Efficient pricing algorithms for exotic derivatives

ISBN 978 90 5170 909 4

This book is no. 437 of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

Efficient pricing algorithms for exotic derivatives

Efficiënte waarderingsalgoritmen voor exotische derivaten

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof. dr. S.W.J. Lamberts

en volgens besluit van het College voor Promoties

De openbare verdediging zal plaatsvinden op

vrijdag 21 november 2008 om 16.00 uur

door

ROGER LORD
geboren te Geleen

ERASMUS UNIVERSITEIT ROTTERDAM

Promotiecommissie

Promotor:	Prof.dr. A.A.J. Pelsser
Overige leden:	Prof.dr. D.J.C. van Dijk Prof.dr. A.C.F. Vorst Prof.dr. P.P. Carr

Preface

When I started thinking about pursuing a PhD in mathematical finance, it was soon clear to me that I wanted to combine academia with the financial industry. The combination of Erasmus University Rotterdam, where Antoon Pelsser was my supervisor, and the Derivatives Research & Validation Team at Rabobank International enabled me to do this. I would like to take this opportunity to thank Antoon for his guidance.

At Rabobank International I am particulary indebted to Sacha van Weeren and Maarten Rosenberg for creating my position. I would like to thank Sacha for the stimulating environment his team provided, the pointers to many good articles, and especially for the fact that, while annoying at the time, he made me explain every step in every derivation I made.

I will now chronologically walk through the various chapters of this thesis, and thank everyone where thanks are due. Without a doubt I will in the process forget some people – they are thanked too. Initially the focus of my thesis was meant to be on exotic interest rate derivatives, but my interests drifted when Jeroen van der Hoek, a former colleague from Cardano Risk Management, pointed me towards Curran's approximation for Asian options. This led to the foundation of Chapter 7, for which I thank Jeroen.

The next chapter was inspired by a presentation Antoon Pelsser gave at the Unfinished Manuscripts seminar in Rotterdam. This eventually led to our joint publication in Chapter 8. I would like to thank Antoon for his ideas for this paper and the opportunity he gave me to present out work at the Quantitative Methods in Finance conference in Sydney.

Somewhere along the line I came into contact with Christian Kahl, with whom it was an honour and great pleasure to work. I would like to take this opportunity to thank Christian for the many ideas we generated in Amsterdam and Wuppertal, which resulted in Chapters 3 and 4, and also for the invaluable feedback he has given on Chapter 6.

The final two chapters, chronologically that is, Chapters 5 and 6, would certainly have been different had I not been allowed to supervise students, both at the university and at the bank. Thanks go out to all my students. Three of them I will mention by name. Firstly ManWo Ng, whose Master's thesis was an inspiring first look at topics covered in Chapters 3 and 4. Secondly Frank Bervoets, whose work inspired the CONV method in Chapter 5. And finally, Remmert Koekkoek, without whom the full truncation method from Chapter 6 would never have existed. I am also grateful to my other co-authors, Kees Oosterlee and Fang Fang in Chapter 5, and Dick van Dijk in Chapter 6, who I also thank for the great opportunity he gave me to attend the Fourth World Congress of the Bachelier Finance Society in Tokyo.

In general I am grateful to the Econometric Institute and the Tinbergen Institute for their financial support, which enabled me to attend several conferences. I am much obliged to several people at Erasmus University Rotterdam: Antoon Pelsser, Martin Martens and Dick van Dijk for allowing me to teach several courses on risk management and option pricing, Michiel de Pooter and Francesco Ravazzolo for creating a pleasant atmosphere when I was in Rotterdam, and finally the secretarial staff from both the Econometric Institute and the Tinbergen Institute for their assistance.

At Rabobank International I would like to thank all colleagues for creating an enjoyable working environment, in particular past and present colleagues from the Derivatives Research & Validation Team, the Risk Research Team (now Quantitative Risk Analytics) and the Credit Risk

Modelling team. My former direct colleagues at the Derivatives Research & Validation Team are thanked for enduring my many talks and technical reports, and for their feedback: Natalia Borovykh, Vladimir Brodski, Richard Dagg, Freddy van Dijk, Dejan Janković, Abdel Lantere, Maurice Lutterot, Herwald Naaktgeboren, Thomas Pignard and Erik van Raaij.

My friends are thanked for all the necessary distractions they provided. A special thank you is in place for Christian Kahl and Peter van de Ven for agreeing to be my paranymphs and therefore agreeing to wear a dress suit at my viva voce.

Last, but not least, I would like to thank my parents for their constant and invaluable support throughout my entire education. And of course Hellen, who I am certain is cherishing the additional time I will now be able to spend with her and our beloved Sam. The saying is true, no man succeeds without a good woman behind him.

Roger Lord
24 September 2008
Richmond, Greater London

Contents

Chapter 1

Introduction[1]

Options and derivatives came into existence long before the Nobel-prize winning papers of Black and Scholes [1973] and Merton [1973] launched the field of financial engineering. Option contracts had already been mentioned in ancient Babylonian and Greek times, though the first recorded case of organised futures trading occurred in Japan during the 1600s, when Japanese feudal lords sold their rice for future delivery in a market called *cho-ai-mai*, literally "rice trade on book". Around the same time, the Dutch started trading futures contracts and options on tulip bulbs at the Amsterdam stock exchange during the tulip mania of the early 1600s. The first formal futures and options exchange, the Chicago Board of Trade opened in 1848, but it was not until the opening of the Chicago Board Options Exchange in 1973, one month prior to the publication of the Black-Scholes paper, that option trading really took off.

Prior to the discovery of the Black-Scholes formula, investors and speculators would have had to use heuristic methods and their projections of the future to arrive at a price for a derivative. Attempts had been made to arrive at an option pricing formula, starting with Bachelier [1900], but all lacked the crucial insight of Black, Scholes and Merton that, under certain assumptions, the risk of an option can be fully hedged by dynamically investing in the underlying asset of that option. If one assumes that no arbitrage opportunities exist in financial markets, the price of any option must therefore be equal to the price of its replicating portfolio. This discovery, together with the arrival of hand-held calculators and, later, personal computers, made the derivatives market into the large industry it is today.

With the replication argument in hand, more exotic structures could be priced. One of the necessary requirements for such a price to make sense is that within the option pricing model, the prices of simpler, actively traded instruments, such as forward contracts and European options, coincide with their market price. It became apparent that this was not the case in the Black-Scholes model, and that the assumption that the underlying asset follows a geometric Brownian motion with constant, possibly time-dependent, drift and volatility was inappropriate. If this assumption were true, inverting the Black-Scholes formula with respect to the volatility for options with different strikes, but the same maturity, should yield approximately the same implied volatility. This is not the case. Much of the research within mathematical finance has therefore focused on alternative stochastic processes for the underlying asset, such that the prices of traded European options are more closely, if not perfectly, matched. To price an exotic option one then:

1. Chooses a model which is both economically plausible and analytically tractable;
2. Calibrates the model to the prices of traded vanilla options;
3. Prices the exotic option with the calibrated model, using appropriate numerical techniques.

This thesis is mainly concerned with the second and third steps in this process. Practitioners demand fast and accurate prices and sensitivities. As the financial models and option contracts used in practice are becoming increasingly complex, efficient methods have to be developed to

[1] Background information for this chapter has been used from Bernstein [1996], Dunbar [2000] and Teweles and Jones [1999].

cope with such models. All but one chapter of this thesis are therefore dedicated to the efficient pricing of options within so-called affine models, using methods ranging from analytical approximations to Monte Carlo methods and numerical integration.

The analytically tractable class of affine models, encompassing the Black-Scholes model, many stochastic volatility models including Heston's [1993] model, and exponential Lévy models, is described in Chapter 2. In order for a model to be practically relevant for the pricing of exotic options, it must allow for a fast and accurate calibration to plain vanilla option prices. This is indeed one of the appealing features of affine models. As for many affine models the characteristic function is known in closed-form, European option prices can be computed efficiently via Fourier inversion techniques.

Chapter 3 deals with a problem that occurs in the evaluation of the characteristic function of many affine models, among which the stochastic volatility model of Heston [1993]. Its characteristic function involves the evaluation of the complex logarithm, which is a multivalued function. If we restrict the logarithm to its principal branch, as is done in most software packages, the characteristic function can become discontinuous, leading to completely wrong option prices if options are priced by Fourier inversion. In this chapter we prove there is a formulation of the characteristic function in which the principal branch is the correct one. Similar problems in other models are also discussed.

Chapter 4 deals with the Fourier inversion that is used to price European options within the class of affine models. In Chapter 2 it is already shown that shifting the contour of integration along the complex plane allows for different representations of the inverse Fourier integral. In this chapter, we present the optimal contour of the Fourier integral, taking into account numerical issues such as cancellation and explosion of the characteristic function. This allows for fast and robust option pricing for virtually all levels of strikes and maturities, as demonstrated in several numerical examples.

The next three chapters are concerned with the actual pricing of exotic options, the third step we outlined above. Chapter 5 is mainly concerned with the pricing of early exercise options, though the presented algorithm can also be used for certain path-dependent options. The method is based on a quadrature technique and relies heavily on Fourier transformations. The main idea is to reformulate the well-known risk-neutral valuation formula by recognising that it is a convolution. The resulting convolution is dealt with numerically by using the Fast Fourier Transform (FFT). This novel pricing method, which we dub the Convolution method, CONV for short, is applicable to a wide variety of payoffs and only requires the knowledge of the characteristic function of the model. As such the method is applicable within many affine models.

Chapter 6 focuses on the simulation of square root processes, in particular within the Heston stochastic volatility model. Using an Euler discretisation to simulate a mean-reverting square root process gives rise to the problem that while the process itself is guaranteed to be nonnegative, the discretisation is not. Although an exact and efficient simulation algorithm exists for this process, at present this is not the case for the Heston stochastic volatility model, where the variance is modelled as a square root process. Consequently, when using an Euler discretisation, one must carefully think about how to fix negative variances. Our contribution is threefold. Firstly, we unify all Euler fixes into a single general framework. Secondly, we introduce the new full truncation scheme, tailored to minimise the upward bias found when pricing European options. Thirdly and finally, we numerically compare all Euler fixes to other recent schemes. The choice of fix is found to be extremely important.

Chapter 7 is a slight departure from the previous chapters in that we focus solely on the Black-Scholes model. Some methods from this chapter can however be applied to the whole class of affine models, as shown in Lord [2006]. This chapter focuses on the pricing of European Asian options, though all approaches considered are readily extendable to the case of an Asian basket option. We consider several methods for evaluating the price of an Asian option, and contribute to them all. Firstly, we show the link between several PDE methods. Secondly, we show how a closed-form expression can be derived for Curran's and Rogers and Shi's lower bound for the general case of multiple underlyings. Thirdly, we considerably sharpen Thompson's [1999a,b] upper bound such that it is tighter than all known upper bounds. Finally, we consider analytical approximations and combine the traditional moment matching approximations with Curran's conditioning approach. The resulting class of partially exact and bounded approximations can be proven to lie between a sharp lower and upper bound. In numerical examples we demonstrate that they outperform all current state-of-the-art bounds and approximations.

Finally, in Chapter 8 we consider a completely different topic, namely the properties of correlation matrices of term structure data, which can be used as inputs within term structure option pricing models, such as interest rate models. The first three factors resulting from a principal components analysis of term structure data are in the literature typically interpreted as driving the level, slope and curvature of the term structure. Using slight generalisations of theorems from total positivity, we present sufficient conditions under which level, slope and curvature are present. These conditions have the nice interpretation of restricting the level, slope and curvature of the correlation surface. It is proven that the Schoenmakers-Coffey correlation matrix also brings along such factors. Finally, we formulate and corroborate our conjecture that the order present in correlation matrices causes slope.

Chapter 2

Affine models

In this section we discuss the class of affine models which will be used frequently in the next chapters. This class of models was pioneered by Duffie and Kan [1996] in a term structure context, and subsequently analysed in great detail in Duffie, Pan and Singleton [2000] and Duffie, Filipović and Schachermayer [2003]. The popularity of this class can be explained by both its modelling flexibility and its analytical tractability, which greatly facilitates the estimation and calibration of such models. Prominent members of this class are the Black-Scholes model [1973], the term structure models of Vašiček [1977] and Cox, Ingersoll and Ross [1985], as well as Heston's [1993] stochastic volatility model.

Our exposition of affine processes will be based on Duffie, Pan and Singleton, who consider affine jump-diffusion processes. We slightly relax their assumptions as we allow the jump processes to have infinite activity. This setup is not quite as general as the regular affine processes considered by Duffie, Filipović and Schachermayer, but is sufficient for our purposes. In Section 2.1 we briefly describe Lévy processes, the building blocks of the affine models we consider in Section 2.2. The motivation behind using more general Lévy processes than the Brownian motion with drift is the fact that the Black-Scholes model is not able to reproduce the volatility skew or smile present in most financial markets. Over the past few years it has been shown that several exponential Lévy models are, at least to some extent, able to reproduce the skew or smile. In Section 2.3 we demonstrate how the discounted characteristic function can be derived for affine processes, and conclude in Section 2.4 by showing how European options can be priced by inverting the characteristic function.

2.1. Lévy processes

A Lévy process, named after the French mathematician Paul Lévy, is a continuous-time stochastic process with stationary independent increments. Its most well-known examples are Wiener processes or Brownian motions, and Poisson processes. To be precise, a càdlàg, adapted, real-valued process L(t) on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with L(0) = 0, is a Lévy process if:

1. it has independent increments;
2. it has stationary increments;
3. it is stochastically continuous, i.e. for any $t \geq 0$ and $\varepsilon > 0$ we have:

$$\lim_{s\to t}\mathbb{P}\big(|L(t)-L(s)|>\varepsilon\big)=0 \tag{2.1}$$

Each Lévy process can be characterised by a triplet (μ, σ, ν), representing the drift, diffusion and jump component of the process. We have $\mu \in \mathbb{R}$, $\sigma \geq 0$ and ν a measure satisfying $\nu(0) = 0$ and:

$$\int_{\mathbb{R}} \min(1, |x|^2)\ \nu(dx) < \infty \tag{2.2}$$

The Lévy measure carries useful information about the path properties of the process. For example, if $\nu(\mathbb{R}) < \infty$, then almost all paths have a finite number of jumps on every compact interval. In this case the process is said to have finite activity. This is the case for e.g. compound Poisson processes. In contrast, if $\nu(\mathbb{R}) = \infty$, the process is said to have infinite activity.

In terms of the Lévy triplet the characteristic function of the Lévy process for $u \in \mathbb{R}$ equals:

$$\begin{aligned}\phi(u) &= \mathbb{E}\left[\exp\left(iu\,L(t)\right)\right] = \exp(\ln\psi(u)) \\ &= \exp\left(t\left(i\mu u - \tfrac{1}{2}\sigma^2 u^2 + \int_{\mathbb{R}} (e^{iux} - 1 - iux1_{[|x|<1]})\,\nu(dx)\right)\right)\end{aligned} \tag{2.3}$$

the celebrated Lévy-Khinchin formula. The exponent $\psi(u)$ in (2.3) is referred to as the Lévy or characteristic exponent.

Lévy processes are strongly connected to infinitely divisible distributions. A random variable X is said to have an infinitely divisible distribution, if for all $n \in \mathbb{N}$ there exist i.i.d. random variables $X_1^{(1/n)}, \ldots, X_n^{(1/n)}$ such that:

$$X \stackrel{d}{=} X_1^{(1/n)} + \ldots + X_n^{(1/n)} \tag{2.4}$$

Equivalently, the characteristic function of X then satisfies:

$$\phi_X(u) = \left(\phi_{X^{(1/n)}}(u)\right)^n \tag{2.5}$$

The probability law of a random variable X is infinitely divisible if and only if its characteristic function can be written in the form of the Lévy-Khinchin formula.

In the subsequent chapters we will often deal with exponential Lévy models, where the asset price is modelled as an exponential function of a Lévy process L(t):

$$S(t) = S(0)\,\exp\left(L(t)\right) \tag{2.6}$$

Let us assume the existence of a bank account M(t) which evolves according to $dM(t) = r\,M(t)\,dt$, r being the deterministic and constant risk-free rate. As is common in most models nowadays we assume that (2.6) is formulated under the risk-neutral measure. For ease of exposure we also assume the asset pays a a deterministic and continuous stream of dividends, measured by the dividend rate q. To ensure the reinvested relative price $e^{qt}S(t) / M(t)$ is a martingale under the risk-neutral measure, we require:

$$\phi(-i) = \mathbb{E}\left[\exp\left(L(t)\right)\right] = e^{(r-q)t} \tag{2.7}$$

which is satisfied if we choose the drift μ as:

$$\mu = r - q - \tfrac{1}{2}\sigma^2 - \int_{\mathbb{R}} \left(e^x - 1 - x1_{[|x|<1]}\right)\nu(dx) \tag{2.8}$$

For more background information we refer the interested reader to Cont and Tankov [2004] for an extensive manuscript on the usage of Lévy processes in a financial context and to Sato [1999, 2001] for a detailed analysis of Lévy processes in general. Papapantoleon [2006] provides a good short introduction to the applications of Lévy processes in mathematical finance.

2.2. Affine processes

Let[2] $\mathbf{X}$ be a Markov process in the domain $D \subset \mathbb{R}^n$ satisfying:

$$d\mathbf{X}(t) = \boldsymbol{\mu}(\mathbf{X}(t))dt + \boldsymbol{\sigma}(\mathbf{X}(t))d\mathbf{W}(t) + d\mathbf{Z}(t) \tag{2.9}$$

Here $\mathbf{W}$ is a standard Brownian motion in $\mathbb{R}^n$, $\boldsymbol{\mu}$: $D \to \mathbb{R}^n$, $\boldsymbol{\sigma}$: $D \to \mathbb{R}^{n\times n}$ and $\mathbf{Z}$ is a pure jump Lévy process of infinite activity, whose jumps are described by the Lévy measure $\boldsymbol{\nu}$ on $\mathbb{R}^n$ and arrive with intensity $\{\lambda(\mathbf{X}(t)) : t \geq 0\}$ for some λ: $D \to \mathbb{R}^+$. Most models encountered in the literature either combine a Brownian motion with a jump process of finite activity, or leave out the Brownian motion and use a jump process of infinite activity. However, it is obviously possible to combine both. For notational convenience we do not allow μ, σ, λ and ν to depend on time. All results are easily extended to accommodate for time dependency in these functions.

In addition to the process in (2.9) we specify the short rate as a function r: $D \to \mathbb{R}$. The money market or bank account is then defined as:

$$dM(t) = r(\mathbf{X}(t))M(t)\,dt \tag{2.10}$$

The process $\mathbf{X}$ is affine if and only if:

- $\boldsymbol{\mu}(\mathbf{x}) = \mathbf{m}_0 + \mathbf{m}_1\mathbf{x}$, for $(\mathbf{m}_0, \mathbf{m}_1) \in \mathbb{R}^n \times \mathbb{R}^{n\times n}$;
- $\boldsymbol{\sigma}(\mathbf{x})\boldsymbol{\sigma}(\mathbf{x})^T = \boldsymbol{\Sigma}_0 + \sum_{i=1}^{n} \boldsymbol{\Sigma}_{1i}x_i = \boldsymbol{\Sigma}_0 + \boldsymbol{\Sigma}_1\mathbf{x}$, for $(\boldsymbol{\Sigma}_0, \boldsymbol{\Sigma}_1) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n\times n}$;
- $\lambda(\mathbf{x}) = \ell_0 + \boldsymbol{\ell}_1^T\mathbf{x}$, for $(\ell_0, \boldsymbol{\ell}_1) \in \mathbb{R} \times \mathbb{R}^n$.

where $\boldsymbol{\Sigma}_1 = (\boldsymbol{\Sigma}_{11} \;\; \cdots \;\; \boldsymbol{\Sigma}_{1n})$ and $\boldsymbol{\Sigma}_1\mathbf{x}$ is interpreted as a vector inner product. Finally, we also assume the short rate or discount function is affine:

- $r(\mathbf{x}) = r_0 + \mathbf{r}_1^T\mathbf{x}$, for $(r_0, r_1) \in \mathbb{R} \times \mathbb{R}^n$.

In words, for the process to be affine, we require both its instantaneous drift, variance and jump intensity to be at most affine combination of the factors. Finally, we also assume the short rate is an affine combination of the factors.

2.2.1. Heston's stochastic volatility model

As an example, let us consider a model that will feature prominently in the rest of this thesis - the Heston stochastic volatility model. Heston [1993] proposed the following model, where the stochastic volatility is modelled by the same square-root process that is used for the short rate in the Cox-Ingersoll-Ross model:

$$\begin{aligned} dS(t) &= \mu(t)S(t)dt + \sqrt{v(t)}\,S(t)\,dW_S(t) \\ dv(t) &= -\kappa(v(t) - \theta)dt + \omega\sqrt{v(t)}\,dW_v(t) \end{aligned} \tag{2.11}$$

[2] As a matter of notation, vectors and matrices will be typeset in bold.

In this stochastic differential equation (SDE) S represents the asset, whose stochastic variance v is modelled as a mean-reverting square root process. The Brownian motions are correlated with correlation coefficient ρ. Leaving the interpretation of the parameters to later chapters, we focus on the characterisation of (2.11) as an affine process. Clearly, if we take (S,v) as the state variables, the Heston model is not affine. The drift will be affine in these state variables, the variance will however not be, as we can see by calculating the instantaneous variance of the stock price, which is equal to $v(t)S(t)^2$. If we however consider (x,v), with x = ln S, as the state variables, we do obtain an affine process. Applying Itō's lemma, and rewriting the SDE in terms of independent Brownian motions, we obtain:

$$\begin{aligned} dx(t) &= \left(\mu - \tfrac{1}{2} v(t)\right)dt + \sqrt{v(t)}\, dW_1(t) \\ dv(t) &= -\kappa(v(t) - \theta)dt + \rho\omega\sqrt{v(t)}\, dW_1(t) + \omega\sqrt{(1-\rho^2)\, v(t)}\, dW_2(t) \end{aligned} \tag{2.12}$$

Using the notation of the previous section for the drift matrix, we obtain:

$$\mathbf{m}_0 = \begin{pmatrix} \mu \\ \kappa\theta \end{pmatrix} \quad \mathbf{m}_1 = \begin{pmatrix} 0 & -\tfrac{1}{2} \\ 0 & -\kappa \end{pmatrix} \tag{2.13}$$

The instantaneous variance-covariance matrix is equal to:

$$\boldsymbol{\sigma}(\mathbf{x})\boldsymbol{\sigma}(\mathbf{x})^{T} = \begin{pmatrix} \sqrt{v} & 0 \\ \rho\omega\sqrt{v} & \omega\sqrt{(1-\rho^2)v} \end{pmatrix} \begin{pmatrix} \sqrt{v} & 0 \\ \rho\omega\sqrt{v} & \omega\sqrt{(1-\rho^2)v} \end{pmatrix}^{T} = \begin{pmatrix} v & \rho\omega v \\ \rho\omega v & \omega^2 v \end{pmatrix} \tag{2.14}$$

leading to:

$$\boldsymbol{\Sigma}_0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \boldsymbol{\Sigma}_1 = \left(\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & \rho\omega \\ \rho\omega & \omega^2 \end{pmatrix} \right) \tag{2.15}$$

The interest rate process is taken constant and since there are no jumps, there is no need to specify the jump intensity and measure.

2.3. The discounted characteristic function

Duffie et al. [2000] studied the discounted extended characteristic function of affine processes, which is defined as:

$$\phi(\mathbf{u}, t, T, \mathbf{X}(t)) \equiv M(t) \cdot \mathbb{E}_t\left[\frac{\exp\left(i\mathbf{u}^T \mathbf{X}(T)\right)}{M(T)}\right] = P(t,T) \cdot \mathbb{E}_t^T\left[\exp\left(i\mathbf{u}^T \mathbf{X}(T)\right)\right] \tag{2.16}$$

where $\mathbf{u} \in \mathbb{C}^n$. We attached a subscript t to the expectation operator to indicate that the expectation is being taken with respect to the information set at time t. The second expectation is taken under the T-forward measure, induced by taking P(·,T), the zero-coupon bond maturing at time T, as the numeraire asset. The characteristic function in (2.16) divided by P(t,T) will be referred to as the forward characteristic function, though we will omit the term forward if the context is clear. Note that when $\mathbf{u} = 0$, we obtain P(t,T).

Contrary to the usual characteristic function, which takes a real-valued argument, the extended characteristic function is not always defined. We have, in shorthand notation:

$$|\phi(\mathbf{u})| = |\mathbb{E}[e^{i\mathbf{u}^T\mathbf{X}(T)}]| \leq \mathbb{E}[|e^{i\mathbf{u}^T\mathbf{X}(T)}|] = \varphi(i\,\mathrm{Im}(\mathbf{u})) \tag{2.17}$$

so that the strip of regularity of the extended characteristic function, $\Lambda_\mathbf{X}$, is defined by:

$$\Lambda_\mathbf{X} = \{\mathbf{u} \in \mathbb{C}^n \,|\, \phi(i\,\mathrm{Im}(\mathbf{u})) < \infty\} \tag{2.18}$$

In the following we will only deal with extended characteristic functions, so that we drop the term extended.

The question we consider here is how to solve the characteristic function of an affine process. From the Feynman-Kac formula for jump processes (see e.g. Cont and Tankov [2004, Proposition 12.5]), we know that ϕ should obey the following partial integro-differential equation (PIDE):

$$\begin{aligned} &\frac{\partial\phi}{\partial t} + \left(\frac{\partial\phi}{\partial\mathbf{x}}\right)^T \boldsymbol{\mu} + \tfrac{1}{2}\mathrm{tr}\left(\frac{\partial^2\phi}{\partial\mathbf{x}\partial\mathbf{x}^T}\boldsymbol{\sigma\sigma}^T\right) \\ &+ \lambda\int_{\mathbb{R}^n}\left(\phi(t,\mathbf{x}+\mathbf{z}) - \phi(t,\mathbf{x}) - \mathbf{z}^T 1_{[|\mathbf{z}|\leq 1]}\frac{\partial\phi}{\partial\mathbf{x}}\right)\boldsymbol{\upsilon}(d\mathbf{z}) = r\phi \end{aligned} \tag{2.19}$$

with the obvious boundary condition

$$\phi(\mathbf{u}, T, T, \mathbf{X}(t)) = \exp(i\mathbf{u}^T\mathbf{X}(T)) \tag{2.20}$$

The solution to (2.19)-(2.20), under certain technical regularity conditions (see Duffie et al. [2000, Proposition 1]), is given by:

$$\phi(\mathbf{u}, t, T, \mathbf{x}) = \exp(A(\mathbf{u}, t, T) + \mathbf{B}(\mathbf{u}, t, T)^T\mathbf{x}) \tag{2.21}$$

where A and $\mathbf{B}$ are respectively $\mathbb{R}$ and $\mathbb{R}^n$-valued functions. For the proof and regularity conditions we refer the interested reader to the article itself. From (2.21) we can deduce that:

$$\frac{\partial\phi}{\partial t} = \left(\frac{dA}{dt} + \left(\frac{d\mathbf{B}}{dt}\right)^T\mathbf{x}\right)\phi \qquad \frac{\partial\phi}{\partial\mathbf{x}} = \mathbf{B}\phi \qquad \frac{\partial^2\phi}{\partial\mathbf{x}\mathbf{x}^T} = \mathbf{B}\mathbf{B}^T\phi \tag{2.22}$$

Using the well-known identity that tr(**AB**) = tr(**BA**) if **AB** and **BA** are well-defined, we deduce:

$$\mathrm{tr}\left(\frac{\partial^2\phi}{\partial\mathbf{x}\mathbf{x}^T}\boldsymbol{\sigma\sigma}^T\right) = \phi\cdot\mathrm{tr}(\mathbf{B}\mathbf{B}^T(\boldsymbol{\Sigma}_0 + \boldsymbol{\Sigma}_1\mathbf{x})) = \phi\cdot\mathbf{B}^T\boldsymbol{\Sigma}_0\mathbf{B} + \phi\cdot\mathbf{B}^T\boldsymbol{\Sigma}_1\mathbf{x}\mathbf{B} \tag{2.23}$$

Inserting the functional form (2.21) into the PIDE in (2.19), and dividing by ϕ yields:

$$\begin{aligned}&\frac{\mathrm{dA}}{\mathrm{dt}}+\left(\frac{\mathrm{d}\mathbf{B}}{\mathrm{dt}}\right)^{\mathrm{T}}\mathbf{x}+\mathbf{B}^{\mathrm{T}}(\mathbf{m}_0+\mathbf{m}_1\mathbf{x})+\tfrac{1}{2}\left(\mathbf{B}^{\mathrm{T}}\boldsymbol{\Sigma}_0\mathbf{B}+\mathbf{B}^{\mathrm{T}}\boldsymbol{\Sigma}_1\mathbf{x}\mathbf{B}\right)\\&+(\ell_0+\boldsymbol{\ell}_1^{\mathrm{T}}\mathbf{x})\int_{\mathbb{R}^n}(e^{\mathbf{B}^{\mathrm{T}}z}-1-\mathbf{z}^{\mathrm{T}}1_{[|\mathbf{z}|\le 1]}\mathbf{B})\nu(\mathrm{d}\mathbf{z})=\mathrm{r}_0+\mathbf{r}_1^{\mathrm{T}}\mathbf{x}\end{aligned}\tag{2.24}$$

Let us denote $\theta(\mathbf{u})\equiv\int_{\mathbb{R}^n}(e^{i\mathbf{u}^{\mathrm{T}}z}-1-i\mathbf{u}^{\mathrm{T}}\mathbf{z}1_{[|\mathbf{z}|\le 1]})\nu(\mathrm{d}\mathbf{z})$ for $\mathbf{u}\in\mathbb{C}^n$, the jump-transform, provided it is well-defined. Now, since (2.24) should hold for every $\mathbf{x}$, it must also hold for $\mathbf{x}=0$. This yields an ordinary differential equation (ODE) for A. Using this relation, we can find an additional n ODEs for each element of $\mathbf{B}$ by setting $\mathbf{x}=(1,0,\dots,0)^{\mathrm{T}}$, …, $\mathbf{x}=(0,0,\dots,1)^{\mathrm{T}}$:

$$\begin{aligned}&\frac{\mathrm{dA}}{\mathrm{dt}}+\mathbf{B}^{\mathrm{T}}\mathbf{m}_0+\tfrac{1}{2}\mathbf{B}^{\mathrm{T}}\boldsymbol{\Sigma}_0\mathbf{B}+\ell_0\theta(-i\mathbf{B})=\mathrm{r}_0\\&\frac{\mathrm{d}\mathbf{B}}{\mathrm{dt}}+\mathbf{m}_1^{\mathrm{T}}\mathbf{B}+\tfrac{1}{2}\mathbf{B}^{\mathrm{T}}\boldsymbol{\Sigma}_1\mathbf{B}+\boldsymbol{\ell}_1\theta(-i\mathbf{B})=\mathbf{r}_1\end{aligned}\tag{2.25}$$

where with a slight abuse of notation we denote $\mathbf{B}^{\mathrm{T}}\boldsymbol{\Sigma}_1\mathbf{B}$ for the vector in $\mathbb{C}^n$ with i^{th} element equal to $\mathbf{B}^{\mathrm{T}}\boldsymbol{\Sigma}_{1i}\mathbf{B}$. The boundary conditions on ϕ carry over to A and $\mathbf{B}$: $A(\mathbf{u},T,T)=0$ and $\mathbf{B}(\mathbf{u},T,T)=i\mathbf{u}$. In general, the solutions to A and $\mathbf{B}$ will have to be found by numerically solving the ODEs in (2.25), for example by using the Runge-Kutta method, see e.g. Press, Teukolsky, Vetterling and Flannery [2007]. In these cases it can be very advantageous to choose a jump measure such that θ can easily be computed. Models for which the characteristic function can be solved in closed-form of course of course have a large advantage over models for which the ODEs in (2.25) have to be solved numerically. The characteristic functions of the models we consider will be supplied when we start using the models in later chapters. For example, the characteristic function for the Heston model from Section 2.2.1 is derived in Chapter 3.

2.4. Characteristic functions and option pricing

The first option pricing model in the literature that provided semi-analytical option prices by means of inverting the characteristic function of the underlying asset was the stochastic volatility model of Heston [1993]. Prior to Heston, Stein and Stein [1991] had utilised Fourier inversion techniques to calculate the stock price distribution in their stochastic volatility model. Whereas Heston's approach is directly applicable to any model where the characteristic function of the logarithm of the asset is known, Stein and Stein's approach relies heavily on the independence of the stochastic volatility process and the asset itself. Since Heston's seminal paper, the pricing of European options by means of Fourier inversion has become more and more commonplace.

Before demonstrating how European options can be priced by means of inversion techniques, we recall that the risk-neutral valuation theorem states that the forward price of a European call option on a single asset S can be written as:

$$C(S(t),K,T)=\mathbb{E}[(S(T)-K)^+]\tag{2.26}$$

where V denotes the value, T the maturity and K the strike price of the call. The expectation is taken under the T-forward probability measure. As (2.26) is an expectation, it can be calculated