Modelling and Simulation of Stochastic Volatility in Finance

Christian Kahl

DISSERTATION.COM

Boca Raton

Modelling and Simulation of Stochastic Volatility in Finance

Dissertation.com
Boca Raton, Florida
USA • 2008

ISBN: 1-58112- 383-3
13-ISBN: 978-1-58112-383-8

Modelling and simulation of stochastic volatility in finance

Dissertation

zur Erlangung des akademischen Grades eines

Doktor der Naturwissenschaften (Dr. rer. nat.)

dem Fachbereich C - Mathematik und Naturwissenschaften -
der Bergischen Universität Wuppertal vorgelegt von

Dipl. Math. Christian Kahl

Promotionsausschuß:

Gutachter/Prüfer:	Prof. Dr. Michael Günther
Gutachter/Prüfer:	Dr. Peter Jäckel
Gutachter/Prüfer:	Prof. Dr. Ansgar Jüngel
Prüfer:	Prof. Dr. Andreas Frommer

Dissertation eingereicht am:	2. Februar 2007
Tag der Disputation:	20. April 2007

Danksagung/Acknowledgement

Danksagung:

Die vorliegende Dissertation entstand aus einer Kooperation der Bergischen Universität Wuppertal und der Product Development Group von ABN Amro London. Daher möchte ich mich zuerst und ganz besonders bei meinen Betreuern Prof. Dr. Michael Günther und Dr. Peter Jäckel für die intensive und erfolgreiche Zusammenarbeit bedanken.

Peter Jäckels fachliche Kompetenz und sein unermüdlicher Einsatz haben die vorliegenden Forschungsergebnisse erst möglich gemacht. Sowohl die Doktorarbeit, als auch meine persönliche Entwicklung haben stark von seinem Einfluss profitiert, wofür ich mich ganz herzlich bedanken möchte. Der Dank fällt umso herzlicher aus, da ich weiß, welche Mühen es gekostet hat trotz ständigen Zeitmangels immer Zeit zu finden, sich meiner Arbeit zu widmen.

Mein Dank richtet sich auch an die "Product Development Group" von ABN Amro London unter der ehemaligen Leitung von Dr. Bernd van Linder und nun Dr. Andrew Greaves. Es fällt mir schwer, einzelne Mitglieder dieses Teams besonders hervorzuheben, dennoch möchte ich bei den Personen bedanken, die mein Arbeitsumfeld maßgeblich geprägt haben. Dr. Ioannis Chryssochoos hat entscheidend dazu beigetragen, meine anfänglichen Probleme mit C++ und Microsoft Visual Studio zu bewältigen, und auch sonst immer Zeit gefunden, meine Fragen zu beantworten. Es war mir eine große Freude, mit Dr. Andrew Betteridge und Dr. Mark Ryten mathematische Probleme zu diskutieren und ich möchte mich auch bei all denen bedanken, die ich an dieser Stelle unerwähnt gelassen habe.

Michael Günther gebührt mein ganz persönlicher Dank dafür, daß er mir diese Dissertation im Rahmen der Kooperation mit ABN Amro London ermöglicht hat. Insbesondere möchte ich mich dabei für sein Vertrauen und die Möglichkeit, meine Forschungsschwerpunkte frei zu gestalten, bedanken. Michael hat mir in Wuppertal ein optimales Forschungsumfeld zur Verfügung gestellt und auch auf persönlicher Ebene war es mir eine Freude mit ihm zusammenzuarbeiten.

Bei der Arbeitsgruppe "Angewandte Mathematik/Numerische Analysis" möchte ich mich für das angenehme Arbeitsumfeld bedanken. Ganz besonders danke ich hier Dr. Andreas Bar-

tel und Cathrin van Emmerich für intensive und hilfreiche Diskussionen. Auch Dr. Roland Pulch hat sich immer die Zeit genommen meine nicht durchweg sinnvollen Fragen zu beantworten. Nicht unerwähnt bleiben dürfen die anderen derzeitigen und ehemaligen Mitglieder der Arbeitsgruppe Stephanie Knorr, Dr. Michael Striebel und Kai Tappe.

An dieser Stelle möchte ich auch meine beiden Co-Autoren Roger Lord und Prof. Dr. Henri Schurz erwähnen. Es war mir eine große Freude mit ihnen zusammenzuarbeiten.

Dank richten möchte ich auch an meine Familie, die es geschafft hat, mich schon in jungen Jahren für Mathematik und Naturwissenschaften zu interessieren und zu begeistern. Danke, daß Ihr mir das Studium der Mathematik ermöglicht habt!

Acknowledgement:

The present dissertation bases on a cooperation of the University of Wuppertal and the Product Developement Group of ABN Amro London. Hence foremost and special thanks goes to my supervisors Prof. Dr. Michael Günther and Dr. Peter Jäckel for the intensive and successful collaboration.

Peter Jäckel's professional capacity and unfatiguing commitment are the foundations of these research results. Both this thesis as well as my personal progress have strongly benefited from his impact. Thank you very much Peter.

I further like to address my thanks to the "Product Developement Group" of ABN Amro London under the former direction of Dr. Bernd von Linder and now Dr. Andrew Greaves. Though it is difficult to highlight some members of this team I would like to thank in particular those, who significantly contributed to my working life in London. Dr. Ioannis Chryssochoos helped me to overcome my initial difficulties with C++ and Microsoft Visual Studio and was always available to answer my questions. Moreover it was a great pleasure to discuss mathematical problems with Dr. Andrew Betteridge and Dr. Mark Ryten and furthermore I want to thank all those who I did not mentioned explicitely.

Michael Günther has my sincere thanks for supervising this cooperation project with ABN Amro London. I want to thank him particularly for the confidence he put in me and the freedom he gave me to choose my research focus. Michael provided me with an optimal research environment in Wuppertal and it was a pleasure to work with him.

Thanks also goes to the working group "Applied Mathematics/Numerical Analysis" for the enjoyable working enviroment. I particularly thank Dr. Andreas Bartel and Cathrin van Emmerich for all the intensive and helpful discussions. Dr. Roland Pulch always found the time to answer my questions. I do not want to leave unmentioned the other current and former mem-

bers of this working group. Thank you, Stephanie Knorr, Dr. Michael Striebel and Kai Tappe.

It is my honour to thank my co-authors Roger Lord and Prof. Dr. Henri Schurz. It has been a pleasure to work together with you.

Last but not least I want to thank my family for all the support and to inspire the enthusiasm for mathematics and science in me. Thanks for giving me the opportunity to study mathematics!

Wuppertal, Februar 2007 Christian Kahl

Table of Contents

List of Figures

List of Tables

Mathematical notation

$1_{\{\dots\}}$	indicator function
$:=$	defined as
$\approx$	approximately
$\boldsymbol{A}^T$	transpose of the matrix $\boldsymbol{A} \in \mathbb{R}^{n\times m}$
$\boldsymbol{x}^T \cdot \boldsymbol{y}$	inner product, $\boldsymbol{x}^T \cdot \boldsymbol{y} = \sum_{i=1}^{n} x_i\, y_i$
$\boldsymbol{x}^T$	transpose of the vector $\boldsymbol{x} \in \mathbb{R}^n$
$\mathring{=}$	short hand notation for the correlation of Wiener processes, i.e. $W \cdot Z \mathring{=} \rho$ stands for $\mathrm{d}W \cdot \mathrm{d}Z = \rho\,\mathrm{d}t$
$\delta(x)$	Dirac delta distribution
$\delta_{i,j}$	Kronecker symbol
e	base of the natural logarithm (Napier's constant)
$\mathsf{E}[X]$	expectation of X, i.e. $\int X \mathrm{d}\mathbb{P}_X$
$\hat{f}(z)$	Fourier transform of f
$\mathrm{Im}(z)$	imaginary part of the complex variable z
$\langle f \rangle_\varphi$	expectation with respect to the distribution density φ
$\ln(x)$	natural logarithm, i.e. the solution of $x = \mathrm{e}^z$
$\mathbb{C}$	complex numbers
$\mathbb{N}$	positive integers
$\mathbb{P}(A)$	probability of A
$\mathbb{P}_X$	distribution induced by the random variable X
$\mathbb{R}$	real numbers
$\mathbb{R}^+$	positive real numbers

$\mathbb{Z}$	integers
$\mathcal{C}(U, V)$	continuous functions $f : U \to V$
$\mathcal{C}^k(U, V)$	functions $f : U \to V$, with k-th derivative in $\mathcal{C}(U, V)$
$\mathcal{N}(\mu, \sigma^2)$	normal distribution with mean μ and standard deviation σ
$\mathcal{O}(n)$	Landau or asymptotic notation
$\otimes$	tensor product
$\overline{z}$	complex conjugate
$\phi(u, t)$	characteristic function, $\phi(u, t) = \mathsf{E}\left[\mathrm{e}^{iuX_t}\right]$
$\mathrm{Re}(z)$	real part of the complex variable z
σ_X	standard deviation of X, $\sigma_X = \sqrt{\mathsf{V}[X]}$
$\sim$	distributed according to
$\mathrm{Black}(S, K, \sigma, T)$	Black formula for a European call option, $\mathrm{Black}(S, K, \sigma, T) = S \cdot N\left(\frac{\ln\left(\frac{S}{K}\right)}{\sigma\sqrt{T}} + \frac{\sigma\sqrt{T}}{2}\right) - K \cdot N\left(\frac{\ln\left(\frac{S}{K}\right)}{\sigma\sqrt{T}} - \frac{\sigma\sqrt{T}}{2}\right)$
$\mathrm{Cor}\left[X, Y\right]$	linear correlation between X and Y
$\mathrm{Cov}\left[X, Y\right]$	linear covariance between X and Y
$\varphi(x)$	standard normal distribution density, $\varphi(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}$
$\mathsf{V}[X]$	variance of X, $\mathsf{V}[X] = \mathsf{E}\left[(X - \mathsf{E}[X])^2\right]$
$f^{(k)}(x)$	short hand notation for the k-th derivative of f, i.e. $\frac{\partial^k f}{\partial x^k}(x)$
$f^{-1}(x)$	inverse function, i.e. $f^{-1}(f(x)) = x$
h_n	Hermite polynomial of degree n
i	$\sqrt{-1}$
$I_{(1)}^{s,t}$	Wiener integral on the interval $[s, t]$, i.e. $\int_s^t \mathrm{d}W_u$
$I_\alpha^{s,t}$	multiple Wiener integral on the interval $[s, t]$ with multi-indices α
$N(x)$	cumulative normal distribution function, $N(x) = \int_{-\infty}^{x} \varphi(z)\mathrm{d}z$
W_t	Brownian motion
$x \vee y$	maximum of x and y
$x \wedge y$	minimum of x and y

CHAPTER I

Introduction

The famous Black-Scholes model [BS73] was the starting point of a new financial industry and has been a very important pillar of all options trading since. One of its core assumptions is that the volatility of the underlying asset is constant. It was realised early that one has to specify a dynamic on the volatility itself to get closer to market behaviour. There are mainly two aspects making this fact apparent. Considering historical evolution of volatility by analysing time series data one observes erratic behaviour over time. Secondly, backing out implied volatility from daily traded plain vanilla options, the volatility changes with strike. The most common realisations of this phenomenon are the implied volatility smile or skew. The natural question arises how to extend the Black-Scholes model appropriately.

Any modelling approach aiming at this problem must fulfil some fundamental requirements. The range of model implied volatilities must be flexible enough to reflect market behaviour. In order to be of practical relevance it must additionally allow the fast and accurate pricing of European options to accomplish calibration within reasonable computation time. The approach of driving volatility itself via a stochastic process turned out to be as intuitive as successful. Depending on the particular choice of the stochastic process for volatility, these models allow different dynamics of implied volatility and thereby mapping of market behaviour.

This thesis focus on the different problems arising within the context of stochastic volatility from a modelling as well as from a numerical point of view. At the very heart of modelling volatility is the choice of a sensible process. Hereby it is of fundamental importance to bear in mind analytical properties such as positivity of volatility. Moreover one usually observes a mean-reverting behaviour of volatility over time and might want to address the analytical tractability, i.e. a closed form solution of the distribution density or its characteristic function. Once a model is chosen one needs to address the problem of calculating or approximating the price of plain vanilla options in order to calibrate the model to market data.

The thesis is divided into six chapters and we illustrate their dependency structure in figure 1.1. The second chapter provides an introduction about the stochastic volatility models

discussed throughout this thesis. We mainly focus on two different modelling approaches. First we introduce stochastic volatility models where volatility is given by a transformed Ornstein-Uhlenbeck process. Here, we concentrate on analysing the analytical properties with respect to different choices of transformation functions. The second class of stochastic volatility models we consider are affine (jump) diffusion models, having the appealing feature that the characteristic function of the underlying asset is known in closed form. In a first step, we discuss a stable numerical implementation of the characteristic function which is a rather challenging problem involving an appropriate handling of multivalued functions such as the complex logarithm. In chapter five, we use the closed form solution of the characteristic function for semi-analytical option pricing of plain vanilla options.

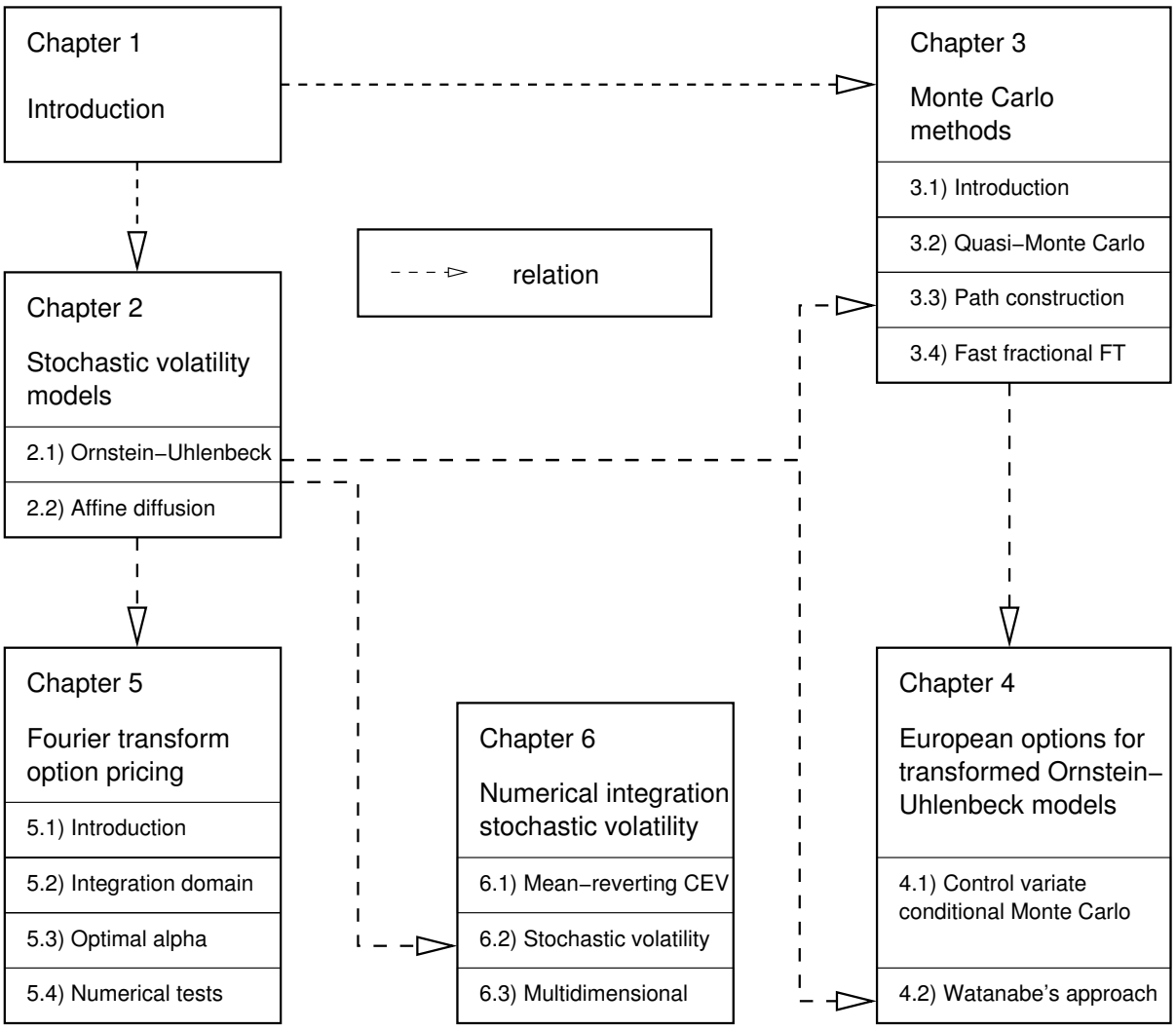


Figure 1.1: Reader's guidance.

The third chapter deals with Monte Carlo simulation methods. At the very heart of this chapter stands the optimal path construction of Wiener and Ornstein-Uhlenbeck processes within a quasi-Monte Carlo framework. In section 3.3, we compare different path construction

methods with regard to their *variability explained* and their computational complexity. Using a fast fractional Fourier transformation helps to reduce the computational effort for discrete and approximative spectral path construction methods of Wiener and Ornstein-Uhlenbeck processes as outlined in section 3.4.

Next we discuss the computation and approximation of European option prices for transformed Ornstein-Uhlenbeck stochastic volatility models, where we mainly concentrate on two different approaches. On the one hand, we use the efficient path construction of Ornstein-Uhlenbeck processes developed in chapter 3 to derive a fast and accurate control variate conditional Monte Carlo method. On the other hand, we develop asymptotic approximations for European option prices within transformed Ornstein-Uhlenbeck models using Watanabe's approach. Due to the generality of this method we extend the stochastic volatility model by adding a local volatility function which provides greater flexibility when calibrated to market data.

In chapter five, we study the semi-analytical option pricing for affine (jump) diffusion models using Fourier inversion techniques. Fourier inversion allows to price European options solely based on the knowledge of the characteristic function of the underlying asset and the payoff function by the aid of the Plancherel/Parseval equality and applying residual calculus. The numerical implementation of the inverse Fourier integral can be accomplished by transforming the unbounded integration domain to a finite interval using the limiting behaviour of the characteristic function. Furthermore we show that it is of fundamental importance for a fast and robust implementation to choose an appropriate integration contour in the complex plane. Numerical tests provide further evidence for the effectiveness of this approach.

The sixth chapter deals with numerical integration schemes for stochastic volatility models which are required to price path dependent options. Since the volatility process does not explicitly depend on the underlying asset, the integration of a stochastic volatility model can be divided into two separate steps. First we concentrate on the numerical integration of the stochastic volatility process itself. Based on the knowledge of the full volatility path, we focus on the simulation of the underlying asset in section 6.2. Using the strong approximation error as a measure to compare different approximation schemes, we derive a fast and efficient approximation method based on an interpolation of the drift and the decorrelated diffusion. Section 6.3 discusses multidimensional stochastic volatility models, where we particularly focus on how to complete a non-fully specified correlation matrix such that we obtain a symmetric positive semi-definite matrix.

In chapter seven, we summarise the results developed and discussed throughout this thesis and we provide an outlook about further research opportunities. Appendices A and B supplement the previous work.

CHAPTER II

Stochastic volatility models

The first publications on stochastic volatility models [Sco87, Wig87, HW88] were ahead of their time: the required computer power to use these models in a simulation framework was simply not available, and analytical solutions could not be found. One of the first articles that provided semi-analytical solutions was published by Stein and Stein [SS91]. An unfortunate feature of that model was that it did not give enough flexibility to represent observable market prices, i.e. it did not provide enough degrees of freedom for calibration. To make matters even worse the Stein and Stein model allows volatility to become negative which is a rather undesirable feature. In 1993, Heston [Hes93] published the first model that allowed for a reasonable amount of calibration freedom permitting semi-analytical solutions. Various other stochastic volatility models have been published since, and computer speed has increased significantly.

Throughout the following we consider stochastic volatility models to be given as a two-dimensional system of stochastic differential equations

$$\mathrm{d}S_t = \mu S_t \,\mathrm{d}t + \eta \cdot v_t^p \, S_t \,\mathrm{d}W_t \,, \tag{2.1}$$

$$\mathrm{d}v_t = a(v_t)\,\mathrm{d}t + b(v_t)\,\mathrm{d}Z_t \,, \tag{2.2}$$

with scaling parameter $\eta \in \mathbb{R}^+$. We assume the driving processes W_t and Z_t to be correlated Brownian motions satisfying $\mathrm{d}W_t \cdot \mathrm{d}Z_t = \rho\,\mathrm{d}t$. The parameter p allows us to specify either instantaneous variance for $p = 1$, or the instantaneous volatility $p = 1/2$. In order to keep the stochastic volatility process v rather general, we will further specify the dynamics of the drift a and diffusion b later on.

The outline of this chapter is as follows. First we introduce so called *transformed Ornstein-Uhlenbeck* models where the stochastic volatility process is, according to its nomenclature, given by a deterministic transformation of a standard Ornstein-Uhlenbeck process. One appealing feature of these model is that we can guarantee the positivity of the stochastic volatility by choosing an appropriate transformation function.

In a second section 2.2, we draw our attention to affine diffusion models. These models are particularly attractive since the characteristic function of the transition density of the underlying asset is known in closed form, which allows for semi-analytical solutions of European option prices by the aid of Fourier inversion techniques. The numerical implementation of the

characteristic function is rather challenging since the characteristic function is of oscillating nature and involves the evaluation of multivalued functions. In chapter 5, we continue the discussion with the implementation of the Fourier inversion, in order to calculate prices of plain vanilla options.

2.1 Transformed Ornstein-Uhlenbeck models

Definition 2.1 An Ornstein-Uhlenbeck [UO30] process can be defined as the solution of the scalar stochastic differential equation

$$\mathrm{d}y_t = -\kappa y_t \,\mathrm{d}t + \alpha\sqrt{2\kappa}\,\mathrm{d}Z_t \;, \tag{2.3}$$

with Brownian motion Z_t, initial value $y(0) = y_0$ and positive constants $\kappa, \alpha \in \mathbb{R}^+$.

The origin of this process goes back to Uhlenbeck and Ornstein's publication [UO30] in which they describe the velocity of a particle that moves in an environment with friction. Doob [Doo42] first treated this process purely mathematically and expressed it in terms of a stochastic differential equation. In modern financial mathematics, the use of Ornstein-Uhlenbeck processes is almost commonplace. The attractive features of an Ornstein-Uhlenbeck process are that, whilst it provides a certain degree of flexibility over its auto-correlation structure, it still allows for the full analytical treatment of a standard Gaussian process.

The parametrisation (2.3) permits a complete separation between the mean reversion speed and the variance of the *limiting* or *stationary distribution* of the process [FPS99]. The solution of (2.3) is

$$y_t = \mathrm{e}^{-\kappa t}\left(y_0 + \int_0^t \mathrm{e}^{\kappa u}\alpha\sqrt{2\kappa}\,\mathrm{d}Z_u\right) \;, \tag{2.4}$$

with initial time $t_0 = 0$. In other words, the stochastic process at time t is Gaussian with

$$y_t \sim \mathcal{N}\left(y_0\mathrm{e}^{-\kappa t}, \alpha^2\left(1 - \mathrm{e}^{-2\kappa t}\right)\right) \;, \tag{2.5}$$

and thus the stationary distribution ($t \to \infty$) is Gaussian with variance α^2: a change in parameter κ requires no rescaling of α if we wish to hold the long-term uncertainty in the process unchanged. It is straightforward to extend the above results to the case when $\kappa(t)$ and $\alpha(t)$ are functions of time [Jäc02]. Throughout the following we assume the process (2.3) to start in $y(0) = 0$. Since the variance of the driving Ornstein-Uhlenbeck process is the main criterion that determines the uncertainty in volatility for the financial underlying process, all further considerations are primarily expressed in terms of

$$\eta(t) \;:=\; \alpha \cdot \sqrt{1 - \mathrm{e}^{-2\kappa t}} \;. \tag{2.6}$$

The Ornstein-Uhlenbeck process is particularly attractive to be used as a stochastic volatility driver since it is full analytical tractable. Despite this it is not suited for modelling the

stochastic volatility itself since the process is fluctuating around zero which is a rather undesirable feature. The idea is to apply a monotonic transformation function which was first proposed in [Sco87, equation (7)] using the exponential function. The model is intuitively very appealing: for any future point in time, volatility has a lognormal distribution which is a very comfortable distribution for practitioners in financial mathematics. Alas, though, recent research [AP04] has cast a shadow on this model's analytical features. It appears that, in its full continuous formulation, the log-normal volatility model can give rise to unlimited higher moments of the underlying financial asset. However, as has been discussed and demonstrated at great length for the very similar phenomenon of infinite futures returns when short rates are driven by a lognormal process [HW93, SS94, SS97a, SS97b], this problem vanishes as soon as the continuous process model is replaced by its discretised approximation which is why lognormal volatility models remain numerically tractable in applications.

In order to avoid this problem altogether, we introduce an alternative to the exponential transformation function which is given by a simple hyperbolic form. In the following, we refer to

$$f_{\text{lin}}(y) \;:=\; y+1\,, \qquad\qquad \sigma_{\text{lin}}(y) \;:=\; \sigma_0 \cdot f_{\text{lin}}(y) \tag{2.7}$$

as the *linear volatility* transformation also known as Stein-Stein [SS91] or Schöbel-Zhu model [SZ99], to

$$f_{\text{exp}}(y) \;:=\; \mathrm{e}^{y}\,, \qquad\qquad \sigma_{\text{exp}}(y) \;:=\; \sigma_0 \cdot f_{\text{exp}}(y) \tag{2.8}$$

as the *exponential volatility* transformation also known as Scott's model [Sco87], and to

$$f_{\text{hyp}}(y) \;:=\; y+\sqrt{y^2+1}\,, \qquad\qquad \sigma_{\text{hyp}}(y) \;:=\; \sigma_0 \cdot f_{\text{hyp}}(y) \tag{2.9}$$

as the *hyperbolic volatility* transformation as introduced by Kahl and Jäckel [KJ06]. The densities of these processes are given by

$$\psi(\sigma,\sigma_0,\eta) \;=\; \frac{\varphi\left(f^{-1}(\sigma/\sigma_0),\eta\right)}{\mathrm{d}\sigma/\,\mathrm{d}y} \tag{2.10}$$

with

$$f_{\text{lin}}^{-1}(\sigma/\sigma_0) \;=\; \sigma-\sigma_0\,, \tag{2.11}$$

$$\mathrm{d}\sigma_{\text{lin}}/\,\mathrm{d}y \;=\; \sigma_0 \tag{2.12}$$

$$f_{\text{exp}}^{-1}(\sigma/\sigma_0) \;=\; \ln\left(\sigma/\sigma_0\right)\,, \tag{2.13}$$

$$\mathrm{d}\sigma_{\text{exp}}/\,\mathrm{d}y \;=\; \sigma_{\text{exp}} \tag{2.14}$$

$$f_{\text{hyp}}^{-1}(\sigma/\sigma_0) \;=\; \left(\sigma/\sigma_0-\sigma_0/\sigma\right)/\,2\,, \tag{2.15}$$

$$\mathrm{d}\sigma_{\text{hyp}}/\,\mathrm{d}y \;=\; \frac{2\sigma_0\sigma_{\text{hyp}}^2}{\sigma_0^2+\sigma_{\text{hyp}}^2} \tag{2.16}$$

and

$$\varphi(y,\eta) \;:=\; \frac{\mathrm{e}^{-\frac{1}{2}\left(\frac{y}{\eta}\right)^2}}{\eta\cdot\sqrt{2\pi}}\,. \tag{2.17}$$

The hyperbolic transformation has been chosen to match the exponential form as closely as possible near the origin, and only to differ significantly in the regions of lower probability given for $|{}^{y}\!/_{\eta}| > 1$.

Even though the linear transformation (2.7) does not guarantee the volatility to remain positive (see figure 2.2), we included this model to our discussion since it allows for semi-analytical solutions for pricing plain-vanilla options by the aid of Fourier inversion.

In figure 2.1, we compare the functional form of the linear, exponential and hyperbolic transformation function.

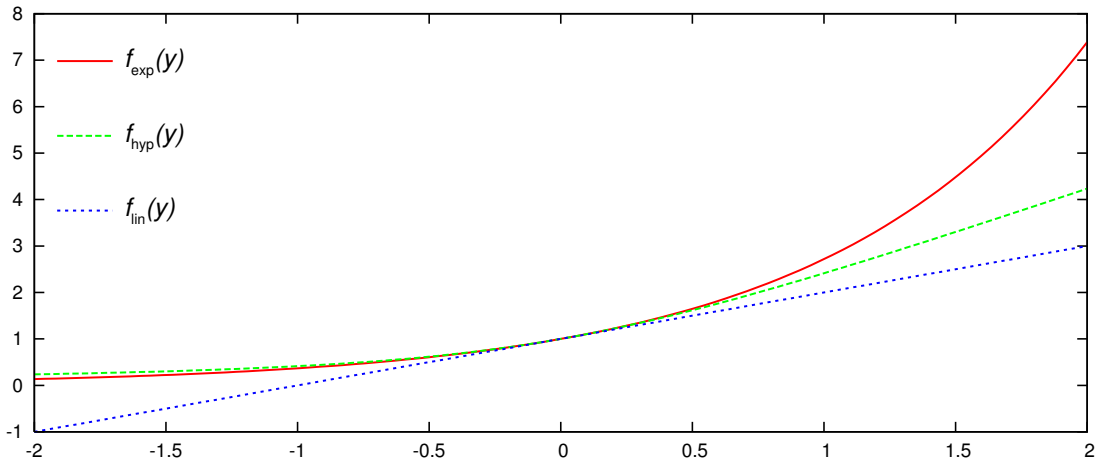


Figure 2.1: Linear, exponential and hyperbolic transformation functions.

2.1.1 Linear vs. Exponential vs. Hyperbolic transformation

In figure 2.2 we compare the densities of the transformed Ornstein-Uhlenbeck process. One clearly sees that the linear transformation has a positive probability for negative values. At first glance on linear scale one cannot recognise any reasonably difference between the hyperbolic and exponential transformation. On a logarithmic scale, the differences become clear: the exponential transformation has significantly fatter tails than the hyperbolic transformation. Thus the probability of very low as well as very high values is noticeable lower for the hyperbolic transformation.

Returning to the analytical form of the density function (2.10), it is interesting to note that, given that $y = f^{-1}({}^{\sigma}\!/_{\sigma_0})$ is Gaussian, the volatility distribution implied by the inverse hyperbolic transformation (2.15) is nearly Gaussian for large values of instantaneous volatility $\sigma \gg 0$ since we have

$$\sigma_{\text{hyp}} \;\approx\; 2\sigma_0\cdot y \quad \text{for} \quad \sigma_{\text{hyp}} \gg 0\,. \tag{2.18}$$

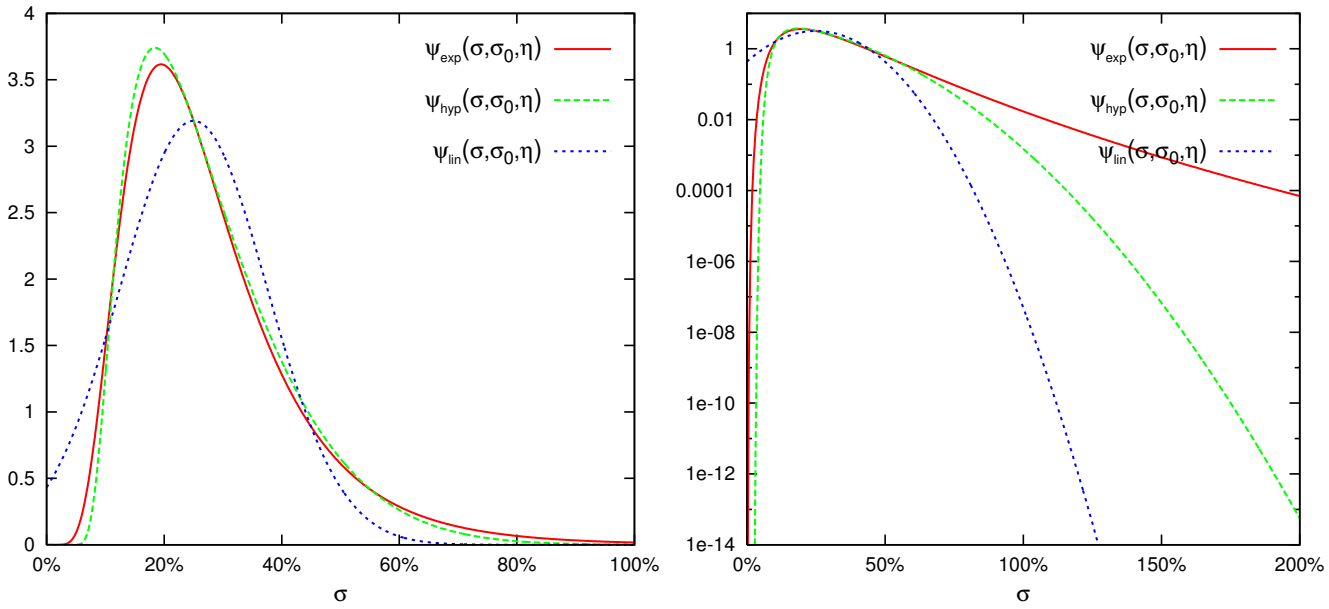


Figure 2.2: Densities of instantaneous volatility using the linear, exponential and the hyperbolic transformation of the driving Ornstein-Uhlenbeck process for $\sigma_0 = 25\%$ and $\eta = 1/2$ on a linear (left) and logarithmic (right) scale. In order to illustrate the distinctly different tails of the distribution for the hyperbolic and exponential transformation on a logarithmic scale, we only plotted the positive part of the density, even though the linear transformation has a positive probability to become negative which is already evident on a linear scale.

This feature is particularly desirable since it ensures that the tail of the volatility distribution at the higher end is as thin as the Gaussian process itself, and thus no moment explosions are to be feared for the underlying. Conversely, for small values of instantaneous volatility $\sigma \ll 1$, the volatility distribution implied by the hyperbolic volatility model is nearly *inverse* Gaussian because of

$$\sigma_{\text{hyp}} \approx -\sigma_0/(2\,y) \quad \text{for} \quad \sigma_{\text{hyp}} \to 0\,. \tag{2.19}$$

In a certain sense, the hyperbolic model can be seen as a blend of an inverse Gaussian model at the lower end of the distribution, and a Gaussian density at the upper end, with their respectively thin tails. In contrast, exponential volatility is simply lognormally distributed, which in turn gives rise to distinctly fatter tails than the normal (at the high end) or inverse normal (at the low end) density. We concretise this behaviour in the following lemma.

Lemma 2.2 The transformation methods of the Ornstein-Uhlenbeck process have the following behaviour in the tails of the distribution

1. the *linear transformation* is Gaussian distributed at the upper and lower end,
2. the *exponential transformation* is lognormal distributed at the upper and lower end,
3. the *hyperbolic transformation* is Gaussian distributed at the upper end and inverse-Gaussian at the lower end.

To obtain even more information about the analytical behaviour of the transformed Ornstein-Uhlenbeck process, we derive the driving stochastic differential equation using Itô's lemma.

Lemma 2.3 The stochastic differential equation for the linear transformed Ornstein-Uhlenbeck process is given by

$$\mathrm{d}\sigma = \kappa\,(\sigma_0 - \sigma)\,\mathrm{d}t + \sigma_0\alpha\sqrt{2\kappa}\,\mathrm{d}Z\ , \tag{2.20}$$

thus by setting $\sigma_0 = \theta$ and $\alpha = \frac{\sigma_{\mathrm{SZ}}}{\sigma_0\sqrt{2\kappa}}$, we obtain Schöbel-Zhu's original SDE [SZ99, equation (3)].

Lemma 2.4 The stochastic differential equation for the exponential transformed Ornstein-Uhlenbeck process is given by

$$\mathrm{d}\sigma = \kappa\sigma\left(\alpha^2 - \ln(\sigma/\sigma_0)\right)\mathrm{d}t + \alpha\sigma\sqrt{2\kappa}\,\mathrm{d}Z\ . \tag{2.21}$$

This corresponds to the original SDE used by Scott [Sco87, equation (7)].

Lemma 2.5 The stochastic differential equation for the hyperbolic transformed Ornstein-Uhlenbeck process is given by

$$\mathrm{d}\sigma = -\kappa\sigma\frac{\sigma^6 + \sigma^4\sigma_0^2 - (8\alpha^2+1)\sigma^2\sigma_0^4 - \sigma_0^6}{\left(\sigma^2+\sigma_0^2\right)^3}\,\mathrm{d}t\ +\ \sigma\alpha\sqrt{8\kappa}\frac{\sigma\sigma_0}{\sigma^2+\sigma_0^2}\,\mathrm{d}Z\ , \tag{2.22}$$

as first mentioned by Kahl and Jäckel [KJ06, equation (36)].

Particularly the diffusion terms of these equations are of great help for the construction of efficient numerical integration schemes for the whole two-dimensional stochastic volatility model which we will discuss in section 6.2. In order to obtain further insight we analyse the moments of the different transformation methods

2.1.2 Moments

Since the hyperbolic Ornstein-Uhlenbeck process is designed to match the exponential as close as possible we can expect that the moments reflect this behaviour. Throughout the following we assume $y \sim \mathcal{N}(0, \eta^2)$.

Lemma 2.6 The moments of the *exponential transformation* are given by

$$\mathsf{E}[(f_{\exp}(y))^m]\ =\ \mathrm{e}^{\frac{1}{2}m^2\eta^2}\ . \tag{2.23}$$

The moments of the linear and hyperbolic transformation both involves Kummer's hypergeometric function ${}_1F_1(a,b,z)$ and the confluent hypergeometric function $U(a,b,z)$.

Definition 2.7 Kummer's hypergeometric function has the following integral representation [AS84, 13.2.1]

$$ {}_1F_1(a,b,z) = \frac{\Gamma(b)}{\Gamma(a)\Gamma(b-a)} \int_0^1 \mathrm{e}^{zt} t^{a-1}(1-t)^{b-a-1}\,\mathrm{d}t \tag{2.24} $$

and an integral representation of the confluent hypergeometric function is given by [AS84, 13.2.5]

$$ U(\alpha,\beta,\gamma) \;=\; \frac{1}{\Gamma(\alpha)} \int_0^\infty s^{\alpha-1}(1+s)^{\beta-\alpha-1}\mathrm{e}^{-\gamma s}\,\mathrm{d}s\;. \tag{2.25} $$

Lemma 2.8 The moments of the *hyperbolic transformation* are given by

$$ \mathsf{E}[(f_{\text{hyp}}(y))^m] \;=\; \tfrac{2^{\frac{3}{2}n}}{\sqrt{4\pi}}\eta^n\Gamma\left(\tfrac{1+n}{2}\right){}_1F_1(-\tfrac{n}{2},1-n,\tfrac{1}{2\eta^2}) + \tfrac{2^{-\frac{3}{2}n}}{\sqrt{4\pi}}\eta^{-n}\Gamma\left(\tfrac{1-n}{2}\right){}_1F_1(\tfrac{n}{2},1+n,\tfrac{1}{2\eta^2})\;, \tag{2.26} $$

where we have a simpler expression for the first two moments

$$ \mathsf{E}[f_{\text{hyp}}(y)] \;=\; \sqrt{2}\cdot\eta\cdot U(-\tfrac{1}{2},0,\tfrac{1}{2\eta^2})\;, \qquad \mathsf{E}\left[(f_{\text{hyp}}(y))^2\right] \;=\; 1+2\eta^2\;. \tag{2.27} $$

Proof: The first moment follows directly from the integral representation of the confluent hypergeometric function (2.25) by setting $s = y^2$, $\alpha := 1/2$, $\beta := 2$, and $\gamma := 1/2\eta^2$ and the use of the relationship [AS84, 13.1.29]

$$ U(\alpha,\beta,\gamma) \;=\; \gamma^{1-\beta}\cdot U(1+\alpha-\beta,2-\beta,\gamma)\;. \tag{2.28} $$

To calculate the second moment we can use the symmetry and second moment of the Gaussian distribution. □

Lemma 2.9 The moments of the *linear transformation* are given by

$$ \mathsf{E}[(f_{\text{lin}}(y))^n] = \begin{cases} \frac{2^{-1+\frac{n}{2}}\eta^{n-1}\left(2\sqrt{2}\,\Gamma\left(1+\frac{n}{2}\right){}_1F_1\left(\frac{1-n}{2},\frac{3}{2},-\frac{1}{2\eta^2}\right)\right)}{\sqrt{\pi}} & \text{if } n \text{ is odd}\;, \\ \frac{2^{-1+\frac{n}{2}}\eta^{n}\left(2\,\Gamma\left(\frac{1+n}{2}\right){}_1F_1\left(-\frac{n}{2},\frac{1}{2},-\frac{1}{2\eta^2}\right)\right)}{\sqrt{\pi}} & \text{if } n \text{ is even}\;, \end{cases} \tag{2.29} $$

where the first moments can be calculated straightforwardly as

$$ \mathsf{E}[f_{\text{lin}}(y)] \;=\; 1\;, \qquad \mathsf{E}\left[(f_{\text{lin}}(y))^2\right] \;=\; 1+\eta^2\;. \tag{2.30} $$

More revealing than the closed form for the moments of the respective transformation functions is an analysis based on their Taylor expansions

$$ f_{\text{exp}}(y) \;=\; 1+y+\frac{y^2}{2}+\mathcal{O}\left(y^3\right)\;, \tag{2.31} $$

$$ f_{\text{hyp}}(y) \;=\; 1+y+\frac{y^2}{2}+\mathcal{O}\left(y^4\right)\;. \tag{2.32} $$

Thus, for both of these functions,

$$(f_{\dots}(y))^n \;=\; 1 + n \cdot y + \left[\binom{n}{2} + \frac{n}{2}\right] \cdot y^2 + \mathcal{O}\left(y^3\right) \;. \tag{2.33}$$

Since y is normally distributed with mean 0 and variance η^2 as defined in (2.6) and since all odd moments of the Gaussian distribution vanish, this means that for both the exponential and the hyperbolic transformation we have

$$\mathsf{E}[(f_{\dots}(y))^n] \;=\; 1 + \left[\binom{n}{2} + \frac{n}{2}\right] \cdot \eta^2 + \mathcal{O}\left(\eta^4\right) \;. \tag{2.34}$$

The implication of (2.34) is that *all* moments of the exponential and the hyperbolic transformation function agree up to order $\mathcal{O}(\eta^3)$.

2.1.3 Correlation

Next we analyse the correlations of the different transformation methods in order to provide further insights about their interconnections. Whilst the moments of the hyperbolic and exponential transformation agree up to third order it turns out that distributions of the linear and hyperbolic transformation are even closer due to their similarity in the upper tail.

Proposition 2.10 The correlation of the *linear* and *hyperbolic transformation* is given by

$$\mathrm{Cor}\left[f_{\mathrm{lin}}(y), f_{\mathrm{hyp}}(y)\right] = \frac{\eta}{\sqrt{\left(1 + 2\eta^2 - 2\eta^2\, U\left(-\frac{1}{2}, 0, \frac{1}{2\eta^2}\right)^2\right)}} \;, \tag{2.35}$$

with $y \sim \mathcal{N}(0, \eta^2)$. Furthermore we have the following limiting behaviour

$$\lim_{\eta\to\infty} \mathrm{Cor}\left[f_{\mathrm{lin}}(y), f_{\mathrm{hyp}}(y)\right] = \sqrt{\frac{\pi}{2(\pi-1)}} \approx 0.85 \;. \tag{2.36}$$

Proof: We already calculated the expectation and the variance of the *linear* and *hyperbolic transformation* in (2.29) and (2.26). Putting together these results we obtain

$$\begin{aligned} \mathrm{Cov}\left[f_{\mathrm{lin}}(y), f_{\mathrm{hyp}}(y)\right] &= \mathsf{E}\left[x\left((x + \sqrt{x^2+1}) - \sqrt{2}\eta\, U\left(-\frac{1}{2}, 0, \frac{1}{2\eta^2}\right)\right)\varphi(x,\eta)\right] \\ &= \eta^2 \;, \end{aligned} \tag{2.37}$$

due to the symmetry of the density φ (2.17) of the normal distribution. Combining (2.30), (2.27) and (2.37) verifies (2.35). Moreover, since

$$\lim_{x\to\infty} U\left(-\frac{1}{2}, 0, \frac{1}{x}\right) = \frac{1}{\pi} \tag{2.38}$$

we obtain (2.36). □

Proposition 2.11 The correlation of the *linear* and *exponential transformation* function is given by

$$\mathrm{Cor}\left[f_{\mathrm{lin}}(y), f_{\exp}(y)\right] = \frac{\eta}{\sqrt{\exp\left(\eta^2\right) - 1}}\,, \tag{2.39}$$

with $y \sim \mathcal{N}(0, \eta^2)$. Therefore we obtain in the limit

$$\lim_{\eta \to \infty} \mathrm{Cor}\left[f_{\mathrm{lin}}(y), f_{\exp}(y)\right] = 0\,. \tag{2.40}$$

Proof: The first moments of the *exponential transformation* are given in equation (2.23) of lemma 2.6. Furthermore we obtain for the covariance of both transformation methods

$$\begin{aligned}
\mathrm{Cov}\left[f_{\mathrm{lin}}(y), f_{\exp}(y)\right] &= \mathsf{E}\left[((1+x)-1)\left(\mathrm{e}^{x} - \mathrm{e}^{\frac{\eta^2}{2}}\right)\varphi(x,\eta)\right] \\
&= \frac{\mathrm{e}^{\frac{\eta^2}{2}}\eta}{\sqrt{2\pi}} \int\limits_{\mathbb{R}} x\mathrm{e}^{-\frac{(x-\eta)^2}{2}}\,\mathrm{d}x \\
&= \mathrm{e}^{\frac{\eta^2}{2}}\eta^2\,. \qquad (2.41)
\end{aligned}$$

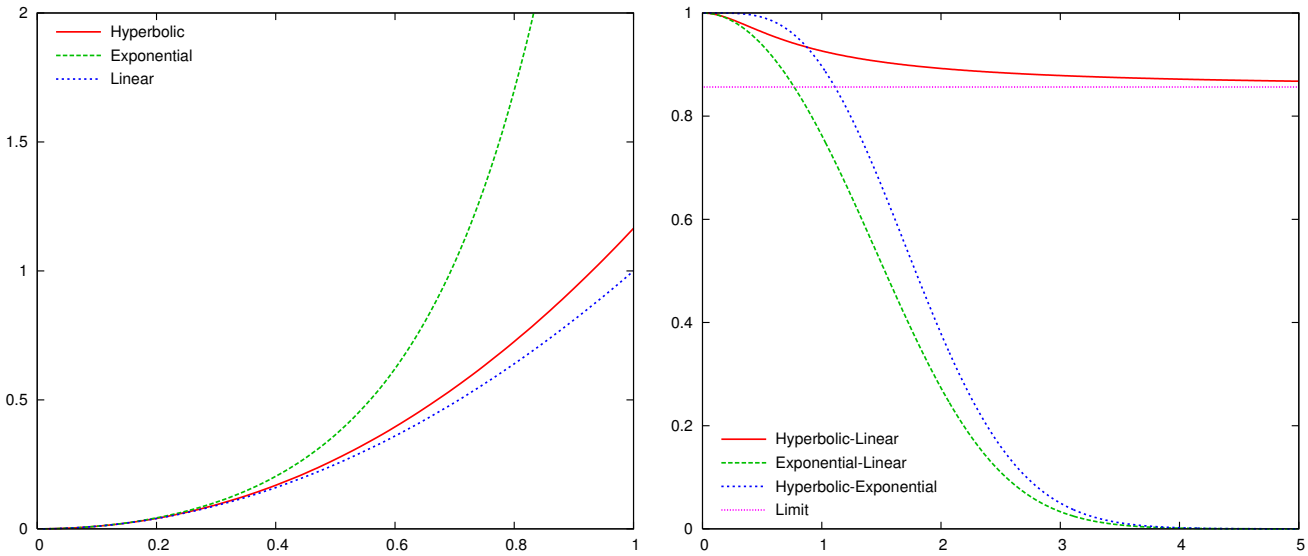
Combining (2.30), (2.23) and (2.41) leads to (2.39). □

Remark 2.12 The correlation of the *hyperbolic* and *exponential transformation* function is computed numerically, since we could not derive an analytical solution.

In figure 2.3, we present the variances and correlations of the different transformation methods. Comparing the variances in figure 2.3 (A) already indicates that the *linear* and *hyperbolic transformation* are more similar with increasing uncertainty η of the normal distribution. This becomes even more evident considering the correlations in figure 2.3 (B), even though for small values of $\eta < 1$ we have the highest linear dependency between the *exponential* and *hyperbolic transformation*. Bearing in mind that η represents the variance of the driving Ornstein-Uhlenbeck process (2.5) and (2.6) we expect that the high correlation of the *linear* and *hyperbolic transformation* reflects in a high codependency of plain vanilla option prices in the *linear* and *hyperbolic* Ornstein-Uhlenbeck model. Since the *linear* Ornstein-Uhlenbeck model admits closed form solutions for European options, we use this model as a control variate for the *hyperbolic* in a Monte Carlo simulation in section 4.1.

2.2 Affine diffusion stochastic volatility models

In this section, we discuss another type of modelling stochastic volatility: affine diffusion models. These models have the appealing feature that the characteristic function of the underlying asset's marginal distribution is known in closed form. This result is not restricted to the modelling of a two-dimensional system of stochastic differential equations but can be generalised to the characterisation of Lévy processes in n-dimensions. Duffie, Pan and Singleton [DPS00] first noticed that the computation of the characteristic function of an



(A) (B)

Figure 2.3: (A) variance and (B) correlation of the different transformation methods (y-axis) over η (x-axis) as given in equation (2.6).

affine diffusion is reduced to solving a system of ordinary differential equations of Riccati type. This result has been generalised to linear-quadratic models by Gaspar [Gas04] and Cheng and Scaillet [CS05]. In general one cannot expect to solve this system of Riccati equation analytically such that solutions have to be computed numerically.

The focus of this section is not to derive another affine diffusion model but to analyse how to implement the standard affine diffusion stochastic volatility models such as Heston [Hes93] and Schöbel-Zhu [SZ99]. Despite the fact that the characteristic function is known in closed form, the implementation is a rather challenging problem since the characteristic function is typically oscillating and involves the evaluation of a complex logarithm which is a multivalued function.

To concretise the characteristic features of an affine diffusion model, we follow the lead by Duffie, Pan and Singleton [DPS00] and Duffie, Filipovic and Schachermeyer [DFS03] and consider the following general diffusion model

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{\mu}(\boldsymbol{X}_t)\,\mathrm{d}t + \boldsymbol{\sigma}(\boldsymbol{X}_t)\,\mathrm{d}\boldsymbol{W}_t\,, \tag{2.42}$$

with $\boldsymbol{X} = (X_1, \ldots, X_n)$ being an n-dimensional stochastic process, $\boldsymbol{\mu} \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n)$, $\boldsymbol{\sigma} \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^{n\times n})$ and $\boldsymbol{W}$ a standard Brownian motion in $\mathbb{R}^n$.

The aim is to show that under certain conditions on the structure of the drift $\boldsymbol{\mu}$ and the diffusion $\boldsymbol{\sigma}$ the characteristic function

$$\phi(\boldsymbol{u}, x, t) := \mathsf{E}\left[\mathrm{e}^{i\cdot\boldsymbol{u}\cdot\boldsymbol{X}_t}\right]\,, \tag{2.43}$$

with $\boldsymbol{u} = (u_1, \ldots, u_n)$ is given as a solution of a system of Riccati equations. By virtue of Feynman-Kac's theorem we know that equations (2.42) and (2.43) are equivalent to solving

$$\frac{\partial\phi(u,x,t)}{\partial t} + \boldsymbol{\mu}(x)\frac{\partial\phi(u,x,t)}{\partial x} + \frac{1}{2}\mathrm{tr}\left[\boldsymbol{\sigma}(x)\cdot\boldsymbol{\sigma}(x)^T\frac{\partial^2\phi(u,x,t)}{\partial x^2}\right] = 0\,, \tag{2.44}$$

where $\mathrm{tr}(A)$ is the trace of a matrix A and $\frac{\partial \phi(u,x,t)}{\partial x} = \left(\frac{\partial \phi}{\partial x_1}, \frac{\partial \phi}{\partial x_2}, \ldots, \frac{\partial \phi}{\partial x_n}\right)$ is the vector of the partial derivatives. The boundary condition of (2.44) is given as

$$\phi(u, x, 0) = \mathrm{e}^{i \cdot u \cdot x} \ . \tag{2.45}$$

where x denotes the initial value of $\boldsymbol{X}$. Next we specify the notation of an affine diffusion:

Definition 2.13 A diffusion process is called affine if the following conditions are satisfied

$$\boldsymbol{\mu}(x) = K_0 + \sum_{i=1}^{n} K_{(1,i)} \cdot x_i \ , \qquad \text{for } K_0 \in \mathbb{R}^n \text{ and } K_{(1,i)} \in \mathbb{R}^n \tag{2.46}$$

and

$$\left(\boldsymbol{\sigma}(x) \cdot \boldsymbol{\sigma}^T(x)\right) = H_0 + \sum_{i=1}^{n} H_{(1,i)} \cdot x_i \ , \qquad \text{for } H_0 \in \mathbb{R}^{n \times n} \text{ and } H_{(1,i)} \in \mathbb{R}^{n \times n} \ . \tag{2.47}$$

The key result of Duffie, Pan and Singleton [DPS00] is that the characteristic function of an affine diffusion model is given as the solution of a system of Riccati equations, which we concretise in the following proposition:

Proposition 2.14 Let $\boldsymbol{X}$ be an affine diffusion with drift and diffusion specified in definition 2.13, then the characteristic function is given as

$$\phi(u, x, t) = \mathrm{e}^{A(t,u) + \boldsymbol{B}(t,u)^T \cdot x} \ , \tag{2.48}$$

where the functions $A \in \mathcal{C}^1(\mathbb{R}, \mathbb{C})$ and $\boldsymbol{B} \in \mathcal{C}^1(\mathbb{R}, \mathbb{C}^n)$ satisfy the complex valued ordinary differential equations

$$\frac{\partial B_i(t)}{\partial t} = K_{(1,i)}^T \cdot \boldsymbol{B}(t) + \frac{1}{2} \boldsymbol{B}(t)^T \cdot H_{(1,i)} \cdot \boldsymbol{B}(t) \ , \tag{2.49}$$

for $i = 1, \ldots, n$ and

$$\frac{\partial A(t)}{\partial t} = K_0^T \cdot \boldsymbol{B}(t) + \frac{1}{2} \boldsymbol{B}(t)^T \cdot H_0 \cdot \boldsymbol{B}(t) \ , \tag{2.50}$$

subject to the boundary conditions $\boldsymbol{B}(0) = i \cdot \boldsymbol{u}$ and $A(0) = 0$.

Proof: Applying the suggested solution (2.48) to the partial differential equation (2.44) and comparing the components with respect to the powers of x verifies this result. □

Note that in the case of stochastic volatility models we are, at least for plain vanilla options, only interested in the characteristic function with respect to the underlying financial driver usually the first element X_1 of the vector valued stochastic process $\boldsymbol{X}$ (2.42). We show in section 2.2.3 that knowing the full multivariate characteristic function (2.43) is particularly usefull for calculating the characteristic function conditionally on a future fixing date.

2.2.1 Characteristic function of affine diffusion stochastic volatility models

The affine diffusion stochastic volatility model we consider here is characterised by the following 2-dimensional system of stochastic differential equations

$$\mathrm{d}S(t) = rS(t)\,\mathrm{d}t + \eta \cdot v(t)^p S(t)\,\mathrm{d}W_t\,, \tag{2.51}$$

$$\mathrm{d}v(t) = \kappa(\theta - v(t))\mathrm{d}t + \omega v(t)^{1-p}\,\mathrm{d}Z_t\,, \tag{2.52}$$

with correlated Brownian motions $\mathrm{d}W_t \cdot \mathrm{d}Z_t = \rho\,\mathrm{d}t$. In order to categorise these models as affine diffusions, we have to transform the underlying asset (2.51) to logarithmic coordinates by virtue of Itô's formula. Setting $p = 0$, we have the standard Black-Scholes [BS73] model, whilst for $p = 1/2$ we obtain the Heston [Hes93] stochastic volatility model. The third model of affine structure is the Stein-Stein [SS91] and Schöbel-Zhu [SZ99] model with $p = 1$. Even though this model is actually non-affine in its original coordinates (2.52), we can augment the state vector by v^2 such that it is affine in $\ln(S)$, v and v^2 as recently became evident from the study of quadratic models by Gaspar [Gas04] and Cheng and Scaillet [CS05]. The latter paper showed that any linear-quadratic jump-diffusion model is equivalent to an affine jump-diffusion model with an extended state vector.

Using the previous results about affine diffusion models, we now derive the univariate characteristic functions of the underlying financial asset

$$\phi(u, x, t) := \mathsf{E}\left[\mathrm{e}^{i \cdot u \cdot \ln F_t}\right]\,, \tag{2.53}$$

where $F_t = S_0 \mathrm{e}^{rt}$ denotes the forward price of the underlying asset. We start with the easiest example, the Black-Scholes model [BS73], where we even know the distribution density of the underlying asset in closed form.

Lemma 2.15 The Black-Scholes model is affine in $X_t = \ln F_t$,

$$\mathrm{d}X_t = -\frac{1}{2}\eta^2\,\mathrm{d}t + \eta\,\mathrm{d}W_t\,, \tag{2.54}$$

with $K_0 = -\frac{1}{2}\eta^2$, $K_{(1,1)} = 0$, as well as $H_0 = \eta^2$ and $H_{(1,1)} = 0$, leading to the following system of ordinary differential equations

$$\frac{\partial B(t)}{\partial t} = 0\,, \tag{2.55}$$

$$\frac{\partial A(t)}{\partial t} = -\frac{1}{2}\eta^2 B(t) + \frac{1}{2}\eta^2 B(t)^2\,, \tag{2.56}$$

subject to the boundary conditions $B(0) = iu$ and $A(0) = 0$. The solution can be derived as

$$B(t) = iu\,, \tag{2.57}$$

$$A(t) = \left(-\frac{1}{2}\eta^2 u(u+i)\right) t\,, \tag{2.58}$$

such that the characteristic function is given by

$$\phi(u, x, t) = \mathrm{e}^{iu \cdot x - \frac{1}{2}\eta^2 u(u+i)t}\,. \tag{2.59}$$

Next we show how to derive the characteristic function of the Heston model.

Lemma 2.16 The Heston model is affine in $X_t = \ln F_t$ and v_t,

$$\mathrm{d}X_t = -\frac{1}{2} v_t \,\mathrm{d}t + \sqrt{v_t}\,\mathrm{d}W_t \,, \tag{2.60}$$

$$\mathrm{d}v_t = \kappa(\theta - v_t)\,\mathrm{d}t + \omega\sqrt{v_t}\,\mathrm{d}Z_t \,, \tag{2.61}$$

with the auxiliary variables given by

$$K_0 = \begin{pmatrix} 0 \\ \kappa\theta \end{pmatrix} \,, \qquad K_{(1,1)} = 0 \,, \qquad K_{(1,2)} = \begin{pmatrix} -\frac{1}{2} \\ -\kappa \end{pmatrix} \,, \tag{2.62}$$

as well as

$$H_0 = 0 \,, \qquad H_{(1,1)} = 0 \,, \qquad H_{(1,2)} = \begin{pmatrix} 1 & \rho\omega \\ \rho\omega & \omega^2 \end{pmatrix} \,. \tag{2.63}$$

Thus we have to solve the following system of ordinary differential equations

$$\frac{\partial B_1(t)}{\partial t} = 0 \,, \tag{2.64}$$

$$\frac{\partial B_2(t)}{\partial t} = \hat{\alpha}(u) - \beta(u) B_2(t) + \gamma B_2(t)^2 \,, \tag{2.65}$$

$$\frac{\partial A(t)}{\partial t} = \kappa\theta B_2(t) \tag{2.66}$$

subject to the boundary conditions $B_1(0) = iu$, $B_2(0) = 0$ and $A(0) = 0$. The auxiliary variables we introduced are

$$\hat{\alpha}(u) = -\frac{1}{2} u(i+u) \,, \quad \beta(u) = \kappa - \rho\omega u i \quad \text{and} \quad \gamma = \frac{1}{2}\omega^2 \,. \tag{2.67}$$

The solution can be derived as

$$B_1(t) = iu \,, \tag{2.68}$$

$$B_2(t) = \frac{\beta(u) + D(u)}{\omega^2} \left(\frac{\mathrm{e}^{D(u)t} - 1}{c(u)\mathrm{e}^{D(u)t} - 1} \right) \,, \tag{2.69}$$

$$A(t) = \frac{\kappa\theta}{\omega^2} \left((\beta(u) + D(u))\, t - 2 \ln \left(\frac{c(u)\mathrm{e}^{D(u)t} - 1}{c(u) - 1} \right) \right) \,, \tag{2.70}$$

where

$$c(u) = \frac{\beta(u) + D(u)}{\beta(u) - D(u)} \,, \tag{2.71}$$

and

$$D(u) = \sqrt{\beta(u)^2 - 4\hat{\alpha}(u)\gamma} \,. \tag{2.72}$$

We further refer to the reciprocal of c as

$$G(u) = \frac{\beta(u) - D(u)}{\beta(u) + D(u)} \,. \tag{2.73}$$

Remark 2.17 The Riccati equation for B_2 is essential when analysing the problem of moment explosion of the characteristic function as first observed by Andersen and Piterbarg [AP04].

Proof:[Lemma 2.16] According to

$$\boldsymbol{\sigma} \cdot \boldsymbol{\sigma}^T = \begin{pmatrix} v & \rho\omega \cdot v \\ \rho\omega \cdot v & \omega^2 \cdot v \end{pmatrix} , \tag{2.74}$$

we can directly compute H_0, $H_{(1,1)}$ and $H_{(1,2)}$, and by the aid of proposition 2.14 set up the system of Riccati equations. The solution of the first differential equation (2.68) is obvious. Recasting the second Riccati equation (2.69) as

$$\frac{\partial B_2(t)}{\partial t} = \hat{c} \cdot (B_2 - a)(B_2 - b) , \tag{2.75}$$

immediately leads to the following solution

$$B_2(t) = ab\frac{\mathrm{e}^{(a-b)\hat{c}t} - 1}{a\mathrm{e}^{(a-b)\hat{c}t} - b} , \tag{2.76}$$

where the roots of the Riccati equation are given by

$$a = \frac{(\beta(u) + D(u))}{\omega^2} , \quad b = \frac{(\beta(u) - D(u))}{\omega^2} , \quad \hat{c} = \gamma , \tag{2.77}$$

with

$$D(u) = \sqrt{\beta(u)^2 - 4\hat{\alpha}(u)\gamma} . \tag{2.78}$$

Hence we can rewrite the solution as

$$B_2(t) = \frac{\beta(u) + D(u)}{\omega^2}\frac{\mathrm{e}^{D(u)t} - 1}{c(u)\mathrm{e}^{D(u)t} - 1} . \tag{2.79}$$

The solution for A now follows immediately

$$\begin{aligned} \int_0^t B_2(s)\,\mathrm{d}s &= \frac{\beta(u) + D(u)}{\omega^2}\int_0^t \frac{\mathrm{e}^{D(u)s} - 1}{c(u)\mathrm{e}^{D(u)s} - 1}\,\mathrm{d}s \\ &= \omega^{-2}\left((\beta(u) + D(u))t - 2\ln\left(\frac{c(u)\mathrm{e}^{D(u)t} - 1}{c(u) - 1}\right)\right) , \end{aligned} \tag{2.80}$$

which completes the proof. □

Remark 2.18 Evaluation of the characteristic function in the Heston model mainly involves two different kind of problems. The first is the fact that the characteristic function is typically of oscillatory nature, a problem which we will discuss in greater length in chapter 5. The second problem is that the characteristic function requires the evaluation of a complex logarithm in equation (2.80) which is a multivalued function. In the next section, we show

that continuity of the characteristic function can be preserved by either applying the so called *rotation count algorithm* as introduced by Kahl and Jäckel [KJ05] or by rearranging terms in the representation of the characteristic function as suggested by Duffie, Pan and Singleton [DPS00], Schoutens, Simons and Tistaert [SST04] and Gatheral [Gat06], such that taking the principal branch of the complex logarithm guarantees continuity as outlined and discussed by Lord and Kahl [LK06, LK07a] and Albrecher, Mayer, Schoutens and Tistaert [AMST06].

Remark 2.19 It is numerically more stable to rewrite (2.69) as

$$B_2(t) = \frac{\beta(u) - D(u)}{\omega^2} \left(\frac{1 - \mathrm{e}^{-D(u)t}}{1 - G(u)\mathrm{e}^{-D(u)t}} \right) , \tag{2.81}$$

since we assume that $\mathrm{Re}(D(u)) > 0$, such that the exponential terms can cause numerical difficulties as $\lim_{u\to\infty} |D(u)| = \infty$, which is shown in the proof of proposition 5.10.

The characteristic function of the Schöbel-Zhu model can be deduced in an analogous manner. Despite the fact that the original approach of Schöbel-Zhu used different arguments to derive the characteristic function, it turns out that the affine diffusion approach allows to find a representation which is similar to the characteristic function of the Heston model. Thus solving the problem of complex discontinuities in the Heston models also solves this problem for the Schöbel-Zhu model.

Lemma 2.20 The Schöbel-Zhu model is affine in $X_t = \ln F_t$, v_t and $z_t = v_t^2$,

$$\mathrm{d}X_t = -\frac{1}{2} z_t \,\mathrm{d}t + \sqrt{z_t}\,\mathrm{d}W_t , \tag{2.82}$$

$$\mathrm{d}v_t = \kappa(\theta - v_t)\,\mathrm{d}t + \omega\,\mathrm{d}Z_t , \tag{2.83}$$

$$\mathrm{d}z_t = \left(-2\kappa z_t + 2\kappa\theta v_t + \omega^2\right)\mathrm{d}t + 2\omega\sqrt{z_t}\,\mathrm{d}Z_t . \tag{2.84}$$

We can compute the necessary auxiliary parameters to calculate the system of ordinary differential equations

$$K_0 = \begin{pmatrix} 0 \\ \kappa\theta \\ \omega^2 \end{pmatrix} , \quad K_{(1,1)} = 0 , \quad K_{(1,2)} = \begin{pmatrix} 0 \\ -\kappa \\ 2\kappa\theta \end{pmatrix} , \quad K_{(1,3)} = \begin{pmatrix} -\frac{1}{2} \\ 0 \\ -2\kappa \end{pmatrix} , \tag{2.85}$$

as well as

$$H_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \omega^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} , \; H_{(1,2)} = \begin{pmatrix} 0 & \rho\omega & 0 \\ \rho\omega & 0 & 2\omega^2 \\ 0 & 2\omega^2 & 0 \end{pmatrix} , \; H_{(1,3)} = \begin{pmatrix} 1 & 0 & 2\omega\rho \\ 0 & 0 & 0 \\ 2\omega\rho & 0 & 4\omega^2 \end{pmatrix} \tag{2.86}$$

and $H_{(1,1)} = 0$. Hence the system of ordinary differential equation is given by

$$\frac{\partial B_1(t)}{\partial t} = 0 \,, \tag{2.87}$$

$$\frac{\partial B_2(t)}{\partial t} = 2\kappa\theta B_3(t) + \left(\gamma B_3(t) - \frac{1}{2}\beta(u)\right) \cdot B_2(t) \,, \tag{2.88}$$

$$\frac{\partial B_3(t)}{\partial t} = \hat{\alpha}(u) - \beta(u) B_3(t) + \gamma B_3(t)^2 \,, \tag{2.89}$$

$$\frac{\partial A(t)}{\partial t} = \kappa\theta B_2(t) + \frac{1}{2}\omega^2 B_2(t)^2 + \omega^2 B_3(t) \,, \tag{2.90}$$

subject to the boundary conditions $\boldsymbol{B}(0) = (iu, 0, 0)$. The auxiliary variables we introduced are

$$\hat{\alpha} = -\frac{1}{2}u(i+u) \,, \qquad \beta(u) = 2(\kappa - \rho\omega u i) \qquad \text{and} \quad \gamma = 2\omega^2 \,. \tag{2.91}$$

Remark 2.21 The system of ordinary differential equations in the Schöbel-Zhu model is closely related to the Heston one. Following remarks by both Heston and Schöbel-Zhu, we know that when $\theta = 0$, the Schöbel-Zhu model collapses to a particular instance of the Heston model as can be see from equation (2.84) - the variance then has a mean-reversion speed of 2κ, a volatility of variance equal to 2ω and a mean-reversion level of $\omega^2/2\kappa$.

Proposition 2.22 If we denote the characteristic function of the Heston model by ϕ_{Heston} the Schöbel-Zhu characteristic function becomes

$$\phi_{\text{SZ}}(u, S_0, \sigma_0, \kappa, \theta, \omega, \rho, t) = \phi_{\text{Heston}}(u, S_0, \sigma_0^2, 2\kappa, \omega^2/2\kappa, 2\omega, \rho, t) \cdot \mathrm{e}^{A_z + B_2 \cdot v(0)} \,, \tag{2.92}$$

where A_z follows the ordinary differential equation

$$\frac{\partial A_z(t)}{\partial t} = \kappa\theta B_2(t) + \frac{1}{2}\omega^2 B_2(t)^2 \,. \tag{2.93}$$

Tedious though straightforward manipulations show that A_z and B_2 can be solved in closed form as:

$$B_2(t) = \kappa\theta \frac{\beta(u) - D(u)}{D(u)\omega^2} \frac{\left(1 - \mathrm{e}^{-\frac{1}{2}D(u)t}\right)}{1 - G(u)\mathrm{e}^{-D(u)t}} \,, \tag{2.94}$$

$$A_z(t) = \frac{(\beta(u) - D(u))\kappa^2\theta^2}{2D(u)^3\omega^2} \left(\beta(u)(D(u)t - 4) + D(u)(D(u)t - 2) + \hat{A}(t)\right) , \tag{2.95}$$

$$\hat{A}(t) = \frac{4\mathrm{e}^{-\frac{1}{2}D(u)t}\left(\frac{D(u)^2 - 2\beta(u)^2}{\beta(u) + D(u)}\mathrm{e}^{-\frac{1}{2}D(u)t} + 2\beta(u)\right)}{1 - G(u)\mathrm{e}^{-D(u)t}} \,. \tag{2.96}$$

Proof: The solution for B_2 can be directly verified by inserting (2.94) into the defining ordinary differential equation (2.88), whilst the proof of A_z requires to integrate B_2 over t due to equation (2.93). □

Remark 2.23 By recognising a special case of the Heston model within the Schöbel-Zhu model it is immediately clear that the discontinuities can be avoided in the same way as in the Heston model. Moreover we can also link the asymptotics of both models using the relationship (2.92).

2.2.2 Continuous complex logarithm

An easy way to avoid the problem of complex discontinuities is to integrate the ordinary differential equations in (2.66) numerically, as this would automatically lead to the correct and continuous solution. A comparative advantage of the Heston and Schöbel-Zhu model is, that it has a closed-form characteristic function, something which distinctly reduces the computational effort. From a computational point of view, it is therefore certainly worthwhile to investigate how to avoid discontinuities in the closed-form solution. By taking a closer look at the characteristic function, it is clear that two multivalued functions are present, both of which could cause complex discontinuities. The first candidate is the square root used in $D(u)$ as defined in equation (2.78). It turns out that the characteristic function is even in D, so that we will from here on use the common convention that the real part of the square root is nonnegative. The complex logarithm used in equation (2.80) is the second candidate.

Definition 2.24 Let $z = a + ib$ a complex variable with $a, b \in \mathbb{R}$. Throughout the following, we denote the real and imaginary part $z_r = \mathrm{Re}(z) = a$ and $z_i = \mathrm{Im}(z) = b$. Furthermore let $z = re^{i\theta}$ be the representation of a complex variable in polar coordinates. We use the notation $z_r = |z| = r$ for the radius and $z_\theta = \arg(z) = \theta$ for the phase or argument.

Definition 2.25 The logarithm of a complex variable $z = a+ib = re^{i(t+2\pi n)}$ with $t \in [-\pi, \pi)$, $r \in \mathbb{R}^+$ and $n \in \mathbb{Z}$ is defined as

$$\ln z = \ln(r) + i(t + 2\pi n) \; . \tag{2.97}$$

In this case the branch cut of the complex logarithm is $(-\infty, 0]$, and the logarithm is discontinuous along it.

Since there are numerous ways to rewrite solution (2.80), we only consider these two in the following:

Definition 2.26 The original formula of Heston, which we derived in proposition 2.16 is referred to as **formulation 1**:

$$A(t,u) = \frac{\kappa\theta}{\omega^2}\left(\left(\beta(u) + D(u)\right)t - 2\ln\left(\Psi_1\left(t,u\right)\right)\right) \; , \quad \Psi_1(t,u) = \frac{c(u)\mathrm{e}^{D(u)t} - 1}{c(u) - 1} \; . \tag{2.98}$$

Definition 2.27 We refer to **formulation 2** as the numerically more stable representation

$$A(t,u) = \frac{\kappa\theta}{\omega^2}\left(\left(\beta(u) - D(u)\right)t - 2\ln\left(\Psi_2\left(t,u\right)\right)\right) \; , \quad \Psi_2(t,u) = \frac{G(u)\mathrm{e}^{-D(u)t} - 1}{G(u) - 1} \; . \tag{2.99}$$

Remark 2.28 Kahl and Jäckel developed the so called *rotation count algorithm* to preserve continuity of the complex logarithm in (2.98) whilst (2.99) was used in Duffie, Pan and Singleton [DPS00], Schoutens, Simons and Tistaert [SST04] and Gatheral [Gat05].

To obtain a first impression of the problem of discontinuities of the complex logarithms, we show the trajectory of $\Psi_1(t,u)$ in the complex plane as u varies from 0 to ∞ in figure 2.4.

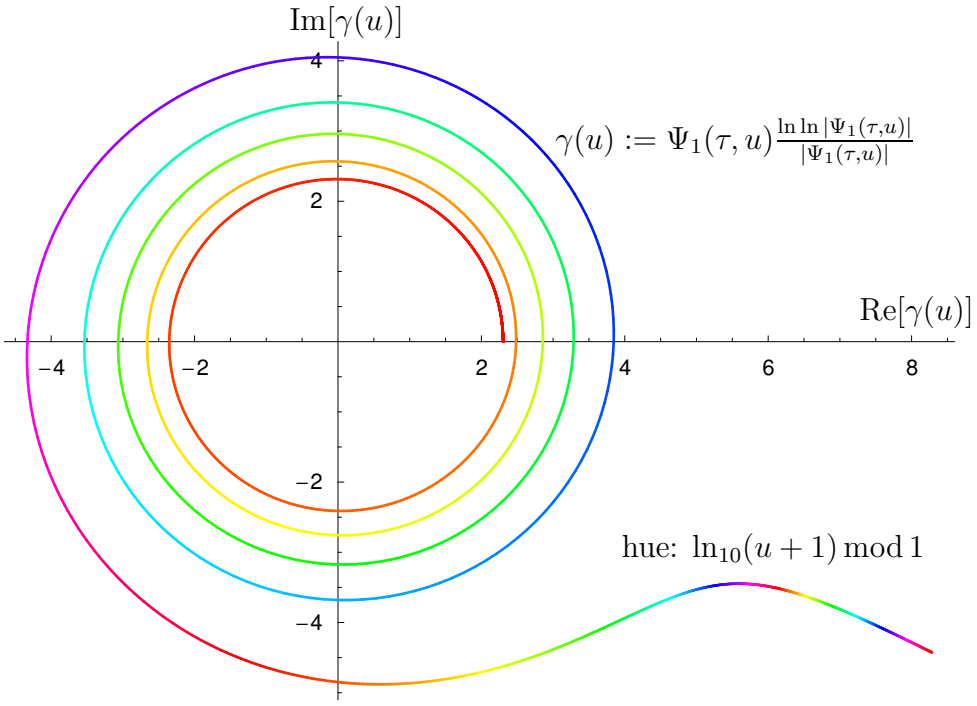


Figure 2.4: Part of the trajectory of the function $\Psi_1(\tau,u)$ (2.98) in the complex plane has the structural shape of a spiral which gives rise to repeated crossing of the negative real axis. Here: $\rho = -0.8$, $\kappa = 1$, $\omega = 2$, time to maturity $\tau = 10$ and $u \in [10^{-6}, 10^3]$.

Every crossing of the negative real axis corresponds to a discontinuity of the characteristic function as shown in figure 2.5.

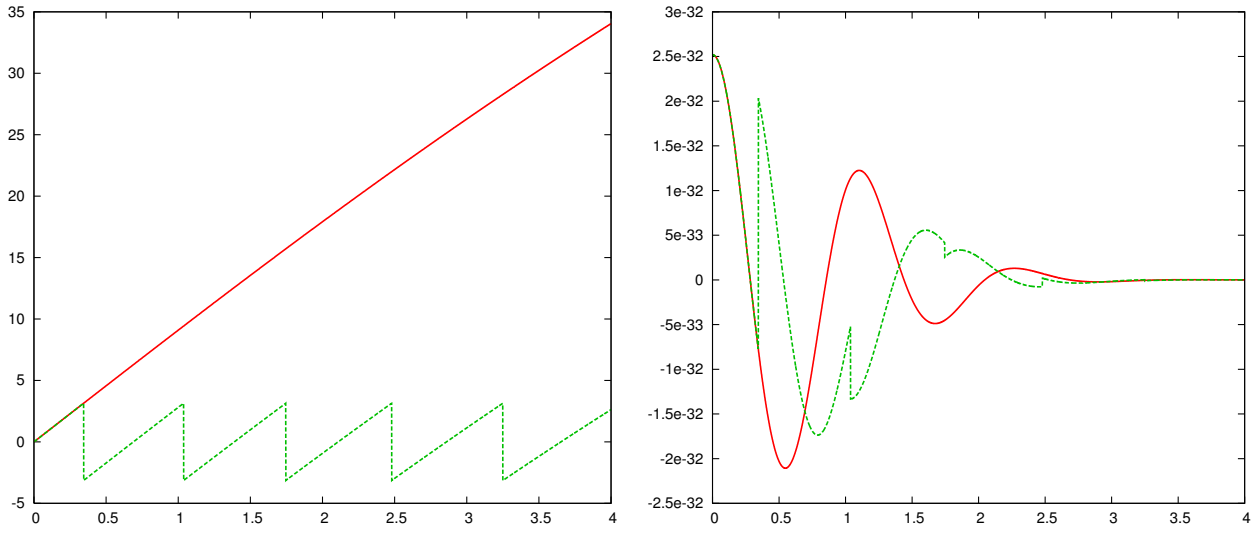
The problem of complex discontinuities was first encountered by Schöbel and Zhu. They mention: "*Therefore we implemented our formula carefully keeping track of the complex logarithm along the integration path. This leads to a smooth CF ...*". This is an involved technical procedure which would make the implementation of this model rather complicated.

Fortunately, there is a much simpler procedure to guarantee the continuity of $\Psi_1(t,u)$: the *rotation count algorithm*, which we present in the following. To show the full capability of this algorithm we keep the notation as general as possible. Therefore let $a, b \in \mathcal{C}(\mathbb{R}, \mathbb{C})$ continuous complex functions and $c \in \mathbb{R}$ such that

$$d(u) = a(u)\mathrm{e}^{b(u)} + c\,. \tag{2.100}$$

One can think of the function d as either the numerator or denominator of the critical function Ψ_1 in equation (2.98). So let us state the main principle to preserve continuity:

Proposition 2.29 Algorithm 1 on page 23 is able to preserve continuity if the function $a(u)\mathrm{e}^{b(u)}$ does not cross the real axis in $[\min(0,-c), \max(0,-c)]$.



(A) (B)

Figure 2.5: Complex discontinuities in the Heston model. Parameters taken from Duffie et al. [DPS00] implied from market data of S&P500 index options: $\kappa = 6.21$, $\omega = 0.61$, $\rho = -0.7$, $\theta = 0.019$, $v(0) = 0.010201$, $F = 1$ and $\tau = 20$. (A): The principal argument/phase of Ψ_1 (2.98) with (red solid line) and without (green dashed line) correction. (B): Characteristic function with and without correction.

Require: $a, b \in \mathbb{C}$, $c \in \mathbb{R}$
1: Calculate the phase interval of ae^b as $n = \text{int}\left[\frac{1}{2\pi}\left(a_\theta + b_i + \pi\right)\right]$, equal to 0 if it is in $[-\pi, \pi)$
2: set $d_r = |ae^b + c|$
3: set $d_\theta = \arg\left(ae^b + c\right) + 2\pi n$ where arg denotes the principal argument

Algorithm 1: Rotation count algorithm.

Proof: Using

$$a(u)\mathrm{e}^{b(u)} = a_r \mathrm{e}^{b_r} \mathrm{e}^{i(a_\theta + b_i)} \,, \tag{2.101}$$

we see that if $a_\theta + b_i \notin [-\pi, \pi)$ the principal argument of ae^b is not the correct one. Thus we have to correct it as done in step 1 of algorithm 1. □

Corollary 2.30 One can compute the complex logarithm of the function Ψ_1 in equation (2.98) by setting

$$d_1(u,t) = c(u)\mathrm{e}^{D(u)t} - 1 \,, \tag{2.102}$$

$$d_2(u) = c(u) - 1 \,, \tag{2.103}$$

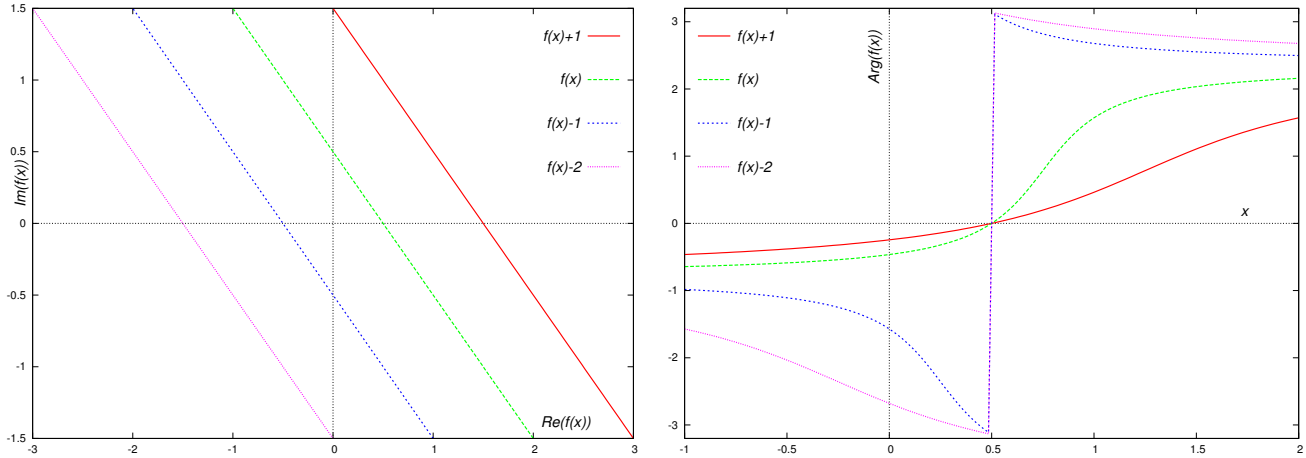
such that

$$\ln\left(\Psi_1(t,u)\right) = \ln\left(\frac{|d_1(u,t)|}{|d_2(u)|}\right) + i\left(d_{1,\theta}(u,t) - d_{2,\theta}(u)\right) \,. \tag{2.104}$$

The remaining problem is now to check that the premise of the *rotation count algorithm* is true. Thus we have to verify that neither the function $c(u)\mathrm{e}^{D(u)t}$ nor $c(u)$ do cross the real axis in the interval $[0,1]$. In order to visualise the problem of adding a constant, we consider the following complex function

$$f(x) = (1-x) - \left(\frac{1}{2} - x\right) i \,, \tag{2.105}$$

where $x \in \mathbb{R}$. Figure 2.6 compares the graphs of $f(x)+1$, $f(x)$, $f(x)-1$ and $f(x)-2$ in the complex plane, as well as their principal arguments:



(A) (B)

Figure 2.6: Figure (A) shows the real and imaginary part of the function (2.105). In figure (B) the argument of all functions is compared. The function $f(x)$ crosses the real axis in $[0,1]$ thus we recognise a jump in the difference of $\arg(f(x)) - \arg(f(x)-1)$ for $x = 1/2$.

Whilst $f(x)+1$ and $f(x)$ do have a continuous principal argument, we obtain a jump in the argument of $f(x)-1$ and $f(x)-2$. The lesson from this stylised example is clear: subtracting 1 can cause a discontinuity in the phase. So, if the *rotation count algorithm* applied to the evaluation of $\Psi_1(t,u)$ does not yield any discontinuities, it must be that subtracting 1 from the numerator (2.102) and denominator (2.103) indeed does not change the phase of both $c(u)$ and $c(u)\mathrm{e}^{D(u)t}$. From the example we can conclude that these discontinuities can only arise when $c(u)$ or $c(u)\mathrm{e}^{D(u)t}$ cross the imaginary axis, and their real parts are in the interval $[0,1]$. The following observation is essential that this cannot happen.

Lemma 2.31 If

$$\rho \le \kappa/\omega \,, \quad \text{or} \quad \mathrm{Im}(u) \ge -\kappa/(\rho\omega) \ \text{ and } \ \kappa/\omega < \rho \le 2\kappa/\omega \,, \tag{2.106}$$

we have

$$|c(u)| = \left|\frac{\beta(u) + D(u)}{\beta(u) - D(u)}\right| \ge 1 \,, \tag{2.107}$$

where the inequality is strict, except where $\mathrm{Re}(u) = 0$ or $\rho = 2\kappa/\omega$.

Proof: The proof can be found in the appendix B. □

We only need one additional lemma before we can prove the main theorem of this section:

Lemma 2.32 It is only possible that

$$c(u)\mathrm{e}^{D(u)t} = 1 \,, \quad \text{or } c(u) = 1 \tag{2.108}$$

if $D(u) = 0$ which further requires $\mathrm{Re}(u) = 0$. In this case neither Ψ_1 nor Ψ_2 will cross the negative real line and should be evaluated as

$$\lim_{D(u)\to 0} \Psi_1(t,u) = \lim_{D(u)\to 0} \Psi_2(t,u) = 1 + \beta(u)t \,. \tag{2.109}$$

Proof: Equation (2.109) can be verified by applying l'Hôpital's rule. □

Theorem 2.33 The *rotation count algorithm* can be applied to the characteristic function of the Heston model for the parameter configuration (2.106) given in lemma 2.31 where we exclude the case $\rho = 2\kappa/\omega$.

Proof: Since we use the convention that the real part of $D(u)$ is positive, we obtain

$$\left|c(u)\mathrm{e}^{D(u)t}\right| = |c(u)|\,\mathrm{e}^{D(u)t} \geq |c(u)| \geq 1 \,, \tag{2.110}$$

where the first equality directly follows from the convention that the real part of D is positive and $t > 0$. If the inequality is strict none of the critical terms can cross the real axis in $[0, 1]$. Since we excluded the case $\rho = 2\kappa/\omega$, we only have to consider $\mathrm{Re}(u) = 0$. Moreover we see from the definition of c in (2.71) that $|c(u)| = 1$ implies $D(u) = 0$. Thus we have to evaluate $\Psi_1(u,t)$ as in equation (2.109). □

Remark 2.34 Calibrating the Heston model to a market given implied volatility surface, the restriction (2.106) is almost always satisfied, i.e. for equity markets one typically finds $-1 < \rho < -0.7$. In the interest rate setting the correlation ρ is either negative, in order to match the downward sloping implied volatility skew, or it is set to zero to simplify the change of measure[1]. Extensive numerical tests suggest that the *rotation count algorithm* works even if the restriction (2.106) is not satisfied.

Even though theorem 2.33 verified the *rotation count algorithm* applied to the formulation 1, we still have to discuss formulation 2. It turns out that within this representation of the characteristic function the principal argument of the complex logarithm is always the correct one and we do not have to apply the *rotation count algorithm* at all.

Theorem 2.35 If the requirements (2.106) of lemma 2.31 are satisfied, the principal argument of the complex logarithm in formulation (2.99) is the right one.

[1] See e.g. footnote 1 in Andersen and Andreasen [AA02].