UNIVERSITÀ DEGLI STUDI DI MILANO &
UNIVERSITÀ CATTOLICA DEL SACRO CUORE

SCUOLA DI DOTTORATO IN
FISICA, ASTROFISICA E FISICA APPLICATA

DIPARTIMENTO
MATEMATICA E FISICA

DOTTORATO DI RICERCA IN
FISICA, ASTROFISICA E FISICA APPLICATA
Ciclo XXIII

Adiabatic Conditions for Financial Systems

Settore scientifico disciplinare FIS/03

Tesi di Dottorato di:
Luca Spadafora

Coordinatore: Prof. Marco Bersanelli
Tutore: Prof. Fausto Borgonovi

A.A. 2009-2010

Contents

Acknowledgments

E ancora una volta mi ritrovo a scrivere la pagina più letta delle tesi, dopo tre anni di lavoro, fatiche, cambiamenti e colpi di scena.
Certamente il mio primo ringraziamento va al prof. Fausto Borgonovi, la persona che più di tutte mi ha supportato (con la u) nella produzione di questo lavoro e mi ha sopportato (con la o) nei momenti di maggiore di difficoltà, quando ostacoli di ogni tipo hanno rischiato più volte di bloccare questo progetto di ricerca. Alla fine, l'abbiamo vinta Noi! Sicuramente porterò sempre con me il metodo scientifico che mi è stato insegnato e cercherò di applicarlo in qualunque ambito, evidenziandone sempre la sua importanza.
Similmente, vorrei ringraziare il dott. Antonio Mangini e il dott. Andrea Marziali che sin dall'inizio hanno creduto in questo progetto, adoperandosi attivamente per la sua buona riuscita e cercando in tutti i modi di superare gli ostacoli, non solo burocratici, che si sono interposti durante la sua realizzazione. A loro va la mia gratitudine anche per essermi stati da guida, mentre muovevo i primi passi nel mondo della finanza, favorendo un mio rapido inserimento e assecondando le mie "follie quantitative".
Inoltre vorrei ringraziare chi, nonostante i continui cambiamenti a cui la vita ci sottopone, continua ad esserci, sempre e comunque: Daniele, irriducibile variabile stocastica e generatore di caos domestico e Gabriele che inspiegabilmente riesce ancora a dormire nonostante il "rumore".
Un ringraziamento particolare va anche a tutti coloro che si occupano di ricerca e promuovono la divulgazione della cultura. A loro, che tenacemente continuano a lavorare con passione nonostante le continue difficoltà e i sempre più scarsi finanziamenti rivolti alla ricerca, va la mia stima ed il mio ringraziamento per l'esempio che ogni giorno forniscono con le loro azioni, nella timida speranza che la miopia di alcune scelte non influisca in modo devastante sulle generazioni future.
Ai miei genitori e a tutti i miei familiari, infine, dedico questo mio lavoro, a suggellare la mia profonda riconoscenza per la loro continua fiducia che mi ha supportato nel raggiungimento di tutti i miei traguardi, accompagnandomi lungo un percorso da sempre ispirato al valore della cultura.

I would like to thank dr. G. P. Berman for his support and his nice hospitality at LANL. It was an amazing experience and I will remember it forever.

Chapter 1

Introduction

In this work we show how to use concepts that come from physics to get a suitable description of financial world, in accordance with the main ideas of a new interdisciplinary field: Econophysics. In particular, the rest of this work is organized as follows.
In Chapter 2 we present our motivations for studying Econophysics. After a personal and general description of the financial world, we explain what are the main aspects of Econophysics and why it is important for the future development of financial and economics research. In particular we focus our attention on the need for economics of a scientific revolution as claimed by J.-P. Bouchaud [1].
In Chapter 3 we make a general introduction on Econophysics and on the main mathematical and statistical instruments used to model financial world. In particular, we focus our attention on stochastic calculus and probability theory, always showing the relationship between physical and financial models. At the end of the Chapter, we also show how these instruments can be used to describe actual data that come from financial markets and we introduce the main financial concepts like market efficiency.
In Chapter 4 the main financial models for the option pricing are presented. In particular we focus our attention on Black-Scholes model and the related concept of implied volatility. For the latter we give a statistical description of the so called volatility smile effect and an intuitive and quantitative interpretation of this phenomenon.
In Chapter 5 we show how a physical concept like the adiabatic transition can be efficiently exploited to get an alternative interpretation of the volatility smile effect and how this approach can be useful to get a new fitting procedure of the volatility smile to avoid arbitrage opportunities. This and the following Chapter represent the original part of this work and our contribution to the research in this topic.

In the Chapter 6, we further extend our adiabatic approach to get a suitable characterization of the implied probability function of financial returns. In particular we focus our attention on the exponential decay of the distribution and we develop a new algorithm to get a fitting function of the volatility smile coherent with the model hypothesis and the historical relation between implied volatility and decay of the tails of the distribution.
Finally in the Chapter 7 we present our conclusions.

Chapter 2

Motivations

2.1 Introduction

The word *Econophysics* was coined by H. Eugene Stanley in the mid 1990s to describe a interdisciplinary field that makes use of the basics ideas, mathematical models and scientific method of Physics to better understand phenomena related to economy. Indeed, a strong relation between economy and complex systems can be found, and, in particular, Statistical Physics seems to be the natural field for the description of financial markets and related phenomena in so much as someone considers economy (as social activities) as a simple generalization of Physics described by the same mathematical model. From this point of view, a Unification Theory should be formulated that could describe all kinds of interactions, not only the physical interactions, but also the human and social ones. In our opinion, this vision is surely fascinating and a little bit visionary and it could be considered as a good philosophical motivation for the development of Econophysics.
In the rest of this short Chapter, we will try to give some other more practical motivations for the study of Econophysics, showing why it could play a role for the future development of economy, considering also personal working experiences.

2.2 Do We Really Need Econophysics?

After about twenty years from its "invention", the term "Econophysics" is, in our opinion, still unknown at financial world practitioners or it is considered a theoretical matter without any practical applications. In the worst cases, people consider Econophysics a forcing thing where unconventional models are used to describe phenomena that are not of interest neither for finance

nor for Physics.
To find a job position, the situation is probably more dramatic as the skills required are typically related to the understanding of *conventional* models like Black-Scholes, Heston or the knowledge of *standard* procedures as the VaR estimation by parametric approach or historical simulation. This attitude can be justified if one considers that, in general, it is not an easy task for big financial institutions to change a pricing model or a standard procedure for risk evaluation as this could mean to change the organization and the work of many people with big consequences on costs and budget. This physiologically induces a sort of inertia and changes are actuated just when they are considered necessaries, i.e. when they could have a deep impact on the P&L (Profit and Loss) of the bank or the financial institution. That explains why a matter that is intrinsically *unconventional* like Econophysics could be not well valued during the job searching activity.
Unfortunately, also from an academic point of view, the situation is not better. In fact there is a sort of reluctance by the side of physicists or economists to recognize Econophysics as something related to respectively Physics or Economy; in our opinion, this is typical for new subjects and probably Econophysics is still too young to be accepted as a new field of Physics or finance. In addiction, one should consider that similar subjects, like Quantitative Finance or Econometrics, already exist with an analogous field of application. This generates more confusion and the understanding of what we exactly mean for Econophysics is almost impossible.
If the criticisms showed above concern mainly the usefulness of Econophysics, there are many others, on more fundamental level, that are related to all the economics fields that make use of the mathematics to model some financial system. All the criticisms can be summed up quoting the sentence of a colleague: *the human being is not Nature!* By this sentence, people against the use of mathematics in economics matter mean that it is not possible to model the behavior of people by mathematics as it have been done in Physics to describe Nature, simply because people can change their opinion abruptly and without any rationality and this fact cannot be modeled by any rational approach. So, from this point of view, mathematics cannot be suitable to describe financial world because of the intrinsic irrational nature of the human being. This position could seem really poor and unfounded to the eyes of a statistical physicist that is used to model complex systems with chaotic and "irrational" behaviors and, indeed, it is; but it shows that the main source of criticism to the math-like subjects related to finance concerns the *method* used and not the details of model itself. From this point of view, Econophysics plays an important role as it tries to apply models that come from Physics to financial matters, exploiting the same scientific method that

Physics uses to model Nature. In our opinion, the most important contribution that Econophysics can give to finance is related to the method used: scientific and rigorous as in Physics, with a deep attention to actual data, as physicists do for the data of the experiments. From this point of view, Econophysics can have a really important and active role in the Scientific Revolution claimed in [1]. This aspect will be treated in the next Section.

2.3 Economics Needs a Scientific Revolution

Physicists[...] have learned to be suspicious of axioms and models. If empirical observation is incompatible with the model, the model must be trashed or amended, even if it is conceptually beautiful or mathematically convenient. So many accepted ideas have been proven wrong in the history of Physics that physicists have grown to be critical and queasy about their own models. Unfortunately, such healthy scientific revolutions have not yet taken hold in economics, where ideas have solidified into dogmas that obsess academics as well as decision-makers high up in government agencies and financial institutions. These dogmas are perpetuated through the education system: teaching reality, with all its subtleties and exceptions, is much harder than teaching a beautiful, consistent formula. Students do not question theorems they can use without thinking. Though scores of physicists have been recruited by financial institutions over the last few decades, these physicists seem to have forgotten the methodology of natural sciences as they absorbed and regurgitated the existing economic lore, with no time or liberty to question its foundations.

By these words J.-P. Bouchaud in [1] explains what is, in his opinion, the relation between Physics and Economics and why economics needs a scientific revolution. From his point of view, the most important thing that should be changed in financial, and in general, in economics world is the method and in doing this, Physics could be a good advisor. Correctly, in our opinion, he talks about a *revolution* to underline the need to completely change the way of thinking at the economy, with a completely new attitude. Theoretical and practical aspects should be included in the same theory based on empirical observations so that the big break between *thinkers* and *doers*, as suggested by Derman in [2] could be reduced and a quantitative and solid knowledge of financial markets could be motivated. To conclude, we quote the words used by J.-P. Bouchaud in [1] to underline the need to include a natural science education in the economics sciences:

[...] Most of all, there is a crucial need to change the mindset of those working in economics and financial engineering. They need to move away from what Richard Feynman called Cargo Cult Science: a science that fol-

lows all the apparent precepts and forms of scientific investigation, while still missing something essential. An overly formal and dogmatic education in the economic sciences and financial mathematics are part of the problem. Economic curriculums need to include more natural science. The prerequisites for more stability in the long run are the development of a more pragmatic and realistic representation of what is going on in financial markets, and to focus on data, which should always supersede perfect equations and aesthetic axioms.

Chapter 3

Basic Concepts on Econophysics

3.1 Introduction

This Chapter is organized in three parts. In the first part we give our definition of Econophysics and we show which are the main relationships between Physics and Financial issues. In particular we give a brief description of the main topics of Econophysics. In the second part, we introduce the basics notions about Probability Theory that are used in Statistical Physics and can be useful for a statistical description of price dynamics and other financial variables. Finally, in the third part of the Chapter, we show some empirical analysis of the statistics of real prices, defining the main concepts related the financial market description, i.e. efficient market hypothesis and arbitrage opportunities.

3.2 What is Econophysics?

Econophysics is [...] *"a neologism that denotes the activities of physicists who are working on economics problems to test a variety of new conceptual approaches deriving from the physical sciences"*. This definition is given in [3] the book that probably gives the most complete description of the topics treated by Econophysics and it is considered as a milestone for all the physicists that are interested to economics concepts and in the application of Physics in multidisciplinary field. In this definition it is understood that there *exists* a strong relation between some physical system and the economical one, at least from the conceptual point of view. This is indeed the fact that is proved by the growing number of physicist that is involved in

the analysis of economic systems and the significant number of papers of relevance to economics that is now published on physics journal. In fact, financial markets are well-defined complex systems and concepts such as power-law distributions, correlations, scaling, unpredictable time series and random process can be efficiently used to describe them. On the other hand, physicists have achieved important results in the field of phase transitions, statistical mechanics, disordered systems and nonlinear dynamics, so, in our opinion, it is somehow natural to try to understand if the same concepts can be applied in some other different field that seems to be so similar to the statistical point of view.
Another really important question that could arise from the previous definition of Econophysics is: *how Econophysics is different from the other subjects that study financial markets?* As in all classification issues, it is really hard to give a precise answer to this question and, in particular, to trace a border line between similar subjects, like *Quantitative Finance, Financial Engineering, Econometrics, etc.* At this point, we just want to stress that probably the main difference is the methodology that physicists follow to carry on their analysis, in particular the emphasis that they put on the empirical analysis of economic data and the desire to find a good fundamental description of market dynamics starting from first principles and a few parameters, sharing the common theoretical background of Statistical Physics.

3.3 Pioneering Approches

Interestingly enough, the first use of a power-law distribution and the first mathematical formalization of a random walk took place in the social sciences before than in the natural sciences. The first concept was introduced when the social economist Pareto investigated the statistical character of the wealth of individuals in a stable economy and modeled its distribution by

$$y \sim x^{-\nu}, \tag{3.1}$$

where y is the number of people having income x or greater and ν is an exponent that Pareto estimated to be 1.5 [4]. Pareto noticed that this result was quite general and applicable to different nations. It is interesting that the concept of a power-law distribution is counterintuitive, because it may lack any characteristic scale; this is related to the non convergence of the integral that evaluates the magnitude of the typical fluctuation. For this reason the use of power-law distributions was prevented until the emergence of two new paradigms:

- the application of the power-law distributions to several problems in probability theory. In particular the work of Lévy [5] and Mandelbrot [6].
- the study of phase transitions, in which it was introduced the concepts of scaling for thermodynamic functions and correlation functions.

The random walk concept was firstly formalized in a doctoral thesis, titled "Théorie de la spéculation", by Bachelier [7] in 1900, where it was discussed the problem of the pricing of the options, a particular financial contract whose value depends on another asset that is called underlying. Bachelier determined the probability of price changes by writing down what is now called the Chapman-Kolmogorov equation and recognizing that, what is now called a Wiener process, satisfies the diffusion equation. If today we reconsider Bachelier's thesis, we could notice that it lacks rigor in some of its mathematical and economic points and, in particular, the Gaussian distribution for the price changes was not sufficiently motivated. Besides, Bachelier made his statistical analysis starting from price increments, whereas economists are mainly interested in logarithmic changes of price. Nonetheless, his work is considered pioneering and somehow revolutionary.

3.4 The Modern Econophysics

After more than 100 years, from the pioneering works of Pareto and Bachelier, Econophysics evolved a lot and now there are many important areas of physics research dealing with financial and economics systems. In particular, after 1980, the electronic trading was introduced also for foreign exchange market and a great amount of data became available to test financial theories, following the same methodology that physicists use to check their theories. Even if, as already said in the previous Section, it is unfeasible to discern between different but similar subjects, for Econophysics, one can recognize the following areas of interest:

- *Analysis of stochastic process of price changes of financial assets.* The main objective of the studies in this area is to find a good statistical description of the price changes, i.e. the shape of their distribution, the temporal memory and the higher-order statistical properties. This requires a really careful analysis of historical series of price paying attention to details and to the data cleaning. A good description of price dynamic from a statistical point of view should also deal with the problematic characterization of the second moment and the so-called fat-tails of the distribution. From this point of view, the techniques used in Statistical Physics related to the Levy flights and Levy distributions seem to be the natural set up for this kind of problems [8].

- *Theoretical models for financial markets.*
  The main goal of this area of studies is to find a good model for the description of the so-called market microstructure, namely the statistical description of the market starting from the single order of the trader and his behavior. The relation with Statistical Physics is particularly evident if, for example, one considers the Ising model assuming that the values ± 1 related to the components of the system are not describing the spin of a particle, but the order to buy or to sell given by a trader at a certain moment. In this case, the "magnetization" of the system represents the general consensus of the system for a price increase or decrease. The main idea related to these kind of models is that it should be possible to get a general statistical description of an economic system just making some (noisy) assumption about the behavior of the single trader and the correlation between traders, as it was already done for statistical thermodynamics. These kind of models are called agent-based and have a great importance for the theoretical understanding of financial market and, in practical applications, for forecasting activities.

- *Rational option pricing of a derivative product.*
  One of the main goals of financial activities is estimating what is the rational price that can be associated to a particular product. This activity can be really hard because of the great complexity of some financial product whose value is connected to another financial product (underlying) in a really complex way. This kind of products are called *financial derivatives* and they had a really important role also in the recent world financial crisis. In 1973, Black and Schöles published one of the most important work about rational pricing of financial derivatives and from that moment a sort of revolution started in financial activities to continuously improve the models for option pricing to get a better estimation of the *fair value* of any financial derivatives. Despite of the huge amount of publications addressed to this kind of problems and the large amount of people with experience in hard sciences hired by banks and financial institutions, nowadays there is a great number of financial derivatives that are still priced without a suitable model and a rational estimation of the risks related to these financial products. This lack of rational knowledge has, in our opinion, a key role also in the

recent financial world crisis and justifies the interest of many physicists in this field. In this case the relation with Physics is due to the need to model the price dynamic of the underlying as a stochastic process like a Brownian motion or a Levy flight. The solution of stochastic differential equations (SDE) or partial differential equations (PDE), like the Fokker-Planck equation, has a central role for this activity.

- *Risk Management*
  As already said in the previous point, the risk estimation and control are really important activities for banks and financial institutions and it is also regulated by the Basel Committee on Banking Supervision (Basel II). For this reason a good estimation of risk is compulsory for every financial activities. As already stressed in the previous Section the distribution of price changes and financial returns are not Gaussian and a good estimation of the typical order of magnitude of the fluctuation of these variables is not an easy task. A suitable theory for the statistics of extreme and rare events (related to huge losses of money) is necessary for this activity together with a good estimation of the correlations between different assets. In this case, a good example of application of Statistical Physics for this kind of problems can be found in [9] where random matrix theory is used to get a correct estimation of the correlation of portfolio of assets, fundamental for risk estimation. This theory can be easily applied in practical situations with good results.

This list of areas of interest can be extended indefinitely and it cannot be considered completed. With this short description, we just want to give an idea of which are the main activities related to the field of Econophysics and how Physics can be exploited in financial modeling. In our opinion, it is important to stress that, although a lot of analogies can be found between the Physics of complex systems and financial modeling, it is unrealistic to look for a perfect correspondence between the two subjects and, in general, from this point of view, forcing relationships should be avoided. In our opinion, the Econophysics researcher should remember that in general people are not interested in finding Physics in Economy, but in finding good and practical solutions for Economy, using a scientific methodology that can be inspired by Physics. The application of this scientific method is probably the most important aspect that justifies why physicists should be involved in financial matters.

3.5 Basic notions of Probability Theory and Physics for Financial Modeling

In this Section we want to introduce the basic concepts of Probability Theory and Physics that can be useful for a correct interpretation of the financial models and the statistical description of the main financial variables that will be presented in the following Section and in the next Chapters.

3.5.1 Probability distributions, typical value and deviations

Randomness has a central role in Nature, in Finance as in many aspects of our life [10, 11]. In [12], it is shown how extreme and rare events has influenced dramatically the life of the author and how a rational decision process cannot leave this intrinsic uncertainty apart. This aspect is also present in Science with the introduction of Quantum Mechanics that forced scientists to abandon Laplace's deterministic vision of Nature accepting a sort of less certain description of the world as a *statistical* system. From this revolutionary vision of the Universe, the need to *measure*, mathematically speaking, this uncertainty in an appropriate mathematical framework and the consequent birth of the modern probability theory.

In order to describe a random process X for which the result is a real number, one uses a probability density $P(x)$, such that the probability that X is within a small interval of width $\mathrm{d}x$ around $X = x$ is equal to $P(x)\mathrm{d}x$. Starting from this notation the probability that X is between a and b is given by the integral of $P(x)$ between a and b

$$\mathcal{P}(a < X < b) = \int_a^b P(x)\mathrm{d}x. \tag{3.2}$$

The function $P(x)$ is a density and in this sense it depends on the units used to measure X. In order to be a probability density in the usual sense, $P(x)$ must be non-negative ($P(x) \geq 0$ for all x) and must be normalized, i.e. the integral of $P(x)$ over the whole range of possible values for X must be equal to one

$$\int_{x_m}^{x_M} P(x)\mathrm{d}x = 1, \tag{3.3}$$

where x_m (x_M) represents the smallest (largest) value which X can take. If X is not bounded one can take $x_m = -\infty$ and $x_M = +\infty$ in Eq. (3.3). An equivalent way of describing the distribution of X is to consider its cumulative distribution $P_<(x)$, defined as

$$P_<(x) = \int_{-\infty}^{x} P(u)\mathrm{d}u, \tag{3.4}$$

where $P_<(x)$ takes values between zero and one, and is monotonically increasing with x. Similarly, one defines $P_>(x) = 1 - P_<(x)$. If these probabilities are based on some information set I_t, then $P(x)$ is called conditional density. The dependence on I_t is formally denoted by $P(x|I_t)$. If the $P(x)$ is not based on any particular information, the I_t term is dropped and the density is written as $P(x)$.
It is quite natural to talk about the "typical" value of a random process and there are at least three intuitive definition for it: the *mean*, the *median* and the *most probable* . The most probable value x^* corresponds to the maximum of the function $P(x)$ (notice that x^* needs not to be unique), the median x_{med} is such that the probabilities that X be greater or less than this particular value are equal, namely $\mathcal{P}_<(x_{med}) = \mathcal{P}_>(x_{med}) = 1/2$, and the mean or *expected value* of X, which we will denote by $< x >$ or by $E[x]$, is the average of all possible values of X, weighted by the corresponding probability:

$$E[x] =< x >= \int xP(x)\mathrm{d}x. \tag{3.5}$$

Once defined the typical value of a random process, one could be interested in the "typical" order of magnitude of the fluctuations around this typical value. The definition of this second quantity can be more complicated than the definition of the first one and, in particular, once a definition is given, the meaning of the value that represents this quantity can be misleading and badly interpreted; one should pay attention to the underlying probability distribution. Some detail will be given in the following Section, when the Risk Management activity will be described in more depth.
Two possible definition for this quantity are the *mean absolute deviation* (MAD):

$$E_{abs} = \int |x - x_{med}|P(x)\mathrm{d}x \tag{3.6}$$

and the *root mean square* (RMS) σ (or, in financial terms, the *volatility*) that is given by the square root of the *variance*:

$$\sigma^2 =< (x- < x >)^2 >= \int (x- < x >)^2 P(x)\mathrm{d}x. \tag{3.7}$$

Since the variance has the dimension of x squared, its square root (the RMS, σ) gives the order of magnitude of the fluctuations around the mean value.

3.5.2 Moments and Characteristic Function

The definition of RMS can be somehow generalized and one can introduce the so-called *moments* of the distribution:

$$< x^n >= \int x^n P(x) \mathrm{d}x, \tag{3.8}$$

where $< x^n >$ is called the n-moment of the distribution $P(x)$. From the definition (3.8), it is clear that the first moment represents the mean of X, and the variance is related to the second moment ($\sigma^2 =< x^2 > - < x >^2$). Considering that the mean represents the typical value of the stochastic process X and the volatility gives the order of magnitude of the fluctuations around the mean, intuitively, one could expect that a complete specification of all the n-moments (for n very large) gives a more complete description of the stochastic process X and, in particular of its distribution $P(x)$. Indeed, this relation between moments and distribution of a stochastic process can be made more explicit introducing the so called *characteristic function*, defined as

$$P(z) = \int e^{izx} P(x) \mathrm{d}x, \tag{3.9}$$

so the explicit relation between the distribution $P(x)$ and its characteristic function can be obtained taking the inverse Fourier transform:

$$P(x) = \frac{1}{2\pi} \int e^{-izx} P(z) \mathrm{d}z. \tag{3.10}$$

It can be easily shown that the moments of $P(x)$ can be obtained through successive derivatives of the characteristic function at $z = 0$

$$< x^n >= (-i)^n \frac{\mathrm{d}^n}{\mathrm{d}z^n} P(z)\Big|_{z=0}. \tag{3.11}$$

This equation makes evident the relation between the distribution $P(x)$ and the moments $< x^n >$ by mean of the characteristic function. Analogously to the moments, one can define the *cumulants* c_n of a distribution as the successive derivatives of the logarithm of its characteristic function:

$$c_n = (-i)^n \frac{\mathrm{d}^n}{\mathrm{d}z^n} log(P(z))\Big|_{z=0}. \tag{3.12}$$

As one could expect from their definitions, there is a polynomial relation between the moments and the cumulants; for example: $c_2 =< x^2 > - < x >^2$. Finally one can define the *normalized cumulants* λ_n

$$\lambda_n = \frac{c_n}{< x^n > - < x >^n}. \tag{3.13}$$

Often one uses the third and the fourth normalized cumulants, called respectively the *skewness* and the *kurtosis*:

$$\begin{aligned} \lambda_3 &= \tfrac{<(x-<x>)^3>}{<x^3>-<x>^3} \\ \lambda_4 &= \tfrac{<(x-<x>)^4>}{<x^4>-<x>^4} - 3. \end{aligned} \tag{3.14}$$

These definitions of cumulants may look arbitrary, but they give indeed important information about the characteristic of a distribution. In fact, for example, a Gaussian distribution is characterized by the fact that all the cumulants of order greater than two are identically zero. So, considering the cumulants of order higher than two (and in particular the kurtosis) one can have an idea of the "distance" between a given distribution and a Gaussian. In general the skewness is interpreted as a measure of the asymmetry of the distribution and the kurtosis a measure of the amplitude of its tails. In particular if a distribution has a kurtosis less than zero it means that its center is thiner than a Gaussian and consequently its tails are fatter. This kind of distributions are called leptokurtic.
Notice finally that the moments (or cumulants) of a given distribution do not always exist. A necessary condition for the nth moment to exist can be derived requiring that the integral in Eq. (3.8) would converge for $|x|$ going towards infinity: the distribution density $P(x)$ should decay faster than $1/|x|^{n+1}$. In particular if one considers distribution densities that are behaving asymptotically as a power-law, with exponent $1 + \mu$

$$P(x) \sim \frac{\mu A^\mu}{|x|^{1+\mu}} \text{ for } x \to \pm\infty, \tag{3.15}$$

where A^μ is a generic constant depending on μ, then all moments such that $n \geq \mu$ are infinite. Obviously, this fact could generate some problems for the definition of the typical order of magnitude of the fluctuations of the process.

3.5.3 Some Typical and Useful Distributions

In this Section we want to introduce the main distributions typically used in financial applications and their main characteristics.

Gaussian Distribution

Surely the most famous and probably the most used distribution in many statistical field is the Gaussian distribution. We have already mentioned its

characteristics and we want to present them here in a more formal way.
A Gaussian distribution is defined by:

$$P_G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp{(-\frac{(x-m)^2}{2\sigma^2})}, \tag{3.16}$$

where m and σ^2 are the two parameters of the distribution and represent respectively the mean and the variance of the distribution. As already mentioned before, all the moments greater than two are equal to zero and this fact can be easily understood considering the Gaussian characteristic function:

$$P_G(z) = \exp{(-\frac{\sigma^2 z}{2} + imz)}. \tag{3.17}$$

The logarithm of this function is a second order polynomial for which all derivatives of order larger than two are zero. This distribution is particularly important because, applying the central limit theorem (see next Section), it can be shown that, if the number of random variables is large enough, the central part of many distributions, can be well approximated by a Gaussian distribution.

Log-Normal Distribution

Log-Normal Distribution is strongly related to the Gaussian distribution and this relation is self-evident if one compare Eq. (3.16) with its analytical expression:

$$P_{LN}(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\ln^2(x/x_0)}{2\sigma^2}\right), \tag{3.18}$$

where σ, x_0 are the parameters of the distribution. The importance of this distribution is due to its relation with the log normal Brownian motion which is the approximation in the Black and Schöles model for the price dynamic. According to this model the log normal distribution represents the distribution of the prices (S) of the underlying, as a consequence, the distribution of financial returns ($\ln(S)$) it is assumed to be a Gaussian. Differently from the Gaussian distribution, the log normal distribution is asymmetric, and predicts that large positive jumps are more frequent than large negative jumps.

Lévy Distribution

The importance of Lévy distribution is related to the context of the Central Limit Theorem (CLT) (See Sect. 3.5.5). The Gaussian distribution, it can be shown that the Lévy distribution can be a good approximation of the central

part of many distributions if the number of stochastic variables is large. It is said an attractor of stable non -Gaussian distributions; the exact meaning of this expression will be clarified in the following.
An important constitutive property of the Lévy distribution is the power-law behavior for large arguments, often called *Pareto tails*

$$L_\mu(x) \sim \frac{\mu A_\pm^\mu}{|x|^{1+\mu}} \text{ for } x \to \pm\infty, \tag{3.19}$$

where $0 < \mu < 2$ and $A_\pm^\mu$ are two constants that represents the *tails amplitude* or the *scale parameters* as they give the order of magnitude of large fluctuations of x. Notice that for $\mu \leq 2$ the variance of a Lévy distribution is formally infinite and, if $\mu \leq 1$, even the mean or the MAD fail to exist. From this point of view the interpretation of the standard deviation as the typical order of magnitude of the fluctuations loses its meaning and the two constants $A_\pm^\mu$ should be used to this purpose. If $A_+^\mu = A_-^\mu$ the distribution is symmetric. Unfortunately, an analytical expression of symmetric Lévy distributions is not know, except for $\mu = 1$, which correspond to a Cauchy (or Lorentzian) distribution:

$$L_1(x) = \frac{A}{x^2 + \pi^2 A^2}. \tag{3.20}$$

In all the other symmetric cases, Lévy distributions can be identified by its characteristic function:

$$L_\mu(z) = \exp\left(-a_\mu |z|^\mu\right), \tag{3.21}$$

where a_μ is a constant proportional to the tail parameter A^μ. Notice that for $\mu = 2$ we recover the Gaussian distribution. For asymmetric cases, Eq. (3.21) can be generalized and one gets:

$$L_\mu^\beta(z) = \exp\left[-a_\mu |z|^\mu \left(1 + i\beta \tan(\mu\pi/2)\frac{z}{|z|}\right)\right] \; (\mu \neq 1), \tag{3.22}$$

where $\beta = (A_+^\mu - A_-^\mu)/(A_+^\mu + A_-^\mu)$ is the *asymmetry parameter*. For practical applications, it can be useful to limit the power-law behavior of the Lévy distribution in a certain regime $x \ll 1/\alpha$, where α is generic parameter that defines the amplitude of the region. Beyond this region, one could ask that the distribution decays exponentially as $\exp(-\alpha x)$, so that the Pareto's tails are "truncated" for large values of x. So, it can be obtained a generalization of the Lévy distributions which accounts for this exponential cut-off, in terms of its characteristic function:

$$L_\mu^{(t)}(z) = \exp\left[-a_\mu \frac{(\alpha^2 + z^2)^{\mu/2}\cos(\mu \arctan(|z|/\alpha)) - \alpha^\mu}{\cos(\pi\mu/2)}\right], \tag{3.23}$$

where $1 \leq \mu \leq 2$. This distribution is called Truncated Lévy Distribution (TLD) and can be really useful for practical financial applications.

3.5.4 Convolutions and Sums of Random Variables

When one think at a random process, generally is considering the dynamic generated by a multiple extraction of the random variables that generate the process itself. It is quite natural, given the probability density function of a random variables, to be interested in the distribution of the sum of many independent random variables. To simplify the problem we consider the sum of two independent random variables $X = X_1 + X_2$ distributed respectively according to $P_1(x_1)$ and $P_2(x_2)$. The probability that X is equal to x is given by all the combinations of x_1 and x_2 that summed give x. Since $x = x_1 + x_2$ and two variables are independent, the joint probability that $X_1 = x_1$ and $X_2 = x_2$ is equal to $P_1(x_1)P_2(x - x_1)$, so that:

$$P(x, N = 2) = \int P_1(x')P_2(x - x')\mathrm{d}x' = P_1 \star P_2. \tag{3.24}$$

Eq. (3.24) tells us that to obtain the distribution of the sum of two independent random variables, one needs to evaluate the convolution of the two distributions of the two variables. This equation can be generalized to the sum of N random variables:

$$P(x, N) = \int P_1(x_1^{'}) \dots P_{N-1}(x_{N-1}^{'})P_N(x - x_1^{'} - \dots - x_{N-1}^{'})\Pi_{i=1}^{N-1}\mathrm{d}x_i^{'}. \tag{3.25}$$

A really important theorem shows that the convolution is a simple product in the Fourier space; for example Eq. (3.24) can be written in the Fourier space as:

$$P(z, N = 2) = \int e^{iz(x-x'+x')} \int P_1(x')P_2(x - x')dx'dx = P_1(z)P_2(z). \tag{3.26}$$

This observation simplify in many cases the calculations of the distribution generated by the sum of many random variables.

3.5.5 Central Limit Theorem

As already explained above, the importance of Gaussian distributions and Lévy distributions is due essentially to their capability to describe the central part of any distribution convoluted with itself a large number of times; in other words they are attractors of these convoluted distributions. In fact, in general, if one adds random variables distributed according to an arbitrary law $P_1(x_1)$, one constructs a random variable which has a different probability distribution given by the convolution $P(x, N) = [P_1(x_1)]^{\star N}$, where the notation $(\cdot)^{\star N}$ means that one has to take the convolution N-times. Instead, it can be shown that Gaussian and Lévy distributions are *stable* or *self-affine*, that is the shape of the convoluted distribution is the same as the elementary one. In more formal words, a distributions is stable (self-affine) if one can find a translation and dilatation of x such that the two laws coincide:

$$P(x, N)\mathrm{d}x = P_1(x_1)\mathrm{d}x_1 \quad \text{where} \quad x = a_N x_1 + b_N, \tag{3.27}$$

where a_N, b_N represent the operations of dilatation and translation. It can be shown that the family of all possible stable distributions coincide with the Lévy distributions definition given in Eq. (3.22) where Gaussian is the special case $\mu = 2, \beta = 0$. Stability is reflected on the moments estimation; in fact, it can be shown that for Lévy symmetric distribution where $\mu < 2$ one finds:

$$< |x|^q >^{\frac{1}{q}} \propto AN^{\frac{1}{\mu}} \qquad q < \mu, \tag{3.28}$$

where $A = A_- = A_+$ because we are considering symmetric distributions. This relation shows the order of magnitude of the fluctuations of N random variables taken from a Lévy symmetric distribution is a factor $N^{1/\mu}$ larger than the fluctuations of *a* Lévy random variable. However, once this factor is taken into account, the probability distributions are identical. Because of this property, Lévy distributions are considered *fixed points* in the field of distributions. The *central limit theorem* (CLT) shows that these distributions are also *attractors* for the others, in the sense that any distribution convoluted with itself a large number of times finally converges towards a stable law. The classical formulation of the CLT deals with sums of independent and identically distributed (iid) random variables of finite variance. In this case it can be shown that all these distributions converge towards a Gaussian one:

$$\lim_{N\to\infty} \mathcal{P}\left(u_1 \leq \frac{x - mN}{\sigma\sqrt{N}} \leq u_2\right) = \int_{u_1}^{u_2} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \mathrm{d}u, \tag{3.29}$$

for all finite u_1, u_2. This formulation can be extended to distributions with infinite variance. In this case the attractor is non a Gaussian, but, in general, a Lévy distribution. It is important to notice that Eq. (3.29) refers to the central part of a generic distribution of finite variance and tell us that this part can be approximated by a Gaussian. In practical situations one could be interested in what happens in the tails of the distribution, for example for risk estimation. In this cases, the application of the CLT is meaningless and the Gaussian approximation cannot be invoked. The main hypothesis ensuring the validity of the Gaussian CLT are:

- X_i must be independent random variables, or at least not "too correlated".
- X_i need not necessarily be identically distributed but one must require that the variance of all these random variables is not too dissimilar.
- Formally CLT applies only when N is infinite. For practical applications one should estimate which is region of the random variable where the Gaussian approximation is still valid and the minimum required N.
- As already explained CLT does not tell anything about the tails of the distributions.

One could be interested in defining more quantitatively which is the region of validity of the CLT, at least for Gaussian case. To do this we define X as the sum of N iid random variables of mean m and variance σ^2 and considering the relation (3.28) we define the rescaled variable

$$U = \frac{X - Nm}{\sigma\sqrt{N}}. \tag{3.30}$$

The CLT can be restated in the following way

$$\lim_{N\to\infty} \mathcal{P}_>(u) = \mathcal{P}_{G>}(u), \tag{3.31}$$

where $\mathcal{P}_{G>}(u)$ is the cumulated density function of a Gaussian distribution:

$$\mathcal{P}_{\mathcal{G}>}(u) = \int_u^{+\infty} \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)\mathrm{d}x. \tag{3.32}$$

One could try to estimate for a fixed N what is the region $|u| \ll u_0(N)$ for which the Gaussian approximation is still valid, assuming that the elementary distribution $P_1(x_1)$ decreases faster than a power-law when $|x_1| \to \infty$ such that all the moments are finite. In this case all the cumulants of P are finite and one can obtain an expansion in powers of $N^{-1/2}$ of the difference $\Delta\mathcal{P}_>(u) = \mathcal{P}_>(u) - \mathcal{P}_{\mathcal{G}>}(u)$

$$\Delta\mathcal{P}_>(u) \sim \frac{\exp(-u^2/2)}{\sqrt{2\pi}} \left(\frac{Q_1(u)}{N^{1/2}} + \frac{Q_2(u)}{N} + \ldots + \frac{Q_k(u)}{N^{k/2}} + \ldots\right), \tag{3.33}$$

where $Q_k(u)$ are polynomial functions which can be expressed in terms of the normalized cumulants λ_n (Eq, (3.13)). In particular we focus on the first two terms on r.h.s.:

$$Q_1(u) = \frac{1}{6}\lambda_3(u^2 - 1) \tag{3.34}$$

and

$$Q_2(u) = \frac{1}{72}\lambda_3^2 u^5 + \frac{1}{8}(\frac{1}{3}\lambda_4 - \frac{10}{9}\lambda_3^2)u^3 + (\frac{5}{24}\lambda_3^2 - \frac{1}{8}.\lambda_4)u \tag{3.35}$$

In general, for asymmetric distribution $P_1(x_1)$, $\lambda_3 \neq 0$, so the leading term is given by Eq. (3.34). Requiring that when $u \sim 1$ ($x - mN \sim \sigma\sqrt{N}$) the Gaussian approximation is suitable means that $Q_1(u) \sim N^{1/2}$, so we obtain the condition: $N \gg N^* = \lambda_3^2$. For large u, the relative error is obtained by dividing Eq. (3.33) by $P_{G>}(u) \sim \exp(-u^2/2)/(u\sqrt{2\pi})$, so one obtains the following condition:

$$\lambda_3 u^3 \ll N^{1/2} \text{ i.e. } |x - mN| \ll \sigma\sqrt{N}\left(\frac{N}{N^*}\right)^{1/6}. \tag{3.36}$$

The last equation shows that the central region where the Gaussian approximation is valid grows as $N^{2/3}$. Following a similar procedure, it can be evaluated the amplitude of the Gaussian region for a completely symmetric distribution ($\lambda_3 = 0$). In this case the leading term is given by Eq. (3.35) and the conditions are $N \gg N^* = \lambda_4$ and

$$\lambda_4 u^4 \ll N \text{ i.e. } |x - mN| \ll \sigma\sqrt{N}\left(\frac{N}{N^*}\right)^{1/4}. \tag{3.37}$$

In this case the central region extends over a region of width $N^{3/4}$.

3.5.6 Correlations

A really natural assumption in financial modeling is that random variables are independent and identically distributed. This assumption is, in many cases a good proxy of reality, but sometimes it requires further specification. In general we define the *correlation coefficient* of two variables X, Y as

$$\rho = \frac{< XY > - < X > - < Y >}{2\sigma_X\sigma_Y}, \tag{3.38}$$

where σ_X, σ_Y are the standard deviations of X and Y. In addiction, a standard results in Statistics tell us that the standard deviation of the variable $Z = X + Y$ is given by:

$$\sigma_Z = \sqrt{\sigma_X^2 + \sigma_Y^2 + 2\rho\sigma_X\sigma_Y}. \tag{3.39}$$

Eq. (3.39) can be really useful for practical application if one wants to evaluate the total risk of a portfolio made by different assets correlated and characterized by different levels of fluctuations. This estimation will be shown in

the following.
The concept of correlation can be relevant if one considers a stochastic process where the random variables that generate it are correlated. In this case we define the correlation function as

$$C_{i,j} =< x_i x_j > -m^2. \tag{3.40}$$

Assuming that the process is *stationary*, i.e. that $C_{i,j}$ depends only on $|i-j|$: $C_{i,j} = C(|i-j|)$, with $C(\infty) = 0$, the variance of the sum can be expressed in term of the correlation function as:

$$< x^2 >= \sum_{i,j=1}^{N} C_{i,j} = N\sigma^2 + 2N \sum_{l=1}^{N} (1 - \frac{l}{N}) C(l), \tag{3.41}$$

where $\sigma^2 = C(0)$. From this expression it is clear that if $C(l)$ decays faster than $1/l$ the sum over l tends to a constant for large N and the CLT is still valid. On the contrary, if $C(l)$ decays for large l as a power-law $l^{-\nu}$ with $\nu < 1$ the second term on the r.h.s. in Eq. (3.41) becomes relevant and the CLT cannot be applied straightforwardly.

3.5.7 Brownian Motion and Other Stochastic Process

To describe the dynamics of the market price, it is quite common to assume that its movements can be well described by the sum of two terms, one that represents the drift and is related to the time increment and one that represents the impossibility to forecast in a deterministic manner the price evolution and it is related to a random term in the equation. For that reason, we will describe theories in therms of the general stochastic differential equation:

$$\mathrm{d}S = A(S,t)\mathrm{d}t + B(S,t)\mathrm{d}W, \tag{3.42}$$

where $\mathrm{d}S$ represents the increment of the variable S, $\mathrm{d}t$ is the time increment and $\mathrm{d}W = W(t+1) - W(t)$ represents the increment (*Wiener increment*) of the stochastic variable $W(t)$. This increment is Normally distributed with zero mean and unit variance. Finally, $A(S,t)$ and $B(S,t)$ are two generic functions that represents respectively that drift term and the diffusion term. As we will show in the next Chapter, a complete specification of the functions A and B implies the specification of a model for some financial variable (price or volatility) and it represents the first step to get a good description of reality.
In this work, we will not give an exhaustive description of all the matters related to stochastic equations, we will just try to give a short summary of the main aspects of this topic. A more formal treatment of stochastic calculus can be found in [13, 14].

Itô's Lemma

One of the fundamental aspects of stochastic equations is that the quadratic variation of the Wiener increment increases with time:

$$\int_0^t (\mathrm{d}W)^2 = t, \tag{3.43}$$

where the stochastic integral is defined as the limit extension of a discrete random process

$$S(t) = \int_0^t f(\tau) dW(\tau) = \lim_{n\to\infty} \sum_{j=1}^{n} f(t_{j-1}) \left(W(t_j) - W(t_{j-1})\right), \tag{3.44}$$

where $f(t)$ is a generic function, $W(t_j)$ is a random variable extracted at the discrete time t_j and n is the number of the time discretization (see [13]). In a less formal notation Eq. (3.43) an be written as:

$$\mathrm{d}W^2 = \mathrm{d}t. \tag{3.45}$$

This relation has a great importance in differential calculus, because it generates some problems in the definition of the incremental rate that is, indeed, divergent. In addiction, because of Eq. (3.45), in evaluating a first order Taylor expansion in time, one has to consider also the second order terms in dW. By these considerations, it is intuitively justified the so-called *Itô Lemma*, that represents the stochastic generalization of the first order Taylor expansion for the stochastic process (3.42)

$$dF = \frac{\partial F}{\partial t}\mathrm{d}t + \frac{\partial F}{\partial S}\mathrm{d}S + \frac{1}{2}B^2\frac{\partial^2 F}{\partial S^2}\mathrm{d}t, \tag{3.46}$$

where $F(S,t)$ is a generic function of S and t. The last term on the r.h.s. represents the Itô correction to the common Taylor expansion and it is due to the stochastic nature of Eq. (3.42). By Eq. (3.46), given a stochastic equation for a certain variable $S(t)$, it is possible to determine the stochastic equation related to any function of $S(t)$. This fact is particularly important for the pricing of financial derivatives, when, given a certain model for the price of underlying, one can evaluate the stochastic equation related to the derivative. In the following we will make an extensive use of this Lemma.

Brownian Motion with Drift

One of the basic and, probably, most important example of random walk is the simple Brownian motion but with a drift:

$$\mathrm{d}S = \mu\mathrm{d}t + \sigma\mathrm{d}W, \tag{3.47}$$

where μ is a constant and represents the drift. The point to note about this stochastic equation is that S can be negative. This random walk would therefore not be a good model for many financial quantities, such as interest rates or equity prices that are supposed to be positive. This stochastic differential equation can be integrated exactly to get:

$$S(t) = S(0) + \mu t + \sigma(W(t) - W(0)). \tag{3.48}$$

Lognormal Random Walk

A second example of random walk is similar to the above but, in this case, the drift and the randomness scale with S:

$$\mathrm{d}S = \mu S \mathrm{d}t + \sigma S \mathrm{d}W. \tag{3.49}$$

In this case, if S starts out positive it can never become negative; the closer that S goes to zero, the smaller the increment $\mathrm{d}S$. This property can be easily derived if one considers the function $F = \ln(S)$ and applies the Itô's Lemma:

$$\mathrm{d}F = \frac{\partial F}{\partial S}\mathrm{d}S + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 F}{\partial S^2}\mathrm{d}t = (\mu - \frac{1}{2}\sigma^2)\mathrm{d}t + \sigma \mathrm{d}W. \tag{3.50}$$

This shows that F follows a Brownian motion and so can range between minus and plus infinity even if it cannot reach these limits in a finite time. As a consequence S cannot reach zero or infinity in a finite time. From Eqs. (3.48,3.50) and from the definition of F it follows that:

$$S(t) = S(0)e^{(\mu - 1/2\sigma^2)t + \sigma(W(t) - W(0))}. \tag{3.51}$$

Orstein-Uhlenbeck and CIR Random Walk

Some financial variables like volatility can be well described by the so-called mean reverting processes. These processes are defined so that they continuously move around their mean. An example is given by the *Ornstein-Uhlenbeck process*, defined as:

$$\mathrm{d}S = \theta(\mu - S(t))\mathrm{d}t + \sigma \mathrm{d}W, \tag{3.52}$$

where θ, μ, σ are three positive constants of the model. From the definition it is clear that if the variable $S(t)$ is greater than μ the drift term of the equation is negative and so it decreases the value of S, while if $S(t)$ is less than μ the drift term is positive and the value of $S(t)$ is increased. The net result is that the variable $S(t)$ continuously oscillates around the mean value μ. From this interpretation it follows that θ represents the velocity at which such trajectories will regroup around μ and σ gives the amplitude of the fluctuations around the mean. The solution of Eq. (3.52) is given by:

$$S(t) = S(0)e^{-\theta t} + \mu(1 - e^{-\theta t}) + \int_0^t \sigma e^{\theta(s-t)} \mathrm{d}W(s), \tag{3.53}$$

from which the mean and the variance can be evaluated

$$\begin{aligned} E[S(t)] &= S(0)e^{-\theta t} + \mu(1 - e^{-\theta t}) \\ Var[S(t)] &= \tfrac{\sigma^2}{2\theta}(1 - e^{-2\theta t}). \end{aligned} \tag{3.54}$$

Furthermore, it can be shown that the Orstein-Uhlenbeck process admits for really large time ($t \to +\infty$) a Gaussian stationary solution where μ represents the average of the process and $\sigma^2/(2\theta)$ its variance.

As a generalization, one can consider a sum of squared Orstein-Uhlenbeck processes. In this case one obtains a *CIR process*, named after its creators John C. Cox, Jonathan T. Ingersoll, and Stephen A. Ross [15], defined as:

$$\mathrm{d}S = \theta(\mu - S(t))\mathrm{d}t + \sigma\sqrt{S(t)}\mathrm{d}W, \tag{3.55}$$

where θ, μ, σ are the parameters of the model and have the same meaning of the Orstein-Uhlenbeck process. Provided that $2\theta\mu > \sigma^2$, the process has a stationary gamma distribution with scale parameter $\sigma^2/(2\theta)$.

Fokker-Planck and Kolmogorov Equations

In the preceding paragraphs, we showed some examples of stochastic equations that can be useful for financial modeling. It can be shown that for each of these equations there is an equation that governs the time evolution of probability density function of the considered stochastic process. This equation can be directly deduced from the general structure of the stochastic equation. The easiest way to deduce this equation is to consider a trinomial approximation to the continuous-time random walk. This approximation implies that in a time step δt, there are only three possibilities for the evolution of the stochastic variable, it can go up or down of a certain amount δS or stay at the same level. The probabilities of a rise and fall are respectively ϕ^+ and ϕ^-, so the probability of staying at the same level comes from the normalization condition $1 - \phi^+ - \phi^-$. From this quantities is it possible to evaluate the mean of the change in S after the time step:

$$E[dS] = \phi^+\delta S + (1 - \phi^+ - \phi^-) \cdot 0 + \phi^-(-\delta S) = (\phi^+ - \phi^-)\delta S \tag{3.56}$$

and the variance:

$$Var[dS] = \delta S^2(\phi^+ + \phi^- - (\phi^+ - \phi^-)^2). \tag{3.57}$$

To the leading order, the mean of the change in the continuous-time version of the random walk is, from Eq. (3.42)

$$E[dS] = A(S,t)\delta t \tag{3.58}$$

and the variance is

$$Var[dS] = B(S,t)^2\delta t. \tag{3.59}$$

To make Eqs. (3.56, 3.57) consistent with Eqs. (3.58, 3.59) we impose:

$$\phi^+(S,t) = \frac{1}{2}\frac{\delta t}{\delta S^2}(B(S,t)^2 + A(S,t)\delta S) \tag{3.60}$$

and

$$\phi^-(S,t) = \frac{1}{2}\frac{\delta t}{\delta S^2}(B(S,t)^2 - A(S,t)\delta S). \tag{3.61}$$

In addiction, because of the diffusive properties of the process we impose the scaling relation:

$$\delta S \sim \sqrt{\delta t}. \tag{3.62}$$

Now it is possible to find the equations for the transition probability density function. In our trinomial walk we can only get the point S' at the time t', if S at the time t assumes one of these three values three values: S', $S' + \delta S$ or $S' - \delta S$. The probability of being at S' at time t' is related to the probabilities of being at the previous three values and moving in the right direction:

$$\begin{aligned} P(S',t'|S,t) &= \phi^-(S'+\delta S, t'-\delta t)P(S'+\delta S, t'-\delta t|S,t) \\ &+ (1-\phi^-(S',t'-\delta t)-\phi^+(S',t'-\delta t))P(S',t'-\delta t|S,t) \\ &+ \phi^+(S'-\delta S,t'-\delta t)P(S'-\delta S,t'-\delta t|S,t). \end{aligned} \tag{3.63}$$

We can easily expand each of the terms in Taylor series about the point S', t'. For example:

$$P(S'+\delta S, t|S,t) \sim P(S',t'|S,t) + \delta S\frac{\partial P}{\partial S'} + \frac{1}{2}\delta S^2\frac{\partial^2 P}{\partial S'^2} - \delta t\frac{\partial P}{.}\partial t' + \dots \tag{3.64}$$

We will omit the rest of the details, but the result is:

$$\frac{\partial P}{\partial t'} = \frac{1}{2}\frac{\partial^2}{\partial S'^2}(B(S',t')^2P) - \frac{\partial}{\partial S'}(A(S',t')P). \tag{3.65}$$

This is the *Fokker-Planck* or *forward Kolmogorov equation*, a forward parabolic partial differential equation, requiring initial condition at time t and to be solved for $t' > t$. For example considering the log normal random walk (Eq. (3.49)), the Fokker-Planck equation becomes:

$$\frac{\partial P}{\partial t'} = \frac{1}{2}\frac{\partial^2}{\partial S'^2}(\sigma^2 S'^2 P) - \frac{\partial}{\partial S'}(\mu S' P). \tag{3.66}$$

Considering the initial condition:

$$P(S', t|S, t) = \delta(S' - S), \tag{3.67}$$

where δ represents the Dirac's function, a solution of Eq. (3.66) is given by:

$$P(S', t'|S, t) = \frac{1}{\sigma S'\sqrt{2\pi(t'-t)}} \exp\left(\frac{-(\ln(S'/S) - (\mu - 1/2\sigma^2)(t'-t))^2}{2\sigma^2(t'-t)}\right). \tag{3.68}$$

Analogously one can derive the equation governing the probabilities of reaching a specified final state from various initial states. In this case the equation is given by:

$$\frac{\partial P}{\partial t} = \frac{1}{2}B(S,t)^2\frac{\partial^2 P}{\partial S^2} - B(S,t)^2\frac{\partial P}{\partial S}. \tag{3.69}$$

This is the so-called *backward Kolmogorov equation*.

3.5.8 Adiabatic Description of Physical Systems

In this Section we want to introduce a particular technique that is used to describe physical systems that are characterized by a slow varying parameter that specifies the properties of the system. We will refer to these kind of systems as *adiabatic systems*. In the following Chapters we will show that this way of modeling systems can be efficiently used to describe financial systems that are governed by a slow varying parameter like implied volatility to get an efficient calibration procedure of the model.
From a physical point of view, the adiabatic approach can be well described by the following example. We consider a mechanical system executing a finite motion in one dimension and characterized by some parameter λ which specifies the properties of the system or of the external field in which it is placed and let us suppose that λ varies slowly (*adiabatically*) with time as the result of some external action. By "slow" variation we mean one in which λ varies only slightly during the period T of the motion:

$$T\frac{\mathrm{d}\lambda}{\mathrm{d}t} \ll \lambda. \tag{3.70}$$

If λ were constant, the system would be closed and would execute a strictly periodic motion with a constant energy E and a fixed period $T(E)$. When the parameter λ is variable, the system is not closed and its energy is not conserved, but, because of our slow varying hypothesis, we could expect that also the rate of change of the energy $\dot{E}$ will also be small. This idea can be extended to other kind of systems described by an adiabatic parameter. In general, one could expect that if the rate of variation of the parameter is sufficiently slow, the properties of the model will not change a lot, equivalently to the energy for a physical system. So, from this point of view one could be interested in defining quantitatively *when* a system is adiabatic and what happens when it is not adiabatic. For the physical system considered, one could expect that if the rate of change of the energy is averaged over the period T and the "rapid" oscillations of its value are thereby smoother out, the resulting value $\dot{E}$ determines the rate of steady slow variation of the energy of the system, and this rate will be proportional to the rate of change $\dot{\lambda}$ of the parameter. In other words, one could expect that $\dot{E}$ will be some function of λ and that this dependence can be expressed as the constancy of some combination of E and λ. We will refer to this quantity as *adiabatic invariant*. Let $H(p, p; \lambda)$ the Hamiltonian of the system, that depends o the parameter λ. The rate of change of the energy of the system is

$$\frac{\partial H}{\partial t} = \frac{\partial H}{\partial \lambda}\frac{\partial \lambda}{\partial t}. \tag{3.71}$$

The expression on the right depends not only on the slowly varying quantity λ but also on the rapidly varying canonical variables q and p. To get the steady variation of energy, one must to take the average n the time period of the motion. Since λ varies only slowly, we can take $\dot{\lambda}$ outside the averaging:

$$< \frac{\mathrm{d}E}{\mathrm{d}t} >= \frac{\mathrm{d}\lambda}{\mathrm{d}t} < \frac{\partial H}{\partial \lambda} >, \tag{3.72}$$

where $\partial H/\partial\lambda$ has to be averaged considering p and q as variables, and not λ. In a more explicit form, one has

$$< \frac{\partial H}{\partial \lambda} >= \frac{1}{T}\int_0^T \frac{\partial H}{\partial \lambda}\mathrm{d}t. \tag{3.73}$$

In addiction, one should notice that, according to Hamilton equations $\dot{q} = \partial H/\partial p$ and $\mathrm{d}t = \mathrm{d}q/(\partial H/\partial p)$, the time period T can be written as:

$$T = \int_0^T \mathrm{d}t = \oint \frac{\mathrm{d}q}{\partial H/\partial p}, \tag{3.74}$$

where $\oint$ denote an integration over the complete range of variation of the co-ordinate during the period T. Thus Eq. (3.72) becomes

$$< \frac{\mathrm{d}E}{\mathrm{d}t} >= \frac{\mathrm{d}\lambda}{\mathrm{d}t}\frac{\oint(\partial H/\partial\lambda)\mathrm{d}q/(\partial H/\partial p)}{\oint \mathrm{d}q/(\partial H/\partial p)}. \tag{3.75}$$

Now, we have to make some considerations about the variable dependences of the system. As already mentioned, because of the operation of averaging, λ can be consider constant, and along such a path also the Hamiltonian has a constant value E. As a consequence the momentum can be defined as a function of the co-ordinate q and of the two independent constant parameters E and λ, namely $p = p(q; E, \lambda)$. So differentiating the equation $H(q, p, \lambda) = E$ with respect to λ we have $\partial J/\partial\lambda + (\partial H/\partial p)(\partial p/\partial\lambda) = 0$, or

$$\frac{\partial H/\partial\lambda}{\partial H/\partial p} = -\frac{\partial p}{\partial\lambda}. \tag{3.76}$$

Substituting this in the numerator of Eq. (3.75) and writing the integrand in the denominator as $\partial p/\partial E$, we obtain

$$< \frac{\mathrm{d}E}{\mathrm{d}t} >= -\frac{\mathrm{d}\lambda}{\mathrm{d}t}\frac{\oint(\partial p/\partial\lambda)\mathrm{d}q}{\oint(\partial p/\partial E)\mathrm{d}q} \tag{3.77}$$

or

$$\oint(\frac{\partial p}{\partial E} < \frac{\mathrm{d}E}{\mathrm{d}t} > +\frac{\partial p}{\partial\lambda}\mathrm{d}\lambda\mathrm{d}t)\mathrm{d}q = 0. \tag{3.78}$$

Finally this may be written as

$$< \frac{\mathrm{d}I}{\mathrm{d}t} >= 0, \tag{3.79}$$

where

$$I = \frac{1}{2\pi}\oint p\mathrm{q} \tag{3.80}$$

and the integral is taken over the path for given E and λ. Eq. (3.79) shows that, if our hypothesis are respected, an adiabatic system can be characterized by a variable I which is constant when the parameter λ varies, i.e. I is an *adiabatic invariant*. The existence of this constant is a consequence of the relation between the rate of change of the energy in the system and the rate of change of the variable λ, that "controls" these changes.
For general and interdisciplinary applications it is of crucial importance to understand *when* the system can be considered adiabatic. In fact let us suppose that we have a system (not necessarily a physical system) that is controlled by a parameter λ. As a consequence of the previous treatment, we know that if the parameter does not vary too much, also the system itself does not change a lot, because it can be described by a similar energy level. In our vision, this means that the mathematical model that correctly describes the system for a certain value of λ can be used to describe the system also for other value of the adiabatic parameter, until the adiabatic hypothesis are

satisfied. So, knowing when the system is adiabatic tell us when a particular simplified model can be applied with a good level of precision and when it should be avoided. In the following Chapters a concrete application of these ideas will be shown.

3.6 Statistics of Real Prices

In this Section we want to analyze the statistic properties of the main variables of financial markets. In particular we want to describe the price dynamic as a process using the mathematical tools described in the previous Section. In addiction, we want to give an general description about price formation and analyze the main characteristics of market microstructure.

3.6.1 Prices, Financial Returns and Distributions

To get a model of financial market, it is quite natural to analyze firstly the statistical properties of the prices. In particular, if one wants to describe the price evolution in terms of Wiener *increments*, it is quite natural to consider as the relevant variable some function of price *increment*. In the whole modern financial literature, it is postulated that the relevant variable is not the *price increment* itself, $\delta S = S(t+\tau) - S(t)$ where $S(t)$ is the price at time t and τ represent the unit of time, but rather the *logarithmic financial return* $z = \ln(S(t+\tau)/S(t))$. Notice that if $S(t+\tau)/S(t) \sim 1$, the definition of logarithmic financial return is equivalent to the common definition of percentage variation, i.e. $\ln(S(t+\tau)/S(t)) \sim (S(t+\tau) - S(t))/S(t)$. The hypothesis to take the logarithmic return as the relevant variable can be tested considering actual market data and the standard deviation of the price increment $\sqrt{<\delta S^2>|_S}$ conditioned to a certain value of price S. If the logarithmic hypothesis is correct, one should get a dependence on the price; $\sqrt{<\delta S^2>|_S} = \sigma_1 S$, where σ_1 is a constant and represents the RMS of log returns. The results of this analysis are shown in [16]. In general it seems that the log return hypothesis is correct only if one considers a long historical series, otherwise one should consider a combination of the two variables as a function of the time scale. Except for different specifications, in the following we will assume that our model has to describe a really long time series, so the logarithmic return represents the right choice as relevant variable.
A typical way to test the hypothesis of a stochastic model of price dynamic is to check if actual returns distribution is well represented by the model itself and, in particular, one is interested in the tails of the distribution. From empirical analysis it was observed that, if the unit of time short (up

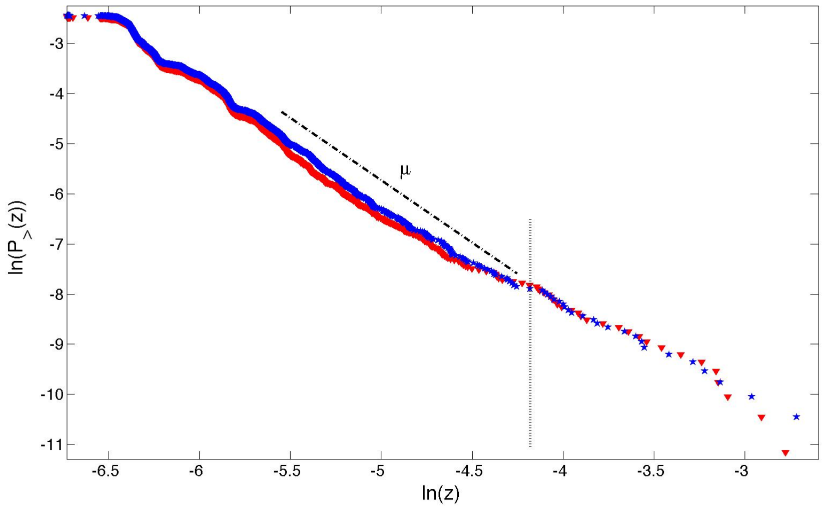

Figure 3.1: Cumulative distribution $\mathcal{P}_>(z)$ (for $z > 0$) and $\mathcal{P}_<(z)$ (for $z < 0$) considering the historical price series of ENI from 01/01/2008 to 01/06/2010 (Bloomberg data-provider) with 1 min of time lag. The vertical line shows the regions where the power-law decay and the exponential decay are suitable.

to one day), returns distributions exhibit a power-law decay in the tails (the so-called fat-tails) for a large region of z (Fig. 3.1).

However, for really large values of $|z|$, the tails seems to be well approximated by an exponential decay. These results seem to be almost robust and recurrent for many asset classes, so, for these time units, a good description of reality seems to be given by a stochastic process involving a truncated Lévy distribution (Eq. (3.23)). If one considers greater units of time something changes and it seems that a good description can be reached using exponential or Gaussian distributions (see [3, 39] where this scaling behavior is well investigated).

3.6.2 Risk Management

The knowledge of the distribution of financial returns has a crucial role in the risk estimation and, in general, in Risk Management activities. In general, in practical estimation of the risk, common assumptions are:

- Financial returns are distributed as a Gaussian

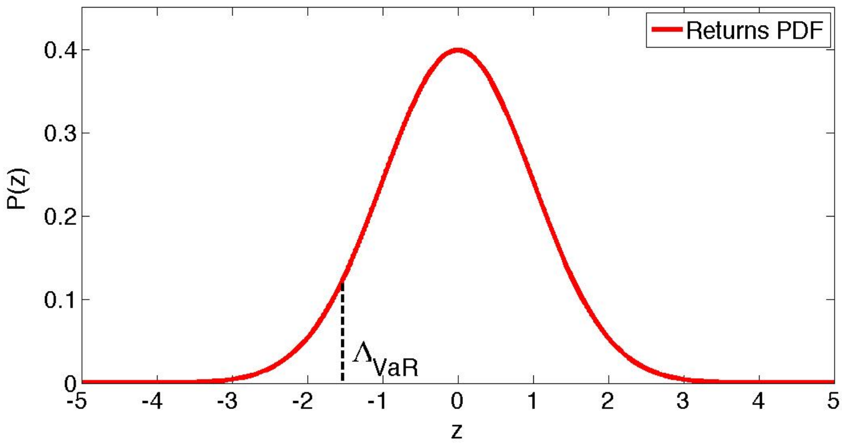

Figure 3.2: We show a graphical representation of the VaR for a theoretical Gaussian distribution.

- The variance of the returns scales linearly with time: $\sigma^2(T) = \sigma_\tau^2 T$, where σ_τ^2 is the unit time variance and T is measured in τ-unit of time
- The relation between assets can be described considering the linear correlation of returns using Eq. (3.38), so, an estimation of the total risk related to a portfolio can be obtained by generalizing Eq. (3.39):

$$\sigma_P^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} \rho_{i,j}\alpha_i\alpha_j\sigma_i\sigma_j, \tag{3.81}$$

  where N represents the number of assets in the portfolio, $\rho_{i,j}$ is the correlation matrix between the asset i and the asset j with $\rho_{i,i} = 1$, α_i is the amount of money invested in the asset i and σ_i is the standard deviation of of returns of the asset i.

Using these assumptions one is able to evaluate the *Value-at-Risk* (VaR) of a generic portfolio, that should represent the market risk that is related to the holding of the portfolio. In a more formal way, the VaR is defined as the level of loss Λ_{VaR} corresponding to a certain probability of loss $\mathcal{P}_{VaR}$ over time time interval τ:

$$\int_{-\infty}^{-\Lambda_{VaR}} P_\tau(z)\mathrm{d}z = \mathcal{P}_{VaR}, \tag{3.82}$$

where $P_\tau(z)$ is the PDF of returns (Fig. 3.2).

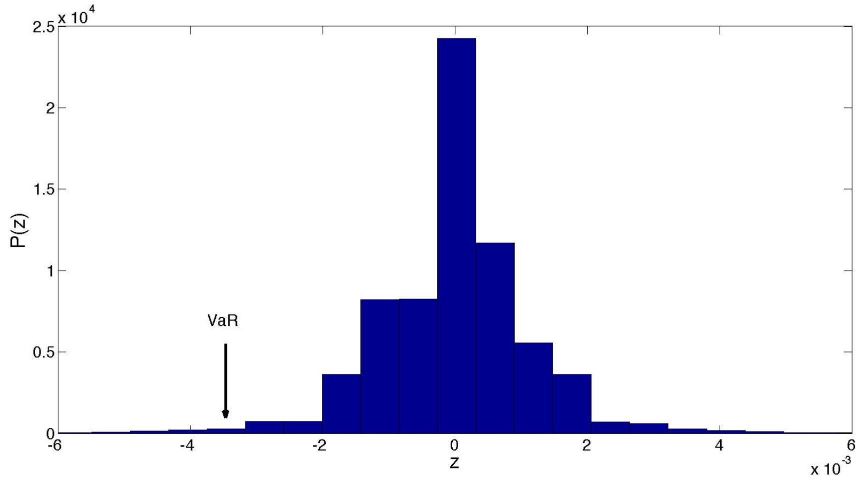

Figure 3.3: We show a graphical representation of the returns distribution (not normalized) obtained considering the historical price series of ENI from 01/01/2008 to 01/06/2010 (Bloomberg data-provider) with 1 min of time lag. The arrow shows our historical VaR estimation at 1% confidence level.

For example, if one fix $\mathcal{P}_{VaR} = 1\%$ and $\tau = 1$ day, it means that a loss greater than Λ_{VaR} happens only every 100 days. If the preceding assumptions hold, VaR estimation of a portfolio can be made easily by measuring historical standard deviation of assets returns and the correlations between assets. Then, using Eq. (3.81), the total variance of the portfolio can be evaluated and, multiplying it by the suitable scaling factor, one can obtain the estimation of VaR at the require level of confidence $\mathcal{P}_{VaR}$. Finally, using the second assumption one can rescale the VaR at the right time unit.
This is a really simple (and maybe naive) estimation of the value-at-risk that can be properly extended for the estimation of the risk of much more complex financial instruments like derivatives. In literature more sophisticated approaches that generalize the assumptions related to Gaussian distribution of returns can be found that can improve VaR estimation. For portfolios that do not include complex financial instruments, a good solution that is insensitive on the choice of the distribution is the so-called historical estimation. In this case, the value-at-risk is extracted directly from the historical series of returns considering a fixed percentile (Fig. 3.3).

In this case any assumption on volatility and correlation estimation are not needed and the only problem could arise in managing the huge amount

of data necessary for the estimation.
VaR definition is subjected to many criticisms, some of them are reported in the following:

- VaR does not take into account the shape of the tails, but it considers the possible losses up to a certain level of confidence. From this point of view, two portfolios with different distributions of losses could be classified by the same risk.
- VaR does not take into account the value of the maximal loss *inside* the period τ. For example, if we consider $\tau = 1$ day, with VaR estimation we do not have any information about the maximal loss during the day, but just at the end of the day.
- VaR does not take into account the fact that losses can accumulate on consecutive time intervals τ, leading to an overall loss which might substantially exceed Λ_{VaR}.

Nonetheless these criticisms, VaR is still considered a fixed point and its definitions is generally accepted in financial world; currently it is used by many financial institutions for risk estimation.

3.7 Efficient Market Hypothesis and Some Elements of Market Microstructure

One of the most accepted paradigm in market description is the so-called *Efficient Market Hypothesis* (EMH). The underlying idea is that the market is highly efficient in the determination of the most rational price of the traded asset. More precisely, a market is said to be efficient if all the available information is instantly processed when it reaches the market and it is immediately reflected in a new value of prices of the assets traded. This hypothesis is strongly related to the impossibility to forecast future price fluctuations starting from the analysis of the historical series of price. Samuelson [18] showed mathematically that assuming the rational behavior of market participants and market efficiency, it is possible to demonstrate there is a relation between the expected future value of the price $S(t+1)$ and its past history:

$$E[S(t+1)|S(0), S(1) \dots S(t)] = S(t), \tag{3.83}$$

i.e. price evolution can be represented by a martingale stochastic process. The notion of martingale is, intuitively, a probabilistic model of the "fair game", that is to say, a game where gains and losses cancel and the future

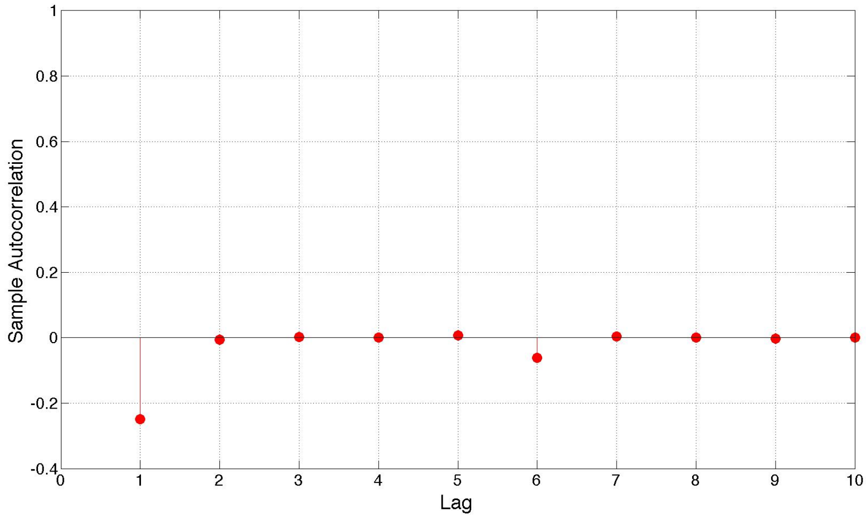

Figure 3.4: We show the autocorrelation as a function of the time lag. We considered a tick-by-tick historical series of EURUSD from 01/01/2007 to 10/06/2010. A correlation of 0.25 do exist for the first time lag. For the other time lags, autocorrelations are not statistically significant.

expected wealth is equal to the present one.
EMH was checked by empirical observations and theoretical considerations that showed that price changes are difficult if not impossible to predict if one starts from time series of price changes. In particular it can be shown that autocorrelation of price increments is substantially negligible for unit time of the order of minutes and this justifies the use of random walk in modeling financial markets.
For really short units of time, correlations are not negligible and are seen as local violations of EMH (Fig. 3.4).

In general, different assumptions should be taken in consideration to obtain a suitable description of realty, in particular the so-called market microstructure. A typical effect related to market microstructure is the negative autocorrelation of price increments, measured for many financial assets [19]. It can be shown that this correlation holds just for the first time lag if one consider a tick-by-tick dataset; from this fact the name *first order* correlation. For the other time lags, autocorrelations mainly lie within the 95% confidence interval of an identical and independent Gaussian distribution and so they can be neglected. For this effect, at least three explanations were proposed [20]:

- The traders have different and diverging opinions about the price and that generate opposite fluctuations of the price.
- Negative correlation is a direct consequence of the market marking tendency to skew the spread in a certain direction to control the evolution of the price
- When direction of the price is not clear, some bank systematically publish overestimated bid-ask spread that can generate negative correlations.

An other interesting aspect of market microstructure is related to a typical transaction cost: the so called bid-ask spread, that is defined as the difference between the best price to buy and to sell at a particular time. In fact, in real market there is not just one price, as in low frequency financial models, but there exist a book of prices that represents, at a particular time t, all the orders of the traders (buyers and sellers) that are waiting for an opposite order to be executed. An example of a typical order book is shown in Table (3.1).

Amount Bid	**Bid**	**Ask**	**Amount Ask**
13295	9.7700	9.7800	21345
28978	9.7650	9.7850	38151
22897	9.7600	9.7900	47930
34798	9.7550	9.7950	16663
58500	9.7500	9.8000	23571

Table 3.1: First five levels of a book. For every levels it is shown the two price and the quantity of shares available for that price. The data in the Table refers to FIAT on 19/08/2010 at 15.58

In this case, the first five levels of price of the book are represented. This book tells to an hypothetical trader that it is possible to buy 21345 shares of FIAT at 9.78 , 38151 shares at 9.785, 47930 shares at 9.79 and so on. Analogously, the hypothetical trader could sell 13295 shares of FIAT at 9.77 , 28978 shares at 9.765, 22897 shares at 9.76. Alternatively, the trader can introduce his own order in the book and wait for an other order that could match his offer. In the tables, the bid-ask spread is represented by the difference between the best buy price 9.78 and best sell price 9.77 and it is interpreted as the friction faced by traders with the most direct access to the market. A good model of the bid-ask spread can be of fundamental importance for every high frequency strategies as it represents a suitable estimation of the transaction costs that a trader should face to exploit an autocorrelation-based

strategy. In [16] it is shown that a quite robust relation holds between the bid-ask spread and the instantaneous volatility of the market estimated on a tick-by-tick time frame. This relation can be interpreted as the need of market maker of increasing their gain given by bid-ask spread to cover their risk exposure due to the volatility of the market. Interestingly enough, it seems that a good estimation of this risk can be obtained by a tick-by-tick volatility instead of the conventional volatility estimated on physical units of time.
Negative autocorrelation of price increments and the relation between bid-ask spread and volatility are just two examples of market microstructure effects that should be taken in consideration in high-frequency modeling activities; most of these effects are still lacking of an exhaustive interpretation.

3.8 Conclusions

In this Chapter we presented a general introduction to Econophysics. After a brief description of the history of Econophysics we presented the many mathematical instruments that come from Physics and are used to model financial markets. In particular we focused on Central Limit Theorem, Brownian Motion and Adiabatic Conditions; these topics will be recurrent in the following Chapters to develop the option pricing models. In the last part of the Chapter, we discussed about the main recurrent hypothesis on financial market and when these hypothesis are suitable for the description of financial data.

Chapter 4

The Volatilities of Financial Markets

4.1 Introduction

In this Chapter we want to introduce the basic concepts about financial derivatives and the most important mathematical models to get the fair value of these complex financial instruments. In particular we focus our attention on the most famous model in the options pricing world: the Black-Schöles (BS) [21] model and its generalizations. In doing this, we also discuss one of the main problem related to the application of this model in practical situations, namely the calibration of the parameters and the so called volatility smile (VS) effect. In particular we emphasize the statistical properties of the volatility implied by the BS model, its relation with other kind of volatility (historical and local) and the main problems related to its estimation. At the end of this Chapter we stress the importance of implied volatility for practical applications and its crucial role for hedging strategies and the pricing of exotic derivatives.

4.2 Introduction to Financial Derivatives

Financial derivatives are complex financial instruments that, in the last decades, increased their importance in the financial world and had a central role also in the recent financial world crisis. From this point of view, getting a reliable estimation of the fair value of many financial instruments is still an open problem of great importance and justifies the great amount of publications that in the recent years try to develop new models more and more suitable for options pricing, risk estimation, trading strategies, etc. Be-

fore describing the main mathematical models for options pricing we need to clarify what we exactly mean for financial derivatives and give a general introduction about this topic.
J. Hull, in one of the most important book about quantitative finance [22] defines a derivative as: *a financial instrument whose value depends on (or derives from) the values of other, more basic, underlying variables. Very often the variables underlying derivatives are the prices of traded assets. A stock option, for example, is a derivative whose value is dependent on the price of a stock. However, derivatives can be dependent on almost any variable, from price of hogs to the amount of snow falling at a certain ski resort.* In particular, in this thesis, we focus on options derivatives, which are special contracts between a buyer and a seller that gives to the buyer of the option the right, but not the obligation, to buy or to sell a specified asset (underlying) on or before the option's expiration time, at an agreed price (the strike price). From this definition it is clear there is a sort of asymmetry between the two parties of the contract: in fact, the buyer of the option can *choose* to buy or sell the underlying, unlike the seller of the option; because of this asymmetry the buyer has to pay a premium to the seller that corresponds to the value of the option. It is straightforward to determine what is fair value of the call and put options at the time to maturity starting from their definition. Let us firstly consider the call option case. At the time to maturity the buyer of the option has to decide if exercising his right of buying the option or not. A rational decision about what to do depends on the difference between the spot price of the underlying at the time to maturity and the strike price fixed by the contract. If the spot price is greater than the strike price for the buyer of the call is gainful to exercise his right; his gain from this operation will be given by the difference of the two prices. On the contrary, if the spot price is less than the strike price, there is no reason for the buyer to exercise his right because it would be more convenient to buy the underlying directly on the market. In this second case the value of the option is zero. These two situations can be summarized by the formula

$$C(S,T) = \max(S(T) - K, 0), \tag{4.1}$$

where $S(T)$ is the spot price of the underlying at the time to maturity T and K is the strike price. By a similar argumentation, it can be shown that the payoff of a put option at the time to maturity is

$$P(S,T) = \max(K - S(T), 0). \tag{4.2}$$

The main objective of the option pricing theory is to evaluate what is the fair value of the option at the time $t < T$, given the time to maturity T, the strike price K and the current spot price of the underlying S_0. Option derivatives can be classified by the time when they can be exercised, the type of payoff and other details related to the kind of contract. In the following we report the main definitions about option derivative.

- *Call Option:* gives the holder the right to buy the underlying asset by a certain date for a certain price.
- *Put Option:* gives the holder the right to sell the underlying asset by a certain date for a certain price.
- *American Option:* can be exercised at any time up to the expiration date.
- *European Option:* can be exercised only on the expiration date itself.
- *Bermudan Option:* can be exercised only at some particular fixed time before the expiration date.
- *Barrier Option:* An option contract that may only be exercised when the underlying asset reaches some barrier price. A barrier option may either be a *knock-in* or a *knock-out*. A knock-in may only be exercised when the underlying asset rises above or falls below (depending on the particular terms) the barrier price. On the other hand, a knock-out automatically expires when the underlying asset rises above or falls below the barrier price. It is important to note that the barrier price is distinct from the exercise price, though, theoretically, they may be set at the same amount.
- *Vanilla Option:* option with simple payoff, namely Call and Put options.
- *Exotic Option:* option with complex payoff. It cannot be classified as call and put option.
- *Option at the money (ATM)*: when the spot price is similar to the strike price.
- *Option out of money (OTM)*: when the spot price is less than the strike price.
- *Option in the money (ITM)*: when the spot price is greater than the strike price.

4.3 The Black, Schöles and Merton Model

4.3.1 Standard Derivation

In 1973, Black and Schöles published what is considered the most famous model in option pricing theory [21] and it represents a really simple and effective answer for the fair value estimation of a European call and put option. The derivation of their formula is really simple and makes use of the stochastic mathematical tools shown in the previous chapter. The main hypothesis of the BS model are:

- It is possible to borrow and lend cash at a known constant risk-free interest rate.
- There are no transaction costs or taxes.
- The stock does not pay a dividend.
- All securities are perfectly divisible.
- There are no restrictions on short selling.
- There is no arbitrage opportunity.
- Options use the European exercise terms, which dictate that options may only be exercised on the day of expiration.
- The price of the underlying follows a log normal Brownian motion

  $$\mathrm{d}S = \mu S \mathrm{d}t + \sigma S \mathrm{d}W, \tag{4.3}$$

  where S is the price of the underlying and $\mathrm{d}S$ its infinitesimal increment, $\mathrm{d}t$ is the time increment, μ is a constant that represent the drift term, σ is the constant diffusion term and $\mathrm{d}W$ is the Wiener increment, a stochastic term distributed as a Gaussian of zero mean and variance equal to dt.

Starting from these hypothesis it is possible to build an imaginary portfolio composed by a long position of a call option (the derivation is similar for a put option) and a short position of a certain amount Δ of the underlying:

$$\Pi(S,t) = C(S,t) - \Delta(S,t)S(t). \tag{4.4}$$

Assuming that Δ does not change over the time interval dt, we differentiate Eq. (4.4)

$$\mathrm{d}\Pi(S,t) = \mathrm{d}C(S,t) - \Delta(S,t)\mathrm{d}S(t). \tag{4.5}$$

Applying Itô's Lemma to the differential $\mathrm{d}C(S,t)$, with $\mathrm{d}S$ given by Eq. (4.3), we find

$$\mathrm{d}C = \frac{\partial C}{\partial t}\mathrm{d}t + \frac{\partial C}{\partial S}\mathrm{d}S + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2}\mathrm{d}t \tag{4.6}$$

and so, substituting the last Equation in Eq. (4.5), we obtain:

$$\mathrm{d}\Pi(S,t) = \frac{\partial C}{\partial t}\mathrm{d}t + \frac{\partial C}{\partial S}\mathrm{d}S + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2}\mathrm{d}t - \Delta \mathrm{d}S. \tag{4.7}$$

From Eq. (4.7) it is clear that by choosing $\Delta = \frac{\partial C}{\partial S}$ any source for randomness can be cancelled making the theory insensitive to the random fluctuations of the market. In addiction, by this choice of Δ, it is cancelled also the dependence of the drift term μ, that is difficult to estimate because it requires a forecast of real market evolution. In financial terms Δ is called the *delta* of the option and represents the amount of underlying that the writer of the option should buy to cancel the market risk (*delta hedging*). Since the stochastic term has been removed from the equation governing the evolution of Π, we require

$$\mathrm{d}\Pi(S,t) = \Pi(S,t) r \mathrm{d}t, \tag{4.8}$$

where r represents the risk free interest rate, namely the rate of interest for an investment without risk. Eq. (4.7) become

$$\frac{\partial C}{\partial t} + \frac{\partial C}{\partial S} rS + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} = rC. \tag{4.9}$$

This is a partial differential equation (PDE) that needs boundary conditions to be solved. This condition is given by the payoff of the call option at the time to maturity which is given by Eq. (4.1). Eq. (4.9) can be solved using different techniques [23] and the solution is (Fig. 4.1):

$$C(S,t) = S(t)N(d_1) - Ke^{-r(T-t)}N(d_2), \tag{4.10}$$

where

$$\begin{aligned} d_1 &= \frac{\ln(S(t)/K) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{(T-t)}}, \\ d_2 &= d_1 - \sigma\sqrt{(T-t)}, \\ N(x) &= \tfrac{1}{\sqrt{2\pi}} \textstyle\int_{-\infty}^{x} \; dz \; e^{-z^2/2}. \end{aligned} \tag{4.11}$$

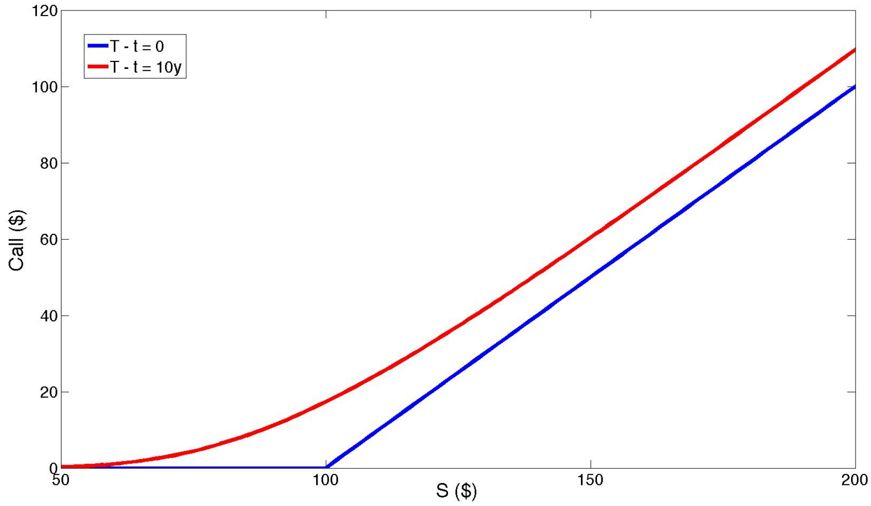

Figure 4.1: We show the value of a call option as a function of the spot price $S(t)$, considering the strike price $K = 100\$$, the volatility $\sigma = 0.1$, the risk free rate $r = 0.01$ and a time to maturity of 10 years (red line). For comparison we show also the value of the option at the maturity with the same parameters (blue line).

The structure of Eq. (5.3) is similar to the payoff (4.1) except for the fact that in this case the spot price and the discounted strike price are weighted in a probabilistic manner by the cumulative of the gaussian distribution. The solution of the BS model depends on five parameters: $K, T, r, S(t), \sigma$. Four of these five parameters can be easily obtained, in fact: the strike price (K) and the time to maturity (T) are written in the call contract, the spot price ($S(t)$) can be taken from the market and the risk free parameter (r) can be evaluated considering, for example the LIBOR rate for the fixed time to maturity. The only real free parameter of the model is σ that represents the random fluctuation of the market and it should be estimated by some calibration procedure that will be presented in the next Chapters. Using a similar demonstration, it can be shown that the fair value for a put option is given by:

$$H(S,t) = SN(d_1) - Ke^{-r(T-t)}N(d_2). \tag{4.12}$$

There is a really useful formula that relates the prices of a call and a put option. This relationship is model-free and follows from the trivial fact that

$S_T - K = \max(S(T) - K, 0) - \max(K - S(T), 0)$

$$C(t) - H(t) = S(t) - e^{-r(T-t)}K. \tag{4.13}$$

Eq. (4.13) is known as *put-call parity.*

4.3.2 Risk Neutral Approach

There is another way to derive BS formula starting from a really general approach that is based on *risk neutral pricing.* The main idea of this approach is founded on the fact that we are not interested to forecast the market, we just want to know what is the *fair value* of an option assuming that all the source of randomness of the market can be cancelled buying or selling a certain amount of the underlying (*delta-hedging*). In the previous Section we showed that to cancel this randomness we had to impose $\Delta = \frac{\partial C}{\partial S}$, that is exactly the amount of underlying to buy or sell to cover the risk due to the market fluctuations. By this operation we moved from a market risk world to a risk free world where no forecast about market trend is needed and the drift term μ of Eq. (4.3) is substituted by the risk free parameter r. Intuitively the risk neutral approach is based on the assumption that the fair value of an option can be obtained just considering the (discounted) expectation value of the relative payoff considering a *new risk free probability measure*, i.e. the probability measure Q that transform the discounted process $O(t + u) = e^{-ru}O(t)$, where $u > 0$ and $O(t)$ represent the value of an option at the time t, in a martingale process [14]

$$O(t) = \mathrm{E}^Q[O(T)|\mathcal{F}_t], \tag{4.14}$$

where $\mathcal{F}_t$ represents the filtration at the time t, i.e. all the information available at the time t. By the *fundamental theorem of asset pricing* such a measure Q exists *if and only if* the market is arbitrage-free, i.e. by this theorem, if the market is arbitrage-free, it is possible to describe the time evolution of an asset price as a martingale process under a certain "synthetic" probability. It can be shown that the probability that satisfies Eq. (4.14) (under the assumption that the portfolio is self-financing, see next paragraph) for the log normal process (4.3) is given by:

$$\ln(S(T)/S(t)) \sim N(r - \frac{1}{2}\sigma^2(T-t), \sigma^2(T-t)), \tag{4.15}$$

where $N(\alpha, \beta)$ is the normal distribution with mean α and variance β. Notice that Eq. (4.15) is independent on the drift term μ and depends on the risk free parameter r whence the name risk neutral distribution. Considering the boundary condition (4.1), Eq. (4.14) implies for a call option:

$$\begin{aligned} C(t) &= \mathrm{E}^Q[e^{-r(T-t)}\max(S(T)-K,0)|\mathcal{F}_t] \\ &= e^{-r(T-t)}\int_0^\infty \max(S(T)-K,0)P(S(T)|S_t)\mathrm{d}S(T), \end{aligned} \tag{4.16}$$

where $P(S(T)|S_t)$ represents the risk neutral probability distribution of the price of the underlying and we assumed that the risk free term r is constant. Using Eqs.(4.16 - 4.15) it is possible to obtain again Eq. (5.3) that shows the equivalence of the risk neutral approach for option pricing. Notice that Eq. (4.16) is valid for every risk neutral distribution modeling the underlying price in a risk free world and in this sense it is a generalization of the BS equation (4.9).

4.3.3 Accuracy of the BS Model

In Section 4.3.1 the source of randomness in the market has been neglected imposing that

$$\Delta = \frac{\partial C}{\partial S}, \tag{4.17}$$

then Δ could be dependent on S. This assumption is in contrast with Eq. (4.5) where we assumed that Δ is constant. So, strictly speaking, our derivation of the BS formula is not coherent and one could wonder what is the accuracy related to this simplification. To better understand the problem, we consider again the derivation of the BS equation, taking into account Eq. (4.17). Our theoretical portfolio is

$$\Pi(t) = C(t) - \frac{\partial C}{\partial S}S(t). \tag{4.18}$$

Then, differentiating yields

$$\mathrm{d}\Pi(t) = \frac{\partial C}{\partial t}\mathrm{d}t + \frac{\partial C}{\partial S}\mathrm{d}S - \frac{\partial C}{\partial S}\mathrm{d}S - \frac{\partial S}{\partial t}\mathrm{d}\frac{\partial C}{\partial S}, \tag{4.19}$$

where the last term in r.h.s. represents the differential form of Δ. Applying Itô's Lemma to the Δ we get

$$\mathrm{d}\frac{\partial C}{\partial S} = \frac{\partial^2 C}{\partial S \partial t}\mathrm{d}t + \frac{\partial^2 C}{\partial S^2}\mathrm{d}S + \frac{1}{2}\frac{\partial^3 C}{\partial S^3}\sigma^2 S^2 \mathrm{d}t. \tag{4.20}$$

Thus, the formal differential form of Π is given by

$$\mathrm{d}\Pi = (\frac{\partial C}{\partial t}\mathrm{d}t + \frac{\partial C}{\partial S}\mathrm{d}S) - \frac{\partial C}{\partial S}\mathrm{d}S - S\left[\left[\frac{\partial^2 C}{\partial S \partial t} + \frac{\partial^2 C}{\partial S^2}\mu S + \frac{\partial^2 C}{\partial S^2}\right]\mathrm{d}t + \frac{\partial^2 C}{\partial S^2}\sigma S \mathrm{d}W\right]. \tag{4.21}$$

This equation can be simplified if one differentiates Eq. (4.9) with respect to S

$$\frac{\partial^2 C}{\partial S \partial t} + \frac{\partial^2 C}{\partial S^2} rS + \frac{1}{2}\frac{\partial^3 C}{\partial S^3}\sigma^2 S^2 + \sigma^2 S \frac{\partial^2 C}{\partial S^2} = 0 \tag{4.22}$$

and uses this relation to eliminate most of the unwanted terms in Eq. (4.21). Unfortunately this simplification is not enough and, to make Eq. (4.21) consistent with Eq. (4.5), one should require

$$S^2 \frac{\partial^2 C}{\partial S^2}(\sigma dW + (\mu - r)\mathrm{d}t) = 0, \tag{4.23}$$

which will not hold in general.
Fortunately, even if formally speaking our derivation of the BS equation is not satisfactory, the method we used still gives us the correct partial differential equation. The answer is in the additional term on the l.h.s. in the Eq. (4.23). This term has nonzero expectation under the true probability measure but it can be shown that the following relation holds for the risk free distribution (4.15)

$$\mathrm{E}^Q[S^2 \frac{\partial^2 C}{\partial S^2}(\sigma \Delta W + (\mu - r)\Delta t)] \sim 0, \tag{4.24}$$

where in this case Δ represents the finite differential increment (not to be confused with the Δ for the randomness cancellation). Thus, in small intervals, the extra cost associated with the portfolio Π has zero expectation and this fact justifies the coherence of our derivation. It is interesting that the average is taken with respect to the synthetic risk-neutral measure and not with respect to the real-life probability.

4.4 Implied Volatility and Volatility Smile

In this Section, we want to introduce the main concept that is the object of all this work, namely the *implied volatility* (IV). Typically, traders on option markets and practitioners consider the BS model as a zeroth order approximation that takes into account the main features of options prices. To get a pricing closer to the actual data, they consider the volatility as a parameter that can be adjusted considering the inverse problem

$$C^{\mathrm{BS}}(S(t), t, K, T, \sigma) - C^{\mathrm{mkt}}(t) = 0, \tag{4.25}$$

where C^{BS} is the value of the call option given by the BS formula and C^{mkt} represent the observed market price of the call option. Eq. (4.25) defines

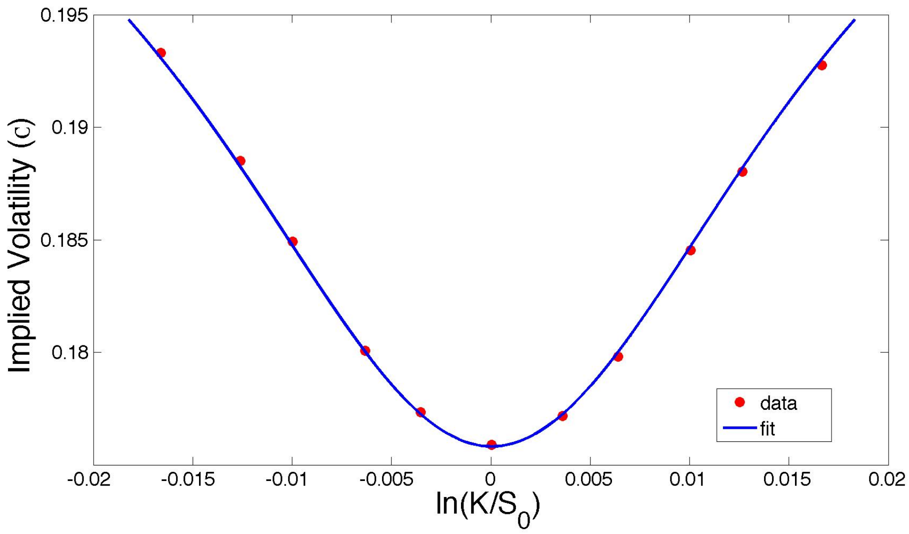

Figure 4.2: We show a typical volatility smile as a function of the logarithm of the strike price K. S_0 represent the (fixed) spot price at the time $t = 0$.

implicitly the value of the volatility σ that is, using the words of Rebonato [24]: the wrong number to put in the wrong formula to obtain the right price. In this way a more reliable value of the volatility (*implied volatility*) can be obtained and used to price more complex options for which analytical solutions are not available. The value of the implied volatility depends on the value of the strike, K, in a well-known characteristic curve called the *smile volatility* (typically for foreign currency options) whose shape is approximately parabolic and symmetric (see Fig. 4.2), or *skew volatility* (typically for equity options) when asymmetric effects dominate [25, 26, 27, 28].

An intuitive explanation of this shape can be found if an actual returns distribution is considered. In fact, it is well known that the tails of the returns PDF are not Gaussian but exhibit a power law decay (fat tails) [29, 30]. On the contrary, BS model assumes that the PDF of returns is Gaussian thus underestimating the actual probability of rare events. To compensate for this model deficiency, one has to consider the greater implied volatility for strike out of the money then for strike at the money.

To get the right price for every call (and put) option one has to choose different values of the IV for every combination of the strike price and the time to maturity. Thus IV is in fact a mapping from time, strike prices and expiry days to $\mathbb{R}^+$

$$\sigma : (t, K, T) \to \sigma_t(K, T). \tag{4.26}$$

This mapping is called *implied volatility surface* (IVS).
A central question about IVS is: *what does the IVS imply for practice?* In two points are raised to answer to this question. The first one is that IV summarizes all the information needed for the pricing in one single entity, hence it is common practice to quote options in terms of IV. From this point of view the BS formula is a simple way to map every option to the same number dependent on the strike price. For this purpose it is not necessary to believe in the BS model. It simply acts as a computational tool insuring a common language among traders. The second point, which is negative, is that for each strike price and each time to maturity we have a different IV, so a different BS model is applied. This generates a lot of difficulties in managing a portfolio of options, from the hedging and exotic pricing point of view.
Commonly, IV is interpreted as the future average volatility of the market and represents the traders and pratictioners vision. In the following it will be shown that this interpretation is of fundamental importance for pratical applications and to get a good calibration procedure of the BS model.

4.4.1 Quantitative Interpretation of the Implied Volatility

As already explained in the previous Section often the BS formula is interpreted as a zeroth order approximation of a model that has to be adapted to real life situations by changing an *a priori* constant parameter to get an approximate theory that describes more subtle effects. In this case the subtle parameter is the volatility and the subtle effect is due to the fact that the distributions of the financial returns are not Gaussian but exhibit fat-tails. We already presented an intuitive explanation of the recurrent shape of the IV as a function of the strike price. In this Section we want to explain it in a more quantitative manner following the lines in [16]. First of all, we assume that the price process is *additive* rather than multiplicative, namely the real independent random variable is not the logarithmic return $(\ln(S(t)/S_0)$ as required by BS model but the price increment itself $(S(t+1) - S(t))$. Eq. (4.16) is still valid and we can get a new Gaussian pricing formula:

$$C_G = \int_K^{+\infty} (S - K) \frac{1}{\sqrt{2\pi T S_0^2 \sigma^2}} \exp\left(-\frac{(S - S_0)^2}{2S_0^2\sigma^2 T}\right) dS. \tag{4.27}$$

Eq. (4.27) can also be derived directly from BS formula in the small maturity limit, where the relative price variations are small allowing one to write $y = \ln(S(T)/S_0) \sim (S(T) - S_0)/S_0$. In this limit Gaussian and log-normal

distributions become really similar.
Then, let us notice that Eq. (4.16) through integration by parts can be written as

$$C = \int_K^{+\infty} \mathcal{P}_>(S, T|S_0, 0)dS, \tag{4.28}$$

where we defined

$$\mathcal{P}_>(u) = 1 - \mathcal{P}_<(u) = \int_{-\infty}^{u} P(x)dx, \tag{4.29}$$

where $P(x)$ represents a generic distribution of the variable x. In particular if the distribution is Gaussian we define

$$\mathcal{P}_{\mathcal{G}>}(u) = \int_u^{+\infty} \frac{1}{\sqrt{2\pi}} \exp{(-x^2/2)}dx. \tag{4.30}$$

If the distribution $P(x)$ is decreasing faster than algebraic when $|x| \to \infty$ (and such that all the moments are finite), Eq. (4.28) can be written as an expansion of cumulants function, using the relation derived in 3.5.5

$$C = C_G + \sigma S\sqrt{T}\frac{e^{-u_K^2/2}}{\sqrt{2\pi}}\left(\frac{\lambda_3}{6\sqrt{N}}u_K + \frac{\lambda_4}{24N}(u_K^2 - 1) + \frac{\lambda_3^2}{72N}(u_K^4 - 6u_s^2 + 3) + \dots\right), \tag{4.31}$$

where C_G represents a Gaussian pricing formula of Eq. (4.27), λ_k is the k-normalized cumulant, $u_k = (K - S_0)/\sqrt{\sigma^2 S_0 T}$ and N represents the number of the stochastic variables that generate the price evolution such that $T = N\tau$ where τ is the unit of time. Eq. (4.31) shows the contributions on call pricing due to the skewness (λ_3) and the kurtosis (λ_4) of the distribution of the price respect to the pricing given by a Gaussian distribution. Neglecting the term associated to the skewness, we get

$$\Delta C = \frac{\lambda_4 \tau}{24T}\sqrt{\frac{DT}{2\pi}} \exp\left(-\frac{(K - S_0)^2}{2DT}\right)\left(\frac{(K - S_0)^2}{DT} - 1\right), \tag{4.32}$$

where $\Delta C = C - C_G$ and $D = \sigma^2 S_0^2$. On the other hand, we can estimate the variation of Eq. (4.27) when the volatility changes by a small quantity $\delta D = 2\sigma S_0^2 \delta\sigma$

$$\delta C = \delta\sigma S_0 \sqrt{\frac{T}{2\pi}} \exp\left[-\frac{(K - S_0)^2}{2DT}\right]. \tag{4.33}$$

Comparing Eqs. (4.32, 4.33) it is possible to understand how to change the volatility in a Gaussian context to reproduce the kurtosis effect:

$$\Sigma = \sigma + \delta\sigma = \sigma\left[1 + \frac{\lambda_4}{24N}(\frac{(K - S_0)^2}{DT} - 1)\right]. \tag{4.34}$$

Eq. (4.34) shows the effect on the implied volatility of Gaussian pricing model if the price increments distribution has a kurtosis different than the Gaussian PDF. The parabolic shape of the formula justifies quantitatively the smile aspect of the IVS and creates a relationship between the fat-tails characteristic shape of actual price increments distribution and the increasing implied volatility for options out of the money.

4.4.2 Static Properties of the Smile Volatility

In the previous Sections we defined the IV and we described intuitively and quantitatively the VS effect. Now we want to analyze the main properties of this variable in order to understand how to model it coherently and to get a better pricing calibration.

Bounds on the Slope

From the general fact that European call prices are monotonically decreasing and puts are monotonically increasing functions of the strike price, it is possible to get boundary conditions on the slope of the smile. Assuming that $K_1 < K_2$ from the previous observation, we have for a fixed time to maturity

$$C(K_1) \geq C(K_2), \;\; P(K_1) \leq P(K_2) \tag{4.35}$$

and due to an observation in [31] this can be improved to

$$C(K_1) \geq C(K_2), \;\; \frac{P(K_1)}{K_1} \leq \frac{P(K_2)}{K_2}. \tag{4.36}$$

Now, assuming an explicit dependence of volatility on strike prices, we get

$$\frac{\partial C}{\partial K} = \frac{\partial C^{BS}}{\partial K} + \frac{\partial C^{BS}}{\partial \sigma}\frac{\partial \sigma}{\partial K} \leq 0, \tag{4.37}$$

so

$$\frac{\partial \sigma}{\partial K} \leq -\frac{\partial C^{BS}/\partial K}{\partial C^{BS}/\partial \sigma}. \tag{4.38}$$

Repeating the differentiation for the second inequality in (4.36), we get

$$\frac{\partial \sigma}{\partial K} \geq \frac{P^{BS}/K - \partial P^{BS}/\partial K}{\partial P^{BS}/\partial \sigma}. \tag{4.39}$$

Putting together Eqs. (4.38 - 4.39) and using the BS formula for call and put options, we get

$$-\frac{N(-d_1)}{\sqrt{T-t}KG(d_1)} \leq \frac{\partial \sigma}{\partial K} \leq \frac{N(d_2)}{\sqrt{T-t}KG(d_2)}, \tag{4.40}$$

where $G(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)$. Eq. (4.40) gives an estimation of the upper and the lower bound for the volatility smile to avoid arbitrage opportunities. These bounds has been derived starting from monotonicity considerations about call and put prices and are very broad estimates of the slope of the smile function in practical situations.

Large Strike Behavior

It is possible to characterize the behavior of the volatility smile for large strike price, imposing again the monotonicity of the BS formula for call and put option. Defining

$$x = \frac{K}{e^{r(T-t)}S(t)}, \tag{4.41}$$

for sufficiently large $|x| > x^*$, one has [32]

$$\sigma(x,T) < \sqrt{\frac{2|x|}{T}}. \tag{4.42}$$

This estimation of the limit behavior of the smile volatility can be made more precise to get a quantitative relation between the slope coefficients of the asymptotes of the implied variance function and the number of finite moments in the underlying distribution of the price $S(t)$. Intuitively this results can be explained considering the fact that the IV smile must carry the same information as the underlying risk neutral transition probability, so it should hold a relationship between the asymptotic behavior of the smile and the tail behavior of the risk neutral PDF of financial returns. The general result is that the smile should not grow faster than $\sqrt{|x|}$ and, furthermore, it should not grow slower than $\sqrt{|x|}$ unless one assumes that the distribution of $S(T)$ has finite moments of all the orders. These results has a great importance for practical applications where the extrapolation of the IVS is necessary beyond the values at which options are typically observed. The choice of the extrapolation is a delicate matter especially for exotic option pricing that depends significantly on the specific extrapolation method.

Empirical Regularities

Analogously to the statistical properties of financial returns distributions shown in the previous Chapter, for IVS too, are observed some empirical regularities that can be summarized as follows:

- For short time to maturities the smile is very pronounced, while it becomes more and more shallow for longer T.

- The smile function achieves its minimum when the strike price is ATM. This behavior can be justified considering the quantitative interpretation of the volatility smile (see Sect. 4.4.1).
- The dependence of the IV on the time to maturity (term structure) is increasing, but may also display a humped profile, especially during high market fluctuations
- OTM put regions display an higher levels of IV than OTM call regions
- The standard deviation of IV (volatility of volatility) is bigger for short maturity options and monotonically declining with time to maturity
- Financial returns of the underlying asset and IV increments are negatively correlated, indicating the so called *leverage effect*.
- IV appears to be mean-reverting
- Shocks across IVS are highly correlated, so the dynamics of IVS can be modeled by a small number of driving factors

All these regularities justify the recent interest in IV dynamics and should be taken in consideration in every modeling attempt.

4.5 Problems with IV

Essentially, the concepts of IV and IVS were introduced to make the results of the BS model more similar to reality as a perturbative fine-tuning approach. Unfortunately, these perturbations generates some ambiguity in the hedging, in the risk management and in the pricing activities that could creates some problem in practical applications. Firstly, let us consider the hedging problem. The standard BS model application requires to buy or sell a certain amount of the underlying asset to cancel the source of randomness of the market, according to the value of the delta, that measures the sensitivity of call price on the price of the underlying. This procedure is well defined when the implied volatility is constant, but it is less clear when one considers a volatility as a function of the strike price. In fact one should be aware that the IV is not necessarily equal to the *hedging volatility*, defined as

$$\frac{\partial C^{BS}}{\partial S}(S(t), t, K, T, \sigma) - \frac{\partial C^{\text{mkt}}}{\partial S} = 0, \tag{4.43}$$

which is the volatility that equates the BS delta with the delta of true, but unknown, pricing model and represented by the second derivative in Eq. 4.43.

From the risk management point of view other difficulties appear when different BS models are applied for different strikes. In this case one may wonder if the delta risks across different strikes can simply be added to control the overall risk in the option book or another more suitable procedure should be followed.
The situation is unclear also in the pricing activity that involves exotic options, for example an ITM knock-out option with strike price K and barrier $L > K$. In this case it is not clear which IV should be used for pricing, if the one at K, the one at L or some average of both.
All these problems can be solved using some pricing model that includes in its structure the smile effect. Some of these models will be presented shortly in the following Sections.

4.6 Local Volatility

The concept of *local* volatility (LV) was introduced by Dupire [33] and further developed by Derman and Kani [34]. Intuitively one may think at local volatility, denoted by $\sigma_{K,T}$, as the market's consensus of the instantaneous volatility for a market level K at some future date T. The ensemble of such estimates for a collection of market levels and future dates, analogously to the implied volatility, is called *local volatility surface* (LVS). The concept of local volatility was introduced to get a reliable estimation of the fair value of the option derivatives that could include in the model the VS effect, to avoid the problems presented in the previous Section.
One starts assuming that the asset price dynamic can be modeled by the following stochastic process

$$\mathrm{d}S(t) = \mu(S(t), t)S(t)\mathrm{d}t + \sigma_I(S(t), t, \cdot)\mathrm{d}W(t), \tag{4.44}$$

where $\mu(\cdot, \cdot)$ is the drift term and $\sigma_I(S(t), t, \cdot)$ represents the *instantaneous* volatility that follows some stochastic process, possibly depending on some function of $S(t)$ and t and other possible variables. The local variance $\tilde{\sigma}^2_{K,T}(S(t), t)$ is defined as the risk-neutral expectation of the square instantaneous volatility conditional on $S(T) = K$ and the time t information $\mathcal{F}_t$

$$\tilde{\sigma}^2_{K,T}(S(t), t) = \mathrm{E}^Q\left[\sigma^2_I(S(t), t, \cdot)|S(T) = K, \mathcal{F}_t\right], \tag{4.45}$$

where E^Q is the expectation operator under the risk neutral measure Q. Then, the LV is given by

$$\tilde{\sigma}_{K,T}(S(t), t) = \sqrt{\tilde{\sigma}^2_{K,T}(S(t), t)}. \tag{4.46}$$

This definition of local volatility assumes that the market's view on future volatility can be expressed by the expectation of the instantaneous volatility so that all sources of risk from the stochastic volatility can be integrated out. Therefore, if the instantaneous volatility is deterministic in spot price and time, i.e. $\sigma_I(S(t), t, \cdot) = \sigma_I(S(t), t)$, both concepts, instantaneous and local variance, coincide

$$\begin{aligned} \tilde{\sigma}^2_{K,T}(S(t), t) &= \mathrm{E}^Q\left[\sigma^2_I(S(t), t, \cdot)|S(T) = K, \mathcal{F}_t\right] \\ &= \mathrm{E}^Q\left[\sigma^2_I(S(t), t)|S(T) = K, \mathcal{F}_t\right] = \sigma^2_I(K, T). \end{aligned} \tag{4.47}$$

By this assumption, one can get an estimation of local volatility modeling instantaneous volatility and vice versa.
A good estimation of local volatility can be obtained by the so called *Dupire formula* that can be derived applying the Itô's Lemma to Eq. (4.16)

$$\tilde{\sigma}^2_{K,T}(S(t), t) = 2\frac{\frac{\partial C(t,K,T)}{\partial T} + \delta C(t, K, T) + rK\frac{\partial C(t,K,T)}{\partial K}}{K^2 \frac{\partial^2 C(t,K,T)}{,} \partial K^2} \tag{4.48}$$

that can also be expressed as a function of the IV

$$\tilde{\sigma}^2_{K,T}(S(t), t) = \frac{\frac{\sigma}{(T-t)} + 2\frac{\partial\sigma}{\partial T} + 2Kr\frac{\partial\sigma}{\partial K}}{K^2\left\{\frac{1}{K^2\sigma(T-t)} + 2\frac{d_1}{K\sigma\sqrt{T-t}}\frac{\partial\sigma}{\partial K} + \frac{d_1 d_2}{\sigma}(\frac{\partial\sigma}{\partial K})^2 + \frac{\partial^2\sigma}{\partial K^2}\right\}}, \tag{4.49}$$

where d_1, d_2 are the parameters of the BS solution. Eq. (4.49) shows the relation between IV (σ) and LV ($\tilde{\sigma}$) and, because of Eq. (4.47), with instantaneous volatility (σ_I). So, using this relation, one can develop a model coherent with VS effect using an instantaneous volatility that is a function of time and the spot price level. Moreover Eq. (4.49) gives a exhaustive criterion to generate this function starting from the implied volatility that can be directly observed on the market. We want to stress that this approach does not introduce any additional source of stochastic noise to model the volatility and so, it does not require additional assumptions on the market price of risk.
Obvioulsy, there is still the problem that the implied volatility is not a known continuous function of strike and maturity, but only known at certain points. Then, to get the local volatility function from Eq. (4.49), some method has to be used to interpolate and extrapolate the given IV points to get a continuous differentiable function that could describe the IVS. This problem will be the topic of the following Chapters, where we show how methods that come from Statistical Physics can be applied to get a suitable description of the IVS.

4.7 Stochastic Volatility Models

Local volatility models try to stay close to the Black-Schöles model by introducing more flexibility into the volatility and this is one of the main reasons of criticism. A different approach is followed by the so called Stochastic Volatility Models that try to obtain a closer description of actual market data introducing a new source of randomness related to the volatility. Since volatility is not a tradable asset, this implies that the market is incomplete, as such there is a whole family of risk-neutral pricing measures, unlike in the constant volatility case when there is only one. So the description of the market is not unique as it depends on the choice if a parameter that cannot be determined directly from the market. In these cases the market is called incomplete. The literature about stochastic volatility models is really wide [35, 36, 37], here we just want to summarize the results of one of the most important stochastic volatility model that is applied by many financial institutions and it seems to be one of the candidate to replace, in the future, the BS model: the *Heston model* [36]. According to this model, asset price dynamic can be described by the following system of equations

$$\begin{array}{rcl} \mathrm{d}S & = & \mu S \mathrm{d}t + \sigma S \mathrm{d}W_1 \\ \mathrm{d}\sigma^2 & = & \gamma(\theta - \sigma^2)\mathrm{d}t + k\sigma \mathrm{d}W_2, \end{array} \tag{4.50}$$

where μ, γ, θ, k are constants and the parameters of the model, $\mathrm{d}W_1, \mathrm{d}W_2$ are the two Wiener's increments that have correlation ρ and, as usual, $S(t)$ represents the price of the asset at time t and σ is the stochastic volatility. In this case we are assuming that the volatility follows a mean-reverting CIR process (See Sect. 3.5.7) where θ is the long-time mean of σ^2, γ is the rate of relaxation of the mean, and k represents the variance of σ^2. By this process we are implicitly supposing that the stationary probability distribution of the variance is given by the gamma distribution

$$\Pi(\sigma^2) = \frac{\alpha^\alpha}{\Gamma(\alpha)} \frac{\sigma^{2(\alpha-1)}}{\theta^\alpha} e^{\alpha\sigma^2/\theta}, \tag{4.51}$$

where $\alpha = 2\gamma\theta/k^2$. By the definition of an imaginary portfolio similar to the one described for the derivation of the BS formula, but using a combination of two different options, it is possible to obtain the following equation:

$$\frac{\partial C}{\partial t} + rS\frac{\partial C}{\partial S} + (\gamma(\theta - \sigma^2) - \lambda(S, \sigma^2, t))\frac{\partial C}{\partial \sigma^2} + \mathcal{D}C - rC = 0, \tag{4.52}$$

where r is the risk free rate, $\lambda(S, \sigma^2, t)$ is a generic function that represents the *price of the volatility risk* and $\mathcal{D}$ is an operator defined as

$$\mathcal{D} = \frac{1}{2}\sigma^2 S^2 \frac{\partial^2}{\partial S^2} + \frac{1}{2}k^2\sigma^2 \frac{\partial^2}{\partial(\sigma^2)^2} + \sigma^2 S k \rho \frac{\partial^2}{\partial S \partial \sigma^2}. \tag{4.53}$$

In his paper, Heston redefined the price of the volatility risk as $\lambda(S, \sigma^2, t) = \lambda\sigma^2$, where λ is a constant of the problem and solved Eq. (4.52) guessing, by analogy with the BS solution, a solution of the form

$$C(S, \sigma, t) = SP_1 - Ke^{-r(T-t)}P_2, \tag{4.54}$$

where P_1, P_2 are two unknown functions to be determined. In [36] it is shown that it is possible to get a semi-closed analytical solution of the problem given by:

$$P_j(x, \sigma^2, T; \ln(K)) = \frac{1}{2} + \frac{1}{\pi}\int_0^{+\infty} Re\left[\frac{e^{-i\phi\ln(K)}f_j(x, v, T; \phi)}{i\phi}\right]\mathrm{d}\phi, \tag{4.55}$$

where $x = \ln(S)$, K is the strike price, T the time to maturity, $j = 1, 2$ and

$$f_j(x, \sigma^2, t; \phi) = \exp\left(C(T-t, \phi) + D(T-t, \phi)v + i\phi x\right), \tag{4.56}$$

where

$$C(T-t, \phi) = r\phi i(T-t) + \frac{a}{k^2}\left\{(b_j - \rho k\phi i + d)(T-t) - 2\ln\left[\frac{1-ge^{d(T-t)}}{1-g}\right]\right\}, \tag{4.57}$$

$$D(T-t, \phi) = \frac{b_j - \rho k\phi i + d}{k^2}\left[\frac{1-e^{d(T-t)}}{1-ge^{d(T-t)}}\right] \tag{4.58}$$

and

$$g = \frac{b_j - \rho k\phi i + d}{b_j - \rho k\phi i - d}, \tag{4.59}$$

$$d = \sqrt{(\rho k\phi i - b_j)^2 - k^2(2u_j\phi i - \phi^2)}, \tag{4.60}$$

$$u_1 = 1/2,\ u_2 = -1/2,\ a = \gamma\theta,\ b_1 = \gamma + \lambda - \rho k,\ b_2 = \gamma + \lambda. \tag{4.61}$$

To estimate the fair value of a call option starting from this model, one just needs to calibrate the six parameters of the model $(\rho, \gamma, \theta, k, r, \lambda)$ and to evaluate the integral in Eq. (4.55) using numerical approximations. A description of how to calibrate Heston model can be found, for example in [38]. Differences between Heston and BS pricing model are essentially due to the different probability distributions implied by the two models. In particular in [39] is shown that the returns PDF of Heston model has exponential tails and this fact can better describe actual returns distribution with a time lag greater than one day. The two main parameters that govern the differences in the pricing are ρ and k. As shown in [36] the first has an important effect on the skewness of the distribution so that a positive skewness is associated with increases in the prices of OTM options relative to ITM options; the latter increases the kurtosis of the spot return, increasing, as a consequence, the value of the options OTM with respect to value obtained by the BS model. The main drawbacks related to the Heston model are essentially due to the difficulty of the calibration process (in particular the evaluation of the complex integral in Eq. (4.55) could be laborious) and to the introduction of a second stochastic process that could have important consequences on the hedging and risk management activities.

4.8 Conclusions

In this Chapter we introduced the main concepts about option pricing modeling. After giving some definitions about derivatives and in particular about options, we introduced the most famous model for the option pricing: the Black-Schöles model. We derived its formula to price and European call (and put) option by the construction of an imaginary portfolio and, equivalently using the so called risk neutral approach. In this way, we showed which are basic concepts related to this model and in particular how to move into a risk free world deleting markets noise by using the delta hedging approach. Then we presented the main concept of this work: the implied volatility. We showed its statistical properties and, in particular, the smile effect, with an intuitive and quantitative interpretation. Finally we focus on the main problems related to the implied volatility and on some possible extension of the BS model to fix them: the local volatility and the stochastic volatility models. We showed that the first approach is closer to the BS framework and can be developed requiring a good fitting procedure of the IVS. The second approach introduces an additional stochastic equation for the description of the price dynamics obtaining a good description of realty starting from a more fundamental point of view. The price to pay is a more complex calibration procedure, losing the completeness of the original Black-Schöles model since the stochastic volatility in asset prices cannot be traded. In the following Chapters we will focus on the first approach and in particular we will deal with the problem of getting a good fitting procedure of the IVS that represents the basis for the application of local volatility models.

Chapter 5

Adiabaticity Conditions for Volatility Smile in Black-Schöles Pricing Model

5.1 Introduction

In this Chapter we show how some concepts, that come from physics can be applied to financial world to obtain new models with concrete applications. These models, presented in [40], were developed starting from an unconventional point of view that finds its inspiration from physics world and this fact confirms that a relationship between physical and financial model should exist and can be exploited by both the areas. On the other hand we want to stress that this relationship can be used as a "source of inspiration" from a methodological point of view and that is unrealistic, in our opinion, to look at finance as a straightforward application of physical models.
In particular, in this Chapter we derive the distribution function of financial returns using the Black-Schöles expression for the call pricing with volatility in the form of a volatility smile. We show that this approach based on a volatility smile leads to relative minima for the distribution function ("bad" probabilities) never observed in real data and, in the worst cases, negative probabilities. We show that these undesirable effects can be eliminated by requiring "adiabatic" conditions on the volatility smile. Finally, we formulate an algorithm that can be used in practical situations to get a good fit of the volatility smile, avoiding negative probabilities and, as a consequence, the arbitrage opportunities.

5.2 The Inverse Problem for the Stock Price Distribution Function and Arbitrage Opportunities

As already mentioned in the previous Chapters, one of the simplest "products" on the derivative financial market is the European call (put) option. Considering the risk neutral approach, the price of the European call option, $C \equiv C(S_T, K, T, r)$, is defined by

$$C = e^{-rT} \int_K^\infty (S_T - K) P(S_T) dS_T, \tag{5.1}$$

where S_T is the stock price at time $t = T$, K is the strike price of the option, T is the expiration time (time to maturity) of the option, r is the interest rate and $P(S_T) \geq 0$ is the probability density function (PDF) of the stock prices in a "risk-neutral world" ($\int_0^\infty P(S_T) dS_T = 1$).

Eq. (5.1) is too general since it does not place any restrictions on the underlying stock price distribution function, $P(S_T)$. To calculate explicitly the option price, C, using Eq. (5.1), one must know the PDF, $P(S_T)$. Consequently, one must make some assumptions about the stock prices, as already seen in the previous Chapters. An important achievement in the theory of option pricing is the Black-Schöles (BS) theory which gives analytic solutions for the European call and put options [21].

In particular, for the European call option, a solution of the BS equation is given by Eq. (5.1), if one assumes for the PDF, $P(S_T)$, a log-normal distribution,

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2(T-t)}} \exp\left[-\frac{(x + \sigma^2(T-t)/2)^2}{2\sigma^2(T-t)}\right], \tag{5.2}$$

where $x = \ln(S_T/S(t)) - r(T-t)$ is the logarithmic return deprived of the risk-free component, $S(t)$ is the stock price at time t and σ is the stock price volatility. For seek of simplicity, in the following we consider $t = 0$ and we define $S_0 \equiv S(t = 0)$. Substituting Eq. (5.2) in Eq. (5.1) an explicit expression for the price of the European call option which satisfies the BS equation [21] is obtained,

$$C^{\text{BS}} = S_0 N(d_1) - K e^{-rT} N(d_2), \tag{5.3}$$

where

$$\begin{aligned} d_1 &= \frac{\ln(S_0/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}, \\ d_2 &= d_1 - \sigma\sqrt{T}, \\ N(x) &= \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} dz\; e^{-z^2/2}. \end{aligned} \tag{5.4}$$

There are some problems with the expressions for C given by Eqs. (5.1-5.3). Indeed, one can derive *any* option price from Eq. (5.1), using different assumptions about the PDF, $P(S_T)$. To derive from Eq. (5.1) a result for C which will even approximately coincide with the real market price, C^{M}, one must specify a distribution function for *future* stock prices, $P(S_T)$. On the other hand, the expression given by Eq. (5.3) is (a) too specific, and (b) derived using rather strong restrictions. Namely, the BS model does not account for correlations of returns, x and, moreover, the volatility, σ, and the interest rate, r, are not well-defined parameters (given the actual data). As a result, the expression, C^{BS}, often does not coincide (even approximately) with the corresponding market option price, C^{M}. Useful approaches have been developed which partially solve the problems mentioned above.

We shall mention here one analysis which is related to that presented below. This analysis deals with building "implied trees" [41, 42, 33, 43, 44]. There are many variations of this approach, but the main idea is based on the solution of the inverse problem: a search for a stock price model that corresponds to the real market prices of options, C^{M}. A more restricted problem is to search for a stock price model that effectively deals with the volatility smile. In this case, one starts with the BS formula (5.3), (even for American options) but instead of choosing a fixed volatility, σ = constant, one uses the dependence, $\sigma = \sigma(K)$ (volatility smile). To some extent, this dependence corresponds to the "real behavior" of the volatility, σ, if one wants to use Eq. (5.3) as the "zeroth approximation" for option pricing.

There are still some problems with this implied volatility. For example, the corresponding "implied" stock prices, S_T, can have "bad" (negative) probabilities which must be eliminated (Fig. (5.1)).

Despite the simplicity of the calibration procedure, a solution to this kind of problem for trees generation was proposed in [45, 46]. In general, another way to avoid the problem of "bad" probabilities is to find some constraint on the general shape of $\sigma(K)$ via a nonparametric or parametric specification of the functional form of the volatility smile. In the first case, volatility smiles are computed requiring some smoothness conditions derived from arbitrage-

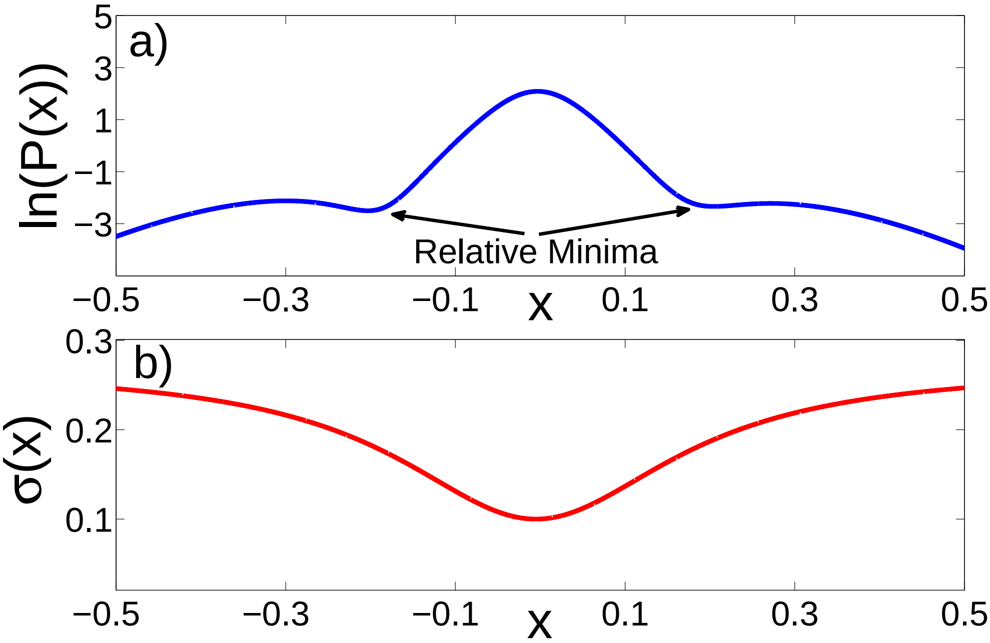

Figure 5.1: a) Log-distribution given by Eq. (5.12) for the volatility smile given by Eq. (5.18), with the following parameters: $g = 0.1, T = 0.5, n = 0.04, \chi = 2.7$. b) Volatility smile related to the PDF shown in a).

free considerations [47, 48, 49]. This kind of algorithms has the advantage to be based on more general assumptions on volatility smile and PDFs of return but, as a consequence, they increase the complexity of the problem and the computational time of the calibration. On the contrary, using a parametric approach, one needs to specify the expression of the volatility smile to get a parametric restriction on the implied volatility space that can be used straightforwardly in the fitting procedure, generally decreasing the computational time for calibration.

In this Chapter, we discuss the inverse problem for Eq. (5.1) using the following parametric approach. First, using Eq. (5.1), we build the PDF for future stocks prices and returns from the empirical data for the market option prices, $C(K)$. Second, we build the returns PDF using for $C(K)$ a BS expression, C^{BS}, with a volatility, $\sigma = \sigma(K)$, in the form of the volatility smile. In particular, we show that the condition of the absence for relative minima in the PDF of returns, (or elimination of "bad" probabilities) leads to the condition of "adiabaticity" for the volatility smile. The term "adiabatic" comes from statistical physics and is related to the slowness of the variation of a parameter λ that specifies the properties of a system or an external field. In fact from a physical point of view, it can be shown that if in a system one introduces a small perturbation (λ) compared to the characteristics period of the motion T, namely:

$$T\frac{d\lambda}{dt} \ll \lambda \tag{5.5}$$

the rate of the change of the energy of the system will be also small [50]. In the same spirit we assume that our parameter λ is represented by the implied volatility σ and we look for an adiabatic condition that leads to a small perturbation of the implied Gaussian distribution that derives from the BS model with constant volatility, eliminating the problem of negative probabilities.

First of all, we derive an explicit expression for the PDF for the future stock prices, $P(S_T)$ and for logarithmic returns $P(x)$. In Eq. (5.1) the PDF, $P(S_T)$, can be rather arbitrary but it is natural to assume that $P(S_T)$ does not depend on the strike price, K. According to Eq. (5.1), the option price, C, is expressed explicitly through the strike price, K. Differentiating C in Eq. (5.1) twice with respect to K, we have [51]

$$P(S_T) = e^{rT}\frac{\partial^2 C(K)}{\partial K^2}\Big|_{K=S_T}. \tag{5.6}$$

In Eq. (5.6), we indicate only the dependence $C(K)$ in the option prices. In particular, applying Eq. (5.6) to C^{BS} given in Eq. (5.3) we derive a distribution function, $P^{\mathrm{BS}}(S_T)$, which we present in the form

$$\begin{aligned} P^{\mathrm{BS}}(S_T) \equiv e^{rT}\frac{\partial^2 C^{\mathrm{BS}}(K)}{\partial K^2}\Big|_{K=S_T} &= \frac{1}{\sqrt{2\pi\sigma^2 T}S_T}\times \\ \times \exp\left(-\frac{(\ln(S_T/S_0)-(rT-\sigma^2/2T))^2}{2\sigma^2 T}\right). \end{aligned} \tag{5.7}$$

Analogously, the distribution of returns for the Black-Schöles model is Gaussian, as expected:

$$P^{\mathrm{BS}}(x) = \frac{1}{\sqrt{2\pi\sigma^2 T}}\exp\left[-\frac{(x+x_0)^2}{2\sigma^2 T}\right], \tag{5.8}$$

where $x_0 = \sigma^2/2T$.

We can try to consider the inverse problem substituting the dependence, $\sigma = \sigma(K)$, in Eq. (5.3) and evaluating the distribution of future stocks price and returns, applying Eq. (5.6). After differentiation we get

$$P(S_T) = \frac{F(S_T; S_0, r, T, \sigma)}{\sqrt{2\pi\sigma^2 T} S_T} \times$$
$$\times \exp\left[-\frac{(\ln(S_T/S_0) - (rT - x_0))^2}{2\sigma^2 T}\right], \tag{5.9}$$

where we defined

$$F(S_T; S_0, r, T, \sigma) = \left[1 + S_T \frac{\dot{\sigma}}{\sigma}(rT - \ln(S_T/S_0))\right]^2$$
$$-\frac{(\dot{\sigma}\sigma T S_T)^2}{4} + \dot{\sigma}\sigma T S_T + S_T^2 \sigma \ddot{\sigma} T,$$
$$\dot{\sigma} = \left.\frac{\partial \sigma}{\partial K}\right|_{K=S_T}, \tag{5.10}$$
$$\ddot{\sigma} = \left.\frac{\partial^2 \sigma}{\partial K^2}\right|_{K=S_T}.$$

Obviously, we can get the expression for the distribution of returns by a simple change of variable

$$x \equiv \ln\left.\left(\frac{K}{S_0}\right)\right|_{K=S_T} - rT, \tag{5.11}$$

so that

$$P_\sigma(x) = \frac{1}{\sqrt{2\pi\sigma^2 T}} \exp\left[-\frac{(x + x_0)^2}{2\sigma^2 T}\right] F(x; T, \sigma), \tag{5.12}$$

where, in a similar way, we have defined:

$$F(x; T, \sigma) = (1 - \frac{\sigma'}{\sigma} x)^2 - \frac{(\sigma' \sigma T)^2}{4} + \sigma\sigma'' T,$$
$$\sigma' = \frac{\partial \sigma}{\partial x}, \tag{5.13}$$
$$\sigma'' = \frac{\partial^2 \sigma}{\partial x^2}.$$

From Eq. (5.13) it is clear that if σ is constant Eq. (5.9-5.12) are the distributions for the standard Black-Schöles model; we will call them "zeroth-approximation" distributions. If $\sigma \neq$ const, the term $F(x; r, T, \sigma)$ could

"perturb" the relative zero approximation (Gaussian) giving rise to distributions that cannot fit real data. As we will show here, it is possible to get distributions with relative minima (not observed in real returns distributions) and, in the worst case, negative probability.
The negative probability aspect in option pricing theory is particularly relevant because it is related to arbitrage opportunity, namely the opportunity to get a risk free profit. In fact even if from a mathematical point of view negative probabilities are impossible, from a financial perspective it can be shown from arbitrage theorem [14] that there is a strong relation between negative probability and arbitrage opportunity. A simple example of this relation is given considering the discrete approximation of the second derivative in Eq. (5.6):

$$\frac{\partial^2 C^{\mathrm{BS}}(K)}{\partial K^2} = \frac{C(K) - 2C(K+\Delta) + C(K+2\Delta)}{\Delta^2}, \tag{5.14}$$

where Δ represents the discrete increment. At the numerator on the right hand of the Eq. (5.14) we have an imaginary portfolio of three call options of different strikes: $K, K+\Delta, K+2\Delta$. Considering that at the time to maturity T the payoff of a call option is $\max(S(T)-K, 0)$, it is clear that at that time the value of the portfolio is greater or equal to zero (See Fig. 5.2).

If for some $t < T$ the implied distribution is negative as well the value of the imaginary portfolio, we know for sure that this portfolio will increment its value when it reaches the time to maturity passing from a negative value to a positive one. This gain can be greater to the one allowed by the risk free rate, so it is, in principle, a simple example of static arbitrage opportunity.

5.3 The Volatility Smile: Real Market Data

As already stressed in the previous Chapter, typically, traders on option markets and practitioners consider the volatility as a parameter that can be adjusted taking into account the inverse problem given by Eq. (5.3) and the real price of call and put options. In this way a more reliable value of the volatility (implied volatility) can be obtained and it can be used to price more complex options for which analytical solutions are not available. In this Chapter, we focus our attention on the volatility smile of foreign currency options and we neglect any skew effects [52]. To perform our analysis we consider the volatility smile as a function of the Δ of the option (defined by Eq. (5.15)), the time to maturity, T, and the currency considered. We consider specific days, for which volatility is not affected by the skew effect, and we use Bloomberg as data provider. In the BS model, the Δ of a call

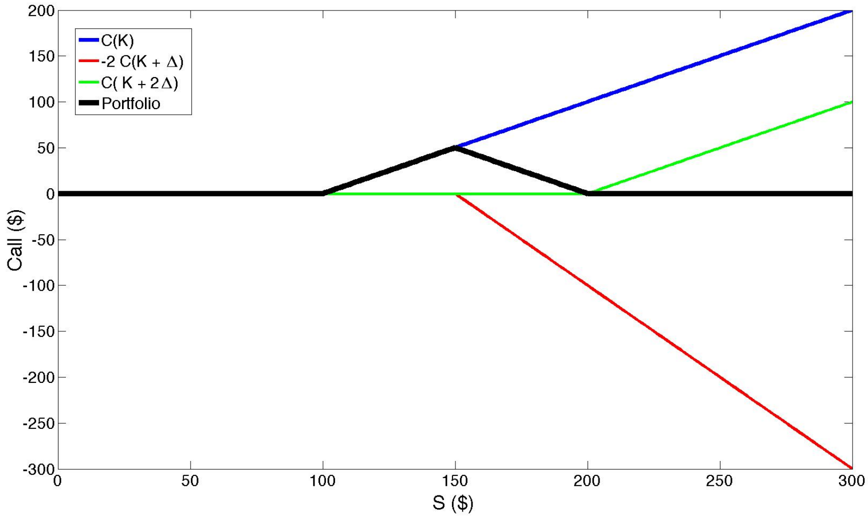

Figure 5.2: We show the payoff of the imaginary portfolio (black) given by Eq. (5.14) and its three call options (gree, blue, red). The value of the portfolio at the maturity is always greater or equal to zero.

option is defined as

$$\Delta = \frac{\partial C}{\partial S_T} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_1} dz \; e^{-z^2/2}. \tag{5.15}$$

Inverting this relation it is possible to get an expression for x:

$$x = \sigma^2/2T - \sigma\sqrt{T}\Phi^{-1}(\Delta), \tag{5.16}$$

where $\Phi^{-1}(x)$ is the inverse of the error function

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} \; dt. \tag{5.17}$$

In Fig. 5.3 we show an example of volatility smile in terms of our variables and a suitable fit given by the function

$$\sigma(x) = g \left[1 + (\chi - 1) \frac{(x + g^2 T/2)^2}{(x + g^2 T/2)^2 + n} \right], \tag{5.18}$$

where g, χ, n are fitting parameters. In this case, g represents the minimum of the volatility smile, $\sqrt{n}$ is the half width at the half height, while $g(\chi - 1)$

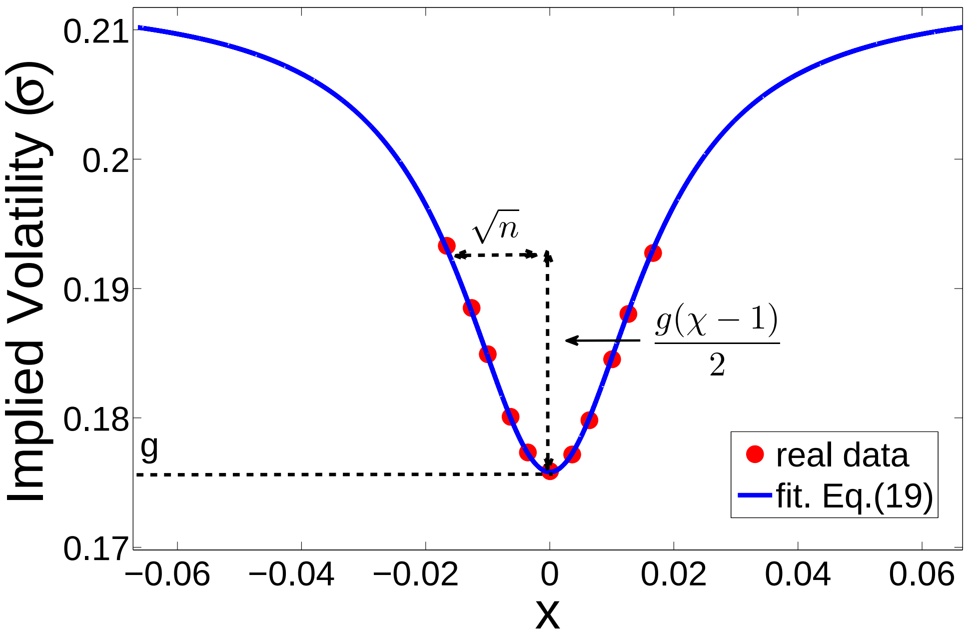

Figure 5.3: Typical volatility smile and the relative fit obtained with Eq. (5.18). The parameters of the fit are: $g = 0.1758(5)$, $\chi = 1.20(9)$, $n = 0.00030(9)$. We get the real data using Bloomberg provider and they refer to the AUDUSD currency with time to maturity $T = 1/365$ years.

represents the height of the smile. In particular χ is the ratio between the limiting value of σ and g as x approaches ∞. In this way the variation of σ is bounded between g and $g\chi$. If we consider the volatility smile as an effect due to the fat-tails of returns distribution, we could expect that the minimum of the smile should be found in correspondence of the center of the returns distribution, where there is no need to change the value of implied volatility because of the fat tails effect. Considering also Eq. (5.8), it follows that the average value of x is

$$\langle x \rangle = -\frac{\sigma^2 T}{2}, \tag{5.19}$$

so one expects that the minimum of the implied volatility occurs at $x = -g^2T/2$ as required by our fitting function.

Repeating many times the interpolation procedure considering different values for T and currencies (Table 5.1), we can determine typical parameters that can fit a wide range of volatility smiles; in the following we will use this information to check our results. Let us notice that the following relation

Currency	**Maturities (days)**	**Date**
AUDUSD, EURCHF	$1, 7, 14, 21, 30$	21/10/2009
EURGBP, EURJPY	$60, 90, 120, 180, 270$	01/02/2010
EURUSD, GBPUSD	$360, 540, 720, 1080$	01/04/2010
USDCAD, USDCHF		

Table 5.1: Dataset for Volatility Smile

between n, g and T holds

$$n \propto Tg^2, \tag{5.20}$$

as shown in Fig. 5.4. This gives a scaling rule that can be used to determine the range of n, fixing T and g. Our intuitive explanation of this equation is really simple and it is related to the fact that the PDF of returns is not Gaussian but exhibits fat/exponential tails. Indeed, while the term $\sqrt{n}$ gives the order of magnitude of the volatility amplitude, $g\sqrt{T}$ represents the minimum of the implied volatility (which can be considered as the unperturbed standard deviation of the PDF of returns). Therefore Eq. (5.20) suggests that when x is about $2-3$ times the standard deviation of the returns distribution (namely in the tails) the implied volatility should be increased to fatten up the PDF of returns.

5.4 First Approximation of the Volatility Smile: the Squared Well

In this Section we show qualitatively the reason why there is a relative minimum in the returns distribution and why an adiabatic approach can describe the problem of avoiding these "bad" probabilities. To keep the problem simple, we consider, as a first step, a volatility smile modelled by a squared well defined as follows

$$\sigma(x) = \begin{cases} \sigma_1 & \text{for} \quad |x| < x_1 \\ \sigma_2 & \text{otherwise} \end{cases}, \tag{5.21}$$

where $\sigma_2 > \sigma_1$ and x_1 are positive constants. See Fig. 5.5a,b. The distribution functions corresponding to two values of σ_i, $i = 1, 2$ are shown in Fig. 5.5c,d. Indicating as $\pm x_1^c$ the abscissa of the intersections between the

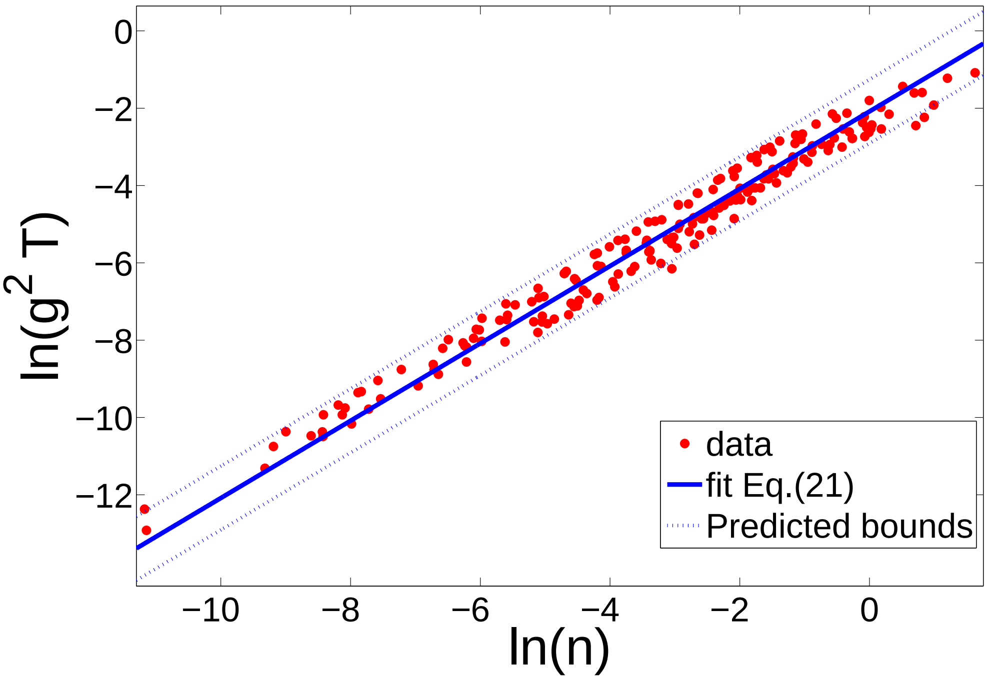

Figure 5.4: Relation between the parameters n, g, T. We fit 72 volatility symmetrical smiles (Bloomberg) considering different currency (EURUSD, AUDUSD, EURCHF, EURGBP, EURJPY, GBPJPY, GBPUSD, USDCAD, USDCHF, USDJPY) and time to maturity (1 day, 1-3 weeks, 1, 2, 3, 4, 6, 9 months, 1, 1.5, 2, 3, 4, 5 years) with the function (5.18). The data are from the days 21/10/2009 and 01/02/2010. We also show the best linear fit $\ln(g^2T) = \ln(n) + c$, where $c = -1.95(12)$.

two distributions, it is clear that a sufficient condition for avoiding spurious minima is $x_1 < x_1^c$. A rough estimation of x_1^c, ignoring the term x_0, usually small, is

$$x_1^c = \sigma_1\sqrt{T}\sqrt{\frac{2\chi^2\ln\chi}{\chi^2-1}}, \tag{5.22}$$

where $\chi = \sigma_2/\sigma_1$. Therefore, a sufficient condition to avoid minima in the PDF is to use, as a fitting function a square well depending on the parameters x_1, σ_1, σ_2, such that $x_1 < x_1^c$. Therefore, a standard fitting procedure of the volatility smile with a square well, constrained by the condition $x_1 < x_1^c$ solves the problem of avoiding spurious minima, even if it is very rough.

If we consider a volatility smile with a continuous variation from σ_1 to σ_2,

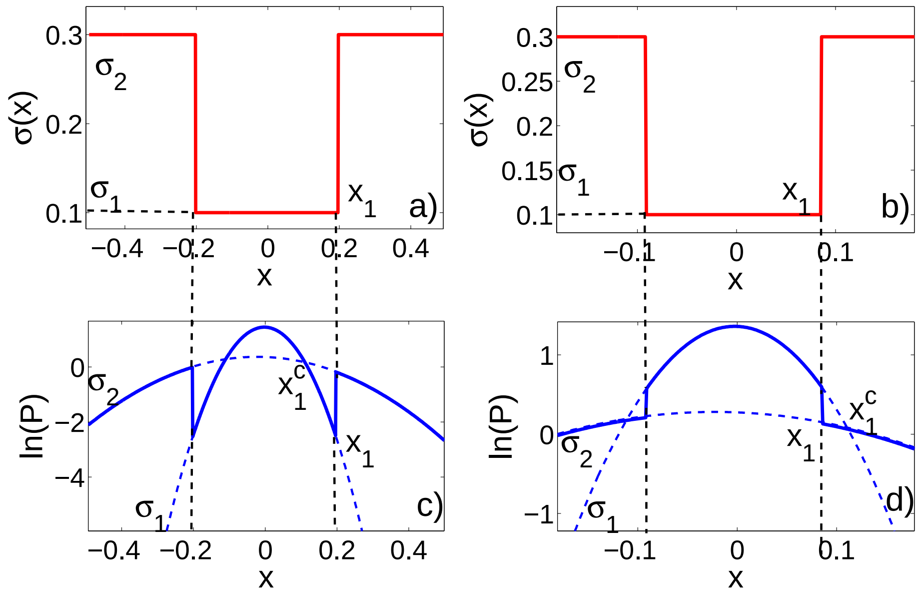

Figure 5.5: a,b) Discontinuous squared well as a fit of the volatility smile. c,d) The corresponding PDF with (c) or without (d) spurious minima.

we can get, instead of a discontinuity, the onset of a relative minimum. The latter can be avoided if the variation between σ_1 and σ_2 is slow enough so that the connection between the two PDFs takes place keeping constant the sign of the first derivative of the distribution during the whole transition. So there will be a critical "speed" of the transition that will generate zero derivative points which will not correspond to the maximum of the distribution. In this case the variable related to the time is x_1, while $\chi = \sigma_2/\sigma_1$ can be identified as a "distance". To be more precise, one should consider that for $x < x_1^{\mathrm{c}}$, $P_{\sigma_1}(x) \geq P_{\sigma_2}(x)$, so that the effective "time" should be: $x_1 - x_1^{\mathrm{c}}$.

We will define an adiabatic transition as one whose passage from σ_1 to σ_2 is sufficiently slow, so that relative minima in the distribution are not generated. In this same spirit, we define the critical adiabatic parameter $\chi_{\mathrm{c}}(n, g, T)$ as the minimum value of χ that generates a minimum in the distribution, Eq. (5.12). This effect has been shown in Fig. (5.6), modelling the volatility smile using Eq. (5.18), so that $\sigma_1 = g$, $\sigma_2 = g\chi$. We fix the parameters g, T and we vary n and χ, seeking relative minima in the PDF. The lines in Fig. 5.6, obtained respectively at fixed T (a) and fixed g (b), divide the plane of parameters into two regions: to the left of the lines the

PDF has spurious minima, while this does not happen in the region to the right of it. It is then clear that for a given set of fixed parameters (g, T), there is a relation between χ and n that allow one to obtain a PDF without minima (minima are not observed in real data). The main goal of this paper is to determine a simple relation that determines whether or not the parameters of a volatility smile fit are consistent with real returns distribution and if they could give a reliable option pricing.

5.5 Numerical and Theoretical Results

In this Section we show our numerical and theoretical results about the relation between the set of parameters n, g, T and the critical adiabatic parameter χ_c. Using a numerical simulation, we kept fixed n, g, T and we continuously increased the parameter χ until we found a zero-derivative point for some $x \neq -g^2T/2$. In this way we could determine numerically the critical χ_c. We repeated this approach for a wide range of the parameters values, as shown in Table 5.2, where we used the parameter $\rho = n/g^2T$ instead of n, due to

	min	**max**
g	0.03	0.5
ρ	2.5	10
T (years)	1/365	4

Table 5.2: Range of the parameters of the numerical simulations

the scaling relation (5.20). In order to obtain the relation $\chi_c = f_T(n, g)$, we use the following fit function

$$f_T(n, g) = \alpha \left(\frac{n}{g^2T} \right)^{\beta} - \gamma \sqrt{T} g \left(\frac{n}{g^2T} \right)^{\delta}. \tag{5.23}$$

This has been obtained assuming that the value of the critical parameter χ_c depends on the rescaled "time" of the transition (in our model given by ρ). We also consider a further term $\gamma\sqrt{T} g \rho^{\delta}$ to take into account the time correction, x_c, as explained in Sec. (5.4). In this case we make explicit the dependence of the time correction on T and g as suggested by σ_2 in Eq. (5.22). In Eq. (5.23), $\alpha, \beta, \gamma, \delta$ are the fitting parameters whose values are given in Table 5.3.

In Fig. 5.7 we show the result of our fit for a few selected values of g and T.

The whole procedure can thus be summarized as follows :

α	1.4373 ± 0.0002
β	0.2787 ± 0.0006
γ	0.1738 ± 0.0002
δ	0.4683 ± 0.0006
mean squared errors	1×10^{-5}

Table 5.3: Fitting parameters and relative errors.

- The real volatility smile, usually given for a fixed T can be fitted by a function dependent on three parameters g, n, χ, as indicated in Eq. (5.18) and the optimal values $g^{\mathrm{opt}}, n^{\mathrm{opt}}, \chi^{\mathrm{opt}}$ are returned.
- The optimal values $g^{\mathrm{opt}}, n^{\mathrm{opt}}$ are inserted in Eq. (5.23), with $\alpha, \beta, \gamma, \delta$ given in Table II and a critical $\chi_{\mathrm{c}} = f_T(n^{\mathrm{opt}}, g^{\mathrm{opt}})$ obtained.
- If $\chi^{\mathrm{opt}} < \chi_{\mathrm{c}}$ then we know that relative minima in the PDF do not exist. Otherwise we should perform a fitting procedure to the volatility smile using Eq. (5.18), constrained by $\chi \leq \chi_{\mathrm{c}}$.

An example of the previous procedure has been shown in Fig. 5.8 where the PDF with unwanted minima and the "corrected " one is shown together with the corresponding fitting curve to the volatility smile. As one can see the price to pay in order to get a smooth PDF is very small: the two fitting curves for the real volatility smile are similar, but the PDF has, in the latter case a more realistic behavior.

5.6 Conclusions

In this Chapter, we started from the pricing equation of the Black-Schöles model for a European call and we analyzed a suitable generalization to include the volatility smile effect. Then we considered the inverse problem and the relative returns distribution, Eq. (5.12), varying the typical parameters of the volatility smile. We showed that, for some values of the parameters, it is possible to get relative minima in the returns distribution (bad distribution) that are never observed in real distributions. We demonstrated that bad distributions can be eliminated by requiring adiabatic constraints (intuitively justified with the example of the squared well) on the volatility smile and we gave a numerical formula to determine the value of the adiabatic critical parameter, χ_{c}. In this way we provide an easy-to-use tool to determine if a volatility smile fit is consistent with the general requirement for probabilities

$(P(x) \geq 0)$ and if it can generate a suitable returns distribution. This kind of problems are also discussed in [47, 48, 49] using a nonparametric approach. In these cases volatility smiles are computed requiring some smoothness conditions derived from arbitrage-free considerations. These algorithms have on one side the advantage to be based on more general assumptions on volatility smile and PDFs of return, while, on the other, the disadvantage to increase the complexity of the problem and the computational time of the calibration. Moreover our approach is based on the explicit knowledge of the fitting function of the volatility smile and the implied PDF of returns. In this way it is possible to get a globally parametric restriction on the implied volatility space that can be used straightforwardly in the fitting procedure, simplifying the calibration procedure. Interestingly, this restriction is derived using theoretical principles that come from physics.

A reliable estimate of the implied volatility has application in the risk management activities and in the pricing of exotic derivatives, where, in general, the implied volatility is an input of more complex models.

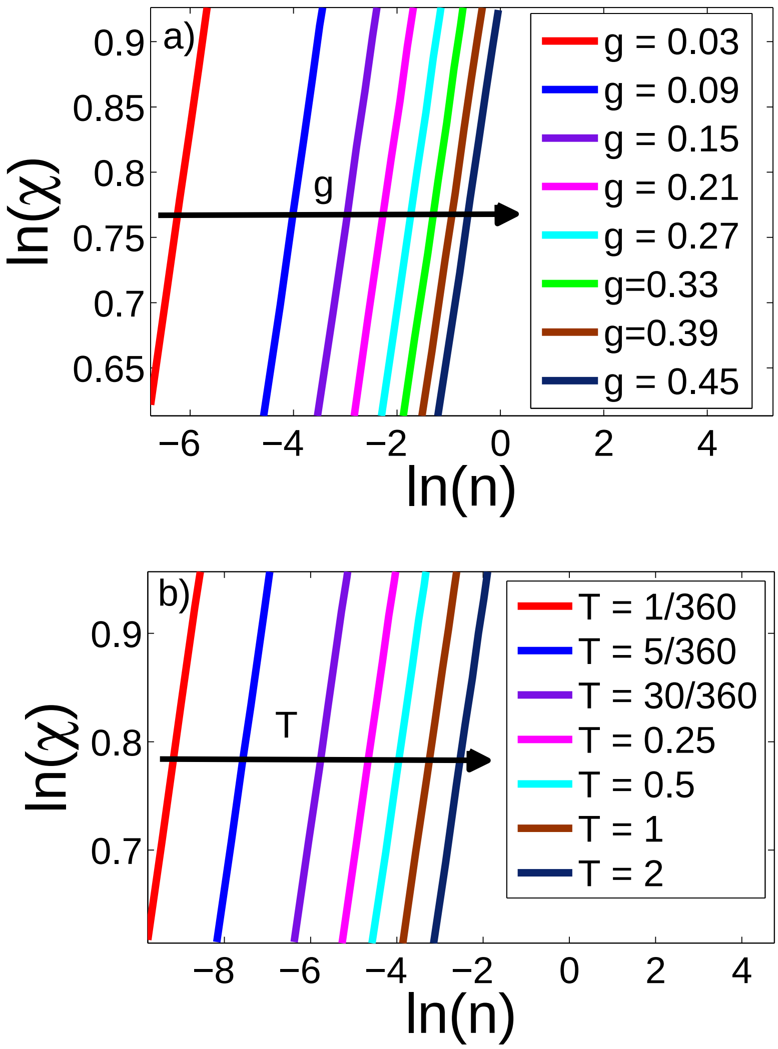

Figure 5.6: a) Critical ratio $\chi_c = \sigma_2/\sigma_1$ as a function of n for fixed $T = 0.5$, and different g as indicated in the legend. The arrow indicates the direction of growing g. In the region to the right of the lines the PDF do not have minima, while in the left hand region it has. b) The same as a) but for fixed $g = 0.1$ and different T values. The arrow indicates the direction of growing T.

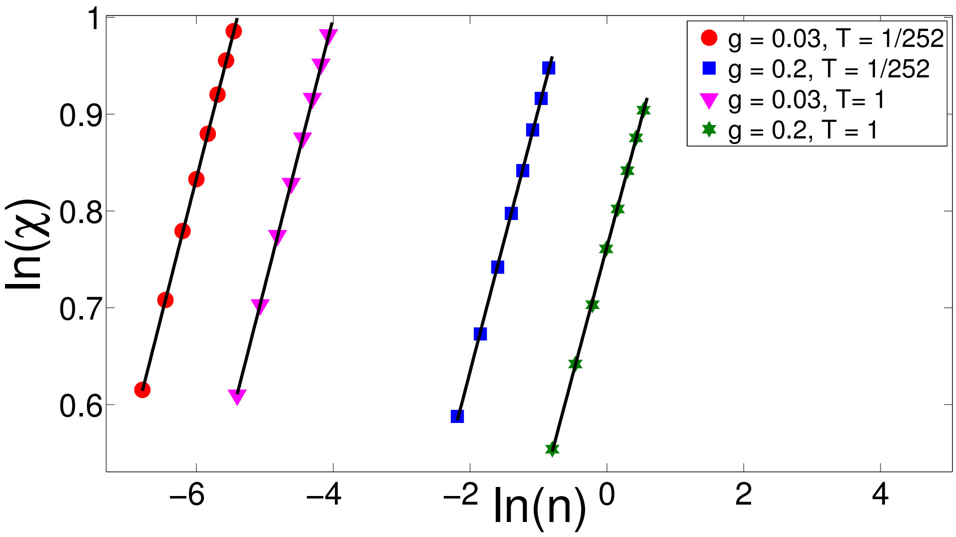

Figure 5.7: Critical adiabatic parameter as a function of n for few selected pairs of values of (g, T) as indicated in the legend. The points are numerical data, straight lines are the result of fitting procedure.

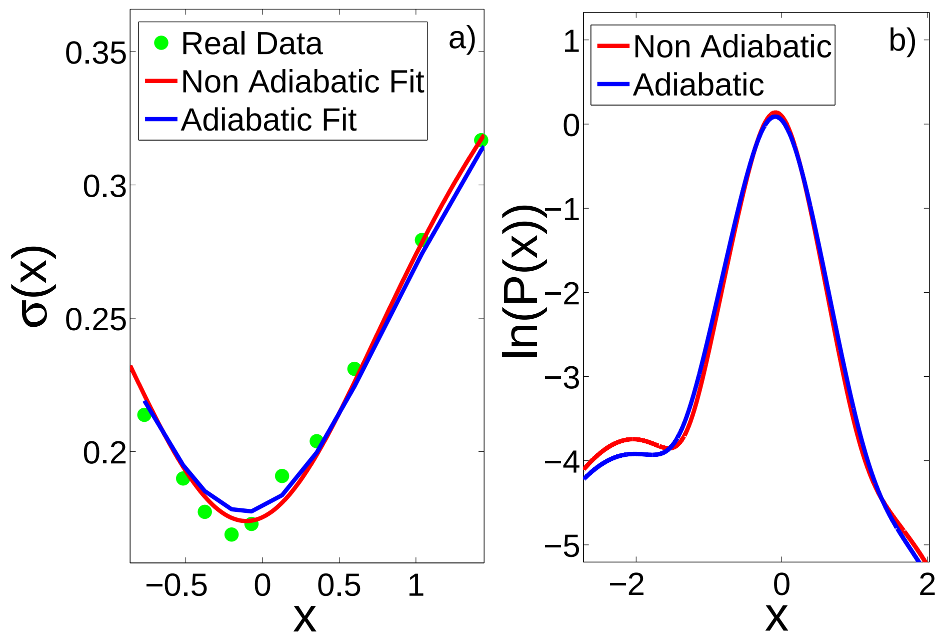

Figure 5.8: a) Volatility smile as a function the returns x. Dots indicate real data, red curve is the non adiabatic fit, while the blue one represents the adiabatic (constrained) fit. b) PDF of returns for the two curves indicated in a).

Chapter 6

Do your Volatility Smiles take care of extreme events?

6.1 Introduction

In the previous Chapter it is shown a new calibration procedure that can be obtained using an adiabatic approach to avoid arbitrage opportunities. The main idea is that there exists a parameter (represented by the implied volatility) that "perturbs slowly" the shape of probability density function (PDF) of returns, so controlling this parameter is possible to avoid arbitrage opportunities. In this Chapter we extend this adiabatic approach and we study we study how to characterize PDF of returns with a small perturbation of the parameter σ to get a suitable description of actual data [53], coherent from a theoretical point of view. In particular, in the following, it is shown the importance of this calibration procedure from the risk management point of view and its relevance in the risk estimation.

6.2 Volatility Smile: Analysis of Actual Market Data

In this Section, as already done in the previous Chapter, we focus our attention on the volatility smile (VS) of foreign currency options and we neglect the skew effect [52]. To perform our analysis we consider the same dataset (Table 6.1) described in the previous Chapter and the same fitting function:

$$\sigma(x) = g\left[1 + (\chi - 1)\frac{(x + g^2T/2)^2}{(x + g^2T/2)^2 + n}\right], \tag{6.1}$$

where g, χ, n are fitting parameters. As already mentioned, g represents the minimum of the volatility smile, $\sqrt{n}$ is the half width at the half height, while $g(\chi - 1)$ represents the height of the smile. In particular χ is the ratio between the limiting value of σ as x approaches ∞ and g . In this way the variation of σ is bounded between g and $g\chi$. In addiction we require the the minimum of the implied volatility to correspond to the average value of returns distribution $x = -g^2T/2$ as required by our fitting function.

Repeating the fitting procedure considering the volatility smile for different days, currencies and time to maturity T (Table 6.1),

Currency	**Maturities (days)**	**Date**
AUDUSD, EURCHF	$1, 7, 14, 21, 30$	21/10/2009
EURGBP, EURJPY	$60, 90, 120, 180, 270$	01/02/2010
EURUSD, GBPUSD	$360, 540, 720, 1080$	01/04/2010
USDCAD, USDCHF		

Table 6.1: Dataset for VS

as already observed in [40], the following relation between n, T, g holds:

$$\sqrt{n} = cg\sqrt{T}, \tag{6.2}$$

where $c = 2.65(28)$ is a fitting parameter.

6.3 Importance of VS in Risk Estimation

In the previous Chapter, we derived the analytical expression of the implied distribution of financial returns, considering the dependence $\sigma = \sigma(K)$:

$$P_\sigma(x) = \frac{1}{\sqrt{2\pi\sigma^2 T}} \exp\left[-\frac{(x + x_0)^2}{2\sigma^2 T}\right] F(x; T, \sigma), \tag{6.3}$$

where we have defined:

$$
\begin{aligned}
x &\equiv \ln\left(\frac{K}{S_0}\right)\Big|_{K=S_T} - rT, \\
F(x;T,\sigma) &= (1-\frac{\sigma'}{\sigma}x)^2 - \frac{(\sigma'\sigma T)^2}{4} + \sigma\sigma'' T, \\
\sigma' &= \frac{\partial\sigma}{\partial x}, \\
\sigma'' &= \frac{\partial^2\sigma}{\partial x^2}.
\end{aligned} \tag{6.4}
$$

It is also helpful to define the implied complementary cumulative distribution function (CCDF) of financial returns as:

$$E(x) = 1 - \int_{-\infty}^{x} P_\sigma(y)dy. \tag{6.5}$$

Eq. (6.3) shows that there is a strong relation between VS and the PDF of financial returns. From another point of view, Eq. (6.3) should be seen as a warning that shows how similar fits of a VS could imply strong differences in the implied returns of the PDF with obvious consequences, for example, on the risk estimation. If one considers, for example, the two curves (red and blue) in Fig. (6.1), it is clear that even if the two lines are close to the actual data, the differences in the decay of the two distributions can be relevant with important consequences for the risk estimation procedure.

One could consider, for example, the estimation of the risk using the standard VAR (value-at-risk) measure [16], defined as

$$\mathcal{P_{VAR}} = \int_{-\infty}^{-\Lambda_{VAR}} P(x)dx, \tag{6.6}$$

where Λ_{VAR} represents our estimation of the maximum potential loss with a fixed confidence level given by $\mathcal{P_{VAR}}$ and $P(x)$ is a generic function that represents reuturns PDF. In this Chapter, we consider $P(x) = P_\sigma(x)$ and $\mathcal{P_{VAR}} = 1\%$ as a standard value for the confidence level; this means we can expect a loss less than or equal to Λ_{VAR} in the 99% of the cases.
For the distributions in Fig. (6.1), we get $\Lambda^{red}_{VAR} = 5.23\%$ and $\Lambda^{blue}_{VAR} = 5.06\%$, so the difference in the VAR estimation using the two different fits is about 3.27%. To have an idea of the order of magnitude of the error, one should consider that for the flat smile (BS) in the figure, we get $\Lambda^{BS}_{VAR} = 4.8\%$ and the difference with the other VAR estimation is about $5\% - 8\%$.

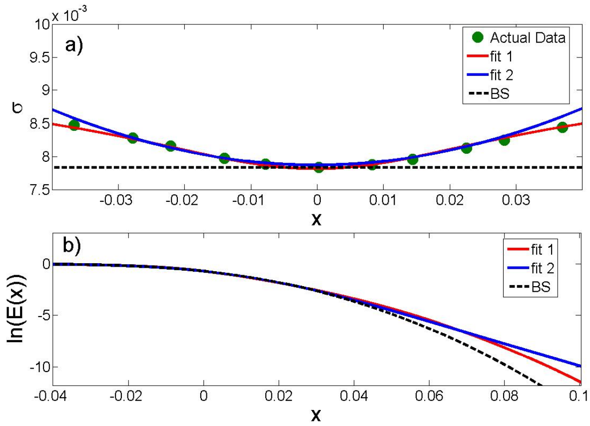

Figure 6.1: Comparison of two suitable approximations for the VS (red and blue) (a) and their CCDFs (b). As evident, even if the two curves can be close to the actual data, the differences in the Value at Risk estimation can be relevant. For comparison we also show the case of a completely flat smile (black) and its Gaussian distribution.

From this example it is clear that there is some arbitrariness in the fitting parameters of the VS function that can generate significant differences in the description of the implied returns distribution, with important consequences, for example, from the risk estimation point of view. So the importance of getting a reliable fitting procedure consistent with the theoretical aspects, as already stressed in [54].
In this framework, we focus our attention on the generalized BS model by considering VS effect and we try to characterize the decay of the tails of the implied distribution of returns as a function of the fitting parameters of the VS, to get a suitable procedure for the smile fitting coherent with the historical observed decay of the actual returns PDFs. As already shown, a suitable characterization of the implied distributions decay can have a fundamental importance, for example, for the risk estimation.

6.4 Relation between VS and the tails of PDF of financial returns

In this Section, we want to establish a simple relation between the parameters of the fitting function Eq. (6.1) and the decay of the tails of the implied distribution of returns, Eq. (6.3). To better understand what we mean for

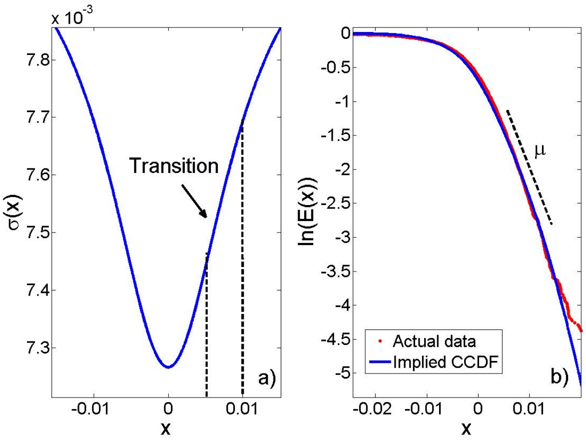

Figure 6.2: a) We show the transition region of the VS $\mathcal{R}$ and b) the exponential decay approximation for the CCDF of returns in the same region.

"decay of the tails", we need to analyze the structure of the Eqs.(6.1, 6.3). First of all, it is important to notice that $\sigma(x)$ is a bounded function

$$g \leq \sigma(x) \leq g\chi.$$

The whole process can be seen as a continuous transition from the a minimum value g to a limit value $g\chi$ reached for large enough returns, x. From the PDF point of view, we can think of the VS as a continuous transition between two Gaussian distributions with different standard deviations, g and $g\chi$. So, due to our choice of the VS fitting function, we already know that for large x values the tails of the implied distribution behaves as a Gaussian distribution. Nonetheless, there is a region of x, namely the *region of the transition*, not described by a Gaussian, since in this case σ is not constant. In Section 6.2 we have already discussed the order of magnitude of x for this region: $x \sim \sqrt{n} = 2.6g\sqrt{T}$ which corresponds to the tail of the distribution. So, even if we know that for really large x the implied distribution is a Gaussian, the region that can be related to the tails of actual returns distributions is the region of transition and this is the region we are going to study in details. Looking at a typical implied distribution of returns on a semilog plot it seems reasonable to approximate the region of the transition by a straight line, as shown in Fig. (6.2).

This approximation is equivalent to assume that the tails distributions of financial returns have an exponential decay, $\exp(-\mu|x|)$, where μ is the factor that characterize the tail. This fact finds confirmation in our real data analysis and it is coherent with results shown in [39].

The main goal of this Section is to establish a relation between the parameter of decay, μ, and the fitting parameters of the smile, g, χ, n. The procedure we consider is straightforward and it is described in the following.
First of all, we fixed the range of the parameters repeating many times the fitting procedure and considering the data set described in Section 6.2. In Table 6.2, we show the range of the parameters that we used to perform our simulations (we used the parameter $\rho = n/(g^2 T)$ instead of n due to the scaling relation Eq. (6.2)).

	min	**max**
g	0.03	0.5
ρ	2.5	10
T (days)	1	1080
χ	1.01	3

Table 6.2: Range of the parameters of the numerical simulations.

Using this range of parameters, we consider the implied CCDF of returns, derived from Eq. (6.3), and we fit the region of transition considering an exponential decay, $\exp(-\mu|x|)$, where μ is the fitting parameter. In this way we get for every set of the parameters in the Table 6.2 the corresponding decay parameter, μ. We define the region of transition as $\mathcal{R} = \{x|\sqrt{n}/2 \leq x \leq \sqrt{n}\}$; in this way, if $A = g\chi + g$ represents the height of the VS, we are considering the region from the 20% to the 50% of the total height.
Our goal is to find a relation between μ and the three parameters of the VS. First of all, let us fix $\chi = 1$, so that the VS is completely flat. In this case we know that the distribution is Gaussian, $F(x, T, \sigma) = 1$ and the parameter μ should be thought of an approximation of an exponential decay. In this case, μ can be easily estimated as:

$$\mu = \frac{\Delta y}{\Delta x} = \frac{\ln(\mathcal{P}(\sqrt{n}) - \ln(\mathcal{P}(\sqrt{n}/2))}{\sqrt{n}/2}, \tag{6.7}$$

where $\mathcal{P}$ is the CCDF of P defined in Eq. (6.3). Performing some calculations we get:

$$\mu_1 = \frac{2}{g\sqrt{T}} f(\rho), \tag{6.8}$$

where, the function $f(\rho)$, is defined by,

$$f(\rho) = \frac{1}{\sqrt{\rho}} \ln \left[\frac{1 - \mathrm{erf}(\frac{1}{2}\sqrt{\frac{\rho}{2}})}{1 - \mathrm{erf}(\sqrt{\frac{\rho}{2}})} \right], \tag{6.9}$$

and has the following asymptotic expansion:

$$f(\rho) \simeq \begin{cases} \sqrt{\rho} & \text{if } \rho \longmapsto +\infty \\ 1/2\sqrt{\pi} & \text{if } \rho \longmapsto 0. \end{cases} \tag{6.10}$$

Let us now discuss the case $\chi \neq 1$: in the light of the adiabatic interpretation presented in [40], we expect that on increasing χ, the PDF will present, soon or later a minimum. This means that the PDF should be flatter than before, so that μ should decrease. This is coherent with our physical interpretation of the VS as a small perturbation of a theoretical system represented by a Gaussian distribution. Increasing the order of magnitude of the perturbation, here represented by the parameter χ, we get a PDF of returns increasingly different from the Gaussian until the adiabatic limit of the perturbation is violated. After that point the system cannot be described by a perturbative approach.

For simplicity, let us assume the simple inverse proportionality:

$$\mu_\chi = \frac{\mu_1}{\chi} = \frac{2}{\chi g \sqrt{T}} f(\rho). \tag{6.11}$$

Relation (6.11) has been checked in Fig. (6.3) where we plot the real parameter μ obtained by our simulation *vs* the parameter μ_χ given by (6.11): the agreement is within a 2% of mean squared error.

6.5 A New Recipe to Fit the Volatility Smile

In this Section we show how to include the information given by the formula (6.11) on the decay of the CCDF of the financial returns to get a suitable fit of the VS coherent from theoretical point of view. Firstly, to do this we need to analyze what is the ordinary interpretation of the implied volatility of the BS model and its relation with historical volatility. Implied volatility is usually interpreted as the future volatility of the market and represents the traders and practitioners vision. From this point of view historical volatility can be interpreted as a peculiar realization of this vision at some particular time period. So, in general, there will be a mismatch between historical volatility and implied volatility and this fact is reflected on historical and implied PDF of returns. Therefore, to use properly the information

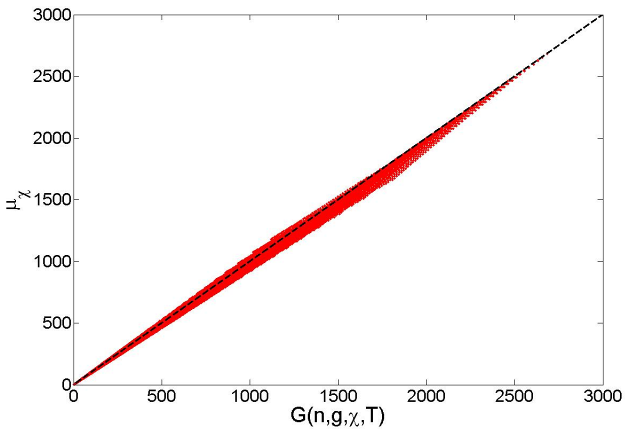

Figure 6.3: We show the relation between the decay parameter μ, given by numerical simulation and the estimation given by Eq. (6.11). As reference, we also show the (dotted) line $\mu = \mu_\chi$.

on the decay of the historical distribution, we need at first the scaling relation between the volatility and the decay of the distribution. This relation can be estimated from historical series of currencies (AUDUSD, EURCHF, EURGBP, EURJPY, EURUSD, GBPUSD, USDCAD, USDCHF, time period 2001-2010) using the following procedure. We consider different time lag ($T = 1, 10, 100$ days) and build different historical series of returns. We divide each series into subgroups of at least 300 elements and we evaluated the standard deviation of each group. To evaluate the decay we consider the CCDF of returns using the procedure described in [16] and we fit the tail decay using a straight line in a semi log plot. We repeat this procedure for any subgroup and for any currency to make explicit the relation between μ_H and σ_H. In Fig. (6.4) we show our results superimposed with a suitable fitting function

$$\sigma_H = \frac{C_1}{\mu_H}, \tag{6.12}$$

where $C_1 = 1.6 \pm 0.5$ is a fitting parameter. Let us observe that this is in quite good agreement with an exponential PDF for returns, since in that case one would have $\sigma_H = 2/\mu_H$.

Eq. (6.12) makes explicit the relation between μ_H and σ_H (their product should be a constant ≈ 1.6) and gives us the opportunity to exploit the information on the historical decay of the PDF of financial returns to get a suitable fit of the VS. The procedure can be summarized as follow:

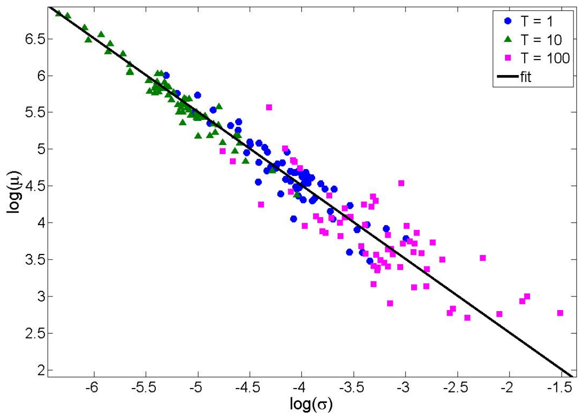

Figure 6.4: Relation (independent on the time lag T) between μ_H and σ_H considering three different time lag for the returns ($T = 1, 10, 100$). We also show the best linear fit $\ln(\mu_H) = \ln(\sigma_H) + ln(C_1)$, where $C_1 = 1.6 \pm 0.5$.

- Using the historical price series we determine the decay and the standard deviation of the financial returns, respectively: μ_H, σ_H.
- Identifying the product $\mu_H \sigma_H$ with $g\sqrt{T}\mu_\chi$ and using our estimation, Eq. (6.11), we can obtain one of three fitting parameters, e.g. χ, describing the VS, as a function of the other two (g, n) :

$$\chi = \frac{2}{\mu_H \sigma_H} f(n/g^2 T). \tag{6.13}$$

Following this approach, we reduce the number of free parameters for the smile fitting, fixing implicitly the right decay of the PDF of returns. As already stressed in Section 6.3, the need of getting a suitable fit for the VS coherent also with the theoretical aspects of the model, is really important in many Risk Management activities and could lead to significant differences in risk estimation.

For example in Fig. (6.5), we compare the PDF of returns obtained by a standard fitting procedure of VS (unconditional fit) with the one obtained following the procedure described before (conditional fit). As evident, even if the two fitting procedures give similar curve for the VS, the effect on the VAR estimation are of the order of 10%.

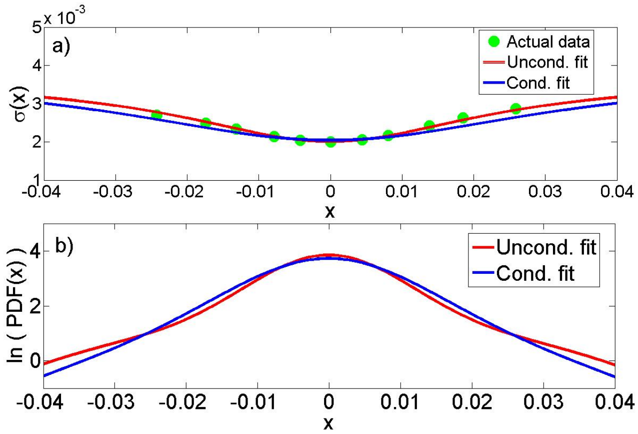

Figure 6.5: We compare unconditional fit of the VS (a) and the implied PDF (b) for a particular dataset (EURJPY, $T = 30$ days, downloaded on 21/10/2009 15:37) with the conditional one.

6.6 Conclusions

We started from the pricing equation of the Black-Schöles model for an European call and we considered the effect of the VS correction on the implied PDF. Our approach comes from statistical physics and it is related to the adiabatic interpretation in [40]. We showed that similar fits of a VS could imply strong differences on the implied returns PDF with obvious consequences on the risk estimation. To obtain a stronger fitting procedure for the VS that can be compatible with the theoretical aspects of the model we first derived a relation between the exponential decay of the CCDF of returns and the parameters of the fitting function of the smile. Then, we exploit this relation to get a new fitting procedure that can be compatible with the historical data. An interesting case is shown in Fig. (6.6) where we compare the PDF of returns obtained by a standard fitting procedure of VS (unconditional fit) with the one obtained following the procedure described before (conditional fit). In this case the time to maturity is large, $T = 2520$, so we cannot get σ^H and μ^H directly from the dataset but we extrapolate their values considering the relation $\mu \propto 1/\sqrt{T}$ and $\sigma \propto \sqrt{T}$. As evident, the unconditional fit generates an implied PDF with a relative minima never observed in actual data [40], on the contrary the conditional fit generates a PDF more "regular" that seems suitable for the description of actual PDF of returns. The price to pay in order to get a smooth PDF is related to the error for the smile fitting: the horizontal amplitude of the conditional fit is

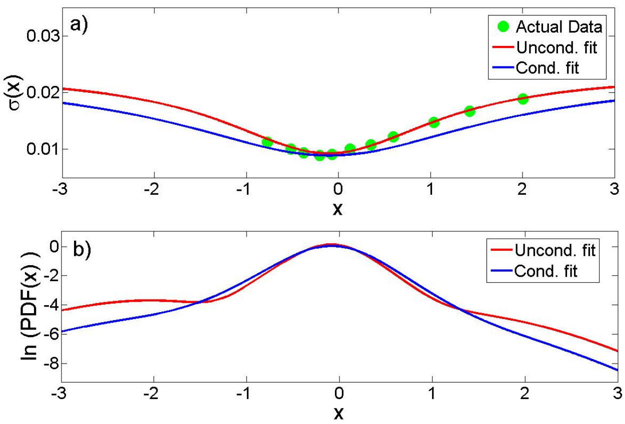

Figure 6.6: We compare unconditional fit of the volatility smile (a) and the implied PDF (b) for a particular dataset (EURJPY, $T = 2520$ days, downloaded on 21/10/2009 15:37) with the conditional one. To fix the historical decay parameters, μ_H, σ_H, we find the values from actual data considering the relation $\mu \propto 1/\sqrt{T}$ (Eq. 6.12) and the standard scaling $\sigma \propto \sqrt{T}$. As evident, the PDF of conditional procedure does not present spurious minima and gives a distribution suitable to describe actual data.

higher than the one required to get a suitable fit. This can be explained assuming that market makers overreact to extreme events when the time to maturity is large, estimating the volatility in a way that is not compatible with historical data. Besides, conditional fit is compatible with the skewness reduction claimed in [54] to get a smile fitting more suitable to the historical data.

In conclusion we provide a new tool for the VS fitting that can be used to get a more coherent estimation of the parameters of fitting function, compatible with historical series and theoretical aspects of the model. A reliable estimate of the implied volatility has application in the risk management activities and in the pricing of exotic derivatives, where, in general, the implied volatility is an input of more complex models.

Chapter 7

Conclusions

In this work we showed how to use concepts that come from physics to get a suitable description of financial world, in accordance with the main ideas of a new interdisciplinary field: Econophysics. In particular, after a wide introduction about what is Econophysics and which are the main motivations that make this subject relevant in the economic world, we introduced the main mathematical instruments useful to treat the main financial problems, namely the Stochastic Calculus and the Probability Theory. In doing this, we also showed how these instruments can be used to describe actual data that come from financial markets.
Then, we focus our attention on an aspect that is peculiar of option derivatives: the implied volatility of the Black-Schöles pricing model and the volatility smile effect. In particular we gave a statistical description of this effect and an intuitive and quantitative interpretation of this phenomenon. Finally, we focus our attention on the problem of getting a suitable fit of the volatility smile, coherent to the theoretical hypothesis of the underling model. To do this, we proposed a new interpretation of the volatility smile effect, starting from the physical concept of adiabatic transition, making a parallelism between a physical system and a financial one. So, starting from a physical interpretation of the problem, we could derive some *adiabatic conditions* to avoid arbitrage opportunities in the pricing of option derivatives. The use of these conditions was summed up in an algorithm that can be applied in practical situations. Then, we further extended this methodology to characterize the implied probability density function of financial returns and in particular its tails. We showed how this characterization can be relevant in the risk estimation activities and, how to get a fitting procedure theoretically coherent with the Black-Schöles model and the historical exponential decay of the implied PDF. In conclusion, by this work we showed how physical methodology can be applied to finance and how this approach can be

relevant to get practical and concrete results.

Bibliography

[1] J.-P. Bouchaud, *Economics Needs a Scientific Revolution*, Nature, **455**, 1181-1181 (2008).

[2] E. Derman, *My Life as a Quant: Reflections on Physics and Finance*, Wiley (2004).

[3] R. N. Mantegna, H. E. Stanley, *An Introduction to Econophysics: Correlations and Complexity in Finance*, Cambridge University Press (1999).

[4] V. Pareto, *Cours d' Économie Politique: Nouvelle édition par G.-H. Bousquet et G. Busino*, Geneva (1964).

[5] P. Levy, *Calcul des Probabilites*, Gauthier-Villars, Paris (1925).

[6] B. B. Mandelbrot, *Fractal Geometry of Nature*, W. H. Freeman, San Francisco (1982).

[7] L. Bachelier, *Theorie de la Speculation*, Annales Scientifiques de l'Ecole Normale Superieure, **3**, 17 (1900).

[8] J. Voit, *The Statistical Mechanics of Financial Markets*, Springer-Verlag (2001).

[9] J.-P. Bouchaud, M. Potters, *Financial Applications of Random Matrix Theory: a short review*, http://arxiv.org/pdf/0910.1205 (2009).

[10] C. W. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, Springer-Verlag (1997).

[11] N. N. Taleb, *Fooled by Randomness: The Hidden Role of Chance in Life and in the Markets*, W. W. Norton & Company (2001).

[12] N. N. Taleb, *The Black Swan*, Random House Trade Paperbacks (2010).

[13] W. Feller, *An Introduction to Probability Theory and Its Applications*, Wiley, NY (1968).

[14] S. N. Neftci, *An Introduction to the Mathematics of Financial Derivatives*, Academic Press (2000).

[15] J. C. Cox, J. E. Ingersoll, S.A. Ross *A Theory of the Term Structure of Interest Rates*, Econometrica **53**, 385-407 (1985).

[16] J. -P. Bouchaud, M. Potter, *Theory of Financial Risk - From Statistical Physics to Risk Management*, University Press, Cambridge (2000).

[17] A. A. Dragulesco and V. M. Yakovenko, *Probability distribution of returns in the Heston model with stochastic volatility*, Quantitative Finance, **2**, 443-453 (2002).

[18] P. A. Samuelson, *Proof that Properly Anticipated Prices Fluctuate Randomly*, Industrial Management Rev **6**, 41-45 (1965).

[19] C. A. E. Goodhart, L. Figliuoli, *Every minute counts in financial markets*, Journal of International Money and Finance, **10**, 23-52 (1991).

[20] M. M. Dacorogna, R. Gençay, U. Müller, R. B. Olsen, O. V. Pictet, *Introduction to High-Frequency Finance*, Academic Press (2001).

[21] F. Black, M. Schöles, *The Pricing of Options and Corporate Liabilities*, Journal of Political Economy **81**, 637-654 (1973).

[22] J. C. Hull, *Options, Futures and Other Derivatives*, Prentice Hall, Upper Saddle Rivel, New Jersey (2003).

[23] P. Willmott, *Derivatives*, John Wiley and Sons, Chinchester, New York, Weinheim, Brisbane, Singapore, Toronto (1998).

[24] R. Rebonato, *Volatility and Correlation: In the Pricing of Equity, FX and Interest-Rate Options*, Wiley (1999).

[25] G. Brown, C. Randall, *If the Skew Fits*, Risk Magazine **12** (4), 62-65 (1999).

[26] R. Tompkins, *Implied Volatility Surfaces: Uncovering the Regularities for Options on Financial Futures*, The European Journal of Finance **7**, 198-230 (2001).

[27] K. Toft, B. Prucyk, *Options on Leveraged Equity: Theory and Empirical Tests*, Journal of Finance **52**(3), 1151-1180 (1997).

[28] J. Campa, K. Chang, R. Reider, *Implied Exchange Rate Distributions: Evidence from OTC Option Markets*, Journal of International Money and Finance **17** (1), 117-160 (1998).

[29] V. F. Pisarenko, D. Sornette, *New Statistic for Financial Returns Distributions: Power-Law or Exponential?*, Physica A **366**, 387-400 (2006).

[30] P. Gopikrishnan, M. Meyer, L. A. N. Amaral, H. E. Stanley, *Inverse Cubic Law for the Distribution of Stock Price Variations*, European Physical Journal B **3** (2), 139-140 (1998).

[31] J. Gatheral, *The volatility skew: Arbitrage constraints and asymptotic be- havior*, Technical report, Merill Lynch (1999).

[32] R. Lee, *The Moment Formula for Implied Volatility at Extreme Strikes*, Mathematical Finance, **14**, 469-480 (2004).

[33] B. Dupire, *Pricing with a Smile*, Risk Magazine **7** (1), 18-20, (1994).

[34] E. Derman, I. Kani, *Stochastic implied trees: Arbitrage pricing with stochastic term and strike structure of volatility*, International Journal of Theoretical and Applied Finance **1**, 61-110 (1998).

[35] E. Stein, J. Stein, *Stock price distributions with stochastic volatility : an analytic approach*, The Review of Financial Studies, **4**, 727-752 (1991).

[36] S. L. Heston, *A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options*, The Review of Financial Studies, **6**, 2, 327-343 (1993).

[37] P. Hagan, D. Kumar, A. Lesniewski and D. Woodward, *Managing Smile Risk*, Wilmott Magazine, **1**, 84-108 (2002).

[38] S. Mikhailov, U. Nögel, *HestonÕs Stochastic Volatility Model: Implementation, Calibration, and some Extensions*, Wilmott, 74-79 (2003).

[39] A. A. Dragulesco and V. M. Yakovenko, *Probability distribution of returns in the Heston model with stochastic volatility*, Quantitative Finance, **2**, 443-453 (2002).

[40] L. Spadafora, G. B. Berman, F. Borgonovi, *Adiabaticity Conditions for Volatility Smile in Black-Schöles Pricing Model*, in press on EPJ B (2010).

[41] J. C. Jackwerth, *Option Implied Risk-Neutral Distributions and Implied Binomial Trees: A literature Review*, Journal of Derivatives **7**, 66- 82 (1999).

[42] E. Derman, I. Kani, *Riding on a Volatility Smile*, Risk Magazine **7** (2), 32-39, (1994).

[43] E. Derman, I. Kani, *The volatility Smile and its Implied Tree*, Quantitative Strategies Research Notes, Goldman Sachs, New York, NY, (1994).

[44] M. Rubinstein, *Implied Binomial Trees*, Journal of Finance **49**, 771-818, (1994).

[45] J. Jackwerth, *Generalized Binomial Trees*, Journal of Derivatives **5** (2), 7-17, (1997).

[46] T. Herwing, *Construction of Arbitrage-Free Implied Trees: A New Approach*, Lecture Notes in Economics and Mathematical System **571**, 11-38, (2006).

[47] M. R. Fengler, *Arbitrage-Free Smoothing of the Implied Volatility Surface*, Quantitative Finance, **9**, 417-428 (2009).

[48] Y. Ait-Sahalia, J. Duarte, *Nonparametric option pricing under shape restrictions*, Journal of Econometrics, **116**, 9-47 (2003).

[49] N. Kahale, *An Arbitrage-free Interpolation of Volatilities*, Risk Magazine, **17**, 102-106 (2004).

[50] L. D. Landau, E. M. Lifshitz, *Mechanics*, Butterworth-Heinenann, Oxford, UK (2000).

[51] A. M. Malz, *Option-Implied Probability Distribution and Currency Excess Returns*, FRB of New York Staff Report **32** (1997).

[52] U. Kirchner, *Market Implied Probability Distributions and Bayesian Skew Estimation*, arXiv:/0911.0805 (2009).

[53] L. Spadafora, G. B. Berman, F. Borgonovi, *Do your volatility smiles take care of extreme events?*, arXiv:/1010.2184 (2010).

[54] S. Ciliberti, J. -P. Bouchaud, M. Potter, *Smile dynamics - a theory of the implied leverage effect*, arXiv:/0809.3375v1 (2008).