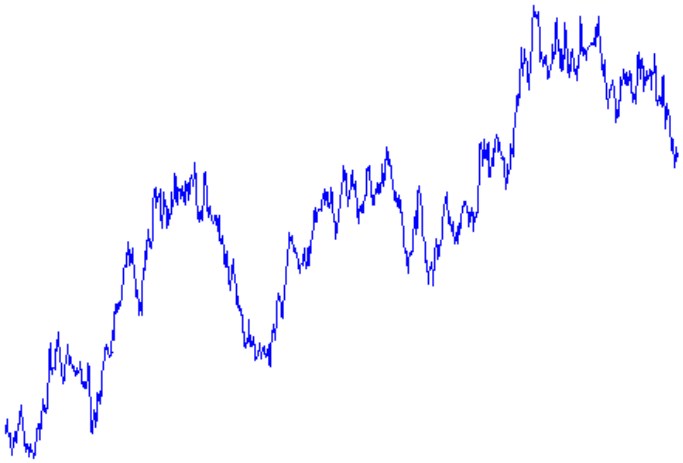
Option Pricing via the FFT

and its Application to Calibration



ManWo Ng
M.W.Ng@EWI.TUDelft.NL
M.Sc. Thesis
Delft University of Technology, Delft, The Netherlands
March 2005

To everyone that supported me during the materialization of my 'American Dream'.

Foreword

This thesis has been submitted in partial fulfillment of the requirements of the Master's degree in Applied Mathematics at Delft University of Technology, Delft, The Netherlands. The reader is supposed to be conversant with simple Black-Scholes option pricing theory.

Without the support of several people, my Master's degree won't have come that fast. For this, I want to express my sincere gratitude to Dr. J.A.M. van der Weide and Dr. F.M. Dekking, who assisted me extensively during my first steps in the world of mathematics. These steps were indeed backbreaking for a mechanical engineer.

The research for this thesis has been done at the Modeling and Research, Credit and Derivatives Department of Rabobank International, Utrecht, The Netherlands. Initially, there were plenty 'logistical' problems. Later on, things went much smoother. Overall, I have to thank Sacha van Weeren for giving me this great opportunity. The flexibility Sacha shows is also something I greatly appreciate. Furthermore, I am indebted to Roger Lord for proofreading all my drafts, and for the numerous hints he gave during the course of the research. Of course, colleagues ranging from Freddy van Dijk who ordered me a comfortable chair to Scientific Workplace guru Erik van Raaij were crucial as well in the successful completion of the thesis. At the side of the University, I want to thank my thesis supervisor Dr. J.A.M. van der Weide for the helpful comments he gave and Dr. A. Almendral for proofreading and the enlightening discussions on complex analysis. My gratitude also goes to the rest of the graduation committee: Dr. F.M. Dekking and Dr. C. Oosterlee.

An immense challenge that paralleled this research was my Ph.D. applications in the US. Indeed, there were moments that they tended to overshadow everything. Fortunately, there is a happy ending. Without the following supportive individuals (to whom I am much beholden), things would have been quite different *almost surely*:

Dr. F.M. Dekking
Dr. J.A.M. van der Weide
Dr. C. Roos (Delft University of Technology)
Dr. D. Morton (University of Texas at Austin)
Dr. M. Crawford (University of Texas at Austin)

Utrecht, March 2005 ManWo Ng

Contents

1 Introduction

Consider the problem of valuing a European call under different assumptions of the underlying asset's model. Various techniques have been devised to provide an answer to this problem. For example, one can resort to Monte-Carlo techniques to simulate sample paths of the asset. Averaging a sufficiently large number of realized payoffs then yields the price required, see for example Glasserman (2003). One can also attempt to derive a pricing partial differential equation which can be readily solved using numerical methods. A champion of this approach is Wilmott (1998). Yet another method is based on Fourier analysis, which is the subject of this thesis.

Two Fourier techniques exist in literature. Both of them rely on the availability of the so-called characteristic function of the logarithm of the stock price. Indeed, for a wide class of stock models characteristic functions have been obtained in closed form even if the risk-neutral densities (or probability mass functions) themselves are not available explicitly. The beauty behind this approach is that the Fourier methods are rather modular in the sense that option prices under several stock models are obtained simply by substituting the respective characteristic function in some predefined integral. Equally important, the fast Fourier transform (FFT) can be utilized for the computational efforts required.

Due to the tremendous speed of the FFT, an important application emerges which is the calibration of stock models: We assume that a certain parametric stock model governs the dynamics of the underlying asset and we try to minimize the differences between European option prices from this stock model and those implied by the market, yielding a set of parameters for the stock model. Presumably, the calibrated stock model might then be used to price more exotic options.

The outline of this thesis is as follows. In Chapter 2 we briefly review the FFT, without which Fourier techniques would be far less attractive. Readers familiar with (discrete) Fourier analysis might directly jump to Chapter 3 in which the Fourier transform methods for option pricing are discussed extensively. The numerical implementation and its accompanying errors will be discussed as well. Chapter 4 presents the effectiveness of Fourier techniques for the Black-Scholes model. Since we have an analytic expression for the option price in this model, we can easily determine the accuracy of the Fourier techniques. Plenty of numerical results will be provided. Although it is incontestable that the Black-Scholes model forms the foundation of modern finance, certain underlying assumptions are questionable however. As an example, the volatility is supposed to be a known constant (or some known deterministic function of time), which is contradicting empirical evidence. The Heston stochastic volatility model corrects this flaw. No closed form expression is available for the option price in this model. Hence the determination of the accuracy of the Fourier techniques will be more cumbersome in this case, but much will be possible. The Heston stochastic volatility model and its numerical results are the topic of Chapter 5. Chapter 6 considers the application we hinted at above: calibration of stock models. Specifically, we used real market data to calibrate the Heston stochastic volatility model. Different objective functions and algorithms will be examined. Conclusions and recommendations can be found in Chapter 7.

2 The fast Fourier transform (FFT)

2.1 Introduction

This chapter provides a brief, but complete, discussion on Fourier transforms needed for the rest of the thesis. Specifically, much of our attention will be directed to the discrete Fourier transform (DFT) and its evaluation via the fast Fourier transform (FFT). Readers with a sufficiently strong background in this matter might directly jump to the next chapter.

The outline of this chapter is as follows. First we will present the (continuous) Fourier transform and some of its fundamental properties. Characteristic functions as encountered in probability theory are introduced hereafter. A fundamental relation with Fourier transforms will be established. Second, the discrete counterpart of Fourier transform - the discrete Fourier transform (DFT) - will be examined. Finally, the fast Fourier transform (FFT) is presented as an efficient way to evaluate DFTs.

2.2 Fourier Transforms

The provision of exact conditions for the existence of Fourier transforms seems to be a rather tedious job. Here we will be satisfied with the following sufficient conditions: If $f : \mathbb{R} \to \mathbb{R}$ is in L^1, i.e.,

$$\int_{-\infty}^{\infty} |f(t)|\, dt < \infty,$$

and if f is continuous, then the Fourier transform of f is defined by

$$\widehat{f}(\omega) = \int_{-\infty}^{\infty} e^{i\omega t} f(t) dt,$$

where $\omega \in \mathbb{R}$. The original function f can be recovered from its Fourier transform by inversion provided that $\widehat{f}$ satisfies the above mentioned conditions for f :

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega t} \widehat{f}(\omega) d\omega.$$

Certain properties of $f(t)$ will have consequences for the specifics of $\widehat{f}(\omega)$. Table 2.1 summarizes some of them. Here $\overline{()}$ denotes complex conjugation.

$f(t)$ is real $\Rightarrow \widehat{f}(-\omega) = \overline{\left[\widehat{f}(\omega)\right]}$
$f(t)$ is imaginary $\Rightarrow \widehat{f}(-\omega) = -\overline{\left[\widehat{f}(\omega)\right]}$
$f(t)$ is even $\Rightarrow \widehat{f}(\omega)$ is even
$f(t)$ is odd $\Rightarrow \widehat{f}(\omega)$ is odd
$f(t)$ is real and even $\Rightarrow \widehat{f}(\omega)$ is real and even
$f(t)$ is real and odd $\Rightarrow \widehat{f}(\omega)$ is imaginary and odd
$f(t)$ is imaginary and even $\Rightarrow \widehat{f}(\omega)$ is imaginary and even
$f(t)$ is imaginary and odd $\Rightarrow \widehat{f}(\omega)$ real and odd

Table 2.1: Some relations between $f(t)$ and its Fourier transform

One of the most valuable properties of Fourier transforms is that convolution in the time domain (or t-domain) reduces to multiplication in the frequency domain (or ω-domain): Let $g(t)$ and $h(t)$ be two functions whose Fourier transforms are given by $\widehat{g}(\omega)$ and $\widehat{h}(\omega)$, respectively. The convolution of $g(t)$ and $h(t)$, denoted as $(g * h)(t)$, is then given by

$$(g * h)(t) = \int_{-\infty}^{\infty} g(\tau)\, h(t-\tau)\, d\tau.$$

(Note that the order of convolution is immaterial, i.e. $g * h = h * g$.) By a change of variable one can easily show that the Fourier transform of $(g * h)(t)$ is given by $\widehat{g}(\omega)\,\widehat{h}(\omega)$.

Closely related to Fourier transforms are characteristic functions that one frequently encounters in probability theory. To give the definition, let X be a random variable defined on some probability space $(\Omega, \mathcal{F}, P)$ with cumulative distribution function F. The characteristic function (ch.f.) of X is then given by

$$\phi(t) = \mathbb{E}e^{itX} = \int_{-\infty}^{\infty} e^{itx} dF(x),$$

where $t \in \mathbb{R}$. When X has a density, then $\phi = \widehat{f}$ or in words: the ch.f. of a continuous random variable X equals the Fourier transform of the density of X.

This is all what we want to say about (continuous) Fourier transforms. More on this fascinating topic can be found in Goldberg (1961). We will proceed to its discrete counterpart, the discrete Fourier transform.

2.3 Discrete Fourier Transforms (DFT)

Consider the vectors $F, f \in \mathbb{C}^N$:

$$F = \begin{pmatrix} F_1 \\ F_2 \\ : \\ F_{N-1} \\ F_N \end{pmatrix}, f = \begin{pmatrix} f_1 \\ f_2 \\ : \\ f_{N-1} \\ f_N \end{pmatrix}.$$

Let $\omega_N = e^{-\frac{2\pi i}{N}}$, the DFT-matrix $W \in \mathbb{C}^{N \times N}$ is then defined as

$$W = \begin{pmatrix} 1 & 1 & 1 & .. & 1 \\ 1 & \omega_N & \omega_N^2 & .. & \omega_N^{N-1} \\ 1 & \omega_N^2 & \omega_N^4 & .. & \omega_N^{2(N-1)} \\ : & : & : & & : \\ 1 & \omega_N^{(N-1)} & \omega_N^{2(N-1)} & .. & \omega_N^{(N-1)(N-1)} \end{pmatrix},$$

that is,

$$W_{nk} = \omega_N^{(n-1)(k-1)}.$$

The discrete Fourier transform (DFT) F of f is then given by the matrix multiplication:

$$F = Wf, \tag{2.1}$$

or equivalently,

$$F_k = \sum_{n=1}^{N} f_n \omega_N^{(n-1)(k-1)}$$

for $k = 1, ..., N$. One can easily verify that the inverse of the DFT-matrix is given by

$$W^{-1} = \frac{1}{N} \begin{pmatrix} 1 & 1 & 1 & .. & 1 \\ 1 & \omega_N^{-1} & \omega_N^{-2} & .. & \omega_N^{-(N-1)} \\ 1 & \omega_N^{-2} & \omega_N^{-4} & .. & \omega_N^{-2(N-1)} \\ : & : & : & & : \\ 1 & \omega_N^{-(N-1)} & \omega_N^{-2(N-1)} & .. & \omega_N^{-(N-1)(N-1)} \end{pmatrix} = \frac{1}{N} W^*. \tag{2.2}$$

Hence we see that the columns of W form an orthogonal basis of $\mathbb{C}^N$. The operation of recovering f from F is called the inverse discrete Fourier transform (IDFT). Note that the DFT-matrix is a so-called Vandermonde matrix. Furthermore, it is Hermitian ($W_{nk} = W_{kn}^*$) and symmetric ($W = W^T$).

In literature, some variations on the above DFT-definition exist. As an example, the indices might be chosen to range from $-N/2 + 1$ to $N/2$. Differences might also be observed with regard to the scaling factor N in (2.2). Moreover, sometimes one refers to (2.1) as the *inverse* Fourier transform. Apart from these differences, all definitions satisfy certain properties that make the whole idea of discrete Fourier analysis useful. In fact, many - if not all - properties of the (continuous) Fourier transform carry over to the DFT. As we won't use any of them, we will not go into the specifics hereof. Briggs & Henson (1995) considered DFT-properties in detail.

The DFT arises naturally in many situations. Its emergence in this thesis is via the approximation of a Fourier integral. This will be illustrated extensively in the next chapter. Since it is just an approximation, errors will in general be introduced. This will also be scrutinized in the chapter to follow. Other situations in which the DFT emerges is the approximation of Fourier coefficients in the context of Fourier series. Moreover, trigonometric interpolation gives rise to the DFT as well. Detailed descriptions of these latter two applications can be found in Briggs & Henson (1995).

2.4 Fast Fourier Transform (FFT)

Recall our definition of the DFT:

$$F = Wf, \tag{2.3}$$

where

$$F = \begin{pmatrix} F_1 \\ F_2 \\ : \\ F_{N-1} \\ F_N \end{pmatrix}, f = \begin{pmatrix} f_1 \\ f_2 \\ : \\ f_{N-1} \\ f_N \end{pmatrix}$$

and

$$W = \begin{pmatrix} 1 & 1 & 1 & .. & 1 \\ 1 & \omega_N & \omega_N^2 & .. & \omega_N^{N-1} \\ 1 & \omega_N^2 & \omega_N^4 & .. & \omega_N^{2(N-1)} \\ : & : & : & & : \\ 1 & \omega_N^{(N-1)} & \omega_N^{2(N-1)} & .. & \omega_N^{(N-1)(N-1)} \end{pmatrix}.$$

This is a matrix multiplication, which requires about N^2 (complex) multiplications and N^2 (complex) additions. Hence the number of arithmetic operations is of order N^2, i.e. $O\left(N^2\right)$. This way of thinking was prevalent for a very long time. To be precise, it was 1965 when Cooley & Tukey (1965) showed that it was in fact possible to have the DFT evaluated with $O\left(N \log_2 N\right)$ arithmetic operations. Figure 2.1 illustrates the huge differences between $O\left(N^2\right)$ and $O\left(N \log_2 N\right)$. This $O\left(N \log_2 N\right)$ algorithm is called the fast Fourier transform (FFT). (As a matter of fact, efficient methods for evaluating the DFT have already been devised as long ago as in 1805 by Gauss. However, the world was dormant until 1965.)

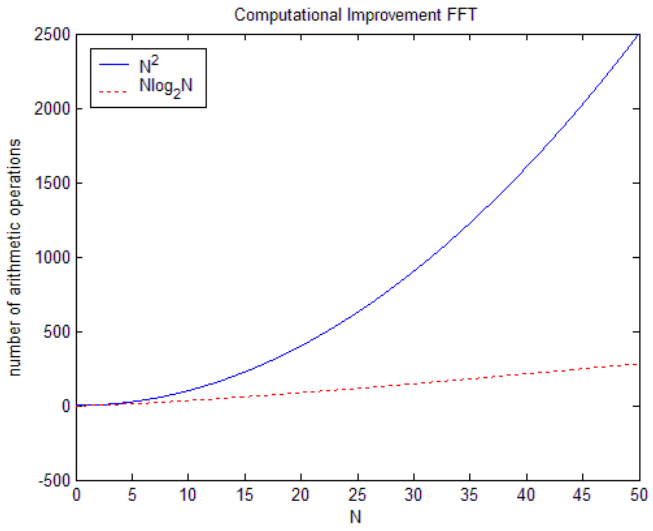


Figure 2.1: Huge differences between $O\left(N^2\right)$ and $O\left(N \log_2 N\right)$.

There is no single FFT algorithm, a whole plethora of tailor-made FFT algorithms exists. The main characteristic feature that is shared by all of them is that we have an $O\left(N \log N\right)$ algorithm. An extensive discussion of a large variety of FFT-algorithms can be found in Van Loan (1992). Next we present two main FFT-algorithms. The first is known as a decimation-in-time algorithm, which originates from Cooley and Tukey. The second algorithm comes by the name decimation-in-frequency, an invention by Gentleman and Sande. As said, many more algorithms exist. The bottom-line is that in case one has (2.3) or some other form of the DFT, one can rapidly evaluate it.

2.4.1 Decimation-in-time Algorithms

In our view, this algorithm is best illustrated when we switch to the sum representation of the DFT:

$$F_k = \sum_{n=1}^{N} f_n \omega_N^{(n-1)(k-1)}, \tag{2.4}$$

for $k = 1, ..., N$. First, let us assume that N is some power of 2:

$$N = 2^L,$$

where $L \in \mathbb{N}$. (Below we shall demonstrate that this is always possible by a technique called zero-padding.) Now, define $x_n = f_{2n-1}$ and $y_n = f_{2n}$ for $n = 1, ..., N/2$, i.e. $\{x_n\}$ and $\{y_n\}$ are the odd and even subsequences of $\{f_n\}$, respectively. Then we can write (2.4) as

$$F_k = \sum_{n=1}^{N/2} \left(x_n \omega_N^{(2n-2)(k-1)} + y_n \omega_N^{(2n-1)(k-1)} \right), \tag{2.5}$$

for $k = 1, ..., N$. The very trick behind this FFT algorithm lies in the simple but crucial equality

$$\omega_N^{(2n-2)(k-1)} = \omega_{N/2}^{(n-1)(k-1)}.$$

With this observation, (2.5) reduces to

$$F_k = \sum_{n=1}^{N/2} x_n \omega_{N/2}^{(n-1)(k-1)} + \omega_N^{k-1} \sum_{n=1}^{N/2} y_n \omega_{N/2}^{(n-1)(k-1)} = X_k + \omega_N^{k-1} Y_k$$

which is - apart from the factor ω_N^{k-1}- the sum of two DFTs, each of length $N/2$. Hence we can write

$$\begin{aligned} F_k &= X_k + \omega_N^{k-1} Y_k \\ F_{k+N/2} &= X_{k+N/2} + \omega_N^{k-1+N/2} Y_{k+N/2,} \end{aligned}$$

where $k = 1, ..., N/2$. By noting that $X_{k+N/2} = X_k$, $Y_{k+N/2} = Y_k$ and $\omega_N^{N/2} = -1$, the above reduces to

$$\begin{aligned} F_k &= X_k + \omega_N^{k-1} Y_k \\ F_{k+N/2} &= X_k - \omega_N^{k-1} Y_{k,}, \end{aligned} \tag{2.6}$$

which are so-called butterfly relations. They simply tell us how the DFT of a sequence can be recovered from the DFTs of its odd and even subsequences. Of course, we need not to stop here. The same procedure can be repeated using the sequences $\{x_n\}$ and $\{y_n\}$, until we have sequences of length 1. The DFT of such a sequence is 'easily' computed: it is the sequence itself. The butterfly relations then give a recipe to obtain the DFT of the original sequence. Hence there are two main phases in this algorithm:

- a reordering phase, in which the sequence is split into odd and even subsequences. This phase comes to an end when we have N sequences of length 1.
- a combine phase, in which the butterfly relations are utilized to recover the DFT of the original sequence. This phase comes to an end when we obtain one sequence of length N.

To obtain the number of arithmetic operations needed, note that the combine-stage consists of $L = \log_2 N$ steps, each of which gives rise to the evaluation of N butterfly relations, or equivalently, $N/2$ pairs. From (2.6) we see that each pair requires one multiplication and two additions, giving a total of $\frac{N}{2} \log_2 N$ multiplications and $N \log_2 N$ additions. Thus this algorithm is $O(N \log_2 N)$.

In the just presented algorithm, we assumed that N was a power of 2. In practice, the sequence fed into the DFT might be of odd length. One might think that in this case the above algorithm is not applicable. This thought is only partially correct. Next we illustrate a

technique called zero-padding which makes sequences of odd length suitable (by adding zeros to the sequence) for the decimation-in-time algorithm just illustrated. To this end, let $\{f_m\}$ denote a sequence defined for $m = 1, ..., M$. Suppose we need a sequence of even length, say length N, where $N > M$. If we allow the new sequence to be an integer multiple of the original sequence, i.e. $N/M > 1$ and integer, a feasible strategy is then to pad $\{f_m\}$ with zeros until we obtain the desired length N sequence: Let $\{g_n\}$ denote the padded sequence, then

$$g_n = \begin{cases} f_n & for\ n = 1, ..., M \\ 0 & for\ n = M+1, ..., N \end{cases} .$$

The DFT of this padded sequence is given by G_k:

$$G_k = \sum_{n=1}^{N} g_n \omega_N^{(n-1)(k-1)} = \sum_{m=1}^{M} f_m \omega_N^{(m-1)(k-1)} \tag{2.7}$$

for $k = 1, ..., N$. However, since

$$\omega_N = \omega_M^{M/N},$$

equation (2.7) reduces to

$$G_k = \sum_{m=1}^{M} f_m \omega_M^{(m-1)(k-1)M/N} = \sum_{m=1}^{M} f_m \omega_M^{(m-1)([(k-1)M/N+1]-1)} = F_{(k-1)M/N+1},$$

or equivalently,

$$F_k = G_{(k-1)N/M+1}$$

for $k = 1, ..., M$. As nothing comes for free, zero-padding has a cost as well. This cost lies in the determination of DFTs of sequences of at least double length compared to the original sequence.

The reader might now surmise that sequences of length 2^L, where $L \in \mathbb{N}$, are crucial to the successes of the FFT. This is definitely false, as we will illustrate next. In fact, when $N = 2^L$ (the so-called radix-2 case) we have the above Cooley-Tukey algorithm. However, nothing prevents us from using the splitting idea to sequences of length $N = r^L$, where $r \in \mathbb{N}$. This is the general radix-r case. Each case just gives rise to other butterfly relations. For example, let $r = 3$. In order to determine

$$F_k = \sum_{n=1}^{N} f_n \omega_N^{(n-1)(k-1)} \tag{2.8}$$

for $k = 1, ..., N$, we split $\{f_n\}$ into three subsequences $x_n = f_{3n-2}$, $y_n = f_{3n-1}$ and $z_n = f_{3n}$ where $n = 1, ..., N/3$. Thus we can write (2.8) as

$$F_k = \sum_{n=1}^{N/3} \left(x_n \omega_N^{(3n-3)(k-1)} + y_n \omega_N^{(3n-2)(k-1)} + z_n \omega_N^{(3n-1)(k-1)} \right)$$

for $k = 1, ..., N$. Using the same trick as before, we have

$$F_k = \sum_{n=1}^{N/3} x_n \omega_{N/3}^{(n-1)(k-1)} + \omega_N^{(k-1)} \sum_{n=1}^{N/3} y_n \omega_{N/3}^{(n-1)(k-1)} + \omega_N^{2(k-1)} \sum_{n=1}^{N/3} z_n \omega_{N/3}^{(n-1)(k-1)},$$

which is the sum of three DFTs, each of length $N/3$. Hence we have

$$\begin{aligned} F_k &= X_k + \omega_N^{(k-1)} Y_k + \omega_N^{2(k-1)} Z_k \\ F_{k+N/3} &= X_{k+N/3} + \omega_N^{(k+N/3-1)} Y_{k+N/3} + \omega_N^{2(k+N/3-1)} Z_{k+N/3} \\ F_{k+2/3N} &= X_{k+2/3N} + \omega_N^{(k+2/3N-1)} Y_{k+2/3N} + \omega_N^{2(k+2/3N-1)} Z_{k+2/3N} \end{aligned}$$

for $k = 1, ..., N/3$, which after further simplifications yields the following butterfly relations

$$\begin{aligned} F_k &= X_k + \omega_N^{(k-1)} Y_k + \omega_N^{2(k-1)} Z_k \\ F_{k+N/3} &= X_k + \omega_N^{N/3} \omega_N^{(k-1)} Y_k + \omega_N^{2/3N} \omega_N^{2(k-1)} Z_k \\ F_{k+2/3N} &= X_k + \omega_N^{2/3N} \omega_N^{(k-1)} Y_k + \omega_N^{N/3} \omega_N^{2(k-1)} Z_k \end{aligned} \tag{2.9}$$

for $k = 1, ..., N/3$. Let's count the number of arithmetic operations in this case. The combine stage now consists of $L = \log_3 N$ steps, each of which gives rise to the evaluation of N butterfly relations. From (2.9) we see that each triple butterfly relations requires six multiplications and six additions, giving a total of $2N \log_3 N$ multiplications and $2N \log_3 N$ additions. Hence we have an $O(N \log_3 N)$ algorithm. One can easily see that the radix-2 case is the fastest among all decimation-in-time algorithms.

2.4.2 Decimation-in-frequency Algorithms

We have seen that decimation-in-time algorithms were characterized by an initial reordering stage, which was followed by a combine stage. This section presents algorithms for which these two phases are reversed: combining is performed first, reordering follows hereafter.

For the time being, let

$$N = 2^L,$$

where $L \in \mathbb{N}$. In order to compute

$$F_k = \sum_{n=1}^{N} f_n \omega_N^{(n-1)(k-1)}$$

for $k = 1, ..., N$, we split the sequence $\{f_n\}$ into its first and second half:

$$\begin{aligned} F_k &= \sum_{n=1}^{N/2} \left(f_n \omega_N^{(n-1)(k-1)} + f_{n+N/2} \omega_N^{(n+N/2-1)(k-1)} \right) \\ &= \sum_{n=1}^{N/2} \left(f_n + (-1)^{k-1} f_{n+N/2} \right) \omega_N^{(n-1)(k-1)}, \end{aligned}$$

where we used the fact that

$$\omega_N^{(k-1)N/2} = (-1)^{k-1}.$$

Hence the odd coefficients of the DFT become

$$\begin{aligned} F_{2k-1} &= \sum_{n=1}^{N/2} \left(f_n + f_{n+N/2} \right) \omega_N^{(n-1)(2k-2)} \\ &= \sum_{n=1}^{N/2} \left(f_n + f_{n+N/2} \right) \omega_{N/2}^{(n-1)(k-1)}, \end{aligned}$$

where we again used the crucial property $\omega_N^2 = \omega_{N/2}$. For the even coefficients of F_k we have

$$\begin{aligned} F_{2k} &= \sum_{n=1}^{N/2} \left(f_n - f_{n+N/2}\right) \omega_N^{(n-1)(2k-1)} \\ &= \sum_{n=1}^{N/2} \left(f_n - f_{n+N/2}\right) \omega_N^{(n-1)2(k-1)} \omega_N^{n-1} \\ &= \sum_{n=1}^{N/2} \left[\left(f_n - f_{n+N/2}\right) \omega_N^{n-1}\right] \omega_{N/2}^{(n-1)(k-1)}. \end{aligned}$$

Now, define the butterfly relations:

$$\begin{aligned} x_n &= f_n + f_{n+N/2} \\ y_n &= \left(f_n - f_{n+N/2}\right) \omega_N^{(n-1)}, \end{aligned}$$

where $n = 1, ..., N/2$, then we have

$$\begin{aligned} F_{2k-1} &= X_k \\ F_{2k} &= Y_k, \end{aligned}$$

for $k = 1, ..., N/2$. Thus we once again have reduced the problem of determining the DFT of a sequence of length N to the evaluation of two DFTs, each of length $N/2$. This procedure is repeated until we have sequences of length 1, the DFTs of which are just the sequences themselves. In this algorithm by Gentleman & Sande (1966), the resulting DFT coefficients will be scrambled. A reordering stage brings them back to their natural order. Because of this last reordering stage in the Fourier or frequency domain, this type of algorithms is categorized as decimation-in-frequency. A simple count of the number of arithmetic operations involved brings us to an $O(N \log_2 N)$ algorithm. Like with decimation-in-time algorithms, it is not necessary for N to be a power of 2. Modifications for the radix-r case are readily made. Zero-padding can be applied as well.

3 Option Valuation via Fourier Transforms

3.1 Introduction

Consider the problem of valuing a European call under different assumptions of the underlying asset's model. Various techniques have been devised to provide an answer to this problem. For example, one can resort to Monte-Carlo techniques to simulate sample paths of the asset. Averaging a sufficiently large number of realized payoffs then yields the price required, see for example Glasserman (2003). One can also attempt to derive a pricing partial differential equation which can be readily solved using numerical methods. A champion of this approach is Wilmott (1998). Yet another method is based on Fourier analysis, which is the subject of the current chapter.

Two methods based on Fourier analysis exist in literature. Both of them rely on the availability of the characteristic function of the logarithm of the stock price. Indeed, for a wide class of stock models characteristic functions have been obtained in closed form even if the risk-neutral densities (or probability mass functions) themselves are not available explicitly. Characteristic functions have been derived in for example Madan, Carr & Chang (1998), Heston (1993) and Zhu (2000).

The first of these Fourier methods is actually the application of the Gil-Palaez inversion formula to finance. This idea originates from Heston (1993). However, singularities in the integrand prevent it to be an accurate method. The second, more recent technique, was first proposed by Carr & Madan (1999). It ensures that the Fourier transform of the call price exist by the inclusion of a damping factor. Moreover, Fourier inversion can be accomplished by the fast Fourier transform (FFT) in this case. The tremendous speed of the FFT allows option prices for a huge number of strikes to be evaluated very rapidly.

The outline of this chapter is as follows. First we discuss the Gil-Palaez inversion formula and its application to option pricing. Then Carr-Madan's method will be examined in detail. In order for these methods to be applicable in practice, discretization is required. After presenting the discretization scheme, we discuss the errors involved. Specifically, we will present methods for quantifying them. Our error analysis is based on the adaptive Simpson's rule which will also be reviewed.

3.2 Gil-Palaez Inversion

The first Fourier pricing method we discuss is heavily based on the Gil-Palaez inversion formula, see Gil-Palaez (1951). Heston (1993) was the first to utilize Gil-Palaez inversion in option pricing theory. Numerous other authors have discussed this method as well. Bakshi & Madan (2000) presents a slightly different perspective compared to Heston. It is this latter approach that we will outline here. We begin with the Gil-Palaez inversion formula which is stated in Proposition 1.

Proposition 1 *Gil-Palaez Inversion Formula. Let $F(x)$ be the cumulative distribution function of some random variable X. Furthermore, let*

$$\phi(t) = \int_{-\infty}^{\infty} e^{itx} dF(x)$$

be the associated characteristic function. Then we have

$$\frac{1}{2}\left[F(x) + F(x-)\right] = \frac{1}{2} + \lim_{\delta \downarrow 0, T \uparrow \infty} \int_{\delta}^{T} \frac{e^{itx}\phi(-t) - e^{-itx}\phi(t)}{2\pi i t} dt.$$

Proof. First note that by definition

$$\begin{aligned} e^{itx}\phi(-t) - e^{-itx}\phi(t) &= e^{itx}\int_{-\infty}^{\infty} e^{-ity}dF(y) - e^{-itx}\int_{-\infty}^{\infty} e^{ity}dF(y) \\ &= \int_{-\infty}^{\infty}\left[e^{it(x-y)} - e^{it(y-x)}\right]dF(y). \end{aligned} \tag{3.1}$$

Now, divide the integrand in (3.1) by $2\pi it$, we then have

$$\frac{e^{it(x-y)} - e^{it(y-x)}}{2\pi it} = \frac{\sin(t(x-y))}{\pi t},$$

which easily follows from Euler's formula. Moreover, from calculus we know that

$$\int_0^{\infty}\frac{\sin(\alpha x)}{x}dx = \frac{\pi}{2}\mathrm{sgn}(\alpha),$$

hence we have

$$\int_0^{\infty}\frac{\sin(t(x-y))}{\pi t}dt = \frac{1}{\pi}\int_0^{\infty}\frac{\sin(t(x-y))}{t}dt = \frac{1}{2}\mathrm{sgn}(x-y).$$

For the moment, assume that Fubini's theorem and the dominated convergence theorem hold, then we get

$$\begin{aligned} \lim_{\delta\downarrow 0, T\uparrow\infty}\int_{\delta}^{T}\frac{e^{itx}\phi(-t) - e^{-itx}\phi(t)}{2\pi it}dt &= \lim_{\delta\downarrow 0, T\uparrow\infty}\int_{\delta}^{T}\left[\int_{-\infty}^{\infty}\frac{\sin(t(x-y))}{\pi t}dF(y)\right]dt \\ &= \int_{-\infty}^{\infty}\left[\int_0^{\infty}\frac{\sin(t(x-y))}{\pi t}dt\right]dF(y) \\ &= \int_{-\infty}^{\infty}\left[\frac{1}{2}\mathrm{sgn}(x-y)\right]dF(y) \end{aligned}$$

$$\begin{aligned} &= \frac{1}{2}\left(\int_{(-\infty,x)}dF(y) - \int_{(x,\infty)}dF(y)\right) \\ &= \frac{1}{2}\left[F(x-) - (1-F(x))\right] \\ &= \frac{1}{2}\left[F(x) + F(x-)\right] - \frac{1}{2}. \end{aligned}$$

Indeed, we can interchange the order of integration since

$$\begin{aligned}
\left| \int_{\delta}^{T} \left[\int_{-\infty}^{\infty} \frac{\sin\left(t\left(x-y\right)\right)}{\pi t} dF\left(y\right) \right] dt \right| &= \left| \int_{\delta}^{T} \frac{e^{itx}\phi\left(-t\right) - e^{-itx}\phi\left(t\right)}{2\pi it} dt \right| \\
&\leq \int_{\delta}^{T} \left| \frac{e^{itx}\phi\left(-t\right) - e^{-itx}\phi\left(t\right)}{2\pi it} \right| dt \\
&\leq \int_{\delta}^{T} \frac{1}{2\pi t} \left[\left|\phi\left(-t\right)\right| + \left|\phi\left(t\right)\right| \right] dt \\
&= \int_{\delta}^{T} \frac{1}{2\pi t} \left[\left|\overline{\phi\left(t\right)}\right| + \left|\phi\left(t\right)\right| \right] dt \\
&= \int_{\delta}^{T} \frac{1}{\pi t} \left|\phi\left(t\right)\right| dt \\
&\leq \int_{\delta}^{T} \frac{1}{\pi t} dt < \infty,
\end{aligned}$$

here we used $\phi(-t) = \overline{\phi(t)}$. Moreover, interchanging limit and integral is allowed by dominated convergence since we have

$$\left| \int_{\delta}^{T} \frac{\sin\left(t\left(x-y\right)\right)}{\pi t} dt \right| \leq \int_{\delta}^{T} \frac{1}{\pi t} dt = \frac{1}{\pi} \left[\log\left[T\right] - \log\left[\delta\right] \right],$$

and

$$\frac{1}{\pi} \left[\log\left[T\right] - \log\left[\delta\right] \right] \int_{-\infty}^{\infty} dF\left(y\right) = \frac{1}{\pi} \left[\log\left[T\right] - \log\left[\delta\right] \right].$$

■

In order to express the option price of a European call in terms of inverse Fourier transforms, we need one additional result. This is given in the next lemma.

Lemma 2 *We have the equality:*

$$\frac{e^{itx}\phi(-t) - e^{-itx}\phi(t)}{2\pi it} = -\frac{1}{\pi} \operatorname{Re}\left(\phi(t) \frac{e^{-itx}}{it} \right).$$

Proof. Let $g(t) = \operatorname{Re}\left\{\phi(t)\right\}$ and $h(t) = \operatorname{Im}\left\{\phi(t)\right\}$. From the equality $\phi(-t) = \overline{\phi(t)}$ we then have

$$\phi(-t) = g(t) - ih(t).$$

Therefore

$$\begin{aligned}\frac{e^{itx}\phi(-t)-e^{-itx}\phi(t)}{2\pi it} &= \frac{e^{itx}\left(g(t)-ih(t)\right)-e^{-itx}\left(g(t)+ih(t)\right)}{2\pi it}\\ &= \frac{g(t)\left(e^{itx}-e^{-itx}\right)-ih(t)\left(e^{itx}+e^{-itx}\right)}{2\pi it}\\ &= \frac{2ig(t)\sin(tx)-i2h(t)\cos(tx)}{2\pi it}\\ &= \frac{g(t)\sin(tx)-h(t)\cos(tx)}{\pi t}\end{aligned}$$

On the other hand,

$$\begin{aligned}\phi(t)\frac{e^{-itx}}{it} &= \left(g(t)+ih(t)\right)\frac{e^{-itx}}{it}\\ &= \left(h(t)-ig(t)\right)\frac{e^{-itx}}{t}\\ &= \frac{1}{t}\left(h(t)-ig(t)\right)\left(\cos(tx)-i\sin(tx)\right)\\ &= \frac{1}{t}\left(h(t)\cos(tx)-g(t)\sin(tx)-i\left(g(t)\cos(tx)+h(t)\sin(tx)\right)\right),\end{aligned}$$

Thus

$$\operatorname{Re}\left(\phi(t)\frac{e^{-itx}}{it}\right)=\frac{\left[h(t)\cos(tx)-g(t)\sin(tx)\right]}{t},$$

or equivalently,

$$-\frac{1}{\pi}\operatorname{Re}\left(\phi(t)\frac{e^{-itx}}{it}\right)=-\frac{\left[h(t)\cos(tx)-g(t)\sin(tx)\right]}{\pi t},$$

completing the proof. ■

The Gil-Palaez inversion formula can now be written as

$$\frac{1}{2}\left[F\left(x\right)+F\left(x-\right)\right]=\frac{1}{2}-\frac{1}{\pi}\int_0^\infty \operatorname{Re}\left(\phi(t)\frac{e^{-itx}}{it}\right)dt.$$

From now on, we will assume that $F\left(x\right)$ is continuous for the sake of notational convenience. Therefore, Gil-Palaez further simplifies to

$$F\left(x\right)=\frac{1}{2}-\frac{1}{\pi}\int_0^\infty \operatorname{Re}\left(\phi(t)\frac{e^{-itx}}{it}\right)dt. \tag{3.2}$$

Now, let $C_T\left(K\right)$ denote the price of a European call with strike K and time to expiry T. Before stating how $C_T\left(K\right)$ can be expressed in terms of characteristic functions, recall that the arbitrage-free price of a vanilla call is given by the discounted expected payoff, provided that the expectation is taken with respect to a risk-neutral measure. Moreover, under the risk-neutral measure the discounted stock price becomes a martingale. More on this can be found in Björk (1998).

Proposition 3 *Assume we live in an arbitrage-free world. Then*

$$C_T\left(K\right)=S_0\Pi_1-Ke^{-rT}\Pi_2,$$

where

$$\Pi_1 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\frac{\phi_T(u-i)}{\phi_T(-i)}\frac{e^{-iu\log K}}{iu}\right)du \tag{3.3}$$

$$\Pi_2 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\phi_T(u)\frac{e^{-iu\log K}}{iu}\right)du, \tag{3.4}$$

and in which

$$\phi_T(u) = \mathbb{E}e^{iu\log S_T}, \tag{3.5}$$

where S_T denotes the stock price at time T. Here all expectations are taken with respect to a risk-neutral measure $\mathbb{Q}$.

Proof. By risk neutral valuation we have

$$\begin{aligned} C_T(K) &= e^{-rT}\mathbb{E}\left[\max\left\{S_T - K, 0\right\}\right] \\ &= e^{-rT}\mathbb{E}\left[S_T - K; S_T > K\right] \\ &= e^{-rT}\mathbb{E}\left[S_T; S_T > K\right] - Ke^{-rT}\mathbb{Q}\left\{S_T > K\right\}, \end{aligned}$$

where r denotes the risk-free rate. Define

$$S = \log S_T,$$

then $\phi_T(u)$ as defined in (3.5) is the characteristic function of S. By (3.2) we have that

$$\mathbb{Q}\left\{S_T > K\right\} = \mathbb{Q}\left\{S > \log K\right\} = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\phi_T(u)\frac{e^{-iu\log K}}{iu}\right)du.$$

Now, define a new probability measure $\mathbb{P}$:

$$d\mathbb{P} \equiv \frac{S_T}{\mathbb{E}S_T}d\mathbb{Q},$$

then

$$\mathbb{E}\left[S_T; S_T > K\right] = \int_{\{S_T>K\}} S_T d\mathbb{Q} = \mathbb{E}S_T \int_{\{S_T>K\}} \frac{S_T}{\mathbb{E}S_T}d\mathbb{Q} = \mathbb{E}S_T\mathbb{P}\left\{S_T > K\right\}.$$

The characteristic function of S under $\mathbb{P}$ is

$$\int_\Omega e^{iuS}d\mathbb{P} = \int_\Omega e^{iuS}\frac{S_T}{\mathbb{E}S_T}d\mathbb{Q} = \frac{\mathbb{E}\left[e^{iuS+S}\right]}{\mathbb{E}S_T} = \frac{\mathbb{E}\left[e^{i(u-i)S}\right]}{\mathbb{E}S_T} = \frac{\phi_T(u-i)}{\phi_T(-i)},$$

where the last equality follows by noting that $\mathbb{E}S_T = \phi_T(-i)$. By (3.2) we get

$$\mathbb{P}\left\{S_T > K\right\} = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\frac{\phi_T(u-i)}{\phi_T(-i)}\frac{e^{-iu\log K}}{iu}\right)du,$$

therefore

$$e^{-rT}\mathbb{E}\left[S_T; S_T > K\right] = e^{-rT}\mathbb{E}S_T\left[\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\frac{\phi_T(u-i)}{\phi_T(-i)}\frac{e^{-iu\log K}}{iu}\right)du\right].$$

Using the fact that under a risk-neutral measure the discounted stock price becomes a martingale $\big($hence $e^{-rT}\mathbb{E}S_T = S_0\big)$ yields the desired result. ■

Note the singularities at $u = 0$ in the integrands of (3.3) and (3.4). Thus care has to be exercised in the numerical evaluation of these integrals. Numerical results will be presented in Section 4.2.

3.3 Carr-Madan Inversion

The Fourier inversion technique illustrated in this section was first proposed by Carr & Madan (1999). It ensures that the Fourier transform of European option prices exist by the inclusion of an exponential damping factor. Moreover, singularities will be removed by this damping factor.

Let S_T denote the price at maturity of the underlying asset of a European call with strike K. Furthermore, define $S \equiv \log(S_T)$ whose associated risk neutral density is given by $q_T(s)$. Then the Fourier transform of $q_T(s)$, or equivalently the characteristic function of S, can be written as

$$\phi_T(u) = \int_{-\infty}^{\infty} e^{ius} q_T(s) ds.$$

Now let $k \equiv \log(K)$, risk-neutral valuation then yields

$$\begin{aligned} C_T(K) &= e^{-rT}\mathbb{E}\left[\max\left\{S_T - K, 0\right\}\right] \\ &= e^{-rT}\mathbb{E}\left[\max\left\{e^S - e^k, 0\right\}\right] \\ &= e^{-rT}\int_{-\infty}^{\infty} \max\left\{e^s - e^k, 0\right\} q_T(s) ds \\ &= e^{-rT}\int_{k}^{\infty} \left(e^s - e^k\right) q_T(s) ds, \end{aligned}$$

in which the expectation is taken with respect to some risk-neutral measure. Since

$$\lim_{K\to 0} C_T(K) = \lim_{k\to -\infty} C_T(e^k) = S_0,$$

we see that $C_T(e^k)$ is not in L^1, as $C_T(e^k)$ does not tend to zero for $k \to -\infty$. Now, consider the modified call price

$$c_T(k) \equiv e^{\alpha k} C_T(e^k)$$

where $\alpha > 0$. Below we will show that under a certain assumption we have $c_T(k) \in L^1$, the space of integrable functions. For now, assume that the Fourier transform of $c_T(k)$ is well-defined:

$$\psi_T(v) \equiv \widehat{c}_T(v) = \int_{-\infty}^{\infty} e^{ivk} c_T(k) dk. \tag{3.6}$$

Inverting gives

$$c_T(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ivk} \psi_T(v) dv,$$

or

$$C_T(K) = \frac{e^{-\alpha \log K}}{2\pi} \int_{-\infty}^{\infty} e^{-iv \log K} \psi_T(v) dv = \frac{e^{-\alpha \log K}}{\pi} \operatorname{Re} \left\{ \int_0^{\infty} e^{-iv \log K} \psi_T(v) dv \right\}, \tag{3.7}$$

where the last equality follows from the observation that

$$\int_{-\infty}^{\infty} e^{-iv \log K} \psi_T(v) dv = \int_0^{\infty} e^{-iv \log K} \psi_T(v) dv + \int_{-\infty}^{0} e^{-iv \log K} \psi_T(v) dv,$$

and where the second term on the right-hand side can be rewritten as

$$\begin{aligned} \int_{-\infty}^{0} e^{-iv \log K} \psi_T(v) dv &= \int_0^{\infty} e^{iu \log K} \psi_T(-u) du \\ &= \int_0^{\infty} \overline{e^{-iu \log K} \psi_T(u)} du \\ &= \overline{\int_0^{\infty} e^{-iv \log K} \psi_T(v) dv}, \end{aligned}$$

yielding the claim. Note that we have a nice closed-form for the Fourier transform of $c_T(k)$:

$$\begin{aligned} \psi_T(v) &= \int_{-\infty}^{\infty} e^{ivk} c_T(k) dk \\ &= \int_{-\infty}^{\infty} e^{ivk} e^{\alpha k} C_T(e^k) dk \\ &= \int_{-\infty}^{\infty} e^{ivk} e^{\alpha k} \left(e^{-rT} \int_k^{\infty} \left(e^s - e^k\right) q_T(s) ds \right) dk \\ &= e^{-rT} \int_{-\infty}^{\infty} q_T(s) \left(\int_{-\infty}^{s} e^{(iv+\alpha)k} \left(e^s - e^k\right) dk \right) ds \\ &= e^{-rT} \int_{-\infty}^{\infty} q_T(s) \left(e^s \int_{-\infty}^{s} e^{(iv+\alpha)k} dk - \int_{-\infty}^{s} e^{(iv+\alpha+1)k} dk \right) ds \\ &= e^{-rT} \int_{-\infty}^{\infty} q_T(s) \left(e^s \left[\frac{1}{iv+\alpha} e^{(iv+\alpha)k} \right]_{-\infty}^{s} - \left[\frac{1}{iv+\alpha+1} e^{(iv+\alpha+1)k} \right]_{-\infty}^{s} \right) ds. \end{aligned} \tag{3.8}$$

Since for $\alpha > 0$

$$\lim_{k\to-\infty}\left|e^{(iv+\alpha)k}\right| = \lim_{k\to-\infty}\left|e^{(iv+\alpha+1)k}\right| = \lim_{k\to-\infty} e^{(\alpha+1)k} = 0,$$

(3.8) reduces to

$$\psi_T(v) = e^{-rT}\int_{-\infty}^{\infty} q_T(s)\left[\frac{e^{(iv+\alpha+1)s}}{iv+\alpha} - \frac{e^{(iv+\alpha+1)s}}{iv+\alpha+1}\right]ds \tag{3.9}$$

$$= \frac{e^{-rT}\phi_T\left(v-(\alpha+1)i\right)}{\alpha^2+\alpha-v^2+i\left(2\alpha+1\right)v}, \tag{3.10}$$

where the last equality follows from straightforward calculus.

Above we have indicated that $c_T(k) \in L^1$. In fact, one simple assumption is enough to ensure this.

Lemma 4 *Let $\alpha > 0$. The Fourier transform of $c_T(k)$ exists (i.e. $c_T(k) \in L^1$) if $\mathbb{E}S_T^{\alpha+1} < \infty$.*

Proof. First note that $\mathbb{E}S_T^{\alpha+1} < \infty$ implies

$$\psi_T(0) < \infty, \tag{3.11}$$

since

$$|\psi_T(0)| = \frac{e^{-rT}\left|\phi_T\left(-(\alpha+1)i\right)\right|}{\alpha^2+\alpha} = \frac{e^{-rT}\mathbb{E}S_T^{\alpha+1}}{\alpha^2+\alpha},$$

where the last equality follows from

$$\left|\phi_T\left(-(\alpha+1)i\right)\right| = \left|\mathbb{E}e^{[-(\alpha+1)i]i\log S_T}\right| = \left|\mathbb{E}e^{(\alpha+1)\log S_T}\right| = \mathbb{E}S_T^{\alpha+1}.$$

However, we also have the equality

$$\psi_T(0) = \int_{-\infty}^{\infty} c_T(k)dk,$$

which follows easily from (3.6). Combining this with (3.11) completes the proof. ■

As a final remark, note that we have only considered the pricing of vanilla calls. Of course, one can obtain prices of vanilla puts by using the put-call parity. However, one can also easily obtain the price $P_T(K)$ of a vanilla put by Carr-Madan inversion by choosing a negative value for α, see Lee (2004).

3.4 Discretization

A characteristic that is shared by the two Fourier methods is the inversion of some Fourier transform. In general, this will not be tractable analytically. A numerical approach is necessary. In doing so we will try to give a formulation that is susceptible to the fast Fourier transform (FFT). Here we will adopt

$$F_k = \sum_{n=1}^{N} f_n \omega_N^{(n-1)(k-1)},\ k = 1, ..., N \tag{3.12}$$

where $\omega_N = e^{-\frac{2\pi i}{N}}$ to be the definition of the discrete Fourier transform (DFT). MATLAB provides an efficient FFT-algorithm for this definition.

Recall that Gil-Palaez inversion gives rise to two probabilities that are expressed in terms of (inverse) Fourier transforms:

$$\Pi_1 = \frac{1}{2} + \frac{1}{\pi} \operatorname{Re} \left\{ \int_0^\infty \left(\frac{\phi_T(v-i)}{\phi_T(-i)} \frac{e^{-ivk}}{iv} \right) dv \right\} \tag{3.13}$$

$$\Pi_2 = \frac{1}{2} + \frac{1}{\pi} \operatorname{Re} \left\{ \int_0^\infty \left(\phi_T(v) \frac{e^{-ivk}}{iv} \right) dv \right\}. \tag{3.14}$$

On the other hand, Carr-Madan inversion yields

$$C_T(k) = \frac{e^{-\alpha k}}{\pi} \operatorname{Re} \left\{ \int_0^\infty e^{-ivk} \psi_T(v) dv \right\}. \tag{3.15}$$

(Note that strictly speaking, we have to write $C_T(K)$ instead of $C_T(k)$. However, we will stick to this for notational convenience, which has also been done in Carr & Madan (1999).) In order to be able to deal with these two methods at once, we will only focus on the integrals of these expressions at first, i.e. we are interested in computing the indefinite integral

$$\int_0^\infty e^{-ivk} \psi_T(v) dv, \tag{3.16}$$

where $\psi_T(v)$ equals

$$\frac{\phi_T(v-i)}{iv\phi_T(-i)}, \frac{\phi_T(v)}{iv} \text{ or } \frac{e^{-rT}\phi_T\left(v-(\alpha+1)i\right)}{\alpha^2+\alpha-v^2+i\left(2\alpha+1\right)v},$$

depending on whether we have (3.13), (3.14) or (3.15), respectively.

Now, define $g(v) \equiv e^{-ivk}\psi_T(v)$, the trapezoidal rule then yields

$$\int_0^A g(v)dv \approx \frac{\Delta v}{2} \left[g(v_1) + 2\sum_{n=2}^{N-1} g(v_n) + g(v_N) \right] \tag{3.17}$$

$$= \Delta v \left[\sum_{n=1}^{N} g(v_n) - \frac{1}{2} \left[g(v_1) + g(v_N) \right] \right], \tag{3.18}$$

where $A = N\Delta v$. (Note that we have truncated the interval of integration. The resulting truncation error will be discussed in the next section.) Now, let

$$v_n = (n-1)\,\Delta v \tag{3.19}$$

where $n = 1, ..., N$. Furthermore, we let

$$k_u = -b + \Delta k\,(u-1)\,, \tag{3.20}$$

where $u = 1, ..., N$, to be the grid in the k-domain. The constant $b \in \mathbb{R}$ can be tuned such that the grid is laid around at-the-money strikes, since we are mainly interested in option prices with these particular strikes. Substituting (3.19) and (3.20) in (3.17) (recall that $g(v) \equiv e^{-ivk}\psi_T(v)$) yields

$$\begin{aligned}\int_0^A g(v)dv &\approx \Delta v \left[\sum_{n=1}^{N} e^{-i[(n-1)\Delta v][-b+\Delta k(u-1)]}\psi_T(v_n) - \frac{1}{2}\left[g(v_1)+g(v_N)\right]\right] \\ &= \Delta v \left[\sum_{n=1}^{N} e^{-i\Delta v\Delta k(n-1)(u-1)}e^{i(n-1)b\Delta v}\psi_T(v_n) - \frac{1}{2}\left[g(v_1)+g(v_N)\right]\right]\end{aligned}$$

by setting

$$\Delta v\Delta k = \frac{2\pi}{N},$$

we have - apart from the correcting term - the form of (3.12). Hence FFT can be utilized to evaluate the above sum, *provided that* $g(v_1) = g(0)$ *exists*. This is certainly not the case when one resorts to Gil-Palaez inversion: at $v = 0$ we have singularities in the integrands! Of course, one can use other definitions of the DFT that avoid function evaluations in the origin. However, we will not attempt to do so. The reason is twofold. First, if we find a suitable definition, we have the additional task of implementing an associated FFT-algorithm (or find some software that has implemented this particular form of the DFT). We prefer to use the readily available algorithm as presented in MATLAB. Second, due to the singularities Gil-Palaez inversion will be far less accurate than Carr-Madan inversion as will be illustrated in Chapter 4. Consequently, we just give the final result for Carr-Madan inversion:

$$C_T(k_u) \approx \frac{e^{-\alpha k_u}}{\pi}\,\mathrm{Re}\left\{\Delta v\left[\sum_{n=1}^{N}\omega_N^{(n-1)(k-1)}\left[e^{iv_nb}\psi_T(v_n)\right] - \frac{1}{2}\left[g(v_1)+g(v_N)\right]\right]\right\},$$

where we have imposed the condition

$$\Delta v\Delta k = \frac{2\pi}{N}.$$

Note that the trapezoidal rule has been wrongly represented in Carr & Madan (1999). Instead of the trapezoidal rule, we can also apply the more accurate Simpson's rule. Along the same lines as above, one can easily show that in this case we have

$$C_T(k_u) \approx \frac{e^{-\alpha k_u}}{\pi}\,\mathrm{Re}\left\{\frac{\Delta v}{3}\left[\sum_{n=1}^{N}\omega_N^{(n-1)(k-1)}e^{iv_nb}\psi_T(v_n)\left(3+(-1)^n-\delta_{n-1}\right) - g_S\right]\right\},$$

where

$$g_S = g\left(v_{N-1}\right) + 4g\left(v_N\right).$$

Furthermore, δ_{j-1} denotes the Kronecker delta function that equals one whenever $j = 1$.

3.5 Errors

3.5.1 Introduction

Numerical techniques give rise to errors, so does the numerical scheme illustrated in the previous section. Apart from round-off errors, the total error can be subdivided into the following three categories:

1. Truncation error;
2. Discretization error;
3. Interpolation error.

In this section we will discuss these errors one by one. Our error analysis will be heavily based on the adaptive Simpson's rule, which will first be reviewed. Based on this adaptive quadrature rule we will then make statements about all three types of errors.

The reason to resort to numerical techniques in error analysis is generality. To see this, note that the above Fourier methods are rather modular in the sense that option prices under several stock models are obtained simply by substituting the corresponding characteristic function $\phi_T(\omega)$ in the respective integrands. If we were able to determine (sharp) error bounds analytically, each stock model would presumably give rise to other (cumbersome) expressions. Note that we have used 'were able', since the characteristic functions involved - while available in closed-form - are typically very awkward to handle analytically. This all contrasts with the numerical approach we present, which is far less sensitive to the particular ch.f. at hand.

3.5.2 The Adaptive Simpson's rule

Statements about discretization and truncation errors will be made using the adaptive Simpson's rule. We first review this advanced quadrature rule. In the subsequent sections we then show how it can be applied to determine error bounds.

Recall that the (non-composite) Simpson's rule reads

$$\int_a^b f(x)dx = S(a,b) - \frac{h^5}{90} f^{(4)}(\varsigma), \tag{3.21}$$

for some $\varsigma \in (a,b)$ and in which

$$S(a,b) = \frac{h}{3}\left[f(a) + 4f(a+h) + f(b)\right].$$

Moreover, we have $h = (b-a)/2$. If we divide the interval $[a,b]$ into two equal subintervals, the composite Simpson's rule then yields

$$\int_a^b f(x)dx = S\left(a, \frac{a+b}{2}\right) + S\left(\frac{a+b}{2}, b\right) - \left(\frac{h}{2}\right)^4 \frac{(b-a)}{180} f^{(4)}(\varsigma') \tag{3.22}$$

for some $\varsigma' \in (a,b)$ and where

$$S\left(a, \frac{a+b}{2}\right) = \frac{h}{6}\left[f(a) + 4f\left(a + \frac{h}{2}\right) + f(a+h)\right]$$

$$S\left(\frac{a+b}{2}, b\right) = \frac{h}{6}\left[f(a+h) + 4f\left(a + \frac{3h}{2}\right) + f(b)\right].$$

For the sake of brevity we will henceforth simply refer to the composite Simpson's rule as Simpson's rule. Now equate (3.21) with (3.22) to obtain

$$S(a,b) - \frac{h^5}{90} f^{(4)}(\varsigma) = S\left(a, \frac{a+b}{2}\right) + S\left(\frac{a+b}{2}, b\right) - \frac{1}{16}\left(\frac{h^5}{90}\right) f^{(4)}(\varsigma'). \tag{3.23}$$

If we assume $f^{(4)}(\varsigma) \approx f^{(4)}(\varsigma')$ (the more this holds, the better the result), then we have

$$\frac{h^5}{90} f^{(4)}(\varsigma) \approx \frac{16}{15}\left[S(a,b) - S\left(a, \frac{a+b}{2}\right) - S\left(\frac{a+b}{2}, b\right)\right].$$

Substituting this result into (3.22) gives

$$\left|\int_a^b f(x)dx - S\left(a, \frac{a+b}{2}\right) - S\left(\frac{a+b}{2}, b\right)\right| \approx \frac{1}{15}\left|S(a,b) - S\left(a, \frac{a+b}{2}\right) - S\left(\frac{a+b}{2}, b\right)\right|,$$

hence if

$$\left|S(a,b) - S\left(a, \frac{a+b}{2}\right) - S\left(\frac{a+b}{2}, b\right)\right| < 15\varepsilon, \tag{3.24}$$

then

$$\left|\int_a^b f(x)dx - S\left(a, \frac{a+b}{2}\right) - S\left(\frac{a+b}{2}, b\right)\right| < \varepsilon.$$

Whenever (3.24) does not hold for a sufficiently small value of ε, the adaptive Simpson's rule prescribes to apply Simpson's rule to each of the subintervals $\left[a, \frac{a+b}{2}\right]$ and $\left[\frac{a+b}{2}, b\right]$. Then we determine whether Simpson's rule gives us an accuracy of $\varepsilon/2$ over each of these subintervals. If so, we have found an ε-accurate approximation to the integral at hand; if not, we further subdivide the intervals to repeat the procedure. If the accuracy criterion is only met in one of the subintervals, then the interval where it fails to hold, is further subdivided. For these subintervals we then require an accuracy of $\varepsilon/4$. Each time an interval is halved, we will have an error that is 16 times smaller which can be seen from (3.23), whereas we only require an increased accuracy factor of 2. Although examples can be constructed in which a predefined accuracy can never be met, most practical cases do show convergence. An efficient implementation of this algorithm can for example be found in MATLAB.

3.5.3 Truncation Error

The first type of error that exists in both Gil-Palaez and Carr-Madan inversion is the truncation error. Basically both inversion methods can be reduced to the numerical determination of quantities of the form

$$\int_0^\infty e^{-ivk}\psi_T(v)dv, \tag{3.25}$$

where $\psi_T(v)$ depends on the inversion method used. In Section 3.4 we have truncated the interval of integration, i.e. we approximated the integral

$$\int_0^A e^{-ivk}\psi_T(v)dv.$$

The error obtained in this way is called the truncation error. As we anticipated in Section 3.4, Carr-Madan inversion is the more accurate approach. Hence we will not make any efforts to determine the truncation error for Gil-Palaez inversion. We only present a way to do this for Carr-Madan inversion. First note that

$$|\psi_T(v)| = \frac{e^{-rT}\left|\phi_T\left(v - (\alpha+1)i\right)\right|}{\sqrt{\left(\alpha^2 + \alpha - v^2\right)^2 + (2\alpha+1)^2 v^2}},$$

where

$$|\phi_T(v-(\alpha+1)i)| = \left|\mathbb{E}e^{ivs_T}S_T^{\alpha+1}\right| \leq \mathbb{E}S_T^{\alpha+1},$$

therefore

$$|\psi_T(v)| \leq \frac{c}{1+v^2}$$

where $c = e^{-rT}\mathbb{E}S_T^{\alpha+1}$. Recall that the foregoing expectation can be obtained via a simple function evaluation since

$$\phi_T(-(\alpha+1)i) = \mathbb{E}S_T^{\alpha+1}.$$

Hence we have the (rather loose) bound for the truncation error E_T :

$$|E_T| = \frac{e^{-\alpha k}}{\pi}\left|\int_A^\infty e^{-ivk}\psi_T(v)dv\right| \leq \frac{e^{-\alpha k}}{\pi}\int_A^\infty |\psi_T(v)|\,dv \leq \frac{e^{-\alpha k}}{\pi}\frac{c}{A},$$

which can be made smaller than $\varepsilon > 0$ by choosing

$$A > \frac{e^{-\alpha k}}{\pi}\frac{c}{\varepsilon}. \tag{3.26}$$

This upper limit of integration is far too rough. For all cases observed a much smaller interval of integration is needed to have the ε-accurate result. To see this, we adopt the following procedure which has been inspired by the peaked nature of the integrand in (3.25). Figure 3.1 depicts a typical integrand met in Carr-Madan inversion.

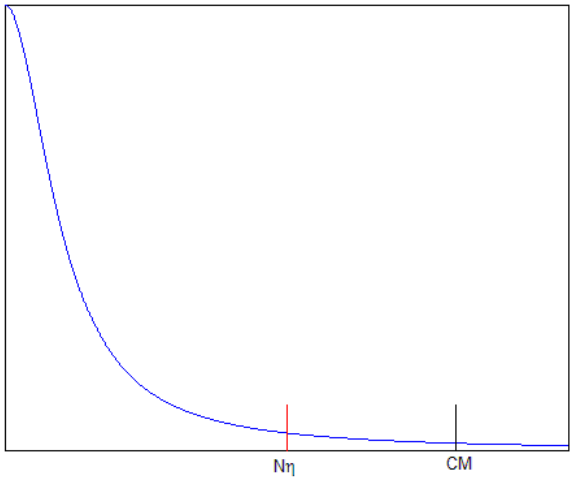


Figure 3.1: A typical peaked integrand one meets in Carr-Madan inversion.

The value CM represents the bound given in (3.26), whereas $N\Delta v$ is the actual upper limit

of integration. Hence in this case we have for the truncation error

$$\begin{aligned} E_T &= \int_{N\Delta v}^{CM} e^{-ivk}\psi_T(v)dv + \int_{CM}^{\infty} e^{-ivk}\psi_T(v)dv \\ &< \int_{N\Delta v}^{CM} e^{-ivk}\psi_T(v)dv + \varepsilon. \end{aligned} \tag{3.27}$$

For the Black-Scholes model (see Chapter 4) and the Heston model (see Chapter 5) there is a rapid decay of the integrand. For typical settings of the numerical integration procedure we would have

$$\int_{N\Delta v}^{CM} e^{-ivk}\psi_T(v)dv << \varepsilon,$$

which for example can be verified using the adaptive Simpson's rule. Therefore, for the Black-Scholes and Heston models equation (3.27) would reduce to

$$E_T < \varepsilon.$$

3.5.4 Discretization Error

In our attempt to evaluate the integral

$$\int_0^A g(v)\,dv,$$

where $g(v) = e^{-ivk}\psi_T(v)$ we resorted to quadrature. More precisely, we used the composite Simpson's rule:

$$\int_0^A g(v)\,dx = \frac{h}{3}\left[g(v_1) + 4g(v_2) + 2g(v_3) + 4g(v_4) + ... + 2g(v_{N-2}) + 4g(v_{N-1}) + g(v_N)\right] + E_N$$

where

$$E_N = \frac{A}{180}(\Delta v)^4 g^{(4)}(\mu),\ \mu \in [0, A].$$

A derivation of this result can for example be found in Burden & Faires (1998). The term

$$E_N = \frac{A}{180}(\Delta v)^4 g^{(4)}(\mu)$$

is called the discretization error. A first year calculus student might immediately suggest to solve the problem

$$\max_{0\le v\le A}\left|g^{(4)}(v)\right|$$

for an error bound on E_N. This is not a viable strategy however. Suppose that we have a closed-form expression for $g^{(4)}(v)$, then we would have to subject this expression to an optimization

procedure for many values of k (recall that $g(v) = e^{-ivk}\psi_T(v)$). This optimization process might be very cumbersome.

Other results for the discretization error exist. See for example Davis & Rabinowitz (1984) for an excellent discussion hereof. However, all these bounds involve the bounding of some higher order derivative, which is merely a consequence of the so-called Lagrange underlying the quadrature rules we have seen. (More on Lagrange can be found in Burden & Faires (1998).) Consequently, we will resort to the adaptive Simpson's rule for the determination of discretization errors.

As indicated in Section 3.5.2, the adaptive Simpson's rule is able to give ε-accurate values for most integrals seen in practice. Thus we simply use this advanced quadrature rule to evaluate the integral

$$\int_0^A e^{-ivk}\psi_T(v)dv$$

with a guaranteed accuracy of $\varepsilon > 0$. The difference between this value and the one obtained from the Simpson's rule is then the desired discretization error.

3.5.5 Error

Recall from Section 3.4 that the grid in the k-space is given by

$$k_u = -b + \Delta k\,(u-1)\,,$$

where $u = 1, ..., N$. Furthermore, the associated grid in the K-domain is exponential:

$$K_u = e^{-b+\Delta k(u-1)}, \tag{3.28}$$

where $u = 1, ..., N$. Due to this exponential grid, one can easily see that in general one cannot obtain option prices for a set of pre-specified strikes within a single computation. To resolve this awkward issue some form of is required. Here we choose cubic spline interpolation to be the interpolation technique used. The main reason to adopt this interpolation technique is that it is a highly accurate method; this will preserve the high accuracy of the results obtained from Simpson's integration. Many software packages (MATLAB inclusive) have implemented this interpolation technique efficiently. A drawback is that it is relatively slow compared to the FFT. Detailed discussions on cubic spline interpolation can be found in Burden & Faires (1998) and Press, Tuekolsky, Vetterling & Flannery (1992).

The procedure to determine the interpolation error is as follows. Based on option prices on the exponential grid (3.28), we apply cubic spline interpolation to determine option prices on intermediate strikes. Now the adaptive Simpson's rule comes into play: we simply compare the interpolated price with the value returned by the adaptive Simpson's rule to determine the interpolation error involved. Here we just anticipate that for typical grid-settings, cubic spline interpolation does a very good job. Numerical evidence can for example be found in Section 4.3.3, where we present numerical results for the Black-Scholes model.

4 The Black-Scholes Model

This chapter presents a case study in which we examine the efficacies of the Fourier inversion methods mentioned in the foregoing chapter in the world of Black and Scholes. The availability of an analytic expression for the option price in this widely known model was the prime motivation to do so. First we present the ch.f. in the Black-Scholes model, which will be derived at a relatively high pace. Readers who are not familiar with this lognormal model should consult Black & Scholes (1973), Björk (1998) or Wilmott (1998). We assume that there are no dividends throughout. Next, numerical results will be presented for both Gil-Palaez and Carr-Madan inversion. The results on Gil-Palaez inversion will be rather brief as it turns out to be a numerically non-attractive approach. On the other hand, a much more detailed analysis will be presented for the highly accurate Carr-Madan technique. Enlightening comments conclude this chapter.

4.1 The characteristic function

The dynamics of the stock price $S(t)$ in a risk-neutral Black-Scholes world follows geometric Brownian motion:

$$dS(t) = rS(t)\,dt + \sigma S(t)\,dW(t),$$

where r, σ and $W(t)$ denote the interest rate, the volatility of the stock price and Brownian motion, respectively. We remark that the interest rate and volatility are allowed to be functions of time. Utilizing Itô's formula we can solve for S_T explicitly:

$$S_T = e^{\left(\left[r-\frac{1}{2}\sigma^2\right]T+\log S_0+\sigma W(T)\right)},$$

from which we can see that S_T is lognormally distributed. Hence for the characteristic function $\phi_T(u)$ of $\log S_T$ we have

$$\phi_T(u) = e^{i\left(\log S_0+\left(r-\sigma^2/2\right)T\right)u-\sigma^2 T u^2/2}.$$

4.2 Results on Gil-Palaez Inversion

Recall from Section 3.2 that the option price $C_T(K)$ of a vanilla call with time to expiry T and strike K can be expressed as

$$C_T(K) = S_0\Pi_1 - Ke^{-rT}\Pi_2,$$

where

$$\Pi_1 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\frac{\phi_T(v-i)}{\phi_T(-i)}\frac{e^{-iv\log K}}{iv}\right)dv \tag{4.1}$$

$$\Pi_2 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\phi_T(v)\frac{e^{-iv\log K}}{iv}\right)dv, \tag{4.2}$$

and in which

$$\phi_T(v) = \mathbb{E}e^{iv\log S_T}.$$

In a Black-Scholes world we have

$$\phi_T(v) = e^{i\left(\log S_0+\left(r-\sigma^2/2\right)T\right)v-\sigma^2 T v^2/2}.$$

From Section 3.4 we also know that the integrals in (4.1) and (4.2) are not amenable to the FFT as available in MATLAB as the integrals have singularities in the origin. Whether the FFT is applied or not, basically, we just numerically integrate certain integrals with the aid of the Simpson's rule. The only difference lies in the speed at which a set of integrals is evaluated; the FFT does not have any influence on the accuracy at all. Having said this, let us apply the Simpson's rule to get the option price for $K = 100$, the other parameters being

$$S = 100, T = 1, \sigma = 0.15, r = 0.05.$$

See Figure 4.1 for plots of the resulting integrands. The smallest depicted value is $v = 0.12$. Note that while the integrands are not defined in the origin, they do not explode. In fact, the imaginary parts of $\frac{\phi_T(v-i)}{\phi_T(-i)} \frac{e^{-iv \log K}}{iv}$ and $\phi_T(v) \frac{e^{-iv \log K}}{iv}$ grow much more rapidly near the origin.

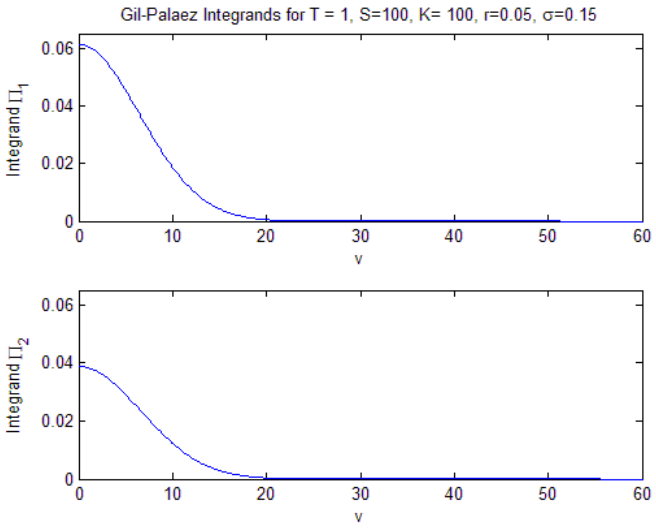


Figure 4.1: Typical integrands involved in Gil-Palaez Inversion.

Due to the fact that the integrands are not defined in the origin, it is difficult to detail the errors. Here we just show some results for certain grid settings and anticipate that based on these settings, Carr-Madan inversion would yield much better results (see the next section). We lay the following grid for the numerical integration by Simpson's rule:

$$v_j = j\Delta v, \ j = 1, ..., N-1 \tag{4.3}$$

where

$$\Delta v = \frac{2\pi}{N\Delta k}$$

for some $\Delta k \in \mathbb{R}$. Table 4.1 reports results for $\Delta k = 0.01$ on the order of magnitude of the absolute relative error $|C_{GP} - C_{BS}|/S$, denoted by $O(|C_{GP} - C_{BS}|/S)$, where C_{GP} and C_{BS} represent option prices as obtained from Gil-Palaez inversion and the Black-Scholes formula, respectively. Table 4.2 shows results for $\Delta k = 0.025$. The primary reason for the relatively large errors is because the grid (4.3) basically discards the first parts of the integrands (the singularities necessitate us to do so). Obviously, results become more and more accurate as $\Delta v \to 0$. The rate of convergence is far too slow for Gil-Palaez inversion to be useful in practice however.

N	Δv	$O\left(\left\|C_{GP}-C_{BS}\right\|/S\right)$
2^{10}	0.6136	10^{-3}
2^{11}	0.3068	10^{-3}
2^{12}	0.1534	10^{-3}
2^{13}	0.0767	10^{-4}
2^{14}	0.0383	10^{-4}

Table 4.1: Numerical results on Gil-Palaez Inversion for $\Delta k = 0.01$

N	Δv	$O\left(\left\|C_{GP}-C_{BS}\right\|/S\right)$
2^{10}	0.2454	10^{-3}
2^{11}	0.1227	10^{-4}
2^{12}	0.0614	10^{-4}
2^{13}	0.0307	10^{-4}
2^{14}	0.0153	10^{-4}

Table 4.2: Numerical results on Gil-Palaez Inversion for $\Delta k = 0.025$

This is all we want to say about Gil-Palaez inversion. Let us now examine numerical results from the more accurate Carr-Madan inversion.

4.3 Results on Carr-Madan Inversion

As anticipated, Carr-Madan inversion is far more accurate than inversion by Gil-Palaez. Equally important, it is formulated in such a way that it is suitable for the application of the FFT (for example as available in MATLAB). Therefore, we will present more details for the practical implementation of this technique. First we will examine the integrands involved in a rather qualitative way. Then we present a discussion on some ways to choose α, the very parameter that made the integrands singularity-free. We conclude with numerical results.

4.3.1 The Black-Scholes Integrand

Recall that Carr-Madan inversion states that for the option price we have

$$C_T(k) = \frac{e^{-\alpha k}}{\pi} \int_0^\infty \mathrm{Re}\left\{e^{-ivk}\psi_T(v)\right\} dv \tag{4.4}$$

where

$$\psi_T(v) = \frac{e^{-rT}\phi_T\left(v-(\alpha+1)i\right)}{\alpha^2+\alpha-v^2+i\left(2\alpha+1\right)v}$$

and where $\phi_T(\cdot)$ denotes the characteristic function of the $\log S_T$. For the Black-Scholes model the integrand in equation (4.4) reduces to

$$\frac{\exp\left(-\frac{\sigma^2 T v^2}{2}+\frac{\alpha^2\sigma^2 T}{2}+\alpha s+\alpha T r+\frac{\sigma^2 T\alpha}{2}+s\right)}{\alpha^4+2\alpha^3+2\alpha^2v^2+\alpha^2+2\alpha v^2+v^4+v^2} g\left(\alpha,k,r,s,\sigma,T,v\right), \tag{4.5}$$

where

$$g\left(\alpha,k,r,s,\sigma,T,v\right) = \left(\alpha^2+\alpha-v^2\right)\cos\left(\left[k-c\right]v\right) - v\left(2\alpha+1\right)\sin\left(\left[k-c\right]v\right) \tag{4.6}$$

and in which

$$c = \sigma^2 T\alpha + s + Tr + \frac{\sigma^2 T}{2}. \tag{4.7}$$

Several interesting observations can be inferred from these expressions. To begin with, the Black-Scholes integrand decreases rapidly with v. Furthermore, we see from (4.6) and (4.7) that by increasing any of the parameters σ, T, α, s and r, we get a more fluctuating integrand. Moreover, the magnitudes of these fluctuations get larger which can be seen from the exponential term in (4.5). Pictures can considerably support us in understanding these observations. The most striking observations are visualized next. Unless stated otherwise, the following plots are generated based on the parameters:

$$S = 100, K = 100, T = 1, \sigma = 0.4, r = 0.05, \alpha = 3.5.$$

In fact, for practical ranges of the above parameters, only (the interplay of) T, S/K and α have noticeable influences on the integrand. Figure 4.2 shows Black-Scholes integrands for $T = 1$ (upper) and $T = 10$ (lower). As anticipated, we get more fluctuations and larger functional values.

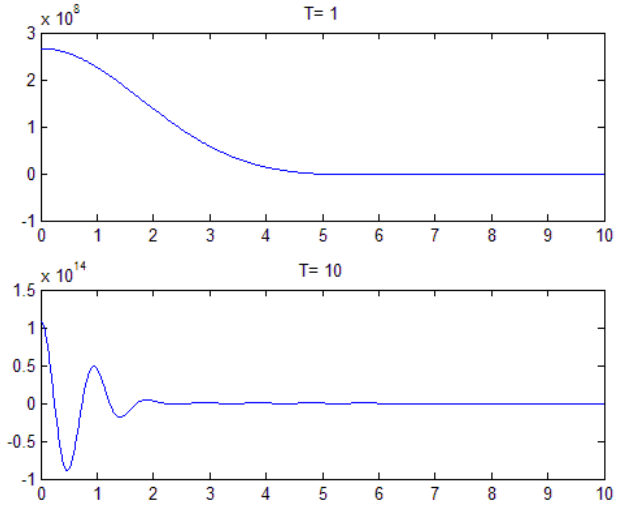


Figure 4.2: The influence of T on the Black-Scholes integrand. Upper: $T = 1$, lower: $T = 10$.

The strike k (or K if you will) appears solely in the sine and cosine terms in (4.6). Since

$$K \to 0 \Leftrightarrow k \to -\infty,$$

we see that both the cosine and sine terms will fluctuate rapidly as K tends to zero. This will cause the integrand to be extremely oscillatory, while the absolute values do not grow in magnitude. Nonetheless, this is sufficient to pose a huge problem from a quadrature point of view. The same is true when $K \to \infty$. This latter case is of less practical interest however. In fact, it is the so-called moneyness S/K that determines the oscillatory nature of the integrand. Figure 4.3 summarizes.

The last parameter to be considered is α. Recall that α is not a parameter of the Black-Scholes model; it is a degree of freedom that is available in Carr-Madan inversion. Momentarily,

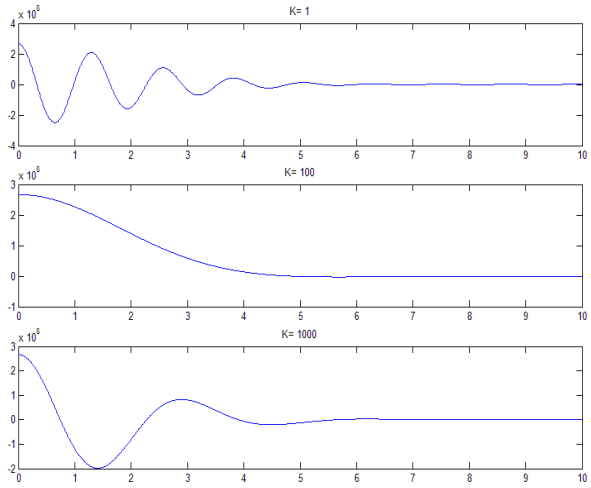


Figure 4.3: The influence of K on the Black-Scholes integrand. Upper: $K = 1$, middle: $K = 100$, lower: $K = 1000$.

the only thing that matters is that we have a choice for α. Because of this, we have a means of manipulating the integrand to our convenience. Specifically, from the above expressions for the Black-Scholes integrand one can easily see that α enables us to control the intensity of the fluctuations and the magnitude of the functional values. Figure 4.4 shows that small values of α reduce the largest attainable function value considerably. Moreover, oscillations are reduced as well. Ways to choose α are the topic of the next section.

4.3.2 On the choice of α

At this point it is inevitable to comment on the choice of α. From the previous section it should be clear that $\alpha > 0$ should ideally be a function of σ, T, s, r and k. Whenever certain combinations of these parameters give rise to cumbersome integrands, α can then (theoretically) be utilized to make them more tractable. Presumably, a small value of α is favorable since this reduces both the oscillations and the magnitudes hereof. However, choosing α too small can turn the integrand into a sort of impulse function, which is not tractable at all from a numerical integration point of view. This follows from the fact that in the origin $v = 0$, the Black-Scholes integrand equals

$$\frac{\exp\left(\frac{\alpha^2\sigma^2 T}{2} + \alpha s + \alpha T r + \frac{\sigma^2 T\alpha}{2} + s\right)}{\alpha\left(\alpha + 1\right)}, \tag{4.8}$$

therefore we see that

$$\lim_{\alpha \to 0} \frac{\exp\left(\frac{\alpha^2\sigma^2 T}{2} + \alpha s + \alpha T r + \frac{\sigma^2 T\alpha}{2} + s\right)}{\alpha\left(\alpha + 1\right)} = \infty.$$

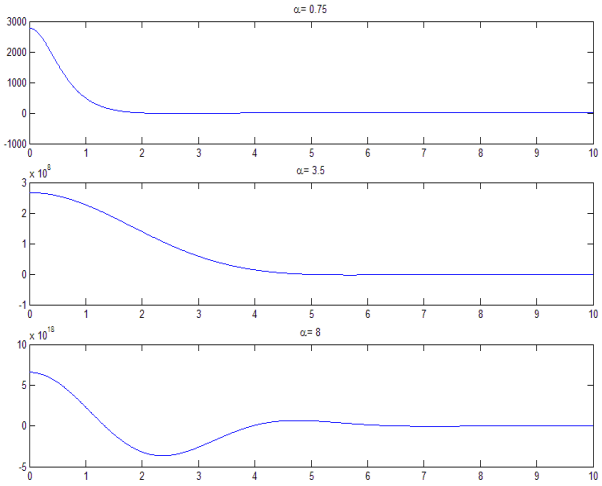


Figure 4.4: The influence of α on the Black-Scholes integrand. Upper: $\alpha = 0.75$, middle: $\alpha = 3.5$, lower: $\alpha = 8$.

Note that (4.8) also tends to infinity for $\alpha \to \infty$. On the other hand, for $v > 0$ and by letting $\alpha \to 0$, the integrand becomes:

$$\frac{\exp\left(-\frac{\sigma^2 T v^2}{2} + s\right)}{v^4 + v^2}\left[-v^2 \cos\left(\left[k - c\right] v\right) - v \sin\left(\left[k - c\right] v\right)\right],$$

which decreases very fast as a function of v because of the exponential term. Hence we see that the Black-Scholes integrand indeed resembles more and more the impulse function. See Figure 4.5 for a pictorial representation of the situation.

Carr & Madan (1999) suggest to use the condition

$$\mathbb{E} S_T^{\alpha+1} < \infty \tag{4.9}$$

to determine an upper bound for α. One fourth of the maximum value of α such that (4.9) holds is allegedly a good choice. However, all moments are finite in a Black-Scholes setting. Hence this strategy is of no use here.

Another heuristic is given by Schoutens, Simons & Tistaert (2004). They proposed the value $\alpha = 0.75$ for a whole range of stock models. Although Black-Scholes is not included in these models, we nevertheless examine the effectiveness of this value.

Now we present a more rational strategy to choose α. Before stating it, note that the Black-Scholes integrand attains its maximum at $v = 0$, which is proved in the following lemma.

Lemma 1 *Let $v \geq 0$. The Black-Scholes integrand*

$$\mathrm{Re}\left\{e^{-ivk} \frac{e^{-rT}\phi_T\left(v - (\alpha+1)i\right)}{\alpha^2 + \alpha - v^2 + i\left(2\alpha + 1\right)v}\right\},$$

where

$$\phi_T(v) = e^{i\left(\log S_0 + \left(r - \sigma^2/2\right)T\right)v - \sigma^2 T v^2/2}$$

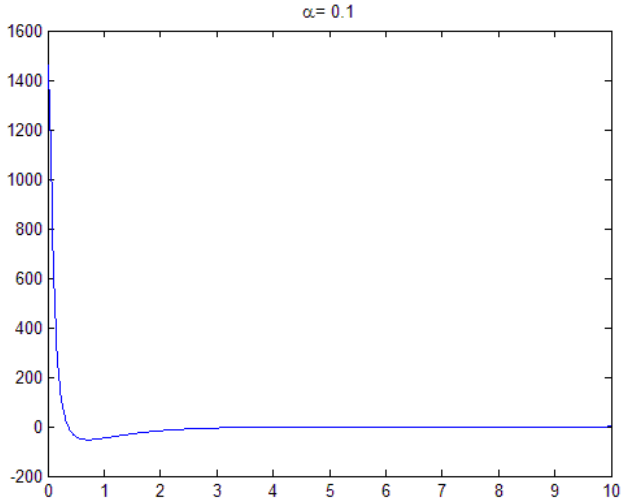


Figure 4.5: The Black-Scholes integrand resembles more and more the impulse function as $\alpha \to 0$. For the integrand depicted we have $S = 100$, $K = 100$, $T = 1$, $\sigma = 0.4$, $r = 0.05$.

attains its maximum at $v = 0$.

Proof. From (4.8) we see that the statement is equivalent with

$$\mathrm{Re}\left\{e^{-ivk}\psi_T(v)\right\} \leq \frac{\exp\left(\frac{\alpha^2\sigma^2 T}{2} + \alpha s + \alpha T r + \frac{\sigma^2 T\alpha}{2} + s\right)}{\alpha\left(\alpha+1\right)}$$

for $v \geq 0$. This follows since

$$\left|\mathrm{Re}\left\{e^{-ivk}\psi_T(v)\right\}\right| \leq \left|e^{-ivk}\psi_T(v)\right| = \left|\psi_T(v)\right|,$$

where

$$\left|\psi_T(v)\right| = \frac{e^{-rT}\left|\phi_T(v-(\alpha+1)i)\right|}{\left|\alpha^2+\alpha-v^2+i\left(2\alpha+1\right)v\right|}.$$

Moreover,

$$\begin{aligned}\left|\phi_T(v-(\alpha+1)i)\right| &= \left|e^{i\left(s+\left(r-\sigma^2/2\right)T\right)(v-(\alpha+1)i))-\sigma^2T(v-(\alpha+1)i))^2/2}\right| \\ &= \exp\left(\frac{\alpha^2\sigma^2T}{2} + \alpha s + \alpha T r + \frac{\sigma^2T\alpha}{2} + s + rT - \frac{\sigma^2Tv}{2}\right),\end{aligned}$$

which yields

$$\begin{aligned}
|\psi_T(v)| &= \frac{\exp\left(\frac{\alpha^2\sigma^2 T}{2} + \alpha s + \alpha T r + \frac{\sigma^2 T\alpha}{2} + s - \frac{\sigma^2 T v}{2}\right)}{\left|\alpha^2 + \alpha - v^2 + i\left(2\alpha + 1\right)v\right|} \\
&= \frac{\exp\left(\frac{\alpha^2\sigma^2 T}{2} + \alpha s + \alpha T r + \frac{\sigma^2 T\alpha}{2} + s - \frac{\sigma^2 T v}{2}\right)}{\left|\left(v - \left(\alpha + 1\right)i\right)\left(v - \alpha i\right)\right|} \\
&\leq \frac{\exp\left(\frac{\alpha^2\sigma^2 T}{2} + \alpha s + \alpha T r + \frac{\sigma^2 T\alpha}{2} + s\right)}{\left|\left(v - \left(\alpha + 1\right)i\right)\left(v - \alpha i\right)\right|} \\
&\leq \frac{\exp\left(\frac{\alpha^2\sigma^2 T}{2} + \alpha s + \alpha T r + \frac{\sigma^2 T\alpha}{2} + s\right)}{\alpha\left(\alpha + 1\right)}
\end{aligned}$$

completing the proof. ■

In order to determine a good value for α, we propose to (numerically) minimize the maximum of the integrand, i.e. to solve the following optimization problem:

$$\min_{\alpha > 0} \frac{\exp\left(\frac{\alpha^2\sigma^2 T}{2} + \alpha s + \alpha T r + \frac{\sigma^2 T\alpha}{2} + s\right)}{\alpha\left(\alpha + 1\right)},$$

which intuitively would yield a nice integrand in the sense that both large variations in function values as well as oscillations are reduced. Note that in this strategy we have discarded the dependence of α on k. To give a flavor of the function to be minimized we present a plot of it, see Figure 4.6 where $r = 0.05, T = 1, \sigma = 0.15$ and $S = 100$. One possible way to solve the optimization problem is "setting the derivative equal to zero". One can easily show that in this case the problem reduces to finding zeros of a (cumbersome) third order polynomial. In fact, this is the strategy we (successfully) followed. Let us now turn to numerical results.

4.3.3 Numerical Results

In Section 3.4 we have imposed the relation

$$\Delta v \Delta k = \frac{2\pi}{N} \tag{4.10}$$

between the grids in the v-domain and the k-domain in order to be able to utilize the FFT to evaluate the Fourier integrals involved. Equation (4.10) is oftentimes referred to as a reciprocity relation. The reason is obvious: for a fixed value of N, reducing Δv would result in an increase of Δk and vice versa. Practically, we have the following situation. Accurate integration requires a fine grid in the v-domain; by (4.10) this would result in a relatively coarse grid in the k-domain, which is undesired since in general - as indicated in Section 3.5.5 - interpolation is required. Increasing N remedies this problem.

Unless stated otherwise, in this section we have used the following grid spacings to obtain the results presented:

$$\Delta k = 0.025, \ \Delta v = \frac{2\pi}{N\Delta k}, \tag{4.11}$$

where $N = 2^{11} = 2048$. Following the procedure as described in Section 3.5.3, one can easily verify that the truncation error is negligible for the above grid settings. Hence the total error only consists of the discretization error in case we don't apply interpolation.

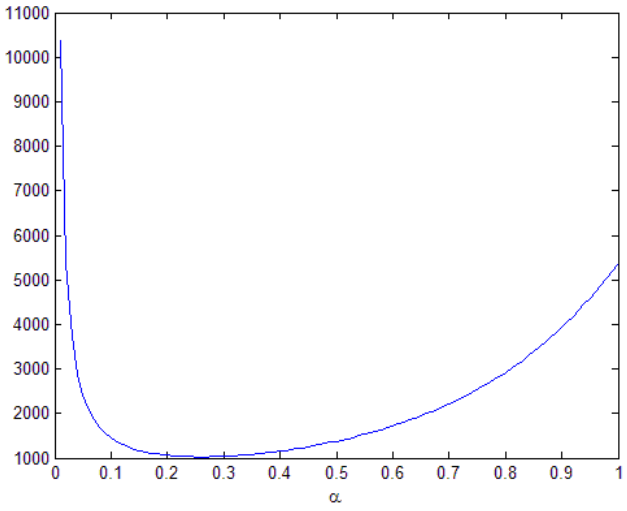


Figure 4.6: A typical function one has to face when the maximum of the Black-Scholes integrand is to be minimized.

In order to determine the best choice for α (0.75 or the minmax-value), we have performed the following experiment. We randomly selected σ, T, S and r from the typical ranges $[0.15, 0.5]$, $[1, 20]$, $[0, 100]$ and $[0, 0.15]$, respectively. The maximum absolute relative error $|C_{CM} - C_{BS}|/S$ is then determined over the range $\frac{1}{2}S \leq K \leq 2S$. ($C_{CM}$ is the price obtained by Carr-Madan inversion.) Figure 4.7 shows 100 realizations of this experiment when $\alpha = 0.75$. The minmax-strategy yields Figure 4.8.

A quick glance is enough to observe the substantial differences in accuracy. The explanation for this rather remarkable result is that the α resulting from the minmax-strategy is sufficiently small to give a much more rapidly decreasing integrand, which Simpson's rule cannot handle that well for the specified grid settings. A typical case is illustrated in Figure 4.9, where we have plotted two integrands as a function of v. Here the minmax-value for α was equal to 0.29.

From Section 4.3.1 we know that we might expect reduced accuracy when S/K is small (or large) or when T is large, since in these cases we have a rather oscillatory integrand. Let us make this precise here. It turns out that only certain combinations of S/K, T and σ cause problems. For example, the settings $T = 20$ and $\sigma = 0.4$ result in relative errors of at most 1 for small strikes. See Figure 4.10 in which we have plotted the absolute relative error versus the strike. Here $S = 100$ and $r = 0.15$. Doubling N to 2^{12} reduces the error significantly. For the above settings the error will then be of order 10^{-12}.

Less problematic but still interesting is that for extremely low values for the maturities (smaller than 10^{-5}) we seem not be able to get higher accuracies than 10^{-4}, even if N is rather large (2^{14}). This might be explained by the fact that small values for T give rise to rather peaked integrands. See Figure 4.11 for an example.

In the above, no interpolation has been applied. Using the grid spacings (4.11), cubic spline interpolation turns out to be highly accurate. In fact, interpolated values are typically one order of magnitude less accurate. In order to present some numerical evidence, let $S = 100, T = 1, \sigma = 0.4, r = 0.15, \alpha = 0.75$. The upper part of Figure 4.12 then shows the analytical Black-Scholes price (solid line) and the cubic spline interpolation values as obtained from interpolation based

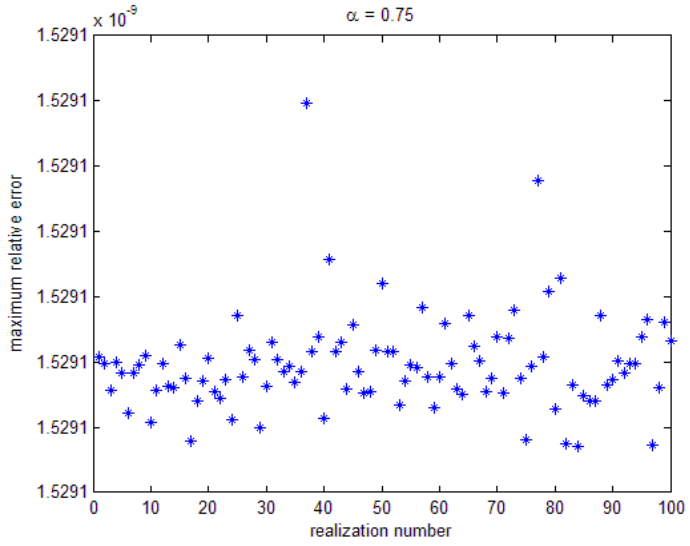


Figure 4.7: Maximum relative error for 100 realizations when $\alpha = 0.75$.

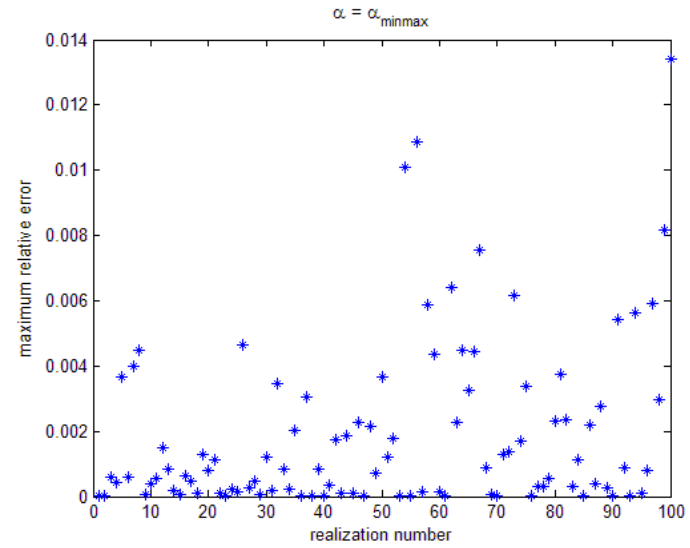


Figure 4.8: Maximum relative error for 100 realizations when α is chosen according to the minmax-strategy.

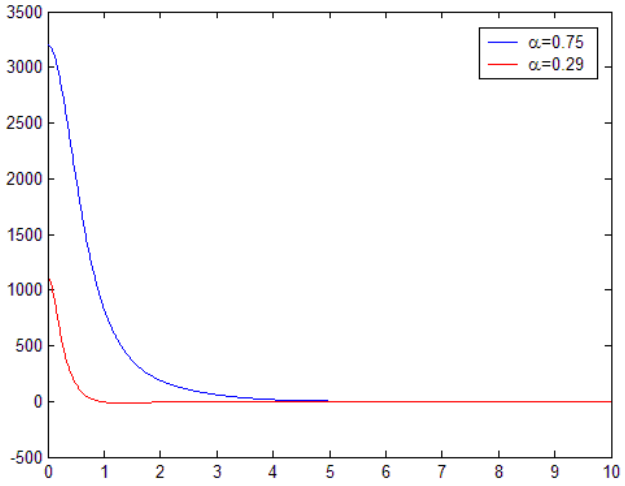


Figure 4.9: We see a much more rapidly decreasing integrand when we set α to be equal to the minmax-value (0.29). The other parameters for this plot are: $r = 0.05$, $S = 100$, $T = 2$, $\sigma = 0.4$, $K = 100$.

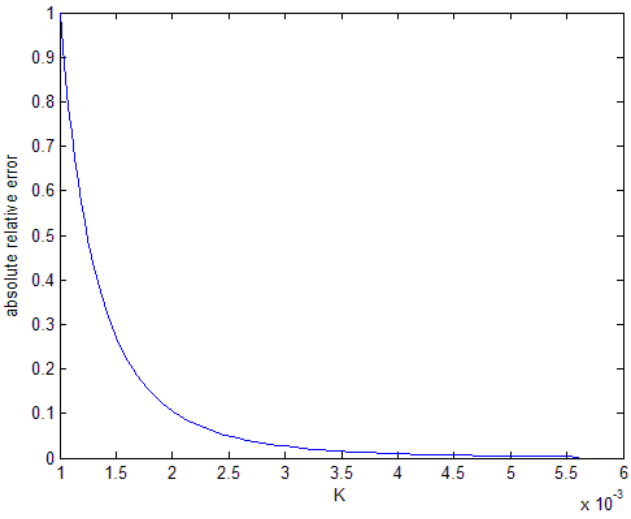


Figure 4.10: Large values for T and σ yield large relative errors.

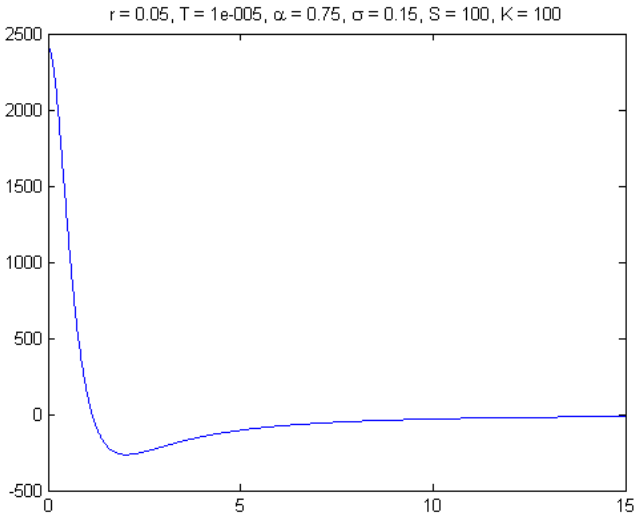


Figure 4.11: A rather peaked integrand as a result of a small value for T.

on the grid spacings (4.11) (circles). The lower part depicts the associated absolute relative error, which is one order of magnitude lower than the non-interpolated-values (cf. Figure 4.7). Similar results can be observed for other parameter combinations.

4.4 Concluding remarks

There is only one conclusion possible: Carr-Madan inversion outperforms Gil-Palaez inversion dramatically. This can be ascribed to the singularities in the integrands involved in Gil-Palaez inversion. Carr-Madan inversion on the other hand, is singularity-free. In order to achieve this, a parameter α has been introduced (of course, the major reason to introduce α was to make the Fourier transform of the modified option price convergent). Choosing $\alpha = 0.75$ typically gives highly accurate results (relative error of order 10^{-9}). It is also computationally less intensive than the minmax-value. Therefore, we recommend to take this value for the Black-Scholes model. The aforementioned relative error has been obtained using the grid spacings:

$$\Delta k = 0.025, \ \Delta v = \frac{2\pi}{N\Delta k}, \tag{4.12}$$

where $N = 2^{11} = 2048$. For these grid settings, the truncation error is negligible. The interpolation error is small as well (typically 10^{-8}). To illustrate the effectiveness of these settings, suppose that we set $N = 2^{10}$ instead, the discretization and interpolation errors will then increase significantly (more than thousand times has been observed.) Certain combinations of parameters (we mention especially the combination of a small K (10^{-4}), large T (>20) and high volatility σ (>0.4)) yields less tractable integrands, either because of the oscillatory nature or the very peaked integrand. Increasing N remedies a lot in this case.

Finally, in this chapter we have often seen the designation "peaked integrand". Obviously, the Simpson's rule (and its associated uniform grid) is not particularly designed to handle these type of functions. However, a uniform grid is required for the application of the FFT! Hence

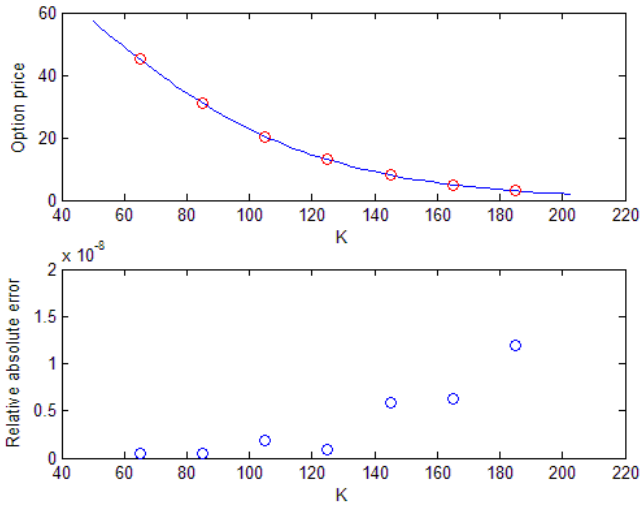


Figure 4.12: Cubic spline interpolation errors. Upper: Here the true Black-Scholes price (solid line) is depicted together with interpolated values. Lower: The resulting errors in the interpolated prices.

there is a trade-off: Whenever we want to apply the FFT (say Simpson's rule) to integrate peaked integrals we have to choose an overly fine grid in order to obtain a certain accuracy (the FFT remains fast though). On the other hand, when we don't apply the FFT, a much coarser grid can be chosen (as illustrated below). Numerical integration is then relatively fast, but is it faster than the FFT? The amazing answer is yes, at least, if we take into account that in practice we are interested in relatively few strikes, say 10, per time to maturity T. The direct integration of these 10 integrals is then faster than the FFT *plus* an additional application of cubic spline interpolation. In fact, it is the latter interpolation step that slows down the whole process. To illustrate this discussion, consider the problem of determining the call prices for vanilla's with the ten strikes $K = 50, 60, 70, ..., 130, 140$. The other parameters are given by $r = 0.05, T = 1, \sigma = 0.2, S = 100$. Recall that the we have the integrand

$$C_T(k) = \frac{e^{-\alpha k}}{\pi} \int_0^\infty \mathrm{Re}\left\{e^{-ivk}\psi_T(v)\right\} dv \tag{4.13}$$

where

$$\psi_T(v) = \frac{e^{-rT}\phi_T\left(v - (\alpha+1)i\right)}{\alpha^2 + \alpha - v^2 + i\left(2\alpha+1\right)v}.$$

Note that $\psi_T(v)$ does not depend on k, which is also beneficial for (repeated) direct integration since $\psi_T(v)$ can be precomputed. Now perform the change of variable

$$v = \sinh\left(y\right), \ y \in \mathbb{R}.$$

Equation (4.13) can then be rewritten as

$$C_T(k) = \frac{e^{-\alpha k}}{\pi} \int_0^{\infty} \mathrm{Re}\left\{ e^{-ik\sinh(y)} \psi_T(\sinh(y)) \right\} \cosh(y)\, dy. \tag{4.14}$$

Figure 4.13 shows the resulting integrand for the above choices of the parameters ($K = 100$). It also depicts the associated unstretched integrand.

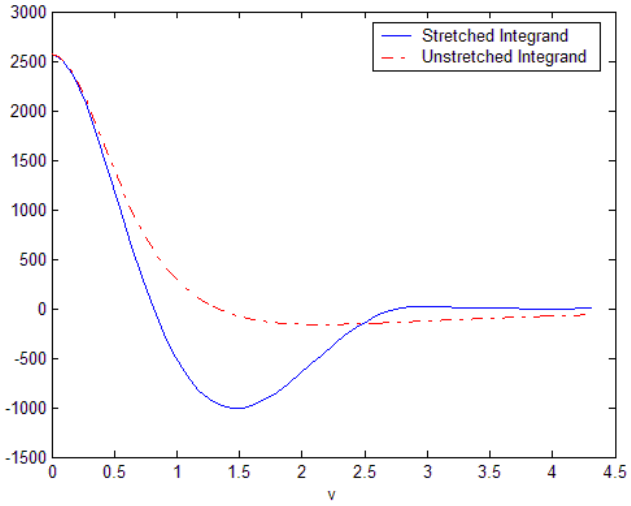


Figure 4.13: By a change of variable we get a less peaked integrand. Here $K = 100$.

Now we apply ten times (we have ten different strikes) Simpson's rule to evaluate (4.14), with $N = 50$ and $\Delta v = 0.09$ (the truncation error is then negligible). Results are shown in Figure 4.14. Using the recommended grid settings in (4.12), the application of Carr-Madan inversion (cubic spline interpolation inclusive) yields similar results (the results differ at most one order of magnitude). However, when we compare the time the FFT and cubic spline interpolation need together, versus that needed by the ten direct integrations, we must conclude that the direct integrations are about four times as fast. In fact, the FFT itself is faster than the direct integrations; it is the interpolation step - which is always needed in practice - that makes Carr-Madan inversion relatively slow. Of course, one can use other types of interpolation that run faster. For example, if we use linear interpolation the two approaches would require about the same time. However, relative errors would reduce to 10^{-5} (which might still be acceptable for some practitioners). Naturally, for the direct integrations, one can expect the same type of problems (e.g. as a result of a small S/K ratio) as in Carr-Madan inversion.

Finally, the reader is challenged to find better change-of-variable-transformations that would further speed up the direct integrations.

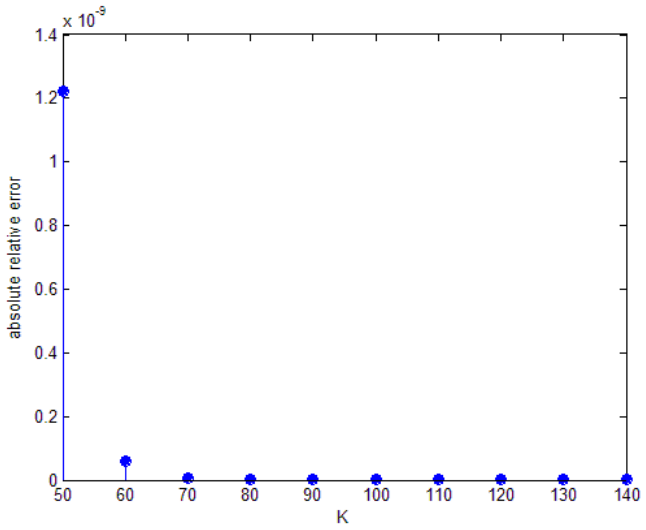


Figure 4.14: The absolute relative error after a change of variable.

5 The Heston Stochastic Volatility Model

5.1 Why Stochastic Volatility

It is incontestable that the Black-Scholes model forms the foundation of modern finance. Certain underlying assumptions are questionable however. To name a few: in the Black-Scholes world delta hedging is supposed to be a continuous process. Furthermore, no transaction costs are associated with this rebalancing of the portfolio. Also, parameters like the volatility are supposed to be known constants (or known deterministic functions of time). For a detailed discussion of these defects and ways to overcome them, see Wilmott (1998).

In this chapter we will present the Heston stochastic volatility stock model (or Heston model for short) that allows volatility to be random. The idea to model volatility as a random variable comes from practical data that indicate the highly variable and unpredictable nature of volatility. Moreover, return distributions under stochastic volatility models typically have fatter tails than their lognormal counterpart, hereby being more realistic. But the most cited argument to allow volatility to be random involves implied volatilities: Let C_{market} denotes the market price of a vanilla call with strike K and time to maturity T. Then we can ask ourselves what volatility σ should be substituted into the Black-Scholes formula in order to recover the market price, i.e. we want to solve for σ in the equation

$$C_{BS}\left(r,t,S,K,T,\sigma\right) = C_{market}, \tag{5.1}$$

where $C_{BS}\left(r,t,S,K,T,\sigma\right)$ denotes the Black-Scholes price of a vanilla call. The σ obtained in this way is called the implied volatility. One easily sees that if

$$\max\left\{S - Ke^{-r(T-t)}, 0\right\} \leq C_{market} \leq S,$$

which simply means that we have an arbitrage-free (Black-Scholes) price, then (5.1) can be solved

for a unique value of the implied volatility since the vega $\partial C_{BS}/\partial\sigma$ is strictly increasing for $t < T$:

$$\frac{\partial C_{BS}}{\partial \sigma} = \frac{Se^{-d_1^2/2}\sqrt{T-t}}{\sqrt{2\pi}} > 0,$$

where

$$d_1 = \frac{\log\left(\frac{S}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}.$$

Now consider options on the same underlying, the same time to expiry T, but with various strikes K. Imagine what we would get if we plot implied volatilities versus strikes under the Black-Scholes model. Indeed, we would get a constant implied volatility, independent of the strike. This is however (never) observed in the real world. Typically, we would get curves as shown in Figure 5.1. Oftentimes this curve is called a smile. And stochastic volatility models do give rise to these smiles!

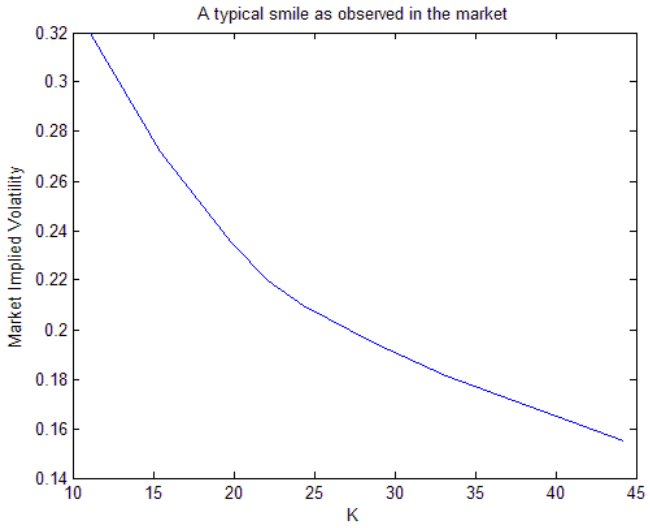


Figure 5.1: A smile based on market data of ING options from Jan 12, 2005.

Stochastic volatility models also introduce new (awkward) issues. For example, financial markets in a stochastic volatility environment are incomplete, which - loosely speaking - means that derivative securities cannot be perfectly replicated with just a stock and a bond (contrary to the Black-Scholes world). This is of particular interest to hedgers. Another characteristic of incomplete markets is that there is no unique risk-neutral measure (again as opposed to the Black-Scholes world). However, it is outside the scope of this thesis to discuss these and other interesting issues. For a discussion on incompleteness in the Heston model, see Heath & Schweizer (2000).

Finally, there seems to be no generally accepted stochastic volatility model. The Heston stochastic volatility model that we discuss is among the more popular models. In Fouque, Papanicolaou & Ronnie Sircar (2001) one can find popular stochastic volatility models.

5.2 The Heston Model and its Pricing PDE

In this section we present the Heston model. Specifically, we will derive the pricing partial differential equation (PDE), which forms the basis of the derivation of the characteristic function in the next section. The Heston model was first introduced in Heston (1993). Here we follow Gatheral (2003*a*).

In the Heston model we have two stochastic differential equations, one for the underlying asset price $S(t)$ and one for the variance $v(t)$ of $\log S(t)$:

$$\begin{aligned} dS(t) &= \mu(t) S(t) dt + \sqrt{v(t)} S(t) dW_1 \\ dv(t) &= -\lambda(v(t) - \overline{v}) dt + \eta \sqrt{v(t)} dW_2. \end{aligned}$$

Here $\lambda \geq 0$, $\overline{v} \geq 0$ and $\eta > 0$ are called the speed of mean reversion, the mean level of variance and the volatility of the volatility, respectively. Furthermore, the Brownian motions W_1 and W_2 are assumed to be correlated with correlation coefficient ρ. The SDE for the variance can be recognized as a mean-reverting square root process, a process originally proposed by Cox, Ingersoll & Ross (1985) to model the spot interest rate. One can easily see why it is called mean-reverting: if $v(t)$ exceeds its mean $\overline{v}$, the term $-\lambda(v(t) - \overline{v}) dt$ drives the variance back to the mean. The same holds when $v(t)$ is below its mean. Based on the so-called classical Feller boundary classification criteria (see for example Karlin & Taylor (1981)) one can show that under certain assumptions $v(t)$ as defined by the above SDE cannot attain negative values.

Now we derive a pricing PDE for the Heston model. To do so, recall from Itô's lemma that for every function $V(S_1, S_2, t)$ - twice differentiable with respect to S_1 and S_2, and once with respect to t - we have

$$dV = \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S_1} dS_1 + \frac{\partial V}{\partial S_2} dS_2 + \frac{1}{2} b_1^2 \frac{\partial^2 V}{\partial S_1^2} dt + \rho b_1 b_2 \frac{\partial^2 V}{\partial S_1 \partial S_2} dt + \frac{1}{2} b_2^2 \frac{\partial^2 V}{\partial S_2^2} dt,$$

in which

$$\begin{aligned} dS_1 &= a_1(S_1, S_2, t) dt + b_1(S_1, S_2, t) dW_1 \\ dS_2 &= a_2(S_1, S_2, t) dt + b_2(S_1, S_2, t) dW_2, \end{aligned}$$

where W_1 and W_2 are correlated with correlation coefficient ρ. To proceed, consider a self-financing portfolio with value Π consisting of an option with value $V(S, v, t)$, $-\Delta$ units of the underlying asset and, in order to hedge the risk associated with the random volatility, $-\Delta_1$ units of another option with value $V_1(S, v, t)$. Hence

$$\Pi = V - \Delta S - \Delta_1 V_1.$$

Moreover, using Itô's lemma and collecting terms yields

$$\begin{aligned} d\Pi &= \left(\frac{\partial V}{\partial t} + \frac{1}{2} v S^2 \frac{\partial^2 V}{\partial S^2} + \rho \eta S v \frac{\partial^2 V}{\partial S \partial v} + \frac{1}{2} v \eta^2 \frac{\partial^2 V}{\partial v^2} \right) dt \\ &\quad - \Delta_1 \left(\frac{\partial V_1}{\partial t} + \frac{1}{2} v S^2 \frac{\partial^2 V_1}{\partial S^2} + \rho \eta S v \frac{\partial^2 V_1}{\partial S \partial v} + \frac{1}{2} v \eta^2 \frac{\partial^2 V_1}{\partial v^2} \right) dt \\ &\quad + \left(\frac{\partial V}{\partial S} - \Delta_1 \frac{\partial V_1}{\partial S} - \Delta \right) dS + \left(\frac{\partial V}{\partial v} - \Delta_1 \frac{\partial V_1}{\partial v} \right) dv. \end{aligned}$$

Note that if we let

$$\begin{aligned} \frac{\partial V}{\partial S} - \Delta_1 \frac{\partial V_1}{\partial S} - \Delta &= 0 \\ \frac{\partial V}{\partial v} - \Delta_1 \frac{\partial V_1}{\partial v} &= 0 \end{aligned} \tag{5.2}$$

we get a risk-free portfolio. In order to eliminate arbitrage opportunities, the return of this risk-free portfolio must equal the (deterministic) risk-free rate of return r:

$$\begin{aligned} d\Pi &= \left(\frac{\partial V}{\partial t}+\frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2}+\rho\eta Sv\frac{\partial^2 V}{\partial S\partial v}+\frac{1}{2}v\eta^2\frac{\partial^2 V}{\partial v^2}\right)dt \\ &\quad -\Delta_1\left(\frac{\partial V_1}{\partial t}+\frac{1}{2}vS^2\frac{\partial^2 V_1}{\partial S^2}+\rho\eta Sv\frac{\partial^2 V_1}{\partial S\partial v}+\frac{1}{2}v\eta^2\frac{\partial^2 V_1}{\partial v^2}\right)dt \\ &= r\Pi dt = r\left(V-\Delta S-\Delta_1 V_1\right)dt. \end{aligned} \tag{5.3}$$

Using (5.2) we can rewrite (5.3) to become:

$$\begin{aligned} &\frac{\frac{\partial V}{\partial t}+\frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2}+\rho\eta Sv\frac{\partial^2 V}{\partial S\partial v}+\frac{1}{2}v\eta^2\frac{\partial^2 V}{\partial v^2}+rS\frac{\partial V}{\partial S}-rV}{\frac{\partial V}{\partial v}} \\ = &\frac{\frac{\partial V_1}{\partial t}+\frac{1}{2}vS^2\frac{\partial^2 V_1}{\partial S^2}+\rho\eta Sv\frac{\partial^2 V_1}{\partial S\partial v}+\frac{1}{2}v\eta^2\frac{\partial^2 V_1}{\partial v^2}+rS\frac{\partial V_1}{\partial S}-rV_1}{\frac{\partial V_1}{\partial v}}. \end{aligned}$$

This equation leads to the conclusion that both left and right-hand sides are equal to some function g that does not make reference to any dependent variable, i.e. g only depends on the independent variables S, v and t. Setting $g=\lambda\left(v-\overline{v}\right)-\theta v$ yields a special case of a so-called affine diffusion process. For this class of processes, the pricing PDE is tractable analytically. In this case we hence have

$$\frac{\partial V}{\partial t}+\frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2}+\rho\eta Sv\frac{\partial^2 V}{\partial S\partial v}+\frac{1}{2}v\eta^2\frac{\partial^2 V}{\partial v^2}+rS\frac{\partial V}{\partial S}-rV=\left(\lambda\left(v-\overline{v}\right)-\theta v\right)\frac{\partial V}{\partial v},$$

defining $\lambda^{'}=\lambda-\theta$, $\lambda^{'}\overline{v}^{'}=\overline{\lambda v}$ gives

$$\frac{\partial V}{\partial t}+\frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2}+\rho\eta Sv\frac{\partial^2 V}{\partial S\partial v}+\frac{1}{2}v\eta^2\frac{\partial^2 V}{\partial v^2}+rS\frac{\partial V}{\partial S}-rV=\lambda^{'}\left(v-\overline{v}^{'}\right)\frac{\partial V}{\partial v}. \tag{5.4}$$

Subsequently, we will drop the primes on $\lambda^{'}$and $\overline{v}^{'}$(the so-called risk-adjusted parameters) for notational convenience. As a final remark, see Heath & Schweizer (2000) for a different choice of the function g.

5.3 The Characteristic Function

Heston (1993) was the first to derive the ch.f. in the Heston model. This was done in a Gil-Palaez inversion framework. Here we present the derivation in a Carr-Madan setting. To start with, we simplify the pricing PDE in (5.4) by defining the forward option price $V_u\left(x(t),v\left(t\right),t\right)$

$$V_u\left(x(t),v\left(t\right),t\right)\equiv e^{r(T-t)}V\left(S(t),v\left(t\right),t\right),$$

in which

$$x\left(t\right)=\log\left[\frac{e^{r(T-t)}S\left(t\right)}{K}\right].$$

Since

$$\begin{aligned}
\frac{\partial V}{\partial t} &= e^{-r(T-t)}\left[\frac{\partial V_u}{\partial x}\frac{\partial x}{\partial t}+\frac{\partial V_u}{\partial t}\right]+re^{-r(T-t)}V_u \\
&= e^{-r(T-t)}\left[-r\frac{\partial V_u}{\partial x}+\frac{\partial V_u}{\partial t}+rV_u\right] \\
\frac{\partial V}{\partial S} &= e^{-r(T-t)}\frac{\partial V_u}{\partial x}\frac{\partial x}{\partial S}=e^{-r(T-t)}\frac{1}{S}\frac{\partial V_u}{\partial x} \\
\frac{\partial^2 V}{\partial S^2} &= e^{-r(T-t)}\left[\frac{\partial^2 V_u}{\partial x^2}\left(\frac{\partial x}{\partial S}\right)^2+\frac{\partial V_u}{\partial x}\frac{\partial^2 x}{\partial S^2}\right] \\
&= e^{-r(T-t)}\frac{1}{S^2}\left[\frac{\partial^2 V_u}{\partial x^2}-\frac{\partial V_u}{\partial x}\right] \\
\frac{\partial^2 V}{\partial S\partial v} &= e^{-r(T-t)}\frac{1}{S}\frac{\partial^2 V_u}{\partial x\partial v} \\
\frac{\partial V}{\partial v} &= e^{-r(T-t)}\frac{\partial V_u}{\partial v} \\
\frac{\partial^2 V}{\partial v^2} &= e^{-r(T-t)}\frac{\partial^2 V_u}{\partial v^2},
\end{aligned}$$

equation (5.4) yields (after dividing by the discount factor):

$$\frac{\partial V_u}{\partial t}+\frac{1}{2}v\left[\frac{\partial^2 V_u}{\partial x^2}-\frac{\partial V_u}{\partial x}\right]+\rho\eta v\frac{\partial^2 V_u}{\partial x\partial v}+\frac{1}{2}v\eta^2\frac{\partial^2 V_u}{\partial v^2}=\lambda\left(v-\overline{v}\right)\frac{\partial V_u}{\partial v}.$$

Finally, define $\tau = T - t$, then we get the so-called forward equation

$$-\frac{\partial V_u}{\partial \tau}+\frac{1}{2}v\left[\frac{\partial^2 V_u}{\partial x^2}-\frac{\partial V_u}{\partial x}\right]+\rho\eta v\frac{\partial^2 V_u}{\partial x\partial v}+\frac{1}{2}v\eta^2\frac{\partial^2 V_u}{\partial v^2}=\lambda\left(v-\overline{v}\right)\frac{\partial V_u}{\partial v}. \tag{5.5}$$

Next we present three lemmas that culminate in the characteristic function of the Heston model. Before presenting the lemmas, we need to refer to advanced results from Duffie, Pan & Singleton (2000). They indicated that for affine diffusion processes the characteristic function of $x\left(T\right)$ (hence not of $\log S_T$) is of the form

$$f\left(x,v,\tau,\omega\right)=\exp\left(A\left(\omega,\tau\right)+B\left(\omega,\tau\right)v+C\left(\omega,\tau\right)x\right). \tag{5.6}$$

Here we have defined $v \equiv v\left(t\right)$ and $x \equiv x\left(t\right)$. Moreover, the characteristic function must satisfy the initial condition

$$f\left(x,v,0,\omega\right)=\exp\left(i\omega x\left(T\right)\right),$$

which in turn implies that $A\left(\omega,0\right)=B\left(\omega,0\right)=0$ and $C\left(\omega,0\right)=i\omega$ for all ω. Since $f\left(x,v,\tau,\omega\right)\equiv\mathbb{E}^{\mathbb{Q}}\left[e^{i\omega x(T)}\right]$, where $\mathbb{Q}$ is some risk-neutral measure, we can see $f\left(x,v,\tau,\omega\right)$ as an (imaginary) claim which must therefore satisfy (5.5). Substituting (5.6) in (5.5) and using the initial condition for $C\left(\omega,\tau\right)$ one can easily show that (5.6) simplifies to

$$f\left(x,v,\tau,\omega\right)=\exp\left(A\left(\omega,\tau\right)+B\left(\omega,\tau\right)v+i\omega x\right). \tag{5.7}$$

The solutions to $A\left(\omega,\tau\right)$ and $B\left(\omega,\tau\right)$ are provided by the next two lemmas. Note that we did not use $\frac{\partial}{\partial\tau}$ for the sake of notational convenience.

Lemma 1 *The functions $A(\omega,\tau)$ and $B(\omega,\tau)$ in (5.7) satisfy the following system of ordinary differential equations (ODEs):*

$$\begin{aligned}\frac{dA}{d\tau} &= aB,\ A(\omega,0)=0 &(5.8)\\ \frac{dB}{d\tau} &= \alpha-\beta B+\gamma B^2,\ B(\omega,0)=0 &(5.9)\end{aligned}$$

for $\omega\in\mathbb{R}$.

Proof. As indicated above, $f(x,v,\tau,\omega)$ satisfies (5.5). Substituting (5.7) into (5.5) yields

$$-f\left(\frac{\partial A}{\partial\tau}+\frac{\partial B}{\partial\tau}v\right)+\frac{1}{2}vf\left(-\omega^2-i\omega\right)+\rho\eta vBfi\omega+\frac{1}{2}\eta^2vB^2f-\lambda(v-\overline{v})Bf=0,$$

or

$$-\frac{\partial A}{\partial\tau}+\lambda\overline{v}B+v\left[-\frac{\partial B}{\partial\tau}-\frac{1}{2}\left(\omega^2+i\omega\right)+(\rho\eta i\omega-\lambda)B+\frac{1}{2}\eta^2B^2\right]=0. \tag{5.10}$$

Now define

$$\begin{aligned}a &= \lambda\overline{v}\\ \alpha &= -\frac{1}{2}\left(\omega^2+i\omega\right)\\ \beta &= \lambda-\rho\eta i\omega\\ \gamma &= \frac{1}{2}\eta^2,\end{aligned}$$

then (5.10) simplifies to

$$-\frac{\partial A}{\partial\tau}+aB+v\left[-\frac{\partial B}{\partial\tau}+\alpha-\beta B+\gamma B^2\right]=0,$$

which is a first order polynomial in v. In order for the equation to hold, both coefficients must vanish, i.e.

$$\frac{\partial A}{\partial\tau}=aB,$$

and

$$\frac{\partial B}{\partial\tau}=\alpha-\beta B+\gamma B^2.$$

Hence we have the following system of ODEs

$$\begin{aligned}\frac{dA}{d\tau} &= aB,\ A(\omega,0)=0\\ \frac{dB}{d\tau} &= \alpha-\beta B+\gamma B^2,\ B(\omega,0)=0,\end{aligned}$$

completing the proof. ■

Lemma 2 *The solution to the system of ODEs as specified by (5.8) and (5.9) is given by*

$$A(\omega,\tau)=\frac{\lambda\overline{v}}{\eta^2}\left[\tau(\lambda-\rho\eta i\omega-D)-2\log\left(\frac{1-Ge^{-D\tau}}{1-G}\right)\right] \tag{5.11}$$

$$B(\omega,\tau)=\frac{1}{\eta^2}\left[\frac{1-e^{-D\tau}}{1-Ge^{-D\tau}}\right](\lambda-\rho\eta i\omega-D), \tag{5.12}$$

where

$$D = \sqrt{\left(\lambda - \rho\eta i\omega\right)^2 + \left(\omega^2 + i\omega\right)\eta^2}$$
$$G = \frac{\lambda - \rho\eta i\omega - D}{\lambda - \rho\eta i\omega + D}.$$

Proof. First we solve for B:

$$\frac{dB}{d\tau} = \alpha - \beta B + \gamma B^2 = \gamma\left(B - r_+\right)\left(B - r_-\right),$$

with

$$r_\pm = \frac{\beta \pm D}{2\gamma},$$

where

$$D = \sqrt{\left(\beta^2 - 4\alpha\gamma\right)}.$$

Separating variables gives

$$\frac{1}{\gamma\left(B - r_+\right)\left(B - r_-\right)} dB = d\tau,$$

which is equivalent to

$$\left[\frac{1/\left(r_+ - r_-\right)}{\gamma\left(B - r_+\right)} - \frac{1/\left(r_+ - r_-\right)}{\gamma\left(B - r_-\right)}\right] dB = d\tau.$$

Integrating on both sides gives

$$\frac{\log\left(B - r_+\right)}{\gamma\left(r_+ - r_-\right)} - \frac{\log\left(B - r_-\right)}{\gamma\left(r_+ - r_-\right)} = \tau + c_B. \tag{5.13}$$

Since

$$\gamma\left(r_+ - r_-\right) = D$$

and $B\left(\omega, 0\right) = 0$, we have

$$\frac{\log\left(-r_+\right)}{D} - \frac{\log\left(-r_-\right)}{D} = c_B,$$

hence

$$e^{-c_B D} = \frac{r_-}{r_+} \equiv G.$$

Solving for B in (5.13) yields

$$B\left(\omega, \tau\right) = \left[\frac{1 - e^{-D\tau}}{1 - Ge^{-D\tau}}\right] r_-$$
$$= \frac{1}{\eta^2}\left[\frac{1 - e^{-D\tau}}{1 - Ge^{-D\tau}}\right]\left(\lambda - \rho\eta i\omega - D\right)$$

Now we are able to solve for $A\left(\omega, \tau\right)$:

$$A\left(\omega, \tau\right) = a\int B d\tau = a\left[r_-\tau - \frac{1}{\gamma}\log\left(1 - Ge^{-D\tau}\right)\right] + c_A.$$

From the initial condition $A\left(\omega, 0\right) = 0$, the constant of integration c_A follows:

$$c_A = \frac{a}{\gamma}\log\left(1 - G\right),$$

which leads to

$$\begin{aligned} A(\omega,\tau) &= a\left[\tau r_- - \frac{1}{\gamma}\log\left(\frac{1-Ge^{-D\tau}}{1-G}\right)\right] \\ &= \lambda\overline{v}\left[\tau\frac{\beta-D}{2\gamma} - \frac{1}{\gamma}\log\left(\frac{1-Ge^{-D\tau}}{1-G}\right)\right] \\ &= \lambda\overline{v}\left[\tau\frac{\beta-D}{\eta^2} - \frac{2}{\eta^2}\log\left(\frac{1-Ge^{-D\tau}}{1-G}\right)\right] \\ &= \frac{\lambda\overline{v}}{\eta^2}\left[\tau(\lambda-\rho\eta i\omega-D) - 2\log\left(\frac{1-Ge^{-D\tau}}{1-G}\right)\right], \end{aligned}$$

completing the proof. ■

Finally, we arrive at the desired result.

Lemma 3 *In the Heston model, the characteristic function $\phi_T(\omega)$ of $\log S_T$ is given by*

$$\begin{aligned} \phi_T(\omega) &= \exp(i\omega[\log[S_0] + rT]) \\ &\quad \cdot\exp\left(\frac{v_0}{\eta^2}\left[\frac{1-e^{-DT}}{1-Ge^{-DT}}\right](\lambda-\rho\eta i\omega-D)\right) \\ &\quad \cdot\exp\left(\frac{\lambda\overline{v}}{\eta^2}\left[T(\lambda-\rho\eta i\omega-D) - 2\log\left(\frac{1-Ge^{-DT}}{1-G}\right)\right]\right), \end{aligned}$$

where D and G are as defined in the previous lemma.

Proof. Substituting (5.11) and (5.12) in (5.7) and defining $S(t) = S$ yields

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}\left[e^{i\omega x(T)}\right] &= \exp(i\omega x) \\ &\quad \cdot\exp\left(\frac{v}{\eta^2}\left[\frac{1-e^{-D\tau}}{1-Ge^{-D\tau}}\right](\lambda-\rho\eta i\omega-D)\right) \\ &\quad \cdot\exp\left(\frac{\lambda\overline{v}}{\eta^2}\left[\tau(\lambda-\rho\eta i\omega-D) - 2\log\left(\frac{1-Ge^{-D\tau}}{1-G}\right)\right]\right), \end{aligned}$$

in which

$$\begin{aligned} \exp(i\omega x) &= \exp\left(i\omega\left[\log\left(\frac{S\exp(r\tau)}{K}\right)\right]\right) \\ &= \exp(i\omega\log[S])\exp(i\omega r\tau)\exp\left(i\omega\log\left[\frac{1}{K}\right]\right). \end{aligned}$$

Note that we are interested in $\mathbb{E}^{\mathbb{Q}}\left[e^{i\omega\log[S(T)]}\right]$. Substituting $t = T$ in the previous equation and taking expectations gives

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}\left[e^{i\omega x(T)}\right] &= \mathbb{E}^{\mathbb{Q}}\left[\exp(i\omega\log S_T)\exp\left(i\omega\log\left[\frac{1}{K}\right]\right)\right] \\ &= \exp\left(i\omega\log\left[\frac{1}{K}\right]\right)\mathbb{E}^{\mathbb{Q}}\left[\exp(i\omega\log S_T)\right], \end{aligned}$$

hence

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}\left[e^{i\omega\log S_T}\right] &= \exp\left(-i\omega\log\left[\frac{1}{K}\right]\right)\mathbb{E}^{\mathbb{Q}}\left[e^{i\omega x(T)}\right] \\ &= \exp\left(i\omega\left[\log S + r\tau\right]\right) \\ &\quad\cdot\exp\left(\frac{v}{\eta^2}\left[\frac{1-e^{-D\tau}}{1-Ge^{-D\tau}}\right](\lambda-\rho\eta i\omega-D)\right) \\ &\quad\cdot\exp\left(\frac{\lambda\overline{v}}{\eta^2}\left[\tau(\lambda-\rho\eta i\omega-D)-2\log\left(\frac{1-Ge^{-D\tau}}{1-G}\right)\right]\right),\end{aligned}$$

yielding the stated. ■

5.4 When is $\mathbb{E}S_T^{\alpha+1} < \infty$?

In the previous section we have derived that in a Heston world the option price of a vanilla call at $t = 0$ is given by

$$C_T(k) = \frac{e^{-\alpha k}}{\pi}\,\mathrm{Re}\left\{\int_0^\infty e^{-ivk}\frac{e^{-rT}\phi_T\left(v-(\alpha+1)i\right)}{\alpha^2+\alpha-v^2+i\left(2\alpha+1\right)v}dv\right\},$$

in which the characteristic function is given by

$$\begin{aligned}\phi_T(\omega) &= \exp\left(i\omega\left[\log\left[S_0\right]+rT\right]\right) \\ &\quad\cdot\exp\left(\frac{v_0}{\eta^2}\left[\frac{1-e^{-DT}}{1-Ge^{-DT}}\right](\lambda-\rho\eta i\omega-D)\right) \\ &\quad\cdot\exp\left(\frac{\lambda\overline{v}}{\eta^2}\left[T(\lambda-\rho\eta i\omega-D)-2\log\left(\frac{1-Ge^{-DT}}{1-G}\right)\right]\right),\end{aligned}$$

and where

$$\begin{aligned}D &= \sqrt{\left(\lambda-\rho\eta i\omega\right)^2+\left(\omega^2+i\omega\right)\eta^2} \\ G &= \frac{\lambda-\rho\eta i\omega-D}{\lambda-\rho\eta i\omega+D}.\end{aligned} \tag{5.14}$$

Here we have defined $v(0) = v_0$ and $S(0) = S_0$. From Section 3.3 we know that we must have $\mathbb{E}S_T^{\alpha+1} < \infty$ in order to be able to use Carr-Madan inversion at all. This section shows what values of α are allowed. To do so, we have to enter the fascinating realm of complex analysis. In fact, basic knowledge of common analytic functions is enough. Readers deficient in complex analysis can for example consult Ahlfors (1979), Knopp (1945) or Knopp (1947).

To begin with, recall that

$$\phi_T\left(-(\alpha+1)i\right) = \mathbb{E}S_T^{\alpha+1},$$

where $\phi_T(\cdot)$ is the ch.f. of $\log S_T$. By inspection, we see that the Heston ch.f. is defined (and hence finite) in $-(\alpha+1)i$ provided that this point is not on the branch cut of the complex logarithm. (Let $z = re^{i\theta} \in \mathbb{C}$, then the complex logarithm $\log(z)$ is defined as

$$\log(z) = \log|z| + i\left(\theta+2k\pi\right) = \log|z| + i\arg(z),$$

where $k \in \mathbb{Z}$. First we have to make sure that $\log(z)$ defines a function, i.e. it must not be multivalued. This can be accomplished by imposing the condition $-\pi \le \arg(z) < \pi$. However,

the complex logarithm obtained in this way is not analytic since it is not continuous. One can easily verify that by imposing the more restrictive condition $-\pi < \arg(z) < \pi$, or equivalently $z \notin (-\infty, 0]$ (this interval is a so-called branch cut), an analytic function is obtained. Alternatively, $-\pi + 2k\pi < \arg(z) < \pi + 2k\pi$ for some $k \in \mathbb{Z}$ can be chosen as well. We will stick with standard practice however. The single-valued complex logarithm is sometimes denoted as $\mathrm{Log}(z) : \mathbb{C}\backslash(-\infty, 0] \to \mathbb{C}$ with $\mathrm{Log}(z) = \log|z| + i\mathrm{Arg}(z)$ and $-\pi < \mathrm{Arg}(z) < \pi$. We will not pursue this notation here. Just keep in mind that in the subsequent, $\log(z)$ denotes the above referred to analytic function. With this definition, one can easily verify that identities such as $\log(z_1) - \log(z_2) = \log\left(\frac{z_1}{z_2}\right)$ continue to hold for $z_1, z_2 \in \mathbb{C}$.)

So let us examine the term $\log\left(\frac{1-Ge^{-DT}}{1-G}\right) = \log\left(1 - Ge^{-DT}\right) - \log(1-G)$. Substituting $-(\alpha+1)i$ in $\log\left(1 - Ge^{-DT}\right)$ and requiring that this point is not on the branch cut $(-\infty, 0]$ yields the condition:

$$1 - G(-(\alpha+1)i)\, e^{-D(-(\alpha+1)i)T} > 0,$$

or equivalently,

$$G(-(\alpha+1)i)\, e^{-D(-(\alpha+1)i)T} < 1, \tag{5.15}$$

where we used the fact that $1 - Ge^{-DT}$ is real on the imaginary axis. Naturally $T > 0$, and by inspection we have $G(-(\alpha+1)i) \leq 1$, implying that the condition (5.15) is satisfied if

$$D(-(\alpha+1)i) > 0.$$

Now, define $b \equiv -(\alpha+1)$. From (5.14) we hence get the condition

$$\eta^2\left(\rho^2 - 1\right)b^2 + \eta\left(2\lambda\rho - \eta\right)b + \lambda^2 > 0,$$

the left-hand-side being a concave quadratic function in b. This inequality is satisfied whenever

$$\frac{1}{2}\frac{-2\lambda\rho + \eta - \sqrt{\eta^2 - 4\lambda\rho\eta + 4\lambda^2}}{(\rho^2-1)\,\eta} < b < \frac{1}{2}\frac{-2\lambda\rho + \eta + \sqrt{\eta^2 - 4\lambda\rho\eta + 4\lambda^2}}{(\rho^2-1)\,\eta},$$

or equivalently,

$$-\frac{1}{2}\frac{-2\lambda\rho + \eta + \sqrt{\eta^2 - 4\lambda\rho\eta + 4\lambda^2}}{(\rho^2-1)\,\eta} - 1 < \alpha < -\frac{1}{2}\frac{-2\lambda\rho + \eta - \sqrt{\eta^2 - 4\lambda\rho\eta + 4\lambda^2}}{(\rho^2-1)\,\eta} - 1 \tag{5.16}$$

For the $\log(1-G)$ term we get exactly the same range for α since in this case we must impose

$$G(-(\alpha+1)i) < 1.$$

This is true if

$$D(-(\alpha+1)i) > 0,$$

which is exactly the condition that appeared when we were handling the $\log\left(1 - Ge^{-DT}\right)$ term!

The last term to be considered is the square-root since it is defined in terms of the complex logarithm:

$$\begin{aligned} D &= \sqrt{(\lambda - \rho\eta i\omega)^2 + (\omega^2 + i\omega)\,\eta^2} \\ &= e^{\frac{1}{2}\log\left((\lambda-\rho\eta i\omega)^2 + \left(\omega^2+i\omega\right)\eta^2\right)}, \end{aligned}$$

Along the same lines as above, one can easily show that we again arrive at condition (5.16).

5.5 The Heston Integrand

In this section we will present a pictorial survey of the Heston integrand

$$\mathrm{Re}\left\{e^{-ivk}\frac{e^{-rT}\phi_T\left(v-(\alpha+1)i\right)}{\alpha^2+\alpha-v^2+i\left(2\alpha+1\right)v}\right\},$$

where ϕ_T is the Heston ch.f. More specifically, we will give a hint what parameter values might give rise to difficulties when it comes to numerical integration. Unless stated otherwise, the parameters used to generate the pictures in this section are $v_0 = 0.15^2, T = 2, S = 100, r = 0.05, \alpha = 0.75, K = 100, \eta = 0.3, \lambda = 1, \overline{v} = 0.2^2, \rho = -0.3$. Broadly speaking, the Heston integrand behaves much like the Black-Scholes integrand does, in the sense that it is the same parameters that cause difficulties to numerical integration.

Low times to maturity result in a fluctuating Heston integrand. The same is true when T is relatively large. Several integrands have been plotted in Figure 5.2. Along the horizontal axes we have set out v. This holds for all figures in this section.

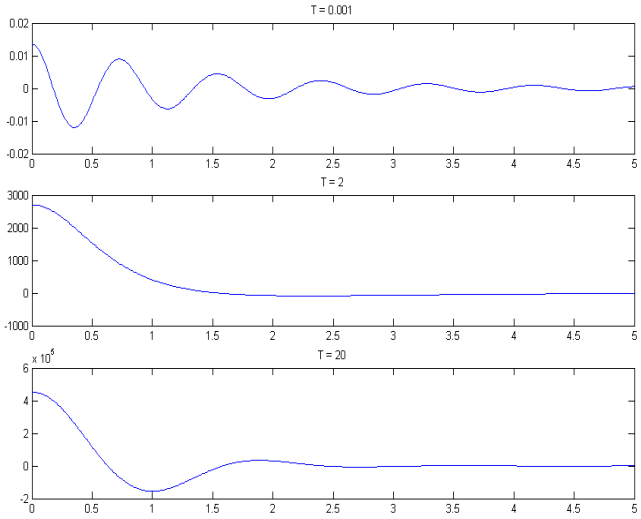


Figure 5.2: Both small and large values for T result in fluctuations.

Options for which the quotient S/K is much larger than one also result in cumbersome integrands. The moneyness being much smaller than one poses far less difficulties. This is illustrated in Figure 5.3.

As expected, choosing α too small makes the integrand resemble an impulse function. Setting α relatively large, magnifies the function value in $v = 0$. Figure 5.4 summarizes these observations.

Finally, we remark that Heston-specific parameters do not influence the behavior of the integrand in a remarkable way.

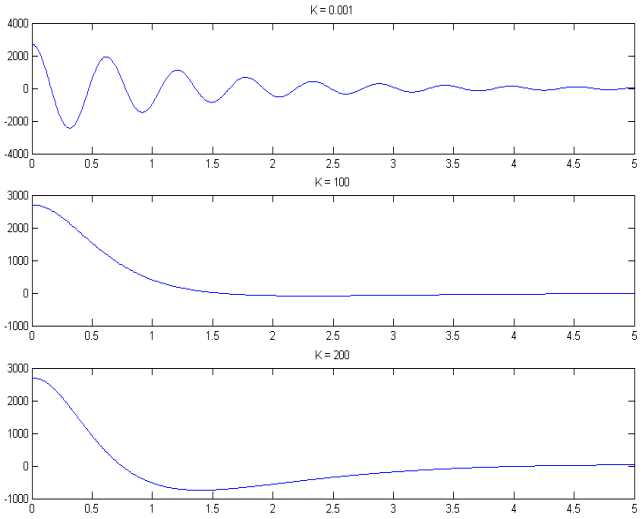


Figure 5.3: Far in-the-money and out-the-money options have oscillatory integrands.

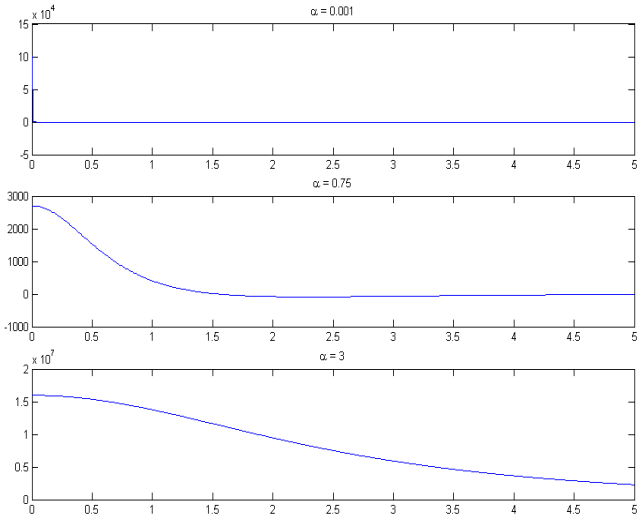


Figure 5.4: The influence α has on the nature of the integrand.

5.6 Numerical results

5.6.1 On the choice of α

Discouraged by the effectiveness of the minmax-strategy for choosing α in the Black-Scholes framework, we won't consider it here (as noted in the previous section, there is similar behavior between the Heston and Black-Scholes integrands). Here we again consider the 'magic' value $\alpha = 0.75$. Furthermore, we scrutinize the suggestion made by Carr & Madan (1999). Recall that they suggest to use the condition

$$\mathbb{E}S_T^{\alpha+1} < \infty \tag{5.17}$$

to determine a suitable value for α. One fourth of the maximum value of α such that (5.17) holds is allegedly a good choice. We will henceforth denote this value by α_{CM}. In fact, this suggestion has most probably not been made with the Heston model in mind, however, let us examine its effectiveness nonetheless. From Section 5.4 we know that

$$-\frac{1}{2}\frac{-2\lambda\rho+\eta+\sqrt{\eta^2-4\lambda\rho\eta+4\lambda^2}}{(\rho^2-1)\,\eta}-1<\alpha<-\frac{1}{2}\frac{-2\lambda\rho+\eta-\sqrt{\eta^2-4\lambda\rho\eta+4\lambda^2}}{(\rho^2-1)\,\eta}-1. \tag{5.18}$$

Obviously, one fourth of the right-most term in (5.18) is a suitable candidate to be α_{CM}. However, one example is enough to demonstrate that α_{CM} does not necessarily yield accurate results in the Heston framework. To see this, let $r = 0.1, T = 3, v_0 = 0.03, S = 100, \lambda = 1, \overline{v} = 0.04, \eta = 0.4, \rho = -0.6$. Carr-Madan inversion with

$$\Delta k = 0.025,\ \Delta v = \frac{2\pi}{N\Delta k},$$

where $N = 2^{11} = 2048$ then yields absolute relative errors $|C_{CM} - C_H|/S$ of order 10^{-2}; here C_H denotes the Heston price. Note that for the aforementioned parameters, $\mathbb{E}S_T^{\alpha+1}$ is indeed finite (it is equal to 194.4). Moreover, we have $\alpha_{CM} = 0.07$. On the other hand, selecting α to be 0.75 gives relative errors of order 10^{-9}. The situation is depicted in Figure 5.5. In the upper part we have plotted the strike on the horizontal axis against the relative error for $\alpha = \alpha_{CM} = 0.07$ on the vertical axis. The lower figure shows results for $\alpha = 0.75$. Similar differences in accuracy can be observed for other parameter settings.

5.6.2 More Results

This section presents numerical results for the Heston model. The grids we have used are given by

$$\Delta k = 0.025,\ \Delta v = \frac{2\pi}{N\Delta k},$$

where $N = 2^{11} = 2048$. Moreover, we haven chosen $\alpha = 0.75$. For these settings, the truncation error is typically negligible (In fact, Lee (2004) presents analytical error bounds for the Heston model. However, it is based on a different formulation of the DFT). First, we have performed the following experiment. We randomly selected parameters from the ranges

$$\begin{aligned} v_0 &\in \left[0.15^2, 0.6^2\right],\ T \in [0.08, 20],\ S = [0, 100],\ r \in [0, 0.15], \\ \lambda &\in [0, 100], \overline{v} \in \left[0.15^2, 0.6^2\right], \eta \in [0.15, 0.6], \rho \in [-1, 1]. \end{aligned}$$

The maximum relative error $|C_{CM} - C_H|/S$ is then determined over the range $\frac{1}{2}S \le K \le 2S$. Figure 5.6 shows 50 realizations of this experiment.

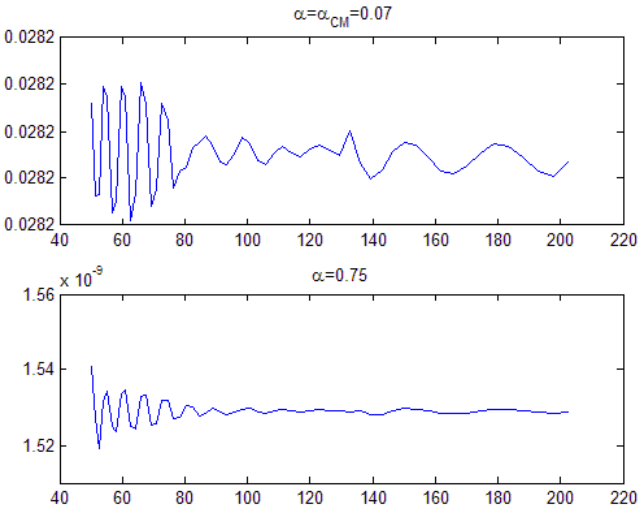


Figure 5.5: Upper: relative errors for $\alpha = \alpha_{CM} = 0.07$. Lower: $\alpha = 0.75$.

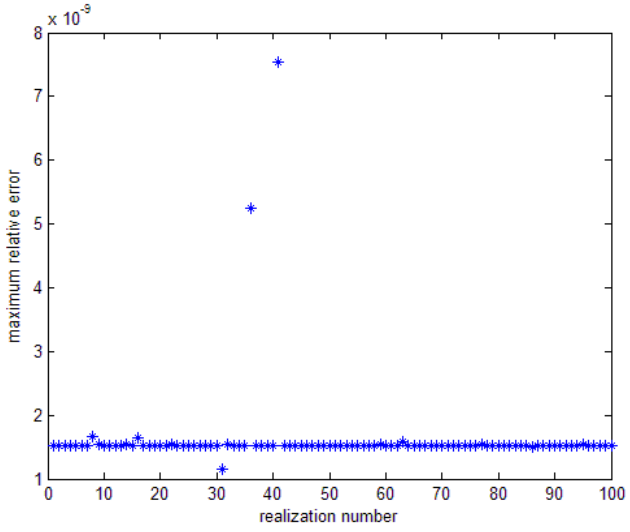


Figure 5.6: Maximum relative errors for 100 realizations.

As mentioned in Section 5.5, far in-the-money-options are difficult to price via Carr-Madan inversion. In fact, for the above grid settings we typically get relative errors of order 1 for small strikes (10^{-2}). When we double the number of points to $N = 2^{12}$ relative errors reduce to 10^{-5}. In case we have a small value for T (10^{-3}), relative errors will be relatively large as well (10^{-5} for $N = 2^{11}$). Again, increasing N is of help. Things are even worse when we both have a small value for T (10^{-3}) and $K(10^{-3})$: relative errors of order 10^{-2} for $N = 2^{11}$. However, these cases are more of theoretical than of practical interest.

Like in the Black-Scholes model, changes of variable can also be applied to the Heston integrand. For practical purposes, direct integrations will then pop up as a true competitor to Carr-Madan inversion (recall that it is the interpolation step that slows down the process). Finally, we remark that cubic spline interpolation again reduces the relative error with one order of magnitude.

6 Calibration of the Heston Model

6.1 Introduction

An important application of Carr-Madan inversion lies in the calibration of stock models, since option prices can be calculated rapidly. One possible approach is based on market prices for European options. Then we assume that a certain parametric stock model governs the dynamics of the underlying asset and we try to minimize - for example - the differences between option prices from this stock model and those implied by the market, yielding a set of parameters for the stock model. Obviously, the calibrated stock model can then be used to price more exotic options as has been done in Schoutens et al. (2004).

In general, we can state the calibration problem as

$$\min_{\theta} f\left(C_{CM}\left(\theta\right), C_{market}\right),$$

where C_{CM} denotes an n-dimensional vector of option prices generated by Carr-Madan inversion. Typically, it contains prices for various strikes and times to expiry. Surely, it depends on a set of parameters θ, the parameters of the underlying stock model. Its market counterpart is given by the n-dimensional market-prices-vector C_{market} , i.e. the i-th entry in the vector C_{market} gives the market price for the option with strike and time to maturity as used in the calculation of the i-th entry in the vector $C_{CM}\left(\theta\right)$. Finally, $f : \mathbb{R}^n \to \mathbb{R}$ denotes the objective function which is to be minimized with the aid of some optimization algorithm.

In this chapter the Heston stochastic volatility model is calibrated to market data. As noted in Section 5.1, stochastic volatility models are not perfect. As a consequence, instead of working very hard to obtain perfect fits to market data - if possible at all - we will put limits on the computational resources we want to spend on the calibration of this model. This will be a 60 seconds time limit. Based on market data of ING options we then calibrate the Heston model. Specifically, we will consider the efficacy of both random and deterministic algorithms. The deterministic algorithms used come from MATLAB. Also, we will examine several types of objective functions, that will in some sense be unified by translating them to a global measure involving volatilities. Our main focus will be on numerical results throughout. Therefore, theoretical issues concerning the algorithms will not be discussed. The random algorithms are extensively discussed in Spall (2003); information on the deterministic algorithms can be found in Bazarraa, Sherali & Shetty (1993) and MATLAB's help section.

The outline of this chapter is as follows. First we examine what types of smile the Heston model is able to produce. Second, we briefly consider the algorithms and objective functions used in the calibration process. References will be given where appropriate. Third, a systematic calibration of the Heston model is presented, in which we compare the various algorithms and objective functions where possible. Concluding remarks end this section.

6.2 The way Heston smiles

Option prices are not quoted directly. Rather, market practice is to quote (Black-Scholes) implied volatilities rather than option prices themselves. One can obtain the associated market prices by simply substituting the quoted volatilities in the Black-Scholes formula. (Once again, we see how entrenched the Black-Scholes model is in the financial world.) In this section we will examine what smile curves can arise in a Heston environment. This knowledge will enable us to make statements about the suitability of the Heston model for the market data at hand. Along this journey we will also reveal the specific effects each Heston parameter has on the smile curve.

Throughout this section, we used

$$r = 0.15, T = 1, S = 22.1, \alpha = 0.75,$$

unless stated otherwise.

We start with the influence the speed of mean reversion λ has on the smile curve. As illustrated in Figure 6.1, increasing the value for λ flattens the curve. In the picture presented we see that the curve is lifted. The opposite movement can also be achieved, by for example letting $T = 10$.

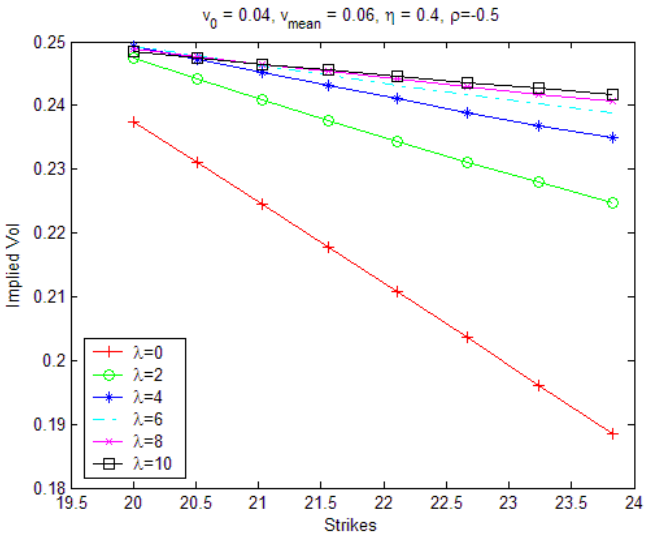


Figure 6.1: The speed of mean reversion λ rotates and shifts the smile curve.

Increasing the initial variance v_0 of $\log S_T$ moves the curve upward. Unlike the speed of mean reversion, there is no dependence on T whatsoever. This is illustrated in Figure 6.2.

The mean of the variance $\overline{v}$ has a similar effect as v_0: the curve is being lifted with increasing $\overline{v}$ (Figure 6.3). However, now we have a dependence on T. The larger this value the flatter the curves become as can be seen in Figure 6.4.

A clockwise rotation of the smile can be achieved by increasing the value of η. When T is relatively small this effect is most striking. Longer times to maturity will render this effect less visible. In this latter case the curve is flattened. These observations are pictorially summarized in Figures 6.5 and 6.6.

Despite that the smile curve rotates, the resulting slope did not change sign however. The last parameter to be considered, the correlation coefficient, makes this transition possible. See Figure 6.7. Once again a large value for T reduces this effect, as illustrated in Figure 6.8.

The above figures suggest that smiles in the Heston model are - roughly speaking - linear. In fact, smile-like curves can be obtained as well. This can be accomplished by certain combinations of parameters. Figure 6.9 presents a smile obtained by letting λ to be small and η to be large. This shape has also been called 'hockeystick' for obvious reasons.

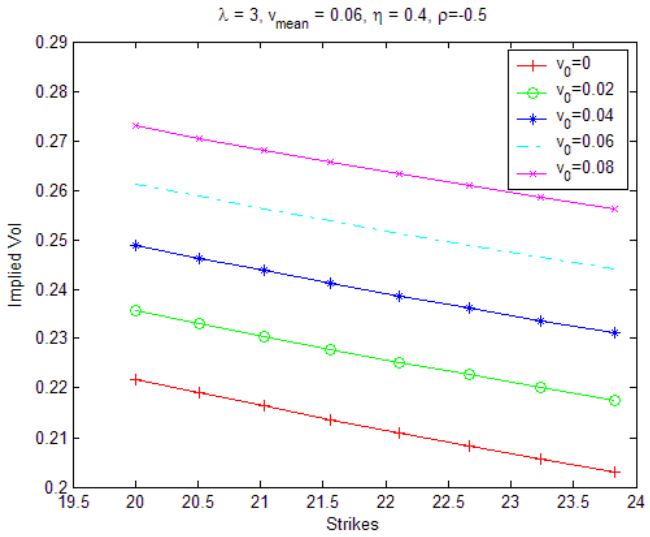


Figure 6.2: Increasing v_0 moves the curve upward.

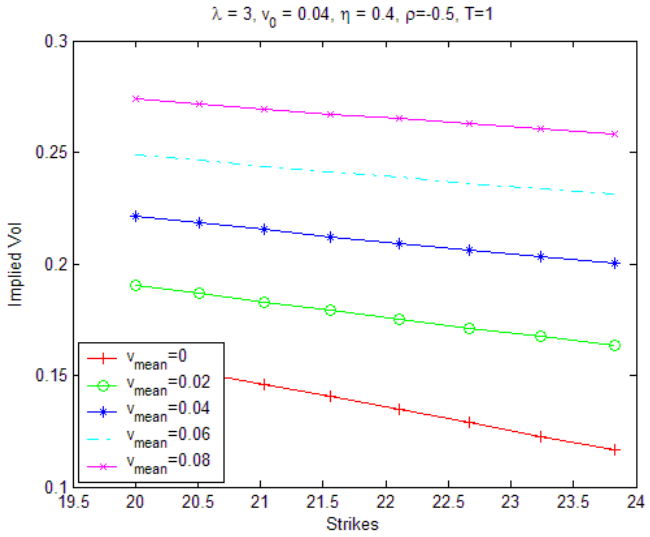


Figure 6.3: Increasing $\overline{v}$ lifts the smile curve.

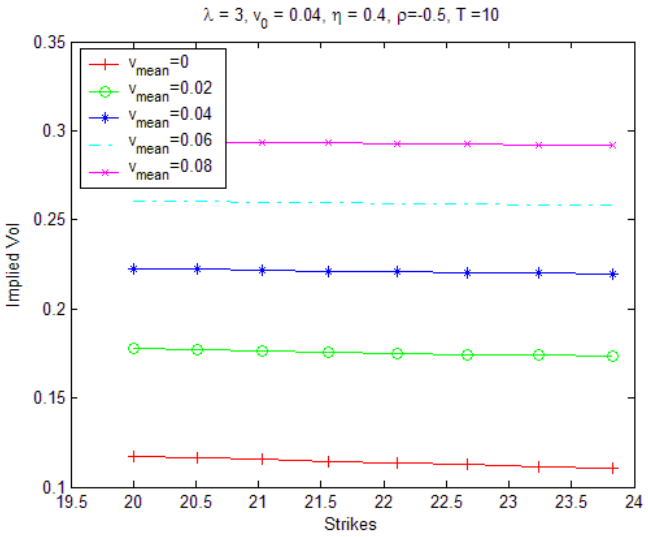


Figure 6.4: We have rather flat curves when T is relatively large.

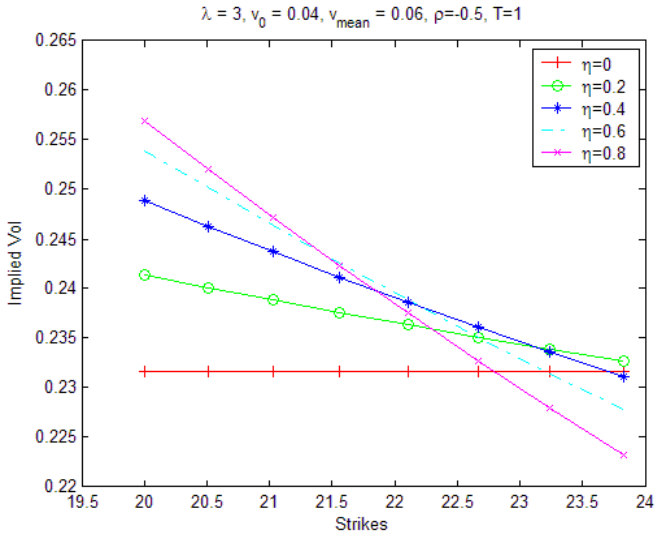


Figure 6.5: A clockwise rotation can be achieved by increasing the volatility of volatility η.

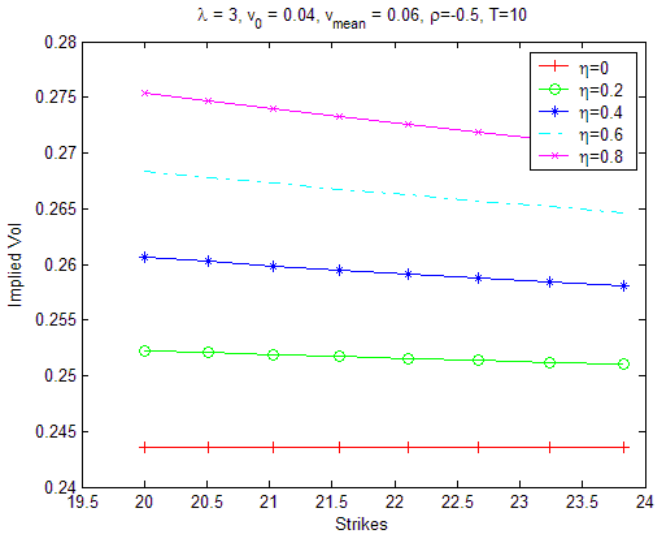


Figure 6.6: The clockwise rotation is being masked by a large value for T.

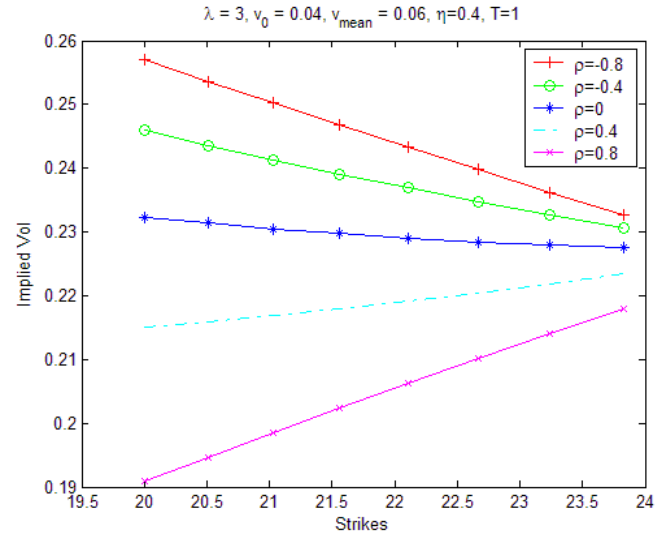


Figure 6.7: The correlation coefficient makes a change of slope possible.

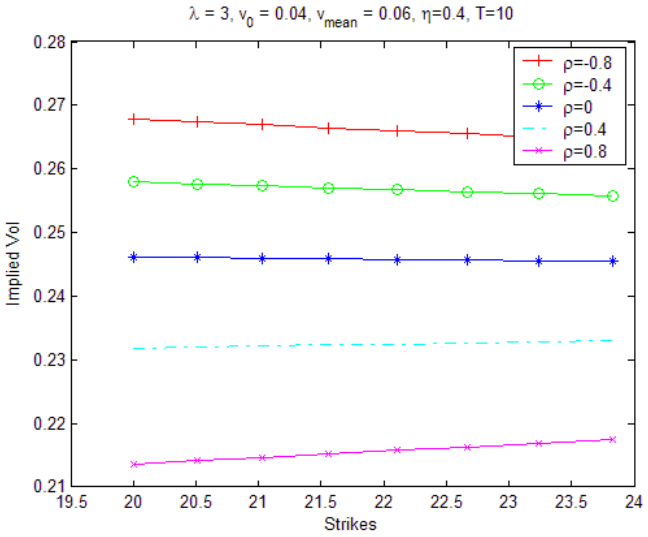


Figure 6.8: A large value for T flattens the smile curves considerably.

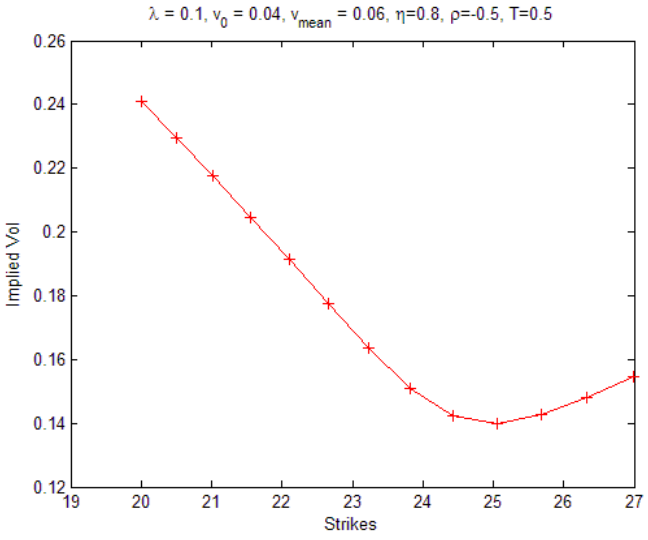


Figure 6.9: A smiling Heston

6.3 Algorithms and Objective Functions

6.3.1 Algorithms

The algorithms considered can be categorized as either being random or deterministic. Random algorithms determine subsequent search directions according to a mechanism involving randomness, while their deterministic counterparts do this in a non-random fashion. Among the advantages of stochastic algorithms are robustness and ease of implementation. These nice properties mainly derive from the fact that no information on the derivative is needed in order to run the algorithms. In fact, derivatives need not even to exist. Simple function evaluations are sufficient, no matter how cumbersome the objective function is. On the other hand, there is the demanding problem to prove convergence of these algorithms. As indicated in Spall (2003), (asymptotic) results on convergence can oftentimes only be proven for very restricted cases, if they exist at all. Hence we will not make any incursions into this territory. Convergence is simply assumed. Moreover, cumbersome constraints cannot straightforwardly be dealt with by stochastic algorithms. Finally, they present the user an extra burden in the sense that these algorithms have degrees of freedom the user has to specify himself. Henceforth we will refer to this as subjectivity. We will consider two random algorithms: localized random search and simulated annealing.

Localized random search is based on the idea that smaller objective function values might be found in the neighborhood of the point with smallest function value found so far. Let θ_k denotes the set of parameters at iteration k, pseudo-code for this algorithm then reads:

Localized Random Search Algorithm

step 0 Set $k = 0$. Choose an initial feasible θ_0. Evaluate the objective value $f(\theta_0)$.

step 1 Create a random vector $\Delta\theta_k$. If $\theta_k + \Delta\theta_k$ is not feasible, generate a new $\Delta\theta_k$. If it is feasible, define $\theta_{new}(k+1) \equiv \theta_k + \Delta\theta_k$ and evaluate $f(\theta_{new}(k+1))$.

step 2 If $f(\theta_{new}(k+1)) < f(\theta_k + \Delta\theta_k)$, set $\theta_{k+1} = \theta_{new}(k+1)$, else set $\theta_{k+1} = \theta_k$.

step 3 Stop if some stopping criteria is met, else set $k = k + 1$ and go to step 1.

Subjectivity is introduced in step 1 of the algorithm: Apart from the fact that we have to choose a distribution for $\Delta\theta_k$ - something that would correspond to for example choosing steepest descent in a gradient based algorithm, we sometimes also have to specify ranges for the distribution, as in our case. Indeed, we have let $\Delta\theta_k$ to be uniformly distributed. But what range should the i-th component of this vector have? Although numerical experiments are of help when answering this question, subjectivity remains. We followed recommendations made in Spall (2003). Our MATLAB-code is available upon request.

Even more subjective is simulated annealing. This algorithm is the mathematical analogue of the cooling of substances while they (may) reach a (desirable) state of minimum energy. Reaching this optimal state is dependent on a sufficiently slow cooling process. The following Boltzmann-Gibbs probability distribution from statistical mechanics plays an important role in it:

$$P\{Energy\ State = x\} = c_T \exp\left(-\frac{x}{c_B T}\right). \tag{6.1}$$

Here $c_T > 0$ is a normalizing constant, $c_B > 0$ denotes the Boltzmann constant and T is the temperature of the system. Note that at higher temperatures the system is more likely to be in higher energy states. The way this idea returns in the algorithm is that while running the

algorithm (i.e. while T is decreasing), the objective function value might increase with a certain probability (i.e. higher energy states might be reached), hereby having the potential of getting out a local minimum. The random mechanism that determines whether a higher energy state is chosen is given by the following. Let $e_{current}$ and e_{new} denote the current energy state and the new energy state, respectively. If $e_{new} < e_{current}$, then the system moves to e_{new}. However, if $e_{new} \geq e_{current}$, then the system moves to e_{new} with probability

$$\exp\left(-\frac{e_{new} - e_{current}}{c_B T}\right),$$

which is known as the Metropolis criterion. One can show that under certain assumptions we in fact have a Markov chain and that the limiting distribution is given by (6.1). Having outlined the underlying ideas we now give the pseudo-code:

Simulated Annealing Algorithm

step 0 Choose an initial temperature T and a feasible $\theta_{current}$. Evaluate value $f(\theta_{current})$.

step 1 Relative to $\theta_{current}$, determine randomly a new θ_{new}. Evaluate $f(\theta_{new})$.

step 2 If $f(\theta_{new}) < f(\theta_{current})$, then $\theta_{current} = \theta_{new}$. Go to step 3.
If $f(\theta_{new}) \geq f(\theta_{current})$, generate a number U on $[0,1]$ that is uniformly distributed.
If $U \leq \exp\left(-\frac{f(\theta_{new}) - f(\theta_{current})}{c_B T}\right)$, then $\theta_{current} = \theta_{new}$, else $\theta_{current}$ remains.

step 3 Repeat steps 1 and 2 until the user-specified number of iterations for T is met.

step 4 Lower T and go to step 1, or when some stopping criterion is met.

Among the most subjective issues that arise we have: Choosing the initial temperature, the determination of θ_{new} in step 1, the number of iterations per T and the way temperature decays. Spall (2003) discusses and gives references to various methods as available in literature.

As for the deterministic algorithms, we will resort to MATLAB. For more information on the functions **fminsearch** (in fact, this algorithm originates from stochastic search. See Spall (2003)), **fmincon** and **lsqnonlin**, please consult MATLAB's help section. A well-known problem with these algorithms is the specification of a good starting point from which the searching process starts. Bad starting points might slow down the optimization process tremendously.

6.3.2 Objective Functions

To set the stage, note that in the Heston model the set of model parameters is given by $\theta = \{\lambda, \rho, \eta, \overline{v}, v_0\}$. Furthermore, let y^i denote the i-th element of the vector y. Following Schoutens et al. (2004), we consider the objective functions RMSE, MSE and AAE. In addition, we will also examine the maximum relative absolute error (MARE):

1. Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{\left[C^i_{CM}(\theta) - C^i_{market}\right]^2}{n}}$$

2. Mean Square Error (MSE):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left[C^i_{CM}(\theta) - C^i_{market}\right]^2$$

3. Average Absolute Error (AAE):

$$AAE = \frac{1}{n}\sum_{i=1}^{n}\left|C^i_{CM}(\theta) - C^i_{market}\right| \tag{6.2}$$

4. Maximum Absolute Relative Error (MARE)

$$MARE = \max_i \frac{|C_{CM}(\theta) - C_{market}|}{C_{market}}$$

In order to give one meaning to these various objective functions, we will translate all objective function values obtained to a global measure VWAEV (Vega Weighted Absolute Error in the Volatilities):

$$VWAEV = \left(\frac{1}{V}\sum_{i=1}^{n}\frac{\partial C^i_{market}}{\partial \sigma^i_{market}}\left|\sigma^i_{CM} - \sigma^i_{market}\right|\right)\cdot 100, \tag{6.3}$$

where

$$V = \sum_i \frac{\partial C^i_{market}}{\partial \sigma^i_{market}}.$$

Furthermore, σ_{CM} denotes the implied volatilities that correspond to $C_{CM}\left(\theta'\right)$, θ' being the set of calibrated parameters. The vector σ_{market} gives the implied volatilities as determined by the market. Note that there is a first order relation between (6.2) and (6.3): To see this, do the following Taylor expansion:

$$C_{BS}(\sigma_{CM}) = C_{BS}(\sigma_{market}) + \frac{\partial C_{BS}}{\partial \sigma}(\sigma_{CM} - \sigma_{market}) + ...,$$

which gives

$$C_{BS}(\sigma_{CM}) - C_{BS}(\sigma_{market}) \approx \frac{\partial C_{BS}}{\partial \sigma}(\sigma_{CM} - \sigma_{market}).$$

6.4 Calibrating the Heston World

6.4.1 Numerical Results

In the introduction of this chapter we have mentioned that we will put restrictions on the computational effort we want to spend. A reasonable restriction is 60 seconds on a Pentium 4, 2.80GHz PC with 1024 MB internal memory. The reported objective function values are all median values corresponding to three repetitions of the same optimization algorithm. Each run however will be initiated with another (randomly chosen) starting point. (Note that we do not perform any so-called regularization, which would say guarantee uniqueness of the optimal point. This would be far outside the scope of this thesis. The results obtained do provide a flavor of the effectiveness of the Heston model however. Regularization is for example discussed in Jackson, Suli & Howison (1998).) Note that we do not calibrate the model per value of T, i.e. the vector C_{market} will in general contain market prices for various values of T. The reason to do

this is simple: Suppose we calibrate the model to vanilla calls with $T = 1$ month. What if we subsequently want to price options having $T = 2$ months? The inclusion of various values for T remedies this problem somewhat. In general, we can say that we calibrate 'to what we need'.

We start with the random algorithms. Tables 6.1 through 6.3 compare localized random search with simulated annealing when we calibrate the Heston model using 70 ING calls. The spot price of the underlying stock was 22.1 Euros and strikes vary from 11.05 to 44.20, while the times to maturity range from 1 month to 10 years. The market implied volatilities are presented in Appendix A. The corresponding Heston parameters $\lambda, v_0, \overline{v}, \eta, \rho$ are also recorded below. We see that in this case localized random search tends to be more accurate within the 60 seconds allotted. Note the substantial differences when it comes to the resulting parameters. In terms of the global measure $VWAEV$ we get the results as shown in Table 6.4. We see that the application of the localized random search algorithm to AAE yields the smallest $VWAEV$-value. In Figure 6.10 we have plotted both the market volatilities (crosses) as well as the calibrated (Black-Scholes) implied volatilities (circles) for this best combination. On row one we have $T = 1$ month and $T = 3$ months on the left and right side, respectively. On the second row we have $T = 6$ months and $T = 1$ year (again left and right, respectively). Continuing in this fashion, we have for the subsequent rows $T = 2$ years and $T = 3$ years, $T = 4$ years and $T = 5$ years, $T = 7$ years and $T = 10$ years.

	Localized Random Search	Simulated Annealing
RMSE	1.5801	1.5654
MSE	0.3338	0.7253
AAE	1.0865	1.6654
MARE	0.9944	1.0103

Table 6.1: Objective values for the random algorithms

	$\lambda, v_0, \overline{v}, \eta, \rho$
RMSE	0.1008,0.0447,0.1282,0.1921,-0.3671
MSE	0.3424,0.0807,0.1112,0.5262,-0.7299
AAE	4.8659,0.0280,0.0526,0.5775,-0.8358
MARE	1.9391,0.0298,0.0386,0.5378,-0.8144

Table 6.2: Parameters obtained through calibration via localized random search

	$\lambda, v_0, \overline{v}, \eta, \rho$
RMSE	0.7453,0.0472,0.1457,1.6240,-0.6357
MSE	4.1659,0.1146,0.0453,0.9325,-0.2870
AAE	2.4155,0.1432,0.0560,0.8567,-0.9887
MARE	7.7506,0.0034,0.0262,0.6590,-0.2755

Table 6.3: Parameters obtained through calibration via simulated annealing

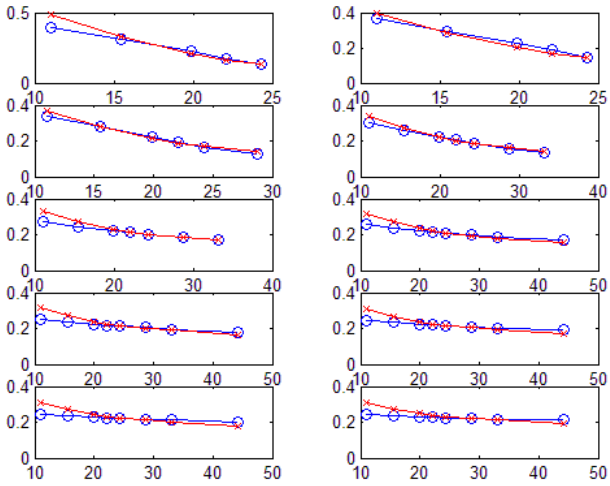


Figure 6.10: Market implied volatilities (crosses) and calibrated volatilies (circles) for the combination localized random search and AAE.

	Localized Random Search	Simulated Annealing
RMSE	1.5105	1.5171
MSE	1.4869	2.4036
AAE	1.4223	1.9330
MARE	5.1984	6.8867

Table 6.4: VWAEV-values for the random algorithms

Now we turn to the deterministic algorithms. To be precise, we will consider the functions **fmincon** (constrained optimization), **fminsearch** (unconstrained optimization) and **lsqnonlin** (nonlinear least squares) as available in MATLAB. Note that the algorithms cannot be compared with each other straightforwardly: **fminsearch** requires the objective function to be rewritten since we have a constrained optimization problem. In addition, **lsqnonlin** can only be used whenever the objective function is given by MSE. Hence the algorithms will only be compared in very few cases. (We can of course compare them to the stochastic algorithms above.) Note that these algorithms can terminate within 60 seconds as some stopping criterion can be met within this time span. This can be partly avoided by choosing loose stopping criteria. We can however not guarantee that this never happens as it depends on the choice of the initial point from which the searching process starts. One should hence bear in mind that the results presented below are obtained using *approximately* 60 seconds. On the other hand, there is the possibility that the starting point is a bad one in the sense that the duration of the calibration process exceeds the time limit greatly, as will be illustrated below.

Let us consider **fmincon** first. To illustrate the point just made, let

$$\lambda = 7.6210, v_0 = 0.1263, \overline{v} = 0.0251, \eta = 0.5196, \rho = -0.1106$$

to be the initial point, then **fmincon** will have great difficulty in solving the optimization problem no matter which of the objective functions is selected. In this case the optimization process has

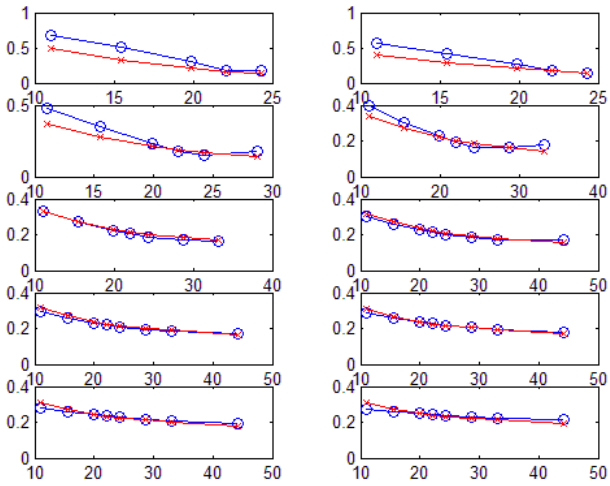


Figure 6.11: Market implied volatilities (crosses) and calibrated volatilies (circles) for the combination **fmincon** and AAE.

not finished yet after 12 minutes! On top of that, bad starting points might yield unrealistic parameters which was particularly true for MARE, e.g. we have observed η to be equal to 640! Below we will present a way to overcome this problem. Calibration results are shown in Tables 6.5 and 6.6. These results are all based on 'good' starting points.

	Objective Values	$\lambda, v_0, \overline{v}, \eta, \rho$
RMSE	0.8724	2.9588,0.0372,0.0689,1.2792,-0.6670
MSE	4.3625	0.0844,0.0504,0.0740,0.3212,-0.7481
AAE	0.7303	6.5140,0.0445,0.0668,2.7121,-0.6703
MARE	0.9999	0.9843,0.0063,0.0158,0.0178,-0.9583

Table 6.5: Calibration results for **fmincon**

	$VWAEV$
RMSE	0.9069
MSE	5.6753
AAE	0.8844
MARE	11.0738

Table 6.6: VWAEV-values for **fmincon**

Again we see that using the objective function AAE gives the smallest (global) objective value. The resulting volatilities are shown in Figure 6.11.

Next we consider **lsqnonlin**. By changing certain settings we can in fact get three different algorithms: a trust region method, Levenberg-Marquardt and Gauss-Newton method. Again, see the MATLAB manual for details. Representative objective values and the associated parameters

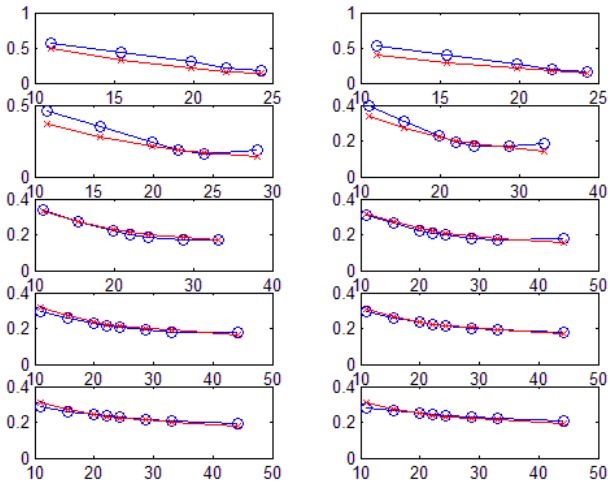


Figure 6.12: Market implied volatilities (crosses) and calibrated volatilies (circles) for **lsqnonlin** (trust region algorithm).

are presented in Tables 6.7 and 6.8.

	Trust Region	Levenberg-Marquardt	Gauss-Newton
MSE	0.7552	0.7639	0.7686

Table 6.7: Objective values for **lsqnonlin**

	$\lambda, v_0, \overline{v}, \eta, \rho$
Trust Region	2.7064,0.0549,0.0702,1.4329,-0.6480
Levenberg-Marquardt	3.2996,0.0562,0.0698,1.7102,-0.6489
Gauss-Newton	3.6965,0.0578,0.0695,1.8979,-0.6456

Table 6.8: Parameters obtained through calibration via **lsqnonlin**

From these tables we can see that there are only minor differences, both in terms of the objective function values as well as the parameters. This is also true in terms of the global performance measure:

	Trust Region	Levenberg-Marquardt	Gauss-Newton
MSE	0.8708	0.8729	0.8742

Table 6.9: VWAEV-values for **lsqnonlin**

Figure 6.12 shows the resulting volatilities for the trust region method.

Finally we have **fminsearch**. In order to be able to use this unconstrained optimization algorithm, we have to redefine the variables/parameters. One way to deal with nonnegative

variables is by taking the exponent of them. For example, $\lambda > 0$ can be replaced the quantity

$$\exp(c_1\lambda_{new}) > 0,$$

where $\lambda_{new} \in \mathbb{R}$. The constant c_1 is chosen such that unrealistic values for the parameters are avoided. The correlation coefficient $-1 \leq \rho \leq 1$ can also be handled with the aid of the exponential function:

$$-1 \leq \frac{2}{1+\exp(c_2\rho_{new})} - 1 \leq 1,$$

where $\rho_{new} \in \mathbb{R}$. The constant c_2 serves the same purpose as c_1. Values for the constants and calibration results are shown in Tables 6.10 and 6.11. Of course, no guarantee can be given that the recorded values for the constants c_1 and c_2 always do their jobs. They were effective in the cases we have seen however.

	Objective Values	$\lambda, v_0, \overline{v}, \eta, \rho$	c_1, c_2
RMSE	0.8627	0.7497,0.0408,0.0734,0.3863,-0.7173	0.05,0.02
MSE	0.1660	3.0949,0.0398,0.0585,0.9038,-0.6603	0.05,0.01
AAE	0.7033	2.0186,0.0452,0.0687,0.8903,-0.6913	0.001,0.005
MARE	0.4741	1.8686,0.0294,0.0756,0.7515,-0.7698	0.001,0.005

Table 6.10: Calibration results for **fminsearch**

	$VWAEV$
RMSE	1.4322
MSE	1.0856
AAE	0.8687
MARE	1.1390

Table 6.11: VWAEV-values for **fminsearch**

In Figure 6.13 we have plotted the volatilities obtained with respect to AAE.

6.4.2 Concluding Remarks

Based on the above results we can conclude that - in terms of the global performance measure $VWAEV$ - **fminsearch** together with AAE yields the smallest objective value. However, **fminsearch** introduces the constants c_1 and c_2, that has to be tuned for each specific set of market data. Due to this 'offline' step, one would certainly violate the 60 seconds time limit. The pair **fmincon**-AAE, again purely based on the $VWAEV$-value, is just slightly worse than the above 'best' combination. However, there aren't any auxiliary efforts in this case. Hence when considering both $VWAEV$-values and the overall time needed, we are tempting to nominate the pair **fmincon**-AAE to be the best, in the sense just defined.

As pointed out, the determination of a good starting point from which the searching process starts, might present problems for deterministic algorithms. Now we will present numerical evidence that doing this via a random algorithm will work. In fact, we have experimented with the strategy of choosing an initial point with the aid of the localized random search algorithm. We used 100 iterations (about 20 seconds) of the algorithm to come up with an initial point. **Fmincon** then uses this point to optimize AAE. Numerical evidence suggests that this is indeed a wise strategy. Unrealistic parameters were not observed. In addition, in some cases even

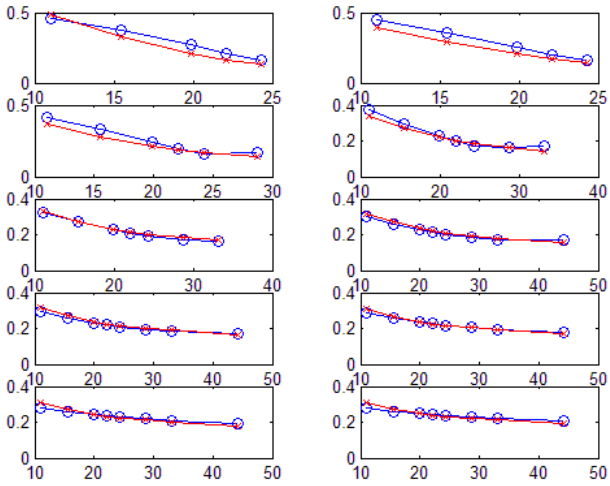


Figure 6.13: Market implied volatilities (crosses) and calibrated volatilies (circles) for the combination **fminsearch** and AAE.

smaller objective values were returned within 60 seconds. Note that we have lifted the 60 seconds restriction. In this way one can gain an idea of how well the Heston model actually is in describing reality. Table 6.12 shows 8 realizations.

	Objective Values	**fmincon** and localized random search	100 iterations
AAE	$VWAEV$	$\lambda, v_0, \overline{v}, \eta, \rho$	Overall Time (seconds)
0.7308	0.8980	4.9966,0.0338,0.0669,2.0460,-0.6763	93
0.6956	0.8583	1.1989,0.0452,0.0702,0.5808,-0.6845	58
0.5923	0.6564	0.1283,0.0555,0.1141,0.2311,-0.6888	138
0.7311	0.8957	5.4214,0.0345,0.0669,2.2382,-0.6704	92
0.5958	0.6569	0.1455,0.0559,0.1066,0.2384,-0.6807	136
0.7166	0.8849	2.3946,0.0430,0.0666,0.9980,-0.6688	144
0.5976	0.6633	0.1505,0.0553,0.1014,0.2228,-0.6961	65
0.6995	0.8671	1.1444,0.0468,0.0685,0.4893,-0.7296	129

Table 6.12: Calibration results for **fmincon**, the initial point being provided by 100 iterations of the localized random search algorithm.

In fact, similar results can be observed when we only use 10 iterations (about 2 seconds) of the random search algorithm. See Table 6.13 in which we again show 8 realizations.

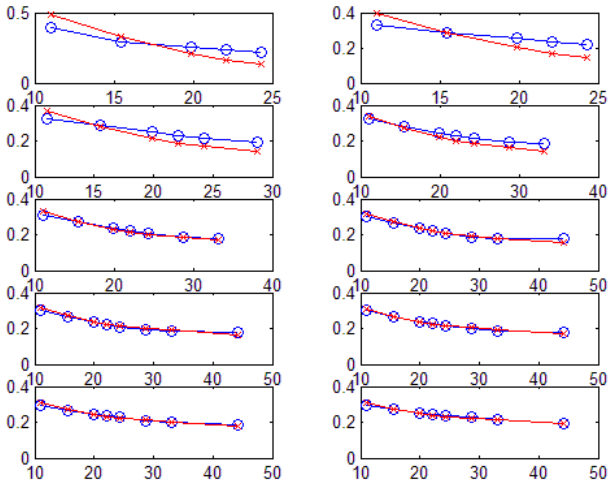


Figure 6.14: Market implied volatilities (crosses) and calibrated volatilies (circles) for the combination **fmincon** and 10 iterations of the localized random search algorithm.

	Objective Values	**fmincon** and localized random search	10 iterations
AAE	$VWAEV$	$\lambda, v_0, \overline{v}, \eta, \rho$	Overall Time (seconds)
0.7352	0.8856	9.6839,0.0463,0.0666,4.0025,-0.6688	27
0.7299	0.8842	6.3326,0.0442,0.0668,2.6395,-0.6702	46
0.5923	0.6587	0.1298,0.0553,0.1139,0.2305,-0.6926	119
0.7325	0.8858	7.7903,0.0434,0.0668,3.2446,-0.6682	43
1.1450	1.2008	0.1851,0.0525,0.2331,0.9100,-0.4442	120
0.7329	0.8852	8.1639,0.0436,0.0668,3.4136,-0.6659	41
0.7124	0.8847	2.4648,0.0403,0.0673,1.0005,-0.7027	61
0.7345	0.8844	9.0890,0.0456,0.0667,3.8154,-0.6607	67

Table 6.13: Calibration results for **fmincon**, the initial point being provided by 10 iterations of the localized random search algorithm.

For the third realization above , we have Figure 6.14. At first sight one might be surprised about the fact that this is in fact the best fit we have seen thus far (in terms of the $VWAEV$-value), because of the relatively large errors for small values of T. However, recall that $VWAEV$ is a vega-weighted measure, and options with short times to expiry have small vega's. Hence option prices are less sensitive to errors in the volatility. In Gatheral (2003*b*) it is indicated that options with short times to expiry are best handled by stock processes with jumps.

Market smile curves surely also can aid us in choosing a good initial point for deterministic algorithms, since we know what parameter values cause certain smiles in the Heston model. However, we will not attempt to embark on this art, say. Instead, we recommend using the localized random search algorithm.

Finally, we remark that the presented results are roughly in line with those presented in Schoutens et al. (2004). In their paper several stock models have been calibrated to market data,

the Heston model being one of them. They noted that the resulting fits are very satisfactory. Furthermore, the calibrated stock models were applied to the pricing of several exotic options. Not surprisingly, each stock model results in different prices for the exotics. It would be an interesting exercise to answer the following question: suppose we have two (different) calibrated sets of parameters that both yield (approximately) the same objective values. Will these two sets of parameters give (approximately) the same prices for the various exotics? If so, it will presumably have major consequences for the use of calibrated (to simple vanilla options) stock models to the pricing of more exotic options.

7 Conclusions and Recommendations

In this thesis we have presented a major complement to existing literature on option pricing via Fourier inversion: We have collected the two existing Fourier methods in a single thesis, and more importantly, we have presented plenty numerical results to prove or disprove their effectiveness. To see that we really add a contribution, consider the following. Zhu (2000) discusses Gil-Palaez inversion in detail for various asset dynamics. However, no mention is made about the numerical difficulties in the approximations of the integrals. Carr & Madan (1999) have provided - as far as it fits in a research paper - numerical insights into the Fourier techniques. However, no mention is made about interpolation that is inevitable in practice. Schoutens et al. (2004) did not mention this crucial step either. To the best of our knowledge, they deserve credit for indicating the application of Carr-Madan inversion to calibration. However, no mention is made about potential problems around calibration. These omissions - and other practical issues around the actual application of the Fourier methods - are all included in our exposition.

We have examined the practical applicability of the two existing Fourier inversion methods in a Black-Scholes and a Heston world. Both Fourier techniques can be reduced to the approximation of some Fourier integral. Singularities in the origin of the integrands in Gil-Palaez inversion prevents the Simpson's rule to be accurate. On top of that, FFT-algorithms in popular software packages cannot be applied for the computational work involved.

Carr-Madan inversion on the other hand, is singularity-free. Furthermore, the FFT can be applied for the rapid pricing of options. However, the application of the FFT comes at a cost since it requires a uniform grid, i.e. only quadrature rules with uniform grids can be utilized to approximate the arising integrals. This itself will in general not be a problem. It is the combination of a uniform grid and a peaked integrand (which is typically involved in Carr-Madan inversion) that complicates matter, as we cannot concentrate our grid points in the peaked part of the integrand, which ideally should be the case. Consequently, the number of grid points N (the grid spacing Δv) will be unnecessary large (unnecessary small). Moreover, by the reciprocity relation

$$\Delta v \Delta k = \frac{2\pi}{N}$$

we typically would have a $\log K$-domain that is far too large. The exponential grid in the K-domain forces us to apply some form of interpolation in practice. The time this step consumes is typically much longer than an application of the FFT requires. In fact, if we only want to evaluate say ten strikes per time to maturity - which is usually the case in practice - direct integrations (after a change of variable) will be about four times as fast as the FFT *plus* an additional application of cubic spline interpolation, if similar accuracies are required. For linear interpolation, the overall computational time needed will be approximately equal. However, the relative error would reduce from 10^{-8} (cubic spline interpolation) to 10^{-5} (linear interpolation), which might be still acceptable for some practitioners. Of course, in this case there will be no need to resort to Carr-Madan inversion.

In Carr-Madan inversion, choosing $\alpha = 0.75$ and using the grid spacings

$$\Delta k = 0.025, \ \Delta v = \frac{2\pi}{N \Delta k},$$

where $N = 2^{11} = 2048$, typically gives relative (discretization) errors of order 10^{-9}. Cubic spline interpolation reduces the relative error to 10^{-8}. Truncation errors will be negligible for the above grid settings. Certain parameter combinations (we especially mention large values for S/K (10^6), a large T (> 20) together with a high volatility σ (> 0.4)) yields less accurate results. Increasing N is of help.

We have calibrated the Heston model to 70 ING vanilla calls. Moreover, we compared several objective functions and algorithms (deterministic versus random). A combination of these two types of algorithms gives the best results in terms of the $VWAEV$ measure. More specifically, we recommend using the localized random search algorithm for the determination of a starting point (say using 10 iterations), from which the deterministic search process starts. This deterministic search is preferably done by the constrained optimization algorithm **fmincon** as available in MATLAB. As for the associated objective function, we recommend the average absolute error function AAE. Finally, we remark that the Heston model manages to fit market data quite reasonably (smallest observed $VWAEV$-value: 0.6587). Fits are worse for shorter times to maturities (say ≤ 1 year).

Throughout this thesis we have followed Carr & Madan (1999) by evaluating integrals using the Simpson's rule. As there are other more accurate uniform-grid-quadratures, improvements might be found in this direction. Specifically, we would like to have quadratures that are able to give accurate results with a relatively large grid spacing Δv, since in this case (by the reciprocity relation) Δk can be chosen smaller for a fixed value of N. Linear interpolation might then be sufficiently accurate, even for academics. Naturally, the quadrature rule must be formulated in a way that is suitable for the FFT.

The (scientific) determination of a value for α presents another challenge. Here we just used the 'magic' value $\alpha = 0.75$ suggested by Schoutens et al. (2004). One would like to have a value that allows the grid spacing Δv to be relatively large. Using exactly the same argument as above, Carr-Madan inversion could then be faster and more accurate than direct integrations, especially when one has more strikes per time to maturity.

Time constraints did not allow us to study more than two stock models. More advanced models should be considered, especially those allowing jumps. From Schoutens et al. (2004) we see that certain normal inverse Gaussian models with stochastic time are rather promising.

Finally, we want to mention other directions in which fast option pricing techniques have been sought recently. In Chourdakis (n.d.) it is claimed that the so-called fractional fast Fourier transform is about 45 times as fast as the FFT. In Den Iseger & Oldenkamp (2005) Laplace inversion is discussed. A systematic comparison of all these methods is still lacking in literature however. The fastest is still to be determined.

A Market Quotes ING Calls

In Section 6.4 the Heston model was calibrated to market data. Specifically, we have used market data of ING calls on January 12, 2005. The spot price of the stock was $S = 22.1$ Euros. The exact data are presented in Table 1.1. For example, we see that when $K/S \cdot 100 = 50$, the (Black-Scholes) implied volatility for $T = 1$ month is equal to $48.87/100 = 0.4887$. As another example, if $K/S \cdot 100 = 150$, the (Black-Scholes) implied volatility for $T = 10$ years is equal to $21.32/100 = 0.2132$.

	1m	3m	6m	1y	2y	3y	4y	5y	7y	10y
50	48.87	39.78	36.83	34.14	32.85	31.95	31.51	30.93	30.55	30.56
70	33.07	28.83	28.14	27.24	27.37	27.16	27.16	26.89	26.94	27.36
90	21.02	20.47	21.51	21.97	23.19	23.51	23.84	23.81	24.19	24.92
100	15.92	16.93	18.7	19.75	21.43	21.97	22.44	22.51	23.02	23.88
110	13.4	14.7	16.92	18.34	20.31	20.99	21.55	21.69	22.29	23.23
130			14.14	16.13	18.55	19.46	20.16	20.39	21.13	22.2
150				14.22	17.04	18.14	18.96	19.28	20.14	21.32
200						15.5	16.56	17.05	18.15	19.55

Table 1.1: Quoted (Black Scholes) implied volatilities

However, there are dividend payments in future that should be incorporated in the option prices. This has been done via *Sophis*, an internal system of Rabobank International. The final discounted option prices (in Euros) are shown in Tables 1.2 and 1.3. Table 1.4 presents the discount factors used.

	1m	3m	6m	1y	2y	3y
50	11.06923972	11.10940828	10.72855528	10.45709137	10.04487921	9.695932624
70	6.656971659	6.718057932	6.41044248	6.340608421	6.3574989	6.351424449
90	2.264428371	2.468670294	2.482554089	2.801762582	3.319239301	3.624364669
100	0.424329287	0.804840096	1.046555593	1.497126009	2.157660537	2.555245094
110	0.002301507	0.090755563	0.286162451	0.672421552	1.326149305	1.754705
130			0.002467601	0.065309982	0.398991481	0.735629447
150				0.00153215	0.081986458	0.254204108
200						0.006570802

Table 1.2: Discounted option prices (1)

	4y	5y	7y	10y
50	9.812222065	10.26374006	11.15382502	12.36626115
70	6.710099815	7.314340687	8.490465247	10.08096182
90	4.151482786	4.833144344	6.178708585	8.031324759
100	3.114936414	3.795788514	5.167525671	7.099039486
110	2.307890322	2.963013277	4.322764002	6.292796992
130	1.182835029	1.723423954	2.960036573	4.907938211
150	0.541411598	0.930673575	1.955741952	3.773711268
200	0.044367939	0.139962345	0.589592237	1.837421864

Table 1.3: Discounted option prices (2)

	discount factor
1m	0.998258865
3m	0.994639238
6m	0.989107139
1y	0.977194804
2y	0.95048351
3y	0.921111771
4y	0.889709543
5y	0.857916315
7y	0.791077563
10y	0.692152808

Table 1.4: Discount factors

References

Ahlfors, L. (1979), *Complex Analysis. An Introduction to the Theory of Analytic Functions of One Complex Variable*, McGraw-Hill Publishing Company.

Bakshi, G. & Madan, D. (2000), 'Spanning and derivative security valuation', *Journal of Financial Economics* **55**, 205–238.

Bazarraa, M., Sherali, H. & Shetty, C. (1993), *Nonlinear Programming. Theory and Algorithms*, John Wiley and Sons.

Björk, T. (1998), *Arbitrage Theory in Continuous Time*, Oxford University Press.

Black, F. & Scholes, M. (1973), 'The pricing of options and corporate liabilities', *Journal of Political Economy* **81**, 637–659.

Briggs, W. L. & Henson, V. E. (1995), *The DFT. An Owner's Manual for the Discrete Fourier Transform*, SIAM.

Burden, L. R. & Faires, J. D. (1998), *Numerical Methods*, Pacific Grove: Brooks/Cole.

Carr, P. & Madan, D. (1999), 'Option valuation using the fast fourier transform', *Journal of Computational Finance* **3**, 463–520.

Chourdakis, K. (n.d.), *http://www.theponytail.net/pdf/Research/Fractional_FFT.pdf* .

Cooley, J. W. & Tukey, J. W. (1965), 'An algorithm for the machine computation of the complex fourier series', *Mathematics of Computation* **19**, 297–301.

Cox, J., Ingersoll, J. & Ross, S. (1985), 'A theory of the term structure of interest rates', *Econometrica* **53**, 385–408.

Davis, P. J. & Rabinowitz, P. (1984), *Methods of Numerical Integration*, Academic Press, Inc.

Den Iseger, P. & Oldenkamp, E. (2005), 'Pricing guaranteed return rate products and average price options', *Journal of Computational Finance* **forthcoming**.

Duffie, D., Pan, J. & Singleton, K. (2000), 'Tranform analysis and asset pricing for affine jump diffusions', *Econometrica* **68**, 1343–1376.

Fouque, J.-P., Papanicolaou, G. & Ronnie Sircar, K. (2001), *Derivatives in Financial Markets with Stochastic Volatility*, Cambridge University Press.

Gatheral, J. (2003*a*), 'Lecture 1: Stochastic volatility and local volatility', *Case Studies in Financial Modelling Course Notes. Courant Insitute in Mathematical Sciences, http://www.math.nyu.edu/fellows_fin_math/gatheral/lecture1_2003.pdf* .

Gatheral, J. (2003*b*), 'Lecture 2: Fitting the volatility skew', *Case Studies in Financial Modelling Course Notes. Courant Insitute in Mathematical Sciences, http://www.math.nyu.edu/fellows_fin_math/gatheral/lecture2_2003.pdf* .

Gentleman, W. & Sande, G. (1966), 'Fast fourier transforms for fun and profit', *Proc. 1966 Fall Joint Computer Conf. AFIPS* **29**.

Gil-Palaez, J. (1951), 'Note on the inversion theorem', *Biometrika* **38**, 481–482.

Glasserman, P. (2003), *Monte Carlo Methods in Financial Engineering*, Springer.

Goldberg, R. R. (1961), *Fourier Transforms*, Cambridge University Press.

Heath, D. & Schweizer, M. (2000), 'Martingales versus pdes in finance: An equivalence result with examples', *Journal of Applied Probability* **37**, 947–957.

Heston, S. (1993), 'A closed-form solution for options with stochastic volatility with applications to bond and currency options', *Review of Financial Studies* pp. 327–343.

Jackson, N., Suli, E. & Howison, S. (1998), 'Computation of deterministic volatility surfaces', *Journal of Computational Finance* **2**, 5–32.

Karlin, S. & Taylor, H. (1981), *A Second Course in Stochastic Processes*, Academic Press.

Knopp, K. (1945), *Theory of Functions. Part One: Elements of the General Theory of Analytic Functions*, New York Dover Publications.

Knopp, K. (1947), *Theory of Functions. Part Two: Applications and Continuation of the General Theory*, New York Dover Publications.

Lee, R. (2004), 'Option pricing by transform methods: Extensions, unifications, and error control', *Journal of Computational Finance* **7**.

Madan, D., Carr, P. & Chang, E. C. (1998), 'The variance gamma process and option pricing', *European Finance Review* **2**, 79–105.

Press, W. H., Tuekolsky, S. A., Vetterling, W. T. & Flannery, B. P. (1992), *Numerical Recipes in C. The Art of Scientific Computing*, Cambridge University Press.

Schoutens, W., Simons, E. & Tistaert, J. (2004), 'A perfect calibration! now what!', *Wilmott Magazine* .

Spall, J. (2003), *Introduction to Stochastic Search and Optimization. Estimation, Simulation, and Control*, John Wiley and Sons.

Van Loan, C. (1992), *Computational Frameworks for the Fast Fourier Transform*, SIAM.

Wilmott, P. (1998), *Derivatives. The Theory and Practice of Financial Engineering*, John Wiley and Sons.

Zhu, J. (2000), *Modular Pricing of Options. An Application of Fourier Analysis*, Springer.