Univerzita Karlova v Praze
Matematicko – fyzikální fakulta

DIPLOMOVÁ PRÁCE

Petr Zahradník

Jednorozměrné difusní stochastické diferenciální rovnice s aplikacemi ve finanční matematice

Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: prof. RNDr. Josef Štěpán, DrSc.

Studijní program: Matematika, obor pravděpodobnost, matematická statistika a ekonometrie

2010

Děkuji vedoucímu své práce, prof. Josefu Štěpánovi, za zajímavé téma a trpělivé vedení této práce. Děkuji všem svým blízkým za vřelou a trvalou podporu během celého studia.

Prohlašuji, že jsem svou diplomovou práci napsal samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce a jejím zveřejňováním.

V Praze dne 5.8.2010 Petr Zahradník

Contents

Název: Jednorozměrné difusní stochastické diferenciální rovnice s aplikacemi ve finanční matematice
Autor: Petr Zahradník
Katedra (ústav): Katedra pravděpodobnosti a matematické statistiky
Vedoucí diplomové práce: prof. RNDr. Josef Štěpán, DrSc.
e-mail vedoucího práce: stepan@karlin.mff.cuni.cz

Abstrakt: Předmětem této práce je využití pokročilých metod teorie pravděpodobnosti a částečně i matematické analýzy na určité partie finanční matematiky. V první kapitole jsou shrnuty potřebné poznatky z teorie pravděpodobnosti. V druhé kapitole jsou postupně zmíněny základy teorie jednorozměrných difusních stochastických diferenciálních rovnic. Jsou zformulovány potřebné výsledky ohledně existence a jednoznačnosti řešení i ve slabém smyslu, je zkonstruováno řešení Engelbertovy – Schmidtovy rovnice a je důkladně zkoumán Fellerův test exploze. Třetí kapitola se zabývá Dirichletovým problémem a jeho aplikací na oceňování finančních opcí včetně implementace. Poslední, čtvrtá, kapitola je určena využití znalostí z předchozích částí textu k odvození některých zajímavých vlastností Coxova – Ingersollova – Rossova modelu.

Klíčová slova: Dirichletův problém, finanční opce, CIR proces.

Title: One – dimensional diffusion stochastic differential equations with applications to financial mathematics
Author: Petr Zahradník
Department: Department of Probability and Mathematical Statistics
Supervisor: prof. RNDr. Josef Štěpán, DrSc.
Supervisor's e-mail address: stepan@karlin.mff.cuni.cz

Abstract: In this thesis, the aim is to employ some of the advanced probability and calculus techniques to financial mathematics. In the first chapter some major facts from continuous – time probability theory are presented. In the second chapter, one – dimensional stochastic differential equations are introduced, we touch upon the questions of existence and uniqueness of solutions in full generality, construct a weak solution to the Engelbert – Schmidt equation and thoroughly present a known procedure called a Feller's test for explosions. In chapter three, focus is directed to a brief presentation of the well known Dirichlet problem. The problem is also interpreted financially, applied to options valuation and related approximations are implemented. The fourth, final, chapter concentrates on the Cox – Ingersoll – Ross model. Techniques derived in the second and third chapters are employed to thoroughly study the model properties.

Keywords: Dirichlet problem, financial options, CIR process.

Chapter 1

Introductory Findings

1.1 Introduction

In this thesis, the aim is to employ some of the advanced probability and calculus techniques to financial mathematics. In the first chapter, continuous – time stochastic processes are treated and some major facts from continuous – time stochastic calculus are presented in the versions which are used in the chapters to follow. In the second chapter, one – dimensional stochastic differential equations are introduced, we touch upon the questions of existence and uniqueness of solutions in full generality, construct a weak solution to the Engelbert – Schmidt equation and thoroughly present a known procedure called a Feller's test for explosions. This test connects a behaviour of some solutions to deterministic problems to an explosion property of related stochastic problems. In chapter three, focus is directed to the well known Dirichlet problem and it is shown how this problem can be financially interpreted and profited. This allows an introduction of financial options and a derivation of some interesting conclusions. Two arising pricing applications are also established, employing numerical and simulation techniques. The fourth, final, chapter concentrates on the Cox – Ingersoll – Ross model. Techniques derived in the second and third chapters are employed to thoroughly study the model properties. The results presented in the fourth chapter, or at least their majority are known, but to the best of the author's knowledge are not presented or treated explicitly in the literature published.

A reference for the first chapter is [8], where all results are thoroughly explained and proved. The reader is assumed to be familiar with a few of the basic notions and concepts in probability theory. To make the thesis self – contained in terms of the mathematical language, the main concepts are briefly reviewed.

1.2 Elementary Probability Review

We herein work with a *σ-field* $\mathcal{F}$, a *measurable space* $(\Omega,\mathcal{F})$, a *filtration* $(\mathcal{F}_t)_{t\geq 0}$, a *probability* P and a triple (a quadruple, respectively) $(\Omega,\mathcal{F},P)$ (or $\{\Omega,\mathcal{F},(\mathcal{F}_t),P\}$) called a *(filtered) probability space.*

We assume that our probability spaces are *complete* and respective filtrations meet the *usual conditions.* A mapping X from Ω to $\mathbf{R}$ which is $\mathcal{F}/\mathcal{B}(\mathbf{R})$ – *measurable* is called a (real – valued) *random variable.* A random variable induces a measure $P^X(B) = P[X \in B]$, $B \in \mathcal{B}(\mathbf{R})$, which is called a *probability distribution.*

Given a sub – σ–*field* $\mathcal{F}_0 \subset \mathcal{F}$, there exists a (up to a null set) unique *conditional expectation* of X relative to $\mathcal{F}_0$, which is denoted by $E(X \mid \mathcal{F}_0)$.

A family $\{X_t, t \geq 0\}$ of random variables defined on a filtered probability space is called a *stochastic process.* We always require the process, referred to as X_t, to be *adapted.* For our purposes however, it is natural to assume our processes to further meet somewhat stricter conditions – be *progressively measurable*:

Definition 1.1 (Progressive Measurability)**.** A stochastic process defined on $(\Omega,\mathcal{F},(\mathcal{F}_t),P)$ is said to be $\mathcal{F}_t$ – *progressively measurable*, if for every $A \in \mathcal{B}(\mathbf{R})$:

$$\{(s,\omega) : s \in [0,t], \omega \in \Omega, X_s(\omega) \in A \ \} \in \mathcal{B}([0,t]) \otimes \mathcal{F}_t$$

It can be proved that a *cadlag* (right – continuous with left limits) process is progressively measurable. It should be noted, that when we talk about continuity of a process in this text, we always mean continuity of the process' sample paths.

A mathematical description of outcomes of a fair game is called a *martingale*:

Definition 1.2 (Martingale)**.** An adapted stochastic process defined on $(\Omega,\mathcal{F},(\mathcal{F}_t),P)$ is called an $\mathcal{F}_t$ – *martingale* if it is integrable (i.e. $E|X_t| < \infty \quad \forall t \geq 0$) and

$$E[X_t|\mathcal{F}_s] = X_s \quad \text{a.s.} \quad \forall 0 \leq s \leq t.$$

A precise formulation of a random time at which we choose to take an action based on a history of the game is represented by a notion of a *stopping time*:

Definition 1.3 (Stopping Time)**.** A nonnegative random variable τ is said to be an $\mathcal{F}_t$ – *stopping time* if

$$\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t \quad \forall t \geq 0.$$

There is an important and intuitively clear concept of information obtained before a stochastic time τ:

Definition 1.4 (pre–τ σ–algebra)**.** Let τ be an $\mathcal{F}_t$ – stopping time. We call

$$\mathcal{F}_\tau = \{F \in \mathcal{F} : F \cap [\tau \leq t] \in \mathcal{F}_t\}$$

a *pre–τ σ–algebra.*

To extend the martingale property from well – behaved (especially in the sense of integrability) to more general processes we use a method called *localization.*

Definition 1.5 (Local Martingale)**.** An adapted process X_t is a *local martingale* in L^p if there exists a sequence of stopping times (*a localization sequence*) $0 \leq \tau_1 \leq \tau_2 \leq \ldots$ with $\tau_n \uparrow \infty$ a.s., such that the stopped process $X_{t \wedge \tau_n}$ is a L^p – martingale for all n.

In every measure theory course, the dominated convergence theorem plays a central role. When talking about martingales, domination is not the best concept, *uniform integrability* is. It is more general and accurate as it provides us with convergence in L^1, see [20], pp. 314–315, and allows us easily change the order of limit and integration.

Definition 1.6 (Uniform Integrability)**.** A stochastic process $\{X_t, t \geq 0\}$ is *uniformly integrable* if it satisfies:

$$\lim_{M\to\infty} \sup_{t\geq 0} E|X_t|\mathbf{1}_{[|X_t|>M]} = 0,$$

where $\mathbf{1}_{[.]}$ is the indicator function.

A very important theorem states that an agent playing a fair game cannot bias his outcomes from the game by choosing when to quit without cheating; his decision being based on a history of the game only:

Theorem 1.7 (Optional Sampling)**.** *Let X_t be a continuous uniformly integrable martingale. Let τ and ν be $\mathcal{F}_t$ – stopping times with $\tau \leq \nu$. Then*

$$X_\nu, X_\tau \in L_1, \qquad E[X_\nu|\mathcal{F}_\tau] = X_\tau \quad a.s.$$

To mathematically describe an irregular motion of pollen grains in liquid, which was observed by Robert Brown in the beginning of 19^{th} century and now is widely used as a theoretical concept in many mathematical applications, we need the following properties:

Definition 1.8 (Brownian Motion)**.** A stochastic process X_t with continuous *sample paths* starting at 0 is a *Brownian motion*, if it has *independent increments* and

$\mathcal{L}(X_t - X_s) = N(0, t-s) \quad \forall 0 \leq s < t$, i.e. a *Gaussian* process with mean zero and variance $(t-s)$.

The mathematical concept of Brownian motion is often referred to as a *Wiener process* in honour of Norbert Wiener, who introduced much of the measure theoretic concepts related to Brownian motion. Brownian motion is not only a special stochastic process with good properties. It is actually, in a certain sense, the only real continuous local martingale which matters. To see statements making such a proposition precise, refer to [16], chapter 4, section 3.

1.3 Stochastic Calculus

With the elementary terms in mind, let us proceed to slightly more advanced parts, specifically those used throughout the thesis. The following theorem is actually a corollary to a much more powerful Doob – Meyer decomposition theorem.

Theorem 1.9 (Doob – Meyer)**.** *Let X_t be a continuous local martingale. Then there exists a unique adapted nondecreasing continuous process $\langle X\rangle_t$ such that $\langle X\rangle_0 = 0$ and $X_t^2 - \langle X\rangle_t$ is a continuous local martingale.*

Definition 1.10 (Quadratic Variation)**.** If X_t is a local martingale with continuous sample paths, then the unique adapted process from Theorem 1.9 denoted by $\langle X\rangle_t$ is called a *quadratic variation* process of a local martingale.

It turns out that the quadratic variation process is a limit in probability of a sequence of processes possessing an important intuitive meaning:

$$\langle X\rangle_t = P - \lim_{n\to\infty} \sum_{k=0}^{n-1} (X_{t^n_{k+1}} - X_{t^n_k})^2, \quad \text{where} \quad 0 = t^n_0 < t^n_1 < \ldots < t^n_n = t, |t^n_{k+1} - t^n_k| \to 0.$$

A *quadratic covariation* of two processes X, Y is a process

$$\langle X, Y\rangle_t = \frac{1}{4}(\langle X+Y\rangle_t - \langle X-Y\rangle_t).$$

We will also often use the following proposition, which is again rather a corollary to the fact, that we can distinguish true martingales from local martingales by their quadratic variation:

Theorem 1.11. *Let M_t be a continuous local martingale, $M_0 = 0$ and let τ be an $\mathcal{F}_t$– stopping time. If $E\langle M\rangle_\tau < \infty$ then $M_{t\wedge\tau}$ and $M^2_{t\wedge\tau} - \langle M\rangle_{t\wedge\tau}$ are uniformly integrable martingales.*

It is time to proceed to a definition of a crucial concept. Up to now we were able to integrate with respect to sufficiently smooth integrators, precisely integrators with bounded *variation*. Here we want to use continuous local martingales as integrators and because the only a.s. nonconstant (i.e. interesting as integrators) continuous local martingales are of unbounded variation, a new tool is needed.

Definition 1.12 (Stochastic Integral)**.** Let X_t be a real continuous local martingale and Ψ_t a progressively measurable process such that

$$\int_0^t \Psi_s^2 \mathrm{d}\langle X\rangle_s < \infty \quad \text{a.s. } \forall t \geq 0.$$

Then we can define a process $I_X(\Psi)_t$, $I_X(\Psi)_0 = 0$ called a *stochastic integral*, such that for every continuous local martingale Y_t:

$$\langle I_X(\Psi), Y\rangle_t = \int_0^t \Psi_s \mathrm{d}\langle X, Y\rangle_s; \quad \text{moreover,} \quad \langle I_X(\Psi)\rangle_t = \int_0^t \Psi_s^2 \mathrm{d}\langle X\rangle_s.$$

We denote such a process $I_X(\Psi)_t$ by $\int_0^t \Psi_s \mathrm{d}X_s$.

Now, when we know what stochastic integral is, let us mention a straightforward generalization of the chain rule known from elementary calculus:

Theorem 1.13 (Stochastic Chain Rule)**.** *Let M_t be a continuous local martingale and G_t, H_t square integrable progressively measurable processes. Let $N_t = \int_0^t G_s \mathrm{d}M_s$. Then $G_t H_t$ is a square integrable progressively measurable process and:*

$$\int_0^t H_s \mathrm{d}N_s = \int_0^t H_s G_s \mathrm{d}M_s.$$

However, that was only an ouverture to what is sometimes called a stochastic version of the analytical chain rule too. It is a dominant theorem in stochastic calculus that provides us with a tool to actually compute stochastic integrals. To formulate it, we let a process X_t be a *continuous semimartingale*, that is, there exist a continuous local martingale M_t, $M_0 = 0$, and a continuous process of finite variation V_t, $V_0 = 0$, such that $X_t = X_0 + V_t + M_t$. Given it exists, by Doob Meyer decomposition theorem such a decomposition is unique. We know how to integrate with respect to continuous semimartingales: by the decomposition, it breaks down to two integrals we are well familiar with. Therefore, it is time to formulate the following pervasive theorem.

Theorem 1.14 (Itô Formula)**.** *Let $f \in C^2(\mathbf{R}^d)$ and X_t be a continuous semimartingale $\forall t \geq 0$ taking values in $\mathbf{R}^d$. Then the process $f(X_t)$ is a continuous semimartingale $\forall t \geq 0$ and it holds that:*

$$f(X_t) = f(X_0) + \int_0^t \sum_{i=1}^d \frac{\partial}{\partial x^i} f(X_s) \mathrm{d}X_s^i + \frac{1}{2}\sum_{i=1}^d \sum_{j=1}^d \int_0^t \frac{\partial^2}{\partial x^i \partial x^j} f(X_s) \mathrm{d}\langle X^i, X^j\rangle_s$$

Remark 1.15. We will come to a point when we want to consider a certain one – dimensional process X_t and a function of its graph, say $f(Y_t)$, where $Y_t = [t, X_t]^T$, $f \in C^2(\mathbf{R}^2)$ and ^{T} stands for "transpose". Itô formula says the following:

$$f(t, X_t) = f(Y_t) = f(Y_0) + \int_0^t \sum_{i=1}^{2} \frac{\partial}{\partial y^i} f(Y_s) \mathrm{d}Y_s^i + \frac{1}{2} \sum_{i=1}^{2} \sum_{j=1}^{2} \int_0^t \frac{\partial^2}{\partial y^i \partial y^j} f(Y_s) \mathrm{d}\langle Y^i, Y^j \rangle_s$$
$$= f(0, X_0) + \int_0^t \frac{\partial}{\partial s} f(s, X_s) \mathrm{d}s + \int_0^t \frac{\partial}{\partial x} f(s, X_s) \mathrm{d}X_s + \frac{1}{2} \int_0^t \frac{\partial^2}{\partial x \partial x} f(s, X_s) \mathrm{d}\langle X \rangle_s. \quad (1.1)$$

As was already envisaged, there exist very useful theorems in literature providing us with representations of continuous local martingales via Brownian motion. We will not be using these explicitly, however, the first cornerstone to prove such facts should be due to its general usefulness mentioned. It is the well known Lévy theorem, which has the following one dimensional formulation:

Theorem 1.16 (Lévy)**.** *Let X_t be a real continuous stochastic process, $X_0 = 0$. Then X_t is a Brownian motion if and only if X_t and $X_t^2 - t$ are continuous martingales.*

The forthcoming famous theorem due to Cameron, Martin and Girsanov is important in the general theory of stochastic processes. It states the key result that if Q is a measure absolutely continuous with respect to P, then every P – semimartingale is a Q – semimartingale. This plays a crucial role in infinite – dimensional analysis. Here, we present a specific formulation in one dimension:

Theorem 1.17 (Girsanov)**.** *Let B_t be an $\mathcal{F}_t$– Brownian motion on $\{\Omega, \mathcal{F}, (\mathcal{F}_t), \mathrm{P}\}$. Let X_t be a progressively measurable process and fix T such that*

$$P\left(\int_0^t X_s \mathrm{d}s < \infty\right) = 1 \qquad \forall 0 \le t \le T.$$

Define

$$G_t = \epsilon\left(\int X \mathrm{d}B\right)_t,$$

where $\epsilon\left(\int X\mathrm{d}B\right)_t$ is the stochastic exponential (or Doléans exponential) of X_t with respect to B_t, i.e.

$$G_t = \exp\left\{\int_0^t X_s \mathrm{d}B_s - \frac{1}{2}\int_0^t X^2 \mathrm{d}s\right\}, \qquad 0 \le t \le T,$$

and suppose that $EG_T = 1$. Let Q be a probability measure on $\mathcal{F}_T$ with density (Radon – Nikodým derivative) G_T, that is, let $\mathrm{d}Q = G_T \mathrm{d}P$. Define

$$B_t^Q = B_t - \int_0^t X_s \mathrm{d}s, \qquad 0 \le t \le T.$$

Then B_t^Q is an $\mathcal{F}_t$ – Brownian motion on the original probability space equipped with measure Q.

Remark 1.18. It is important to note why Girsanov theorem also plays one of the central roles in financial mathematics. It shows how to convert from the physical measure P, which describes the probability that an underlying instrument (such as a share price or an interest rate) will take a particular value, to the risk – neutral measure Q. That means a very useful tool for evaluating the value of derivatives on the underlying. An approach using the measure Q, usually called *risk – neutral pricing*, often offers a very simple and elegant way to prove even results seemingly demanding a technical, difficult approach. For a striking example, see an alternative proof of the Nobel price winning Black – Scholes formula in [25] pp. 218–220.

The proof of the following theorem relies on a nice technique using a *number of upcrossings*. It can be found in [23], p.176.

Theorem 1.19 (Doob's Supermartingale Convergence Theorem). *Let X_t be a cadlag supermartingale. Suppose further that $EX_t \leq K < \infty \quad \forall t$. Then*

$$\lim_{t\to\infty} X_t \quad \textit{exists finite almost surely.}$$

1.4 A Note on Financial Time Series

We conclude the introductory chapter with the following note.

It is tempting to say, from the *Central limit theorem* vaguely at least, that once we have a lot of independent agents in the financial system, their common behaviour should result in *Gaussian* properties of the financial time series. Also, because *Normal* distribution has very tractable and analytically feasible properties, Gaussian processes have played a central role even in the modern stochastic finance, the Black – Scholes formula being the front – runner. Unfortunately, many empirical studies have shown that the Gaussian distribution does not fit the financial returns series data very well. One of the most painful examples was the *Gaussian copula* function, which has been blindly overused in estimating risk and pricing complicated baskets of assets, or the widespread Value – At – Risk measure, which makes a good sense only when the silent assumption of normality is made. The recent credit – crisis in USA again showed that extremal returns are much more likely than a Gaussian distribution would suggest. It followed as matter of fact, that Gaussian copula even entered the main – stream media in USA as "The formula that killed Wall street", see [24]. Therefore, even though in this thesis Gaussian distribution appears trough out too, in real world modeling a greater care should be taken and often distributions with heavier tails should be chosen. Recent theoretical financial literature demonstrates such efforts, there is a renewed interest concentrated on jump processes, which were originally proposed already in 1976 by Cox and Ross in [7], or an *NIG* distribution is applied, for example. In practice, from the author's experience at an American bank it seems that a simple approach using *a Student distribution* as well as GARCH models are very popular.

Chapter 2

One – Dimensional Diffusion Stochastic Differential Equations

2.1 General Properties

We will consider an equation of the following type:

$$\mathrm{d}X_t = \mu(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t; \quad X_0 = x. \tag{2.1}$$

Such an equation is called a *One – Dimensional Diffusion Stochastic Differential Equation* and should be interpreted as follows: Given $x \in \mathbf{R}$ and Borel – measurable functions μ, σ: $\mathbf{R} \to \mathbf{R}$, we have the following stochastic integral equation:

$$X_t^x = x + \int_0^t \mu(X_s^x)\mathrm{d}s + \int_0^t \sigma(X_s^x)\mathrm{d}B_s \quad \forall t \geq 0. \tag{2.2}$$

The convergence of integrals in (2.2) is ensured if

$$\int_0^t (|\mu(X_s^x)| + |\sigma(X_s^x)|^2)\mathrm{d}s < \infty \quad \text{a.s.}$$

In this chapter, we partially treat the questions of existence and uniqueness of processes satisfying (2.1), referred to as *solutions*, study their properties and show some methods to find them. All stochastic differential equations (and respective processes as their solutions) in this chapter are assumed to be one – dimensional without further notice. Note, that when the coefficients μ and σ are Lipschitz continuous, it is a common practice to call processes satisfying (2.1) *Itô diffusions*. In this text, when we talk about diffusions we mean Itô diffusions. Also, when there can't be any confusion and should it help readability, we often drop the suffix ^{x} in X_t^x.

Remark 2.1. Some authors define a process to be a *diffusion* when it meets conditions concerning the first and second infinitesimal moments:

$$E(X_{t+h} - X_t|\mathcal{F}_t) = \mu(X_t)h + o(h) \quad \text{a.s.},$$
$$\mathrm{var}(X_{t+h} - X_t|\mathcal{F}_t) = \sigma^2(X_t)h + o(h) \quad \text{a.s. when } h \downarrow 0.$$

These conditions give an important interpretation to the coefficients μ and σ and can be extended to (2.1).

Remark 2.2. When postulating equation (2.1), we assumed the coefficients to depend on the state x only. The general time – inhomogeneous case can be reduced to our situation by formally considering a space – time process, see remark 3.1 or for example [20], p. 220.

Solutions to (2.1) may not exist for all times t – in that case we can find ourselves in two situations. The first is usually called an *explosion*, when the solution tends to infinity. The second, in the case of coefficients μ and σ defined on a subinterval of a real line only, could be referred to as an *exit* (trough a finite boundary). To be able to treat these cases together, we define the coefficients μ and σ rather generally on an interval

$$I = (l, r); \qquad -\infty \leq l < r \leq \infty,$$

and formally introduce a notion of an *exit time* e:

Definition 2.3 (Exit time)**.** Let $I = (l, r), -\infty \leq l < r \leq \infty$, be an interval. By an *exit time* of a solution X_t^x we mean a random variable

$$e^X = \lim_{n\to\infty} \tau_n; \quad \tau_n = \inf\{t; X_t^x \notin [l_n, r_n]\}; \qquad (2.3)$$
$$l < l_n < r_n < r; \quad l_n \downarrow l, \quad r_n \uparrow r.$$

It needs only a little verification that the limit does not depend on a specific choice of $\{l_n\}$ and $\{r_n\}$. Again, wherever possible, we drop the suffix ^{X} in e^X.

When the solution stays in I forever, i.e. $e = \infty$ a.s., we call the endpoints of I *unattainable*. When $I = \mathbf{R}$ and $e = \infty$ a.s., it is common to call such a solution *global*.

A question of solvability of (2.1) is of course fundamental. There are two concepts related to the existence of a solution and either one has a well suited concept of uniqueness attached. A *weak solution* offers much more general assumptions on the drift and volatility coefficients and, as stressed out in [16], p.300, its uniqueness mode leads naturally to a *strong Markov property*.

Definition 2.4 (Weak Solution)**.** A triple $[(\Omega,\mathcal{F},P),\{X_t^x\},\{\mathcal{G}_t, B_t\}]$ is called a *weak solution* (*up to an exit time* e) on an interval $I = (l, r)$ of a diffusion equation (2.1), when $\mathcal{G}_t$ is a filtration, B_t a $\mathcal{G}_t$ – Brownian motion and the following holds:

$$X_t^x \text{ is a continuous, progressively measurable, } [l, r] - \text{valued process; } X_0^x = x; \quad x \in I, \qquad (2.4)$$

and with τ_n defined above in definition 2.3. Note that trough out the thesis $\wedge$ and $\vee$ signs mean minimum and maximum, respectively. We have:

$$\int_0^{t\wedge\tau_n} (|\mu(X_s^x)| + |\sigma(X_s^x)|^2)\mathrm{d}s < \infty \quad \text{a.s. } \forall t > 0, \quad \forall n \geq 1, \qquad (2.5)$$

$$X_{t\wedge\tau_n}^x = x + \int_0^{t\wedge\tau_n} \mu(X_s^x)\mathrm{d}s + \int_0^{t\wedge\tau_n} \sigma(X_s^x)\mathrm{d}B_s \quad \forall t > 0 \quad \text{a.s. } \forall n \geq 1. \qquad (2.6)$$

Definition 2.5 (Strong Solution)**.** Given $\{\Omega, \mathcal{F}, (\mathcal{F}_t), P\}$ and an $\mathcal{F}_t$ – Brownian motion B_t, a *strong solution* (*up to an exit time* e) on an interval $I = (l, r)$ of (2.1) is a process X_t^x satisfying (2.4),(2.5) and (2.6).

When there can't be any misunderstanding, even a weak solution may be denoted by X_t^x. We only have to bear in mind that in such a case, the filtration and Brownian motion representations are not fixed but come as a part of the solution.

Definition 2.6 (Uniqueness in Law)**.** We say that a *uniqueness in law* (*up to an exit time*) for solutions of (2.1) holds, if whenever $[\{\Omega, \mathcal{F}, (\mathcal{F}_t), P\}, \{X_t^x\}, B]$ and $[\{\Omega, \mathcal{F}, (\widetilde{\mathcal{F}}_t), P\}, \{Y_t^x\}, \widetilde{B}]$ are two weak solutions, the laws of the processes X_t and Y_t coincide.

When the uniqueness in law holds, we may also say that a solution is *weakly unique*.

Definition 2.7 (Pathwise Uniqueness)**.** We say that a *pathwise uniqueness* (*up to an exit time*) for solutions of (2.1) holds, if whenever $[\{\Omega, \mathcal{F}, (\mathcal{F}_t), P\}, \{X_t^x\}, \mathrm{B}]$ and $[\{\Omega, \mathcal{F}, (\mathcal{F}_t), P\}, \{Y_t^x\}, \mathrm{B}]$ are two solutions defined on the same filtered probability space with the same Brownian motion, then $e^X = e^Y$ a.s. and:

$$P\left\{\omega \in \Omega; X_t^x(\omega) = Y_t^x(\omega) \quad \forall t \in [0, e^X(\omega))\right\} = 1.$$

At this instant, we state two results about existence and uniqueness which are suited to our problems. In dimension one, the theory concerning weak solutions is developed even much further, see for example [16], section 5.5., pp.329–342.

Theorem 2.8 (Yamada, Watanabe)**.** *Let $I = \mathbf{R}$ and μ and σ be continuous. Suppose that there exist a constant $C > 0$ and a strictly increasing function h: $\mathbf{R}_+ \to \mathbf{R}_+$ with $\int_0^{0+\epsilon} h^{-2}(x)\mathrm{d}x = \infty$ such that $\forall x, y \in I$:*

$$\begin{aligned} |\mu(x) - \mu(y)| &\leq C|x - y| \\ |\sigma(x) - \sigma(y)| &\leq h(|x - y|) \\ |\mu(x)| + |\sigma(x)| &\leq C(1 + |x|). \end{aligned} \tag{2.7}$$

Then global strong existence and pathwise uniqueness hold.

Proof. We shall combine several results in [13] together. Assumption (2.7) serves to prove that $E[X_t]^2 < \infty$ and hence a.s. $e = \infty$ as in the proof of Theorem IV.2.4., pp. 164–165. Pathwise uniqueness is proved in Theorem IV.3.2., pp. 168–169. Theorem IV.2.3. says that a weak solution exists. A part of Theorem IV.1.1. on the other hand says that pathwise uniqueness and weak existence imply strong existence. □

Theorem 2.9 (Krylov)**.** *Let $I = \mathbf{R}$ and μ and σ be Borel – measurable and bounded. Suppose there exists a constant ϵ such that*

$$|\sigma(x)| \geq \epsilon > 0 \quad \forall x \in I.$$

Then global weak existence and uniqueness in law hold.

Proof. A general multi – dimensional version of this theorem is proved in [17]. □

In the preceding chapter, we mentioned the famous Girsanov theorem 1.17, which allows us to study equations without drift whose solutions are (local) martingales. There is another possibility to remove drift, this time without changing the probability measure, as we can see in the following important Example 2.10:

Example 2.10 (Scale Function). Let X_t be such that the coefficients μ and σ meet the assumptions of Theorem 2.9. Suppose $s(x) \in C^2(\mathbf{R})$ and set $Y_t = s(X_t)$. Then Itô formula yields:

$$dY_t = \left(\frac{1}{2}\sigma^2(X_t)s''(X_t) + \mu(X_t)s'(X_t)\right)\mathrm{d}t + \sigma(X_t)s'(X_t)\mathrm{d}B_t.$$

For Y_t to be a local martingale, we need the dt term to disappear, therefore we need:

$$Ls(x) \equiv \frac{1}{2}\sigma^2(x)s''(x) + \mu(x)s'(x) = 0, \tag{2.8}$$

where "≡" stands for "define". A solution of such a second order ordinary differential equation, assuming $\sigma^{-2}\mu$ locally integrable, after a little computation stands as follows:

$$s(x) = \int_c^x \exp\left\{-\int_c^y \frac{2\mu(u)}{\sigma(u)^2}\mathrm{d}u\right\}\mathrm{d}y. \tag{2.9}$$

The function s, which is unique up to adding (or multiplying with) a constant, is called a *scale function*. There are good reasons for this name as we shall see further. Now, s is strictly increasing and strictly positive and as such has an inverse function. By Itô formula 1.14 and a Chain rule 1.13, we can verify that X_t solves (2.1) if and only if $s(X_t)$ solves the *Engelbert – Schmidt* equation:

$$\mathrm{d}Y_t = g(Y_t)\mathrm{d}B_t; \quad g(y) = \sigma(s^{-1}(y))s'(s^{-1}(y)), \qquad Y_0 = s(x). \tag{2.10}$$

This equation has a weakly unique weak solution under very mild conditions, see for example the Krylov theorem 2.9.

Engelbert – Schmidt type equations are of the most simple and best known equations. There is an elegant way how to solve such equations – by directly constructing a solution via a random time change as shown in the following Example 2.11. This example is inspired by Lemma V.28.7 in [23], p.179. A part of the flow of our thoughts is really only a replication of a proof of the well known Dambis – Dubins – Schwarz (*DDS*) theorem in [16], pp. 174–175.

Example 2.11 (Solution by Random Time Change). We are constructing a weak solution of the following equation.

$$\mathrm{d}X_t = \sigma(X_t)\mathrm{d}B_t; \qquad X_0 = 0. \tag{2.11}$$

where σ is a continuous function such that it meets the assumptions of the Krylov theorem 2.9. From that theorem we have a weak existence of a solution that is unique in law.

First, fix a stochastic basis $(\Omega,\mathcal{F},P)$ with a filtration $\mathcal{F}_t$ and a Brownian motion Y_t. Define a stochastic process

$$X_t = \int_0^t \sigma(Y_s)^{-1}\mathrm{d}Y_s; \qquad X_0 = 0.$$

Because $E\left[\int_0^t \sigma(Y_s)^{-2}\mathrm{d}s\right]$ is finite for every $t \geq 0$, X_t is a square integrable continuous $\mathcal{F}_t$ – martingale. Since

$$\langle X\rangle_t = \int_0^t \sigma(Y_s)^{-2}\mathrm{d}s,$$

it is clear that, almost surely, $\langle X\rangle_t$ is a strictly increasing continuous function of t and that $\langle X\rangle_\infty = \infty$. Define

$$\tau_t = \inf\{s > 0 : \langle X\rangle_s > t\}. \tag{2.12}$$

The random variable τ_t is actually due to the continuity of $\langle X\rangle_s$ an $\mathcal{F}_t$ – stopping time, which is continuous and strictly increasing. Hence we can define an inverse. (The inverse is correctly defined. We are not in the general case, as in the afore remembered *DDS* proof, where a need of a pseudo – inverse arises):

$$\tau_t^{-1} = \langle X\rangle_{\tau_t} = t.$$

Now, let us fix $0 \leq s_1 < s_2$ and define

$$\begin{aligned} \tilde{L}_t &= X_{\tau_{s_2} \wedge t} \qquad \tilde{R}_t = X^2_{\tau_{s_2} \wedge t} - \langle X \rangle_{\tau_{s_2} \wedge t}, \\ L_t &= X_{\tau_t} \qquad R_t = X^2_{\tau_t} - \langle X \rangle_{\tau_t} = L_t^2 - t, \end{aligned}$$

and a pre–τ_t σ–algebra $\mathcal{F}_{\tau_t}$ as in definition 1.4. Since our original filtration $\mathcal{F}_t$ meets the usual conditions, so does $\mathcal{F}_{\tau_t}$. Also, because $\tau_0 = 0$, it follows that $\mathcal{F}_{\tau_0} = \mathcal{F}_0$. Due to Theorem 1.11, we know that $\tilde{L}_t$ and $\tilde{R}_t$ are uniformly integrable martingales. Using the Optional Sampling theorem 1.7, we immediately get

$$\begin{aligned} E(L_{s_2} - L_{s_1} | \mathcal{F}_{\tau_{s_1}}) &= E(X_{\tau_{s_2}} - X_{\tau_{s_1}} | \mathcal{F}_{\tau_{s_1}}) = E(\tilde{L}_{s_2} - \tilde{L}_{s_1} | \mathcal{F}_{\tau_{s_1}}) \\ &= 0; \qquad \text{and similarly,} \\ E(R_{s_2} - s_2 | \mathcal{F}_{\tau_{s_1}}) &= 0. \end{aligned}$$

Therefore, we see that L_t and R_t are $\mathcal{F}_{\tau_t}$ – martingales. If we were able to show that L_t had continuous sample paths, above proved results would according to Lévy theorem 1.16 imply, that L_t was an $\mathcal{F}_{\tau_t}$ – Brownian motion. Indeed, continuity follows from the fact that both X_t and τ_t have continuous sample paths. If the latter weren't true, a clear reasoning for why the continuity of L_t sample paths would still hold could be found in the proof of the proposition IV.1.13 in [22], p.126. Further, take the probability space $(\Omega,\mathcal{F},P)$ equipped with the new filtration $\mathcal{F}_{\tau_t}$ and the $\mathcal{F}_{\tau_t}$ – Brownian motion X_{τ_t}. We claim:

$$[(\Omega, \mathcal{F}, P), \{Y_{\tau_t}\}, \{\mathcal{F}_{\tau_t}, X_{\tau_t}\}] \qquad \text{is a weak solution to the equation (2.11).} \tag{2.13}$$

To prove (2.13) we refer to the properties of a stochastic integral – Theorem 3.2.10 in [16], pp. 139–140 – for the first equality; and the stochastic Chain rule 1.13 for the second. The two remaining equalities are trivial:

$$\begin{aligned} \int_0^t \sigma(Y_{\tau_t}) \mathrm{d}X_{\tau_t} &= \int_0^{\tau_t} \sigma(Y_t) \mathrm{d}X_t = \int_0^{\tau_t} \sigma(Y_t)\sigma(Y_t)^{-1} \mathrm{d}Y_t \\ &= \int_0^{\tau_t} 1 \mathrm{d}Y_t = Y_{\tau_t}. \end{aligned}$$

We have hence constructed a weak solution to equation (2.11).

Remark 2.12. We should visualize what has been done. In Figure 2.12 on the upper left side, there are simulated sample paths of a solution to

$$\mathrm{d}X_t = \max[\epsilon, \sin(X_t)]\mathrm{d}B_t; \qquad X_0 = 0; \quad \epsilon = 0.05. \tag{2.14}$$

Below on the lower left, relevant Brownian motion sample paths are constructed. There are 10^7 time steps used in the simulation.

The nature of our solution construction in Example 2.11, the time – change, is not only an enthralling probabilistic method, it is also a method which can find vast application opportunities. Since the final Chapter 4 concentrates on an interest rate model, we shall also vaguely indicate how this time – change could be applied when studying intraday prices of German government bond futures, namely the *Bund*.

Futures contracts are nowadays the most liquid examples of derivative securities – securities whose value depends on more basic assets, usually called *underlyings*. These are in case of futures usually stocks indices or interest rates. Futures mean for their holder an obligation to buy the underlying at a predetermined price, which is the market price at the time of the agreement so that initially no cash changes hands, at a prespecified future time called *expiry*.

Since the trading activity varies during the day and hence the price process has wildly time – varying characteristics, it might look demanding to find a trustworthy form of a stochastic differential equation to model such prices. However, the method of a time – change can help avoid any similar obstacles. Imagine that τ_t from (2.12) were a time when a trade occurred. In other words, time would go very quickly when the trading activity is high and vice versa. The upper and lower right of the following Figure 2.12 show, that such a time – change makes it visually possible for the intraday Bund prices to be modeled by a *CIR* process (for details on a CIR process refer to Chapter 4). Of course, there are economic fundamental reasons for the CIR model to be a reasonable framework. The time – change actually only tells us how to calibrate the model when we want to make use of high – frequency data, where the observations are not equidistant in time.

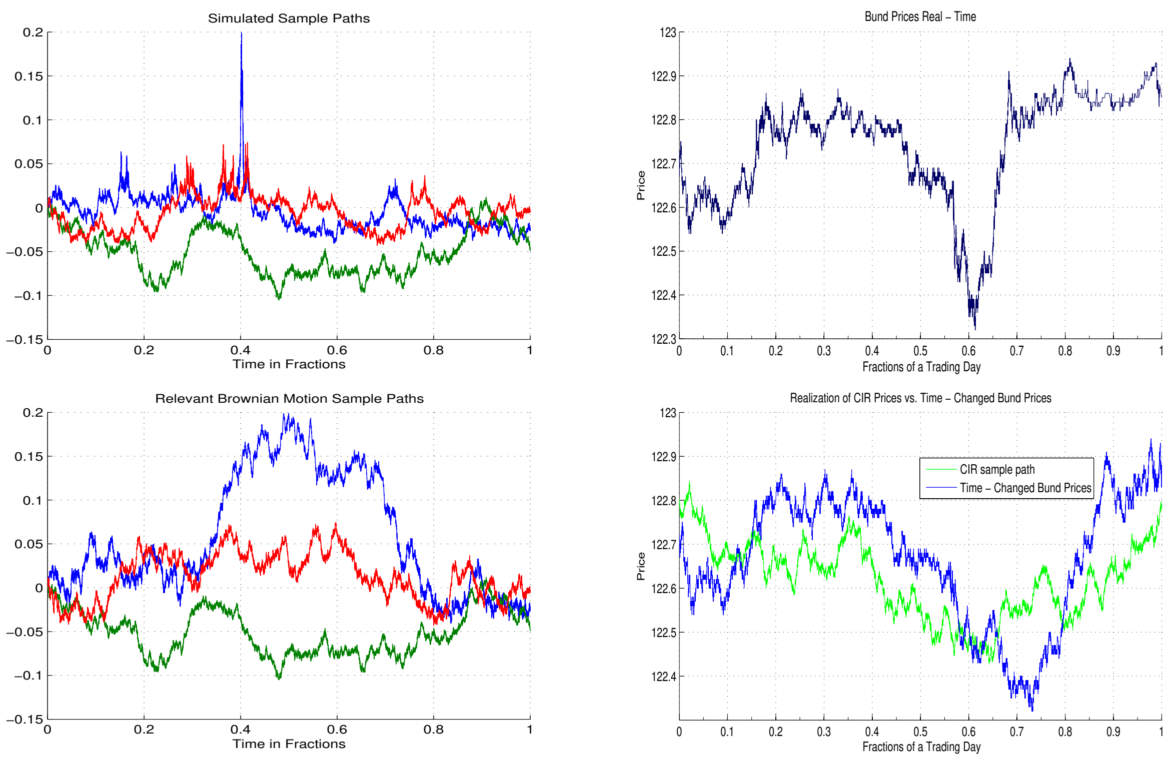

Figure 2.1: The left two figures are connected to (2.14). The right two are Intraday Bund prices on 1st September 2009, real – time and time – changed. The time changed price process is plotted in one graph with a realization of a CIR process with accordingly adjusted parameters. Source: Author's computations and a part of a data sample from www.deutsche – boerse.com.

2.2 Some Specific Results on Exit Times

Herein we present two important theorems that treat the question of a solution behaviour near the endpoints of the state space and can be found for example in [13], pp. 362–366. They are known together as a *Feller's test for explosions*. We are interested in the necessary and sufficient conditions for the *solution* to exit an interval. The theorems we are interested in can be proved using the so –

called martingale techniques. These techniques are widespread in the field of financial mathematics and hence we shall first explore them in their simplest form.

Let W_t^x be a Brownian motion starting from x, i.e.:

$$\mathrm{d}W_t = \mathrm{d}B_t; \qquad W_0 = x,$$

and let τ be an exit time of W_t^x from an interval (A, B), $0 < A < x < B < \infty$. Suppose we are interested in computing, or at least finding properties of, $\mathrm{E}[h(\tau, W_\tau^x)]$. A natural straightforward idea is to try to find

$$f(t, x) = g(t, x) + h(t, x),$$

such that $M_t = f(t, W_t)$ is a local martingale and $\mathrm{E}[g(t, W_t)]$ is known – for if it were known how to compute $\mathrm{E}[g(t, W_t)]$, it would be easy to investigate $\mathrm{E}[h(t, W_t)]$. In the following example 2.13 we replicate the fact that t is a *compensator* of W_t^2 in the sense of the *Doob–Meyer* decomposition 1.9 and compute the expected exit time $\mathrm{E}\tau$.

Example 2.13. Choose $g(t, x) = ax^2 + bx + c$ and set $f = t + ax^2 + bx + c$. For $f(t, W_t)$ to be a local martingale we use the Itô formula:

$$\mathrm{d}f(t, W_t) = (a + 1)\mathrm{d}t + (2aW_t + b)\mathrm{d}W_t$$

to deduct that $a = -1$. Choosing $b = c = 0$, $f(t, x) = t - x^2$ makes $M_t \equiv f(t, W_t)$ a martingale. We will see in Lemma 2.17 tailored to Brownian motion that $\mathrm{E}\tau < \infty$ and therefore, $M_{t\wedge\tau}$ is uniformly integrable. Hence the Optional Sampling theorem 1.7 applies and we get that:

$$\mathrm{E}[f(\tau, W_\tau)] = \mathrm{E}[\lim_{t\to\infty} f(\tau \wedge t, W_{\tau\wedge t})] = \lim_{t\to\infty} \mathrm{E}[f(\tau \wedge t, W_{\tau\wedge t})] = \mathrm{E}[f(0, W_0)];$$
$$\mathrm{E}[f(\tau, W_\tau)] = \mathrm{E}\tau - \mathrm{E}[W_\tau^2] = \mathrm{E}[f(0, W_0)] = -x^2.$$

Because $E[W_\tau^2] = x(A + B) - AB$, we conclude that:

$$\mathrm{E}\tau = x(A + B) - AB - x^2.$$

We repeat the steps from example 2.13 several times below, however, it shall never be as transparent as above – the technical details may blur such an elegant approach. For practitioners, it turns out that such an approach can solve a majority of problems they have to cope with. A good example is a book [3], where one can find solutions to a vast number practical modeling issues beginning with "Find a martingale..." throughout.

After this revision we proceed to more demanding parts. We start with an open interval

$$I = (l, r); \qquad -\infty \le l < r \le \infty; \quad x \in I,$$

and assume that the coefficients $\mu(x)$ and $\sigma(x)$ defined on I are such that the stochastic differential equation (2.1) has a weak solution up to an exit time e which is unique in law. For example, it is enough to let $\mu(x)$ and $\sigma(x)$ be continuous as is shown in [13], Theorem IV.2.3 p.159. Because we will be treating exit times of the solutions to (2.1), we first have to establish that when the exit time e as we defined it in definition 2.3 is finite almost surely, the solution X_t at the time e almost surely has a limit. Such a result is straightforward, once it is proved that the sample paths of weak solutions to (2.1) are really continuous functions with the property that once they reach a boundary of I, they stay there forever. Such a proof, which the author finds very technical and quite demanding, can be found in [13], lemma IV.2.1., pp. 160–162. The intuition that the existence of such a limit is established is, however, from the proof obvious.

Let us now recall the scale function (2.9) and prove the following easy lemma.

Lemma 2.14. *The limit behaviour of the scale function, i.e. finiteness or infiniteness of*

$$s(x) = \int_c^x \exp\left\{-\int_c^y \frac{2\mu(u)}{\sigma(u)^2}\mathrm{d}u\right\}\mathrm{d}y$$

by the endpoints of I does not depend on the constant $c \in I$.

Proof. Let us stress the dependence of $s(x)$ on c by adjusting the notation for a short while and let $a < c$ without loss of generality:

$$\begin{aligned} s_a(x) &= \int_a^x \exp\left\{-\int_a^y \frac{2\mu(u)}{\sigma(u)^2}\mathrm{d}u\right\}\mathrm{d}y \\ &= s_a(c) + \int_c^x \exp\left\{-\int_a^c \frac{2\mu(u)}{\sigma(u)^2}\mathrm{d}u - \int_c^y \frac{2\mu(u)}{\sigma(u)^2}\mathrm{d}u\right\}\mathrm{d}y \\ &= s_a(c) + s_c(x)s'_a(c). \end{aligned}$$

Therefore $s_a(x)$ is finite if and only if $s_c(x)$ is finite. Since $s'_a(c)$ is a positive number, when one side diverges, the other side diverges with the same sign, which proves our assertion. □

At this point, we must find solutions of two types of deterministic differential equations we will need further. We start with

$$Lv = v; \qquad v'(c) = 0; \quad v(c) = 1, \tag{2.15}$$

where L is as in (2.8). To find a solution to such an equation, we first have to find a solution to the following equation:

$$Lv_n = v_{n-1}; \qquad v'_n(c) = 0; \quad v_n(c) = 0, \quad n \in \mathbf{N}, \tag{2.16}$$

which we search for as a solution to a system of differential equations

$$\begin{aligned} h'(x) + 2\sigma^{-2}(x)\mu(x)h(x) &= 2\sigma^{-2}(x)v_{n-1}(x); \qquad h(c) = 0; \\ v'_n(x) &= h(x); \qquad v_n(c) = 0. \end{aligned} \tag{2.17}$$

Equation (2.17) can be solved using standard methods. First we find a fundamental solution using a homogeneous equation, which – as we have already seen – is the derivative of the scale function (2.9) $s'(x)$. Next we use the *variation of constants* formula to conclude, that

$$h(x) = s'(x)\int_c^x [s'(y)]^{-1} 2\sigma^{-2}(y)v_{n-1}(y)\mathrm{d}y.$$

And hence it follows that

$$\begin{aligned} v_n(x) &= \int_c^x s'(z)\int_c^z [s'(y)]^{-1} 2\sigma^{-2}(y)v_{n-1}(y)\mathrm{d}y\mathrm{d}z \\ &= \int_c^x s'(z)\int_c^z v_{n-1}(y)\mathrm{d}m(y)\mathrm{d}z, \end{aligned}$$

where $\mathrm{d}m(y) = \frac{2\mathrm{d}y}{s'(y)\sigma^2(y)}$ is usually called a *speed measure* and $v_0(x) = 1$. The above cited *DDS* theorem says that every nonconstant continuous local martingale is a time – changed Brownian motion. A speed measure says how such a change of clock affects exit times from an interval. Note that $v_n(x)$ is increasing on (c, r) and decreasing on (l, c). It can be quickly verified that none of the properties we are interested in depend on a specific choice of c, which we verify below for $n = 1$ only. Because the function $v_1(x)$ will play an important role throughout the chapter, we shall reserve a letter for it and let $k(x) \equiv v_1(x)$.

Lemma 2.15. *The limit behaviour of the function*

$$k(x) = \int_c^x s'(z) \int_c^z \mathrm{d}m(y)\mathrm{d}z$$

by the endpoints of I does not depend on the constant $c \in I$.

Proof. Similarly to the proof of lemma 2.14, let us stress the dependence of $k(x)$ on c by adjusting the notation for a short while and assume without loss of generality that $a < c$:

$$\begin{aligned} k_c(x) &= \int_c^x s'(z) \int_c^z \mathrm{d}m(y)\mathrm{d}z \\ k_a(x) &= \int_a^x s'(z) \int_a^z \mathrm{d}m(y)\mathrm{d}z = \int_a^c s'(z) \int_a^z \mathrm{d}m(y)\mathrm{d}z + \int_c^x s'(z) \int_a^z \mathrm{d}m(y)\mathrm{d}z \\ &= k_a(c) + k_c(x) + \int_c^x s'(z) \int_a^c \mathrm{d}m(y)\mathrm{d}z = k_a(c) + k_c(x) + \int_a^c \mathrm{d}m(y) \int_c^x s'(z)\mathrm{d}z \\ &= k_a(c) + k_c(x) + k_a'(c)s(x). \end{aligned}$$

Keeping in mind Lemma 2.14 we further need to verify the following almost obvious implications:

$$\begin{aligned} k(r-) < \infty \quad &\Rightarrow s(r-) < \infty; \\ k(l+) < \infty \quad &\Rightarrow s(l+) > -\infty. \end{aligned} \tag{2.18}$$

It is a straightforward reasoning: Let $\epsilon > 0$, $c + \epsilon < x < r$, and remember that for such an x both $s'(z)$ and $m(y)$ are positive. Therefore,

$$\begin{aligned} k(x) &\geq \int_{c+\epsilon}^x s'(z) \int_c^{c+\epsilon} \mathrm{d}m(y)\mathrm{d}z \\ &\geq (s(x) - s(c+\epsilon)) \int_c^{c+\epsilon} \mathrm{d}m(y), \end{aligned}$$

which is a proof of (2.18). Now, if $k_a(r-)$ is finite, so is $s(r-)$ and hence $k_c(r-)$ must be finite too. The second endpoint behaviour is obtained analogously. □

In the following, we take advantage of a real analysis lemma from [16] pp. 347–358:

Lemma 2.16. $\sum_{n=0}^{\infty} v_n(x)$ *converges uniformly on compact subsets of I to a differentiable function* $v(x) = \sum_{n=0}^{\infty} v_n(x)$ *which satisfies the equation* (2.15)*. Moreover,*

$$1 + k(x) \leq v(x) \leq \exp(k(x)). \tag{2.19}$$

Proof. By definition, $1 + k(x) \leq v(x)$. We show by induction that

$$v_n(x) \leq \frac{k^n(x)}{n!}.$$

Indeed, it is true for $n = 0$ and we assume it holds for a fixed n. Assuming $c < x$, we already clarified that $k(x)$ is nondecreasing. Hence,

$$\begin{aligned} v_{n+1}(x) &= \int_c^x s'(z) \int_c^z \frac{2v_n(y)}{s'(y)\sigma^2(y)}\mathrm{d}y\mathrm{d}z \leq \int_c^x s'(z) \int_c^z \frac{2k^n(y)}{n!s'(y)\sigma^2(y)}\mathrm{d}y\mathrm{d}z \\ &\leq \int_c^x \frac{s'(z)k^n(z)}{n!} \int_c^z \frac{2}{s'(y)\sigma^2(y)}\mathrm{d}y\mathrm{d}z = \int_c^x \frac{k^n(z)}{n!}\mathrm{d}k(z) \\ &= \frac{k^{n+1}(z)}{(n+1)!}. \end{aligned}$$

Consequently, we have (2.19) and also $v'_{n+1}(x) \le \frac{k'(x)k^n(x)}{n!}$. From the dominated convergence criteria we deduct that both $\sum_n v_n$ and $\sum_n v'_n$ converge absolutely on I and (2.16) implies that $\sum_n v''_n$ converges absolutely on I too, to an integrable function. Formally integrating $\sum_n v''_n$ term by term we see that it is the second derivative of $v(x) = \sum_n v_n(x)$. □

Another result we will use is a solution to:

$$Lw = g; \qquad w(a) = 0; \quad w(b) = 0. \tag{2.20}$$

Taking our previous steps into account, using variation of constants:

$$w'(x) = s'(x)w'(0) + s'(x)\int_a^x [s'(y)]^{-1} 2\sigma^{-2}(y)g(y)\mathrm{d}y.$$

Thus, including both boundary values and changing the order of integration we get the following equalities:

$$\begin{aligned} w(x) &= \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (2.21)\\ &= \int_a^x \int_a^y \frac{2s'(y)g(z)}{s'(z)\sigma^2(z)}\mathrm{d}z\mathrm{d}y - \frac{s(x)-s(a)}{s(b)-s(a)}\int_a^b\int_a^y \frac{2s'(y)g(z)}{s'(z)\sigma^2(z)}\mathrm{d}z\mathrm{d}y \\ &= \int_a^x \int_z^x \frac{2s'(y)g(z)}{s'(z)\sigma^2(z)}\mathrm{d}y\mathrm{d}z - \frac{s(x)-s(a)}{s(b)-s(a)}\int_a^b\int_z^b \frac{2s'(y)g(z)}{s'(z)\sigma^2(z)}\mathrm{d}y\mathrm{d}z \\ &= \int_a^x \frac{2(s(x)-s(a))g(z)}{s'(z)\sigma^2(z)}\mathrm{d}z - \frac{s(x)-s(a)}{s(b)-s(a)}\int_a^b \frac{2(s(b)-s(z))g(z)}{s'(z)\sigma^2(z)}\mathrm{d}z \\ &= \int_a^b \frac{(s(b)-s(x\vee z))(s(x\wedge z)-s(a))}{s(b)-s(a)}\frac{2(-g(z))}{s'(z)\sigma^2(z)}\mathrm{d}z \\ &\equiv \int_a^b G_{a,b}(x,z)(-g(z))\mathrm{d}m(y), \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (2.22)\end{aligned}$$

where $G_{a,b}$ may be called a *Green's function.* This name refers to a classical analytical approach, see for example [20], pp.196–198.

We now move back to our weak solution of (2.1) on I, where we assume $\mu(x)$ and $\sigma(x) > 0$ continuous.

Lemma 2.17. *X_t^x exits every compact subinterval of I in a finite expected time.*

Proof. Let $[a,b] \subsetneq I$. Define:

$$\begin{aligned}\tau &= \inf\left[t \ge 0, X_t^x \notin (a,b)\right];\\ \upsilon_n &= \inf\left[t \ge 0, \int_0^t \sigma^2(X_s^x)\mathrm{d}s \ge n\right], \quad n = 1,2,\ldots\end{aligned}$$

and find w such that (2.20) holds with $g = -1$. From (2.22) we have an explicit form of w. It holds that w is continuous, bounded and nonnegative on I. Using Itô formula and the fact that $\{\upsilon_n\}_n$ goes to infinity almost surely, for an arbitrary t we can proceed using the local martingale property of stochastic integrals, the continuity of solutions and the Lebesgue dominated convergence theorem:

$$\begin{aligned}
\mathrm{d}w(X_t^x) &= Lw(X_t^x)\mathrm{d}t + w'(X_t^x)\sigma(X_t^x)\mathrm{d}B_t;\\
w(X_{t\wedge\tau\wedge\upsilon_n}) &= w(X_0) - (t\wedge\tau\wedge\upsilon_n) + \int_0^{t\wedge\tau\wedge\upsilon_n} w'(X_s)\sigma(X_s)\mathrm{d}B_s;\\
\mathrm{E}[w(X_{t\wedge\tau\wedge\upsilon_n})] &= w(x) - \mathrm{E}[t\wedge\tau\wedge\upsilon_n];\\
\lim_{n\to\infty}\mathrm{E}[w(X_{t\wedge\tau\wedge\upsilon_n})] &= w(x) - \lim_{n\to\infty}\mathrm{E}[t\wedge\tau\wedge\upsilon_n];\\
\mathrm{E}[w(X_{\lim_{n\to\infty} t\wedge\tau\wedge\upsilon_n})] &= w(x) - \mathrm{E}\left[\lim_{n\to\infty}(t\wedge\tau\wedge\upsilon_n)\right];\\
\mathrm{E}(t\wedge\tau) &= w(x) - \mathrm{E}[w(X_{t\wedge\tau})] \le w(x) < \infty.
\end{aligned}$$

As t was arbitrary, we now let $t\to\infty$ to obtain $\mathrm{E}\tau \le w(x) < \infty$ which proves the assertion. □

In the following, note that $f(x+)$ or $f(x-)$ mean a shorthand for a limit of f evaluated at x from the right or left, respectively.

Theorem 2.18 (Sufficient Condition for Unattainability)**.** *Let $s(x)$ be as in (2.9).*
(i) If $s(l+) = -\infty$ and $s(r-) = \infty$, then for every x the process X_t^x is recurrent:

$$P(e=\infty) = P\left(\overline{\lim_{t\uparrow\infty}} X_t^x = r\right) = P\left(\underline{\lim_{t\uparrow\infty}} X_t^x = l\right) = 1. \tag{2.23}$$

(ii) If $s(l+) > -\infty$ and $s(r-) = \infty$, then for every x the process X_t^x has a limit almost surely and:

$$P\left(\lim_{t\uparrow e} X_t^x = l\right) = P\left(\sup_{t<e} X_t^x < r\right) = 1. \tag{2.24}$$

(iii) If $s(l+) = -\infty$ and $s(r-) < \infty$, then for every x the process X_t^x has a limit almost surely and:

$$P\left(\lim_{t\uparrow e} X_t^x = r\right) = P\left(\inf_{t<e} X_t^x > l\right) = 1. \tag{2.25}$$

(iv) If $s(l+) > -\infty$ and $s(r-) < \infty$, then for every x the process X_t^x has a limit almost surely and:

$$P\left(\lim_{t\uparrow e} X_t^x = l\right) = 1 - P\left(\lim_{t\uparrow e} X_t^x = r\right) = \frac{s(r-)-s(x)}{s(r-)-s(l+)}. \tag{2.26}$$

Proof. Let again $[a,b] \subsetneq I$, $x\in(a,b)$ and $\tau = \inf[t\ge 0, X_t^x \notin (a,b)]$. Using the approach we are already well familiar with – Itô formula and switching integration and a limit – from (2.9) we get the following equalities:

$$\begin{aligned}
s(X_{t\wedge\tau}^x) - s(x) &= \int_0^{t\wedge\tau} s'(X_s^x)\sigma(X_s^x)\mathrm{d}B_s;\\
s(x) &= \mathrm{E}s(X_{t\wedge\tau}^x);
\end{aligned} \tag{2.27}$$

upon letting $t\to\infty$:

$$\begin{aligned}
s(x) &= \mathrm{E}s(X_\tau^x) = s(a)P(X_\tau^x = a) + s(b)P(X_\tau^x = b)\\
&= s(a)P(X_\tau^x = a) + s(b)(1 - P(X_\tau^x = a)),
\end{aligned}$$

and we have:

$$P(X_\tau = b) = \frac{s(x)-s(a)}{s(b)-s(a)}; \quad P(X_\tau = a) = \frac{s(b)-s(x)}{s(b)-s(a)}. \tag{2.28}$$

(i) Suppose $s(r-) = -s(l+) = \infty$. Then

$$\lim_{a \downarrow l} P(X_\tau = b) = 1.$$

Clearly, $P(\sup_{t<e} X_t \geq b) \geq P(X_\tau = b)$. Hence,

$$1 \geq P\left(\sup_{t<e} X_t \geq b\right) \geq \lim_{a \downarrow l} P\left(X_\tau = b\right) = 1.$$

Since b was fixed but arbitrary, letting $b \uparrow r$ we have:

$$P\left(\sup_{t<e} X_t = r\right) = 1.$$

A similar dual argument shows $P(\inf_{t<e} X_t = l) = 1$. Suppose there exists a set of positive probability where $e < \infty$. Now on this set we know that $\lim_{t \uparrow e} X_t^x$ exists and is equal to l or r. But that is impossible indeed, X is continuous and in that case both $\{\inf_{t<e} X_t = l\}$ and $\{\sup_{t<e} X_t = r\}$ cannot simultaneously have probability one. Therefore, $e = \infty$ almost surely and (2.23) is proved.

(ii) As above, it is still true that $P(\inf_{t<e} X_t = l) = 1$. Now let $a \downarrow l$ and from (2.28) we obtain:

$$P[\exists t < e : X_t^x = b] = \frac{s(x) - s(l+)}{s(b) - s(l+)}.$$

We want to prove that $P(\sup_{t<e} X_t < r) = 1$. Clearly, $P(\sup_{t<e} X_t \leq r) = 1$, so to reach a contradiction we assume that $P(\sup_{t<e} X_t = r) = K > 0$. In other words:

$$\forall \epsilon > 0 \quad P(\omega \in \Omega : \quad \exists t_0 : |X_{t_0}(\omega) - r| < \epsilon) \geq K.$$

Because $\lim_{b \uparrow r} P[\exists t < e : X_t^x = b] = 0$, it is possible to choose ϵ_0 such that:

$$P[\exists t < e : X_t^x = r - \epsilon_0] = \frac{s(x) - s(l+)}{s(r - \epsilon_0) - s(l+)} < K.$$

Therefore:

$$\begin{aligned} K &> P(\exists 0 \leq t_0 < e : X_{t_0} = r - \epsilon_0) \\ &\geq P(\exists 0 \leq t_0 < e : X_{t_0} > r - \epsilon_0) \\ &= P(\exists 0 \leq t_0 < e : |X_{t_0} - r| < \epsilon_0) \\ &\geq K. \end{aligned}$$

This contradiction proves $P(\sup_{t<e} X_t < r) = 1$. It follows that the only candidate for $\lim_{t \uparrow e} X_t^x$ is $\inf_{t<e} X_t$ and to establish (2.24) it suffices to show that the limit almost surely exists. Remember from the definition of exit time, that we have defined $\tau_n = \inf\{t; X_t^x \notin [l_n, r_n]\}$. Now, $s(l+)$ is finite and recall we could have chosen l_1 such that $x > l_1$. Therefore, the process $Y_{t \wedge \tau_n}^x = s(X_{t \wedge \tau_n}^x) - s(l+)$ is a nonnegative martingale. Letting $n \to \infty$, by *Fatou's lemma* $Y_{t \wedge e}^x$ is a nonnegative supermartingale. As such, by Theorem 1.19, $Y_{t \wedge e}^x$ has an almost surely finite limit. Moreover, because $s(x)$ is continuous increasing, it has a continuous inverse. We conclude that $\lim_{t \uparrow e} X_t$ must exists almost surely too, thus

$$P\left(\inf_{0 \leq t < e} X_t = \lim_{t \uparrow e} X_t^x = l\right) = 1.$$

(iii) (2.25) follows by dual arguments to (ii).

(iv) We obtain (2.26) by similar arguments as in (ii) and (iii), limiting $a \downarrow l$ and $b \uparrow r$ in (2.28). $\square$

Theorem 2.19 (Necessary and Sufficient Conditions for Unattainability)**.**

$$\text{(A) } P(e=\infty)=1 \quad \forall x \in I \quad \text{if and only if} \quad k(r-)=k(l+)=\infty.$$

Moreover,

$$\text{(B) } P(e<\infty)=1 \quad \forall x \in I$$

if and only if one of the following cases holds:

$$\text{(1) } k(r-)<\infty \text{ and } k(l+)<\infty; \tag{2.29}$$

$$\text{(2) } k(r-)<\infty \text{ and } s(l+)=-\infty; \tag{2.30}$$

$$\text{(3) } k(l+)<\infty \text{ and } s(r-)=\infty. \tag{2.31}$$

And in case (1) we actually have $\mathrm{E}e<\infty$*.*

Proof. To prove (A), let $v(x)$ be defined by (2.15) and let again $l<l_n<x<r_n<r$ and $\tau_n=\inf\{t; X_t^x \notin [l_n,r_n]\}$. We proceed similarly as above, by Itô lemma and considering the properties of $v(x)$:

$$\begin{aligned} \mathrm{d}e^{-t\wedge\tau_n}v(X^x_{t\wedge\tau_n}) =& e^{-t\wedge\tau_n}v'(X^x_{t\wedge\tau_n})\sigma(X^x_{t\wedge\tau_n})\mathrm{d}B_{t\wedge\tau_n} \\ &+ e^{-(t\wedge\tau_n)}\left[(-v(X^x_{t\wedge\tau_n}))+(Lv)(X^x_{t\wedge\tau_n})\right]\mathrm{d}(t\wedge\tau_n) \\ =& e^{-t\wedge\tau_n}v'(X^x_{t\wedge\tau_n})\sigma(X^x_{t\wedge\tau_n})\mathrm{d}B_{t\wedge\tau_n} \end{aligned}$$

and hence $e^{-t\wedge\tau_n}v(X^x_{t\wedge\tau_n})$ is a nonnegative martingale. Letting $n\to\infty$ and using Fatou's lemma, because v is strictly positive we see that $e^{-t\wedge e}v(X^x_{t\wedge e})$ is a nonnegative supermartingale and as such has an almost surely finite limit for $t\to\infty$, again by Theorem 1.19. If $k(r-)=k(l+)=\infty$, then by the first inequality in (2.19), $v(r-)=v(l+)=\infty$. Consequently, $P(e<\infty)>0$ is impossible, because $\lim_{t\to\infty} e^{-t\wedge e}v(X^x_{t\wedge e})=\infty$ on the set where $e<\infty$. It follows, that:

$$P(e=\infty)=1 \quad \forall x \in I.$$

For the converse, assume for example $k(l+)<\infty$, which immediately yields $s(l+)>-\infty$.
Because the assumptions talk about limits of $k(x)$ only, in the spirit of lemma 2.15 we can assume $l<x<c$ and set

$$T=\inf\{t\geq 0; X_t=c\}.$$

A continuous process $e^{-t\wedge e\wedge T}v(X^x_{t\wedge e\wedge T})$ is clearly a bounded martingale and it follows that for $t\to\infty$ it converges almost surely to a finite limit. Therefore,

$$v(x)=\mathrm{E}\left[\exp\{-(e\wedge T)\}v(X^x_{e\wedge T})\right]=v(l+)\mathrm{E}\left[\exp\{-e\}\mathrm{I}_{\{e<T\}}\right]+v(c)\mathrm{E}\left[\exp\{-T\}\mathrm{I}_{\{e>T\}}\right]. \tag{2.32}$$

Assuming $e=\infty$ almost surely, the preceding equality (2.32) yields:

$$v(x)=v(c)\mathrm{E}\left[\exp\{-T\}\mathrm{I}_{\{e>T\}}\right]\leq v(c).$$

But this is a contradiction to the fact, that $v(y)$ strictly decreases on $[x,c]$. We have just proved $P(e=\infty)<1$.

We shall now prove (B) and start with necessity. We assume that $P(e<\infty)=1$ $\forall x\in I$ and suppose none of (1), (2), (3) hold. From (A) we know that either $k(l+)$ or $k(r-)$ is finite. We will suppose $k(l+) < \infty$, the latter would be treated similarly. Now with these assumptions, thanks to what we have already proved, we are in the situation where:

$$k(r-) = \infty; \quad v(r-) = \infty \quad s(r-) < \infty; \quad s(l+) > -\infty.$$

Hence, we meet the criteria of Theorem 2.18, case (iv), and $1 > C \equiv P(\lim_{t\uparrow e} X_t^x = r) > 0$.
As we know from the proof of (A), $\exp\{-(t \wedge e)\}v(X_{t\wedge e})$ is a nonnegative supermartingale with an almost surely finite limit. Because $v(r-) = \infty$, e must be equal to infinity on the set of probability C. This shows $P(e < \infty) < 1$ and we contradict our initial assumption. Hence, if $P(e < \infty) = 1$ holds, so must at least one of (1), (2), (3).

We finally prove the sufficiency in (B). If (1) holds, then we have finiteness of $s(l-)$ and $s(r+)$ and in analogy with what we have already done in (2.22), we can define:

$$G(x, y) = \frac{(s(r-) - s(x \vee y))(s(x \wedge y) - s(l+))}{s(r-) - s(l+)}; \quad (x, y) \in I^2;$$

$$M(x) = \int_l^r G(x, y)\mathrm{d}m(y)$$

Now, M satisfies (2.20) with $g = -1$ on the whole I. Replicating lemma 2.17, we conclude that $\mathrm{E}e = M(x) < \infty$ and the very last statement of the theorem as well as (1) are justified.

Next we shall assume (2). Set a sequence of stopping times $\tau_n = \inf\{t; X_t = l_n\}$, where l_n is the sequence from the definition of an exit time 2.3 and $\tau_r = \inf\{t; X_t = r\}$. Clearly, $\lim_{n\to\infty} \tau_n \wedge \tau_r = e$. Because both $v(r-)$ and $v(l_n)$ are finite, we obtain similarly to the result of (1) that $E[\tau_n \wedge \tau_r] < \infty$. Because $s(r-) < \infty$ and $s(l+) = -\infty$, we find ourselves in the case (iii) of Theorem 2.18. Therefore, because $P(\inf_{t<e} X_t^x > l) = 1$, for sufficiently large n, we must have $\tau_n = \infty$ almost surely. Consequently:

$$P(e < \infty) = P(\tau_r < \infty) = 1.$$

Case (3) can be obtained dually to (2) by accordingly interchanging the roles of l and r. □

Remark 2.20. On this spot an example of how to use Theorem 2.18 and 2.19 should be made. It is postponed, however, to Chapter 4, where these theorems are used to deduce strict positiveness of a CIR process, see Theorem 4.1.

Chapter 3

Boundary Value Problems and Options Valuation

There is a rich interplay between probability theory and classical analysis. In the preceding chapter, we used the knowledge of solutions to deterministic problems to conclude assertions about solutions to stochastic differential equations. In this chapter, however, we explore the relationship between stochastic differential equations and deterministic differential equations somewhat from the other side. We summarize an extremely elegant theory showing that certain deterministic problems can be solved via diffusion functionals and that theses solutions are in some sense unique. We are still interested in one – dimensional diffusions X_t^x, but we may also work with their graphs, which results in exploring two – dimensional processes. Therefore, in this chapter, we generally work with a d – dimensional stochastic differential equation, where in (2.1) we replace our coefficients with $\mu(x) \in \mathbf{R}^d$, $\sigma(x) \in \mathbf{R}^{d\times d}$ and B_t stands for a d – dimensional Brownian motion.

Remark 3.1. When applying our results to the graphs $Y_t^y = Y_t^{s,x} = [s+t, X_t^x]$ of one – dimensional diffusions X_t^x, as said, we may turn concretely to the use of a two – dimensional equation of a special form:

$$\mathrm{d}Y_t = \mu(Y_t)\mathrm{d}t + \sigma(Y_t)\mathrm{d}B_t,$$

where:

$$\mathrm{d}X_t = \tilde{\mu}(X_t)\mathrm{d}t + \tilde{\sigma}(X_t)\mathrm{d}\tilde{B}_t$$

is a one – dimensional equation as in (2.1) and:

$$\mu(Y_t) = \begin{pmatrix} 1 \\ \tilde{\mu}(X_t) \end{pmatrix}; \sigma(Y_t) = \begin{pmatrix} 0 & 0 \\ 0 & \tilde{\sigma}(X_t) \end{pmatrix}; \quad \text{and} \quad B_t \text{ is two – dimensional.}$$

We consider Itô diffusions exclusively, i.e. global strong and pathwise unique solutions to stochastic differential equations where the coefficients μ and σ are Lipschitz continuous on $\mathbf{R}^d$.

We focus on building an intuition in a very interesting and profoundly explored field first studied by Richard Feynman, who was treating the subject as a physicist without introducing a proper mathematical rigour. It was Mark Kac, who formulated the interplay in a mathematical language and since then many generalizations were made. To be able to quickly outline the connections to finance and how the financial language treats this topic, we often stick to quite strict assumptions and therefore the theorems are not very strong. The proofs in this chapter, when presented, are rather sketches of proofs and relevant literature is cited.

3.1 Preliminary Results

What is desperately needed in many applications and what is often one of the most basic assumptions in modeling real world, is the important *Markov property* of an Itô diffusion: the future behaviour of an Itô diffusion, given what has happened until time time t, is the same as the behaviour based on knowledge of the state X_t^x and a "restart" of the process at X_t^x. We give a precise mathematical formulation of an even stronger result in the following Theorem 3.2, which states that the relation holds even if the time t is replaced by a stopping time. We adjust the notation to stress the context of Markov processes: for F Borel – measurable, $P^x(X_t \in F) \equiv P(X_t^x \in F)$, as well as $E^x(f(X_t)) \equiv E(f(X_t^x))$.

Theorem 3.2 (Strong Markov Property). *Let f be a bounded Borel – measurable function on $\mathbf{R}^d$ and τ an a.s. finite $\mathcal{F}_t$ – stopping time. Then*

$$E^x[f(X_{\tau+h})|\mathcal{F}_t](\omega) = E^{X_\tau(\omega)}[f(X_h)] \quad a.s. \quad \forall h \geq 0.$$

The right hand side of the equation above reads as the function $E^y[f(X_h)]$ evaluated at $y = X_\tau(\omega)$.

Proof. To be found in [20], pp. 118–119. □

Definition 3.3 (Hitting Distribution). Let X_t^x be an Itô diffusion. Let H be an open measurable set in $\mathbf{R}^d$ such that $x \in H$. Put

$$\tau_H = \inf\{t : X_t^x \notin H\}.$$

Then we define the *hitting distribution* μ_H^x of X_t^x on the boundary ∂H by

$$\mu_H^x(F) = P^x(X_{\tau_H} \in F) \qquad \forall F \subset \partial H.$$

Let G be a subset of H. One important consequence of the strong Markov property is that the expected value of $f(X_{\tau_H}^x)$ can be computed as an integral over all values $y \in \partial G$ with respect to the just defined hitting distribution of X_t^x on ∂G.

Theorem 3.4 (Mean Value Property). *Let X_t be an Itô diffusion and let $G \subset\subset H \subset \mathbf{R}^d$ be measurable such that $x \in G$. Assume further, that $f(x)$ is a bounded measurable function. Then the function $\phi(x) \equiv E^x[f(X_{\tau_H})]$ satisfies the* mean value property*:*

$$\phi(x) = E^x[f(X_{\tau_H})] = E^x\left[E^{X_{\tau_G}}[f(X_{\tau_H})]\right] = \int_{\partial G} \phi(y) \mathrm{d}\mu_G^x(y).$$

Proof. See [20], pp. 119–121. □

We have already seen throughout chapter 2 – for the first time we touched this topic when defining a scale function in example 2.10 – that it is sometimes absolutely fundamental to associate a differential operator to an Itô diffusion X_t. To establish the relationship, let us first denote

$$\begin{aligned} Lf(x) &= \sum_{i=1}^d \mu_i(x) \frac{\partial}{\partial x_i} f(x) + \frac{1}{2} \sum_{i,j=1}^d m_{ij}(x) \frac{\partial^2}{\partial x_i x_j} f(x) \\ &= \sum_{i=1}^d \mu_i(x) \frac{\partial}{\partial x_i} f(x) + \frac{1}{2} \sum_{i,j=1}^d \left(\sigma\sigma^T\right)_{ij}(x) \frac{\partial^2}{\partial x_i x_j} f(x), \end{aligned} \tag{3.1}$$

where $0 < d < \infty$ is the integer dimension we work in. This notation uses tacitly, that once we have a nonnegative definite matrix $M(x) = [(m_{ij})(x)]$, it has a nonnegative square root, which we denote by $[\sigma_{ij}(x)]$, see for example the introductory section in chapter 6 in [10]. The connection between a d – dimensional Itô diffusion and a partial differential operator L from (3.1) is provided by the following two concepts.

Definition 3.5 (Generator and Characteristic Operator of an Itô Diffusion)**.** Let X_t be an Itô diffusion in $\mathbf{R}^d$. The *generator* A of X_t is defined by:

$$A_X f(x) = \lim_{t \downarrow 0} \frac{E^x[f(X_t)] - f(x)}{t}; \qquad x \in \mathbf{R}^d. \tag{3.2}$$

The *characteristic operator* $\mathcal{A}$ is defined in a manner similar to (3.2), but rather more generally:

$$\mathcal{A}_X f(x) = \lim_{n \to \infty} \frac{E^x[f(X_{\tau_{U_n}})] - f(x)}{E^x[\tau_{U_n}]}, \tag{3.3}$$

where the open sets U_n satisfy $U_n \subset U_{n-1}$, $\cap_n U_n = x$ and $\tau_U = \inf\{t > 0; X_t \notin U\}$. The sets of functions such that the limits exist for all $x \in \mathbf{R}^d$ are denoted by $\mathbf{D}_A$ and $\mathbf{D}_{\mathcal{A}}$ respectively. Again, wherever possible we drop the subscript $_X$.

It is once again an application of Itô formula that proves the following theorem, for a detailed discussion see [20], pp. 122–124.

Theorem 3.6 (Dynkin)**.** *Let X_t^x be an Itô diffusion in $\mathbf{R}^d$. Let $f \in C^2(\mathbf{R}^d)$ with compact support , further denoted $f \in C_0^2(\mathbf{R}^d)$, and let τ be a stopping time such that $E^x\tau$ is finite. With these assumptions, it holds that $f \in \mathbf{D}_A$ and*

$$E^x[f(X_\tau)] = f(x) + E^x\left[\int_0^\tau Lf(X_s)\mathrm{d}s\right] = f(x) + E^x\left[\int_0^\tau Af(X_s)\mathrm{d}s\right]. \tag{3.4}$$

Remark 3.7. The above theorem holds even for more general f, μ and σ. We only have to make sure, that the local martingales appearing in the proof are true martingales and that the integrals make sense. If τ were an exit time from a bounded subsets, it would allow us to drop the compact support assumption, for example.

Proof. We know that stochastic integrals are local martingales and that continuous functions on $\mathbf{R}^d$ are bounded on compact subsets. Bounded local martingales are true martingales. Therefore, by Itô formula, the first equality is obtained. The second equality follows from the fact, that Itô diffusions have continuous sample paths and that due to the boundedness of L we can switch the order of computing an expectation and a limit. By the definition of A, $Af(x) = Lf(x)$ must hold. □

For our purposes, it is important that the concept of $\mathcal{A}f(x)$ coincides with $Lf(x)$ too, this time for all functions $f \in C^2(\mathbf{R}^d)$. This result is obtained easily from the definition of $\mathcal{A}$ and the Dynkin formula 3.6, see [20], Theorem 7.5.4, page 127.

3.2 Dirichlet Problem

In this section, we finally treat the Dirichlet problem, which is a popular name to the following type of equation:

Let D be an open connected set in $\mathbf{R}^d$ and let $f \in C(\partial D)$ be bounded. Find $u \in C^2(D)$ such that:

$$Lu(x) = 0 \qquad \forall x \in D; \tag{3.5}$$

$$\lim_{x \to y} u(x) = f(y) \qquad \forall x \in D; \quad y \in \partial D, \tag{3.6}$$

where L is a *semi – elliptic* operator. An operator is semi – elliptic, if the eigenvalues of the matrix $[(\sigma\sigma^T)_{ij}(x)]$ are non – negative.

The Dirichlet problem was amongst the first *Boundary Value Problems* to be studied. One idea of how to find its solution comes from the established relationship between an Itô diffusion and a partial differential operator: One must find an Itô diffusion whose generator A coincides with L. To be able to do this, $\mu(x)$ and $M(x) = \sigma\sigma^T(x)$ must satisfy conditions so that the existence of a global pathwise unique solution to a multi – dimensional stochastic differential equation holds. For example, $\mu(x)$ must be Lipschitz continuous and $M(x)$ needs be bounded with continuous bounded second derivatives, see [10], pp. 129–131.

Thus, let X_t be such a unique solution to

$$\mathrm{d}X_t = \mu(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t. \tag{3.7}$$

The power of the probabilistic approach is that we can immediately write down a very likely candidate for the solution $u(x)$ of (3.5) and (3.6):

$$u(x) = \mathrm{E}^x[f(X_{\tau_D})\mathbf{1}_{[\tau_D < \infty]}]. \tag{3.8}$$

It turns out, that under our assumptions, if there exists a bounded solution to (3.5), it has the representation (3.8). Therefore we have *uniqueness* of solutions to the Dirichlet problem:

Theorem 3.8. *Suppose $\tau_D < \infty$ a.s. for all $x \in D$. Then if $u \in C^2(D)$ is a bounded solution to the Dirichlet problem, it holds that*

$$u(x) = \mathrm{E}^x[f(X_{\tau_D})]. \tag{3.9}$$

Proof. This is a straightforward application of the Dynkin formula 3.6 and the Lebesgue's dominated convergence theorem. Details are in [20], pp.178–179. □

The second answer to be found is when do the solutions to the Dirichlet problem exist. Or in other words, explore when $u(x)$ is a solution. This is generally a difficult topic, and it is often needed to completely leave probability and borrow general theorems from the theory of partial differential equations. For $L = \Delta$, where Δ stands for the Laplace operator, the answer is however complete even in the probabilistic language and can be found in [16], pp.243–250. The answer about existence is reached with the help of a solution to the probabilistic counterpart to the Dirichlet problem, the *Stochastic Dirichlet Problem*.

Definition 3.9 (X-harmonic Functions)**.** Let f be a locally bounded measurable function on D. We call f *X – harmonic* in D, when

$$f(x) = \mathrm{E}^x\left[f(X_{\tau_U})\right] \tag{3.10}$$

holds for all $x \in D$ and all bounded open sets U such that $\bar{U} \subset D$.

The knowledge of Dynkin formula 3.6 allows us to state the following lemma without proof:

Lemma 3.10. *If f is X – harmonic in D, then $\mathcal{A}f = 0$. If moreover $f \in C^2(D)$, then the other implication holds too.*

Finally, the following theorem establishes existence and uniqueness of solutions to the stochastic parallel of the Dirichlet problem.

Theorem 3.11 (Solution of the Stochastic Dirichlet Problem)**.** *Let ϕ be a bounded measurable function on ∂D.*
(i) Set $u(x) = \mathrm{E}^x[\phi(X_{\tau_D})]$. Then

$$u(x) \text{ is } X \text{ – harmonic and } \quad \lim_{t\uparrow\tau_D} u(X_t^x) = \phi(X_{\tau_D}^x) \quad a.s.; \quad x \in D.$$

(ii) If on the other hand g is a bounded function on D such that it is X – harmonic and

$$\lim_{t\uparrow\tau_D} g(X_t^x) = \phi(X_{\tau_D}^x) \quad a.s.; \quad x \in D,$$

then $g(x) = u(x)$.

Proof. In [20], p. 184. □

It turns out that the Dirichlet problem cannot have a solution such that (3.6) holds necessarily for all the boundary points ∂D. From physics we know that there exist examples of boundary points which are in some sense irregular. The probabilistic treatment of this fact is the following

Definition 3.12 (Regular Points)**.** A point $y \in \partial D$ is called X – *regular* for D (or simply regular) if $P^y(\tau_D = 0) = 1$.

It might seem that irregular points are all those y which have $0 \leq P^y(\tau_D = 0) < 1$. Due to the Blumenthal zero – one law, see [16] p.94, however, we have that if a point is not regular, than $P^y(\tau_D = 0) = 0$.

Example 3.13 (Regular Points). All points in one dimension, degenerate processes left aside, are regular. Even in two – dimensions, irregular points are constructed rather artificially, see for example the punctured disc example in [20], p.187. Once we have a space – time process, the time component of course runs in one direction only and hence irregular points appear too. In dimension three, however, an interesting behaviour can occur. One famous example is called a Lebesgue's thorn (spine). Vaguely speaking, it postulates that a potential on a sharp enough edge can not be computed via averaging over the potential in the vicinity, as an intuition in our probability approach could suggest. For details, refer to [16], p. 249.

When is a solution to a Stochastic Dirichlet Problem also a solution to the Generalized Dirichlet Problem, where the condition (3.6) is required to hold for regular points of the boundary only? It requires a decent knowledge of partial differential equations to follow the thoughts in [20], pp.188–190, where a partial answer to such a question is given. First it must be determined, under what conditions it is true that if the Generalized Dirichlet Problem has a solution, it is the one obtained as a solution to the stochastic version. Then further assumptions have to be made to ensure the existence of a solution to the Generalized Dirichlet Problem. For our purposes it suffices to know, that once the operator L in (3.5) is uniformly elliptic, i.e. all its eigenvalues are bounded away from zero in D, then:

$$\begin{aligned} u(x) = \mathrm{E}[\phi(X_{\tau_D})] \quad & \text{is a solution to the Generalized Dirichlet Problem in D, i.e.} \\ Lu(x) = 0 \qquad & \forall x \in D; \\ \lim_{x\to y} u(x) = f(y) \qquad & \forall x \in D; \quad \text{and for all regular } y \in \partial D. \end{aligned} \tag{3.11}$$

The general approach chosen embraces even the parabolic – type equations were the operator L is *elliptic*. For a rather robotic, but systematic and quite intuitive exposure of the similarities and differences between elliptic and parabolic equations, see chapter 4 in [9].

Thus, treating both elliptic and parabolic equations at once, we talk about all partial differential equations to be found throughout the field of financial mathematics. Emphasis is put on the fact, that we can work with the graphs of diffusions of interest. In elementary finance, what we are interested in, is pricing an agreement which depends on the evaluation of a risky price of a financial asset only, or it may depend on time too, or in the most realistic case, it depends also on a comparison of the return to a risk – free interest rate. How do the generators of the processes of a price, of a space – time graph of the price, and of the space – time graph of the discounted price look like? Fix $[x, T, d] \in \mathbf{R}_+ \times \mathbf{R} \times \mathbf{R}$ and define $D_t = d - \int_0^t r(X_s)\mathrm{d}s$. First, let the financial agreement be such that its value is a function $f \in C_0^2(\mathbf{R})$ of a price X_t, which is a solution to (2.1). We know, that a generator of such a process for such an f coincides with the differential operator L in (3.1) in dimension one:

$$Lf(x) = \left[\mu(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x)\right]. \tag{3.12}$$

It is obtained via Itô formula, that the generator of the space – time graph process $Y_t = [X_t^x, T - t]$ is for $\phi \in C_0^2(\mathbf{R}^2)$:

$$\hat{L}\phi(x,t) = L\phi(x,t) - \dot{\phi}(x,t) = \left[\mu(x)\phi'(x,t) + \frac{1}{2}\sigma^2(x)\phi''(x,t) - \dot{\phi}(x,t)\right], \tag{3.13}$$

and by analogous computation, we also obtain a generator of $Z_t = [X_t^x, T - t, D_t]$, for $\varphi \in C_0^2(\mathbf{R}^3)$:

$$\begin{aligned}\tilde{L}\varphi(x,t,d) &= \hat{L}\varphi(x,t,d) - r(x)\frac{\partial}{\partial d}\varphi(x,t,d)\\ &= \mu(x)\varphi'(x,t,d) + \frac{1}{2}\sigma^2(x)\varphi''(x,t,d) - \dot{\varphi}(x,t,d) - r(x)\frac{\partial}{\partial d}\varphi(x,t,d),\end{aligned}$$

which for a special $\varphi(x,t,d) \equiv e^d\phi(x,t)$ has a familiar form:

$$\tilde{L}e^d\phi(x,t) = e^d\left[\mu(x)\phi'(x,t) + \frac{1}{2}\sigma^2(x)\phi''(x,t) - \dot{\phi}(x,t) - r(x)\phi(x,t)\right], \tag{3.14}$$

where as usual: $f'' = \frac{\partial^2 f}{\partial x^2}$; $f' = \frac{\partial f}{\partial x}$; $\dot{f} = \frac{\partial f}{\partial t}$.

With this in mind, it is easy to conclude that by Theorem 3.8 we have also established uniqueness of solutions to (a version of) the Kolmogorov backward equation. This equation in its simplest form is usually studied to analytically obtain transition densities, see for example [16], pp. 282–283. The theorems treating existence and uniqueness of solutions to the Kolmogorov backward equation are named after the pioneers Feynman and Kac. Because in (3.11) we have existence for uniformly elliptic operators only and hence we haven't included the space – time process, we shall concentrate below on the more intuitive case of the Kolmogorov backward equation, where time and space variables are treated separately. We prove a version which requires only a compilation of facts we have already treated. For more general versions, consult for example [10], pp. 144–147. For the sake of completeness we state one of the more powerful versions in Remark 3.15.

Theorem 3.14 (Feynman – Kac)**.** *Let $f \in C_0^2(\mathbf{R}^n)$ and $r \in C(\mathbf{R}^n)$ positive. Then there exists a solution to*

$$\dot{u}(x,t) = Au(x,t) - r(x)u(x,t); \qquad t > 0, \quad x \in \mathbf{R}^n \tag{3.15}$$

$$u(x,0) = f(x); \qquad x \in \mathbf{R}^n, \tag{3.16}$$

and this solution can be represented as:

$$u(x,t) = E^x\left[\exp\left(-\int_0^t r(X_s)\mathrm{d}s\right) f(X_t)\right]. \tag{3.17}$$

If, moreover, there is a bounded solution $v(x,t) \in C^{2,1}(\mathbf{R}^n \times \mathbf{R})$ *to* (3.15)*, than* $v(x,t) = u(x,t)$*.*

Proof. Since we meet the assumptions of Dynkin formula 3.6, we will borrow facts from its proof and verify that $u(x,t)$ is a solution. We know that $u(x,t)$ is differentiable with respect to t. Further, we use the strong Markov property 3.2 of Itô diffusions to compute:

$$\begin{aligned}
&\frac{1}{r}\left(E^x[u(X_r,t)] - u(x,t)\right) \\
&\quad = \frac{1}{r}E^x\left\{E^x\left[f(X_{t+r})\exp(-\int_r^{t+r} r(X_s)\mathrm{d}s)|\mathcal{F}_t\right] - E^x\left[\exp\left(-\int_0^t r(X_s)\mathrm{d}s\right)f(X_t)\right]\right\} \\
&\quad = \frac{1}{r}\left([u(x,t+r)-u(x,t)] + E^x\left\{\exp\left(-\int_0^{t+r} r(X_s)\mathrm{d}s\right) f(X_{t+r})\left[\exp\left(\int_0^r r(X_s)\mathrm{d}s\right) - 1\right]\right\}\right).
\end{aligned}$$

Since the very last term under expectation is bounded, we can switch limit and expectation to conclude, that:

$$\frac{1}{r}\left(E^x[u(X_r,t)] - u(x,t)\right) \to \frac{\partial}{\partial t}u(x,t) + r(x)u(x,t).$$

This proves that $u(x,t)$ really is a solution to (3.15). To prove the uniqueness part, we turn to Dynkin formula again. Let $v(x,t)$ be the bounded solution to (3.15). Remember that, for Z_t defined above equation (3.14) and for $\varphi(x,t,d) = e^d v(x,t)$, by (3.14) and (3.15) we have:

$$A_Z\varphi(x,t,d) = \tilde{L}e^d v(x,t) = 0,$$

and it follows from Dynkin formula, that for every $R > 0$, if we set $\tau_R = \inf\{t : |Z_t| > R\}$:

$$E^{x,s,0}\left[\int_0^{t\wedge\tau_R} \tilde{L}\varphi(Z_r)\mathrm{d}r\right] + \varphi(x,s,0) = E^{x,s,0}[\varphi(Z_{t\wedge\tau_R})];$$

$$\varphi(x,s,0) = E^{x,s,0}\left[\exp\left(-\int_0^{t\wedge\tau_R} r(X_r)\mathrm{d}r\right) v(X_{t\wedge\tau_R}, s-(t\wedge\tau_R))\right].$$

Hence, because $v(x,t)$ is bounded, letting $R \to \infty$ and choosing $s = t$:

$$\begin{aligned}
v(x,s) = \varphi(x,s,0) &= E^{x,s,0}\left[\exp\left(-\int_0^s r(X_r)\mathrm{d}r\right) v(X_s,0)\right] \\
&= E^{x,s,0}\left[\exp\left(-\int_0^s r(X_r)\mathrm{d}r\right) f(X_s)\right] = u(x,s).
\end{aligned}$$

□

Remark 3.15 (Feynman – Kac formula II). The case where the initial and boundary conditions are mixed should be at least briefly mentioned. Note that the form of the solution is highly instructive in how the importance of the initial and boundary conditions varies as we change the width or length of the time – space cylinder. Here, to avoid discussion of regularity of the boundary, assume D bounded with a C^2 boundary ∂D. Let L be as in (3.1). Consider the mixed initial – boundary value problem:

$$\begin{aligned}
\dot{u}(x,t) &= Lu(x,t) - r(x)u(x,t) + \varphi(x,t); \qquad T > t \geq 0, \quad x \in D, \\
u(x,T) &= f(x); \qquad x \in D, \\
u(x,t) &= \phi(x,t); \qquad T > t \geq 0, \quad x \in \partial D.
\end{aligned} \tag{3.18}$$

Then under suitable conditions, see [10] p.146, there is a unique solution to (3.18) which is given by:

$$\begin{aligned} u(x,t) = \mathrm{E}^{x,t}&\left\{\phi\left(X_\tau,\tau\right)\exp\left[-\int_t^\tau r(X_s,s)\mathrm{d}s\right]\mathbf{1}_{[\tau<T]}\right\} \\ +\mathrm{E}^{x,t}&\left\{f\left(X_T\right)\exp\left[-\int_t^\tau r(X_s,s)\mathrm{d}s\right]\mathbf{1}_{[\tau=T]}\right\} \\ +\mathrm{E}^{x,t}&\left\{\int_t^\tau \varphi\left(X_s,s\right)\exp\left[-\int_t^s r(X_r,r)\mathrm{d}r\right]\mathrm{d}s\right\}. \end{aligned}$$

where as usual, τ is the first time $[X_t, t]$ leaves $D \times [0, T)$.

3.3 Financial Options

Option contracts, or simply options, are standard examples of derivative securities. In this text, we usually consider *put* options, which provide their holder with a right to sell the underlying (thereby exercise the option) at a predetermined price, referred to as *strike* or an *exercise price*, at (or before) a prespecified future time called *expiry*. Options which can be exercised any time before expiry are called *American*, otherwise they are *European*.There are two main concepts used in option pricing theory: the so – called *partial differential equation, or PDE* pricing and the *risk neutral or martingale* valuation.

The PDE pricing is an older approach and was originally developed for European type options by hedging, arbitrage and self – financing arguments with little rigourous probability in the seminal paper [2] by Fisher Black and Myron Scholes. This paper surprisingly does not even mention Itô formula, since it works with infinitely small increments and uses a vague argument only referring to a stochastic integrals manuscript by Henry McKean. This argument made precise is, however, actually a part of the proof of the Itô formula. It simply says that, due to infinite variation of Brownian motion, we cannot neglect the second derivative with respect to x in the Taylor expansion. Robert Merton, who worked closely with Black and Scholes and discussed their yet unpublished article with them, did indeed put the whole idea into a continous – time stochastic analysis framework in his article [19] (which was published actually a few weeks earlier than [2]). The result of Black and Scholes can be, omitting all the assumptions which can be found in the original article, formulated as follows:

Theorem 3.16. *Assume an investor can invest into a riskless bond D_t or a risky asset X_t, whose prices are driven by the following one – dimensional stochastic differential equations respectively:*

$$\begin{aligned} \mathrm{d}D_t &= rD_t\mathrm{d}t; \qquad D_T = 1; \\ \mathrm{d}X_t &= \mu X_t\mathrm{d}t + \sigma X_t \mathrm{d}B_t; \qquad X_0 > 0. \end{aligned} \tag{3.19}$$

Then the price C of a European call option on the asset X_s, at time t prior to maturity T, where $X_t = x$ and the strike is K, is a solution to the following second – order parabolic differential equation:

$$\begin{aligned} rxC'(x,t) + \frac{1}{2}\sigma^2 x^2 C''(x,t) + \dot{C}(x,t) - rC(c,t) = 0; \qquad (x,t) \in [0,T) \times (0,\infty); \\ C(x,T) = (x-K)^+; \qquad C(0,t) = 0. \end{aligned} \tag{3.20}$$

This solution exists, is unique and can be obtained analytically.

This approach can be extended to accommodate more general price processes or initial – boundary conditions. One can then turn to numerical schemes to solve the PDEs. For the flag – ship of this approach, refer to [26].

Martingale approach, on the other hand, followed three years later in [7] by John Cox and Stephen Ross. We already mentioned it in connection to the Girsanov theorem 1.17. Cox and Ross realized that Black and Scholes reasoning implies that pricing the options should not depend on the risk preferences of investors involved, and deduced that option pricing formulas can be easily computed as expectations in a risk – neutral world. Their insight allowed them to compute an option price in the setting (3.19) by:

$$C(x,t) = \tilde{\mathrm{E}}^{x,t}\left[e^{-r(T-t)}(X_T-K)^+\right], \tag{3.21}$$

where $\tilde{\mathrm{E}}$ means computing expectation in a risk – neutral world, such a world where $D_{T-t}^{-1}X_t$ is a martingale. In a more general setting without known interest rates, a term *risk – neutral* world can be rather unclear and one has to stick to a risk premium theory. That is out of our scope, however, once it is agreed on what the risk premium should look like, Girsanov theorem 1.17 applies. The martingale approach leads to computing expected values. Of course, usually this is not analytically feasible and we have to employ Monte – Carlo simulation methods.

Clearly, the theory developed in the previous section says that the prices computed either way are exactly the same, since the solution to the initial value problem (3.20) can be represented by (3.21). Virtually any option pricing involves a Dirichlet problem. The underlying PDE or the underlying stochastic process are chosen from a few models of which the geometric Brownian motion is by far the most popular. The boundary conditions, i.e. the conditions on the payoff of the options, can vary a lot. Consider, for example, the most simple barrier options, which earn a payoff, or on the contrary go worthless, upon the underlying hitting a certain barrier. This barrier is usually agreed upon in the beginning, but generally can be time – varying. Valuing such an option only means adding a boundary condition to the already stated problem and hence solving a mixed initial – boundary problem with the same PDE.

American options are different, since the boundaries appearing in the valuation procedure are not known a priori. When accommodating Black and Scholes hedging arguments, we have to be ready to pay off the option at all times. It suffices, however, to compute the initial capital needed to hedge oneself against an exercise at an *optimal exercise time*. This naturally leads to, as shown in [15], studying a demanding field of optimal stopping problems, the Dirichlet problem per se is not enough. A notationally and technically quite difficult treatment of optimal stopping can be found in [20], chapter 10. However, the author finds some articles directly connected to American options more accessible, for example the already cited [15] or its successor [14] form a great start to the field. These articles were the pathfinders to price the American put. Their authors build on a geometric Brownian motion model, which simplifies the computations and allows a derivation of some quasi – analytical formulas. These divide the price of an American put option into two terms. The first is the price of an equivalent European put option and the second is the so – called *early exercise premium*.

It should be noted, that a good introduction to American put option pricing may be provided by perpetual American put options, which are theoretical contracts whose time to maturity is always infinity. Despite the fact that such contracts are not traded they are very interesting – also because in some models their price can be completely determined analytically. The price of an American put when the underlying asset follows a geometric Brownian motion can be found in [25], p.351.

Finite expiry American put options from a much more difficult problem. Since the decision depends on time to maturity too, it is not a priori clear where will the exercise level lay. Jacka has shown in [14], that the optimal stopping problem gives rise to a so – called *free boundary* problem and proved that this particular problem has a unique solution, which has an analytically tractable lower boundary. Unfortunately, pricing options even in the most simple Black – Scholes framework seems to belong to the majority of free – boundary problems whose solutions cannot be found analytically and must be approximated numerically. These discretization procedures can again be divided into two categories, numerically solving the parabolic free boundary PDE or iteratively solving the optimal stopping problem via simulation.

3.4 Some Approximations to Price American Put

To see, how those discretization approaches can be employed in practice, we applied two algorithms to be found in the literature: a well known numerical *Implicit Finite Difference Method, IFD*, taken from [12], pp.420–423, and a quite recent simulation method called by its authors Longstaff and Schwartz *Least – Squares Monte Carlo, LSMC*, which originates from the instantly famous article [18]. The algorithms will not be given here, the reader is referred to the original articles. The algorithms are quite easy to implement, the only impediment could of course be the time and space efficiency. As pointed out in Appendix in Section A.2, the source code given there tries to address the time efficiency, the space efficiency would have to be treated only for extremely long maturities.

It should be highlighted that the use of the LSMC algorithm for pricing options with Markovian payoff functions is not very reasonable. Since the simulation methods often collapse on not good enough random number generators, it may be needed to use quite many realizations with different seeds to obtain, via averaging, good estimates of expected values. What results from sampling many paths of course cannot compete with usually very quickly converging Implicit Finite Differences. However, this disadvantage of LSMC is more than offset, when we have to value a path – dependent claim. It is not known to the author how to accommodate the numerical schemes to such a situation. In contrast, the LSMC algorithm doesn't have to be adjusted at all. In the following Figure 3.4 we show a perfect match of the prices once the number of simulated paths was risen to 3.000 as well as a fine graph of prices of an American put computed via IFD even on quite a scarce grid.

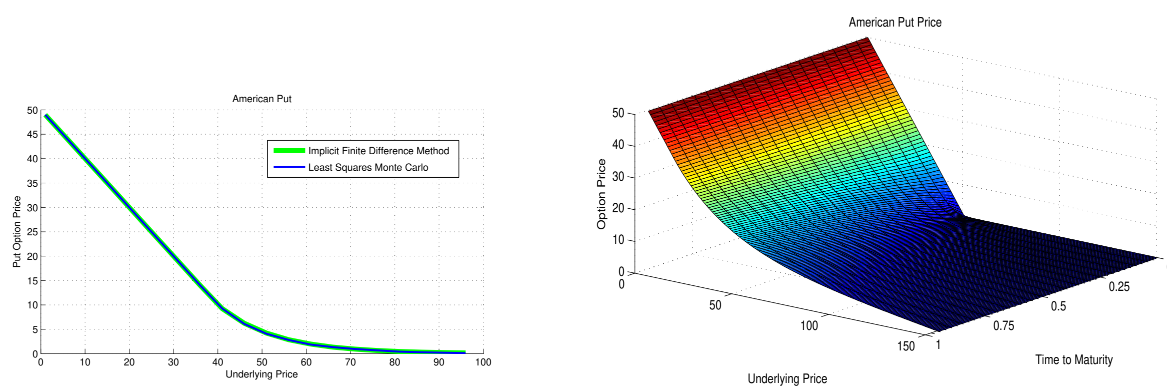

Figure 3.1: The left figure is a comparison of American put prices computed via LSMC and IFD. The right figure shows American put prices as a function of both the underlying and the time to maturity. Source: Author's computations A.2.

Chapter 4

The Case of Cox – Ingersoll – Ross Model

The *Cox – Ingersoll – Ross* model used to be one of the most favourite interest rate models. The parameters of the model can be set so that the model possesses many favourable features, but it also has a few modeling shortfalls, the major one being the impossibility to fit the model to observed yield curves. This caused the practitioners to replace it with some more general multi – factor models. However, the model still plays an important role in modeling volatility, this time named after Steven Heston. For practical issues connected to Cox – Ingersoll – Ross model, please see the third chapter of [3]. The original paper in which the model appeared related to finance is [5], where the authors solved an agent's optimal consumption and investment problem. This article was published together with its direct extension [6], which further analyzes the model properties and computes some relevant financial results, see Section 4.3.

4.1 Basic Properties of the Cox – Ingersoll – Ross Model

In this section we will use the results obtained in the preceding chapters and some other martingale techniques to explore the properties of the Cox – Ingersoll – Ross model. We concentrate on as a solution to:

$$\mathrm{d}X_t = (a + bX_t)\mathrm{d}t + c\sqrt{|X_t|}\mathrm{d}B_t; \quad X_0 = x > 0; \tag{4.1}$$

where $a > 0$; $b < 0$; $c > 0$ and $2a/c^2 \geq 1$. For financial applications it makes sense to restrict the coefficients this way and remove the absolute value a posteriori, see below. Since the coefficients meet the assumptions of Theorem 2.8, for every given initial value $x \geq 0$ the equation (4.1) admits a global strong pathwise unique solution, thereafter often referred to as the *CIR process*. We would like to compute the Laplace transforms of the process and its first exit times. First, however, to see that with these coefficients the process may be financially relevant as a model of interest rates, volatility etc., we explore attainability of zero and see that the coefficients are such that the process is strictly positive.

First, we slightly digress and stress an important fact which is used in [21] and explored further in this chapter – the fact that a CIR process can be represented with the help of the Girsanov theorem 1.17 or by a space – time transformation as a squared Bessel process of a dimension $4a/c^2$. Bessel processes of an integer dimension d are simply processes of an Euclidean norm of a d – dimensional Brownian motion. For Brownian motion it is well – known, that it reaches zero for every starting point x with a positive probability only when it is one – dimensional (the probability is actually equal to one). A Brownian motion in $\mathbf{R}^d$, $d \geq 2$, may get very close to the origin, but almost surely never hits it. Using different approach this is similar to (4.2) and (4.4), which say that $4a/c^2$ must be greater or equal to 2 for the CIR process to stay strictly positive.

Theorem 4.1. *Let $x > 0$ and set $e(\omega) = \inf\{t; X_t^x(\omega) = 0\}$. Then the following implications are valid:*

$$\textit{If } \frac{c^2}{2} > a \geq 0, \textit{ then } P^x(e < \infty) > 0; \tag{4.2}$$

$$\textit{If moreover } b \leq 0, \textit{ then } P^x(e < \infty) = 1. \tag{4.3}$$

$$\textit{If on the other hand } \frac{c^2}{2} \leq a, \textit{ then } P^x(e = \infty) = 1. \tag{4.4}$$

Proof. We will be using our developed machinery and apply Theorem 2.18 and 2.19. It is a straightforward algebra:

$$\begin{aligned}
s(x) &= \int_1^x \exp\left\{-\int_1^y \frac{2(a+bz)}{c^2 z}\mathrm{d}z\right\}\mathrm{d}y \\
&= \int_1^x \exp\left\{\frac{-2}{c^2}(a\log y + b(y-1))\right\}\mathrm{d}y = e^{\frac{2b}{c^2}}\int_1^x y^{\frac{-2a}{c^2}} e^{\frac{-2by}{c^2}}\mathrm{d}y \\
k(x) &= \int_1^x \exp\left\{-\int_1^y \frac{2(a+bz)}{c^2 z}\mathrm{d}z\right\}\left\{\int_1^y \exp\left[\int_1^z \frac{2(a+bu)}{c^2 u}\mathrm{d}u\right]\frac{2}{c^2 z}\mathrm{d}z\right\}\mathrm{d}y \\
&= \frac{2}{c^2}\int_x^1 e^{-\frac{2bz}{c^2}} z^{-\frac{2a}{c^2}}\left(\int_z^1 e^{\frac{2by}{c^2}} y^{\frac{2a}{c^2}-1}\mathrm{d}y\right)\mathrm{d}z
\end{aligned}$$

Now it is not demanding to see that $s(0+) = -\infty$ or $s(0+)$ finite according to whether $\frac{2a}{c^2} \geq 1$ or not. Also, $s(\infty) = \infty$ in the cases when $b < 0$ or $b = 0$; $\frac{2a}{c^2} \leq 1$. Moreover, if $\frac{2a}{c^2} < 1$ then $k(0+) < \infty$. Now (4.2) follows from (ii) in Theorem 2.18, (4.3) follows from (3) in Theorem 2.19 and (4.4) follows from (i) in 2.18. □

Heuristically, when we want to model processes like interest rates or volatility, it makes sense to set $a > 0$ and $b < 0$ so that there is a positive level to which the process tends to revert (in the sense of remark 2.1). Also, because it makes little sense for such processes to reach the origin, the choice of $2a/c^2 \geq 1$ in (4.1) is justified.

It is impossible to derive a closed form solution to (4.1) as is noted in [25] pp. 151–153. The same book proposes a way to derive a distribution of the solution on pp. 286–288. We will head another direction, derive a Laplace transform of the solution and follow up with a computation of the first two moments of the CIR process. For that we shall employ a martingale approach again – being once more lead to one of the rare cases of partial differential equations that are solvable analytically. Consider the following equation:

$$\dot{u}(t,x) = Lu(t,x); \qquad u(0,x) = e^{-\lambda x}, \tag{4.5}$$

where $u \in C_b^{1,2}([0,\infty)\times[0,\infty))$, i.e. u is in $C^{1,2}$ and bounded. Remember this equation is called the *Kolmogorov backward equation* and we have seen it in the preceding chapter. Recall also that if

$Lu = \triangle u$, it reduces to a *heat equation*, which derives its name from the fact that if the units are suitably chosen, the temperature profile at 0 is given by $f(x)$ and $u(t,x)$ really gives the temperature at x at time t.

Lemma 4.2.

$$u(t,x) = \left[\frac{\lambda c^2}{2b}(e^{bt}-1)+1\right]^{-\frac{2a}{c^2}} \exp\left\{-\frac{\lambda e^{bt}x}{\frac{\lambda c^2}{2b}(e^{bt}-1)+1}\right\} \tag{4.6}$$

is a unique solution to the equation (4.5).

Proof. Uniqueness of solutions to equation (4.5) is treated in Theorem 3.14. We will find a solution in a manner similar to a few examples in [25], Chapter 6. We guess and subsequently verify that the solution is of the form

$$u(t,x) = h(t)e^{-g(t)x}. \tag{4.7}$$

It suffices to find functions $g(t), h(t)$ such that:

$$\frac{c^2}{2}h(t)g(t)^2 + (a+bx)(-g(t))(h(t)) = \dot{h}(t) + h(t)(-\dot{g}(t))x,$$

which reduces to a system:

$$\begin{aligned} &1) \quad -ag(t)h(t) = \dot{h}(t); \qquad h(0)=1,\\ &2) \quad \frac{c^2}{2}g(t)^2 - bg(t) = -\dot{g}(t); \qquad g(0)=\lambda. \end{aligned}$$

First we solve 2):

$$t = \ln\left(\frac{Kg(t)}{b-\frac{c^2}{2}g(t)}\right)^{1/b} \quad \Rightarrow \quad g(t) = \frac{be^{bt}}{K+\frac{c^2}{2}e^{bt}}; \quad g(0)=\lambda \quad \Rightarrow \quad K = \frac{b-\lambda c^2/2}{\lambda};$$

$$g(t) = \frac{\lambda e^{bt}}{\frac{\lambda c^2}{2b}(e^{bt}-1)+1}.$$

And we proceed to solve 1):

$$g(t) = -\frac{1}{a}\dot{\ln} h(t) \quad \Rightarrow \quad h(t) = \exp\left(-a\int_0^t g(s)\mathrm{d}s\right) = \left[\frac{\lambda c^2}{2b}(e^{bt}-1)+1\right]^{-\frac{2a}{c^2}}.$$

□

If u satisfies (4.5), then for a fixed but arbitrary $t_0 \in (0,\infty)$, $v(t,X_t) = u(t_0-t,X_t)$ is a continuous nonnegative uniformly integrable martingale on $[0,t_0)$ and:

$$\mathrm{E}\left[e^{-\lambda X_{t_0}}\right] = \mathrm{E}\left[\lim_{t\to t_{0_-}} u(t_0-t,X_t)\right] = \lim_{t\to t_{0_-}} \mathrm{E}\left[u(t_0-t,X_t) = u(t_0,x)\right]. \tag{4.8}$$

We have proved the main part of the following proposition 4.3. To complete the proof it only remains to apply the fact that we can easily compute moments of the process from its Laplace transform.

Proposition 4.3.

$$f^*(\lambda) \equiv \mathrm{E}\left[e^{-\lambda X_t}\right] = \left[\frac{\lambda c^2}{2b}(e^{bt}-1)+1\right]^{-\frac{2a}{c^2}} \exp\left\{-\frac{\lambda e^{bt}x}{\frac{\lambda c^2}{2b}(e^{bt}-1)+1}\right\}; \tag{4.9}$$

$$\mathrm{E}[X_t] = -f^{*'}(0) = \left(e^{bt}x + \frac{a}{b}(e^{bt}-1)\right);$$

$$var(X_t) = f^{*''}(0) - \left(f^{*'}(0)\right)^2 = \frac{a}{2}\left(\frac{c}{b}(e^{bt}-1)\right)^2 + \frac{xc^2}{b}(e^{2bt}-e^{bt}).$$

Proof. It remains to compute the derivatives of $f^*(\lambda)$. For that, we denote

$$g(\lambda) \equiv \left[\frac{\lambda c^2}{2b}(e^{bt_0}-1)+1\right]^{-\frac{2a}{c^2}};$$

$$h(\lambda) \equiv \frac{\lambda e^{bt_0}x}{\frac{\lambda c^2}{2b}(e^{bt_0}-1)+1}.$$

Clearly, $g(0)=1$, $h(0)=0$. Also:

$$f^{*'}(\lambda) = e^{-h(\lambda)}(g'(\lambda) - g(\lambda)h'(\lambda));$$

$$f^{*''}(\lambda) = e^{-h(\lambda)}(g''(\lambda) - 2g'(\lambda)h'(\lambda) + g(\lambda)[(h'(\lambda))^2 - h''(\lambda)]);$$

$$g'(\lambda) = -\frac{a}{b}(e^{bt}-1)\left(\frac{\lambda c^2}{2b}(e^{bt}-1)+1\right)^{-\frac{2a}{c^2}-1}; \qquad g'(0) = -\frac{a}{b}(e^{bt}-1);$$

$$g''(\lambda) = \frac{ac^2}{2b^2}(e^{bt}-1)^2\left(\frac{\lambda c^2}{2b}(e^{bt}-1)+1\right)^{-\frac{2a}{c^2}-2}\left(\frac{2a}{c^2}+1\right); \qquad g''(0) = \frac{ac^2}{2b^2}(e^{bt}-1)^2\left(\frac{2a}{c^2}+1\right);$$

$$h'(\lambda) = \frac{e^{bt}x}{\left(\frac{\lambda c^2}{2b}(e^{bt}-1)+1\right)^2}; \qquad h'(0) = e^{bt}x;$$

$$h''(\lambda) = \frac{-2e^{bt}x}{\left(\frac{\lambda c^2}{2b}(e^{bt}-1)+1\right)^3}\frac{c^2}{2b}(e^{bt}-1); \qquad h''(0) = -2e^{bt}x\frac{c^2}{2b}(e^{bt}-1).$$

We have therefore computed that:

$$f^{*'}(0) = -\left(xe^{bt} + \frac{a}{b}(e^{bt}-1)\right);$$

$$\begin{aligned} f^{*''}(0) &= (xe^{bt})^2 + \left(\frac{a^2}{b^2} + \frac{ac^2}{2b^2}\right)\left(e^{bt}-1\right)^2 + \left(\frac{2ax}{b} + \frac{xc^2}{b}\right)e^{bt}(e^{bt}-1) \\ &= \left(xe^{bt} + \frac{a}{b}(e^{bt}-1)\right)^2 + \frac{a}{2}\left(\frac{c}{b}(e^{bt}-1)\right)^2 + \frac{xc^2}{b}(e^{2bt}-e^{bt}), \end{aligned}$$

and the assertions of the proposition easily follow. □

Remark 4.4. Until now, to be able to use the theorems proved in Chapter 2, Section 2.2, it was needed to assume that x, the starting point of our CIR diffusions, was strictly positive. With these assumptions we were able to prove strict positiveness of the whole process. Nonetheless, this will not be enough in the next section, where we require strict positiveness of the sample paths $\forall t > 0$ a.s. even in the case of $x = 0$. With help of the expression (4.9), however, we can also show that

$$P(X_t^0 > 0 \quad \forall t > 0) = 1. \tag{4.10}$$

Indeed, for a fixed but arbitrary $t > 0$:

$$\lim_{\lambda\to\infty} \mathrm{E}\left[e^{-\lambda X_t^0}\right] = \lim_{\lambda\to\infty}\left[\mathrm{E}(e^{-\lambda X_t^0}|X_t^0 \leq 0)P(X_t^0 \leq 0) + \mathrm{E}(e^{-\lambda X_t^0}|X_t^0 > 0)P(X_t^0 > 0)\right]$$
$$= \lim_{\lambda\to\infty}\left[\frac{\lambda c^2}{2b}(e^{bt_0}-1)+1\right]^{-\frac{2a}{c^2}} \exp\left\{-\frac{\lambda e^{bt_0}0}{\frac{\lambda c^2}{2b}(e^{bt_0}-1)+1}\right\} = 0;$$

which gives $P(X_t^0 \leq 0) = 0$. Because of (4.4) we immediately conclude that

$$P(X_{s+t}^0 > 0 \quad \forall s \geq 0) = 1.$$

Since t was arbitrary, we have proved (4.10).

4.2 First Exit Times of CIR Processes

In the following section, we move to the computation of the Laplace transform of the first exit time of a CIR process. We closely follow and develop the ideas in [11] and [21]. Göing – Jaeschke and Yor in [11] build on [21] to derive a Laplace transform of hitting times of a radial Ornstein – Uhlenbeck process. Clearly, it is a trivial observation that if we have a Laplace transform of the time a radial Ornstein – Uhlenbeck process reaches x, we have a Laplace transform of the time a *squared* radial Ornstein – Uhlenbeck process reaches x^2. By scaling it is straightforward to move from squared radial Ornstein – Uhlenbeck processes to CIR processes. Therefore, we could have used the results presented by Göing – Jaeschke and Yor in [11]. Nevertheless, we rather slightly modify and extend their approach to directly derive the hitting time distribution of a CIR process. Let us start with the following very useful representation, which was mentioned in the beginning of this chapter. For the origins of this representation please refer to [21] and references cited therein.

Proposition 4.5. *Let Y_t be a squared Bessel process of a dimension $\delta = 4a/c^2$, which means*

$$\mathrm{d}Y_t = 4a/c^2\mathrm{d}t + 2\sqrt{Y_t}\mathrm{d}B_t; \qquad Y_0 = x; \quad t \geq 0. \tag{4.11}$$

Define X_t as

$$X_t = e^{bt}Y_{\left(\frac{c^2}{4b}(1-e^{-bt})\right)} \tag{4.12}$$

Then X_t satisfies the equation (4.1)*, in other words, X_t is a CIR process with coefficients a, b and c.*

Proof. We have:

$$\mathrm{d}Y_t = \mu(Y_t)\mathrm{d}t + \sigma(Y_t)\mathrm{d}B_t \quad \Rightarrow \quad \mathrm{d}Y_{f(t)} = \mu(Y_{f(t)})f'(t)\mathrm{d}t + \sigma(Y_{f(t)})\mathrm{d}B_{f(t)}$$
$$= \mu(Y_{f(t)})f'(t)\mathrm{d}t + \sigma(Y_{f(t)})\sqrt{f'(t)}\mathrm{d}B_t.$$

Hence $Y_{\left(\frac{c^2}{4b}(1-e^{-bt})\right)}$ and X_t satisfy

$$\mathrm{d}Y_{\left(\frac{c^2}{4b}(1-e^{-bt})\right)} = ae^{-bt}\mathrm{d}t + 2\sqrt{Y_{\left(\frac{c^2}{4b}(1-e^{-bt})\right)}}\sqrt{\frac{c^2}{4}e^{-bt}}\mathrm{d}B_t;$$
$$\mathrm{d}X_t = \mathrm{d}\left(e^{bt}Y_{\left(\frac{c^2}{4b}(1-e^{-bt})\right)}\right) = \left(a + be^{bt}Y_{\left(\frac{c^2}{4b}(1-e^{-bt})\right)}\right)\mathrm{d}t + ce^{bt}\sqrt{Y_{\left(\frac{c^2}{4b}(1-e^{-bt})\right)}e^{-bt}}\mathrm{d}B_t$$
$$= (a + bX_t)\mathrm{d}t + c\sqrt{X_t}\mathrm{d}B_t.$$

Since equations (4.1) and (4.11) admit strong pathwise unique solutions, we have established a correspondence between the two families of solutions. □

Our aim is to find an explicit form of

$$\mathrm{E}\left[e^{-\lambda\tau_y^x}\right]; \qquad \text{where } \tau_y^x \equiv \inf\{t \geq 0 : X_t^x = y\}; \quad x, y \geq 0;$$

which we search for as a Laplace transform of the first time y is hit by the process (4.12). Analogously to [11] and to a few examples above, we are eager to find a function $g(\lambda, x)$ such that $Z_t = g(\lambda, X_t)e^{-\lambda t}$ is a local martingale with respect to the filtration generated by $Y_{\left(\frac{c^2}{4b}(1-e^{-bt})\right)}$. To be able to use Theorem 1.11 it is enough to check that $\mathrm{E}\langle Z\rangle_{\tau_y^x}$ is almost surely finite. Once we verify this, by the Optional Sampling Theorem 1.7 it holds that

$$\mathrm{E}\left[g(\lambda, X_{\tau_y^x}^x)e^{-\lambda\tau_y^x}\right] = g(\lambda, x), \tag{4.13}$$

which yields a clear motivation to compute $g(\lambda, x)$ explicitly:

$$\mathrm{E}\left[e^{-\lambda\tau_y^x}\right] = \frac{g(\lambda, x)}{g(\lambda, y)}. \tag{4.14}$$

If we set $u = \frac{c^2}{4b}(1 - e^{-bt})$, we would get $e^{-\lambda t} = (1 - 4bu/c^2)^{\lambda/b}$ and $X_t = (1 - 4bu/c^2)^{-1}Y_u$, therefore it is equivalent to find g such that $g(\lambda, (1 - 4bu/c^2)^{-1}Y_u)(1 - 4bu/c^2)^{\lambda/b}$ is a local martingale with respect to the filtration generated by Y_u, or g such that $g(\lambda, (1-4bu/c^2)^{-1}R_u^2)(1-4bu/c^2)^{\lambda/b}$, where R_u is a Bessel (Bessel, not *squared* Bessel) process, is a local martingale with respect to the filtration generated by R_u. The latter form means, that

$$h(r, t) = g(\lambda, (1 - 4bt/c^2)^{-1}r^2)(1 - 4bt/c^2)^{\lambda/b}$$

via Itô formula 1.14 solves a partial differential equation (parabolic with an elliptic operator L or semi – elliptic when we consider $\hat{L}$):

$$\hat{L}h(r, t) = Lh(r, t) + \dot{h}(r, t) = \frac{1}{2}h''(r, t) + \frac{\delta - 1}{2r}h'(r, t) + \dot{h}(r, t) = 0; \tag{4.15}$$

where as usual: $h''(r, t) = \frac{\partial^2 h(r,t)}{\partial r^2}$; $h'(r, t) = \frac{\partial h(r,t)}{\partial r}$; $\dot{h}(r, t) = \frac{\partial h(r,t)}{\partial t}$ and $\delta = 4a/c^2$. That is, recalling Definition 3.9 and Lemma 3.10, $h(r, t)$ is an $\hat{R}_t$-harmonic function, where $\hat{R}_t$ is a a stochastic process of a graph of the Bessel process R_t. Let us denote $\nu = \delta/2 - 1 = 2a/c^2 - 1$. We pursue an idea to which we are indebted to Göing – Jaeschke and Yor in [11] p.323, where they search for $h(r, t)$ in a specific form:

$$h(r, t) = \Gamma(\nu + 1)2^{\nu} \int_0^{\infty} I_{\nu}(ur)e^{-u^2 t/2}r^{-\nu}f(u)\mathrm{d}u, \tag{4.16}$$

where f is an L^1- function, I_ν is a modified Bessel function of the first kind with an index ν:

$$I_{\nu}(z) = \left(\frac{z}{2}\right)^{\nu} \sum_{k=0}^{\infty} \frac{\left(\frac{z}{2}\right)^{2k}}{k!\Gamma(\nu + k + 1)} \tag{4.17}$$

and

$$\Gamma(z) = \int_0^{\infty} t^{z-1}e^{-t}\mathrm{d}t = \frac{1}{z}\prod_{n=1}^{\infty} \frac{(1 + 1/n)^z}{1 + z/n} \tag{4.18}$$

is a classical Gamma function. It should therefore hold, that it is enough to search for $h(r,t)$ in the form:

$$h(r,t) = \varphi(\nu) \int_0^\infty \Psi_\nu(ur) e^{-u^2 t/2} r^{-\nu} f(u) \mathrm{d}u, \tag{4.19}$$

where φ is a continuous strictly positive function and Ψ_ν is a solution to modified Bessel equation (4.20):

$$x^2 \Psi''(x) + x\Psi'(x) - (x^2 + \nu^2)\Psi(x) = 0. \tag{4.20}$$

As we see in the following lemma, this is indeed the case.

Lemma 4.6. *Let Ψ_ν be a solution to modified Bessel equation* (4.20)*. Then $h(r,t)$ defined by* (4.19) *is a solution to* (4.15)*.*

Proof. We only have to compute the derivatives of $h(r,t)$:

$$\begin{aligned}
h'(r,t) &= \varphi(\nu) r^{-\nu-1} \int_0^\infty [(-\nu)\Psi_\nu(ur) + \Psi'_\nu(ur)ur] e^{-u^2 t/2} f(u) \mathrm{d}u; \\
h''(r,t) &= \varphi(\nu) r^{-\nu-2} \int_0^\infty [\nu(\nu+1)\Psi_\nu(ur) - 2\nu\Psi'_\nu(ur)ur + \Psi''_\nu(ur)(ur)^2] e^{-u^2 t/2} f(u) \mathrm{d}u; \\
\dot{h}(r,t) &= \varphi(\nu) r^{-\nu-2} \frac{1}{2} \int_0^\infty \Psi_\nu(ur)(-(ru)^2) e^{-u^2 t/2} f(u) \mathrm{d}u,
\end{aligned}$$

to see that

$$\begin{aligned}
&\frac{1}{2} h''(r,t) + \frac{\delta - 1}{2r} h'(r,t) + \dot{h}(r,t) = \\
&\quad = \frac{1}{2}\varphi(\nu) r^{-\nu-2} \int_0^\infty e^{-u^2 t/2} f(u) [\Psi''_\nu(ur)(ur)^2 + \Psi'_\nu(ur)(ur) - (\nu^2 + (ur)^2)\Psi_\nu(ur)] \mathrm{d}u \\
&\quad = 0,
\end{aligned}$$

which proves the assertion. □

There are two linearly independent strictly monotone convex solutions to (4.20). The afore mentioned modified Bessel function of the first kind denoted by I_ν and a modified Bessel function of the second kind, usually denoted by K_ν. I_ν is continuous on $[0, \infty)$ with limit equal to zero at 0 and growing exponentially at ∞. On the contrary, K_ν explodes at 0 and has a finite limit at ∞. To meet the desired boundedness of our local martingale $h(R_t, t)$ for $t \le \tau_y^x$, we therefore have to distinguish between two cases of exit times:

$$\tau_y^x \equiv \inf\{t \ge 0 : X_t^x = y\}; \quad 0 \le x < y < \infty; \tag{4.21}$$

$$\tau_y^x \equiv \inf\{t \ge 0 : X_t^x = y\}; \quad 0 < y < x < \infty. \tag{4.22}$$

It is easy to verify that in the first case we really have to use I_ν. This case we thoroughly analyze in all detail. In the second case we shall turn to replacing I_ν by K_ν and everything will follow analogously. By substitution we have the following:

$$g\left(\lambda, \frac{r^2}{1 - 4bt/c^2}\right) = (1 - 4bt/c^2)^{-\lambda/b} \Gamma(\nu+1) \left(\frac{2}{r}\right)^\nu \int_0^\infty I_\nu(ur) e^{-u^2 t/2} f(u) \mathrm{d}u;$$

$$g(\lambda, y) = \left(1 - \frac{4bt}{c^2}\right)^{-\frac{\lambda}{b} - \frac{\nu+1}{2}} \Gamma(\nu+1) \left(\frac{2}{\sqrt{y}}\right)^\nu \int_0^\infty I_\nu(\sqrt{y} z) e^{-v(z,t)} f\left(\frac{z}{\sqrt{1 - 4bt/c^2}}\right) \mathrm{d}z, \tag{4.23}$$

where $v(z,t) = \frac{z^2tc^2}{2(c^2-4bt)} = \frac{z^2c^2}{8b}\left(\frac{1}{1-4bt/c^2}-1\right)$. Denote $w(t) = (1-4bt/c^2)$. Clearly, the right hand side of (4.23) should depend on y, λ, ν, b and c only. In other words:

$$w(t)^{-\frac{\lambda}{b}-\frac{\nu+1}{2}}e^{-\frac{z^2c^2}{8bw(t)}}f\left(\frac{z}{\sqrt{w(t)}}\right)$$

should *not* depend on t. It is not difficult to guess that if we choose $f(x) = e^{\frac{xc^2}{8b}}x^{-\frac{2\lambda}{b}-\nu-1}k$, where k is an arbitrary constant, then it is true that $f \in L^1([0,\infty))$ and

$$w(t)^{-\frac{\lambda}{b}-\frac{\nu+1}{2}}e^{-\frac{z^2c^2}{8bw(t)}}f\left(\frac{z}{\sqrt{w(t)}}\right) = kz^{-\frac{2\lambda}{b}-\nu-1}.$$

Hence we can conclude that

$$g(\lambda,y) = k\Gamma(\nu+1)\left(\frac{2}{\sqrt{y}}\right)^{\nu}\int_0^{\infty} I_\nu(\sqrt{y}z)e^{\frac{z^2c^2}{8b}}z^{-2\lambda/b-\nu-1}\mathrm{d}z.$$

Now, to derive the expression for $g(\lambda,0)$, we must make an easy observation that

$$\begin{aligned}\lim_{\epsilon\downarrow 0}\frac{I_\nu(z\sqrt{\epsilon})}{\sqrt{\epsilon^\nu}} &= \lim_{\epsilon\downarrow 0}\left(\frac{z}{2}\right)^{\nu}\sum_{k=0}^{\infty}\frac{\left(\frac{\sqrt{\epsilon}z}{2}\right)^{2k}}{k!\Gamma(\nu+k+1)}\\ &= \left(\frac{z}{2}\right)^{\nu}\frac{1}{\Gamma(\nu+1)}.\end{aligned}$$

Then it is straightforward to see

$$\begin{aligned}g(\lambda,0) = \lim_{\epsilon\downarrow 0} g(\lambda,\epsilon) &= k\Gamma(\nu+1)2^\nu\int_0^{\infty}\left(\frac{z}{2}\right)^{\nu}\Gamma^{-1}(\nu+1)e^{\frac{z^2c^2}{8b}}z^{-2\lambda/b-\nu-1}\mathrm{d}z\\ &= k\int_0^{\infty}e^{\frac{z^2c^2}{8b}}z^{-2\lambda/b-1}\mathrm{d}z\\ &= \frac{k4b}{c^2}\int_0^{\infty}e^{-t}\left(-\frac{8bt}{c^2}\right)^{-\lambda/b-1}\mathrm{d}t\\ &= \frac{k}{2}\left(-\frac{8b}{c^2}\right)^{-\lambda/b}\Gamma\left(-\frac{\lambda}{b}\right).\end{aligned}$$

We have prepared everything for the consequential lemma:

Lemma 4.7. *The Laplace transform of the stopping time $\tau_y^0 \equiv \inf\{t \geq 0 : X_t^0 = y\}$; $0 < y < \infty$; i.e. of the first time the CIR process X_t^0 starting at 0 hits a positive level y is*

$$\mathrm{E}\left[e^{-\lambda\tau_y^0}\right] = \frac{1}{\phi(-\lambda/b,\nu+1;-2by/c^2)}, \tag{4.24}$$

where the confluent hypergeometric function $\phi(a,b;z)$ is defined in Definition A.2.

Proof. There is only little work to be done. The first equality below is implied by (4.14), which was thoroughly explained. The rest follows from an integral transform mentioned in [11] which originates from [1], section 4.16, equation (20) and from the definition of the Whittaker's function $M_{k,\mu}(z)$ in Definition A.4:

$$\begin{aligned}
\mathrm{E}\left[e^{-\lambda\tau_y^0}\right] &= \frac{g(\lambda,0)}{g(\lambda,y)} \\
&= \frac{\frac{1}{2}\left(-\frac{8b}{c^2}\right)^{-\lambda/b}\Gamma\left(-\frac{\lambda}{b}\right)}{\Gamma(\nu+1)(\frac{2}{\sqrt{y}})^{\nu}\int_0^\infty I_\nu(\sqrt{y}z)e^{\frac{z^2c^2}{8b}}z^{-2\lambda/b-\nu-1}\mathrm{d}z} \\
&= \frac{y^{\nu/2}\left(-\frac{\sqrt{-2b}}{c}\right)^{-2\lambda/b}\Gamma\left(-\frac{\lambda}{b}\right)2^{-2\lambda/b-\nu-1}}{\Gamma(\nu+1)\frac{\Gamma(-\frac{\lambda}{b})e^{-yb/c^2}\left(\frac{\sqrt{-2b}}{c}\right)^{-2\lambda/b-\nu-1}2^{-2\lambda/b-\nu-1}}{\sqrt{y}\Gamma(\nu+1)}} \\
&= \frac{\sqrt{y}^{\nu+1}\left(\frac{\sqrt{-2b}}{c}\right)^{\nu+1}e^{yb/c^2}}{M_{1/2(2\lambda/b+\nu+1),\nu/2}\left(\frac{-2by}{c^2}\right)} = \frac{1}{\phi(-\lambda/b,\nu+1;-2by/c^2)}.
\end{aligned}$$

□

The following theorem can be proved directly, or as an easy corollary of Lemma 4.7 and the Strong Markov Property of Itô diffusions 3.2. For $0 < x < y$ the following almost surely holds:

$$\tau_y^0 = \tau_x^0 + \tau_y^x, \tag{4.25}$$

where the stopping times τ_x^0 and τ_y^x are independent.

Theorem 4.8. *The Laplace transform of the stopping time $\tau_y^x = \inf\{t \geq 0 : X_t^x = y\}; 0 < x < y < \infty$; i.e. of the first time the CIR process X_t^x starting at x hits a positive level y, $y > x$, is*

$$\mathrm{E}\left[e^{-\lambda\tau_y^x}\right] = \frac{\phi(-\lambda/b,\nu+1;-2bx/c^2)}{\phi(-\lambda/b,\nu+1;-2by/c^2)}. \tag{4.26}$$

Proof. Because of (4.25) and lemma 4.7:

$$\begin{aligned}
\mathrm{E}\left[e^{-\lambda\tau_y^x}\right] = E\left[e^{-\lambda(\tau_y^0-\tau_x^0)}\right] &= \frac{E\left[e^{-\lambda\tau_y^0}\right]}{E\left[e^{-\lambda\tau_x^0}\right]} \\
&= \frac{\phi(-\lambda/b,\nu+1;-2bx/c^2)}{\phi(-\lambda/b,\nu+1;-2by/c^2)}.
\end{aligned}$$

□

We want to turn to the question when the CIR process X_t^x hits y, where $x > y > 0$. As was foreseen, we are in a different situation – before being killed at τ_y^x, the CIR process is bounded away from zero but we certainly have no control over the supremum. We therefore use a different construction of $g(\lambda, x)$ as is justified by Lemma 4.6. With

$$\hat{g}(\lambda,x) = k\Gamma(\nu+1)\left(\frac{2}{\sqrt{x}}\right)^{\nu}\int_0^\infty K_\nu(\sqrt{x}z)e^{\frac{z^2c^2}{8b}}z^{-2\lambda/b-\nu-1}\mathrm{d}z,$$

it is clear that if R_t is a Bessel process,

$$X_t^x \equiv \hat{g}(\lambda,(1-4bt/c^2)^{-1}R_t^2)(1-4bt/c^2)^{\lambda/b}$$

is a local martingale. Moreover, it is bounded for $0 < t < \tau_y^x$, where τ_y^x is as in (4.22). Therefore, inspired by (4.14), we may again turn to Theorem 1.11 and the Optional Sampling Theorem 1.7 to conclude that

$$\mathrm{E}\left[e^{-\lambda \tau_y^x}\right] = \frac{\hat{g}(\lambda, x)}{\hat{g}(\lambda, y)}. \tag{4.27}$$

It therefore requires only a knowledge of a few more special functions to believe the following theorem:

Theorem 4.9. *The Laplace transform of the stopping time $\tau_y^x = \inf\{t \geq 0 : X_t^x = y\}$; $0 < y < x < \infty$; i.e. of the first time the CIR process X_t^x starting at x hits a positive level y, $y < x$, is*

$$\mathrm{E}\left[e^{-\lambda \tau_y^x}\right] = \frac{\psi(-\lambda/b, \nu+1; -2bx/c^2)}{\psi(-\lambda/b, \nu+1; -2by/c^2)}. \tag{4.28}$$

where the confluent hypergeometric function of the second kind $\psi(a, b; z)$ is defined in Definition A.3.

Proof. The following steps proceed similarly to the proofs of the two propositions above. We make use of an integral representation in [1], section 4.16, equation (37), and the definition of the Whittaker's function of the second kind $W_{k,\mu}(z)$ as in Definition A.5.

$$\begin{aligned}
\mathrm{E}\left[e^{-\lambda \tau_y^x}\right] &= \left(\frac{y}{x}\right)^{\nu/2} \frac{\int_0^\infty K_\nu(\sqrt{x}z) e^{\frac{z^2 c^2}{8b}} z^{-2\lambda/b - \nu - 1} \mathrm{d}z}{\int_0^\infty K_\nu(\sqrt{y}z) e^{\frac{z^2 c^2}{8b}} z^{-2\lambda/b - \nu - 1} \mathrm{d}z} \\
&= \left(\frac{y}{x}\right)^{\frac{\nu+1}{2}} e^{\frac{b(y-x)}{c^2}} \frac{W_{1/2(2\lambda/b+\nu+1),\nu/2}\left(\frac{-2bx}{c^2}\right)}{W_{1/2(2\lambda/b+\nu+1),\nu/2}\left(\frac{-2by}{c^2}\right)} \\
&= \frac{\psi(-\lambda/b, \nu+1; -2bx/c^2)}{\psi(-\lambda/b, \nu+1; -2by/c^2)}.
\end{aligned}$$

□

The computations made, when applied, can be significantly simplified by utilizing approximations. The Gamma function can be reduced to an easy to compute form using *Stirling formula*, for example.

4.3 Options on CIR Assets

It is computationally quite lengthy, but not difficult, to price bonds and options on bonds in the CIR model. The computations do not differ too much from above and are easier than in the previous section, as well as they are already stated explicitly in the literature. To make the presentation about the CIR model complete, we cite the results.

A fair price of a zero – coupon bond in CIR model is to be found in [5]. A CIR bond European put option price is computed in [6]. There is even a quasi – analytical formula for a finite – expiration American put, which is computed using the same results from [21] that we used above, to be found in [4].

We can also study the properties of a CIR asset barrier option. Since we know the Laplace transform of the hitting time from theorems 4.8 and 4.9 and the conditional distribution of a CIR process is known too, it should be only a straightforward application of these results to compute the Laplace transform of the up – and – out call with respect to time, for example.

Appendix A

A.1 Special Functions

In this section we define some special functions used in Chapter 4. The importance of these functions has been well known for more than one hundred years. Since their definitions are somewhat lengthy, they were moved to the very end of the thesis so that the thesis remains self – contained but the definitions don't distract the reader from the main points of the text.

Definition A.1 (Rising Factorial)**.** We define the j-th rising factorial $(r)_j$ as:

$$(r)_0 = 1; \qquad (r)_j = \frac{\Gamma(r+j)}{\Gamma(r)} = r(r+1)\dots(r+j-1), \quad j = 1, 2, \dots$$

The functions whose definitions follow are two linearly independent solutions to the Kummer's differential equation:

$$\frac{\mathrm{d}^2 w}{\mathrm{d}z^2} + (b-z)\frac{\mathrm{d}w}{\mathrm{d}z} - aw = 0.$$

Remember, that this equation shows the operator $\tilde{L}$ from (3.14) in one of its most simple nontrivial cases.

Definition A.2 (Kummer's Confluent Hypergeometric Function)**.**

$$\phi(a, b; z) = \sum_{j=0}^{\infty} \frac{(a)_j z^j}{(b)_j j!}; \qquad \text{where} \quad b \neq 0, -1, -2, \dots$$

Definition A.3 (Tricomi's Confluent Hypergeometric Function of the Second Kind)**.**

$$\psi(a, b; z) = \frac{\Gamma(1-b)}{\Gamma(1+a-b)}\phi(a, b; z) + \frac{\Gamma(b-1)}{\Gamma(a)}\phi(1+a-b, 2-b; z).$$

On the other hand, Whittaker's functions are special solutions of Whittaker's equation, a modified form of the Kummer's equation introduced to make formulas involving the solutions more symmetric. Whittaker's equation is the following:

$$\frac{d^2 w}{dz^2} + \left(-\frac{1}{4} + \frac{\kappa}{z} + \frac{1/4 - \mu^2}{z^2}\right) w = 0.$$

The solutions make use of the confluent hypergeometric functions defined in A.2 and A.3

Definition A.4 (Whittaker's Function)**.**

$$M_{\kappa,\mu}(z) = \exp(-z/2)\, z^{\mu+\frac{1}{2}} \phi\left(\mu - \kappa + \frac{1}{2}, 1 + 2\mu; z\right)$$

Definition A.5 (Whittaker's Function of the Second Kind)**.**

$$W_{\kappa,\mu}(z) = \exp(-z/2)\, z^{\mu+\frac{1}{2}} \psi\left(\mu - \kappa + \frac{1}{2}, 1 + 2\mu; z\right)$$

A.2 Source Code

In this section a Matlab source code for the two algorithms mentioned in section 3.4 are provided. The implementation can only be improved on the line denoted by **, where the scarceness of the matrix *price* could be exploited. The code attempts to profit from the Matlab strength in handling matrices, therefore it is optimized to use as few loops as possible and handle boolean vectors or solve linear equations instead. The *.m files are provided as a part of the electronic version of this thesis.

The Longstaff – Schwartz algorithm Least – Squares Monte Carlo is implemented in the following function *lsamput_PZ*.

```
function lsap_price = lsamput_PZ(T, dT, K, r, s, x, nPath)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% American Put Option -- Longstaff Schwartz
% lsamput_PZ(T, dT , K, r, s, x, nPath)
% lsamput_PZ(1, .1, 100, 0.1, 0.9, 110, 10000)
%
% Computes the price of an American put by Longstaff -
% Schwartz: Least Squares Monte Carlo
%
% reference: Valuing American Options by Simulation : A Simple
% Least Squares Approach; Longstaff, Schwartz, 2001
%
% author: Petr Zahradnik, July 2010
%
% nPath                  : number of paths to simulate
% dT                     : time btwn exrcse points 'Bermuda like'
% T                      : maturity date
% x                      : initial price
% K                      : strike price
% r                      : risk free rate correspondiung to dT
% s                      : volatility corresponding to dT
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

nStep      = ceil(T/dT);     % no. of steps
dt = T/nStep;                % new timestep

S =  gbmpaths_PZ(r, s, x, T, nStep, nPath);
                             % get simulated GBM paths
```

```
price = zeros(nPath,nStep);  % initialize cashflow matrix
price(:,nStep) = max(0,K-S(:,nStep));  % put option payoff

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

for iStep = (nStep-1):-1:1   % price obtained backwards "Bermuda-like"

    y = max(0,K-S(:,iStep)); % payoff of put option
    itm = (y>0);             % where in-the-money
    n_itm = sum(itm);        % no. of in the money positions
    y_ex = y(itm);           % possible excercise value
    X   = S(itm,iStep);      % such underlying prices where put is itm
**  Y = (exp(-r*dt).^(1:(nStep-iStep))*price(itm,(iStep+1):nStep)')';
                             % discount the fut cashflows to time iStep
    cnst = ones(n_itm,1);
    L0X = exp(-X/2);         % weighted Laguerre polynomials
    L1X = L0X.*(1-X);
    L2X = L0X.*(1-2*X+X.^2/2);

    A = [cnst L0X L1X L2X ]; % basis functions Laguerre pols
    [U,W,V] = svd(A);        % Least-Square Regression
    b = V*(W\(U'*Y));        % b = (X'X)\X'y
    y_co = A*b;              % continuation values
    y_ex(y_ex < y_co) = 0;   % possible exc vals
                             % when not excercised become zero
    price(itm,iStep) = y_ex; % excercise rule, new prices into cash flow
    price(price(:,iStep)>0,(iStep+1):nStep) = 0;%if excercised, than later
                             % positive excercise cash flows impossible
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

lsap_price = sum(exp(-r*dt).^(0:(nStep-1))*price')/nPath;
```

The classical Implicit Finite Difference method is implemented in the following function *ifdamput_PZ*.

```
function ifdap_price = ifdamput_PZ(Smax, dS, T, dT, K, r, s)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% American Put Option -- Implicit Finite Differences
% ifdamput_PZ(Smax, dS, T, dT, K, r, s)
% ifdamput_PZ(100, 5, 1, 0.05, 50, 0.1, 0.5)
%
% Computes the price of an American put
% by implicit finite difference method
%
% reference : J. C. Hull, Options, Futures, and Other Derivatives
% 6th Ed., Chap 17, pp. 419-423
```

```
% author: Petr Zahradnik, July 2010
%
% Smax  : maximum stock price
% dS    : increment of stock price
% T     : maturity date
% dT    : time step
% K     : strike price
% r     : risk free rate
% s     : volatility
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

nPrice = ceil(Smax/dS);   % round toward positive infinity
nStep = ceil(T / dT);     % no. of prices and time steps
ds = Smax / nPrice;
dt = T / nStep;
J = 1:(nPrice-1);

a = .5*r*dt*J - .5*s^2*dt*J.^2;  % a,b,c as in the algorithm
b = 1 + s^2*dt*J.^2 + r*dt;      % .^2 matrix square
c = -.5*r*dt*J - .5*s^2*dt*J.^2;

A = diag(b) + diag(a(2:(nPrice-1)), -1) + diag(c(1:(nPrice-2)), 1);

price = zeros(nStep+1, nPrice+1);% includes one boundary condition
                                 % if S = Smax then price = 0
price(nStep+1, :) = max(K - (0:ds:Smax), 0);% second bndr condition
price(:, 1) = K;                 % third boundary condition

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

for iStep = nStep:-1:1
    y = price(iStep+1, 2:nPrice)';
    y(1) = y(1) - a(1)*K;        % solve the simultaneous equations
    price(iStep, 2:nPrice) = (A\y)'; % \ right divide matrix
    price(iStep, :) = max(K - (0:ds:Smax), price(iStep,:));
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

ifdap_price = price;
```

Bibliography

[1] H. Bateman and A. Erdélyi. *Tables of Integral Transforms; Vol. 1*. McGraw – Hill, New York, 1st edition, 1953.

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

[3] D. Brigo and F. Mercurio. *Interest Rate Models – Theory and Practice*. Springer, Berlin Heidelberg New York, 2nd edition, 2007.

[4] M. Chesney, R.J. Elliott, and R. Gibson. Analytical solutions for the pricing of American bond and yield options. *Mathematical Finance*, 3(3):277–294, 1993.

[5] J.C. Cox, J.E. Ingersoll, and S.A. Ross. An intertemporal general equilibrium model of asset prices. *Econometrica*, 53(2):363–384, 1985.

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

[7] J.C. Cox and S.A. Ross. The valuation of options for alternative stochastic processes. *Journal of Financial Economics*, (3):145–166, 1976.

[8] J. Dupačová, J. Hurt, and J. Štěpán. *Stochastic Modeling in Economics and Finance*. Kluwer Academic Publishers, Dordrecht Boston London, 1st edition, 2002.

[9] R. Durrett. *Stochastic Calculus*. CRC Press, Boca Raton New York London Tokyo, 1st edition, 1996.

[10] A. Friedman. *Stochastic Differential Equations and Applications; Vol. 1*. Academic Press, New York San Francisco London, 1st edition, 1975.

[11] A. Göing Jaeschke and M. Yor. A survey and some generalizations of Bessel processes. *Bernoulli*, 9(2):313–349, 2003.

[12] J.C. Hull. *Options, Futures, and Other Derivatives*. Pearson Prentice Hall, New Jersey, 6th edition, 2006.

[13] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*. North – Holland Publishing, Amsterdam New York Oxford, 1st edition, 1981.

[14] S.D. Jacka. Optimal stopping and the American put. *Mathematical Finance*, 1(2):1–14, 1991.

[15] I. Karatzas. On the pricing of American options. *Applied Mathematics and Optimization*, 17:37–60, 1988.

[16] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus.* Springer, Berlin Heidelberg New York, 2nd edition, 1991.

[17] N.V. Krylov. On Itô's stochastic integral equations. *Theory of Probability and Its Applications*, 14(2):330–336, 1969.

[18] F.A. Longstaff and E.S. Schwartz. Valuing american options by simulation: A simple least – squares approach. *The Review of Financial Studies*, 14(1):113–147, 2001.

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

[20] B. Øksendal. *Stochastic Differential Equations.* Springer, Berlin Heidelberg New York, 6th edition, 2007.

[21] J. Pitman and M. Yor. A decomposition of Bessel bridges. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 59:425–457, 1982.

[22] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion.* Springer, Berlin Heidelberg New York, 3nd edition, 1999.

[23] L.C.G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales.* Cambridge University Press, Cambridge, 2nd edition, 2008.

[24] F. Salmon. The formula that killed Wall Street. *Wired Magazine*, 17.03, 2009.

[25] S.E. Shreve. *Stochastic Calculus for Finance II: Continuous – Time Models.* Springer, Berlin Heidelberg New York, 2nd edition, 2008.

[26] P. Wilmott, J. Dewynne, and S. Howison. *Option Pricing.* Oxford Financial Press, Oxford, 1st edition, 1994.