Utrecht University

Mathematical Sciences

Master thesis

American Options with Guarantee

Option Valuation by the Optimal Stopping Problem for Lévy Processes

Author:
Alexander Heinrich Wilhelm Borchert

Student number:
4109953

Supervisor:
Dr. Karma Dajani

Second reader:
Dr. Paul Zegeling

March 20, 2018

Contents

Introduction

This thesis is about the pricing problem of perpetual American options with guarantee. This is an option that can be exercised at any time and for which a specific payoff is always guaranteed. The most common approach is to model the underlying stock price process by geometric Brownian motion. We are modelling the underlying stock price process by the exponential function of a spectrally negative Lévy process. This is a more general approach. Spectrally negative Lévy processes have the advantage that they allow negative jumps. This is especially interesting for risk management questions since this approach allows to model strong price decreases in a short time; for example the crash of a stock. One requirement will be that the payoff function is non-decreasing. So the theory can be applied for example on call options but not on put options. The value of an American option with guarantee is given by

$$v(x) = \sup_{\tau \in \mathcal{T}} \mathbb{E}_x\Big(e^{-r\tau}\big(f(X_\tau) \vee k(x)\big)\Big), \tag{1}$$

where $\mathcal{T}$ is the set of stopping times, r the interest rate, f the payoff function, $k(x)$ the guarantee, $(X_t)_t$ the process of the stock price and x the present value of the stock. Throughout the whole thesis we assume that the interest rate is a positive constant. This thesis relies heavily on the paper [1] by Albrecht Irle and Sören Christensen. The important result of their paper which we are interested in is [1, Theorem 5.5]. In this thesis this is Theorem 3.3.2. The theorem gives an analytic formula for the price of the option which relies on the so called scale functions of the spectrally negative Lévy process. The formula is given by

$$v(x) = \sup_{a<b}\left(k(x)Z(x-a) + \frac{g_x(b) - g_x(a)Z(b-a)}{W(b-a)}W(x-a)\right), \tag{2}$$

where W and Z are the scale functions of the stock price process $(X_t)_t$ and $g_x(y) := g(y) \vee k(x)$. The function Z can be derived by ordinary integration over W. So the crucial element for using equation (2) is to find the scale function W.

Structure and Content of the Thesis

In our first chapter we are interested in giving a rigorous introduction of $\mathbb{E}_x$ which means the expectation in case that the process $(X_t)_t$ starts at x. A proper definition of $\mathbb{E}_x$ can be found in the context of Markov processes by defining a family of probability measures. We are introducing the notation $\mathbb{E}_x$ for Markov processes and prove that Lévy processes are Markov processes. Furthermore,

we are giving the definition of spectrally negative Lévy processes and prove some properties we will need. As we have mentioned, Theorem 3.3.2 depends on the scale functions of spectrally negative Lévy processes. For this reason we introduce them in Chapter 2 and derive an analytic solution of the scale function of Brownian motion. Moving on, in Chapter 3 we review in detail the paper [1]. The main result is Theorem 3.3.2 as mentioned above. In Chapter 4 we test how good equation (2) can be used for computations of a call option. The crucial point in formula (2) is the scale function W. So the problem reduces to finding W. Chapter 4 has two aims. First we want see if equation (2) gives good computational results. In order to do that, we test it on the case that the price process is a scaled Brownian motion with drift. We will be using the analytic solution of the scale functions W and Z which we found in Section 2.3. Second, even though we are modelling the stock in this chapter with the well studied Brownian motion, the call option with guarantee is not very well studied. So the results are interesting in themselves and we take a closer look on them. The last thing we do is a little investigation of the case of more general process. We give a first idea how the scale function can be approximated in this case. The approximation are then compared with the analytic solution which we found.

Contribution of the Thesis

This thesis can be seen as literature review of [1] with some additional contributions. We want to mention four of them. The results found in Chapter 1 are well known to experts in the field but it is very hard to find actual proofs in the literature. We choose to include all the necessary proofs, which is our first contribution. The second contribution is Lemma 2.3.2 in which we derive an analytic solution for the scale functions of scaled Brownian motion with drift. Furthermore, it is not clear in [1] for which payoff functions a solution for (1) exists. We prove in Lemma 3.2.2 and Lemma 3.2.1 that it is enough for the payoff functions to be continuous. This is our third contribution. Our fourth contribution is to show in Chapter 4 that the theory can be used successfully for computations and to give a better understanding of perpetual American call options with guarantee which are not well researched.

Chapter 1

Lévy Processes

The main topic of this thesis is optimal stopping for Lévy processes. A typical notation in this context is $\mathbb{E}_x X_t$ and it means the expectation of a process $(X_s)_s$ at time t if the process starts at point x. In the context of Lévy processes most authors are working with this terminology quite intuitively. Our approach is to give a technical accurate definition of this expression and show that it is justified to work with it in the way one would assume. An accurate definition can be found in the context of Markov processes where a family of probability measures $\mathbb{P}_x$ with the property $\mathbb{P}_x(X_0 = x) = 1$ is defined. We will first introduce Markov processes and give a rigorous definition of the notation $\mathbb{E}_x X_t$. In the second section we introduce Lévy processes and show that there exists a family of probability measures $\mathbb{P}_x$ with the right properties. In the third section we show that Lévy processes are Markov processes. In the fourth section we show useful properties for Lévy processes which we will need in the following chapters. In the last section we introduce spectrally negative Lévy processes. All proofs in this chapter are original by the author of this thesis.

1.1 Markov Processes

In this section we introduce Markov processes in a very general way. This will help us in the following to apply results of Markov processes to Lévy processes. The state space of our processes we will be working on has the following form: Let $I_1, ..., I_n \in \mathcal{B}(\mathbb{R})$ with $0 \in I_1, ..., 0 \in I_n$, $E := I_1 \times ... \times I_n$ where $\mathcal{B}(\mathbb{R})$ is the Borel σ-algebra. Furthermore, we consider as σ-algebra on E the Borel sets $\mathcal{E} := \mathcal{B}(E)$ and $\|\cdot\|$ be the Euclidean norm on E. So $\mathcal{E}$ is the restriction of the Borel σ-algebra of $\mathbb{R}^n$ on E. In the following all measure spaces $(E, \mathcal{E})$ are assumed to be of this form and whenever a norm is mentioned it is the Euclidean norm.
To define Markov processes we first have to introduce transition kernels and transition functions:

Definition 1.1.1 (Transition kernel)**.** *Let $(E, \mathcal{E})$ be a measurable space. A transition kernel on $(E, \mathcal{E})$ is a function $p : E \times \mathcal{E} \to [0, 1]$ such that*

- *for every $x \in E$, the function $B \mapsto p(x, B)$ is a probability measure on $(E, \mathcal{E})$ and*

- *for every $B \in \mathcal{B}(\mathbb{R})$, the function $x \mapsto p(x, B)$ is $\mathcal{E}/\mathcal{B}([0,1])$-measurable.*

Definition 1.1.2 (Transition function)**.** *Let $(E, \mathcal{E})$ be a measurable space and for $x \in E$ let δ_x be the Dirac measure. Furthermore, let $(p_{s,t}) = \{p_{s,t} : 0 \leq s \leq t\}$ be a collection of transition kernels on $(E, \mathcal{E})$.*

- *We say that $(p_{s,t})$ is a transition function if and only if*

$$p_{t,t}(x, \cdot) = \delta_x \quad (x \in E, t \geq 0)$$

and the Chapman-Kolmogorov equation

$$p_{s,t}(x, B) = \int p_{u,t}(y, B) p_{s,u}(x, dy) \quad (s \leq u \leq t)$$

is fulfilled.

- *Furthermore, if*

$$p_{s,t} = p_{0,t-s} \quad \text{for all } 0 \leq s \leq t$$

is satisfied, we say that $(p_{s,t})$ is a homogeneous transition function and we can identify $(p_{s,t})$ with $(p_t)_t = \{p_t : t \geq 0\}$ where $p_t := p_{t,0}$. In this case the Chapman-Kolmogorov equation becomes

$$p_{t+s}(x, B) = \int p_s(y, B) p_t(x, dy). \tag{1.1}$$

We introduce now Markov processes.

Definition 1.1.3 (Markov Process)**.** *Let $(X_t)_t$ be an $(E, \mathcal{E})$-valued stochastic process defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, $(p_t)_t$ be a homogeneous transition function and $\{\mathbb{P}_x : x \in \mathbb{R}\}$ be a collection of probability distributions on $(\Omega, \sigma(\{X_s : s \geq 0\}))$. Furthermore let be $\mathcal{F}_t := \sigma(X_u : 0 \leq u \leq t)$. For any random variable $Y : \Omega \to \mathbb{R}$ and $x \in E$ we write $\mathbb{E}_x Y$ for the expectation of Y with respect to the probability space $(\Omega, \mathcal{F}, \mathbb{P}_x)$ and we write $\mathbb{E}$ instead of $\mathbb{E}_0$. Furthermore, we define for a Borel measurable and bounded function $f : E \to \mathbb{R}$ and $s, t \geq 0$ the functions*

$$\mathbb{E}_{X_s} f(X_t) : \Omega \to \mathbb{R}, \quad \omega \mapsto \mathbb{E}_{X_s(\omega)} f(X_t)$$

and

$$p_t f : E \to \mathbb{R}, \quad x \mapsto \int f(y) dp_t(x, dy).$$

Furthermore we use the notation

$$\mathbb{E}_x(f(X_{t+s})|X_s) := \mathbb{E}_x(f(X_{t+s})|\sigma(X_s)). \tag{1.2}$$

We say that $(X_t)_t$ is an $(E, \mathcal{E})$-valued homogeneous Markov process on $(\Omega, \mathcal{F}, \mathbb{P})$ with transition function $(p_t)_t$ and distribution $(\mathbb{P}_x)_{x \in \mathbb{R}}$ if and only if

- $\mathbb{P}_x(X_0 = x) = 1$ *for every $x \in E$ and*

- *for every bounded and Borel-measurable functions $f : E \to \mathbb{R}$ and every pair of real numbers $s, t \geq 0$ there exists a Borel-measurable function $g : E \to \mathbb{R}$ such that*

$$\mathbb{E}_x(f(X_{t+s})|\mathcal{F}_s) = g(X_s) = \mathbb{E}_x(f(X_{t+s})|X_s) = p_t f(X_s) = \mathbb{E}_{X_s} f(X_t) \quad \mathbb{P}_x - a.s. \tag{1.3}$$

  for every $x \in E$.

Remark. *The expert reader will notice that the four equations in 1.3 are just equivalent forms of the Markov property normally found in the literature. See for example [2], [3], [6] and [9]. By defining Markov processes in the above way we are able to use results for Markov processes from these different references in the following chapters. The definition above is now a rigorous introduction of the notation $\mathbb{E}_x$. Especially in the context of Lévy processes there is normally only an intuitive explanation of this notation. For example $\mathbb{E}X_t$ is then introduced as the expectation of X_t if $(X_t)_t$ starts a zero. We should of course always keep this intuition in mind when we work with the equations (1.3). For example the intuitive understanding of the equation*

$$p_t f(X_s) = \mathbb{E}_{X_s} f(X_t) \quad \mathbb{P}_x - a.s.$$

would be: The random variable $p_t f(X_s)$ is equal to the random variable $\mathbb{E}_{X_s} f(X_t)$ in the case that the process $(X_t)_t$ starts at x.

The definition of transition kernels appears to be very technical, but now that we have introduced Markov processes, we can give a good intuitive understanding of it. The idea is that $p_t(x, B)$ is expressing the probability that the process hits the set B at time t if it started at time zero at the point x. The intuitive idea of a Markov process is that every change of the process at any time is independent of all events that occurred before in time. Combining these two ideas means that $p_t(x, B)$ is expressing the probability that the process hits the set B at some time $t + s$ if it had been at point x at time s. This is shown in the following lemma.

Lemma 1.1.1. *For an $(E, \mathcal{E})$-valued Markov process $(X_t)_t$ with transition function $(p_u)_u$, $s, t \geq 0$ and $B \in \mathcal{E}$ we have $p_t(x, B) = \mathbb{P}(X_{t+s} \in B | X_s = x)$ if $\mathbb{P}(X_s = x) > 0$.*

Proof. We have for every $x \in E$

$$\big(p_t 1_B\big)(x) = \int 1_B(y) p_t(x, dy) = p_t(x, B). \tag{1.4}$$

It follows for $x \in E$ that

$$\begin{aligned}
\mathbb{P}(X_{t+s} \in B | X_s = x) &= \frac{\mathbb{P}(X_{t+s} \in B, X_s = x)}{\mathbb{P}(X_s = x)} \\
&= \frac{\mathbb{E}\Big(1_B(X_{t+s}) 1_{\{x\}}(X_s)\Big)}{\mathbb{P}(X_s = x)} \\
&= \frac{\mathbb{E}\Big(\mathbb{E}\big(1_B(X_{t+s})|X_s\big) 1_{\{x\}}(X_s)\Big)}{\mathbb{P}(X_s = x)} \\
&\overset{\text{Markov}}{=} \frac{\mathbb{E}\Big(\big(p_t 1_B\big)(X_s) 1_{\{x\}}(X_s)\Big)}{\mathbb{P}(X_s = x)} \\
&\overset{(1.4)}{=} \frac{\mathbb{E}\Big(p_t(X_s, B) 1_{\{X_s = x\}}\Big)}{\mathbb{P}(X_s = x)} \\
&= p_t(x, B).
\end{aligned}$$

□

1.2 Lévy Processes

In this section we introduce Lévy processes and show that there is a very natural choice for probability measures $\mathbb{P}_x$ such that $\mathbb{P}_x(X_0 = x) = 1$.

Definition 1.2.1. *Let $I_1, ..., I_n \in \mathcal{B}(\mathbb{R})$ with $0 \in I_1, ..., 0 \in I_n$, $E = I_1 \times ... \times I_n$ and $\mathcal{E} = \mathcal{B}(E)$. An $(E, \mathcal{E})$-valued Lévy process is an $(E, \mathcal{E})$-valued stochastic process $(X_t)_t$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that*

- $\mathbb{P}(X_0 = 0) = 1$
- *For all $0 \leq s \leq t$, $X_t - X_s$ is independent of $\{X_u : u \leq s\}$*
- *For all $0 \leq s \leq t$, $X_t - X_s \overset{d}{=} X_{t-s}$*
- *Almost every path ist right-continuous and its left limits exist,*

where the d above the equal sign means that both random variables have the same probability distribution.
We also say in the following that a path of a stochastic process is cadlag if it is right-continuous and has left limits.

For a given Lévy process $(X_t)_t$ we would like to have a family of probability measure $\{\mathbb{P}_x : x \in E\}$ such that $\mathbb{P}_x(X_0 = x) = 1$ and for every Borel measurable function f

$$\mathbb{E}_x f(X_{t_1}, ..., X_{t_n}) = \mathbb{E} f(X_{t_1} + x, ..., X_{t_n} + x).$$

Unfortunatettly, such a family does not always exist. But there is an easy solution to this problem. It is possible to extend the Lévy process to a larger domain for which such a family of probability measures exists. We prove this in the following lemma.

Lemma 1.2.1. *Let $(X_t)_t$ be an $(E,\mathcal{E})$-valued Lévy process defined on a probability space $(\Omega,\mathcal{F},\mathbb{P})$. Then there exists a $(E,\mathcal{E})$-valued Lévy process $(X'_t)_t$ defined on a probability space $(\Omega',\mathcal{F}',\mathbb{P}')$ and a unique collection of probability measures $\{\mathbb{P}'_x : x \in E\}$ on $(\Omega',\mathcal{F}')$ such that for every $t_1,...,t_n \geq 0$, $a \in E$ and every Borel measurable function $f : E \to \mathbb{R}$ which is bounded or nonnegative*

$$\mathbb{E}f(X_{t_1},...,X_{t_n}) = \mathbb{E}'f(X'_{t_1},...,X'_{t_n}) \tag{1.5}$$

and

$$\mathbb{E}'_x f(X'_{t_1},...,X'_{t_n}) = \mathbb{E}'f(X'_{t_1}+x,...,X'_{t_n}+x), \tag{1.6}$$

where $\mathbb{E}$ means the expectation with respect to $(\Omega,\mathcal{F},\mathbb{P})$ and $\mathbb{E}'_x$ means the expectation with respect to $(\Omega',\mathcal{F}',\mathbb{P}'_x)$.

Proof. The uniqueness of the probability measures $\mathbb{P}'_x$ follows simply from the fact that the sets $\{X'_{t_1} \in A_1,...,X'_{t_n} \in A_n\}$ form a π-system of $\sigma(X'_u : u \geq 0)$.
Now let us consider the induced canonical process

$$\Omega' := E^{\mathbb{R}_0^+}$$

and for $n \in \mathbb{N}, t_1,...,t_n \geq 0, A_1,...,A_n \in \mathcal{E}$ we define the set

$$Pass(t_1,...,t_n,A_1,...,A_n) := \{\omega' \in \Omega' : \omega'(t_1) \in A_1,...,\omega'(t_n) \in A_n\}$$

of all functions in Ω' which are passing A_i at time t_i for $i = 1,...,n$.
We define now the σ-algebra $\mathcal{F}'$ on Ω' by

$$\mathcal{F}' := \sigma\big(\{Pass(t_1,...,t_n,A_1,...,A_n) : n \in \mathbb{N}, t_1,...,t_n \geq 0, A_1,...,A_n \in \mathcal{E}\}\big).$$

We see that for $n,m \in \mathbb{N}$, $s_1,...s_n,t_1,...,t_m \geq 0$ and $A_1,...,A_n,B_1,...B_m \in \mathcal{E}$ we have

$$\begin{aligned} &Pass(s_1,...s_n,A_1,...,A_n) \cap Pass(t_1,...,t_m,B_1,...,B_m) \\ &\quad = Pass(s_1,...s_n,t_1,...,t_m,A_1,...,A_n,B_1,...B_m) \end{aligned} \tag{1.7}$$

and that these sets form a generating π-system of $\mathcal{F}'$.
We consider now the function

$$X : \Omega \to \Omega', \quad \omega \mapsto (X_t(\omega) : t \geq 0),$$

where we write $(X_t(\omega) : t \geq 0)$ for the function which maps t to $X_t(\omega)$.
We have

$$\begin{aligned} X^{-1}(Pass(t_1,..,t_n,A_1,...,A_n)) &= \{\omega \in \Omega : X(\omega) \in Pass(t_1,..,t_n,A_1,...,A_n)\} \\ &= \{\omega \in \Omega : (t \mapsto X_t(\omega)) \in Pass(t_1,...,t_n,A_1,...,A_n)\} \\ &= \{\omega \in \Omega : X_{t_1}(\omega) \in A_1,...,X_{t_n}(\omega) \in A_n\} \\ &= \bigcap_{i=1}^{n} \{X_{t_i} \in A_i\} \in \mathcal{F}. \end{aligned} \tag{1.8}$$

So X is $\mathcal{F}/\mathcal{F}'$-measurable. Furthermore, we define for every $t \geq 0$

$$X'_t : \Omega' \to E, \quad \omega' \mapsto \omega'(t).$$

Since, for every $A \in \mathcal{E}$

$$\begin{aligned}(X'_t)^{-1}(A) &= \{\omega' \in \Omega' : X'_t(\omega') \in A\} \\ &= \{\omega' \in \Omega' : \omega'(t) \in A\} \\ &= Pass(t, A) \in \mathcal{F}', \end{aligned} \tag{1.9}$$

it follows that X'_t is $\mathcal{F}'/\mathcal{E}$-measurable. Now let us define for every $x \in E$ the shift operator

$$s_x : \Omega' \to \Omega', \quad \omega' \mapsto (t \mapsto \omega'(t) + x).$$

We have for $t_1, ..., t_n \geq 0$ and $A_1, ..., A_n \in \mathcal{E}$

$$\begin{aligned} s_x^{-1}\big(Pass(t_1, ..., t_n, A_1, ..., A_n)\big) &= \{\omega' \in \Omega' : s_x(\omega') \in Pass(t_1, ..., t_n, A_1, ..., A_n)\} \\ &= \{\omega' \in \Omega' : \underbrace{\big(s_x(\omega')\big)(t_1)}_{=\omega'(t_1)+x} \in A_1, ..., \underbrace{\big(s_x(\omega')\big)(t_n)}_{=\omega'(t_n)+x} \in A_n\} \\ &= \{\omega' \in \Omega' : \omega'(t_1) \in A_1 - x, ..., \omega'(t_n) \in A_n - x\} \\ &= Pass(t_1, ..., t_n, A_1 - x, ..., A_n - x). \end{aligned} \tag{1.10}$$

So s_x is $\mathcal{F}'/\mathcal{F}'$-measurable and $s_x \circ X$ is $\mathcal{F}/\mathcal{F}'$-measurable.
Now let $\mathbb{P}'_x$ be the law of $s_x \circ X$, *i.e.*, $\mathbb{P}'_x$ is a probability measure on $(\Omega', \mathcal{F}')$.
We have for $t_1, ..., t_n \geq 0$ and $A_1, ..., A_n \in \mathcal{E}$

$$\begin{aligned} \{X'_{t_1} \in A_1, ..., X'_{t_n} \in A_n\} &= \bigcap_{i=1}^{n} \{X'_{t_i} \in A_i\} \\ &\overset{(1.9)}{=} \bigcap_{i=1}^{n} Pass(t_i, A_i) \\ &\overset{(1.7)}{=} Pass(t_1, ..., t_n, A_1, ..., A_n). \end{aligned} \tag{1.11}$$

We have then

$$\begin{aligned} \mathbb{P}'_0(X'_{t_1} \in A_1, ..., X'_{t_n} \in A_n) &\overset{(1.11)}{=} \mathbb{P}'_0(Pass(t_1, ..., t_n, A_1, ..., A_n)) \\ &= \mathbb{P}(X^{-1}(Pass(t_1, ..., t_n, A_1, ..., A_n))) \\ &\overset{(1.8)}{=} \mathbb{P}(X_{t_1} \in A_1, ..., X_{t_n} \in A_n) \end{aligned} \tag{1.12}$$

and equation (1.5) follows by the standard machinery which we explain in the remark after the proof.
Furthermore, we have

$$\begin{aligned} \mathbb{P}'_x\big(X'_{t_1} \in A_1, ..., X'_{t_n} \in A_n\big) &\overset{(1.11)}{=} \mathbb{P}'_x\big(Pass(t_1, ..., t_n, A_1, ..., A_n)\big) \\ &= \mathbb{P}\big(\underbrace{(s_x \circ X)^{-1}}_{X^{-1}(s_x^{-1})}(Pass(t_1, ..., t_n, A_1, ..., A_n))\big) \\ &\overset{(1.10)}{=} \mathbb{P}\big(X^{-1}(Pass(t_1, ..., t_n, A_1 - x, ..., A_n - x))\big) \\ &\overset{(1.8)}{=} \mathbb{P}\big(X_{t_1} \in A_1 - x, ..., X_{t_n} \in A_n - x\big) \\ &\overset{(1.12)}{=} \mathbb{P}'_0\big(X'_{t_1} \in A_1 - x, ..., X'_{t_n} \in A_n - x\big) \\ &= \mathbb{P}'_0\big(X'_{t_1} + x \in A_1, ..., X'_{t_n} + x \in A_n\big). \end{aligned}$$

and equation (1.6) follows by the standard machinery.
That (X'_t) is an $(E,\mathcal{E})$-valued Lévy process defined on the probability space $(\Omega',\mathcal{F}',\mathbb{P}')$ follows by equation (1.5) and the fact that $(X_t)_t$ is a Lévy process. □

So from now on we assume that for every Lévy process there exists a collection of probability measures $\{\mathbb{P}_x : x \in E\}$ such that

$$\mathbb{E}_x f(X_{t_1},...,X_{t_n}) = \mathbb{E} f(X_{t_1}+x,...,X_{t_n}+x).$$

We have used in the last proof the standard machinery which we explain in the following remark.

Remark (Standard machinery)**.** *With standard machinery we mean a standard procedure in probability theory which can be used for many proofs. The procedure works along the following steps. First a statement is proved for indicator functions, then one extends the result by linearity arguments to nonnegative simple functions. Then the Monotone Convergence Theorem is used to extend the result to nonnegative measurable functions. In the last step the claim is proved for measurable functions by splitting them into their positive and negative part. Both parts are then nonnegative measurable functions and the result from the step before can be applied.*

1.3 Lévy Processes are Markov Processes

We have seen in the last section that we can assume that for every Lévy process $(X_t)_t$ there exists a collection of probability measures $\{\mathbb{P}_x : x \in E\}$ such that

$$\mathbb{E}_x f(X_{t_1},...,X_{t_n}) = \mathbb{E} f(X_{t_1}+x,...,X_{t_n}+x)$$

for every Borel measurable function f which is nonnegative or bounded. We will introduce now also a transition function and show that a Lévy process is a Markov process with respect to this transition function and the probability measures $\mathbb{P}_x$.

Definition 1.3.1. *Let $(X_t)_t$ be an $(E,\mathcal{E})$-valued Lévy process defined on a probability space $(\Omega,\mathcal{F},\mathbb{P})$ and $0 \le s \le t$.*
We define the function

$$p_{s,t} : E \times \mathcal{E} \to \mathbb{R}, \quad (x,B) \mapsto \mathbb{P}(X_t - X_s + x \in B).$$

For a measurable function $f : E \to \mathbb{R}$ with $\mathbb{E}|f(X_t - X_s + x)| < \infty$ we define the function

$$g^f_{s,t} : E \to \mathbb{R}, \quad x \mapsto \mathbb{E} f(X_t - X_s + x).$$

By [7, Theorem 8.12 (iv)] $g^f_{s,t}$ is Borel measurable and for $x \in E$ and $B \in \mathcal{E}$ we have

$$\begin{aligned} p_{s,t}(x,B) &= \mathbb{P}(X_t - X_s + x \in B) \\ &= \mathbb{E} 1_B(X_t - X_s + x) \\ &= g^{1_B}_{s,t}(x) \end{aligned} \tag{1.13}$$

Lemma 1.3.1. *For an $(E,\mathcal{E})$-valued Lévy process $(X_t)_t$ defined on a probability space $(\Omega,\mathcal{F},\mathbb{P})$ and $s \leq t$ the function $p_{s,t}$ is a transition kernel*

Proof. By definition $p_{s,t}(x,\cdot)$ is the probability distribution of $X_t - X_s + x$ and this is a probability distribution on $(\mathbb{R},\mathcal{B}(\mathbb{R}))$. The second property of a transition kernel holds by equation (1.13) and the fact that $g_{s,t}^{1_B}$ is Borel-measurable. □

Lemma 1.3.2. *For an $(E,\mathcal{E})$-valued Lévy process $(X_t)_t$ defined on a probability space $(\Omega,\mathcal{F},\mathbb{P})$ the collection of transition kernels $(p_{s,t})$ is a homogeneous transition function.*

Proof. Because $(X_t)_t$ is a Lévy process we know that $X_t - X_s \overset{d}{=} X_{t-s}$ for $0 \leq s \leq t$. We have for $x \in E$ and $B \in \mathcal{E}$

$$p_{s,t}(x,B) = \mathbb{P}(X_t - X_s \in B - x) = \mathbb{P}(X_{t-s} - X_0 \in B - x) = p_{t-s,0}(x,B).$$

So we can define

$$p_t := p_{t,0}.$$

Now we show the Chapman-Kolmogorov equation (1.1) is fulfilled. In order to do so, let us fix some $0 \leq s \leq t$, $x \in E$ and define

$$\nu' : \mathcal{E} \to [0,1], \quad A \mapsto p_{t+s}(x,A) = \nu'(A)$$

and

$$\nu : \mathcal{E} \to [0,1], \quad A \mapsto \int p_s(y,A)p_t(x,dy) = \nu(A).$$

We see that the Chapman-Kolmogorov equation is equivalent to $\nu' = \nu$. So we finish the proof now by showing that ν and ν' are probability measures on $(E,\mathcal{E})$ which are identical on a generating π-system which implies that they are identical on the whole of $\mathcal{E}$.
That ν' is a probability measure follows from Lemma 1.3.1. That $\nu(\emptyset) = 0$ and $\nu(\Omega) = 1$ is easy to see. To see the σ-additivity of ν we take some $A = \bigcup_{n\in\mathbb{N}} A_n$ with $(A_n)_n \subseteq \mathcal{E}$ being a sequence of disjoint sets. We have

$$\nu(A) = \int p_s(y,A)p_t(x,dy) = \int \left(\sum_{n\in\mathbb{N}} p_s(y,A_n)\right) p_t(x,dy)$$

$$\overset{\text{Mon. Conv.}}{=} \sum_{n\in\mathbb{N}} \left(\int p_s(y,A_n)p_t(x,dy)\right) = \sum_{n\in\mathbb{N}} \nu(A_n).$$

So ν is a probability measure as well.
By definition of Lévy processes there exists an $n \in \mathbb{N}$ and $I_1,...,I_n \in \mathcal{B}(\mathbb{R})$ with $E = I_1 \times \cdots \times I_n$. We assume without loss of generality that $I_1 = \cdots = I_n = \mathbb{R}$. We know that the collection of sets

$$\{(-\infty,b_1] \times \cdots \times (-\infty,b_n] : b_1,...,b_n \in \mathbb{R}\}$$

is a generating π-system of the Borel σ-algebra $\mathcal{E} = \mathcal{B}(\mathbb{R}^n)$. Let us fix now some $b = (b_1,...,b_n) \in E$ and set $B := (-\infty,b_1] \times ... \times (-\infty,b_n]$. We have to show that $\nu(B) = \nu'(B)$.

Let F_s be the distribution function of X_s, μ_t the distribution of $X_t + x$ and F_Z the distribution function of $(X_t + x) + X_s$. We have then the well known distribution formula for sums

$$F_Z(b) = \int F_s(b-y)\mu_t(dy). \tag{1.14}$$

Furthermore, we have

$$\mu_t(A) = \mathbb{P}(X_t + x \in A) = p_t(x, A),$$

i.e. we have

$$\int f(y)\mu_t(dy) = \int f(y)p_t(x, dy) \tag{1.15}$$

for every measurable function $f : E \to \mathbb{R}$. Furthermore we have

$$F_s(b-y) = \mathbb{P}(X_s \leq b - y) = \mathbb{P}(X_s \in B - y) = p_s(y, B). \tag{1.16}$$

Since $(X_t)_t$ is a Lévy process we have

$$\begin{aligned} X_{t+s} + x &= X_{t+s} - X_s + X_s + x \\ &\stackrel{d}{=} X_t + X_s + x. \end{aligned} \tag{1.17}$$

We can calculate now

$$\nu'(B) = p_{t+s}(x, B) = \mathbb{P}(X_{t+s} + x \in B) \stackrel{(1.17)}{=} \mathbb{P}((X_t + x) + X_s \in B) = F_Z(b)$$

$$\stackrel{(1.14)}{=} \int \underbrace{F_s(b-y)}_{\stackrel{(1.16)}{=} p_s(y,B)} \mu_t(dy) \stackrel{(1.15)}{=} \int p_s(y, B)p_t(x, dy) = \nu(B).$$

and this finishes the proof.

□

So equation (1.13) reads then

$$p_t(x, B) = g_t^{1_B}(x)$$

and it can be shown in a straightforward manner by using the standard machinery that for every Borel measurable function $f : E \to R$ which is bounded or non-negative

$$(p_t f)(x) = g_{s,s+t}^f(x) \tag{1.18}$$

for every $x \in E$ and arbitrary $0 \leq s \leq t$.

Lemma 1.3.3. *Let $(X_t)_t$ be an $(E, \mathcal{E})$-valued Lévy process defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then $(X_t)_t$ is a homogeneous Markov process with transition function $(p_t)_t$ and distribution $(\mathbb{P}_x)_{x\in\mathbb{R}}$*

Proof. We have to prove that

$$\mathbb{E}_x(f(X_{t+s})|\mathcal{F}_s) = g(X_s) = \mathbb{E}_x(f(X_{t+s})|X_s) = p_t f(X_s) = \mathbb{E}_{X_s} f(X_t). \quad \mathbb{P}_x - a.s.$$

First let us take $A, B \in \mathcal{E}$, $0 \leq u \leq s \leq t$, $h_1(x,y) := 1_A(x-y)$ and $h_2(x,y,z) := 1_A(x-y)1_B(z)$. We have

$$\begin{aligned}\mathbb{P}(X_t - X_s \in A) &= \mathbb{P}((X_t + x) - (X_s + x) \in A) \\ &= \mathbb{E}h_1(X_t + x, X_s + x) \\ &= \mathbb{E}_x h_1(X_t, X_s) \\ &= \mathbb{P}_x(X_t - X_s \in A) \end{aligned} \tag{1.19}$$

and it follows

$$\begin{aligned}\mathbb{P}_x(X_t - X_s \in A, X_u \in B) &= \mathbb{E}_x h_2(X_t, X_s, X_u) \\ &= \mathbb{E}h_2(X_t + x, X_s + x, X_u + x) \\ &= \mathbb{P}(X_t - X_s \in A, X_u + x \in B) \\ &= \mathbb{P}(X_t - X_s \in A)\mathbb{P}(X_u + x \in B) \\ &\overset{(1.19)}{=} \mathbb{P}_x(X_t - X_s \in A)\mathbb{P}_x(X_u \in B).\end{aligned}$$

This means that $X_t - X_s$ is independent of $\mathcal{F}_s = \sigma(X_u : 0 \leq u \leq s)$ with respect to $\mathbb{P}_x$.
So $X_{s+t} - X_s$ is independent of $\mathcal{F}_s$ with respect to $\mathbb{P}_x$ and X_s is of course $\mathcal{F}_s$-measurable. Thus, we have by [7, Theorem 8.12 (iv)] for a Borel-measurable function $f : E \to \mathbb{R}$ for $\mathbb{P}_x$-almost every $\omega \in \Omega$

$$\mathbb{E}_x(f(X_{t+s} - X_s + X_s)|\mathcal{F})(\omega) = \mathbb{E}_x(f(X_{t+s} - X_s + X_s(\omega))). \tag{1.20}$$

It follows

$$\begin{aligned}\mathbb{E}_x(f(X_{t+s}|\mathcal{F}_s))(\omega) &= \mathbb{E}_x(f(X_{t+s} - X_s + X_s)|\mathcal{F})(\omega) \\ &\overset{(1.20)}{=} \mathbb{E}_x(f(X_{t+s} - X_s + X_s(\omega))) \\ &= \mathbb{E}(f(X_{t+s} - X_s + X_s(\omega))) \\ &= g^f_{s,s+t}(X_s(\omega)) \\ &= (g^f_{s,s+t}(X_s))(\omega).\end{aligned}$$

We get

$$g^f_{s,s+t}(X_s) = \mathbb{E}_x(f(X_{t+s})|X_s) \quad \mathbb{P}_x\text{-a.s.}$$

completely analogously.
Furthermore, the equation

$$p_t f(X_s) = g^f_{s,s+t}(X_s)$$

follows simply by equation (1.18).
The last equation can be seen by

$$\begin{aligned}\big(\mathbb{E}_{X_s} f(X_t)\big)(\omega) &= \mathbb{E}_{X_s(\omega)} f(X_t) \\ &= \mathbb{E}f(X_t + X_s(\omega)) \\ &= g^f_{0,t}(X_s(\omega)) \\ &\overset{(1.18)}{=} (p_t f)(X_s(\omega)) \\ &= \big((p_t f)(X_s)\big)(\omega)\end{aligned}$$

□

1.4 Properties of Lévy Processes

Lemma 1.4.1. *If $(X_t)_t$ is an $(E, \mathcal{E})$-valued Lévy process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then $(Y_t)_t$ given by $Y_t(\omega) = (t, X_t(\omega))$ is an $(\mathbb{R}_+ \times E, \mathcal{B}(\mathbb{R}_+) \times \mathcal{E})$-valued Lévy process on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.*

Proof. • The first property we see by

$$\mathbb{P}(Y_0 = 0) = \mathbb{P}((0, X_0) = (0, 0)) = \mathbb{P}(X_0 = 0) = 1.$$

We consider now for every $t \geq 0$ the inclusion

$$e_t : E \to \mathbb{R}_0^+ \times E, \quad x \mapsto (t, x).$$

This is a measurable function and we have for $0 \leq s \leq t$ and $A \in \mathbb{R}_0^+ \times \mathcal{E}$

$$\begin{aligned} \{Y_t - Y_s \in A\} &= \{(t - s, X_t - X_s) \in A\} \\ &= \{e_{t-s}(X_t - X_s) \in A\} \\ &= \{X_t - X_s \in e_{t-s}^{-1}(A)\} \end{aligned}$$

and by choosing $s = 0$ we also have

$$\{Y_t \in A\} = \{X_t \in e_t^{-1}(A)\}.$$

With these two equations the other Lévy properties for $(Y_t)_t$ follow simply from the Lévy properties of $(X_t)_t$:

- For $0 \leq u \leq s \leq t$ and $A, B \in \mathbb{R}_0^+ \times \mathcal{E}$ we have

$$\begin{aligned} \mathbb{P}(Y_t - Y_s \in A, Y_u \in B) &= \mathbb{P}(X_t - X_s \in e_{t-s}^{-1}(A), X_u \in e_u^{-1}(B)) \\ &= \mathbb{P}(X_t - X_s \in e_{t-s}^{-1}(A))\mathbb{P}(X_u \in e_u^{-1}(B)) \\ &= \mathbb{P}(Y_t - Y_s \in A)\mathbb{P}(Y_u \in B) \end{aligned}$$

and this proves that $Y_t - Y_s$ and Y_u are independent.

- Furthermore we have

$$\begin{aligned} \mathbb{P}(Y_t - Y_s \in A) &= \mathbb{P}(X_t - X_s \in e_{t-s}^{-1}(A)) \\ &= \mathbb{P}(X_{t-s} \in e_{t-s}^{-1}(A)) \\ &= \mathbb{P}(Y_{t-s} \in A). \end{aligned}$$

- Let $t_0 \geq 0$, $(s_n)_n$ a non-decreasing sequence converging to t_0, $(t_n)_n$ a non-increasing sequence converging to t_0 and $\omega \in \Omega$ such that $t \mapsto X_t(\omega)$ is cadlag. Then we have

$$Y_{t_n}(\omega) = (t_n, X_{t_n}(\omega)) \xrightarrow{n \to \infty} (t_0, X_{t_0}(\omega))$$

and we also have

$$\lim_{n \to \infty} Y_{s_n}(\omega) = \lim_{n \to \infty} (s_n, X_{s_n}(\omega)) = (t_0, \lim_{n \to \infty} X_{s_n}(\omega))$$

and that means that the left limit of $t \mapsto Y_t(\omega)$ exists at t_0 and so $(Y_t)_t$ is almost sure cadlag. So $(Y_t)_t$ is a Lévy process.

□

The following lemma will be used in Chapter 4 to prove a convergence.

Lemma 1.4.2. *Let $(X_t)_t$ be a real-valued Lévy process defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, $K \geq 0$ be a constant and $k : \mathbb{R} \to \mathbb{R}$ be a continuous function. Furthermore, $\mathcal{T}$ be the set of finite stopping times with respect to the filtration generated by $(X_t)_t$. We have then for $x \in \mathbb{R}$*

$$\left| \sup_{\tau \in \mathcal{T}} \left(\mathbb{E}_x \Big(\left(e^{X_\tau} - K \right)^+ \vee k(x) \Big) \right) - \sup_{\tau \in \mathcal{T}} \left(\mathbb{E}_x \Big(\left(e^{X_\tau} - K \right)^+ \Big) \right) \right| \leq k(x).$$

Proof. For every $\omega \in \Omega$ and $\tau \in \mathcal{T}$ we have

$$\begin{aligned} \left(e^{X_t(\omega)} - K \right)^+ + \left| k(x) \right| &\geq \left(e^{X_\tau(\omega)} - K \right)^+ \vee k(x) \\ &\geq \left(e^{X_\tau(\omega)} - K \right)^+ \geq 0 \end{aligned}$$

and it follows

$$\begin{aligned} |k(x)| &= \left| \left(e^{X_\tau(\omega)} - K \right)^+ + k(x) - \left(e^{X_\tau(\omega)} - K \right)^+ \right| \\ &\geq \left| \left(e^{X_\tau(\omega)} - K \right)^+ \vee k(x) - \left(e^{X_\tau(\omega)} - K \right)^+ \right| \end{aligned}$$

So taking expectation and supremum of both sides we get

$$\sup_{\tau \in \mathcal{T}} \mathbb{E}_x \left(\left| \left(\left(e^{X_\tau} - K \right)^+ \vee k(x) \right) - \left(e^{X_\tau} - K \right)^+ \right| \right) \leq k(x). \tag{1.21}$$

Next we want to use that for any two collections of real numbers $\{a_i : i \in I\}$ and $\{b_i : i \in I\}$ with I an index set we have

$$\sup\{a_i : i \in I\} - \sup\{b_i : i \in I\} \leq \sup\{a_i - b_i : i \in I\}. \tag{1.22}$$

To see this we fix an $i_0 \in I$ and see that

$$a_{i_0} = (a_{i_0} - b_{i_0}) + b_{i_0} \leq \sup\{a_i - b_i : i \in I\} + \sup\{b_i : i \in I\}.$$

So $\sup\{a_i - b_i : i \in I\} + \sup\{b_i : i \in I\}$ is an upper bound of $\{a_i : i \in I\}$ and we have

$$\sup\{a_i : i \in I\} \leq \sup\{a_i - b_i : i \in I\} + \sup\{b_i : i \in I\}$$

and subtraction of $\sup\{b_i : i \in I\}$ on both sides gives inequality (1.22).
We can get now our inequality

$$\begin{aligned} &\left| \sup_{\tau \in \mathcal{T}} \left(\mathbb{E}_x \Big(\left(e^{X_\tau} - K \right)^+ \vee k(x) \Big) \right) - \sup_{\tau \in \mathcal{T}} \left(\mathbb{E}_x \Big(\left(e^{X_\tau} - K \right)^+ \Big) \right) \right| \\ &= \sup_{\tau \in \mathcal{T}} \left(\mathbb{E}_x \Big(\left(e^{X_\tau} - K \right)^+ \vee k(x) \Big) \right) - \sup_{\tau \in \mathcal{T}} \left(\mathbb{E}_x \big((e^{X_\tau} - K)^+ \big) \right) \\ &\overset{(1.22)}{\leq} \sup_{\tau \in \mathcal{T}} \left(\mathbb{E}_x \Big(\left(e^{X_\tau} - K \right)^+ \vee k(x) \Big) - \mathbb{E}_x \Big(\left(e^{X_\tau} - K \right)^+ \Big) \right) \\ &= \sup_{\tau \in \mathcal{T}} \mathbb{E}_x \left(\left| \left(\left(e^{X_\tau} - K \right)^+ \vee k(x) \right) - \left(e^{X_\tau} - K \right)^+ \right| \right) \\ &\overset{(1.21)}{\leq} k(x). \end{aligned}$$

□

1.5 Spectrally negative Lévy Processes

The big difference between Lévy processes and Brownian motion is that the paths of Lévy processes do not have to be continuous and can contain jumps. The theory we will develop in the next chapter is not working if we allow positive jumps for our process. This is why we introduce now Lévy processes with only negative jumps. Such processes are called in the literature spectrally negative Lévy processes. See for example [2].

Definition 1.5.1 (Positive jump)**.** *We say that a function $f : \mathbb{R} \to \mathbb{R}$ does not have a positive jump at $x_0 \in \mathbb{R}$ if and only if $\lim_{x \uparrow x_0} f(x) \geq \lim_{x \downarrow x_0} f(x)$.*

Definition 1.5.2 (Spectrally negative Lévy process)**.** *A spectrally negative Lévy process is a real valued Lévy process where each path does not have positive jumps. We say that a process $(X_t)_t$ is a spectrally positive Lévy process if and only if $(-X_t)_t$ is a negative Lévy process.*

Chapter 2

The Scale Functions W and Z

In this chapter we introduce the scale functions W and Z. They can be a useful tool in the context of spectrally negative Lévy processes. The reason we are introducing them is Theorem 3.3.2 of the following chapter. The theorem gives an analytic solution for a perpetual American call option which depends on the scale functions. The definition of the scale function consists of the inverse Laplace transform of a function and the Laplace exponent of a Lévy process. For this reason we intoduce this terminology first in this chapter.

In the first section we define the inverse Laplace transform and show that it is well defined. We also derive an inverse Laplace transform for a specific function which we will use in the third section. In the second section we define the Laplace exponent of a Lévy process and derive the Laplace exponents of a scaled Brownian motion with drift and of a compound Poisson process. In the third section we introduce the scale functions W and Z for a spectrally negative Lévy process and derive an analytic solution for them in case that the underlying process is a scaled Brownian motion with drift. Theorem 2.1.1 is inspired by [10]. The proof of Lemma 2.2.1 uses [2, page 4]. Lemma 2.2.2 and Lemma 2.2.3 are standard results and their proofs are straightforward calculations. Many similiar proofs can be found in the literature. All other proofs are original by the author of this thesis.

2.1 The Laplace Transform

In this section we define the inverse Laplace transform and give an example which we will need in the next section. But first we have to define the Laplace transform.
We recall that for a measurable function $\xi : \Omega \to \mathbb{C}$ its integral (if it exists) is defined by

$$\int \xi(x)dx = \int \Re(\xi(x))dx + i \int \Im(\xi(x))dx.$$

Now we define the Laplace operator.

Definition 2.1.1 (Laplace transform)**.** *Let $f : [0, \infty) \to \mathbb{R}$ be a Borel measurable function. Then we define the constant*

$$c_f := \inf\{s \in \mathbb{R} : \int_0^\infty |e^{-st} f(t)| dt < \infty\}$$

and the set

$$D_f := \{s \in \mathbb{C} : \int_0^\infty |e^{-st} f(t)| dt < \infty\}.$$

We call the function

$$\mathcal{L}(f) : D_f \to \mathbb{C}, \quad s \mapsto \int_0^\infty e^{-ts} f(t) dt$$

the Laplace transform of f.

Let $f : [0, \infty) \to \mathbb{R}$ be a Borel measurable function. Now let $s_0 \in D_f$ and $s \in \mathbb{C}$ with $\Re(s_0) \leq \Re(s)$. Then we have

$$\begin{aligned}
\int_0^\infty |e^{-st} f(t)| dt &= \int_0^\infty |e^{-\Re(s)t} f(t)| \underbrace{|e^{-i\Im(s)t}|}_{=1} dt \\
&\leq \int_0^\infty |e^{-\Re(s_0)t} f(t)| \\
&= \int_0^\infty |e^{-\Re(s_0)t} f(t)||e^{-i\Im(s_0)t}| dt \\
&= \int_0^\infty |e^{-s_0 t} f(t)| dt < \infty.
\end{aligned}$$

So we have $s \in D_f$.
In the following we indentify $\mathbb{R}^2$ with $\mathbb{C}$ and work with the space which is respectively convenient. We have

$$D_f = \{s \in \mathbb{R} : \int_0^\infty |e^{-st} f(t)| dt < \infty\} \times \mathbb{R}.$$

So the domain of $\mathcal{L}(f)$ can only have the form $(c_f, \infty) \times \mathbb{R}$ or $[c_f, \infty) \times \mathbb{R}$.
Furthermore we see that $\mathcal{L}$ is linear in the sense that for given $f, g : [0, \infty) \to \mathbb{R}$ continuous and $a, b \in \mathbb{R}$ we have for every $s \in D_f \cap D_g$

$$\begin{aligned}
\mathcal{L}(af + bg)(s) &= \int_0^\infty e^{-ts}(af(t) + bg(t)) dt \\
&= a \int_0^\infty e^{-ts} f(t) dt + b \int_0^\infty e^{-ts} g(t)) dt \\
&= a\mathcal{L}(f)(s) + b\mathcal{L}(g)(s). \qquad (2.1)
\end{aligned}$$

The next lemma shows that the inverse Laplace transform is well defined. It is known as Lerchs theorem and our proof is inspired by [10].

Lemma 2.1.1 (Lerchs Theorem)**.** *Let $f, g : [0, \infty) \to \mathbb{R}$ be continuous functions with $c_f, c_g < \infty$ and assume there exists a $c > \max\{c_f, c_g, 1\}$ such that for every nonnegative $x \in \mathbb{R}$*

$$\mathcal{L}(f)(x) = \mathcal{L}(g)(x).$$

Then we have $f \equiv g$, i.e., $f(t) = g(t)$ for every $t \geq 0$.

Proof. Let $h = f - g$. So we have to prove $h \equiv 0$. First we see that for every $x \geq c$

$$\int_0^\infty |e^{-tx}h(t)|dt \leq \int_0^\infty |e^{-tx}f(t)|dt + \int_0^\infty |e^{-tx}g(t)|dt < \infty \tag{2.2}$$

and

$$\begin{aligned}\mathcal{L}(h)(x) &= \int_0^\infty e^{-tx}h(t)dt \\ &= \int_0^\infty e^{-tx}f(t)dt - \int_0^\infty e^{-tx}g(t)dt \\ &= \mathcal{L}(f)(x) - \mathcal{L}(g)(x) = 0.\end{aligned} \tag{2.3}$$

Now let us define

$$H : (0,1] \to \mathbb{R}, \quad t \mapsto t^c h(-\log(t))$$

We have then

$$\begin{aligned}H \equiv 0 \Leftrightarrow\ & \text{For every } t \in (0,1] : H(t) = 0 \\ \overset{t^c \geq 0}{\Leftrightarrow}\ & \text{For every } t \in (0,1] : h(-\log(t)) = 0 \\ \Leftrightarrow\ & \text{For every } t \in [0,\infty) : h(t) = 0 \\ \Leftrightarrow\ & h \equiv 0.\end{aligned}$$

So we only have to prove that $H \equiv 0$.
We see that $H \in L^1$ by

$$\begin{aligned}\int_0^1 |H(t)|dt &= \int_0^1 |t^c h(-\log(t))|dt \\ &\overset{\text{Mon. Conv.}}{=} \lim_{a \nearrow \infty} \int_{e^{-a}}^{e^{-0}} |t^c h(-\log(t))|dt \\ &\overset{\text{Int. by Subst.}}{=} \lim_{a \nearrow \infty} \int_0^a |(e^{-t})^c h(t) e^{-t}|dt \\ &= \lim_{a \nearrow \infty} \int_0^a |e^{-t(c+1)}h(t)|dt \\ &\overset{\text{Mon. Conv.}}{=} \int_0^\infty |e^{-t(c+1)}h(t)|dt \overset{(2.2)}{<} \infty.\end{aligned}$$

So we have for every $n \in \mathbb{N}$

$$\begin{aligned}\int_0^1 t^n H(t)dt &\overset{\text{Dom. Conv. } H \in L^1}{=} \lim_{a \to \infty} \int_{e^{-a}}^{e^{-0}} t^n H(t)dt \\ &\overset{\text{Int. by Subst.}}{=} \lim_{a \to \infty} \int_0^a (e^{-t})^n \underbrace{H(e^{-t})}_{= e^{-tc}h(t)} e^{-t}dt \\ &= \lim_{a \to \infty} \int_0^a e^{-t(c+n+1)}h(t)dt \\ &\overset{\text{Dom. Conv. 2.2}}{=} \int_0^\infty e^{-t(c+n+1)}h(t)dt \\ &= \mathcal{L}(h)(c+n+1) \overset{2.3}{=} 0.\end{aligned}$$

From the above, it follows that for every polynomial

$$p : [0,1] \to \mathbb{R}, \quad t \mapsto \sum_{n=1}^{N} a_n t^n \quad (a_1, ..., a_n \in \mathbb{R})$$

we have

$$\int_0^1 H(t)p(t) = \sum_{n=1}^{N} a_n \int_0^1 t^n H(t)dt = 0. \tag{2.4}$$

By the Weierstrass Approximation Theorem there exists a sequence of polynomials $(p_n)_n$ defined on $[0,1]$ with

$$\|H - p_n\|_\infty \xrightarrow{n\to\infty} 0.$$

So we get

$$\begin{aligned}\int_0^1 H^2(t)dt &= \int_0^1 H(t)(H(t) - p_n(t) + p_n(t)) \\ &= \int_0^1 H(t)(H(t) - p_n(t))dt + \underbrace{\int_0^1 p_n(t)H(t)dt}_{\overset{(2.4)}{=}0} \\ &\leq \|H - p_n\|_\infty \int_0^1 H(t)dt \xrightarrow{n\to\infty} 0\end{aligned}$$

So we have $\int_0^1 H^2(t)dt = 0$ and since H is continuous it follows that $H \equiv 0$ and we are done.

□

Definition 2.1.2 (Inverse Laplace transform)**.** *Let $D \subseteq \mathbb{C}$ contains a strip of the real line, i.e., there exists $a \in \mathbb{R}$ such that $[a, \infty) \subseteq D$ and let $F : D \to \mathbb{C}$ be a continuous function. If there exists a continuous function $f : [0, \infty) \to \mathbb{R}$ with $\mathcal{L}(f)(x) = F(x)$ for every $x \geq a$, then we write $f = \mathcal{L}^{-1}(F)$ and call f the inverse Laplace transform of F.*

The inverse Laplace operator is linear in the following sense:
Consider

$$F : D_f \subseteq \mathbb{C} \to \mathbb{C}$$

and

$$G : D_g \subseteq \mathbb{C} \to \mathbb{C}$$

where $\mathcal{L}^{-1}(F) := f$ and $\mathcal{L}^{-1}(G) := g$ exist and they are both continuous and let $a, b \in \mathbb{R}$ be arbitrary. We have then for every $x > \max\{c_f, c_g\}$

$$\mathcal{L}(a\mathcal{L}^{-1}(F) + b\mathcal{L}^{-1}(G))(x) \overset{(2.1)}{=} (aF + bG)(x)$$

So by Lemma 2.1.1 follows

$$\mathcal{L}^{-1}(aF + bG) = a\mathcal{L}^{-1}(F) + b\mathcal{L}^{-1}(G).$$

The following example will be useful in the next section.

Example 1. *For $a \in \mathbb{R} \setminus \{0\}$, $b, c \in \mathbb{R}$ with $b^2 - 4ac \neq 0$ we define*

$$D := \begin{cases} (\frac{-b-\sqrt{b^2-4ac}}{2a}, \infty) \times \mathbb{R}, & \text{if } b^2 - 4ac < 0 \\ (\frac{-b+\sqrt{b^2-4ac}}{2a}, \infty) \times \mathbb{R}, & \text{if } b^2 - 4ac > 0 \end{cases}$$

and

$$F : D \to \mathbb{C}, \quad s \mapsto \frac{1}{as^2 + bs + c}.$$

Since we know that $\frac{-b+\sqrt{b^2-4ac}}{2a}$ is the root of as^2+bs+c with the largest real part, we see that F is well defined. We want to find the inverse Laplace transform of F. For this we do the following partial fraction decomposition of F:

$$\begin{aligned}
\frac{1}{as^2 + bs + c} &= \frac{4a}{4a^2s^2 + 4abs + 4ac} \\
&= \frac{4a}{4a^2s^2 + 4abs + b^2 - (\sqrt{b^2 - 4ac})^2} \\
&= \frac{4a}{(2as + b)^2 - (\sqrt{b^2 - 4ac})^2} \\
&= \frac{4a}{\Big((2as + b) + \sqrt{b^2 - 4ac}\Big)\Big((2as + b) - \sqrt{b^2 - 4ac}\Big)} \\
&= \frac{2a}{\sqrt{b^2 - 4ac}} \cdot \frac{2\sqrt{b^2 - 4ac}}{(2as + b + \sqrt{b^2 - 4ac})(2as + b - \sqrt{b^2 - 4ac}} \\
&= \frac{2a}{\sqrt{b^2 - 4ac}} \cdot \frac{(2as + b + \sqrt{b^2 - 4ac}) - (2as + b - \sqrt{b^2 - 4ac})}{(2as + b + \sqrt{b^2 - 4ac})(2as + b - \sqrt{b^2 - 4ac}} \\
&= \frac{2a}{\sqrt{b^2 - 4ac}} \cdot \left(\frac{1}{2as + b - \sqrt{b^2 - 4ac}} - \frac{1}{2as + b + \sqrt{b^2 - 4ac}} \right) \\
&= \frac{1}{\sqrt{b^2 - 4ac}} \cdot \left(\frac{1}{s + \frac{b-\sqrt{b^2-4ac}}{2a}} - \frac{1}{s + \frac{b+\sqrt{b^2-4ac}}{2a}} \right) \\
&= \frac{1}{\sqrt{b^2 - 4ac}} \cdot \left(\frac{1}{s - \frac{-b+\sqrt{b^2-4ac}}{2a}} - \frac{1}{s - \frac{-b-\sqrt{b^2-4ac}}{2a}} \right) \qquad (2.5)
\end{aligned}$$

Since the inverse Laplace transform is linear our problem reduces to finding the inverse Laplace transform of

$$G : (d, \infty) \times \mathbb{R}, \quad s \mapsto \frac{1}{s - d}$$

for some $d \in \mathbb{R}$. If we define now

$$g : [0, \infty) \to \mathbb{R}, \quad t \mapsto e^{dt}$$

Then we get for every $s \in (d, \infty) \times \mathbb{R}$

$$\begin{aligned}\mathcal{L}(g)(s) &= \int_0^\infty e^{-st} e^{td} dt \\ &= \int_0^\infty e^{(d-s)t} dt \\ &\overset{Dom.\ Conv}{=} \lim_{a\to\infty} \int_0^a e^{(d-s)t} dt \\ &= \lim_{a\to\infty} \frac{1}{d-s}[e^{(d-s)a} - e^{(d-s)\cdot 0}] \\ &\overset{\Re(d-s)<0}{=} \frac{1}{d-s}.\end{aligned}$$

So we have $g = \mathcal{L}^{-1}(G)$. *So by putting this together with (2.5) and using the linearity of* $\mathcal{L}^{-1}$ *we get for every* $t \geq 0$

$$\mathcal{L}^{-1}(F)(t) = \frac{1}{b^2 - 4ac}\Big(e^{(\frac{-b+\sqrt{b^2-4ac}}{2a})t} - e^{(\frac{-b-\sqrt{b^2-4ac}}{2a})t}\Big)$$

2.2 The Laplace Exponent

The reason that we are introducing now the Laplace exponent of a Lévy process is Lemma 2.3.1 of the next section. We will see that we need the Laplace exponent to find a nice solution for the value function of the stopping problem. Since we will be interested in the next chapters in processes $(X_t)_t$ consisting of Brownian motion and Poisson process we derive in this section the Laplace exponents for these cases.
Before we define the Laplace exponent we present the following lemma which motivates the definition.

Lemma 2.2.1. *Let* $(Y_t)_t$ *be a Lévy process. Then we have for every* $t \geq 0$

$$\log\Big(\mathbb{E}e^{Y_t}\Big) = t\log\Big(\mathbb{E}e^{Y_1}\Big).$$

Proof. Let $t \geq 0$ be a real number and n be a natural number. We can write Y_t as

$$Y_t = Y_{\frac{t}{n}} + \sum_{k=1}^{n-1}\Big(Y_{\frac{(k+1)t}{n}} - Y_{\frac{kt}{n}}\Big). \tag{2.6}$$

By using the second and third property of Lévy processes we get

$$\begin{aligned}\mathbb{E}\big(\exp(Y_t)\big) &\overset{(2.6)}{=} \mathbb{E}\Bigg(\exp\Big(Y_{\frac{t}{n}}\Big)\prod_{k=1}^{n-1}\Big(\exp\Big(Y_{\frac{(k+1)t}{n}} - Y_{\frac{kt}{n}}\Big)\Big)\Bigg) \\ &\overset{\text{second}}{=} \mathbb{E}\Big(\exp\Big(Y_{\frac{t}{n}}\Big)\Big)\prod_{k=1}^{n-1}\Big(\mathbb{E}\Big(\exp\Big(Y_{\frac{(k+1)t}{n}} - Y_{\frac{kt}{n}}\Big)\Big)\Big) \\ &\overset{\text{third}}{=} \mathbb{E}\Big(\exp\Big(Y_{\frac{t}{n}}\Big)\Big)\prod_{k=1}^{n-1}\Big(\mathbb{E}\Big(\exp\Big(Y_{\frac{t}{n}}\Big)\Big)\Big) \\ &= \Big(\mathbb{E}\Big(\exp\big(Y_{\frac{t}{n}}\big)\Big)\Big)^n.\end{aligned}$$

Taking the logarithm on both sides gives

$$\log\left(\mathbb{E}e^{Y_t}\right) = n\log\left(\mathbb{E}\Big(\exp\left(X_{\frac{t}{n}}\right)\Big)\right). \tag{2.7}$$

Now let m be a natural number. If we substitute t by m in equation (2.7) we get

$$\log\left(\mathbb{E}e^{Y_m}\right) = n\log\left(\mathbb{E}\Big(\exp\left(X_{\frac{m}{n}}\right)\Big)\right). \tag{2.8}$$

If we substitute t by m and also n by m in equation (2.7) we get

$$\log\left(\mathbb{E}e^{Y_m}\right) = m\log\left(\mathbb{E}\Big(\exp\left(X_1\right)\Big)\right). \tag{2.9}$$

Now we put together equation (2.8) and (2.9) and devide by n on both sides. We get

$$\frac{m}{n}\log\left(\mathbb{E}\Big(\exp\left(X_1\right)\Big)\right) = \log\left(\mathbb{E}\Big(\exp\left(X_{\frac{m}{n}}\right)\Big)\right). \tag{2.10}$$

Now let $t_0 \geq 0$ be a real number and $(q_n)_n$ be a nonincreasing sequence of rational numbers which is converging to $t)_0$. Because Lévy processes are almost sure right-continuous we have

$$X_{q_n} \xrightarrow{n\to\infty} X_{t_0} \quad \text{a.s..} \tag{2.11}$$

We get now by the Dominated Convergence Theorem (DCT)

$$\begin{aligned}
\log\left(\mathbb{E}e^{Y_{t_0}}\right) &\overset{\text{DCT}}{=} \lim_{n\to\infty}\log\left(\mathbb{E}e^{Y_{q_n}}\right) \\
&\overset{(2.10)}{=} \lim_{n\to\infty} q_n\log\left(\mathbb{E}e^{Y_1}\right) \\
&= t_0\log\left(\mathbb{E}e^{Y_1}\right)
\end{aligned}$$

□

So if we know $\log\left(\mathbb{E}e^{Y_1}\right)$, we know $\log\left(\mathbb{E}e^{Y_t}\right)$ for every $t \geq 0$. This motivates the following definition.

Definition 2.2.1 (Laplace exponent)**.** *Let $Y : \Omega \to \mathbb{R}$ be a random variable. Then we define the Laplace exponent ψ of Y to be the function*

$$\psi : [0,\infty) \to \mathbb{R}, \quad \theta \mapsto \log\left(\mathbb{E}e^{\theta Y}\right).$$

For a Lévy process $(Y_t)_t$ we call the Laplace exponent of Y_1 the Laplace exponent of $(Y_t)_t$.

We define now the (compound) Poisson process and derive its Laplace exponent.

Definition 2.2.2 (Poisson process)**.** *A Poisson process $(N_t)_t$ with parameter λ is a Lévy process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ which is a counting process, i.e.,*

- *for every $t \geq 0$ N_t takes only values in $\mathbb{N}_0$*

- *for every* $\omega \in \Omega$ $t \mapsto N_t(\omega)$ *is nondecreasing*
- *for every* $\omega \in \Omega$ $\lim_{s \downarrow t} N_s(\omega) \leq \lim_{s \uparrow t} N_s(\omega)$

and for every $0 \leq s \leq t$ $N(t) - N(s) \sim Pois(\lambda(t-s))$*, i.e., for every* $k \in \mathbb{N}_0$

$$\mathbb{P}(N_t - N_s = k) = \frac{(\lambda(t-s))^k}{k!} e^{-\lambda(t-s)}.$$

We call a stochastic process $(X_t)_t$ *on* $(\Omega, \mathcal{F}, \mathbb{P})$ *a compound Poisson process with parameter* $\lambda > 0$ *if there exists a Poisson process* $(N_t)_t$ *with parameter* λ *and a sequence of i.i.d. random variables* $(\xi_i)_{i \in \mathbb{N}}$ *independent of* $(N_t)_t$ *so that for every* $\omega \in \Omega$ $X_t(\omega) = \sum_i^{N_t(\omega)} \xi_i(\omega)$*.*

Lemma 2.2.2 (Laplace exponent of a Poisson process)**.** *Let* $(X_t)_t$ *be a compound Poisson process, i.e.,* $X_t = \sum_{i=1}^{N(t)} \xi_i$ *for* $(\xi_i)_{i \in \mathbb{N}}$ *a sequence of independent and identical distributed random variables with distribution* F *and* $(N_t)_t$ *a Poisson process independent of* ξ_i*. Then* $(X_t)_t$ *has the Laplace exponent*

$$\psi(\theta) = \log(\mathbb{E}e^{\theta X_1}) = \lambda(\int_{\mathbb{R}} e^{\theta x} F(dx) - 1).$$

Proof. We first consider that

$$\mathbb{E}1_{\{N_1 = n\}} = \mathbb{P}(N_1 = n) = \frac{\lambda^n}{n!} e^{-\lambda} \tag{2.12}$$

and

$$\begin{aligned} \mathbb{E}(\exp(\theta \sum_{i=1}^{n} \xi_i)) &= \mathbb{E}(\prod_{i=1}^{n} \exp(\theta \xi_i)) \\ &\overset{(\xi_i)_i \text{ ind.}}{=} \prod_{i=1}^{n} \mathbb{E}(\exp(\theta \xi_i)) \\ &\overset{(\xi_i)_i \text{ id. distr.}}{=} (\mathbb{E}(\exp(\theta \xi_i)))^n. \end{aligned} \tag{2.13}$$

So we can calculate

$$\begin{aligned}
\exp(\psi(\theta)) &= \mathbb{E}(\exp(\theta X_1)) \\
&\overset{\text{Mon. Conv.}}{=} \Big(\sum_{n=0}^{\infty} \mathbb{E}(\exp(\theta X_1) 1_{\{N_1=n\}}) \\
&= \sum_{n=0}^{\infty} \mathbb{E}(\exp(\theta \sum_{i=1}^{N_1} \xi_i) 1_{\{N_1=n\}}) \\
&\overset{N_1 \text{ind. of} \xi_i}{=} \sum_{n=0}^{\infty} \mathbb{E}(\exp(\theta \sum_{i=1}^{n} \xi_i)) \mathbb{E}(1_{\{N_1=n\}}) \\
&\overset{2.2,2.13}{=} \sum_{n=0}^{\infty} \mathbb{E}(e^{\theta \xi_1})^n \frac{\lambda^n}{n!} e^{-\lambda} \\
&= e^{-\lambda} \sum_{n=0}^{\infty} \frac{(\lambda \mathbb{E} e^{\theta \xi_1})^n}{n!} \\
&= e^{-\lambda} e^{\lambda \mathbb{E} e^{\theta \xi_1}} \\
&= e^{\lambda(\mathbb{E}(e^{\theta \xi_1})-1)} \\
&= e^{\lambda(\int_{\mathbb{R}} e^{\theta x} F(dx)-1)}
\end{aligned}$$

and taking the logarithm gives the desired result. □

We define now Brownian motion and derive its Laplace exponent.

Definition 2.2.3 (Brownian motion). *We call a Lévy process $(B_t)_t$ on $(\Omega, \mathcal{F}, \mathbb{P})$ Brownian motion if*

- *for almost every $\omega \in \Omega$ $t \mapsto B_t(\omega)$ is continuous and*
- *for $t \geq s \geq 0$ $B_t - B_s \sim N(0, t-s)$.*

Lemma 2.2.3 (Laplace exponent of a Brownian motion). *Let $(X_t)_t$ be a scaled Brownian motion with linear drift, i.e.,*

$$X_t = sB_t + \gamma t \quad \text{for } s, \gamma \in \mathbb{R} \text{ and } (B_t)_t \text{ Brownian motion.}$$

Then $(X_t)_t$ has the Laplace exponent

$$\psi(\theta) = \log(\mathbb{E} e^{\theta X_1}) = \frac{1}{2} s^2 \theta^2 + \gamma \theta.$$

Proof. Let us fix some θ and write $z := \theta s$. Furthermore we consider the function $g(x) = \frac{x}{\sqrt{2}} - \frac{z}{\sqrt{2}}$ and we see that $g'(x) = \frac{1}{\sqrt{2}}$, $\lim_{x \to -\infty} g(x) = -\infty$ and

$\lim_{x\to\infty} g(x) = \infty$. We then have

$$\begin{aligned}
e^{-\theta\gamma}\exp(\psi(\theta)) &= e^{-\theta\gamma}\mathbb{E}e^{\theta(sB_1+\gamma)} \\
&= \mathbb{E}e^{zB_1} \\
&= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{zx}e^{-\frac{1}{2}x^2}dx \\
&= \frac{1}{\sqrt{2\pi}}e^{\frac{1}{2}z^2}\int_{\mathbb{R}} e^{-(\frac{1}{2}x^2-zx+\frac{1}{2}z^2)}dx \\
&= \frac{1}{\sqrt{2\pi}}e^{\frac{1}{2}z^2}\int_{\mathbb{R}} e^{-(\frac{x}{\sqrt{2}}-\frac{z}{\sqrt{2}})^2}dx \\
&= \frac{1}{\sqrt{\pi}}e^{\frac{1}{2}z^2}\int_{\mathbb{R}} e^{-g(x)^2}g'(x)dx \\
&= e^{\frac{1}{2}z^2}.
\end{aligned}$$

If we multiply now both sides with $e^{\theta\gamma}$ and take the logarithm we get

$$\begin{aligned}
\psi(\theta) &= \log(e^{\theta\gamma}e^{\frac{1}{2}z^2}) \\
&= \theta\gamma + \frac{1}{2}z^2 \\
&= \theta\gamma + \frac{1}{2}\theta^2 s^2.
\end{aligned}$$

□

Corollary 2.2.1. *Let $(Y_t)_t$ be a stochastic process which consists of a compound Poisson process component, a scaled Brownian motion component and some drift, i.e., there is a compound Poisson process $(X_t)_t$ with parameter $\lambda > 0$ and jump distribution F, a Brownian motion $(B_t)_t$ independent of $(X_t)_t$ and parameters $s, \gamma \in \mathbb{R}$ such that for every $t \geq 0$*

$$Y_t = X_t + sB_t + \gamma t.$$

Then the Laplace exponent ψ of $(Y_t)_t$ is given by

$$\psi(\theta) = \lambda\Big(\int_{\mathbb{R}} e^{\theta x}F(dx) - 1\Big) + \frac{1}{2}s^2\theta^2 + \gamma\theta.$$

Proof. Let ψ_1 be the Laplace exponent of $(X_t)_t$ and let ψ_2 be the Laplace exponent of $(sB_t + \gamma t)_t$. We have then

$$\begin{aligned}
\psi(\theta) &= \log\Big(\mathbb{E}e^{\theta Y}\Big) = \log\Big(\mathbb{E}e^{\theta X_t}e^{\theta(sB_t+\gamma t)}\Big) \\
&\overset{X_t, B_t \text{ indep.}}{=} \log\Big(\mathbb{E}\Big(e^{\theta X_t}\Big)\Big(\mathbb{E}e^{\theta(sB_t+\gamma t)}\Big)\Big) \\
&= \log\Big(\mathbb{E}\Big(e^{\theta X_t}\Big)\Big) + \log\Big(\Big(\mathbb{E}e^{\theta(sB_t+\gamma t)}\Big)\Big) \\
&= \psi_1(\theta) + \psi_2(\theta) \\
&\overset{\text{Lemma } 2.2.2, 2.2.3}{=} \lambda\Big(\int_{\mathbb{R}} e^{\theta x}F(dx) - 1\Big) + \frac{1}{2}s^2\theta^2 + \gamma\theta.
\end{aligned}$$

□

2.3 The Scale Functions for a Lévy Process W and Z

We need the following Lemma to define the scale functions.

Lemma 2.3.1. *Let $(X_t)_t$ be a spectrally negative Lévy process with Laplace exponent ψ and $r > 0$. Furthermore, we define the constant*

$$\phi(r) := \sup\{\lambda \geq 0 : \psi(\lambda) = r\}.$$

Then, there exists a unique continuous function $W^{(r)} : \mathbb{R} \to \mathbb{R}$ with $W(x) = 0$ for every negative $x \in \mathbb{R}$ such that

$$\mathcal{L}\left(W^{(r)}|_{[0,\infty)}\right)(t) = \frac{1}{\psi(t) - r}$$

for every $t > \phi(r)$.

Proof. [2, Theorem 8.1] □

Definition 2.3.1 (Scale functions)**.** *Let $(X_t)_t$ be a spectrally negative Lévy process and $r > 0$. Now let $W = W^{(r)}$ be the function defined in Lemma 2.3.1. Furthermore let us define the function*

$$Z^{(r)} = Z : \mathbb{R} \to \mathbb{R}, \quad x \mapsto 1 + r \int_0^x W(y)dy.$$

We call then W and Z the scale functions of $(X_t)_t$.

Remark. *The reader might wonder why we give the same name for W and Z. We do this to be consistent with the literature. See for example [1] and [2]. In this thesis we use exclusively the notation W and Z. So this way there should be no confusion which function is meant.*

We can use now the results from the last two sections for scaled Brownian motion with drift to derive its scale functions W and Z.

Lemma 2.3.2. *Let the process $(X_t)_t$ be given by $X_t = sW_t + \gamma t$ with γ, s being real constant, $(W_t)_t$ being a Brownian motion and $s \neq 0$. Then the scale functions W and Z for $(X_t)_t$ are given by*

$$W(t) = \begin{cases} 0, & t < 0 \\ \frac{1}{d}\left(e^{c_1 t} - e^{c_2 t}\right) & t \geq 0 \end{cases} \tag{2.14}$$

and

$$Z(t) = \begin{cases} 1, & t < 0 \\ 1 + \frac{r}{d}\left(\frac{e^{c_1 t} - 1}{c_1} - \frac{e^{c_2 t} - 1}{c_2}\right) & t \geq 0 \end{cases}$$

for $d := \sqrt{\gamma^2 + 2rs^2}$, $c_1 := \frac{-\gamma + d}{s^2}$ and $c_2 := \frac{-\gamma - d}{s^2}$.

Proof. We first derive W. That $W \equiv 0$ on $(-\infty, 0]$ follows by definition of the scale function.

Now let us define $a := \frac{1}{2}s^2$, $b = \gamma$ and $c := -r$. By Lemma 2.2.3 we know that the Laplace exponent of $(X_t)_t$ is given by

$$\psi(\theta) = \frac{1}{2}s^2\theta^2 + \gamma\theta = a\theta^2 + b\theta. \tag{2.15}$$

We have then

$$b^2 - 4ac = \gamma^2 + \underbrace{2rs^2}_{>0} > 0$$

and

$$a = \frac{1}{2}s^2.$$

So, we can define the subset of $\mathbb{C}$

$$D := \left(\frac{-b + \sqrt{b^2 - 4ac}}{2a}, \infty\right) \times \mathbb{R},$$

the function

$$F : D \to \mathbb{C}, \quad s \mapsto \frac{1}{as^2 + bs + c}$$

and the function

$$f : [0, \infty) \to \mathbb{R}, \quad t \mapsto \frac{1}{\sqrt{b^2 - 4ac}}\left(e^{\frac{-b+\sqrt{b^2-4ac}}{2a}} - e^{\frac{-b-\sqrt{b^2-4ac}}{2a}}\right).$$

We have then

$$\sqrt{b^2 - 4ac} = \sqrt{\gamma^2 + 2rs^2} = d$$

and it follows

$$f(t) = \frac{1}{d}\left(e^{\frac{-\gamma+d}{s^2}t} - e^{\frac{-\gamma-d}{s^2}t}\right) \tag{2.16}$$

$$= \frac{1}{d}\left(e^{c_1 t} - e^{c_2 t}\right) \tag{2.17}$$

Let us consider the constant

$$\phi(r) := \sup\{\lambda \geq 0 : \psi(\lambda) = r\} = \frac{-b + \sqrt{b^2 - 4ac}}{2a}.$$

By Example 1 we know that for every $t > \phi(r)$

$$\mathcal{L}(f)(t) = F(t) = \frac{1}{\psi(t) - r}$$

and by Lemma 2.3.1 We have

$$W_{[0,\infty)} = f.$$

Now we derive Z. Since $r > 0$ and $s \neq 0$ we see that $d \neq 0$. Now let us check that c_1 and c_2 are unequal zero as well. For the case $\gamma \geq 0$ we have

$$2rs^2 > 0 \Rightarrow \gamma^2 + 2rs^2 > \gamma^2 \Rightarrow d = \sqrt{\gamma^2 + 2rs^2} > \gamma \Rightarrow c_1 = \frac{d - \gamma}{s^2} > 0$$

and

$$2rs^2 > 0 \Rightarrow d = \sqrt{\gamma^2 + 2rs^2} > 0 \Rightarrow -d < 0 \Rightarrow c_2 = \frac{-\gamma - d}{s^2} < 0.$$

For the case $\gamma < 0$ we have

$$\gamma < 0 \leq \sqrt{\gamma^2 + 2rs^2} = d \Rightarrow c_1 = \frac{d - \gamma}{s^2} > 0$$

and

$$2rs^2 > 0 \Rightarrow \gamma^2 + 2rs^2 > \gamma^2 \Rightarrow d = \sqrt{\gamma^2 + 2rs^2} > \sqrt{\gamma^2} = -\gamma$$
$$\Rightarrow c_2 = \frac{-\gamma - d}{s^2} < 0.$$

So we can calculate now for $t \geq 0$

$$\begin{aligned} Z(t) &= 1 + r \int_0^t W(s) ds \\ &= 1 + \frac{r}{d} \int_0^t \left(e^{c_1 s} - e^{c_2 s} \right) ds \\ &= 1 + \frac{r}{d} \left[\frac{1}{c_1} e^{c_1 s} - \frac{1}{c_2} e^{c_2 s} \right]_0^t \\ &= 1 + \frac{r}{d} \left(\frac{e^{c_1 t} - 1}{c_1} - \frac{e^{c_2 t} - 1}{c_2} \right) \end{aligned}$$

and for $t < 0$ we have

$$Z(t) = 1 + r \int_0^t \underbrace{W(s)}_{=0} ds = 1.$$

□

In case that our process contains additionally a compound Poisson process we have derived an analytic form of its Laplace exponent ψ in section two. But unfortunately we cannot find the inverse Laplace transform of $\frac{1}{\psi(t) - r}$. So in this case we have to approximate the scale functions W and Z numerically.

Chapter 3

Perpetual American Options with Guarantee and Optimal Stopping

A perpetual American call option is a contract which gives the buyer (the owner or holder of the option) the right, but not the obligation, to buy a specific asset for a specified (strike) price at any time in the future. Let X_t be the value of the asset at time $t \geq 0$ and K be strike price of the option. If at some time $t \geq 0$ the value of the asset is larger than the strike price we could exercise the option. That means we could buy the asset for K and sell it for X_t. We would make a profit of $X_t - K$. We call this amount of profit the payoff. So in case of call option we have the payoff $f(X_t)$ with $f(x) := (x-K)^+$ and we call f the payoff function. Furthermore we make in our model the assumptoin that there exist also a risk-free investment possibility. That means we assume that if we are able to invest the amount $x \geq 0$ at time 0 and receive the amount xe^{-rt} at time t without risk, with $r > 0$ the interest rate being a constant.
There are different approaches to find the value of American options. One is the approach of the risk-neutral measure [9, Chapter 5.2] and optimal stopping. As can be seen for example by [9, Chapter 8.5.2], the value of the option is given by

$$\sup_{\tau \in \mathcal{T}} \mathbb{E}_x\Big(e^{-r\tau}\mathbb{E}_x(f(X_\tau))\Big)$$

where $\mathcal{T}$ is the set of stopping times. The most common approach to model the asset is by using geometric Brownian motion. One drawback of this approach is that Brownian motion is a continuous process and does not allow jumps. To avoid this restriction we will work with a broader class of processes, namely Lévy processes. Furthermore, we will not only consider the case of the call option where the payoff is defined by $f(x) := (x-K)^+$ but assume just that f is continuous and non-decreasing. Another generalisation we will make is that we allow the option to have a guaranteed amount of money that will be paid in any case when the option is exercised.
In the following section we consider the stopping problem for the general case that $(X_t)_t$ is given by a Markov process and present an existence theorem. In section two we use this theorem to show that we have a solution of the stopping

problem in the case of a Lévy process. In section three we show that we can find a nice solution to the stopping problem if we restrict ourselves to to spectrally negative Lévy processes. In the fourth section we derive a system of differential equations which could be used for computations.
This chapter is highly motivated by the paper [1]. Lemma 3.2.1, Theorem 3.3.1 and Theorem 3.3.2 can be found in [1][page 249-252]. The proofs in our thesis of these three theorem are just more detailed and are different in some parts. All other theorems in this chapter are original by the author of this thesis. The idea of chapter 4 is also from [1][page 249-252] but the calculations are not shown in the paper.

3.1 Optimal Stopping for Markov Processes

Let $(X_t)_t$ be an $(E, \mathcal{E})$-valued homogeneous Markov process with transition function $(p_t)_t$ and distribution $(\mathbb{P}_x)_{x\in E}$ defined on a measure space $(\Omega, \mathcal{F})$. Furthermore assume that the filtration $(\mathcal{F}_t)_{t\geq 0}$ generated by $(X_t)_t$ satisfies the usual conditions. This means that any $\mathbb{P}$-null set of $\mathcal{F}$ is contained in $\mathcal{F}_0$ and that the filtration $(\mathcal{F}_t)_{t\geq 0}$ is right-continuous, *i.e.* $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$ for every $t \geq 0$. Let $X_\infty : \Omega \to E$ be a random variable and $\mathcal{T}$ be the set of finite stopping times with respect to $(\Omega, (\mathcal{F}_t)_t)$. We assume that all paths of $(X_t)_t$ are right-continuous with left limits (i.e. they are cadlag). Furthermore, let there be a measurable function $G : E \to \mathbb{R}$, the gain function, that satisfies $G(X_\infty) = 0$ and

$$E_x(\sup_{0\leq t\leq\infty} |G(X_t)|) < \infty \tag{3.1}$$

for every $x \in \mathbb{R}$. So for every $\tau \in \mathcal{T}$

$$E_x|(G(X_\tau))| < \infty. \tag{3.2}$$

We call then the function

$$V : E \to \mathbb{R}, \quad x \mapsto \sup_{\tau\in\mathcal{T}} E_x(G(X_\tau)) \tag{3.3}$$

the value function. The stopping problem is now the task to find an analytic form of V and a stopping time τ_0 such that

$$V(x) = \sup_{\tau\in\mathcal{T}} E_x(G(X_\tau)) = E_x(G(X_{\tau_0})) \tag{3.4}$$

for every $x \in \mathbb{R}$.
We call

$$C = \{x \in E : V(x) > G(x)\}$$

the continuation set and

$$D = \{x \in E : V(x) = G(x)\}$$

the stopping set. The intuition behind these sets is the following. If $(X_t)_t$ starts at $x_0 \in E$ and we have $V(x_0) = G(x_0)$, then the trivial stopping time which stops at zero is optimal. Now let us assume that $(X_t)_t$ starts at $y_0 < x_0$ instead but still $V(x_0) = G(x_0)$ and that $(X_t)_t$ hits x_0 at some time $t_0 > 0$. Since the

process is Markov it has no relevance what happened between the times zero and t_0. So there is actually no difference to our first scenario and it is optimal to stop at t_0. This intuitive reasoning leads to the idea that

$$\tau_D(\omega) := \inf\{t \geq 0 : X_t(\omega) \in D\}$$

might be an optimal stopping time. That this under some conditions is indeed the case we will see in the following.

Definition 3.1.1. *A function $f : E \to \mathbb{R}$ is upper semi-continuous (usc) at a point $x_0 \in E$ if and only if for every $\epsilon > 0$ exists $\delta > 0$ such that $\|x - x_0\| < \delta$ implies $f(x) \leq f(x_0) + \epsilon$.*
A function $f : E \to \mathbb{R}$ is lower semi-continuous (lsc) at a point $x_0 \in E$ if and only if for every $\epsilon > 0$ exists $\delta > 0$ such that $\|x - x_0\| < \delta$ implies $f(x_0) \leq f(x) + \epsilon$.
We say that f is usc (lsc) if and only if f is usc (lsc) at every point.

Lemma 3.1.1. *For V lower semi-continuous and G upper semi-continuous, D is closed and consequently τ_D is a stopping time.*

Proof. Let $(x_n)_n \subset D$ be a converging sequence with limit $x_0 \in E$. Since V is lsc, for every $m \in \mathbb{N}$ exists a $\delta_V^m > 0$ such that

$$|x - x_0| < \delta_V^m \text{ implies } V(x_0) \leq V(x) + \frac{1}{m} \tag{3.5}$$

and because G is usc, there exist for every $m \in \mathbb{N}$ a $\delta_G^m > 0$ such that

$$|x - x_0| < \delta_G^m \text{ implies } G(x) \leq G(x_0) + \frac{1}{m}. \tag{3.6}$$

Now we define $\delta^m := \min\{\delta_V^m, \delta_G^m\}$. Since $(x_n)_n$ converges to x_0 there exists a subsequence $(x_{n_m})_m$ of $(x_n)_n$ with

$$|x_{n_m} - x_0| < \delta^m \text{ for every } m \in \mathbb{N}.$$

So we have for every $m \in \mathbb{N}$

$$V(x_0) \overset{(3.5)}{\leq} V(x_{n_m}) + \frac{1}{m} \overset{(x_n)_n \subseteq D}{=} G(x_{n_m}) + \frac{1}{m} \overset{(3.6)}{\leq} G(x_0) + \frac{2}{m}.$$

Letting m go to infinity and considering that $V \geq G$ by definition of V (just stopping at time zero) we get $V(x_0) = G(x_0)$. So $x_0 \in D$ and this means that D is closed.

Now let us define for every $n \in \mathbb{N}$ the compact sets $D_n := D \cap [-n, n]$ and the random variables

$$\tau : \Omega \to \mathbb{R}, \quad \omega \mapsto \inf\{t \geq 0 : X_t(\omega) \in D_n\}.$$

We have then for every $\omega \in \Omega$

$$\begin{aligned}\omega \in \{\tau_D \leq t\} &\Leftrightarrow \tau_D(\omega) \leq t \\ &\Leftrightarrow \inf\{s \geq 0 : X_s(\omega) \in D\} \leq t \\ &\Leftrightarrow \text{There exists } 0 \leq s_0 \leq t \text{ such that } X_{s_0}(\omega) \in D \\ &\Leftrightarrow \text{There exists } n \in \mathbb{N} \text{ and } 0 \leq s_0 \leq t \text{ such that } X_{s_0}(\omega) \in D_n \\ &\Leftrightarrow \text{There exists } n \in \mathbb{N} \text{ such that } \inf\{s \geq 0 : X_s(\omega) \in D_n\} \leq t \\ &\Leftrightarrow \text{There exists } n \in \mathbb{N} \text{ such that } \tau_{D_n}(\omega) \leq t \\ &\Leftrightarrow \text{There exists } n \in \mathbb{N} \text{ such that } \omega \in \{\tau_{D_n} \leq t\} \\ &\Leftrightarrow \omega \in \bigcup_{n\in\mathbb{N}} \{\tau_{D_n} \leq t\}\end{aligned} \tag{3.7}$$

and it follows

$$\{\tau_D \leq t\} = \bigcup_{n\in\mathbb{N}} \{\tau_{D_n} \leq t\}. \tag{3.8}$$

Because all paths of $(X_t)_t$ are cadlag, D_n is closed and $(\mathcal{F}_t)_t$ satisfies the usual conditions it follows by [6, Lemma II.75.1] that all τ_{D_n} are stopping times. So we have for every $n \in \mathbb{N}$ and $t \geq 0$

$$\{\tau_{D_n} \leq t\} \in \mathcal{F}_t$$

and since sigma algebras are closed under countable union follow for ever $t \geq 0$

$$\{\tau_D \leq t\} \overset{(3.8)}{=} \bigcup_{n\in\mathbb{N}} \{\tau_{D_n} \leq t\} \in \mathcal{F}_t$$

and this proves that τ_D is a stopping time.

□

Lemma 3.1.2. *If V is lsc, G usc and $\mathbb{P}(\tau_D < \infty) = 1$ for every $x \in \mathbb{R}$, then*

$$V(x) = \mathbb{E}_x G(X_{\tau_D})$$

for every $x \in E$.

Proof. A proof is given in [3, Chapter 1, Lemma 2.9]. □

3.2 Optimal Stopping for Lévy Processes

In the last section we have discussed the general case that our process $(X_t)_t$ is a Markov process. In this chapter we consider now the more special case that the process $(X_t)_t$ is a real-valued Lévy process. We are interested in the stopping problem where a fraction $k(x)$ of our starting value x is guaranteed, *i.e.*, our payoff is at least $k(x)$. So the stopping problem we are interested in has the form

$$v(x) = \sup_{\tau\in\mathcal{T}} \mathbb{E}_x\Big(e^{-r\tau}\big(g(X_\tau) \vee k(x)\big)\Big) \tag{3.9}$$

where $x \in \mathbb{R}$, $\mathcal{T}$ is the set of finite stopping times (with respect to the natural filtration of $(X_t)_t$), $g, k : \mathbb{R} \to \mathbb{R}$ are continuous non-decreasing function and

$r > 0$. To ensure that v is well defined, *i.e.*, that for every $x \in \mathbb{R}$ $v(x) \in \mathbb{R}$ we make the assumption

$$\mathbb{E}_x(\sup_{t\geq 0} e^{-rt}|g(X_t)|) < \infty \tag{3.10}$$

for every $x \in \mathbb{R}$. If we compare this stopping problem (3.9) with the stopping problem (3.4) from last section we spot two structural differences. First we see that the payoff in (3.9) depends not only on the value of the process when it stops at X_τ but also on the starting value x of the process at time zero and second we see that the payoff has the extra term $e^{-r\tau}$. We deal with the second difference in the next two lemmas by extending the process to $(t, X_t)_t$. To deal with the first difference we consider instead of the stopping problem (3.9) a whole family of stopping problems depending on $x \in \mathbb{R}$

$$\begin{aligned} v_x(y) &:= \sup_{\tau\in\mathcal{T}} E_y(e^{-r\tau}(g(X_\tau) \vee k(x))) \\ &= \sup_{\tau\in\mathcal{T}} E_y(e^{-r\tau} g_x(X_\tau)) \end{aligned} \tag{3.11}$$

with $g_x(y) := g(y) \vee k(x)$.
So our approach is that for every starting value $x \in \mathbb{R}$ we consider the stopping problem (3.11), find a solution for it and consequently find a solution for (3.9) at x since $v_x(x) = v(x)$ as we can see.
For every $x \in \mathbb{R}$ we define the stopping set

$$S_x := \{y \in \mathbb{R} : v_x(y) = g_x(y)\}. \tag{3.12}$$

The following lemmas gives us a candidate to solve the stopping problem, namely

$$\tau_{S_x}(\omega) := \inf\{t \geq 0 : X_t(\omega) \in S_x\}$$

In the following we will always assume that

$$\mathbb{P}_y(\tau_{S_x} < \infty) = 1 \quad \text{for every } y, x \in \mathbb{R}.$$

To make some proofs more readable we will work with the assumption that $\tau_{S_x}(\omega) < \infty$, for every $\omega \in \Omega$. It should always be clear that this simplification does not change the proof.
We will see in the following two lemmas that g being continuous is a sufficient condition for τ_{S_x} being an optimal stopping time.

Lemma 3.2.1. *Let $Y_t = (t, X_t)$ be the process with values in $(E, \mathcal{E})$ with $E = \mathbb{R}_+ \times \mathbb{R}$ and $\mathcal{E} = \mathcal{B}(\mathbb{R}_+ \times \mathbb{R})$ the Borel sets of E, $\hat{g} : E \to \mathbb{R}$ defined by $\hat{g}(t, x) := e^{-rt} g(x)$,*

$$w : E \to \mathbb{R}, \quad (t, x) \mapsto \sup_{\tau\in\mathcal{T}} \mathbb{E}_{(t,x)}\hat{g}(Y_\tau)$$

and

$$S_w = \{(t, x) \in E : \hat{g}(t, x) = w(t, x)\}.$$

Furthermore, we assume that g is continuous. Then we have $w(t, x) = \mathbb{E}_{(t,x)}\hat{g}(Y_{\tau_{S_w}})$ for every $(t, x) \in E$.

Proof. First we recall that $(Y_t)_t$ is a Lévy process by Lemma 1.4.1.
We define for every $\tau \in \mathcal{T}$ the function

$$f_\tau : E \to \mathbb{R}, \quad (t, x) \mapsto \mathbb{E}_{(t,x)}\hat{g}(Y_\tau).$$

To see that the f_τ are continuous let us fix some $\tau \in \mathcal{T}$ and consider a sequence $(t_n, x_n)_n \subseteq E$ converging to some $(t_0, x_0) \in E$. Since $(x_n)_n$ is converging, it is bounded and hence there exist $c, d \in \mathbb{R}$ with $c \leq x_n \leq d$ for every $n \in \mathbb{N}$. Now let us define the random variables

$$H_n(\omega) := \hat{g}(Y_\tau(\omega) + (t_n, x_n)),$$

$$H(\omega) := \hat{g}(Y_\tau(\omega) + (t_0, x_0))$$

and

$$Z(\omega) := \max\{e^{-r\tau}|g(X_\tau(\omega) + c)|, e^{-r\tau}|g(X_\tau(\omega) + d)|\}.$$

Because $\hat{g}$ is continuous we have for every $\omega \in \Omega$

$$H_n(\omega) \xrightarrow{n\to\infty} H(\omega).$$

To see that the sequence of random variables $(H_n)_n$ is dominated by the random variable Z let us fix some $n \in \mathbb{N}$ and $\omega \in \Omega$. We first consider the case $g(X_\tau(\omega) + x_n) \geq 0$.
We have then

$$\begin{aligned} |H_n(\omega)| &= |\hat{g}(Y_\tau(\omega) + (t_n, x_n))| \\ &= e^{-r(\tau(\omega)+t_n)}|g(X_\tau(\omega) + x_n)| \\ &\leq e^{-r\tau(\omega)}g(X_\tau(\omega) + x_n) \\ &\leq e^{-r\tau(\omega)}g(X_\tau(\omega) + d) \\ &\leq |Z(\omega)|. \end{aligned}$$

Now we consider the case $g(X_\tau(\omega) + x_n) \leq 0$.
Because g is nondecreasing we have then

$$0 \geq g(X_\tau(\omega) + x_n) \geq g(X_\tau(\omega) + c).$$

This gives

$$\begin{aligned} |g(X_\tau(\omega) + x_n)| &= -g(X_\tau(\omega) + x_n) \\ &\leq -g(X_\tau(\omega) + c) \\ &= |g(X_\tau(\omega) + c)| \end{aligned} \tag{3.13}$$

and it follows

$$\begin{aligned} |H_n(\omega)| &= e^{-r(\tau(\omega)+t_n)}|g(X_\tau(\omega) + x_n)| \\ &\leq e^{-r\tau(\omega)}|g(X_\tau(\omega) + x_n)| \\ &\overset{(3.13)}{\leq} e^{-r\tau(\omega)}|g(X_\tau(\omega) + c)| \\ &\leq |Z(\omega)|. \end{aligned}$$

So Z indeed dominates $(H_n)_n$. That Z is integrable we see by

$$\begin{aligned} \mathbb{E}|Z| &\leq \mathbb{E}\big(e^{-r\tau}|g(X_\tau + c)|\big) + \mathbb{E}\big(e^{-r\tau}|g(X_\tau + d)|\big) \\ &= \mathbb{E}_c\big(e^{-r\tau}|g(X_\tau)|\big) + \mathbb{E}_d\big(e^{-r\tau}|g(X_\tau)|\big) \\ &\leq \mathbb{E}_c\big(\sup_{t\geq 0} e^{-rt}|g(X_t)|\big) + \mathbb{E}_d\big(\sup_{t\geq 0} e^{-rt}|g(X_t)|\big) \overset{(3.10)}{<} \infty. \end{aligned}$$

So we get by the Dominated Convergence Theorem

$$\begin{aligned} f_\tau(t_n, x_n) &= \mathbb{E}_{(t_n,x_n)}\hat{g}(Y_\tau) \\ &= \mathbb{E}\hat{g}(Y_\tau + (t_n, x_n)) \\ &= \mathbb{E}H_n \xrightarrow{n\to\infty} \mathbb{E}H \quad \text{(Dominant Convergence)} \\ &= \mathbb{E}\hat{g}(Y_\tau + (t_0, x_0)) \\ &= \mathbb{E}_{(t_0,x_0)}\hat{g}(Y_\tau) \\ &= f_\tau(t_0, x_0). \end{aligned}$$

So f_τ is continuous.
We show now that w is lower semi-continuous. First we see that for every $(t, x) \in E$

$$w(t, x) = \sup_{\tau\in\mathcal{T}} f_\tau(t, x).$$

Now let us fix some $(t_0, x_0) \in E$ and $\epsilon > 0$. By the definition of supremum exists $\tau_0 \in \mathcal{T}$ with

$$w(t_0, x_0) \le f_{\tau_0}(t_0, x_0) + \frac{\epsilon}{2}.$$

Because f_{τ_0} is continuous there exists $\delta > 0$ so that

$$f_{\tau_0}(t_0, x_0) \le f_{\tau_0}(t, x) + \frac{\epsilon}{2}$$

for every $(t, x) \in E$ with $\|(t, x) - (t_0, x_0)\| < \delta$.
So we have then for every $(t, x) \in E$ with $\|(t, x) - (t_0, x_0)\| < \delta$

$$w(t_0, x_0) \le f_{\tau_0}(t_0, x_0) + \frac{\epsilon}{2} \le f_{\tau_0}(t, x) + \epsilon \le w(t, x) + \epsilon.$$

So w is lsc and the result follows by Lemma 3.1.2.

□

Lemma 3.2.2. *Let g be continuous. Then we have for every $x, y \in \mathbb{R}$,*

$$v_x(y) = \mathbb{E}_y e^{-r\tau_{S_x}} g_x(X_{\tau_{S_x}}),$$

i.e. for every $x \in \mathbb{R}$ τ_{S_x} is an optimal stopping time for the stopping problem (3.11).

Proof. Let us fix some $x \in \mathbb{R}$ and simply write $S = S_x$, $g = g_x$ and $\tau_S = \tau_{S_x}$. We define as in Lemma 3.2.1 $Y_t := (t, X_t)$

$$\hat{g}(t, y) := e^{-rt} g(y) \quad (t \ge 0, y \in \mathbb{R}),$$

$$w(t, y) := \sup_{\tau\in\mathcal{T}} \mathbb{E}_{t,y}\hat{g}(Y_\tau)$$

and

$$S_w := \{(t, y) \in \mathbb{R}_+ \times \mathbb{R} : w(t, y) = \hat{g}(t, y)\}.$$

We have then for $(t,y) \in \mathbb{R}_+ \times \mathbb{R}$

$$\begin{aligned} w(t,y) &= \sup_{\tau \in \mathcal{T}} \mathbb{E}_{t,y} \hat{g}(Y_\tau) \\ &= \sup_{\tau \in \mathcal{T}} \mathbb{E} \hat{g}(Y_\tau + (t,y)) \\ &= \sup_{\tau \in \mathcal{T}} \mathbb{E} \hat{g}(\tau + t, X_\tau + y) \\ &= \sup_{\tau \in \mathcal{T}} \mathbb{E} e^{-r(\tau+t)} g(X_\tau + y) \\ &= e^{-rt} \sup_{\tau \in \mathcal{T}} \mathbb{E} e^{-r\tau} g(X_\tau + y) \\ &= e^{-rt} \sup_{\tau \in \mathcal{T}} \mathbb{E}_y e^{-r\tau} g(X_\tau) \\ &= e^{-rt} v(y). \end{aligned} \tag{3.14}$$

With this we see that

$$(t,y) \in S_w \Leftrightarrow w(t,y) = \hat{g}(t,y) \overset{(3.14)}{\Leftrightarrow} e^{-rt} v(y) = e^{-rt} g(y)$$

$$\Leftrightarrow v(y) = g(y) \Leftrightarrow y \in S.$$

So we have $S_w = \mathbb{R}_+ \times S$ and that gives

$$\{Y_t \in S_w\} = \{(t, X_t) \in \mathbb{R}_+ \times S\} = \{X_t \in S\}$$

and finally

$$\begin{aligned} \tau_{S_w} &= \inf\{t \geq 0 : Y_t \in S_w\} \\ &= \inf\{t \geq 0 : X_t \in S\} \\ &= \tau_S \end{aligned}$$

We know by Lemma 3.2.1 that

$$w(t,y) = \mathbb{E}_{t,y} \hat{g}(Y_{\tau_{S_w}}). \tag{3.15}$$

So we get now for arbitrary $t \geq 0$

$$\begin{aligned} v(y) &\overset{(3.14)}{=} e^{rt} w(t,y) \\ &\overset{(3.15)}{=} e^{rt} \mathbb{E}_{t,y} \hat{g}(Y_{\tau_{S_w}}) \\ &= e^{rt} \mathbb{E} \hat{g}(Y_{\tau_S} + (t,y)) \\ &= e^{rt} \mathbb{E} e^{-r(\tau_S + t)} g(X_\tau + y) \\ &= \mathbb{E} e^{-r\tau_S} g(X_\tau + y) \\ &= \mathbb{E}_y e^{-r\tau_S} g(X_\tau). \end{aligned}$$

□

So we have proved that a solution of the stopping problem exists if g is continuous.

In the following we assume that v_x is continuous. Since τ_{S_x} is an optimal stopping time for our problem, we want now to understand the structure of S_x better. For this we define

$$a_x := \sup\{a \in S_x : a \leq x\}$$

and

$$b_x := \inf\{a \in S_x : a \geq x\}.$$

The following lemma gives us very useful information about the structure of S_x.

Lemma 3.2.3. *(a) We have $S_x \cap (-\infty, x] = (-\infty, a_x]$ with possibly $a_x = -\infty$ and $(-\infty, -\infty] := \emptyset$.*

(b) We have $b_x \in S_x \cap [x, \infty)$, implying that $b_x < \infty$.

Proof. (a) In case that $S_x \cap (-\infty, x] = \emptyset$ we just have $a_x = \sup \emptyset = -\infty$. Now let us consider the case $S_x \cap (-\infty, x] \neq \emptyset$:

Suppose first that y is an element of $S_x \cap (-\infty, x]$. Then $y \leq \sup\big(S_x \cap (-\infty, x]\big) = a_x$, *i.e.*, $y \in (-\infty, a_x]$.
Let us now show the other direction. Because $S_x \cap (-\infty, x]$ is an non-empty set of real numbers with an upper bound, $a_x = \sup(S_x \cap (-\infty, x]) \in \mathbb{R}$ exists. Since g is continuous, g_x is continuous as well and because v_x is also continuous the set

$$S_x = \{y \in \mathbb{R} : v_x(y) = g_x(y)\} = (v_x - g_x)^{-1}(\underbrace{\{0\}}_{\text{closed}})$$

is closed. So

$$\{a \in S_x : a \leq x\} = S_x \cap (-\infty, x]$$

is closed as well and contains its supremum a_x (which exists because the set has an upper bound), *i.e.*, $a_x \in S_x \cap (-\infty, x]$. Now take $y \leq a_x$. Then for every $\tau \in \mathcal{T}$ we have

$$\mathbb{E}_y e^{-r\tau} g_x(X_\tau) = \mathbb{E} e^{-r\tau} g_x(X_\tau + y) \overset{g \text{ nondecr.}}{\leq} \mathbb{E} e^{-r\tau} g_x(X_\tau + a_x)$$

$$= \mathbb{E}_{a_x} e^{-r\tau} g_x(X_\tau) \leq v_x(a_x) \overset{a_x \in S_x}{=} g_x(a_x) = g(a_x) \vee g(x)$$

$$\overset{g \text{ nondecr.}}{=} g(x) \leq g(x) \vee g(y) = g_x(y).$$

Taking supremum over all stopping times $\tau \in \mathcal{T}$ on both sides (right side is independent of the stopping time) gives then

$$v_x(y) \leq g_x(y).$$

Since trivially $g_x(y) \leq v_x(y)$ by definition of v_x follows $y \in S_x$. So because $y \leq a_x \leq x$ we have $y \in S_x \cap (-\infty, x]$ and (a) is proved.

(b) We first prove $S_x \cap [x, \infty) \neq \emptyset$ by contradiction. So let us assume that $S_x \cap [x, \infty) = \emptyset$. That means that all elements of S_x are smaller than x.

So since for every $\omega \in \Omega$ $X_{\tau_{S_x}}(\omega) \in S_x$ we have $X_{\tau_{S_x}} < x$ everywhere. So we have for every $\omega \in \Omega$

$$(g_x(X_{\tau_{S_x}}))(\omega) = g_x(X_{\tau_{S_x}}(\omega)) = g(x) \vee g(X_{\tau_{S_x}}(\omega)) \overset{g \text{ nondecr.}}{=} g(x). \quad (3.16)$$

Since by definition $\mathbb{P}_x(X_0 = x) = 1$ and $x \notin S_x$ we have $\mathbb{P}_x(X_0 \in S_x) = 0$. So we have $\tau_{S_x} > 0$ $\mathbb{P}_x$-a.s. and consequently $e^{-r\tau_{S_x}} < 1$ $\mathbb{P}_x$-a.s. and it follows

$$\mathbb{E}_x e^{-r\tau_{S_x}} < 1. \quad (3.17)$$

Putting this together we get

$$v_x(x) = \mathbb{E}_x e^{-r\tau_{S_x}} \underbrace{g_x(X_{\tau_{S_x}})}_{\overset{(3.16)}{=} g(x)} = g(x)\mathbb{E}_x e^{-r\tau_{S_x}} \overset{(3.17)}{<} g(x) \leq g_x(x) \quad (3.18)$$

and this is a contradiction because obviously $v_x(x) \geq g_x(x)$ (We just have to consider stopping at zero). So $S_x \cap [x, \infty)$ is a non-empty set of real numbers which is bounded from below by x. So $b_x = \inf(S_x \cap [x, \infty)) \in \mathbb{R}$. We have seen in (a) that S_x is a closed set and so $S_x \cap [x, \infty)$ is closed as well and contains its infimum, *i.e.*, $b_x \in S_x \cap [x, \infty)$.

□

Lemma 3.2.3(a) shows that the part of S_x which lies on the left hand side of x is an interval. So we have a very nice structure there. But unfortunately the part of S_x which lies on the right hand side of x could have a very complicated structure without any helpfull properties. We only know by Lemma 3.2.3(b) that this part of S_x is not empty and that b_x is its smallest element. To understand better what this means for our optimal stopping time τ_{S_x}, let us consider the example $x = 1$ and $S_x = (-\infty, 0.5] \cup [2.5, 3.5]$. We have then $a_x = 0.5$ and $b_x = 2.5$. Let us first consider the case that $(X_t)_t$ has continuous paths (for example Brownian Motion). If $(X_t)_t$ hits first $[2.5, 3.5]$ (Figure 3.1(a)) we have $\tau_{S_x} = \inf\{t \geq 0 : X_t = b_x\}$. If $(X_t)_t$ hits first $(-\infty, 0.5]$ (Figure 3.1(b)) we have $\tau_{S_x} = \inf\{t \geq 0 : X_t \leq a_x\}$. So for a continuous process we get the nice form for the stopping time

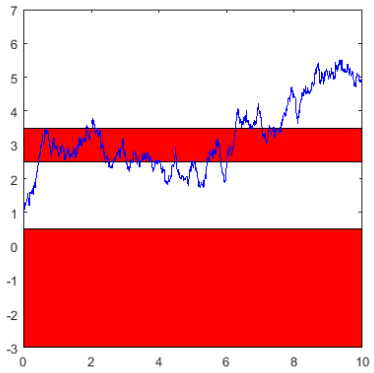
$$\tau_{S_x} = \inf\{t \geq 0 : X_t \leq a_x \text{ or } X_t = b_x\}.$$

But we want to consider Lévy processes which do not have continuous paths. So instead of just crossing continuously a_x or b_x it is possible for $(X_t)_t$ to jump into S_x without crossing a_x or b_x. In case $(X_t)_t$ jumps first below a_x (Figure 3.1(c)), we know that it hits S_x by Lemma 3.2.3(a). So in this case we also have as in the continuous case above that $\tau_{S_x} = \inf\{t \geq 0 : X_t \leq a_x\}$. But we do not know any such properties for $S_x \cap [b_x, \infty)$. So if the process makes a jump above b_x it could happen that it is not hitting S_x (Figure 3.1(d)) and we have $\tau_{S_x} \neq \inf\{t \geq 0 : X_t \geq b_x\}$. So together we have

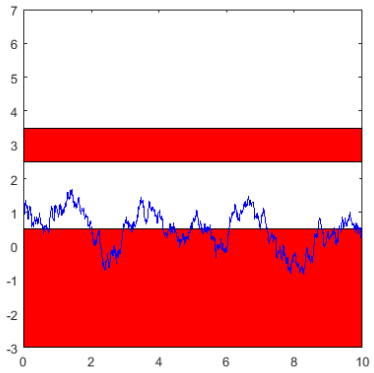
$$\tau_{S_x} \neq \inf\{t \geq 0 : X_t \leq a_x \text{ or } X_t \geq b_x\}.$$

So for general Lévy processes we do not have a nice form for τ_{S_x}. Because of this we will in the following only consider processes $(X_t)_t$ that do not have positive jumps, *i.e.*, we consider only spectrally negative Lévy processes. We will see in

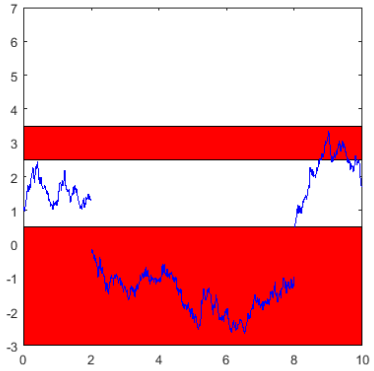
Figure 3.1: First hitting times. In (a) and (b) we see a continuous process hitting a_x and b_x. In (c) the process hits S_x by a jump down. In (d) the process jumps above b_x without hitting S_x.



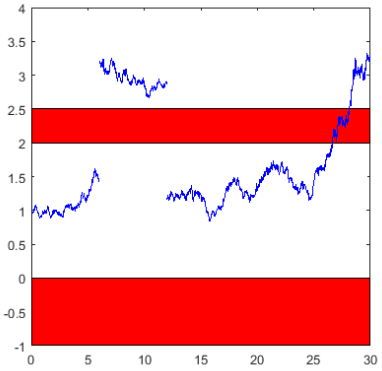
(a) Continuous; hitting b_x



(b) Continuous; hitting a_x



(c) Negative jump



(d) Positive jump

the next section that with this restriction we have as in the continuous case the nice form for our optimal stopping time

$$\tau_{S_x} = \inf\{t \geq 0 : X_t \leq a_x \text{ or } X_t = b_x\}. \tag{3.19}$$

3.3 Optimal Stopping for Spectrally negative Lévy Processes

Let $(X_t)_t$ be a real valued spectrally negative Lévy process defined on a probbility space $(\Omega, \mathcal{F}, \mathbb{P})$. Everything else is as in the last section.
Let us consider some starting point $x \in \mathbb{R}$ of the process and some $a \leq x \leq b$. We define the two-sided stopping time

$$\tau_{a,b} := \inf\{t \geq 0 : X_t \leq a \text{ or } X_t = b\}$$

and its corresponding value

$$v_{a,b}(x) := \mathbb{E}_x(e^{-r\tau_{a,b}} g_x(X_{\tau_{a,b}})).$$

For any $a \in \mathbb{R}$ we define also the one-sided stopping times

$$\tau_a(\omega) := \inf\{t \geq 0 : X_t(\omega) = a\},$$

$$\tau_a^+(\omega) := \inf\{t \geq 0 : X_t(\omega) \geq a\}$$

and

$$\tau_a^-(\omega) := \inf\{t \geq 0 : X_t(\omega) \leq a\}.$$

We need the following simple lemma.

Lemma 3.3.1. *Let $f : [0, \infty) \to \mathbb{R}$ be a right-continuous function and let A be a non-empty subset of $\mathbb{R}$ which is closed. Then*

$$\min\{t \geq 0 : f(t) \in A\}$$

exists.

Proof. Let us define

$$Z := \{t \geq 0 : f(t) \in A\}.$$

Because $Z \neq \emptyset$ is bounded from below its infimum exists. By definition of infimum exists a sequence $(t_n)_n \subseteq Z$ with $t_n \downarrow \inf Z$ for $n \to \infty$. So because f is right-continuous we have then

$$f(t_n) \xrightarrow{n\to\infty} f(\inf Z)$$

and because Z is closed we have $\inf Z \in Z$. So $\min Z$ exists. □

With this we can prove now the following lemma.

Lemma 3.3.2. *Let $(X_t)_t$ be a real valued spectrally negative Lévy process defined on a probbility space $(\Omega, \mathcal{F}, \mathbb{P})$. For $a_x = -\infty$ we have $\tau_{S_x} = \tau_{b_x}$ and for $a_x > -\infty$ we have $\tau_{S_x} = \tau_{a_x,b_x}$*

Proof. By Lemma 3.2.3 we know that

$$S_x = (-\infty, a_x] \cup \{b_x\} \cup B \tag{3.20}$$

for some $B \subseteq \mathbb{R}$ with $b \geq b_x$ for every $b \in B$.
Let us fix now an $\omega \in \Omega$. We define

$$C := \{t \geq 0 : X_t(\omega) \in S_x\}.$$

We know that

$$\tau_{S_x}(\omega) < \infty$$

So we have $C \neq \emptyset$. By Lemma 3.3.1 we have

$$\tau_{S_x}(\omega) = \inf C = \min C \in C$$

and so

$$X_{\tau_{S_x}}(\omega) = X_{\tau_{S_x}(\omega)}(\omega) \in S_x.$$

Now assume that $X_{\tau_{S_x}}(\omega) \in B$.
Then exists $b \in B$ with $X_{\tau_{S_x}}(\omega) = b$. Let us define now

$$D := \{t \geq 0 : X_t(\omega) \geq b_x\}.$$

Again by Lemma 3.3.1 exists $t_0 := \min D$. So we have

$$\lim_{t \uparrow t_0} X_t(\omega) \leq b_x. \tag{3.21}$$

Since $X_{\tau_{S_x}}(\omega) = b > b_x$, we have $\tau_{S_x}(\omega) \in D$ and so $t_0 \leq \tau_{S_x}(\omega)$. If $t_0 < \tau_{S_x}(\omega)$, we have $X_{t_0}(\omega) \notin S_x$ and for $t_0 = \tau_{S_x}(\omega)$ we have

$$X_{t_0}(\omega) = X_{\tau_{S_x}}(\omega) = b \neq b_x.$$

So in both cases we have $X_{t_0}(\omega) \neq b_x$. Because $t_0 = \min D \in D$, we have $X_{t_0}(\omega) > b_x$ and because $t \mapsto X_t(\omega)$ is right-continuous it follows

$$\lim_{t \downarrow t_0} X_t(\omega) = X_{t_0}(\omega) > b_x. \tag{3.22}$$

So (3.21) and (3.22) together gives

$$\lim_{t \uparrow t_0} X_t(\omega) \leq b_x < \lim_{t \downarrow t_0} X_t(\omega).$$

That means that $t \mapsto X_t(\omega)$ has a positive jump at t_0 and this is a contradiction to our assumptions. So $X_{\tau_{S_x}}(\omega) \notin B$ and it follows $X_{\tau_{S_x}}(\omega) \in (-\infty, a_x] \cup \{b_x\}$ and it follows

$$\begin{aligned}
\tau_{S_x}(\omega) &= \inf\{t \geq 0 : X_t(\omega) \in S_x\} \\
&= \inf\{t \geq 0 : X_t(\omega) \in (-\infty, a_x] \cup \{b_x\}\} \\
&= \begin{cases} \tau_{b_x}(\omega), & \text{for } a_x = -\infty \\ \tau_{a_x, b_x}(\omega), & \text{for } a_x > -\infty. \end{cases}
\end{aligned}$$

□

So we know now that the optimal stopping time is a one-sided or two-sided stopping time. In the the next theorem we prove that in case of a positive guarantee the optimal stopping time is a two-sided stopping time and is given by equation (3.19).

Theorem 3.3.1. *Let $(X_t)_t$ be a real valued spectrally negative Lévy process defined on a probbility space $(\Omega, \mathcal{F}, \mathbb{P})$. For $x \in \mathbb{R}$ and $k(x) > 0$ we have $a_x > -\infty$ and the optimal stopping time is $\tau_{S_x} = \tau_{a_x,b_x}$. Furthermore, the value function is then given by $v(x) = v_{a_x,b_x}(x)$.*

Proof. We only have to prove that $a_x > -\infty$. In this case, $\tau_{S_x} = \tau_{a_x,b_x}$ by Lemma 3.3.2, and we get for every $x \in \mathbb{R}$

$$\begin{aligned} v(x) = v_x(x) &\overset{\text{Lemma 3.2.2}}{=} \mathbb{E}_x\big(e^{-r\tau_{S_x}} g_x(X_{\tau_{S_x}})\big) \\ &= \mathbb{E}_x\big(e^{-r\tau_{a_x,b_x}} g_x(X_{\tau_{a_x,b_x}})\big) = v_{a_x,b_x}(x). \end{aligned}$$

We prove now $a_x > -\infty$ by contradiction. So let us assume that $a_x = -\infty$. Then by Theorem 3.3.2 we have

$$\tau_{S_x} = \tau_{b_x}. \tag{3.23}$$

Now let $e_r : \Omega \to \mathbb{R}$ be an exponential distributed random variable with parameter r which is "completely independent" of $(X_t)_t$ and its starting value, meaning that

$$\mathbb{P}_y(e_r > a) = e^{-ra} \quad \text{for every } y, a \in \mathbb{R} \tag{3.24}$$

and

$$\mathbb{P}_y(e_r \in A, \tau_0 \in B) = \mathbb{P}_y(e_r \in A)\mathbb{P}_y(\tau_0 \in B) \tag{3.25}$$

for every $y \in \mathbb{R}$ and $A, B \in \mathcal{B}(\mathbb{R})$.
Since our distributions depend on the starting value of $(X_t)_t$ we consider for every random variable $Y : \Omega \to \mathbb{R}$ the distribution $\mathbb{P}_y^Y$ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ defined by

$$\mathbb{P}_y^Y(A) := \mathbb{P}_y(Y \in A) \quad \text{for } A \in \mathcal{B}(\mathbb{R}).$$

We have for $z \in \mathbb{R}$

$$e^{-rz} \overset{(3.24)}{=} \mathbb{P}_{y-b_x}(e_r > z) = \mathbb{E}_{y-b_x}\big(1_{\{e_r > z\}}\big) = \int 1_{\{w>z\}} \mathbb{P}^{e_r}_{y-b_x}(dw) \tag{3.26}$$

We have then for $y \in \mathbb{R}$

$$\begin{aligned} \mathbb{E}_y\big(e^{-r\tau_{S_x}}\big) &\overset{(3.23)}{=} \mathbb{E}_y\big(e^{-r\tau_{b_x}}\big) \\ &= \mathbb{E}_{y-b_x}\big(e^{-r\tau_0}\big) \\ &= \int e^{-rz} \mathbb{P}^{\tau_0}_{y-b_x}(dz) \\ &\overset{(3.26)}{=} \int \Bigg(\int \underbrace{1_{\{w>z\}}}_{\geq 0} \mathbb{P}^{e_r}_{y-b_x}(dw) \Bigg) \mathbb{P}^{\tau_0}_{y-b_x}(dz) \\ &\overset{\text{Fubini}}{=} \int 1_{\{w>z\}} \mathbb{P}^{e_r}_{y-b_x} \times \mathbb{P}^{\tau_0}_{y-b_x}(d(w,z)) \\ &\overset{(3.25)}{=} \int 1_{\{w>z\}} \mathbb{P}^{(e_r,\tau_0)}_{y-b_x}(d(w,z)) \\ &= \mathbb{E}_{y-b_x}\big(1_{\{e_r > \tau_0\}}\big) \\ &= \mathbb{P}_{y-b_x}\big(e_r > \tau_0\big) \end{aligned} \tag{3.27}$$

We define now the maximum process of $(X_t)_t$ by

$$\bar{X}_t(\omega) := \sup\{X_s(\omega) : 0 \leq s \leq t\}.$$

We have for $\omega \in \Omega$

$$\begin{aligned}
\omega \in \{e_r > \tau_0\} &\Leftrightarrow e_r(\omega) > \inf\{t \geq 0 : X_t(\omega) = 0\} \\
&\Leftrightarrow \exists 0 \leq t_0 \leq e_r(\omega) : X_{t_0}(\omega) = 0 \\
&\Rightarrow \bar{X}_{e_r}(\omega) = \sup\{X_t(\omega) : 0 \leq t \leq e_r(\omega)\} \geq X_{t_0}(\omega) = 0 \\
&\Rightarrow \omega \in \{\bar{X}_{e_r} \geq 0\}
\end{aligned}$$

and it follows

$$\{e_r > \tau_0\} \subseteq \{\bar{X}_{e_r} \geq 0\} \tag{3.28}$$

So together we get then

$$\begin{aligned}
\mathbb{E}_y\big(e^{-r\tau_{S_x}}\big) &\overset{(3.27)}{=} \mathbb{P}_{y-b_x}(e_r > \tau_0) \\
&\overset{(3.28)}{\leq} \mathbb{P}_{y-b_x}(\bar{X}_{e_r} \geq 0) \\
&= \mathbb{P}_0(\bar{X}_{e_r} \geq b_x - y) \xrightarrow{y \to -\infty} \mathbb{P}_0(\emptyset) = 0.
\end{aligned} \tag{3.29}$$

Furthermore we have

$$\begin{aligned}
k(x) &\leq \mathbb{E}_y\big(e^{-r\cdot 0}(g(y) \vee k(x))\big) \qquad \text{(correlates to stopping at zero)} \\
&\leq \sup_{\tau \in \mathcal{T}} \mathbb{E}_y\big(e^{-r\tau}(g(X_\tau) \vee k(x))\big) \\
&= v_x(y) \\
&\overset{\text{Lemma 3.2.2}}{=} \mathbb{E}_y\big(e^{-r\tau_{S_x}}(g(\underbrace{X_{\tau_{S_x}}}_{\overset{(3.23)}{=} X_{\tau_{b_x}} = b_x}) \vee k(x))\big) \\
&= g_x(b_x)\mathbb{E}_y(e^{-r\tau_{S_x}}).
\end{aligned}$$

By (3.29) the right hand side converges to zero for $y \to -\infty$ and so we get our contradiction

$$0 < k(x) \leq 0.$$

□

So we have for every starting point $x \in \mathbb{R}$

$$v(x) = \sup_{a \leq x \leq b} v_{a,b}(x) \tag{3.30}$$

In the next two proposition we show that there is a nice expression of $v_{a,b}$ if we use the scale functions W and Z.

Proposition 3.3.1. *Let $(X_t)_t$ be a real-valued spectrally negative Levy process, $r > 0$, $a < x < b$ and W, Z be the scale functions for $(X_t)_t$. Then we have*

$$\mathbb{E}_x(e^{-r\tau_b^+} 1_{\{\tau_b^+ < \tau_a^-\}}) = \frac{W(x-a)}{W(b-a)} \tag{3.31}$$

and

$$\mathbb{E}_x(e^{-r\tau_a^-} 1_{\{\tau_a^- < \tau_b^+\}}) = Z(x-a) - W(x-a)\frac{Z(b-a)}{W(b-a)}. \tag{3.32}$$

Proof. [2, Theorem 8.1] □

Theorem 3.3.2. *Let $(X_t)_t$ be a real-valued spectrally negative Levy process, $r > 0$, $a < x < b$ and W, Z be the scale functions for $(X_t)_t$. For $k(x) \geq g(x)$ we have*

$$v_{a,b}(x) = k(x)Z(x-a) + \frac{g_x(b) - g_x(a)Z(b-a)}{W(b-a)}W(x-a).$$

Proof. We have for every $y \leq x$

$$k(x) \geq g(x) \geq g(y)$$

and so follows

$$g_x(y) = g(y) \vee k(x) \tag{3.33}$$

We calculate

$$\begin{aligned}
v_{a,b}(x) &= \mathbb{E}_x(e^{-r\tau_{a,b}} g_x(X_{\tau_{a,b}})) \\
&\overset{(3.33)}{=} \mathbb{E}_x(e^{-r\tau_a^-} g_x(a) 1_{\{\tau_a^- < \tau_b^+\}}) + \mathbb{E}_x(e^{-r\tau_b^+} g_x(b) 1_{\{\tau_b^+ < \tau_a^-\}}) \\
&= k(x)\mathbb{E}_x(e^{-r\tau_a^-} 1_{\{\tau_a^- < \tau_b^+\}}) + g_x(b)\mathbb{E}_x(e^{-r\tau_b^+} 1_{\{\tau_b^+ < \tau_a^-\}}) \\
&\overset{(3.31),(3.32)}{=} k(x)\left(Z(x-a) - W(x-a)\frac{Z(b-a)}{W(b-a)}\right) + g_x(b)\frac{W(x-a)}{W(b-a)} \\
&= k(x)Z(x-a) + \frac{g_x(b) - g_x(a)Z(b-a)}{W(b-a)}W(x-a).
\end{aligned}$$

□

3.4 The Optimal Stopping Boundaries as Solution to an ODE

We derive now a system of ordinary differential equations for which

$$f : \mathbb{R} \to \mathbb{R}^2, \quad x \mapsto (a_x, b_x)$$

is a solution. Let us define

$$F : \mathbb{R}^3 \to \mathbb{R}, \quad (x, a, b) \mapsto v_{a,b}(x).$$

Let $x_0 \in \mathbb{R}$ and suppose τ_{a_0,b_0} is an optimal stopping time if the process starts at x_0. So (a_0, b_0) is a maximum point of

$$(a, b) \mapsto F(x_0, a, b)$$

So

$$\frac{\partial F}{\partial a}(x_0, a_0, b_0) = \frac{\partial F}{\partial b}(x_0, a_0, b_0) = 0.$$

If we define

$$G(x, a, b) := \begin{bmatrix} \frac{\partial F}{\partial a}(x, a, b) \\ \frac{\partial F}{\partial b}(x, a, b) \end{bmatrix}$$

we have $G(x_0, a_0, b_0) = 0$. Now suppose that the matrix

$$\begin{bmatrix} \frac{\partial G_1}{\partial a}(x_0, a_0, b_0) & \frac{\partial G_1}{\partial b}(x_0, a_0, b_0) \\ \frac{\partial G_2}{\partial a}(x_0, a_0, b_0) & \frac{\partial G_2}{\partial b}(x_0, a_0, b_0) \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 F}{\partial a^2}(x_0, a_0, b_0) & \frac{\partial^2 F}{\partial a \partial b}(x_0, a_0, b_0) \\ \frac{\partial^2 F}{\partial a \partial b}(x_0, a_0, b_0) & \frac{\partial^2 F}{\partial b^2}(x_0, a_0, b_0) \end{bmatrix}$$

is invertible. Then, by the Implicit Function Theorem, there exists a neighbourhood V_0 of (a_0, b_0) and a unique continuously differentiable function $f : U_0 \to V_0$ with $f(x_0) = (a_0, b_0)$ and we have for all $x \in U_0, (a, b) \in V_0$

$$G(x, (a, b)) = 0 \Leftrightarrow f(x) = (a, b).$$

Total differentiation gives now

$$DG = \begin{bmatrix} \frac{\partial G_1}{\partial x} & \frac{\partial G_1}{\partial a} & \frac{\partial G_1}{\partial b} \\ \frac{\partial G_2}{\partial x} & \frac{\partial G_2}{\partial a} & \frac{\partial G_2}{\partial b} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 F}{\partial x \partial a} & \frac{\partial^2 F}{\partial a^2} & \frac{\partial^2 F}{\partial a \partial b} \\ \frac{\partial^2 F}{\partial x \partial b} & \frac{\partial^2 F}{\partial a \partial b} & \frac{\partial^2 F}{\partial b^2} \end{bmatrix}$$

and

$$D \begin{bmatrix} x \\ f(x) \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{\partial f_1}{\partial x} \\ \frac{\partial f_2}{\partial x} \end{bmatrix}.$$

By using this and the chain rule we get

$$\begin{aligned} D(G(x, f(x))) &= (DG)(x, f(x)) \circ (D(x, f(x))) \\ &= \begin{bmatrix} \frac{\partial^2 F}{\partial x \partial a} & \frac{\partial^2 F}{\partial a^2} & \frac{\partial^2 F}{\partial a \partial b} \\ \frac{\partial^2 F}{\partial x \partial b} & \frac{\partial^2 F}{\partial a \partial b} & \frac{\partial^2 F}{\partial b^2} \end{bmatrix} (x, f(x)) \circ \begin{bmatrix} 1 \\ \frac{\partial f_1}{\partial x} \\ \frac{\partial f_2}{\partial x} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial^2 F}{\partial x \partial a} \\ \frac{\partial^2 F}{\partial x \partial b} \end{bmatrix} (x, f(x)) + \begin{bmatrix} \frac{\partial^2 F}{\partial a^2} & \frac{\partial^2 F}{\partial a \partial b} \\ \frac{\partial^2 F}{\partial a \partial b} & \frac{\partial^2 F}{\partial b^2} \end{bmatrix} (x, f(x)) \circ \begin{bmatrix} \frac{\partial f_1}{\partial x}(x) \\ \frac{\partial f_2}{\partial x}(x) \end{bmatrix}. \end{aligned}$$

So at point $x = x_0$ we get

$$0 = \begin{bmatrix} \frac{\partial^2 F}{\partial x \partial a} \\ \frac{\partial^2 F}{\partial x \partial b} \end{bmatrix} (x_0, f(x_0)) + \begin{bmatrix} \frac{\partial^2 F}{\partial a^2} & \frac{\partial^2 F}{\partial a \partial b} \\ \frac{\partial^2 F}{\partial a \partial b} & \frac{\partial^2 F}{\partial b^2} \end{bmatrix} (x_0, f(x_0)) \circ \begin{bmatrix} \frac{\partial f_1}{\partial x}(x) \\ \frac{\partial f_2}{\partial x}(x) \end{bmatrix}.$$

and this is equivalent to the system of ordinary differential equations

$$- \begin{bmatrix} \frac{\partial^2 F}{\partial a^2} & \frac{\partial^2 F}{\partial a \partial b} \\ \frac{\partial^2 F}{\partial a \partial b} & \frac{\partial^2 F}{\partial b^2} \end{bmatrix}^{-1} (x_0, f(x_0)) \begin{bmatrix} \frac{\partial^2 F}{\partial x \partial a} \\ \frac{\partial^2 F}{\partial x \partial b} \end{bmatrix} (x_0, f(x_0)) = \begin{bmatrix} \frac{\partial f_1}{\partial x}(x) \\ \frac{\partial f_2}{\partial x}(x) \end{bmatrix}.$$

This system of differential equations is a nice result and could be used for computations; for example for improved performance of the option price calculation.

Chapter 4

Results for Analytic Inverse Laplace Transform

In this chapter we do some numerical investigation of our theory on the example of Brownian motion with drift. There are two motivations for this chapter. First, we check if we get correct results for the analytic solution of the scale function. We need this to check if the approximations of the scale function in the next chapter give correct results. This is done in Section 2 and 4. Second, even though perpetual American call options modelled by geometric Brownian motion are well studied, this is not the case for perpetual American call options with guarantee. So this is also interesting in geometric Brownian motion case. We do a little investigation of this in Section 3.
We consider the same setting as in chapter 3, only that we assume now that the process $(X_t)_t$, which is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, is a scaled Brownian motion with drift and that we have the payoff of a call option. So by using the same notation as in the last chapter the stock price process $(X_t)_t$ is given by

$$X_t = \sigma B_t + \gamma t$$

where $(B_t)_t$ is a Brownian motion and σ, γ are real constants with $\sigma > 0$. Since we are considering a call option the payoff function has the form

$$g(x) = \left(e^x - K\right)^+$$

for some constant (the strike price) $K \geq 0$. We are having again a risk-free investment possibility with constant interest rate $r > 0$. Now we have to say a few words about the so called risk-neutral measure $\mathbb{Q}$, which is a very important tool in financial mathematics. We assume the measure $\mathbb{P}$ above to be the real world measure. That means that for every $t \geq 0$, the probability distribution of X_t under $\mathbb{P}$ is the one we would estimate by real world data. The probability measure $\mathbb{Q}$ on the other hand is an artificial probability measure on $(\Omega, \mathcal{F})$. The reason we are interested in it is that under the risk-neutral measure the process $(X_t)_t$ has the form

$$X_t = \sigma B_t + (r - \frac{1}{2}\sigma^2)t.$$

where $(B_t)_t$ is a new Brownian motion (with respect to $\mathbb{Q}$) and that the value of the option is given by

$$\sup_{\tau \in \mathcal{T}} \mathbb{E}_x^{\mathbb{Q}} \Big(e^{-r\tau} \big(g(X_\tau) \vee k(x) \big) \Big) \tag{4.1}$$

where $\mathbb{Q}$ at the expectation means that we have the expectation with respect to $\mathbb{Q}$. See for example [9, Chapter 5.2]. We see that equation (4.1) is exactly the same as equation (3.9) of chapter 3. We could say that we were in chapter 3 actually the whole time working under the risk-neutral measure. Having that said, we will continue now working under the risk neutral measure, i.e. we assume our process to be given by $X_t = \sigma B_t + (r - \frac{1}{2}\sigma^2)t$. Furthermore, we will nevertheless just write write $\mathbb{P}$ and $\mathbb{E}$ to be consistent with the notation of the previous chapters.
We know by Theorem 3.3.2 and equation (3.3.2) that the value of the option is given by

$$v(x) = \sup_{a<x<b} v_{a,b}(x) \tag{4.2}$$

where $v_{a,b}(x)$ is given by

$$v_{a,b}(x) = k(x)Z(x-a) + \frac{g_x(b) - g_x(a)Z(b-a)}{W(b-a)} W(x-a).$$

Furthermore, we have by Lemma 2.3.2 an analytic solutions for the scale functions W and Z of $(X_t)_t$. (The case that no analytic solution for the scale functions exists will be discussed in the next chapter). So to calculate the value of the option the only approximation we have to do is the approximation of the supremum. We do this by the Matlab function fmincon. To make more clear that we only have approximations of $v_{a,b}(x)$ we will also use the notation $F(x,a,b)$. So $F(x,a,b)$ means the computational approximation of $v_{a,b}(x)$.
In the first section we check if condition (3.10) of Section 3.1 is satisfied. In the second section we make our first computations of the value of the option. The value of the option is well known for the case of no guarantee. So we test our calculations by letting the guarantee go to zero. We do this to test if our code is working properly. In the third section we have a closer look on the optimal boundaries for the optimal stopping time. For this we produce a 3D plot with Mathematica and give an interpretation. This is interesting because in practice not only the value of the option is important but also the optimal time to exercise it. This is especially interesting since perpetual American call options with guarantee are not well researched. As we have mentioned, we approximate the supremum in our calculations by the Matlab function fmincon. This functions needs a starting value from where it starts searching for a maximum. In Section 4 we take a closer look on this and check if a wrong choice of the starting value can lead to wrong results.
The two proofs of Lemma 4.1.1 and 4.1.2 are original by the author of this thesis. Nevertheless, Lemma 4.1.1 is a standard result and there should exist a lot of literature with proofs of it. For figure 4.2 Mathematica has been used. All other plots and calculations are done by Matlab.

4.1 Conditions

In this section we check when condition (3.10) of Section 3.1 is fulfilled. To do so we first need the following lemma that allows us to exchange the supremum and the exponential function. Since it can be difficult to find sometimes references for simple results as the following we just present our own proof.

Lemma 4.1.1. *Let $f : R \to R$ be a continuous, monotone increasing function and let $\{a_t : t \geq 0\}$ be a collection of real numbers. Then we have*

$$\sup\{f(a_t) : t \geq 0\} = f(\sup\{a_t : t \geq 0\}). \tag{4.3}$$

Proof. By definition of the supremum $\sup\{a_t : t \geq 0\} - \frac{1}{n}$ is not an upper bound of $\{a_t : t \geq 0\}$ for every $n \in \mathbb{N}$. So for every $n \in \mathbb{N}$ there exists a $t_n \geq 0$ such that

$$\sup\{a_t : t \geq 0\} - \frac{1}{n} \leq a_{t_n} \leq \sup\{a_t : t \geq 0\}.$$

So $a_{t_n} \xrightarrow{n\to\infty} \sup\{a_t : t \geq 0\}$ and we get

$$f(\sup\{a_t : t \geq 0\}) = f(\lim_{n\to\infty} a_{t_n}) \overset{f \text{ cont.}}{=} \lim_{n\to\infty} f(a_{t_n}).$$

So we only have to prove that

$$\lim_{n\to\infty} f(a_{t_n}) = \sup\{f(a_t) : t \geq 0\}.$$

First we see that for every $n \in \mathbb{N}$ we have $f(a_{t_n}) \leq \sup\{f(a_t) : t \geq 0\}$ and it follows $\lim_{n\to\infty} f(a_{t_n}) \leq \sup\{f(a_t) : t \geq 0\}$.
We show the other inequality by contradiction. So let us suppose that

$$\lim_{n\to\infty} f(a_{t_n}) < \sup\{f(a_t) : t \geq 0\}.$$

Then $\lim_{n\to\infty} f(a_{t_n})$ is not an upper bound of $\{f(a_t) : t \geq 0\}$ and so there exists a $t_0 \geq 0$ with $\lim_{n\to\infty} f(a_{t_n}) < f(a_{t_0})$. Consequently there exists a constant $c > 0$ and some $N_1 \in \mathbb{N}$ such that for every $n \geq N_1$

$$f(a_{t_n}) + c < f(a_{t_0}). \tag{4.4}$$

Since $a_{t_n} \xrightarrow{n\to\infty} \sup\{a_t : t \geq 0\}$ and f is continuous there exists an $N_2 \in \mathbb{N}$ such that for every $n \geq N_2$

$$f(a_{t_n}) + c \geq f(\sup\{a_t : t \geq 0\}). \tag{4.5}$$

So for $N := \max\{N_1, N_2\}$ we have

$$f(a_{t_0}) \overset{f \text{ mon.}}{\leq} f(\sup\{a_t : t \geq 0\}) \overset{(4.5)}{\leq} f(a_{t_N}) + c \overset{(4.4)}{<} f(a_{t_0})$$

and this is a contradiction.

□

We can prove now the following lemma which is original by the author of this thesis.

Lemma 4.1.2. *Let $(X_t)_t$ be a scaled Brownian motion with drift defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, i.e.*

$$X_t = \sigma B_t + (r - \frac{1}{2}\sigma^2)t$$

with the constants $\sigma > 0$ and $r > 0$. Furthermore, we define the function

$$g : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \left(e^x - K\right)^+$$

with $K \geq 0$. Then condition (3.10)

$$\mathbb{E}_x(\sup_{t\geq 0} e^{-rt}|g(X_t)|) < \infty$$

is satisfied if $\sigma > 1$.

Proof. Let us define the random variable

$$Y(\omega) := \sup\{\sigma B_t(\omega) - \frac{1}{2}\sigma^2 t : t \geq 0\}.$$

Then we have by [4, Proposition 6.8.1] for $x \geq 0$

$$\mathbb{P}(Y > x) = \mathbb{P}(\frac{Y}{\sigma} > \frac{x}{\sigma}) \overset{6.8.1}{=} e^{-\sigma x}.$$

Since

$$Y \geq X_0 = 0 \quad \mathbb{P} - \text{a.s.}$$

we have

$$\mathbb{P}(Y > x) = \begin{cases} e^{-\sigma x}, & \text{if } x \geq 0 \\ 1, & \text{if } x < 0 \end{cases} \tag{4.6}$$

So Y is exponential distributed with parameter σ. By fixing some $x \in \mathbb{R}$ and considering the monotone increasing function $f(y) := e^{y+\sigma x}$ we can calculate now

$$\begin{aligned}
\mathbb{E}_x\left(\sup_{t\geq 0} e^{-rt}|g(X_t)|\right) &= \mathbb{E}_x\left(\sup_{t\geq 0} e^{-rt}\left(e^{\sigma B_t + (r-\frac{1}{2}\sigma^2)t} - K\right)^+\right) \\
&\leq \mathbb{E}\left(\sup_{t\geq 0} e^{\sigma B_t - \frac{1}{2}\sigma^2 t + \sigma x}\right) \\
&= \mathbb{E}\left(\sup_{t\geq 0} f\left(\sigma B_t - \frac{1}{2}\sigma^2 t\right)\right) \\
&\overset{\text{Lemma } 4.1.1}{=} \mathbb{E}\left(f\left(\sup_{t\geq 0}\left(\sigma B_t - \frac{1}{2}\sigma^2 t\right)\right)\right) \\
&= \mathbb{E}(f(Y)) \\
&\overset{(4.6)}{=} \sigma \int_0^\infty f(y) e^{-\sigma y} dy \\
&= \sigma \int_0^\infty e^{y+\sigma x} e^{-\sigma y} dy \\
&= \sigma e^{\sigma x} \int_0^\infty e^{(1-\sigma)y} dy
\end{aligned}$$

and we see that the last term is finite if $\sigma > 1$. □

Unfortunately, in reality we have quite often $\sigma < 1$. We will nevertheless apply our theory (also for the case $\sigma < 1$).

4.2 A Convergence Test

For the case that our option does not have a guarantee, the value of the option is e^x. This is proved for example in [5, Proposition 11.11]. So we have

$$\sup_{\tau \in \mathcal{T}} \mathbb{E}_x\Big(\big(e^{-r\tau}(e^{X_\tau} - K)^+\big)\Big) = e^x.$$

Now let $\{k_n : n \in \mathbb{N}\}$ be a sequence of guarantee functions that converges pointwise to zero. By equation (4.2) and Lemma 1.4.2 we have

$$V_n(x) := \sup_{\tau \in \mathcal{T}} \mathbb{E}_x\Big(e^{-r\tau}\big((e^{X_\tau} - K)^+ \vee k_n(x)\big)\Big) \xrightarrow{n\to\infty} e^x.$$

To test our theory numerically we consider the guarantee functions

$$k_n(y) := \frac{g(y)}{n} = \frac{(e^y - K)^+}{n}$$

and fix as interest rate $r = 0.01$, as volatility of the stock $\sigma = 0.2$ and as strike price $K = 10$. We test the theory now for the four cases $x = \log(20)$, $x = \log(30)$, $x = \log(40)$ and $x = \log(50)$. The result can be seen in Figure 4.1. On the x-axis we see n going from 10^2 to 10^5 with a step size of 30 and on the y-axis we see the respective values of $V_n(x)$ and as predicted by our theory the values converge to $20, 30, 40$ and 50 respectively. So this is a very good sign that our implementation of the theory is working. We also see the little spikes in the plots. We keep in mind that Matlab has to find for each n-value the x-axis the supremum $\sup_{a<x<b} v_{a,b}(x)$. So we conclude that Matlab function fmincon does not find the actual maximum at these points. We see that all spikes are pointed downwards. If we would be searching for minima the spikes would show in the other direction at the points where the correct maximum would not be found. So we see that we have to be careful with the approximations of fmincon. We will have a closer look on this in Section 4. Except for these values, we have the convergence we were expecting. So our implementation seems to work.

4.3 The Stopping Boundaries

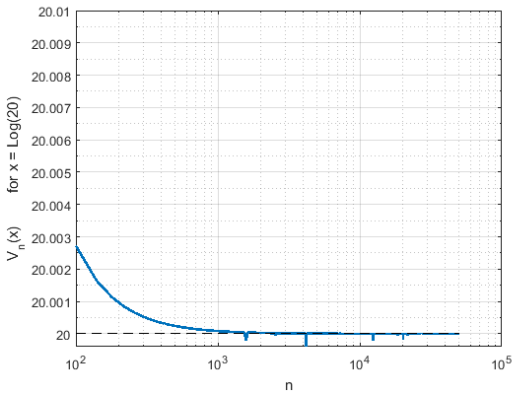
So far we were only talking about the value of the option. But when one possesses an American option the question arises when to exercise it. We know by Theorem 3.3.1 that the optimal stopping time is of two-sided form, *i.e.* it is of the form

$$\tau_{a,b} = \inf\{t \geq 0 : X_t \leq a \text{ or } X_t = b\}.$$

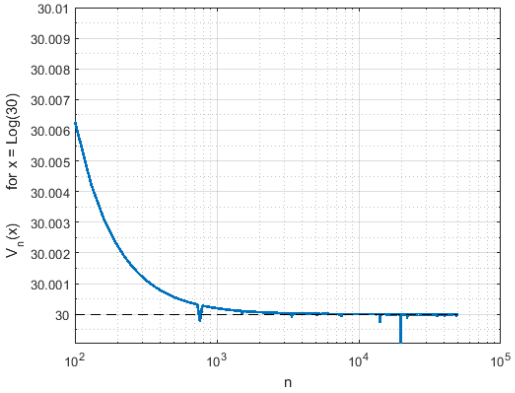
and so the value of the option is given by the supremum over all possible two-sided stopping times, *i.e.*

$$v(x) = \sup_{a<x<b} \mathbb{E}_x\Big(e^{-r\tau_{a,b}}\mathbb{E}_x(g(X_{\tau_{a,b}}) \vee k(x))\Big). \tag{4.7}$$

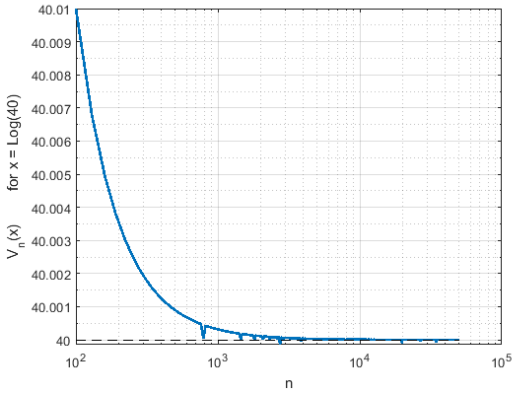
Figure 4.1: Guarantee converging to zero. In all four plots we have on the x-axis incresing natural numbers n and on the y-axis we see the corresponding values $V_n(x)$ with respect to the guarantee $k_n(x)$. In (a) we have the case $x = \log(20)$, in (b) $x = \log(30)$, in (c) $x = \log(40)$ and in (d) $x = \log(50)$.



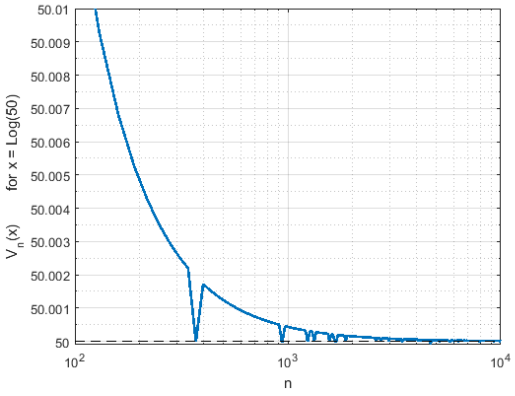
(a)



(b)



(c)



(d)

Consequently, the exercise strategy consist of two constants $a < b$ such that the option gets exercised if the stock value goes below a or above b. For every such strategy we have the expected outcome

$$v_{a,b}(x) = \mathbb{E}_x\Big(e^{-r\tau}\mathbb{E}_x(g(X_\tau) \vee k(x))\Big) \tag{4.8}$$

and the optimal choices for a and b are the ones who maximise $v_{a,b}(x)$. For a normal perpetual call option with Brownian motion it is well known that it is optimal never to exercise. See for example [9, Section 8.5.1]. So the boundaries are then $a = -\infty$ and $b = \infty$. Our case with guarantee is not really researched. So it is interesting to see what the optimal boundaries in this case are. We use Mathematica to produce a 3D plot of $v_{a,b}(x)$. We fix as interest rate $r = 0.01$, as volatility $\sigma = 0.15$, as strike price $K = 10$ and as starting value of the process $x = \log(50)$. The result is Figure 4.2. This is a 3D plot with the two variables a, b on the horizontal x-axis and y-axis. On the perpendicular z-axis we see the outcome $v_{a,b}(x)$. All three pictures (a), (b) and (c) show the same plot just from three different angles. Let us first consider the upper bound b. We see that the expected outcome $v_{a,b}(x)$ is increasing in b, *i.e.* the higher the chosen upper boundary is, the higher the expected outcome becomes. We know by Lemma 3.2.3 that the optimal boundary b is finite. But in Figure 4.2 it appears to be quite large. The explanation is the following. Let us say our stock has a high value at some time t_0. Then it is very unlikely that it will drop to a value where the guarantee is needed. So our option is in that case very similar to a perpetual American call option without guarantee. As we have explained above it is optimal to never exercise a perpetual American call option. That explains that the optimal upper boundary b is very large. Let us now consider the lower bound a. We see that $v_{a,b}(x)$ has a clear maximum in a around 2.9 and the values decrease on the left and on the right of it. So there is a clear lower bound and when it is crossed by the stock the guarantee should be exercised. We can explain this by the following. If we exercise the option at some time t we receive the payoff $(e^{X_t} - K) \vee k(x)$. Now we have in our model positive interest rate. So a specific amount of payoff becomes less valuable with time. This is reflected in the e^{-rt}-term in equation (3.9) and (4.1). So the real value we are receiving at time t is

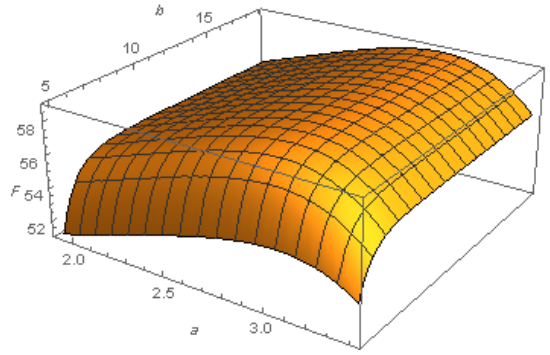
$$e^{-rt}\big((e^{X_t} - K) \vee k(x)\big)$$

Let us consider the case that the value of the stock is close to zero at a time $t_0 > 0$. Since we assume the increase of $(X_t)_t$ normal distributed with the variance dependent on time, it is likely that it will take a long time until $X_t - K$ becomes larger than the guarantee. So the real value $e^{-rt}(e^{X_t} - K)$ we get then by exercising will likely be smaller than $e^{-rt_0}k(x)$. So in this scenario the option should be exercised at time t_0 for the guaranteed payoff $k(x)$. This explains why there is a clear optimal lower bound a for which the guarantee should be exercised once the stock price crosses it.

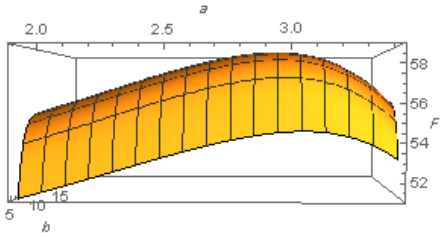
4.4 Starting Values for Fmincon

We are working in Matlab with the function fmincon. For fmincon we have to choose a starting value (a_0, b_0) from which on fmincon starts searching for a couple (a_{max}, b_{max}) such that $v_{a_{max},b_{max}}(x)$ is a maximum of the function

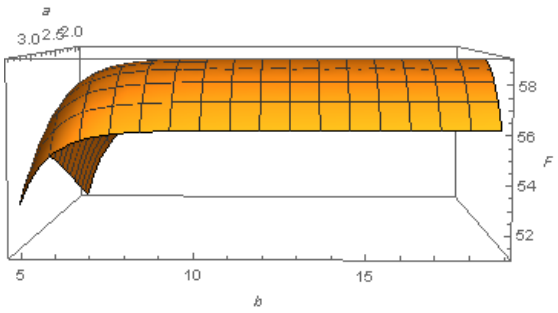
Figure 4.2: The optimal stopping boundaries a and b. We see the same plot from three different angles. On the x-axis we see a, on the y-axis we have b and on the z-axis we see $v_{a,b}(x)$.



(a)



(b)



(c)

$(a, b) \mapsto v_{a,b}(x)$ for some fixed x. (Actually fmincon searches for minima: so we multiply the function by -1 in the Matlab code). We have seen by the spikes in Figure 4.1 that fmincon does not always find the right maximum. Since the right outcome of fmincon depends on the chosen a_0 and b_0 we are interested how stable fmincon works in regards to these starting values. To see this we fix the parameters $K = 10$, $n = 2$, $r = 0.01$, $\sigma = 0.2$ and $x = log(60)$. The guarantee is as in the last section given by

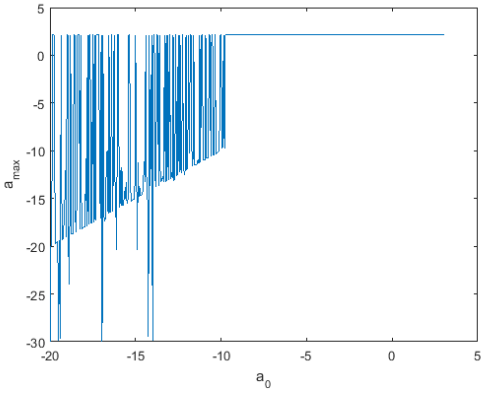
$$k_n(y) := \frac{g(y)}{n} = \frac{(e^y - K)^+}{n}.$$

In Figure 4.3 we fix $b_0 = 27$ as starting value for fmincon. We have as variable in all three plots on the x-axis a_0. In plot (a) we see on the y-axis the optimal a_{max} values fmincon finds. We see a strong fluctuation for starting values a_0 between -20 and -10. Between -10 and 5 the result are completely constant at the value $a_{max} = 2.1203$. We keep in mind that we are considering here always the logarithm of the stock value. So -10 stands for a very small value of the stock. That means that we are actually talking about the values between e^{-20} and e^{-10}. So for very small starting values a_0 we can get bad results. But for reasonable starting values fmincon finds the correct result $a_{max} = 2.1203$. This is also consistent with our results of Figure 4.2. There we saw that there is a clear maximum of the lower bound a_{max} which fmincon can find. In plot (c) we see on the y-axis the values $F(x, a_{max}, b_{max})$. We see as in plot (b) that the corresponding values for $F(x, a_{max}, b_{max})$ are also very unstable and wrong between -20 and -10. But if we choose a_0 between -10 and 5 we get the correct result. In plot (b) we see on the y-axis the optimal b_{max} values fmincon finds. We see as in in (a) and (c) a strong fluctuation between -20 and -10. But now the result are not becoming stable between -10 and 5. To understand this we have to look back to Figure 4.2. We saw there that there is not a clear maximal boundary b_{max}. For the optimal values it only seems to be important that the upper bound is high. This is what we are seeing here. The optimal values for b_{max} which fmincon is finding between -10 and 5 lies between 20 and 30. So as long we are working with a reasonable choice for a_0 the algorithm works.

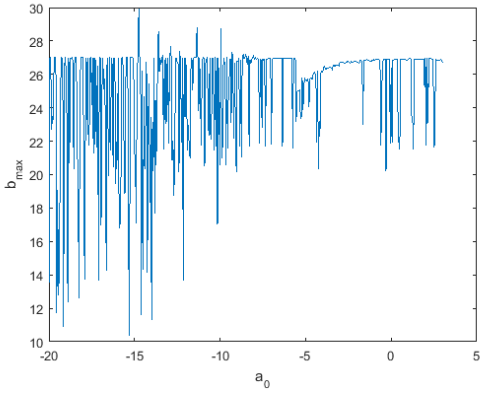
Now we plot the same three plots just this time with b_0 as variable and a_0 being fixed. We fix $a_0 = 2$. This is very close to to the actual maximum $a_{max} = 2.1203$ which we found in Figure 4.3. All other parameters are the same as in Figure 4.3. The result we can see in Figure 4.4. On the x-axis we have in all three plots the variable b_0.

In plot (a) we see on the y-axis the optimal a_{max} values fmincon finds. We see a very small fluctuation around the correct value $a_{max} = 2.1203$. So we get good results along the whole interval between between 5 and 30. In plot (c) we see on the y-axis the values $F(x, a_{max}, b_{max})$. We see as in plot (b) that the corresponding values for $F(x, a_{max}, b_{max})$ are also all very close to the correct result. In plot (b) we see on the y-axis the optimal b_{max} values fmincon finds. Different to (a) and (c) we have here a strong fluctuation between $b_{max} = 20$ and $b_{max} = 30$. So again we see that the values for b_{max} only have to be big but that the exact value does not influence the result.
In conclusion we can say that we have to be a little careful with the choice of a_0 and for b_0 we should choose a large number between 10 and 30.

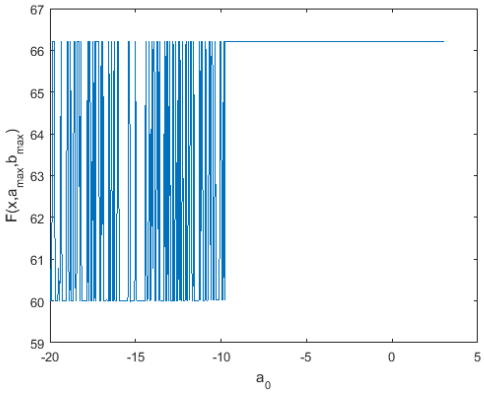
Figure 4.3: Fmincon for fixed $b_0 = 27$. In all three plots we have as variable on the x-axis the starting value a_0 for fmincon. In (a) we see on the y-axis which a_{max} fmincon finds. In (b) we see on the y-axis which b_{max} fmincon finds. In (c) we see on the y-axis the correscponding values $F(x, a_{max}, b_{max})$.



(a)

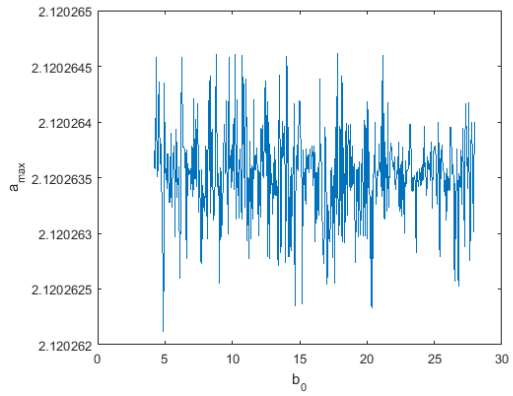


(b)

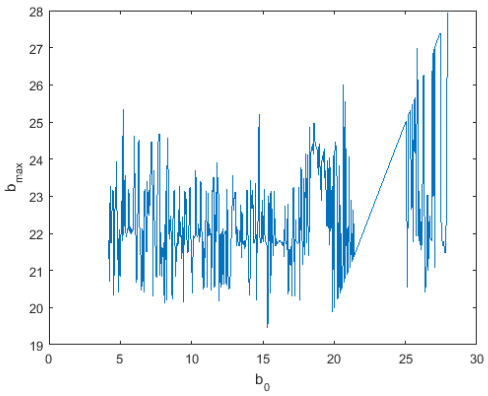


(c)

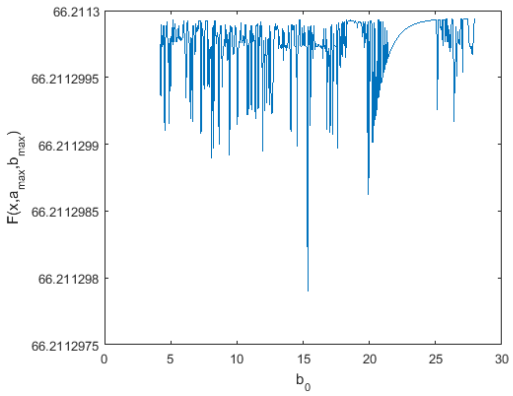
Figure 4.4: Fmincon for fixed $a_0 = 2$. In all three plots we have as variable on the x-axis the starting value b_0 for fmincon. In (a) we see on the y-axis which a_{max} fmincon finds. In (b) we see on the y-axis which b_{max} fmincon finds. In (c) we see on the y-axis the correscponding values $F(x, a_{max}, b_{max})$.



(a)



(b)



(c)

4.5 Approximated Inverse Laplace Transform

In the last section we used the analytic solution for the scale function which we found by Lemma 2.3.1. For more general processes there does not have to exist an analytic solution. For most processes it is easy to find their Laplace exponent ψ. The difficulty lies in finding the inverse Laplace transform of $\frac{1}{\psi - r}$. In this section we want to do a little investigation if this can be done in principle. For this we try to approximate the inverse Laplace transform of $\frac{1}{\psi - r}$ where ψ is the Laplace exponent of the scaled Brownian motion with drift. We compare then the results with the analytic solution of the scale function W to measure the error.

We first try the Euler algorithm. It is based on the paper [11] which uses the Gaver-Stehfest algorithm, a version of the Fourier-series method with Euler summation, and a version of the Talbot algorithm, which is based on deforming the contour in the Bromwich inversion integral. The used Matlab code can be found at [12]. We plot the error of W when we approximate it by the Euler algorithm, *i.e.*, we compare it with the analytic values we know. The result we see in Figure 4.5. In (a) we have $r = 0.01, \sigma = 0.2$, in (b) we have $r = 0.05, \sigma = 0.1$, in (c) we have $r = 0.1, \sigma = 0.2$ and in (d) we have $r = 0.1, \sigma = 0.6$. We can see two things: If we have a close look we see that all the graphs are not starting at zero but some value smaller than 1. The reason for that is that the Euler algorithm produces errors for these x values. The second thing we spot is that the error stays in all three plots reasonably small until around $x = 18$. But this is actually good enough since x is the logarithm of the actual value of the stock and stock values of e^{18} do not occur in reality. Furthermore, we have seen in Section 4.3 that an upper bound of 18 is more than enough.

Now we try out an algorithm invented by Hollenbeck, K. J. which can be found in [13]. We make four plots with exactly the same constant for r and σ as in Figure 4.5. The result we see in Figure 4.6. We see that the error is rapidly increasing for increasing x, but that we are getting good results for x close to zero. Unfortunately, for x smaller 0.15 we also get errors for the Hollenbeck algorithm. But this is a step forward. We can combine now the Hollenbeck and Euler algorithm. From 0.15 to 1 we calculate W by the Hollenbeck algorithm and from 1 till 14 we use the Euler algorithm.

In Figure 4.7 we have a look at the shape of W. In red colour we see the exact graph of W produced by the analytic formula. In turquoise colour we see the approximation of W by Hollenbeck algorithm. Under close inspection, the graph of W looks smooth. So it might be an idea to get the last missing part of W by spline interpolation. In Figure 4.8 we have again in red colour the exact graph of W produced by the analytic formula. In turquoise colour we see this time the approximation of W by the Hollenbeck algorithm with a spline interpolation on the interval [0,0.15]. We see that this works very good and that the approximation fits very nicely to the exact graph. This is only a first investigation and additional work and adjustment is required to test it on jump processes. But it is a good sign that the approximation works for Brownian motion. So this section can be seen as an outlook for future work. So because we get the value of the option by the formula

$$v_{a,b}(x) = k(x)Z(x-a) + \frac{g_x(b) - g_x(a)Z(b-a)}{W(b-a)}W(x-a),$$

we have found a method to approximate the option value for processes which can contain jumps.

In Figure 4.7 we have a look at the shape of W. In red colour we see the exact graph of W produced by the analytic formula. In turquoise colour we see the approximation of W by Hollenbeck algorithm. Under close inspection, the graph of W looks smooth.

Figure 4.5: Error of W approximated by Euler algorithm. In all four plots we have on the x-axis the real numbers x in the interval $[0, 20]$ and on the y-axis the error of the approximation of the scale function $W(x)$. In (a) we have the case $r = 0.01, \sigma = 0.2$, in (b) we have $r = 0.05, \sigma = 0.1$, in (c) we have $r = 0.1, \sigma = 0.2$ and in (d) we have $r = 0.1, \sigma = 0.6$.

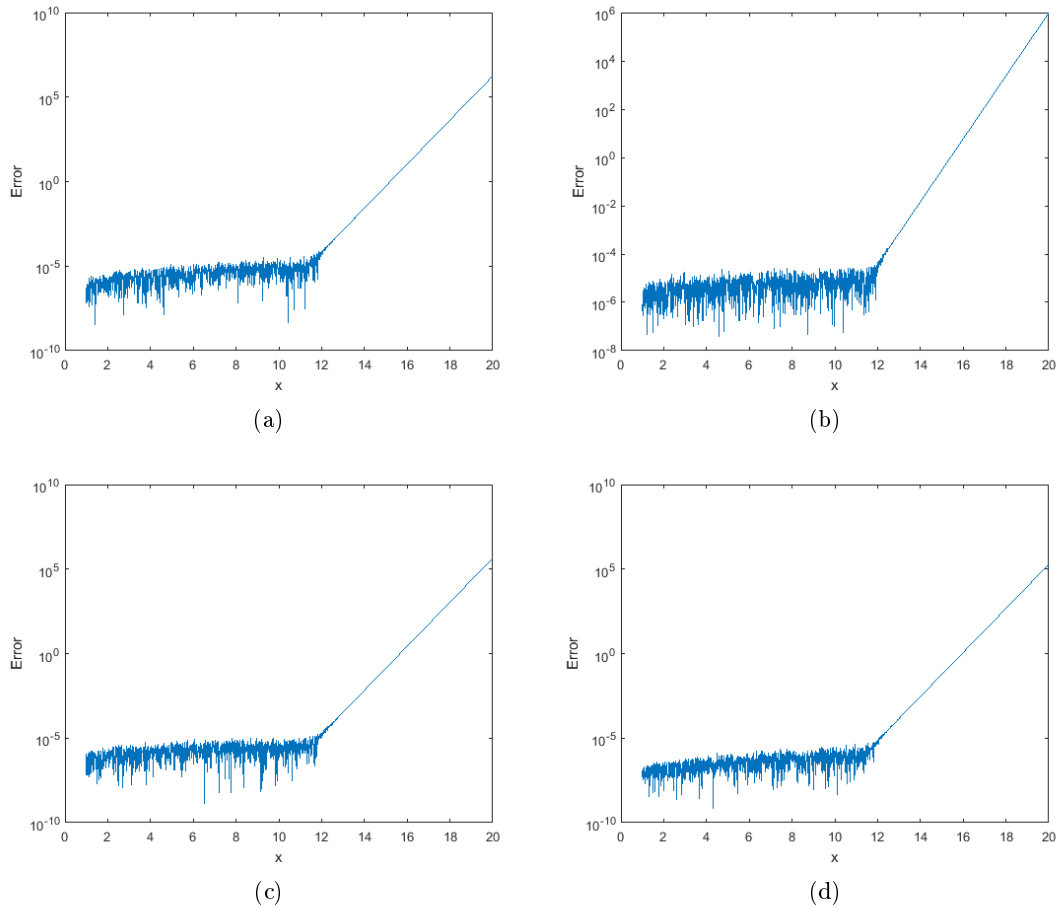
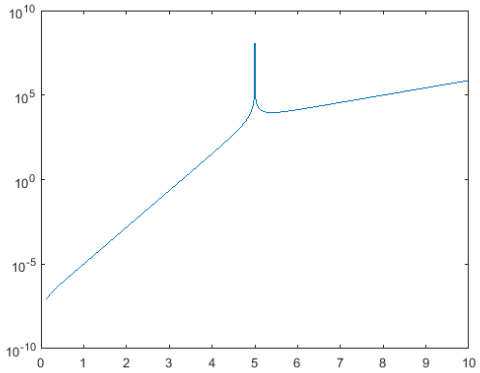
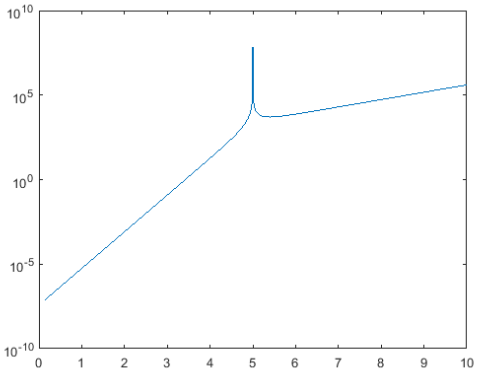


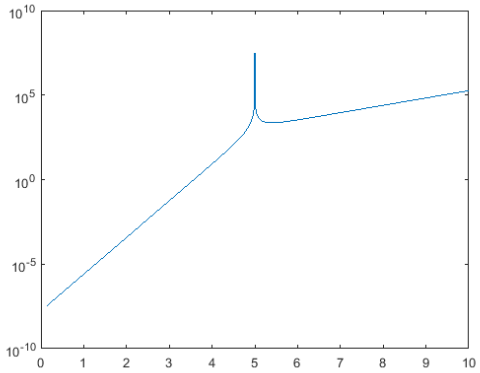
Figure 4.6: Error of W approximated by Euler and Hollenbeck algorithm. In all four plots we have on the x-axis the real numbers x in the interval $[0, 10]$ and on the y-axis the error of the approximation of the scale function $W(x)$. In (a) we have the case $r = 0.01, \sigma = 0.2$, in (b) we have $r = 0.05, \sigma = 0.1$, in (c) we have $r = 0.1, \sigma = 0.2$ and in (d) we have $r = 0.1, \sigma = 0.6$.



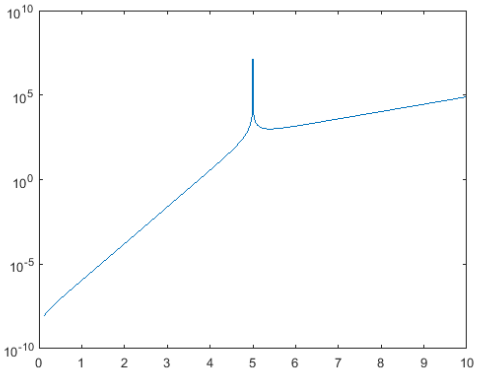
(a)



(b)

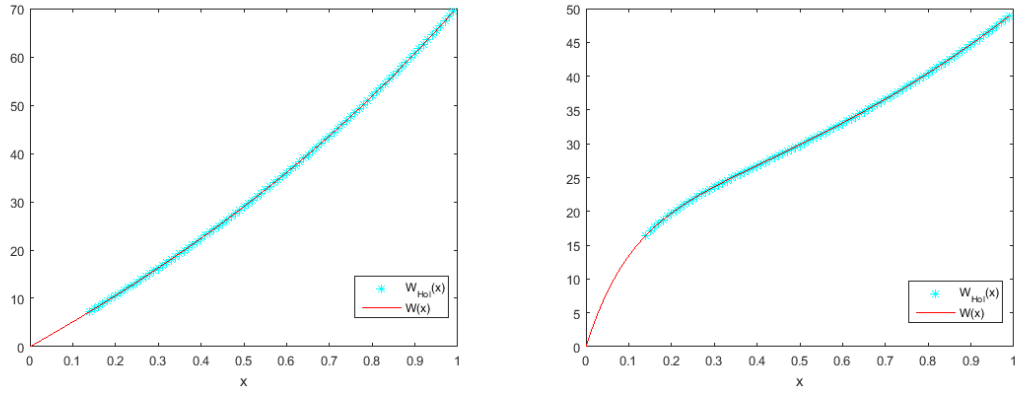


(c)



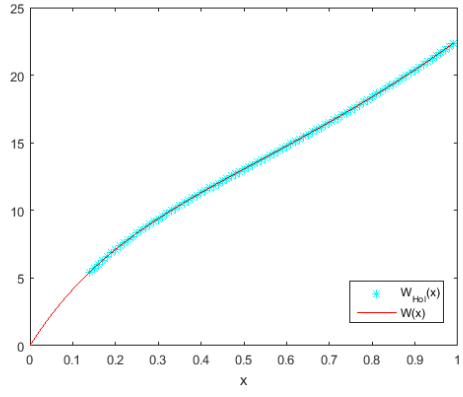
(d)

Figure 4.7: W approximated by Hollenbeck algorithm. In all four plots we have on the x-axis the real numbers x in the interval $[0, 1]$ and on the y-axis we see in red color $W(x)$ computed by the the analytic formula and in turquois color the the values $W(x)$ approximated by the Hollenbeck algorithm. In (a) we have the case $r = 0.01, \sigma = 0.2$, in (b) we have $r = 0.05, \sigma = 0.1$, in (c) we have $r = 0.1, \sigma = 0.2$ and in (d) we have $r = 0.1, \sigma = 0.6$.

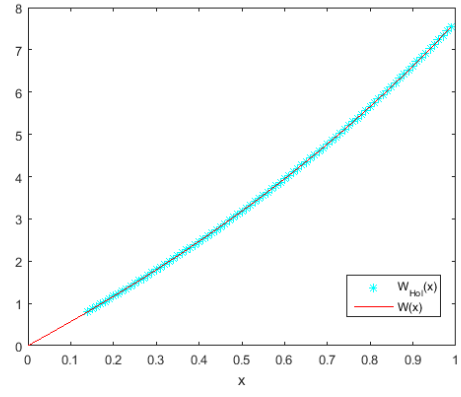


(a)

(b)

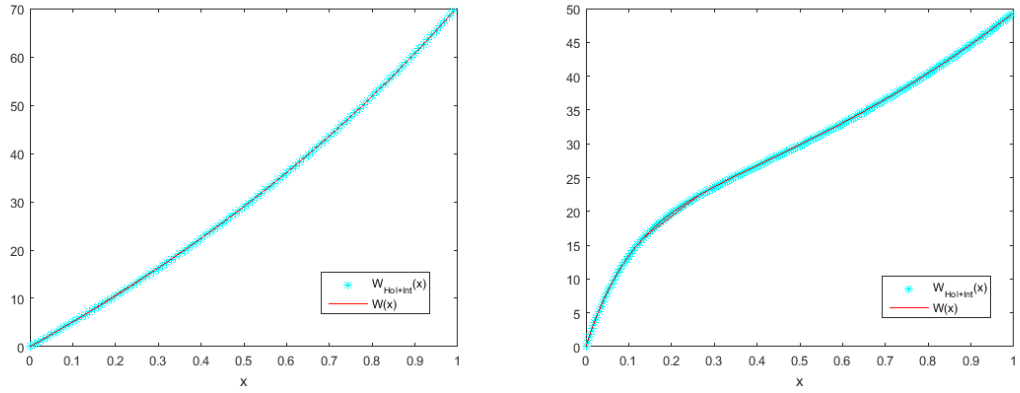


(c)



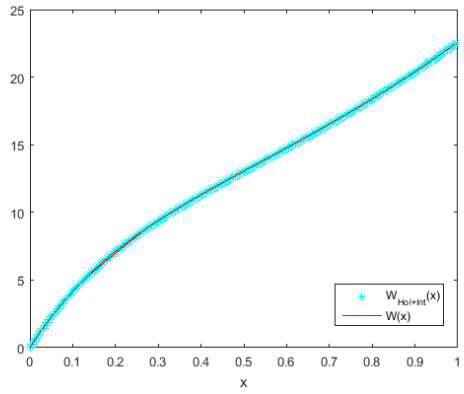
(d)

Figure 4.8: W approximated by spline and Hollenbeck algorithm. In all four plots we have on the x-axis the real numbers x in the interval $[0, 1]$ and on the y-axis we see in red color $W(x)$ computed by the the analytic formula. In in turquois color we see the the values $W(x)$ approximated by the Hollenbeck algorithm and spline interpolation. In (a) we have the case $r = 0.01, \sigma = 0.2$, in (b) we have $r = 0.05, \sigma = 0.1$, in (c) we have $r = 0.1, \sigma = 0.2$ and in (d) we have $r = 0.1, \sigma = 0.6$.

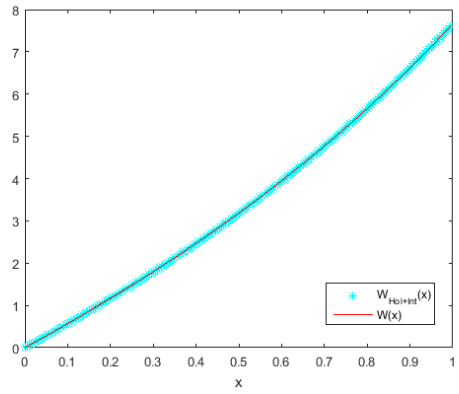


(a)

(b)



(c)



(d)

Conclusions

In this thesis we had a detailed analysis of the evaluation of perpetual American options with non-decreasing payoff function and guarantee. We saw in Chapter 1 how the notation for optimal stopping problems is rigorously introduced and that a Lévy process is a Markov process in a very general way. Furthermore, we have identified how its transition function has to be chosen and we showed that we can assume without loss of generality the existence of a family of probability measures $\{\mathbb{P}_x : x \in \mathbb{R}\}$ under which the Lévy process starts with different values $x \in \mathbb{R}$. We used this to give a rigorous derivation of Theorem 3.3.2 in Chapter 3. This theorem tells us that the value of the option is given by

$$v(x) = \sup_{a<x<b} v_{a,b}(x),$$

with

$$v_{a,b}(x) = k(x)Z(x-a) + \frac{g_x(b) - g_x(a)Z(b-a)}{W(b-a)}W(x-a).$$

Since the scale function Z can easily be derived by integration over W, the crucial part to use the formula is to find the scale function W. We have derived in Chapter 2 an analytic solution for W in the case that the process $(X_t)_t$ is modelled by a scaled Brownian motion with drift. In chapter 4 we received good computational results for this case. So we found a formula for the value of an perpetual American call option with guarantee when the stock price is modelled by scaled Brownian motion with drift. For the case that the process $(X_t)_t$ is modelled by a Brownian motion with jumps we could not derive an analytic solution for the scale function. In this case we have to find W by approximating the inverse Laplace transform of

$$\frac{1}{\psi - r},$$

where ψ is the Laplace exponent of $(X_t)_t$. We did already a first approach for this in Section 4.5, where we tested an approximation method in the Brownian motion case. An interesting question for further research would be to investigate how good this can actually be applied to jump processes.

Bibliography

[1] Christensen, S., & Irle, A. (2013). American Options with guarantee–A class of two-sided stopping problems. Statistics & Risk Modeling, 30(3), 237-254.

[2] Kyprianou, A. (2006). Introductory lectures on fluctuations of Lévy processes with applications. Springer Science & Business Media.

[3] Peskir, G., & Shiryaev, A. N. (2006). Optimal stopping and free-boundary problems. Springer Science & Business Media.

[4] S. Resnick (1992). Adventures in stochastic processes. Birkhuser Boston, Inc.

[5] N. Privault. Financial Mathematics. Notes on stochastic finance. Nanyang Technological University. Retrieved from `http://www.ntu.edu.sg/home/nprivault/indext.html.`

[6] Rogers, L. C. G., & Williams, D. (1994). Diffusions, Markov processes and martingales: Volume 1, Itô calculus (Vol. 1). Cambridge university press.

[7] Spreij, P. J. (2012). Measure theoretic probability. UvA Course Notes.

[8] FM Spieksma (2016). An Introduction to Stochastic Processes in Continuous Time. VU Course Notes.

[9] Shreve, S. E. (2004). Stochastic calculus for finance II: Continuous-time models (Vol. 11). Springer Science & Business Media.

[10] `https://math.stackexchange.com/questions/1090201/equality-of-laplace-transform.`(2018)

[11] `http://www.columbia.edu/~ww2040/allpapers.html.`

[12] Tucker McClure. `https://nl.mathworks.com/matlabcentral/fileexchange/39035-numerical-inverse-laplace-transform?focused=5251794&tab=function`.

[13] Hollenbeck, K. J. (1998) INVLAP.M: A matlab function for numerical inversion of Laplace transforms by the de Hoog algorithm, `http://www.isva.dtu.dk/staff/karl/invlap.html.` (2018)