SAPIENZA
UNIVERSITÀ DI ROMA

Some Extensions of the Black-Scholes and Cox-Ingersoll-Ross Models

Dipartimento Memotef: Modelli per l'Economia e la Finanza

Dottorato di Ricerca in Matematica per le Applicazioni Economico/Finanziarie – XXXI Ciclo

Candidate

Michele Bufalo
ID number 1702550

Thesis Advisor

Prof. Rosa Maria Mininni

A thesis submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Mathematics for Economic/Financial Applications

October 2018

Thesis defended on 25 February 2019
in front of a Board of Examiners composed by:

- Prof. Marusca De Castris
- Prof. Fabrizio Durante
- Prof. Barbara Martini

(chairman)

Some Extensions of the Black-Scholes and Cox-Ingersoll-Ross Models
Ph.D. thesis. Sapienza – University of Rome

This thesis has been typeset by LaTeX and the Sapthesis class.

Author's email: michele.bufalo@uniroma1.it

Dedicated to
my family

Contents

Introduction

In this thesis we will study some financial problems concerning the option pricing in complete and incomplete markets and the bond pricing in the short-term interest rates framework. We start from well known models in pricing options or zero-coupon bonds, as the Black-Scholes model [12] (see also R. C. Merton [67], [68]) and the Cox-Ingersoll-Ross model [25] and study some their generalizations. In particular, in the first part of the thesis, we study a generalized Black-Scholes equation to derive explicit or approximate solutions of an option pricing problem in incomplete market where the incompleteness is generated by the presence of a non-traded asset. As explained by L. E. O. Svensson in [89]: "In the literature of international finance, the existence of income from non-traded assets seems to be the rule rather than the expectation. The existence of non-traded assets could be a result of asset market imperfections, which in turn are caused by the usual reasons: transaction costs, moral hazard, legal restrictions, etc. As examples we can think of an individual who cannot trade claims to his future wages for obvious moral hazard reasons; a government which cannot trade claims to future tax receipts; or a country which cannot trade claims to its gross domestic product (GDP) in world capital markets". Finding solutions to portfolio problems in a continuous time model when there is some income from non-traded assets, can be seen as the problem of pricing and hedging in incomplete markets, where incompleteness is generated just by non-traded assets that prevent the creation of perfectly replicating portfolio. Thus the evaluation based on replication and no arbitrage assumptions is no longer possible and new strategies to price and hedge derivatives that are written on such securities are needed (see, for example, M. Musiela and T. Zariphopoulou [71] and references therein). A pricing methodology is based on utility maximization criteria which produce the so called indifference price. By this approach the price is not determined with respect to the risk neutral measure as in a complete setting, but with respect to an indifference measure: "describing the historical behavior of the non-traded asset. In fact, it refers to which is defined as the closest to the risk neutral one and, at the same time, capable of measuring the unhedgeable risk, by being defined on the filtration of the Brownian motion used for the modelling of the non-traded asset dynamics" (M. Musiela and T. Zariphopoulou [71, p.3]). The indifference pricing problem and the underlying utility-optimization problem are well characterized by martingale duality results (see M. Frittelli [37]), by stochastic differential equations (see R. Rouge and N. El Karoui [85]) and by non-linear partial differential equations (see L. A. Bordag and R. Frey [13]). The utility-based price and the hedging strategy can be described by the solution of a partial differential equation, in analogy to the Black Scholes model, but it is more difficult to obtain the explicit solution in specific models. M. Musiela and T. Zariphopoulou in [71] derived the indifference price of a European claim, written exclusively on the non-traded asset, as a non-linear expectation of the derivative's payoff under an appropriate martingale measure. Our aim (see also [18]) is to give a closed form representation of the indifference price by using the analytic tool of (C_0) semigroup theory, which allows to study the evolution in time of some problems coming from Mathematical Physics, Mathematical Finance and other applied sciences (see e.g. K. J. Engel and R. Nagel [32], J. A. Goldstein [44], and J. A. Goldstein et al. [42]). We will focus on the abstract Cauchy problem associated with operators of the type

$$\mathcal{L}u = \frac{1}{2}(cx+d)^2u'' + [(\gamma x + \delta) - \theta(cx+d)]u', \tag{0.1}$$

acting on the space of all real-valued continuous functions in a suitable real interval having finite limits at the endpoints.

The second part of the thesis deals with the problem of forecasting future interest rates from observed financial market data by a suitable partition of the dataset. We propose a new numerical methodology for the CIR framework, which we call the *CIR# model*, that well fits the term structure of short interest rates as observed in a real market. J. C. Cox, J. E. Ingersoll and S. A. Ross [25] proposed a term structure model, well known as the CIR model, to describe the price of discount zero-coupon bonds with various maturities under no-arbitrage condition. This model generalizes the Vasicek model [92] to the case of non constant volatility and assumes that the evolution of the underlying short term interest rate is a diffusion process, i.e. a continuous Markov process, unique solution to the following stochastic differential equation (SDE)

$$dr(t) = [k(\theta - r(t)) - \lambda(t, r(t))]dt + \sigma\sqrt{r(t)}dW(t), \tag{0.2}$$

with initial condition $r(0) = r_0 > 0$. $(W(t))_{t\geq 0}$ denotes a standard Brownian motion under the risk neutral probability measure, intended to model a random risk factor. The interest rate process $r = (r(t))_{t\geq 0}$ solution to (0.2) is usually known as the *CIR process* or *square root process*. The SDE (0.2) is classified as a one-factor time-homogeneous model, because the parameters k, θ and σ, are time-independent and the short interest rate dynamics is driven only by the market price of risk $\lambda(t, r(t)) = \lambda r(t)$, where λ is a constant. Therefore the SDE (0.2) is composed of two parts: the "mean reverting" drift component $k[\theta - r(t)]$, which ensures the rate $r(t)$ is elastically pulled towards a long-run mean value $\theta > 0$ at a speed of adjustment $k > 0$, and the random component $W(t)$, which is scaled by the standard deviation $\sigma\sqrt{r(t)}$. The volatility parameter is denoted by $\sigma > 0$.

The paths of the CIR process never reach negative values and their behaviour depends on the relationship between the three constant positive parameters k, θ, σ. Indeed, it can be shown (see, e.g., M. Jeanblanc, M. Yor and M. Chesney [54, Chapter III]) that if the condition $2k\theta \geq \sigma^2$ is satisfied, then the interest rates $r(t)$ are strictly positive for all $t > 0$, and, for small $r(t)$, the process rebounds as the random perturbation dampens with $r(t)$ vanishing to zero. Furthermore, the CIR process belongs to the class of processes satisfying the "affine property", i.e., the logarithm of the characteristic function of the transition distribution of such processes is an affine (linear plus constant) function with respect to their initial state (for more details the reader can refer to D. Duffie, D. Filipović and W. Schachermayer [30, Section 2]). As a consequence, the non-arbitrage price of a discount zero coupon bond with maturity $T > 0$ and underlying interest rate dynamics described by a CIR process, is given by

$$P(T-t, r(t)) = A(T-t)e^{-B(T-t)r(t)}, \quad t \in [0, T], \tag{0.3}$$

where $A(\cdot)$ and $B(\cdot)$ are deterministic functions (see [25]). The final condition is $P(T, r(T)) = 1$, which corresponds to the nominal value of the bond conventionally set equal to 1 (monetary unit). Since bonds are commonly quoted in terms of yields rather than prices, the formula (0.3) allows derivation of the yield-to-maturity curve

$$\begin{aligned} Y(T-t, r(t)) &= -\ln P(t, T, r(t))/(T-t) \\ &= [B(T-t)r(t) - \ln(A(T-t))]/(T-t), \quad t \in [0, T], \end{aligned}$$

which tends to the asymptotic value $\Upsilon = 2k\theta/(\gamma + k + \lambda)$ as $T \to \infty$, where $\gamma = \sqrt{(k+\lambda)^2 + 2\sigma^2}$.

The CIR model became very popular in finance among practitioners because it does not allow for negative rates and introduces a rate dependent volatility. Further the CIR model was attractive for its relatively handy implementation and analytical tractability. Other applications of the model (0.2) include stochastic volatility modelling in option pricing problems (see e.g. G. Orlando and G. Taglialatela [79], S. L. Heston [49]), or default intensities in credit risk (see D. Duffie [29]).

However, the CIR model fails as a satisfactory calibration to market data since it depends on a small number of constant parameters, k, θ and σ. As explained by D. Brigo and F. Mercurio [15, Section 3.2]: "the zero coupon curve is quite likely to be badly reproduced, also because some typical shapes, like that of an inverted yield curve, may not be reproduced by the model,... no matter the values of the parameters in the dynamics that are chosen". J. C. Cox, J. E. Ingersoll and S. A. Ross [25], instead, mentioned explicitly that the model can "produce only normal, inverse or humped shapes" . The seeming contradiction lies in the practical implementation.

As proved by M. Keller-Ressel and T. Steiner [59, Theorem 3.9 and Section 4.2] the yield-to-maturity curve of any time-homogenous, affine one-factor model is either normal (i.e., a strictly increasing function of $T-t$), humped (i.e., with one local maximum and no minimum on $]0, \infty[$) or inverse (i.e., a strictly decreasing function of $T-t$) (see Figure 0.1). For the CIR process, the yield is normal when $r(t) \leq k\theta/(\gamma - 2(k+\lambda))$, while it is inverse when $r(t) \geq k\theta/(k+\lambda)$. For intermediate values the yield curve is humped. However, R. A. Carmona and M. R. Tehranchi [20, Section 2.3.5] explained that: "Tweaking the parameters can produce yield curves with one hump or one dip (a local minimum), but it is very difficult (if not impossible) to calibrate the parameters so that the hump/dip sits where desired. There are not enough parameters to calibrate the models to account for observed features contained in the prices quoted on the markets" (see Figure 0.2).

Thus the need for more sophisticated models fitting closely to the currently-observed yield curve, which could take into account multiple correlated sources of risk as well as shocks and/or structural changes of the market, has lead to the development of

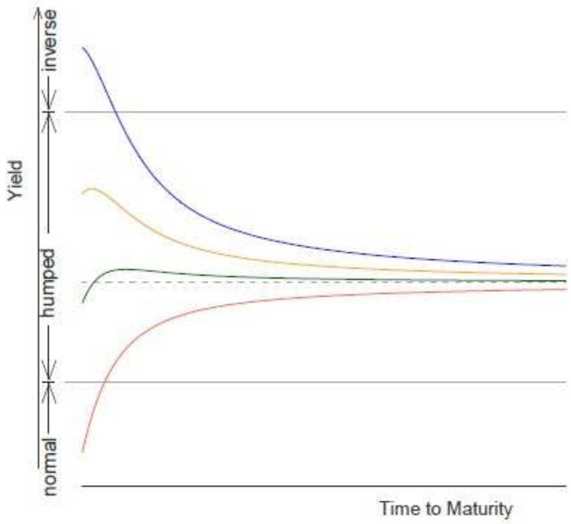
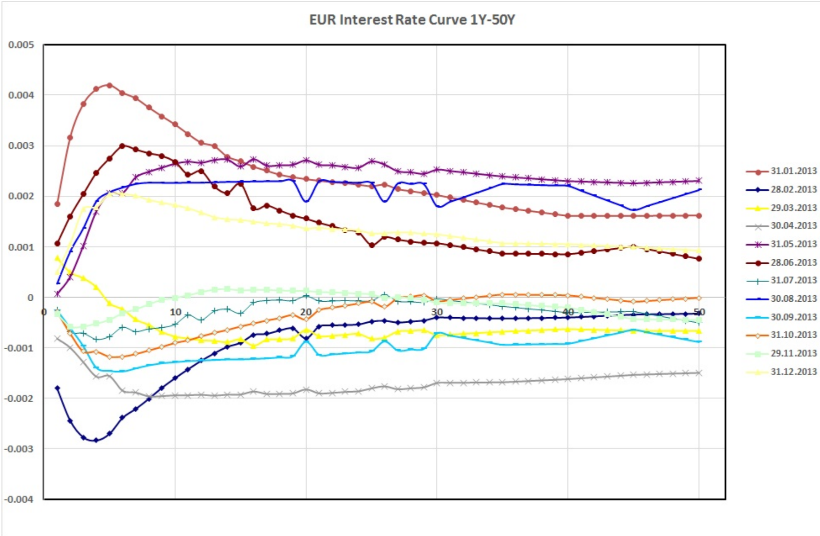
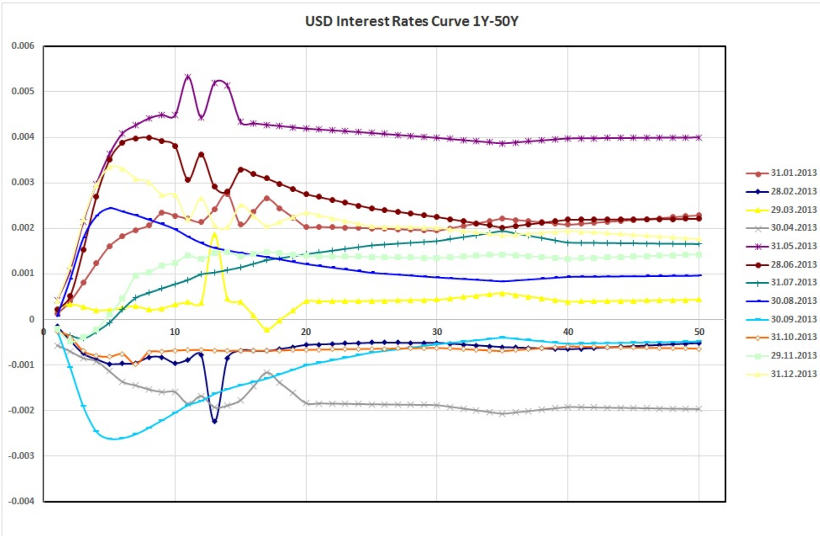


Figure 0.1. Inverse, humped or normal yield-to-maturity curves (M. Keller-Ressel and T. Steiner [59, Figure 1]).

a number of papers for pricing interest rate derivatives based on stochastic short rates interest models that generalize the classical CIR paradigm. Among the best known we mention: the Hull-White model [50] based on the idea of considering time-dependent coefficients; the Chen three-factor model [22]; the CIR++ model by D. Brigo and F. Mercurio [15] that considers short rates shifted by a deterministic function chosen to fit exactly the initial term structure of interest rates; the jump diffusion JCIR model (see D. Brigo and F. Mercurio [17]) and JCIR++ by D. Brigo and N. El-Bachir [16] where jumps are described by a time-homogeneous Poisson process; the CIR2 and CIR2++ two-factor models (see D. Brigo and F. Mercurio [17]). Very recently, L. Zhu [93], in order to incorporate the default clustering effects, proposed a CIR process with jumps modelled by a Hawkes process (which is a point process that has self-exciting property and the desired clustering effect), M. Moreno et al. [70] presented a cyclical square-root model, and A. R. Najafi et al. [73], [74] proposed some extensions of the CIR model where a mixed fractional Brownian motion applies to display the random part of the model.
Note that all the above cited extensions preserve the positivity of interest rates, in some cases through reasonable restrictions on the parameters. But the financial crisis of 2008 and the ensuing quantitative easing policies brought down interest rates, as a consequence of reduced growth of developed economies, and accustomed markets to unprecedented negative interest regimes under the so called "new normal". As observed in K. C. Engelen [33] and BIS (2015) [11]: "Interest rates have been extraordinarily low for an exceptionally long time, in nominal and inflation-adjusted terms, against any benchmark"(see Figure 0.3). "Between December 2014 and end-May 2015, on average around $2 trillion in global long-term sovereign debt, much of it issued by euro area sovereigns, was trading at negative yields", "such yields are unprecedented. Policy rates are even lower than at the peak of the Great Financial Crisis in both nominal and real terms. And in real terms they have now



(a) EUR term structure through 2013.



(b) USD term structure through 2013.

Figure 0.2. EUR and USD term structure (1Y-50Y) as observed on monthly basis from January to December 2013.

been negative for even longer than during the Great Inflation of the 1970s. Yet, exceptional as this situation may be, many expect it to continue". "Such low rates are the most remarkable symptom of a broader malaise in the global economy: the economic expansion is unbalanced, debt burdens and financial risks are still too high, productivity growth too low, and the room for manoeuvre in macroeconomic policy too limited. The unthinkable risks becoming routine and being perceived as the new normal."

Therefore, the need for adjusting short term interest rate models for negative rates has become an additional characteristic that a "good" model should possess. It is worth noting that the main drawback of the Vasicek model [92], which allows for

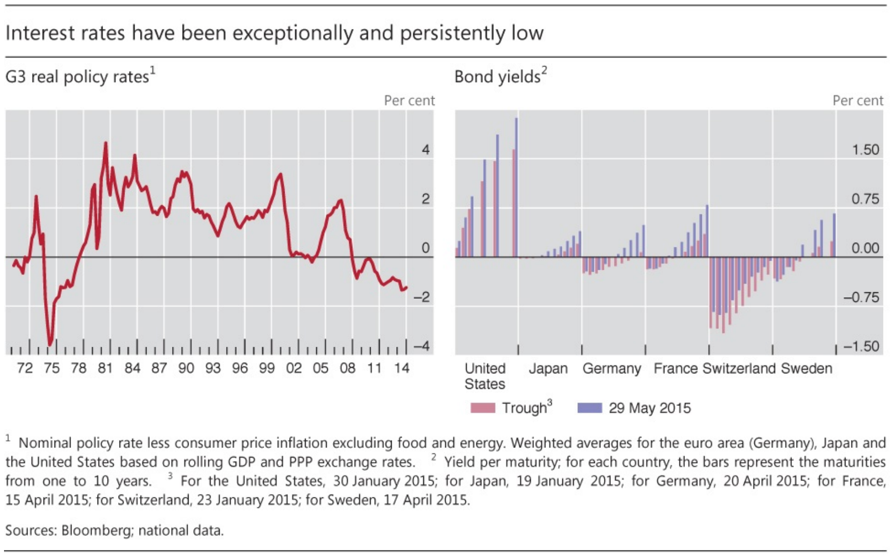


Figure 0.3. BIS 85th Annual Report 2015.

negative interest rates, is that the conditional volatility of changes in the interest rate is constant, independent on the level of it, and this may unrealistically affect the prices of bonds (see L. C. G. Rogers [84]). For this reason, the Vasicek model is unused by practitioners.
Hence, let us recap the main issues of the original CIR model:

- **i.** Negative interest rates are precluded;
- **ii.** The diffusion term in (0.2) goes to zero when $r(t)$ is small (in contrast with market data);
- **iii.** The instantaneous volatility σ is constant (in real life σ is calibrated continuously from market data);
- **iv.** There are no jumps (e.g. caused by government fiscal and monetary policies, by release of corporate financial results, etc.);
- **v.** There is not a satisfactory calibration at each time to market data since it depends on a small number of constant parameters;
- **vi.** Risk premia are linear with interest rates (false if credit worthiness of a counterparty and market volatility are considered);
- **vii.** The change in interest rates depends only on the market risk.

Our contribute is to provide a new methodology that gives an answer to points **i.-v.** by preserving the structure of the original CIR model to describe the dynamics of spot interest rates observed in financial markets.
The thesis is organized as follows:

- Chapter 1 is dedicated to the (C_0) semigroup theory. We recall first of all the main theorems and results (see K. J. Engel and R. Nagel [32] and J. A.

Goldstein [44]). Then we describe some financial application related to the semigroup governing the Black-Scholes and CIR partial differential equations (see G. R. Goldstein et al. [40], J. A. Goldstein et al. [42], [43]);

- Chapter 2 describes the results reported in on our paper [18]. A generalized Black-Scholes type operator is studied by using the (C_0) semigroup theory to derive explicit or approximate solutions of an option pricing problem in incomplete market proposed by M. Musiela and T. Zariphopoulou in [71]. In particular, approximate solutions expressed in terms of a generalized Feynman-Kac type formula are given when explicit closed-form solutions are not available;

- Chapter 3 and Chapter 4 introduce a new methodologies to forecast future interest rates from observed financial market data by a suitable partition of the dataset. In Chapter 3 (see also [76] and [78]) we will focus on forecasting future expected interest rates in the framework of the CIR model. In particular, we improve the performances of the model by a suitable partion of the available market data samples and a proper translation of the interest rate values. This to shift to positive values negative or near-to-zero interest rates and to capture all the statistically significant changes of variance, respectively. Moreover, to calibrate the parameters model, we use the martingale estimating functions approach (see B. M. Bibby et al. [10, Example 5.4]) which provides better results than the usual maximul likelhood method for discretely observed diffusion processes. Empirical results on observed market data with different maturities in EUR currency are also shown and compared with the results obtained by the Exponentially Weighted Moving Average model.
  Chapter 4 describes a new model, named $CIR\#$ model (see also [77]) in which we will show that this model improves the results obtained in Chapter 3. To this end, we again consider a suitable partition and and a proper translation to positive values of the available market data samples. The innovation in our procedure is to replace the classical Brownian motion as a noise source perturbing the time dynamics of the interest rate process, with the standardized residuals of the "optimal" ARIMA model selected for each sub-sample partitioning the whole observed data sample. In this way, we calibrate the parameters and simulate exact trajectories of the CIR process by a strong convergent discretization scheme. Finally we forecast the future interest rates and compare our results with the ones obtained by the classical CIR model. The considered dataset is available from IBA [52] and consider both EUR and USD currencies;

- Chapter 5 considers the problem of solving abstract Cauchy problems in the non-autonomous case, i.e. when the coefficients of the models are time-dependent. We will introduce the basic semigroup theory to solve this type of problems involving the concept of evolution families. Furtehr, we consider the approximation splitting formulas provided by A. Batkai et al. [8], which are very helpful when the explicit solution can not be directly computed. We study the non-autonomous abstract Cauchy problem associated with the Black-Scholes equation and give some hints to study the non-autonomous CIR problem.

Chapter 1

Semigroups of Operators and Applications in Finance

This chapter will summarize some important results about the operator semigroup theory that will be also used in Chapter 2 and Chapter 5. The aim of this analytic tool is to solve the abstract Cauchy problem (ACP)

$$\begin{cases} u_t(t,x) = Au(t,x), & t \geq 0,\, x \in J, \\ u(0,x) = u_0(x), & x \in J, \end{cases} \tag{1.1}$$

in some spaces of continuous functions on $J \subseteq \mathbb{R}$, where A is a suitable differential operator and $u_0 \in D(A) \subseteq X$ is fixed.
In Section 1.1 we present the standard notations and the classic theorems, while Section 1.2 concerns the main financial applications (see for instance [40], [42] and [43]).

1.1 Semigroups of Operators Theory

In this section, based on [32], [44], we recall some basic definitions and properties of functional analysis and the main results concerning semigroups of linear operators. Let $(X, \|\cdot\|)$ be a Banach space on $\mathbb{K}$ (i.e. $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$) and $T : X \to X$ a linear operator.

Definition 1.1.1. The linear operator T is bounded, and it is denoted by $T \in \mathcal{L}_{\mathbb{K}}(X)$, if there exists $M \geq 0$ such that for all $x \in X$ one has $\|T(x)\| \leq M\|x\|$.

Let $A : D(A) \to X$ a linear operator, with $D(A) \subseteq X$, and introduce the following subspaces

$$R(A) = \{y \in X | \exists x \in D(A)\ such\, that\, y = A(x)\},$$

$$\mathcal{G}(A) = \{(x,y) \in X \times X | x \in D(A), y = A(x)\},$$

$$\rho(A) = \{\lambda \in \mathbb{K} | (\lambda I - A)\, is\, bijective\},$$

where I denotes the identity operator. $D(A)$, $R(A)$, $\mathcal{G}(A)$, $\rho(A)$ are called respectively the domain, the rank, the graph and the resolvent of the operator A. Moreover, the operator $R(\lambda, A) = (\lambda I - A)^{-1}$ is said to be the resolvent operator associated to A.

Definition 1.1.2. The operator A is closed if $\mathcal{G}(A)$ is a closed set of $X \times X$. In particular, the following characterization holds

$$(A \textit{ is closed}) \iff (\forall (x_n)_{n\in\mathbb{N}} \in D(A) \textit{ such that } \lim_{n\to+\infty} x_n = x \in X,$$

$$\lim_{n\to+\infty} A(x_n) = y \in X \Rightarrow x \in D(A),\ A(x) = y).$$

If A is closed, then the closure $\overline{A}$ coincides with A (for more details see [32, Appendix B]).

Now, we introduce the definition of a (C_0) semigroup and its generator.

Definition 1.1.3. ((C_0) Semigroup of Operators)
Given $(T(t))_{t\geq 0} \in \mathcal{L}_{\mathbb{K}}(X)$, if the following properties hold

i $T(0) = I$,

ii $\forall t, s \in \mathbb{R}_+$, $T(s+t) = T(t) \cdot T(s)$,

$(T(t))_{t\geq 0}$ is said to be a semigroup of operators on X.
The semigroup $(T(t)_{t\geq 0}$ is strongly continuous, or (C_0), if for all $t_0 \in \mathbb{R}_+$ one has

$$\lim_{t\to t_0} \|T(t)f - T(t_0)f\|,$$

for all $f \in X$. Finally, if all these properties hold for $\mathbb{R}$ istead of $\mathbb{R}_+$ (with $\mathbb{R}_+ = [0, +\infty)$), $(T(t))_{t\in\mathbb{R}}$ is said to be a (C_0) group of operators on X.

Definition 1.1.4. (Semigroup Generator)
Let $(T(t))_{t\geq 0}$ be a (C_0) semigroup on X. Define the subspace

$$D(A) = \left\{ f \in X \middle| \exists \lim_{t\to 0^+} \frac{T(t)f - f}{t} \in X \right\},$$

and define

$$Af = \lim_{t\to 0^+} \frac{T(t)f - f}{t}.$$

The operator $(A, D(A))$ (i.e. A with domain $D(A)$) is said to be the generator of $(T(t))_{t\geq 0}$.

Remark 1.1.5. The definition of the generator $(A, D(A))$ of a (C_0) group on X is analogous to Definition 1.1.4. In particular, given a (C_0) group $(T(t))_{t\in\mathbb{R}}$ with generator $(A, D(A))$ we can define $T_+(t) = T(t)$ for $t \geq 0$ and $T_-(t) = T(-t)$ for $t < 0$. Then it is clear that $T_+(t)$ and $T_-(t)$ are (C_0) semigroups with generators $(A, D(A))$ and $(-A, D(A))$, respectively. Therefore, if $(A, D(A))$ is the generator of a group, then both $(A, D(A))$ and $(-A, D(A))$ generate (C_0) semigroups. Moreover, the converse of this statement is also true (for more details see [32, Section 2.3]).

Now, we recall the definition of some important Banach spaces, that we will use in the rest of this chapter and in Chapters 2 and 5. Let $J \subseteq \mathbb{R}$, we define:

- $C(J)$ the space of all real-valued continuous functions f defined on J,

- $BUC(J)$ the space of all real-valued, bounded and continuous functions f defined on J,
- $C^k(J)$ ($k \in \mathbb{N}$, $k \geq 1$) the space of all real-valued continuous and k-times differentiable functions f defined on J, where each derivative is continuous,
- $C_0(J)$ the space of all real-valued continuous functions defined on $\overline{J}$, which vanish at ∂, where ∂ denotes the point at infinity, if J is not compact. Observe that if $J = [0, +\infty)$, we have $\overline{J} = J \cup \partial = [0, +\infty]$, and $C_0(J) = \{f \in C(\overline{J}) | \lim_{x \to +\infty} f(x) = 0\}$.
- $C(\overline{J})$ is the space of all real-valued continuous functions f on J, having finite limits at the endpoints not included in J,

In particular, given a Banach space X, we define

- $C(J, X)$ the space of all X-valued continuous functions f defined on J,
- $BUC(J, X)$ the space of all X-valued bounded and continuous functions f defined on J,
- $C_0(J, X)$ the space of all X-valued continuous functions f defined on $\overline{J}$, which vanish at ∂, where ∂ denotes the point at infinity, if J is not compact.
- $C^k(J, X)$ ($k \in \mathbb{N}$, $k \geq 1$) the space of all X-valued continuous and k-times differentiable functions f defined on $\overline{J}$, where each derivative is continuous,

Example 1.1.6. Let $X = C_0(\mathbb{R}_+)$, the Banach space endowed with the norm $\|f\| = \sup_{x \in \mathbb{R}_+} |f(x)|$. For all $t \geq 0$, $f \in X$, $x \in \mathbb{R}_+$, define $T(t) : X \to X$ such that $(T(t)f)(x) = f(x + t)$. It is simple to see that $T(t)$ is a linear and bounded operator for all t. We have that $(T(t))_{t \geq 0}$ is a (C_0) semigroup on X with generator $(A, D(A))$, where

$$Af = f', \quad D(A) = \{f \in X | f' \in X\}.$$

Example 1.1.7. Let $A \in \mathcal{L}_{\mathbb{K}}(X)$. It is easy to prove that the series $\sum_{n=0}^{+\infty} \frac{t^n A^n}{n!}$ converges. Denote $e^{tA} = \sum_{n=0}^{+\infty} \frac{t^n A^n}{n!} \in \mathcal{L}_{\mathbb{K}}(X)$, for all $t \geq 0$. Thus, the family of operators $(T(t))_{t \geq 0}$ defined as $T(t) = e^{At}$ is a (C_0) semigroup on X, and its generator coincides with A (for more details see [32, Section 1.2]).

Proposition 1.1.8. *Let $(T(t))_{t \geq 0}$ be a (C_0) semigroup on X with generator $(A, D(A))$ on X. The following properties hold*

i. $\exists M \geq 1$, $\omega \in \mathbb{R}$ *such that* $\forall t \geq 0$, $\|T(t)\| \leq Me^{\omega t}$. *If $M = 1$ the semigroup is quasi-contractive. If $M = 1$ and $\omega = 0$ the semigroup is contractive;*

ii. $\forall t \geq 0$, $T(t)D(A) \subset D(A)$;

iii. $f \in D(A) \iff \exists \frac{d}{dt}(T(t)f) = T(t)A(f)(= AT(t)f)$, *for all* $t \geq 0$;

iv. $(A, D(A))$ *is densely defined, i.e. the closure* $\overline{D(A)} = X$;

v. $(A, D(A))$ *is closed;*

vi. The semigroup $(T(t))_{t\geq 0}$ identifies uniquely its generator $(A, D(A))$.

Proof. See [32, Chapter I, Proposition 5.3 and Chapter II, Lemma 1.2, Theorem 1.4]. □

We recall that we can solve our problem (1.1) by the semigroup theory.

Definition 1.1.9. Given $A : D(A) \to X$ a linear operator, consider the following abstract Cauchy problem (ACP)

$$\begin{cases} u_t(t,x) = Au(t,x), \\ u(0,x) = u_0(x), \end{cases} \tag{1.2}$$

with $u_0 \in D(A)$. The function $u = u(t,x)$, $t \geq 0$ is called solution of (1.2) if

i. $u \in D(A)$;

ii. the map $t \in \mathbb{R}_+ \mapsto u(t,\cdot) \in D(A)$ is of class C^1;

iii. u satisfies (1.2).

Theorem 1.1.10. *(Existence and Uniqueness of a Solution of (ACP))*
Consider the (ACP) (1.2), if A is the generator of a (C_0) semigroup $(T(t))_{t\geq 0}$ on X, then the (ACP) admits a unique solution given by

$$u(t,x) = T(t)u_0(x). \tag{1.3}$$

Proof. See [32, Chapter II, Proposition 6.2]. □

Now, we recall the well-known Hille-Yosida Theorem for the contractive case (for more details and generalizations, see [32, Section 3.2])

Theorem 1.1.11. *(Hille-Yosida Theorem)*
Let A be a linear and densely defined operator. Then the following statements are equivalent:

i. $(A, D(A))$ generates a (C_0) contraction semigroup on X;

ii. $(A, D(A))$ is closed and for all $\lambda > 0$, one has $\lambda \in \rho(A)$ and $\|R(\lambda, A)\| \leq 1/\lambda$.

In particular, if i. or ii. is true, one has

$$R(\lambda, A) = \int_0^{+\infty} e^{-\lambda s} T(s) f \, ds,$$

for all $f \in X$, $\lambda \in \rho(A)$.

Proof. See [32, Chapter II, Theorem 3.5]. □

A further charaterization for the generation can be given in the case of dissipative operators defined as follows (for more details, see [32, Section 2.3]).

Definition 1.1.12. The operator A is *dissipative* if for all $\lambda > 0$, $u \in D(A)$ one has $\lambda\|u\| \leq \|(\lambda I - A)u\|$.

Definition 1.1.13. The operator A is $m-$ *dissipative* if it is dissipative and $\rho(A) \cap (0, +\infty) \neq \emptyset$.

Theorem 1.1.14. *(Lumer-Phillips Theorem)*
Let A be a dissipative operator such that $\overline{D(A)} = X$. Then the following statements are equivalent:

i. $(\overline{A}, D(\overline{A}))$ *generates a* (C_0) *contraction semigroup on* X*;*

ii. *there exists* $\lambda > 0$ *such that* $\overline{R(\lambda I - A)} = X$.

Proof. See [32, Chapter II, Theorem 3.15]. □

Theorem 1.1.15. *The operator $(A, D(A))$ generates a (C_0) contraction semigroup on X if and only if A is closed, densely defined and $m-$ dissipative.*

Proof. See [44, Chapter I, Theorem 3.3]. □

Another important characterization for the generation is given by using Feller theory, which links the semigroup theory with the stochastic differential equations (SDE) theory (for more details see, for instance [19], [90]).

Definition 1.1.16. (Feller Semigroup)
Let us consider a locally compact, separable, metric space (K, m) and define $K_\partial = K \cup \partial$ where ∂ is the point at infinity if K is not compact. A (C_0) semigroup $(T(t))_{t \geq 0} \in \mathcal{L}_{\mathbb{R}}(C(K_\partial))$ is a Feller semigroup on $C(K_\partial)$ if it satisfies the following property

$$(f \in C(K_\partial),\, 0 \leq f \leq 1\, on\, K_\partial) \iff (0 \leq T(t)f \leq 1,\, t \geq 0\, on\, K_\partial,\, T(t)1 = 1^1,\, t \geq 0).$$

It is clear that every Feller semigroup is also a positive semigroup, i.e. $T(t)f \geq 0$ for all $f \geq 0$, $t \geq 0$.
If we leave out the (C_0)-hypothesis in the above definition, we obtain a Markov semigroup.

Definition 1.1.17. A Markov transition probability function on K is a measure p_t, $t \geq 0$, on K that is uniformly stochastically continuous, i.e. for all E compact subset of K and $\varepsilon > 0$ there exists $U_\varepsilon(x)$, an ε-neighborhood of x, such that one has $\lim_{t \to 0^+} \sup_{x \in E}(1 - p_t(x, U_\varepsilon(x)) = 0$.

Theorem 1.1.18. *The following statements are equivalent:*

i. $(p_t)_{t \geq 0}$ *is a uniformly stochastically continuous* (C_0)*-transition function on* K*, such that for all* $s > 0$ *and* E *compact subset of* K*,* $\lim_{x \to \partial} \sup_{0 \leq t \leq s} p_t(x, E) = 0$*;*

ii. *for all* $f \in C_0(K)$*,*
$$T(t)f(x) = \int_K p_t(x, dy) f(y),$$
is a Feller semigroup.

[1] 1 represents the constant function equals to 1, in these cases.

In particular, any Feller semigroup (and hence any corresponding family of uniformly stochastically continuous (C_0) transition functions) is uniquely associated to a suitable operator $(A, D(A))$, called the generator of the Feller semigroup, defined as follows

$$D(A) = \left\{ u \in C_0(K) \middle| \exists \lim_{t \to 0^+} \frac{T(t)u - u}{t} \in C_0(K) \right\},$$

$$Au = \lim_{t \to 0^+} \frac{T(t)u - u}{t},$$

for any $u \in D(A)$.

Proof. See [90, Theorem 9.2.3]. □

Remark 1.1.19. The link with the SDE theory can be summarized in this way. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ a filtered probability space, and consider $(X_t)_{t\geq 0}$ a Markov process with state space $(K, \mathcal{B}(K))$ (for more details see e.g. [75], [88]), with dynamics

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \quad X_0 = x, \tag{1.4}$$

where W_t denotes the Brownian motion and μ, σ are the drift and diffusion functions respectively.
Let $\mathbb{P}^x(X_t \in B) = \mathbb{P}(X_t \in B | X_0 = x)$ for any $B \in \mathcal{B}(K)$, and

$$\mathbb{E}^x[f(X_t)] = \int_K f(y)\mathbb{P}^x(X_t \in dy), \tag{1.5}$$

for any $f \in C_0(K)$, $x \in K$. A *Feller process* is a Markov process whose transition semigroup $T(t)f(x) = \mathbb{E}^x[f(X_t)]$ is a Feller semigroup. Clearly the related transition function is

$$p_t(x, B) = \mathbb{P}(X_t \in B | X_0 = x),$$

for all $t \geq 0$, $x \in K$, $B \in \mathcal{B}(K)$.
Finally, the generator $(A, D(A))$ of $(T(t))_{t\geq 0}$ coincides with the infinitesimal generator of the process $(X_t)_{t\geq 0}$, that is defined by

$$Au(x) = \lim_{t \to 0^+} \frac{\mathbb{E}^x[u(X_t)] - u(x)}{t},$$

and $D(A)$ is the set of all u for which tha above limit exists for all $x \in \mathbb{R}$. One has that

$$Au = \frac{1}{2}\sigma^2 u'' + \mu u',$$

for any $u \in D(A)$. One can show that any compactly-supported C^2 function u lives in $D(A)$.
So, we can conclude that if $(X_t)_{t\geq 0}$ is a process which satisfies the SDE (1.4) then a function $u(t, x)$ defined by (1.5) for all $x \in K$, $t \in \mathbb{R}_+$, has to solve the PDE $u_t(x) = Au(x)$.
In particular, if we consider the following (ACP)

$$\begin{cases} u_t(t, x) = Au(t, x) - ru(t, x), & (t, x) \in \mathbb{R}_+ \times K \\ u(0, x) = f(x), & x \in K \end{cases}$$

for all $r \geq 0$, its solution is given by the Feller semigroup

$$u(t,x) = \mathbb{E}^x[e^{-rt}f(X_t)]. \tag{1.6}$$

The above result is well known as the Feynman-Kac Theorem (see, for instance [87, Chapter VI, Section 4]).

Now, we recall a generation theorem based on the so called *Feller classification* (see [32, Chapter VI, Section 4]). An equivalent classification for the boundary endpoints is described in Appendix 6.1.

Definition 1.1.20. (Feller Classification)
If $J = (r_1, r_2)$ is a real interval, with $-\infty \leq r_1 < r_2 \leq +\infty$, let A be a second order differential operator of the type $Au = a(x)u'' + b(x)u'$, where a, b are real continuous functions on J such that $a(x) > 0$ for any $x \in J$. Introduce the following functions

$$W(x) = exp\Big(-\int_{x_0}^{x} \frac{b(s)}{a(s)} ds\Big), \tag{1.7}$$

$$Q(x) = \frac{1}{a(x)W(x)} \int_{x_0}^{x} W(s) ds, \tag{1.8}$$

$$R(x) = W(x) \int_{x_0}^{x} \frac{1}{a(s)W(s)} ds, \tag{1.9}$$

where $x \in J$, x_0 is fixed in J. The boundary point r_2 is said to be

i. *regular* if $Q \in L^1(x_0, r_2)$ and $R \in L^1(x_0, r_2)$;

ii. *exit* if $Q \notin L^1(x_0, r_2)$ and $R \in L^1(x_0, r_2)$;

iii. *entrance* if $Q \in L^1(x_0, r_2)$ and $R \notin L^1(x_0, r_2)$;

iv. *natural* if $Q \notin L^1(x_0, r_2)$ and $R \notin L^1(x_0, r_2)$.

Analogous definitions can be given for r_1 by considering the interval (r_1, x_0) instead of (x_0, r_2).

Theorem 1.1.21. *The operator A with the so called maximal domain*

$$D_M(A) = \{u \in C(\bar{J}) \cap C^2(J) | Au \in C(\bar{J})\},$$

generates a Feller semigroup on $C(\bar{J})$ if and only if r_1 and r_2 are of entrance or natural type.
Moreover, the operator A with the so called Wentzell domain

$$D_W(A) = \Big\{u \in C(\bar{J}) \cap C^2(J) \Big| \lim_{\substack{x \to r_1 \\ x \to r_2}} Au(x) = 0\Big\},$$

generates a Feller semigroup on $C(\bar{J})$ if and only if both the endpoints r_1 and r_2 are not of entrance type.

Proof. See [32, Chapter VI, Theorem 4.15, Theorem 4.18]. □

The next lemma will be useful in the sequel to study the behaviour of the boundary points according to the Feller classification.

Lemma 1.1.22. *Fix $x_0 \in J = (r_1, r_2)$, then we have*

i) $W \notin L^1(x_0, r_2)$ *implies* $R \notin L^1(x_0, r_2)$*;*

ii) $R \in L^1(x_0, r_2)$ *implies* $W \in L^1(x_0, r_2)$*;*

iii) $(mW)^{-1} \notin L^1(x_0, r_2)$ *implies* $Q \notin L^1(x_0, r_2)$*;*

iv) $Q \in L^1(x_0, r_2)$ *implies* $(mW)^{-1} \in L^1(x_0, r_2)$*.*

Analogously for r_1.

Proof. See [32, Chapter VI, Remark 4.10]. □

Remark 1.1.23. By some interchanges between the integration variables, the integrals of Q and R can be written as

$$\begin{aligned}\int_{x_0}^{r_2} Q(x)\,dx &= \int_{x_0}^{r_2}\left(\int_{x_0}^{x} W(s)\,ds\right)(m(x)W(x))^{-1}\,dx \\ &= \int_{x_0}^{r_2}\left(\int_{x}^{r_2} (m(s)W(s))^{-1}\,ds\right)W(x)\,dx \qquad (1.10)\end{aligned}$$

and

$$\begin{aligned}\int_{x_0}^{r_2} R(x)\,dx &= \int_{x_0}^{r_2}\left(\int_{x_0}^{x} (m(s)W(s))^{-1}\,ds\right)W(x)\,dx \\ &= \int_{x_0}^{r_2}\left(\int_{x}^{r_2} W(s)\,ds\right)(m(x)W(x))^{-1}\,dx \qquad (1.11)\end{aligned}$$

for any fixed $x_0 \in (r_1, r_2)$. Analogously for r_1.

Among all possible (C_0) semigroups, the most regular class is the class of analytic semigroups that we will define below.

Definition 1.1.24. (Analytic semigroup) Let $\delta \in (0, \pi/2]$ and define the sector of angle δ, $S_\delta = \{\lambda \in \mathbb{C} || arg(\lambda)| < \delta\}/\{0\}$. Let $(T(z))_{z\in S_\delta \cup \{0\}} \in \mathcal{L}_{\mathbb{C}}(X)$. $(T(z))_{z\in S_\delta \cup \{0\}}$ is said be an analytic semigroup if

i. $T(0) = I$ and $\forall z_1, z_2 \in S_{\delta\cup\{0\}}$, $T(z_1 + z_2) = T(z_1) \cdot T(z_2)$;

ii. the map $z \in S_\delta \mapsto T(z)$ is analytic;

iii. $lim_{z\to 0} T(z)f = f$, $\forall z \in S_{\delta'}$, $\delta' \in (0, \delta)$, $f \in X$.

Theorem 1.1.25. *Let $(A, D(A))$ be a linear operator on a complex Banach space X. The following statements are equivalent:*

i. $(A, D(A))$ *generates an analytic semigroup on* X*;*

ii. there exists $\theta \in (0, \pi/2)$ *such that* $e^{\pm i\theta} A$ *generate* (C_0) *bounded semigroups on* X.

Proof. See [32, Chapter II, Theorem 4.6]. □

The advantage to have the generation of an analytic semigroup is given by the possibility to take the initial value of (ACP) (1.1) in the whole space X (for more details, see [32, Chapter II, Section 4.a]).
In many concrete situations, the linear operator associated to an evolution equation is given as the sum of several terms of different type. We analyze the case of bounded perturbation (see [32, Chapter III, Section 1]).

Theorem 1.1.26. *Let* $(A, D(A))$ *be generator of a* (C_0) *semigroup* $(T(t))_{t\geq 0}$ *on* X, *such that* $\|T(t)\| \leq Me^{\omega t}$ *for all* $t \geq 0$ *and some* $M \geq 1$, $\omega \in \mathbb{R}$. *Consider* $B \in \mathcal{L}_{\mathbb{K}}(X)$. *Then, the operator* $(A+B, D(A))$ *generates a* (C_0) *semigroup* $(S(t))_{t\geq 0}$ *on* X, *such that* $\|S(t)\| \leq Me^{(\omega+M\|B\|)t}$ *for all* $t \geq 0$.

Proof. See [32, Chapter III, Theorem 1.3]. □

Definition 1.1.27. Let $(A, D(A))$ and $(B, D(B))$ be linear operators with $D(A) \subset D(B)$. B is an *A-bounded perturbation* if there exist $a, b \in \mathbb{R}_+$ such that for all $u \in D(A)$ one has $\|Bu\| \leq a\|Au\| + b\|u\|$. In particular the element

$$a_0 = inf\{a \in \mathbb{R}_+ | \exists b \in \mathbb{R}_+ \, such\, that \, \forall u \in D(A)\, , \|Bu\| \leq a\|Au\| + b\|u\|\}$$

is called the A-bound of B.

Theorem 1.1.28. *Let* $(A, D(A))$ *be the generator of a* (C_0) *contraction semigroup and assume* $(B, D(B))$ *to be dissipative and A-bounded with A-bound* $a_0 < 1$. *Then* $(A + B, D(A))$ *generates a* (C_0) *contraction semigroup.*

Proof. See [32, Chapter III, Theorem 2.7]. □

Now, we recall the main useful approximation theorems (see [24], [32, Section 3.5]).

Theorem 1.1.29. *(Chernoff Product Formula)*
Consider a function $V : \mathbb{R}_+ \to \mathcal{L}_{\mathbb{R}}(X)$ *satisfying* $V(0) = I$ *and* $\|[V(t)]^m\| \leq M$, *for all* $t \geq 0$, $m \in \mathbb{N}$ *and some* $M \geq 1$. *Assume that*

$$Af = \lim_{h\to 0} \frac{V(h)f - f}{h}$$

exists for all $f \in D \subset X$, *with* D, $(\lambda I - A)D$ *dense subspaces in* X *for some* $\lambda > 0$. *Then the closure of* A *generates a* (C_0) *contraction semigroup* $(T(t))_{t\geq 0}$ *on* X *given by*

$$T(t)f = \lim_{h\to +\infty} [V(t/n)]^n f,$$

for all $f \in X$ *and uniformly for* t *in compact intervals.*

Proof. See [32, Chapter III, Theorem 5.2]. □

Theorem 1.1.30. *(Trotter Product Formula)*
Let $(T(t))_{t\geq 0}$ and $(S(t))_{t\geq 0}$ be (C_0) semigroup satisfying the stability condition

$$\|[T(t/n)S(t/n)]^n\| \leq Me^{\omega t},$$

for all $t \geq 0$, $n \in \mathbb{N}$, $n \geq 1$ and for constants $M \geq 1$, $\omega \in \mathbb{R}$. Let $(A, D(A))$ and $(B, D(B))$ be the generators of the semigroups $(T(t))_{t\geq 0}$ and $(S(t))_{t\geq 0}$ respectively, and consider the operator $(A + B, D)$, where $D = D(A) \cap D(B)$. Assume that D, $(\lambda I - A - B)D$ are dense in X for some $\lambda > \omega$. Then the closure of $(A + B)$ generates a (C_0) semigroup $(U(t))_{t\geq 0}$ on X given by

$$U(t)f = \lim_{h\to+\infty}[T(t/n)S(t/n)]^n f, \tag{1.12}$$

for all $f \in X$ and uniformly for t in compact intervals.

Proof. See [32, Chapter III, Corollary 5.8]. □

It remains to give some results about the explicit representation of the semigroup, which give an analytic representation of the solution of (ACP) (1.1). The most used result is given by the Romanov's formula (see [44, Chapter II, Section 8]).

Definition 1.1.31. A (C_0) cosine function on a Banach space X is a family $(C(t))_{t\in\mathbb{R}}$ of linear bounded operators on X satisfying the following properties

i. $C(0) = I$;

ii. for all $t, s \in \mathbb{R}$, $C(t + s) + C(t - s) = 2C(t) \cdot C(s)$;

iii. for all $f \in X$, $C(\cdot)f \in C(\mathbb{R}, X)$.

The generator $(A, D(A))$ of a cosine function C is the operator $A = C''(0)$, with domain

$$D(A) = \{f \in X | C(\cdot)f \in C^2(\mathbb{R}, X)\}.$$

Theorem 1.1.32. *If $(A, D(A))$ generates a (C_0) group $(S(t))_{t\in\mathbb{R}}$ on X, then $(A^2, D(A^2))$ generates a (C_0) cosine function given by*

$$C(t) = \frac{S(t) + S(-t)}{2}, \ t \in \mathbb{R}.$$

Moreover, $(A^2, D(A^2))$ generates a (C_0) semigroup $(T(t))_{t\geq 0}$ given by

$$T(t)f = 2\int_0^{+\infty} C(y)f\, p(t, y)dy, \tag{1.13}$$

($T(0) = I$, if $t = 0$) for any $f \in X$, and with

$$p(t, y) = \frac{1}{\sqrt{4\pi t}}e^{-\frac{y^2}{4t}}, \ t > 0, \tag{1.14}$$

the density distribution function of a normally distributed random variable with zero mean and variance $2t$. In addition, if X is a complex space, then $(T(t))_{t\geq 0}$ is an analytic semigroup in the right half plane.

Proof. See [44, Chapter II, Section 8]. □

Remark 1.1.33. We note that A is an A^2-bounded perturbation (with $a_0 = 0$). Thus if $(A, D(A))$ is the generator of the (C_0) semigroup $(S(t))_{t\geq 0}$ on X, for any $\theta \in \mathbb{R}$, the operator $(A^2 + \theta A, D(A^2))$ generates the (C_0) semigroup on X given by the product $(T(t)S(\theta t))_{t\geq 0}$, as a consequence of Theorem 1.1.28.

Remark 1.1.34. Note that the integral in (1.13) can be rewritten as

$$\int_0^{+\infty} (S(y) + S(-y))\, f(x)\, p(t,y)dy$$

$$= \int_0^{+\infty} S(y)\, f(x)\, p(t,y)dy + \int_0^{+\infty} S(-y)\, f(x)\, p(t,y)dy,$$

for all $x \in J$, $t \geq 0$. By changing the variable $w = -y$ in the above second integral, and by (1.14), the semigroup $(T(t))_{t\geq 0}$ has the following explicit representation

$$T(t)f(x) = \int_{-\infty}^{+\infty} S(y)\, f(x)\, p(t,y)dy.$$

1.2 Financial Applications

In this section we will present some applications of the semigroup theory. In particular we will focus on the well-known Black-Scholes-Merton model and Cox-Ingersoll-Ross model, described in the next Subsection 1.2.1 and 1.2.2 respectively.
It is worth noting that in [7], [38], [46] and [21] (see also references therein) the generation of analitic semigroup for a family of degenerate elliptic operators arising in Financial Mathematics have been studied. In particular, their results are given both in $L^p, p \geq 1$ and in weighted Sobolev spaces ([7], [38], [46]) or in the space of unifomly and bounded continuous function on $\mathbb{R}^d$ ([21], by using a probabilistic approach), but with conditions and results different from those herein obtained. Indeed, as observed in the Introduction, the main goal in this thesis (see Chapter 2) is to study the generation of strongly continuous semigroups on suitable spaces of continuous functions on a real interval, and their explicit or approximate representation, which is useful from a numerical point of view.

1.2.1 The semigroup governing the Black-Scholes equation

In the financial literature the Black-Scholes-Merton equation (see [12], [68]) plays a fundamental role, because solved the problem of pricing European options. This equation was studied from different points of view, especially in the stochastic process theory (see, for instance [37], [63], [88]). The semigroup approach gives more results about the solution of the Black-Scholes equation and its properties (see, for instance [4], [43], [31]). In [4] the shape-preserving properties of a semigroup generated by a particular second-order differential operator are investigated, with particular emphasis on the higher order convexity, Lipschitz classes and asymptotic behaviour. The asymptotic behaviour of the Black-Scholes semigroup is studied also in [5], by the analysis of a suitable approximate sequences of positive linear operators in a setting of weighted continuous function spaces. Moreover, since the

initial value of the Cauchy problem associated to the Black-Scholes-Merton equation is an unbounded function, the corresponding semigroup is also studied in spaces of continuous functions wich may grow at $+\infty$ and is proved that it is chaotic for $t \geq 0$, $s > 1$ and $s\sigma > 1$ (see [31]). Several exstensions of the Black-Scholes equation was considered involving e.g. dividend payments, transaction costs, or illiquid and incomplet markets (see, for instance [9], [28], [13], [18]).
Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ a filtered probability space. Consider a dynamic market environment with a riskeless bond B with maturity T, satisfying

$$dB_t = rB_t dt, \quad B_0 = 1,$$

being r the (fixed) risk-free interest rate, and a risky asset whose price S is a geometric Brownian motion, i.e. it is driven by the following SDE

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad S_0 = s_0 > 0, \tag{1.15}$$

where μ and σ are the drift and the volatility of the process and $(W_t)_{t\geq 0}$ is the standard Brownian motion. An European call option is the right to buy one share of stock (S) at a given strike price K at the expiration time $t = T$. It is assumed that the bond can be purchased at the (fixed) risk-free interest rate r, and the option should be priced to avoid arbitage. Hedging consists in duplicating the option by a portfolio of changing holdings of the risk-free bond and of the stock. The price function $v(t,x)$ is the proper price of the option at time $t = 0$, given $S_t = x > 0$. Hence we have to find $v(0,x)$. Black-Scholes and Merton showed that the price v satisfies the following backward parabolic problem

$$\begin{cases} v_t(t,x) + \frac{1}{2}\sigma^2 x^2 v_{xx}(t,x) + rxv_x(t,x) - rv(t,x) = 0, \quad t \in [0,T],\, x > 0, \\ v(T,x) = (x-K)^+, \quad x > 0, \\ v(t,0) = 0, \quad t \in [0,T], \end{cases} \tag{1.16}$$

where the second and third equation of (1.16) represent the terminal condition of the call option and the price boundary condition, respectively. If we let $u(t,x) = v(T-t,x)$, then u satisfies the forward parabolic problem

$$\begin{cases} u_t(t,x) = \frac{1}{2}\sigma^2 x^2 u_{xx}(t,x) + rxu_x(t,x) - ru(t,x) = 0, \quad t \in [0,T],\, x > 0, \\ u(0,x) = (x-K)^+, \quad x > 0, \\ u(t,0) = 0, \quad t \in [0,T]. \end{cases} \tag{1.17}$$

Let X be the Banach space $C[0,+\infty]$. The problem (1.17) can be rewritten as the following abstract Cauchy problem (ACP)

$$\begin{cases} u_t(t,x) = \mathcal{L}u(t,x), \quad t \geq 0,\, x > 0 \\ u(0,x) = u_0(x), \quad x > 0 \\ u(t,0) = 0, \quad t \geq 0; \end{cases} \tag{1.18}$$

with a general initial condition u_0, and for any $t \geq 0$, the operator $\mathcal{L}$ is given by

$$\mathcal{L}u(x) = \frac{1}{2}\sigma^2 x^2 u''(x) + rxu'(x) - ru(x).$$

First of all we give the generation result for the operator $\mathcal{L}$, i.e. the well-posedness theorem for the problem (1.18).

Theorem 1.2.1. *The operator $\mathcal{L}$ with the maximal domain*

$$D_M(\mathcal{L}) = \{u \in X \cap C^2(0, +\infty) | \mathcal{L}u \in X\}, \tag{1.19}$$

generates a (C_0) semigroup on X.

Proof. The proof is based on the Feller classification (see Definition 1.1.20, Theorem 1.1.21). Consider the operator

$$\mathcal{L}_0 u(x) = \frac{\sigma^2}{2} x^2 u''(x) + rxu'(x)$$

Let $r_1 = 0$, $r_2 = +\infty$ and choose $x_0 = 1$. For all $x \in \mathbb{R}_+$, we can compute

$$W(x) = exp\Big(-\int_1^x \frac{2rs}{\sigma^2 s^2} ds\Big) = e^{-k\, lnx} = \frac{1}{x^k},$$

where $k = \frac{2r}{\sigma^2}$. Now, we have that

$$Q(x) = \frac{2}{\sigma^2 x^{2-k}} \int_1^x \frac{1}{s^k}\, ds = \begin{cases} \frac{2}{\sigma^2(1-k)}\Big(\frac{1}{x} - x^{k-2}\Big), \; if\, k \neq 1, \\ \frac{2}{\sigma^2}\frac{lnx}{x}, \quad if\, k = 1, \end{cases}$$

and

$$R(x) = \frac{2}{\sigma^2 x^k} \int_1^x \frac{1}{s^{2-k}}\, ds = \begin{cases} \frac{2}{\sigma^2(k-1)}\Big(\frac{1}{x} - \frac{1}{x^k}\Big), \; if\, k \neq 1, \\ \frac{2}{\sigma^2}\frac{lnx}{x}, \quad if\, k = 1. \end{cases}$$

Now, if $k = 1$, we have

$$\int_1^{+\infty} Q(x)\, dx = \int_1^{+\infty} R(x)\, dx = \frac{2}{\sigma^2}\int_1^{+\infty} \frac{lnx}{x}\, dx = \Big[\frac{ln^2 x}{\sigma^2}\Big]_1^{+\infty} = +\infty,$$

and

$$\int_0^1 Q(x)\, dx = \int_0^1 R(x)\, dx = \frac{2}{\sigma^2}\int_0^1 \frac{lnx}{x}\, dx = \Big[\frac{ln^2 x}{\sigma^2}\Big]_0^1 = -\infty,$$

hence, the endpoints 0 and $+\infty$ are of a natural type if $k = 1$. Otherwise, if $k \neq 1$, we have

$$\int_1^{+\infty} Q(x)\, dx = \frac{2}{\sigma^2(1-k)}\int_1^{+\infty} \frac{1}{x} - x^{k-2}\, dx$$

$$\begin{cases} = \frac{2}{\sigma^2(1-k)}\Big[lnx - \frac{x^{k-1}}{k-1}\Big]_1^{+\infty} = +\infty, \; if\, k < 1 \\ \simeq \frac{2}{\sigma^2(k-1)}\int_1^{+\infty} x^{k-2}\, dx = +\infty, \; if\, k > 1, \end{cases}$$

$$\int_1^{+\infty} R(x)\, dx = \frac{2}{\sigma^2(k-1)}\int_1^{+\infty} \frac{1}{x} - \frac{1}{x^k}\, dx$$

$$\begin{cases} = \frac{2}{\sigma^2(k-1)}\left[lnx - \frac{x^{1-k}}{1-k}\right]_1^{+\infty} = +\infty, \quad if\, k > 1, \\ \simeq \frac{2}{\sigma^2(1-k)} \int_1^{+\infty} \frac{1}{x^k}\, dx = +\infty, \quad if\, k < 1, \end{cases}$$

and analogously $Q(x) \notin L^1(0,1)$, $R(x) \notin L^1(0,1)$, i.e. $0, +\infty$ are endpoints of natural type even if $k \neq 1$. Then, we can conclude that the operator $\mathcal{L}_0$ with the maximal domain generates a Feller semigroup on X, for every $k \geq 1$. Finally, since rI is a bounded perturbation, the whole operator $\mathcal{L} = \mathcal{L}_0 - rI$ with the maximal domain generates a (C_0) semigroup on X (see Theorem 1.1.26). □

The next result gives an alternative proof of the generation by $\mathcal{L}$ and provides an explicit representation of the (ACP) solution (1.18).

Theorem 1.2.2. *The operator $\mathcal{L}$ with the maximal domain $D_M(\mathcal{L})$ defined in* 1.19 *generates a (C_0) contraction semigroup on X given by*

$$T(t)f(x) = e^{-rt} \int_{-\infty}^{+\infty} f\left(xe^{(r-\sigma^2/2)t - y\sigma/\sqrt{2}}\right) p(t,y)dy, \tag{1.20}$$

*with $p(t,y)$ defined in (*1.14*), for all $t \geq 0$, $x \in \mathbb{R}_+$ and $f \in X$.*

Proof. See [43, Section 3].
We start to rewrite our operator $\mathcal{L}$ as

$$\mathcal{L} = G^2 + \gamma G + \delta I, \tag{1.21}$$

where

$$Gu(x) = \nu x u'(x),$$
$$D(G) = \{u \in X \cap C^1(0,+\infty) | xu' \in X\}, \tag{1.22}$$

and

$$G^2u(x) = \nu^2 x^2 u''(x) + \nu^2 x u'(x),$$
$$D(G^2) = \{u \in D(G) | Gu \in D(G)\} = \{u \in X \cap C^2(0,+\infty) | \mathcal{L}u \in X\}, \tag{1.23}$$

with $\nu = \sigma/\sqrt{2}$, $\delta = -r$ and $\gamma = -(\delta + \nu^2)/\nu$.
The operator $(G, D(G))$ generates a (C_0) contraction semigroup on X given by

$$e^{tG}f(x) = f(xe^{\nu t})$$

for all $x \in \mathbb{R}_+$, $t \geq 0$ and $f \in X$ (see [32, Chapter VI]). Then, $(G^2, D(G^2))$ generates a (C_0) semigroup on X given by

$$e^{tG^2}f(x) = \int_{-\infty}^{+\infty} e^{-yG} f(x)\, p(t,y)dy,$$

and the perturbed operator $(G^2 + \gamma G, D(G^2))$ generates the following (C_0) semigroup on X

$$\int_{-\infty}^{+\infty} e^{-yG} e^{\gamma t G} f(x)\, p(t,y)dy,$$

for all $x \in \mathbb{R}_+$, $t \geq 0$, $f \in X$ (see Remark (1.1.33)). Moreover, the multiplication operator $M = \delta I$ generates a (C_0) semigroup on X given by the funtion $e^{tM} = e^{\delta t}$.

Hence, the whole operator $(\mathcal{L}, D(G^2))$ generates a (C_0) (and analytic) semigroup on X given by

$$T(t)f(x) = e^{\delta t}\int_{-\infty}^{+\infty} e^{-yG}e^{\gamma tG}f(x)\,p(t,y)dy = e^{-rt}\int_{-\infty}^{+\infty} f\big(xe^{(r-\sigma^2/2)t-y\sigma/\sqrt{2}}\big)\,p(t,y)dy,$$

for all $x \in \mathbb{R}_+$, $t \geq 0$, $f \in X$. We note that the domain $D(G^2)$ coincides with the maximal domain $D_M(\mathcal{L})$. In particular, it is easy to see that $(T(t))_{t\geq 0}$ is also contractive on X. $\square$

Remark 1.2.3. By Theorem 1.1.10, the solution of (ACP) (1.18) has the form $u(t,x) = T(t)u_0(x)$, with $(T(t))_{t\geq 0}$ defined in (1.20). Since the semigroup $(T(t))_{t\geq 0}$ is analytic in the right half plane, we can take u_0 in the whole space X (and not only in $D(\mathcal{L})$). However, u_0 could not be in the space X considered. This is the case of the call option whose payoff is unbounded at infinity. For this reason, we can let $u_M = min\{u_0, M\}$, for $M \geq 1$, so that u_M belongs to X.

Remark 1.2.4. We can compare the solution (1.20) obtained via semigroup theory with the classic Black-Scholes formula for European call option, derived in a probabilistic context (see e.g. [88, Section 5.2]). Indeed, the solution of (1.16) is directly derived by the Feynman-Kac Theorem

$$\mathbb{E}_{\mathbb{Q}}[e^{-r(T-t)}(S_T - K)^+|\mathcal{F}_t], \tag{1.24}$$

where $\mathbb{Q}$ is the risk-neutral measure under which the solution of the SDE (1.15) is given by

$$S_t = S_0\,exp\Big(\sigma\widetilde{W}_t + (r - \frac{\sigma^2}{2})t\Big),$$

and $\widetilde{W}_t$ is the Brownian motion under the measure $\mathbb{Q}$. Thus, we may write

$$\begin{aligned} S_T &= S_t\,exp\Big(\sigma(\widetilde{W}_T - \widetilde{W}_t) + (r - \frac{\sigma^2}{2})(T-t)\Big) \\ &= S_t\,exp\Big(-\sigma\sqrt{\tau}Y + (r - \frac{\sigma^2}{2})\tau\Big), \end{aligned}$$

where Y is the standard normal random variable

$$Y = -\frac{\widetilde{W}_T - \widetilde{W}_t}{\sqrt{\tau}},$$

and τ is the time to expiration $(T-t)$.
Consider an European call option with strike price K, maturity T and payoff

$$G(x) = (x-K)^+ = \begin{cases} x - K, & if\, x \geq K \\ 0, & if\, x < K. \end{cases} \tag{1.25}$$

Therefore, (1.24) becomes

$$\frac{1}{\sqrt{2\pi}}e^{-r\tau}\int_{-\infty}^{+\infty}\Big(xexp\Big(-\sigma\sqrt{\tau}y + (r - \frac{\sigma^2}{2})\tau\Big) - K\Big)^+ e^{-\frac{y^2}{2}}dy, \tag{1.26}$$

where x is a realization of S_t.

1.2.2 The semigroup governing the Cox-Ingersoll-Ross equation

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ a filtered probability space. Cox, Ingersoll and Ross [25] derived the pricing formula of a zero-coupon bond under the assumption that the short-term interest rate is a diffusion process, called the $square - root$ process, satisfying the following stochastic differential equation (SDE)

$$dr_t = \widetilde{k}(\widetilde{\theta} - r_t)dt + \sigma\sqrt{r_t}d\widetilde{W}_t, \ \ r_0 > 0,$$

where $\widetilde{k}(\widetilde{\theta} - r_t)$ defines a mean reverting drift pulling the interest rate towards its long-term value $\widetilde{\theta}$ with a speed of adjustment equal to $\widetilde{k}$, and $\widetilde{W}_t$ is the Brownian motion. In the risk-adjusted economy, the dynamics is supposed to be given by

$$dr_t = (\widetilde{k}(\widetilde{\theta} - r_t) - \lambda r_t)dt + \sigma\sqrt{r_t}dW_t, \ \ r_0 > 0$$

where

$$W_t = \widetilde{W}_t + \frac{\lambda}{\sigma}\int_0^t \sqrt{r_s}\, ds,$$

is a Brownian motion under the martingale measure $\mathbb{Q}$ with λ being the market price of risk. Setting

$$k = \widetilde{k} + \lambda, \ \ \theta = \widetilde{k}(\widetilde{\theta}/k),$$

the SDE describing the $\mathbb{Q}$-dynamics of the square-root process r is

$$dr_t = k(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t, \ \ r_0 > 0. \tag{1.27}$$

Cox et al. in [25] derived the bond price P, with face value 1, that satisfies the following backward parabolic problem

$$\begin{cases} P_t(t,x) + \frac{\sigma^2}{2}xP_{xx}(t,x) + k\theta P_x(t,x) - kxP_x(t,x) - xP(t,x) = 0, \ \ t \in [0,T], \, x \in \mathbb{R}_+, \\ P(T,x) = 1, \ \ x \in \mathbb{R}_+, \end{cases}$$

where T is the maturity date and x is a realization of r_t. If we let $u(t,x) = P(T-t,x)$, then $u(t,x)$ satisfies the forward parabolic problem

$$\begin{cases} u_t(t,x) = \frac{\sigma^2}{2}xu_{xx}(t,x) + k\theta u_x(t,x) - kxu_x(t,x) - xu(t,x) = 0, \ \ t \in [0,T], \, x \in \mathbb{R}_+, \\ u(0,x) = 1, \ \ x \in \mathbb{R}_+. \end{cases} \tag{1.28}$$

Let X be the Banach space $C[0,+\infty]$. The problem (1.28) can be rewritten as the following abstract Cauchy problem (ACP)

$$\begin{cases} u_t(t,x) = \mathcal{L}u(t,x), \ \ t \geq 0, \, x \in \mathbb{R}_+ \\ u(0,x) = u_0(x), \ \ x \in \mathbb{R}_+, \end{cases} \tag{1.29}$$

where u_0 is the initial data, and for any $t > 0$, the operator $\mathcal{L}$ is given by

$$\mathcal{L}u(x) = \frac{1}{2}\sigma^2 x u''(x) + k(\theta - x)u'(x) - xu(x).$$

Proposition 1.2.5. *The homogeneous operator $\mathcal{L}_0$ defined as*

$$\mathcal{L}_0 u(x) = \frac{1}{2}\sigma^2 x u''(x) + k(\theta - x)u'(x),$$

with the maximal domain

$$D(\mathcal{L}_0) = \{u \in X \cap C^2(0, +\infty) | \mathcal{L}_0 u \in X\},$$

generates a Feller semigroup on X.

Proof. See [54, Chapter VI, Section 3] and Example 6.1.8 of Appendix 6.1.
We start to observe that the square-root process r_t satisfies the following relation

$$r_t = e^{-kt}\rho\Big(\frac{\sigma^2}{4k}(e^{kt} - 1)\Big),$$

so, it can be viewed as a squared Bessel process $\rho(t)$, which is the unique solution of the following SDE

$$d\rho(t) = \frac{4k\theta}{\sigma^2}dt + 2\sqrt{\rho(t)}dW_t.$$

Note that the state space of the process $\rho(t)$ is the interval $[0, +\infty)$. Then, the behavior of the square-root process r_t at the boundary points 0 and $+\infty$ can be studied by knowing the behavior of the process ρ_t at these endpoint. It can be showed that the endpoint 0 is of entrance type if $k\theta \geq \frac{\sigma^2}{4}$, and the endpoint $+\infty$ is of natural type. Hence, the assertion follows from Theorem 1.1.21. □

Remark 1.2.6. The generation result for the CIR operator is a more difficult task than the Black-Schols operator. The main issue is given by the potential term $-xu$. This term represents an unbounded pertutbation, so, we can not apply Theorem 1.1.26 as it is for the Black-Scholes operator.
Moreover, this perturbation forces us to let $X = C_0[0, +\infty)$ instead of $C[0, +\infty]$. Since X is not dense in $C[0, +\infty]$, $\mathcal{L}$ is not densely defined in $C[0, +\infty]$, but its restriction to X will be densely defined on X.
Moreover, by probabilistic results, it is known that the unique solution of (1.29) is given by the Feynman-Kac Theorem (see e.g. [54, Section 6.3])

$$u(t, x) = \mathbb{E}^x_{\mathbb{Q}}\Big[exp\Big(-\int_0^t r_s\, ds\Big)u_0(r_t)\Big], \tag{1.30}$$

for $t \geq 0$, with initial value $r_0 = x > 0$. The expectation is taken with respect to the martingale measure $\mathbb{Q}$.
By the semigroup theory one can prove that $\mathcal{L}$ is m-dissipative and densely defined on X (see [40]) and so the generation follows from Theorem 1.1.15.

The next Theorem 1.2.8 states the generation result for the operator $\mathcal{L}$ and provides an approximate (semi-explicit) representation for the solution of problem (1.29), by applying a generalized Lie-Trotter-Daletskii type formula given in the following proposition.

Proposition 1.2.7. *Let A_i be a densely defined and m-dissipative operator on a Banach space X, such that A_i generates a (C_0) contraction semigroup T_i on X, for $i = 1, ..., n$. Suppose that the closure $\overline{A} = \overline{A_1 + ... + A_n}$ is densely defined and m-dissipative, and let T be the semigroup that it generates. Let $J_1, .., J_q$ be a piecewise disjoint decomposition of $\{1, ..., n\}$ into q nonempty subsets, $2 \leq q \leq n$. Let*

$$\mathcal{A}_h = \overline{\sum_{j \in J_h} A_j}, \quad h = 1, ..., q.$$

Then, $\mathcal{A}_h$ generate a (C_0) contraction semigroup S_h on X, with

$$S_h(t)f = \lim_{m \to +\infty} \left(T_{j_1}\left(\frac{t}{m}\right) \cdots T_{j_l}\left(\frac{t}{m}\right)\right)^m f, \tag{1.31}$$

for all $t \geq 0$ $f \in X$ and with $J_h = \{j_1, ..., j_l\}$. Furthermore,

$$T(t)f = \lim_{n \to +\infty} \left(S_1\left(\frac{t}{n}\right) \cdots S_q\left(\frac{t}{n}\right)\right)^n f, \tag{1.32}$$

for all $t \geq 0$ $f \in X$. The convergence is uniform for t in compacta.

Proof. See [40, Section 2, Proposition 1].
If $i = 2$, the proposition reduces to Theorem 1.1.30. If $i > 2$, the representation (1.31) follows from Theorem 1.1.29 writing $V(t) = T_{j_1}(t) \cdots T_{j_l}(t)$. Then (1.32) is a consequence of setting $V(t) = S_1(t) \cdots S_q(t)$. Thus, this proposition formally extends Theorem 1.1.29 but has essentially the same proof, and so it is simply a variant of the standard product formula. □

Theorem 1.2.8. *Assume that $k\theta \geq \frac{\sigma^2}{4}$ and $k \neq 0$. Then, the operator $\mathcal{L}$ with domain*

$$D(\mathcal{L}) = \{u \in X \cap C^2(0, +\infty) | \mathcal{L}u \in X\}, \tag{1.33}$$

generates a (C_0) contraction semigroup on X given by

$$T(t)u_0(x) = \lim_{n \to +\infty} u_n(t, x), \tag{1.34}$$

uniformly in $x \in [0, \infty)$ and for t in bounded intervals of $[0, +\infty)$, where $n = 2^k$, $k \in \mathbb{N}$. In particular, the approximate solution u_n can be written as

$$u_n(t, x) = \underbrace{\int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty}}_{n\, times} L(t, n, \sigma, \{y_i\}_{1 \leq i \leq n}, x, u_0) \prod_{i=1}^{n} p\left(\frac{t}{n}, y_i\right) dy_1 ... dy_n. \tag{1.35}$$

The expression of the above function L is given in Appendix 6.2.

Proof. This proof is based on the recent work [40] and summarizes its main results. We start to rewrite the operator $\mathcal{L}$ as

$$\mathcal{L} = G^2 + Q + P,$$

where

$$Gu(x) = \nu\sqrt{x}u'(x),\ \ D(G) = \{u \in X \cap C^1(0,+\infty) | \sqrt{x}u' \in X\},$$

$$Qu(x) = \alpha u'(x) + \beta x u'(x),\ \ D(Q) = \{u \in X \cap C^1(0,+\infty) | u', xu' \in X\},$$

$$Pu(x) = -xu(x),\ \ D(P) = \{u \in X | xu \in X\},$$

with $\nu = \frac{\sigma}{\sqrt{2}} > 0$, $\alpha = (k\theta - \frac{\nu^2}{2}) \geq 0$, $\beta = -k \neq 0$, for any $x \in \mathbb{R}_+$. In particular

$$G^2u(x) = \nu^2 x u''(x) + \frac{\nu^2}{2}u'(x),\ \ D(G^2) = \{u \in D(G) | Gu \in D(G)\}$$

$$= \{u \in X \cap C^2(0,+\infty) | \sqrt{x}u', \sqrt{x}(\sqrt{x}u')' \in X\}.$$

The operators $(G^2, D(G^2))$, $(Q, D(Q))$ and $(P, D(P))$ generate the (C_0) contraction semigroups $(e^{tG^2})_{t\geq 0}$, $(e^{tQ})_{t\geq 0}$ and $(e^{tP})_{t\geq 0}$ on X (see [40, Section 3, Lemma 1] and [40, Section 5, Lemma 3]) respectively, given by

$$e^{tG^2}f(x) = \int_{-\infty}^{+\infty} f\left(\left(\sqrt{x} + \frac{\nu y}{2}\right)^2\right) p(t,y)dy, \tag{1.36}$$

with p defined in (1.14),

$$e^{tQ}f(x) = f\left(e^{t\beta}x + \frac{\alpha}{\beta}(e^{t\beta} - 1)\right), \tag{1.37}$$

and

$$e^{tP}f(x) = e^{-tx}f(x),$$

for any $x \in \mathbb{R}_+$, $t \geq 0$, $f \in X$.
Now, the closure of $Q + P$ with domain $D(Q + P) = \{u \in D(Q) | Pu \in X\}$ generates a (C_0) contraction semigroup on X given by

$$e^{t(Q+P)}f(x) = exp\left(-\frac{1}{\beta}\left((e^{t\beta} - 1)\left(\frac{\alpha}{\beta} + x\right) - \alpha t\right)\right) f\left(e^{t\beta}x + \frac{\alpha}{\beta}(e^{t\beta} - 1)\right) \tag{1.38}$$

(see [40, Section 5, Theorem 4]). Finally, the operator $\mathcal{L}$ with domain $D(\mathcal{L}) = D(G^2) \cap D(Q + P) = \{u \in D(G^2) | Qu, Pu \in X\}$ (coincident with the domain (1.33)) is densely defined and m-dissipative on X (see [40, Theorem 3, Section 4]). Hence, by Proposition 1.2.7, the closure $\overline{\mathcal{L}}$ generates a (C_0) contraction semigroup on X given by the following product formula

$$T(t)u_0(x) = \lim_{n\to+\infty} u_n(t,x) = \lim_{n\to+\infty} \left(e^{\frac{t}{n}G^2} e^{\frac{t}{n}(Q+P)}\right)^n u_0(x),$$

that leads to formula (1.35) (see Appendix 6.2). □

Remark 1.2.9. Formula (1.34) is called a "generalized Feynman-Kac type formula". In fact, it is the analytic form of the classic Feynman-Kac formula (1.30). The (C_0) property of the semigroup $(T(t))_{t\geq 0}$ on $C_0[0, +\infty)$ had been established in [30], but there it is expressed in a very different context using affine processes. Note that the initial data $u_0 = 1$ by referring to the CIR problem (1.28), is not included in the space $C_0[0, +\infty)$. However, it is possible to define the following sequence

$$u_m(x) = \begin{cases} 1, & if\, x \in [0, m], \\ m + 1 - x, & if\, x \in [m, m+1], \\ 0, & if\, x \in [m+1, +\infty], \end{cases} \tag{1.39}$$

$u_m \in X$ and clearly $u_m(x) \to u_0(x)$, for all $x \geq 0$, as $m \to +\infty$. Thus the approximate solutions $(T(t)u_m)(x)$ given by (1.34) and (1.35) with initial condition u_m converges to

$$u(t, x) = \lim_{n \to +\infty} \underbrace{\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty}}_{n\, times} L(t, n, \nu, \{y_i\}_{1 \leq i \leq n}, x, 1) \prod_{i=1}^{n} p\left(\frac{t}{n}, y_i\right) dy_1 \ldots dy_n,$$

as $m \to +\infty$, for all $x \in R_+$.

Chapter 2

A Semigroup Approach to Generalized Black-Scholes Type Equations in Incomplete Markets

In this chapter an option pricing problem in incomplete market is studied by an analytic point of view. In particular, the results reported in the paper [18] are herein illustrated. The incompleteness is generated by the presence of a non-traded asset. The aim is to use the semigroup theory in order to prove existence and uniqueness of solutions to generalized Black-Scholes type equations that are non-linearly associated with the price of European claims written exclusively on non-traded assets. Then, analytic expressions of the solutions are obtained. An approximate representation in terms of a generalized Feynman-Kac type formula is derived in cases where an explicit closed form solution is not available. Numerical applications and examples are also given showing an excellent agreement between our theoretical results and the numerical tests.
The chapter is organized as follows. After a short discussion in Section 2.1, in Section 2.2 we will focus on initial value Cauchy problems associated with operators of the type

$$\mathcal{L}u = \frac{1}{2}(cx+d)^2u'' + [(\gamma x + \delta) - \theta(cx+d)]u',$$

acting on the space of all real-valued continuous functions in a suitable unbounded real interval having finite limits at the endpoints. Here the domain of $\mathcal{L}$ depends on the coefficients of u', u''. More generation results are provided, giving an analytic expression of the solution to the associated Cauchy problem. In particular, approximate solutions expressed in terms of a generalized Feynman-Kac type formula are given when explicit closed forms are not available. This is a valuable result since very few exact solution formulas to Cauchy problems associated with financial models are available. Some numerical applications and examples are also presented in Section 2.3 to compare the approximate solutions with benchmark formulas in the literature. These applications confirm the accuracy and the fast convergence of the proposed approximation formulas, thus showing that an approximate representation of the indifference price is particularly helpful in cases where exact pricing formulas fail because their coefficients cannot be explicitely computed.

2.1 Analytic Problem

We consider the market environment with two risky assets assumed by M. Musiela and T. Zariphopoulou in [71]; the price S of the traded asset is a geometric Brownian motion, i.e. it is the unique solution of the following SDE

$$dS_\tau = \mu S_\tau d\tau + \sigma S_\tau d\widetilde{W}_\tau, \quad 0 \le t < \tau, \tag{2.1}$$

with initial value $S_t = s > 0$, $\mu \in \mathbb{R}$, $\sigma > 0$. The dynamics of the level Y of the non-traded risky asset is described by a general diffusion process satisfying the SDE

$$dY_\tau = b(Y_\tau, \tau)\, d\tau + a(Y_\tau, \tau)\, dW_\tau, \quad 0 \le t < \tau, \tag{2.2}$$

with initial value $Y_t = y \in \mathbb{R}$, where the coefficients $b(\cdot,\cdot)$ and $a(\cdot,\cdot)$ satisfy enough regularity for (2.2) to have a unique (strong) solution. The processes $\{\widetilde{W}_\tau, \tau \ge 0\}$ and $\{W_\tau, \tau \ge 0\}$ are standard one-dimensional Brownian motions defined on a given filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_\tau)_{\tau \ge 0}, \mathbb{P})$, where $\mathcal{F}_\tau$ is the $\sigma-$algebra generated by $\{\widetilde{W}_u, W_u,\ 0 \le u \le \tau\}$. The Brownian motions are correlated with instantaneous correlation coefficient $\rho \in (-1, 1)$. It is assumed that the derivative to be priced is a European claim written exclusively on the non-traded asset, whose payoff at the maturity $T > t$ is of the form $\varepsilon = G(Y_T)$, being G a bounded function. Moreover, trading occurs in the time horizon $[t, T]$ and only between the risky asset with price S and a riskless bond $B = 1$ with maturity T, yielding constant interest rate $0 \le r < \mu$. Without any loss of generality, we assume that $r = 0$.
In the framework described above, the individual risk preferences are modelled via an exponential utility function

$$U(x) = -e^{-\eta x}, \quad x \in \mathbb{R}, \tag{2.3}$$

where $\eta > 0$ is the *risk aversion* parameter. Then according to the approach to pricing based on the comparison of maximal expected utility payoffs, Musiela and Zariphopoulou [71] derived the writer's indifference price of a European derivative with payoff $\varepsilon = G(Y_T)$ in the following closed form (see [71, Theorem 3])

$$h(y,t) = \frac{1}{\eta(1-\rho^2)} \ln \mathbb{E}_{\mathbb{Q}}[e^{\eta(1-\rho^2)G(Y_T)} \mid Y_t = y], \tag{2.4}$$

for $(y,t) \in \mathbb{R} \times [0,T]$, where the pricing indifference measure $\mathbb{Q}$ is defined as follows

$$\mathbb{Q}(A) = \mathbb{E}_{\mathbb{P}}\Big[exp\Big(-\rho\frac{\mu}{\sigma}W_T - \frac{1}{2}\rho^2\frac{\mu^2}{\sigma^2}T\Big)I_A\Big], \quad A \in \mathcal{F}_T^W, \tag{2.5}$$

where $\mathcal{F}_T^W$ is the augmented σ-algebra generated by W_τ, $0 \le \tau \le T$.

Remark 2.1.1. In the the reduced complete model in which one trades the asset S and the bond B, and where the non-traded asset Y is not taken into account, the pricing measure in (2.5) is replaced by the following unique martingale measure

$$\widetilde{\mathbb{Q}}(A) = \mathbb{E}_{\mathbb{P}}\Big[\frac{d\widetilde{\mathbb{Q}}}{d\mathbb{P}}I_A\Big] = \mathbb{E}_{\mathbb{P}}\Big[exp\Big(-\frac{\mu}{\sigma}\widetilde{W}_T - \frac{1}{2}\frac{\mu^2}{\sigma^2}T\Big)I_A\Big], \quad A \in \mathcal{F}_T^{\widetilde{W}},$$

where $F_T^{\widetilde{W}}$ is the augmented σ-algebra generated by $\widetilde{W}_\tau, 0 \le \tau \le T$.
Thus the indifference measure $\mathbb{Q}$ defined in (2.5) satisfies (see [72, Theorem 2.2])

$$\mathbb{Q}(A) = \mathbb{E}_{\mathbb{P}}\left[\mathbb{E}_{\mathbb{P}}\left[\frac{d\widetilde{\mathbb{Q}}}{d\mathbb{P}}\middle|\mathcal{F}^{W_T}\right]I_A\right],$$

where $\mathbb{E}_{\mathbb{P}}[d\widetilde{\mathbb{Q}}/d\mathbb{P}|\mathcal{F}^{W_T}]$ is the projection of the Radon-Nikodym density $d\widetilde{\mathbb{Q}}/d\mathbb{P}$ onto the Brownian motion $W_\tau, \tau \ge 0$. Thus, the indifference probability measure $\mathbb{Q}$ is the closest measure to the risk-neutral one $\widetilde{\mathbb{Q}}$. Note that in the perfect correlated case, i.e. when $\rho^2 = 1$, we have that $\mathbb{Q} = \widetilde{\mathbb{Q}}$ (see [72, Theorem 2.3]).

In particular, under the measure $\mathbb{Q}$, Y_τ satisfies the following SDE

$$dY_\tau = \big(b(Y_\tau, \tau) - \rho\frac{\mu}{\sigma}a(Y_\tau, \tau)\big)d\tau + a(Y_\tau, \tau)d\overline{W}_\tau, \tag{2.6}$$

where $\overline{W}_\tau = W_\tau + \rho\frac{\mu}{\sigma}\tau$ is a Brownian motion. Hence, $(Y_\tau)_{\tau\ge 0}$ is a diffusion process with the infinitesimal generator

$$\frac{\partial}{\partial\tau} + \frac{1}{2}a^2(y)\frac{\partial^2}{\partial y^2} + \big(b(y) - \rho\frac{\mu}{\sigma}\,a(y)\big)\frac{\partial}{\partial y}$$

(see [71, Theorem 2] or [72, Theorem 2.1]).

Remark 2.1.2. In complete arbitrage-free markets every financial instrument can be replicated. Thus by using self-financing and replicating portfolio strategies, the arbitrage-free representation of a contingent claim at a time $0 \le t < T$ is uniquely determined as the conditional expectation of the discounted payoff function under the unique risk-neutral martingale measure (see, e.g., [87, Section VI.1]).
As mentioned in the Introduction, in incomplete arbitrage-free markets financial instruments are not, in general, perfectly replicable. Further, asset pricing will depend on the utility function of investors. Thus a specific martingale measure, which is defined as the closest to the risk-neutral one, must be determined by a certain optimality criteria to price a contingent claim. Frittelli in [37] showed that if the minimal entropy martingale measure exists, it is unique and is equivalent to $\mathbb{P}$.

Our aim herein is to give an explicit representation for the indifference pricing function $h(y,t)$ that, by (2.4), can be written as

$$h(y,t) = \frac{1}{\eta(1-\rho^2)}\ln w(y,t),$$

where

$$w(y,t) = \mathbb{E}_{\mathbb{Q}}[e^{\eta(1-\rho^2)G(Y_T)} \mid Y_t = y]. \tag{2.7}$$

Under the usual regularity for the Feynman-Kac approach (see, for example, [75, Section 8.2]), $w(y,t)$ solves the Cauchy problem

$$\begin{cases} \frac{\partial w}{\partial t} + \frac{1}{2}a^2(y,t)\frac{\partial^2 w}{\partial y^2} + \big(b(y,t) - \rho\frac{\mu}{\sigma}a(y,t)\big)\frac{\partial w}{\partial y} = 0, & (y,t) \in \mathbb{R}\times[0,T], \\ w(y,T) = e^{\eta(1-\rho^2)G(y)}, & y \in \mathbb{R}, \end{cases} \tag{2.8}$$

where $y = Y_t$ is a dummy variable for any $t \in [0,T]$.
To give an explicit representation of $h(y,t)$, we will prove existence and uniqueness

of the solution to problem (2.8) by a semigroup approach. Observe that, in a perfect correlation between the traded and the non-traded asset, i.e. when $\rho^2 = 1$, and, in addition, the coefficients in (2.2) are $b(y,t) = \mu y$ and $a(y,t) = \sigma y$, the market becomes complete and the indifference price h reduces to the usual Black-Scholes model (see [72, Theorem 2.3]).
With the variable change $u(y,t) = w(y, T-t)$ the problem (2.8) can be transformed from a backward to a forward parabolic problem

$$\begin{cases} \frac{\partial u}{\partial t} = \frac{1}{2} a^2(y,t) \frac{\partial^2 u}{\partial y^2} + \big(b(y,t) - \rho \frac{\mu}{\sigma} a(y,t)\big) \frac{\partial u}{\partial y}, & (y,t) \in \mathbb{R} \times [0,T], \\ u(y,0) = e^{\eta(1-\rho^2)G(y)}, \quad y \in \mathbb{R}. \end{cases} \tag{2.9}$$

Once we consider (2.9), we may do so for $0 \leq t < +\infty$.
The coefficients $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ are assumed of the type

$$a(Y_\tau, \tau) = c_\tau Y_\tau + d_\tau, \quad b(Y_\tau, \tau) = \gamma_\tau Y_\tau + \delta_\tau, \tag{2.10}$$

where $c_\tau, d_\tau, \gamma_\tau, \delta_\tau$ are suitable functions depending on τ. This assumption is not restrictive from a financial point of view, because it is often possible to reduce $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ to (2.10) (for more details the reader can refer to O. S. Farad [35]).
In the sequel, we shall focus the discussion on the autonomous case, i.e. we assume

$$a(Y_\tau, \tau) \equiv a(Y_\tau) = c\, Y_\tau + d, \quad b(Y_\tau, \tau) \equiv b(Y_\tau) = \gamma\, Y_\tau + \delta,$$

for any $\tau \geq 0$, with c, d, γ, $\delta \in \mathbb{R}$. Under this assumption, obviously, the SDE (2.6) has a unique strong solution (see, for example, [75, Section 5.2]).

Remark 2.1.3. We notice that assumption (2.10) is suggested by M. Musiela and T. Zariphopoulou in [72]. Indeed, they indicate as possible candidate for the dynamics of the non-traded asset a class of diffusion processes for which $a(y) = d$, $b(y) = \gamma y + \delta$, with $d, \gamma, \delta \in \mathbb{R}$. We shall examine this particular case in Subsection 2.2.2.

2.2 Semigroup Approach

Let $J = (r_1, r_2)$ be a real interval with $-\infty \leq r_1 < r_2 \leq +\infty$ and $C(\bar{J})$ be the space of all real-valued continuous functions in J having finite limits at the endpoints r_1, r_2. We consider the following abstract Cauchy problem

$$(ACP) \qquad \begin{cases} u_t = \mathcal{L}u, & \text{in } J \times [0, +\infty), \\ u(x,0) = g(x), & x \in J, \end{cases} \tag{2.11}$$

where $\mathcal{L}$ is a differential operator of the type $\mathcal{L}u = m(x)u'' + q(x)u'$ acting on $C(\bar{J})$, with the maximal domain

$$D_M(\mathcal{L}) = \{u \in C(\bar{J}) \cap C^2(J) |\, \mathcal{L}u \in C(\bar{J})\}. \tag{2.12}$$

For our purposes

$$\mathcal{L}u = \frac{1}{2} a^2(x) u'' + (b(x) - \theta a(x))u',$$

with $a(x) = cx + d$, $b(x) = \gamma x + \delta$, and $c, d, \gamma, \delta, \theta \in \mathbb{R}$ constant parameters. Here $\theta = \rho\mu/\sigma$, where the parameters ρ, μ and σ are defined in Section 2.1. Thus we can write

$$\mathcal{L}u = \frac{1}{2}(cx+d)^2 u'' + ((\gamma x + \delta) - \theta(cx+d))u'. \tag{2.13}$$

The initial condition of (2.11) is

$$g(x) = e^{\eta(1-\rho^2)G(x)}, \quad x \in J, \tag{2.14}$$

where $\eta > 0$ is the risk aversion parameter and $G \in C(\bar{J})$. We recall a well-known result that will be useful later (see [32, Chapter VI, Section 4]).

Lemma 2.2.1. *Let us consider $J = (r_1, r_2)$ and $\Gamma = (s_1, s_2)$, with $-\infty \leq r_1 < r_2 \leq +\infty$, $-\infty \leq s_1 < s_2 \leq +\infty$. Let $\Phi : J \to \Gamma$ be a bijective function with inverse Φ^{-1}, such that $\Phi \in C^2(J)$, $\Phi'(x) > 0$ and for any $v \in C(\bar{\Gamma})$, define $T_\Phi(v) = v \circ \Phi$. Then $T_\Phi(v) \in C(\bar{J})$. Moreover T_Φ is an invertible bounded linear operator from $C(\bar{\Gamma})$ to $C(\bar{J})$ such that $\|T_\Phi\| \leq 1$ and $(T_\Phi)^{-1} = T_{\Phi^{-1}}$. Further, s_i has the same type of Feller classification of r_i, $i = 1, 2$.*

Proof. Observe that, if $v \in C(\bar{\Gamma})$, then $v \circ \Phi \in C(J)$, moreover

$$\lim_{x \to r_1} (v \circ \Phi)(x) = \lim_{z \to s_1} v(z) \in \mathbb{R}.$$

Analogously,

$$\lim_{x \to r_2} (v \circ \Phi)(x) = \lim_{z \to s_2} v(z) \in \mathbb{R}.$$

Hence, $T_\Phi(v) \in C(\bar{J})$. It is clear that T_Φ is a linear and

$$\|T_\Phi(v)\|_\infty = \|v \circ \Phi\|_\infty \leq \|v\|_\infty.$$

This yields that T_Φ is bounded with $\|T_\Phi\| \leq 1$. Moreover, T_Φ is bijective with inverse $T_{\Phi^{-1}}$. $\square$

Remark 2.2.2. Under the assumptions of Lemma 2.2.1, according to [32, Chapter II, Section 2.a], if $(T(t))_{t \geq 0}$ is a (C_0) semigroup on $C(\bar{J})$ having $(A, D(A))$ as generator, then $(S(t))_{t \geq 0}$ given by $S(t) = T_{\Phi^{-1}} \circ T(t) \circ T_\Phi$, $t \geq 0$, is a (C_0) semigroup on $C(\bar{\Gamma})$ with generator $(B, D(B))$. Here $B = T_{\Phi^{-1}} \circ A \circ T_\Phi$, and $D(B) = \{g \in C(\bar{\Gamma}) | T_\Phi(g) \in D(A)\}$. Moreover, if $(T(t))_{t \geq 0}$ is a positive contraction semigroup on $C(\bar{J})$ (i.e.

$$(f \in C(\bar{J}),\, f \geq 0) \Rightarrow (T(t)f \geq 0,\, t \geq 0,\, \|T(t)\| \leq 1),$$

then $(S(t))_{t \geq 0}$ is a positive contraction semigroup on $C(\bar{\Gamma})$.

2.2.1 Generation of a (C_0) semigroup and its explicit representation: the case $\delta c = d\gamma$

In order to solve the (ACP) (2.11), we will start by showing the existence of a (C_0) semigroup generated by $\mathcal{L}$ on $C(\bar{J})$, under the additional assumption $\delta c = d\gamma$.
A preliminary result is presented in the following.

Lemma 2.2.3. *Let us assume that* c, d, γ, δ *satisfy*

$$(\delta c = d\gamma) \ \ and \ \ (c,d) \neq (0,0). \tag{2.15}$$

Then the operator $(\mathcal{L}, D_M(\mathcal{L}))$ *has one of the following expressions, for any* $u \in D_M(\mathcal{L})$,

$$\mathcal{L}u = \frac{1}{2}c^2x^2u'' + k_1xu', \ \ if\ d = 0, \tag{2.16}$$

$$\mathcal{L}u = \frac{1}{2}d^2u'' + k_2u', \ \ if\ c = 0, \tag{2.17}$$

$$\mathcal{L}u = \frac{1}{2}d^2(kx+1)^2u'' + k_3(kx+1)u', \ \ if\ c \neq 0\, and\, d \neq 0. \tag{2.18}$$

Here $k_1 = \gamma - \theta c$, $k_2 = \delta - \theta d$, $k = \frac{d}{c}$, $k_3 = k_2$ *if* $\gamma \neq 0$ *and* $k_3 = -\theta d$ *if* $\gamma = \delta = 0$. *Moreover, for* $\gamma \neq 0$ *and* $d \neq 0$ *we have* $k = \frac{d}{c} = \frac{\gamma}{\delta}$.

Proof. We examine the following cases

$$(i)\ d = 0, \qquad (ii)\ d \neq 0.$$

In the case (i), due to (2.15), we deduce that $c \neq 0$ and $\delta = 0$. Hence, for any $u \in D_M(\mathcal{L})$, we obtain

$$\mathcal{L}u = \frac{1}{2}c^2x^2u'' + (\gamma - \theta c)\, x\, u'$$

and hence, $\mathcal{L}$ has the form (2.16).
In the case (ii), we have to examine the following subcases

$$(ii)_1\ \gamma = c = 0, \qquad (ii)_2\ \gamma = \delta = 0, \qquad (ii)_3\ \gamma \neq 0.$$

In the subcase $(ii)_1$, we obtain

$$\mathcal{L}u = \frac{1}{2}d^2u'' + (\delta - \theta d)u'.$$

Thus $\mathcal{L}$ is of the form (2.17).
In the subcase $(ii)_2$, we can assume $c \neq 0$, otherwise we come back to the case $(ii)_1$. Hence

$$\mathcal{L}u = \frac{1}{2}(cx+d)^2u'' - \theta(cx+d)u' = \frac{1}{2}d^2(kx+1)^2u'' - \theta d(kx+1)u',$$

with $k = \frac{d}{c}$ and $\mathcal{L}$ is of the type (2.18).
In the subcase $(ii)_3$, $\delta \neq 0$ and $c \neq 0$, being $\gamma \neq 0$ and $d \neq 0$. Hence

$$\begin{aligned}\mathcal{L}u &= \frac{1}{2}d^2(\frac{d}{c}x+1)^2u'' + \left[\delta\Big(\frac{\gamma}{\delta}x+1\Big) - \theta d\Big(\frac{d}{c}x+1\Big)\right]u' \\ &= \frac{1}{2}d^2(kx+1)^2u'' + (\delta - \theta d)(kx+1)u',\end{aligned}$$

where $k = \frac{d}{c} = \frac{\gamma}{\delta}$. Thus $\mathcal{L}$ is of the type (2.18). □

We are now in a position to state our first main result of this section.

Theorem 2.2.4. *Assume that* c, d, γ, δ, $\theta \in \mathbb{R}$ *satisfy* (2.15) *and the operator* $\mathcal{L}$ *is defined by* (2.13) *with domain* $D_M(\mathcal{L})$ *given by* (2.12). *Then the operator* $(\mathcal{L}, D_M(\mathcal{L}))$ *generates a positive* (C_0) *contraction semigroup on* $C(\bar{J})$. *Here* $J = (0, +\infty)$ *if* $\mathcal{L}$ *is of the type (*2.16*),* $J = \mathbb{R}$ *if* $\mathcal{L}$ *is of the type (*2.17*) and* $J = (-\frac{d}{c}, +\infty)$ *if* $c > 0$ *(resp.* $J = (-\infty, -\frac{d}{c})$ *if* $c < 0$*) if* $\mathcal{L}$ *is of the type (*2.18*).*

Proof. As consequence of Lemma 2.2.3, the operator $\mathcal{L}$ takes the form (2.16) or (2.17) or (2.18). In the case $\mathcal{L}$ is of the type (2.16), then, according to [42, Theorem 3.2], the operator $\mathcal{L}$ with domain $D_M(\mathcal{L})$ generates a positive (C_0) contraction semigroup on $C[0, +\infty]$. In the case $\mathcal{L}$ is of the type (2.17), the assertion follows from [42, Section IV] and $J = \mathbb{R}$.
Finally, let us focus on $\mathcal{L}$ of the form (2.18). Let us proceed with the change of variable

$$z = \Phi(x) = kx + 1, \tag{2.19}$$

where $k = \frac{d}{c}$. If $c > 0$ (analogous arguments work for $c < 0$), then the operator $\mathcal{L}$ acting on $C[-\frac{d}{c}, +\infty]$ can be transformed in the operator $\widetilde{\mathcal{L}}$ acting on $C[0, +\infty]$, where

$$\widetilde{\mathcal{L}}v = \frac{1}{2}\tilde{k}^2 z^2 v'' + k_3 kzv', \tag{2.20}$$

with $\tilde{k} = kd$, and k_3 defined as in Lemma 2.2.3. We observe that $D(\widetilde{\mathcal{L}})$ has the same expression of $D_M(\mathcal{L})$ where J is replaced by $(0, +\infty)$ (see [42, Theorem 3.2]). Consequently, $\widetilde{\mathcal{L}}$ is of the type (2.16) on $C[0, +\infty]$. Thus, according to the above arguments, $\widetilde{\mathcal{L}}$ generates a positive (C_0) contraction semigroup on $C[0, +\infty]$, where $D_M(\widetilde{\mathcal{L}})$ is obtained easily from $D_M(\mathcal{L})$. Hence, Lemma 2.2.1 and Remark 2.2.2 imply the assertion. □

Remark 2.2.5. From Theorem 1.1.21 follows that $D_M(\mathcal{L})$ coincides with

$$D_W(\mathcal{L}) = \left\{ u \in C(\bar{J}) \cap C^2(J) \middle| \lim_{\substack{x \to r_1 \\ x \to r_2}} \mathcal{L}u(x) = 0 \right\},$$

provided that the assumptions of Theorem 2.2.4 hold.

The next step is to give an *explicit* representation of the solution to (2.11).

Theorem 2.2.6. *Under the assumptions of Theorem* 2.2.4*, fixed* $g \in C(\bar{J})$ *and for any* $(x, t) \in J \times [0, +\infty)$ *the explicit solution to the problem (*2.11*) is given by*
Case 1: *If* $d = 0$*,*

$$u(x,t) = \int_{-\infty}^{+\infty} g(e^{\frac{c}{\sqrt{2}}(\psi t + y)} x) p(t, y) dy, \tag{2.21}$$

with $J = (0, +\infty)$ *and* $\psi = \frac{\sqrt{2}}{c}\gamma - \sqrt{2}\theta - \frac{c}{\sqrt{2}}$.
Case 2: *If* $c = \gamma = 0$*,*

$$u(x,t) = \int_{-\infty}^{+\infty} g\left(x + \frac{d}{\sqrt{2}}(\omega t + y)\right) p(t, y) dy, \tag{2.22}$$

with $J = \mathbb{R}$ *and* $\omega = \frac{\sqrt{2}}{d}(\delta - \theta d)$.
Case 3: *If* $d \neq 0$ *and* $c \neq 0$,

$$u(x,t) = \int_{-\infty}^{+\infty} g\big(e^{\frac{k}{\sqrt{2}}(\chi t+y)}(kx+1)\big)p(t,y)dy, \tag{2.23}$$

with $J = \left(-\frac{d}{c}, +\infty\right)$ *if* $c > 0$ *(resp.* $J = \left(-\infty, -\frac{d}{c}\right)$ *if* $c < 0$*) and* $\chi = \frac{\sqrt{2}}{d}\, k_3 - \frac{\tilde{k}}{\sqrt{2}}$.
In all cases

$$p(t,y) = \frac{1}{2\sqrt{\pi t}} e^{-\frac{y^2}{4t}}, \qquad t > 0,\ y \in \mathbb{R}. \tag{2.24}$$

Proof. **Case 1.** According to Lemma 2.2.3, if $d = 0$ the operator $\mathcal{L}$ can be written as in (2.16). Thus define the operator $Gu = \dfrac{c}{\sqrt{2}}\, xu'$ with domain

$$D(G) = \{u \in C[0,+\infty] \cap C^1(0,+\infty) | u', xu' \in C[0,+\infty]\}. \tag{2.25}$$

Hence, the square of G is given by

$$G^2 u = \frac{c}{\sqrt{2}} x \left(\frac{c}{\sqrt{2}} xu'\right)' = \frac{c^2}{2} x^2 u'' + \frac{c^2}{2} xu',$$

with domain

$$\begin{aligned} D(G^2) &= \{u \in D(G) | Gu \in D(G)\} \\ &= \{u \in C[0,+\infty] \cap C^2(0,+\infty) | u', xu', x(xu')' \in C[0,+\infty]\}. \end{aligned}$$

Therefore, for any $u \in D(G^2)$, the operator $\mathcal{L}$ can be written as

$$\mathcal{L}u = G^2 u + \psi G u, \tag{2.26}$$

with $\psi = \frac{\sqrt{2}}{c}\gamma - \sqrt{2}\theta - \frac{c}{\sqrt{2}}$. Notice that $(G, D(G))$ generates a (C_0) group on $C[0,+\infty]$ and, according to [42, Section 3], it is a suitable perturbation of $(G^2, D(G^2))$. Hence, by [44, Chapter II, Section 8], $(\mathcal{L}, D(G^2))$ generates the following (C_0) contraction semigroup

$$\begin{aligned} T(t)g(x) &= \int_0^{+\infty} \big(g\big(e^{\frac{c}{\sqrt{2}}(\psi t+y)}x\big) + g\big(e^{\frac{c}{\sqrt{2}}(\psi t-y)}x\big)\big)p(t,y)dy \\ &= \int_{-\infty}^{+\infty} g\big(e^{\frac{c}{\sqrt{2}}(\psi t+y)}x\big)\, p(t,y)\, dy \end{aligned}$$

for any $t \geq 0$, and therefore the representation (2.21) follows.
Case 2: If $c = \gamma = 0$, then the operator $\mathcal{L}$ can be written as in (2.17). Let us define

$$Gu = \frac{d}{\sqrt{2}} u', \quad D(G) = \{u \in C(\bar{\mathbb{R}}) | u' \in C(\mathbb{R})\}. \tag{2.27}$$

Then, the square of G is given by

$$G^2 u = \frac{d}{\sqrt{2}} \left(\frac{d}{\sqrt{2}} u'\right)' = \frac{d^2}{2} u'', \tag{2.28}$$

with domain

$$D(G^2) = \{u \in D(G) | Gu \in D(G)\} = \{u \in C(\overline{\mathbb{R}}) \cap C^2(\mathbb{R}) | u', u'' \in C(\overline{\mathbb{R}})\}. \quad (2.29)$$

Thus, for any $u \in D(G^2)$, $\mathcal{L}$ can be written as

$$\mathcal{L}u = \frac{1}{2}c^2x^2u'' + k_1xu' = G^2u + \omega Gu, \quad (2.30)$$

with $\omega = \frac{\sqrt{2}}{d}(\delta - \theta d)$. The operator $(G, D(G))$ generates a (C_0) group on $C[0, +\infty]$ and according to [42, Section 4], it is a suitable perturbation of $(G^2, D(G^2))$. Then $(\mathcal{L}, D(G^2))$ generates the following (C_0) contraction semigroup

$$\begin{aligned} T(t)g(x) &= \int_0^{+\infty} \Big(g\Big(x + \frac{d}{\sqrt{2}}(\omega t + y)\Big) + g\Big(x + \frac{d}{\sqrt{2}}(\omega t - y)\Big)\Big)p(t, y)\, dy \\ &= \int_{-\infty}^{-\infty} g\Big(x + \frac{d}{\sqrt{2}}(\omega t + y)\Big)\, p(t, y)\, dy, \end{aligned}$$

and therefore the representation (2.22) follows.
Case 3: If $c \neq 0$, $d \neq 0$ the operator $\mathcal{L}$ can be written as in (2.18). By the change of variable (2.19), $\mathcal{L}$ can be transformed into the operator $\widetilde{\mathcal{L}}$ defined in (2.20), acting on $C[0, +\infty]$. Hence, let us define

$$Gv = \frac{\tilde{k}}{\sqrt{2}}zv', \quad D(G) = \{v \in C[0, +\infty] \cap C^1(0, +\infty) | v', zv' \in C[0, +\infty]\}.$$

Thus, the square of G is given by

$$G^2v = \frac{\tilde{k}}{\sqrt{2}}z(\frac{\tilde{k}}{\sqrt{2}}zv')' = \frac{\tilde{k}^2}{2}z^2v'' + \frac{\tilde{k}^2}{2}zv',$$

with domain

$$\begin{aligned} D(G^2) &= \{v \in D(G) | Gv \in D(G)\} \\ &= \{v \in C[0, +\infty] \cap C^2(0, +\infty) | v', zv', z(zv')' \in C[0, +\infty]\}. \end{aligned}$$

Therefore, for any $v \in D(G^2)$, $\widetilde{\mathcal{L}}$ can be written as

$$\widetilde{\mathcal{L}}v = G^2v + \chi Gv, \quad (2.31)$$

where $\chi = \frac{\sqrt{2}}{d}k_3 - \frac{\tilde{k}}{\sqrt{2}}$. Analogous arguments as for the Case 1 lead to state that $(\mathcal{L}, D(G^2))$ generates the following (C_0) contraction semigroup

$$T(t)g(x) = \int_{-\infty}^{+\infty} g\big(e^{\frac{k}{\sqrt{2}}(\chi t + y)}(kx + 1)\big)\, p(t, y)\, dy$$

for any $t \geq 0$ and $x \in J = \big(-\frac{d}{c}, +\infty\big)$ if $c > 0$ (resp. $x \in J = \big(-\infty, -\frac{d}{c}\big)$ if $c < 0$), and therefore the representation (2.23) follows.

□

2.2.2 Generation of a (C_0) semigroup and its approximate representation: the case $c = 0$, $\gamma \neq 0$

In this section we prove the existence and uniqueness of the solution to (2.11) in the case $c = 0$, $\gamma \neq 0$. Further, by considering similar arguments to those used in [40], we derive an approximate formula for the solution. Notice that the assumption $c = 0$ implies $d > 0$ because the diffusion coefficient $a(x) = cx + d$ in the SDE (2.6) must be positive.

Theorem 2.2.7. *Assume $c = 0$, $\gamma \neq 0$, $d > 0$ and $\delta, \theta \in \mathbb{R}$, so that $\mathcal{L}$, with maximal domain $D_M(\mathcal{L})$ defined in* (2.12)*, takes the form*

$$\mathcal{L}u = \frac{1}{2}d^2u'' + (\gamma x + (\delta - \theta d))\, u'. \tag{2.32}$$

*Then, for any $g \in C(\overline{\mathbb{R}})$, there exists a unique solution to the (ACP) (*2.11*).*

Proof. First of all, we note that the operator $\mathcal{L}$ in (2.32) can be rewritten as

$$\mathcal{L}u = \frac{1}{2}d^2u'' + \gamma\left(x + \frac{(\delta - \theta d)}{\gamma}\right) u'. \tag{2.33}$$

Thus, by the change of variable

$$z = \Phi(x) = x + \frac{(\delta - \theta d)}{\gamma} \tag{2.34}$$

the operator (2.33) is transformed in the following operator

$$\widetilde{\mathcal{L}}u = \frac{1}{2}d^2v'' + \gamma z v'. \tag{2.35}$$

Hence, to prove the existence and uniqueness of the solution to (ACP), it is sufficient to show that the boundary endpoints $\pm\infty$ are of entrance or natural type (see Theorem 1.1.21).

We start to study the boundary point $-\infty$. For sake of simplicity, we set $z_0 = -1$ and compute (see Definition 1.1.20)

$$W(z) = \exp\left(-\int_{-1}^{z} \frac{2\gamma s}{d^2}\, ds\right) = \exp\left(-\frac{\gamma}{d^2}(z^2 - 1)\right) = e^{-\alpha} e^{\alpha\, z^2},$$

with $\alpha = -\frac{\gamma}{d^2} \neq 0$. Then

$$\int_{-\infty}^{-1} W(z)\, dz = \frac{e^{-\alpha}}{2\alpha}\int_{-\infty}^{-1} \frac{2\alpha z e^{\alpha z^2}}{z}\, dz = \frac{e^{-\alpha}}{2\alpha}\int_{-\infty}^{-1} \frac{1}{z}\frac{d}{dz}(e^{\alpha z^2})\, dz.$$

Observe that, by twice integration by parts,

$$\int \frac{1}{z}\frac{d}{dz}(e^{\alpha z^2})\, dz = \frac{1}{z}e^{\alpha z^2} + \frac{1}{2\alpha z^3}e^{\alpha z^2} + \frac{3}{4\alpha^2}\int \frac{1}{z^5}\frac{d}{dz}(e^{\alpha z^2})\, dz,$$

and therefore, as $z \to -\infty$

$$\int \frac{1}{z}\frac{d}{dz}(e^{\alpha z^2})\, dz \simeq \frac{1}{z}e^{\alpha z^2}. \tag{2.36}$$

Moreover, by noting that in our case $m(z) = d^2/2$, so that $(m(z)W(z))^{-1} = \frac{2e^\alpha}{d^2} e^{-\alpha z^2}$, we compute

$$\int_{-\infty}^{z} (m(s)W(s))^{-1}\, ds = \frac{2e^\alpha}{d^2} \int_{-\infty}^{z} e^{-\alpha s^2}\, ds = -\frac{e^\alpha}{\alpha d^2} \int_{-\infty}^{z} \frac{1}{s}\frac{d}{ds}(e^{-\alpha s^2})\, ds.$$

With similar arguments as before it follows that, as $s \to -\infty$

$$\int \frac{1}{s}\frac{d}{ds}(e^{-\alpha s^2})\, ds \simeq \frac{1}{s} e^{-\alpha s^2}. \tag{2.37}$$

Assume $\alpha > 0$ (i.e. $\gamma < 0$). Thus by (2.36) we can conclude that $W \notin L^1(-\infty, z_0)$ and hence, by Lemma 1.1.22, $R \notin L^1(-\infty, z_0)$.
Further, by (2.37) and Remark 1.1.23, we compute

$$\begin{aligned}\int_{-\infty}^{-1} Q(z)\, dz &= \int_{-\infty}^{-1} \Big(\int_{-\infty}^{z} (m(s)W(s))^{-1}\, ds \Big) W(z)\, dz \\ &\simeq -\frac{1}{\alpha d^2} \int_{-\infty}^{-1} \frac{1}{z} e^{-\alpha z^2} e^{\alpha z^2}\, dz = \frac{1}{\alpha d^2} \int_{1}^{+\infty} \frac{1}{z}\, dz = +\infty,\end{aligned}$$

and therefore $Q \notin L^1(-\infty, z_0)$.
Assume now $\alpha < 0$ (i.e. $\gamma > 0$). Thus by (2.37) we can conclude that $(m(z)W(z))^{-1} \notin L^1(-\infty, z_0)$ and hence, by Lemma 1.1.22, $Q \notin L^1(-\infty, z_0)$.
Further, by (2.36) and Remark 1.1.23 we compute

$$\begin{aligned}\int_{-\infty}^{-1} R(z)\, dz &= \int_{-\infty}^{-1} \Big(\int_{-\infty}^{z} W(s)\, ds \Big) (m(z)W(z))^{-1}\, dz \\ &\simeq \frac{1}{\alpha d^2} \int_{-\infty}^{-1} \frac{1}{z} e^{\alpha z^2} e^{-\alpha z^2}\, dz = -\frac{1}{\alpha d^2} \int_{1}^{+\infty} \frac{1}{z}\, dz = +\infty,\end{aligned}$$

and therefore $R \notin L^1(-\infty, z_0)$. We can then conclude that $-\infty$ is of natural type for all $\alpha \neq 0$.
We consider now the endpoint $+\infty$, set $z_0 = 1$ for sake of simplicity. Similar calculations as carried out for $-\infty$ allow to conclude that $+\infty$ is of natural type too, for all $\alpha \neq 0$.
By Lemma 2.2.1 the boundary points $\pm\infty$ have the same type of Feller classification with respect to the operator $\mathcal{L}$. This complete the proof. □

The Lie-Trotter-Daletskii product formula (see Proposition 1.2.7) implies the following theorem.

Theorem 2.2.8. *By referring to Theorem* 2.2.7*, the solution to the (ACP) (*2.11*) admits the following approximate formula*

$$u(x,t) = \lim_{n \to +\infty} u_n(x,t),$$

uniformly in $x \in \mathbb{R}$ *and for* t *in bounded intervals of* $[0, \infty)$*. Here* $u_n(\cdot,\cdot), n \geq 1$*, is a sequence of approximate solutions given by*

$$u_n(x,t) = \underbrace{\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty}}_{n\, times} L_0(t, n, \{y_j\}_{1 \leq j \leq n}, x, g) \prod_{j=1}^{n} p\Big(\frac{t}{n}, y_j\Big) dy_1 \ldots dy_n, \tag{2.38}$$

where

$$L_0(t,n,\{y_j\}_{1\le j\le n},x,g) = g\left[\left(x+\frac{(\delta-\theta d)}{\gamma}\right)e^{\gamma t}+\xi\sum_{j=1}^{n} y_j e^{\gamma(n-j)t/n}\right], \qquad (2.39)$$

with $n=2^k,\ k\in\mathbb{N},\ \xi=\frac{d}{\sqrt{2}}$, *and* $p(t,y)$ *is defined in (*2.24*).*

Proof. Consider the operator $\widetilde{\mathcal{L}}$ defined in (2.35), obtained from $\mathcal{L}$ by the change of variable (2.34). Moreover, consider the operator $(G,D(G))$ defined in (2.27) and the operator

$$G_1 v=\gamma z v',\quad D(G_1)=\{v\in C(\bar{\mathbb{R}})\cap C^1(\mathbb{R})|zv'\in C(\bar{\mathbb{R}})\}.$$

The square $(G^2,D(G^2))$ defined in (2.28), (2.29) represents the well known heat operator that generates a (C_0) contraction semigroup on $C(\bar{\mathbb{R}})$ defined by

$$U(t)f(z)=\int_{-\infty}^{+\infty} f(z+\xi y)p(t,y)\,dy, \qquad (2.40)$$

for $t\ge 0$, $z\in\mathbb{R}$, $f\in C(\overline{\mathbb{R}})$ (see, e.g. [43]), with $\xi=\frac{d}{\sqrt{2}}$ and $p(t,y)$ given in 2.24.
Further, the operator $(G_1,D(G_1))$ generates a (C_0) semigroup of isometries on $C(\overline{\mathbb{R}})$ given by

$$V(t)f(z)=f(ze^{\gamma t}), \qquad (2.41)$$

for $t\ge 0$, $z\in\mathbb{R}$, $f\in C(\overline{\mathbb{R}})$.
Then, by the Lie-Trotter-Daletskii formula we can conclude that the closure of the operator $(\widetilde{\mathcal{L}},D(G^2)\cap D(G_1))$ generates a (C_0) semigroup $(T(t))_{t\ge 0}$ on $C(\overline{\mathbb{R}})$ defined by

$$T(t)g(z)=\lim_{n\to+\infty}\left[U(t/n)V(t/n)\right]^n g(z), \qquad (2.42)$$

uniformly in $z\in\mathbb{R}$ and for t in bounded intervals of $[0,+\infty)$. Hence, the solution to the (ACP) (2.11) is given by

$$u(z,t)=T(t)g(z)=\lim_{n\to+\infty}\left[U(t/n)V(t/n)\right]^n g(z).$$

Denote $u_n(z,t)=\left[U(t/n)V(t/n)\right]^n g(z)$. To compute the approximate solutions $u_n(\cdot,\cdot), n\ge 1$, we proceed by steps. The details are given in Appendix 6.3. Finally, by Lemma 2.2.1 the formula (2.39) holds. □

Remark 2.2.9. If $\gamma<0$, the operator $\tilde{\mathcal{L}}$ in (2.35) represents the infinitesimal generator of the Ornstein-Uhlenbeck process. The corresponding (C_0) semigroup on $C_0(\overline{\mathbb{R}})$ has the following explicit representation (see e.g. [66, Chapter 12])

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} g\left(ze^{\gamma t}+\xi\sqrt{\frac{e^{2\gamma t}-1}{\gamma}}y\right)e^{-y^2/2}dy, \qquad (2.43)$$

for $t\ge 0$, $z\in\mathbb{R}$, $g\in C_0(\overline{\mathbb{R}})$. In Section 2.3 we will show that the approximate solutions (2.38) converge numerically to (2.43).

Remark 2.2.10. It is worth noting that in the case $c = 0$ the non-traded asset level Y is an affine process since both the drift b and the square of the diffusion coefficient a in (2.2) are time-homogeneous affine functions in $y = Y_t$.

$$\begin{cases} b(y,t) = b(y) = \gamma y + (\delta - \theta d) \\ a^2(y,t) = a^2(y) = d^2, \end{cases}$$

for any γ, δ, θ, $d \in \mathbb{R}$. This condition is equivalent to state that the problem (2.8) admits an affine term structure (for more details on affine processes and affine term structures the reader can refer, for example, to [29] or [30]) and then, its solution is of the form

$$w(y,t) = e^{B(t,T)-A(t,T)y}, \quad (y,t) \in \mathbb{R} \times [0,T], \tag{2.44}$$

where A and B are deterministic functions satisfying the following differential equations

$$\frac{\partial A(t,T)}{\partial t} = -\gamma A(t,T), \tag{2.45}$$

$$\frac{\partial B(t,T)}{\partial t} = -\frac{d^2}{2} A^2(t,T) + A(t,T)(\delta - \theta d), \tag{2.46}$$

obtained by plugging the partial derivatives of w into the parabolic equation in (2.8). For fixed T, equations (2.45) and (2.46) are uniquely solvable ODEs when the final conditions $A(T,T)$ and $B(T,T)$ are known. The final condition in (2.8) implies that $A(T,T)$ and $B(T,T)$ must satisfy the following equation

$$e^{B(T,T)-A(T,T)y} = e^{\eta(1-\rho^2)G(Y_T)}, \tag{2.47}$$

which can be explicitly solved if and only if the payoff function $G(Y_T)$ is a polynomial, as shown in the next Example 2.2.11.
Hence the closed form (2.44) is a useful explicit representation of the solution to problem (2.8) provided that the deterministic functions A and B are explicitly known. For this reason, the approximation formula (2.38) for the case $c = 0$ may be considered a helpful alternative to the Feynman-Kac formula (2.7).

Example 2.2.11. We consider a European option that conveys the opportunity, but not the obligation, to sell an underlying asset at time $t > 0$ for some fixed price $K > 0$. This is known as a put option; the corresponding call option to buy the asset may be treated similarly. Assuming that the option is written exclusively on the non-traded asset Y, its payoff at the maturity date $T > t$ is

$$G(Y_T) = (K - Y_T)^+ = \begin{cases} K - Y_T, & if\ K > Y_T, \\ 0, & if\ K \le Y_T. \end{cases}$$

Thus equation (2.47) becomes

$$e^{B(T,T)-A(T,T)y} = e^{r(K-y)^+} = \begin{cases} e^{r(K-y)}, & if\ K > y, \\ 1, & if\ K \le y, \end{cases}$$

with $r = \eta(1-\rho^2)$ and $Y_T = y \in \mathbb{R}$, and hence

$$\begin{cases} B(T,T) = rK, \quad A(T,T) = r, \quad if\, K > y, \\ B(T,T) = 0, \quad A(T,T) = 0, \quad\quad if\, K \le y. \end{cases}$$

In the case $K > y$ (from (2.7) we deduce that w reduces to the constant function 1 when $K \le y$), the functions A and B solve the following systems

$$\begin{cases} \dfrac{dA(t,T)}{dt} = -\gamma A(t,T) \\ A(T,T) = r, \end{cases} \tag{2.48}$$

$$\begin{cases} \dfrac{dB(t,T)}{dt} = (\delta - \theta d)A(t,T) - \dfrac{d^2}{2}A^2(t,T) \\ B(T,T) = rK. \end{cases} \tag{2.49}$$

Since (2.48) is a simple linear ODE, for fixed T, we immediately obtain

$$A(t,T) = re^{\gamma(T-t)}.$$

Plugging this expression into the so called Riccati equation (2.49) and integrating in the interval $[t,T]$

$$B(t,T) = r(\delta - \theta d)\int_t^T e^{\gamma(T-s)}ds - \frac{r^2 d^2}{2}\int_t^T e^{2\gamma(T-s)}ds,$$

we obtain

$$B(t,T) = -\frac{r(\delta-\theta d)}{\gamma}(1 - e^{\gamma(T-t)}) + \frac{r^2 d^2}{4\gamma}(1 - e^{2\gamma(T-t)}) + rK.$$

2.2.3 Generation of a (C_0) semigroup and its approximate representation: the case $c \neq 0$

In this section we will prove a generation result and an approximation formula for the solution to the (ACP) (2.11) when the condition (2.15) on the parameters c, d, γ, δ, θ is not verified and $c \neq 0$.
Thus the operator $\mathcal{L}$ is defined by

$$\mathcal{L}u = \frac{1}{2}(cx+d)^2 u'' + ((\gamma x + \delta) - \theta(cx+d))u', \tag{2.50}$$

Fix $c > 0$ (analogous arguments work for $c < 0$) and consider the change of variable $z = \Phi(x) = cx + d$, where Φ is the isomorphism that maps $C(-\frac{d}{c}, +\infty)$ into $C(0,+\infty)$. Then $\mathcal{L}$ is transformed in the following operator

$$\widetilde{\mathcal{L}}v = \alpha_0 z^2 v'' + (\alpha_1\, z + \alpha_2)\, v', \tag{2.51}$$

where

$$\alpha_0 = \frac{c^2}{2} > 0, \quad \alpha_1 = (\gamma - \theta\, c) \in \mathbb{R}, \quad \alpha_2 = (\delta\, c - \gamma\, d) \neq 0 \tag{2.52}$$

(if $\alpha_2 = 0$, then condition (2.15) holds).

We are now in position to prove the main result of this section.

Theorem 2.2.12. *Assume* $c \neq 0$, d, γ, δ, $\theta \in \mathbb{R}$ *and consider the operator* $\widetilde{\mathcal{L}}$ *defined in* (2.51) *with the parameters given in* (2.52). *Denote by*

$$h_1 = -\frac{\alpha_1}{\alpha_0} \in \mathbb{R}, \; h_2 = \frac{\alpha_2}{\alpha_0} \neq 0. \tag{2.53}$$

Then

i. for any $h_1 \in \mathbb{R}$ *and* $h_2 > 0$, *the operator* $\widetilde{\mathcal{L}}$ *with maximal domain*

$$D_M(\widetilde{\mathcal{L}}) = \{v \in C[0,+\infty] \cap C^2[0,+\infty) | \, \widetilde{\mathcal{L}}v \in C[0,+\infty]\},$$

generates a positive (C_0) *contraction semigroup on* $C[0,+\infty]$.

ii. for any $h_1 \in \mathbb{R}$ *and* $h_2 < 0$, *the operator* $\widetilde{\mathcal{L}}$ *with Wentzell domain*

$$D_W(\widetilde{\mathcal{L}}) = \left\{ v \in C[0,+\infty] \cap C^2[0,+\infty) \Big| \lim_{\substack{z\to 0 \\ z\to+\infty}} \widetilde{\mathcal{L}}v(z) = 0 \right\},$$

generates a positive (C_0) *contraction semigroup on* $C[0,+\infty]$.

Proof. Consider $J = (0,+\infty)$. According to Theorem 1.1.20, the proof is based on the Feller classification of the endpoints $0, +\infty$. First, we study the boundary point $z = +\infty$. Set $z_0 = 1$ for sake of simplicity. For any $z > 0$, we compute (see Definition 1.1.21)

$$\begin{aligned} W(z) &= \exp\Big(-\int_1^z \frac{(\alpha_1 s + \alpha_2)}{\alpha_0 s^2} \, ds \Big) = \exp\Big(-\frac{1}{\alpha_0} \Big[\alpha_1 \ln z - \frac{\alpha_2}{z} + \alpha_2 \Big] \Big) \\ &= e^{-h_2} z^{h_1} e^{h_2/z}, \end{aligned}$$

and

$$(m(z)W(z))^{-1} = \frac{e^{h_2}}{\alpha_0} \frac{e^{-h_2/z}}{z^{h_1+2}}.$$

with $h_1 = -\frac{\alpha_1}{\alpha_0} \in \mathbb{R}$ and $h_2 = \frac{\alpha_2}{\alpha_0} \neq 0$.
Assume $h_2 > 0$. Since $e^{h_2/z} > 1$ for all $z > 0$, we obtain

$$\begin{aligned} \int_z^{+\infty} W(s) \, ds &= e^{-h_2} \int_z^{+\infty} s^{h_1} e^{h_2/s} \, ds \\ &> e^{-h_2} \int_z^{+\infty} s^{h_1} \, ds = \begin{cases} +\infty, & if \; h_1 \geq -1 \\ -\dfrac{e^{-h_2} \, z^{h_1+1}}{h_1+1}, & if \, h_1 < -1. \end{cases} \end{aligned} \tag{2.54}$$

Hence $W \notin L^1(z_0,+\infty)$ if $h_1 \geq -1$ and therefore, by Lemma 1.1.22, $R \notin L^1(z_0,+\infty)$. If $h_1 < -1$, by Remark 1.1.23 and (2.54) we have

$$\begin{aligned} \int_1^{+\infty} R(z) \, dz &= \int_1^{+\infty} \Big(\int_z^{+\infty} W(s) \, ds \Big) (m(z)W(z))^{-1} \, dz \\ &> -\frac{1}{\alpha_0(h_1+1)} \int_1^{+\infty} z^{h_1+1} \, \frac{e^{-h_2/z}}{z^{h_1+2}} \, dz \\ &> -\frac{1}{\alpha_0(h_1+1)} \int_1^{+\infty} \frac{1}{z} \, dz = +\infty, \end{aligned}$$

where the last inequality follows by observing that $e^{-h_2/z} < 1$ for $z > 1$ Moreover, for any $z > 0$ we compute

$$\begin{aligned}
\int_z^{+\infty} (m(s)W(s))^{-1}\,ds &= \frac{e^{h_2}}{\alpha_0}\int_z^{+\infty} \frac{e^{-h_2/s}}{s^{h_1+2}}\,ds \\
&\qquad (e^{-h_2/s} > e^{-h_2/z} \text{ for all } s > z) \\
&> \frac{e^{h_2}}{\alpha_0} e^{-h_2/z}\int_z^{+\infty} \frac{1}{s^{h_1+2}}\,ds = \begin{cases} +\infty, & if\ h \le -1 \\ \frac{e^{h_2}}{\alpha_0(h_1+1)}\frac{e^{-h_2/z}}{z^{h_1+1}}, & if\ h > -1. \end{cases}
\end{aligned} \tag{2.55}$$

Hence $(m\,W)^{-1} \notin L^1(z_0,+\infty)$ if $h_1 \le -1$ and therefore, by Lemma 1.1.22, $Q \notin L^1(z_0,+\infty)$.
If $h_1 > -1$, by Remark 1.1.23 and (2.55) we have

$$\begin{aligned}
\int_1^{+\infty} Q(z)\,dz &= \int_1^{+\infty}\Big(\int_z^{+\infty} (m(s)W(s))^{-1}\,ds\Big)W(z)\,dz \\
&> \frac{1}{\alpha_0(h_1+1)}\int_1^{+\infty} \frac{e^{-h_2/z}}{z^{h_1+1}}\, z^{h_1}e^{h_2/z}\,dz = \frac{1}{\alpha_0(h_1+1)}\int_1^{+\infty}\frac{1}{z}\,dz = +\infty.
\end{aligned}$$

We have then proved that $+\infty$ is natural for any $h_2 > 0$ and $h_1 \in \mathbb{R}$.
Now, assume $h_2 < 0$. Observe that $e^{h_2/s} > e^{h_2/z}$ for all $s > z > 0$. Thus

$$\begin{aligned}
\int_z^{+\infty} W(s)\,ds &= e^{-h_2}\int_z^{+\infty} e^{h_2/s}s^{h_1}\,ds \\
&> e^{-h_2}e^{h_2/z}\int_z^{+\infty} s^{h_1}\,ds = \begin{cases} +\infty, & if\ h \ge -1 \\ -\frac{e^{-h_2}}{h_1+1}\frac{e^{h_2/z}}{z^{h_1+1}}, & if\ h < -1. \end{cases}
\end{aligned}$$

By Lemma 1.1.22, that implies $R \notin L^1(z_0,+\infty)$ if $h_1 \ge -1$. If $h_1 < -1$, with similar calculations as in the case $h_2 > 0$, we obtain that

$$\int_1^{+\infty} R(z)\,dz = +\infty,$$

and therefore, we can conlude that $R \notin L^1(z_0,+\infty)$ for any $h_2 < 0$ and $h_1 \in \mathbb{R}$.
Moreover, by observing that $e^{-h_2/z} > 1$ for any $z > 0$, we have

$$\begin{aligned}
\int_z^{+\infty} (m(s)W(s))^{-1}\,ds &= \frac{e^{h_2}}{\alpha_0}\int_z^{+\infty} \frac{e^{-h_2/s}}{s^{h_1+2}}\,ds \\
&> \frac{e^{h_2}}{\alpha_0}\int_z^{+\infty} \frac{1}{s^{h_1+2}}\,ds = \begin{cases} +\infty, & if\ h_1 \le -1 \\ \frac{e^{h_2}}{\alpha_0(h_1+1)}\frac{1}{z^{h_1+1}}, & if\ h_1 > -1. \end{cases}
\end{aligned}$$

Then, by Lemma 1.1.22 it follows that $Q \notin L^1(z_0,+\infty)$ when $h \le -1$.
If $h > -1$, with similar calculations as in the case $h_2 > 0$, we obtain that

$$\int_1^{+\infty} Q(z)\,dz = +\infty.$$

Hence, the endpoint $+\infty$ is natural when $h_2 < 0$ and $h_1 \in \mathbb{R}$. Therefore, we have proved that $+\infty$ is natural for all $h_2 \neq 0$ and $h_1 \in \mathbb{R}$.
We study now the behaviour at the boundary point 0.
First, assume $h_2 > 0$. Without any loss of generality, set $z_0 = 1$. Observe that $e^{h_2/s} > e^{h_2/z}$ for any $0 < s < z$. Thus

$$\int_0^z W(s)\,ds > e^{-h_2}e^{h_2/z}\int_0^z s^{h_1}\,ds = \begin{cases} +\infty, & if\ h_1 \leq -1 \\ \frac{e^{-h_2}}{h_1+1}\,e^{h_2/z}z^{h_1+1}, & if\ h_1 > -1, \end{cases} \tag{2.56}$$

which implies $R \notin L^1(0, z_0)$ when $h_1 \leq -1$. If $h_1 > -1$, by (2.56) we obtain

$$\begin{aligned}\int_0^1 R(z)\,dz &= \int_0^1 \Big(\int_1^z W(s)\,ds\Big)(m(z)W(z))^{-1}\,dz \\ &> \frac{1}{\alpha_0(h_1+1)}\int_0^1 \frac{e^{h_2/z}e^{-h_2/z}z^{h_1+1}}{z^{h_1+2}}\,dz = \frac{1}{\alpha_0(h_1+1)}\int_0^1 \frac{1}{z}\,dz = +\infty,\end{aligned}$$

and therefore, $R \notin L^1(0, z_0)$ for any $h_2 > 0$ and $h_1 \in \mathbb{R}$. Moreover, for any $z > 0$

$$0 < \int_0^z (m(s)W(s))^{-1}\,ds < +\infty.$$

Hence $Q \in L^1(0, z_0)$ or $Q \notin L^1(0, z_0)$ that is, the endpoint 0 may be of entrance or natural type for any $h_2 > 0$ and $h_1 \in \mathbb{R}$. Then from Theorem 1.1.21 the assertion *i.* holds.
Assume now $h_2 < 0$. Since $e^{-h_2/s} > e^{-h_2/z}$ for $0 < s < z$, we compute

$$\begin{aligned}\int_0^z (m(z)W(z))^{-1}\,dz &= \frac{e^{h_2}}{\alpha_0}\int_0^z \frac{e^{-h_2/s}}{s^{h_1+2}}\,ds \\ &> \frac{e^{h_2}}{\alpha_0}e^{-h_2/z}\int_0^z \frac{1}{s^{h_1+2}}\,ds = \begin{cases} +\infty, & if\ h_1 \geq -1 \\ -\frac{e^{h_2}}{\alpha_0(h_1+1)}\frac{e^{-h_2/z}}{z^{h_1+1}}, & if\ h_1 < -1, \end{cases}\end{aligned}$$

wich implies $Q \notin L^1(0, z_0)$ when $h_1 \geq -1$. If $h_1 < -1$

$$\begin{aligned}\int_0^1 Q(z)\,dz &= \int_0^1 \Big(\int_0^z (m(s)W(s))^{-1}\,ds\Big)W(z)\,dz \\ &> -\frac{1}{\alpha_0(h_1+1)}\int_0^1 \frac{e^{-h_2/z}\,z^{h_1}\,e^{h_2/z}}{z^{h_1+1}}\,dz = -\frac{1}{\alpha_0(h_1+1)}\int_0^1 \frac{1}{z}\,dz = +\infty;\end{aligned}$$

and therefore, $Q \notin L^1(r_1, x_0)$ for any $h_2 < 0$ and $h_1 \in \mathbb{R}$. Moreover, for any $z > 0$,

$$0 < \int_0^z W(s)\,ds < +\infty,$$

Hence $R \in L^1(0, z_0)$ or $R \notin L^1(0, z_0)$ that is, 0 may be of exit or natural type when $h_2 < 0$ and $h_1 \in \mathbb{R}$. This implies that both the boundary points 0 and $+\infty$ are not of entrance type for any $h_2 < 0$ and $h_1 \in \mathbb{R}$. Therefore, from Theorem 1.1.21, the assertion *ii.* holds. $\square$

Remark 2.2.13. Consider $J = \left(\frac{-d}{c}, +\infty\right)$ if $c > 0$ (resp. $J = \left(-\infty, \frac{-d}{c}\right)$ if $c < 0$). From Remark 2.2.2 and Theorem 2.2.12 follows that

i) for any $h_1 \in \mathbb{R}$ and $h_2 > 0$, the operator $(\mathcal{L}, D_M(\mathcal{L}))$ generates a positive (C_0) contraction semigroup on $C(\bar{J})$;

ii) for any $h_1 \in \mathbb{R}$ and $h_2 < 0$, the operator $(\mathcal{L}, D_W(\mathcal{L}))$ generates a positive (C_0) contraction semigroup on $C(\bar{J})$,

where $\mathcal{L}$ is given in (2.50) and h_1, h_2 are defined in (2.53). We recall that the case $h_2 = 0$ corresponds to condition (2.15). Then, for any $g \in C(\bar{J})$, there exists a unique solution to the (ACP) (2.11).

Remark 2.2.14. If $c > 0$, $\alpha_1 < 0\,(\Rightarrow h_1 > 0)$, $\alpha_2 > 0\,(\Rightarrow h_2 > 0)$, the operator $\widetilde{\mathcal{L}}$ in (2.51) represents the infinitesimal generator of a diffusion process considered by Brennan and Schwartz [14] and, successively, by Courtadon [26] to model the dynamic behaviour of interest rates for the valuation of default-free bonds and prices of European options written on default-free bonds.

As in the previous section, the Lie-Trotter-Daletskii product formula implies the following result.

Theorem 2.2.15. *Assume $c \neq 0$, d, $\gamma\,\delta$, $\theta \in \mathbb{R}$. Consider $J = \left(\frac{-d}{c}, +\infty\right)$ if $c > 0$ (resp. $J = \left(-\infty, \frac{-d}{c}\right)$ if $c < 0$). Then the unique solution u to (ACP) (2.11) admits the following approximate formula*

$$u(x,t) = \lim_{n \to +\infty} u_n(x,t),$$

uniformly for $x \in J$ and for t in bounded intervals of $[0,\infty)$. The sequence $u_n(\cdot,\cdot), n \geq 1$, of approximate solutions is given by

$$u_n(x,t) = \underbrace{\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty}}_{n\, times} L(t,n,\{y_j\}_{1\leq j\leq n}, x, g) \prod_{j=1}^{n} p\left(\frac{t}{n}, y_j\right) dy_1 \ldots dy_n, \tag{2.57}$$

where

$$L(t,n,\{y_j\}_{1\leq j\leq n}, x, g) = \\ g\left(e^{\beta(\sum_{i=1}^{n} y_i + \zeta t)}\left(x + \frac{d}{c}\right) - \frac{d}{c} + \frac{\alpha_2 t}{nc} \sum_{j=1}^{n} e^{\left[\beta\left(\sum_{i=j}^{n} y_i + \left(\frac{n-j+1}{n}\right)\zeta t\right)\right]} \right), \tag{2.58}$$

for $g \in C(\bar{J})$, with $n = 2^k$, $k \in \mathbb{N}$. The parameters α_1, α_2 are defined in (2.52), $\zeta = \sqrt{2}\left(\frac{\alpha_1}{c} - \frac{c}{2}\right)$, and $p(t,y)$ is given in (2.24).

Proof. We start by rewriting the operator $\mathcal{L}$ defined in (2.50) as follows

$$\begin{aligned} \mathcal{L}u &= \frac{1}{2}(cx+d)^2 u'' + ((\gamma - \theta c)x + (\delta - \theta d))u' \\ &= \frac{1}{2}(cx+d)^2 u'' + \left(\left(\frac{\gamma}{c} - \theta\right)(cx+d) + \left(\delta - \frac{\gamma d}{c}\right)\right)u' \\ &= \frac{1}{2}(cx+d)^2 u'' + \left(\frac{\alpha_1}{c}(cx+d) + \frac{\alpha_2}{c}\right)u'. \end{aligned} \tag{2.59}$$

Let us introduce the operator

$$Gu = \frac{cx+d}{\sqrt{2}}u', \quad D(G) = \{u \in C(\bar{J}) \cap C^1(J) | u', xu' \in C(\bar{J})\}.$$

Then the square of G is given by

$$G^2 u = \frac{cx+d}{\sqrt{2}}\left(\frac{cx+d}{\sqrt{2}}u'\right)' = \frac{1}{2}(cx+d)^2 u'' + \frac{c}{2}(cx+d)u',$$

with domain

$$\begin{aligned} D(G^2) &= \{u \in D(G) | Gu \in D(G)\} \\ &= \{u \in C(\bar{J}) \cap C^2(J) | u', xu', x(xu')' \in C(\bar{J})\}. \end{aligned}$$

Hence, the operator $\mathcal{L}$ can be written in the form

$$\mathcal{L}u = P_1 u + P_2 u, \tag{2.60}$$

where

$$P_1 u = G^2 u + \zeta G u, \quad D(P_1) = D(G^2),$$

with $\zeta = \sqrt{2}(\frac{\alpha_1}{c} - \frac{c}{2})$, and

$$P_2 u = \frac{\alpha_2}{c}u', \quad D(P_2) = \{u \in C(\bar{J}) \cap C^1(J) | u' \in C(\bar{J})\}.$$

According to [40, Section 5, Lemma 3], $(G, D(G))$ generates a (C_0) group of isometries, $(S(t))_{t\geq 0}$, on $C(\bar{J})$, given by

$$S(t)f(x) = e^{tG}f(x) = f\left(e^{\beta\, t}x + \frac{d}{c}(e^{t\beta} - 1)\right), \tag{2.61}$$

with $\beta = \frac{c}{\sqrt{2}}$, for any $f \in C(\bar{J})$, $t > 0$, $x \in J$. Indeed, $(G, D(G))$ and $(-G, D(G))$ are respectively generators of the (C_0) semigroups $(S_+(t))_{t\geq 0}$ and $(S_-(t))_{t\geq 0}$ on $C(\bar{J})$, where $S_+(t) = S(t)$ and $S_-(t) = S(-t)$ for $t > 0$. The first result is proved by [40, Section 5, Lemma 3], the second one can be proved analogously.
Thus, $(P_1, D(P_1))$ generates a (C_0) contraction semigroup (see [40, Sections 3-4] and [44, Chapter I, Section 9, and Chapter II, Section 8]) given by

$$\begin{aligned} U(t)f(x) &= \int_0^{+\infty} S(y)\, S(\zeta\, t)f(x)p(t,y)dy + \int_0^{+\infty} S(-y)\, S(\zeta\, t)f(x)p(t,y)\, dy \\ &= \int_{-\infty}^{+\infty} S(\zeta\, t + y)f(x)p(t,y\,)dy \\ &= \int_{-\infty}^{+\infty} f\left[e^{\beta(\zeta\, t+y)}x + \frac{d}{c}(e^{\beta(\zeta\, t+y)} - 1)\right] p(t,y)\, dy, \end{aligned} \tag{2.62}$$

for any $f \in C(\bar{J})$, $t > 0$, $x \in J$, and $p(t,y)$ is defined in (2.24).
Further, it is well known that $(P_2, D(P_2))$ generates the (C_0) contraction semigroup given by

$$V(t)f(x) = f\left(x + \frac{\alpha_2}{c}t\right), \tag{2.63}$$

for any $f \in C(\bar{J})$, $t > 0$, $x \in J$. Finally, by the Lie-Trotter-Daletskii product formula the operator $(\mathcal{L}, D(G^2))$ generates a (C_0) semigroup $(T(t))_{t\geq 0}$ on $C(\bar{J})$ defined as

$$T(t)g(x) = \lim_{n\to+\infty} \Big[U(t/n)V(t/n)\Big]^n g(x), \tag{2.64}$$

uniformly in $x \in J$ and for t in bounded intervals of $[0, +\infty)$. Hence, the solution to the (ACP) (2.11) is given by

$$u(x,t) = T(t)g(x) = \lim_{n\to+\infty} [U(t/n)V(t/n)]^n g(x),$$

for any $g \in C(\bar{J})$, $t > 0$, $x \in J$. Denote $u_n(x,t) = [U(t/n)V(t/n)]^n g(x)$. Thus $u_n(\cdot,\cdot), n \geq 1$, is a sequence of approximating solutions whose explicit expression (2.57)-(2.58) is computed in the Appendix 6.4. $\square$

2.3 Numerical Evaluations

In this section we focus on some numerical applications and examples related to the approximate solutions given in (2.38) and (2.57). It is well known that very few exact solution formulas to Cauchy problems associated with financial models are available. Therefore, one can choose among different computational methods: finite difference methods, finite element methods, finite volume methods, spectral methods, etc...(the reader can refer to [1] and references therein), which are in general very slow.

It is worth noting that formulas (2.38) and (2.57) seems to be hard to numerically calculate because of the high computational cost for solving a n-dimensional integral when n is large. Some numerical methods have been proposed in literature to faster evaluate a multidimensional integral (see, for example, [80]). However, since the function $p(t,y)$ defined in (2.24) is the probability density of a normal distribution, $N(0, 2t)$, with mean 0 and variance $2t$, the integration variables $\{y_j\}_{1\leq j\leq n}$ are n independent realizations of a normal distribution $N(0, 2t/n)$. Thus, the n-dimensional integrals (2.38) and (2.57) are nothing but the conditional expected value of a function of n independent and normally distribuited random variables $Y_j \sim N(0, 2t/n)$, that can be estimated by the Monte Carlo integration method.

In the following examples we compute the approximate solutions for both cases $c = 0$ and $c \neq 0$, when the payoff G refers to some well known fixed income derivatives. In particular, we consider a put European option with maturity T and strike price K, whose payoff is given by $G(x) = (K - x)^+$ (where $x \equiv Y_T$). Hence, the initial condition of (2.11), defined in (2.14), can be written as

$$g(x) = e^{r(K-x)^+}, \tag{2.65}$$

with $r = \eta(1 - \rho^2)$.

Example 2.3.1. (The case $c = 0$)

As observed at the beginning of this section, the solution to the (ACP) (2.11) admits the approximate formula (Theorem 2.2.12)

$$\begin{aligned} u(x,t) &= \lim_{n\to+\infty} u_n(x,t) \\ &= \lim_{n\to+\infty} \underbrace{\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty}}_{n\,times} L_0(t,n,\{y_j\}_{1\le j\le n},x,g) \prod_{j=1}^{n} p\Big(\frac{t}{n},y_j\Big)dy_1 \dots dy_n \\ &= \lim_{n\to+\infty} \mathbb{E}_p[L_0(t,n,\{Y_j\}_{1\le j\le n},x,g)], \end{aligned} \tag{2.66}$$

for any $x \in \mathbb{R}$, $t \geq 0$ and $g \in C(\overline{\mathbb{R}})$, where the function L_0 is defined in (2.39). Assuming that the initial condition g is of type (2.65), L_0 may be written as

$$L_0(t,n,\{y_j\}_{1\le j\le n},x,g) = \exp\Big[r\Big(K - \Big(x + \frac{(\delta-\theta d)}{\gamma}\Big)e^{\gamma t} - \xi \sum_{j=1}^{n} y_j e^{\gamma(n-j)t/n}\Big)^+\Big]. \tag{2.67}$$

As shown in Remark 2.2.10, the solution to the (ACP) (2.11) is also given by the affine type solution formula derived in Example 2.2.11 by taking into account the time transformation $t \to (T-t)$. Therefore, the affine type solution (ATS) is as follows

$$u(t,x) = e^{B(t)-A(t)x}, \quad Y_0 = x \tag{2.68}$$

with

$$\begin{cases} A(t) = re^{\gamma t}, \\ B(t) = \dfrac{r^2 d^2}{4\gamma}(e^{2\gamma t}-1) - \dfrac{r(\delta-\theta d)}{\gamma}(e^{\gamma t}-1) + Kr. \end{cases}$$

Further, as observed in Remark 2.2.9, if the parameter $\gamma < 0$, the solution to (2.11) can be also expressed in closed form by the Ornstein-Uhlenbeck semigroup as follows

$$u(t,x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\Big[r\Big(K - \Big(x + \frac{(\delta-\theta d)}{\gamma}\Big)e^{\gamma t} - \xi\sqrt{\frac{e^{2\gamma t}-1}{\gamma}}y\Big)^+\Big] e^{-y^2/2} dy, \tag{2.69}$$

for $t \geq 0$, $z \in \mathbb{R}$, $g \in C_0(\overline{\mathbb{R}})$. Using basic probability calculations one can easily show that the term $\sum_{j=1}^{n} y_j e^{\gamma(n-j)t/n}$ in (2.67) converges to $\sqrt{\frac{e^{2\gamma t}-1}{\gamma}}y$ as $n \to +\infty$, where y is a realization of a standard normal distribution $N(0,1)$.

Figure 2.1 plots the behaviour of both the function w solving the parabolic problem (2.8) or, equivalently, (2.9), and the corresponding indifference price h, defined in (2.4), for a put European option, under the following suitably chosen parameter values

$$\begin{array}{ccccc} d = 0.2016, & \gamma = -0.9593, & \delta = 0.3209, & \mu = 0.0380, & \sigma = 0.0300, \\ \rho = 0.8, & \eta = 2, & Y_0 = 8, & K = 12, & T = 8. \end{array}$$

The function w is computed via the approximate formula (2.66), for $n = 2^k$ with $k = 1, 3, 5$, by applying a Monte Carlo integration method. In particular, when $k = 5$, the plots in Figure 2.1 show an accurate goodness-of-fit to the curves obtained, respectively, by the closed form ATS (2.68) and the OU type solution (2.69).

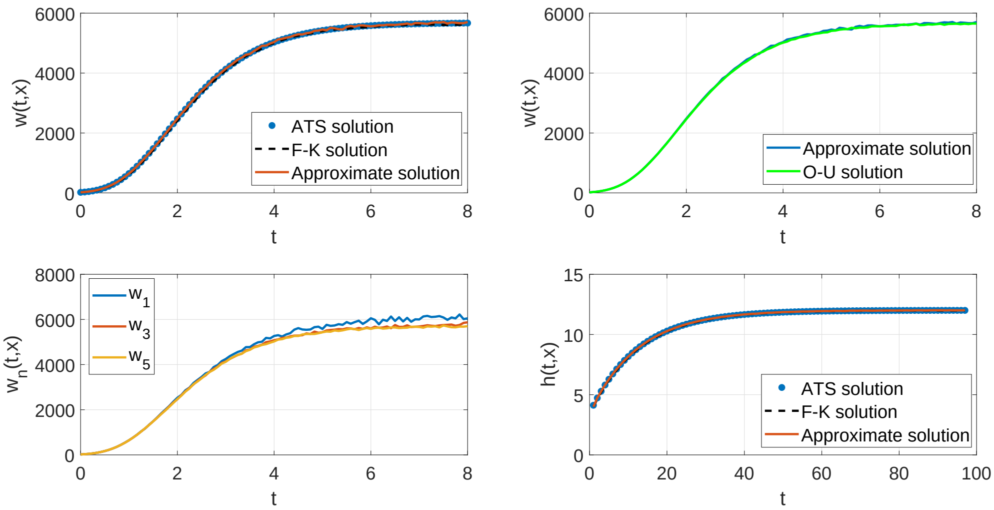


Figure 2.1. **Top line:** approximate solution w_n $(n = 2^5)$ versus the ATS solution *(left plot)* and the OU type solution *(right plot)*. **Bottom line:** approximate solutions for $n = 2^k$, $k = 1, 3, 5$ *(left plot)*; indifference price h of a put European option corresponding to the function w computed, respectively, by the approximate solution w_n $(n = 2^5)$ and the ATS solution *(right plot)*.

Example 2.3.2. (The case $c \neq 0$)
Let $c > 0$ and consider the interval $J = (-\frac{d}{c}, +\infty)$. By Theorem 2.2.15, the solution to the (ACP) (2.11) admits the approximate formula

$$\begin{aligned} u(x,t) &= \lim_{n\to+\infty} u_n(x,t) \underbrace{\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty}}_{n\,times} L(t,n,\{y_j\}_{1\le j\le n},x,g) \prod_{j=1}^{n} p\left(\frac{t}{n}, y_j\right) dy_1 \dots dy_n \\ &= \lim_{n\to+\infty} \mathbb{E}_p[L(t,n,\{Y_j\}_{1\le j\le n},x,g)]. \end{aligned} \tag{2.70}$$

for any $x \in J$, $t \geq 0$ and $g \in C(\bar{J})$. Assuming the initial condition g to be of type (2.65), the function L, defined in (2.58), can be written as

$$L(t,n,\{y_j\}_{1\le j\le n},x,g) =$$
$$\exp\left[r\left(K - e^{\beta(\sum_{i=1}^{n} y_i + \zeta t)}\left(x + \frac{d}{c}\right) + \frac{d}{c} - \frac{\alpha_2 t}{nc} \sum_{j=1}^{n} e^{\left[\beta\left(\sum_{i=j}^{n} y_i + \left(\frac{n-j+1}{n}\right)\zeta t\right)\right]}\right)^+\right].$$

As in Example 2.3.1, Figure 2.2 shows the behaviour of the function w and of the corresponding indifference price h of a put European option, under the following suitably chosen parameter values

$$\begin{array}{cccccc} c = 0.0300, & d = -0.0300, & \gamma = -0.9593, & \delta = 0.3209, & \mu = 0.0380, & \sigma = 0.0300, \\ \rho = 0.8, & \eta = 2, & Y_0 = 8, & K = 12, & T = 8. & \end{array}$$

The function w is computed via the approximate formula (2.70) for $n = 2^k$ with $k = 1, 3, 5, 7, 9$, by applying a Monte Carlo integration method. In particular, when $k = 9$, an accurate goodness-of-fit to the curve obtained by the classical Feynman-Kac formula (2.7) is showed.

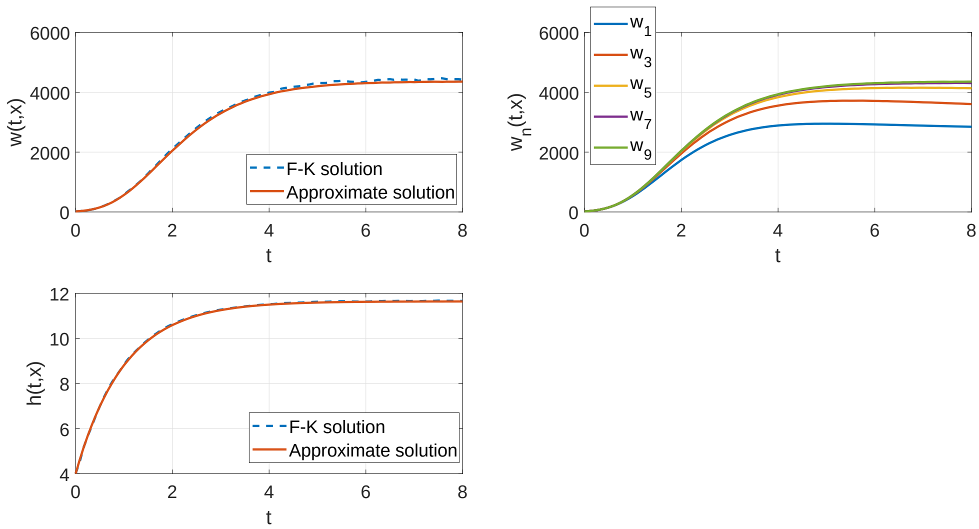


Figure 2.2. Top line: approximate solution w_n $(n = 2^9)$ versus the classical Feynman-Kac solution *(left plot)*; approximate solutions for $n = 2^k,\ k = 1, 3, 5, 7, 9$ *(right plot)*. **Bottom line:** indifference price h of a put European option corresponding to the function w computed, respectively, by the approximate solution w_n $(n = 2^9)$ and the Feynman-Kac solution.

Chapter 3

On the Forecast of Expected Short Interest Rates in the CIR Model

In this chapter, based on [76] and [78], we propose a new methodology to forecast future short-term interest rates from observed financial market data by using the original CIR model (1.27), even if interest rates are negative. This new approach preserves the analytical tractability of the CIR model that has become inadequate to describe the term structure of interest rates for all the reasons explained in the Introduction, especially after the 2008 financial crisis. The performance of the new approach, tested on monthly data, provides a fitting close to market interest rates on different maturities. It is shown how the proposed methodology overcomes both the usual challenges (e.g. simulating regime switching, clustered volatility, skewed tails, etc.) as well as the new ones added by the current market environment characterized by low to negative interest rates.
This chapter is organized as follows. Section 3.1 explains the reasons behind the idea to implement a new methodology. Section 3.2 presents the model in full detail. Finally, Section 3.3 shows the empirical results on observed market data for different maturities.

3.1 Material and Method

3.1.1 Dataset

Our dataset records monthly EUR interest rates (spanning from 31 December 2010 to 29 July 2016) with maturities 1/360A, 30/360A, 60/360A,..., 360/360A and 1Y,..., 50Y (i.e. at 1 day (overnight), 30 days, 60 days,...., 360 days and 1 Year,...,50 Years) available from IBA [1] [52]. For our convenience we have split the dataset in two Datasets: money market (Dataset I) and short- to long-term interest rates (Dataset II). It is well known that the CIR model is adequate to describe the short-term

[1] ICE Benchmark Administration, Data Vendor Codes.

interest rates, but we will show that our procedure can be used also for maturities of Dataset II.

Table 3.1. Monthly EUR interest Rates: the Dataset

	Dataset I				Dataset II			
	Maturity							
Date	1/360A	30/360A	⋯	360/360A	1Y	2Y	⋯	50Y
31.12.2010	0.606	0.788	⋯	1.507	1.311	1.557	⋯	3.306
31.01.2011	1.231	0.895	⋯	1.644	1.582	2.012	⋯	3.482
⋮	⋮	⋮	⋯	⋮	⋮	⋮	⋯	⋮
29.07.2016	-0.397	-0.371	⋯	-0.049	-0.201	-0.215	⋯	0.632

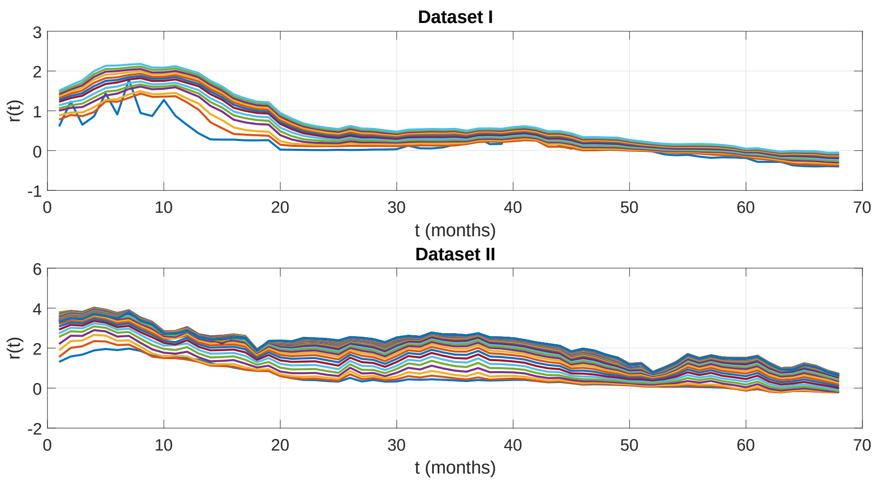


Figure 3.1. Datasets I and II

In Table 3.1, each column lists a sample of $n = 68$ monthly observed EUR interest rates with a set maturity; each row shows interest rates on different maturities observed at a fixed monthly date.
The plots in Figure 3.1 represent the columns of Dataset I and II, so they are different from the yield curves (term structure) by plotting the rows. From Dataset I it is evident that the short-term rates become permanently negative after 2014 (as from March 2015). However, sample data from Dataset II also show a downward trend. In [76] we carried out a qualitative analysis of the dataset and showed tha the most challenging task is to fit short-term interest rates with maturities in Dataset I, due to the largest presence of next-to-zero and/or negative spot rate values. For this reason, we start to examine samples of interest rates with maturity from Dataset II. In this chapter we limit ourselves to estimate and forecast future expected interest rates over a fixed time horizon. In Chapter 4 we will focus on the problem of forecasting exact future interest rates values based on rolling windows of market data.

3.2 The Model

As explained in the Introduction, many models have been proposed for fixing the shortcomings in the classical CIR model. Despite the great progress in this matter, all these models assume the interest rates to be always non-negative, in some cases to the detriment of reasonable restrictions on the parameters and/or analytical tractability.

3.2.1 Procedure and Accuracy

To solve challenges **i.-ii.** mentioned in the Introduction, of the original CIR model, we translate the observed interest rates by a suitable scalar parameter such that the the eventually negative or near-to-zero market rates are shifted to positive values and the diffusion term in (1.27) is not dampened by the proximity to zero but fully reflects the same level of volatility present on the market. Moreover, in order to catch clusters of volatility and jumps in financial time series for short-rates due to a mixture of probability distributions, we partition the observed data sample into sub-samples with a Normal or non-central Chi-square distribution by using an appropriate technique described in Section 3.2.2. This should allow to overcome the critical issue point **iv.** in the Introduction. Finally, to dealing with issue **iii.** we calibrate, for each sub-sample, the parameters of the CIR model to the observed interest rates (Section 3.2.4).
In order to measure the accuracy of our approach, we compute the square-root of the mean square error (RMSE) ε defined as

$$\varepsilon = \sqrt{\frac{1}{n}\sum_{h=1}^{n} e_h^2}, \tag{3.1}$$

where $e_h = r_h - \widehat{r}_h$ denotes the residual between the market interest rate r_h and the corresponding fitted value $\widehat{r}_h$. In our case the fitted values are the expected short-term interest rates estimated through the numerical procedure described in the following subsections, which are compared to market data in Section 3.3.
Finally, the expected values of future next-month interest rates based on fixed size rolling windows are estimated by applying the procedure described in Section 3.3.1.

3.2.2 Step 1 - Dataset Partition

As observed in Section 3.2.1, the novelty in our procedure consists in partioning the observed samples of interest rates into suitable sub-samples to take account of multiple jumps and changes in the volatility. The sub-samples are chosen according to the data empirical probability distribution, which is unknown and clearly different from the Chi-square conditional distribution or the Gamma stationary distribution of the square-root process $(r(t))_{t\geq 0}$.
Notice that in the literature there exist several approaches for detecting multiple changes in the probability distribution of a stochastic process or a time series (see, for instance, M. Lavielle [64], M. Lavielle and G. Teyssiere [65], J. Bai and P. Perron [6]). We adopt the numerical partition herein described into sub-samples following a Normal or a non-central Chi-square distribution.

Partition with Normal Distribution

In [76] we hypothesized the empirical distribution of the observed data sample to be a mixture of normal distributions, with jumps and changes in the standard deviation, for the presence of negative interest rate values. This hypothesis is appropriate because the dynamics of the form (1.27) is obtained from a squared Gaussian model (see, for example, L. C. G. Rogers [84]). Our idea was, therefore, to divide the data sample into a number of sub-samples each coming from an appropriate normal distribution. The goodness-of-fit to a normal distribution was performed by the Lilliefors test (at a 5% significance level) as an improvement on the Kolomogorov-Smirnov test when the population mean and standard deviation are not known, but instead are estimated from the sample data. In this work we have implemented a *forward* procedure that starts by considering the first four data of the original sample, say $(r_1, \ldots, r_4)$, and performs the Lilliefors test until the first normally distributed sub-sample, say $(r_1, \ldots, r_{n_1})$, with $n_1 \geq 4$, is found. Then, the procedure is applied to the remaining sequence $(r_{n_1+1}, \ldots, r_{n_1+4})$ until the second normally distributed sub-sample $(r_{n_1+1}, \ldots, r_{n_2})$, with $n_2 \geq n_1 + 4$, is found, and so on up to partition the entire data sample into m normally distributed sub-samples, namely $(r_1, ..., r_{n_1})$, $(r_{n_1+1}, ..., r_{n_2}), \ldots, (r_{n_{m-1}+1}, ..., r_{n_m})$, $n_m \leq n$.[2] Table 3.2 summarizes the *forward* segmentation procedure.

Table 3.2. "Forward" procedure

```
1. Initialize h=4;
2. run the Lilliefors test on the interest rate vector r(1:h);
3. while the null hypothesis is not rejected
4. h=h+1;
5. run the Lilliefors test on r(1:h);
6. end
7. set n(1)=h;
8. initialize i=1;
9. while n(i)<length(r)
10. h=n(i)+4;
10. repeat steps 2-6 for r(n(i)+1:h) and find n(i+1);
11. if length(r)-n(i+1)<4
12. set resti=r(n(i+1)+1:length(r));
13. break
14. else
15. set i=i+1;
16. end
17. end
```

Further, observe that in performing the Lilliefors test it could happen that the p-value is greater than the chosen significance level (5%), but differs from it by no more than 10^{-2}. Thus, in this case, the Johnson transformation is applied to ensure that each sub-sample follows a normal distribution. The Johnson's method consists

[2]The Matlab's *lillietest* function performs the test only for samples of a size greater than or equal to 4. It is therefore possible that the last three data of the entire sample, namely r_{n-2}, r_{n-1}, r_n, are excluded from partitioning at most.

in transforming a non-normal random variable X to a standard normal variable Z as follows

$$Z = \gamma + \delta f\left(\frac{X-\xi}{\lambda}\right), \quad \lambda,\ \delta > 0 \tag{3.2}$$

where f must be a monotonic function of X with the same range of values of the standardized random variable $(X-\xi)/\lambda$, where ξ and λ are respectively the mean and the standard deviation of X. The parameters δ and γ reflect respectively the skewness and kurtosis of f. The algorithm to estimate the four parameters γ, δ, λ and ξ, and performs the appropriate transformation is available as a Matlab Toolbox written by D. L. Jones [56].
To apply the Johnson's method to our case, the market interest rates in each sub-sample have been transformed by (3.2) to m sub-samples with standard normal distribution that is, for any $k = 1, \ldots, m$,

$$z_h = \gamma + \delta f\left(\frac{r_h - \mu^k}{\sigma^k}\right), \quad h = n_{k-1}+1, .., n_k \ (n_0 = 0),$$

where μ^k, σ^k denote respectively the sample mean and standard deviation of the k-th sub-group, and, finally, to m sub-samples with normal distribution $N(\mu^k, \sigma^k)$ as follows

$$r_h = \sigma^k z_h + \mu^k, \quad h = n_{k-1}+1, .., n_k \ (n_0 = 0).$$

Partition with Non-Central Chi-Square Distribution

As an alternative to the previous hypothesis of a mixture of normal distributions, the empirical distribution of the analysed data sample may be assumed a mixture of non-central Chi-square distributions. This hypothesis is justified from the transition probability density of the original CIR model, as mentioned above. Since the non-central Chi-square distribution admits only positive values, to partition the original sample into sub-samples following a non-central Chi-square distribution, the market observed interest rates have to be first shifted to positive values by (3.4). The partitioning procedure is analogous to that described in Table 3.2, and the Kolmogorov-Smirnov test is performed to test (at a 5% significance level) the goodness-of-fit test of the m sub-samples to a non central Chi-square distribution.

3.2.3 Step 2 - Interest Rates Shift

As mentioned in Section 3.2.1 an important step of the procedure consists in translating market interest rates to positive values to eliminate negative/near-zero values and to not dampen the volatility. A possible transformation could be of affine type, that is

$$r_{shift}(t) = \hat{\mu} + \hat{\sigma}\, r_{real}(t), \quad t \in [0, T], \tag{3.3}$$

where $\hat{\mu}$ and $\hat{\sigma}$ are respectively the sample mean and the sample standard deviation of the process $\{r_{real}(t), t \in [0, T]\}$ observed at n discretized time points of $[0, T]$, for a fixed maturity T. In our opinion the transformation (3.3) does not represent the best choice due to some possible complications (e.g., persistence of negative values, worse fitting, changes in short interest rates dynamics). Therefore, among different options, we have preferred to consider the following transformation

$$r_{shift}(t) = r_{real}(t) + \alpha, \quad t \in [0, T], \tag{3.4}$$

where α is a deterministic positive quantity. This translation leaves unchanged the stochastic dynamics of the interest rates i.e., for any time t, $dr_{shift}(t) = dr_{real}(t)$. There are many values that could be assigned to α, but we believe that the most appropriate choice is the $99th$ percentile of the empirical interest rates probability distribution. If the translation (3.4) is not adequate to move negative interest rates to corresponding positive values, which means further negative values are between the $99th$- and the $100th$-percentile, we can set α equal to the $1st$-percentile of the empirical distribution. In this case (3.4) becomes

$$r_{shift}(t) = r_{real}(t) - \alpha.$$

3.2.4 Step 3 - Calibration

In order to estimate interest rates from the CIR model, the involved parameters k, θ, σ in (1.27) need to be calibrated to the market interest rates. Among many approaches existing in the literature to estimate the parameters of the CIR model (see, for instance, M. Poletti Laurini and L. K. Hotta [81] and references therein), we considered the MATLAB implementation of the maximum likelihood (ML) estimation method for the CIR process proposed by K. Kladìviko [60], and the estimating function approach for ergodic diffusion models introduced in B. M. Bibby et al. [10]). As we will show in Section 3.3, the latter method has turned out to be very useful in obtaining optimal estimators for the parameters of discretely sampled diffusion-type models whose likelihood function is usually not explicitly known. In [10, Example 5.4] the authors constructed an approximately optimal estimating function for the CIR model, from which they derived the following explicit estimators of the three parameters k, θ, σ based on a sample of n observed market spot rates $(r_1, \ldots, r_n)$:

$$\begin{aligned}
\hat{k}_n &= -\ln\left(\frac{(n-1)\sum_{i=2}^n r_i/r_{i-1} - (\sum_{i=2}^n r_i)(\sum_{i=2}^n r_{i-1}^{-1})}{(n-1)^2 - (\sum_{i=2}^n r_{i-1})(\sum_{i=2}^n r_{i-1}^{-1})}\right), \\
\hat{\theta}_n &= \frac{1}{(n-1)}\sum_{i=2}^n r_i + \frac{e^{-\hat{k}_n}}{(n-1)(1-e^{-\hat{k}_n})}(r_n - r_1), \\
\hat{\sigma}_n^2 &= \frac{\sum_{i=2}^n r_{i-1}^{-1}(r_i - r_{i-1}e^{-\hat{k}_n} - \hat{\theta}_n(1-e^{-\hat{k}_n}))^2}{\sum_{i=2}^n r_{i-1}^{-1}((\hat{\theta}_n/2 - r_{i-1})e^{-2\hat{k}_n} - (\hat{\theta}_n - r_{i-1})e^{-\hat{k}_n} + \hat{\theta}_n/2)/\hat{k}_n}.
\end{aligned} \tag{3.5}$$

Remark 3.2.1. These estimators exist provided that the expression for $e^{-\hat{k}_n}$ is strictly positive (the authors observed that this happens with a probability tending to one as $n \to \infty$).

It is worth noting that all the inferential methods for diffusion-type models available in literature are applicable to the CIR model only when the sampled interest rates are nonnegative. Thus, in the presence of negative/near-zero interest rate values, as shown in Figure 3.1, the calibration of the unknown parameters can only be made after shifting spot rates to positive values by using the transformation (3.4).

3.3 Empirical Results

In order to test the performance of the methodology herein proposed, some empirical investigations have been done using two data samples of the dataset reported in Table 3.1. As observed in Section 3.1.1, we started to examine a market data sample from Dataset II. We considered a sample consisting of $n = 68$ monthly observed market interest rates with maturity $T = 30$Y. Figure 3.2 shows (green line) that the observed interest rates are all positive with next-to-zero values in the tail.
In this section we are interested in estimating the expected CIR interest rates by applying the proposed numerical procedure to the entire observed data sample. To this end, for the calibration of the CIR parameters (k, θ, σ) to market data, we have to decide which of the estimation methods introduced in Section 3.2.4 provides the better estimates. Thus, we first shift all the observed rates away from zero. Then we apply the two estimation methods and compare their efficiency by computing the corresponding RMSE (3.1). For convergence to the ML estimates, Kladìviko [60] suggests to use the Ordinary Least Squares (OLS) regression method for initial parameter estimates. These initial values are listed in Table 3.3.

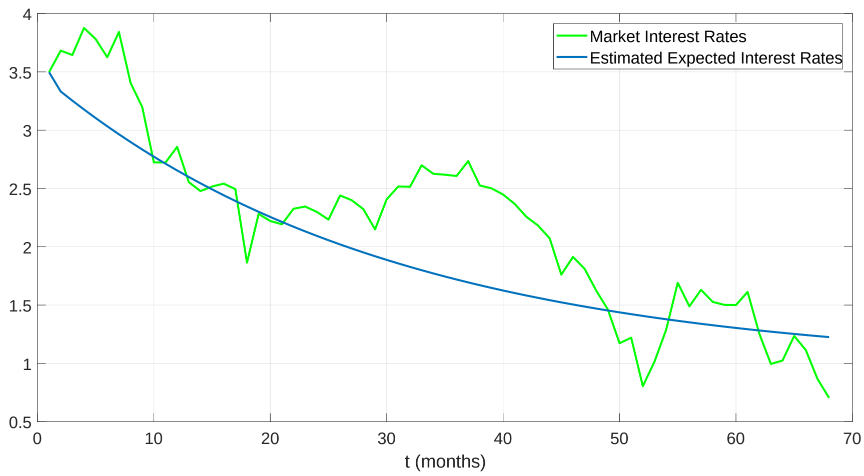


Figure 3.2. Estimated expected interest rates (blue line) versus $n = 68$ monthly observed EUR interest rates (green line) with maturity $T = 30$Y from Dataset II in Table 3.1.

In our case, the optimization process was completed after 74 steps giving the ML estimates

$$(\hat{k}_n, \hat{\theta}_n, \hat{\sigma}_n) = \arg\max \ln L(k,\, \theta,\, \sigma),$$

where $L(\cdot)$ is the log-likelihood function of the CIR process (see [60, formula (7)]) maximized over its parameter space $\Omega = [0, \infty)^3$.
To compare the efficiency of both the estimation methods,Table 3.3 lists the calculated parameter estimates $(\hat{k}_n,\, \hat{\theta}_n,\, \hat{\sigma}_n)$ and the corresponding RMSE ε . A better performance of the estimating function approach, to which a significantly smaller value of RMSE corresponds, is clearly evident. Note also that the ML implementation is more time consuming compared with the estimating function approach, which exactly computes the parameter estimate from equations (3.5).

Table 3.3. Estimates of the parameter vector (k, θ, σ) and the corresponding RMSE ε for a sample of $n = 68$ monthly observed interest rates with maturity $T = 30$Y from the Dataset II in Table 3.1.

	ML estimation		**Estimating Function Method**
	initial value	ML estimate	optimal value
$\hat{k}_n$	0.3300	0.3217	0.0278
$\hat{\theta}_n$	4.5799	4.5201	4.5799
$\hat{\sigma}_n$	0.2832	0.2881	0.0830
ε		1.5453	0.4391

Then the optimal estimate vector computed by (3.5) has been used to calculate the estimated expected interest rates using the following conditional expectation formula availabe in closed form for the square-root process $(r(t))_{t\geq 0}$

$$\mathbb{E}[r(t)|r(s)] = \theta + (r(s) - \theta)e^{-k(t-s)}, \quad 0 \leq s < t. \tag{3.6}$$

The initial value has been set equal to the first value in the observed (shifted) data sample.
Figure 3.2 compares the original (no shifted) market data sample with the corresponding sequence $\widehat{r}_{exp}$ of the estimated expected CIR interest rates shifted back by using the inverse of the transformation (3.4).
To improve the results showed in Figure 3.2 in terms of fitting closely to market data, we implemented a numerical algorithm based on the following main steps summarizing the procedure described in Sections 3.2.2-3.2.4:

1. Partition the whole sample into m Normal/non-central Chi-square distributed sub-groups;
2. Shift each sub-group to positive values by using the translation formula (3.4);
3. Apply the Johnson's transformation if needed (only in the case of normally distributed sub-samples);
4. Calibrate the parameters of the CIR interest rate process r to each sub-group by applying the estimating function method described in Section 3.2.4 and generate a sequence, $\widehat{r}_{exp}$, of estimated expected CIR interest rates, which are shifted back by using the inverse transformation of (3.4).

Again, we considered the above mentioned monthly observed data sample with long-term maturity (T=30Y). From Step 1 we obtained a partition of the sample into $m = 4$ normally distributed sub-groups, namely $(r_1, ..., r_{22})$, $(r_{23}, ..., r_{50})$, $(r_{51}, ..., r_{59})$, $(r_{60}, ..., r_{67})$, and into $m = 7$ sub-groups with non central Chi-square distribution, namely $(r_1, ...r_{11})$, $(r_{12}, ..., r_{21})$, $(r_{22}, ..., r_{28})$, $(r_{29}, ..., r_{37})$, $(r_{38}, ..., r_{46})$, $(r_{47}, ..., r_{56})$, $(r_{57}, ..., r_{65})$. Note that the values r_{68}, in the first case, and (r_{66}, r_{67}, r_{68}), in the second case, were left out the partitioning. We then applied Steps 2-4 to each sub-group for both the partitions. Tables 3.4 and 3.5 list the RMSE computed for each sub-group, namely ε_k, with $k = 1, ..., m$, and the total RMSE $\widetilde{\varepsilon}$ computed over the whole sample as a weighted mean of the ε_k, that is

$$\widetilde{\varepsilon} = \sqrt{\sum_{k=1}^{m} \frac{n_k}{n} \sum_{h=1}^{n_k} e_h^2}. \tag{3.7}$$

Figures 3.3 and 3.4 compare the plots of the estimated expected interest rates with the original (no shifted) market data. It is evident a better fitting to market data when a partitioning into non-central Chi-square distributed sub-samples was considered.

Table 3.4. RMSEs after partitioning (with normal distribution) of a sample of $n = 68$ monthly EUR interest rates with maturity $T = 30$Y from the Dataset II in Table 3.1.

	Normal distribution			
	Subgroups			
	$1st$	$2nd$	$3rd$	$4th$
ε_k	0.4995	0.3821	0.2734	0.2090
$\widetilde{\varepsilon}$		1.3438		

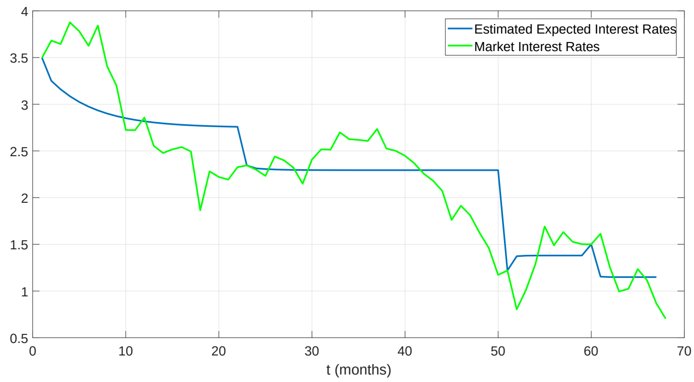


Figure 3.3. Estimated expected interest rates (blue line) versus market rates (green line) after segmentation with the Normal distribution for a data sample of $n = 68$ monthly EUR interest rates with maturity $T = 30$Y from Dataset II in Table 3.1.

Table 3.5. RMSEs after partitioning (with non-central Chi-square distribution) of a sample of $n = 68$ monthly EUR interest rates with maturity $T = 30$Y from Dataset II in Table 3.1.

Non-central Chi-square distribution							
	Subgroups						
	$1st$	$2nd$	$3rd$	$4th$	$5th$	$6th$	$7th$
ε_k	0.5151	0.2008	0.0633	0.0765	0.1514	0.2448	0.1822
$\widetilde{\varepsilon}$				0.7136			

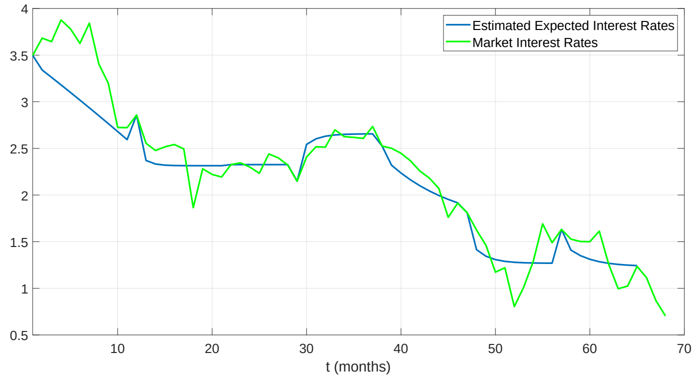


Figure 3.4. Estimated expected interest rates (blue line) versus market rates (green line) after segmentation with the non-central Chi-square distribution for a data sample of $n = 68$ monthly EUR interest rates with maturity $T = 30$Y from Dataset II in Table 3.1.

The second tested data sample consists of $n = 68$ monthly EUR interest rates in a money market, with maturity $T = 30/360$A from Dataset I in Table 3.1. From Step 1 of the above described numerical procedure, the entire sample has been partitioned into $m = 5$ normally distributed sub-groups, namely $(r_1, ..., r_{27})$, $(r_{28}, ..., r_{35})$, $(r_{36}, ..., r_{42})$, $(r_{43}, ..., r_{49})$, $(r_{50}, ..., r_{67})$, and into $m = 10$ non-central Chi-square distributed sub-groups, namely $(r_1, ...r_8)$, $(r_9, ..., r_{18})$, $(r_{19}, ..., r_{24})$, $(r_{25}, ..., r_{30})$, $(r_{31}, ..., r_{36})$, $(r_{37}, ..., r_{42})$, $(r_{43}, ..., r_{48})$, $(r_{49}, ..., r_{54})$, $(r_{55}, ..., r_{60})$, $(r_{61}, ..., r_{66})$. In this case, the values r_{68} and (r_{67}, r_{68}), respectively, were left out the partitioning.
The results listed in Tables 3.6 and 3.7 as well as the plots in Figures 3.5 and 3.6, show a better fitting to the observed market interest rates when a partitioning into non-central Chi-square distributed sub-samples was considered.

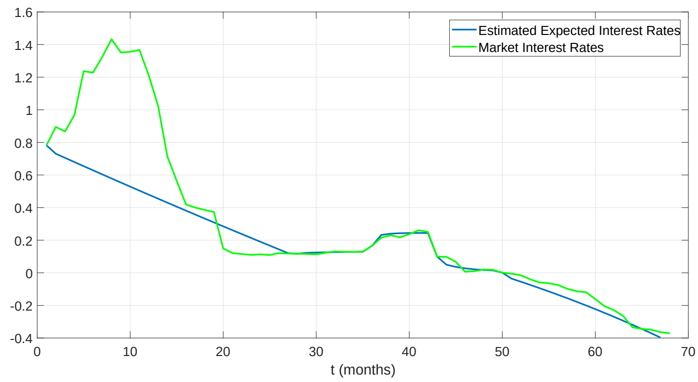


Figure 3.5. Expected simulated interest rates (blue line) versus market rates (green line) after segmentation with Normal Distribution for a data sample of $n = 68$ monthly EUR interest rates with maturity $T = 30/360A$ from Dataset I in Table 3.1.

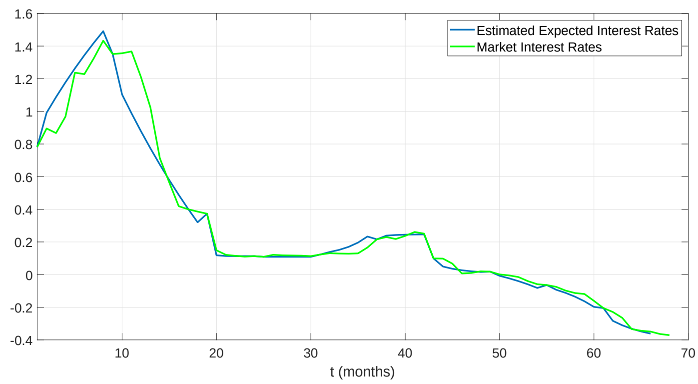


Figure 3.6. Expected simulated interest rates (blue line) versus market rates (green line) after segmentation with non-central Chi-square Distribution for a data sample of $n = 68$ monthly EUR interest rates with maturity $T = 30/360A$ from Dataset I in Table 3.1.

Table 3.6. RMSEs after partitioning (with normal distribution) of a sample of $n = 68$ monthly EUR interest rates with maturity $T = 30/360A$ from Dataset I in Table 3.1.

	Normal distribution				
	Subgroups				
	$1st$	$2nd$	$3rd$	$4th$	$5th$
ε_k	0.4387	0.0050	0.0138	0.0234	0.0440
$\widetilde{\varepsilon}$			1.6470		

Table 3.7. RMSEs after partitioning (with non central Chi-square distribution) of a sample of $n = 68$ monthly EUR interest rates with maturity $T = 30\text{Y}$ from Dataset II in Table 3.1.

	Non-central Chi-square distribution									
	Subgroups									
	$1st$	$2nd$	$3rd$	$4th$	$5th$	$6th$	$7th$	$8th$	$9th$	$10th$
ε^k	0.1267	0.1979	0.0129	0.0076	0.0432	0.0129	0.0252	0.0179	0.0270	0.0294
$\widetilde{\varepsilon}$					0.1323					

3.3.1 Forecasting Expected Interest Rates

In the section we apply the proposed numerical procedure to forecast future expected values of market interest rates.

To explain our idea we refer to the first data sample considered in the previous section ($n = 68$ monthly observed EUR interest rates with maturity $T = 30\text{Y}$). We consider a fixed size window of 12 real interest rates that is rolled through time, each month adding a new rate and taking off the oldest one. The length of this window (12 months) is the historical period over which we forecasted the next-month expected spot rate value. Steps 2-4 of the numerical procedure are applied to the historical market interest rates. Note that Step 1 of the proposed procedure cannot be (always) applied in case of low-frequency data, as in this case of monthly observed data, because the size of the historical data sample is small. Further, the calibration of the CIR parameters with sample data of smaller size than 12 is not always possible when the optimal estimating function method described in Section 3.1.1 is applied (see Remark 3.2.1).

The resulting forecasted next-month expected values computed by the formula (3.6) are plotted in Figure 3.7. Moreover, their values show a better performance with respect to the sequence of expected future rates computed by the Exponentially Weighted Moving Average (EWMA) model [3]. The EWMA is a weighting scheme to estimate future values averaging on historical data with weights that decrease exponentially at a rate λ throughout as the observations are far in the past (the reader can refer, for example, to J. C. Hull [51, Chapter II]). The EWMA has been shown to be powerful for prediction over a short horizon and track closely the

[3] We used the Matlab function *tsmovavg*, parameter "e".

volatility as it changes. Indeed recent interest rates movement is the best predictor of future movement as it is not conditioned on a mean level of volatility. Finally, we applied the proposed numerical procedure and the EWMA model to all data samples (63 maturities) available from Dataset I and II in Table 3.1 to forecast future next-month expected interest rates. Note that we considered a rolling window of 14 real interest rates because the calibration of the CIR model parameters to market data described requires at least a historical period of 14 months for most available market samples. Figure 3.8 compares the corresponding RMSE values computed by our numerical algorithm and by the EWMA model (the vertical black line differentiates the samples in Dataset I from the ones in Dataset II). The error analysis shows that our procedure provides a better fitting of the predicted expected interest rates to real data than the EWMA model.

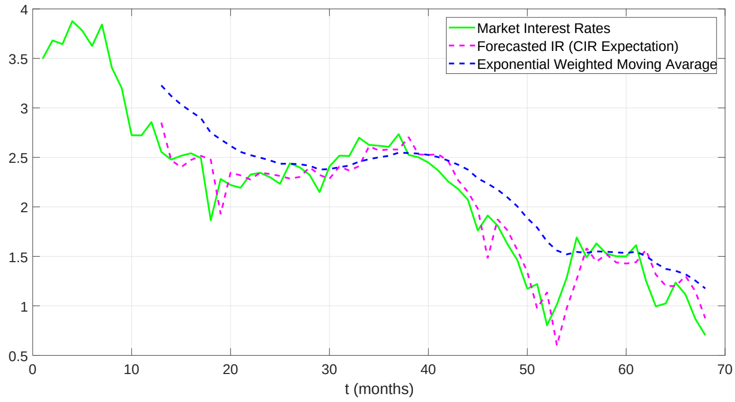


Figure 3.7. Forecast of the expected next-month interest rates based on a rolling window of 12 real data: $n = 68$ monthly EUR interest rates with maturity T=30Y (green line); CIR future expected interest rates (magenta dashed line); EWMA predicted values (blu dotted line).

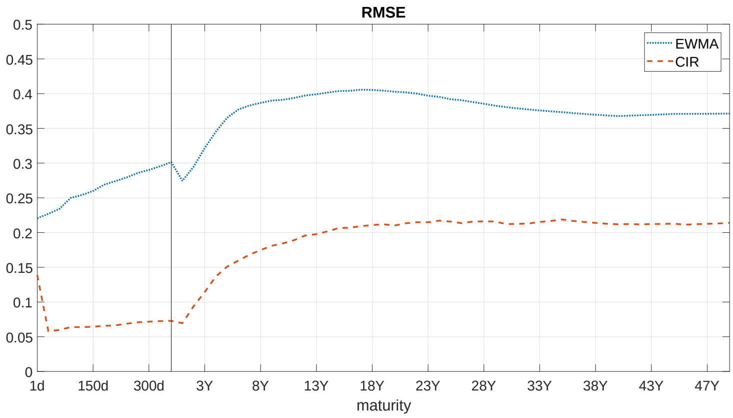


Figure 3.8. Error analysis for future next-month expected interest rates based on a rolling window of 14 real data: RMSE values computed by the proposed numerical procedure (dashed orange line) vesus RMSE values computed by the EWMA model (blue line). The vertical black line differentiates samples in Dataset I from samples in Dataset II.

Remark 3.3.1. In an ongoing research work our aim is to consider sample data of larger size, e.g. weekly observed market interest rates (unavailable at this time). In the case of larger dataset, indeed, the methodology here proposed necessarily requires a rolling window of variable size. Roughly speaking, the partitioning of the available data sample will be useful to determine the size of the historical period (i.e. of the more recent sub-sample) over which calibrate the CIR parameters and forecast future interest rates.

Chapter 4

A Revised Approach to CIR Short-Term Interest Rates Model: the CIR# Model

In this chapter a new accessible methodology, named CIR# model, is described (see [77]). We will show that this model improves the results obtained in Chapter 3. In Section 4.1 the reasons behind our idea to propose a new approach in the CIR framework is explained and a short description of it is provided. Section 4.2 presents the numerical procedure in full details and tests it to market data. Finally Section 4.3 shows the powerful of the model in forecasting future next interest rates values based on rolling windows of market data for EUR and USD currencies.

4.1 The CIR# model

In the following we will illustrate our original approach, but first let us recap the main issues of the CIR model, as explaned in the Introduction:

- **i** Negative interest rates are precluded;
- **ii** The diffusion term in (1.27) goes to zero when $r(t)$ is small (in contrast with market data);
- **iii** The instantaneous volatility σ is constant (in real life σ is calibrated continuously from market data);
- **iv** There are no jumps (e.g. caused by government fiscal and monetary policies, by release of corporate financial results, etc.);
- **v.** There is not a satisfactory calibration at each time to market data since it depends on a small number of constant parameters;

As already explained in Chapter 3, our aim is to give an answer to points **i.-v.**. In particular, our task herein is to improve the results obtained in Chapter 3 by estimating and forecasting exact (not expected) interest rates values by preserving the structure of the original CIR model (1.27). In the remainder of this section

we will summarize the main steps our methodology. The first step consists in partitioning the available market data sample into sub-samples - not necessarily of the same size - by an ANOVA test in order to capture all the statistically significant changes of variance in market spot rates and consequently, to give an account of jumps (see Section 4.2.1). This should allow to overcome the critical issue pointed out in **iv.**. After that, as already explained in Chapter 3, to overcome challenges **i.-ii.**, the observed market rates are properly translated to shift them away from zero or negative values (see Section 4.2.2).
The second step consists in fitting an "optimal" - as explained in Section 4.2.4 - ARIMA model to each sub-sample of market data. To ensure that the residuals of the chosen "optimal" ARIMA model for each sub-sample look like Gaussian white noise, the Johnson's transformation [55] is applied to the standardized residuals (see Section 4.2.3).
As a third step, the parameters k, θ, σ in the CIR model (1.27) are calibrated to the (eventually) shifted market interest rates by estimating them for each sub-sample of available data, as explained in Section 4.2.3 (this allows to overcome the issue **iii.**). For this purpose, trajectories of the CIR process $(r(t))_{t\geq 0}$ are simulated by a strong convergent discretization scheme. The innovation in our procedure is to replace the standard Brownian motion, as a noise source perturbing the interest rates dynamics, with the standardized residuals of the "optimal" ARIMA model selected for each sub-sample. As a result, exact CIR fitted values to market data are calculated and the computational cost of the numerical procedure is considerably reduced. It is worth noting that to determine each "optimal" ARIMA model further parameters needed to be estimated in addition to (k, θ, σ). That allows to solve issue **v.**. Finally, the estimated interest rates are shifted back and compared to market data. As a measure of goodness-of-fit to the available market data, we compute:

- the statistics R^2 given by the following expression (see [62])

$$R^2 = 1 - \frac{\sum_{h=1}^{m} (e_h - \overline{e})^2}{\sum_{h=1}^{m} (r_h - \overline{r})^2}, \tag{4.1}$$

  where $e_h = r_h - \widehat{r}_h$ denotes the residual between the observed market interest rate r_h and the corresponding fitted value $\widehat{r}_h$, evaluated on a data sample of size $m \geq 2$. Furthermore, $\overline{e}$ and $\overline{r}$ denote the sample mean of e_h and r_h, respectively;

- the square root of the mean square error (RMSE)

$$\varepsilon = \sqrt{\frac{1}{m}\sum_{h=1}^{m} e_h^2}. \tag{4.2}$$

The datasets tested in this chapter records EUR and USD interest rates for a total number of 63 maturities (see Section 3.1.1 for more details). In particular Table 4.1 and Figure 4.1 represent the columns of Dataset I and Dataset II in the case of USD currency.

Table 4.1. Monthly USD interest Rates: the Dataset

	Dataset I				Dataset II			
			Maturity					
Date	1/360A	30/360A	$\cdots$	360/360A	1Y	2Y	$\cdots$	50Y
31.12.2010	0.251	0.260	$\cdots$	0.780	0.441	0.784	$\cdots$	4.059
31.01.2011	0.235	0.260	$\cdots$	0.781	0.430	0.784	$\cdots$	4.259
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
29.07.2016	0.412	0.495	$\cdots$	1.432	0.813	0.870	$\cdots$	1.724

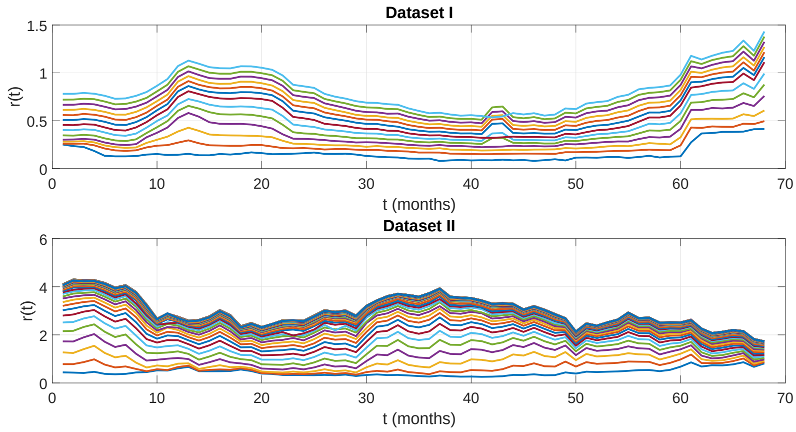


Figure 4.1. Dataset I and II (USD).

4.2 Numerical Implementation and Empirical Analysis

4.2.1 ANOVA test with a fixed segmentation

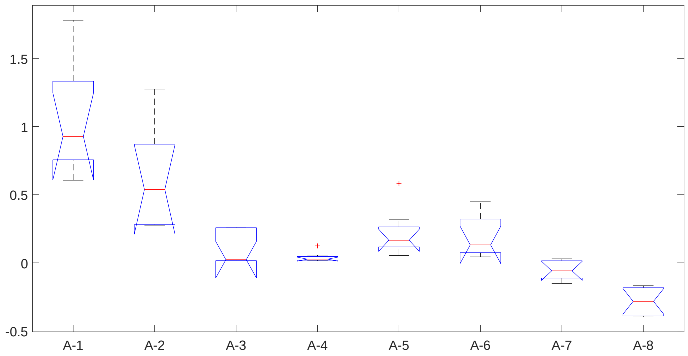
As mentioned in Section 4.1, the first step of the CIR# model consits in partitioning the observed market data sample into sub-samples, which we call *groups*, by a one-way ANOVA analysis to highlight statistically significant changes of variance in market rates and so to give an account of possible jumps. The main difficulty concerns the choice of the optimal partition into groups to apply the ANOVA test; we had to take into account both the size (the smaller the group is, the more refined is the analysis) and the ability to capture any jumps (the larger the group, the better in terms of statistical significance).

As an example, we consider the data sample consisted of n=68 EUR interest rates with 1 day (overnight) maturity from Dataset I in Table 3.1. After several tests, we decided to segment the whole sample into eight groups each of size $m = 8$ or a multiple thereof (except for the last group, obviously). The results of the one-way ANOVA test are reported in Table 4.2. The *p-value* (Prob>F) of $8.00796 \cdot 10^{-19}$ indicates a statistically significant difference between groups.

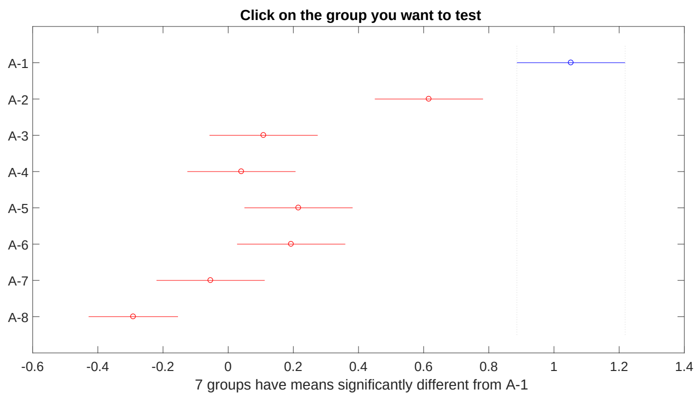
Table 4.2. The ANOVA Table shows the between-groups (Groups) and the within-groups (Error) variation. "SS" is the sum of squares and "df" means degrees of freedom associated to SS. MS indicates the mean squared error, i.e. the estimate of the error variance. The value of the F-statistic is given by the ratio of the mean squared errors.

Source	SS	df	MS	F	Prob>F
Groups	10.8783	7	1.55405	34.71	8.00796e-19
Error	2.6862	60	0.04477		
Total	13.5645	67			

Furthermore, the boxplot (Figure 4.2.a)) and a multiple comparison test performed on the eight groups (Figure 4.2.b)) have suggested partitioning the data sample into the following four groups of observations 1–8, 9–16, 17–56, 57–68.



a)



b)

Figure 4.2. a). Boxplot; **b).** Multiple comparison test.

4.2.2 Jumps fitting by translation

As observed in the Introduction and in Chapter 3, the CIR model (1.27) does not fit negative interest rates and normal/high volatility when the rate value is small.

Thus, the translation formula (3.4) is applied in the presence of near zero/negative interest rate values in the observed market data.
The translation is applied after carrying out a check on each group partitioning the original data sample. The check consists in calculating the harmonic mean, that is more robust than the arithmetic one in the presence of extreme values, and in verifying whether it is smaller than a constant value chosen arbitrarily small (e.g. 10^{-2}). If this happens in at least one group, then the whole sample is translated.

4.2.3 Sub-optimal ARIMA models

The second step of our procedure consists in deriving the best fitting ARIMA (p, i, q) model to each group of interest rates partitioning the observed market data sample. Thus we start by selecting, for each group, a set of ARIMA (p, i, q) models satisfying the following "sub-optimal" conditions:

1. Absence of both autocorrelation (AC) and partial autocorrelation (PAC) in the time series[1];
2. Absence of unit roots (stationarity of the time series);
3. Normally distributed standardized residuals;
4. $R^2_{ARIMA} > 0.5$,

where R^2_{ARIMA} denotes the statistics R^2, defined in (4.1), computed for the ARIMA (p, i, q) model. We look for only the indices $i \in \{0, 1, 2\}$ and $p, q \in \{1, 2, 3\}$.
As mentioned, to ensure that the residuals of the selected ARIMA (p, i, q) models look like a Gaussian white noise, the Johnson's transformation [55] described in Section 3.2.2, is applied to the standardized residuals. The normally distributed standardized ARIMA residuals will be used in the sequel to calibrate the parameters (k, θ, σ) in the CIR model to the (eventually shifted) market interest rates.

Calibration of CIR parameters

Consider the jth-group partitioning the available market data sample, which we assume to be of length n_j. The calibration of the CIR parameters in the group is performed as follows

1. The volatility σ is estimated by the group standard deviation, namely $\hat{\sigma}_j$;
2. The long-run mean parameter θ is estimated by the group mean, namely $\hat{\theta}_j$;
3. The speed of mean reversion k is estimated by that value, say $\hat{k}_j$, solving the following minimization problem:

$$\min_{k>0} S_j(k) = \min_{k>0} \sqrt{\frac{\sum_{h=1}^{n_j} (u^j_h(k) - \overline{u}^j(k))^2}{n_j - 1}}. \tag{4.3}$$

[1]If this condition is not verified, we can require just the absence of autocorrelation.

For any $k > 0$, we define

$$u_h^j(k) = r_h^j(k) - r_h^j, \quad h = 1, \cdots, n_j, \tag{4.4}$$

being r_h^j the market (shifted) interest rate value, $r_h^j(k)$ the corresponding simulated CIR interest rate value expressed as a function of the unknown parameter k, and $\overline{u}^j(k)$ denotes the sample mean of $u_h^j(k)$. The $r_h^j(k)$ values are calculated by applying the strong convergent Milstein discretization ([69]) to the SDE (1.27). D. Brigo and F. Mercurio in [17, Section 22.7] showed that the Milstein scheme converges in a much better way than other numerical schemes for the CIR process. It reads as

$$r_{h+1}^j(k) = r_h^j(k) + k(\hat{\theta}_j - r_h^j)\,\Delta + \hat{\sigma}_j\sqrt{r_h^j\Delta}\; Z_{h+1}^j + \frac{(\hat{\sigma}_j)^2}{4}\,[(\sqrt{\Delta}\; Z_{h+1}^j)^2 - \Delta], \tag{4.5}$$

where Δ is the time step and Z_{h+1}^j are the normally distributed standardized residuals of each ARIMA (p, i, q) model satisfying the "sub-optimal" conditions 1.-4. for the jth-group. Thus to simulate trajectories of the CIR process, the random terms in the simulation scheme (4.5) are generated by the standardized residuals Z_{h+1}^j instead of a standard Brownian motion, as is usual. After calculation of the estimates $(\hat{k}_j,\, \hat{\theta}_j,\, \hat{\sigma}_j)$, the CIR fitted values to the observed marked rates in the jth-group are computed by the simulation scheme (4.5) as follows

$$\hat{r}_{h+1}^j = \hat{r}_h^j + \hat{k}_j(\hat{\theta}_j - \hat{r}_h^j)\,\Delta + \hat{\sigma}_j\sqrt{\hat{r}_h^j\Delta}\, Z_{h+1}^j + \frac{(\hat{\sigma}_j)^2}{4}\,[(\sqrt{\Delta}\, Z_{h+1}^j)^2 - \Delta], \tag{4.6}$$

where Δ and Z_{h+1}^j are as before. We would emphasize that the partitioning of the available data sample in groups affects the estimation of the parameters (k, θ, σ), which are so locally calibrated to each group.
To measure the goodness-of-fit, the statistics R^2 is computed. For sake of clarity, in the sequel we will denote by R^2_{CIR} the statistics (4.1) when referring to the CIR model.

4.2.4 Optimal ARIMA-CIR model

For each group j, the "optimal" ARIMA(p, i, q) model providing the best CIR fitting to market data will be chosen among the selected sub-optimal ARIMA (p, i, q) models, as described in Section 4.2.3, satisfying the following additional conditions:

5. The ARIMA (p, i, q) minimizes the Bayesian Information Criterion (BIC) matrix whose rows and columns are, respectively, the corresponding p and q lags (*BIC condition*);

6. $R^2_{CIR} > 0.5$.

Therefore we define the following sets of candidate ARIMA models:

$$\mathcal{I}_{AC} = \{(p, i, q) \,|\, \text{ARIMA}(p, i, q) \text{ satisfies conditions 1.– 4. and 6.}\}$$

and

$$\mathcal{I}_{ACB} = \{(p, i, q)|\, \text{ARIMA}(p, i, q) \text{ satisfies conditions 1.– 6.}\}\,.$$

Obviously, $\mathcal{I}_{ACB} \subset \mathcal{I}_{AC}$.
Last but not least, the “optimal" ARIMA model is chosen in the above defined classes as the model minimizing the RMSE ε_j

$$\min_{\hat{r}^j} \varepsilon_j = \min_{\hat{r}^j} \sqrt{\frac{1}{n_j} \sum_{h=1}^{n_j} (r_h^j - \hat{r}_h^j)^2}, \tag{4.7}$$

where the minimum is computed with respect to all the CIR fitted values vectors, $\hat{r}^j$, simulated for the jth-group by (4.6).
Table 4.3 summarizes the main steps of the algorithm to select the optimal ARIMA (p, i, q) model for each group j, and returns as output: the matrix of indices (p, i, q) belonging to the sets $\mathcal{I}_{AC}$ and $\mathcal{I}_{ACB}$, the corresponding CIR fitted values vector $\hat{r}^j$ computed by (4.6), and the associated values of the statistics $(R^2_{CIR})_j$ and ε_j.

Table 4.3. ARIMA-CIR algorithm

Step 1: verify if $check1 = 1$ for the jth-group;
Step 2: if $check1 = 1$ verify if $check2 = 1$ for the current group;
Step 3: if $check2 = 1$ print the output. Else, reduce the size n^j of the current group to $(n^j - m)$ where $m = 8$.
Step 4: repeat **Step 1-Step 3** for the remaining observations in the current group.
Step 5: return to **Step 1** for the group $j + 1$.

Note that $check1$ and $check2$ refer to conditions 1.– 5. and 6., respectively. Their value is equal to 1 if those conditions are satisfied.
It is worth noting that to compute the "optimal" ARIMA model for each group, the parameters (p, i, q), in addition to those of the CIR model, need to be calibrated to the observed market data. This, in some way, addresses the issue **v.**. We applied the ARIMA-CIR algorithm to the $n = 68$ monthly observed EUR interest rates with 1 day (overnight) maturity, considered in Section 4.2.2. We recall that the ANOVA analysis suggested to partition the data sample into four groups of observations: 1–8, 9–16, 17–56, 57–68. Table 4.4 shows in detail the outputs for this sample. The group containing the observations 17–56 has been futher segmented into three sub-groups of size $m = 8$ or a multiple thereof: 17–32, 33–48 and 49–56. The triplets (p, i, q) identified by a rectangle in Table 4.4, indicates the "optimal" ARIMA model chosen for each group/sub-group (with the smaller ε_j value). As it can be seen, none of these models fulfils the *BIC condition.*
Figure 4.3 reports the qualitative statistical analysis referring to the ARIMA $(1, 1, 2)$ chosen as the "optimal" fitting model for the second group of observations 9-16 (similar plots for the other group/sub-groups are reported in Appendix 6.5).
Figure 4.4 compares the short-term interest rates structure of the analysed market data sample with the corresponding curve of CIR fitted values computed by the simulation scheme (4.6) for each group partitioning the whole data sample.
The market interest rates have been shifted by using a translation of type (3.4). Finally, the CIR fitted values have been shifted back.
The total values of the statistics R^2_{CIR} and the RMSE ε have been computed on the whole sample as a weighted mean of the $(R^2_{CIR})_j$ and ε_j values corresponding to the

Table 4.4. Outputs from the ARIMA-CIR algorithm for the 68 monthly EUR interest rates on overnight maturity

j	group/sub-groups	ARIMA model	$(R^2_{CIR})_j$	ε_j	BIC cond.
1	1–8	$\boxed{(2,0,1)}$	0.8166	$\boxed{0.1643}$ [1]	
		(2,0,2)	0.5930	0.2414	
		(3,0,2)	0.780	0.1865	
		(1,1,1)	0.8166	0.1643	
		(1,2,1)	0.7309	0.2026	
		(1,2,2)	0.7805	0.1845	√
		(3,2,1)	0.7023	0.2104	
2	9–16	(1,0,1)	0.6842	0.2090	
		(2,0,2)	0.7799	0.2588	
		(3,0,2)	0.6418	0.2661	√
		(1,1,1)	0.7378	0.2043	
		$\boxed{(1,1,2)}$	0.8472	$\boxed{0.1554}$	
		(2,1,1)	0.6842	0.2169	
		(2,1,2)	0.7799	0.2012	
		(3,1,2)	0.6418	0.2333	
3	17–32	$\boxed{(1,0,3)}$	0.9174	$\boxed{0.0326}$	
		(3,0,1)	0.5485	0.0646	
4	33–48	$\boxed{(3,0,1)}$	0.6901	$\boxed{0.0833}$	
		(3,1,2)	0.6332	0.1146	
		(3,2,1)	0.6332	0.1146	
5	49–56	(1,0,1)	0.5076	0.0597	
		(1,0,2)	0.6030	0.0526	
		(2,0,1)	0.5702	0.0577	
		(3,0,1)	0.7648	0.0483	
		(3,0,2)	0.6240	0.0537	
		(1,1,1)	0.5702	0.0577	
		(2,1,1)	0.5393	0.0588	
		$\boxed{(3,1,2)}$	0.7715	$\boxed{0.0479}$	
		(1,2,1)	0.5542	0.0582	
		(3,2,1)	0.6987	0.0479	
		(3,2,2)	0.6343	0.0536	√
6	57–68	(1,0,2)	0.5964	0.0752	
		(1,0,3)	0.8080	0.0570	
		(2,0,2)	0.8136	0.0559	
		(3,0,2)	0.7899	0.0577	
		(3,0,3)	0.6560	0.0704	
		(1,1,2)	0.8136	0.0559	
		(1,1,3)	0.8006	0.0614	
		$\boxed{(2,1,1)}$	0.8239	$\boxed{0.0486}$	
		(2,1,2)	0.8542	0.0525	
		(2,1,3)	0.8936	0.0551	
		(3,1,2)	0.9023	0.0511	
		(3,1,3)	0.8000	0.0580	

[1] For a more exact comparison we use the numeric format long.

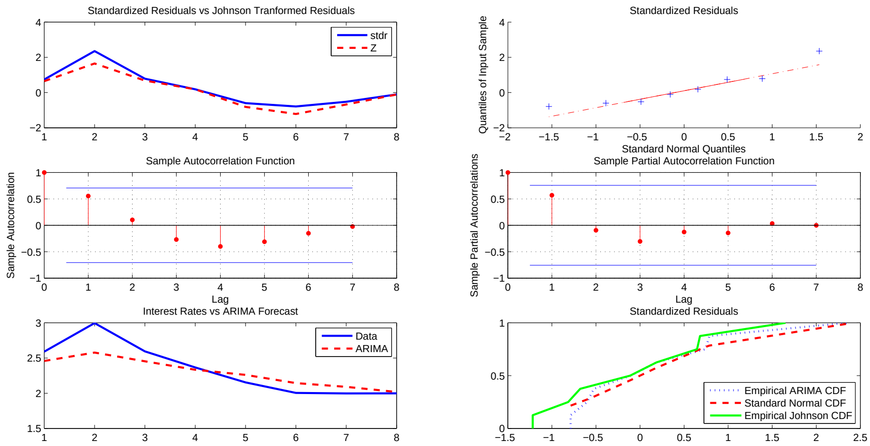


Figure 4.3. Qualitative statistical analysis related to the group 9–16. **Top line:** ARIMA $(1, 1, 2)$ standardized residuals versus Johnson's transformed residuals (*left*); Q-Q normal plot for the ARIMA $(1, 1, 2)$ standardized residuals (*right*). **Middle line:** AC plot (*left*) and PAC plot (*right*). **Bottom line:** real interest rates versus ARIMA $(1, 1, 2)$ fitted values (*left*); comparison among the cumulative distribution function (CDF) of the standard normal distribution, the empirical CDF of ARIMA $(1, 1, 2)$ standardized residuals and the empirical CDF of the residuals after the Johnson's transformation (*right*).

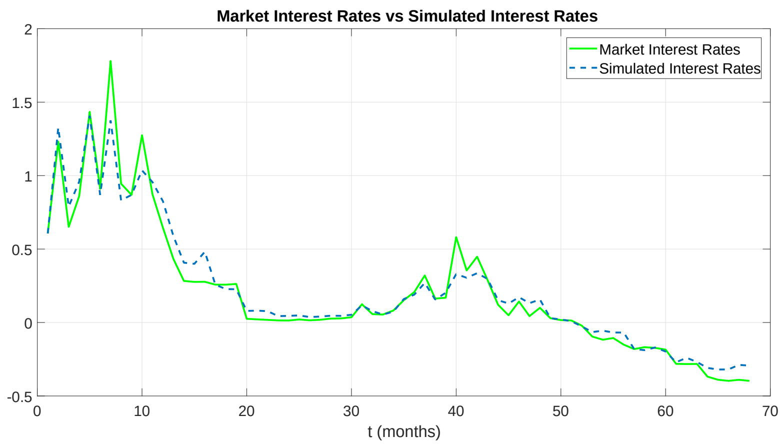


Figure 4.4. Monthly EUR interest rates with T=1 day (overnight) maturity versus CIR fitted rates

"optimal" ARIMA model chosen for each group/sub-group, i.e.

$$\widetilde{R}^2_{CIR} = \sum_{j=1}^{J} \frac{n_j}{n} (R^2_{CIR})_j,$$

$$\varepsilon = \sqrt{\sum_{j=1}^{J} \frac{n_j}{n} \sum_{h=1}^{n_j} (r_h^j - \hat{r}_h^j)^2}. \tag{4.8}$$

Their values for the analysed data sample are: $R^2_{CIR} = 0.8101$, $\varepsilon = 0.2922$. The Appendix 6.6 reports the CIR parameters estimates and the plots of the function $S_j(k)$, defined in (4.3), for all groups/sub-groups.
This strategy allows us to get an exact trajectory of CIR fitted values instead of a curve averaged over 100000 simulated trajectories. Consequently, the computational cost is considerably reduced.

4.2.5 The change points detection problem

As explained in Section 4.2.1, the main difficulty in partitioning the observed data sample concerns the choice of the optimal segmentation to detect abrupt changes in the variance of the interest rates dynamics. In the literature there exist several approaches for detecting multiple changes in the probability distribution of a stochastic process or a time series such as sequential analysis (i.e., "online" methods), clustering based on maximum likelihood estimation (i.e. "offline" methods), minimax change detection, etc. (see, for example, [6], [64], [65], [48], [2] and [3]).

Table 4.5. Numerical scheme for change points detection by Lavielle's algorithm

- compute $v(1 : end)$ the array of change points detected in the real data array x by the Lavielle method
- set $l = v(1)$;
- initialize $xstart = 1$, $xend = l$;
  ($xstart$, $xend$ denote the first and last component of a partitioning group at each processing cycle)
- initialize $j = 1$;
- set $smax = xstart + 1$ (each group must have a minimum length equal to 2);[1]
- **while** $l < v(end)$ & $l \neq smax - 1$
- compute $check\ 1$ and the matrix L
(L is the matrix of possible ARIMA (p, i, q) for $x(xstart : xend)$);
- **if** $check1 = 1$
- compute $check2$;
- **if** $check2 = 1$
- compute ε_j and $(R^2_{CIR})_j$;
- let $ex(j) = l$;
  (ex is the array of the rescaled change points, see Figure 4.5);
- set $xstart = ex(j) + 1$;
- set $j = j + 1$;
- set $l = v(j)$;
- set $xend = l$;
- **else**
- set $l = l - 1$;
- set $xend = l$;
- **end**
- **end**
- **if** $l = smax$
- **if** $length(v) \leq j + 1$
- set $l = v(j + 1)$;
- **else** $break$;
- **end**
- **end**
- **end**

[1] Some statistical tests (involved in $check1$) require a minimum sample length equal to 6, so one can also set $smax = xstart + 5$.

We decided to implement the Matlab algorithm proposed by Lavielle in [64] for the detection of changes in the variance, which allows to partition the data sample analysed in the previous section into the following six groups of observations: 1–13, 14–19, 20–30, 31–39, 40–52, 53–68. However, taking into account that our CIR# model is based on the combination of an ARIMA model and the original CIR model, the above detected change points reported in Figure 4.5, namely 13, 19, 34, 39, 52, have to be adjusted according to the numerical scheme described in Table 4.5.
Table 4.6 lists the results from the ARIMA-CIR algorithm after application of the change point detection algorithm. Figure 4.6 plots the market interest rates versus the CIR# fitted values according to results in Table 4.6. We found that the total $R^2_{CIR} = 0.7584$ and the total RMSE $\varepsilon = 0.4159$. As before, for all j, the errors ε_j are at most of the order of 10^{-1}.

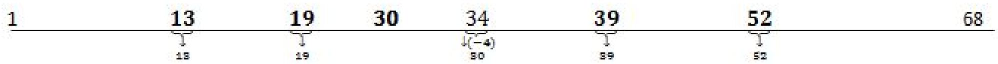


Figure 4.5. Scheme for change points detection from the algorithm in Table 4.5

Table 4.6. Outputs from the ARIMA-CIR algorithm applied to $n = 68$ monthly EUR interest rates with T=1 day (overnight) maturity (after application of the Lavielle method)

j	Groups	ARIMA model	$(R^2_{CIR})_j$	ε_j	BIC cond.
1	1–13	$\boxed{(2,1,1)}$	0.6223	$\boxed{0.2251}$	
		(3,1,1)	0.6117	0.2258	
		(3,1,2)	0.5985	0.2299	
2	14–19	(1,0,1)	0.8842	0.0048	
		(2,0,1)	0.7960	0.0062	
		(2,0,2)	0.8814	0.0047	
		$\boxed{(2,0,3)}$	0.8795	$\boxed{0.0047}$	
		(1,1,1)	0.8291	0.0058	
		(1,1,2)	0.8821	0.0048	
		(1,2,1)	0.7623	0.0068	
		(2,1,1)	0.8421	0.0055	
3	20–30	(3,1,2)	0.6345	0.0087	$\surd$
		(3,1,3)	0.6193	0.0089	
		(1,2,3)	0.6452	0.0099	
		(2,2,2)	0.5178	0.0096	$\surd$
		(2,2,3)	0.5777	0.0089	
		$\boxed{(3,2,2)}$	0.6369	$\boxed{0.0085}$	

j	Groups	ARIMA model	$(R^2_{CIR})_j$	ε_j	BIC cond.
4	31–39	(1,0,1)	0.7165	0.0444	
		(1,0,2)	0.6895	0.0465	
		(1,0,3)	0.6815	0.0469	
		(1,1,1)	0.7168	0.0439	
		(1,1,2)	0.7247	0.0415	
		(1,1,3)	0.5409	0.0546	
		(2,1,1)	0.6876	0.0458	
		$\boxed{(2,1,2)}$	0.7478	$\boxed{0.0404}$	
		(2,1,3)	0.6778	0.0449	
		(1,2,3)	0.5911	0.0507	
		(2,2,3)	0.5817	0.0509	
		(3,2,2)	0.6379	0.0473	
5	40–52	$\boxed{(1,0,2)}$	0.8841	$\boxed{0.1050}$	
		(2,0,1)	0.8868	0.1196	
		(2,0,2)	0.7985	0.1317	√
		(2,0,3)	0.7979	0.1315	
		(3,0,1)	0.8944	0.1178	
		(3,0,2)	0.8186	0.1230	
		(3,0,3)	0.8786	0.1158	
		(1,1,1)	0.6359	0.1653	
		(1,1,2)	0.6004	0.1752	√
		(1,1,3)	0.6436	0.1595	
		(2,1,1)	0.8459	0.1287	
		(2,1,2)	0.5547	0.1783	
		(2,2,1)	0.6694	0.1637	
		(3,2,3)	0.7306	0.1630	
6	53–68	$\boxed{(1,0,3)}$	0.8111	$\boxed{0.0683}$	

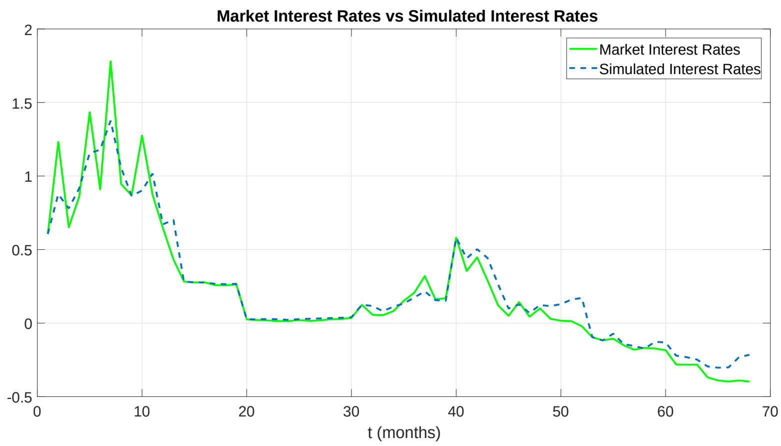


Figure 4.6. Monthly EUR interest rates with T=1 day (overnight) maturity versus CIR# fitted rates (after application of the Lavielle method)

4.3 Forecast of future interest rates

In this section we will address the CIR# model's progress on future interest rate forecasts from a window of observed market data. It is worth noting that we decided to impose the most challenging conditions by modelling the shortest part of the yield curve (e.g. the overnight rate) and using only a handful of number of observations. For instance, with monthly data we have found that $m = 8$ observations are sufficient for a good calibration. Thus we start to consider a fixed size window of 8 real interest rates that is rolled through time, each month adding the new rate and taking off the oldest rate. The length of this window (8 months) is the historical period over which we forecast the next-month spot rate value. The numerical procedure described in Section 4.2 has been applied to forecast future monthly interest rates. We would like to focus for a moment on some difficulties experienced in implementing the CIR# model to forecast future interest rate values:

1. In the case of low-frequency rates and with fixed "a priori" rolling windows of small length, as in our case, the partition of the data sample may not (always) be performed. When higher-frequency data are available (e.g. weakly interest rates), rolling windows of variable size can be determined by a segmentation of the historical data sample (Section 4.2.1) and a change point detection (Section 4.2.5). That we will be our task in an ongoing research.

2. It is better to calibrate the long-run mean parameter θ to the historical data as an exponential moving average (EMA). Indeed, the EMA places a greater weight and significance on the most recent interest rate values. Thus, it reacts more significantly to recent interest rate changes than the sample mean, which applies an equal weight to all observations in the historical period.

3. To forecast future interest rates from a window of historical data of length m, say $w = \{r_{h+1}, r_{h+2}, ..., r_{h+m}\}$, $h \geq 0$, we have first to calibrate the CIR parameters (k, θ, σ) and the parameters (p, i, q) corresponding to the "optimal" ARIMA model to the sample w. Denote the parameter estimates by $(\widehat{k}_w, \widehat{\theta}_w, \widehat{\sigma}_w, \widehat{p}_w, \widehat{i}_w, \widehat{q}_w)$. Then, a future interest rate value is predicted by the simulation scheme (4.6) as follows:

$$\begin{aligned}\widehat{r}_{h+m+s} &= r_{h+m} + \widehat{k}_w(\widehat{\theta}_w - r_{h+m})\Delta \\ &+ \widehat{\sigma}_w\sqrt{r_{h+m}\Delta}\widehat{Z}_{h+m+s} + \frac{(\widehat{\sigma}_w)^2}{4}[(\sqrt{\Delta}\widehat{Z}_{h+m+s})^2 - \Delta], \quad s \geq 1, \end{aligned} \tag{4.9}$$

   where $\widehat{Z}_{h+m+s}$ is an estimate of the unknown standardized residual, Z_{h+m+s}, corresponding to the increment of time s in the interest rate dynamics. The estimate of Z_{h+m+s} is computed as follows:

   i. calculate the difference

$$d = \widehat{r}_{h+m+s} - \widehat{r}^{ARIMA}_{h+m+s}, \tag{4.10}$$

   where $\widehat{r}^{ARIMA}_{h+m+s}$ and $\widehat{r}_{h+m+s}$ are estimates of the unknown interest rate r_{h+m+s} predicted, respectively, by the ARIMA $(\widehat{p}_w, \widehat{i}_w, \widehat{q}_w)$ and the expectation closed formula (3.6) for the CIR model, i.e.

$$\widehat{r}_{h+m+s} = \mathbb{E}[r_{h+m+s}|r_{h+m}] = \widehat{\theta}_w + (r_{h+m} - \widehat{\theta}_w)e^{-\widehat{k}_w s}, \quad s \geq 1;$$

ii. the difference d is standardized with respect to the sample mean and the sample standard deviation of the residual array $\{r_j - r_j^{ARIMA} | j = h+1, ..., h+m\} \cup \{d\}$.
Finally, the Johnson transformation is applied to obtain the standard normally distributed residuals

$$Z = \{Z_j | j = h+1, ..., h+m\} \cup \{\widehat{Z}_{h+m+s}\}.$$

The algorithm to forecast a future interest rate value is summarized in Table 4.7. The predicted yield curve for monthly EUR and USD interst rates with overnight maturity is shown in Figure 4.7 and 4.8, respectively, and compared with the corresponding market observed data.
It is evident that the predicted next-month spot rates computed by the CIR# model follow the market trend. Moreover, the values of R^2 and RMSE ε, computed to measure the goodness-of-fit of forecast interest rates to real data, are respectively 0.8741 and 0.1120 for the EUR currency and 0.9216 and 0.0263 for the USD currency.

Table 4.7. CIR# forecast algorithm

-Denote by $(\widehat{k}_w, \widehat{\theta}_w, \widehat{\sigma}_w, \widehat{p}_w, \widehat{i}_w, \widehat{q}_w)$ the parameters' model calibrated to the rolling window $w = [r_{h+1}, ..., r_{h+m}]$;
-denote by res the array of the ARIMA $(\widehat{p}_w, \widehat{i}_w, \widehat{q}_w)$ residuals;
-fix $s \geq 1$ and compute $\widehat{r}_{h+m+s}^{ARIMA}$;
-compute the expected future interest value $\widehat{r}_{h+m+s}$ by the CIR model;
-set $d = \widehat{r}_{h+m+s} - \widehat{r}_{h+m+s}^{ARIMA}$;
-set $res\_new = [res, d]$;
-compute the sample mean $\widehat{\mu}_{res}$ and the sample standard deviation $\widehat{\sigma}_{res}$ of the sample $res\_new = [res, d]$;
-apply the Johnson transformation to the standardized array: $\frac{res\_new - \widehat{\mu}_{res}}{\widehat{\sigma}_{res}}$;
-denote by Z the array of the standard normally distributed residuals $\{Z_j | j = h+1, ..., h+m\} \cup \{\widehat{Z}_{h+m+s}\}$;
-forecast the future interest rate by the simulation scheme (4.6).

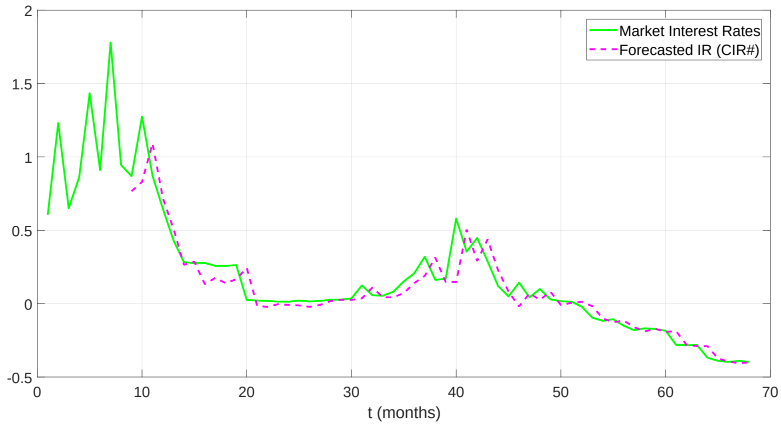


Figure 4.7. Forecast of the next-month interest rate based on a rolling window of 8 real data: monthly EUR interest rates with maturity T=1 day (overnight) versus predicted next-month interest rates.

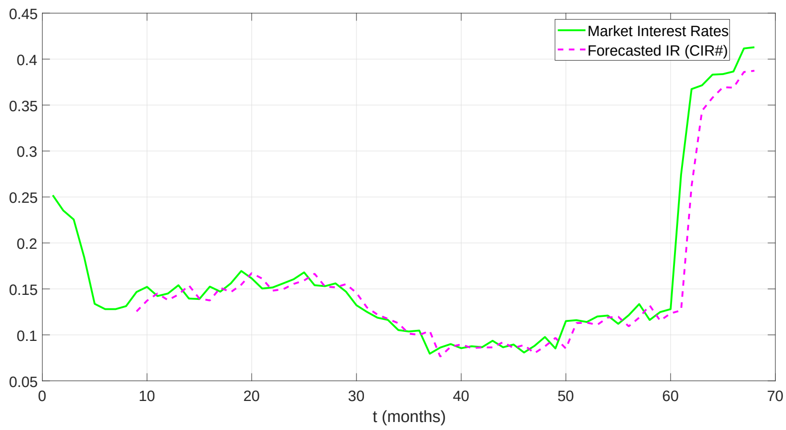


Figure 4.8. Forecast of the next-month interest rate based on a rolling window of 8 real data: monthly USD interest rates with maturity T=1 day (overnight) versus predicted next-month interest rates.

4.3.1 CIR# forecasts versus CIR forecasts

In this section we would like to point out the CIR# improvements in forecast as compared to the original CIR model. Note that future next-month interest rates are predicted by the CIR model using the expectation closed formula (3.6) after the CIR parameters (k, θ, σ) have been calibrated to the market historical data. This is done by applying the martingale estimating function method [10] described in Section 3.2.4. In this case to ensure the existence of such estimates (see Remark 3.2.1) the length of a rolling window must be greater than the window size of $m = 8$ observation required by the CIR# model. Figures 4.9 and 4.10 compare the future next-month values predicted by the CIR# with the ones forecasted by the classical CIR model, showing a better fitting to the real market available data.

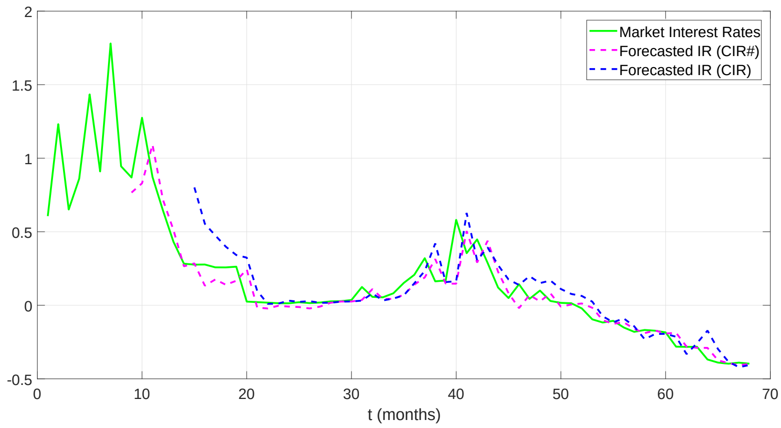


Figure 4.9. CIR# versus CIR: forecast of future next-month interest rates: monthly EUR interest rates with overnight maturity compared with future next-month interest rates predicted by the CIR# model based on a rolling window of $m = 8$ market data, and future next-month interest rates predicted by the classical CIR model based on a rolling window of $m = 14$ market data.

Figures 4.11, 4.12, 4.13 and 4.14 show the statistics R^2 and $RMSE$ values computed for all (63) maturities of EUR and USD datasets by the proposed CIR# and the original CIR model, respectively. The vertical black line separates Dataset I from Dataset II (see Tables 3.1 and 4.1). The plotted results show in all cases a better performance (bigger R^2 values and smaller $RMSE$ values) of the CIR# model.

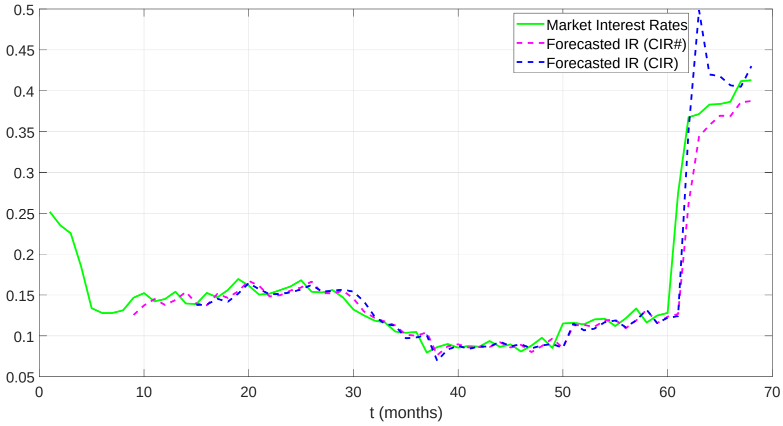


Figure 4.10. CIR# versus CIR: forecast of future next-month interest rates: monthly USD interest rates with overnight maturity compared with future next-month interest rates predicted by the CIR# model based on a rolling window of $m = 8$ market data, and future next-month interest rates predicted by the classical CIR model based on a rolling window of $m = 14$ market data.

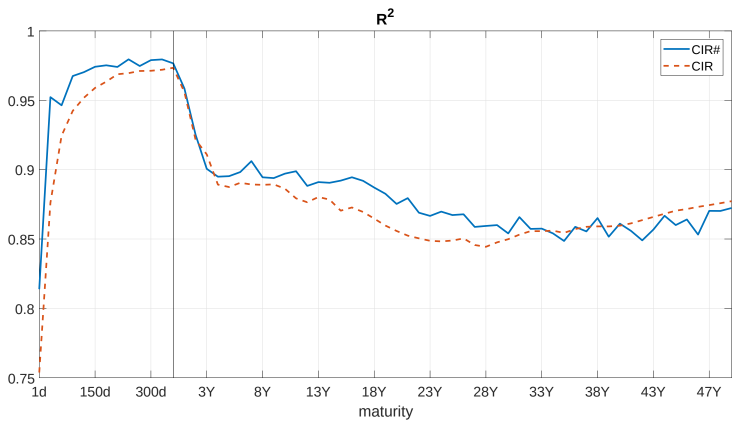


Figure 4.11. Statistics for EUR dataset: $R^2_{CIR\#}$ versus R^2_{CIR}, computed on a rolling window of size $m = 14$.

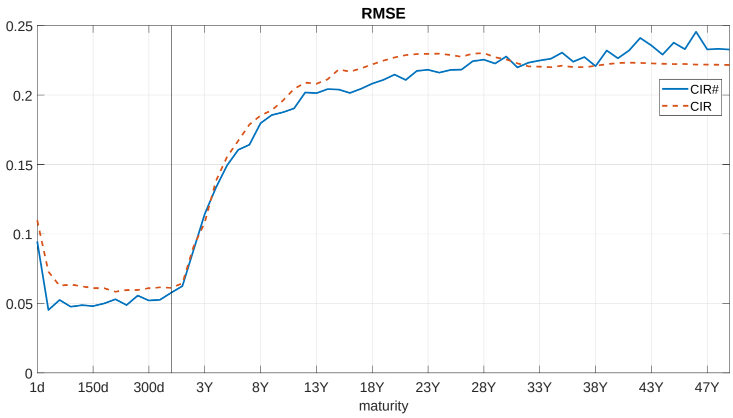


Figure 4.12. Statistics for EUR dataset: $RMSE_{CIR\#}$ versus $RMSE_{CIR}$, computed on a rolling window of size $m = 14$.

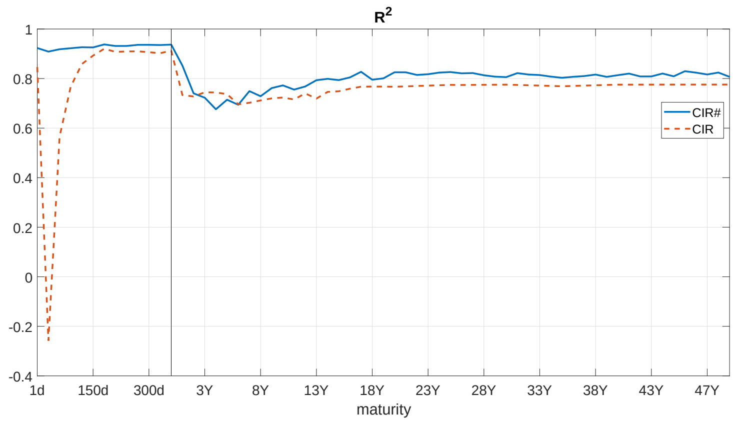


Figure 4.13. Statistics for USD dataset: $R^2_{CIR\#}$ versus R^2_{CIR}, computed on a rolling window of size $m = 14$.

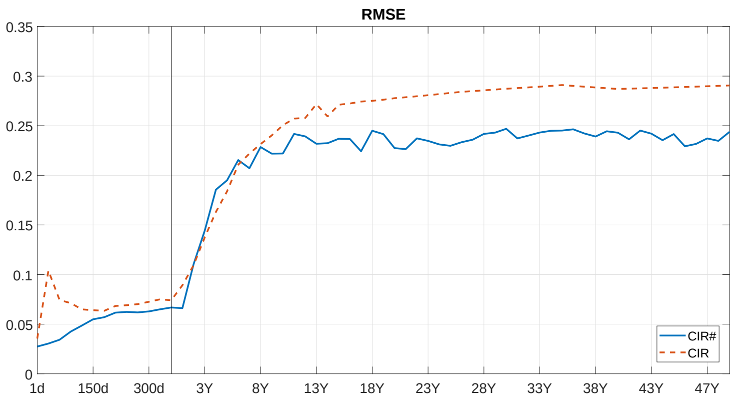


Figure 4.14. Statistics for USD dataset: $RMSE_{CIR\#}$ versus $RMSE_{CIR}$, computed on a rolling window of size $m = 14$.

Chapter 5

Evolution Families for Non-Autonomous Cases

In this chapter we will study the Cauchy problem related to the Black-Scholes model in the non-autonomous case, i.e. when the coefficients in the corresponding PDE are time dependent. Hence we start to consider the non-autonomous abstract Cauchy problem (nACP) of type

$$\begin{cases} u_t(t) = A(t)u(t), & t \geq s \in \mathbb{R} \\ u(s) = u_0 \in X, \end{cases}$$

where $A(t)$ represents a family of linear operators and u_0 is the initial data at time s, which belongs to a given Banach space X. In Section 5.1 we will introduce the basic theory to solve the above (nACP) (see, for instance [32, Section 9, Chapter VI] and [44, Section 13, Chapter II]). In particular Section 5.2 contains the approximation formulas provided by A. Batkai et al. [8], which are very helpful when the explicit solution can not be directly computed. Sometimes, by suitable transformations, the (nACP) may be reduced to another problem whose solution is known. This is the case of the Black-Scholes model with time-dependent deterministic coefficients, as will be described in Section 5.3. The existence of a unique evolution family (see Definition 5.1.2) solving more general (nACP), which are of interest in Mathematical Finance, is still an open problem. This will be a challenging task in a forthcoming research, starting from studying the generalized CIR problem (1.29) in the non-autonomous case and in suitable Banach spaces.

5.1 Evolution Families

We start by considering the following (nACP) on a given Banach space X

$$\begin{cases} u_t(t) = A(t)u(t), & t \geq s \in \mathbb{R} \\ u(s) = u_0 \in X, \end{cases} \tag{5.1}$$

where $(A(t), D(A(t)))$ is a family of (unbounded) linear operators on X.

Definition 5.1.1. Fixed $s \in \mathbb{R}$, $u_0 \in X$, a (classical) solution of (5.1) is a continuous function $u : [s, +\infty) \to X$, such that $u \in C^1([s, +\infty), X)$, $u(t) \in D(A(t))$ for all $t \geq s$ and satisfies (5.1).
The Cauchy problem (5.1) is called *well − posed* (on regularity subspaces $(Y_s)_{s \in \mathbb{R}}$, with exponentially bounded solution) if

i. (existence and uniqueness) for all $s \in \mathbb{R}$ there are dense subspaces Y_s of X such that $Y_s \subseteq D(A(s))$ and for every $y \in Y_s$ there is a unique classical solution $t \mapsto u(t; s, y)$;

ii. (continuous dependence) for $s_n \to s$ and $y_n \in Y_{s_n} \to y \in Y_s$, we have

$$\|\widehat{u}(t; s_n, y_n) - \widehat{u}(t; s, y)\| \to 0,$$

uniformly for t in compact subsets of $\mathbb{R}$, where

$$\widehat{u}(t; s, y) = \begin{cases} u(t; s, y), & if\, s \leq t \\ y, & otherwise. \end{cases}$$

iii. (exponential boundedness) there exist constants $M \geq 1$ and $\omega \in \mathbb{R}$ such that

$$\|u(t; s, y)\| \leq M e^{\omega(t-s)} \|y\|,$$

for all $y \in Y_s$ and $t \geq s$.

Definition 5.1.2. A family of bounded linear operators $(U(t, s))_{t \geq s}$ on a Banach space X is called a (strongly continuous) evolution family if

i. $U(t, r)U(r, s) = U(t, s)$ and $U(t, t) = I$, for all $t \geq r \geq s \in \mathbb{R}$;

ii. the mapping $(t, s) \mapsto U(t, s)$ is strongly continuous, with $t \geq s \in \mathbb{R}$.

We say that $(U(t, s))_{t \geq s}$ solves the Cauchy problem (5.1) (on spaces Y_t) if for all $s \in \mathbb{R}$ there are dense subspaces Y_s of X, such that the function $t \mapsto u(t; s, y) = U(t, s)y$ is a solution of (5.1), for $y \in Y_s$.

Definition 5.1.3. An evolution family $(U(t, s))_{t \geq s}$ is called evolution family solving the (nACP) (5.1) if for all $s \in \mathbb{R}$ the regularity subspace

$$Y_s = \{y \in X | t \mapsto U(t, s)y\, solves\, (nACP)\, (5.1)\},$$

is dense in X.

Proposition 5.1.4. *The Cauchy problem (*5.1*) is well-posed if and only if there exists a unique evolution family $(U(t, s))_{t \geq s}$ solving (*5.1*).*

Definition 5.1.5. Let $(U(t, s))_{t \geq s}$ be an evolution family on a Banach space X. We define $\mathcal{X} = C_0(\mathbb{R}, X)$, the space of continuous functions from $\mathbb{R}$ to X vanishing at infinity, normed by $\|f\| = \sup_{t \in \mathbb{R}} \|f(t)\|$ for all $f \in \mathcal{X}$, and the corresponding *evolution semigroup* $(\mathcal{T}(t))_{t \geq s}$ in the following way

$$(\mathcal{T}(t)f)(s) = U(s, s - t)f(s - t), \tag{5.2}$$

for $f \in \mathcal{X}$, $t \geq s \in \mathbb{R}$. It easy to check that $(\mathcal{T}(t))_{t\geq s}$ is a (C_0) semigroup on $\mathcal{X}$. We denote its generator by $(\mathcal{G}, D(\mathcal{G}))$.
In particular (5.2) can be also written as

$$(\mathcal{T}(t)f)(s) = U(t, t-s)\mathcal{R}(t)f(s).$$

$(\mathcal{R}(t)f)(s) = f(s-t)$ represents the (right) translation semigroup on $\mathcal{X}$, for all $t \geq s$, whose generator is the differentiation operator, say $-\frac{d}{ds}$ with domain

$$\mathcal{X}_d = D\left(-\frac{d}{ds}\right) = \{f \in C^1(\mathbb{R}, X) | f, f_s \in \mathcal{X}\}.$$

We can recover the evolution family from the evolution semigroup by choosing a function $f \in \mathcal{X}$, with $f(s) = u_0$, to obtain

$$U(t,s)u_0 = (\mathcal{R}(s-t)\mathcal{T}(t-s)f)(s),$$

for every $s \in \mathbb{R}$ and $t \geq s$. For a family $(A(s), D(A(s)))_{s\geq 0}$ of unbounded operators on X we consider the corresponding multiplication operator $(A(\cdot), D(A(\cdot)))$ on the space $\mathcal{X}$ defined as follows

$$(A(\cdot)f)(s) = A(s)f(s), \quad \forall s \in \mathbb{R},$$

$$D(A(\cdot)) = \{f \in \mathcal{X} | f(s) \in D(A(s)), \forall s \in \mathbb{R}, \textit{ and each function } s \mapsto A(s)f(s) \textit{ belongs to } \mathcal{X}\}.$$

Theorem 5.1.6. *Problem* (5.1) *is well-posed for the family* $(A(s), D(A(s)))_{s\in\mathbb{R}}$ *on a Banach space* X *if and only if there exists a unique evolution semigroup* $(\mathcal{T}(t))_{t\geq s}$ *with generator* $(\mathcal{G}, D(\mathcal{G}))$ *and an invariant core* $\mathcal{D} \subset \mathcal{X}_d \cap D(\mathcal{G})$ *such that* $\mathcal{G}f + f_s = A(\cdot)f$, *for all* $f \in \mathcal{D}$.

There exist several sufficient conditions for well-posedness. We present here the assumptions proposed by A. Batkai et al. [8] in the parabolic case.

Assumption 5.1.7. Assume that

i. The domain $D(A(t))$ is dense in X and is indipendent of t;

ii. for each $t \in \mathbb{R}$, $A(t)$ generates an analytic semigroup $e^{\cdot A(t)}$ such that $\|e^{\cdot A(t)}\| \leq Me^{\omega \cdot}$ for some constants $M \geq 1$, $\omega \leq 0$. Moreover, for all $t \in \mathbb{R}$ the resolvent $R(\lambda, A(t))$ exists for all $\lambda \in \mathbb{C}$ with $Re\, \lambda > 0$, there is a constant $M \geq 1$ such that
$$\|R(\lambda, A(t))\| \leq \frac{M}{|\lambda| + 1};$$

iii. there exist constants $L \geq 0$ and $\alpha \in (0, 1]$ such that
$$\|(A(t) - A(s))A(0)^{-1}\| \leq L|t-s|^{\alpha},$$
for all t, $s \in \mathbb{R}$.

Then the problem (5.1) is well-posed on $D(A(t))$.

5.2 Product Formulas

We start to consider the autonomous Cauchy problem on a Banach space X

$$\begin{cases} u_t(t) = (A+B)u(t), & t \geq s \in \mathbb{R} \\ u(s) = u_0 \in X, \end{cases} \tag{5.3}$$

where the operators A, B, and the closure of $(A+B)$ are supposed to be generators of strongly continuous semigroups T, S, and U, respectively. The operator splitting consists in recovering the solution semigroup U using the semigroups T and S. For splitting procedures we mention the most used Trotter-Kato approximation defined by

$$u_n(t) = [T(t/n)S(t/n)]^n u_0,$$

for $t \geq s$ and $n \in \mathbb{N}, n \geq 1$, whose convergence is guaranteed by Theorem 1.1.30. However, we notice that Theorem 1.1.30 does not provide any information a priori about the rate of convergence p. For convenience, we can denote by $(F(t))_{t\geq s}$ the family $(T(t)S(t))_{t\geq s}$. F is said to be of order $p > 0$, if for u_0 fixed in a suitably subset of X there is $C > 0$ such that for all $t \in [s, \tau]$, $\tau > 0$, it follows

$$\|F(t/n)^n u_0 - U(t)u_0\| \leq C/n^p.$$

To obtain error estimates for diffusion problems, we apply the following result by T. Jahnke and C. Lubich [53, Theorem 2.1], which relies on commutator bounds.

Theorem 5.2.1. *Let A be the generator of a (C_0) contraction semigroup on X and B a bounded operator on X. Assume that there exists a $\alpha \geq 0$ such that*

$$\|[A, B]v\| \leq c\|(-A)^\alpha v\|,$$

for all v in some dense subspace D of $D((-A)^\alpha)$ which is invariant under the semigroup $e^{t(A+B)}$ generated by $(A+B)$. In particular, $[A, B]$ is the commutator operator given by $(AB - BA)$ and $(-A)^\alpha$ represents the fractional power of A (for more details, see e.g. [32, Section 2.5]). Then one has the convergence with order $p = 1$, i.e.

$$\|(e^{\frac{t}{n}B} e^{\frac{t}{n}A})^n v - e^{t(A+B)}v\| \leq \frac{Ct^2}{n} \|(-A)^\alpha v\|.$$

Now, we consider the the Cauchy problem (5.3) in the non-autonomous case, i.e.

$$\begin{cases} u_t(t) = (A(t)+B(t))u(t), & t \geq s \in \mathbb{R} \\ u(s) = u_0 \in X. \end{cases} \tag{5.4}$$

Suppose we are able to solve both the following autonomous Cauchy problems

$$w_t(t) = A(r)w(t), \quad t \geq s, \tag{5.5}$$

$$v_t(t) = B(r)v(t), \quad t \geq s, \tag{5.6}$$

with appropriate initial conditions, for every fixed time $r \in \mathbb{R}$.
In order to find the solution at time $s + t > 0$ with the sequential splitting, let the

time interval $[s, s+t]$ be partitioned in n subintervals with time step $\tau = t/n$. If $u(s) = w_0$ is the initial value of (5.5), compute the solution $w^{(1)}(s+\tau)$ of (5.5) on $[s, s+\tau]$ taking $r = s$. Then, compute the solution $v^{(1)}(s+\tau)$ of (5.6) on $[s, s+\tau]$ taking $w^{(1)}(s+\tau)$ as initial value. Set $u(s+\tau) = v^{(1)}(s+\tau)$ as the initial value of (5.4), repeat the above procedure for n times.
The iterative scheme at step k, with $k \in \{1, 2, ..., n\}$, can be summarized as follows
Step 1. Compute the solution $w^{(k)}(s+k\tau)$ of the problem

$$\begin{cases} w_t^{(k)}(t) = A(s+(k-1)\tau)w^{(k)}(t), \;\; t \in (s+(k-1)\tau, s+k\tau], \\ w^{(k)}(s+(k-1)\tau) = u(s+(k-1)\tau). \end{cases}$$

Step 2. Compute the solution $v^{(k)}(s+k\tau)$ of the problem

$$\begin{cases} v_t^{(k)}(t) = B(s+(k-1)\tau)v^{(k)}(t), \;\; t \in (s+(k-1)\tau, s+k\tau], \\ v^{(k)}(s+(k-1)\tau) = w^{(k)}(s+k\tau). \end{cases}$$

Step 3. Set

$$u(s+k\tau) = v^{(k)}(s+k\tau).$$

Observe that, for any $r \in [0, \tau]$

$$w^{(k)}(s+(k-1)\tau+r) = e^{rA(s+(k-1)\tau)}u(s+(k-1)\tau),$$

and

$$\begin{aligned} v^{(k)}(s+(k-1)\tau+r) &= e^{rB(s+(k-1)\tau)}w^{(k)}(s+(k-1)\tau) \\ &= e^{rB(s+(k-1)\tau)}e^{rA(s+(k-1)\tau)}u(s+(k-1)\tau). \end{aligned}$$

Thus, the split solution $u(s+k\tau)$ can be written by induction as

$$u(s+k\tau) = \prod_{p=0}^{k-1} e^{\tau B(s+p\tau)}e^{\tau A(s+p\tau)}u_0, \tag{5.7}$$

for $k\tau \le t$, $k \in \{1, 2, ..., n\}$ and $u_0 \in X$.
The next result is a general convergence thorem for the problem (5.4).

Theorem 5.2.2. *Consider the non-autonomous Cauchy problem (5.4) and assume that*

i. (Well-Posedness) the problem is well-posed;

ii. (Stability) for all $r \in \mathbb{R}$ the operators $A(r), B(r)$ generate (C_0) semigroups of type (M, ω), with $M \ge 1$, $\omega \in \mathbb{R}$ on X. Thus $(\omega, +\infty) \subset \rho(A(r)) \cap \rho(B(r))$ for all $r \in \mathbb{R}$. Moreover, let

$$\sup_{s\in\mathbb{R}} \left\| \prod_{p=n}^{1} \left(e^{\frac{t}{n}B(s-p\frac{t}{n})} e^{\frac{t}{n}A(s-p\frac{t}{n})} \right) \right\| \le Me^{\omega t};$$

iii. (Continuity) the maps $t \mapsto R(\lambda, A(t))u_0$ and $t \mapsto R(\lambda, B(t))u_0$ are continuous for all $\lambda > \omega$ and $u_0 \in X$.

If $(U(t,s))_{t\geq s}$ denotes the evolution family solving (5.4), with related evolution semigroup $(\mathcal{U}(t))_{t\geq s}$ generated by the closure of the operator $-\frac{d}{ds}+A(\cdot)+B(\cdot)$, then one has the convergence

$$U(t,s)u_0 = \lim_{n\to+\infty} \prod_{p=0}^{n-1} (e^{\frac{t-s}{n}B(s+\frac{p(t-s)}{n})} e^{\frac{t-s}{n}A(s+\frac{p(t-s)}{n})})u_0, \tag{5.8}$$

for all $u_0 \in X$, locally uniformly in s,t, with $t \geq s$.
In partcular $(\mathcal{U}(t))_{t\geq s}$ is given by

$$(\mathcal{U}(t)f)(s) = \lim_{n\to+\infty} \prod_{p=n}^{1} (e^{\frac{t}{n}B(s+\frac{pt}{n})} e^{\frac{t}{n}A(s+\frac{pt}{n})}) f(t-s), \tag{5.9}$$

for all $f \in \mathcal{X}$ in the uniform topology.

Now, suppose we are able to solve both the non-autonomous Cauchy problems

$$w_t(t) = A(t)w(t), \ \ t \geq s, \tag{5.10}$$

$$v_t(t) = B(t)v(t), \ \ t \geq s, \tag{5.11}$$

with appropriate initial conditions.
If $w(s) = w_0$ is the initial value of (5.10), compute the solution $w^{(1)}(s+\tau)$ of (5.5) on $[s, s+\tau]$. Then, compute the solution $v^{(1)}(s+\tau)$ of (5.11) on $[s, s+\tau]$ taking $w^{(1)}(s+\tau)$ as initial value. With $u(s+\tau) = v^{(1)}(s+\tau)$ as the initial value of (5.10), repeat the above procedure for n times.
The algorithm scheme at step k, with $k \in \{1, 2, ..., n\}$ can be summarized as follows
Step 1. Compute the solution $w^{(k)}(s+k\tau)$ of the problem

$$\begin{cases} w_t^{(k)}(t) = A(t)w^{(k)}(t), \ \ t \in (s+(k-1)\tau, s+k\tau], \\ w^{(k)}(s+(k-1)\tau) = u(s+(k-1)\tau), \end{cases}$$

Step 2. Compute the solution $v^{(k)}(s+k\tau)$ of the problem

$$\begin{cases} v_t^{(k)}(t) = B(t)v^{(k)}(t), \ \ t \in (s+(k-1)\tau, s+k\tau], \\ v^{(k)}(s+(k-1)\tau) = w^{(k)}(s+k\tau), \end{cases}$$

Step 3. Set

$$u(s+k\tau) = v^{(k)}(s+k\tau).$$

If $(W(t,s))_{t\geq s}$ and $(V(t,s))_{t\geq s}$ denote the evolution families solving the above problems (5.10) and (5.11), we have

$$w^{(k)}(r) = W(r, s+(k-1)\tau)u(s+(k-1)\tau),$$

and

$$v^{(k)}(r) = V(r, s+(k-1)\tau)w^{(k)}(s+k\tau) = V(r, s+(k-1)\tau)W(s+k\tau, s+(k-1)\tau)u(s+(k-1)\tau),$$

for $r \in [0,\tau]$. Then the split solution $u(s+k\tau)$ can be written by induction as

$$u(s+k\tau) = \prod_{p=0}^{k-1} (V(s+(p+1)\tau, s+p\tau)W(s+(p+1)\tau, s+p\tau))u_0, \tag{5.12}$$

for $k\tau \leq t$, $k \in \mathbb{N}, k \geq 1$ and $u_0 \in X$.
The general convergence result, in this case, is given by the following theorem.

Theorem 5.2.3. *Consider the problem (*5.4*) and assume that*

i. (Well-Posedness) the (nACP) corrisponding to the operators $A(\cdot), B(\cdot)$ and $A(\cdot)+B(\cdot)$ are well-posed;

ii. (Stability) there exist $M \geq 1$ and $\omega \in \mathbb{R}$ such that

$$\sup_{s\geq 0}\left\| \prod_{p=n-1}^{1} \left(V\left(s-\frac{pt}{n}, s-\frac{(p+1)t}{n}\right) W\left(s-\frac{pt}{n}, s-\frac{(p+1)t}{n}\right)\right)\right\| \leq Me^{\omega t},$$

where $(W(t,s))_{t\geq s}$ and $(V(t,s))_{t\geq s}$ denotes the evolution families solving the (nACP) corresponding to $A(\cdot)$ and $B(\cdot)$ respectively. Then, if $(U(t,s))_{t\geq s}$ denotes the evolution family solving the (nACP) corresponding to $A(\cdot)+B(\cdot)$, then one has the convergence

$$\begin{aligned} &U(t,s)u_0 \\ &= \lim_{n\to+\infty} \prod_{p=0}^{n-1} \left(V\left(s+\frac{(p+1)(t-s)}{n}, s+\frac{p(t-s)}{n}\right) W\left(s+\frac{(p+1)(t-s)}{n}, s+\frac{p(t-s)}{n}\right)\right) u_0, \end{aligned} \tag{5.13}$$

for all $u_0 \in X$.
In partcular the associated evolution semigroup $(\mathcal{U}(t))_{t\geq s}$ is given by

$$(\mathcal{U}(t)f)(s) = \lim_{n\to+\infty} \prod_{p=n-1}^{0} \left(V\left(s-\frac{pt}{n}, s-\frac{(p+1)t}{n}\right) W\left(s-\frac{pt}{n}, s-\frac{(p+1)t}{n}\right)\right) f(s-t), \tag{5.14}$$

for all $f \in \mathcal{X}$ in the uniform topology.

For a general result that ensures the well-posedness of (nACP) (5.4), a reader can also see the Kato's Existence Theorem, given by [44, Theorem 13.13].

5.3 Black-Scholes Model with Time-Dependent Coefficients

In this section we consider the Black-Scholes problem with the coefficients, say the volatility σ and the risk-free interest rate r, depending on time. In this way the model becomes more realistic for financial markets. In the recent literature, the non-autonomous problem is treated with different techniques, see e.g. the Stieltjes moment approach [83], the Lie simmetry theory [91], the PDEs approach by using change of variables [63, Section 3.4.1]. In this subsection we will derive the explicit solution to the problem by applying the semigroup approach. Consider the (forward) Black-Scholes equation

$$\begin{cases} u_t(t,x) = \frac{\sigma^2(t)}{2}x^2 u_{xx}(t,x) + r(t)xu_x(t,x) - r(t)u(t,x), \quad t \geq s \geq 0,\ x \in \mathbb{R}_+, \\ u(s,x) = u_0(x), \quad x \in \mathbb{R}_+. \end{cases} \tag{5.15}$$

The state variable x denotes a realization of the stochastic price at time t of the underlying risky asset. The initial data $u_0 \in X$, where we consider $X = C[0, +\infty]$. The problem (5.15) can be rewritten as the following (nACP)

$$\begin{cases} u_t(t,x) = \mathcal{L}(t)u(t,x), & t \geq s \geq 0,\ x \in \mathbb{R}_+, \\ u(s,x) = u_0(x), & x \in \mathbb{R}_+. \end{cases} \tag{5.16}$$

where

$$\mathcal{L}(t)u(x) = \frac{\sigma^2(t)}{2} x^2 u''(x) + r(t)xu'(x) - r(t)u(x).$$

The following result holds.

Theorem 5.3.1. *Assume that $\sigma(\cdot), r(\cdot) \in BUC^1(\mathbb{R}_+)$. For all $t \geq s \geq 0$ the operator $\mathcal{L}(t)$ has domain independent of time*

$$D(\mathcal{L}(t)) = D(\mathcal{L}(0)) = \{u \in X \cap C^2(0, +\infty) | \mathcal{L}(0)u \in X\}, \tag{5.17}$$

that is dense in X. Then the (nACP) (5.16) is well-posed with solution given by

$$\begin{aligned} u(t,x) &= U(t,s)u_0(x) \\ &= e^{-\int_s^t r(\zeta)\,d\zeta} \cdot \int_{-\infty}^{+\infty} u_0\Big(x\, exp\Big(\int_s^t \Big(r(\zeta) - \nu^2(\zeta)\Big)\,d\zeta - y\Big)\Big) q(t,y)\,dy, \end{aligned}$$

where $\nu(t) = \sigma(t)/\sqrt{2}$ and

$$q(t,y) = \frac{1}{\sqrt{2\pi \int_s^t \nu^2(\zeta)\,d\zeta}} exp\Big(-\frac{y^2}{2\int_s^t \nu^2(\zeta)\,d\zeta}\Big). \tag{5.18}$$

Proof. Note that the operators $\mathcal{L}(t)$ can be written as

$$\mathcal{L}(t)u = \nu^2(t)G^2u + \gamma(t)Gu + \delta(t)I, \quad t \geq s \geq 0,$$

where we set

$$\nu(t) = \frac{\sigma(t)}{\sqrt{2}},\ \gamma(t) = r(t) - \nu^2(t),\ \delta(t) = -r(t),$$

and the operators G, G^2, with their respective domains, are defined in (1.22) and (1.23), respectively. Thus, for all $t \geq s \geq 0$, $D(\mathcal{L}(t)) = D(G^2)$.
With similar arguments to those in Theorem 1.2.2 we can conclude that there exists the unique evolution family solving the (nACP) (5.16) having the following explicit representation

$$\begin{aligned} &U(t,s)u_0(x) \\ &= \frac{e^{-\int_s^t r(\zeta)\,d\zeta}}{\sqrt{2\pi}} \cdot \int_{-\infty}^{+\infty} u_0\Big(x\, exp\Big(\int_s^t \Big(r(\zeta) - \nu^2(\zeta)\Big)\,d\zeta - z\sqrt{\int_s^t \nu^2(\zeta)\,d\zeta}\Big) e^{-\frac{z^2}{2}}\,dz. \end{aligned} \tag{5.19}$$

Let

$$y = z\sqrt{\int_s^t \nu^2(\zeta)\,d\zeta},$$

then we have

$$U(t,s)u_0(x) = e^{-\int_s^t r(\zeta)\,d\zeta} \cdot \int_{-\infty}^{+\infty} u_0\Big(x\,exp\Big(\int_s^t \Big(r(\zeta) - \nu^2(\zeta)\Big)\,d\zeta - y\Big)q(t,y)\,dy,$$

where $q(t,y)$ is given in (5.18). □

5.4 Some Hints for the CIR Process with Time-Dependent Coefficients

In this section we consider the following (nACP)

$$\begin{cases} u_t(x) = \mathcal{L}(t)u(x), & t \geq s \geq 0,\ x \in \mathbb{R}_+ \\ u(s,x) = u_0(x), & x \in \mathbb{R}_+, \end{cases} \tag{5.20}$$

where u_0 is the initial data and for any $t \geq s$

$$\mathcal{L}u(x) = \frac{1}{2}\sigma^2(t)xu''(x) + k(\theta(t) - x)u'(x) - xu(x), \tag{5.21}$$

represents the operator related to the CIR problem. We are assuming that the volatility and the lon-run mean parameters are time-dependent, i.e. $\sigma = \sigma(t)$ and $\theta = \theta(t)$ while k is constant. This problem was recentely studied by M. Moreno and F. Platania in [70] in the framework of affine models. They consider consider a cyclical square-root model for the term structure where the short interest rates are pulled back to a certain time-dependent long term level characterized by an harmonic oscillator. Departing from this harmonic oscillator, they assume that the volitility and the mean reversion level are defined as

$$\begin{aligned} \sigma^2(t) &= \sigma_0\, sin^2(\varphi - \omega t), \\ \theta(t) &= \theta_0\, sin^2(\varphi - \omega t), \end{aligned} \tag{5.22}$$

where $\sigma_0, \theta_0, \varphi$ and ω denote the amplitude, offset phase, and temporal frequency, respectively.

In Section 1.2.2 we showed that the solution of the (ACP) related to the CIR problem can be expressed as the limit of a sequence of approximate solutions computed by the Lie-Trotter-Daletskii formula, in Section 5.2 we described the methodology provided by A. Batkai et al [8], to compute similar approximate solutions for non-autonomous Cauchy problems. The existence of an evolution family solving the problem (5.20) and the representation of the corresponding evolution semigroup is under investigation. In this section we give some hints to apply Theorem 5.2.3 to the problem (5.20).

We let $X = C_0[0,+\infty)$ and assume that $\sigma(t), \theta(t) \in BUC^1(\mathbb{R}_+)$. As made in Theorem 1.2.8 we can split the entire operator $\mathcal{L}(t)$ as the sum of the operators $\nu^2(t)G^2$ and $(Q(t)+P)$, where

$$Gu(x) = \sqrt{x}u'(x), \quad G^2u(x) = xu''(x) + \frac{1}{2}u'(x),$$

$$Q(t)u(x) = (\alpha(t) + \beta x)u'(x), \quad Pu(x) = -xu(x),$$

with $\nu(t) = \frac{\sigma(t)}{\sqrt{2}}$, $\alpha(t) = \big(k\theta(t) - \frac{\nu^2(t)}{2}\big)$, $\beta = -k$, for any $x \in \mathbb{R}_+$. It is clear that also $\nu(t), \alpha(t)$ are of class BUC^1. Now, in order to apply Theorem 5.2.3, we have to solve the two following (nACP):

$$\begin{cases} u_t(x) = \nu^2(t)G^2u(x), & t \geq s,\ x \in \mathbb{R}_+ \\ u(s,x) = f(x), & f \in X,\ x \in \mathbb{R}_+, \end{cases} \tag{5.23}$$

and

$$\begin{cases} u_t(x) = (Q(t) + P)u(x), & t \geq s,\ x \in \mathbb{R}_+ \\ u(s,x) = f(x), & f \in X,\ x \in \mathbb{R}_+. \end{cases} \tag{5.24}$$

The evolution family solving problem (5.23) is given by

$$\int_{-\infty}^{+\infty} f([\sqrt{x} + y]^2)q(t,y)\,dy, \tag{5.25}$$

where $q(t,y)$ is defined in (5.18) (see [40, Lemma 1, Section 3]). This solution is analogous to formula (1.36) for the autonomous case, by the substitution

$$\nu^2 \to \frac{1}{t-s}\int_s^t \nu^2(\zeta)\,d\zeta.$$

To solve the second problem (5.24) we start to consider the following sub-problems

$$\begin{cases} u_t(x) = Q(t)u(x), & t \geq s,\ x \in \mathbb{R}_+ \\ u(s,x) = f(x), & f \in X,\ x \in \mathbb{R}_+, \end{cases} \tag{5.26}$$

$$\begin{cases} u_t(x) = Pu(x), & t \geq s,\ x \in \mathbb{R}_+ \\ u(s,x) = f(x), & f \in X,\ x \in \mathbb{R}_+. \end{cases} \tag{5.27}$$

By a direct inspection we see that the evolution family solving problem (5.26) is given by

$$f\Big(e^{(t-s)\beta}x + \int_s^t \alpha(\zeta)e^{\zeta\beta}\,d\zeta\Big), \tag{5.28}$$

while the solution of (5.26) is given by $e^{-(t-s)x}f(x)$. In particular, we note that formula (5.28) is analogous to formula (1.37) for the autonomous case, by replacing[1]

$$\alpha\int_s^t e^{(\zeta-s)\beta}\,d\zeta \to \int_s^t \alpha(\zeta)e^{(\zeta-s)\beta}\,d\zeta. \tag{5.29}$$

Again, by a direct inspection we can see that the evolution family solving problem (5.24) is given by

$$exp\Big(-\Big(x\int_s^t e^{(\zeta-s)\beta}\,d\zeta + \int_s^t \alpha(\zeta)\Big(\int_s^\zeta e^{(q-s)\beta}\,dq\Big)\,d\zeta\Big)\Big)f\Big(e^{(t-s)\beta}x + \int_s^t \alpha(\zeta)e^{(\zeta-s)\beta}\,d\zeta\Big). \tag{5.30}$$

[1]We use the equivalent (integral) form $\int_s^t e^{(\zeta-s)\beta}\,d\zeta$ for the term $\frac{e^{(t-s)\beta}-1}{\beta}$ of (1.37) and (1.38).

We note that the expression (5.30) is analogous to formula (1.38), by (5.29) and replacing

$$\frac{\alpha}{\beta}\left(\int_s^t e^{(\zeta-s)\beta}\,d\zeta-(t-s)\right)\to\int_s^t \alpha(\zeta)\left(\int_s^\zeta e^{(q-s)\beta}\,dq\right)d\zeta. \tag{5.31}$$

Once we have computed the evolution families solving the sub-problems (5.23) and (5.24) we can apply Theorem 5.2.3 to find the approximate solution to problem (5.20) if and only if we are able to prove that it is well-posed. As mentioned above this is under investigation. The most challenging task consists in proving that the operator $\mathcal{L}(t)$ is independent of t, for all t.

Chapter 6

Appendices

6.1 Boundary Classification for Diffusion Processes

In this appendix, we describe the Feller boundary classification introduced by Karlin and Taylor in [58, Chapter XV, Paragraph 6.1] using a proababilistic approach. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ a filtered probability space. Recall that a diffusion process $(X_t)_{t\geq 0}$ is a Markov process with dynamics governed by

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t \;\; X_0 = x_0, \tag{6.1}$$

whose sample paths are continuous functions of $t \geq 0$.

Definition 6.1.1. (Regular Diffusion Processes)
Let $(X_t)_{t\geq 0}$ a diffusion process with state space the interval $J = (r_1, r_2)$.[1] $(X_t)_{t\geq 0}$ is regular if starting from any point in the interior of J any other interior point of J can be reached with positive probability. In other words, for any point ζ in the interior of J, let I_ζ the event that ζ is even reached; then $(X_t)_{t\geq 0}$ is regular if

$$\mathbb{P}(I_\zeta < +\infty | X_0 = x_0) > 0,$$

whenever $r_1 < x_0$, $\zeta < r_2$.

We describe the classification of possible behavior near the boundaries r_1, r_2 concentrating just on the left point r_1, the right will be entirely similar. Consider $\mu(t, x) = \mu(x)$ and $\sigma(t, x) = \sigma(x)$.
We start to define the *scale function*

$$S(x) = \int_{x_0}^{x} s(z)\, dz, \tag{6.2}$$

with

$$s(z) = exp\Big(-\int_{z_0}^{z} \frac{2\mu(\eta)}{\sigma^2(\eta)}\, d\eta\Big), \tag{6.3}$$

where x_0, z_0 are arbitrary fixed point inside (r_1, r_2) whose choice is no relevant. For all $[c, d] \subset (r_1, r_2)$, define the *scale measure* $dS(x)$ on the infinitesimal interval $[x, x + dx]$, where

$$dS(x) = S(x + dx) - S(x) = s(x)dx.$$

[1] J can be also of type $(r_1, r_2], [r_1, r_2), [r_1, r_2]$.

Similarly we introduce the *speed density*

$$m(x) = \frac{1}{\sigma^2(x)s(x)}, \tag{6.4}$$

and its *speed measure* $dM(x) = m(x)dx$.
Finnaly, we define

$$S(r_1, x_0] = \lim_{a \to r_1} S[a, x_0] = \lim_{a \to r_1} \int_a^{x_0} s(z)dz.$$

Lemma 6.1.2. *For all $r_1 < a < x_0 < b < r_2$ and for some $x \in (r_1, r_2)$, we have that*

i. $S(r_1, x_0] < +\infty$ *implies* $\mathbb{P}(I_{r_1+} \leq I_b | X_0 = x_0) > 0$*;*

ii. $S(r_1, x] = +\infty$ *implies* $\mathbb{P}(I_{r_1+} \leq I_b | X_0 = x_0) = 0$*,*

where $I_{r_1+} = \lim_{a \to r_1} I_a \leq +\infty$*.*

In view of Lemma 6.1.2 we define the *attracting* boundary.

Definition 6.1.3. (Attracting Boundary)
The boundary r_1 is of attracting type if $S(r_1, x_0] < +\infty$ for an arbitrary $x_0 \in (r_1, r_2)$.

We define (employing some obvious interchanges of order of integration)

$$\begin{aligned} \Sigma(r_1) &= \lim_{a \to r_1} \int_a^{x_0} S[a, z]\, dM(z) = \int_{r_1}^{x_0} S(r_1, z]\, dM(z) = \int_{r_1}^{x_0} \left(\int_{r_1}^{z} s(\eta)\, d\eta \right) m(z)\, dz \\ &= \int_{r_1}^{x_0} \left(\int_{\eta}^{x_0} m(z)\, dz \right) s(\eta)\, d\eta = \int_{r_1}^{x_0} M[\eta, x_0]\, dS(\eta). \end{aligned} \tag{6.5}$$

Notice that we use the notation $\Sigma(r_1)$ instead of $\Sigma(r_1, x_0)$ since the choice of x_0 is not relevant, $\Sigma(r_1)$ measures the time needed to reach the boundary r_1 starting from an interior point x_0.

Definition 6.1.4. (Attainable/Unattainable Boundary)
The boundary r_1 is said to be

i. *attainable* if $\Sigma(r_1) < +\infty$;

ii. *unattainable* if $\Sigma(r_1) = \pm\infty$.

In particular, r_1 is *nonattracting* if it is unattainable with $S(r_1, x_0] = \pm\infty$.

It easy to verify that $S(r_1, x_0] < +\infty$ whenever $\Sigma(r_1) < +\infty$, i.e. if r_1 is attainable then it is attracting.
The next lemma shows that an attainable boundary can be reached in finite time with positive probability, and the expected time to reach an unattainable boundary is always infinite.

Lemma 6.1.5. *Let $r_1 < x_0 < b < r_2$ with r_1 an attracting boundary. The following statements are equivalent:*

i. $\mathbb{P}(I_{r_1} < +\infty | X_0 = x_0) > 0$;

ii. $\Sigma(r_1) = \int_{r_1}^{x_0} S(r_1, \eta] \, dM(\eta) < +\infty$.

We next introduce the functionals

$$M(r_1, x_0] = \lim_{a \to x_0} M[a, x_0] = \lim_{a \to r_1} \int_a^{x_0} m(z) dz,$$

and

$$N(r_1) = \int_{r_1}^{x_0} M(r_1, z] \, dS(z) = \int_{r_1}^{x_0} S[\eta, x_0] \, dM(\eta). \tag{6.6}$$

$M(r_1, x_0]$ measures the speed of the process near r_1 and $N(r_1)$ measures the time needed to reach an interior point x_0 starting from r_1.
The classification of boundary behavior is based on the values of the four functionals $S(r_1, x_0], \Sigma(r_1), N(r_1)$ and $M(r_1, x_0]$, as described by the next lemma.

Lemma 6.1.6. *The following relations hold between* $S(r_1, x_0], \Sigma(r_1), N(r_1)$ *and* $M(r_1, x_0]$*:*

i. $S(r_1, x_0] = \pm\infty$ *implies* $\Sigma(r_1) = \pm\infty$;

ii. $\Sigma(r_1) < +\infty$ *implies* $S(r_1, x_0] < +\infty$;

iii. $M(r_1, x_0] = \pm\infty$ *implies* $N(r_1) = \pm\infty$;

iv. $N(r_1) < +\infty$ *implies* $M(r_1, x_0] < +\infty$.

Proof. See [58, Chapter XV, Lemma 6.3]. □

The next definition provides the terminology of the Feller boundary classification.

Definition 6.1.7. (Feller Classification)
In the notations defined above, the following boundary classification holds:

i. r_1 is of *regular* type if $S(r_1, x_0] < +\infty$ and $M(r_1, x_0] < +\infty$;

ii. r_1 is of *exit* type if $\Sigma(r_1) < +\infty$ but $M(r_1, x_0] = \pm\infty$;

iii. r_1 is of *entrance* type if $S(r_1, x_0] = \pm\infty$ while $N(r_1) < +\infty$;

iv. r_1 is of *natural* type if $\Sigma(r_1) = \pm\infty$ and $N(r_1) = \pm\infty$.

In particular, it easy to verify that a regular or exit boundary is attracting while an entrance or natural boundary is nonattracting.

Example 6.1.8. Consider the squared Bessel process X_t (of parameter $\delta \geq 0$) with dynamic

$$dX_t = \delta dt + 2\sqrt{X_t} dW_t,$$

and boundaries $r_1 = 0$, $r_2 = +\infty$. In particular, if we let $Y_t = X_t^2$, obtain a Bessel process satisfying the following SDE

$$dY_t = \frac{\delta - 1}{2Y_t} dt + dW_t.$$

With this transformation it is more simple to study the boundaries behavior.
The scale and speed functions are $s(z) = z^{1-\delta}$ and $m(z) = z^{\delta-1}$, hence, given $x_0 = 1$, we have

$$S(x) = \int_1^x z^{1-\delta}\, dz = \begin{cases} \frac{1}{2-\delta} x^{2-\delta}, & if\, \delta \neq 2, \\ ln\, x, & if\, \delta = 2. \end{cases}$$

We start to examine the boundary point 0. Let $x_0 = 1$ for comodity. We have

$$\Sigma(0) = \int_0^1 \Big(\int_z^1 m(\eta)\, d\eta \Big) s(z)\, dz = \int_0^1 \frac{1}{\delta}(1 - z^{\delta}) z^{1-\delta}\, dz$$

$$= \frac{1}{\delta} \int_0^1 (z^{1-\delta}\, dz - 1) \begin{cases} < +\infty, & if\, \delta < 2, \\ = +\infty & if\, \delta \geq 2. \end{cases}$$

$$N(0) = \int_0^1 \Big(\int_\eta^1 s(z)\, dz \Big) m(\eta)\, d\eta = \frac{1}{2-\delta} \int_0^1 (1 - \eta^{2-\delta}) \eta^{\delta-1}\, d\eta$$

$$= \frac{1}{2-\delta} \int_0^1 (\eta^{\delta-1}\, d\eta - \frac{1}{2}) \begin{cases} < +\infty, & if\, \delta > 0, \\ = +\infty & if\, \delta = 0. \end{cases}$$

Hence, the boundary 0 is entrance if $\delta \geq 2$, regular if $0 < \delta < 2$ and exit if $\delta = 0$. It simple to verify that the boundary $+\infty$ is natural for all $\delta \geq 0$.

Remark 6.1.9. Note that Definition (6.1.7) and Definition (1.1.20) are equivalent. This is a consequence of Remark 1.1.19.

6.2 Approximate Solution for the CIR Problem

This appendix contains the calculations given in [40] to obtain the approximate formula (1.35).
For any $t > 0$ consider a uniform partition of $[0, t]$ into $n \in \mathbb{N}$ subintervals of length t/n, say $\Pi_n(t) = \{t_{j,n}\}_{j=0}^n$ $(0 = t_{0,n} < t_{1,n} < \cdots < t_{n,n} = t)$, such that $t_{j,n} = \frac{jt}{n}$, $j = 0, 1, \ldots, n$. Suppose $n = 2^k$ for some $k \in \mathbb{N}$, so that each partition is obtained by bisecting its predecessor.

Step 1. Take $n = 1\,(k = 0)$. Then

$$(e^{t\,G^2}\, e^{t\,(Q+P)})f(x) = \int_{-\infty}^{+\infty} g(S(y_1, x))\, p(t, y_1) dy_1,$$

with $g(x) = e^{Z(t,x)}\, f(R(t, x))$, where we denote

$$S(y, x) = (\sqrt{x} + \frac{\nu\, y}{2})^2, \tag{6.7}$$

$$R(t, x) = e^{t\,\beta}\, x + \frac{\alpha}{\beta}\,(e^{t\,\beta} - 1), \tag{6.8}$$

and

$$Z(t, x) = -\,r \int_0^t R(s, x)\, ds = -\frac{r}{\beta}\Big[\frac{\alpha}{\beta}\,(e^{t\,\beta} - 1)\; + (e^{t\,\beta}\; - 1)\, x - \alpha\, t\Big].$$

Then

$$(e^{t\,G^2}\, e^{t\,(Q+P)})f(x) = \int_{-\infty}^{+\infty} e^{Z(t,S(y_1,x))}\, f(R(t, S(y_1, x)))\, p(t, y_1) dy_1. \tag{6.9}$$

After some calculation, we get

$$R(t, S(y, x)) = R(t, x) + e^{t\beta}\, \nu\, y\, (\sqrt{x} + \frac{\nu\, y}{4}),$$

$$Z(t, S(y, x)) = Z(t, x) -\, r\Big[\frac{(e^{t\beta} - 1)}{\beta}\, \nu\, y\, (\sqrt{x} + \frac{\nu\, y}{4})\Big].$$

Step 2. Take $n = 2\,(k = 1)$. By applying (6.9) and replacing t by $t/2$, we get

$$\begin{aligned}(e^{\frac{t}{2}G^2}\, e^{\frac{t}{2}\,(Q+P)})^2\, f(x) &= (e^{\frac{t}{2}\,G^2}\, e^{\frac{t}{2}\,(Q+P)})\, h(x)\\ &= \int_{-\infty}^{+\infty} e^{Z(t/2,S(y_1,x))}\, h(R(t/2, S(y_1, x)))\, p(t/2, y_1) dy_1,\end{aligned}$$

where, again from (6.9)

$$h(x) = (e^{\frac{t}{2}G^2}\, e^{\frac{t}{2}\,(Q+P)})\, f(x) = \int_{-\infty}^{+\infty} e^{Z(t/2,S(y_2,x))} f(R(t/2, S(y_2, x)))p(t/2, y_2) dy_2.$$

Then

$$\begin{aligned}(e^{\frac{t}{2}G^2}\, e^{\frac{t}{2}\,(Q+P)})^2\, f(x) = &\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{Z(t/2,S(y_1,x))+Z(t/2,S(y_2,R(t/2,S(y_1,x))))}\\ &\cdot f(R(t/2, S(y_2, R(t/2, S(y_1, x)))))\prod_{i=1}^{2} p(t/2, y_i)\, dy_1\, dy_2.\end{aligned} \tag{6.10}$$

After many calculations, we get

$$R(t/2, S(y_2, R(t/2, S(y_1, x)))) = R(t, x) + e^{t\beta}\,\nu\, y_1\Big(\sqrt{x} + \frac{\nu\, y_1}{4}\Big) + e^{t\beta/2}\,\nu\, y_2\Big(\sqrt{R(t/2, S(y_1, x))} + \frac{\nu\, y_2}{4}\Big), \quad (6.11)$$

and

$$Z(t/2, S(y_1, x)) + Z(t/2, S(y_2, R(t/2, S(y_1, x)))) = Z(t, x) - r\Big[\frac{(e^{t\beta}-1)}{\beta}\nu y_1(\sqrt{x} + \frac{\nu y_1}{4}) - \frac{(e^{t\beta/2}-1)}{\beta}\nu y_2(\sqrt{R(t/2, S(y_1, x))} + \frac{\nu y_2}{4})\Big].$$

Step 3. Take $n = 4$ ($k = 2$). By applying (6.10) and replacing $t/2$ by $t/4$, we get

$$(e^{\frac{t}{4}G^2} e^{\frac{t}{4}(Q+P)})^4 f(x) = (e^{\frac{t}{4}G^2} e^{\frac{t}{4}(Q+P)})^2\, h(x) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{Z(t/4, S(y_1,x)) + Z(t/4, S(y_2, R(t/4, S(y_1,x))))} \cdot h(R(t/4, S(y_2, R(t/4, S(y_1, x))))) \prod_{i=1}^{2} p(t/4, y_i)\, dy_1\, dy_2,$$

where, again from (6.10)

$$h(x) = (e^{\frac{t}{4}G^2} e^{\frac{t}{4}(Q+P)})^2\, f(x) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{Z(t/4, S(y_3,x)) + Z(t/4, S(y_4, R(\frac{t}{4}, S(y_3,x))))} \cdot f(R(t/4, S(y_4, R(t/4, S(y_3, x))))) \prod_{i=1}^{2} p(t/4, y_i)\ dy_3\, dy_4.$$

Then

$$(e^{\frac{t}{4}G^2} e^{\frac{t}{4}(Q+P)})^4 f(x) = \underbrace{\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}}_{4 \text{ times}} e^{Z(t/4, S(y_1,x)) + Z(t/4, S(y_2, R(t/4, S(y_1,x)))) + Z(t/4, S(y_3, R(t/4, S(y_2, R(t/4, S(y_1,x))))))} \cdot e^{Z(t/4, S(y_4, R(t/4, S(y_3, R(t/4, S(y_2, R(t/4, S(y_1,x))))))))} \cdot f(R(t/4, S(y_4, R(t/4, S(y_3, R(t/4, S(y_2, R(t/4, S(y_1, x))))))))) \prod_{i=1}^{4} p(t/4, y_i) dy_1 \cdots dy_4. \quad (6.12)$$

Analogous calculations as before show that

$$R(t/4, S(y_4, R(t/4, S(y_3, R(t/4, S(y_2, R(t/4, S(y_1, x))))))))$$
$$= R(t, x) + e^{t\beta}\,\nu\, y_1\Big(\sqrt{x} + \frac{\nu\, y_1}{4}\Big) + e^{3t\beta/4}\nu\, y_2\Big(\sqrt{R(t/4, S(y_1, x))} + \frac{\nu\, y_2}{4}\Big)$$
$$+ e^{2t\beta/4}\nu\, y_3\,\Big(\sqrt{R(t/4, S(y_2, R(t/4, S(y_1, x))))} + \frac{\nu\, y_3}{4}\Big)$$
$$+ e^{t\beta/4}\nu\, y_4\,\Big(\sqrt{R(t/4, S(y_3, R(t/4, S(y_2, R(t/4, S(y_1, x))))))} + \frac{\nu\, y_4}{4}\Big), \quad (6.13)$$

and

$$
\begin{aligned}
&Z(t/4,S(y_1,x))+Z(t/4,S(y_2,R(t/4,S(y_1,x))))\\
&+Z(t/4,S(y_3,R(t/4,S(y_2,R(t/4,S(y_1,x))))))\\
&+Z(t/4,S(y_4,R(t/4,S(y_3,R(t/4,S(y_2,R(t/4,S(y_1,x))))))))\\
&=Z(t,x)-r\Big[\frac{(e^{t\beta}-1)}{\beta}\nu y_1(\sqrt{x}+\frac{\nu y_1}{4})\\
&-\frac{(e^{3t\beta/4}-1)}{\beta}\nu y_2\Big(\sqrt{R(t/4,S(y_1,x))}+\frac{\nu\, y_2}{4}\Big)\\
&-\frac{(e^{2t\beta/4}-1)}{\beta}\,\nu\, y_3\,\Big(\sqrt{R(t/4,S(y_2,R(t/4,S(y_1,x))))}+\frac{\nu\, y_3}{4}\Big)\\
&-\frac{(e^{t\beta/4}-1)}{\beta}\,\nu\, y_4\,\Big(\sqrt{R(t/4,S(y_3,R(t/4,S(y_2,R(t/4,S(y_1,x))))))}+\frac{\nu\, y_4}{4}\Big)\Big].
\end{aligned}
$$

A straightforward but rather tedious induction argument yields for $n=2^k,\ k\in\mathbb{N}$,

$$
\begin{aligned}
&R(t/n,S(y_n,R(t/n,S(y_{n-1},R(t/n,\dots,R(t/n,S(y_1,x)))))))\\
&\quad=R(t,x)+e^{t\beta}\nu y_1\Big(\sqrt{x}+\frac{\nu\, y_1}{4}\Big)+e^{t(n-1)\beta/n}\nu y_2\Big(\sqrt{R(t/n,S(y_1,x))}+\frac{\nu y_2}{4}\Big)\\
&\qquad+e^{t(n-2)\beta/n}\nu\, y_3\Big(\sqrt{R(t/n,S(y_2,R(t/n,S(y_1,x))))}+\frac{\nu\, y_3}{4}\Big)+\cdots\\
&\qquad+e^{t\beta/n}\nu\, y_n\Big(\sqrt{R(t/n,S(y_{n-1},R(t/n,\dots,R(t/n,S(y_1,x)))))}+\frac{\nu\ y_n}{4}\Big),
\end{aligned}
\tag{6.14}
$$

and

$$
\begin{aligned}
&Z(t/n,S(y_1,x))+Z(t/n,S(y_2,R(t/n,S(y_1,x))))\\
&+Z(t/n,S(y_3,R(t/n,S(y_2,R(t/n,S(y_1,x))))))+\cdots\\
&+Z(t/n,S(y_n,R(t/n,S(y_{n-1},R(t/n,\cdots,R(t/n,S(y_1,x)))))))\\
&=Z(t,x)-r\Big[\frac{(e^{t\beta}-1)}{\beta}\,\nu\, y_1\,(\sqrt{x}+\frac{\nu\, y_1}{4})\\
&\quad-\frac{(e^{t(n-1)\beta/n}-1)}{\beta}\,\nu\, y_2\,\Big(\sqrt{R(t/n,S(y_1,x))}+\frac{\nu\, y_2}{4}\Big)\\
&-\frac{(e^{t(n-2)t\beta/n}-1)}{\beta}\,\nu\, y_3\,\Big(\sqrt{R(t/n,S(y_2,R(t/n,S(y_1,x))))}+\frac{\nu\, y_3}{4}\Big)-\cdots\\
&-\frac{(e^{t\beta/n}-1)}{\beta}\,\nu\, y_n\,\Big(\sqrt{R(t/n,S(y_{n-1},R(t/n,\dots,R(t/n,S(y_1,x)))))}+\frac{\nu\, y_n}{4}\Big)\Big].
\end{aligned}
$$

Let

$$
\begin{aligned}
&L(t,n,\nu,\{y_j\}_{1\le j\le n},x,f)=e^{Z(t/n,S(y_1,x))+Z(t/n,S(y_2,R(t/n,S(y_1,x))))}\\
&\cdot e^{Z(t/n,S(y_3,R(t/n,S(y_2,R(t/n,S(y_1,x))))))+\cdots+Z(t/n,S(y_n,R(t/n,S(y_{n-1},R(t/n,\cdots,R(t/n,S(y_1,x)))))))}\\
&\qquad\cdot f(R(t/n,S(y_n,R(t/n,S(y_{n-1},R(t/n,\dots,R(t/n,S(y_1,x)))))))).
\end{aligned}
\tag{6.15}
$$

In particular, the (6.15) can be rewritten as

$$
L(t,n,\nu,\{y_j\}_{1\le j\le n},x,f)=e^{Z_n(t/n,x,y_1,\dots,y_n)}\cdot f(R_n(t/n,x,y_1,...,y_n)),
\tag{6.16}
$$

where

$$R_1(t,x,y_1) = e^{t\beta x} + \frac{\alpha}{\beta}(e^{t\beta} - 1) + e^{t\beta}\nu y_1 S(y_1, x),$$

$$Z_1(t,x,y_1) = -\frac{1}{\beta}\Big((e^{t\beta} - 1)\Big(x + \frac{\alpha}{\beta}\Big) - \alpha t\Big) - \frac{(e^{t\beta} - 1)}{\beta}\nu y_1 S(y_1, x),$$

the functions R_i, Z_i are recursively defined for $i \in \{2, ..., n\}$ as follows

$$\begin{aligned} &R_i(t/i, x, \{y_h\}_{1\le h\le i}) \\ &\quad = R_1(t,x,y_1) + \sum_{j=2}^{i} e^{t\beta(n-j+1)/n}\nu y_j S(y_j, R_{j-1}(t/i, x, \{y_h\}_{1\le h\le (j-1)})), \end{aligned} \tag{6.17}$$

$$\begin{aligned} &Z_i(t/i, x, \{y_h\}_{1\le h\le i}) \\ &= Z_1(t,x,y_1) - \sum_{j=2}^{i} \frac{e^{t\beta(n-j+1)/n} - 1}{\beta}\nu y_j S(y_j, R_{j-1}(t/i, x, \{y_h\}_{1\le h\le (j-1)})). \end{aligned} \tag{6.18}$$

It is evident that the form (6.16) is more convenient for numerical implementations. Finally

$$\begin{aligned} u_n(t,x) &= (e^{\frac{t}{n}G^2} e^{\frac{t}{n}(Q+P)})^n f(x) \\ &= \underbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}_{n\ times} L(t, n, \nu, \{y_j\}_{1\le j\le n}, x, f) \cdot \prod_{j=1}^{n} p\Big(\frac{t}{n}, y_j\Big)\, dy_1 \cdots dy_n, \end{aligned} \tag{6.19}$$

for $n = 2^k$.

6.3 The Case $c = 0$

We refer to Theorem 2.2.8 and consider the (C_0) semigroups U and V defined in (2.40) and (2.41), respectively. Fix $g \in C(\overline{\mathbb{R}}), t > 0, z \in \mathbb{R}$.

Step 1. Take $n = 1$ $(k = 0)$. We have

$$U(t)V(t)g(z) = \int_{-\infty}^{+\infty} g(ze^{\gamma t} + \xi y_1)p(t, y_1)dy_1. \tag{6.20}$$

Step 2: Take $n = 2$ $(k = 1)$. Then

$$\begin{aligned} [U(t/2)V(t/2)]^2 g(z) &= U(t/2)V(t/2)h(z) \\ &= \int_{-\infty}^{+\infty} h(ze^{\gamma t/2} + \xi y_1)p(t/2, y_1)dy_1 \end{aligned}$$

where

$$h(z) = U(t/2)V(t/2)g(z) = \int_{-\infty}^{+\infty} g(ze^{\gamma t/2} + \xi y_2)p(t/2, y_2)dy_2.$$

Hence,

$$
\begin{aligned}
&[U(t/2)V(t/2)]^2\, g(z) \\
&\quad = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g((ze^{\gamma t/2}+\xi y_1)e^{\gamma t/2}+\xi y_2)\prod_{i=1}^{2} p(t/2,y_i)dy_1dy_2 \\
&\quad = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g((ze^{\gamma t}+\xi(e^{\gamma t/2}y_1+y_2))\prod_{i=1}^{2} p(t/2,y_i)dy_1dy_2.
\end{aligned} \tag{6.21}
$$

By induction, we can conclude that, for $n = 2^k,\ k \in \mathbb{N}$,

$$
\begin{aligned}
u_n(z,t) &= [U(t/n)V(t/n)]^n\, g(z) \\
&= \underbrace{\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}}_{n\,times} L_0(t,n,\{y_j\}_{1\le j\le n},z,g)\prod_{i=1}^{n} p\left(\frac{t}{n},y_i\right)dy_1...dy_n,
\end{aligned} \tag{6.22}
$$

where

$$
L_0(t,n,\{y_j\}_{1\le j\le n},z,g) = g\left(ze^{\gamma t}+\xi\sum_{j=1}^{n} y_j e^{\gamma(n-j)t/n}\right).
$$

6.4 The Case $c \neq 0$

We refer to Theorem 2.2.15 and consider the (C_0) semigroups U and V defined in (2.62) and (2.63), respectively.
Assume $c > 0$ and consider $J = (-\frac{d}{c}, +\infty)$. Denote $\beta = c/\sqrt{2}$ and

$$
R(t,x) = e^{t\beta}x + \frac{d}{c}(e^{t\beta}-1). \tag{6.23}
$$

Observe that for any $x_0 \in J,\ R(t,x_0) \in J$. Indeed,

$$
R(t,x_0) = \underbrace{e^{t\beta}\left(x_0+\frac{d}{c}\right)}_{>0} + \left(-\frac{d}{c}\right) > -\frac{d}{c}.
$$

To compute the approximate functions $u_n(\cdot,\cdot), n \geq 1$, given in (2.57) we proceed by steps. Fix $g \in C(\bar{J})$.
Step 1: Take $n = 1$ ($k = 0$). Thus

$$
U(t)V(t)g(x) = \int_{-\infty}^{+\infty} g(R(y_1+\zeta t, x+\alpha_2 t/c))p(t,y_1)dy_1, \tag{6.24}
$$

where

$$
R(y_1+\zeta t,\, x+\alpha_2 t/c) = e^{\beta(y_1+\zeta t)}\left(x+\frac{\alpha_2}{c}t+\frac{d}{c}\right)-\frac{d}{c}. \tag{6.25}
$$

Step 2: Take $n = 2$, $(k = 1)$. By applying (6.25) and replacing t by $t/2$, we get

$$\Big[U(t/2)V(t/2)\Big]^2 g(x) = U(t/2)V(t/2)\, h(x)$$
$$= \int_{-\infty}^{+\infty} h(R(y_1 + \zeta t/2,\ x + \alpha_2\, t/2\, c))p(t/2, y_1)dy_1,$$

where

$$h(x) = U(t/2)V(t/2)\, g(x) = \int_{-\infty}^{+\infty} g(R(y_2 + \zeta t/2,\ x + \alpha_2\, t/2\, c))p(t/2, y_2)dy_2.$$

Hence,

$$\Big[U(t/2)V(t/2)\Big]^2 g(x) =$$
$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(R(y_2+\zeta t/2,\ R(y_1+\zeta t/2,\ x+\alpha_2\, t/2\, c)+\alpha_2\, t/2\, c)\prod_{i=1}^{2} p(t/2, y_i)dy_1 dy_2, \tag{6.26}$$

where

$$R(y_2 + \zeta t/2,\ R(y_1 + \zeta t/2,\ x + \alpha_2\, t/2\, c) + \alpha_2\, t/2\, c)$$
$$= e^{\beta(y_2+\zeta t/2)}\left(e^{\beta(y_1+\zeta t/2)}\left(x + \frac{\alpha_2 t}{2c} + \frac{d}{c}\right) - \frac{d}{c} + \frac{\alpha_2 t}{2c} + \frac{d}{c}\right) - \frac{d}{c}$$
$$= e^{\beta(y_1+y_2+\zeta t)}\left(x + \frac{\alpha_2 t}{2c} + \frac{d}{c}\right) + \frac{\alpha_2 t}{2c} e^{\beta(y_2+\zeta t/2)} - \frac{d}{c}. \tag{6.27}$$

Step 3: Take $n = 4$, $(k = 2)$. Analogously

$$\Big[U(t/4)V(t/4)\Big]^4 g(x) = \Big[U(t/4)V(t/4)\Big]^2 h(x),$$

that is given by (6.26) replacing $t/2$ by $t/4$ and g by h. Moreover, h is also given by (6.26) replacing $t/2$ by $t/4$ and y_1, y_2 by y_3, y_4. So, after many calculations, we can write

$$\Big[U(t/4)V(t/4)\Big]^4 g(x) =$$
$$\underbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}_{4\,times} g(R(y_4 + \zeta t/4, R(y_3 + \zeta t/4,\ R(y_2 + \zeta t/4,$$
$$R(y_1 + \zeta t/4,\ x + \alpha_2\, t/4\, c) + \alpha_2\, t/4\, c) + \alpha_2\, t/4\, c) + \alpha_2\, t/4\, c)\prod_{i=1}^{4} p(t/4, y_i)dy_i,$$

with

$$
\begin{aligned}
&R(y_4+\zeta t/4,\ R(y_3+\zeta t/4,\ R(y_2+\zeta t/4,\ R(y_1+\zeta t/4,\ x+\alpha_2\, t/4\, c)+\\
&\qquad\qquad\qquad\qquad\qquad\qquad + \alpha_2\, t/4\, c)+\alpha_2\, t/4\, c)+\alpha_2\, t/4\, c)\\
&\quad = e^{\beta(\sum_{i=1}^{4} y_i+\zeta t)}\left(x+\frac{d}{c}\right)-\frac{d}{c}+\frac{\alpha_2 t}{4c}\Big[e^{\beta(\sum_{i=1}^{4} y_i+\zeta t)}+e^{\beta(\sum_{i=2}^{4} y_i+3\zeta t/4)}\\
&\qquad\qquad\qquad\qquad\qquad\qquad + e^{\beta(\sum_{i=3}^{4} y_i+\zeta t/2)}+e^{\beta(y_4+\zeta t/4)}\Big]\\
&\quad = R(\sum_{i=1}^{4} y_i+\zeta t,\ x)+\frac{\alpha_2 t}{4c}\sum_{j=1}^{4} e^{\left[\beta\left(\sum_{i=j}^{4} y_i+(\frac{4-j+1}{4})\zeta t\right)\right]}.
\end{aligned}
$$

By induction, we can conclude that, for $n=2^k,\ k\in\mathbb{N}$,

$$
\Big[U(t/n)V(t/n)\Big]^n g(x)=
\underbrace{\int_0^{+\infty}\dots\int_0^{+\infty}}_{n\, times} L(t,n,\{y_j\}_{1\le j\le n},x,g)\ \prod_{j=1}^{n} p\left(\frac{t}{n},y_j\right)dy_1\dots dy_n,
$$

where

$$
L(t,n,\{y_j\}_{1\le j\le n},x,g)=g\left(R(\sum_{i=1}^{n} y_i+\zeta t,\ x)+\frac{\alpha_2 t}{nc}\sum_{j=1}^{n} e^{\left[\beta\left(\sum_{i=j}^{n} y_i+(\frac{n-j+1}{n})\zeta t\right)\right]}\right).
$$

6.5 Qualitative analysis related to Table 4.4

We report the qualitative statistical analysis carried out for each group/sub group according to the results reported in Table 4.4. The qualitative analysis related to the results in Table 4.6 is analogous and so it is omitted.

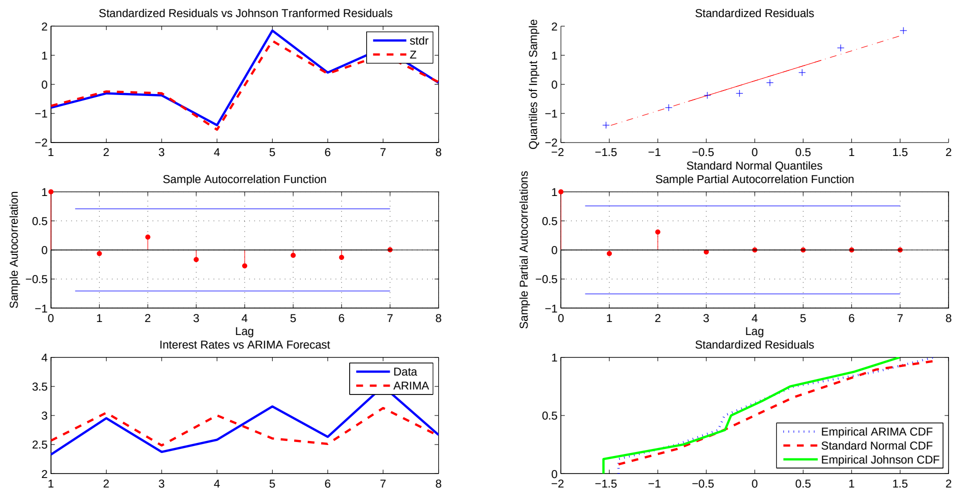


Figure 6.1. Qualitative statistical analysis related to the sub-group 1–8. **Top line:** ARIMA $(2, 0, 1)$ standardized residuals versus Johnson's transformed residuals (*left*); Q-Q normal plot for the ARIMA $(2, 0, 1)$ standardized residuals (*right*). **Middle line:** AC (*left*) and PAC (*right*) plots. **Bottom line:** real interest rates versus ARIMA $(2, 0, 1)$ fitted values (*left*); comparison of the standard normal cumulative distribution function (CDF) with the empirical CDF of ARIMA $(2, 0, 1)$ standardized residuals and of Johnson's transformed residuals (*right*).

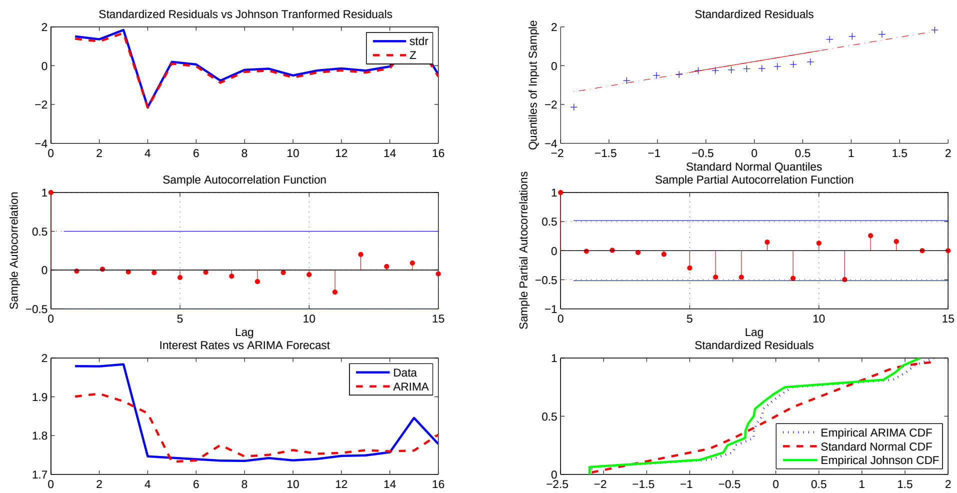


Figure 6.2. Qualitative statistical analysis related to the sub-group 17–32. **Top line:** ARIMA $(1, 0, 3)$ standardized residuals versus Johnson's transformed residuals (*left*); Q-Q normal plot for the ARIMA $(1, 0, 3)$ standardized residuals (*right*). **Middle line:** AC (*left*) and PAC (*right*) plots. **Bottom line:** real interest rates versus ARIMA $(1, 0, 3)$ fitted values (*left*); comparison of the standard normal cumulative distribution function (CDF) with the empirical CDF of ARIMA $(1, 0, 3)$ standardized residuals and of Johnson's transformed residuals (*right*).

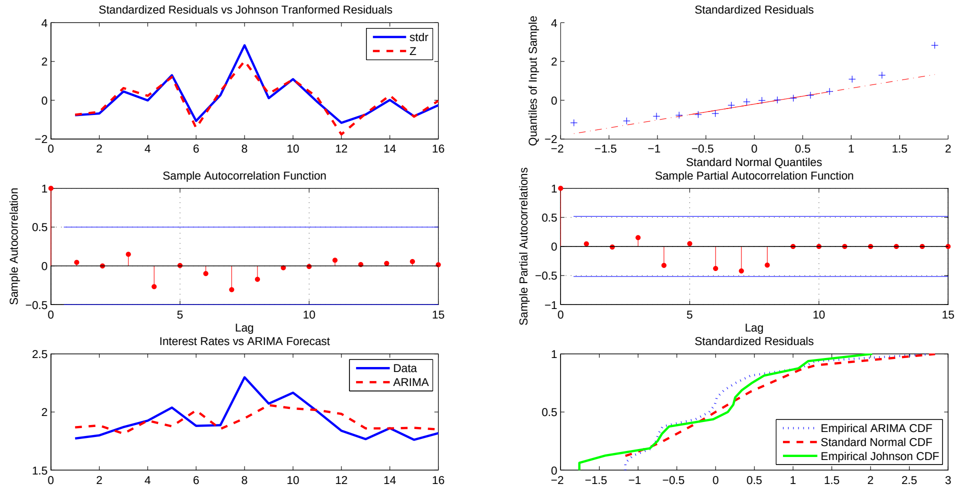


Figure 6.3. Qualitative statistical analysis related to the sub-group 33–48. **Top line:** ARIMA $(3, 0, 1)$ standardized residuals versus Johnson's transformed residuals (*left*); Q-Q normal plot for the ARIMA $(3, 0, 1)$ standardized residuals (*right*). **Middle line:** AC (*left*) and PAC (*right*) plots. **Bottom line:** real interest rates versus ARIMA $(3, 0, 1)$ fitted values (*left*); comparison of the standard normal cumulative distribution function (CDF) with the empirical CDF of ARIMA $(3, 0, 1)$ standardized residuals and of Johnson's transformed residuals (*right*).

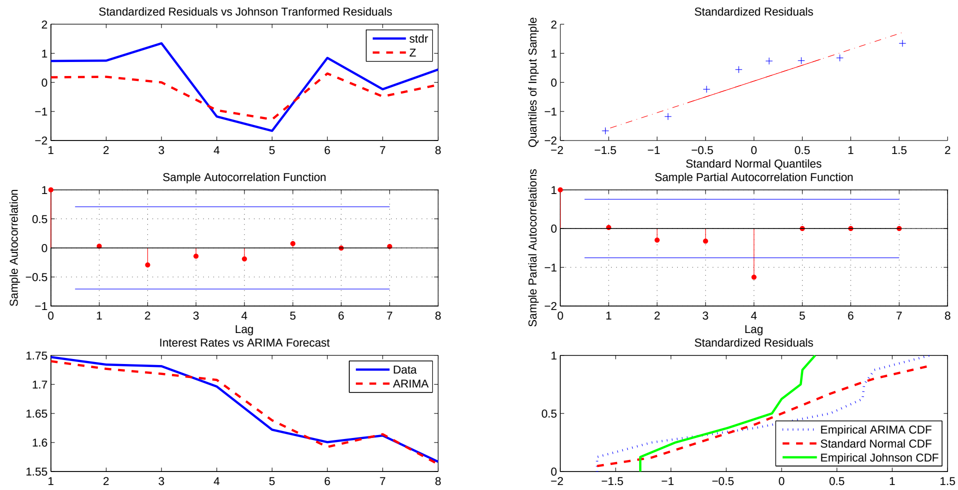


Figure 6.4. Qualitative statistical analysis related to the sub-group 49–56. **Top line:** ARIMA $(3, 1, 2)$ standardized residuals versus Johnson's transformed residuals (*left*); Q-Q normal plot for the ARIMA $(3, 1, 2)$ standardized residuals (*right*). **Middle line:** AC (*left*) and PAC (*right*) plots. **Bottom line:** real interest rates versus ARIMA $(3, 1, 2)$ fitted values (*left*); comparison of the standard normal cumulative function (CDF) with the empirical CDF of ARIMA $(3, 1, 2)$ standardized residuals and of Johnson's transformed residuals (*right*).

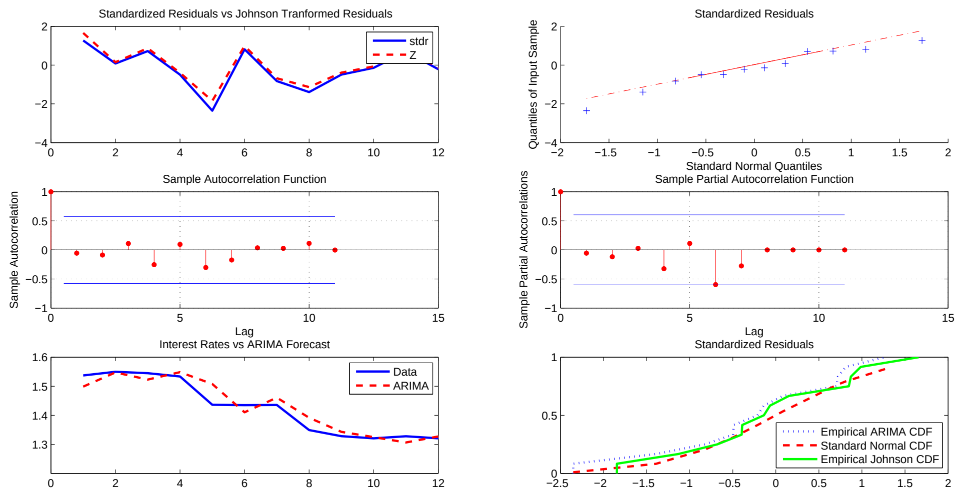


Figure 6.5. Qualitative statistical analysis related to the sub-group 57–68. **Top line:** ARIMA $(2, 1, 1)$ standardized residuals versus Johnson's transformed residuals (*left*); Q-Q normal plot for the ARIMA $(2, 1, 1)$ standardized residuals (*right*). **Middle line:** AC (*left*) and PAC (*right*) plots. **Bottom line:** real interest rates versus ARIMA $(2, 1, 1)$ fitted values (*left*); comparison of the standard normal cumulative distribution function (CDF) with the empirical CDF of ARIMA $(2, 1, 1)$ standardized residuals and of Johnson's transformed residuals (*right*).

6.6 CIR# parameter estimates

We report the estimates of the CIR# parameters k, θ, σ and, in particular, the plots of the function $S_j(k)$ defined in (4.3), corresponding to the selected "optimal" ARIMA models reported in Table 4.4.

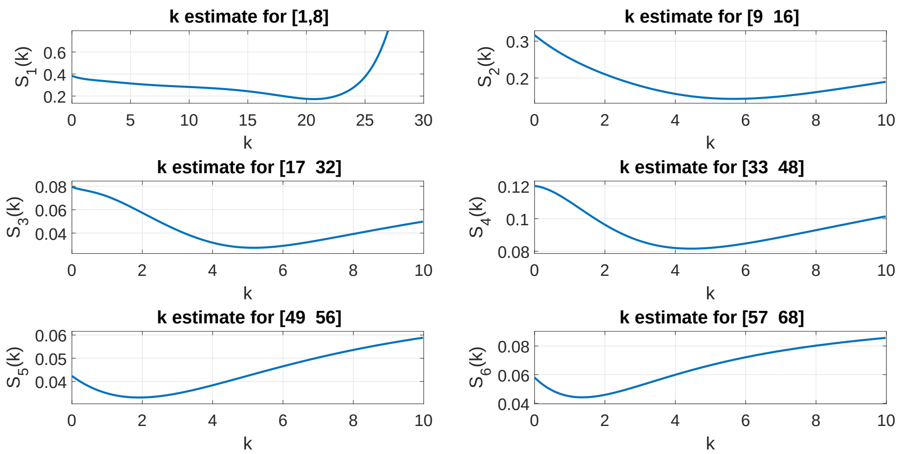


Figure 6.6. Plots of the functions $S_j(k)$ for each group/sub-group

Table 6.1. CIR# parameter estimates based on 68 monthly observed 1-day (overnight) EUR interest rates

j	group/sub-group	$\hat{k}_j$	$\hat{\theta}_j$	$\hat{\sigma}_j$
1	1–8	20.6364	2.7699	0.4027
2	9–16	5.6621	2.3338	0.3663
3	17–32	5.1649	1.7924	0.0954
4	33–48	4.4462	1.9223	0.1546
5	49–56	1.9092	1.6637	0.0709
6	57–68	1.3555	1.4264	0.0958

Ringraziamenti

Giunto al termine di questa meravigliosa avventura, durata tre anni, vorrei ringraziare tutte quelle persone senza delle quali non sarebbe stato possibile fare tutto ciò. Inizio ringraziando il Prof. Raimondo Manca, mio tutor presso la Sapienza, che mi ha da subito apprezzato (pur non conoscendomi bene), dandomi una grande autonomia nel gestire la mia ricerca. Ringrazio poi la Prof.ssa Rosa Maria Mininni che mi ha dato una grande possibilità, offrendosi come mio supervisor e avendo creduto nelle mie potenzialità. Lei è stata la mia guida in tutto questo percorso, mostrandosi sempre disponibile, donandomi tutti i Suoi preziosi insegnamenti e il Suo aiuto. E' stata Lei a presentarmi il Prof. Giuseppe Orlando, nostro coautore, nonchè risk manager la cui esperienza sul campo mi è stata fondamentale per capire tanti aspetti anche pratici della finanza. La Prof.ssa Mininni ha poi favorito la collaborazione con la Prof.ssa Silvia Romanelli, la cui esperienza mi ha fatto crescere molto e per questo è stato un grande onore lavorare con Lei. Infine voglio ringraziare tutti gli altri professori, ricercatori e colleghi che ho incontrato in vari eventi quali corsi, seminari, convegni e scuole estive, da ognuno dei quali ho preso interessanti spunti. In particolare rivolgo un caro abbraccio al mio compagno di dottorato, Fabrizio, che mi ha offerto sempre una spalla nei momenti più opportuni. Infine ringrazio la mia famiglia, onnipresente in ogni momento, bello o meno della mia vita.
Concludo, augurandomi di non terminare mai questo nobile lavoro che è la ricerca scientifica e di guardare sempre avanti a nuove sfide...
Grazie.

Bibliography

[1] Y. Achdou and O. Pironneau, Computational Methods for Option Pricing, *SIAM Frontiers in Applied Mathematics* vol. 30, Philadelphia, PA (2005);

[2] R. P. Adams and D. J. MacKay, Bayesian Online Changepoint Detection, arXiv:0710.3742 (2007);

[3] S. Arlot and A. Celisse, Segmentation of the Mean of Heteroscedastic Data Via Cross-Validation, *Statistics and Computing* 21(4), 613-632 (2011);

[4] A. Attalienti and I. Rasa, Shape-Preserving Properties and Asymptotic Behaviour of the Semigroup Generated by the black-Scholes Operator, *Czechoslovak Mathematical Journal* 58, 457-467 (2008);

[5] A. Attalienti, R.M. Mininni and I. Rasa, Gamma-Type Operators and the Black-Scholes Semigroup, *Journal of Approximation Theory* 163(2), 163-182 (2011);

[6] J. Bai and P. Perron, Computation and Analysis of Multiple Structural Change Models, *Journal of Applied Econometrics* 18(1), 1-22 (2003);

[7] E. Barucci, F. Gozzi and V. Vespri, A Semigroup Approach to No-Arbitrage Pricing Theory, in book *Seminar on Stochastic Analysis, Random Fields and Applications*, Progress in Probability, 45, Birkhauser, Basel (1999);

[8] A. Batkai, P. Csomos, B. Farkas and G. Nickel, Operator Splitting for Non-Autonomous Evolution Equations, *Journal of Functional Analysis* 260, 2163-2190 (2010);

[9] R. Beneder and T. Vorst, Option on Dividend Paying Stocks, in book *Recent Developments in Mathematical Finance*, Shanghai, China, World Scientific (2001);

[10] B. M. Bibby, M. Jacobsen and M. Sørensen, (2010). Estimating Functions for Discretely Sampled Diffusion Type Models. In Y. Ait-Sahalia and L. P. Hansen (Eds.), *Handbook of Financial Econometrics* 1, 203-268 (2010);

[11] BIS, Is the Unthinkable Becoming Routine?, Technical Report, Bank for International Settlements (2015);

[12] F. Black and M. Scholes, The Pricing of Options and Corporate Liabilities, *Journal of Political Economy* 81, 637-654 (1973);

[13] L. A. Bordag and R. Frey, Nonlinear Option Pricing Models for Illiquid Markets: Scaling Properties and Explicit Solutions, book: *Nonlinear Models in Mathematical Finance. New Research Trends in Option Pricing*, Nova Science Publischer, Inc., 103-130 (2009);

[14] M. J. Brennan and E. S. Schwartz, Savings Bonds, Retractable Bonds and Callable Bonds, *Journal of Financial Economics* 5, 67-88 (1977);

[15] D. Brigo and F. Mercurio, A Deterministic-Shift Extension of Analytically-Tractable and Time-Homogeneous Short Rate Models, *Finance and Stochastics* 5, 369-388 (2001);

[16] D. Brigo and N. El-Bachir, Credit Derivatives Pricing with a Smile-Extended Jump Stochastic Intensity Model, Available at SSRN 950208, doi: 10.2139/ssrn.950208 (2007);

[17] D. Brigo and F. Mercurio, *Interest Rate Models-Theory and Practice. With Smile, Inflation and Credit*, Springer-Verlag, Berlin, Heidelberg (2006);

[18] M. Bufalo, R.M. Mininni and S. Romanelli, A Semigroup Approach to Generalized Black-Scholes Type Equation in Incomplete Markets, preprint (2018), submitted;

[19] B. Böttcher, R. Schilling and J. Wang (2013) A Primer on Feller Semigroups and Feller Processes, in: *Lévy Matters III, Lecture Notes in Mathematics* vol. 2099. Springer, Cham (2013);

[20] R. A. Carmona and M. R. Tehranchi, *Interest Rate Models: an Infnite Dimensional Stochastic Analysis Perspective*, Springer-Verlag, Berlin, Heidelberg (2006);

[21] S. Cerrai, Analytic Semigroups and Degenerate Elliptic Operators with Unbounded Coefficients: a Probabilistic Approach, *Journal of Dierential Equations* 166, 151-174 (2000);

[22] L. Chen, *Stochastic Mean and Stochastic Volatility - A Three-Factor Model of the Term Structure of Interest Rates and Its Applications in Derivatives Pricing and Risk Management*, Financial Markets, Institutions and Instruments 5, 1-88 (1996);

[23] W. Cheng, N. Costanzino, J. Liechty, A. Mazzucato and V. Nistor, Closed-Form Asymptotics and Numerical Approximations of 1D Parabolic Equations with Applications to Option Pricing, *SIAM Journal on Financial Mathematics* 2, 901-934 (2011);

[24] P. R. Chernoff, Note on Product Formulas for Operator Semigroups, *Journal of Functional Analysis* 2, 238-242 (1968);

[25] J. C. Cox, J. E. Ingersoll and S. A. Ross, A Theory of the Term Structure of Interest Rates, *Econometrica* 53, 385-407 (1985);

[26] G. Courtadon, The Pricing of Options on Default-Free Bonds, *Journal of Financial and Quantitative Analysis* 17, 75-100 (1982);

[27] S. R. Das and S. Foresi Exact Solutions for Bond and Option Prices with Systematic Jump Risk, *Review of Derivatives Research* 1, 7-24 (1996);

[28] M. H. Davis, V. Panas and T. Zariphopoulou, European Option Pricing with Transaction Costs, *SIAM Journal of Control and Optimization* 31, 470-493 (1993);

[29] D. Duffie, *Dynamic Asset Pricing Theory*, third ed., Princeton University Press (2001);

[30] D. Duffie, D. Filipoviȁ and W. Schachermayer, Affine Processes and Applications in Finance, *Annals of Applied Probability* 13(3), 984-1053 (2003);

[31] H. Emamirad, G. R. Goldstein and J.A. Goldstein, Chaotic Solution for the Black-Sholes Equation, *Proceedings of the American Mathematical Society* 142, 4385-4386 (2014);

[32] K. J. Engel, and R. Nagel, *One-Parameter Semigroup for Linear Evolution Equations*, Springer-Verlag, New York (2000);

[33] K. C. Engelen, The Unthinkable as the New Normal, *The International Economy*, 29(3) (2015);

[34] W. Feller, *An Introduction to Probability Theory and Its Applications*, John Weley & Sons (1971);

[35] O. S. Farad, Linearization and Nonlinear Stochastic Differential Equations with Locally Lipschitz Condition, *Applied Mathematical Sciences* 1(52), 2553-2563 (2007);

[36] F. Fasola, V. M. R. Muggeo and H. Küchenhoff, Heuristic, Iterative Algorithm for Change-Point Detection in Abrupt Change Models, *Computational Statistics*, DOI 10.1007/s00180-017-0740-4 (2017);

[37] M. Frittelli, Introduction to a Theory of Value Coherent with the No-Arbitrage Principle, *Finance and Stochastics* 4, 275-297 (2000);

[38] M. Giuli, F. Gozzi, R. Monte and V. Vespri, Generation of Analytic Semigroup and Domain Characterization for Degenerate Elliptic Operators with Unbounded Coefficients Arising in Financial Mathematics, part II, in H. Amann, W. Arendt, M. Hieber, F. M. Neubrander, S. Nicaise, and J. von Below (Eds.), *Functional Analysis and Evolution Equations: the Gunter Lumer Volume*, 315-330, Birkhauser Verlag, Basel/Switzerland (2007);

[39] P. Glasserman, Monte Carlo Methods in Financial Engineering, Springer Series: *Stochastic Modelling and Applied Probability* 53, Springer, New York (2013);

[40] G. R. Goldstein, J. A. Goldstein, R. M. Mininni and S. Romanelli, The Semigroup Governing the Generalized Cox-Ingersoll-Ross Equation, *Advances in Differential Equations* 21, 235-264 (2016);

[41] J. A. Goldstein, R. M. Mininni and S. Romanelli, Markov Semigroup and Estimating Functions, with Applications to Some Financial Models, *Communications on Stochastic Analysis* 1, 381-391 (2007);

[42] J. A. Goldstein, R. M. Mininni and S. Romanelli, Markov Semigroup and Groups of Operators, *Communications on Stochastic Analysis* 1, 247-262 (2007);

[43] J. A. Goldstein, R. M. Mininni and S. Romanelli, A New Explicit Formula for the Solution of the Black-Merton-Scholes Equation, in *Infinite Dimensional Stochastic Analysis*. In honor of Hui-Hsiung Kuo, Subseries: QP-PQ: Quantum Probability and White Noise Analysis 22 (eds. A. N. Sengupta and P. Sundar), World Scientic, 226-235 (2008);

[44] J. A. Goldstein, *Semigroup of Linear Operators and Applications*, Oxford University Press (1985);

[45] V. Gorovoi and V. Linetsky, Black's Model of Interest Rates as Options, Eigenfunction Expansions and Japanese Interest Rates, *Mathematical Finance* 1, 49-78 (2004);

[46] F. Gozzi, R. Monte and V. Vespri, Generation of Analytic Semigroup and Domain Characterization for Degenerate Elliptic Operators with Unbounded Coefficients Arising in Financial Mathematics, part I, *Differential and Integral Equations* 15(9), 1085-1128 (2002);

[47] T. Graser, Operator Multipliers Generating Strongly Continuous Semigroups, *Semigroup Forum* 55, 68-79 (1997);

[48] R. S. Hacker and A. Hatemi-J, Tests for Causality between Integrated Variables Using Asymptotic and Bootstrap Distributions: Theory and Application, *Applied Economics* 38(13), 1489-1500 (2006);

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

[50] J. C. Hull and A. White, Pricing Interest-Rate-Derivative Securities, *Review of Financial Studies* 3(4), 573-592 (1990);

[51] J. C. Hull, *Options, Futures, and other Derivatives*, 8th ed. Prentice Hall, Boston (2012);

[52] IBA. Data Vendor Codes. ICE Benchmark Administration (IBA) https://www.theice.com/publicdocs/IBA_Quote_Vendor_Codes.xlsx;

[53] T. Jahnke and C. Lubich, Error Bounds for Exponential Operator Splittings, *BIT* 40(4), 735-744 (2000);

[54] M. Jeanblanc, M. Yor and M. Chesney, *Mathematical Methods for Financial Markets*, Springer, New York (2004);

[55] N. L. Johnson, Systems of Frequency Curves Generated by Methods of Translation, *Biometrika* 36, 149-176 (1949);

[56] D. L. Jones, The Johnson Curve Toolbox for Matlab: Analysis of Non-Normal Data Using the Johnson System of Distributions, College of Marine Science, University of South Florida, St. Petersburg, Florida, USA (2014), http://it.mathworks.com/matlabcentral/fileexchange/46123-johnson-curve-toolbox;

[57] C. Kahl and P. Jäckel, Fast Strong Approximation Monte Carlo Schemes for Stochastic Volatility Models, *Quantitative Finance* 6, 513-536 (2006);

[58] S. Karlin and H. M. Taylor, *A Second Course in Stochastic Processes*, Academic Press, Inc. New York (1981);

[59] M. Keller-Ressel and T. Steiner, Yield Curve Shapes and the Asymptotic Short Rate Distribution in Affine One-Factor Models, *Finance and Stochastics* 12(2), 149-172 (2008);

[60] K. Kladıvko, Maximum Likelihood Estimation of the Cox-Ingersoll-Ross Process: the Matlab Implementation, Technical Computing Prague, (2007);

[61] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag: Berlin, Heidelberg (1992);

[62] T. O. Kvalseth, Cautionary Note about R^2, *The American Statistician* 39, 279-285 (1985);

[63] Y. Kwok, *Mathematical Models of Financial Derivatives*, Springer Finance (2008);

[64] M. Lavielle, Using Penalized Contrasts for the Change-Point Problem, *Signal Processing* 85, 1501-1510 (2005);

[65] M. Lavielle and G. Teyssiere, Detection of Multiple Change-Points in Multivariate Time Series, *Lithuanian Mathematical Journal* 46, 287-306 (2006);

[66] A. Lunardi, M. Miranda and D. Pallara, Lecture Notes 19th Internet Seminar 2015/2016: Infinite Dimensional Analysis (2016), http://www.math.kit.edu/iana3/seite/isem/;

[67] R. C. Merton, Lifetime Portfolio Selection under Uncertainty: the Continuous Time Model, *Review of Economic Studies* 51, 247-257 (1969);

[68] R. C. Merton, Theory of Rational Option Pricing, *Journal of Economy and Management Sciences* 4, 141-183 (1973);

[69] G. Milshtein, A Method of Second-Order Accuracy Integration of Stochastic Differential Equations, *Theory of Probability and Its Applications* 23(2), 396-401 (1979);

[70] M. Moreno and F. Platania, A Cyclical Square-Root Model for the Term Structure of Interest Rates, *European Journal of Operational Research* 241(1), 109-121 (2015);

[71] M. Musiela and T. Zariphopoulou, An Example of Indifference Price under Exponential Preferences, *Finance and Stochastics* 8, 229-239 (2004);

[72] M. Musiela and T. Zariphopoulou, Indifference Prices and Related Measures, Technical Report, The University of Texas at Austin (2001) ;

[73] A. R. Najafi and F. Mehrdoust, Bond Pricing Under Mixed Generalized CIR Model with Mixed Wishart Volatility, *Journal of Computational and Applied Mathematics* 319, 108-116 (2017);

[74] A. R. Najafi, F. Mehrdoust and S. Shirinpour, Pricing American Put Option on Zero-Coupon Bond under Fractional CIR Model with Transaction Cost, *Communications in Statistics - Simulation and Computation* 0(0), 1-7 (2017);

[75] B. Øksendal, *Stochastic Differential Equations and Applications: an Introduction with Applications*, 6th edition, Springer, Berlin (2003);

[76] G. Orlando, R. M. Mininni and M. Bufalo, A New Approach to CIR Short-Term Rates Modelling. In M. Mili, F. di Pietro and R. Samaniego Medina (Eds.), *New Methods in Fixed Income Modeling*, Springer International (USA) (2018);

[77] G. Orlando, R. M. Mininni, M. Bufalo, A Revised Approach of the CIR model to forecast short-term interest rates: the CIR# model (2018); working paper;

[78] G. Orlando, R. M. Mininni and M. Bufalo, On the forecast of expected short-term interest rates in the CIR model (2018); submitted;

[79] G. Orlando and G. Taglialatela, A Review on Implied Volatility Calculation, *Journal of Computational and Applied Mathematics* 320, 202-220 (2017);

[80] A. Papageorgiou and J. F. Traub, Faster Evaluation of Multidimensional Integrals, *Computers in Physics*, 1997, 574-578 (2000);

[81] M. Poletti Laurini and L. K. Hotta, Generalized Moment Estimation of Stochastic Differential Equations, *Computational Statistics* 31, 1169-1202 (2016);

[82] K. Park and S. Kim, On Interest Rate Option Pricing with Jump Processes, *International Journal of Engineering and Applied Sciences* 2, 64-67 (2015);

[83] M. R. Rodrigo and R. S. Mamon, Recovery of Time-Dependent Parameters of a Black-Scholes Type Equation: an Inverse Stieltjes Approach, *Journal of Applied Mathematics*, 1-8 (2007);

[84] L. C. G. Rogers, Gaussian errors, *Risk* 9(1), 42-45 (1996);

[85] R. Rouge and N. El Karoui, Pricing via Utility Maximization and Entropy, *Mathematical Finance* 10, 259-276 (2000);

[86] R. Schnaubelt, Lecture Notes Summer Semester 2015: Spectral Theory, http://www.math.kit.edu/iana3/ schnaubelt/media/st-skript15.pdf;

[87] A. N. Shiryaev, *Essential of Stochastic Finance: facts, models, theory*, World Scientific Publishing Co.Pte.Ltd., Singapore (1999);

[88] A. N. Shiryaev, *Stochastic Calculus for Finance II: Continuous-Time models*, Springer-Verlag New York (2004);

[89] Lars E. O. Svensson, Nontraded Assets in Incomplete Markets: Pricing and Portfolio Choice, *European Economic Review* 37, 1149-1168 (1993);

[90] K. Taira, *Diffusion Processes and Partial Differential Equations*, Acaemic Press, San Diego (1988);

[91] K. M Tamizhmani and K. Krishnakumar, Algebraic Resolution of Equations of the Black-Scholes Type with Arbitrary Time-Dependent Parameters, PGL Leach, *Applied Mathematics and Computation* 247, 115-124 (2014);

[92] O. Vasicek, An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics* 5(2), 177-188 (1977);

[93] L. Zhu, (2014) Limit Theorems for a Cox-Ingersoll-Ross Process with Hawkes Jumps, *Journal of Applied Probability* 51(3), 699-712 (2014);