1.1, 1.8]$, we have a well behaved functional that should be integrable without too many difficulties. See Fig. 6.8 (right graph).

To conclude this section, one can say that it is possible to perform the path integrations, provided we limit the range of the $q$ value. This value must be extracted from real data and can only be within a small range, in this case it appears to be between 1 and 2, that is for $q \in [1, 2]$.

### The Solution When $(q-3)^{1+\frac{2}{q-3}}(q-2)^{\frac{2}{q-3}} \longrightarrow (|q-3|)^{1+\frac{2}{q-3}}(|q-2|)^{\frac{2}{q-3}}$.

As just discussed in Section 6.4.1, the function $\gamma(q)$, Eq. (6.134), is not well defined for most values of $q$ leading to a complex value solution and hence a complex path integral.

The problem may be addressed by setting

$$(q-3)^{1+\frac{2}{q-3}}(q-2)^{\frac{2}{q-3}} \longrightarrow (|q-3|)^{1+\frac{2}{q-3}}(|q-2|)^{\frac{2}{q-3}} \tag{6.140}$$

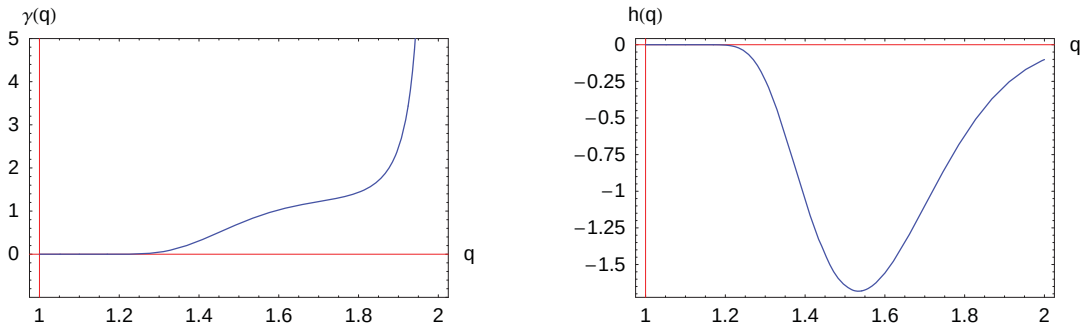


**Figure 6.9. The graph of the functions $\gamma(q)$ and $h(q)$.** The graph of the functions $\gamma(q)$, Eq. (6.134), (left figure) and $h(q)$, Eq. (6.135), (right figure) as a function of $q$ for $q \in [1,2]$ when the terms $(q-3)^{1+\frac{2}{q-3}}\,(q-2)^{\frac{2}{q-3}}$ in Eq (6.134) are set to $(|q-3|)^{1+\frac{2}{q-3}}\,(|q-2|)^{\frac{2}{q-3}}$.

in Eq. (6.134). The function $\gamma(q)$ then becomes

$$\begin{aligned}\gamma(q) \quad &= \quad \pi^{\frac{q}{q-3}}\,(q-3)^{1+\frac{2}{q-3}}\,(q-2)^{\frac{2}{q-3}}\,(q-1)^{1+\frac{1}{q-3}}\,\Gamma\left[\frac{1}{q-1}-\frac{1}{2}\right]^{\frac{2q}{q-3}}\,\Gamma\left[\frac{1}{q-1}\right]^{\frac{2}{q-3}} \\ &\longrightarrow \quad \pi^{\frac{q}{q-3}}\,(|q-3|)^{1+\frac{2}{q-3}}\,(|q-2|)^{\frac{2}{q-3}}\,(q-1)^{1+\frac{1}{q-3}}\,\Gamma\left[\frac{1}{q-1}-\frac{1}{2}\right]^{\frac{2q}{q-3}}\,\Gamma\left[\frac{1}{q-1}\right]^{\frac{2}{q-3}}.\end{aligned}$$

This has the effect of only considering the absolute value of $q-3$ and $q-2$ for all $q$. In that case it is then possible to obtain a graph for $\gamma(q)$ and $h(q)$, that is Eq. (6.134) and Eq. (6.135) respectively, as a function of $q$ on a small interval.

In Fig. 6.9 we show the graph of $\gamma(q)$ (left graph) and $h(q)$ (right graph) as a function of $q$. In these graphs we can see the shape of the curve, and we can also see that $\gamma(q) = 0$ for $q \leq 1.25$. For $q \in [1.25, 1.8]$ the function is stable, while for $q \geq 1.8$ it strongly diverges to a large value. Now looking at the right graph in Fig. 6.9, that is the graph for $h(q)$, on the interval of $q \in [1,2]$ we see a well behaved function. If we look outside this interval $h(q)$ becomes a highly oscillating function for $q < 1$, while remaining finite and non–divergent, at $q = 2$ and $q = 3$ it is zero.

Combining the Eqs. (6.135) and Eq. (6.141) into Eq. (6.138) and Eq. (6.139), setting $\Omega_0 = \Omega_1 = 1$ in $C_1$ and $\Omega_0 = 1$, $\Omega_1 = 2$ in $C_2$, it is possible to graph the coefficients $C_1$ and $C_2$, defined in Eq. (6.138) and Eq. (6.139) respectively, as seen Fig. 6.10.

On the left graph of Fig. 6.10 we show the graph of the coefficient $C_1$ on a large $q$ interval, i.e. $q \in [-5,5]$. From this graph it is very easy to see how the function behaves. We can also see regions where it is possible to use these coefficients. Especially in the region of $q \in [1,2]$ where we see a perfectly smooth function. As for $C_2$, which is graphed

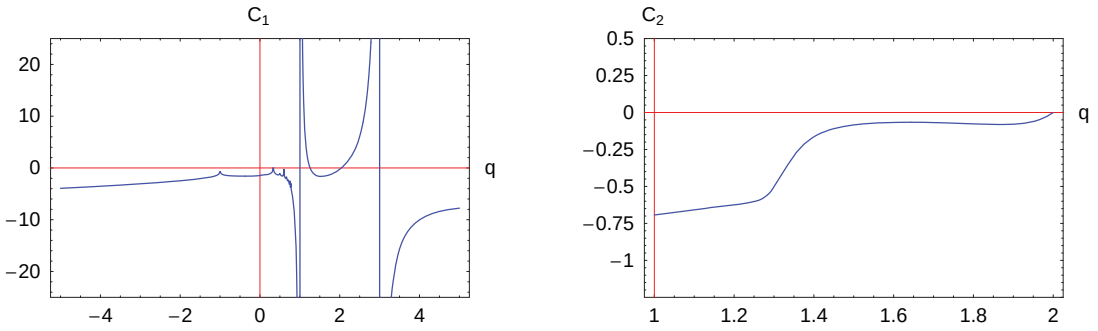


**Figure 6.10. The graph of the coefficients $C_1$ and $C_2$.** The graph of the coefficients $C_1$, Eq. (6.138) (left graph), and $C_2$, Eq. (6.139) (right graph) as a functions $q$ with the functions $\gamma(q)$, Eq. (6.134) and $h(q)$, Eq. (6.135) as a function of $q$ for $q \in [1,2]$ when the terms $(q-3)^{1+\frac{2}{q-3}}\,(q-2)^{\frac{2}{q-3}}$ in Eq (6.134) are set to $(|q-3|)^{1+\frac{2}{q-3}}\,(|q-2|)^{\frac{2}{q-3}}$.

on the right hand side of Fig. 6.10, there too we find a smooth function for $q \in [1,2]$. Although not shown on the figure, for $q \leq 1$ we observed a highly oscillating function with non divergent sharp peaks and for $q \geq 3$, $C_2 = 0$.

We can proceed in graphing $x(t)$ with the above functions as input. In Fig. 6.11 the graph of $x(t)$ is shown on two different $q$ intervals, that is $q \in [1.1, 1.99]$ on the right and $q \in [0,5]$ on the left. On both graphs $x(t)$ is plotted over $t \in [0.01, 10]$. On the left graph we notice that $x(t)$ take a very large value for $q \to 3$. This is due to the $C_1$ and $\gamma(q)$ contributions in the terms involving these two in Eq. (6.136) and Eq. (6.137). Looking at the right graph in Fig. 6.11 we see a smooth and non divergent surface.

In Fig. 6.12 we show the plot of the probability distribution function $P(x)$ for the same $q$ interval but on two different $t$ intervals. In a similar way we obtain a graph for the Lagrangian density, $L(x, \dot{x})$. This is shown in Fig. 6.13.

To summarize, it is possible to avoid the problem of complex integration in the path integral, because of $\gamma(q)$ being not well defined for all $q$, and certainly not for the $q$ value extracted from empirical data in Borland (2002c), and Borland and Bouchaud (2004). However by setting $(q-3)^{1+\frac{2}{q-3}}\,(q-2)^{\frac{2}{q-3}} \longrightarrow (|q-3|)^{1+\frac{2}{q-3}}\,(|q-2|)^{\frac{2}{q-3}}$ in Eq. (6.134) and $\sqrt{\Omega_0 - \gamma(q)\Omega_1\, 16\, h(q)} \longrightarrow \sqrt{|\Omega_0 - \gamma(q)\Omega_1\, 16\, h(q)|}$, in Eq. (6.139), it is possible to obtain smooth surfaces for the solution $x(t)$ of the Euler–Lagrange, Eq. (6.132). Similarly for the probability distribution function, $P(x)$, and consequently for the Lagrangian density $L(x, \dot{x})$ for a small $q$ region, i.e. $q \in [1.1, 1.99]$. As a result it is possible to carry out the integrations of the path integral.

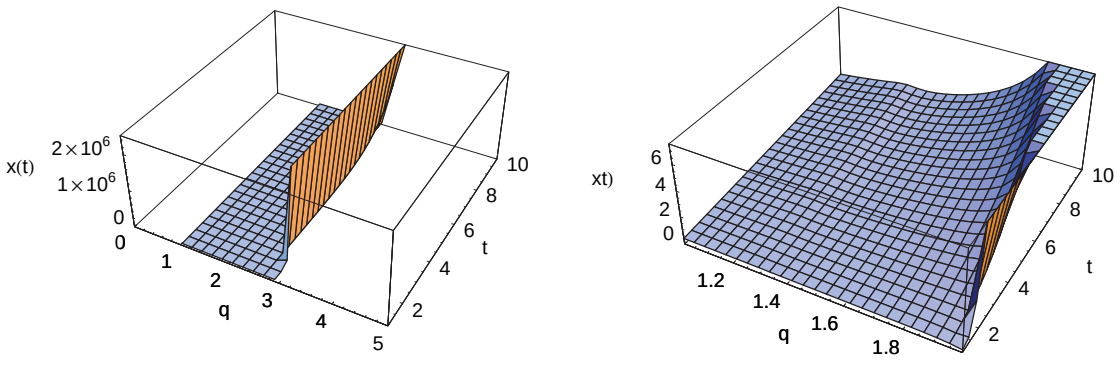


**Figure 6.11. The graph of the solution $x(t)$ of the Euler–Lagrange eqaution.** The graph of the solution $x(t)$ for two different $q$ intervals (left figure $q \in [0,5]$ and for $q \in [1.1, 1.99]$ on the right figure). Here the coefficients $C_1$, Eq. (6.138), and $C_2$, Eq. (6.139), are functions of $q$ with the above functions $\gamma(q)$, Eq. (6.134), and $h(q)$, Eq. (6.135), as a function of $q$ for $q \in [1,2]$ when the terms $(q-3)^{1+\frac{2}{q-3}}\,(q-2)^{\frac{2}{q-3}}$ in Eq (6.134) are set to $(|q-3|)^{1+\frac{2}{q-3}}\,(|q-2|)^{\frac{2}{q-3}}$.

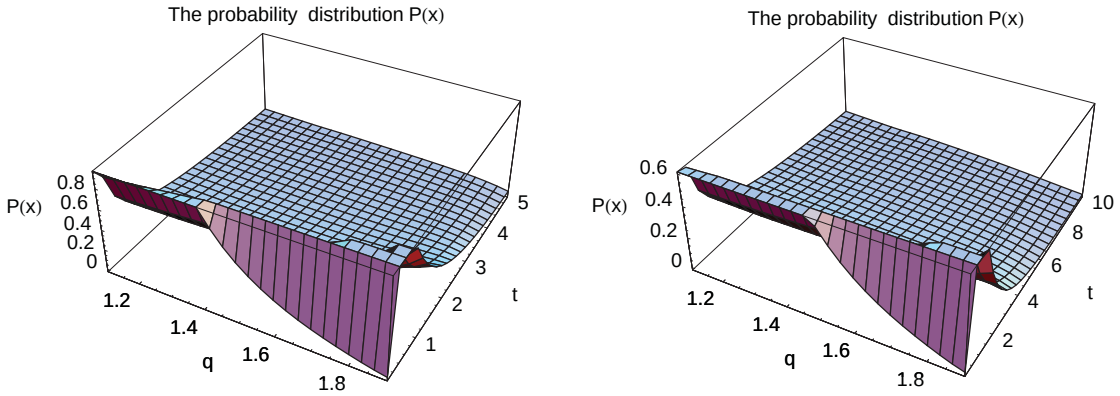


**Figure 6.12. The graph of the solution $x(t)$ of the Euler–Lagrange eqaution.** The graph of the probability distribution function $P(x)$ using the solution $x(t)$ for two different $t$ intervals (left figure $t \in [0,5]$ and $t \in [0,10]$ on the right figure, both figures are plotted over $q \in [1.1, 1.99]$). Here the coefficients $C_1$, Eq. (6.138), and $C_2$, Eq. (6.139), are functions of $q$ with the above functions $\gamma(q)$, Eq. (6.134), and $h(q)$, Eq. (6.135), as a function of $q$ for $q \in [1,2]$ when the terms $(q-3)^{1+\frac{2}{q-3}}\,(q-2)^{\frac{2}{q-3}}$ in Eq (6.134) are set to $(|q-3|)^{1+\frac{2}{q-3}}\,(|q-2|)^{\frac{2}{q-3}}$.

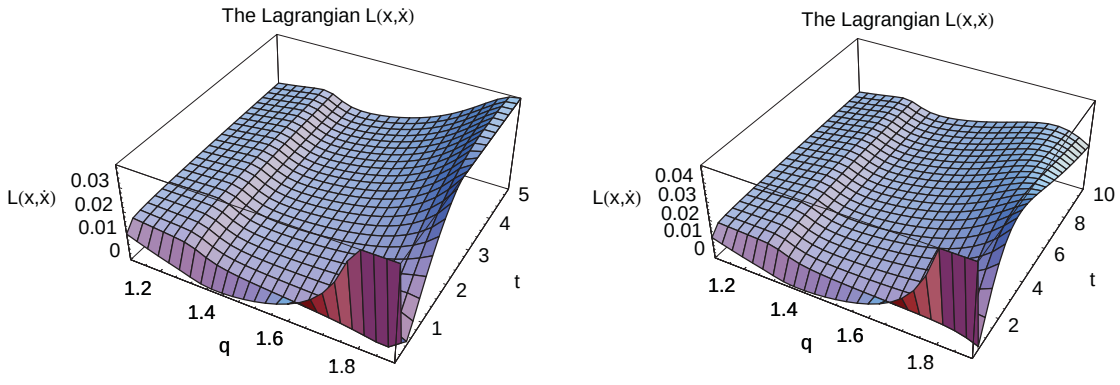


**Figure 6.13. The Lagrangian density.** The graph of the Lagrangian density function $L(x,\dot{x})$ using the solution $x(t)$ and the probability distribution function plotted in Fig. 6.12 for two different $t$ intervals (left figure $t \in [0,5]$ and $t \in [0,10]$ on the right figure, both figures are plotted over $q \in [1.1,1.99]$). Here the coefficients $C_1$, Eq. (6.138), and $C_2$, Eq. (6.139), are functions of $q$ with the above functions $\gamma(q)$, Eq. (6.134), and $h(q)$, Eq. (6.135), as a function of $q$ for $q \in [1,2]$ when the terms $(q-3)^{1+\frac{2}{q-3}}(q-2)^{\frac{2}{q-3}}$ in Eq (6.134) are set to $(|q-3|)^{1+\frac{2}{q-3}}(|q-2|)^{\frac{2}{q-3}}$.

## 6.5 Numerical Approach to the Evaluation of the Path Integral

In this section we use an alternative approach to the method developped in the previous section. Here we exploit the idea of the discretization to see if it is possible to arrive at a solution for the path integral.

By using the Chapman-Kolmogorov equation on a given time interval for a given number of equally spaced intervals combined with the transition probability, which were obtained from the Fokker–Plank equation we can approximate the path integral in that way. In Sec. 6.3.2 we saw how this was done in the simple case of the Black–Scholes–Merton equation, in section. 6.5.1 we attempt to use the same idea for the non–Gaussian model described in Sec. 6.3.3

### 6.5.1 Discretisation of the Path Integral

In earlier section, that is in Sec. 6.3.2, we saw how one could discretize the path integral. Here we use the Chapman–Kolmogorov equation repeadly. In Fig. 6.14, we show the

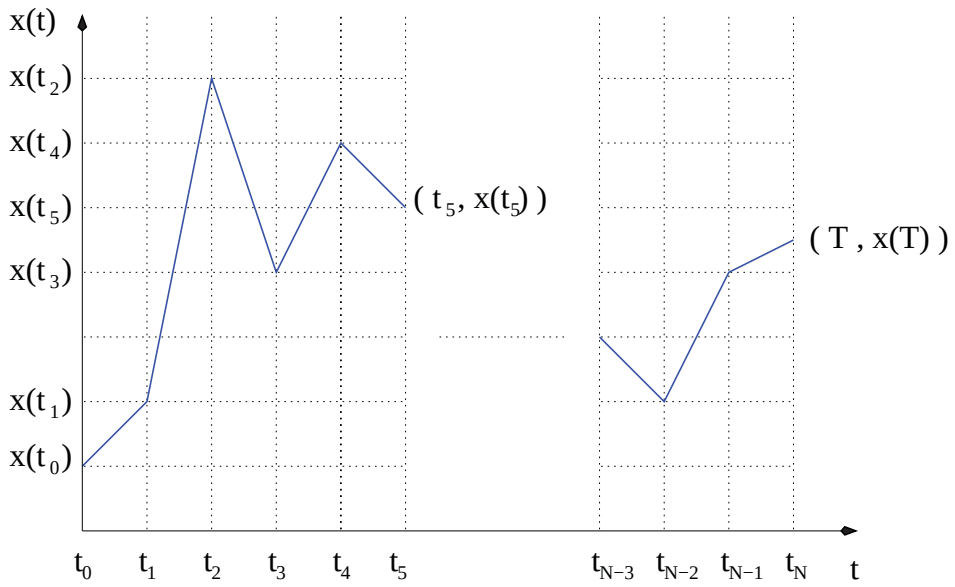


**Figure 6.14. The discrete path with $N-1$ paths.** The discretized path when the number of discrete path is equal to $N-1$. Here the $t_i$ are the time slices at a particular time. Ideally we would take the number of slices to infinity. In that case we obtain the continous limit. The more time slices we have the more accurate the path integral will be.

discretized time interval from $t_0$ to $T$. This corresponds to Eq. (6.27) where the intermediate transition probability are evaluated at each time slice. When the Lagrangian functional is inserted in that equation we obtain, Eq. (6.33).

Let us first consider special cases of $N$. As $N$ increases the number of integrals increases and the more complicated the integrand becomes. Let us start with the case when $N = 1$.

### The Transition Probability When $N = 1$

This is shown in Fig. (6.15). In this case we just have the option price is being given by

$$\begin{aligned}\mathcal{O}(S(t),t) &= \mathrm{E}^{t,x}\left[e^{-r\tau}\mathcal{F}[e^{x(T)}]\right], \quad \tau = T-t. \\ &= \int_{-\infty}^{\infty} dx(T) e^{-r\tau}\mathcal{F}[e^{x(T)}]P(x(T),T|x(t),t), \end{aligned} \tag{6.141}$$

where the transition probabilities are extracted from the Fokker–Planck equation, Eq. (6.70), and is given by Eq. (6.71), that is

$$P(x(T),T|x(t),t) = \frac{1}{Z(T)}\left[1-\beta(T)(1-q)(x(T)-x(t))^2\right]^{\frac{1}{1-q}}. \tag{6.142}$$

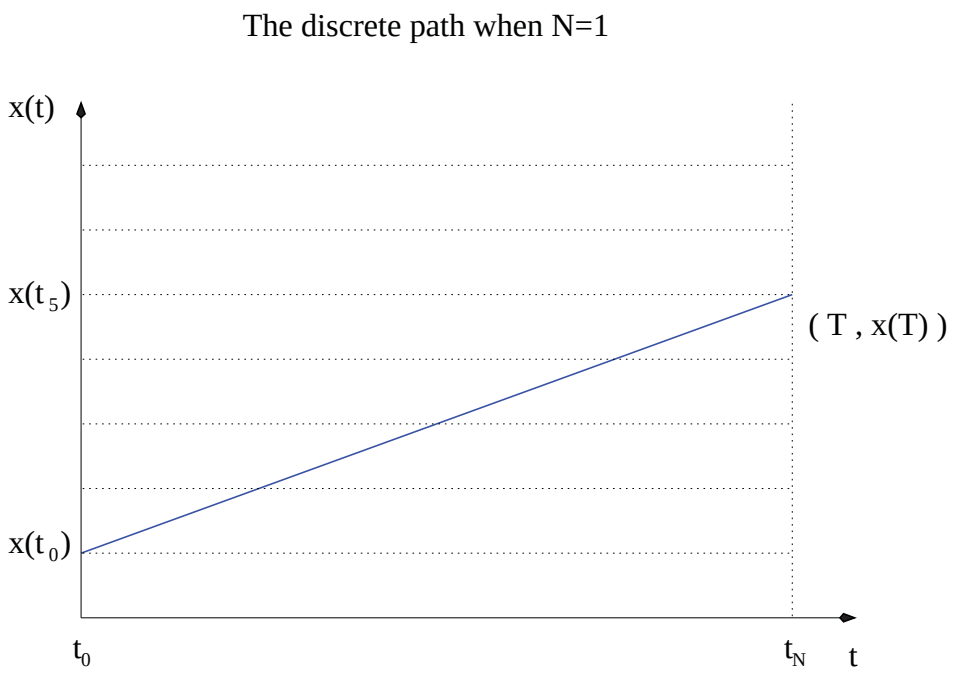


**Figure 6.15. The discrete path when $N = 1$.** The discretized path when the number of discrete path is equal to $N = 1$. This is the lowest number of time slices as possible, in this case the path integral must have the simplest integrand

In this case we have the solution obtained by Borland (2002a), and references therein (Borland 2002b, Borland 1998a)

### The Transition Probability When $N = 2$

We now consider the case when $N = 2$, as shown in Fig. 6.16 which is a two step transition probability. In this case we need to evaluate

$$P(x(T),T|x(t_0),t_0) = \int_{-\infty}^{\infty} dx(t_1) P(x(T),T|x(t_1),t_1) P(x(t_1),t_1|x(t_0),t_0), \quad (6.143)$$

where

$$P(x(T),T|x(t_1),t_1) = \frac{1}{Z(T)}\left[1-\beta(T)(1-q)(x(T)-x(t_1))^2\right]^{\frac{1}{1-q}} \quad (6.144)$$

$$P(x(t_1),t_1|x(t_0),t_0) = \frac{1}{Z(t_1)}\left[1-\beta(t_1)(1-q)(x(t_1)-x(t_0))^2\right]^{\frac{1}{1-q}} \quad (6.145)$$

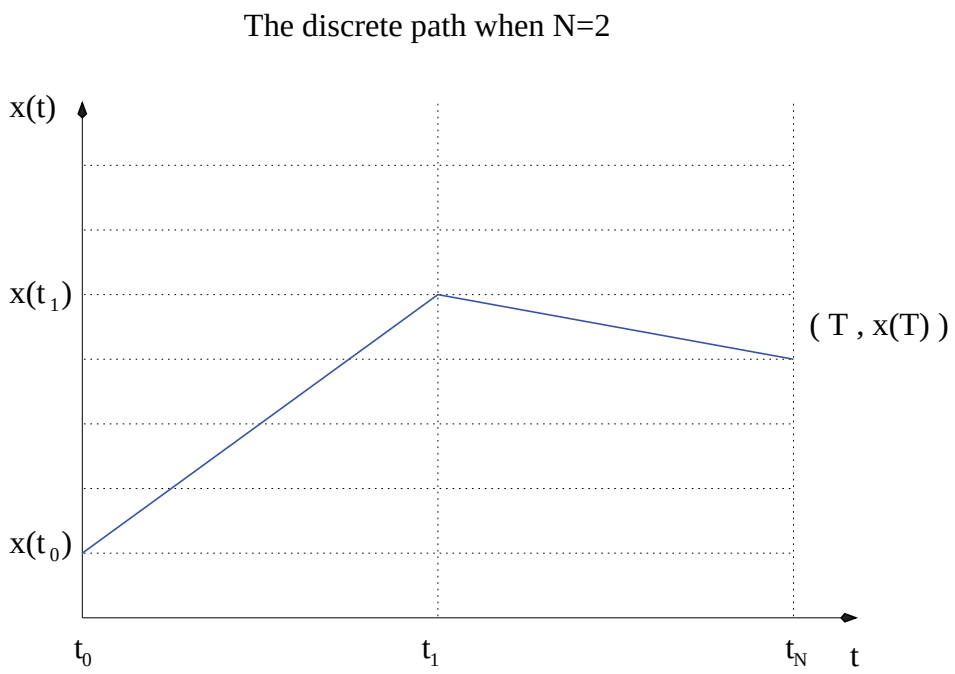


**Figure 6.16. The discrete path when $N = 2$.** The discretized path when the number of discrete path is equal to $N = 2$. This is the next to simplest case with only two time slices.

hence for the case $N = 2$, Eq. (6.142), becomes

$$\begin{aligned} P(x(T),T|x(t_0),t_0) &= \int_{-\infty}^{\infty} dx(t_1)\frac{1}{Z(T)Z(t_1)}\left[\left[1-\beta(T)(1-q)(x(T)-x(t_1))^2\right]\right. \\ &\times \left.\left[1-\beta(t_1)(1-q)(x(t_1)-x(t_0))^2\right]\right]^{\frac{1}{1-q}} \qquad (6.146) \\ &= \int_{-\infty}^{\infty} dx(t_1)\frac{1}{Z(T)Z(t_1)}\left[1\right. \\ &- (1-q)\left[\beta(t_1)(x(t_1)-x(t_0))^2+\beta(T)(x(T)-x(t_1))^2\right] \\ &+ \left.(1-q)^2\left[\beta(T)\beta(t_1)(x(T)-x(t_1))^2(x(t_1)-x(t_0))^2\right]\right]^{\frac{1}{1-q}}. \end{aligned}$$

### The Transition Probability When $N = 3$

We now move onto the case when $N = 3$ as shown in Fig. 6.17. As for the previous two cases we are interested in evaluating

$$\begin{aligned} P(x(T),T|x(t_0),t_0) &= \int_{-\infty}^{\infty} dx(t_2)\int_{-\infty}^{\infty} dx(t_1) \qquad (6.147) \\ &\times P(x(T),T|x(t_2),t_2)P(x(t_2),t_2|x(t_1),t_1)P(x(t_1),t_1|x(t_0),t_0), \end{aligned}$$

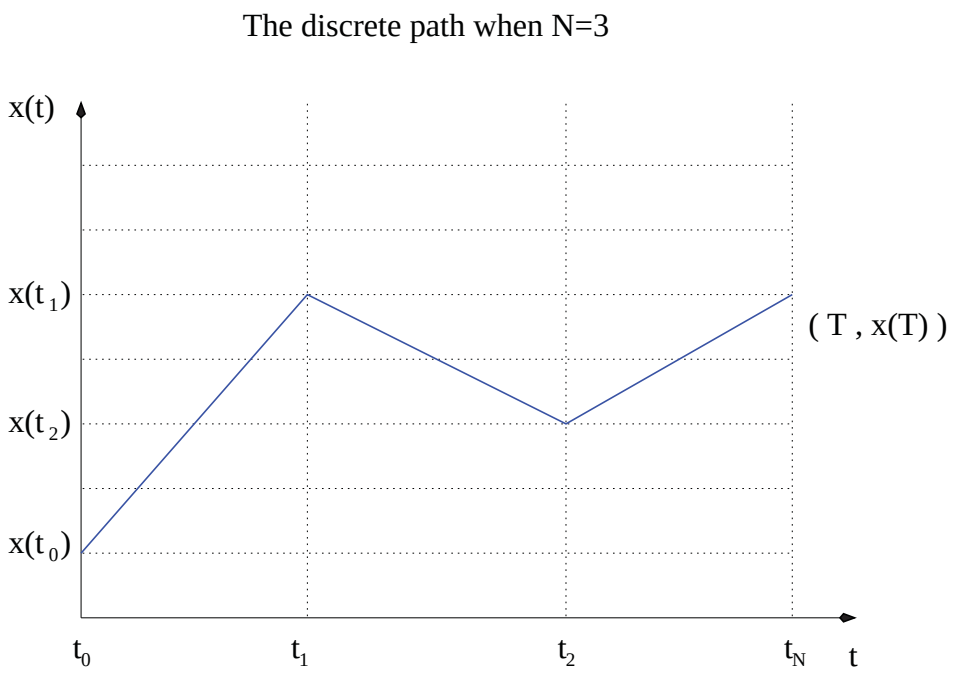


**Figure 6.17. The discrete path when $N = 3$.** The discretized path when the number of discrete path is equal to $N = 3$.

where the transitions probabilities are define in the usual way, these are

$$P(x(T),T|x(t_2),t_2) = \frac{1}{Z(T)}\left[1-\beta(T)(1-q)(x(T)-x(t_2))^2\right]^{\frac{1}{1-q}} \quad (6.148)$$

$$P(x(t_2),t_2|x(t_1),t_1) = \frac{1}{Z(t_2)}\left[1-\beta(t_2)(1-q)(x(t_2)-x(t_1))^2\right]^{\frac{1}{1-q}} \quad (6.149)$$

$$P(x(t_1),t_1|x(t_0),t_0) = \frac{1}{Z(t_1)}\left[1-\beta(t_1)(1-q)(x(t_1)-x(t_0))^2\right]^{\frac{1}{1-q}}. \quad (6.150)$$

Combining these equations together, we obtain

$$
\begin{aligned}
& P(x(T),T|x(t_0),t_0) \\
= & \int_{-\infty}^{\infty} dx(t_2) \int_{-\infty}^{\infty} dx(t_1) \frac{1}{Z(T)Z(t_2)Z(t_1)} \\
\times & \left\{ \left[ 1-\beta(T)(1-q)(x(T)-x(t_2))^2 \right] \left[ 1-\beta(t_2)(1-q)(x(t_2)-x(t_1))^2 \right] \right. \\
\times & \left. \left[ 1-\beta(t_1)(1-q)(x(t_1)-x(t_0))^2 \right] \right\}^{\frac{1}{1-q}} \\
= & \int_{-\infty}^{\infty} dx(t_2) \int_{-\infty}^{\infty} dx(t_1) \frac{1}{Z(T)Z(t_2)Z(t_1)} \\
\times & \left\{ 1-(1-q) \left[ \beta(t_2)(x(t_2)-x(t_1))^2 + \beta(t_1)(x(t_1)-x(t_0))^2 \right. \right. \\
+ & \left. \beta(T)(x(T)-x(t_2))^2 \right] \\
+ & (1-q)^2 \left[ \beta(T)\beta(t_2)(x(T)-x(t_2))^2(x(t_2)-x(t_1))^2 \right. \\
+ & \beta(T)\beta(t_1)(x(T)-x(t_2))^2(x(t_1)-x(t_0))^2 \\
+ & \left. \beta(t_2)\beta(t_1)(x(t_2)-x(t_1))^2(x(t_1)-x(t_0))^2 \right] \\
- & (1-q)^3 \left[ \beta(T)\beta(t_2)\beta(t_1)(x(T)-x(t_2))^2 \right. \\
\times & \left. \left. (x(t_2)-x(t_1))^2(x(t_1)-x(t_0))^2 \right] \right\}^{\frac{1}{1-q}}. \qquad (6.151)
\end{aligned}
$$

### The Transition Probability When $N = 4$

We now move onto the case when $N = 4$ as shown in Fig. 6.18. As for the prvious two cases we are interested in evaluating

$$
\begin{aligned}
P(x(T),T|x(t_0),t_0) = & \int_{-\infty}^{\infty} dx(t_3) \int_{-\infty}^{\infty} dx(t_2) \int_{-\infty}^{\infty} dx(t_1) \qquad (6.152) \\
\times & P(x(T),T|x(t_3),t_3) P(x(t_3),t_3|x(t_2),t_2) \\
\times & P(x(t_2),t_2|x(t_1),t_1) P(x(t_1),t_1|x(t_0),t_0).
\end{aligned}
$$

where the transitions probabilities are define in the usual way, these are

$$
\begin{aligned}
P(x(T),T|x(t_3),t_3) &= \frac{1}{Z(T)} \left[ 1-\beta(T)(1-q)(x(T)-x(t_3))^3 \right]^{\frac{1}{1-q}} \qquad (6.153) \\
P(x(t_3),t_3|x(t_2),t_2) &= \frac{1}{Z(t_3)} \left[ 1-\beta(t_3)(1-q)(x(t_3)-x(t_2))^2 \right]^{\frac{1}{1-q}} \qquad (6.154) \\
P(x(t_2),t_2|x(t_1),t_1) &= \frac{1}{Z(t_2)} \left[ 1-\beta(t_2)(1-q)(x(t_2)-x(t_1))^2 \right]^{\frac{1}{1-q}} \qquad (6.155) \\
P(x(t_1),t_1|x(t_0),t_0) &= \frac{1}{Z(t_1)} \left[ 1-\beta(t_1)(1-q)(x(t_1)-x(t_0))^2 \right]^{\frac{1}{1-q}}. \qquad (6.156)
\end{aligned}
$$

Combining these equations together in the same way as the previous section, we obtain

$$
\begin{aligned}
& P(x(T),T|x(t_0),t_0) \\
= & \int_{-\infty}^{\infty} dx(t_3) \int_{-\infty}^{\infty} dx(t_2) \int_{-\infty}^{\infty} dx(t_1) \frac{1}{Z(T)Z(t_3)Z(t_2)Z(t_1)} \\
\times & \left\{ \left[ 1-\beta(T)(1-q)(x(T)-x(t_3))^2 \right] \left[ 1-\beta(t_3)(1-q)(x(t_3)-x(t_2))^2 \right] \right. \\
\times & \left. \left[ 1-\beta(t_2)(1-q)(x(t_2)-x(t_1))^2 \right] \left[ 1-\beta(t_1)(1-q)(x(t_1)-x(t_0))^2 \right] \right\}^{\frac{1}{1-q}} \\
= & \int_{-\infty}^{\infty} dx(t_2) \int_{-\infty}^{\infty} dx(t_1) \frac{1}{\prod_{i=1}^{N} Z(t_i)} \\
\times & \left\{ 1-(1-q) \left[ \sum_{i=1}^{N} \beta(t_i)(x(t_i)-x(t_{i-1}))^2 \right] \right. \\
+ & (1-q)^2 \left[ \frac{1}{2!} \sum_{i,j=1,i\neq j}^{N} \beta(t_i)\beta(t_j)(x(t_j)-x(t_{j-1}))^2 (x(t_i)-x(t_{i-1}))^2 \right] \\
- & (1-q)^3 \left[ \frac{1}{3!} \sum_{i,j,k=1,i\neq j\neq k}^{N} \beta(t_i)\beta(t_i)\beta(t_k) \right. \\
\times & \left. (x(t_k)-x(t_{k-1}))^2 (x(t_j)-x(t_{j-1}))^2 (x(t_i)-x(t_{i-1}))^2 \right] \\
+ & \left. (1-q)^4 \left[ \prod_{i=1}^{N} \beta(t_i)(x(t_i)-x(t_{i-1}))^2 \right] \right\}^{\frac{1}{1-q}}. \\
= & \int_{-\infty}^{\infty} dx(t_3) \int_{-\infty}^{\infty} dx(t_2) \int_{-\infty}^{\infty} dx(t_1) \frac{1}{\prod_{i=1}^{N} Z(t_i)} \\
\times & \left\{ 1-(1-q) \left[ \sum_{i=1}^{N} \beta(t_i)(x(t_i)-x(t_{i-1}))^2 \right] \right. \\
+ & (1-q)^2 \left[ \frac{1}{2!} \sum_{i,j=1,i\neq j}^{N} \beta(t_i)\beta(t_j)(x(t_j)-x(t_{j-1}))^2 (x(t_i)-x(t_{i-1}))^2 \right] \\
- & (1-q)^3 \left[ \frac{1}{3!} \sum_{j_1,j_2,j_3=1,j_1\neq j_2\neq j_3}^{N} \left( \prod_{l=1}^{3} \beta(t_{j_l})(x(t_{j_l})-x(t_{j_l-1}))^2 \right) \right] \\
+ & \left. (1-q)^4 \left[ \prod_{i=1}^{N} \beta(t_i)(x(t_i)-x(t_{i-1}))^2 \right] \right\}^{\frac{1}{1-q}}. \qquad (6.157)
\end{aligned}
$$

Eq. (6.157) is derived when $N = 4$. The parameter $N$ stands for the number of time slices. In the continuum the number of time slices is infinite, i.e. $N \to \infty$. Therefore in order to evaluate the path integral we must take the limit as $N \to \infty$, i.e. when the path approaches the continuum limit; until then it is just an approximation. Now we obtain an expression in the limit $N \to infty$. The path integral is in fact a repeated

application of the Chapman-Kolmogorov equation. Consequently as an alternative approach to the one discussed in the earlier section, here we approach the problem differently to see if even when we are given the probability transition we can actually evaluate the integrals in the infinite limit.

In general we take the value of $q$ to be between 1 and 2 as was observed by Borland (2002a) and as we saw in Sec. 6.3.3 and Sec. 6.4.1. In this case

$$\lim_{n\longrightarrow\infty}(1-q)^n|_{q\in[1,2],n\in N}=\begin{cases}0 & \text{if } q=1\\ 0 & \text{if } q=1.43\\ \pm 1 & \text{if } q=2.\end{cases} \tag{6.158}$$

If we graph $f(x)=(1-Q)^x$ where $q=1.43$ we see that as $n$ gets large $(1-q)^n$ goes to zero relatively fast. Consequently we can suppress the contribution of higher terms in $q$. Now supposing that we keep the terms up to $(1-q)$ we can approximate the path integral as

$$\begin{aligned}&P(x(T),T|x(t_0),t_0)\\ &=\lim_{N\longrightarrow\infty}\overbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}^{N-1}\frac{\prod_{i=1}^{N}dx(t_i)}{\prod_{i=1}^{N}Z(t_i)}\\ &\quad\times\left\{1-(1-q)\left[\sum_{i=1}^{N}\beta(t_i)(x(t_i)-x(t_{i-1}))^2\right]+\mathcal{O}\left((1-q)^2\right)\right\}^{\frac{1}{1-q}}\\ &\approx\lim_{N\longrightarrow\infty}\overbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}^{N-1}\frac{\prod_{i=1}^{N}dx(t_i)}{\prod_{i=1}^{N}Z(t_i)}\left\{1-(1-q)\left[\sum_{i=1}^{N}\beta(t_i)(x(t_i)-x(t_{i-1}))^2\right]\right\}^{\frac{1}{1-q}}.\end{aligned} \tag{6.159}$$

Eq. (6.159) is an alternative approach to the Lagrangian formulation of the path integral discussed earlier. Here we consider the transition probabilities obtained directly from the Fokker-Planck equation to evaluate the path integral by direct insert and performing a repeated use of the Chapman-Kiolmogorov equation. In the next section we try to evaluate this integral for the case $N=2$ and the case $N=3$. However to obtain an accurate answer one has to obtain a solution when $N\to\infty$.

### 6.5.2 Evaluating the Path Intgeral for $N=2$

In this section we evaluate the Eq. (6.146) up to first order in $(1-q)$. We can perform the integration for $P(x(T),T|x(t_0),t_0)$ as shown in Eq. (F.1). Where the short hand notation are expressed as Eq. (F.2).

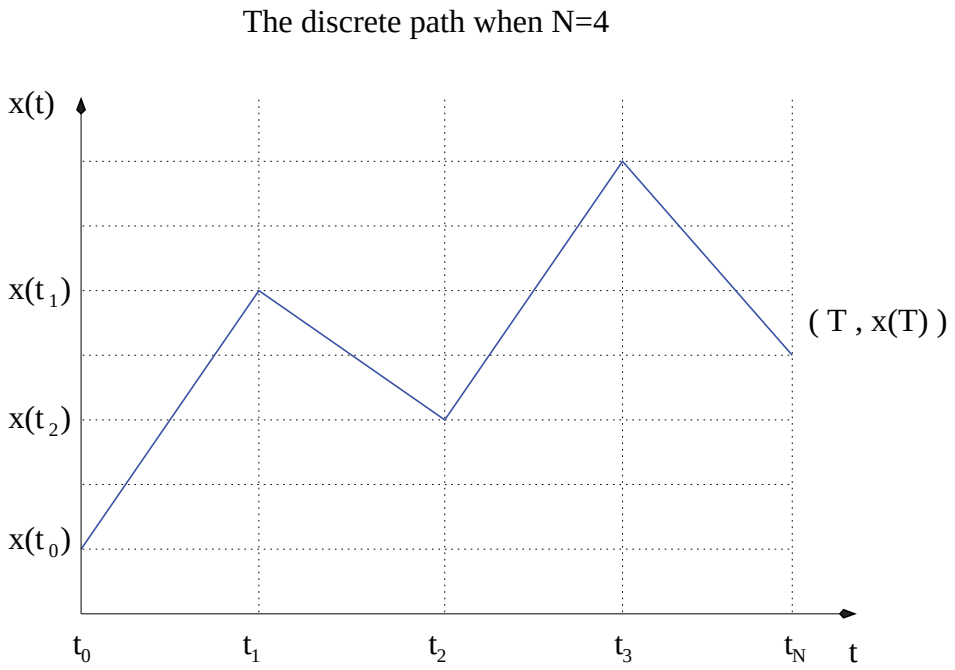


**Figure 6.18. The discrete path when $N = 4$.** The discretized path when the number of discrete path is equal to $N = 4$.

We quickly realize that the resulting integral does not take a simple form. This resulting equation needs to be integrated over $x(T)$ to obtain the option price. This cannot be integrated analytically and numerical approach is required at this point.

Furthermore we also realize that here too the value used in Borland (2002a) for $q$ cannot be used here either because the $P(x(T), T|x(t_0), t_0)$ becomes complex and has discontinuities. Making the integration virtually not possible.

### 6.5.3 Maple Output for the Path Intgeral When $N = 3$

Here we evaluate Eq. (6.151) up to first order in $(1 - q)$. When we insert the expressions for $\beta(t)$ and $Z(t)$ we obtain the following equation, Eq. (6.160)

$$
\begin{aligned}
& P(x(T),T|x(t_0),t_0) \\
= & \int_{-\infty}^{\infty} dx(t_2) \int_{-\infty}^{\infty} dx(t_1) \frac{1}{Z(T)Z(t_2)Z(t_1)} \\
& \left(1-(1-q)\left(\left(\frac{\pi\,\Gamma(\frac{1}{q-1}-\frac{1}{2})^2}{(q-1)\,\Gamma(\frac{1}{q-1})^2}\right)^{(\frac{1-q}{3-q})} ((2-q)\,(3-q)\,t_2)^{(-2\frac{1}{3-q})}\,(x(t_2)-x(t_1))^2\right.\right. \\
& \quad + \left(\frac{\pi\,\Gamma(\frac{1}{q-1}-\frac{1}{2})^2}{(q-1)\,\Gamma(\frac{1}{q-1})^2}\right)^{(\frac{1-q}{3-q})} ((2-q)\,(3-q)\,t_1)^{(-2\frac{1}{3-q})}\,(x(t_1)-x(t_0))^2 \qquad (6.160) \\
& \quad \left.\left. + \left(\frac{\pi\,\Gamma(\frac{1}{q-1}-\frac{1}{2})^2}{(q-1)\,\Gamma(\frac{1}{q-1})^2}\right)^{(\frac{1-q}{3-q})} ((2-q)\,(3-q)\,T)^{(-2\frac{1}{3-q})}\,(x(T)-x(t_2))^2\right)\right)^{(\frac{1}{1-q})}.
\end{aligned}
$$

We can perform the integration for $P(x(T),T|x(t_0),t_0)$ but only for the first time slice, that is only for the $dx(t_1)$ integral as shown in Eq. (F.3). This means that we are left with the second integral which cannot be performed because the integrand becomes intractable after the first integration as shown in Eq. (F.3). Where the short hand notation are expressed as in Eq. (F.4). As a result we cannot complete the calculation and an alternative method for evaluating these integrals must be used.

To summarize in both cases, that is for $N = 2$ and $N = 3$ we have approximated the path to its lowest order in $q$ and in both cases we saw that it led to not tractable solutions. We therefore conclude that in order to carry out these calculations one would need to use alternative approach to evaluate the path integrals this way.

### 6.5.4 Monte Carlo Methods for the Path Integral

Monte Carlo techniques (Glasserman 2003) are usually used when the path integral cannot be evaluated analytically. This is the case in lattice quantum chromodynamics when Monte Carlo methods are used instead of the analytical approach, where millions of integral would be necessary in order to evaluate the path integral for one configuration see Bonnet *et al.* (2000a), and (Bonnet *et al.* 2002b, Bonnet *et al.* 2004, Bonnet *et al.* 2002b, Bonnet *et al.* 2001b, Bonnet *et al.* 2000b, Bonnet *et al.* 2000a, Zhang *et al.* 2004, Bonnet *et al.* 2002c, Bonnet *et al.* 2001a, Rebbi 1983) and all of the references therein. This is

of course humanly impossible even if everybody in a country like China were to work on it eight hours a day for one year, hence the need for a numerical approach, and supercomputing.

### 6.5.5 Pertubation Theory for the Path Integral

Perturbation theory (Peskin and Schroeder 1995, Itzykson and Zuber 1980) is also an alternative approach, and is mostly used in quantum mechanics. In finance the use of perturbation theory becomes a lot more obscure because the perturbation parameter becomes unclear. Nevertheless it can be an alternative approach to the other methods when possible but will only bring an approximation to the problem and will therefore not lead to a more accurate evaluation for the option price. Although it still remains a very nice academic problem it will not improve what we already know about the evaluation of the option price and it will not be able to compete with other more useful methods such as the binomial method. The binomial method is the method of choice in the financial sector, because it is possible to evaluate the option price within seconds. We will therefore not explore the method of perturbation theory in this work and leave it for further studies.

### 6.5.6 Variational Methods for the Path Integral

The variational method was used in quantum electrodynamics by Feynman (1948). The method also works very well in quantum mechanics, but once again in finance it is not so clear on how to apply the method. We will not discuss this method in this thesis. This is left open for further research.

## 6.6 Chapter Summary

In this chapter we explored the use of path integrals in a financial context. In Section 6.1 we gave the Feynman derivation for the path integral applied to quantum mechanics. We then applied this formalism to a financial setting by making direct connections with stochastic differential equations. In Section 6.3 we saw how we can apply this formalism to the Black–Scholes–Merton model in a very efficient way. This is because the formalism in quantum mechanics for a free particle particle is essentially the same as the Black–Scholes–Merton model. That the drift and diffusion terms

are left constant. This formulation leads simple Lagrangian which has the form of a Gaussian functional. These kind of functionals can be easily integrated using Gaussian identities A.7.2. Hence the path path integral can be evaluated without difficulties and making any approximations. At the end of Sec 6.1 it was already pointed out that the Feynman formulation was hill defined and could not be used on every quantum mechanical system and that the path integral could not be evaluated for system with atomic potentials in particular potential of the Coulomb type, that is potentials of the form $V(x) \rightarrow -\frac{1}{|x|}$. In other words these kind of potentials make the Lagrangian functional non–Gaussian.

The formulation of the Lagrangian functional is not unique and really depends on the structure of the SDE, we therefore cannot apply the same Lagrangian functional for all SDE, the formulation of the Lagrangian had to be generalized and this was carried out in Sec. 6.3.1. This formulation is also not unique and must take into account the dependence of drift and diffusion term on time and process variable. These Lagrangian formulations enabled us to write down the Lagrangian functional for non–Gaussian processes like the one in Sec. 6.3.3. Unfortunately these formulations lead to intractable path integrals and we therefore cannot evaluate the path integral this way.

An alternative approach is to use the method of least action—this finds the path that contributes the most in the path integral and the path which minimizes the action functional. We referred to this approach as the instanton method. We can obtain such a path by directly solving the Euler–Lagrange equation. The Euler–Lagrange equation often leads to highly non–trivial, and non–linear differential equations. Nevertheless it was possible to solve such differential equations and obtain a solution. The instanton method is the most promising way in evaluating the path integral when using the Lagrangian method. This was done in Sec. 6.4 for a non–Gaussian model developed by Borland (2002a). Unfortunately in both cases, that is when we make most of the variables constants as in Sec. 6.4.1 and trying to make the problem as trivial as possible or when we consider the general case, as in Sec. 6.4.1, we quickly see that these also lead to intractable Lagrangians. So in order to obtain a solution one would have to turn to numerical solutions. In Section 6.4.1 it was shown that the solution leads to complex integration and lead to an intractable solution as well. In other words, both the Lagrangian formulation and the instanton method both lead to intractability in the solution. We then moved to an alternative approach such as the one described in Section 6.5, which also turned out to be not practical and quickly became intractable as

well. Also shown in Section 6.4.1 the value of q needs to be renormalized when using the path integral within the instanton framework because it does not lie within the range resulting in a complex integrand making integration extremely challenging, if at all possible. That is it may not even lie in the range of $q < 5/3$ (Gaussian regime) or $5/3 < q < 3$ (Lévy regime). This means that the distribution may simply be unknown, i.e. neither Gaussian nor Lévy. In addition to the level of complexity the value obtain in the Borland (2002a) model for $q$ does not correspond the possible value that we can use because the solution becomes complex, therefore forcing us to make further assumptions, as we saw in Sec. 6.4.1. It also means that the value of $q$ needs to be renormalized somehow so that it fits the new range.

On a different approach to the Lagrangian functional approach is to use the transition probabilities directly. This is derived from the Fokker-Planck equation, Eq. (4.196) Sec. 4.5.2. Then by repeated use of the Chapman–Kolmogorov equation, Eq. (6.27) (see Appendix A.5) we can evaluate the path integral for a given number of time slices, $N$. Ideally one wants to take the limit of $N$ to infinity to approximate the continuous limit. The idea there, is to evaluate the Chapman–Kolmogorov equation for $N = 2$ and then increase $N$ to larger values so that pattern can be identify. We can then turn the integrand as a power series of the parameters in question. It is then possible to truncate the expansion to a given order of that parameter. This was carried out in Sec. 6.5.1 where we saw in the case of the Borland (2002a) that we were able to do this with the parameter $q$. This has led us to Eq. (6.159). Furthermore, Eq. (6.159) is the most useful result, because it gives a very compact and very good approximation to the path integral in the limit as $N \to \infty$. This formula should be possible to be evaluated numerically. Since we are interested mostly in analytical solutions for this problem we then tried to evaluate the path integral for very small time slices this was done in Sec. 6.5.1 for $N = 2$ and in Sec. 6.5.1 for $N = 3$. In the first case we were able to avaluate the path integral but in the case $N = 3$ we were not able to carry out the second integral , the one with respect with $dx(t_2)$, and we clearly saw that the integral became intractable. As a result for any higher values of $N$ we can clearly see that the path integral rapidly becomes intratable and it is therefore not possible to obtain an analytical solution this way either.

As future propects in this approach it may be possible to use Eq. (6.159) to turn the non–Gaussian path integral into a Gaussian path integral which would then allow us to derive an analytical solution. This will be the subject of future research and propects.

In the next chapter we turn to a different approach to pricing. We use game theory to evaluate pricing function using real data. In particular we use the minority game which is then extended to the $–Game to simulate the price functions using real data and also as a possible tool for the detection of bubbles in stock markets.

Chapter 7

# Agent models

IT is commonly known in economics that markets follow both positive and/or negative trends, crashes and bubble effects. In general a strong positive trend is followed by a crash. Famous examples of these effects were seen in the recent crash on the NASDAQ (April 2000) and prior to the crash on the Hong Kong market, which was associated with the Asian crisis in the early 1994. In this chapter we use real market data input into a *minority* game with a variable payoff function and a non–linear super exponential model for bubbles, to explore financial bubbles. By changing the payoff function in the minority game we study how one can get the price function to follow the dynamics of a real market.

## 7.1 Introduction

Before the seminal paper on the minority game (Chalet and Zhang 1997), there were a great number of physicists already exploring various economic related issues. Around the late 90's there were various groups who proposed multi–agent models for the stock market (Arthur *et al.* 1997b, Arthur *et al.* 1997a, Caldarelli *et al.* 1997, Lux and Marchesi 1999)—these were important studies that showed that interacting agent models could produce realistic price histories, with crashes, clustered volatility, chronic bubbles, and depression. However the biggest problem with these models was that the relevant features of the interaction were buried under so many parameters that a systematic understanding was unclear. This is mainly because the market mechanisms are intrinsically non–linear, which means small variations in any of the parameters can lead to dramatic changes and one can never be sure which aspect is responsible for which price movement.

In order to get around this problem one has to adopt a completely different strategy. In physics the usual procedure in constructing models is to start from the simplest model, capturing the essential features in question, and to then progressively add complexities to it. A famous example of this is the Ising model (Kersen 1987, Tsvelik 1995), which tries to describe the magnetization in materials.

It is in this spirit of simplicity that led to the creation of a model known as the *minority game*. This model is aimed at having a simple yet rich platform to examine various phenomena arising from financial markets.

The minority game was foreshadowed by the El Farol Bar problem, which was invented by Arthur (1994). The El Farol Bar is an Irish bar in Sante Fe that used to have live Irish music on Thursdays. The problem was that there were usually more people that wanted to go than available seats in the bar. The problem written by Arthur (1994) goes as follows: "there are 100 Irish music lovers but El Farol has only 60 seats. The show is enjoyable when fewer than 60 people show up. What should we do?"

Although this problem appears to be a very simple one, it is commonly known that such everyday life situations pose quite a complex problem for deductive rational agents. The basic idea is to construct a model where the agents resort to "inductive thinking" with a limited number of strategies. In Arthur (1994), we can see that the El Farol bar problem has been successful in demonstrating that inductive reasoning can indeed be quite effective, however a statistical physicist would request the El Farol Bar

problem be explained in much more detail. This of course requires a more precisely defined model. The minority game came about as a way of modelling the dynamics without such fine grained details.

## 7.2 The Minority Game

The main difference between the El Farol bar problem and the minority game, is that El Farol emphasises about the inductive reasoning path to equilibrium while the minority game is mainly concerned with fluctuations around the equilibrium.

In the El Farol problem, the difficulty lies in the definition of an agent's strategies as well as to give a rule that predicts the future attendance given information about the past attendance. Each agent can have more than one strategy, which he ranks according to performance.

Now let us consider the El Farol problem. Supposing that there are $N$ agents and each agent bases his decision on the last $M$ attendances, then the attendance can take $(N+1)$ values each time. This makes $(N+1)^M$ possible combinations of information about the past. If the strategies are based on predictions of the attendance, given the past history as in Arthur *et al.* (1997b), we then have $N+1$ possible predictions for each combination of information and hence $(N+1)^{(N+1)^M}$ possible strategies. One can therefore see that searching through this large set for the best strategy may be a difficult task.

The first step in order to simplify the problem is carried out by observing that the players in the El Farol problem are only interested in going or not going to the bar. They therefore do not have to predict the precise attendance size, but instead only whether it is worthwhile going or not. As a result the number of strategies are then reduced to a much smaller number, that is $2^{(N+1)^M}$. However this is still a very large number depending on $N$. Another question is why would the agents consider the precise attendance size in the past in order to make a binary prediction? This seems to be unnecessary. Information encoding only the past $M$ correct choices should be enough. Doing so reduces the number of available strategies to $2^{2^M}$, which removes the undesirable dependence on $N$.

Now the number of strategies only depends on the number of past steps, moreover we see that if we symmetrize the problem by assuming that the bar can contain half of the players we obtain a model that focuses on the fluctuations in the attendance.

At this point we have a minority game as it was originally defined by Chalet and Zhang (1997). The general idea of the minority game is as follows: at any given time some people have two choices, they make their decisions simultaneously without any kind of communications between them, and those who happen to be in the minority win. In this context it is not in the interest of any agent to behave in the same way as the rest of the agents.

### 7.2.1 The Model

The dynamics of the Minority Game (MG) [20] are defined in terms of the dynamical variables $U_{s,i}(t)$ in discrete time $t \in \mathcal{N}^+$. These are the scores that each agent $i = \{1, .., N\}$ attaches to each other of his possible choices $s = \{1, .., S\}$. Each agent makes a decision $s_i(t)$ with probability

$$\text{Prob}\{s_i(t) = s\} = \frac{\exp[\Gamma_i U_{s,i}(t)]}{\sum_{s'} \exp[\Gamma_i U_{s',i}(t)]} \tag{7.1}$$

where $\Gamma_i > 0$ appears as an "individual inverse temperature". The original MG corresponds to $\Gamma_i = \infty$ (Chalet and Zhang 1997) and was later generalized to $\Gamma_i \equiv \Gamma < \infty$ (Cavagna *et al.* 1999).

The public information variable $\mu(t)$ is given to all agents, it belongs to the set of integers $(1, .., P)$ and can either be the binary encoding of the last $M$ winning choices (Chalet and Zhang 1997) or drawn randomly from a uniform distribution (Cavagna 1999).

The action $a^{\mu(t)}_{s_i(t),i}$ of each agent depends on choices $s_i(t)$ and on $\mu(t)$. The coefficients $a^{\mu}_{s_i,i}$, which are either +1 or -1, are called strategies and play the role of quenched disorder. These are randomly drawn with probability of a $1/2$ for each $i, s$ and $\mu$. They can also be thought of as agents buying (when +1) or selling (when -1) an asset.

On the basis of the outcome

$$A(t) = \sum_{i=1}^{N} a^{\mu(t)}_{s_i(t),i} \tag{7.2}$$

each agent updates his scores according to

$$U_{s,i}(t+1) = U_{s,i}(t) - a^{\mu(t)}_{s_i(t),i} \frac{A(t)}{P}, \tag{7.3}$$

[20] In this thesis we will consider the minority game defined by Chalet and Zhang (1997). When we refer to the Minority Game written in capital letters we are specifically talking about the dynamics defined in the Chalet and Zhang (1997) model.

where $P = 2^M$ is the total number of predictions. The idea of this equation is that agents reward $[U_{s,i}(t+1) > U_{s,i}(t)]$ those strategies that would have predicted the minority sign, i.e. $A(t)/|A(t)|$.

Similar results may be obtained when one considers the case when there is a nonlinear dependence on $A(t)$ i.e. with the dynamics

$$U_{s,i}(t+1) = U_{s,i}(t) - a^{\mu(t)}_{s_i(t),i} \mathrm{sgn}\left[\mathrm{A(t)}\right], \tag{7.4}$$

where the sgn function is the sign function also known as the step function, and is defined as

$$\mathrm{sgn(A(t))} = \begin{cases} +1 & \text{if } A(t) > 0 \\ -1 & \text{if } A(t) < 0. \\ 0 & \text{Otherwise.} \end{cases} \tag{7.5}$$

This leads to qualitatively similar results. A more lengthy discussion may be found elsewhere (Marsili *et al.* 2000, Zhang 1998, Savit *et al.* 1999, Chalet *et al.* 2001).

The source of randomness is in the choice of $\mu(t)$ and by $s_i(t)$. These are fast fluctuating degrees of freedom. As a consequence $U_{s,i}(t)$ is also fast fluctuating and hence the probability with which the agents choose $s_i(t)$ are subject to stochastic fluctuations.

The key parameters is the ratio $\alpha = P/N$ and the two relevant quantities are

$$\sigma^2 = \left\langle A^2(t) \right\rangle \text{ and, } H = \frac{1}{P} \sum_{\mu=1}^{P} \langle A|\ \mu\rangle^2, \tag{7.6}$$

which measure respectively, the fluctuations of attendance $A(t)$, i.e. the smaller $\sigma^2$ is, the larger a typical minority group is—in other words $\sigma^2$ is a reciprocal of the global efficiency of the system and the predictability; here $\langle ... \rangle$ denotes the temporal average over time.

One of the striking properties of this model is the fact that agents cooperate measured by $\sigma^2$. Agents taking random decisions would produce fluctuations equal to $N$ so that agents cooperate if they manage to produce fluctuations lower than $N$. In Fig. 7.1 we show the graph of the global efficiency $\sigma^2/N$ and the predictibility [21] $H/N$ versus the critical parameter $\alpha = 2^M/N$ for a sequence of number of agents varying from 1 to

[21] In this work we follow the same terminology and meaning as in (Chalet and Zhang 1997, Chalet *et al.* 2001, Marsili *et al.* 2000). The term predictibility and global efficiency means the same as in those references and are used right through this chapter.

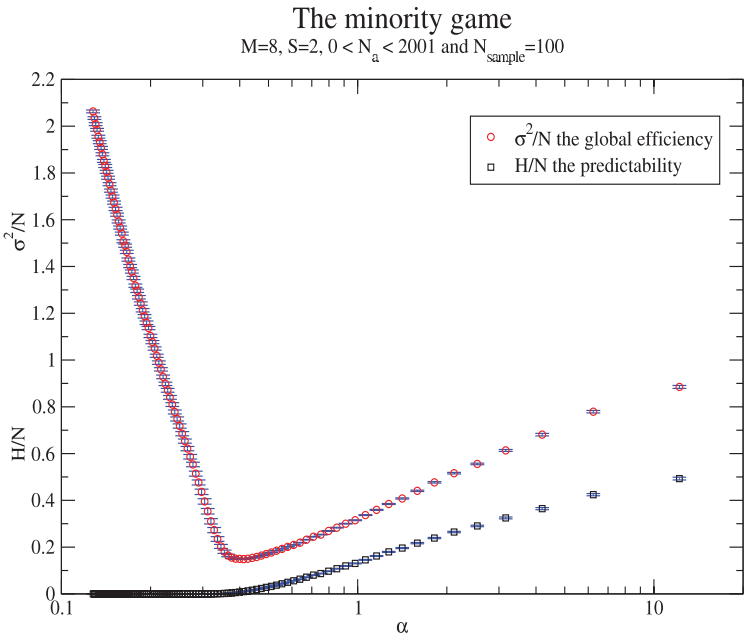


**Figure 7.1. The Global efficiency $\sigma^2/N$ and the predictability $H/N$ versus the critical parameter $\alpha = 2^M/N$.** The global efficiency $\sigma^2/N$ and the predictibility $H/N$ versus the critical parameter $\alpha = 2^M/N$ for a sequence of number of agents varying from 1 to 2001 when $M = 8$ and $S = 2$, in each simulation with $(N)_i$ number of agents it has been ensemble averaged over 100 samples ($N_{\text{sample}} = 100$). In this graph we can clearly see the three different regions, the first one fluctuations rapidly increases beyond the random agents and the game enters what has been called crowded region. At intermediate $\alpha$ the agent are at best coordination with each other, and finally at large $\alpha$ the game is more or less in a random mode.

2001 when $M = 8$ and $S = 2$, in each simulation with $(N)_i$ number of agents and with ensemble averaging over 100 samples ($N_{\text{sample}} = 100$).

In Fig 7.2 we show the graph of the same quantities but this time plotted for two different values of $S$, that is, $S = 2, 4$ and 6. This time the graph is a log–log plot so that we can get a good view of the behaviour of both the global efficiency and the predictability as $S$ varies. In Fig 7.3 and 7.4 the global efficiency and the predictability are graphed respectively.

It was initially pointed out (Chalet and Zhang 1997) that one could observe three different regions in this graph. The first one is found when $\alpha$ is small. In that case there is a large number of agents. In that region fluctuations rapidly increase beyond the level

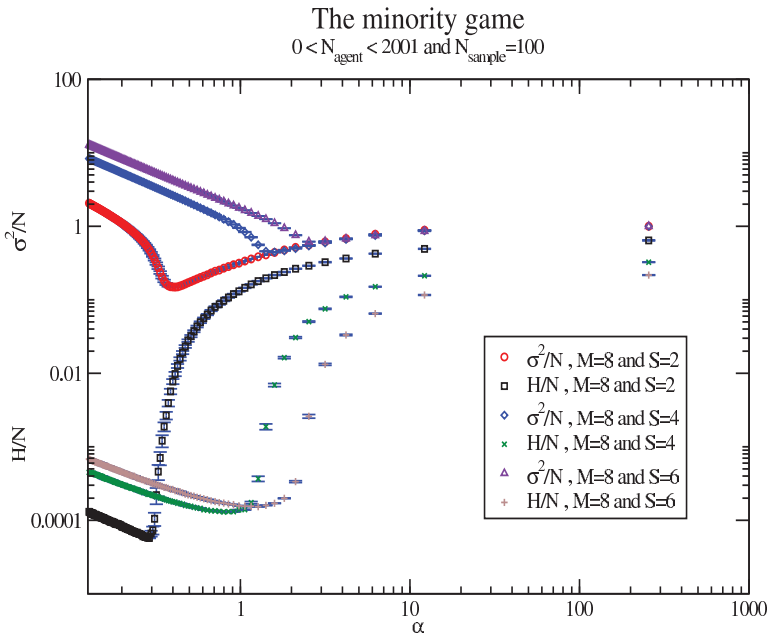


**Figure 7.2. The global efficiency $\sigma^2/N$ and the predictibility $H/N$ versus the critical parameter $\alpha = 2^M/N$.** The global efficiency $\sigma^2/N$ and the predictibility $H/N$ versus the critical parameter $\alpha = 2^M/N$ for a sequence of number of agents varying from 1 to 2001 when $M = 8$ and $S = 2, 4$ and 6, in each simulation with $(N)_i$ number of agents it has been ensemble averaged over 100 samples ($N_{\text{sample}} = 100$). This is the same graph as Fig. 7.1 but this time more different scenarios.

of random agents and the game enters what has been called crowded region since it is reached by keeping $M$ constant and $N$ increasing. In other words the agents display a herding behaviour and produce non–Gaussian fluctuations $\sigma^2 \sim N^2$ (Chalet and Zhang 1997, Chalet *et al.* 2001, Marsili *et al.* 2000).

At intermediate $\alpha$, as $N$ decreases that is, when the game enters into a regime where agents coorporate to reduce fluctuations. In other words, that is when maximal coorperation is achieved.

Now if we go to the region where $\alpha$ is large, which means that $N$ is small, then the outcome is more or less random. That is coorporation slowly disappears and the variance of the outcome tends to the value that would be produced by agents making random decisions. The reason for this is that the information, which agents receive about the past history, is too complex and their behaviour over–fits the fluctuations of past attendance.

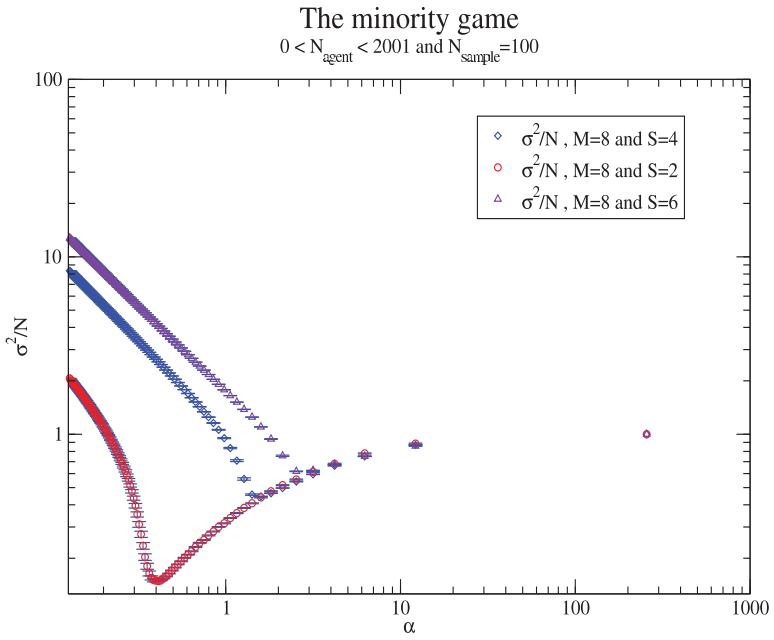


**Figure 7.3. The global efficiency $\sigma^2/N$ versus the critical parameter $\alpha = 2^M/N$ for the Minority Game.** The global efficiency $\sigma^2/N$ versus the critical parameter $\alpha = 2^M/N$ for a sequence of number of agents varying from 1 to 2001 when $M = 8$ and $S = 2, 4$ and 6, in each simulation with $(N)_i$ number of agents it has been ensemble averaged over 100 samples ($N_{\text{sample}} = 100$) for the Minority Game.

When $S$ is varied the crowded region moves to the right, whereas $\sigma^2/N$ for $N \ll 2^M$ seems to collapse on roughly the same curve. The measure of $\sigma^2/N$ is less and less pronounced when $S$ is larger, as shown in Fig. 7.3.

Also shown in Fig. 7.1, 7.2 and Fig. 7.4 is the predictability, which is another quantity of interest in the Minority Game. The predictability is a major issue in finance. It is commonly believed that markets are not efficient markets, violating the Efficient Market Hypothesis (EMH) (Fama 1965). Even in their weakest form (that is all public information on past prices and volumes affects the current price at every time), empirical studies (Zhang 1999, Brock *et al.* 1992) show that there are systematic correlations in most financial markets.

In the case of the Minority Game there are different pieces of information such as the histories, which are common pieces of public information encoding the previous $M$ last minority choices. Another aspect is the memory of the game in Eq. (7.3) with a

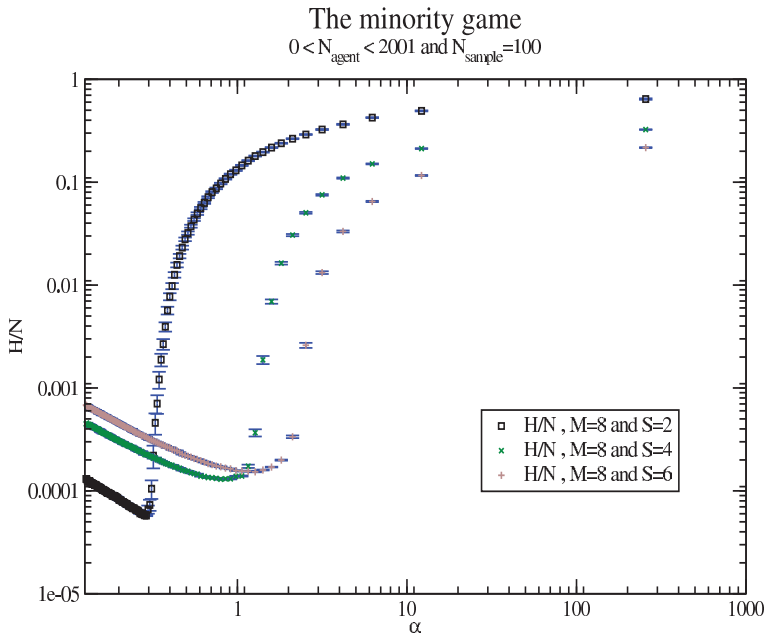


**Figure 7.4. The predictibility $H/N$ versus the critical parameter $\alpha = 2^M/N$.** The predictibility $H/N$ versus the critical parameter $\alpha = 2^M/N$, for a sequence of number of agents varying from 1 to 2001 when $M = 8$ and $S = 2, 4$ and 6, in each simulation with $(N)_i$, agents has been ensemble averaged over 100 samples ($N_{\text{sample}} = 100$).

given payoff function, in the case of the Minority Game it is given by

$$g_i(t) = -a^{\mu(t)}_{s_i(t),i} A(t), \tag{7.7}$$

with $A(t)$ defined as in Eq. (7.2) for the agents.

The scores, given by Eq. (7.3), contain information about the game. The normalized predictability in the Minority Game is calculated from $A(t)$, that is,

$$H = \frac{1}{2^M} \sum_{i=1}^{P} \langle A(t) |\, \mu(t) \rangle^2 . \tag{7.8}$$

At the point where $H$ starts to differ from 0 (at around $\alpha_c \approx 0.34$ for $S = 2$) and starts to increase the system becomes predictable. In statistical physics this is commonly known as a phase transition with symmetry breaking as $\alpha$ varies. For $S = 2$, where $\alpha_c \approx 0.34$ when $\alpha > \alpha_c$, we then have an asymmetric phase. This is when the outcome becomes probabilistically predictable. The computer code which implements the Minority Game can be found in Appendix E.6.2, will all the relevant routines therein. The

code contains various switches for selecting the game and the scenarios whether the users decides to use real data or just simulated games.

In Fig. 7.1 and 7.4 we see a graph of the predictability for $S = 2$, and $M = 8$, and for the number of agents varying from 1 to 2001.

### 7.2.2 The Price Function in the Minority Game

To connect the Minority Game with financial market, one needs to examine the price dynamics. Here we shall focus on a market for a single asset and call $P(t)$, its price function at a time $t$. Let us assume that the price is driven by the difference between the number of shares being bought and sold, called the *excess demand*. This is how the connection has been made with the Minority Game. In the Minority Game we assume that the behaviour of agents is restricted to the two possible actions, that is buy (i.e. $a_i(t) = 1$) and sell (i.e. $a_i(t) = -1$). The $A(t) = \sum_i a_i(t)$, Eq. (7.2), is simply the difference between demand and supply, i.e. the excess demand.

Several price formulation rules can be found in the literature, which link the excess demand $A(t)$ to the price return. The simplest one is to suppose that the price return $r(t)$ depends linearly on $A(t)$ (Farmer 1998),

$$r(t) = \ln\left(\frac{P(t)}{P(t-1)}\right) = \frac{A(t)}{\lambda}, \tag{7.9}$$

where $\lambda$ is sometimes called the liquidity or the market depth (Bouchaud and Cont 1998). This relationship is implicit in many early works, which refer to $\sigma^2$ as price volatility, but a plot of $\ln(P(t)) = \sum_{t' \leq t}^{t} A(t'/\lambda)$ was not shown until the paper by Johnson *et al.* (2000). Eq. (7.9) can be justified in limit order markets, that is markets where people can submit limit orders (Chalet and Zhang 1997, Chalet *et al.* 2001, Marsili *et al.* 2000), which are requests to buy or sell a given quantity of the asset at a given price. Each of these orders can only be executed if there is an opposite matching request. In this way, the quantity and the price of the transaction are fixed and the time when the limit order will be executed is left undetermined. Orders waiting to be executed are stored in the order book. In Fig 7.5 we show the price function for two different values of the liquidity $\lambda = N = 21$ and 41 for two different samples for $t$ up to five hundred ticks. Now supposing that at time $t - \epsilon$, $0 < \epsilon \ll 1$, $N$ market orders of size 1 arrive simultaneously on the market. Assuming that $(N + A)/2$ are buy orders and $(N - A)/2$ are sell orders, it is then possible to match $(N - |A|)/2$ buy and sell orders and to execute them at the current price. This leaves unexecuted $|A|$ orders of one kind. If $A > 0$

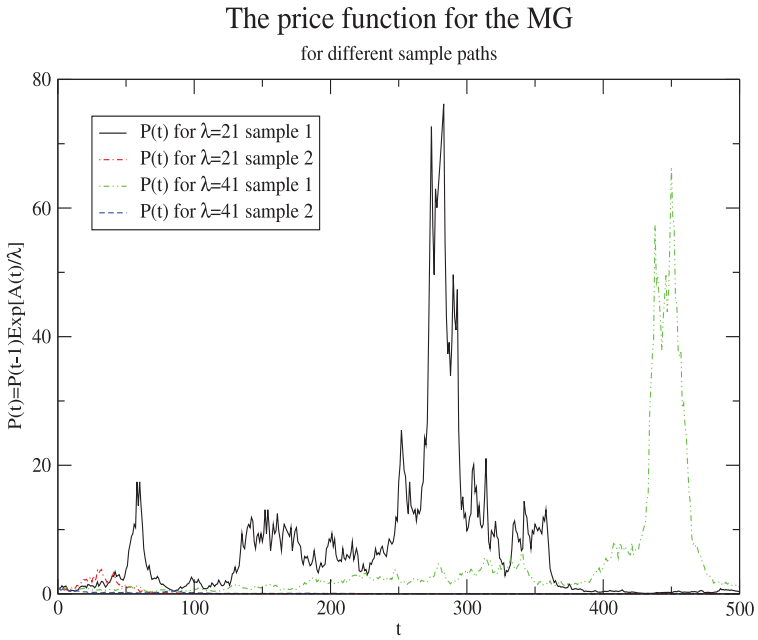


**Figure 7.5. The price function $P(t)$, Eq. (7.9).** The price function $P(t)$, Eq. (7.9), for two different liquidity values $\lambda = N = 21$ and 41 for two different samples for $t$ up to five hundred ticks. This is for simulated data within the Minority Game.

they will be buy orders, else sell orders. There orders will be matched with the best limit orders of the opposite type present in the order book.

Now assuming that there is a uniform density $\lambda$ of limit orders, that is $\lambda$ orders per tick (ticks are evenly spaced), the price will be displaced by a quantity $A/\lambda$, as all the orders between $P(t-1)$ and $P(t) \equiv P(t-1) + A/\lambda$ will be executed. This is what Eq. (7.9) postulates. This process can go on assuming that there are new limit orders that fill the gap between $P(t-1)$ and $P(t)$, restoring a uniform distribution of limit orders. Bouchaud and Potters (2000b) have shown that the assumption of uniform order density of the order book, which is responsible for the linear relationship between $A$ and $r$ is a very rough approximation.

An alternative definition for the price function, under the same assumption as the Minority Game for each agent, is specified as follows. Supposing that $a_i(t) = +1$ means that agent $i$ invests \$1 in order to buy the asset at time $t$, whereas $a_i(t) = -1$ means that he/she sells $1/P(t-1)$ units of assets, where $P(t-1)$ is the price of the last transaction. Then the total demand is $(N+A)/2$ and the total supply is $(N-A(t))/2P(t-1)$

units of asset where $A(t) = \sum_i a_i(t)$. Then the price $P(t)$ is fixed in such a way that the demand matches the supply, that is

$$P(t) = P(t-1)\frac{N + A(t)}{N - A(t)}. \tag{7.10}$$

If $A(t) \ll N$, taking the logarithm of both sides and keeping the leading order terms leads to an expression that is very similar to Eq. (7.9) with $\lambda = N/2$.

Using these two definitions we compare the price time series in the Minority Game. This is shown in Fig. 7.6. The graph of the price function in the Minority Game for the

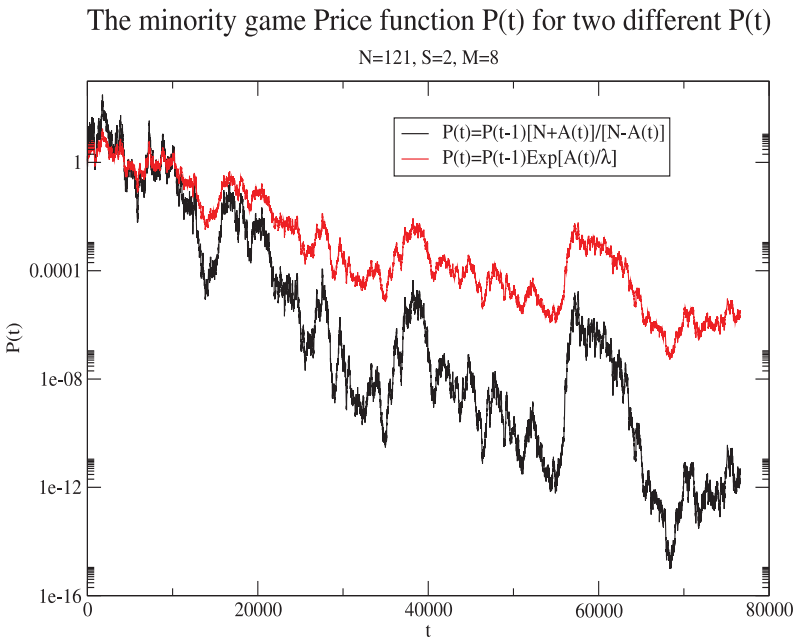


**Figure 7.6. The simulated price function in the Minority Game.** The simulated price function in the Minority Game for the two definition of $P(t)$ given by Eq. (7.9) and Eq. (7.10) for the full Minority Game time ticks. Here $S = 2$, $N = 121 \equiv \lambda$ and $M = 8$. Ignoring the scaling issue we can see that the price function remains stable for large value of time.

two definition of $P(t)$ given by Eq. (7.9) and Eq. (7.10) for the full range of Minority Game time ticks. Here $S = 2$, $N = 121 \equiv \lambda$ and $M = 8$, for the full time time series in the Minority Game when each of the 121 agents have 2 strategies and when the memory is of the order of 8, i.e. $M = 8$ and in Fig. 7.7 for the first 2500 time ticks of the time series. In these two figures the number of agents is $N = 121$ and was used for the liquidity $\lambda$.

Now looking at the two trajectories we can see that Eq. (7.10) gives a higher estimate that Eq. (7.9) while giving very similar trajectories. The graph of the price function in the minority game for the two definition of $P(t)$ given by Eq. (7.9) and Eq. (7.10) for the first 2500 time ticks. Here $S = 2$, $N = 121 \equiv \lambda$ and $M = 8$. These two definitions

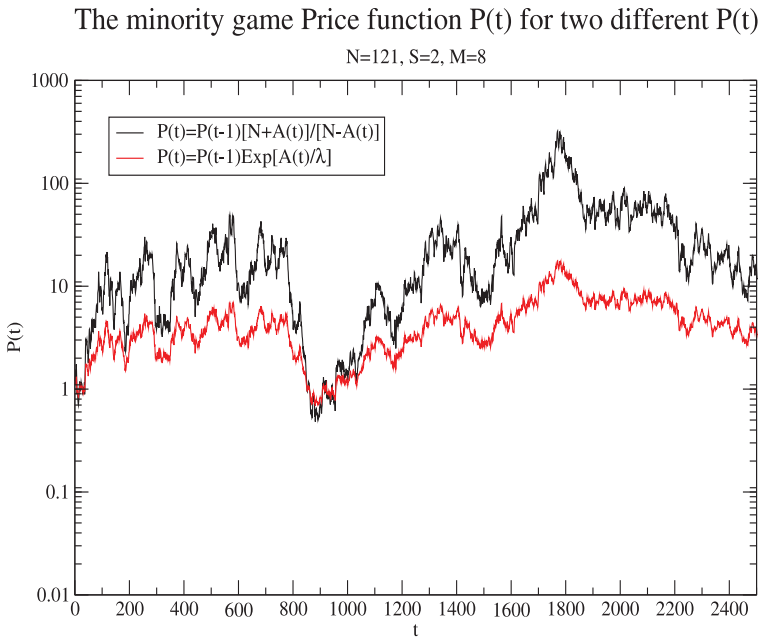


**Figure 7.7. The price function in the Minority Game.** The price function in the Minority Game for the two definition of $P(t)$ given by Eq. (7.9) and Eq. (7.10) for the first 2500 time ticks. Here $S = 2$, $N = 121 \equiv \lambda$ and $M = 8$. This is the same as Fig. 7.6 but this time on a smaller range.

may be compared better when real data is used in the Minority Game with a different payoff, see Section 7.2.3 for later discussion.

As previously mentioned, Fig. 7.5 shows the price time series evolution for Eq. (7.9) for the first 500 time ticks when $\lambda = 21$ and 41, $S = 2$ and $M = 8$ in the Minority Game for 2 different initial configurations or samples. In this figure we can see some sharp peaks. Here at this level the time series either diverges to infinity or converges to 0. This is because we have taken $\lambda = N$ when $\lambda$ should not be taken as a constant and there is also a time scale associated with it. Furthermore, if $\lambda$ is taken as the market depth it is commonly accepted that the market depth is also a time series and varies in time thus taking $\lambda$ to be a constant is partially incorrect.

Finally to make sure that we have not any bias in the time series we have repeated the experiment a number of times and plotted Eq. (7.9) for many different configurations, this is shown in Fig. 7.8. From this graph we can clearly see that each paths are clearly distinct and it therefore shows no bias overall.

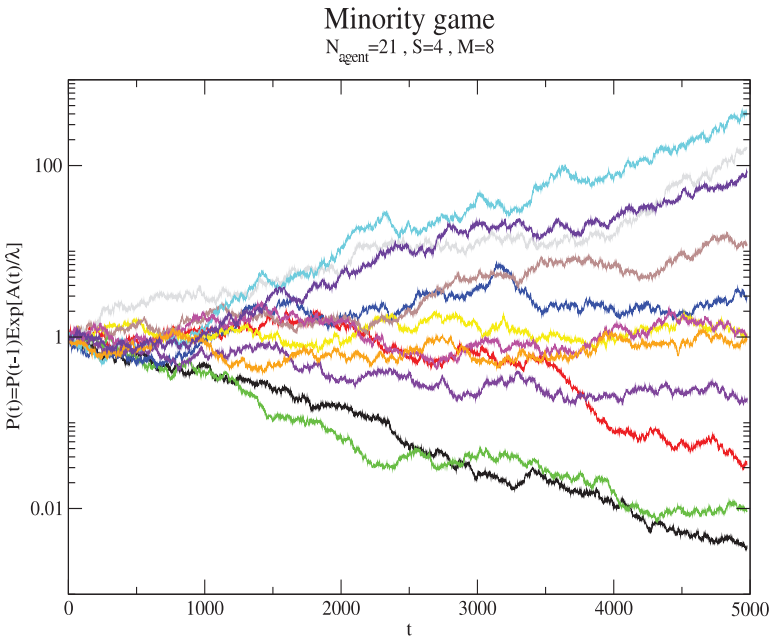


**Figure 7.8. The simulated price function in the Minority Game for different sample paths.** The price function in the Minority Game for $P(t)$ given by Eq. (7.9) for the first 5000 time ticks. Here $S = 2$, $\lambda \propto N = 21$ and $M = 8$ taken over many different initial configuration on a smaller range of time. From this graph we can see how the different paths are not biased.

### 7.2.3 The Dollar Game

We now consider the $–Game, and point the small difference between the Minority Game and the $–Game.

The Minority Game is a repeated game where $N$ agents, have to choose one out two possible alternatives at each step. Each agent, $i$, has a memory of the past. At each time step $t$ every agent decide whether to buy or sell an asset. The agent takes an action $a_i(t) = \pm 1$ where 1 is when buying an asset as opposed to -1 when selling. The

Excess demand $A(t)$ at time $t$ is then given by Eq. (7.2), that is $A(t) = \sum_{i=1}^{N} a_{s_i(t),i}^{\mu(t)}$. The payoff of agent $i$ in the Minority Game is given by Eq. (7.7).

In order to model financial markets, some authors have used the following definition for the return $r(t)$ using the price time series $P(t)$ (Bouchaud and Cont 1998, Farmer 1998)

$$r(t) \equiv \ln[P(t)] - \ln[P(t-1)] = \frac{A(t)}{\lambda}, \tag{7.11}$$

which means that price time series is defined by

$$P(t) = P(t-1) \exp\left[\frac{A(t)}{\lambda}\right]. \tag{7.12}$$

Here the liquidity $\lambda$ is proportional to the number of agents $N$. In the Minority Game the agents predicts the price movements only over the next time step. However, Andersen and Sornette (2003) have shown that in order to know when the price reaches its next local extreme-mum and optimize their gain the agents need to predict the price movement over the next two time steps ahead ($t$ and $t+1$) and they therefore have postulated the correct payoff function to be given by

$$g_i^{\$}(t+1) = a_i(t)A(t+1). \tag{7.13}$$

This small difference in the payoff function is what defines the \$–Game. From now on when we are refering to the \$–Game, we really mean the Minority Game with the payoff function defined by Eq. (7.13). In this case we define the game as the \$–Game.

## 7.3 Financial Bubbles

It is a well established fact of economics that markets follow both positive and/or negative trends, crashes and bubble effects. In general a strong positive trend is followed by a crash, famous examples of these effects were seen in the recent crash on the NASDAQ (April 2000) and prior to the crash in the Hong Kong market, which was associated with the Asian crisis in the early 1994.

A strong positive trend in economics is commonly called a *bubble*. Bubbles can occur in all sorts of different sectors for example in the technology sector, resources sector, housing sector, the music industry or the pharmaceutical sector. So a bubble is really when investors follow the same trend or strategies for a given time (e.g. buying or

selling) for a while until the demand decreases, which may sometime be due sometime to economic slowdown or change of perspectives in economical strategies. At that time the trend usually takes an opposite direction (either a positive trend corresponding to buying then once the bubble has matured everyone starts selling, or vice versa).

A common approach to viewing the market is carried out by assuming that these are complex evolutionary systems that are adaptive and that they are populated by rational agents interacting with each other. These sorts of models are researched at the Santa Fe Institute in New Mexico (le Baron *et al.* 1999, Farmer 1998) as well as other institutions worldwide (Rachlevsky-Reich *et al.* 1999, Hommes 2001, le Baron 2000, Chalet and Zhang 1997).

One of the main problems in most of the models is that they do not capture the characteristic structure of bubbles. However if such effects are actually present in markets (which is commonly accepted that they are) they probably constitute one of the most important facts in explaining and predicting market behaviour with their associated consequences such as large potential losses during crashes and recession following these bubbles.

Since the earlier works on Rational Expectation (RE) bubbles (Farmer 1998, Rachlevsky-Reich *et al.* 1999) the size of the literature on the subject has been growing with theoretical improvements of the original concept and on the empirical detectability of RE bubbles in financial data (Camerer 1989, Adam and Szafarz 1992). At the same time, empirical research has largely concentrated on testing for explosive exponential trends in the time series of asset price and foreign rates (Evans 1991, Woo 1987).

Many RE bubbles produce curves that are not always consistent with economic facts, a major problem is that the appearance of bubbles can be reinterpreted in terms of market fundamentals that are not observed by the researcher. Another suggestion is that if stock prices are not more explosive than dividends then it can be concluded that rational bubbles are not present, since bubbles are taken to generate an explosive component of stock prices (Sornette and Andersen 2002). However periodically collapsing bubbles are not detectable by using standard tests to determine whether stock prices are more explosive or less stationary than dividends (Evans 1991). So in short, the present evidence for an ability to speculate on bubbles remains an unsolved problem.

### 7.3.1 Positive Feedback Model With Multiplicative Noise

In this section, a model to generate the bubble price $B(t)$ is described. This model has been developed by Sornette and Andersen (2002) and has been used in previous studies (Sornette and Andersen 2002, Andersen and Sornette 2004). Here we use the same notation and interpretation as in these references. Readers interested in how the model is derived may see these last two references for further details.

The bubble price model is an hyperbolic stochastic finite–time singularity formula, which transforms a Wiener process into a time series containing no correlation of returns (Campbell *et al.* 1997) long range correlation of volatility (Ding *et al.* 1993), fat–tail of returns distribution (Mandlebrot 1963, Vries 1994, Mantegna and Stanley 1995), apparent multifractality (Mandlebrot 1997, Muzy *et al.* 2001), sharp peak through flat pattern of price peaks (Roehner and Sornette 1998), as well as accelerated speculative bubbles preceding crashes (Johansen *et al.* 1999).

One of the key aspects of this model is that bubbles are growing superexponentially, that is, self growing in time, this leads to power law acceleration, which eventually leads to a singularity as opposed to other bubble models, which are based on exponential growth.

The formulation of the bubble price $B(t)$ is initially constructed from the Black–Scholes–Merton option pricing model (Merton 1990), $dB(t) = \mu B(t)dt + \sigma B(t)dW_t$ with $\mu$ the instantaneous return rate and $\sigma$ the volatility. The Gaussian noise $W_t$ is the standard Wiener process.

The bubble model is generalized as

$$dB(t) = \mu(B(t))B(t)dt + \sigma(B(t))B(t)dW_t - \kappa(t)B(t)dj, \tag{7.14}$$

where $B(t)$ is the price of the bubble, $\mu$ the abnormal return rate above the fundamental return, $\sigma$ is the volatility of the bubble and the jump term $dj$ describes a correction or a crash that may occur with amplitude $\kappa$. The crash amplitude can be a stochastic variable taken from an arbitrary distribution.

Straight after the last crash which becomes the new origin of time 0, $dj$ is reset to 0 and will eventually jump to 1 with hazard rate $h(t)$ with probability $P(t < t_{\text{crash}} < t + dt) = h(t)dt$. The discontinuous process for the jump $dj$ is defined as in Cox *et al.* (Cox *et al.* 1979, Merton 1976a) to define the jump as a discontinuous process. Its

average $\langle dj \rangle$ is given by

$$\langle dj \rangle = 1 \times h(t)dt + 0 \times (1 - h(t)dt) = h(t)dt. \tag{7.15}$$

Using this definition of the hazard rate $h(t)dt = \langle dj \rangle$ we can see for Eq. (7.14) that the expectation, over all possible outcomes since the last crash, leads to

$$\mu(B(t))B(t) - \langle \kappa \rangle B(t)h(t) = 0, \tag{7.16}$$

which gives an expression for the hazard rate

$$h(t) = \frac{\mu(B(t))}{\langle \kappa \rangle}. \tag{7.17}$$

It is possible to generalize Eq. (7.14) by allowing some non–linearity in $\mu(B(t))$ and $\sigma(B(t))$, as shown in Sornette and Andersen (2002) and in Andersen and Sornette (2004):

$$\mu(B(t))B(t) = \frac{m}{2B(t)} \left[B(t)\sigma(B(t))\right]^2 + \mu_0 \left[\frac{B(t)}{B_0}\right]^m, \tag{7.18}$$

$$\sigma(B(t))B(t) = \sigma_0 \left[\frac{B(t)}{B_0}\right]^m. \tag{7.19}$$

Here $B_0, \mu_0, m > 0$, and $\sigma_0$, are respectively four parameters of the model that are a reference scale, an effective drift, the strength of non–linearity and the magnitude of stochastic component which sets the scale of the volatility (i.e. the non–linear positive feedback). The first term in Eq. (7.18) was added for convenience to simplify the Îto calculation of the stochastic differential equation.

Herding is perhaps the most obvious reason that leads to positive non–linear feedback of $\mu(B(t))$ and $\sigma(B(t))B(t)$ on stock prices.

The solution of Eq. (7.14) with Eq. (7.18) and Eq. (7.19) is derived in references such as Sornette and Andersen (2002) and in Andersen and Sornette (2004) and is given by

$$B(t) = \alpha^\alpha \frac{1}{\left(\mu_0 \left[t - t_c\right] - \frac{\sigma_0}{B_0^m} W(t)\right)^\alpha}, \tag{7.20}$$

where $\alpha \equiv 1/m - 1$ and with $t_c = y_0/(m-1)\mu_0$. The critical time $t_c$ is a finite time singularity that is determined by initial conditions with $y_0 = 1/[B^{(m-1)}(t=0)]$, see Appendix in Sornette and Andersen (2002). In Fig. 7.9, the graph of the time series for the bubble defined in Eq. (7.20) versus the time $t$, $0 \leq t \leq 2500$ with fixed parameters $m = 3$, $\mu_0 = 0.01$, $B_0 = y_0 = 1$, $\delta t = 0.0003$ and the critical time $t_c = 1$ for two distinct

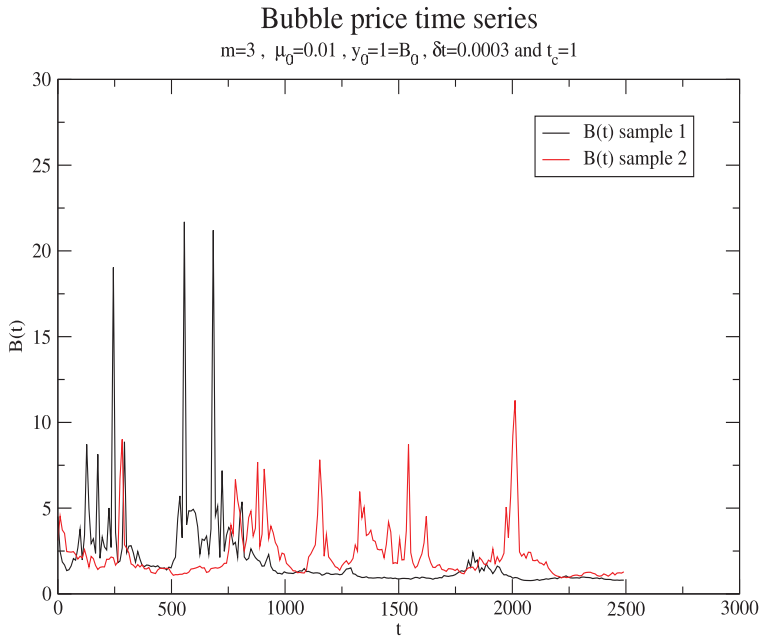


**Figure 7.9. The time series for the bubble defined in Eq. (7.20) versus the time $t$.** The time series for the bubble defined in Eq. (7.20) versus the time $t$, $0 \leq t \leq 2500$ with fixed parameters $m = 3$, $\mu_0 = 0.01$, $B_0 = y_0 = 1$, $\delta t = 0.0003$ and the critical time $t_c = 1$ for two distinct sample path of the Wiener process.

sample path of the Wiener process. That is the graph of Eq. (7.20) versus the time $t$, $0 \leq t \leq 2500$ with fixed parameters $m = 3$, $\mu_0 = 0.01$, $B_0 = y_0 = 1$, $\delta t = 0.0003$ and the critical time $t_c = 1$ for two distinct sample path is shown. In both cases the graphs show some very sharp but finite peaks after a certain time of normal activity.

Note that Eq. (7.20) is correct as long as a crash $dj = 1$ has not occurred, which may happen at any time according to the crash hazard rate $h(t)$, given by Eq. (7.17) determined from non–arbitrage conditions. Here $\langle \kappa \rangle$ is the average amplitude calculated over some pre–determined distribution of $\kappa$. In the deterministic case $\sigma_0 = 0$ reduces to $B(t) \propto 1/[t_c - t]^{1/m-1}$, that is the bubble follows a hyperbolic growth path which would diverge in finite time if not checked by crashes according to Eq. (7.17).

One must note that this hyperbolic growth is a sign of the positive feedback characterized by $m > 1$ of the price $B(t)$ on the return rate $\mu$.

On the other hand if $\sigma \neq 0$ we see that the crash hazard rate grows even further than the bubble price we then do not obtain a singularity. In the limit $1/\alpha \to 0$ ($m \to 1$ in

Eq. (7.20)

$$B(t) = \exp\left[\mu_0 t + \sigma_0 W(t)\right], \tag{7.21}$$

one recovers the standard Black–Scholes–Merton solution.

## 7.4 Minority Game and Dollar Game Price Function With Real Data

In this section, we combine the results from the previous sections to monitor the price function when real data is inserted into the Minority Game with the dollar game payoff. The idea is to see if the agent model does follow the real data trajectories.

From past historical data we can see where bubbles have occurred in the past, and use this information to see how an agent model—such as the Minority Game—will track the real data.

Here we will use the historical price time series of the NASDAQ over a period of about twenty years, that is from October 1984 to late September 2005, Fig. 7.10. Over this period we can clearly see the bubbles due to the technological sector from the mid eighties until the bubble burst in the early 2000. Large growth was then followed by a big crash, where billions of dollars were wiped out off the market.

The other set of data that will be considered will be from the S&P 500 from the late nineties to the present day, see Fig. 7.11, that is over the period of January 1998 to September 2005.

We now use this data to insert it into the Minority Game to see how the game behaves and evolves as a function of time $t$ with two different payoff, that two different dynamical processes. Here the payoff function is updated differently as in the standard Minority Game described in the earlier section. We introduce an extra parameter that looks over a certain time in the past, we call it $T$. It can be understood as a window parameter that can be attributed a certain length. In this setting we update the scores, defined in equation as in Eq. (7.3)

$$\Delta U_{s,i}(t) = \sum_{\kappa}^{t} a^{\mu(j)}_{s_i(t),i} \frac{A(j)}{P}, \quad \text{with } \kappa = \begin{cases} j = t - T + 1 & \text{if } t - T + 1 > 0 \\ j = 1 & \text{if } t - T + 1 \leq 0, \end{cases} \tag{7.22}$$

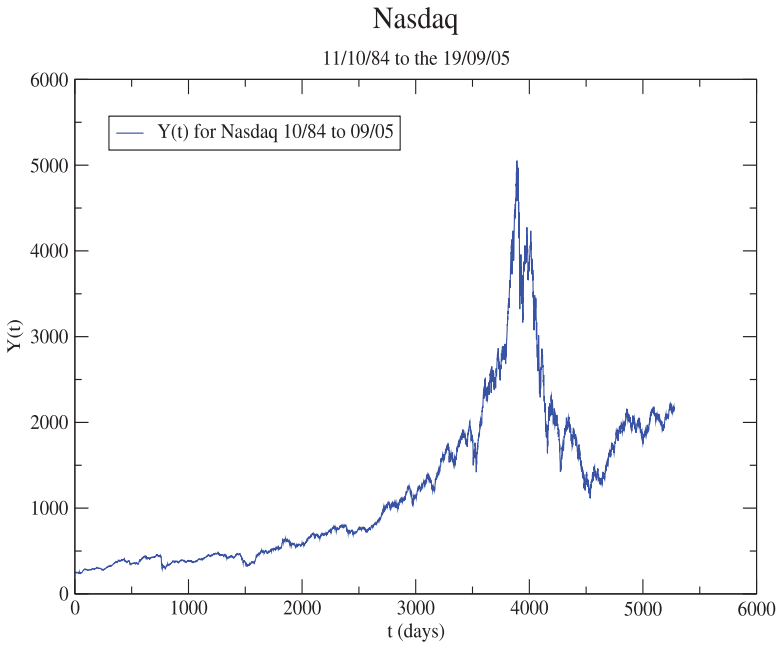


**Figure 7.10. Time series for the NASDAQ.** The graph of the NASDAQ versus the time $t$, $1 \leq t \leq 5283$ over the period of 11/09/84 to the 19/09/05, showing clearly the signs of a bubble over the time.

where $A(t)$ is from the Minority Game strategy selection as described in Sec. 7.2.1. The scores are then updated such as

$$U_{s,i}(t+1) = \Delta U_{s,i}(t). \tag{7.23}$$

The real data is inserted via the action $a^{\mu(t)}_{s_i(t),i}$. In the simulated case the action $a^{\mu(t)}_{s_i(t),i}$ is generated randomly and take the value of $+1$ or $-1$. Using the real data we can generate the evolution of the action functional as we evolve through the real data. Supposing that we denote the real data set by $S(t)$. We set $a^{\mu(t)}_{s_i(t),i} = 1$ when the value goes up, in other words when $S(t+1) > S(t)$ and set $a^{\mu(t)}_{s_i(t),i} = -1$ when the value goes down, that is when $S(t+1) < S(t)$. When the value stays unchanged, $S(t+1) = S(t)$, we flip a coin with equal probability. This is implemented in the code *real.f*90 in Appendix. E.6.4

Using this method we can compare the dynamics of both games. In Fig. 7.12, we show the graph of the price function as a time series for the \$–Game versus the Minority Game in the Minority Game as a function of time $t$ on a linear scale, $1 \leq t \leq 5283$. This is compared to the real data from the NASDAQ over the period of 11/09/84 to

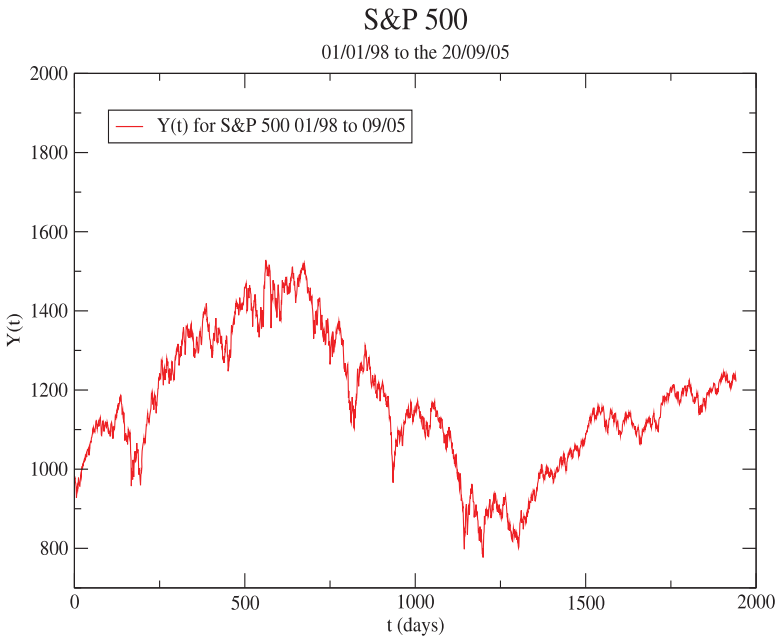


**Figure 7.11. Time series for the S&P 500.** The graph of the S&P 500 versus the time $t$, $1 \leq t \leq 1941$ over the period of 01/01/98 to the 20/09/05, showing clearly the signs of a bubble over the time.

the 19/09/05, showing clearly the signs of a bubble over the time. Here the number of agents $N = 41$ and each agent have $S = 2$ strategies to choose from with a memory of 8, $M = 8$ and with a window size of $T = 100$.

Ignoring the scale factor problem between the games and the real data—something that will need to be resolved later on—we can see that in Fig. 7.12 the \$–Game and the real data follow very similar trajectories as opposed to the Minority Game, which is not sensitive to the existence of a bubble. So in this figure we can see that the \$–Game is significantly more sensitive to the bubble showing clear evidences of lumps and troughs displayed in the real data. There is also clear evidence that there is a scaling problem. This comes from the fact that the liquidity is approximated to be $\lambda \sim N$. The liquidity is however, usually affected, as the market depth is. The market is not constant right through and should be taken as a time series. In Fig. 7.13 we show the graph of the price function as a time series for the \$–Game versus the Minority Game in the Minority Game as a function of time $t$ on a logarithmic scale, $1 \leq t \leq 5283$. This is compared to the real data from the NASDAQ over the period of 11/09/84 to the

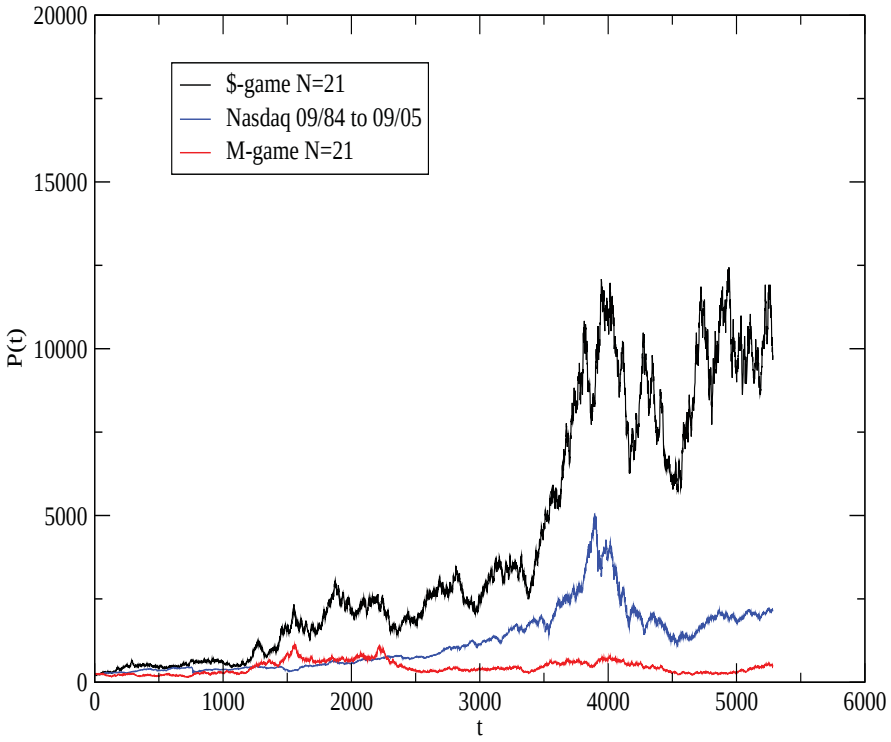


**Figure 7.12. Time series for the NASDAQ.** The graph of the NASDAQ versus the time $t$, $1 \leq t \leq 5283$ over the period of 11/09/84 to the 19/09/05, showing clearly the signs of a bubble over the time. The is compared with the plot of both the Minority Game and the $–Game. We can clearly see that the Minority Game does not sensitive to the existence of the bubble, but the $–Game does.

19/09/05. Here the number of agents $N = 41$ and each agent have $S = 4$ strategies to choose from with a memory of 4, $M = 4$. The window size is $T = 10$.

If we look at Fig. 7.13 when the number of strategies is set to $S = 4$ with a memory of $M = 4$, where we have set $\lambda$ to two different values, we can easily see the dependence of the liquidity over time therefore showing clearly the sign of non constant liquidity over the time series evolution. In this figure the black curve is when the liquidity $\lambda = 10N$ while the red curve is when $\lambda = 15N$. On the other hand one should note that increasing the factor in front of the liquidity does not always bring the curve closer to the real data, sometimes it is the opposite.

We can now compare the dynamics of both payoffs for a different number of agents $N$ and liquidity $\lambda$. In Fig. 7.14 we show on a logarithmic scale the graph of the price function as a time series for the $–Game versus the Minority Game in the Minority Game as a function of time $t$, $1 \leq t \leq 5283$. This is compared to the real data from the NASDAQ (the blue curve) over the period of 11/09/84 to the 19/09/05. Here the

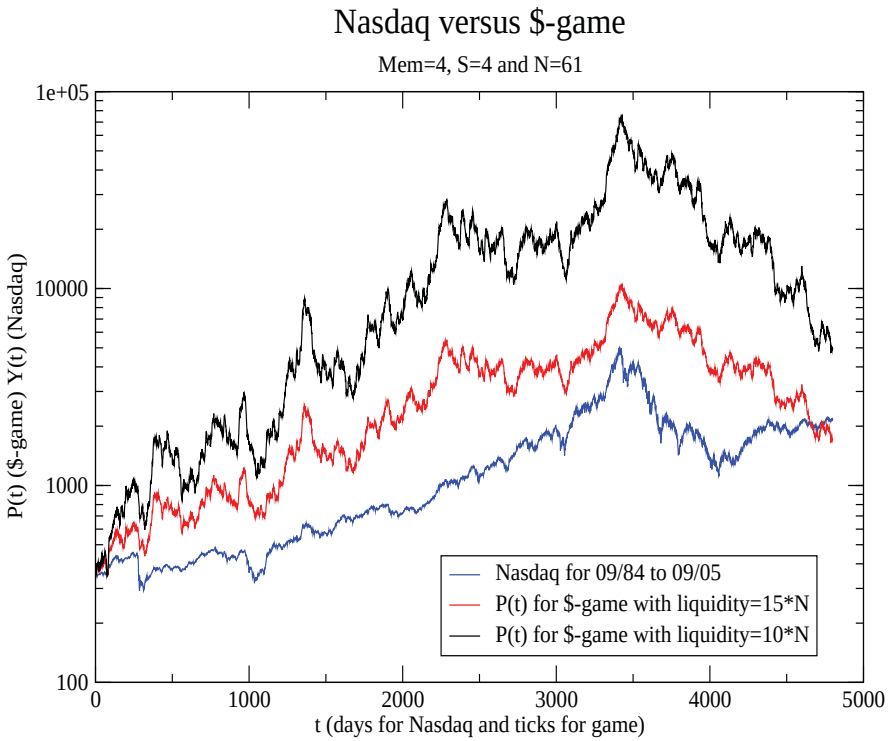


**Figure 7.13. Price function for the $–Game versus the Minority Game in the Minority Game as a function of time $t$.** The price function as a time series for the $–Game in the Minority Game on a logarithmic scale, $1 \leq t \leq 5283$. This is compared to the real data from the NASDAQ over the period of 11/09/84 to the 19/09/05, showing clearly the signs of a bubble over the time. Here the number of agents $N = 41$ and each agent has $S = 4$ strategies to choose from with a memory of 4, $M = 4$. The window size is $T = 10$.

number of agents $N = 21$, 41, and 61 in each games agents have $S = 4$ strategies to choose from with a memory of 4, $M = 4$. The window size is $T = 10$. In Fig. 7.14 we can see that in all cases the dynamics of the Minority Game with the Minority Game payoff does not quite follow those of the real data, contrarily to the $–Game.

As a final test we turn off the dynamics of both games by setting $\Delta U_{s,i}(t) = 0$ which means that the scores do not get updated, and seeing how the game performs on real data, namely on the NASDAQ and S&P500. This is shown in Fig. 7.15 and Fig. 7.16, where we can see that both curves follow each other quite well.

Finally comparing the outcome of both the $–Game and the minority payoff in the Minority Game when the payoff $\Delta U_{s,i}(t) = 0$ gives trajectories that overlap on top of another, see Fig. 7.17, as one would expect.

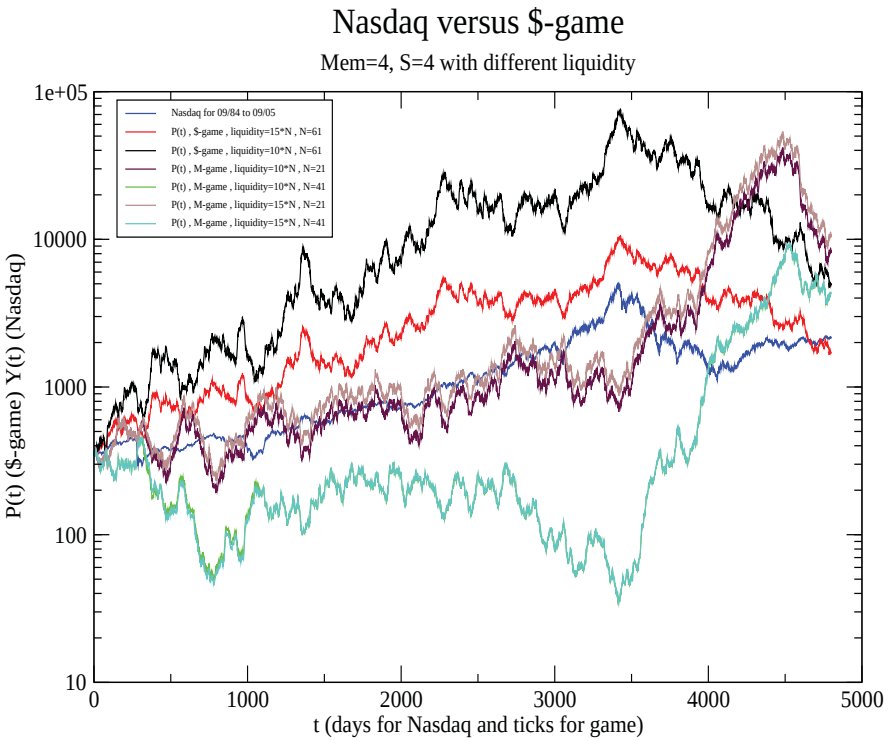


**Figure 7.14. Price function for the $–Game versus the Minority Game in the Minority Game as a function of time $t$.** The price function for the $–Game versus the Minority Game in the Minority Game as a function of time $t$ on a logarithmic scale, $1 \leq t \leq 5283$. This is compared to the real data from the NASDAQ over the period of 11/09/84 to the 19/09/05, showing clearly the signs of a bubble over the time. Here the number of agents $N = 21$, 41, and 61 in each games agents have $S = 4$ strategies to choose from with a memory of 4, $M = 4$. The window size is $T = 10$.

## 7.5 Chapter Summary

In this chapter we used the Minority Game, which is a special class of agent models, to simulate the evolution of the price function using real data. It is well established that the NASDAQ has undergone a major bubble effect, which started during the late 90's and bursting in the early years of this millennium, as shown in Fig 7.10. This is commonly known as the "tech bubble".

Bubble detection and prediction remains an unsolved problem in economics, attempts like the one mentioned in Sec. 7.3.1 can be used to model these phenomena, but from Fig. 7.9 we remark that these models still remain unstable. However by using an agent model like the Minority Game it is possible to mimic the dynamics of the bubbles. It is also clear that the Minority Game does not really follow the dynamics of the real

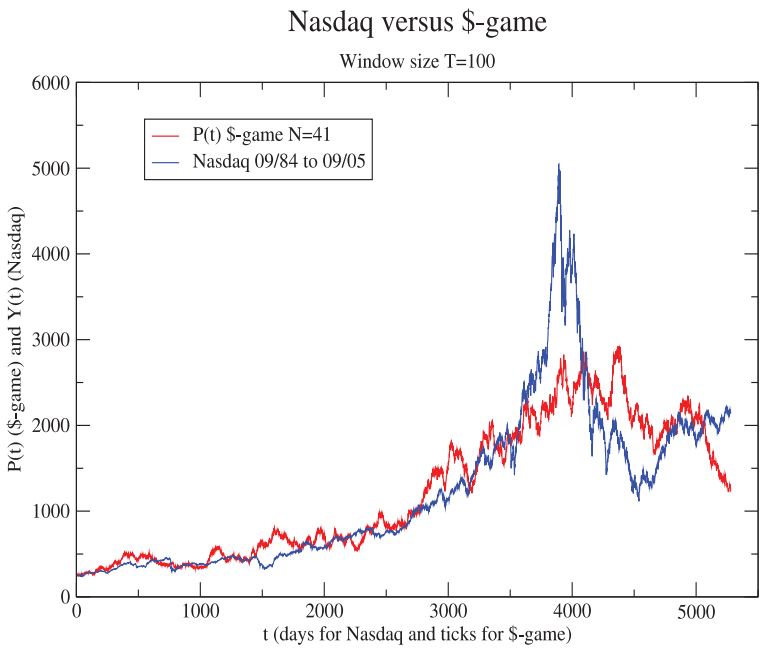


**Figure 7.15. The time series for the NASDAQ versus the $–Game price function in the Minority Game as a function of time $t$.** The time series for the NASDAQ versus the $–Game price function in the Minority Game as a function of time $t$, $1 \leq t \leq 5283$ over the period of 11/09/84 to the 19/09/05, showing clearly the signs of a bubble over the time. Here the number of strategies is $S = 2$ and the memory is $M = 8$.

data and that it is not sensitive to the presence of the bubble, as shown in Fig. 7.12, but the $–Game is a more suitable way to explore the dynamics. The flow in the Minority Game is that the updating of the scores is carried out at the wrong time, and this is what has been corrected by Sornette and Andersen (2002) by the introduction of the $–Game.

In the next chapter we conclude this thesis and explore areas for future work.

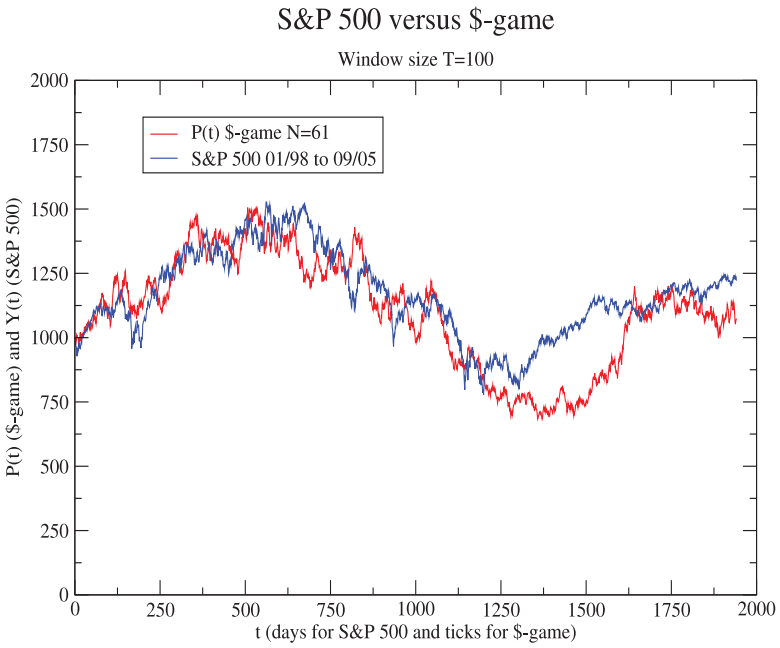


**Figure 7.16. Time series for the S&P 500 versus the $–Game price function in the Minority Game.** The graph of the S&P 500 versus the $–Game price function in the Minority Game as a function of time $t$, $1 \leq t \leq 1941$ over the period of 11/01/98 to the 20/09/05, showing clearly the signs of a bubble over the time. Here the number of strategies is $S = 2$ and the memory is $M = 8$.

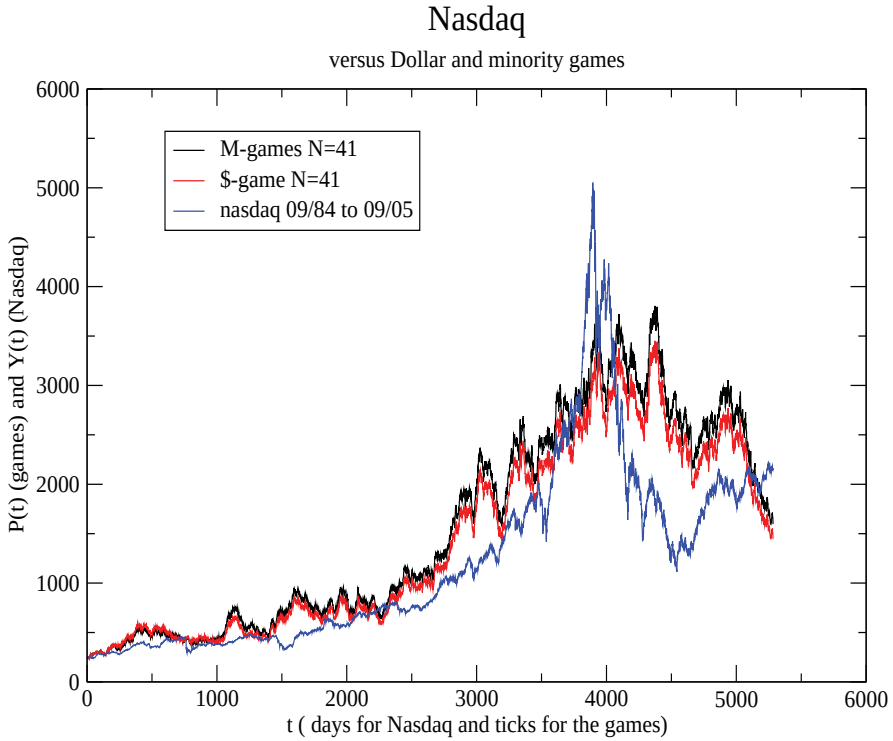


**Figure 7.17. Price function for the $–Game versus the Minority Game when the payoff function are set to 0 in the Minority Game as a function of time $t$.** The price function for the $–Game versus the Minority Game when the payoff function are set to 0 in the Minority Game as a function of time $t$, $1 \leq t \leq 5283$. This is compared to the real data from the NASDAQ over the period of 11/09/84 to the 19/09/05, showing clearly the signs of a bubble over the time.

# Chapter 8

# Conclusion

THE key thrust of this thesis in that we have used path integrals to calculate the option price as an approach to the standard approach using stochastic differential equations. We have used several types of SDE. In this final chapter we summarize the work and bring it to a conclusion, and suggest items for future investigation.

## 8.1 What is the Best Method so Far?

In this thesis we approached option pricing using path integrals to address the aims one to three in Section 1.3. However in order to address these correctly and move beyond the standard Black–Scholes–Merton model where the volatility is held constant, and where there is no memory on the historical data, one needs to address these issues separately so that a more accurate method can arise. We began this thesis by building an array of distributions that can be used to analyze financial data and construct models, this was carried out in Chapter 2.

In that chapter we saw that reasonable fits can be obtained for the log returns, Eq. (2.1), using a Student distribution, Eq. (2.8), for two different data sets namely The NASDAQ and S&P500. This was possible only when $\delta t = 1$ and we saw that as we increased $\delta t$ we lost reliability in the fits, therefore suggesting that for $\delta t > 1$ the Student distribution cannot be used, but distributions such as those described in Sec 2.2.3 would probably fit the data much better because of the extra parameters which take into account the skewness of the distribution and its shape as well as scaling factors.

Nevertheless it would be possible to create relatively accurate models based on the Student distribution.

Modeling the volatility of empirical data sets is not a simple task, one can either do it via stochastic differential equations or via time series analysis. In Chapter 2 we used simple time series analysis to see if we were able to obtain reasonable fits to the data and we saw that it was possible to obtain a rough fit using a simple combination of ARMA and GARCH models for the mean and variance equation.

We also built an array of models that can be used to model the volatility, these include models such as asymmetric GARCH models and the exponential GARCH models known as EGARCH$(p, q)$ or GJR $-$ GARCH$(p, q)$.

In Chapter 3, we briefly reviewed the fundamental building blocks of stochastic calculus. These ideas were used in Chapter 4, where we defined Îto calculus from the ground up and applied its rules and properties to some examples in the context of finance. In particular we were interested in modeling volatility—such models are known as stochastic volatility models. We also explicitly wrote out the equations used solve these stochastic differentials.

Also in this Chapter 4 we focused on two main aspects of the stochastic calculus the first one is its definition and how it is constructed from the Brownian motion, explained

in Chapter 3. Using these definitions and theorems we made contact with one of the most important equation, namely the Îto–Doeblin equation.

The second aspect was related to the evaluation of stochastic differential equations. In general an SDE cannot be evaluated analytically directly and one needs to have some methods for approximating these equations. This is normally carried out using the numerical approaches like the ones derived in Sec. 4.2. These methods can be the only way to obtain some insights into the SDE solutions. We implemented three methods that we have tested on a particular example where the explicit solution is known. These methods are to be used for model construct checking purposes. We found that out of the three methods the strong 1.5 Taylor performed best. The strong 1.5 Taylor expansion is only of order 1.5, but there are higher order methods that have been implemented by other authors. The reader is invited to see the pioneering work carried out by Burrage *et al.* (2000) on the numerical solutions of SDE or by Kloeden and Platen (1992).

Alternatively in order to obtain the transition probabilities one needs to make contact with partial differential equations that are associated with these SDE. This is done using the Kolmogorov equations. Finally the Feynman–Kac formula is the equation that is used in the evaluation of the discounted option price. These were defined in more depth in Chapter 5 where we have reviewed and given specific examples of options.

We used these ideas in the later Chapter 6 to evaluate the option price using path integrals as an alternative approach to the stochastic calculus and aimed at providing a more accurate method for evaluating the option price.

We began Chapter 6 with the simplest case where the option is path independent. In this category we have European options. We also apply path integrals to other types stochastic differential equation.

In that chapter we explored the use of path integrals in a financial context. In Section 6.1 we gave the Feynman derivation for the path integral applied to quantum mechanics. We then applied this formalism to a financial setting by making direct connections with stochastic differential equations. In Section 6.3 we saw how we can apply this formalism to the Black–Scholes–Merton model in a very efficient way. This is because the formalism in quantum mechanics for a free particle particle is essentially the same as the Black–Scholes–Merton model. That is the drift and diffusion terms are left constant. This formulation leads to a simple Lagrangian, which has the form of a Gaussian functional. These kind of functionals can be easily integrated using

Gaussian identities A.7.2. Hence the path integral can be evaluated without difficulties and making any approximations. At the end of Sec 6.1 it was already pointed out that the Feynman formulation was ill defined and could not be used on every quantum mechanical system and that the path integral could not be evaluated for system with atomic potentials in particular potential of the Coulomb type—that is potentials of the form $V(x) \rightarrow -\frac{1}{|x|}$. In other words these kind of potentials make the Lagrangian functional non–Gaussian.

The formulation of the Lagrangian functional is not unique and really depends on the structure of the SDE, we therefore cannot apply the same Lagrangian functional for all SDE, the formulation of the Lagrangian had to be generalized and this was carried out in Sec. 6.3.1. This formulation is also not unique and must take into account the dependence of drift and diffusion term on time and process variable. These Lagrangian formulations enabled us to write down the Lagrangian functional for non–Gaussian processes like the one in Sec. 6.3.3. Unfortunately these formulations led to intractable path integrals and we therefore cannot evaluate the path integral this way.

An alternative approach is to use the method of least action— this finds the path that contributes the most in the path integral and the path which minimizes the action functional. We referred to this approach as the instanton method. We can obtain such a path by directly solving the Euler–Lagrange equation. The Euler–Lagrange equation often leads to highly non–trivial, and non–linear differential equations. Nevertheless it was possible to solve such differential equations and obtain solutions. The instanton method is the most promising way in evaluating the path integral when using the Lagrangian method. This was carried out in Sec. 6.4 for a non–Gaussian model developed by Borland (2002a). Unfortunately in both cases, that is when we make most of the variables constants as in Sec. 6.4.1 in an attempt to make the problem as simple as possible or when we consider the general case, as in Sec. 6.4.1, we quickly see that these also lead to intractable Lagrangians. So in order to obtain a solution one would have to turn to numerical solutions. In addition to the level of complexity the value obtained in the Borland (2002a) model for $q$ does not correspond the possible value that we can use because the solution becomes complex, therefore forcing us to make the further assumptions, as we saw in Sec. 6.4.1. It also means that the value of $q$ needs to be renormalized somehow so that it fits the new range.

On a different approach to the Lagrangian functional approach is to use the transition probabilities directly. This is derived from the Fokker-Planck equation, Eq. (4.196)

Sec. 4.5.2. Then by repeated use of the Chapman–Kolmogorov equation, Eq. (6.27) (see Appendix A.5) we can evaluate the path integral for a given number of time slices, $N$. Ideally one wants to take the limit of $N$ to infinity to approximate the continuous limit. The idea there is to evaluate the Chapman–Kolmogorov equation for $N = 2$ and then increase $N$ to larger values so that the best procedure and pattern can be identified. We can then turn the integrand into a power series of the parameters in question. It is then possible to truncate the expansion to a given order of that parameter. This is carried out in Sec. 6.5.1 where we see in the case of the Borland (2002a) that we are able to do this with the parameter $q$. This leads us to Eq. (6.159), which is Eq. (6.159) is the most useful result, because it gives a very compact and good approximation to the path in the limit as $N \to \infty$. It should be possible to evaluate this formula numerically. Since we are interested mostly in analytical solutions for this problem, we then tried to evaluate the path integral for very small time slices —this is carried out in Sec. 6.5.1 for $N = 2$ and in Sec. 6.5.1 for $N = 3$. In the first case we are able to evaluate the path integral but in the case $N = 3$ we are not able to carry out the second integral, the one with respect with $dx(t_2)$, and we clearly see that the integral becomes intractable. As a result for any higher values of $N$ we can clearly see that the path integral rapidly becomes intractable and it is therefore not possible to obtain an analytical solution this way either.

We then turned to a different approach for pricing. We use game theory to evaluate pricing function using real data. In particular we use the Minority Game, which is then extended to the \$–Game to simulate the price functions using real data and also as a possible tool for the study of bubble dynamics in stock markets.

In that chapter we see the Minority Game, which is a special class of agent models, to simulate the evolution of the price function using real data. It is well established that the NASDAQ has undergone a major bubble effect, which started during the late 90's and bursting in the early years of this millennium, as was shown in Fig 7.10. This is commonly known as the tech bubble.

Bubble detection and prediction remains an unsolved problem in economics, attempts like the one mentioned in Sec. 7.3.1 can be used to model these phenomenon but from Fig. 7.9 we remark that these models still remain unstable. However by using an agent model based on the Minority Game it is possible to mimic the dynamics of the bubbles. It is also clear that the Minority Game itself does not really follow the dynamics of the real data and that it did not detect the presence of the bubble as was shown in Fig. 7.12.

This is because of the way the action is formulated that is the sampling is carried out at time $t$. On the other hand the $–Game is a more accurate way to model the dynamics of the real data because the sampling is carried out at time $t+1$ as shown in Eq. (7.13) instead of time $t$ as in the Minority Game. For that reason we observe in Fig. (7.12) that the $–Game follows the real data more closely than the Minority Game. Hence in the Minority Game the updating in the scores is carried out at the wrong time. This was corrected by Sornette and Andersen (2002) in the $–Game.

In summary the $–game dynamics, that is the $–Game payoff, better describes the real world than the Minority Game. The $–Game payoff appears to be sensitive to bubbles and may be useful in studying bubble dynamics.

## 8.2 Open Questions for Future Work

In reality there is no best method to solve stochastic differential equation Furthermore it is not possible to actually come with a model that fully describes the financial markets, simply because if one was to come up with a strategy, which would give indications of the future directions of the asset price, then everyone in the market would use this strategy. Eventually this would be annihilated by the market, meaning that everyone would apply it, leading for a need to a new set of strategies, as the previous one would not longer be valid. So one must constantly adjust strategies in order for it to match the market behaviors.

In other words it will not be possible to come up with a strategy that will predict the asset price accurately in the long run.

However it is possible to predict trends in asset price of some companies based on their investment strategies, exploration (for resources stocks), business plans, market prospects and acquisition etc...

Portfolio management can be optimized by combining different types of asset in the portfolio to create capital gains and/or defensive strategies when needed, this normally depends on the market trends.

On the other hand it may be possible to predict the direction of the market in a very short run using the history of an asset price. It may be possible to use stochastic differential equations but these would have to have some memory of the past hence these types of SDE will have to fall in the class of the multifractal stochastic differential equations discussed in Sec. 6.3.4, unfortunately we saw that such types of SDE

were intractable when inserted in the path integral. This is because these types of SDE are non–Markovian processes. However it may be possible to use such models at a particular scale. This was not explored in this work and is recommended as a future direction.

In Section 6.3.3 we saw that even for simpler types of SDE, such as the non–Gaussian model, the path integral was intractable. This really reduces the size of the problem set, which can be inserted in the path integral, outside the Gaussian models like the geometric Brownian motion proposed by Black–Scholes–Merton.

As an item for future work it may be possible to use Eq. (6.159) to turn the non–Gaussian path integral into a sum of Gaussian integrals. This is called Gaussian decomposition and is expounded by Anderson and Moore (1979). The Gaussian sum decomposition is possible if the probability distribution function is non–negative for values of the dependent variable and if it integrates to 1 over $\mathcal{R}^n$. There are numerous numerical approaches for approximating an arbitrary probability density by Gaussian sum (Sorenson and Alspach 1971) using non–linear optimization techniques.

This would then allow us to apply the methods on the approach used in this thesis and arrive at an analytical solution when the probability density function satisfies the last two conditions previously mentioned. Furthermore techniques used in optimal filtering (Fomin 1999) may be a useful approach for filtering the stochastic processes and estimating complicated distributions.

In Chapter 7 we show results coming from the agent models that are based on game theory. These kinds of models are the most likely candidates for modelling market behaviour because in general it is much simpler to model the market on a much smaller set of parameters, where each parameter can be tuned in turn, than to develop a stochastic model that captures only a small set of the features observed on the market. In that chapter we see that it is possible to simulate the price function using real data. Note that we only used the real data to determine market direction. The result looks very promising but needs more attention on the issue of scaling, and the interpretation of the liquidity, for example. This is recommended for future work.

In summary if one is to use path integrals in finance one would need to refine the theory a great deal in order to adapt it from what we already know within the sphere of physics to a financial context.

Also further work needs to be carried out in the numerical evaluation of non–Gaussian path integrals using numerical methods, this was the aim of this thesis, but unfortunately we were not able to take it to that stage. Also needed is to improve the Monte Carlo methods to accelerate the path integral simulation, moreover fast numerical approaches for performing the highly non–trivial integrals will also need to be the subject of future works.

## 8.3 Summary of Original Contributions

This thesis has made a number of contributions to the field of econophysics.

- The main innovation in this thesis is the approach to option pricing and the way of combining existing knowledge about distribution theory, time series, stochastic calculus, and path integrals, from different disciplines, combining them all in a fully automated algorithm in order to extract the option price.
- The idea of using scalable window size on real data as an input for model construction to then be inserted into a path integral is an original idea and to our knowledge has not been considered elsewhere.
- The first step in the implementation of this idea was carried out in Chapter 2 where we build an array of distributions that describe the data set in question. From the data set we extract a prescription for the volatility either, from time series analysis or from a stochastic volatility model.
- In Chapter 3 we review the fundamentals of Brownian motion and use these ideas in Chapter 4. In Chapter 4 we implement numerical approaches that estimate the solutions of stochastic differential equations. We use three different numerical schemes and use these as comparison checks to obtain insights for the solution of the stochastic differential model with the given probability distribution for that particular data set contained within that particluar window size. Hence the need for many different distributions to isolate the different aspects of that data set in question.
- Chapter 6 is the main original contribution of this thesis and gives some insights on how to proceed in the evaluation of path integral when using stochastic differential equations which spans beyond the Black–Scholes model, that is, beyond

geometric Brownian motion. It also gives an indication on how to proceed for future work.

- The second main contribution of this thesis in Chapter 7 is through the use of agent models where we have demonstrated an agent model for studying market bubbles. Within that framework we demonstrate the distinction between the dynamics of the Minority Game and the $–Game in the ability to simulate the price function.