UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE MATEMÁTICA

ISCTE
INSTITUTO UNIVERSITÁRIO DE LISBOA
DEPARTAMENTO DE FINANÇAS



# PRICING AMERICAN-STYLE OPTIONS UNDER THE HESTON MODEL USING THE COS APPROXIMATION

Miguel Ângelo Vieira da Piedade

**Mestrado em Matemática Financeira**

Dissertação orientada por:
Professor Doutor João Pedro Vidal Nunes

2020

# Acknowledgements

First and foremost, I would like to express my gratitude to Professor João Pedro Nunes, for his availability on supervising this thesis and for the valuable guidance provided by him throughout these months.

I would also like to show my deepest appreciation for the unconditional love and support from my mother.

Last but not least, I would like to thank my housemates Bruna Rodrigues, Miguel Vicente and Tânia Castillo, for the companionship, encouragement and support that we've shared during these times.

# Resumo

Nesta tese, apresenta-se o método COS: um método numérico para avaliação de opções, baseado na expansão em série de cossenos de Fourier. Foi proposto por Fang e Oosterlee [10] [11], e pode ser eficazmente utilizado para avaliar opções de estilo Bermudiano subjacentes a processos de Lévy ou ao modelo de volatilidade estocástica de Heston.

O ingrediente principal do método é a relação próxima entre a função característica e os coeficientes da expansão em série de cossenos de Fourier da função densidade.

Ao considerar o modelo de Heston, o problema de avaliação bidimensional é abordado combinando a ideia anterior com regras de quadratura de ordem superior no domínio da log-variância.

**Palavras-chave**: Exercício antecipado, séries de cossenos de Fourier, transformada rápida de Fourier, processos de Lévy, modelo de Heston.

# Abstract

In this thesis, we present the COS method: an option pricing numerical method based on Fourier cosine series expansions. It was proposed by Fang and Oosterlee [10] [11], and it can be used to effectively price Bermudan options under Lévy processes or the Heston stochastic volatility model.

The method's key ingredient is the close relationship between the characteristic function and the series coefficients of the Fourier cosine expansion of the density function.

Under the Heston model, the two-dimensional pricing problem is dealt by combining the prior Fourier cosine series insight with high-order quadrature rules in the log-variance dimension.

**Keywords**: Early-exercise, Fourier cosine series, fast Fourier transform, Lévy processes, Heston model.

# Contents

# List of Figures

# List of Tables

# 1. Introduction

The American option pricing problem has been studied intensively over the last decades, giving rise to a myriad of literature focused on the subject. This is a matter of special interest for financial institutions, mainly due to calibration purposes: once model prices are fitted to market prices, model parameters can be extrapolated and used to price more complex instruments that are not priced on the market. Other possible applications for option pricing models are the computation of hedging strategies and the quantification of risk associated to different positions.

Up to this date, proposed methods to solve this problem can be categorized into three groups: Monte Carlo simulation, partial-integro differential equation (PIDE) methods, and numerical integration methods (also often referred to as transform methods). Every method has its own merits and demerits, depending on the financial model that is being considered and on the option's features.

Monte Carlo simulation methods consist in simulating many paths for the state variables and evaluate the option payoff for each path. Discounting and averaging over all paths gives an estimate for the derivative price. While this methodology allows to handle analytically intractable problems, its biggest disadvantage is the computational requirements involved.

PIDE methods rely on transforming the pricing problem from a discounted risk-neutral expectation to an equivalent partial differential equation, through the Feynman-Kac repre-

sentation. Numerical methods are applied to find solutions to these PIDE's, as closed-form solutions are not available in the vast majority of cases.

Finally, numerical integration methods usually involve a transformation of the problem to the Fourier domain. The Carr and Madan method [6] is one of the best known examples of this class and, more recently, Fang and Oosterlee proposed the COS method [9] [10] [11], which is based on the Fourier cosine series expansion. Together with the use of a fast Fourier transform (FFT) algorithm, this family of methods proves to be very efficient when pricing early-exercise options, making them state-of-the-art for calibration purposes at financial institutions.

## 1.1. Overview of this thesis

This thesis is structured as follows. In Subsection 1.2, we introduce some notation used in latter sections and formalize the pricing problem.

In Section 2, we present the financial models and respective characteristic functions that will be latter used in computational experiments.

In Section 3, we present the COS method under Lévy processes. A pricing formula is derived and a computationally efficient FFT-based algorithm is presented.

In Section 4, the COS method under the Heston model is introduced. Similarly to the previous section, a pricing formula and an efficient algorithm are presented.

In Section 5, results for the conducted experiments are discussed. Convergence rates are compared between models and different approaches are tested when pricing American-style options.

Finally, Section 6 outlines a brief summary of this thesis and discusses future work possibilities.

## 1.2. The pricing problem

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$ be a filtered probability space, where the filtration $\mathbb{F} = \{\mathcal{F}(t) : t \geq 0\}$ satisfies the usual conditions and $\mathbb{Q}$ is a risk-neutral measure, as defined in Definition A.2.17. The stochastic processes $\{S(t) : t \geq 0\}$, $\{u(t) : t \geq 0\}$, $\{r(t) : t \geq 0\}$ are $\mathbb{F}$-adapted and represent the evolution in time of the underlying asset price, log-variance of the underlying asset and risk-free interest rate, respectively. Consider a Bermudan option with strike price $K$ and maturity $T$, whose set of exercise dates is defined by $\mathcal{T} = \{t_1, ..., t_M\}$, with $0 = t_0 < t_1 < ... < t_M = T$, and whose exercise value at time $t_m \in \mathcal{T}$ is given by $g(x_m)$, where $x_m := x(t_m) = \ln(S(t_m)/K)$. For simplification purposes, we also introduce the notation $\mathcal{F}_m := \mathcal{F}(t_m)$, $u_m := u(t_m)$ and $r_m := r(t_m)$. Denote the state vector at time $t_m$ by $\mathbf{s_m} = [x_m, \mathbf{s'_m}]$, where $\mathbf{s'_m}$ is a vector of other (stochastic non-constant) market state variables. The time-$t_m$ value of the option contract is equal to

$$v(\mathbf{s_m}, t_m) = \begin{cases} g(x_m), & m = M, \\ \max\{c(\mathbf{s_m}, t_m), g(x_m)\}, & m = M-1, ..., 1 \\ c(\mathbf{s_m}, t_m), & m = 0. \end{cases} \tag{1.1}$$

where

$$g(x) = \max\{\alpha K e^x - \alpha K, 0\} =: \alpha K (e^x - 1)^+, \quad \alpha = \begin{cases} 1, & \text{for a call,} \\ -1, & \text{for a put,} \end{cases} \tag{1.2}$$

and $c(\mathbf{s_m}, t_m)$ is the continuation value at time $t_m$, i.e.

$$c(\mathbf{s_m}, t_m) = \mathbb{E}^{\mathbb{Q}} \left[ \exp\left( - \int_{t_m}^{t_{m+1}} r(s) ds \right) v(\mathbf{s_{m+1}}, t_{m+1}) | \mathcal{F}_m \right]. \tag{1.3}$$

The Bermudan option price at the initial time is given by $v(\mathbf{s_0}, 0)$ and it can be determined

via backward recursion, as described in Algorithm 1.

**Algorithm 1:** Bermudan option valuation process

1 $v(\mathbf{s_M}, t_M) = g(x_M)$;

2 **for** $m = M-1, ..., 1$ **do**

3 $\quad c(\mathbf{s_m}, t_m) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_{t_m}^{t_{m+1}} r(s)ds\right) v(\mathbf{s_{m+1}}, t_{m+1})|\mathcal{F}_m\right]$;

4 $\quad v(\mathbf{s_m}, t_m) = \max\{c(\mathbf{s_m}, t_m), g(x_m)\}$

5 $v(\mathbf{s_0}, 0) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_0^{t_1} r(s)ds\right) v(\mathbf{s_1}, t_1)|\mathcal{F}_0\right]$

Finding a solution to the previous problem requires repeated evaluation of the continuation value. In order to find this value, we need beforehand a mathematical description of how market state variables are assumed to behave through time. Under many frameworks, these state variables cannot be described in terms of a probability density function, as this function contains many complicated functions or is simply not available in closed-form. On the other hand, the characteristic function is often available and has a simpler closed-form expression. Therefore, as a first step to solve this problem, our focus should be on determining the characteristic function of the market state variables under a variety of frameworks. We present some of these valuation frameworks in the next section.

# 2. Financial Models

The aim of this section is to present and describe the financial models that will be used throughout this work, with a focus in the deduction of their respective characteristic functions. It is then divided as follows. In Subsection 2.1 we introduce the class of Lévy processes, discuss their general properties and select a few representatives of interest to demonstrate their applications in asset price modelling. In Subsection 2.2, we incorporate stochastic volatility into our modelling and introduce the Heston (1993) model.

## 2.1. Lévy Processes

In this section we introduce Lévy processes, which are essentially continuous-time versions of random walks. They are central to the study of several fields, such as physics, engineering, economics, and, certainly, mathematical finance. In the latter field, Lévy processes attempt to describe the observed reality of financial markets as accurately as possible. One of the main motivations for the use of this class of processes in asset price modelling is the jump component that they contemplate, meeting the empirical discontinuous behaviour that asset prices manifest.

**Definition 2.1.1.** *A càdlàg stochastic process* $X = \{X(t) : t \geq 0\}$ *defined on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *is a* ***Lévy process*** *if it possesses the following properties:*

*(i)* $\mathbb{P}(X(0) = 0) = 1,$

*(ii)* $X(t+s) - X(t)$ *are independent,* $\quad \forall s, t \geq 0,$

*(iii)* $X(t+s) - X(t) \overset{d}{=} X(s), \quad \forall s, t \geq 0,$

*(iv)* $\lim_{s \to t} \mathbb{P}(|X(t) - X(s)| \geq \epsilon) = 0, \quad \forall \epsilon > 0.$

An example of a Lévy process is the Wiener process, defined in Definition A.2.11, which turns out to be the only Lévy process having continuous sample paths. Its characteristic function is given by

$$\varphi_{W(t)}(u) = \mathbb{E}\left[e^{iuW(t)}\right] = \exp\left(-\frac{1}{2}tu^2\right). \tag{2.1}$$

We present some other important examples of Lévy processes that integrate the structure of more complex processes.

**Definition 2.1.2.** *A non-negative integer-valued stochastic process* $\{N(t) : t \geq 0\}$ *defined on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *is a* ***Poisson process*** *with parameter* $\lambda > 0$ *if it is a Lévy process and, for* $t > 0$, $N(t) \sim Poisson(\lambda t)$.

The Poisson process is, therefore, a non-negative integer counting process with finite unit increments over a period of time. As we will see, it can be used to model the frequency of jumps in the asset price process. The characteristic function of $N(t)$ equals

$$\varphi_{N(t)}(u) = \mathbb{E}\left[e^{iuN(t)}\right] = e^{\lambda t(e^{iu}-1)}. \tag{2.2}$$

**Definition 2.1.3.** *Let* $\{N(t) : t \geq 0\}$ *be a Poisson process with intensity* $\lambda$ *and* $\{Y_i : i \geq 1\}$ *be a sequence of independent and identically distributed random variables with probability density function* $f$ *and characteristic function* $\varphi_Y$. *The process* $\{Z(t) : t \geq 0\}$ *defined by*

$$Z(t) = \sum_{i=1}^{N(t)} Y_i$$

*is called a* ***compounded Poisson process****.*

The compound Poisson process not only incorporates the frequency of jumps, regulated by $N(t)$, but also includes a jump magnitude component, controlled by the sequence $Y$. Note that a Poisson process is a particular case of a compound Poisson process, where $Y_i = 1$ for $i \geq 1$. By conditioning on $N(t)$ and using equation (2.2), the expression for the characteristic function of a compound Poisson process writes

$$\begin{aligned}
\varphi_{Z(t)}(u) &= \mathbb{E}\left[e^{iuZ(t)}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[e^{iu\sum_{i=1}^{N(t)} Y_i} \mid N(t)\right]\right] \\
&= \mathbb{E}\left[\varphi_Y(u)^{N(t)}\right] \\
&= \varphi_{N(t)}\left(\frac{1}{i}\ln(\varphi_Y(u))\right) \\
&= e^{\lambda t(\varphi_Y(u)-1)} \\
&= \exp\left\{\lambda t \int_{\mathbb{R}} (e^{iux}-1) f(x)dx\right\}.
\end{aligned} \tag{2.3}$$

**Definition 2.1.4.** *A real-valued stochastic process* $\{\gamma(t) : t \geq 0\}$ *defined on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *is a* ***gamma process*** *if it is a Lévy process and, for* $t > 0$, $\gamma(t) \sim Gamma(\alpha t, \beta)$.

The gamma process has $\mathbb{P}$-almost surely non-decreasing paths, which makes it suitable to model the times where the relevant trading activity occurs. The characteristic function of a gamma process is given by

$$\varphi_{\gamma(t)}(u) = \mathbb{E}\left[e^{iu\gamma(t)}\right] = \left(1 - \frac{iu}{\beta}\right)^{-\alpha t}. \tag{2.4}$$

A key result in this context is the so-called Lévy-Khintchine formula, which allows us to obtain the characteristic function of any Lévy process.

**Theorem 2.1.1** (**Lévy-Khintchine representation**). *Let* $\{X(t) : t \geq 0\}$ *be a Lévy process on* $\mathbb{R}$. *Then*

$$\begin{aligned} \varphi_{X(t)}(u) &= \mathbb{E}\left[e^{iuX(t)}\right] \\ &= e^{t\psi(u)}. \end{aligned} \tag{2.5}$$

*The function* $\psi$ *is called the characteristic exponent and is expressed analytically by*

$$\psi(u) = iub - \frac{1}{2}\sigma^2 u^2 + \int_{\mathbb{R}} (e^{iux} - 1 - iuxh(x))\nu(dx), \tag{2.6}$$

*where* $b \in \mathbb{R}$, $\sigma \geq 0$, $\nu$ *is a measure on* $\mathbb{R}$ *satisfying*

$$\nu(\{0\}) = 0, \quad \int_{\mathbb{R}} \min\{1, |x|^2\}\nu(dx) < \infty,$$

*and* $h$ *is a real-valued truncation function satisfying*

$$\begin{aligned} h(x) &= 1 + o(|x|) \text{ as } x \to 0, \\ h(x) &= O(1/|x|) \text{ as } x \to \infty. \end{aligned} \tag{2.7}$$

*Proof.* A proof for $h(x) = \mathbf{1}_{|x|<1}$ is given in [7, p. 86]. □

The measure $\nu(A)$ can be interpreted as the average number of jumps over a unit of time whose magnitudes belong to $A$. The choice of function $h$ will depend on the nature of the process. For instance, if the process has finitely many jumps, $h$ can be set equal to 0 and we get, as a result, an integrable function relative to $\nu$. Other common choices for this function are $h(x) = 1$ and $h(x) = \mathbf{1}_{|x|<1}$. An important consequence of this result is that the distribution of a Lévy process is uniquely defined by the triplet $(b, \sigma, \nu)$.

### 2.1.1. Asset Price Modelling

To ensure non-negativity of the asset price, we let it be modelled as an exponential Lévy process, that is,

$$S(t) = S(0)e^{X(t)}, \tag{2.8}$$

or, equivalently, we let the log-price process

$$X(t) := \ln\left(\frac{S(t)}{S(0)}\right)$$

to be modelled as a Lévy process. Under the risk-neutral measure $\mathbb{Q}$, the discounted process $\{e^{-(r-q)t}S(t) : t \geq 0\}$ must be a martingale due to no-arbitrage arguments (see Theorem A.2.2). Then,

$$\mathbb{E}^{\mathbb{Q}}[S(T)|\mathcal{F}(t)] = e^{(r-q)\tau}S(t), \tag{2.9}$$

where $r$ denotes the deterministic risk-free interest rate, $q$ is the continuous dividend yield of the asset and $\tau := T - t$. By the stationarity property of Lévy processes, equation (2.9) can be rewritten as

$$\psi(-i) = r - q. \tag{2.10}$$

#### The Black-Scholes model

The most famous asset pricing model based on an exponential Lévy process is the geometric Brownian motion (GBM), where the log-price $X(t)$ is modelled as a diffusion process with a drift component, i.e.

$$X(t) = bt + \sigma W(t), \tag{2.11}$$

where $b \in \mathbb{R}$ and $\{W(t) : t \geq 0\}$ is a Wiener process. Its characteristic function is given by

$$\begin{aligned} \varphi_{X(t)}(u) &= e^{iubt}\varphi_{W(t)}(u) \\ &= \exp\left\{t\left(iub - \frac{1}{2}\sigma^2 u^2\right)\right\} \\ &= e^{t\psi(u)}. \end{aligned} \tag{2.12}$$

Using equation (2.10), we get the risk-neutral value for the drift

$$b = r - q - \frac{\sigma^2}{2}. \tag{2.13}$$

This model gave rise to the so-called Black-Scholes Merton [2, 18] (BSM) formula, which relates the theoretical fair price of a plain-vanilla European option to other parameters that characterize both the option contract and market conditions.

The original intent was to calculate the fair value of an option using its strike and time to maturity, as well as the risk-free rate, the current underlying asset price and its volatility. However, a second application for this formula emerged. The idea was to invert the formula and compute the implied volatility given the current observable option price and the remaining parameters. If the BSM assumptions were correct, this implied volatility would be the same for options with different strike prices but on the same underlying. In reality though, what we observe is a smile shaped (or skew shaped in some markets) dependency relation between the two. In fact, the very existence of a volatility smile or skew indicates that market participants do not consider that asset returns follow a log-normal distribution. Two of the reasons for this to happen is the observable evidence of jumps and non-constant volatility in asset prices. Therefore, an abundance of models that include at least one of these components have been proposed in the literature.

In the remaining of this subsection we deal with jump models, while in Subsection 2.2 we treat the stochastic volatility case. Lévy jump processes can be further split into two

categories of interest, namely jump-diffusion processes and infinite activity processes.

### Jump-diffusion processes

The paths of a Lévy jump-diffusion process can be described by

$$X(t) = bt + \sigma W(t) + Z(t), \tag{2.14}$$

where $\{Z(t) : t \geq 0\}$ is a compound Poisson process whose jump sizes possess a density function $f$. They can be interpreted as extensions of diffusion processes, punctuated by jumps at random intervals. Here, jumps represent rare events, e.g. market crashes or large drawdowns. Given that all sources of randomness are mutually independent, the characteristic function of $X(t)$ writes

$$\begin{aligned} \varphi_{X(t)}(u) &= e^{iubt}\varphi_{W(t)}(u)\varphi_{Z(t)}(u) \\ &= \exp\left\{t\left(iub - \frac{1}{2}\sigma^2u^2 + \lambda\int_{\mathbb{R}}(e^{iux} - 1)f(x)dx\right)\right\} \\ &= e^{t\psi(u)}. \end{aligned} \tag{2.15}$$

Note that equation (2.15) is a particular case of Theorem 2.1.1, where $h(x) = 0$ and $\nu(dx) = \lambda f(x)dx$.

Merton [19] proposed a Lévy jump-diffusion model for the log-price $X(t)$. Jump sizes are assumed to follow a Normal distribution in this model, with density function given by

$$f(y) = (2\pi\delta^2)^{-\frac{1}{2}}\exp\left(-\frac{(y-\mu)^2}{2\delta^2}\right). \tag{2.16}$$

The characteristic function $\varphi_{X(t)}$ can be recovered from equation (2.15), leading to

$$\varphi_{X(t)}(u) = \exp\left\{t\left(iub - \frac{1}{2}\sigma^2u^2 + \lambda(e^{i\mu u - \frac{1}{2}\delta^2u^2} - 1)\right)\right\}. \tag{2.17}$$

Moreover, by equation (2.10), the risk-neutral value of $b$ is

$$b = r - q - \frac{\sigma^2}{2} - \lambda\left(e^{\mu+\frac{\delta^2}{2}} - 1\right). \tag{2.18}$$

Kou [15] proposed another Lévy jump-diffusion model for the log-price $X(t)$. In this approach, jump sizes are assumed to follow an asymmetric double exponential distribution, which has a density function of the form

$$f(y) = p\eta_1 e^{-\eta_1 y}\mathbf{1}_{\{y\geq 0\}} + (1-p)\eta_2 e^{\eta_2 y}\mathbf{1}_{\{y\leq 0\}}, \tag{2.19}$$

where $\eta_1 > 1, \eta_2 > 0$. Parameter $p$ represents the probability of an upward jump. The characteristic function $\varphi_{X(t)}$ can be obtained once again from equation (2.15), and is given by

$$\varphi_{X(t)}(u) = \exp\left\{t\left(iub - \frac{1}{2}\sigma^2 u^2 + \lambda iu\left(\frac{p}{\eta_1 - iu} - \frac{1-p}{\eta_2 + iu}\right)\right)\right\}. \tag{2.20}$$

Furthermore, by equation (2.10), the risk-neutral value of $b$ is

$$b = r - q - \frac{\sigma^2}{2} - \lambda\left(\frac{p}{\eta_1 - 1} - \frac{1-p}{\eta_2 + 1}\right). \tag{2.21}$$

#### Infinite activity processes

For infinite activity Lévy processes, in any finite time interval there are infinitely many jumps. An example of such processes is the Variance Gamma (VG) process, introduced by Madan et al. [17]. Under this process, the log-price $X(t)$ is modelled as

$$X(t) = bt + \theta\gamma(t) + \sigma W(\gamma(t)), \tag{2.22}$$

where $\gamma(t) \sim Gamma\left(\frac{t}{\kappa}, \frac{1}{\kappa}\right)$. By first conditioning on $\gamma(t)$ and using equations (2.1) and (2.4), we obtain the characteristic function $\varphi_{X(t)}$ for the VG process, given by

$$\begin{aligned}
\varphi_{X(t)}(u) &= \mathbb{E}\left[e^{iuX(t)}\right] \\
&= e^{iubt}\mathbb{E}\left[\mathbb{E}\left[e^{iu(\theta\gamma(t)+\sigma W(\gamma(t)))}|\gamma(t)\right]\right] \\
&= e^{iubt}\mathbb{E}\left[e^{i\gamma(t)\left(u\theta+\frac{iu^2\sigma^2}{2}\right)}\right] \\
&= e^{iubt}\left(1 - iu\theta\kappa + \frac{1}{2}u^2\sigma^2\kappa\right)^{-\frac{t}{\kappa}}.
\end{aligned} \tag{2.23}$$

Hence, the characteristic exponent $\psi$ writes

$$\begin{aligned}
\psi(u) &= \frac{1}{t}\ln\left(\varphi_{X(t)}(u)\right) \\
&= iub - \frac{1}{\kappa}\ln\left(1 - iu\theta\kappa + \frac{1}{2}u^2\sigma^2\kappa\right),
\end{aligned} \tag{2.24}$$

and from equation 2.10, we conclude that the risk-neutral value of $b$ is given by

$$b = r - q + \frac{1}{\kappa}\ln\left(1 - \theta\kappa - \frac{1}{2}\sigma^2\kappa\right). \tag{2.25}$$

## 2.2. The Heston (1993) Model

One of the most widely used stochastic volatility model was proposed by Heston [13] in 1993. The attractiveness of this model comes from its several desirable properties. First, it exhibits analytical tractability concerning its characteristic function. Second, it considers correlation shocks between the underlying asset price process and volatility process, taking the leverage effect found in equity option markets into account. Finally, it models volatility as a mean-reverting process, which is in line with observed market behaviour.

### 2.2.1. Model Specification

In the Heston model, the dynamics of the underlying asset price $S(t)$ and its variance $\nu(t)$ are described, under the real-world probability measure $\mathbb{P}$, by the stochastic differential equations (SDE's)

$$dS(t) = \mu S(t)dt + \sqrt{\nu(t)}S(t)dW_1^{\mathbb{P}}(t), \tag{2.26}$$

$$d\nu(t) = k(\theta - \nu(t))dt + \sigma\sqrt{\nu(t)}dW_2^{\mathbb{P}}(t), \tag{2.27}$$

where $\mu \in \mathbb{R}$, and the three non-negative parameters, $k$, $\theta$, and $\sigma$, represent the speed of mean reversion, the long term mean of the instantaneous variance and the volatility of the variance process, respectively. The standard Wiener processes $\{W_1^{\mathbb{P}}(t) : t \geq 0\}$ and $\{W_2^{\mathbb{P}}(t) : t \geq 0\}$ are assumed to be correlated with correlation given by $\rho$, i.e.,

$$dW_1^{\mathbb{P}}(t)dW_2^{\mathbb{P}}(t) = \rho dt. \tag{2.28}$$

We will also work with the log-price $X(t) := \ln(S(t))$ instead of the price $S(t)$, as the characteristic function of Heston's model is expressed in terms of the former. As so, applying Itô's Lemma on the log-price $X(t)$ leads to

$$dX(t) = \left(\mu - \frac{1}{2}\nu(t)\right)dt + \sqrt{\nu(t)}dW_1^{\mathbb{P}}(t). \tag{2.29}$$

#### Risk-neutral dynamics

In order to be able to price contingent claims under Heston's dynamics, equations (2.29) and (2.27) need to be rewritten under the risk-neutral measure $\mathbb{Q}$, as in Definition A.2.17.

Assuming that the process $\boldsymbol{\lambda} = [\lambda_1, \lambda_2]$ satisfies Novikov's condition and with $\mathbf{W}^{\mathbb{P}} = [W_1^{\mathbb{P}}, W_2^{\mathbb{P}}]^T$, we define the measure $\mathbb{Q}$ as

$$\frac{d\mathbb{Q}}{d\mathbb{P}}|\mathcal{F}(T) = \mathcal{E}(\boldsymbol{\lambda} * \mathbf{W}^{\mathbb{P}})(T). \tag{2.30}$$

By equation (A.19) and assuming the non-existence of arbitrage opportunities, we must have

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}[S(T)] &= \mathbb{E}^{\mathbb{P}}[S(T)\mathcal{E}(\boldsymbol{\lambda} * \mathbf{W}^{\mathbb{P}})(T)] \\ &= S(0)e^{(r-q)T}.\end{aligned} \tag{2.31}$$

Writing equation (2.29) w.r.t. $S(T)$ gives

$$S(T) = S(0)\exp\left\{\mu T - \frac{1}{2}\int_0^T \nu(t)dt + \int_0^T \sqrt{\nu(t)}dW_1^{\mathbb{P}}(t)\right\}. \tag{2.32}$$

Using equation (2.32), we can rewritte condition (2.31) as

$$\mathbb{E}^{\mathbb{P}}\left[\mathcal{E}([\sqrt{\nu}+\lambda_1, \lambda_2] * \mathbf{W})(T)\exp\left\{(\mu-(r-q))T + \int_0^T \sqrt{\nu(t)}\lambda_1(t)dt\right\}\right] = 1, \tag{2.33}$$

and, by Novikov's condition, the only restriction we must impose in order for (2.33) to be verified is that

$$\lambda_1(t) = -\frac{\mu-(r-q)}{\sqrt{\nu(t)}}, \tag{2.34}$$

regardless of the choice for $\lambda_2$. This implies that there are infinite risk-neutral measures and, therefore, the market is incomplete under Heston's framework (see Theorem A.2.3). This is a natural conclusion since one can only trade on the underlying asset and in the money-market account, and so, volatility risk cannot be hedged. To complete the market, one has to add a tradable volatility derivative. Relying on economic arguments, Heston [13] assumes that the market price of volatility risk is proportional to the volatility level, i.e.,

$$\lambda_2(t) = -\lambda\frac{\sqrt{\nu(t)}}{\sigma}, \tag{2.35}$$

for $\lambda \in \mathbb{R}$. By Girsanov's theorem, we then have

$$dW_1^{\mathbb{P}}(t) = dW_1^{\mathbb{Q}}(t) + \lambda_1(t)dt, \tag{2.36}$$

$$dW_2^{\mathbb{P}}(t) = dW_2^{\mathbb{Q}}(t) + \lambda_2(t)dt, \tag{2.37}$$

and so, rewritting equations (2.29) and (2.27) in terms of $W_1^{\mathbb{Q}}$ and $W_2^{\mathbb{Q}}$ gives

$$dX(t) = \left(r - q - \frac{1}{2}\nu(t)\right)dt + \sqrt{\nu(t)}dW_1^{\mathbb{Q}}(t), \tag{2.38}$$

$$d\nu(t) = k^*(\theta^* - \nu(t))dt + \sigma\sqrt{\nu(t)}dW_2^{\mathbb{Q}}(t), \tag{2.39}$$

where

$$k^* = k + \lambda, \tag{2.40}$$

and

$$\theta^* = \frac{k\theta}{k^*}. \tag{2.41}$$

It is also more convenient to work with a model that is expressed in terms of independent Wiener processes. Therefore, if we let $\{Z_1^{\mathbb{Q}}(t) : t \geq 0\}$ and $\{Z_2^{\mathbb{Q}}(t) : t \geq 0\}$ be two independent Wiener processes under measure $\mathbb{Q}$, we can rewritte $dW_1^{\mathbb{Q}}(t)$ and $dW_2^{\mathbb{Q}}(t)$ as

$$dW_1^{\mathbb{Q}}(t) = \sqrt{1-\rho^2}dZ_1^{\mathbb{Q}}(t) + \rho dZ_2^{\mathbb{Q}}(t), \tag{2.42}$$

$$dW_2^{\mathbb{Q}}(t) = dZ_2^{\mathbb{Q}}(t). \tag{2.43}$$

Finally, combining equations (2.38) and (2.42) gives the risk-neutral dynamics for the log-price process:

$$dX(t) = \left(r - q - \frac{1}{2}\nu(t)\right)dt + \sqrt{\nu(t)}\left(\sqrt{1-\rho^2}dZ_1^{\mathbb{Q}}(t) + \rho dZ_2^{\mathbb{Q}}(t)\right), \tag{2.44}$$

while combining equations (2.39) and (2.43) gives the risk-neutral dynamics for the variance process:

$$d\nu(t) = k^*(\theta^* - \nu(t))dt + \sigma\sqrt{\nu(t)}dZ_2^{\mathbb{Q}}(t). \tag{2.45}$$

### 2.2.2. Variance analysis

If the Feller condition

$$2k^*\theta^* \geq \sigma^2$$

is satisfied, the origin becomes inaccessible for the variance process, as shown in [12]. This restriction then ensures that the variance process remains positive as long as $\nu(0) > 0$. However, this condition is hardly verified in calibrations to real market data, and often one finds $2k^*\theta^* < \sigma^2$. This leads to a near-singular behaviour of the variance cumulative distribution function near the origin, i.e., the left tail of the variance density function grows extremely fast in value as it approaches zero. The problem with this behaviour is that it may easily produce significant approximation errors in integration-based option pricing methods, when truncating the integration range of the cumulative distribution function. The solution therefore lies in transforming the problem from the variance domain to the log-variance domain.

#### Transformation to log-variance domain

As indicated in [8, p. 391], with

$$\xi := \frac{2k^*\theta^*}{\sigma^2} - 1, \tag{2.46}$$

$$\zeta := \frac{2k^*}{\sigma^2\left(1 - e^{-k^*\tau}\right)}, \tag{2.47}$$

the random variable $2\zeta\nu_{m+1}$ follows a noncentral chi-square distribution with degree $\xi$ and noncentrality parameter $2\zeta\nu_m e^{-k^*\tau}$, with $\tau = t_{m+1} - t_m$. The transition density function of the variance process $\{\nu(t) : t \geq 0\}$ is therefore given by

$$p_\nu(\nu_{m+1};\nu_m) = \zeta e^{-\zeta\left(\nu_m e^{-k^*\tau} + \nu_{m+1}\right)} \left(\frac{\nu_{m+1}}{\nu_m e^{-k^*\tau}}\right)^{\frac{\xi}{2}} I_\xi\left(2\zeta e^{-\frac{1}{2}k^*\tau}\sqrt{\nu_m\nu_{m+1}}\right), \tag{2.48}$$

where $I_\xi(.)$ denotes the modified Bessel function of the first kind with order $\xi$. Note that the definition interval of $\xi$ is $[-1, +\infty[$. Hence, the Feller condition being violated is equivalent to $\xi \in [-1, 0[$.

The transition density function of the log-variance process $\{u(t) : t \geq 0\}$ writes

$$p_u(u_{m+1}; u_m) = \zeta e^{-\zeta\left(e^{u_m - k^*\tau} + e^{u_{m+1}}\right)} \left(\frac{e^{u_{m+1}}}{e^{u_m - k^*\tau}}\right)^{\frac{\xi}{2}} I_\xi\left(2\zeta e^{-\frac{1}{2}k^*\tau}\sqrt{e^{u_m + u_{m+1}}}\right) e^{u_{m+1}}. \tag{2.49}$$

The new term $e^{u_{m+1}}$ prevents the density's left tail to increase drastically around the origin when $\xi \in [-1, 0]$, causing it to converge to zero instead as $u_{m+1} \to -\infty$, although the decay rate decreases as $\xi \to -1$.

#### Log-variance boundaries of the truncation range

The integration range of the log-variance density needs to be truncated for the sake of its computation, i.e., we have to find boundaries $a_u$ and $b_u$ such that

$$\int_{\mathbb{R}} p_u(u_{m+1}; u_m) du_{m+1} \approx \int_{a_u}^{b_u} p_u(u_{m+1}; u_m) du_{m+1}, \tag{2.50}$$

with a sufficiently small approximation error. Boundaries $a_u$ and $b_u$ can be found via Newton's method, according to some error tolerance $TOL$, so that the iteration stopping criterion is

$$p_u(u_{m+1}; u_m) < TOL,$$

for $u_{m+1} \in \mathbb{R}\backslash[a_u, b_u]$.

Application of Newton's method requires both an expression for the derivative of $p_u$ and

a proper initial guess for interval boundaries. The derivative expression is given by

$$\begin{aligned}\frac{dp_u(u_{m+1};u_m)}{du_{m+1}} = &-\left[(-\zeta e^{u_{m+1}}-\xi-1)I_q(2\sqrt{\zeta e^{u_{m+1}}a})-I_{\xi+1}(2\sqrt{\zeta e^{u_{m+1}}a})\right]\\ &\cdot\zeta e^{-a-\zeta e^{u_{m+1}}+u_{m+1}}\left(\frac{\zeta e^{u_{m+1}}}{a}\right)^{\frac{\xi}{2}},\end{aligned} \tag{2.51}$$

where $a := \zeta e^{u_m - k^*\tau}$. Regarding the initial guesses for the interval boundaries, we first estimate the center of the interval as the natural logarithm of the mean value of the variance process, i.e.,

$$\ln(\mathbb{E}[\nu_{m+1}]) = \ln\left(v_m e^{-k^*\tau} + \theta^*(1-e^{-k^*\tau})\right). \tag{2.52}$$

Because the left tail of $p_u$ usually decays much slower than the right tail and the speed of decay decreases as $\xi \to -1$, the following values prove to be reasonable choices for the initial guesses of the boundaries of the truncation range $[a_u, b_u]$:

$$[a_u^0, b_u^0] = \left[\ln(\mathbb{E}[\nu_{m+1}]) - \frac{5}{1+\xi}, \ln(\mathbb{E}[\nu_{m+1}]) + \frac{2}{1+\xi}\right]. \tag{2.53}$$

### 2.2.3. Joint distribution

For valuation purposes, the joint conditional density function $p_{x,u}(x_{m+1}, u_{m+1}; x_m, u_m)$ needs to be known. Because the variance at a future time is independent of the log-price at the current time, as indicated by equations (2.44) and (2.45), we can rewrite the joint conditional density function as

$$p_{x,u}(x_{m+1}, u_{m+1}; x_m, u_m) = p_{x|u}(x_{m+1}; u_{m+1}, x_m, u_m)p_u(u_{m+1}; u_m). \tag{2.54}$$

The probability density function of the log-variance $p_u(u_{m+1}; u_m)$ is already given by equation (2.49). Although an analytical expression for $p_{x|u}(x_{m+1}; u_{m+1}, x_m, u_m)$ does not exist, we can still deduce its correspondent characteristic function $\varphi_{x|u}$. To get an expression for it,

we solve the system of SDE's formed by equations (2.44) and (2.45) w.r.t. $x_{m+1}$, leading to

$$\begin{aligned} x_{m+1} = x_m + (r-q)\tau + \frac{\rho}{\sigma}(\nu_{m+1} - \nu_m - k^*\theta^*\tau) + \left(\frac{k^*\rho}{\sigma} - \frac{1}{2}\right)\int_{t_m}^{t_{m+1}} \nu(s)ds \\ + \sqrt{1-\rho^2}\int_{t_m}^{t_{m+1}} \sqrt{\nu(s)}dW_2(s). \end{aligned} \tag{2.55}$$

The conditional characteristic function of $x_{m+1}$, $\varphi_{x|u}(\omega; u_{m+1}, x_m, u_m)$, is then equal to

$$\begin{aligned} \varphi_{x|u}(\omega; u_{m+1}, x_m, u_m) &= \mathbb{E}\left[\exp(i\omega x_{m+1})|u_{m+1}, \mathcal{F}_{t_m}\right] \\ &= \exp\left\{i\omega\left[x_m + (r-q)\tau + \frac{\rho}{\sigma}(e^{u_{m+1}} - e^{u_m} - k^*\theta^*\tau)\right]\right\} \\ &\times \Phi\left(\omega\left(\frac{k^*\rho}{\sigma} - \frac{1}{2}\right) + \frac{1}{2}i\omega^2(1-\rho^2); e^{u_{m+1}}, e^{u_m}\right), \end{aligned} \tag{2.56}$$

where $\Phi(\omega; \nu_{m+1}, \nu_m)$ is the characteristic function of the time-integrated variance $\int_{t_m}^{t_{m+1}} \nu(s)ds$, and is given by

$$\begin{aligned} \Phi(\omega; \nu_{m+1}, \nu_m) =& \frac{I_\xi\left[\sqrt{\nu_{m+1}\nu_m}\dfrac{4\gamma(\omega)e^{-\frac{1}{2}\gamma(\omega)\tau}}{\sigma^2\left(1-e^{-\gamma(\omega)\tau}\right)}\right]}{I_\xi\left[\sqrt{\nu_{m+1}\nu_m}\dfrac{4k^*e^{-\frac{1}{2}k^*\tau}}{\sigma^2\left(1-e^{-k^*\tau}\right)}\right]} \cdot \frac{\gamma(\omega)e^{-\frac{1}{2}(\gamma(\omega)-k^*)\tau(1-e^{k^*\tau})}}{k^*\left(1-e^{-\gamma(\omega)\tau}\right)} \\ &\times \exp\left\{\frac{\nu_{m+1}+\nu_m}{\sigma^2}\left(\frac{k^*(1+e^{-k^*\tau})}{1-e^{-k^*\tau}} - \frac{\gamma(\omega)(1+e^{-\gamma(\omega)\tau})}{1-e^{-\gamma(\omega)\tau}}\right)\right\}, \end{aligned} \tag{2.57}$$

where $\gamma(\omega) := \sqrt{(k^*)^2 - 2i\sigma^2\omega}$.

**Remark.** *Further details on the derivation of* $\Phi(\omega; \nu_{m+1}, \nu_m)$ *can be found in [4, p. 229].*

# 3. The Cosine Method under Lévy Processes

In this section, we present the COS method, a numerical approximation based on the Fourier cosine series expansion used to price Bermudan and American-style options, under the assumption that the underlying asset price moves accordingly to a Lévy process. The section is organized as follows. In Subsections 3.1 and 3.2, we deduce closed-form expressions for the Fourier cosine series coefficients of the density and the value functions, respectively. In Subsection 3.3, we present an efficient algorithm for the pricing of Bermudan options, based on a fast Fourier transform, and introduce a formula for the pricing of American-style options based on a Richardson Extrapolation. Over the course of this section, we follow [10] closely.

## 3.1. Pricing formula

We start where we left off in Subsection 1.2. We further assume that the underlying asset price follows a Lévy process, while the asset volatility and the risk-free interest rate are held

constant, i.e., the state vector is given by $\mathbf{s_m} = [x_m]$. As so, equation (1.3) simplifies into

$$\begin{aligned} c(x_m, t_m) &= e^{-r\tau}\mathbb{E}^{\mathbb{Q}}\left[v(x_{m+1}, t_{m+1})|\mathcal{F}_m\right] \\ &= e^{-r\tau}\int_{\mathbb{R}} v(y, t_{m+1}) f(y; x_m) dy, \end{aligned} \tag{3.1}$$

where $r := r(t)$, for $t \geq 0$, and $\tau := t_{m+1} - t_m$.

Following equations (A.30) and (A.31), the transition density function $f$ can be approximated by its Fourier cosine series expansion on a closed interval $[a, a + l] \subset \mathbb{R}$, leading to

$$f(x_{m+1}; x_m) = \sum_{k=0}^{+\infty}{}' A_k(x_m) \cos\left(k\pi \frac{x_{m+1} - a}{l}\right), \tag{3.2}$$

where $\sum'$ indicates that the first term of the series is divided by 2 and with the series coefficients $A_k$ defined by

$$A_k(x_m) = \frac{2}{l}\int_a^{a+l} f(y; x_m) \cos\left(k\pi \frac{y - a}{l}\right) dy. \tag{3.3}$$

Hence, by truncating the integration range, we obtain the approximation

$$c_1(x_m, t_m) = e^{-r\tau}\int_a^{a+l} v(y, t_{m+1}) \sum_{k=0}^{+\infty}{}' A_k(x_m) \cos\left(k\pi \frac{y - a}{l}\right) dy. \tag{3.4}$$

Over the course of this work, the terms $c_i$ and $\hat{c}$ shall represent the $i$-th approximation and the final approximation to the continuation value $c$, respectively. As in [9], we let $a$ and $l$ be taken as

$$\begin{aligned} a &= \kappa_1 - 10\sqrt{\kappa_2 + \sqrt{\kappa_4}}, \\ l &= 20\sqrt{\kappa_2 + \sqrt{\kappa_4}}, \end{aligned} \tag{3.5}$$

where $\kappa_n$ denotes the $n$-th cumulant of $x_M$. Combining equations (3.3) and (3.4) and inter-

changing the summation and integration operators leads to

$$c_1(x_m, t_m) = \frac{l}{2} e^{-r\tau} \sum_{k=0}^{+\infty}{}' A_k(x_m) V_k(t_{m+1}), \tag{3.6}$$

where $V_k(t_m)$ are the Fourier cosine series coefficients of $v(x_m, t_m)$ on $[a, a+l]$, i.e.,

$$V_k(t_m) = \frac{2}{l} \int_a^{a+l} v(x, t_m) \cos\left(k\pi \frac{x-a}{l}\right) dx. \tag{3.7}$$

Due to the rapid decay of coefficients $A_k$ and $V_k$ (see Corollary A.3.1), we further truncate the infinite series in equation (3.6), which results in a new approximation for the continuation value, given by

$$c_2(x_m, t_m) = \frac{l}{2} e^{-r\tau} \sum_{k=0}^{N-1}{}' A_k(x_m) V_k(t_{m+1}), \tag{3.8}$$

and by application of Euler's formula, coefficients $A_k$ can be rewritten as

$$A_k(x_m) = \frac{2}{l} \Re\left\{ \exp\left(-\frac{ik\pi a}{l}\right) \int_a^{a+l} \exp\left(\frac{ik\pi}{l} y\right) f(y; x_m) dy \right\}. \tag{3.9}$$

By extending the integration range to $\mathbb{R}$, we can approximate the integral in the previous equation by the characteristic function (see Definition A.2.5) of $x_{m+1}$ as

$$\int_a^{a+l} \exp\left(\frac{ik\pi}{l} y\right) f(y; x_m) dy \approx \int_{\mathbb{R}} \exp\left(\frac{ik\pi}{l} y\right) f(y; x_m) dy = \phi\left(\frac{k\pi}{l}; x_m\right), \tag{3.10}$$

where $\phi(\omega; x_m) = \mathbb{E}^{\mathbb{Q}}\left[e^{i\omega x_{m+1}} | x_m\right]$.

As a result,

$$A_k(x_m) \approx \frac{2}{l} \Re\left\{ \exp\left(-\frac{ik\pi a}{l}\right) \phi\left(\frac{k\pi}{l}; x_m\right) \right\}, \tag{3.11}$$

and thus,

$$c_3(x_m, t_m) = e^{-r\tau} \sum_{k=0}^{N-1}{}' \Re\left\{ \exp\left(-\frac{ik\pi a}{l}\right) \phi\left(\frac{k\pi}{l}; x_m\right) \right\} V_k(t_{m+1}) =: \hat{c}(x_m, t_m) \tag{3.12}$$

For Lévy processes, the previous expression simplifies to

$$\hat{c}(x_m, t_m) = e^{-r\tau} \sum_{k=0}^{N-1}{}' \Re\left\{ \exp\left(ik\pi \frac{x_m - a}{l}\right) \varphi_{levy}\left(\frac{k\pi}{l}\right)\right\} V_k(t_{m+1}), \tag{3.13}$$

where $\varphi_{levy}\left(\frac{k\pi}{l}\right) := \phi\left(\frac{k\pi}{l}; 0\right)$. Finally, using equation (3.13), we can approximate $v(x_0, 0)$ by

$$\hat{v}(x_0, 0) = \hat{c}(x_0, 0) = e^{-r\tau} \sum_{k=0}^{N-1}{}' \Re\left\{ \exp\left(ik\pi \frac{x_0 - a}{l}\right) \varphi_{levy}\left(\frac{k\pi}{l}\right)\right\} V_k(t_1). \tag{3.14}$$

Through a backwards induction formula, it is possible to recover $V_k(t_m)$ from $V_k(t_{m+1})$ for $m = 1, ..., M-1$, since $V_k(t_M)$ is always known. We present such formula in the next subsection.

## 3.2. Value coefficients

The purpose of this subsection is to derive an induction formula for $V_k(t_m)$, so that the value of $V_k(t_1)$ can be retrieved and used in the pricing formula (3.14).

We start by discussing the final time point $t_M$. Because $v(x_M, t_M) = g(x_M)$, we get from equation (3.7)

$$V_k(t_M) = \begin{cases} G_k(0, a+l), & \text{for a call,} \\ G_k(a, 0), & \text{for a put,} \end{cases} \tag{3.15}$$

where

$$G_k(x_1, x_2) = \frac{2}{l} \int_{x_1}^{x_2} g(x) \cos\left(k\pi \frac{x-a}{l}\right) dx. \tag{3.16}$$

For time points $t_m$, where $m = 1, ..., M-1$, one can approximate $V_k(t_m)$ by

$$\hat{V}_k(t_m) = \frac{2}{l} \int_{a}^{a+l} \hat{v}(x, t_m) \cos\left(k\pi \frac{x-a}{l}\right) dx, \tag{3.17}$$

and splitting the integral into two parts gives

$$\hat{V}_k(t_m) = \begin{cases} \hat{C}_k(a, x_m^*, t_m) + G_k(x_m^*, a+l), & \text{for a call,} \\ G_k(a, x_m^*) + \hat{C}_k(x_m^*, a+l, t_m), & \text{for a put,} \end{cases} \tag{3.18}$$

with

$$\hat{C}_k(x_1, x_2, t_m) = \frac{2}{l} \int_{x_1}^{x_2} \hat{c}(x, t_m) \cos\left(k\pi \frac{x-a}{l}\right) dx. \tag{3.19}$$

The point $x_m^*$ is where the continuation value approximation equals the payoff, i.e., $c(x_m^*, t_m) = g(x_m^*)$. Note that $\alpha x_m^* > 0$ because $c(x, t_m) > g(x)$ for $\alpha x \leq 0$. Also note that the function $\hat{c}$ is bounded and smooth, and the function $g$ is smooth and unbounded for $\alpha x > 0$. Therefore, the root $x_m^*$ exists and we can use Newton's method to determine it. If $x_m^*$ is not on the interval $[a, a+l]$, it is set to the nearest boundary point.

#### $G_k$-functions

For a put option, with $x_2 \leq 0$, we have

$$G_k(x_1, x_2) = \frac{2}{l} \int_{x_1}^{x_2} K(1 - e^x) \cos\left(k\pi \frac{x-a}{l}\right) dx, \tag{3.20}$$

and for a call option, with $x_1 \geq 0$,

$$G_k(x_1, x_2) = \frac{2}{l} \int_{x_1}^{x_2} K(e^x - 1) \cos\left(k\pi \frac{x-a}{l}\right) dx. \tag{3.21}$$

Because $x_m^* < 0$ for put options and $x_m^* > 0$ for call options, the last two expressions can be rewritten as

$$G_k(x_1, x_2) = \frac{2}{l} \alpha K \big[\chi_k(x_1, x_2) - \psi_k(x_1, x_2)\big], \tag{3.22}$$

with

$$\chi_k(x_1, x_2) := \int_{x_1}^{x_2} e^x \cos\left(k\pi \frac{x-a}{l}\right) dx \tag{3.23}$$

and

$$\psi_k(x_1, x_2) := \int_{x_1}^{x_2} \cos\left(k\pi \frac{x-a}{l}\right) dx. \tag{3.24}$$

Integrating $\chi_k(x_1, x_2)$ by parts yields

$$\begin{aligned}\chi_k(x_1, x_2) = \frac{1}{1+\left(\frac{k\pi}{l}\right)^2}\Bigg[&\cos\left(k\pi\frac{x_2-a}{l}\right)e^{x_2} - \cos\left(k\pi\frac{x_1-a}{l}\right)e^{x_1} \\ &+ \frac{k\pi}{l}\sin\left(k\pi\frac{x_2-a}{l}\right)e^{x_2} - \frac{k\pi}{l}\sin\left(k\pi\frac{x_1-a}{l}\right)e^{x_1}\Bigg],\end{aligned} \tag{3.25}$$

while direct integration on $\psi_k(x_1, x_2)$ leads to

$$\psi_k(x_1, x_2) = \begin{cases}\left[\sin\left(k\pi\frac{x_2-a}{l}\right) - \sin\left(k\pi\frac{x_1-a}{l}\right)\right]\frac{l}{k\pi}, & k \neq 0, \\ x_2 - x_1, & k = 0.\end{cases} \tag{3.26}$$

### $C_k$-functions

Combining equations (3.13) and (3.19), and interchanging the integration and summation operators, results in

$$\hat{C}_k(x_1, x_2, t_m) = e^{-r\tau}\Re\left\{\sum_{j=0}^{N-1}{}' \varphi_{levy}\left(\frac{j\pi}{l}\right) V_j(t_{m+1}) \mathcal{M}_{k,j}(x_1, x_2)\right\}, \tag{3.27}$$

with

$$\mathcal{M}_{k,j}(x_1, x_2) := \frac{2}{l}\int_{x_1}^{x_2} \exp\left(ij\pi\frac{x-a}{l}\right)\cos\left(k\pi\frac{x-a}{l}\right)dx. \tag{3.28}$$

Note that $V_j(t_M)$ is given by equation (3.15). Therefore, we get

$$\hat{C}_k(x_1, x_2, t_{M-1}) = e^{-r\tau}\Re\left\{\sum_{j=0}^{N-1}{}' \varphi_{levy}\left(\frac{j\pi}{l}\right) V_j(t_M) \mathcal{M}_{k,j}(x_1, x_2)\right\}. \tag{3.29}$$

For $m = 1, ..., M-2$, we replace $V_j(t_{m+1})$ in equation (3.27) by its approximation (3.18), and rewritte $\hat{C}$ as

$$\hat{C}_k(x_1, x_2, t_m) = e^{-r\tau}\Re\left\{ \sum_{j=0}^{N-1}{}' \varphi_{levy}\left(\frac{j\pi}{l}\right)\hat{V}_j(t_{m+1})\mathcal{M}_{k,j}(x_1, x_2)\right\}. \tag{3.30}$$

To summarize, in vector form and for $m = 1, ..., M-1$ we have

$$\hat{\mathbf{V}}(t_m) = \begin{cases} \hat{\mathbf{C}}(a, x_m^*, t_m) + \mathbf{G}(x_m^*, a+l), & \text{for a call,} \\ \mathbf{G}(a, x_m^*) + \hat{\mathbf{C}}(x_m^*, a+l, t_m), & \text{for a put,} \end{cases} \tag{3.31}$$

with

$$\hat{\mathbf{C}}(x_1, x_2, t_m) = \begin{cases} e^{-r\tau}\Re\{\mathcal{M}(x_1, x_2)\Lambda\}\hat{\mathbf{V}}(t_{m+1}), & m = 1, ..., M-2, \\ e^{-r\tau}\Re\{\mathcal{M}(x_1, x_2)\Lambda\}\mathbf{V}(t_M), & m = M-1, \end{cases} \tag{3.32}$$

where the bold-faced letters denote vectors, for instance $\hat{\mathbf{V}}(.) = \left\{\hat{V}_k(.)\right\}_{k=0}^{N-1}$, $\mathcal{M}(x_1, x_2) = \left\{\mathcal{M}_{k,j}(x_1, x_2)\right\}_{k,j=0}^{N-1}$ and $\Lambda = \left\{\varphi\left(\frac{j\pi}{l}\right)\mathbf{1}_{\{k=j\}}\right\}_{k,j=0}^{N-1}$.

The matrix-vector product in (3.32) has complexity $O(N^2)$. As a result, its direct calculation is not efficient to determine $\hat{\mathbf{V}}(t_m)$. An alternative approach based on a fast Fourier transform (FFT) algorithm, with complexity $O(N\log_2(N))$, is presented in the next subsection.

## 3.3. Efficient algorithm

In the following, a FFT-based algorithm is presented for the efficient computation of $\hat{\mathbf{C}}(x_1, x_2, t_m)$, defined in equation (3.32). We start by introducing the notions of both a circular convolution and a discrete Fourier transform.

**Definition 3.3.1.** *For any two n-dimensional vectors* $\boldsymbol{a} = [a_0, a_1, ..., a_{n-1}]$, $\boldsymbol{b} = [b_0, b_1, ..., b_{n-1}]$,

*we define the circular convolution of* $\boldsymbol{a}$ *and* $\boldsymbol{b}$ *to be a new vector* $\boldsymbol{c} = \boldsymbol{a} \circledast \boldsymbol{b}$ *such that*

$$c_j = \sum_{k=0}^{n-1} a_{(j-k)_n} b_k, \tag{3.33}$$

*for* $j = 0, ..., n-1$ *and where* $(m)_n$ *represents the remainder of the division between* $m$ *and* $n$*.*

**Definition 3.3.2.** *The discrete Fourier transform of a vector* $\boldsymbol{a} = [a_0, a_1, ..., a_{n-1}]$ *is the vector* $\boldsymbol{b} = [b_0, b_1, ..., b_{n-1}]$ *such that*

$$b_k = \sum_{j=0}^{n-1} a_j e^{-i\frac{2\pi}{n}jk} \tag{3.34}$$

*for* $k = 0, ..., n-1$*. We write* $\mathcal{D}(\boldsymbol{a}) = \boldsymbol{b}$*.*

The following useful theorem allows us to relate the previous two concepts.

**Theorem 3.3.1** (**Convolution theorem**)**.** *For any* $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{C}^n$*, the following equality holds:*

$$\mathcal{D}(\boldsymbol{a} \circledast \boldsymbol{b}) = \mathcal{D}(\boldsymbol{a})\mathcal{D}(\boldsymbol{b}). \tag{3.35}$$

*Proof.* By Definitions 3.3.1 and 3.3.2,

$$\begin{aligned}
\mathcal{D}(\mathbf{a} \circledast \mathbf{b})_k &= \sum_{j=0}^{n-1} (\mathbf{a} \circledast \mathbf{b})_j e^{-i\frac{2\pi}{n}jk} \\
&= \sum_{j=0}^{n-1} a_{(j-l)_n} e^{-i\frac{2\pi}{n}jk} e^{i\frac{2\pi}{n}lk} \sum_{l=0}^{n-1} b_l e^{-i\frac{2\pi}{n}lk} \\
&= \sum_{j=0}^{n-1} a_{(j-l)_n} e^{-i\frac{2\pi}{n}jk} e^{i\frac{2\pi}{n}lk} \mathcal{D}(\mathbf{b})_k \\
&= \sum_{m=-l}^{n-l-1} a_{(m)_n} e^{-i\frac{2\pi}{n}(l+m)k} e^{i\frac{2\pi}{n}lk} \mathcal{D}(\mathbf{b})_k \\
&= \sum_{m=0}^{n-1} a_m e^{-i\frac{2\pi}{n}mk} \mathcal{D}(\mathbf{b})_k \\
&= \mathcal{D}(\mathbf{a})_k \mathcal{D}(\mathbf{b})_k.
\end{aligned} \tag{3.36}$$

□

**Efficient computation of the matrix-vector product**

Using the identity $\cos(\alpha) = \frac{1}{2}(\exp(i\alpha) + \exp(-i\alpha))$, we can rewritte equation (3.28) as

$$\begin{aligned}
\mathcal{M}_{k,j}(x_1, x_2) &= \frac{1}{l} \int_{x_1}^{x_2} \left[ \exp\left( i(j+k)\pi \frac{x-a}{l} \right) + \exp\left( i(j-k)\pi \frac{x-a}{l} \right) \right] dx \\
&= -\frac{i}{\pi} \left( \mathcal{M}^C_{k,j}(x_1, x_2) + \mathcal{M}^S_{k,j}(x_1, x_2) \right),
\end{aligned} \tag{3.37}$$

where

$$\mathcal{M}^C_{k,j}(x_1, x_2) := \begin{cases} \frac{1}{l}(x_2 - x_1)\pi i, & k = j = 0, \\ \dfrac{\exp\left( i(j+k)\pi \frac{x_2-a}{l} \right) - \exp\left( i(j-k)\pi \frac{x_1-a}{l} \right)}{j+k}, & \text{otherwise} \end{cases} \tag{3.38}$$

and

$$\mathcal{M}^S_{k,j}(x_1,x_2) := \begin{cases} \dfrac{1}{l}(x_2-x_1)\pi i, & k=j, \\ \dfrac{\exp\left(i(j+k)\pi\frac{x_2-a}{l}\right)-\exp\left(i(j-k)\pi\frac{x_1-a}{l}\right)}{j-k}, & k\neq j. \end{cases} \tag{3.39}$$

Since $\Re\{-iz\} = \Im\{z\}$ for $z \in \mathbb{C}$, replacing equation (3.37) into (3.32) gives

$$\hat{C}(x_1,x_2,t_m) = \frac{e^{-r\tau}}{\pi}\Im\{(\mathcal{M}^C+\mathcal{M}^S)\mathbf{u}(t_{m+1})\} \tag{3.40}$$

where

$$\mathbf{u}(t_{m+1}) = \{u_j\}_{j=0}^{N-1}, \quad u_j := \varphi\left(\frac{j\pi}{l}\right)V_j(t_{m+1}), \quad u_0 = \frac{1}{2}V_0(t_{m+1}). \tag{3.41}$$

The matrices $\mathcal{M}^C = \left\{\mathcal{M}^C_{k,j}(x_1,x_2)\right\}_{k,j=0}^{N-1}$ and $\mathcal{M}^S = \left\{\mathcal{M}^S_{k,j}(x_1,x_2)\right\}_{k,j=0}^{N-1}$ are Hankel and Toeplitz matrices, respectively. Hence, they can be represented as

$$\mathcal{M}^C = \begin{bmatrix} m_0 & m_1 & \dots & m_{N-2} & m_{N-1} \\ m_1 & m_2 & \dots & m_{N-1} & m_N \\ \vdots & & \ddots & & \vdots \\ m_{N-2} & m_{N-1} & \dots & m_{2N-4} & m_{2N-3} \\ m_{N-1} & m_N & \dots & m_{2N-3} & m_{2N-2} \end{bmatrix} \tag{3.42}$$

and

$$\mathcal{M}^S = \begin{bmatrix} m_0 & m_1 & \dots & m_{N-2} & m_{N-1} \\ m_{-1} & m_0 & \dots & m_{N-3} & m_{N-2} \\ \vdots & & \ddots & & \vdots \\ m_{2-N} & m_{3-N} & \dots & m_0 & m_1 \\ m_{1-N} & m_{2-N} & \dots & m_{-1} & m_0 \end{bmatrix} \tag{3.43}$$

with

$$m_j := \begin{cases} \dfrac{1}{l}(x_2 - x_1)\pi i, & j = 0, \\ \dfrac{\exp\left(ij\pi\dfrac{x_2 - a}{l}\right) - \exp\left(ij\pi\dfrac{x_1 - a}{l}\right)}{j - k}, & j \neq 0. \end{cases} \tag{3.44}$$

If we let

$$\mathbf{m}_c = [m_{2N-1}, m_{2N-2}, ..., m_1, m_0]^T, \quad \mathbf{u}_c = [0, ..., 0, u_0, ..., u_{N-1}]^T,$$

be two $2N$-dimensional vectors, by Theorem 3.3.1 the product $\mathcal{M}^C\mathbf{u}$ can be rewritten as

$$\begin{aligned} \mathcal{M}^C\mathbf{u} &= \left\{ \sum_{j=0}^{N-1} m_{j+k}u_j \right\}_{k=0}^{N-1} \\ &= \text{rev}((\mathbf{m}_c \circledast \mathbf{u}_c)[N : 2N]) \\ &= \text{rev}(\mathcal{D}^{-1}(\mathcal{D}(\mathbf{m}_c)\mathcal{D}(\mathbf{u}_c))[N : 2N]), \end{aligned} \tag{3.45}$$

where $\mathbf{x}$[i:j] denotes a new vector with the elements from $i$ to $j$ of vector $\mathbf{x}$ and rev($\mathbf{x}$) denotes a new vector with the elements of $\mathbf{x}$ sorted in reversed order.

By a similar reasoning, if we let

$$\mathbf{m}_s = [m_0, m_{-1}, m_{-2}, ..., m_{1-N}, 0, m_{N-1}, m_{N-2}, ..., m_2, m_1]^T, \quad \mathbf{u}_s = [u_0, ..., u_{N-1}, 0, ..., 0]^T,$$

be two $2N$-dimensional vectors, by Theorem 3.3.1 the product $\mathcal{M}^S\mathbf{u}$ can be rewritten as

$$\begin{aligned} \mathcal{M}^S\mathbf{u} &= \left\{ \sum_{j=0}^{N-1} m_{j-k}u_j \right\}_{k=0}^{N-1} \\ &= (\mathbf{m}_s \circledast \mathbf{u}_s)[0 : N-1] \\ &= \mathcal{D}^{-1}(\mathcal{D}(\mathbf{m}_s)\mathcal{D}(\mathbf{u}_s))[0 : N-1]. \end{aligned} \tag{3.46}$$

In order to further increase the efficiency of the algorithm, it should be noted that:

- $\mathcal{D}(\mathbf{u}_c) = \{s_k\}_{k=0}^{2N-1} \cdot \mathcal{D}(\mathbf{u}_s)$, where $s_k = 1$ for $k$ even and $s_k = -1$ for $k$ odd;
- $m_{-j} = -\overline{m}_j$;
- $m_{j+N} = \dfrac{\exp\left(ij\pi\dfrac{x_2-a}{l}\right)\exp\left(iN\pi\dfrac{x_2-a}{l}\right) - \exp\left(ij\pi\dfrac{x_1-a}{l}\right)\exp\left(iN\pi\dfrac{x_1-a}{l}\right)}{j+N}$.

Therefore, it suffices to compute $\mathcal{D}(\mathbf{u}_s)$ and $m_j$ for $j = 0, ..., N-1$.

The algorithms to calculate $\hat{\mathbf{C}}(x_1, x_2, t_m)$ and $\hat{v}(x_0, 0)$ are described as follows:

**Algorithm 2:** Efficient computation of $\hat{C}(x_1, x_2, t_m)$

Compute $m_j$ for $j = 0, ..., N-1$;

Compute $\mathbf{m}_s(x_1, x_2)$ and $\mathbf{m}_c(x_1, x_2)$;

Compute $\mathbf{u}(t_m)$;

Construct $\mathbf{u}_s$ by appending N zeros to $\mathbf{u}(t_m)$;

$\mathcal{M}^C\mathbf{u} \leftarrow$ first N elements of $\mathcal{D}^{-1}(\mathcal{D}(\mathbf{m}_s)\mathcal{D}(\mathbf{u}_s))$;

$\mathcal{M}^S\mathbf{u} \leftarrow$ reverse{first N elements of $\mathcal{D}^{-1}(\mathcal{D}(\mathbf{m}_c) \cdot \mathbf{s} \cdot \mathcal{D}(\mathbf{u}_s))$};

$\hat{C}(x_1, x_2, t_m) \leftarrow [\exp(-r\tau)/\pi]\,\Im\{\mathcal{M}^C\mathbf{u} + \mathcal{M}^S\mathbf{u}\}$

**Algorithm 3:** Pricing Bermudan options with the COS method

**Initialization**: If $\alpha = 1$, $\mathbf{V}(t_M) \leftarrow \mathbf{G}(0, a+l)$; If $\alpha = -1$, $\mathbf{V}(t_M) \leftarrow \mathbf{G}(a, 0)$;

**for** *m=M-1,...,1* **do**

  Determine $x_m^*$ by Newton's method;

  **if** $\alpha = 1$ **then**

    Compute $\hat{\mathbf{C}}(a, x_m^*, t_m)$ following Algorithm 1;

    Compute $\mathbf{G}(x_m^*, a+l)$

  **if** $\alpha = -1$ **then**

    Compute $\mathbf{G}(a, x_m^*)$;

    Compute $\hat{\mathbf{C}}(x_m^*, a+l, t_m)$ following Algorithm 1

  $\hat{\mathbf{V}}(t_m) \leftarrow \hat{\mathbf{C}}(.) + \mathbf{G}(.)$

Compute $\hat{v}(x_0, 0)$ using $\hat{\mathbf{V}}(t_1)$

Finally, the price of an American option can be obtained by applying a Richardson extrapolation on the prices of a few Bermudan options. Let $v(M)$ represent the value of a Bermudan option with $M$ exercise dates. Using the following 4-point Richardson extrapolation scheme proposed in [10, p. 50], the price of an American option is approximately equal to

$$v_{AM}(d) = \frac{1}{21}\left(64v(2^{d+3}) - 56v(2^{d+2}) + 14v(2^{d+1}) - v(2^d)\right). \tag{3.47}$$

# 4. The Cosine Method under the Heston Model

In this section, we derive the pricing formula for Bermudan options under the Heston model. Since this model assumes stochastic volatility, it requires the evaluation of a two-dimensional integral in order to price the option contract. As we have seen in Subsection 2.2.3, the joint probability density function, which is part of the integration kernel, can be written as a product between the log-variance density function and the conditional density function of the log-price given the log-variance. The former has a closed-form expression, and hence can be integrated using a quadrature rule. On the other hand, because the latter is not known in closed-form, we take advantage of the availability of its Fourier transform and respective connection with its Fourier cosine series expansion to develop a discrete formula as a means to evaluate its integral. We follow [11] closely throughout this section.

## 4.1. Pricing formula

We start by rewriting equation (1.3) in terms of Heston's model specifications, presented in Subsection 2.2. With the state vector given by $\mathbf{s_m} = [x_m, u_m]$, it writes

$$\begin{aligned} c(\mathbf{s_m}, t_m) &= e^{-r\tau}\mathbb{E}^{\mathbb{Q}}\left[v(\mathbf{s_{m+1}}, t_{m+1})|\mathcal{F}_m\right] \\ &= e^{-r\tau}\int_{\mathbb{R}}\int_{\mathbb{R}} v(y,\varsigma,t_{m+1})p_{x,u}(y,\varsigma;x_m,u_m)dyd\varsigma \\ &= e^{-r\tau}\int_{\mathbb{R}}\left[\int_{\mathbb{R}} v(y,\varsigma,t_{m+1})p_{x|u}(y;\varsigma,x_m,u_m)dy\right]p_u(\varsigma;u_m)d\varsigma, \end{aligned} \tag{4.1}$$

where $r := r(t)$, for $t \geq 0$, and the function $p_u$ is given by equation (2.49).

#### Log-price conditional density approximation by Fourier cosine series

The COS method is employed in order to get an approximation for the conditional probability density function $p_{x|u}$.

We start by approximating $p_{x|u}$ by its Fourier cosine series expansion on a closed interval $[a, a+l] \subset \mathbb{R}$, defined as

$$[a, a+l] := \left[\kappa_1 - 12\sqrt{|\kappa_2|}, \kappa_1 + 12\sqrt{|\kappa_2|}\right], \tag{4.2}$$

where $\kappa_n$ denotes the $n$th cumulant of the log-stock process. We then have

$$p_{x|u}(x_{m+1};u_{m+1},x_m,u_m) = \sum_{n=0}^{\infty}{}' P_k(u_{m+1},x_m,u_m)\cos\left(k\pi\frac{x_{m+1}-a}{l}\right), \tag{4.3}$$

where the coefficients $P_k$ are defined as

$$P_k(u_{m+1},x_m,u_m) = \frac{2}{l}\int_a^{a+l} p_{x|u}(y;u_{m+1},x_m,u_m)\cos\left(k\pi\frac{y-a}{l}\right)dy. \tag{4.4}$$

Applying Euler's formula, coefficients $P_k$ can be rewritten as

$$\begin{aligned} P_k(u_{m+1}, x_m, u_m) &= \frac{2}{l}\Re\left\{\int_a^{a+l} p_{x|u}(y; u_{m+1}, x_m, u_m) e^{ik\pi\frac{y-a}{l}} dy\right\} \\ &\approx \frac{2}{l}\Re\left\{\varphi_{x|u}\left(\frac{k\pi}{l}; u_{m+1}, x_m, u_m\right) e^{-ik\pi\frac{a}{l}}\right\}, \end{aligned} \tag{4.5}$$

where $\varphi_{x|u}\left(\omega; u_{m+1}, x_m, u_m\right)$ represents the characteristic function of the log-price given the log-variance at time $t_{m+1}$.

By combining equations (4.3) and (4.5), truncating the series of the former and observing that equation (2.56) implies that $\varphi_{x|u}(\omega; u_{m+1}, x_m, u_m) = e^{iwx_m}\varphi_{x|u}(\omega; u_{m+1}, 0, u_m)$, we get the approximation for the conditional transition density function

$$\begin{aligned} p_{x|u}(x_{m+1}; u_{m+1}, x_m, u_m) \approx {\sum_{k=0}^{N-1}}' \frac{2}{l}\Re\left\{\varphi_{x|u}\left(\frac{k\pi}{l}; u_{m+1}, 0, u_m\right) e^{ik\pi\frac{x_m-a}{l}}\right\} \\ \cdot \cos\left(k\pi\frac{x_{m+1}-a}{l}\right). \end{aligned} \tag{4.6}$$

### Quadrature rule for log-variance

Using approximation (2.50) and since the log-variance transition density function is known analytically, the outer integral in equation (4.1) can be implemented by means of a J-point Gaussian quadrature, which gives

$$\int_{a_u}^{b_u} p_u(\varsigma; u_m) d\varsigma \approx \sum_{j_{m+1}=0}^{J-1} w_{j_{m+1}} p_u(\varsigma_{j_{m+1}}; u_m). \tag{4.7}$$

### Discrete pricing formula

Truncating the integration region in equation (4.1) by $[a_u, b_u] \times [a, a+l]$, we obtain the approximation

$$c_1(\mathbf{s_m}, t_m) = e^{-r\tau}\int_{a_u}^{b_u}\left[\int_a^{a+l} v(y, \varsigma, t_{m+1}) p_{x|u}(y; \varsigma, x_m, u_m) dy\right] p_u(\varsigma; u_m) d\varsigma, \tag{4.8}$$

Combining equations (4.7) and (4.8) leads to

$$c_2(\mathbf{s_m}, t_m) = e^{-r\tau} \sum_{j_{m+1}=0}^{J-1} w_{j_{m+1}} p_u(\varsigma_{j_{m+1}}; u_m) \cdot \left[ \int_a^{a+l} v(y, \varsigma_{j_{m+1}}, t_{m+1}) p_{x|u}(y; \varsigma_{j_{m+1}}, x_m, u_m) dy \right], \tag{4.9}$$

where $w_{j_m}$ represents the weight of the quadrature node $\varsigma_{j_m}$, for $j_m = 0, \ldots, J-1$ and for $m = 0, \ldots, M$.

Replacing $p_{x|u}(y; \varsigma_{j_{m+1}}, x_m, u_m)$ by the COS approximation (4.6) gives

$$\begin{aligned} c_3(\mathbf{s_m}, t_m) &= e^{-r\tau} \sum_{j_{m+1}=0}^{J-1} w_{j_{m+1}} {\sum_{k=0}}' V_{k,j_{m+1}}(t_{m+1}) \Re \left\{ \tilde{\varphi}\left( \frac{k\pi}{l}; \varsigma_{j_{m+1}}, u_m \right) e^{ik\pi \frac{x_m - a}{l}} \right\} \\ &=: \hat{c}(\mathbf{s_m}, t_m), \end{aligned} \tag{4.10}$$

where $V_{k,j_m}(t_m)$ are the Fourier cosine series coefficients of the option value function on $[a, a+l]$ at time $t_m$, defined as

$$V_{k,j_m}(t_m) = \frac{2}{l} \int_a^{a+l} v(x, \varsigma_{j_m}, t_m) \cos\left( k\pi \frac{x-a}{l} \right) dx, \tag{4.11}$$

and with

$$\tilde{\varphi}\left( \frac{k\pi}{l}; \varsigma_{j_{m+1}}, u_m \right) := p_u(\varsigma_{j_{m+1}}; u_m) \varphi_{x|u}\left( \frac{k\pi}{l}; \varsigma_{j_{m+1}}, 0, u_m \right). \tag{4.12}$$

Finally, interchanging summation operators in equation (4.10) yields the discrete formula for the continuation value

$$\hat{c}(\mathbf{s_m}, t_m) = e^{-r\tau} \Re \left\{ {\sum_{k=0}}' \beta_k(u_m, t_m) e^{ik\pi \frac{x_m - a}{l}} \right\}, \tag{4.13}$$

with

$$\beta_k(u_m, t_m) := \sum_{j_{m+1}=0}^{J-1} w_{j_{m+1}} V_{k,j_{m+1}}(t_{m+1}) \tilde{\varphi}\left( \frac{k\pi}{l}; \varsigma_{j_{m+1}}, u_m \right). \tag{4.14}$$

Due to the use of a quadrature rule in the log-variance dimension, we compute on a log-variance grid. The same log-variance grid is employed for all time points, which gives

$$\hat{c}(\mathbf{s_m}, t_m) = e^{-r\tau}\Re\left\{\sum_{k=0}^{N-1}{}' \beta_k(\varsigma_{j_m}, t_m) e^{ik\pi\frac{x_m - a}{l}}\right\}, \tag{4.15}$$

where we redefine $\mathbf{s_m} = [x_m, \varsigma_{j_m}]$. Finally, the option's value at the initial time can be approximated by

$$\hat{v}(\mathbf{s_0}, 0) = \hat{c}(\mathbf{s_m}, 0) = e^{-r\tau}\Re\left\{\sum_{k=0}^{N-1}{}' \beta_k(\varsigma_{j_0}, 0) e^{ik\pi\frac{x_0 - a}{l}}\right\}, \tag{4.16}$$

with

$$\beta_k(\varsigma_{j_0}, 0) = \sum_{j_1=0}^{J-1} w_{j_1} V_{k,j_1}(t_1)\tilde{\varphi}\left(\frac{k\pi}{l}; \varsigma_{j_1}, \varsigma_{j_0}\right). \tag{4.17}$$

Similarly to the single-factor case addressed in Section 3 , it is possible to recover the coefficients $V_{k,j_m}(t_m)$ from $V_{k,j_{m+1}}(t_{m+1})$, for $m = 1, ..., M-1$. We devote the next subsection to this subject.

## 4.2. Value coefficients

The purpose of this section is to derive an induction formula for $V_{k,j_m}(t_m)$, so that the value of $V_{k,j_1}(t_1)$ can be retrieved and used in equations (4.17) and (4.16).

We start by discussing the final time point $t_M$. Because $v(x_M, \varsigma_{j_M}, t_M) = g(x_M)$, we get from equation (4.11)

$$V_{k,j_M}(t_M) = \begin{cases} G_k(0, a+l), & \text{for a call,} \\ G_k(a, 0), & \text{for a put,} \end{cases} \tag{4.18}$$

where $G_k(x_1, x_2)$ is defined in equation (3.16). For time points $t_m$, where $m = 1, ..., M-1$,

coefficients $V_{k,j_m}(t_m)$ can be approximated by

$$\hat{V}_{k,j_m}(t_m) = \frac{2}{l}\int_a^{a+l} \hat{v}(x, \varsigma_{j_m}, t_m)\cos\left(k\pi\frac{x-a}{l}\right)dx, \tag{4.19}$$

and splitting the integral into two parts gives

$$\hat{V}_{k,j_m}(t_m) = \begin{cases} \hat{C}_{k,j_m}(a, x_m^*(\varsigma_{j_m}), t_m) + G_k(x_m^*(\varsigma_{j_m}), a+l), & \text{for a call,} \\ G_k(a, x_m^*(\varsigma_{j_m})) + \hat{C}_{k,j_m}(x_m^*(\varsigma_{j_m}), a+l, t_m), & \text{for a put,} \end{cases} \tag{4.20}$$

where $x_m^*(\varsigma_{j_m})$ is the early-exercise point, i.e., $\hat{c}(x_m^*(\varsigma_{j_m}), \varsigma_{j_m}, t_m) = g(x_m^*(\varsigma_{j_m}))$, and

$$\hat{C}_{k,j_m}(x_1, x_2, t_m) := \frac{2}{l}\int_{x_1}^{x_2} \hat{c}(x, \varsigma_{j_m}, t_m)\cos\left(k\pi\frac{x-a}{l}\right)dx. \tag{4.21}$$

### $\hat{C}_{k,j_m}$-functions

Combining equations (4.15) and (4.21) and interchanging the integration and summation operators leads to

$$\hat{C}_{k,j_m}(x_1, x_2, t_m) = e^{-r\tau}\Re\left\{\sum_{n=0}^{N-1}{}' \mathcal{M}_{k,n}(x_1, x_2)\beta_n(\varsigma_{j_m}, t_m)\right\}, \tag{4.22}$$

where $\mathcal{M}_{k,n}(x_1, x_2)$ is defined in equation (3.37). Because $V_{n,j_M}(t_M)$ is given by equation (4.18), we have

$$\beta_n(\varsigma_{j_{M-1}}, t_{M-1}) = \sum_{j_M=0}^{J-1} w_{j_M} V_{n,j_M}(t_M)\tilde{\varphi}\left(\frac{n\pi}{l}; \varsigma_{j_M}, \varsigma_{j_{M-1}}\right), \tag{4.23}$$

and for $m = 1, \ldots, M-2$, we define the approximation $\hat{\beta}$ as

$$\hat{\beta}_n(\varsigma_{j_m}, t_m) = \sum_{j_{m+1}=0}^{J-1} w_{j_{m+1}} \hat{V}_{n,j_{m+1}}(t_{m+1})\tilde{\varphi}\left(\frac{n\pi}{l}; \varsigma_{j_{m+1}}, \varsigma_{j_m}\right). \tag{4.24}$$

Simply put, in matrix form we have

$$\hat{V}(t_m) = \begin{cases} \hat{C}(a, x_m^*(\varsigma_{j_m}), t_m) + G(x_m^*(\varsigma_{j_m}), a + l), & \text{for a call,} \\ G(a, x_m^*(\varsigma_{j_m})) + \hat{C}(x_m^*(\varsigma_{j_m}), a + l, t_m), & \text{for a put,} \end{cases} \tag{4.25}$$

with $G(.) = \left\{G_k(.)\right\}_{k,j_m=0}^{N-1,J-1}$ and

$$\hat{C}(x_1, x_2, t_m) = e^{-r\tau}\Re\{\mathcal{M}(x_1, x_2) \cdot B'(t_m)\}, \tag{4.26}$$

where $\mathcal{M}(x_1, x_2) = \left\{\mathcal{M}_{k,n}(x_1, x_2)\right\}_{k,n=0}^{N-1}$, and $B'$ indicates that the first row of the matrix $B$ is divided by 2. The matrix $B$ is defined as

$$B(t_m) := [\hat{\beta}_0(t_m), \hat{\beta}_1(t_m), \ldots, \hat{\beta}_{J-1}(t_m)] \tag{4.27}$$

where

$$\hat{\beta}_{j_m}(t_m) = \begin{cases} [\hat{V}(t_{m+1}) \circ \tilde{\Phi}(\varsigma_{j_{m+1}})]\mathbf{w}, & m = 1, \ldots, M-2, \\ [V(t_M) \circ \tilde{\Phi}(\varsigma_{j_M})]\mathbf{w}, & m = M-1, \end{cases} \tag{4.28}$$

with $\tilde{\Phi}(\varsigma_{j_m}) := \left\{\tilde{\varphi}\left(\frac{k\pi}{l}; \varsigma_{j_{m+1}}, \varsigma_{j_m}\right)\right\}_{k,j_{m+1}=0}^{N-1,J-1}$ and $\mathbf{w} = \{w_j\}_{j=0}^{J-1}$. The operator "$\circ$" denotes the Hadamard product between two matrices.

## 4.3. Efficient algorithm

As discussed in Subsection 3.3, the matrix $\mathcal{M}(x_1, x_2)$ can be written as the sum of a Hankel matrix, $\mathcal{M}^C$, and a Toeplitz matrix, $\mathcal{M}^S$. The FFT algorithm, also discussed in that subsection, can therefore be employed to calculate the matrix product in equation (4.26), with complexity $O(N \log_2(N))$.

The backward recursion algorithm is summarized below.

**Algorithm 4:** Pricing Bermudan options with the COS method under the Heston model

**Initialization**:

Determine $a_u$ and $b_u$ by Newton's method (as described in Subsection 2.2.2);

Calculate matrix $V(t_M)$ using equation (4.18);

Calculate matrices $\tilde{\Phi}(\varsigma_j)$, for $j = 0, \ldots, J-1$;

**Main Loop**:

**for** *m=M-1,...,1* **do**

  Determine $x_m^*(\varsigma_{j_m})$ by Newton's method;

  Calculate the first row of $\mathcal{M}^S$ (or, equivalently, the first column of $\mathcal{M}^C$);

  Calculate $\hat{\beta}_{j_m}(t_m)$, for $j_m = 0, \ldots, J-1$, using equation (4.28);

  Divide the first element of $\hat{\beta}_{j_m}(t_m)$ by 2, for $j_m = 0, \ldots, J-1$, giving $\hat{\beta}'_{j_m}(t_m)$;

  Calculate $e^{-r\tau}\Re\{(\mathcal{M}^S + \mathcal{M}^C)\hat{\beta}'_{j_m}(t_m)\}$ (column vectors of $\hat{C}(., t_m)$), for $j_m = 0, \ldots, J-1$;

  Calculate $\hat{V}(t_m)$ using equation (4.25);

Compute $\hat{v}(\mathbf{s_0}, 0)$ using $\hat{V}(t_1)$ in equations (4.17) and (4.16), for $j_0 = 0, \ldots, J-1$;

Use spline interpolation to get $\hat{v}(x_0, u_0, 0)$

Lastly, the price of an American-style option can once more be approximated using the extrapolation scheme given by equation (3.47).

# 5. Numerical Results

In this section, we present the numerical results obtained by using the COS method to price Bermudan and American options, under the models introduced in Section 2. In Subsection 5.1, we analyse the results for Bermudan options with ten exercise dates under Lévy processes. In Subsection 5.2, we proceed the analysis to American options under Heston's model. All experiments were carried out in a MATLAB 9.7 environment, on a machine equipped with an Intel Core i5 Quad-Core 1.4GHz CPU and 6.00GB RAM.

## 5.1. Bermudan Options

In this subsection, we value a set of 10-time exercisable Bermudan put options on a non-dividend paying stock. Strikes and maturities for these options are contained in the sets $\mathcal{K} = \{K \in \mathbb{Z} : 90 \leq K \leq 110\}$ and $\mathcal{T} = \{0.25, 0.5, 1, 2.5, 5\}$, respectively, for a total of 105 options being evaluated. Moreover, we assume that the initial stock price $S_0$ is 100 and that the interest rate $r$ is 10%, while the parameter set for each model under analysis is given in Table A.1.

As a measure of accuracy, we consider the root mean square error (RMSE) value, given by

$$RMSE = \sqrt{\frac{\sum_{K\in\mathcal{K}}\sum_{T\in\mathcal{T}}[v(K,T) - v_{ref}(K,T)]^2}{105}}, \tag{5.1}$$

| Config. | Model | Parameters |
|---|---|---|
| 1 | GBM | $\sigma = 0.2$ |
| 2 | Kou | $\sigma = 0.2$, $\lambda = 10$, $\eta_1 = 50$, $\eta_2 = 25$, $p = 0.3$ |
| 3 | Merton JD | $\sigma = 0.1$, $\lambda = 3$, $\mu = -0.05$, $\delta = 0.05$ |
| 4 | VG | $\sigma = 0.12$, $\theta = -0.14$, $\kappa = 0.1$ |

Table 5.1.: Test parameters for pricing Bermudan options.

| $n$ | Test 1 | | Test 2 | |
|---|---|---|---|---|
| | RMSE | Time (ms) | RMSE | Time (ms) |
| 5 | 7.90E-03 | 3.968 | 2.33E-02 | 4.400 |
| 6 | 4.32E-06 | 4.534 | 1.11E-04 | 5.021 |
| 7 | 5.43E-10 | 4.665 | 3.80E-09 | 5.240 |
| 8 | 5.43E-10 | 5.826 | 2.64E-09 | 7.314 |
| 9 | 5.43E-10 | 6.691 | 2.64E-09 | 8.405 |

Table 5.2.: Results for configurations 1 and 2, with $N = 2^n$.

where $v(K,T)$ is the value obtained by the COS method and $v_{ref}(K,T)$ is the reference value, obtained through the CONV [16] method with $N = 2^{20}$.

Tables 5.2 and 5.3 summarize the results for all four tests. For configurations 1 to 3, the method exhibits a rapid convergence to the reference values.

As illustrated by Figure 5.1, configuration 4 seems to converge relatively slower when compared to the remaining tests. When dealing with European options, the VG density function ceases to be continuous for $T < \kappa$, which causes a slower convergence to the reference value, as reported in [9, p. 17]. In the Bermudan case, this situation is then observed when $\tau < \kappa$, as shown by Figure 5.2.

## 5.2. American Options

We now turn our attention to the pricing of American put options. Prices for these contracts can be obtained either by approximation to a Bermudan option with many exercise dates, or by applying the Richardson extrapolation scheme given by equation (3.47). We analyse the results obtained by both approaches.

| $n$ | Test 3 | | Test 4 | |
|---|---|---|---|---|
| | RMSE | Time (ms) | RMSE | Time (ms) |
| 5 | 7.40E-01 | 4.400 | 3.55E-02 | 4.597 |
| 6 | 1.51E-01 | 5.021 | 4.58E-03 | 4.830 |
| 7 | 2.38E-02 | 5.240 | 9.35E-04 | 5.525 |
| 8 | 1.88E-03 | 7.314 | 1.60E-04 | 6.833 |
| 9 | 1.99E-05 | 8.405 | 3.78E-05 | 8.047 |
| 10 | 1.82E-09 | 16.734 | 1.05E-05 | 18.025 |
| 11 | 1.82E-09 | 27.327 | 2.12E-06 | 27.303 |
| 12 | 1.82E-09 | 44.035 | 4.62E-07 | 46.612 |

Table 5.3.: Results for configurations 3 and 4, with $N = 2^n$.

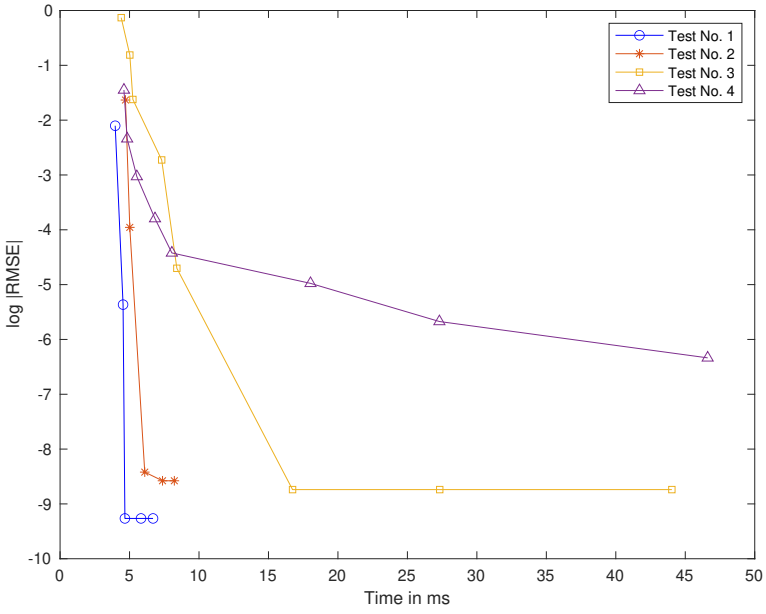


Figure 5.1.: Convergence speed for each configuration.

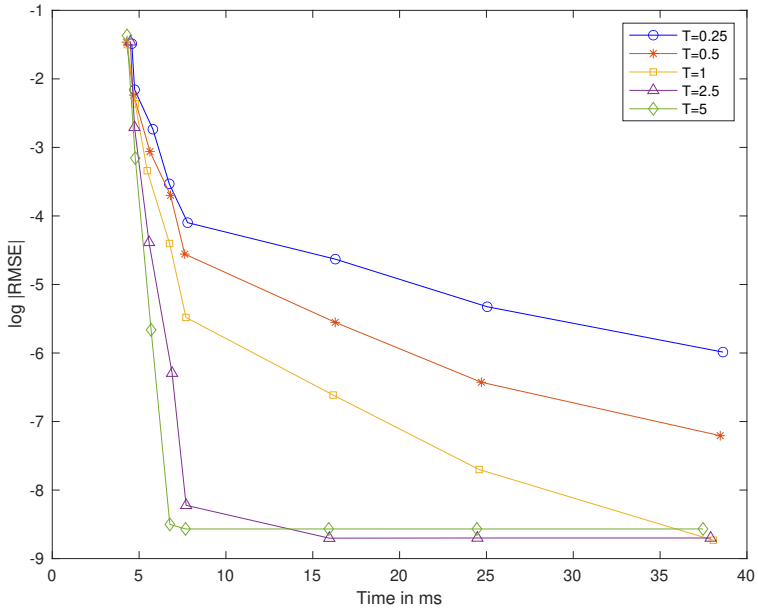


Figure 5.2.: Convergence speed for the VG model under different maturities.

| Config. | Model | Parameters | $\xi$ |
|---|---|---|---|
| 5 | Heston | $k = 5$, $\theta = 0.16$, $\sigma = 0.9$, $\rho = 0.1$, $\nu_0 = 0.25$ | 0.96 |
| 6 | Heston | $k = 1.15$, $\theta = 0.0348$, $\sigma = 0.39$, $\rho = -0.64$, $\nu_0 = 0.0348$ | -0.47 |

Table 5.4.: Model parameters for pricing American options.

The COS method is tested under configuration 1 from Table A.1 and under configurations 5 and 6, from Table 5.4. Reference values for configuration 1 were obtained by application of the binomial method, using 15000 steps, while reference values for configuration 5 are taken from [14, p. 17].

Configuration 6 aims to produce a "realistic" market scenario under Heston's model, with $\xi$ from equation (2.46) taking values on $[-1, 0[$. Parameters for this configuration are taken

| Config. | Model | Parameters |
|---|---|---|
| 5 | Heston | $T = 0.25$, $S = \{9, 10, 11\}$, $K = 10$, $r = 0.1$, $q = 0$ |
| 6 | Heston | $T = 0.25$, $S = \{90, 100, 110\}$, $K = 100$, $r = 0.04$, $q = 0$ |

Table 5.5.: Market parameters for pricing American options.

| $P(N/2)$ | | | Richardson | | |
|---|---|---|---|---|---|
| $n$ | RMSE | Time (ms) | $d$ | RMSE | Time (ms) |
| 6 | 8.99E-02 | 11.035 | 1 | 2.51E-02 | 14.315 |
| 7 | 4.47E-02 | 25.391 | 2 | 7.04E-03 | 27.254 |
| 8 | 2.25E-02 | 58.895 | 3 | 2.92E-03 | 53.064 |
| 9 | 1.13E-02 | 141.238 | 4 | 9.48E-04 | 107.801 |
| 10 | 5.61E-03 | 351.278 | 5 | 1.88E-04 | 222.505 |
| 11 | 2.75E-03 | 943.131 | 6 | 1.34E-04 | 428.556 |

Table 5.6.: Results for American put options under configuration 1 and with $N = 2^n$.

from [11, p. 461].

For configuration 1, we price a set of options with strikes and maturities contained in sets $\mathcal{K}$ and $\mathcal{T}$ from Subsection 5.1 and use the RMSE value from equation (5.1) as a measure of accuracy, whereas for configurations 5 and 6 we price a set of options assuming market parameters as in Table 5.5.

Results for configuration 1 are summarized in Table 5.6, where "$P(N/2)$" indicates that the American put option price is approximated by a Bermudan put option price with $N/2$ exercise dates, and "Richardson" denotes the prices obtained through the Richardson extrapolation (RE) scheme given by equation (3.47), for different values of $d$ and fixing $n = 8$. It is observable that the extrapolation method converges faster and exhibits an increasingly better computational efficiency when compared to the direct approximation (DA) approach.

Results for configurations 5 and 6 are also given by Tables 5.7 and 5.8, respectively. A Gauss-Legendre quadrature rule is applied to the discretised log-variance domain. As $q$ approaches -1, the method demands a finer grid to maintain at least reasonable results, which computationally speaking translates into a challenge when working with a $q$ very close to -1. Nevertheless, the method exhibits a fast error convergence and returns highly accurate results when the Richardson extrapolation scheme is applied.

| $S_0$ | 9 | 10 | 11 | Time (sec) |
|---|---|---|---|---|
| DA; $M = 60$ | -7.13E-04 | -4.96E-04 | -3.16E-04 | 4.267 |
| DA; $M = 80$ | -4.91E-04 | -3.30E-04 | -2.04E-04 | 5.161 |
| RE; $d = 1$ | 1.59E-04 | 1.21E-06 | -9.73E-06 | 6.336 |

Table 5.7.: Error results for configuration 5, with $N = J = 2^7$ and $TOL = 10^{-7}$.

| $S_0$ | 90 | 100 | 110 | Time (sec) |
|---|---|---|---|---|
| DA; $M = 60$ | 9.99570939 | 3.20730142 | 0.92786129 | 125.465 |
| DA; $M = 80$ | 9.99768423 | 3.20774327 | 0.92799714 | 136.114 |
| RE; $d = 1$ | 10.00578528 | 3.20896038 | 0.92834218 | 265.321 |

Table 5.8.: Prices for configuration 6, with $N = J = 2^9$ and $TOL = 10^{-7}$.

# 6. Conclusions

This thesis shows how the COS method can be used to price early-exercise options. The main idea is to consider the discounted risk-neutral expectation as a point of departure and proceed with the calculation by substituting the probability density function by its Fourier cosine series expansion. The series is then truncated and, after exploiting the availability of the characteristic function, we arrive at a semi-analytic formula that, when computed with the help of a FFT algorithm, can be used to price early-exercise options.

When considering Heston's stochastic volatility model, the near-singular behaviour in the left-side tail of the variance density is dealt with by applying a change of variables to the log-variance domain. A quadrature rule is then applied, dividing the log-variance domain into a discrete grid.

The method itself returns the value of a Bermudan option with $M$ exercise dates. In this case, rapid convergence is observed, except for the Variance Gamma model, which displays some dependence on the parameter combination.

Prices for American put options were obtained either by direct approximation, i.e., by computing the value of a Bermudan option with many exercise dates, or by means of a Richardson extrapolation scheme. The issue with the first approach is that values tend to get inaccurate as the time interval between two consecutive dates decreases. The solution lies in considering a larger $N$ value, which in turn increases the invested computational time.

As the Richardson extrapolation approach uses a set of Bermudan options with a reduced number of exercise dates, this proves to be a better way of determining the price of an American option, both in terms of accuracy and speed.

Direct application of the method to price call options returns inaccurate results. This can be attributed to a numerical error that cannot be minimised by working on a finer grid. This problem could then be solved by implementing a put-call parity or duality relation.

As for future work, it would be interesting to see the method extended to multi-factor financial models, as for example the Heston-Hull-White model, which considers both a stochastic volatility and a stochastic interest rate.

# A. Auxiliar Results

## A.1. Measure Theory

**Definition A.1.1.** *Let $X$ be a nonempty set. A collection $\mathcal{F}$ of subsets of $X$ is a* ***$\sigma$-algebra*** *if:*

*(i)* $\emptyset \in \mathcal{F}$,

*(ii)* $A \in \mathcal{F} \implies A^C \in \mathcal{F}$,

*(iii)* $A_n \in \mathcal{F}\ (n \in \mathbb{N}) \implies \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$

*The pair $(X, \mathcal{F})$ is called a* ***measurable space****.*

**Proposition A.1.1.** *Given any collection $\mathcal{C}$ of subsets of $X$, there exists a smallest $\sigma$-algebra $\sigma(\mathcal{C})$ containing $\mathcal{C}$, namely*

$$\sigma(\mathcal{C}) = \bigcap\{\mathcal{G} : \mathcal{G} \ \sigma\text{-algebra on } X,\ \mathcal{C} \subset \mathcal{G}\}.$$

*We call $\sigma(\mathcal{C})$ the* ***$\sigma$-algebra generated by $\mathcal{C}$****.*

**Definition A.1.2.** *The* ***Borel $\sigma$-algebra*** *$\mathcal{B}(\mathbb{R})$ is the $\sigma$-algebra generated by all closed intervals of $\mathbb{R}$.*

**Definition A.1.3.** *Let $(X, \mathcal{F})$ be a measurable space. A function $f : X \to \mathbb{R}$ is called* ***$\mathcal{F}$-measurable*** *if*

$$f^{-1}(B) := \{x \in X : f(x) \in B\} \in \mathcal{F}, \quad \forall B \in \mathcal{B}(\mathbb{R}).$$

**Definition A.1.4.** *Let $X$ be a nonempty set. Given a function $f : X \to \mathbb{R}$, the $\sigma$-algebra $\sigma(f)$ is defined as the smallest $\sigma$-algebra with respect to which $f$ is measurable, and we call it the* ***$\sigma$-algebra generated by $f$****.*

**Definition A.1.5.** *A* ***measure*** *$\mu$ on a measurable space $(X, \mathcal{F})$ is a function $\mu : \mathcal{F} \to [0, +\infty]$, such that*

*(i) $\mu(\emptyset) = 0$,*

*(ii) if $A_n \in \mathcal{F}$ $(n \in \mathbb{N})$ and $A_i \cap A_j = \emptyset$, for $i \neq j$, then*

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n).$$

*We say that $\mu$ is* ***finite*** *if*

$$\mu(X) < \infty.$$

*The triple $(X, \mathcal{F}, \mu)$ is called a* ***measure space****.*

**Definition A.1.6.** *A measure space $(X, \mathcal{F}, \mu)$ is said to be* ***complete*** *if*

$$A \subset B, B \in \mathcal{F} \text{ and } \mu(B) = 0 \implies A \in \mathcal{F}.$$

**Definition A.1.7.** *Let $(X, \mathcal{F}, \mu)$ be a measure space. Given a measurable function $f : X \to$*

$[0, +\infty)$, *the* ***integral*** *of* $f$ *w.r.t.* $\mu$ *over* $X$ *is defined by*

$$\int_X f(x)d\mu(x) = \sup\left\{\int_X \varphi(x)d\mu(x) : \varphi \text{ is a simple function and } 0 \leq \varphi \leq f\right\}. \tag{A.1}$$

**Definition A.1.8.** *Let* $(X, \mathcal{F}, \mu)$ *be a measure space. Given a measurable function* $f : X \to \mathbb{R}$*, the* ***integral*** *of* $f$ *w.r.t.* $\mu$ *over* $X$ *is defined by*

$$\int_X f(x)d\mu(x) = \int_X f^+(x)d\mu(x) - \int_X f^-(x)d\mu(x), \tag{A.2}$$

*where*

$$f^+ = \max[0, f], \quad f^- = \max[-f, 0].$$

**Definition A.1.9.** *Let* $p \in \mathbb{N}$*. The function class* $L^p(X, \mathcal{F}, \mu)$ *(we write* $L^p(X)$ *when there is no ambiguity concerning the measure space) is the class of measurable functions* $f : X \to \mathbb{R}$ *such that*

$$\int_X |f(x)|^p d\mu(x) < \infty. \tag{A.3}$$

*In particular,* $L^1(X)$ *forms the class of* ***integrable*** *functions over* $X$*.*

**Definition A.1.10.** *Let* $f \in L^p(X)$*. The* $\boldsymbol{L^p}$***-norm*** $||f||_p$ *is defined by*

$$||f||_p = \left(\int_X |f(x)|^p d\mu(x)\right)^{\frac{1}{p}}. \tag{A.4}$$

**Definition A.1.11.** *Let* $f, g \in L^2(X)$*. The real number* $(f, g)$ *denotes the* ***inner product*** *between* $f$ *and* $g$*, and is given by*

$$(f, g) = \int_X f(x)g(x)d\mu(x). \tag{A.5}$$

**Definition A.1.12.** *Let* $\mu$ *and* $\nu$ *be two measures on a measurable space* $(X, \mathcal{F})$*. Then,*

*(i)* $\nu$ *is* ***absolutely continuous*** *w.r.t.* $\mu$ *if*

$$\mu(A) = 0 \implies \nu(A) = 0, \quad \forall A \in \mathcal{F}.$$

*We write this as* $\nu \ll \mu$.

*(ii)* $\nu$ *and* $\mu$ *are said to be* ***equivalent*** *if* $\nu \ll \mu$ *and* $\mu \ll \nu$. *We write this as* $\mu \sim \nu$.

**Theorem A.1.1 (Radon-Nikodym).** *Let* $\nu$ *and* $\mu$ *be two finite measures on a measurable space* $(X, \mathcal{F})$, *such that* $\nu \ll \mu$. *Then, there exists a function* $f \in L^1$ *such that*

$$\nu(A) = \int_A f d\mu, \quad \forall A \in \mathcal{F}.$$

*Proof.* See [5, p. 195]. □

**Definition A.1.13.** *The function* $f$ *of Theorem A.1.1 is called the* ***Radon-Nidokym derivative*** *of* $\nu$ *w.r.t.* $\mu$, *and is commonly represented by*

$$f = \frac{d\nu}{d\mu}. \tag{A.6}$$

## A.2. Stochastic processes

**Definition A.2.1.** *A* ***probability measure*** $\mathbb{P}$ *on a measurable space* $(\Omega, \mathcal{F})$ *is a finite measure such that*

$$\mathbb{P}(\Omega) = 1.$$

*The triple* $(\Omega, \mathcal{F}, \mathbb{P})$ *is called a* ***probability space****.*

**Definition A.2.2.** *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a given probability space. A function* $X : \Omega \to \mathbb{R}^n$ *is a* ***random variable*** *if it is* $\mathcal{F}$***-measurable****.*

**Definition A.2.3.** *The **probabiltiy distribution** $\mu_X$ for a random variable $X$ is a measure on a measurable space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, defined by*

$$\mu_X(B) = \mathbb{P}(X^{-1}(B)).$$

**Proposition A.2.1.** *Let $m$ denote the Lebesgue measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, i.e., $m([a,b]) = b-a$. Then, there exists a Borel-measurable function $f \in L^1(\mathbb{R})$, called the **probability density function** of the random variable $X$, such that, for every $B \in \mathcal{B}(\mathbb{R})$,*

$$\mu_X(B) = \int_B f(x)dm(x) =: \int_B f(x)dx.$$

*Proof.* Since $m(A) = 0 \implies \mu_X(A) = 0$, we are under the conditions of Theorem A.1.1 and the result follows. □

**Definition A.2.4.** *Let $X \in L^1(\Omega)$ be a random variable defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with probability density function $f$. Then, the **expected value of $X$** w.r.t. $\mathbb{P}$ is defined to be*

$$\mathbb{E}^{\mathbb{P}}[X] := \int_\Omega X(\omega)d\mathbb{P}(\omega) = \int_{\mathbb{R}} xf(x)dx. \tag{A.7}$$

**Definition A.2.5.** *The **characteristic function**, $\phi_X$, of a random variable $X$, is the Fourier transform of the probability density function $f$, i.e.,*

$$\phi_X(u) := \mathbb{E}\left[e^{iuX}\right] = \int_{\mathbb{R}} e^{iux} f(x)dx. \tag{A.8}$$

**Definition A.2.6.** *The **cumulant generating function**, $K_X$, of a random variable $X$, is defined as*

$$\begin{aligned} K_X(t) &:= \ln\left(\mathbb{E}\left[e^{tX}\right]\right) \\ &= \sum_{n=1}^{\infty} \kappa_n \frac{t^n}{n!}, \end{aligned} \tag{A.9}$$

| Model | Cumulants |
|---|---|
| GBM | $\kappa_1 = \left(b - 0.5\sigma^2\right) t$<br>$\kappa_2 = \sigma^2 t$<br>$\kappa_4 = 0$ |
| Kou | $\kappa_1 = t\left(b + \frac{\lambda p}{\eta_1} + \frac{\lambda(1-p)}{\eta_2}\right)$<br>$\kappa_2 = t\left(\sigma^2 + 2\frac{\lambda p}{\eta_1^2} + 2\frac{\lambda(1-p)}{\eta_2^2}\right)$<br>$\kappa_4 = 24t\lambda\left(\frac{p}{\eta_1^4} + \frac{1-p}{\eta_2^4}\right)$ |
| Merton JD | $\kappa_1 = t(b + \lambda\mu)$<br>$\kappa_2 = t(\sigma^2 + \lambda\mu^2 + \delta^2\lambda$<br>$\kappa_4 = t\lambda\left(\mu^4 + 6\delta^2\mu^2 + 3\delta^4\lambda\right)$ |
| VG | $\kappa_1 = t(b + \theta)$<br>$\kappa_2 = t\left(\sigma^2 + \kappa\theta^2\right)$<br>$\kappa_4 = 3t\left(\sigma^4\kappa + 2\theta^4\kappa^3 + 4\sigma^2\theta^2\kappa^2\right)$ |
| Heston | $\kappa_1 = (r-q)t + \frac{(1-e^{k^*t})(\theta^* - \nu_0)}{2k^*} - \frac{1}{2}\theta^* t$<br>$\kappa_2 = \big(\sigma t k^* e^{-k^*t}(\nu_0 - \theta^*)(8k^*\rho - 4\sigma) + k^*\rho\sigma(1 - e^{-k*t})(16\theta^* - 8\nu_0)$<br>$+2\theta^* k^* t(-4k^*\rho\sigma + \sigma^2 + 4{k^*}^2) + \sigma^2(e^{-2k^*t}(\theta^* - 2\nu_0)$<br>$+\theta^*(6e^{-k^*t} - 7) + 2\nu_0) + 8{k^*}^2(\nu_0 - \theta^*)(1 - e^{-k^*t})\big)/(8{k^*}^3)$ |

Table A.1.: Cumulants of $\ln\left(\frac{S_t}{K}\right)$ for various models.

*where $\kappa_n$ is the n-th cumulant of the random variable $X$.*

**Definition A.2.7.** *The **n-th cumulant**, $\kappa_n$, of a random variable $X$, is the n-th derivative of the cumulant generating function evaluated at zero, i.e.,*

$$\kappa_n := K_X^{(n)}(0). \tag{A.10}$$

**Definition A.2.8.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\mathcal{G} \subset \mathcal{F}$ be a $\sigma$-algebra, and let $X \in L^1(\Omega)$ be a random variable. The **conditional expectation of X given $\mathcal{G}$** under the probability measure $\mathbb{P}$ is a random variable, $\mathbb{E}^{\mathbb{P}}[X|\mathcal{G}]$, such that*

*(i)* $\mathbb{E}^{\mathbb{P}}[X|\mathcal{G}]$ *is* $\mathcal{G}$*-measurable,*

*(ii)* $\mathbb{E}^{\mathbb{P}}[\mathbf{1}_A \mathbb{E}^{\mathbb{P}}[X|\mathcal{G}]] = \mathbb{E}^{\mathbb{P}}[\mathbf{1}_A X], \quad \forall A \in \mathcal{G}.$

**Definition A.2.9.** *A* ***filtration*** *on a measurable space* $(\Omega, \mathcal{F})$ *is a family* $\mathbb{F} = \{\mathcal{F}(t) : t \geq 0\}$ *of* $\sigma$*-algebras* $\mathcal{F}(t) \subset \mathcal{F}$ *such that*

$$0 \leq s < t \implies \mathcal{F}(s) \subset \mathcal{F}(t).$$

**Definition A.2.10.** *Let* $T \subseteq \mathbb{R}_0^+$*. A* ***stochastic process*** $X = \{X_t : t \in T\} := \{X_t(\omega) : t \in T\}$ *on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *is a collection of random variables indexed by a time variable* $t$*.*

**Remark.** *When referring to random variables of a generic stochastic process* $X$*, we shall use interchangeably the notations* $X_t$ *and* $X(t) := X(t, \omega)$*, whether* $T$ *is a countable or an uncountable set.*

**Definition A.2.11.** *A real-valued stochastic process* $W$ *defined on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *is a* ***Wiener process*** *if the following holds:*

*(i)* $\mathbb{P}(W(0) = 0) = 1,$

*(ii)* $W(t+s) - W(t)$ *are independent,* $\quad \forall s, t \geq 0,$

*(iii)* $W(t) \sim \mathcal{N}(0, t), \quad \forall t > 0,$

*(iv)* $W$ *has* $\mathbb{P}$*-almost surely continuous paths.*

**Definition A.2.12.** *Let* $X$ *be a stochastic process on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*. We say that* $\mathbb{F}^X$ *is the* ***filtration generated by X*** *if*

$$\mathcal{F}(t) = \sigma(X(s) : s \leq t), \quad \forall \mathcal{F}(t) \in \mathbb{F}^X.$$

*In order words, $\mathcal{F}(t)$ is the smallest $\sigma$-algebra with respect to which all the variables $\{X(s) : s \leq t\}$ are measurable.*

**Definition A.2.13.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathbb{F}$ be a filtration on $(\Omega, \mathcal{F})$. A stochastic process $X$ is called $\mathbb{F}$-**adapted** if, for each $t \geq 0$, the random variable $X(t)$ is $\mathcal{F}(t)$-measurable. We say that $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a **filtered probability space**.*

**Definition A.2.14.** *Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space.*

*(i) The filtration $\mathbb{F}$ is said to be **complete** if the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is complete and if $\mathcal{F}(0)$ contains all the $\mathbb{P}$-null sets.*

*(ii) The filtration $\mathbb{F}$ is said to satisfy the usual conditions if it is complete and right continuous, i.e.,*

$$\mathcal{F}(t) = \bigcap_{s<t} \mathcal{F}(s), \quad \forall t > 0.$$

**Definition A.2.15.** *Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space and let $X$ be a $\mathbb{F}$-adapted stochastic process of random variables $X(t) \in L^1(\Omega)$. If*

$$\mathbb{E}^{\mathbb{P}}[X(t)|\mathcal{F}(s)] = X(s), \quad \forall s \leq t,$$

*then the process $X$ is a $\mathbb{P}$-**martingale** w.r.t. filtration $\mathbb{F}$.*

**Definition A.2.16.** *A stochastic process $X$, defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, is an **Itô process** if it can be represented as*

$$dX(t) = \mu(t, X(t)) + \sigma(t, X(t))dW(t), \tag{A.11}$$

*i.e.,*

$$X(t) = X(0) + \int_0^t \mu(s, X(s))ds + \int_0^t \sigma(s, X(s))dW(s), \tag{A.12}$$

*where $W$ is a Wiener process generating the filtration $\mathbb{F}$ and the processes $\mu$ and $\sigma$ are $\mathbb{F}$-adapted.*

**Lemma A.2.1.** *Let $X$ be an Itô process, as in Definition A.2.16, and let $f \in C^2(\mathbb{R}_0^+, \mathbb{R})$. Then,*

$$\begin{aligned} f(t, X(t)) = f(0, X(0)) &+ \int_0^t \frac{\partial f(s, X(s))}{\partial s} ds + \int_0^t \frac{\partial f(s, X(s))}{\partial x} dX(s) \\ &+ \frac{1}{2} \int_0^t \frac{\partial^2 f(s, X(s))}{\partial x^2} (dX(s))^2, \end{aligned} \tag{A.13}$$

*or, in differential notation,*

$$\begin{aligned} df(t, X(t)) = &\left[ \frac{\partial f(t, X(t))}{\partial t} + \mu(t, X(t)) \frac{\partial f(t, X(t))}{\partial x} + \frac{1}{2} \sigma^2(t, X(t)) \frac{\partial^2 f(t, X(t))}{\partial^2 x} \right] dt \\ &+ \sigma(t, X(t)) \frac{\partial f(t, X(t))}{\partial x} dW(t). \end{aligned} \tag{A.14}$$

**Lemma A.2.2 (Novikov's condition).** *Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space and let $\boldsymbol{W}^{\mathbb{P}} = [W_1^{\mathbb{P}}, ..., W_d^{\mathbb{P}}]^T$ be a $d$-dimensional Wiener process generating the filtration $\mathbb{F}$. Assume that $\boldsymbol{\lambda} = [\lambda_1, ..., \lambda_d]$ is an $\mathbb{F}$-adapted $d$-dimensional process such that*

$$\mathbb{E}^{\mathbb{P}} \left[ \exp \left\{ \frac{1}{2} \int_0^T ||\boldsymbol{\lambda}(s)||^2 ds \right\} \right] < \infty. \tag{A.15}$$

*Then, the **Doléans Dade exponential** process $\mathcal{E}$, defined by*

$$\mathcal{E}(\boldsymbol{\lambda} * \boldsymbol{W})(t) := \exp \left\{ \int_0^t \boldsymbol{\lambda}(s) d\boldsymbol{W}^{\mathbb{P}}(s) - \frac{1}{2} \int_0^t ||\boldsymbol{\lambda}(s)||^2 ds \right\}, \quad 0 \le t \le T, \tag{A.16}$$

*is a $\mathbb{P}$-martingale w.r.t. filtration $\mathbb{F}$, i.e., $\mathbb{E}^{\mathbb{P}}[\mathcal{E}(\boldsymbol{\lambda} * \boldsymbol{W})(T)] = 1$.*

**Theorem A.2.1 (Girsanov).** *Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space and let $\boldsymbol{W}^{\mathbb{P}} =$*

*$[W_1^{\mathbb{P}}, ..., W_d^{\mathbb{P}}]^T$ be a d-dimensional Wiener process generating the filtration $\mathbb{F}$. Assume that $\boldsymbol{\lambda} = [\lambda_1, ..., \lambda_d]$ is an $\mathbb{F}$-adapted d-dimensional process that satisfies Novikov's condition. Define an equivalent probability measure $\tilde{\mathbb{P}}$ on $(\Omega, \mathcal{F})$, such that the Radon-Nikodym derivative of $\tilde{\mathbb{P}}$ w.r.t. $\mathbb{P}$ at time $T$ is given by*

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}|\mathcal{F}(T) = \mathcal{E}(\boldsymbol{\lambda} * \boldsymbol{W})(T). \tag{A.17}$$

*Then, $\boldsymbol{W}^{\tilde{\mathbb{P}}}(t) := \boldsymbol{W}^{\mathbb{P}}(t) - \int_0^T \boldsymbol{\lambda}(s)ds$ is a d-dimensional Wiener process w.r.t. $\tilde{\mathbb{P}}$ and has the differential representation*

$$d\,\boldsymbol{W}^{\tilde{\mathbb{P}}}(t) = d\,\boldsymbol{W}^{\mathbb{P}}(t) - \boldsymbol{\lambda}(t)dt. \tag{A.18}$$

*Proof.* See [1, p. 165]. □

**Definition A.2.17.** *Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space and let $X$ and $\beta$ be $\mathbb{F}$-adapted stochastic processes describing the evolution in time of the price of a given financial asset and of a money market account, respectively. The probability measure $\mathbb{Q}$ is a **risk-neutral measure** (or, alternatively, an **equivalent martingale measure**) if:*

*(i) $\mathbb{Q} \sim \mathbb{P}$,*

*(ii) the discounted price process $\{\beta^{-1}(t)X(t) : t \geq 0\}$ is a $\mathbb{Q}$-martingale, i.e.,*

$$\mathbb{E}^{\mathbb{Q}}\left[\frac{X(T)}{\beta(T)}|\mathcal{F}(t)\right] = \frac{X(t)}{\beta(t)}, \quad \forall T \geq t. \tag{A.19}$$

**Theorem A.2.2** (**First Fundamental Theorem of Asset Pricing**). *A market is free of arbitrage if and only if there exists a risk-neutral measure $\mathbb{Q}$.*

*Proof.* See [1, p. 33]. □

**Theorem A.2.3** (**Second Fundamental Theorem of Asset Pricing**). *Assume that the market is free of arbitrage. Then, the market is complete if and only if the risk-neutral measure $\mathbb{Q}$ is unique.*

*Proof.* See [1, p. 151]. □

## A.3. Fourier Analysis

**Definition A.3.1.** *A family of continuous functions* $\{f_k\}_{k\in\mathbb{N}}$ *is called an* ***orthogonal family of norm-square*** $L$ *on* $[-L, L]$ *if, for* $m, n \in \mathbb{N}$*, we have*

$$(f_m, f_n) = \begin{cases} 0, & m \neq n, \\ L, & m = n. \end{cases} \tag{A.20}$$

**Proposition A.3.1.** *The families of functions* $\{s_k\}_{k\in\mathbb{N}}$ *and* $\{c_k\}_{k\in\mathbb{N}}$*, with* $s_k(x) := \sin\left(\frac{k\pi}{L}x\right)$ *and* $c_k(x) := \cos\left(\frac{k\pi}{L}x\right)$*, are orthogonal of norm-square* $L$ *on* $[-L, L]$*.*

*Proof.* Following Definition A.1.11,

$$\begin{aligned}
(s_m, s_n) &= \int_{-L}^{L} \sin\left(\frac{m\pi x}{L}\right) \sin\left(\frac{n\pi x}{L}\right) dx \\
&= \frac{1}{2}\int_{-L}^{L} \left[\cos\left(\frac{(m-n)\pi x}{L}\right) - \cos\left(\frac{(m+n)\pi x}{L}\right)\right] dx \\
&= \begin{cases} \frac{1}{2}\left[x\right]_{-L}^{L} - \frac{L}{4m\pi}\left[\sin\left(\frac{2m\pi x}{L}\right)\right]_{-L}^{L}, & m = n, \\ \frac{L}{2(m-n)\pi}\left[\sin\left(\frac{(m-n)\pi x}{L}\right)\right]_{-L}^{L} - \frac{L}{2(m+n)\pi}\left[\sin\left(\frac{(m+n)\pi x}{L}\right)\right]_{-L}^{L}, & m \neq n \end{cases} \\
&= \begin{cases} L, & m = n, \\ 0, & m \neq n, \end{cases}
\end{aligned}$$

and

$$
\begin{aligned}
(c_m, c_n) &= \int_{-L}^{L} \cos\left(\frac{m\pi x}{L}\right)\cos\left(\frac{n\pi x}{L}\right) dx \\
&= \frac{1}{2}\int_{-L}^{L}\left[\cos\left(\frac{(m-n)\pi x}{L}\right) + \cos\left(\frac{(m+n)\pi x}{L}\right)\right] dx \\
&= \begin{cases} \frac{1}{2}\left[x\right]_{-L}^{L} + \frac{L}{4m\pi}\left[\sin\left(\frac{2m\pi x}{L}\right)\right]_{-L}^{L}, & m = n, \\ \frac{L}{2(m-n)\pi}\left[\sin\left(\frac{(m-n)\pi x}{L}\right)\right]_{-L}^{L} + \frac{L}{2(m+n)\pi}\left[\sin\left(\frac{(m+n)\pi x}{L}\right)\right]_{-L}^{L}, & m \neq n \end{cases} \\
&= \begin{cases} L, & m = n, \\ 0, & m \neq n. \end{cases}
\end{aligned}
$$

□

**Proposition A.3.2.** *The functions $c_m$ and $s_n$ are orthogonal for $m, n \in \mathbb{N}$.*

*Proof.* Following Definition A.1.11,

$$
\begin{aligned}
(c_m, s_n) &= \int_{-L}^{L} \cos\left(\frac{m\pi x}{L}\right)\sin\left(\frac{n\pi x}{L}\right) dx \\
&= \frac{1}{2}\int_{-L}^{L}\left[\sin\left(\frac{(m+n)\pi x}{L}\right) - \sin\left(\frac{(m-n)\pi x}{L}\right)\right] dx \\
&= \begin{cases} -\frac{L}{4m\pi}\left[\cos\left(\frac{2m\pi x}{L}\right)\right]_{-L}^{L}, & m = n, \\ -\frac{L}{2(m+n)\pi}\left[\cos\left(\frac{(m+n)\pi x}{L}\right)\right]_{-L}^{L} + \frac{L}{2(m-n)\pi}\left[\cos\left(\frac{(m-n)\pi x}{L}\right)\right]_{-L}^{L}, & m \neq n \end{cases} \\
&= \begin{cases} 0, & m = n, \\ 0, & m \neq n. \end{cases}
\end{aligned}
$$

□

**Theorem A.3.1.** *Let $f \in L^1([-L, L])$. If $f$ is of the form*

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} \left[ a_n \cos\left(\frac{n\pi x}{L}\right) + b_n \sin\left(\frac{n\pi x}{L}\right) \right], \tag{A.21}$$

*then the coefficients $a_n$ and $b_n$ are uniquely determined by the formulas*

$$a_n = \frac{1}{L}\int_{-L}^{L} f(x) \cos\left(\frac{n\pi x}{L}\right) dx = L^{-1}(f, c_n), \quad n \in \mathbb{N}_0, \tag{A.22}$$

$$b_n = \frac{1}{L}\int_{-L}^{L} f(x) \sin\left(\frac{n\pi x}{L}\right) dx = L^{-1}(f, s_n), \quad n \in \mathbb{N}, \tag{A.23}$$

*and we say that the series in (A.21) is called the **Fourier series** of $f$, whereas the coefficients $a_n$ and $b_n$ given by equations (A.22) and (A.23), respectively, are called the **Fourier coefficients** of $f$.*

*Proof.* For $n \in \mathbb{N}_0$, we have

$$(f, c_m) = (\frac{1}{2}a_0, c_m) + \sum_{n=1}^{\infty} \left[ a_n(c_n, c_m) + b_n(s_n, c_m) \right], \tag{A.24}$$

$$(f, s_m) = (\frac{1}{2}a_0, s_m) + \sum_{n=1}^{\infty} \left[ a_n(c_n, s_m) + b_n(s_n, s_m) \right]. \tag{A.25}$$

Using the fact that

$$(c_n, c) = (s_n, c) = 0,$$

for $c \in \mathbb{R}$, and recalling Propositions A.3.1 and A.3.2, we get

$$(f, c_m) = a_m L, \tag{A.26}$$

$$(f, s_m) = b_m L. \tag{A.27}$$

Since $b_0 = 0$, by equation (A.27), equations (A.22) and (A.23) follow. □

**Theorem A.3.2** (**Bessel's inequality**). *Let $f \in L^2([-L, L])$. Assume that the Fourier coefficients $a_n$ and $b_n$ of $f$ exist and are defined as in equations (A.22) and (A.23), respectively. Then,*

$$\frac{1}{2}a_0^2 + \sum_{n=1}^{\infty}(a_n^2 + b_n^2) \leq L^{-1}(||f||_2)^2. \tag{A.28}$$

*Proof.* See [3, p. 209]. □

**Corollary A.3.1** (**Riemann-Lebesgue Lemma**). *The Fourier coefficients $a_n$ and $b_n$ tend to zero as $n \to \infty$.*

*Proof.* By Theorem A.3.2, $a_n^2$ and $b_n^2$ are the $n$-terms of convergent series, so they tend to zero as $n \to \infty$; hence, $a_n$ and $b_n$ must also tend to zero as $n \to \infty$. □

**Definition A.3.2.** *Let $f$ be a function defined on $[0, L]$. The function $f_e$, defined as*

$$f_e(x) = \begin{cases} f(x), & 0 \leq x \leq L, \\ f(-x), & -L \leq x < 0, \end{cases} \tag{A.29}$$

*is called the **even extension** of $f$ on $[-L, L]$.*

**Proposition A.3.3.** *Let $f \in L^1([-L, L])$, with Fourier coefficients given by equations (A.22) and (A.23). If $f$ is even, then:*

*(i)* $a_n = \frac{2}{L}\int_0^L f(x)\cos\left(\frac{n\pi x}{L}\right)dx, \quad n \in \mathbb{N}_0,$

*(ii)* $b_n = 0, \quad n \in \mathbb{N}.$

*Proof.* See [3, p. 240]. □

**Proposition A.3.4.** *Let $f \in L^1([0, L])$ and let $f_e$ be its even extension on $[-L, L]$. The*

***Fourier cosine series*** *expansion of* $f$ *is given by*

$$f(x) = f_e(x) = \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi x}{L}\right), \tag{A.30}$$

*where*

$$a_n = \frac{2}{L}\int_0^L f(x)\cos\left(\frac{n\pi x}{L}\right) dx. \tag{A.31}$$

# Bibliography


[1] T. Björk. *Arbitrage Theory in Continuous Time*. Oxford University Press, 3rd edition, 2009.

[2] F. Black and M. Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81(3):637–654, 1973.

[3] D. Bleecker and G. Csordas. *Basic Partial Differential Equations*. International Press, 2003.

[4] M. Broadie and Ö. Kaya. Exact simulation of stochastic volatility and other affine jump diffusion processes. *Oper. Res.*, 54(2):217–231, 2006.

[5] M. Capiński and E. Kopp. *Measure, Integral and Probability*. Springer, 2nd edition, 2004.

[6] P. P. Carr and D. B. Madan. Option valuation using the fast Fourier transform. *J. Comp. Finance*, 2:61–73, 1999.

[7] R. Cont and P. Tankov. *Financial Modelling with Jump Processes*. Chapman and Hall/CRC, 2004.

[8] J. C. Cox, J. E. Ingersoll, and S.A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–408, 1985.

[9] F. Fang and C. W. Oosterlee. A novel pricing method for European options based on Fourier-cosine series expansions. *SIAM J. Sci. Comput.*, 31(2):826–848, 2008.

[10] F. Fang and C. W. Oosterlee. Pricing early-exercise and discrete barrier options by Fourier-cosine series expansions. *Numerische Mathematik*, 114:27–62, 2009.

[11] F. Fang and C. W. Oosterlee. A Fourier-based valuation method for Bermudan and barrier options under Heston's model. *SIAM J. Financial Math*, 2:439–463, 2011.

[12] W. Feller. Two singular diffusion problems. *Annals of Mathematics*, 54(1):173–182, 1951.

[13] S. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2):327–343, 1993.

[14] K. Ito and J. Toivanen. Lagrange multiplier approach with optimized finite difference stencils for pricing American options under stochastic volatility. *SIAM J. Sci. Comput.*, 31:2646–2664, 2009.

[15] S. G. Kou. A jump-diffusion model for option pricing. *Manage. Sci.*, 48(8):1086–1101, 2002.

[16] R. Lord, F. Fang, F.Bervoets, and C. W. Oosterlee. A fast and accurate fft-based method for pricing early-exercise options under Lévy processes. *SIAM J. Sci. Comput.*, 30:1678–1705, 2008.

[17] D. B. Madan, P. P. Carr, and E. C. Chang. The variance gamma process and option pricing. *European Finance Review*, 2:79–105, 1998.

[18] R. C. Merton. Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1):141–183, 1973.

[19] R. C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1-2):125–144, 1976.