# 1 Fourier Pricing Methods

## 1.1 INTRODUCTION

In recent years, Fourier transform methods have emerged as some of the major methodologies for the evaluation of derivative contracts. The main reason has been the need to strike a balance between the extension of existing pricing models beyond the traditional Black and Scholes setting and a parsimonious stance for the evaluation of prices consistently with the market quotes.

On the one hand, the end of the Black–Scholes world spurred more research on new models of the dynamics of asset prices and risk factors, beyond the traditional framework of normally distributed returns. On the other, restricting the search to the set of processes with independent increments pointed to the use of Fourier transform as a natural tool, mainly because it was directly linked to the characteristic functions identifying such processes.

This book is devoted to the use of Fourier transform methods in option pricing. With respect to the rest of the literature on this topic, we propose a new approach, based on generalized functions. The main idea is that the price of the fundamental securities in an economy – that is, digital options and Arrow–Debreu securities – may be represented as the convolution of two generalized functions, one representing the payoff and the other the pricing kernel.

In this chapter we present the main results of the book. The remaining chapters will then lead the reader through a sort of flashback story over the main steps needed to understand the rationale of Fourier transform pricing methods and the tools needed for implementation.



## 1.2 A GENERAL REPRESENTATION OF OPTION PRICES

The market crash of 19 October 1987 may be taken as the date marking the end of the Black–Scholes era. Even though the debate on evidence that market returns were not normally distributed can be traced back much further in the past, from the end of the 1980s departures from normality have become the usual market environments, and exploiting these departures has even suggested new business ideas for traders. Strategies bound to gain from changes in the skew or higher moments have become the usual tools in every dealing room, and concerns about exposures to changes in volatility and correlation have become a major focus for risk managers.

On the one hand, the need to address the issue of non-Gaussian returns started the quest for new models that could provide a better representation of asset price dynamics; and, on the other, that same need led to the rediscovery of an old idea. According to a model going back to Breeden and Litzenberger (1978), one may recover the risk-neutral probability from the prices of options quoted in the market. Notice that this finding only depends on the requirement to rule out arbitrage opportunities and must hold in full generality for all risk-neutral probability distributions. The idea is that the risk-neutral density can be computed as the second derivative

of the price of options with respect to the strike. More precisely, we have that

$$B(t,T)f_{t,T}(K) \equiv B(t,T)\mathbb{Q}_t(S_T \in \mathrm{d}K) = \frac{\partial^2 \mathcal{P}(S_t;K,T)}{\partial K^2}$$

where $\mathcal{P}(S_t;K,T)$ denotes the put option and $B(t,T)$ is the risk-free discount factor – that is, the value at time $t$ of earning a unit of cash for sure at future time $T$. This is true of all option pricing models. Notice that the no-arbitrage condition immediately leads to characterize $f_{t,T}(x)$ as a density. First, if one assumes to have bought a product paying a unit of cash if $(S_T \in dx)$ and zero otherwise, the price of this product cannot be negative. Second, if one assumes to have bought a set of products paying one unit of cash if $(S_T \in dx)$ in such a way as to cover the all-positive real line $[0,\infty]$, then one must earn one unit of cash for sure, so that we have

$$\int_0^\infty f_{t,T}(x)\,\mathrm{d}x = 1$$

Computing option prices amounts to an evaluation of the integrals of the density above, when it exists. Namely, consider the price of an option paying 1 unit of cash if the value of the underlying asset is lower than $K$ at time $T$. The price of this option, which is called a digital *cash-or-nothing put* option, is

$$\mathcal{P}_{\mathrm{CoN}} = B(t,T)\int_0^K f_{t,T}(x)\,\mathrm{d}x = B(t,T)\mathbb{Q}_t(S_T \le K)$$

Now consider a similar product delivering one unit of asset $S$ in the event $S_T \le K$. This product is called an *asset-or-nothing put* option. Likewise, its price will be

$$\mathcal{P}_{\mathrm{AoN}} = B(t,T)\int_0^K x f_{t,T}(x)\,\mathrm{d}x = B(t,T)\mathbb{E}_t^Q(S_T \mathbf{1}_{[S_T \le K]})$$

where $\mathbb{E}_t^Q(x)$ denotes the conditional expectation taken under probability measure $\mathbb{Q}$ with respect to the information available at time $t$. Consider now the portfolio of a short position on an *asset-or-nothing put* and a long position in $K$ *cash-or-nothing put* options, with same strike price $K$ and same maturity $T$. Then, at time $T$ the value of such a portfolio will be

$$K\mathbf{1}_{[S_T \le K]} - S_T \mathbf{1}_{[S_T \le K]} = \max(K - S_T, 0)$$

which is the payoff of a European put option. The no-arbitrage assumption then requires that the value of the put option at any time $t < T$ should be equal to

$$\mathcal{P}(S_t;K,T) = B(t,T)\left(K\mathbb{Q}_t(S_T \le K) - \mathbb{E}_t^Q(S_T \mathbf{1}_{[S_T \le K]})\right)$$

It is easy to check that the no-arbitrage assumption requires that a digital option paying one unit of cash if, at time $T$, the underlying asset is worth more than $K$ (*cash-or-nothing call*) must have the same value as that of a long position in the risk-free asset and a short position in a *cash-or-nothing put* option. Namely, we must have

$$\mathcal{C}_{\mathrm{CoN}} = B(t,T) - \mathcal{P}_{\mathrm{CoN}} = B(t,T)(1 - \mathbb{Q}_t(S_T \le K)$$

where $\mathcal{C}_{\mathrm{CoN}}$ denotes the *cash-or-nothing call* option. By the same token, and *asset-or-nothing call* option can be replicated by buying a unit of the underlying asset spot while going short

the *asset-or-nothing put*

$$\mathcal{C}_{\text{AoN}} = S_t - B(t,T)\mathbb{E}_t^{\mathcal{Q}}(S_T \mathbf{1}_{[S_T \leq K]})$$

Notice that the value of an *asset-or-nothing call* option must also be equal to

$$\mathcal{C}_{\text{AoN}} = B(t,T)\mathbb{E}_t^{\mathcal{Q}}(S_T \mathbf{1}_{[S_T > K]})$$

so that we have

$$\mathcal{C}_{\text{AoN}} + \mathcal{P}_{\text{AoN}} = B(t,T)\mathbb{E}_t^{\mathcal{Q}}(S_T) = S_t$$

This defines the main property of the probability measure $\mathbb{Q}$. Under this measure, the asset $S$, and every other asset in the economy, is expected to earn the risk-free rate. For this reason, this measure is called *risk-neutral*. Alternatively, if one defines a new variable $Z_t \equiv S_t/B(t,T)$, it is evident that under measure $\mathbb{Q}$ we have

$$Z_t = \mathbb{E}_t^{\mathcal{Q}}(Z_T)$$

and the price of the asset $S$, and every other asset, turns out to be a martingale when measured using the risk-free asset as the numeraire. For this reason, this measure is also called an *equivalent martingale measure* (EMM), where equivalent means that it gives zero measure to the events that have zero measure under the historical measure, and only to those.

Notice that just as for the put option, the price of a call option can be written as a long position in an *asset-or-nothing call* option and a short position in $K$ *cash-or-nothing call* options. Formally,

$$\mathcal{C}(S_t : K, T) = B(t,T)\mathbb{E}_t^{\mathcal{Q}}(S_T \mathbf{1}_{[S_T > K]}) - K\,B(t,T)(1 - \mathbb{Q}_t(S_T \leq K))$$

Notice that by applying a change of numeraire, namely using $S_t$, we can rewrite the *asset-or-nothing* option in the form

$$\mathcal{C}_{\text{AoN}} = B(t,T)\mathbb{E}_t^{\mathcal{Q}}(S_T \mathbf{1}_{[S_T > K]}) = S_t \mathbb{Q}_t^*(S_T > K)$$

where $\mathbb{Q}^*$ is a new probability measure. So, European options can be written in full generality as a function of two probability measures, one denoting the price of a *cash-or-nothing* option and the other pricing the *asset-or-nothing* one. For call options we have then

$$\mathcal{C}(S_t : K, T) = S_t(1 - \mathbb{Q}_t^*(S_T \leq K)) - K\,B(t,T)(1 - \mathbb{Q}_t(S_T \leq K))$$

and for put options

$$\mathcal{P}(S_t : K, T) = -S_t \mathbb{Q}_t^*(S_T \leq K) + K\,B(t,T)\mathbb{Q}_t(S_T \leq K)$$

So, the risk-neutral density completely specifies the price of options for all strikes and maturities.

## 1.3 THE DYNAMICS OF ASSET PRICES

From the discussion above, pricing derivatives in an arbitrage-free setting amounts to selecting a measure endowed with the martingale property. In a complete market, only one measure is sufficient to fit all prices exactly. This implies that all financial products can be exactly replicated by a dynamic trading strategy (all assets are *attainable*). In incomplete markets, the measure must be chosen according to auxiliary concepts, such as mean-variance optimization or the expected utility framework. Concerning this choice, the current presence of liquid option

markets with different strike prices and maturities has added more opportunities to replicate derivative contracts and, at the same time, more information on the shape of the risk-neutral distribution. This has brought about the problem of selection and comparison of the models with the whole set of prices observed on the market – that is, the issue of calibration to market data.

By and large, two main strategies are available. One could try models with a limited number of parameters, but a sufficient number of degrees of freedom to represent the dynamics of assets as consistently as possible with the prices of options. The advantage of this route is that it allows a parsimonious arbitrage-free representation of financial prices and it directly provides dynamic replication strategies for contingent claims. This has to be weighted against the risk of model mis-specification. On the other hand, one could try to give a non-parametric representation of the dynamics, based on portfolios of cash positions and derivative contracts held to maturity. This approach is known as *static replication* and it has the advantage of providing the best possible fit to observed prices. The risk is that some products used for static replication may be illiquid, and their prices inconsistent with the no-arbitrage requirement.

This book is devoted to the first strategy, that is the selection of a convenient fully specified dynamics for the prices of assets. The models reviewed in this book are based on two assumptions that jointly determine what is called the *Efficient Market Hypothesis*. The first is that prices are Markovian, meaning that all information needed to predict future price changes is included in the price currently observed, so that past information cannot produce any improvement in the forecast. The second assumption is that such forecasts are centred around zero, so that price changes are not predictable.

The above framework directly leads to modelling the dynamics of asset prices as processes with *independent increments*. The price, or more precisely the logarithm of it, is assumed to move according to a sequence of shocks such that no shock can be predicted from a previous shock. If one adds that all these shocks have the same distribution – that is, are identically distributed, and finite variance – the standard result, called, the central limit theorem, predicts that these log-changes, when aggregated over a reasonable number of shocks, should be normally distributed, so that the prices should be log-normally distributed. This is the standard model used throughout most of the last century, and named the Black–Scholes model after the famous option pricing formula that is recovered under this assumption.



In the Black–Scholes setting, the logarithm of each asset is then assumed to be driven by a Brownian motion with constant diffusion and drift parameters. Formally, if we denote $X_t \equiv \ln(S_t)$ we have

$$\mathrm{d}X_t = \left(r - \frac{1}{2}\sigma^2\right)\mathrm{d}t + \sigma\,\mathrm{d}W_t$$

where $\sigma$ is the diffusion parameter, $r$ is the instantaneous risk-free rate of return and $W_t$ is a *Wiener process*. The dynamics of price $S$ is then represented by a *geometric Brownian motion*. Notice that this model predicts that all options traded on the market should be consistent with the same volatility figure $\sigma$, for all strike and maturity dates. As discussed before, this prediction is clearly at odds with the empirical evidence gathered from option market prices. In many option markets, prices of *at-the-money* options are consistent with volatility levels different from those implied by *out-of-the-money* and *in-the-money* option prices. Namely, in markets such as foreign exchange and interest rate options, the volatility of both *in* and *out of* the money options is higher than that of *at-the-money* options, producing a phenomenon called the *smile effect*, after the scatter of the relationship between volatility and moneyness

that resembles the image of a smiling mouth. In other markets, such as that of equity options, this relationship is instead generally negative, and it is called *skew*, recalling the empirical regularity that volatility tends to increase in low price scenarios. Moreover, volatility also tends to vary across maturities, generating *term structures of volatility* typical of every market.

The quest for a more flexible representation of the asset price dynamics, consistent with smiles and term structures of volatility, has brought us to dropping either of the two assumptions underlying the Black–Scholes framework. The first is that the assets follow a diffusion process, and the second is the stationarity of the increments of log-prices. So, more general models could be constructed allowing for the presence of jumps in asset price dynamics and for changes in the volatility and the probability of such jumps – that is, intensity. If we stick to processes with independent stationary increments, this defines a class of processes called *Lévy processes*. An effective way to describe these processes is to resort to their *characteristic function*. We recall that the characteristic function of a variable $X_t$ is defined as

$$\phi_{X_t}(\lambda) = \mathbb{E}\left(\mathrm{e}^{i\lambda X_t}\right)$$

A general result holding for all Lévy processes is that this characteristic function may be written as

$$\phi_{X_t}(\lambda) = \mathrm{e}^{-t\psi(\lambda)}$$

where the function $\psi(\lambda)$ is called the *characteristic exponent* of the process. Notice that stationarity of increments implies that the characteristic exponent is multiplied by the time $t$ so that increments of the process over time intervals of the same length have the same characteristic function and the same distribution. A fundamental result is that such a characteristic exponent can be represented in full generality using the so-called *Lévy–Khintchine formula*.

$$\psi(\lambda) = -ia\lambda + \frac{1}{2}\sigma^2\lambda^2 - \int_{-\infty}^{+\infty}\left(\mathrm{e}^{i\lambda x} - 1 - i\lambda x\mathbb{I}_{\{|x|\leq 1\}}\right)\nu(\mathrm{d}x) \qquad \lambda \in \mathbb{R}$$

Every Lévy process can then be represented by a triplet $\{a, \sigma, \nu\}$, which uniquely defines the characteristic exponent. The first two parameters define the diffusion part of the dynamics, namely drift and diffusion. The last parameter is called the *Lévy measure* and refers to jumps in the process. Loosely speaking, the Lévy measure provides a synthetic representation of the contribution of jumps by the product of the instantaneous probability of such jumps, the intensity, and the probability density function of the dimension of jumps. Intuitively, keeping this measure finite requires that relatively large jumps must have finite intensity, while jumps with infinite intensity must have infinitesimal length. The former kind of jumps are denoted as *finite activity*, while the latter are called *infinite activity* and describe a kind of dynamics similar to that of diffusion processes. For further generalization, positive and negative jumps may also be endowed with different specifications.

Stationarity may be a limit for Lévy processes. As a matter of fact, this would imply that the distribution of log-returns on assets over holding periods of the same length should be the same, while in the market we usually see changes in their distribution: typically, we see periods of very huge movements followed by periods of relative calm, a phenomenon which is known as *clustering* of volatility. An intuitive way of moving beyond stationary increments is to assume that both the volatility of the diffusive part and the intensity of jumps change randomly as time elapses. Even the economic rationale for that goes back to a very old stream of literature of the 1970s. Clark (1973) proposed a model to explain the joint dynamics of trading volume and asset prices using subordinated processes. In the field of probability theory, Monroe (1978)

proved that all semi-martingale processes can be represented as Brownian processes evaluated at stochastic times. Heuristically, this means that one can always represent any general process by sampling a Brownian motion at random times. Several *stochastic clocks* may be used to switch from the non-Gaussian process observed at *calendar time* to a Brownian motion. If the stochastic clock is taken to be a continuous process, then the required change of time is its quadratic variation. As an alternative, a stochastic clock can be constructed by any strictly increasing Lévy process: these processes are called *subordinators*. One could also use other variables as proxies for this level of activity of the market. The main idea is in fact to model the process of information arrival to the market: in periods in which the market is hectic and plenty of information flows to the market, business time is moving more quickly, but when the market is illiquid or closed, the pace of time slows down.

In the time change approach, the characteristic function is obtained by a composition of the characteristic exponent of the stochastic clock process and that of the subordinated process. The result follows directly from the assumption that the subordinator is independent of the time-changed process. As an alternative approach, it is possible to remain within the realm of stochastic processes with independent increments by extending the Lévy–Khintchine representation. In this case, the characteristic function becomes

$$\phi_{X_t}(\lambda) = \exp(-\psi_t(\lambda))$$

with characteristic exponent

$$\psi_t(\lambda) = i a_t \lambda - \frac{1}{2}\sigma_t^2 \lambda^2 + \int_{-\infty}^{+\infty} \left(e^{i\lambda x} - 1 - i\lambda x \mathbb{I}_{\{|x|\leq 1\}}\right) \nu_t(dx) \qquad \lambda \in \mathbb{R}$$

Notice that, unlike the case of Lévy processes, $\psi_t(\lambda)$ is no longer linear in $t$. Technical requirements must be imposed on the process governing volatility and the Lévy measure (heuristically, they must not decrease with the time horizon).



## 1.4 A GENERALIZED FUNCTION APPROACH TO FOURIER PRICING

From what we have seen above, a pricing system can be completely represented by a *pricing kernel*, which is the price of a set of digital options at each time $t$. We now formally define the payoff of such options, for all maturities $T > t$. We start by denoting $m \equiv (B(t,T)K)/S_t$ the scaled value of the strike price, where the forward price is used as the scaling variable. This is a natural measure of moneyness of the option. Now, define $k \equiv \ln(m)$ as our key variable representing the strike. We omit the subscript $t$ to the strike for ease of convenience, but notice that at time $T$, $k = \ln(K/S_T)$. Let $X_t = \ln(S_t/B(t,T))$. Then, the *Heaviside function* $\theta(\omega(X_T - X_t - k))$, where $\omega = -1$, defines the event $\{S_T \leq K\}$ and $\omega = 1$ refers to the complementary event. So, in what follows we will refer to the probability measure of the variable $X_T - X_t$, that is, the increment of the process between time $t$ and time $T$, rather than its level at the terminal date. Anyway, since we are concerned with pricing a set of contingent claims at time $t$, when $X_t$ is observed, this will only amount to a rescaling by a known constant.

As for the function $\theta(x)$, we recall its formal definition as

$$\theta(x) = \begin{cases} 1 & x > 0 \\ 0 & x < 0 \end{cases}$$

### 1.4.1 Digital payoffs and the Dirac delta function

In financial terms, the *cash-or-nothing* product can be considered as the limit of a sequence of bull/bear spreads. This limit leads to the derivative of the call option pricing formula with respect to the strike price. It is also easy to check that – in financial terms – just as the digital option is the limit of a sequence of call spreads, the derivative of this option is the limit of a sequence of *butterfly spreads*. In fact, it may be verified by heuristic arguments that the payoff of such a product is a *Dirac delta* function assigning infinite value to the case $S_T = K$ and zero to all other events. Not surprisingly, the price of such a limit product, computed as the expected value under the equivalent martingale measure, is the density, when it exists, of the *pricing kernel*, and it is considered to be the equivalent of Arrow–Debreu prices for asset prices that are continuous variables.

Then, from a financial viewpoint, it is quite natural to consider the *Dirac delta* function as the derivative of the *Heaviside step* function. It is not so from a mathematical viewpoint, unless we introduce the concept of *generalized functions*. Loosely speaking, a generalized function may be defined as a linear functional from an assigned set of functions, called *testing functions* to the set of complex numbers. This set of functions is chosen to be infinitely smooth and with compact support, or with some particular regularity condition on their speed of descent. Formally, if we denote $\varphi(x)$ to be a testing function, a generalized function $f(x)$ is defined through the operator assigning a complex number to the function

$$\langle f, \varphi \rangle \equiv \int_{\mathbf{R}} f(x)\varphi(x)\,\mathrm{d}x$$

Notice that by the main property of the *Dirac delta* function we have that

$$\langle \delta, \varphi \rangle = \varphi(0)$$

Furthermore, by a straightforward application of integration by parts, one may prove that the derivative of the distribution $f(x)$ is

$$\langle f', \varphi \rangle = \int_{\mathbf{R}} f'(x)\varphi(x)\,\mathrm{d}x = -\int_{\mathbf{R}} f(x)\varphi'(x)\,\mathrm{d}x = -\langle f, \varphi' \rangle$$

Now notice what happens if we compute the derivative of the *Heaviside step function* $\theta(x)$. We have

$$\langle \theta', \varphi \rangle = -\langle \theta, \varphi' \rangle = -\int_{\mathbf{R}} \theta(x)\varphi'(x)\,\mathrm{d}x = \phi(0) - \varphi(\infty) = \varphi(0)$$

where we have used bounded support or the rapid descent property of the testing functions. We have then that

$$\langle \theta', \varphi \rangle = \langle \delta, \varphi \rangle$$

and the conjecture based on financial arguments is rigorously proved: in the realm of generalized functions, the derivative of the *Heaviside step function* is actually the *Dirac delta function*.

The strategy followed throughout this book is to remain in the realm of a generalized function to consistently recover the price of options in terms of Fourier transforms.

### 1.4.2 The Fourier transform of digital payoffs

The starting point of our approach is to recover the Fourier transform of the payoff of digital options. This is clearly not defined if the Fourier transform is applied to functions, but it is well defined in the setting of generalized functions.

For a start, we will denote by $\mathcal{F}$ the Fourier transform operator, and by $\overline{\mathcal{F}}$ its inverse, and write

$$\hat{f} = \mathcal{F} f, \quad f = \overline{\mathcal{F}} \hat{f}$$

following the convention:

$$\mathcal{F} f(v) \equiv \int \mathrm{d}u \, \mathrm{e}^{i2\pi uv} f(u)$$
$$\overline{\mathcal{F}} g(v) \equiv \int \mathrm{d}u \, \mathrm{e}^{-i2\pi uv} g(u)$$

We report here the main result concerning the Fourier transform of the digital option that is fully developed and explained in Chapter 5. Let us introduce

$$\delta^+(x) \equiv \frac{i}{2\pi} g^+(x)$$

where

$$g^+(x) = \lim_{\epsilon \to 0^+} \frac{1}{x + i\epsilon}$$

We are now going to show that $\overline{\mathcal{F}}[\delta^+] = \theta$, from which $\mathcal{F}[\theta] = \delta^+$. Since

$$\langle \overline{\mathcal{F}}[\delta^+], \varphi \rangle = \langle \delta^+, \overline{\mathcal{F}}[\varphi] \rangle$$

$$\begin{aligned}
\langle \delta^+, \overline{\mathcal{F}}[\varphi] \rangle &= \frac{i}{2\pi} \lim_{\epsilon \to 0^+} \int \mathrm{d}x \, \mathrm{d}\lambda \frac{\varphi(\lambda)}{x + i\epsilon} \mathrm{e}^{-2\pi i \lambda x} \\
&= \frac{i}{2\pi} \lim_{\epsilon \to 0^+} \int_{-\infty}^{0} \mathrm{d}\lambda \, \varphi(\lambda) \int \mathrm{d}x \frac{1}{x + i\epsilon} \mathrm{e}^{2\pi i |\lambda| x} \\
&\quad + \frac{i}{2\pi} \lim_{\epsilon \to 0^+} \int_{0}^{+\infty} \mathrm{d}\lambda \, \varphi(\lambda) \int \mathrm{d}x \frac{1}{x + i\epsilon} \mathrm{e}^{-2\pi i |\lambda| x} \\
&= \lim_{\epsilon \to 0^+} \int_0^\infty \mathrm{d}\lambda \, \varphi(\lambda) \mathrm{e}^{-2\pi\epsilon\lambda} = \int_0^\infty \mathrm{d}\lambda \, \varphi(\lambda)
\end{aligned}$$

it follows that:

$$\overline{\mathcal{F}}[\delta^+] = \theta$$

Now, it is possible to compute that the distributional value of $g^+(x)$ is p.v. $1/x - i\pi\delta(x)$ (see Example 5.4.3), so that we conclude

$$\begin{aligned}
\mathcal{F}[\theta](v) = \delta^+(v) &= \frac{i}{2\pi} \left( \text{p.v.} \frac{1}{v} - i\pi\delta(v) \right) \\
&= \frac{1}{2}\delta(v) + \frac{i}{2\pi} \text{p.v.} \left( \frac{1}{v} \right)
\end{aligned}$$

where p.v. denotes the principal value and $\delta$ is the *Dirac delta* function.

### 1.4.3 The cash-or-nothing option

We are now going to recover the price of digital *cash-or-nothing* options. We shall treat both the probability distribution $\mathbb{Q}$ and the payoff as generalized functions, and the pricing formula as a convolution of distributions. In this setting, we have already computed the Fourier transform of the payoff. As for the distribution, we assume that we only know its *characteristic function*, which we redefine in a slightly different way, which is useful for computational purposes:

$$\phi_X(v) \equiv \mathbb{E}\left[\mathrm{e}^{i2\pi v X_T}\right] = \int \mathbb{Q}(\mathrm{d}u)\,\mathrm{e}^{i2\pi vu} = \int \mathcal{F}\,\mathrm{d}\,\mathbb{Q} \tag{1.1}$$

Notice that with respect to the usual definition we have simply multiplied the exponent by $2\pi$.

The maths concerning these assumptions is thoroughly discussed in the main body of this book, namely Chapters 5 and 6, so here we stick to essential definitions for the reader who is already familiar with the technique.

Let '$f$', and '$g$' be two generalized functions. The convolution will be denoted as:

$$f \star g \equiv \int \mathrm{d}u\; f(u)g(y-u)$$

If $\mathbb{Q}$ is a (probability) measure, we shall write:

$$(\mathbb{Q} \star g)(y) \equiv \int \mathbb{Q}(\mathrm{d}u)g(y-u)$$

We are interested in the convolution, in a generalized function sense, of the density and the digital payoff function $\theta(x)$.

$$\mathbb{Q}(k) = \mathbb{Q} \star \theta(k) \equiv \int \mathbb{Q}(\mathrm{d}u)\theta(k-u) \tag{1.2}$$

Notice that the main pillar of our approach is the requirement that this convolution of generalized functions be well defined. In Chapter 5, section 8, we give a proof under very weak conditions, which amount to the existence of the first moment of the probability distribution. We now apply the Fourier transform to the convolution and obtain:

$$f \star g = \overline{\mathcal{F}}[(\mathcal{F}f)(\mathcal{F}g)] \tag{1.3}$$

and

$$\langle \mathcal{F}f, \phi\rangle = \langle f, \mathcal{F}\phi\rangle$$

We now use equation (1.3) to compute (1.2 ):

$$\mathbb{Q}(k) = \int \mathrm{d}u\,\mathrm{e}^{-2\pi iku}\phi_X(u)\delta^+(u) \tag{1.4}$$

Replacing the value for $\delta^+$ in equation (1.4) and applying a result that may be found in Chapter 5, Example 5.4.2, we end up with

$$\mathbb{Q}(k) = \frac{1}{2} + \frac{i}{2\pi}\int \frac{\mathrm{d}u}{u}\left[\phi_X(u)\,\mathrm{e}^{-2\pi iuk} - 1\right] \tag{1.5}$$

The above formula is certainly not new (see, for example, Kendall and Stuart, 1977, vol. III). It provides the relationship between the characteristic function and the cumulative probability distribution, which in our case is the pricing kernel of the economy.

The value of a *cash-or-nothing put* option is then given by

$$\mathcal{P}_{\text{CoN}}(k) = B(t,T)\left\{\frac{1}{2} + \frac{i}{2\pi}\int \frac{\mathrm{d}u}{u}\left[\phi_X(u)\,\mathrm{e}^{-2\pi iuk} - 1\right]\right\} \tag{1.6}$$

It is now immediate to obtain the price of the corresponding cash-or-nothing call option. Namely, we have

$$\begin{aligned} 1 - \mathbb{Q}(k) &\equiv 1 - \int \mathbb{Q}(\mathrm{d}u)\theta(k-u) \\ &= 1 - \left\{\frac{1}{2} + \frac{i}{2\pi}\int \frac{\mathrm{d}u}{u}\left[\phi_X(u)\,\mathrm{e}^{-2\pi iuk} - 1\right]\right\} \end{aligned} \tag{1.7}$$

and we immediately obtain

$$\mathcal{C}_{\text{CoN}} = B(t,T)\left\{\frac{1}{2} - \frac{i}{2\pi}\int \frac{\mathrm{d}u}{u}\left[\phi_X(u)\,\mathrm{e}^{-2\pi iuk} - 1\right]\right\} \tag{1.8}$$

### 1.4.4 The asset-or-nothing option

We now extend the analysis to *asset-or-nothing* options. The whole analysis above would of course lead to a result analogous to that obtained for *cash-or-nothing* options. As a matter of fact, we saw before that the two prices are linked by a change of measure. Namely,

$$B(t,T)\mathbb{E}(S_T \mathbf{1}_{[S_T \le K]}) = S_t\, Q^*(k)$$

Under our notation, which is based on the forward price rescaled with respect to the price at time $t$ (that is, $S_t = 1$), the Radon–Nikodym derivative linking the two measures is $S_T$, so that we may write

$$\int \mathbb{Q}^*(\mathrm{d}u) = 1, \quad \mathbb{Q}^*(\mathrm{d}u) = \mathbb{Q}(\mathrm{d}u)\,\mathrm{e}^u \tag{1.9}$$

We may now denote the characteristic function of measure $\mathbb{Q}^*$ as

$$\phi_X^*(k) \equiv \int \mathbb{Q}^*(\mathrm{d}u)\,\mathrm{e}^{2\pi iku} \tag{1.10}$$

and a straightforward computation gives the relationship between the characteristic function of measure $\mathbb{Q}^*$ and that of measure $\mathbb{Q}$:

$$\phi_X^*(k) = \int \mathbb{Q}(\mathrm{d}x)\,\exp\left[2\pi i\left(k - \frac{i}{2\pi}\right)x\right] = \phi_X\left(k - \frac{i}{2\pi}\right) \tag{1.11}$$

*Asset-or-nothing* options may then be computed using the same formalism as *cash-or-nothing* options. Namely, we have

$$\mathcal{C}_{\text{AoN}} = S_t\left\{\frac{1}{2} - \frac{i}{2\pi}\int \frac{\mathrm{d}u}{u}\left[\phi_X\left(u - \frac{i}{2\pi}\right)\mathrm{e}^{-2\pi iuk} - 1\right]\right\} \tag{1.12}$$

for call options and

$$\mathcal{P}_{\text{AoN}} = S_t \left[ \frac{1}{2} + \frac{i}{2\pi} \int \frac{\mathrm{d}u}{u} \left\{ \phi_X \left( u - \frac{i}{2\pi} \right) \mathrm{e}^{-2\pi i u k} - 1 \right] \right\} \tag{1.13}$$

for put options.

### 1.4.5 European options: the general pricing formula

It is now possible to derive a general pricing formula for European options that will be used to calibrate pricing models to market data. Notice that all the information content concerning the dynamics of the risk factor $S$, the underlying asset of our options, is summarized in the function

$$\mathrm{d}(k, \alpha) \equiv \int \frac{\mathrm{d}u}{u} \left( \mathrm{e}^{-2\pi i u k} \phi_X (u - \alpha) - 1 \right) \tag{1.14}$$

We call this function the **characteristic integral** of asset $S$. The probability distribution used in the pricing of all *cash-or-nothing* and *asset-or-nothing* options for all maturities can be synthetically reported with the common notation:

$$D(k, \alpha, \omega) = \frac{1}{2} - \omega \frac{i}{2\pi} \mathrm{d}(k, \alpha) \tag{1.15}$$

Clearly the *cash-or-nothing* case corresponds to $\alpha = 0$ while the *asset-or-nothing* case is covered by $\alpha = i/2\pi$. Furthermore, as stated before, $\omega = 1$ denotes call options, while $\omega = -1$ denotes put.

Adopting this notation for European options, the prices for call or put can be written as:

$$\mathcal{O}(S_t; K, T, \omega) = \omega \left[ S_t D\left( k, \frac{i}{2\pi}, \omega \right) - B(t, T) K D(k, 0, \omega) \right] \tag{1.16}$$

which only depends on the **characteristic integral**. In order to highlight that, the European option pricing formula can be rewritten as

$$\mathcal{O}(S_t; K, T, \omega) = \frac{1}{2} \omega S_t (1 - m) + S_t \frac{i}{2\pi} \left[ \mathrm{d}(k, 0) m - \mathrm{d}\left( k, \frac{i}{2\pi} \right) \right] \tag{1.17}$$

where we recall that $m \equiv B(t, T)K/S_t$ denotes *moneyness* (in the forward price sense). Notice that the **characteristic integral** enters the formula with the same sign for both call and put options. The shape of the smile could then be recovered by using the statistics

$$\frac{\mathcal{C}(m) + \mathcal{P}(m)}{S_t} = \frac{i}{\pi} \left[ \mathrm{d}(k, 0) m - \mathrm{d}\left( k, \frac{i}{2\pi} \right) \right] \tag{1.18}$$

where $\mathcal{C}$ and $\mathcal{P}$ denote call and put options as usual.

Finally, notice that for the *at-the-money forward* option ($m = 1$) we have

$$\mathcal{O}_{\text{AtM}}(S, t; K, T) = S_t \frac{i}{2\pi} \left[ d(0, 0) - \mathrm{d}\left( 0, \frac{i}{2\pi} \right) \right] \tag{1.19}$$

which may be useful to calibrate the term structure of volatility around the most liquid option quotes.

With this general structure we are then ready not only to price options but also to use option prices to back out in a synthetic way all relevant information concerning the dynamics of the underlying assets.

## 1.5 HILBERT TRANSFORM

We are now going to show that the **characteristic integral** defined above can be represented in an alternative way, resorting to what is known as the *Hilbert transform*. This technique was recently applied to the option pricing problem by Feng and Linetsky (2008).

The Hilbert transform $\mathcal{H}f$ of a function $f$ is obtained by performing the convolution of the function with the distribution p.v.$1/x$, in formula:

$$[\mathcal{H}f](y) = \frac{1}{\pi}\int \mathrm{d}x\ f(x)\,\text{p.v.}\left[\frac{1}{y-x}\right]$$

If we call $h(x)$ the tempered distribution:

$$h(x) = \text{p.v.}\frac{1}{\pi x}$$

we may define the Hilbert transform by the alternative notation:

$$\mathcal{H}f = h \star f$$

We can immediately see that the characteristic integral defined above, and yielding the prices of options, can be written in terms of the Hilbert transform

$$\begin{aligned}
\mathbb{Q}(k) &= \int \mathrm{d}u\, \mathrm{e}^{-2\pi iku}\phi_X(u-\alpha)\delta^+(u) \\
&= \int \mathrm{d}u\, \mathrm{e}^{-2\pi iku}\phi_X(u-\alpha)\left(\frac{1}{2}\delta(u) + \frac{i}{2\pi}\text{p.v.}\left(\frac{1}{u}\right)\right) \\
&= \frac{1}{2} + \frac{i}{2\pi}\int \mathrm{d}u\, \mathrm{e}^{-2\pi iku}\phi_X(u-\alpha)\text{p.v.}\left(\frac{1}{u}\right) \\
&= \frac{1}{2} + \frac{1}{2i}[\mathcal{H}f_k](0), \quad \text{where } f_k : u \to \mathrm{e}^{-i2\pi ku}\phi x(u-\alpha).
\end{aligned} \tag{1.20}$$

In order to compute Hilbert transforms of the quantities in which we are interested in the development of this chapter, we anticipate some relations that will be presented in Chapter 5:

$$\text{p.v.}\left(\frac{1}{x}\right) = \frac{1}{x - i\epsilon} - i\pi\delta(x) \ = \frac{1}{x + i\epsilon} + i\pi\delta(x)$$

We then get:

$$[\mathcal{H}f](y) = \frac{1}{\pi}\int \mathrm{d}x\ \frac{f(x)}{y - x - i\epsilon} - if(y)$$

as a general rule to compute the Hilbert transform.

Adopting the usual *hat* notation for the Fourier transform we can write:

$$[\mathcal{H}f](y) = \overline{\mathcal{F}}\left(\hat{h}\hat{f}\right) \tag{1.21}$$

where

$$\hat{h} := \mathcal{F}\left(\text{p.v.}\frac{1}{u}\right)$$

A result that will be needed in the development is the Fourier transform of p.v.$(1/u)$.

**Example 1.5.1** *From the definition of $\hat{h}$ we get:*

$$\begin{aligned}\hat{h}(k) &= \frac{1}{\pi}\int \mathrm{d}x\, \frac{\mathrm{e}^{i2\pi kx}}{x - i\epsilon} - i \\ &= 2i\theta(k) - i \\ &= i\,\mathrm{sign}(k)\end{aligned}$$

*where the "signum" function "sign" is defined by:*

$$\mathrm{sign}(x) = \begin{cases} 1 & x > 0 \\ -1 & x < 0 \end{cases}$$

We now provide a set of examples that should (a) illustrate how to compute the Hilbert transform of functions and (b) lead to a formula that will be paramount in the development of the numerical implementation.

**Example 1.5.2** *Consider the function $\mathrm{e}_\beta : x \to \mathrm{e}^{i2\pi\beta x}$, following the definition we have:*

$$\begin{aligned}[\mathcal{H}\mathrm{e}_\beta](y) &= -i\mathrm{e}^{i2\pi\beta y} + \frac{1}{\pi}\int \mathrm{d}x\, \frac{\mathrm{e}^{i2\pi\beta x}}{y - x - i\epsilon} \\ &= -i\mathrm{e}^{i2\pi\beta y} + 2i\,\mathrm{e}^{i2\pi\beta y}\theta(-\beta) \\ &= -i\mathrm{e}^{i2\pi\beta y}\,\mathrm{sign}(\beta)\end{aligned}$$

*There is also a second method available (as in most cases) to get to the result. The method exploits equation* (1.21)*. We observe that*

$$\hat{h}(u) = i\,\mathrm{sign}(u), \quad \hat{e}_\beta(u) = \delta(u + \beta)$$

*therefore:*

$$\begin{aligned}[\mathcal{H}e_\beta](y) &= \overline{\mathcal{F}}\left(\hat{h}\hat{e}_\beta\right) \\ &= i\int \mathrm{d}u\, e^{-i2\pi uy}\mathrm{sign}(u)\delta(u + \beta) \\ &= -i\mathrm{e}^{i2\pi\beta y}\,\mathrm{sign}(\beta)\end{aligned}$$

We can exploit the linearity of the Hilbert transform and the result in the example above to recover the transform of trigonometric functions.

**Example 1.5.3** *Let $s_m : x \to \sin(mx)$. If we set $\mu = m/2\pi$, from the definition we get:*

$$\begin{aligned}[\mathcal{H}s_m](y) &= \frac{1}{2i}[\mathcal{H}e_\mu](y) - \frac{1}{2i}[\mathcal{H}e_{-\mu}](y) \\ &= -\frac{1}{2}\,\mathrm{e}^{imy}\,\mathrm{sign}(m) + \frac{1}{2}\,\mathrm{e}^{-imy}\,\mathrm{sign}(-m) \\ &= -\cos(my)\,\mathrm{sign}(m)\end{aligned}$$

We are now ready to compute the Hilbert transform of a function that will be crucial in the numerical applications below.

**Example 1.5.4** *Let's consider the function:*

$$\text{sinc}_m : x \to \frac{\sin(mx)}{x}$$

*then:*

$$[\mathcal{H}\text{sinc}_m](y) = -i\frac{\sin(my)}{y} - \frac{1}{\pi}\int \mathrm{d}x\, \frac{\sin(mx)}{x[x-(y-i\epsilon)]}$$

*Exploiting the relation:*

$$\frac{1}{x}\,\frac{1}{x-(y-i\epsilon)} = \frac{1}{y}\left[\frac{1}{x-(y-i\epsilon)} - \frac{1}{x}\right]$$

*we have:*

$$[\mathcal{H}\text{sinc}_m](y) = -i\frac{\sin(my)}{y} - \frac{1}{\pi y}\int \mathrm{d}x\, \sin(mx)\left[\frac{1}{x-(y-i\epsilon)} - \frac{1}{x}\right]$$

*The integral*

$$\int \mathrm{d}x\, \frac{\sin(mx)}{x}$$

*is finite, so among the different ways to compute it, one particularly convenient for us is to replace it with its principal value (since it is finite, its value must coincide with its principal value). With this understanding we get*

$$\begin{aligned}[\mathcal{H}\text{sinc}_m](y) &= \frac{1}{y}[\mathcal{H}\sin](y) - \frac{1}{y}[\mathcal{H}\sin](0)\\ &= \text{sign}(m)\frac{1-\cos(my)}{y}\end{aligned}$$



## 1.6 PRICING VIA FFT

We are now going to address the numerical issues involved in the application of Fourier pricing methods to market data.

It is quite clear that all the numerical work needed to compute prices for vanilla options consists in performing the *characteristic integral* as defined in equation (1.14). For later convenience we will introduce now a change in notation:

$$\mathrm{d}(k,\alpha) = \int_{-\infty}^{+\infty} \mathrm{d}u\, \frac{f(u,k,\alpha)-1}{u} \tag{1.22}$$

or equivalently, in terms of the Hilbert transform,

$$\mathrm{d}(k,\alpha) = \int_{-\infty}^{+\infty} \mathrm{d}u\, f(u,k,\alpha)\,\text{p.v.}\left(\frac{1}{u}\right) \tag{1.23}$$

where

$$f(u,k,\alpha) = \mathrm{e}^{-2\pi iuk}\phi_X(u-\alpha) \tag{1.24}$$

As we shall see in the following sections, there are powerful numerical methods to compute, with great accuracy, the Hilbert transform of a characteristic function.

Having devised a method to compute the integral, the problem is how to compute many of such integrals in a run. This problem is particularly relevant in finance. In fact, nowadays we have plenty of information concerning not only the historical dynamics of market data, but also the *forward-looking* dynamics of the distribution implied by market data, even though the two sources of information are referred to different probability measures and so are not directly comparable. There are many instances, both in time series and cross-section analysis, in which it is required to compute many prices by inversion of the Fourier transform. In this case, a well known technique, called *Fast Fourier Transform* (FFT), is typically applied. At the end of the section we will address how to cast the computation of the characteristic integral in a FFT setting.

### 1.6.1 The sampling theorem

We start now to develop the theory concerning the numerical integration of the *characteristic integral*.

We recall the fundamental relation between a function $p(x)$ and its Fourier transform.

$$p(x) = \int_{-\infty}^{+\infty} \mathrm{d}u\, \hat{p}(u)\, \mathrm{e}^{-i2\pi ux}, \quad \hat{p}(u) = \int_{-\infty}^{+\infty} \mathrm{d}x\, p(x)\, \mathrm{e}^{i2\pi ux}$$

In the language of the previous section, and later as well, $p(x)$ is the p.d.f. of some stochastic process at a given time $t$ and $\hat{p}(u)$ is its characteristic function. More precisely, the variable $x$ would represent the log return of the process. A first step to be taken while landing from the realm of theory to applications is that for any practical purpose we are required to restrict the support of this variable, which is typically taken to be unbounded, to a bounded subset. This justifies the change of notation from $\phi_X(x)$, the characteristic function of the process, to $\hat{p}(x)$ as a characteristic function of the density defined on a bounded support. We want to use the latter as an approximation for the former, so that outside the support of $p(x)$ the value of the true probability distribution function is so close to zero that it can be considered zero for any practical purpose. In other words, we are saying that there exists a value $X_c$ such that:

$$p(x) < \epsilon, \quad |x| > X_c$$

So, if the value of the asset is normalized with respect to its price today, even a modest value of $X_c$ such as 4 means that we give a negligible probability to moves beyond 140% or below 60%, and this may be large enough, particularly if we are not looking at extremely long time intervals and at times of normal volatility.

From this point on, we then substitute $\phi_X(x)$ with a function $p(x)$ such that $p(x) = 0$ for $|x| > X_c$. Therefore the characteristic function is given by:

$$\hat{p}(u) = \int_{-X_c}^{+X_c} \mathrm{d}x\, p(x)\, \mathrm{e}^{i2\pi ux} \tag{1.25}$$

**Example 1.6.1** *To gain some insight into what happens when the p.d.f. is (nearly) zero outside its bounded domain, as described above, we make things very simple and assume that*

$$p(x) = 1, \quad |x| < X_c, \quad p(x) = 0, \quad |x| > X_c$$

*Despite its simplicity, it will turn out that this example is extremely useful, so the reader is well advised to work through it until a good grasp is achieved. Performing the simple integral we obtain:*

$$\hat{p}(u) := \int_{-X_c}^{+X_c} \mathrm{d}x \, \mathrm{e}^{i2\pi ux} = 2X_c \, \mathrm{sinc}(2\pi X_c u)$$

*where we have adopted the definition for the "sinc" functions as:*

$$\mathrm{sinc}(x) = \frac{\sin(x)}{x}$$

Let us now define

$$\Delta := \frac{1}{2X_c}, \quad u_n := n\Delta, \quad \hat{p}_n := \hat{p}(u_n)$$

then we can compute the l.h.s. in equation (1.25) only at the (sampling) values $u_n$

$$\hat{p}_n = \int_{-X_c}^{+X_c} \mathrm{d}x \, p(x) \, \mathrm{e}^{i2\pi nx\Delta}$$

From the theory of Fourier series, we know that we can use the sequence $\{\hat{p}_n\}$ to get back the function $p(x)$, and the inversion formula is given by:

$$p(x) = \frac{1}{2X_c} \sum_{n=-\infty}^{+\infty} \hat{p}_n \, \mathrm{e}^{-i2\pi nx\Delta}, \quad |x| < X_c$$

We can get rid of the explicit constraint $|x| < X_c$ resorting to the indicator function and write:

$$p(x) = \frac{1}{2X_c} \mathbf{1}_{[|x|<X_c]} \sum_{n=-\infty}^{+\infty} \hat{p}_n \, \mathrm{e}^{-i2\pi \Delta nx}$$

The original function $\hat{p}(u)$ can be recovered by applying Fourier transform to $p(x)$:

$$\hat{p}(u) = \frac{1}{2X_c} \sum_{n=-\infty}^{+\infty} \hat{p}_n \int_{-\infty}^{+\infty} \mathrm{d}x \, \mathbf{1}_{[|x|<X_c]} \, \mathrm{e}^{i2\pi x(u-n\Delta)}$$

The remaining integral

$$\int_{-\infty}^{+\infty} \mathrm{d}x \, \mathbf{1}_{[|x|<X_c]} \, \mathrm{e}^{i2\pi x(u-n\Delta)}$$

is nothing but the integral performed in Example 1.6.1 and the result is:

$$\frac{\sin[2\pi X_c(u - n\Delta)]}{\pi(u - n\Delta)}$$

Then, we can conclude that the whole Fourier spectrum of the function $p(x)$ with bounded domain is given by:

$$\hat{p}(u) = \frac{1}{2X_c} \sum_{n=-\infty}^{+\infty} \hat{p}_n \frac{\sin[\,2\pi X_c(u - n\Delta)\,]}{\pi(u - n\Delta)} \tag{1.26}$$

This remarkable result, also known as the sampling theorem, shows that the Fourier transform $\hat{p}(u)$ of a function with bounded domain can be fully known provided it is known at discrete sampling points.

### 1.6.2 The truncated sampling theorem

Numerical approximations will be introduced, replacing the infinite sum with a finite one.

$$\hat{p}^N(u) = \frac{1}{2X_c} \sum_{n=-N}^{+N} \hat{p}_n \frac{\sin[\,2\pi X_c(u - n\Delta)\,]}{\pi(u - n\Delta)} \tag{1.27}$$

We will discuss the type of error introduced by this truncation in the next section, after presenting the final result for the computation of the *characteristic integral*. Presently we limit ourselves to a numerical verification of the accuracy of the truncated sampling theorem. The whole foundation of the approach adopted in this book is that, for many interesting models, the characteristic function is easy to compute. Accordingly we know the exact form of the l.h.s. of equation (1.26) and we are in a position to check the accuracy of the approximation produced by the r.h.s. of equation (1.27) when we select different values for the bound $X_c$ and different values for N.

The measure that we propose to represent the error in the representation of characteristic function consists in looking at the quantity:

$$\mathrm{d}_N(X_c) := \sqrt{\sum_{i=0}^{n} |\varphi_X(-x_{\min} + si) - \hat{p}_N(-x_{\min} + si)|^2}$$

where $n$ is the number of points in which the distance between the two functions is computed, $x_{\min}$ is the lowest value of $x$ where the comparison is made, and $s$ is the increment of $x$ from one point to the latter. This distance is computed for fixed $X_c$ as a function of $N$ and for fixed $N$ as a function of $X_c$. In the former case, provided we select $X_c$ large enough, this will give us insight on the number of Fourier modes needed to achieve the desired accuracy. In the latter case, provided we take $N$ large enough, we may gauge the values of $X_c$ for which the p.d.f. can be considered negligibly small when $|x| > X_c$.

As an example, in Figure 1.1 we look at a simple diffusion model with $\sigma = 0.4423$, $X_c = 4.0$. We see that we reach machine precision with as little as 60 Fourier modes, while in Figure 1.2 we look at the same model but keep fixed the number of Fourier modes at $N = 64$. We see that we can consider negligible the p.d.f. for values $|x| > 4.0$.

In Figures 1.3 and 1.4 we present the same model but with $\sigma = 0.1$. Smaller volatility means a narrower distribution, so we do expect to be able to use a much lower cutoff $X_c$. As we can see, in fact we reach machine precision for $N \geq 50$ and $X_c < 1.0$.

The reader is warmly invited to run the same test for the case $\sigma = 0.4423$, keeping the spatial cutoff at $X_c = 1.0$. It should not come as a surprise that no amount of Fourier modes will be able to reduce the error to acceptable values.

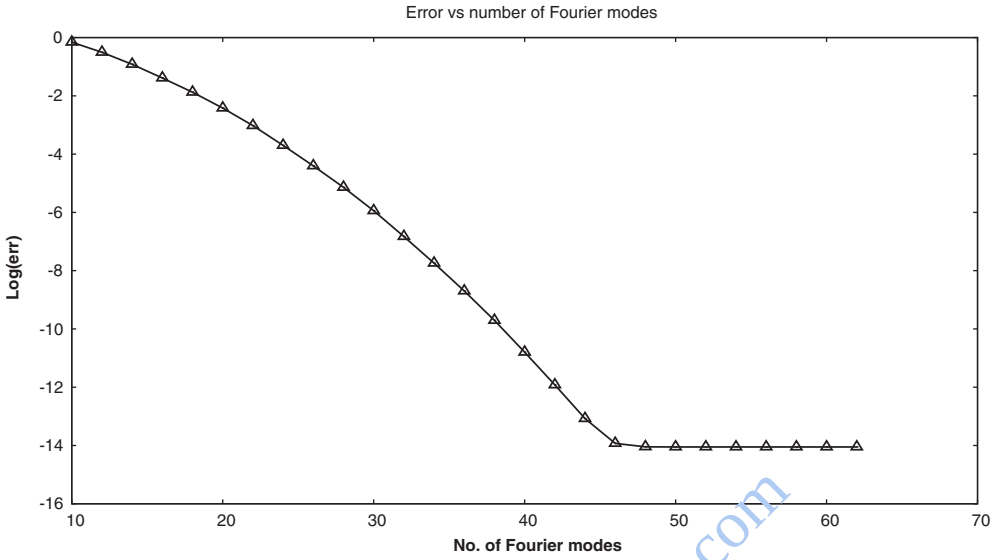


**Figure 1.1** The dependency of the error on the number of Fourier modes for the truncated sampling theorem. The model used is simple diffusion with $\sigma = 0.4423$, $T = 1$ year, the spatial cutoff is $X_c = 4.0$



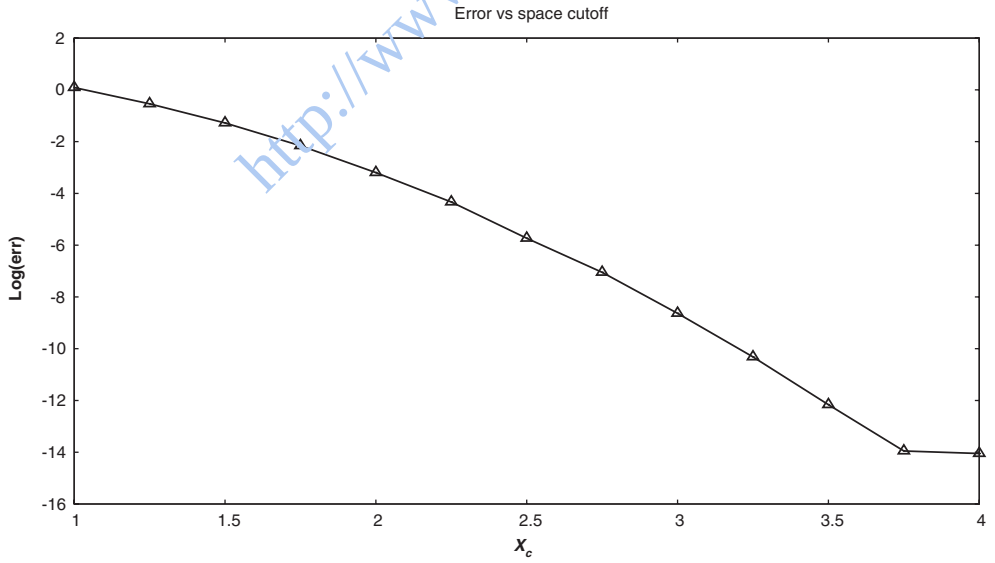


**Figure 1.2** The dependency of the error on the spatial cutoff for the truncated sampling theorem. The model used is simple diffusion with $\sigma = 0.4423$, $T = 1$ year. The number of Fourier modes used is $N = 64$

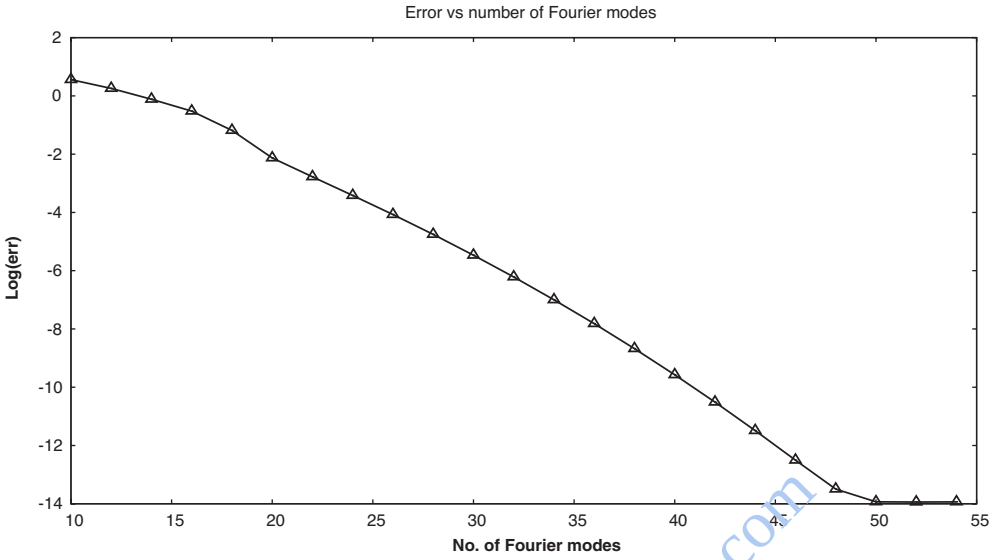


**Figure 1.3** The dependency of the error on the number of Fourier modes for the truncated sampling theorem. The model used is simple diffusion with $\sigma = 0.1$, $T = 1$ year, the spatial cutoff is $X_c = 1.0$

As a final example we run the same test on the Heston model. The results are presented in Figures 1.5 and 1.6. The parameters of the model are detailed in the figure captions. Also in this case we see that with a judicious choice of the spacial cutoff we can reach machine accuracy.



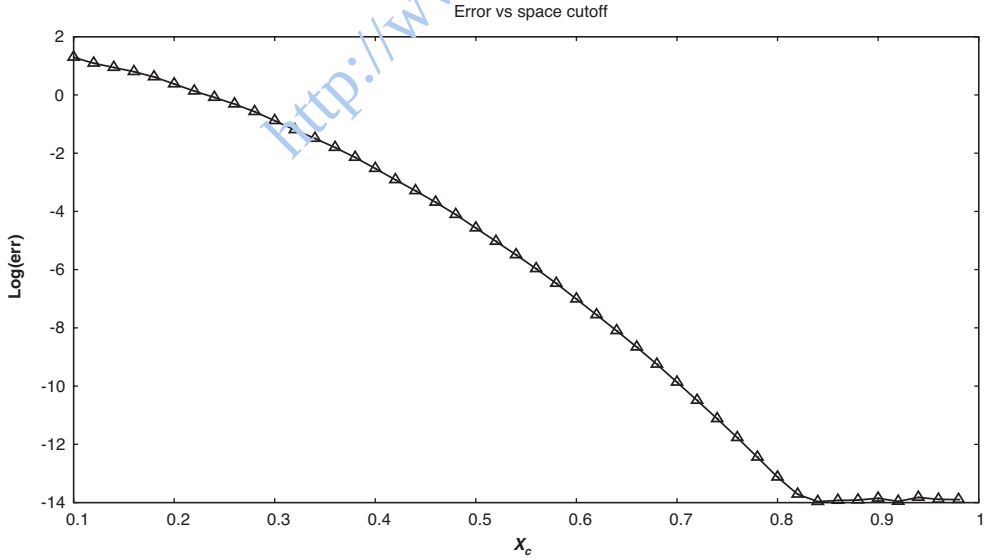


**Figure 1.4** The dependency of the error on the spatial cutoff for the truncated sampling theorem. The model used is simple diffusion with $\sigma = 0.1$, $T = 1$ year. The number of Fourier modes used is $N = 56$

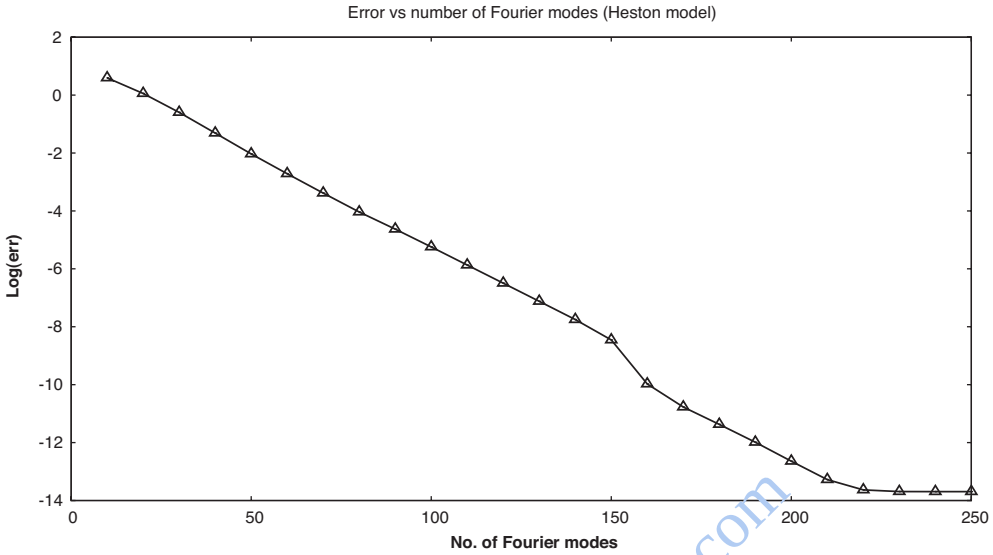


**Figure 1.5** The dependency of the error on the number of Fourier modes for the truncated sampling theorem. The model used is the Heston model with $\eta = 0.256$, $\lambda = 1.481$, $v_0 = 0.2104$, $\overline{v} = 0.1575$, $\rho = -0.8941$, $T = 1$ year. The spatial cutoff is $X_c = 8.0$



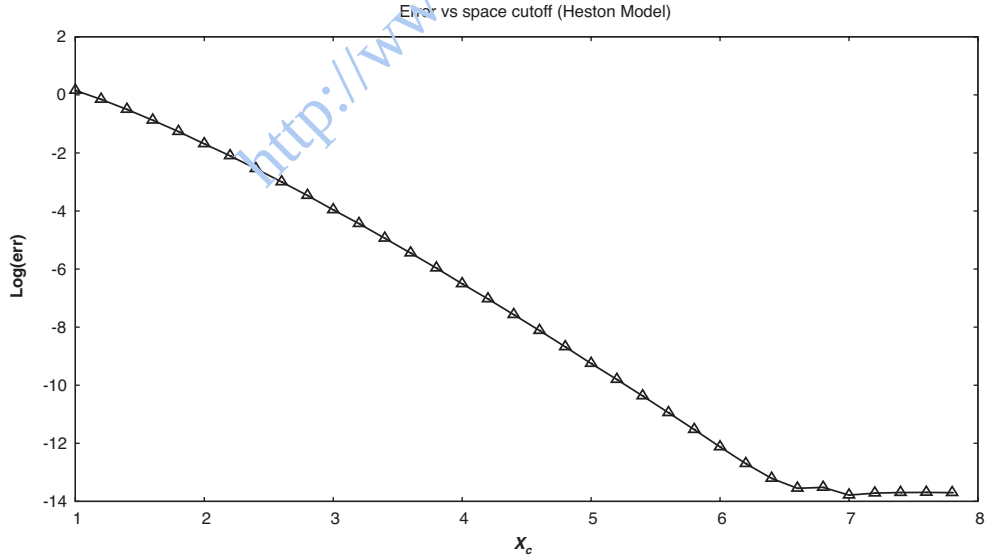


**Figure 1.6** The dependency of the error on the spatial cutoff for the truncated sampling theorem. The model used is the Heston model with $\eta = 0.256$, $\lambda = 1.481$, $v_0 = 0.2104$, $\overline{v} = 0.1575$, $\rho = -0.8941$, $T = 1$ year. The number of Fourier modes used is $N = 256$

### 1.6.3 Why bother?

The wily reader might have nursed a cunning question. If we can get "exactly" the characteristic function of the process under examination, why bother to compute it via its representation given by the sampling theorem? The answer rests on the fact that, for pricing purposes, we need to compute the convolution of the characteristic function with the distribution p.v.$1/u$. In general this cannot be computed in closed form for the characteristic function of most models. A naive numerical integration of that convolution would prove to be highly delicate due to the oscillatory nature of the characteristic function itself, and is emphatically *something that should not be done*.

On the contrary, the sampling theorem gives us a nice and exact representation in term of the "sinc" function, and the convolution of this function with p.v.$1/u$ is something we can compute. In equation (1.24) the function $\phi_X(u-\alpha)$ is the characteristic function of some p.d.f. $p(x)$ with bounded (approximately bounded) support. An immediate result from Fourier transform theory is that

$$\mathrm{e}^{-i2\pi ku}\phi_X(u-\alpha)$$

is the characteristic function of the p.d.f. $p(x-k)$ that is again a p.d.f. with approximately bounded support, so we can resort to the sampling theorem to represent it.

From equations (1.24) and (1.26), and the considerations expressed above, we see that the *characteristic integral* can be written as:

$$\mathrm{d}(k,\alpha) = \frac{1}{2X_c}\sum_{n=-\infty}^{+\infty} \mathrm{e}^{-i2\pi n\Delta k}\phi_X(n\Delta-\alpha)\int_{-\infty}^{+\infty}\mathrm{d}u\,\frac{\sin[\,2\pi X_c(u-n\Delta)\,]}{\pi(u-n\Delta)}\mathrm{p.v.}\left(\frac{1}{u}\right) \tag{1.28}$$

The integral can now be performed. It is recognized as the Hilbert transform of the "sinc" function, and having done this we have disposed of the most delicate part of the numerical integration and are left with an infinite sum over discretely sampled values. Since this sum will be related in a straightforward manner with the sum coming from Fourier series, we will have at our disposal all of the tools to control the accuracy of the approximation introduced in replacing the infinite sum with a finite sum.



### 1.6.4 The pricing formula

The discussion above lead us to the conclusion that the numerical integration of the *characteristic integral* is equivalent to the computation of the r.h.s. of equation (1.28).

Let us concentrate on the integral on the r.h.s.

$$I = \int_{-\infty}^{+\infty}\mathrm{d}u\,\frac{\sin[\,2\pi X_c(u-n\Delta)\,]}{\pi(u-n\Delta)}\left(\mathrm{p.v.}\frac{1}{u}\right)$$

performing the change of variables $u := -v/2\pi X_c + n\Delta$ we get:

$$\begin{aligned} I &= \int_{-\infty}^{+\infty}\mathrm{d}v\,\frac{\sin(\,v\,)}{\pi\,v}\mathrm{p.v.}\left(\frac{1}{n\Delta - v/2\pi XC}\right) \\ &= \left(\frac{\pi}{\Delta}\right)\frac{1}{\pi}\int_{-\infty}^{+\infty}\mathrm{d}v\,\frac{\sin(\,v\,)}{v}\mathrm{p.v.}\left(\frac{1}{\pi n - v}\right) \end{aligned}$$

where in the last equality we have made use of the definition $\Delta = 1/2X_c$, and we recognize the integral on the r.h.s. as the Hilbert transform of the "sinc" function:

$$I = \frac{\pi}{\Delta}[\mathcal{H}\,\mathrm{sinc}_1](n\pi).$$

This is a result that we know from Example 1.5.4 and:

$$I = \frac{\pi}{\Delta}\frac{1-\cos(n\pi)}{n\pi} = \frac{1-(-1)^n}{n\Delta}.$$

Having done this we have achieved an amazingly accurate formula by which to compute the *characteristic integral* numerically

$$\mathrm{d}(k,\alpha) = \frac{1}{2X_c}\sum_{n=-\infty}^{+\infty} \mathrm{e}^{-i2\pi n\Delta k}\phi_X(n\Delta-\alpha)\frac{1-(-1)^n}{n\Delta} \tag{1.29}$$

It is worth stressing once more that, apart from the assumption that the p.d.f. of the process under examination has bounded domain, this is an exact integration formula.

Some approximation arises when we decide to introduce a cutoff in the number of Fourier modes that we use. The terms

$$\mathrm{e}^{-i2\pi n\Delta k}\phi_X(n\Delta-\alpha)$$

are the Fourier coefficients for a function with approximately bounded support. Without proof we will quote the following well-known theorem:

**Theorem 1.6.1** *Let $p(x)$ a function with support in the interval $I_c = [-X_c, X_c]$.*

$$p(x) = 0, \quad x \notin I_c$$

*Let*

$$c_k = \int_{-X_c}^{X_c} \mathrm{d}x\, \mathrm{e}^{i2\pi xk}\, p(x)$$

*If $p(x) \in C^q$ then*

$$\sum_{n=-\infty}^{+\infty} |n^q c_n| < \infty$$

*in particular*

$$\lim_{n\to\infty} n^q c_n = 0$$

The meaning of this theorem is that the truncation error we are going to incur depends on the smoothness property of the p.d.f. of the process under examination (actually, it depends on the smoothness at $x = 0$).

The general smoothness property of a generic model cannot be assessed in advance without knowledge of the model, so the issue concerning truncation errors has to be addressed from case to case relative to each individual model.

For the time being we replace the infinite sum with a truncated sum,

$$\mathrm{d}_N(k,\alpha) = \sum_{n=-N/2}^{+N/2} \mathrm{e}^{-2\pi i n k \Delta} \phi_X(n\Delta - \alpha) \frac{1-(-1)^n}{n\Delta} \tag{1.30}$$

and (7.1), the fundamental pricing equation, is modified accordingly:

$$\mathcal{O}(S_t; K, T, \omega) = \frac{1}{2}\omega S_t(1-m) + S_t \frac{i}{2\pi}\left(\mathrm{d}_N(k,0)m - \mathrm{d}_N\left(k, \frac{i}{2\pi}\right)\right) \tag{1.31}$$

To steer clear of any form of circular reasoning, the assessment of the quality of the numerical approximation embedded in equation (1.31) can be performed with arbitrary accuracy only for models that admit an analytical solution that is NOT obtained by performing a Fourier integral.

For the sake of providing a simple example, we report some results for the Black–Scholes model. More precisely, Figures 1.7 and 1.8 report results for a volatility of $\sigma = 0.4423$, and Figures 1.9 and 1.10 for a volatility of $\sigma = 0.1$.

### 1.6.5 Application of the FFT

The next issue to address concerns the best way to perform the finite sum in equation (1.30). If all we need is just one value of the option at fixed strike, the issue is non-existent, and we simply sum the terms exactly as they are described in equation (1.30). Whenever we need to extract a larger set of results – quite a common situation when calibrating a model – it might be



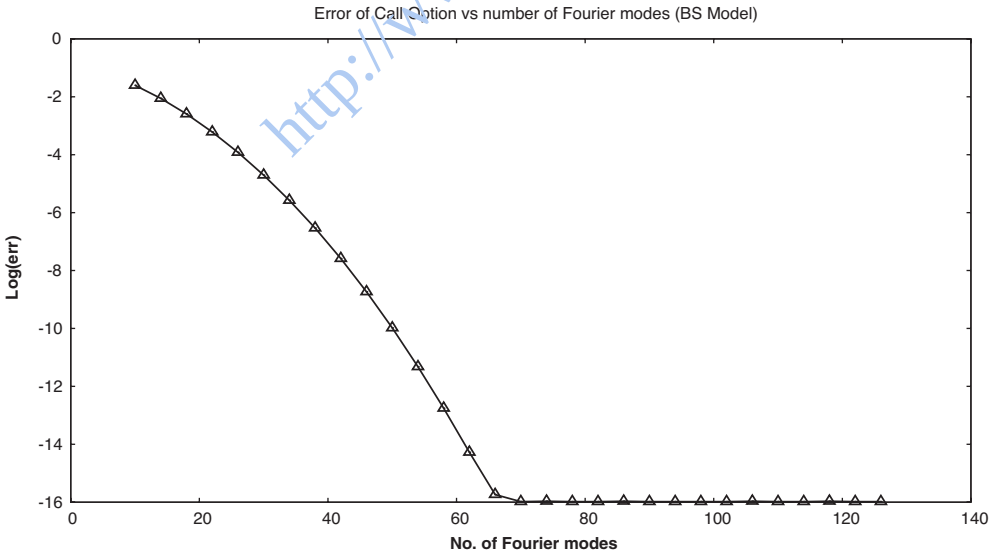


**Figure 1.7** The dependency of the error on the number of Fourier modes used to compute a call option. Parameters are $T = 1$, $K = 1.0$, $\sigma = 0.4423$, $r = 0.05$. The spatial cutoff is $X_c = 6.0$

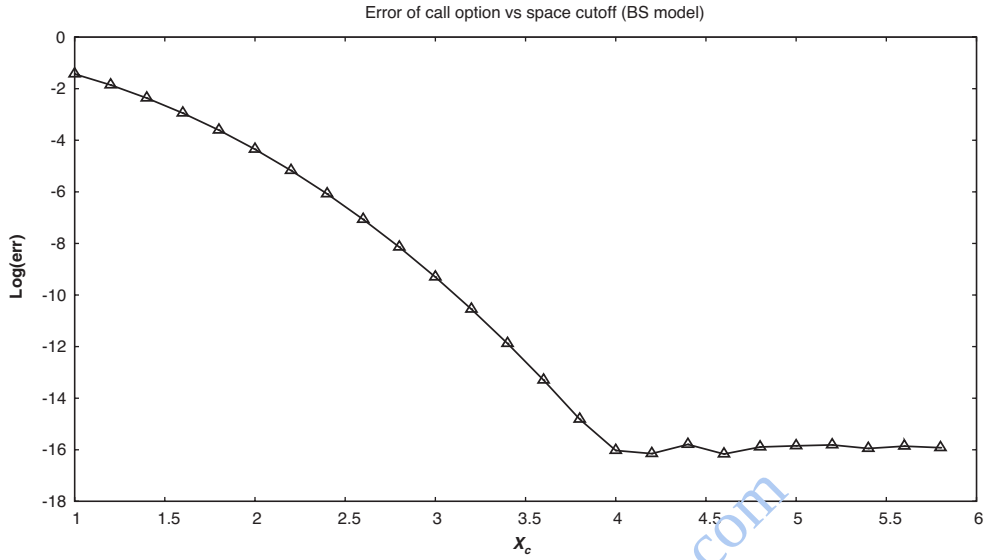


**Figure 1.8** The dependency of the error on the spatial cutoff used to compute a call option. Parameters are $T = 1$, $K = 1.0$, $\sigma = 0.4423$, $r = 0.05$. The number of Fourier modes is 128

convenient to resort to the fast Fourier transform (FFT). The FFT can compute in $o(N \log(N))$ operation the sum in (1.30) for a set on $N$ values $k_1 \ldots, k_N$ of the strike $k$, provided that:

- $N$ can be written as $2^l$.
- We confine the computation to the values $k_q = q/N\Delta$.



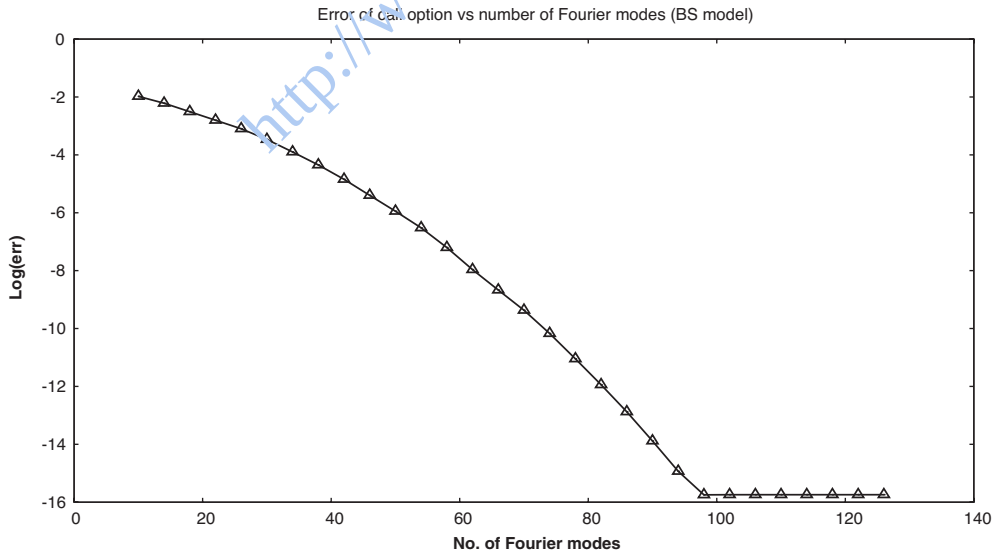


**Figure 1.9** The dependency of the error on the number of Fourier modes used to compute a call option. Parameters are $T = 1$, $K = 1.0$, $\sigma = 0.1$, $r = 0.05$. The spatial cutoff is $X_c = 2.0$

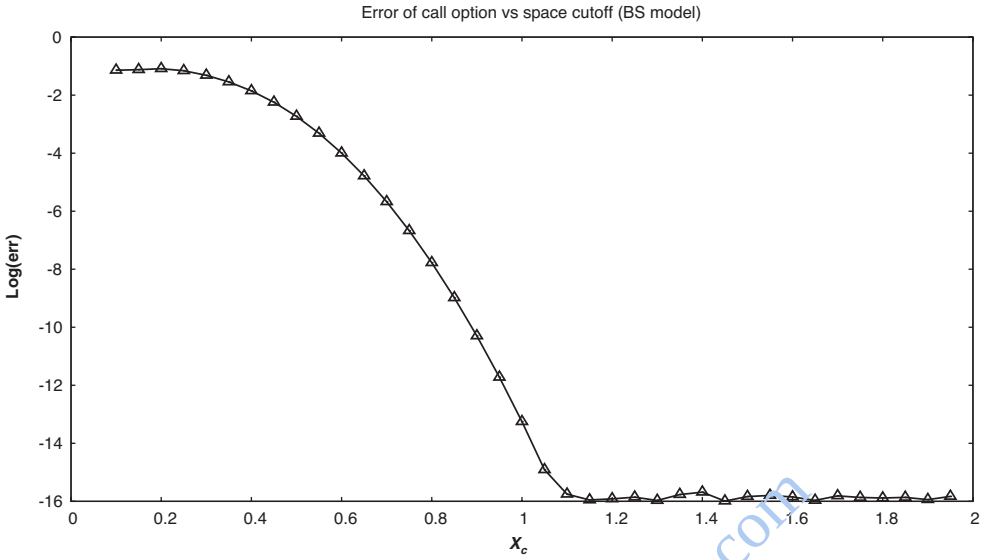


**Figure 1.10** The dependency of the error on the spatial cutoff used to compute a call option. Parameters are $T = 1$, $K = 1.0$, $\sigma = 0.1$, $r = 0.05$. The number of Fourier modes is 128



Let

$$\zeta_n := \phi_X(n\Delta)\frac{1-(-1)^n}{n\Delta}$$

if we consider only FFT compliant strikes we can write:

$$\mathrm{d}_N\left(\frac{q}{N\Delta},\alpha\right) = \sum_{n=-N/2}^{+N/2} \mathrm{e}^{-\frac{2\pi inq}{N}}\zeta_n$$

then we separate positive and negative frequencies:

$$\mathrm{d}_N\left(\frac{q}{N\Delta},\alpha\right) = \sum_{n=-N/2}^{1} \mathrm{e}^{-\frac{2\pi inq}{N}}\zeta_n + \sum_{n=0}^{N/2-1} \mathrm{e}^{-\frac{2\pi inq}{N}}\zeta_n + \mathrm{e}^{-\pi iq}\zeta_{N/2}$$

The last term is clearly zero ($N/2$ is even, in the working hypothesis), and the first term can be written as

$$\sum_{n=-N/2}^{1} \mathrm{e}^{-\frac{2\pi inq}{N}}\zeta_n = \sum_{n=N/2}^{N-1} \mathrm{e}^{-\frac{2\pi i(n-N)q}{N}}\zeta_{n-N} = \sum_{n=N/2}^{N-1} \mathrm{e}^{-\frac{2\pi inq}{N}}\zeta_{n-N}$$

Finally,

$$\mathrm{d}_N\left(\frac{q}{N\Delta},\alpha\right) = \sum_{n=0}^{N-1} \mathrm{e}^{-\frac{2\pi inq}{N}}\tilde{\zeta}_n \tag{1.32}$$

where

$$\tilde{\zeta}_n = \begin{cases} \zeta_n & 0 \le n < N/2 \\ \zeta_{n-N} & N/2 \le n < N \end{cases}$$

$$\zeta_n := \phi_X(n\Delta)\frac{1-(-1)^n}{n\Delta}$$

*Small practical matters*

The periodicity of the sum on the r.h.s. of equation (1.32 ) implies a periodicity of the l.h.s. of the same equation. This observation is what is required to extract the correct frequencies from the FFT sum that in fact turn out to be:

$$k_q = \begin{cases} \frac{q}{N\Delta} & k_q > 0 \\ \frac{q-N}{N\Delta} & k_q < 0 \end{cases}$$

When we use the FFT algorithm for calibration, usually we cannot choose the strikes that we want to calibrate. Our procedure is to compute the array of strikes at the FFT values $k_q$ and interpolate linearly for the desired ones. This opens up the question of whether the FFT strikes are dense enough to populate reasonably the range needed. The best resolution we can achieve is given by:

$$\delta k = \frac{2X_c}{N}$$

The smallest $X_c$ is dictated by the "boundedness" of the domain of the p.d.f. so we cannot use that as a free parameter; therefore if, for a given $N$, the resolution turns out to be too coarse, we have only two options:

- increase the number of Fourier nodes even though the selected $N$ is large enough for the desired accuracy;
- switch to the "fractional FFT" (FFFT) that allows for a different discretization of the FFT strikes.

The fractional FFT turns out to be, on average, four times slower that the straight FFT, so alternatives must be weighted carefully if performance is an issue. The basic ideas underlying the fractional FFT are presented in Appendix F.



## 1.7 RELATED LITERATURE

We provide here a very general review of the literature on Fourier transform applications to option pricing problems. We stick to a mandatory reading list on the subject with a particular focus on aspects of this literature that are related to our approach.

To the best of our knowledge, the gold rush to Fourier transform pricing applications was initiated by Heston (1993). Our approach shares the same philosophy of searching a relationship between the characteristic function of the pricing kernel of an underlying asset. In this sense, our work is also in the line of literature of Bakshi and Madan (2000) and Duffie *etal*. (2000) both of which define a spanning structure of the pricing kernel based on Fourier transforms. We denote by Arrow–Debreu prices the discounted value of the density instead of

the digital options, but that is a mere question of taste, to keep a similarity with the binomial model.

Carr and Madan (1999) proposed a technique to represent the price of a plain vanilla option in terms of Fourier transforms, in such a way as to have a model that was well suited for application of the FFT technique. For this purpose, they addressed the problem of performing the Fourier transform of the payoff function with respect to the strike. This is then substituted in the pricing integral and, by a change of order of integration, produces the price of the option as a function of the characteristic function of the density. Lewis (2001) addressed the problem of computing the Fourier transform of the payoff function in a more general setting. Differently from Carr and Madan (1999), the Fourier transform is computed with respect to the underlying asset. With this technique, Lewis provides a pricing formula that is valid for general payoffs, including the pricing kernels that had represented the focus of the first stream of literature.

Our approach blends most of the features of the literature that we have so brutally reviewed. For one thing, our attention is focused on the pricing kernel, as in the first stream of literature quoted above. For another, our focus is on disentangling the payoff of this digital option from the characteristic function in the pricing formula. Differently from Lewis (2001), we are only interested in the pricing kernel, because our task is to use the model for calibration. While this interest in calibration recalls the contribution by Carr and Madan (1999), our focus is on digital instead of European options, even though we finally obtain pricing formulas for European options that can be applied in a FFT procedure to perform calibration to market data. What we think is original with respect to the literature is that our approach is cast in the framework of generalized functions, in which the Fourier transform of singular functions, such as the payoff of digital options (which is the core of our approach) are well defined, and so is the convolution of these payoffs with pricing density.