Stockholms universitet

Optimal Dividends With Applications To Insurance

Jörgen Persson

Masteruppsats i försäkringsmatematik
Master Thesis in Actuarial Mathematics

Masteruppsats 2021:19
Försäkringsmatematik
September 2021

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Mathematical Statistics
Stockholm University
Master Thesis **2021:19**
http://www.math.su.se

Optimal Dividends With Applications To Insurance

Jörgen Persson*

September 2021

Abstract

The purpose of this thesis is to investigate how the optimal dividend problem relates to insurance mathematics, in particular the Cramér-Lundberg model. The optimal dividend problem will be studied for both restricted and unrestricted dividend rates. We will derive optimal value functions for restricted and unrestricted dividend rates when the reserve dynamics are governed by the Cramér-Lundberg model and its approximation. In order to achieve this aim we will assume that claim sizes are exponentially distributed, this will enable us to find explicit solutions to the Hamillton-Jacobi-Bellman equations.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: jorgen.persson94@gmail.com. Supervisor: Kristoffer Lindensjö.

Acknowledgements

I would like to give a special thanks to my supervisor associate professor Kristofer Lindensjö who despite an ongoing pandemic patiently helped me to understand various results and definitions used in this paper. Moreover, I would like to thank my friends: Per, Jonas and Michaela for their valuable comments and proof reading. I am also highly grateful for the support I have received from my parents Annica and Tommy Persson.

Contents

Notation

Symbol	Explanation
$\mathbb{E}$	Expected value operator.
$\sim$	Distributed as.
$\emptyset$	Empty set.
$\in$	Element relation.
$\forall$	Universal quantifier.
$\exists$	Existential quantifier.
$\cup$	Union.
$\cap$	Intersection
$C^{1,2}$	Denotes differentiability class.
$Po(\lambda)$	Denotes the Poisson distribution.
$N(\mu, \sigma)$	Denotes the normal distribution.
$preim$	Denotes the pre-image of a function.
$\mathbb{I}$	Indicator variable.
$\max$	Maximum operator.
$\inf$	Infimum operator.

1 Introduction

Stochastic control theory is a branch of control theory which studies optimization problems subject to dynamical systems with uncertainty. It has had quite wide application in diverse fields such as operations research, finance, game theory and insurance. Stochastic control heavily relies on probability theory and stochastic calculus as well as deterministic control theory, for example dynamic programming [5].

Stochastic control theory has been widely applied in insurance mathematics and actuarial science due to the stochastic nature of what is studied. When one considers stochastic control in insurance settings, examples of control variables include premium loadings, reinsurance policy, initial free reserves and dividend rates [14]. Of particular importance is the stochastic control problem referred to as the optimal dividend problem which has been important for insurance mathematics and is an active area of research [4].

The optimal dividend problem seeks to find optimal dividends that maximize expected discounted dividends given the stochastic nature of the risk reserves of a insurance company, where the control variable is the dividend rate. In a non-life insurance setting these reserves can be modelled by the Cramér-Lundberg model [12]. In this thesis we will investigate the optimal dividend problem in continuous time for restricted and unrestricted dividend rates i.e., dividends rates which are constrained to a specific interval, and dividends rates which are allowed to vary in an unconstrained way. We will then apply this within the context of the Cramér-Lundberg model, both for a diffusion approximation of said model and without the approximation. In this context we are interested in finding a so called optimal value function which is the supremum of the expected total amount of discounted dividends as a function of initial capital. Given certain regularity assumptions, then the optimal value function is the solution to the Hamilton-Jacobi-Bellman equation(HJB equation). In general solving the HJB is a rather difficult task, and solutions may not be unique, but many times one can find so called "weak solutions" or "viscosity solutions". However, we will in this paper only focus on cases where the optimal value function is sufficiently regular such that unique solutions can be found. Furthermore, if one has found a solution to the HJB equation, it then remains to show that the solution is indeed the optimal value function, this is done via a so called "verification theorem", from which it follows that the solution to HJB equation is the optimal value function. Verification theorems hence play an important role in the optimal dividend problem. Consequently, we will consistently deal with different versions of it. We will in this thesis investigate the HJB equations and verification theorems for the Cramér-Lundberg model, as well as its diffusion approximation.

1.1 Outline and purpose

The thesis will have the following structure:

- In Section 2 we will present definitions and results from probability theory, stochastic calculus and stochastic control.

- In Section 3 we will present and prove important results for the optimal dividend problem in continuous time in the context of insurance. In particular we will show results for the diffusion model and the Cramér-Lundberg model for restricted and unrestricted dividend rates.

In Section 2 we will present important definitions such as filtrations, Brownian motion, Martingales, Itô's lemma and the stochastic control problem. In Section 3 we will present and prove results for optimal dividend problem in a diffusion setting, these results will be connected to finding optimal solutions such as the optimal value function. We will then investigate the Cramér-Lundberg model using a diffusion approximation and without the approximation.

We will endeavour to make this thesis as self-contained as possible. However, some mathematical details will be omitted and not explicitly mentioned. In order to achieve this we will rely on several sources of information for the optimal dividend problem such as the work done by Taksar and Asmussen[4], Asmussen and Steffensen[3, p.355-381], and Jeanblanc et al [10],and Schmidli[14, p.69-97].

2 Preliminaries

In this section we will introduce necessary definitions, lemmas and theorems from probability theory, stochastic calculus and stochastic control. The results in this section will play an important role in constructing and proving results for optimal dividend problem formulated in Section 3 of this thesis. We will divide this section in three subsections. The first subsection we will present definitions and results from probability theory and stochastic processes, in the second we will present results from stochastic calculus and in the third we will present results from stochastic control theory in continuous time.

2.1 Probability theory and stochastic processes

In this subsection we will introduce fundamental definitions and results from probability theory and stochastic process. In order to do this we will need to rely on concepts from measure theory such as σ-algebras and filtrations. The results and definitions in this section can be found in [11], [8], [14]. Some of definitions and results are included in their entirety for the sake of completeness, hence some results will not be necessary in order to appropriately define and solve the optimal control problem.

Remark 2.1 (Terminology and notation)**.** Let Ω be a non-empty set called the *sample space* and let 2^Ω denote the power set i.e. the set of all subsets of Ω. Let the elements of Ω be called the *sample points*, and let the subsets Ω of be called *events*.

Definitions 2.1-2.13 are defined in a quite similar way to [8, p.3-32]. These definitions are used frequently when dealing with probability theory and stochastic calculus, and will be important when defining the optimal dividend problem.

Definition 2.1 (σ-algebra)**.** *Let $\mathcal{F} \subseteq 2^\Omega$ be called a σ-algebra on the sample space Ω if*

1. $\emptyset \in \mathcal{F}$;
2. $A \in \mathcal{F} \implies A^c \in \mathcal{F}$;
3. $A_1, A_2, A_3 \dots \in \mathcal{F} \implies \cup_{k=1}^{\infty} A_k \in \mathcal{F}$.

Where A^c denotes the complement i.e. the set $A^c := \{\omega \in \Omega : \omega \notin A\}$.

Definition 2.2 (σ-algebra generated by an arbitrary family)**.** *The σ-algebra generated by the arbitrary family $\mathcal{O} \subset 2^\Omega$ is*

$$\mathcal{F}_{\mathcal{O}} = \bigcap_i \{\mathcal{F}_i \subset 2^\Omega : \mathcal{O} \subset \mathcal{F}_i\}$$

Where $\mathcal{F}_i$ are σ-algebras. In particular, this means that $\mathcal{F}_{\mathcal{O}}$ is the smallest σ-algebra containing $\mathcal{O}$.

Definition 2.3 (Borel σ-algebra)**.** Let $\Omega = \mathbb{R}^d$ and let $\mathcal{O} = \{B_x(R)\}_{R>0, x\in\mathbb{R}^d} \subset 2^\Omega$ *be a collection of open balls with radius* R *centered at* x*, i.e. the set* $B_x(R) = \{y \in \mathbb{R}^d : |x - y| < R\}$*. The generated* σ*-algebra by* $\mathcal{O}$ *is the Borel* σ*-algebra denoted by* $\mathcal{B}(\mathbb{R}^d)$*, whose elements are called Borel sets.*

Definition 2.4 (Probability measure)**.** *A probability measure is defined as a real-valued function:*

$$\mathbb{P} : \mathcal{F} \to [0, 1],$$

which satisfies the following requirements:

1. $\mathbb{P}(\Omega) = 1$;
2. *For an arbitrary family of disjoint sets* $\{A_k\}_{k\in\mathbb{N}} \subseteq \mathcal{F}$ *we have that:*

$$\mathbb{P}\left(\bigcup_{k\in\mathbb{N}} A_k\right) = \sum_{k\in\mathbb{N}} \mathbb{P}(A_k).$$

The sample space Ω *together with the* σ*-algebra* $\mathcal{F}$ *and the probability measure* $\mathbb{P}$ *is called a probability space i.e.* $(\Omega, \mathcal{F}, \mathbb{P})$*.*

Definition 2.5 (Filtration)**.** *Given a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, a filtration* $\{\mathcal{F}(t)\}_{t\geq 0}$ *is a family of* σ*-algebras such that:*

- $\forall\, t \geq 0 : \mathcal{F}(t) \subseteq \mathcal{F}$;
- $\forall s \leq t : \mathcal{F}(s) \subseteq \mathcal{F}(t)$.

The probability space equipped with a filtration i.e. $(\Omega, \mathcal{F}, \{\mathcal{F}(t)\}_{t\geq 0}, \mathbb{P})$*, is called a filtered probability space.*

Definition 2.6 (Stochastic variable)**.** *A function:* $X : \Omega \to \mathbb{R}$*, is called a stochastic variable if*

$$\forall\, U \in \mathcal{B}(\mathbb{R}) : \; preim_X(U) \in \mathcal{F},$$

where $\mathcal{B}(\mathbb{R})$ is the Borel σ-algebra and the preimage of the Borel set U under X is $preim_X(U) = \{\omega \in \Omega : X(\omega) \in U\}$.

Definition 2.7 (σ-algebra generated by a stochastic variable)**.** *The* σ*-algebra generated by the stochastic variable* X *is the family* $\sigma(X) \subseteq \mathcal{F}$ *given by:*

$$\sigma(X) = \{preim_X(U) \in \mathcal{F} : U \in \mathcal{B}(\mathbb{R})\}.$$

Definition 2.8 (Stochastic process)**.** *A stochastic process* $\{X(t)\}_{t\geq 0}$ *is a family of stochastic variables, where*

$$\forall t \geq 0 : \; X(t) : \Omega \to \mathbb{R}.$$

Let $X(t, \omega)$ *denote the value of* $X(t)$ *for a fixed sample point* ω*. The continuous* ω*-path of the stochastic process is defined as*

$$\begin{aligned} \gamma_X^\omega &: \mathbb{R} \to \mathbb{R}, \\ t &\mapsto X(t, \omega). \end{aligned}$$

A stochastic process in discrete time is defined similarly where the set $t \geq 0$ *is changed by some discrete index set* I *i.e.* $\{X(t)\}_{t\in I}$*.*

Remark 2.2. Note that a stochastic process can be seen as a function of two variables. However, for the sake of convenience we will just denote it as a function of the time-parameter t, i.e. $X(t)$.

Definition 2.9. *Given a filtered probability space* $(\Omega, \mathcal{F}, \{\mathcal{F}(t)\}_{t\geq 0}, \mathbb{P})$*, a filtration* $\{\mathcal{F}_X(t)\}_{t\geq 0}$ *generated by the stochastic process* $\{X(t)\}_{t\geq 0}$ *is defined as*

$$\mathcal{F}_X(t) = \mathcal{F}_{\mathcal{O}(t)},$$

where $\mathcal{O}(t) = \cup_{0\leq s\leq t}\sigma(X(s))$*. That is,* $\mathcal{F}_X(t)$ *is the smallest* σ*-algebra containing the* σ*-algebra generated by* $X(s)$ *for all* $s \leq t$*. Moreover, if* $\forall t \geq 0 : \mathcal{F}_X(t) \subseteq \mathcal{F}(t)$ *then* $\{X(t)\}_{t\geq 0}$ *is said to be adapted to the filtration* $\{\mathcal{F}(t)\}_{t\geq 0}$*.*

Remark 2.3. The filtration generated by the stochastic process can intuitively be seen as the "flow of information" of the stochastic process over time.

The definition of a Cadlag process will be important when defining various optimal control problems, for example the optimal dividend problem relating to the Cramer-Lundberg model in Section 3. The definition is inspired by [14, p.201].

Definition 2.10 (Cadlag process)**.** *A stochastic process* $\{X(t)\}_{t\geq 0}$ *is said to be Cadlag if the* ω*-path is right-continuous with left limits existing. That is,*

- $X(t-) = \lim_{s\uparrow t} X(s)$ exists;
- $X(t+) = \lim_{s\downarrow t} X(s) = X(t)$

A crucial part of the Cramér–Lundberg model is the stochastic process called the Poisson process, which has the following definition [12, p.13].

Definition 2.11 (Poisson process)**.** *A Poisson process is a stochastic process* $\{N(t)\}_{t\geq 0}$ *such that:*

1. $N(0) = 0$;
2. *For any* $u < v < s < t$ *it holds that* $N(t) - N(s)$ and $N(v) - N(u)$ *are independent, i.e. the property of independent increments.*
3. $\forall s < t :\ N(t) - N(s) \sim Po(\lambda(t - s))$;
4. $\{N(t)\}_{t\geq 0}$ *is Cadlag.*

Definition 2.12 (Brownian motion)**.** *A Brownian motion is a stochastic process* $\{W(t)\}_{t\geq 0}$ *such that:*

1. $W(0) = 0$;
2. *For any* $u < v < s < t$ *it holds that* $W(t) - W(s)$ *and* $W(v) - W(u)$ *are independent, i.e. the property of independent increments*
3. $\forall s < t :\ W(t) - W(s) \sim N(0, t - s)$;

4. $\forall \omega_0 \in \Omega$; $\gamma_W^{\omega_0} : \mathbb{R} \to \mathbb{R}$, *is a continuous function of* t

Remark 2.4. Brownian motion is sometimes referred to as the Wiener process.

The next definition will prove useful when defining properties of the Ito integral, and the optimal control problem relating to the Cramer-Lundberg model. The next definition follows from [8, p.63-64].

Definition 2.13 (Martingale)**.** *A stochastic process* $\{M(t)\}_{t\geq 0}$ *which is adapted to the filtration* $\{\mathcal{F}(t)\}_{t\geq 0}$ *is called a martingale if we have that:*

$$\mathbb{E}\left[M(t)|\mathcal{F}(s)\right] = M(s), \quad \forall\, s \leq t,$$

If instead

$$\mathbb{E}\left[M(t)|\mathcal{F}(s)\right] \geq M(s), \quad \forall\, s \leq t,$$

the stochastic process $\{M(t)\}_{t\geq 0}$ *is called a sub-martingale.* If

$$\mathbb{E}\left[M(t)|\mathcal{F}(s)\right] \leq M(s), \quad \forall\, s \leq t,$$

the stochastic process $\{M(t)\}_{t\geq 0}$ *is called a super-martingale.*

Remark 2.5. The intuition behind this definition of a martingale is that the future expected value of a stochastic process given the information up to time s is the same as value of the stochastic process at time s. For further technicalities see, [8, 63-64].

The next theorem will prove useful when dealing with verification theorems in Section 3. This result can be found in. [9, p.57].

Theorem 2.1 (Bounded convergence theorem)**.** *Let* $\{X_n\}_{n\in\mathbb{N}}$ *be a sequence of stochastic variables. If* $|X_n| \leq K$ *and* $X_n \to X$ *as* $n \to \infty$*, it then holds that:*

$$\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}\left[\lim_{n\to\infty} X_n\right].$$

2.2 Stochastic calculus

In this section we will present results and definitions in stochastic calculus that will be necessary in order to formulate and prove the results for the section on the optimal dividend problem. For this section we will rely on several sources, that is, the definitions and results can be found in [11], [5], [8]. Throughout this section we will assume that stochastic quantities are defined on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}(t)\}_{t\geq 0}, \mathbb{P})$.

The next definitions and results will be necessary in order to define the Itô integral and Itô's lemma, which will prove useful in relation to the verification theorems in Section 3. For a complete and rigours treatment and construction of the Itô integral, see[8, p.73-82], however we will follow a construction similar to [15, p.126-136]. Before giving a simple construction of the Itô integral, we will define a specific type of stochastic process, this definition follows from[15, p.126].

Definition 2.14 (Simple stochastic process)**.** *Let*

$$0 \leq t_0 \leq t_1 \cdots \leq t_n = T$$

be a partition of the interval $[0, T]$*. A stochastic process* $\{\Delta(t)\}_{t\geq 0}$ *which is adapted to the filtration* $\{\mathcal{F}(t)\}_{t\geq 0}$ *and is constant in* t *on each subinterval* $[t_j, t_{j+1})$ *is called a simple stochastic process.*

The general idea behind the Itô integral is to give a reasonable interpretation of the following expression:

$$\int_0^T \Delta(t) dW(t),$$

where $\{W(t)\}_{t\geq 0}$ is a Brownian motion, and $\{\Delta(t)\}_{t\geq 0}$ is a simple stochastic process. The evident problem that one encounters with respect the expression above is that the Brownian motion paths are not differentiable with respect to t [15, p.125-126]. We will now proceed with giving a simplified heuristic construction of the Itô integral. First consider:

$$I(t) = \sum_{j=0}^{k-1} \Delta(t_j)(W(t_{j+1}) - W(t_j)) + \Delta(t_k)(W(t) - W(t_k)),$$

where $t_k \leq t \leq t_{k+1}$, and we can choose $t = T$. The stochastic process $\{I(t)\}_{t\geq 0}$ described above is the Itô integral of a simple stochastic process $\{\Delta(t)\}_{t\geq 0}$, that is

$$I(t) = \int_0^t \Delta(s) dW(s).$$

Moreover, it can be shown that the stochastic process above possesses the martingale property(among other properties) i.e. that $\mathbb{E}[I(t)|\mathcal{F}(s)] = I(s)$ [15, p.128]. Evidently, we want to generalise the Itô integral for integrands other than the simple stochastic process $\{\Delta(t)\}_{t\geq 0}$. Let us consider a stochastic process $\{X(t)\}_{t\geq 0}$ which is adapted to the filtration $\{\mathcal{F}(t)\}_{t\geq 0}$, with the property

$$\mathbb{E}\left[\int_0^T X(t)^2 dt\right] < \infty, \forall\, T > 0.$$

In order to make sense of the expression

$$\int_0^T X(t) dW(t),$$

we need to approximate $\{X(t)\}_{t\geq 0}$ by a simple stochastic process $\{\Delta_n(t)\}_{t\geq 0}$, where $0 \leq t_0 \leq t_1 \cdots \leq t_n = T$. It turns out that it is possible to choose a sequence of processes $\{\Delta_n(t)\}_{t\geq 0}$, $n \in \mathbb{N}$ such that

$$\lim_{n\to\infty} \mathbb{E}\left[\int_0^T |\Delta_n(t) - X(t)|^2 dt\right] = 0.$$

Since $\{\Delta_n(t)\}_{t\geq 0}$ is a simple stochastic process the Itô integral

$$I_n(T) = \int_0^T \Delta_n(t)dW(t)$$

is already defined for all $t \leq T$, it can then be shown that the following integral

$$\int_0^t X(u)dW(u) = \lim_{n\to\infty} \int_0^t \Delta_n(u)dW(u)$$

inherits the properties of the Itô integral of simple stochastic processes, e.g. the martingale property. However, further technical requirements are necessary in order to assure that the above limit exists, see [8, p.78] and [15, p.134]. In particular, we must have that I_n is a Cauchy sequence in $L^2(\mathbb{P})$ [13, p.28].

Remark 2.6. We will throughout this thesis utilise the differential form of the Itô integral

$$I(t) = I(0) + \int_0^t X(u)dW(u),$$

that is, $dI(t) = X(t)dW(t)$,[15, p.132].

We will now provide an theorem stating important properties of the Itô integral, in particular the martingale property which will play an important role in Section 3. For further details, see [11, p.91-99] and [8, p.74-82].

Theorem 2.2 (Properties of the Itô integral)**.** *Let $\{X(t)\}_{t\geq 0}$ be a stochastic process adapted to the filtration $\{\mathcal{F}(t)\}_{t\geq 0}$. Moreover, let $\{X(t)\}_{t\geq 0}$ satisfy:*

$$\mathbb{E}\left[\int_0^T X(t)^2 dt\right] < \infty, \forall\, T > 0.$$

Then the Itô integral

$$I(t) = \int_0^t X(t)dW(t)$$

have the following properties:

1. *Linearity: for any stochastic processes $\{X(t)\}_{t\geq 0}$ and $\{Y(t)\}_{t\geq 0}$ adapted to the filtration $\{\mathcal{F}(t)\}_{t\geq 0}$ which satisfies the condition above, then it holds that:*

$$\int_0^t c_1 X(t) + c_2 Y(t)dW(s) = c_1 \int_0^t X(t)dW(s) + c_2 \int_0^t Y(t)dW(s).$$

   for all $c_1, c_2 \in \mathbb{R}$.

2. *Martingale property: it holds that*

$$\mathbb{E}[I(t)|\mathcal{F}(s)] = \mathbb{E}[I(s)] = 0.$$

We will now give a definition of a stochastic differential equation, this will prove important when defining the stochastic control problem. The following definition is constructed from [11, p.126].

Definition 2.15 (Stochastic differential equation)**.** *Let* $\{W(t)\}_{t\geq 0}$ *be a Brownian motion. Consider the equation:*

$$dX(t) = \mu(X(t), t)dt + \sigma(X(t), t)dW(t),$$

where the functions are defined as follows:

$$\mu : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R},$$
$$\sigma : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R},$$

and $\{X(t)\}_{t\geq 0}$ *is an unknown stochastic process. An equation of this form is called a stochastic differential equation(SDE). Furthermore,* $\{X(t)\}_{t\geq 0}$ *is called a strong solution to the stochastic differential equation if*

$$\int_0^t \mu(X(s), s)ds \text{ and } \int_0^t \sigma(X(s), s)dW(s)$$

exists $\forall t > 0$, *and*

$$X(t) = X(0) + \int_0^t \mu(X(s), s)ds + \int_0^t \sigma(X(s), s)dW(s).$$

Remark 2.7. A solution to a stochastic differential equation is often called *a diffusion process*. Furthermore, there are more general SDEs of the form

$$dX(t) = \mu(t)dt + \sigma(t)dW(t),$$

where $\mu(t)$ and $\sigma(t)$ are assumed to be adapted stochastic processes, for further details see [11, p.126], [8, p.83].

Remark 2.8. Note that we contrast "strong" solutions with so called "weak" solutions which are solutions where strong solutions does not exist. That is, we can find solutions to a SDE with less stringent conditions on the coefficients of the SDE [11, p.136].

Definition 2.16 (Brownian motion with drift)**.** *A stochastic process* $\{X(t)\}_{t\geq 0}$ *is called a Brownian motion with drift if it satisfies the following stochastic differential equation:*

$$dX(t) = \mu dt + \sigma dW(t),$$

where $\{W(t)\}_{t\geq 0}$ is a Brownian motion μ *and* σ *are given constants.*

The next theorem will prove crucial when dealing with verification theorems and the HJB equation, hence Itô's Lemma plays an important role in stochastic control theory. For further details on this theorem, see [11, p.105], [5, p.51-52].

Theorem 2.3 (Itô's lemma)**.** *Let f be a $C^{1,2}$- function , and let $\{X(t)\}_{t\geq 0}$ be a stochastic process that satisfies the SDE:*

$$dX(t) = \mu(t)dt + \sigma(t)dW(t)$$

then

$$df(t, X(t)) = \left(\frac{\partial f}{\partial t} + \mu(t)\frac{\partial f}{\partial x} + \frac{\sigma^2(t)}{2}\frac{\partial^2 f}{\partial x^2}\right)dt + \sigma(t)\frac{\partial f}{\partial x}dW(t).$$

A useful version of Itô's lemma is the following. Moreover, we will provide a informal heuristic proof of the following lemma.

Lemma 2.1. *Let f be a C^2 function and let $\{X(t)\}_{t\geq 0}$ be a Brownian motion with drift, it then holds that*

$$df(X(t)) = \left(\mu f'(x) + \frac{1}{2}\sigma^2 f''(x)\right)dt + \sigma f'(x)dW(t),$$

where the integral form is

$$f(X(t)) = f(0) + \int_0^t \left(\mu f'(x) + \frac{1}{2}\sigma^2 f''(x)\right)du + \int_0^t \sigma\frac{\partial f}{\partial x}dW(u).$$

Heuristic Proof. Using the Taylor expansion we can approximate $df(X(t))$ in the following way

$$df(X(t)) = f'(x)dX(t) + \frac{1}{2}f''(x)\left(dX(t)\right)^2 + \frac{1}{6}f'''(x)\left(dX(t)\right)^3\cdots.$$

We then use $dX(t) = \mu dt + \sigma dW(t)$, which gives us

$$df(X(t)) = f'(x)(\mu dt + \sigma dW(t)) + \frac{1}{2}f''(x)\left(\mu dt + \sigma dW(t)\right)^2 + \cdots.$$

By rearranging terms we see that

$$df(X(t)) = \mu f'(x)dt + \frac{1}{2}\sigma^2 f''(x)dW(t)^2 +$$
$$\sigma f'(x)dW(t) + \frac{1}{2}f''(x)\left(\mu^2 dt^2 + 2\mu\, dt\, \sigma\, dW(t)\right) + \cdots$$

We then use the following relations $dt^2 = 0$, $dt\, dW(t) = 0$ and $dW(t)^2 = dt$, [15, p.105], which gives us

$$df(X(t)) = \left(\mu f'(x) + \frac{1}{2}\sigma^2 f''(x)\right)dt + \sigma f'(x)dW(t),$$

note that the higher order terms of the Taylor expansion vanish, since $(dX(t))^3 = \sigma^2 dt(\mu dt + \sigma dW(t))$, etc. From Remark 2.7 we then get

$$f(X(t)) = f(0) + \int_0^t \left(\mu f'(x) + \frac{1}{2}\sigma^2 f''(x)\right)du + \int_0^t \sigma\frac{\partial f}{\partial x}dW(u).$$

This completes the proof. □

2.3 Stochastic control in continuous time

In this subsection we will introduce the stochastic control problem in continuous time as well as important results and concepts that will be used in the next section it formulate the optimal dividend problem. In order to do this we will use the definitions and results from the previous subsections on stochastic processes and stochastic calculus. The definitions, results and notation in this subsection are formulated with respect to [5, p.282-294] and [6].

We will now give a general formal definition of the optimal control problem in continuous time. This will play a foundational role when defining the optimal dividend problem in Section 3. The following definition is constructed from [5, p.283-284,287].

Definition 2.17 (Control problem)**.**

$$\begin{aligned}
\max_u \quad & J(t,x,u) = \mathbb{E}_{t,x}\left[\int_t^T F(s, X^u(s), u(s, X^u(s)))ds + \Phi(X^u(T))\right] \\
\text{subject to} \quad & dX^u(s) = \mu(s, X^u(s), u(s, X^u(s)))ds + \sigma(s, X^u(s), u(s))dW(s), \\
& X(t) = x, \\
& u(s,y) \in \boldsymbol{U},\ \forall\ (s,y) \in [t,T] \times \mathbb{R},
\end{aligned}$$

where $\boldsymbol{U}$ *is some subset of* $\mathbb{R}$*, referred to as a control constraint. Furthermore, the functions are defined as*:

$$\begin{aligned}
&F : \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R} \to \mathbb{R} \\
&\Phi : \mathbb{R} \to \mathbb{R} \\
&\mu : \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R} \to \mathbb{R} \\
&\sigma : \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R} \to \mathbb{R} \\
&u : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}.
\end{aligned}$$

Remark 2.9. We will throughout this thesis use the notation $\mathbb{E}_{t,x}[\cdot] := \mathbb{E}[\cdot | X^u(t) = x, t < T]$.

Remark 2.10. We shall only consider feedback or Markovian controls, meaning that the u used at time t is a function of t and the state x of the controlled process at that time. Moreover, note that the for the "control law" u we have

$$u(t) := u(t, X^u(t)).$$

How these functions are interpreted depends on the context of application. In some constructions the function F can be the utility function, in our case it will be interpreted as the dividend rate. That is, we want to maximise our utility given the dynamics modeled by a stochastic differential equation which can be "steered" or "controlled". We will now define what is meant by "admissible strategies". This will play an important role when defining the restricted and unrestricted optimal dividend problems in Section 3. For further details, see [5, p.283-284].

Definition 2.18 (Admissible strategies)**.** *A "control law" u is called admissible if the following conditions hold:*

- $\forall t \in \mathbb{R}_+ : \ \forall x \in \mathbb{R} : u(t,x) \in \boldsymbol{U}$, *where* $\boldsymbol{U} \subseteq \mathbb{R}$ *is the control constraint*;
- *The stochastic differential equation:*

$$\begin{aligned} &dX^u(s) = \mu(s, X^u(s), u(s))ds + \sigma(s, X^u(s), u(s))dW(s), \\ &X(t) = x, \end{aligned}$$

  has a unique solution for all initial values.

The class of admissible strategies or "control laws" is denoted by $\mathcal{U}$.

Having defined admissible strategies, we can now define the optimal value function. The primary focus of Section 3 will be to find such a function, hence, it will play a crucial role in this thesis.

Definition 2.19 (Value function)**.**

- *The value function J is defined as:*

$$\begin{aligned} &J : \mathbb{R}_+ \times \mathbb{R} \times \mathcal{U} \to \mathbb{R} \\ &J(t,x,u) = \mathbb{E}_{t,x}\left[\int_t^T F(s, X^U(s), u(s))ds + \Phi(X^U(T))\right], \end{aligned}$$

  given the dynamics in Definition 2.17.

- *The optimal value function V is defined as follows*:

$$\begin{aligned} &V : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R} \\ &V(t,x) = \sup_{u \in \mathcal{U}} J(t,x,u). \end{aligned}$$

For the optimal control problems we will make certain assumptions. These assumptions are ad hoc and can vary slightly depending on the context, however, they usually are the following [6, p.10].

Assumption 1. *For the control problem it holds that:*

$$\exists\, \hat{u} \in \mathcal{U} : J(t,x,\hat{u}) = V(t,x),$$

where $\hat{u}$ is called an optimal control law. That is, for the optimal control problem there must exist a control law which solves the problem.

Assumption 2. *The optimal value function V must be contained in some suitable differentiability class, specifically we have that $V \in C^{1,2}$.This is often referred to as "regularity assumptions".*

We will now state the Bellman optimality principle, this is an important result for dynamic programming with respect to the Hamilton-Jacobi-Bellman equation. This result can found in [6, p.11].

Theorem 2.4 (Bellman optimality principle)**.** *If $\hat{u}$ is an optimal control law on the interval $[t,T]$, then $\hat{u}$ is optimal control law on all subinterval, i.e.*

$$\forall v \in [t,T] : J(v,x,\hat{u}) = V(v,x) \implies \forall n \geq t : \forall s \in [n,T] : J(s,x,\hat{u}) = V(s,x).$$

In order to give a concise formulation of the Hamilton-Jacobi-Bellman equation we will define the differential operator, which has the following definition[6, p.7].

Definition 2.20 (Differential operator)**.** *The differential operator $\mathcal{A}^u$ is given by*

$$\mathcal{A}^u = \mu(t,x,u)\frac{\partial}{\partial x} + \frac{1}{2}\sigma(t,x,u)^2\frac{\partial^2}{\partial x^2}.$$

We are now ready to state the Hamilton-Jacobi-Bellman equation in relation to the stochastic control problem. The "HJB" equation will play a crucial role throughout this paper since it is the foundation of stochastic control problems in continuous time. Here we will state the HJB equation as a partial differential equation, however, in the next section we will consider it as an ordinary differential equation. The heuristic proof is a slight elaboration of the argument found in [5, p.288-291]. The proof will utilise Theorem 2.2-2.3 and Assumption 1-2.

Theorem 2.5 (Hamilton-Jacobi-Bellman equation)**.** *If Assumption 1-2 holds, then:*

1. *The optimal value function V satisfies the Hamilton-Jacobi-Bellman equation:*

$$\frac{\partial V}{\partial t}(t,x) + \sup_{u\in U}\{F(t,x,u) + \mathcal{A}^u V(t,x)\} = 0$$
$$V(T,x) = \Phi(x).$$

2. *If the control law is equal to the optimal control law i.e. $u = \hat{u}(t,x)$, it holds that:*

$$\sup_{u\in U}\{F(t,x,u) + \mathcal{A}^u V(t,x)\} = F(t,x,\hat{u}) + \mathcal{A}^{\hat{u}} V(t,x),$$

for all initial values $(t,x) \in [0,T] \times \mathbb{R}$.

Heuristic Proof. We begin by considering the following control law u^*:

$$u^*(s,y) = \begin{cases} u(s,y), & (s,y) \in [t,t+h] \times \mathbb{R} \\ \hat{u}(s,y), & (s,y) \in (t+h,T] \times \mathbb{R}, \end{cases}$$

where $h > 0$(interpreted as an arbitrarily small number), u is some arbitrary control, and $\hat{u}$ is the optimal control law from Assumption 1. That is, in interval $[t,t+h]$ we use the arbitrary control and in $(t+h,T]$ we use the optimal control law.
Given a fixed pair $(t,x) \in (0,T) \times \mathbb{R}$, we consider the "strategies" of using the optimal control law $\hat{u}$ or the control law u^*. We then compute $J(t,x,\hat{u})$ and $J(t,x,u^*)$, where

according to Definition 2.19, it must be the case that $V(t,x) = J(t,x,\hat{u}) \geq J(t,x,u^*)$.

We begin by computing $J(t,x,u^*)$. Let J_1 and J_2 denote functions in the two intervals such that $J_1 + J_2 = J(t,x,u^*)$. On the interval $[t, t+h]$, the value function J_1 is the following:

$$J_1 = \mathbb{E}_{t,x}\left[\int_t^{t+h} F(s, X^u(s), u(s))ds\right].$$

On the interval $[t+h, T]$, the value function J_2 is the following

$$J_2 = \mathbb{E}_{t,x}\left[V(t+h, X^u(t+h))\right],$$

since $X^u(t+h)$ at time $t+h$ is a stochastic variable. We then see that

$$J(t,x,u^*) = \mathbb{E}_{t,x}\left[\int_t^{t+h} F(s, X^u(s), u(s))ds + V(t+h, X^u(t+h))\right].$$

We now again consider the inequality $V(t,x) \geq J(t,x,u^*)$, that is:

$$V(t,x) \geq \mathbb{E}_{t,x}\left[\int_t^{t+h} F(s, X^u(s), u(s))ds\right] + \mathbb{E}_{t,x}\left[V(t+h, X^U(t+h))\right], \tag{2.2}$$

where equality holds if and only if u is optimal control law $\hat{u}$. Given Assumption 2, Ito's lemma for the optimal value function V gives us

$$V(t+h, X^u(t+h)) = V(t,x) +$$
$$\int_t^{t+h}\left(\frac{\partial V}{\partial s}(s, X^u(s)) + \mathcal{A}^u V(s, X^u(s))\right)ds + \int_t^{t+h} \sigma \frac{\partial V}{\partial x}(s, X^u(s))dW(s).$$

Using the martingale property from Theorem 2.2, we see that

$$\mathbb{E}_{t,x}\left[V(t+h, X^U(t+h))\right] =$$
$$V(t,x) + \mathbb{E}_{t,x}\left[\int_t^{t+h}\left(\frac{\partial V}{\partial s}(s, X^U(s)) + \mathcal{A}^u V(s, X^U(s))\right)ds\right]$$

If we then plug $\mathbb{E}_{t,x}\left[V(t+h, X^U(t+h))\right]$ into the inequality from 2.2, we obtain

$$0 \geq \mathbb{E}_{t,x}\left[\int_t^{t+h}\left(F(s, X^U(s), u(s)) + \frac{\partial V}{\partial s}(s, X^U(s)) + \mathcal{A}^u V(s, X^U(s))\right)ds\right]$$

If we divide the expression by h and let $h \to 0$, we get

$$0 \geq F(t,x,u) + \frac{\partial V}{\partial t}(t,x) + \mathcal{A}^u V(t,x),$$

again equality holds if and only if u is the optimal control law $\hat{u}$. This gives us the desired result. $\square$

Theorem 2.6 (Verification theorem)**.** *Given two function $f(t,x)$ and $g(t,x)$. If f solves the Hamilton-Jacobi-Bellman equation:*

$$\frac{\partial f}{\partial t}(t,x) + \sup_{u \in U} \{F(t,x,u) + \mathcal{A}^u f(t,x)\} = 0$$
$$f(T,x) = \Phi(x),$$

and if

$$\sup_{u \in U} \{F(t,x,u) + \mathcal{A}^u f(t,x)\} = F(t,x,g) + \mathcal{A}^g f(t,x),$$

where $u = g(t,x)$, then it follows that:

1. *The optimal value function V is equal to the solution f, that is:*

$$V(t,x) = f(t,x);$$

2. *The optimal control law $\hat{u}$ is equal to g, that is:*

$$\hat{u}(t,x) = g(t,x).$$

Proof. The proof can be found in [5, p.292-293] □

Remark 2.11 (Logic behind the problem)**.** The logic behind stochastic control problem, Hamilton-Jacobi-Bellman equation and the verification theorem is the following:

1. Begin by formulating the optimal control problem.
2. Derive the Hamilton-Jacobi-Bellman equation for the problem , i.e. showing that **if** V is the optimal value function **and** if V is sufficiently regular **then** V solves the HJB equation.
3. **If** f solves the HJB equation **then** $f = V$ i.e. the solution to the HJB equation is the optimal value function. This is, the so called verification step of the argument.

Remark 2.11 will prove highly important for the rest of this thesis, since it summarises the general structure of Section 3 both for the restricted and unrestricted optimal dividend problem. In fact, every argument will follow this logical structure.

3 Optimal dividend problems in insurance

In this section we will present a version of a stochastic control problem, that is, the optimal dividend problem. In the first subsection we will present and prove important results for the optimal dividend problem for the case when the surplus or reserve process is governed by a Brownian motion with drift. In the second subsection we will present some relevant actuarial modelling, in particular the Cramér-Lundberg model of the risk reserve. In the third and last subsection we will formulate the optimal dividend problem in the context of the Cramér-Lundberg model for restricted and unrestricted dividend rates. The results in this section are similar to results given in [4], [10], [2, p.355-381] and [14, p.70-79].

3.1 Optimal dividend problem in a diffusion setting

In this section we will consider the problem of maximising the total discounted dividends given that the reserve dynamics are governed by a stochastic differential equation. In particular we will consider the case where the control variable i.e. the dividends are restricted to a specific interval. The uncontrolled reserve dynamics $\{X(t)\}_{t\geq 0}$, i.e without dividends that we will consider in this section can be formulated as a Brownian motion with drift. That is the uncontrolled reserve dynamics $\{X(t)\}_{t\geq 0}$ satisfies the SDE:

$$dX(t) = \mu dt + \sigma dW(t),$$

where μ is the drift parameter and σ is the variance parameter. We will assume that the controlled reserve dynamics $\{X^U(t)\}_{t\geq 0}$ satisfies the following SDE:

$$dX^U(t) = (\mu - u(X^U(t))dt + \sigma dW(t),$$

where the function $u(\cdot)$ is the dividend rate which is adjusted by the decision maker. The intuition behind this formulation is that the drift μ in the controlled setting is reduced by the dividend rate. In this section this function is only allowed to vary in $[0, u_0]$. Moreover, we will also provide the general logic and intuition behind the problem, which will apply for the rest of this section, see Remark 2.11. We will now give a formal definition of the problem. This definition follows the structure found in [5, p.283,287] and [3, p.367, 370].

Definition 3.1 (Restricted optimal dividend problem)**.**

$$\begin{aligned}
\max_{u(\cdot)} \qquad & J(x,u) = \mathbb{E}\left[\int_0^\tau e^{-\delta t} u(X^U(t))dt\right] \\
\text{subject to} \qquad & dX^U(t) = (\mu - u(X^U(t))dt + \sigma dW(t), \\
& X(0) = x, \\
& u(\cdot) \in [0, u_0],\ u_0 < \infty,\ \forall\, t > 0,
\end{aligned}$$

where $\{X^U(t)\}_{t\geq 0}$ is our controlled reserve process and $\{W(t)\}_{t\geq 0}$ is the standard Brownian motion given on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}(t)\}_{t\geq 0}, \mathbb{P})$ and

$(\mu, \sigma) > (0, 0)$ are our drift parameter and variance parameter. The risk-free interest rate is denoted by $\delta > 0$ and the ruin time is denoted by $\tau = \inf\{t : X^U(t) < 0\}$. The control variable is denoted by $U = (u(x))_{x \geq 0}$ i.e. the dividend rate. Furthermore, we define our control law u as a function:

$$u : \mathbb{R} \to [0, u_0]$$
$$x \mapsto u(x).$$

The aim of this optimization problem is therefore to find such a function that maximises $J(x, u)$.

Remark 3.1. Note that this control problem differs slightly from the control problem defined in Definition 2.17, since the upper limit of integration in the value function is a stochastic variable.

Remark 3.2. We will denote the optimal control by $U^* = (u^*(x))_{x \geq 0}$.Furthermore, U is chosen in a set of *admissible strategies*, see Definition 2.18. We will assume tacitly that $dX^U(t) = (\mu - u(X^U(t))dt + \sigma dW(t)$ has a solution for every (t, x), [3, p.356].

In order to find an optimal control (i.e a dividend rate which maximises expected total discounted dividends) we will need to rely on the so called Hamilton-Jacobi-Bellman equation[3, p.366-370]. The optimal value function V under some ad hoc regularity assumptions will satisfy this ODE. The optimal value function is the solution to the control problem stated in Definition 3.1, in this context $V(x)$ is interpreted as the maximal total discounted dividends. This function is defined as:

$$V(x) := \sup_{u \in [0, u_0]} J(x, u).$$

In particular, we need to show that **if** V is the optimal value function and **if** V is sufficiently regular in the sense of being differentiable, **then** V satisfies the Hamilton-Jacobi-Bellman equation. Furthermore, we need to show that **if** f is a solution to the Hamilton-Jacobi-Bellman equation, **then** $V = f$, i.e. the solution is equal to the optimal value function, this result follows from the verification theorem. Hence, we will have the following structure of argumentation:

1. Deriving the Hamilton-Jacobi-Bellman equation given the stochastic control problem and certain regularity assumptions.

2. Show that if we found a solution to the Hamilton-Jacobi-Bellman then the solution is the optimal value function i.e. the verification theorem.

3. Solving the Hamilton-Jacobi Bellman equation, whose solution is the optimal value function which follows from the verification theorem.

We will now proceed with deriving the Hamilton-Jacobi-Bellman equation. In order to formulate the Hamilton-Jacobi-Bellman equation we first need to define a *differential operator*. The differential operator is defined as follows [2, p.367]:

Definition 3.2 (Differential operator)**.** *Let f be a C^2 function, then the differential operator $\mathscr{L}^u$ is defined as:*

$$\mathscr{L}^u f(x) = (\mu - u(x))f'(x) + \frac{1}{2}\sigma^2 f''(x).$$

We will now state the Hamilton-Jacobi-Bellman equation for this problem. This theorem can be found in[3, p.368].

Theorem 3.1 (Hamilton-Jacobi-Bellman equation)**.** *If the optimal value function V is a C^2 function, then the optimal value function satisfies the following equation:*

$$0 = \sup_{0 \leq u \leq u_0} \left[\mathscr{L}^u V(x) - \delta V(x) + u \right].$$

Proof. A proof can be found in [3, p.368], other sources are [4, p.3-4] and [14, p.98-99]. □

First note that if $x = 0$, we will have that $V(0) = 0$, since this would mean that the company is bankrupt and no dividends can be paid out. From Theorem 3.1 and Definition 3.2 we see that the Hamilton-Jacobi-Bellman equation takes the following form:

$$\begin{aligned} 0 &= \sup_{0 \leq u \leq u_0} \left[(\mu - u)V'(x) + \frac{1}{2}\sigma^2 V''(x) - \delta V(x) + u \right] \\ &= \mu V'(x) + \frac{1}{2}\sigma^2 V''(x) - \delta V(x) + \sup_{0 \leq u \leq u_0} \left[u(1 - V'(x)) \right]. \end{aligned}$$

Furthermore, If $V'(x) > 1$ then $u(1 - V'(x))$ will be a strictly decreasing function of u hence we get the optimal $U^* = 0$. If $V'(x) < 1$ then $u(1 - V'(x))$ is a strictly increasing function of u, hence we get the optimal $U^* = u_0$. That is if $V'(x) < 1$ we will have the maximum dividend rate u_0. We then have the following equations:

$$0 = \begin{cases} \mu V'(x) + \frac{1}{2}\sigma^2 V''(x) - \delta V(x), & V'(x) > 1 \\ (\mu - u_0)V'(x) + \frac{1}{2}\sigma^2 V''(x) - \delta V(x) + u_0, & V'(x) < 1 \end{cases}$$

The first equation is a homogeneous second-order linear ordinary differential equation. The second equation is a non-homogeneous second-order linear ordinary differential equation. Both of which have explicit solutions. Let us formulate the solutions as a lemma [3, p.370]. The following lemma is inspired by [3, p.371] [4, p.4].

Lemma 3.1. *The following linear ordinary differential equations:*

$$\begin{aligned} 0 &= \mu f_0'(x) + \frac{1}{2}\sigma^2 f_0''(x) - \delta f_0(x), \\ 0 &= (\mu - u_0)f'(x) + \frac{1}{2}\sigma^2 f''(x) - \delta f(x) + u_0, \end{aligned}$$

have the solutions given by:

$$
\begin{aligned}
f_0(x) &= C_1 e^{\theta_1 x} + C_4 e^{\theta_2 x}, \\
f(x) &= C_3 e^{\theta_3 x} + C_4 e^{\theta_4 x} + \frac{u_0}{\delta},
\end{aligned}
$$

where:

$$
\begin{aligned}
\theta_{1,2} &= -\frac{\mu}{\sigma^2} \pm \sqrt{\frac{\mu^2}{\sigma^4} + \frac{2\delta}{\sigma^2}}, \\
\theta_{3,4} &= \frac{(u_0 - \mu)}{\sigma^2} \pm \sqrt{\frac{(u_0 - \mu)^2}{\sigma^4} + \frac{2\delta}{\sigma^2}}.
\end{aligned}
$$

Moreover, we have that $\theta_2 < 0 < \theta_1$ *and* $\theta_4 < 0 < \theta_3$.

Proof. The solutions to the characteristic polynomials

$$
\begin{aligned}
0 &= \theta^2 + 2\frac{\mu}{\sigma^2}\theta - \frac{2\delta}{\sigma^2}, \\
0 &= \theta^2 + 2\frac{\mu - u_0}{\sigma^2}\theta - \frac{2\delta}{\sigma^2},
\end{aligned}
$$

are given by:

$$
\begin{aligned}
\theta_{1,2} &= -\frac{\mu}{\sigma^2} \pm \sqrt{\frac{\mu^2}{\sigma^4} + \frac{2\delta}{\sigma^2}} \\
\theta_{3,4} &= \frac{(u_0 - \mu)}{\sigma^2} \pm \sqrt{\frac{(u_0 - \mu)^2}{\sigma^4} + \frac{2\delta}{\sigma^2}}.
\end{aligned}
$$

We can clearly see that $\theta_2 \leq 0$ and $\theta_4 \leq 0$. Hence, $f_0(x) = C_1 e^{\theta_1 x} + C_4 e^{\theta_2 x}$ is the the solutions to the first ODE. The homogeneous solution to the second ODE is given by $f_h(x) = C_3 e^{\theta_3 x} + C_4 e^{\theta_4 x}$, and the particular solution is given by $f_p(x) = \frac{u_0}{\delta}$. Hence, the solution to the second ODE is given by $f(x) = f_h(x) + f_p(x)$. This proves the lemma. □

Now we can proceed to the second step, i.e. the verification step. The proof is a slight elaboration of the second part of the proof found in [3, p.371] and [4, p.6]. The proof will rely on Definition 3.1, Theorem 2.1-2.2, Lemma 2.1 and Theorem 3.1.

Remark 3.3. We will throughout this thesis use the notation $x \wedge y = \min\{x, y\}$.

Theorem 3.2 (Verification theorem)**.** *Suppose that* $f(x)$ *is a* C^2 *function which is an increasing, bounded and positive solution to the HJB equation*

$$
0 = \begin{cases} \mu f'(x) + \frac{1}{2}\sigma^2 f''(x) - \delta f(x), & f'(x) > 1 \\ (\mu - u_0) f'(x) + \frac{1}{2}\sigma^2 f''(x) - \delta f(x) + u_0, & f'(x) < 1 \end{cases}
$$

with $f(0) = 0$, then it follows that $f(x) = V(x)$, i.e. the solution is the optimal value function, and the optimal control law is given by:

$$U^* = \begin{cases} 0 & \text{for } f'(x) > 1 \\ u_0 & \text{for } f'(x) < 1. \end{cases}$$

Proof. Let U be an arbitrary control law. Since f is C^2, and $X^U(0) = x$, it follows from Lemma 2.1, i.e. Ito's lemma that:

$$\begin{aligned} e^{-\delta(\tau\wedge t)} f(X^U(\tau \wedge t)) &= f(x) + \int_0^{\tau\wedge t} e^{-\delta s} \sigma f'(X^U(s)) dW(s) \\ + \int_0^{\tau\wedge t} e^{-\delta s} &\left[(\mu - u(s)) f'(X^U(s)) + \frac{1}{2}\sigma^2 f''(X^U(s)) - \delta f(X^U(s)) \right] ds \\ &\leq f(x) + \int_0^{\tau\wedge t} e^{-\delta s} \sigma f'(X^U(s)) dW(s) - \int_0^{\tau\wedge t} e^{-\delta s} u(s) ds. \end{aligned}$$

Furthermore, as in Theorem 2.5, equality holds if and only if $U = U^*$, that is, if U is the optimal control law. We then rearrange terms and take the expectations on both sides of the equation:

$$\begin{gathered} \mathbb{E}\left[e^{-\delta(\tau\wedge t)} f(X^U(\tau \wedge t))\right] + \\ \mathbb{E}\left[\int_0^{\tau\wedge t} e^{-\delta s} u(s) ds\right] - \mathbb{E}\left[\int_0^{\tau\wedge t} e^{-\delta s} \sigma f'(X^U(s)) dW(s)\right] \leq f(x). \end{gathered}$$

Since f' is a bounded function, this implies that

$$\mathbb{E}\left[\int_0^T \left(e^{-\delta s} \sigma f'(X^U(s))\right)^2 ds\right] < \infty, \ \forall T > 0.$$

Hence, from Theorem 2.2, i.e. the martingale property of the Ito integral, it follows that:

$$\mathbb{E}\left[e^{-\delta(t\wedge\tau)} f(X^U(t \wedge \tau))\right] + \mathbb{E}\left[\int_0^{t\wedge\tau} e^{-\delta s} u(s) ds\right] \leq f(x).$$

If the ruin-time is less than t, i.e. $\tau < t$ then $X^U(\tau) = 0$ which implies that $e^{-\delta\tau} f(X^U(\tau)) = 0$. Since f is a bounded function, it follows from Theorem 2.1 that:

$$\lim_{t\to\infty} \mathbb{E}\left[e^{-\delta(t\wedge\tau)} f(X^U(t \wedge \tau))\right] = \mathbb{E}\left[\lim_{t\to\infty} e^{-\delta(t\wedge\tau)} f(X^U(t \wedge \tau))\right] = 0.$$

It then follows that

$$\mathbb{E}\left[\int_0^{\tau} e^{-\delta s} u(s) ds\right] \leq f(x).$$

If $U = U^*$we get that:

$$V(x) = f(x).$$

This completes the proof. □

We are now ready to formulate a theorem stating the optimal value function under the condition that $u_0 \leq -\delta/\theta_4$. We will show that the value function has the desired properties i.e being bounded, $V'(x) < 1 \; \forall x$, since we are interested in the case when $u = u_0$, and $V(0) = 0$. We know that $V(x)$ needs to be bounded since we must have that:

$$0 \leq V(x) \leq \int_0^\infty e^{-\delta s} u_0 ds = \frac{u_0}{\delta}$$

We are now ready to formulate and prove the existence of an optimal value function for the case when $u_0 \leq -\delta/\theta_4$. The proof of the next theorem will be a version of proofs found in [3, p.372] and [4, p.5]. The proof will rely on Theorem 3.1-3.2 and Lemma 3.1.

Theorem 3.3. *Assume that $u_0 \leq -\delta/\theta_4$. Then the optimal value function is given by*

$$V(x) = \frac{u_0}{\delta}(1 - e^{\theta_4 x}) \tag{3.2}$$

and the optimal control law is given by $U^* = u_0$.

Proof. First we need to show that the solution to HJB equation from Theorem 3.1 is an increasing, bounded and positive solution. Consider the solution to the HJB equation in Lemma 3.1:

$$f(x) = C_3 e^{\theta_3 x} + C_4 e^{\theta_4 x} + \frac{u_0}{\delta}.$$

It is clear that we need $C_3 = 0$, otherwise f would not be bounded, since $\theta_3 > 0$. Furthermore, we need that $f(0) = 0$, this condition gives us $C_4 = -u_0/\delta$. Hence we get

$$f(x) = \frac{u_0}{\delta}(1 - e^{\theta_4 x}).$$

We can clearly see that:

$$\theta_4 < 0 \implies \lim_{x \to \infty} f(x) = \frac{u_0}{\delta}.$$

Moreover, we can see that $f(x)$ is an increasing solution, since

$$\theta_4 < 0 \implies f'(x) = -\theta_4 \frac{u_0}{\delta} e^{\theta_4 x} > 0.$$

If $f(x)$ is a increasing function and $f(0) = 0$ it follows that $f \geq 0$. Furthermore, we need to verify that $f'(x) < 1$ for all $x > 0$, in order for the solution to satisfy the HJB equation with the control law u_0. We can see that:

$$\theta_4 < 0 \implies f''(x) = -\frac{u_0}{\delta}\theta_4^2 e^{-\theta_4 x} < 0, \forall x \geq 0.$$

This shows that the function is concave, which implies that $f'(x)$ is a decreasing function. Hence, $f'(x)$ attains it maximum at $x = 0$, that is $f'(0) = -\theta_4 \frac{u_0}{\delta} > 0$. Furthermore, we can see from the condition $u_0 \leq -\delta/\theta_4$ that:

$$0 < -\theta_4 \frac{u_0}{\delta} \leq -\theta_4 \frac{-\delta/\theta_4}{\delta} = 1.$$

Hence,

$$f(x) = \frac{u_0}{\delta}(1 - e^{\theta_4 x}).$$

is a increasing, bounded positive solution to the equation

$$0 = (\mu - u_0)f'(x) + \frac{1}{2}\sigma^2 f''(x) - \delta f(x) + u_0.$$

It then follows from Theorem 3.2 that:

$$f(x) = V(x).$$

This completes the proof. □

We will now formulate a theorem for the case when $u_0 > -\delta/\theta_4$. The proof is an elaboration of the proof found in [3, p.372] and [4, p.5]. The proof will rely on Theorem 3.1-3.2 and Lemma 3.1.

Theorem 3.4. *Assume that $u_0 > -\delta/\theta_4$. Then the optimal value function is given by:*

$$V(x) = \begin{cases} K_1(e^{\theta_1 x} - e^{\theta_2 x}) & \textit{for } 0 \leq x \leq x_0, \\ u_0/\delta - K_2 e^{\theta_4 x} & \textit{for } x > x_0, \end{cases}$$

where the optimal control law is given by

$$U^* = \begin{cases} 0 & \textit{for } x < x_0 \\ u_0 & \textit{for } x > x_0. \end{cases}$$

The constants are given by:

$$\begin{aligned} K_1 &= \frac{1}{(\theta_1 e^{\theta_1 x_0} - \theta_2 e^{\theta_2 x_0})} \\ K_2 &= \frac{1}{\theta_4 e^{\theta_4 x_0}} \\ x_0 &= \frac{1}{\theta_1 - \theta_2} \log \frac{1 - A\theta_2}{1 - A\theta_1} > 0, \end{aligned}$$

where $A = \frac{u_0}{\delta} + \frac{1}{\theta_4}$.

Proof. From Lemma 3.1 we can see that the following functions satisfy the HJB equation:

$$f(x) = \begin{cases} K_1(e^{\theta_1 x} - e^{\theta_2 x}) & \text{for } 0 \leq x \leq x_0, \\ u_0/\delta - K_2 e^{\theta_4 x} & \text{for } x > x_0 \end{cases}$$

since, we can can choose $C_1 = K_1$, $C_2 = -C_1$, $C_3 = 0$(again to insure that the function is bounded) and $K_2 = -C_4$. Again we see that

$$\lim_{x \to \infty} f(x) = \frac{u_0}{\delta}$$

hence, the solutions is bounded. Moreover, we can see that $u_0/\delta - K_2 e^{\theta_4 x}$ is concave since $\theta_4 < 0$, as in Theorem 3.3. We can also see that $K_1(e^{\theta_1 x} - e^{\theta_2 x})$ is concave by taking the second derivative, and observing that $(\theta_1/\theta_2)^2 \leq 1$. Furthermore, the solution f is C^2, however we need to demonstrate that this still holds at x_0. That is, we need to show that $f''(x_0-) = f''(x_0+)$. From the HJB equation

$$0 = \begin{cases} \mu f'(x) + \frac{1}{2}\sigma^2 f''(x) - \delta f(x), & f'(x) > 1 \\ (\mu - u_0) f'(x) + \frac{1}{2}\sigma^2 f''(x) - \delta f(x) + u_0, & f'(x) < 1. \end{cases}$$

We see that

$$\begin{aligned} f''(x_0-) &= \frac{2}{\delta^2}\left(\delta f(x) - \mu\right) \\ f''(x_0+) &= \frac{2}{\delta^2}\left(\delta f(x) - (\mu - u_0) + u_0\right) \end{aligned}$$

We hence see that $f''(x_0-) = f''(x_0+)$. Since $f(x)$ is concave and C^2, it follows that the derivative is decreasing, this implies that there exists exactly one $x_0 > 0$ with $f'(x_0) = 1$, such that:

$$\begin{cases} f'(x) \geq 1 & \text{for } \ 0 \leq x \leq x_0, \\ f'(x) \leq 1 & \text{for } \ x > x_0. \end{cases}$$

Furthermore, the C^2 property it implies that

$$f(x_0-) = f(x_0+), \quad f'(x_0-) = 1 = f'(x_0+)$$

which gives us the following equations

$$\begin{aligned} K_1(e^{\theta_1 x_0} - e^{\theta_2 x_0}) &= \frac{u_0}{\delta} - K_2 e^{\theta_4 x_0} \\ K_1(\theta_1 e^{\theta_1 x_0} - \theta_2 e^{\theta_2 x_0}) &= 1 \\ -K_2 \theta_4 e^{\theta_4 x_0} &= 1 \end{aligned}$$

Let $A = \frac{u_0}{\delta} + \frac{1}{\theta_4}$ we can then see that

$$\frac{e^{\theta_1 x_0} - e^{\theta_2 x_0}}{\theta_1 e^{\theta_1 x_0} - \theta_2 e^{\theta_2 x_0}} = A,$$

If we solve this equation for x_0 we get

$$x_0 = \frac{1}{\theta_1 - \theta_2} \log \frac{1 - A\theta_2}{1 - A\theta_1}.$$

In order to ensure that $x_0 > 0$, we will now need to demonstrate that $A\theta_1 < 1$, if we rearrange we attain

$$\frac{u_0}{\delta} < \frac{1}{\theta_1} - \frac{1}{\theta_4}.$$

In order to demonstrate this, first note the inequality $\sqrt{a^2+b}-a < b/2a$ for $a, b > 0$, from this inequality and Lemma 3.1 we see that

$$\theta_1 = -\frac{\mu}{\sigma^2} + \sqrt{\frac{\mu^2}{\sigma^4} + \frac{2\delta}{\sigma^2}} = \frac{\sqrt{\mu^2 + 2\delta\sigma^2} - \mu}{\sigma^2} < \frac{\delta}{\mu},$$

$$\theta_4 = \frac{(u_0 - \mu)}{\sigma^2} - \sqrt{\frac{(u_0-\mu)^2}{\sigma^4} + \frac{2\delta}{\sigma^2}} > \frac{\delta}{\mu - u_0}$$

From this we see that

$$\frac{1}{\theta_1} - \frac{1}{\theta_4} > \frac{\mu}{\delta} + \frac{u_0 - \mu}{\delta} = \frac{u_0}{\delta}$$

which shows that $A\theta_1 < 1$, hence $x_0 > 0$. Therefore, f is an increasing, bounded and positive solution to the HJB equation. From Theorem 3.2 it then follows that

$$f(x) = V(x),$$

this completes the proof. $\square$

3.2 The Cramér-Lundberg model

We begin by examining the Cramér-Lundberg model of risk reserve $R(t)$. This model was first introduced by Swedish actuary Filip Lundberg in 1903, and forms the backbone of modern non-life insurance mathematics. We will in this section use the probability space $(\Omega, \mathcal{F}, \{\mathcal{F}(t)\}_{t\geq 0}, \mathbb{P})$ on which all stochastic quantities are defined. We will now give this model a formal definition. The following definition can be found in [12, p.18] and [14, p.220].

Definition 3.3 (Cramér-Lundberg model). *The risk reserve $\{R(t)\}_{t\geq 0}$ is given by*

$$R(t) = x + ct - \sum_{k=1}^{N(t)} Y_k,$$

where c is the premium rate, x is the initial capital and $\{N(t)\}_{t\geq 0}$ is a Poisson process with intensity λ, and the claim occurrence times are denoted by $T_1 < T_2 \ldots ..$, and $T_0 = 0$. The claim sizes are denoted by Y_k has a distribution function denoted by $G(y)$ with $G(0) = 0$. Furthermore, $N, Y_1, Y_2, \ldots$ are independent.

For further details on the Cramér-Lundberg model, see for example [12, p.13-21].
We will now introduce a diffusion approximation of the Cramér-Lundberg model i.e. approximating $R(t)$ by a Brownian motion with drift. This is a necessary step in order to replicate the pattern of the results presented in the previous subsection. We need the Cramér-Lundberg model to be in the form of a Brownian motion with drift . The general idea behind this approximation is to define a sequence of reserve processes as is defined in Definition 3.3 which converge weakly to a Brownian motion with drift. We will not dwell on the technicalities of this approximation, for further information see[1, p.4]. The following result can be found in [3, p.355-356].

Lemma 3.2 (Cramér-Lundberg diffusion approximation). *Given the assumptions of independence defined in Definition 3.3, and some technical conditions, then the risk process* $\{R(t)\}_{t\geq 0}$ *can be approximated as follows*

$$dX(t) = (c - \lambda\mu_G)dt + \sqrt{\lambda\mu_G^{(2)}}dW(t),$$

where μ_G *and* $\mu_G^{(2)}$ *are the first and second moment of* $G(y)$.

3.3 The Cramér-Lundberg model and optimal dividends

We begin by considering the restricted optimal dividend problem for the Cramér-Lundberg model. That is, we want to maximise total discounted dividends given the reserve dynamics described by the Cramér-Lundberg model. We can consider this problem in two ways:

1. We assume that the reserve $R(t)$ as defined in definition 3.3 can be approximated as a diffusion given by Lemma 3.2.
2. We consider the problem directly without approximating $R(t)$.

In the first case we consider the case where the basic idea is to define a sequence of reserve processes which converge weakly to a Brownian motion with drift as in Lemma 3.2. We can then directly use the results given by Theorem 3.3 and Theorem 3.4 in Section 3.1. The only difference being that we change the parameters in Definition 3.1 to $\mu = c - \lambda\mu_G$ and $\sigma = \sqrt{\lambda\mu_G^{(2)}}$. That is, we consider the controlled reserve process as the following

$$dX^U(t) = (c - \lambda\mu_G - u(X^U(t)))dt + \sqrt{\lambda\mu_G^{(2)}}dW(t),$$

where the parameters are given as in Definition 3.3. If we assume that $c > \lambda\mu_G$, then we see that we have the same optimal value function for the diffusion approximation as in Theorem 3.3.

$$V(x) = \frac{u_0}{\delta}(1 - e^{\theta_4 x}), \quad \text{if } u_0 \leq -\delta/\theta_4,$$

where θ_4 is given by

$$\theta_4 = \frac{u_0 - (c - \lambda\mu_G)}{\sqrt{\lambda\mu_G^{(2)}}} - \sqrt{\frac{(u_0 - (c - \lambda\mu_G))^2}{\lambda\mu_G^{(2)}} + \frac{u_0}{\delta}},$$

The roots θ_1, θ_2 and θ_3 are given by the same argument i.e. by changing the parameters. Furthermore, in Theorem 3.4 and its associated assumptions once again we have the same optimal value functions, that is:

$$V(x) = \begin{cases} K_1(e^{\theta_1 x} - e^{\theta_2 x}) & \text{for } 0 \leq x \leq x_0, \\ u_0/\delta - K_2 e^{\theta_4 x} & \text{for } x > x_0, \end{cases}, \quad \text{if } u_0 > -\delta/\theta_4.$$

The second case is substantially more difficult. However, we can find explicit solution given certain assumptions [3, p.374, 375]. Moreover, the verification step in this case will rely on martingale arguments [14, p.VIII], in contrast to the previous section where we used Ito's lemma.

We begin by giving a formal definition to the version of the optimal dividend problem where we do not approximate The Cramér-Lundberg model . This definition is inspired by [14, p.69], and uses notation from [5, p.287]. As before, the set of admissible strategies is constrained to the interval $[0, u_0]$.

Definition 3.4 (Restricted optimal dividend problem:Cramér-Lundberg model)**.**

$$
\begin{aligned}
\max_{U(\cdot)} \qquad & V^U(x) = \mathbb{E}\left[\int_0^\tau e^{-\delta t} U(t) dt\right] \\
\text{subject to} \qquad & X^U(t) = x + ct - \sum_{k=1}^{N(t)} Y_k - \int_0^t U(s) ds, \\
& X(0) = x, \\
& U(t) \in [0, u_0],\ u_0 < \infty,\ \forall\, t > 0,
\end{aligned}
$$

where the coefficients are the same as in the Cramér-Lundberg model defined in Definition 3.3. Furthermore, $\{U(t)\}$ denotes the dividend rate process which is adapted and u_0 denotes the maximum dividends rate.

We will now make a assumption which will simplify the arguments somewhat, in particular with respect to differentiability, see [14, p.71].

Assumption 3. *We will assume that the dividend rate u_0 is less than the premium rate c, i.e. $u_0 < c$.*

As before we have that the optimal value function given by:

$$V(x) = \sup_{U(t) \in [0, u_0]} V^U(x).$$

We now want to demonstrate that the optimal value function has the desirable properties of being continuous and bounded. Before formulating the lemma, we need to define Lipschitz continuity, the following definition is constructed from [7, p.4].

Definition 3.5 (Lipschitz continuity)**.** *A function $F : \mathbb{R} \to \mathbb{R}$ is said to be Lipschitz continuous if there exists a positive constant K such that*

$$|F(x) - F(y)| \leq K|x - y|, \quad \forall x, y \in \mathbb{R},$$

where K is known as the Lipschitz constant.

Lemma 3.3. The optimal value function $V(x)$ is increasing, Lipschitz continuous. Moreover, $V(x)$ is bounded i.e.

$$V(x) \leq \frac{u_0}{\delta}.$$

Proof. The proof that $V(x)$ is bounded is identical to the argument in Section 3.1. The rest of the proof can be found in [14, p.70-71]. $\square$

We will now state the Hamilton-Jacobi-Bellman equation for the problem, which will be necessary in order to derive the optimal value function. However, from Assumption 3, we will only consider the case when $u_0 < c$ which implies that the optimal value function is continuously differentiable [14, p.71].

Theorem 3.5. *If optimal value function $V(x)$ is continuously differentiable, then $V(x)$ satisfies the Hamilton-Jacobi-Bellman equation:*

$$\sup_{0\leq u\leq u_0}\left[(c-u)V'(x)+\lambda\left[\int_0^x V(x-y)\mathrm{d}G(y)-V(x)\right]-\delta V(x)+u\right]=0, \tag{3.4}$$

where $G(y)$ with $G(0)=0$ is the cumulative distribution function of Y i.e the claim size distribution.

Proof. The proof is omitted, it can be found in [14, p.71-74]. $\square$

As in Section 3.1, note that u is equal to either $u=u_0$ when $V'(x)<1$ or $u=0$ when $V'(x)>1$ [14, p.74]. This follows from the fact that the HJB equation is linear in u. We hence get the following equations:

$$0=\begin{cases} cV'(x)+\lambda\left[\int_0^x V(x-y)\mathrm{d}G(y)-V(x)\right]-\delta V(x), & V'(x)>1\\ (c-u_0)V'(x)+\lambda\left[\int_0^x V(x-y)\mathrm{d}G(y)-V(x)\right]-\delta V(x)+u_0, & V'(x)<1. \end{cases}$$

We will now need to proceed with the verification step. That is, we want to show that if $f(x)$ is a solution to the HJB equation then $f(x)$ is the optimal value function i.e. $f(x)=V(x)$. The proof will be a heuristic proof, for a more rigorous proof see [14, p.75-76]. Moreover, this proof will utilise martingale arguments, this distinguishes this verification theorem from Theorem 3.2.

Theorem 3.6 (Verification theorem)**.** *If $f(x)$ is a increasing, bounded and positive solution to the HJB equation:*

$$\sup_{0\leq u\leq u_0}\left[(c-u)f'(x)+\lambda\left[\int_0^x f(x-y)\mathrm{d}G(y)-f(x)\right]-\delta f(x)+u\right]=0,$$

then $f(x)=V(x)$, and an optimal control law is given by:

$$U^*(x)=\begin{cases} 0, & \textit{for } V'(x-)>1\\ u_0, & \textit{for } V'(x-)<1 \end{cases}$$

Heuristic Proof. Let U be an arbitrary control law. From [14, p.75], we have that the process $\{M'(t)\}_{t\geq 0}$ with

$$M'(t)=\sum_{i=1}^{N(\tau\wedge t)}(f(X^U(T_i))-f(X^U(T_i-)))e^{-\delta T_i}$$
$$-\lambda\int_0^{\tau\wedge t}e^{-\delta s}\left(\int_0^{X^U(s)}f(X^U(s)-y)dG(y)-f(X^U(s)\right)ds$$

is a martingale. From the fundamental theorem of calculus we see that:

$$f\left(X^U(T_i-)\right)\mathrm{e}^{-\delta T_i} - f\left(X^U(T_{i-1})\right)\mathrm{e}^{-\delta T_{i-1}} =$$
$$\int_{T_{i-1}}^{T_i-}\left[(c-U(s))\, f'\left(X^U(s)\right) - \delta f\left(X^U(s)\right)\right]\mathrm{e}^{-\delta s}\ \mathrm{d}s.$$

We can rewrite this as

$$f\left(X^U(T_i-)\right)\mathrm{e}^{-\delta T_i} - f\left(X^U(T_{i-1})\right)\mathrm{e}^{-\delta T_{i-1}} =$$
$$\int_{T_{i-1}\wedge\tau\wedge t}^{(T_i\wedge\tau\wedge t)-}\left[(c-U(s))\, f'\left(X^U(s)\right) - \delta f\left(X^U(s)\right)\right]\mathrm{e}^{-\delta s}\ \mathrm{d}s.$$

If we consider the sum in the martingale above as a telescoping sum we can rewrite $\{M'(t)\}_{t\geq 0}$ as

$$M'(t) = f\left(X^U(\tau\wedge t)\right)\mathrm{e}^{-\delta(\tau\wedge t)} - \int_0^{\tau\wedge t}\Bigg((c-U(s))f'(X^U(s))$$
$$+\lambda\int_0^{X^U(s)} f(X^U(s)-y)\mathrm{d}G(y) - (\lambda+\delta)f(X^U(s))\Bigg)\mathrm{e}^{-\delta s}\ \mathrm{d}s,$$

If $U = U^*$ (where we denote $X^{U^*} = X^*$ and $V^{U^*} = V^*$) we know from the HJB equation and the optimal control law $U(x)$ as is given above that:

$$M'(t) = f\left(X^*(\tau^*\wedge t)\right)\mathrm{e}^{-\delta(\tau\wedge t)} + \int_0^{\tau\wedge t} U^*(s)\mathrm{e}^{-\delta s}\ \mathrm{d}s,$$

recall that $\{M'(t)\}_{t\geq 0}$ is a martingale. By taking the expected value:

$$\mathbb{E}\left[f\left(X^*(\tau\wedge t)\right)\mathrm{e}^{-\delta(\tau\wedge t)} + \int_0^{\tau\wedge t} U^*(s)\mathrm{e}^{-\delta s}\ \mathrm{d}s\right],$$

then we letting $t\to\infty$, which shows that

$$\lim_{t\to\infty}\mathbb{E}\left[f\left(X^*(\tau\wedge t)\right)\mathrm{e}^{-\delta(\tau\wedge t)}\right] = 0,$$

since f is a bounded function and $f(X(\tau)) = 0$, as in Theorem 3.2, this follows from Theorem 2.1. Hence, we have that:

$$f(x) = \mathbb{E}\left[\int_0^{\tau\wedge t} U^*(s)\mathrm{e}^{-\delta s}\ \mathrm{d}s\right] = V^{U^*}(x).$$

Moreover , for any arbitrary control law U we have that

$$f(x) \geq \mathbb{E}\left[f\left(X^U(\tau\wedge t)\right)\mathrm{e}^{-\delta(\tau\wedge t)} + \int_0^{\tau\wedge t} U(s)\mathrm{e}^{-\delta s}\ \mathrm{d}s\right].$$

Again, if we let $t \to \infty$ we have that

$$f(x) \geq V^U(x).$$

Which given the optimal control law $U = U^*$ gives us

$$f(x) = V(x),$$

this completes the proof. □

Now that we have stated the HJB equation and proved the verification theorem, we can now proceed to solve the HJB equation, such that the solution has the required properties such as being bounded, increasing etc. For this type of problem there tend to be no explicit solutions. However, for the special case when the distribution of the claim sizes are exponential, one can find explicit solutions for this optimal control problem.

We will now solve the HJB equation. We will do this by formulating its solutions as a theorem. The proof will follow from Theorem 3.5 and the verification theorem Theorem 3.6, with the addition of solving the HJB equation for the special case. The proof relies primarily on the following sources [14, p.77-79], and [3, p. 381].

Theorem 3.7. *If $G(y) = 1 - e^{-\alpha y}$, i.e. the claim sizes are exponentially distributed $Y_i \sim Exp(\alpha)$, and if f is an increasing bounded and positive solution to HJB equation*

$$0 = \begin{cases} cf''(x) - (\lambda + \delta - \alpha c)f'(x) - \alpha\delta f(x), & f'(x) > 1 \\ (c - u_0)f''(x) - (\lambda + \delta - \alpha(c - u_0))f'(x) - \alpha\delta f(x) + \alpha u_0, & f'(x) < 1, \end{cases}$$

then $f(x) = V(x)$ and the optimal value function is given by:

$$V(x) = \begin{cases} \dfrac{e^{\theta_1 x} - \frac{(\lambda+\delta)/c - \theta_1}{(\lambda+\delta)/c - \theta_2} e^{\theta_2 x}}{\theta_1 e^{\theta_1 a} - \frac{(\lambda+\delta)/c-\theta_1}{(\lambda+\delta)/c-\theta_2}\theta_2 e^{\theta_2 a}} & \text{for } x < a \\ \frac{u_0}{\delta} + \frac{1}{\theta_4} e^{\theta_4 (x-a)}, & \text{for } x \geq a, \end{cases}$$

where,

$$a = \frac{1}{\theta_1 - \theta_2} \log\left[\frac{(\lambda+\delta)/c - \theta_1}{(\lambda+\delta)/c - \theta_2}\left(\frac{1 - \theta_2 S}{1 - \theta_1 S}\right)\right], \quad S = \frac{u_0}{\delta} + \frac{1}{\theta_4}$$

where the roots to the characteristic polynomials are given by:

$$\theta_{1,2} = \frac{\lambda + \delta - \alpha c}{2c} \pm \sqrt{\left(\frac{\lambda + \delta - \alpha c}{2c}\right)^2 + \frac{\alpha\delta}{c}},$$

$$\theta_{3,4} = \frac{\lambda + \delta - \alpha(c - u_0)}{2(c - u_0)} \pm \sqrt{\left(\frac{\lambda + \delta - \alpha(c - u_0)}{2(c - u_0)}\right)^2 + \frac{\alpha\delta}{c - u_0}}.$$

Proof. From Theorem 3.5 we attain the HJB equation $\sup_{0\leq u\leq u_0} H(x,u) = 0$, given that $G(y) = 1 - e^{-\alpha y}$, where

$$H(x,u) = (c-u)f'(x) + \lambda\alpha e^{-\alpha x}\int_0^x f(y)e^{\alpha y}dy - \lambda f(x) - \delta f(x) + u$$

We can see that if $f'(x) < 1$ then $u = u_0$ and if $f'(x) > 1$ then $u = 0$, that is, the same argument we used in Section 3.1. For this case we get:

$$\begin{aligned}
\frac{d}{dx}\lambda\alpha e^{-\alpha x}\int_0^x f(y)e^{\alpha(y)}dy &= \lambda\alpha f(x) - \lambda\alpha^2 e^{-\alpha x}\int_0^x f(y)e^{\alpha y}dy \\
&= \lambda\alpha f(x) + \alpha\left[(c-u)f'(x) - \lambda f(x) - \delta f(x) + u\right]. \\
\frac{d}{dx}H(x,u) = (c-u)f''(x) &+ \alpha\left[(c-u)f'(x) - \delta f(x) + u\right] - (\lambda+\delta)f'(x).
\end{aligned}$$

In the first step we use the product rule, and in the second step we use the equality from the HJB equation. By rearranging terms for the cases where $f'(x) > 1$ and $f'(x) < 1$ we get the following ordinary differential equations:

$$0 = \begin{cases} cf''(x) - (\lambda+\delta-\alpha c)f'(x) - \alpha\delta f(x), & f'(x) > 1 \\ (c-u_0)f''(x) - (\lambda+\delta-\alpha(c-u_0))f'(x) - \alpha\delta f(x) + \alpha u_0, & f'(x) < 1. \end{cases}$$

Where the roots to their characteristic polynomials are:

$$\begin{aligned}
\theta_{1,2} &= \frac{\lambda+\delta-\alpha c}{2c} \pm \sqrt{\left(\frac{\lambda+\delta-\alpha c}{2c}\right)^2 + \frac{\alpha\delta}{c}}, \\
\theta_{3,4} &= \frac{\lambda+\delta-\alpha(c-u_0)}{2(c-u_0)} \pm \sqrt{\left(\frac{\lambda+\delta-\alpha(c-u_0)}{2(c-u_0)}\right)^2 + \frac{\alpha\delta}{c-u_0}}.
\end{aligned}$$

We can see that $\theta_2 < 0 < \theta_1$ and $\theta_4 < 0 < \theta_3$. The solutions to the ordinary differential equations are the following:

$$f(x) = \begin{cases} A\mathrm{e}^{\theta_1 x} + B\mathrm{e}^{\theta_2 x} & f'(x) > 1 \\ \frac{u_0}{\delta} + C\mathrm{e}^{\theta_3 x} + D\mathrm{e}^{\theta_4 x}, & f'(x) < 1. \end{cases}$$

As in the proof of Theorem 3.4 we see that we must have that $C = 0$, since otherwise the function f would not be bounded. Moreover, if we plug in $f(x)$ into the original HJB equation $H(0,0) = 0$, we then get

$$B = -\frac{(\lambda+\delta)/c - \theta_1}{(\lambda+\delta)/c - \theta_2}A.$$

Now, consider a point a where $f'(a) = 1$, that is, we get equations:

$$\begin{aligned}
&A\theta_1 e^{\theta_1 a} + B\theta_2 e^{\theta_2 a} = 1, \\
&D\theta_4 e^{\theta_4 a} = 1.
\end{aligned}$$

Hence, we get $D = e^{-\theta_4 a}/\theta_4$, and

$$A = \frac{1}{\theta_1 e^{\theta_1 a} - \frac{(\lambda+\delta)/c-\theta_1}{(\lambda+\delta)/c-\theta_2}\theta_2 e^{\theta_2 a}}$$

As in Theorem 3.4, we can see that $f(x)$ is concave. Because $f(x)$ is concave and differentiable it holds that there exists exactly one point a such that $f'(a) = 1$. Differentiability follows in a similar way as in Theorem 3.4. Hence we get:

$$f(x) = \begin{cases} \frac{e^{\theta_1 x} - \frac{(\lambda+\delta)/c-\theta_1}{(\lambda+\delta)/c-\theta_2} e^{\theta_2 a}}{\theta_1 e^{\theta_1 x} - \frac{(\lambda+\delta)/c-\theta_1}{(\lambda+\delta)/c-\theta_2}\theta_2 e^{\theta_2 a}} & \text{for } x < a \\ \frac{u_0}{\delta} + \frac{1}{\theta_4} e^{\theta_4 (x-a)}, & \text{for } x \geq a. \end{cases}$$

By solving $f'(a-) = f'(a+)$ for a, i.e.

$$\frac{e^{\theta_1 a} - \frac{(\lambda+\delta)/c-\theta_1}{(\lambda+\delta)/c-\theta_2} e^{\theta_2 a}}{\theta_1 e^{\theta_1 a} - \frac{(\lambda+\delta)/c-\theta_1}{(\lambda+\delta)/c-\theta_2}\theta_2 e^{\theta_2 a}} = \frac{u_0}{\delta} + \frac{1}{\theta_4},$$

we get:

$$a = \frac{1}{\theta_1 - \theta_2} \log \left[\frac{(\lambda+\delta)/c - \theta_1}{(\lambda+\delta)/c - \theta_2} \left(\frac{1 - \theta_2 S}{1 - \theta_1 S} \right) \right].$$

Where $S = \frac{u_0}{\delta} + \frac{1}{\theta_4}$, in particular, as in Theorem 3.4, we will have that $a > 0$ if $\theta_1 S < 1$ and $\theta_1 < (\lambda + \delta)/c$. It then follows from Theorem 3.6 that $f(x) = V(x)$. This concludes the proof.

□

Having investigated the case for restricted dividends we will now investigate the optimal dividend problem for the case when we have unrestricted dividends. For the unrestricted case i.e. when $0 \leq u(t) \leq \infty$. The unrestricted case is substantially more difficult. Hence, the analysis presented will be rather heuristic and not very rigorous.

The definition and the goal of the problem is however not very different. The results presented are influenced by [4, p.7-14] and [14, p.79-94].
We will now give a formal definition of the problem. The following definition is inspired by [14, p.79] and uses notation from [5, p.287].

Definition 3.6 (Unrestricted optimal dividend problem:Cramér-Lundberg model)**.**

$$\begin{aligned} \max_{D(\cdot)} \qquad & V^D(x) = \mathbb{E}\left[\int_{0-}^{\tau-} e^{-\delta t} dD(t)\right] \\ \text{subject to} \qquad & X^U(t) = x + ct - \sum_{k=1}^{N(t)} Y_k - D(t), \end{aligned}$$

where the coefficients are the same as in the Cramér-Lundberg model defined in the previous section. Furthermore, $\{D(t)\}_{t\geq 0}$ denotes the accumulated dividend process which is an increasing adapted Càdlàg process and τ is the ruin time. Moreover, $\{X^U(t)\}_{t\geq 0}$ denotes the controlled reserve process.

Remark 3.4. When dealing with the unrestricted optimal dividend problem we will not use Assumption 3, for reasons that we become apparent in Theorem 3.8.

The idea behind letting the dividend process D being Càdlàg is to prevent the possibility of paying out a large amount leading to ruin. We prevent this by not taking dividends at the time of ruin into account, that is, we view the process $\{X^U(t)\}_{t\geq 0}$ as the post-dividend process.
As in the previous section we define the optimal value function as

$$V(x) := \sup_{D\in\mathcal{U}} V^D(x),$$

where $\mathcal{U}$ is the set of admissible strategies. The intuition behind this problem is to consider the problem as defined in Definition 3.4 and letting $u_0 \to \infty$.
The optimal value function in the context of the unrestricted optimal dividend problem is obtained as a limit of the optimal value function for the restricted problem, see Definition 3.4. The following theorem demonstrates this fact [14, p.81].

Theorem 3.8. *If $V_u(x)$ is the optimal value function for the optimal dividend problem in Definition 3.4, when $u_0 = u$, then*

$$\lim_{u\to\infty} V_u(x) = V(x).$$

We can now state the Hamilton-Jacobi-Bellman equation for the unrestricted problem. The idea behind this theorem is to consider the HJB equation as in Theorem 3.5 and letting $u_0 \to \infty$ [14, p.82].

Theorem 3.9 (Hamilton-Jacobi-Bellman equation)**.** *If certain regularity assumptions are fulfilled, then the optimal value function $V(x)$ satisfies the Hamilton-Jacobi-Bellman equation:*

$$\max\left\{cV'(x) + \lambda\int_0^x V(x-y)\mathrm{d}G(y) - (\lambda+\delta)V(x), 1 - V'(x)\right\} = 0,$$

where $G(y)$ with $G(0)$ is the cumulative distribution function of Y.

Proof. The proof is rather complicated and is therefore omitted, a proof can be found [14, p.82-84]. □

When investigating the case when $x = 0$ we see that $cV'(0) = (\lambda+\delta)V(0)$ from inside the maximum operator in the HJB equation above. Moreover, for $V'(x) = 1$ we have that $V(0) = c/(\lambda+\delta)$.
We will now present a so called barrier strategy, The definition of which is derived from [14, p.91]. This strategy will later be applied in the end of this section when deriving the optimal value functions when the claim sizes are exponentially distributed.

Definition 3.7 (Barrier strategy)**.** *A strategy is called a barrier strategy at x_0 if $D(0) = (x-x_0)^+$and $\Delta D(t) = c\mathbb{I}_{X(t)=x_0}$.*

Remark 3.5. The notation used here is defined as follows $x^+ = \max\{x, 0\}$.

We will now investigate a verification theorem for the barrier strategy. The following result can found in[14, p.86-87, 91-92].

Theorem 3.10. *Let $f(x)$ be the solution to the equation*

$$cf'(x) + \lambda \int_0^x f(x-y)\mathrm{d}G(y) - (\lambda+\delta)f(x) = 0$$

with $f(0) = 1$, the optimal value functions of the barrier strategy at x_0 is then given by

$$V_{x_0}(x) = \frac{f(x)}{f'(x_0)}, \quad \textit{if } x \le x_0$$
$$V_{x_0}(x) = \frac{f(x_0)}{f'(x_0)} + x - x_0, \quad \textit{if } x > x_0$$

Proof. The proof relies on similar martingale arguments as in Theorem 3.6, the proof can be found in [14, p.92] □

Remark 3.6 (Construction of the optimal value function)**.** Note that one can construct a optimal value function $V(x)$ in the context of Theorem 3.8 from:

$$V_{x_0}(x) = \frac{f(x)}{f'(x_0)}, \quad \text{if } x \le x_0$$
$$V_{x_0}(x) = \frac{f(x_0)}{f'(x_0)} + x - x_0, \quad \text{if } x > x_0.$$

Where then

$$V(x) = \begin{cases} \frac{f(x)}{f'(a)}, & \text{for } x \le a \\ \frac{f(a)}{f'(a)} + x - a, & \text{for } x > a. \end{cases}$$

Where a is the barrier. For further details, see [14, p.92].

We are now ready to solve the Hamilton-Jacobi-Bellman equation under the assumption that the claim sizes are exponentially distributed. Given the solution to HJB it follows from the verification theorems that the solution is the optimal value function. We will get different solutions depending on the barrier and the parameters connected to the Cramér–Lundberg model. We will formulate this as a theorem. The proof is inspired by [14, p.93-94]. We will now rely on Definition 3.6, Theorem 3.8-3.10, Definition 3.7, and Remark 3.6.

Theorem 3.11. *If $G(y) = 1 - e^{-\alpha y}$, i.e. the claim sizes are exponentially distributed $Y_i \sim Exp(\alpha)$, and if f solves*

$$\max\left\{cf'(x) + \lambda \int_0^x f(x-y)\mathrm{d}G(y) - (\lambda+\delta)f(x), 1 - f'(x)\right\} = 0,$$

then $f(x) = V(x)$, i.e. the barrier strategy at a then yields:

$$V(x) = \begin{cases} \frac{e^{\theta_1 x} - \frac{(\lambda+\delta)/c-\theta_1}{(\lambda+\delta)/c-\theta_2} e^{\theta_2 x}}{\theta_1 e^{\theta_1 a} - \frac{(\lambda+\delta)/c-\theta_1}{(\lambda+\delta)/c-\theta_2}\theta_2 e^{\theta_2 a}}, & x < a \\ \frac{\alpha c - \lambda - \delta}{\alpha\delta} + x - a, & x \geq a. \end{cases}$$

where

$$\theta_{1,2} = \frac{\lambda + \delta - \alpha c}{c} \pm \sqrt{\left(\frac{\lambda+\delta-\alpha c}{c}\right)^2 + \frac{\alpha\delta}{c}},$$
$$a = \frac{1}{\theta_1 - \theta_2} \log\left[\frac{(\lambda+\delta)/c - \theta_1}{(\lambda+\delta)/c-\theta_2}\left(\frac{1-\theta_2 K}{1-\theta_1 K}\right)\right],$$

and $K = \frac{\alpha c - \lambda - \delta}{\alpha\delta}$.

Proof. If we consider the HJB equation from Theorem 3.9:

$$\max\left\{cf'(x) + \lambda\int_0^x f(x-y)\mathrm{d}G(y) - (\lambda+\delta)f(x), 1 - f'(x)\right\} = 0.$$

We see that if $f'(x) > 1$ we get the HJB equation:

$$cf'(x) + \lambda\int_0^x f(x-y)\mathrm{d}G(y) - (\lambda+\delta)f(x) = 0,$$

Moreover, as in Theorem 3.7 we have an a such that $f(x) > 1$ for $x < a$ and $f'(a) = 1$. In particular, if $G(y) = 1 - e^{-\alpha y}$ we get the HJB equation:

$$cf'(x) + \lambda\alpha e^{-\alpha x}\int_0^x f(y)e^{\alpha y}dy - (\lambda+\delta)f(x) = 0. \tag{3.6}$$

We can rewrite this in a similar way as in the proof of Theorem 3.7:

$$cf''(x) - (\lambda + \delta - \alpha c)f'(x) - \alpha\delta f(x) = 0.$$

The solution to this homogeneous ODE is given by

$$f(x) = C_1 \mathrm{e}^{\theta_1 x} + C_2\mathrm{e}^{\theta_2 x}.$$

Hence, we get the same solution as in Theorem 3.7 when $f(x) > 1$, i.e.

$$f(x) = \frac{e^{\theta_1 x} - \frac{(\lambda+\delta)/c-\theta_1}{(\lambda+\delta)/c-\theta_2} e^{\theta_2 x}}{\theta_1 e^{\theta_1 a} - \frac{(\lambda+\delta)/c-\theta_1}{(\lambda+\delta)/c-\theta_2}\theta_2 e^{\theta_2 a}}$$

If rearrange Equation 3.6 when $f'(x) = 1$, then we define $v(x) = f(x)$, we see that:

$$v(x) = \frac{c}{\lambda+\delta} + \frac{\lambda\alpha}{\lambda+\delta}e^{-\alpha x}\int_0^x f(y)e^{\alpha y}dy,$$

by taking the derivative we see that

$$1 = -\frac{\lambda\alpha^2}{\lambda+\delta}e^{-\alpha x}\int_0^x f(x)e^{\alpha x} + \frac{\lambda\alpha}{\lambda+\delta}f(x).$$

Moreover, from above we have that

$$\frac{\lambda\alpha}{\lambda+\delta}e^{-\alpha x}\int_0^x f(y)e^{\alpha y}dy = v(x) - \frac{c}{\lambda+\delta}.$$

Hence, we have that

$$1 = -\alpha\left(f(x) - \frac{c}{\lambda+\delta}\right) + \frac{\lambda\alpha}{\lambda+\delta}f(x),$$

which implies that

$$f(a) = \frac{\alpha c - \lambda - \delta}{\alpha\delta}.$$

By solving

$$\frac{e^{\theta_1 a} - \frac{(\lambda+\delta)/c-\theta_1}{(\lambda+\delta)/c-\theta_2}e^{\theta_2 a}}{\theta_1 e^{\theta_1 a} - \frac{(\lambda+\delta)/c-\theta_1}{(\lambda+\delta)/c-\theta_2}\theta_2 e^{\theta_2 a}} = \frac{\alpha c - \lambda - \delta}{\alpha\delta}$$

for a, we get

$$a = \frac{1}{\theta_1 - \theta_2}\log\left[\frac{(\lambda+\delta)/c-\theta_1}{(\lambda+\delta)/c-\theta_2}\left(\frac{1-\theta_2 K}{1-\theta_1 K}\right)\right],$$

where $K = \frac{\alpha c-\lambda-\delta}{\alpha\delta}$. Having solved the HJB equation and calculated $f(a)$, it follows from Theorem 3.10 that:

$$V_a(x) = \frac{f(x)}{f'(a)} = \frac{e^{\theta_1 x} - \frac{(\lambda+\delta)/c-\theta_1}{(\lambda+\delta)/c-\theta_2}e^{\theta_2 x}}{\theta_1 e^{\theta_1 a} - \frac{(\lambda+\delta)/c-\theta_1}{(\lambda+\delta)/c-\theta_2}\theta_2 e^{\theta_2 a}} \quad x < a,$$

$$V_a(x) = \frac{f(a)}{f'(a)} + x - a = \frac{\alpha c - \lambda - \delta}{\alpha\delta} + x - a, \quad x > a.$$

From Remark 3.6 it then follows that

$$V(x) = \begin{cases} \frac{e^{\theta_1 x} - \frac{(\lambda+\delta)/c-\theta_1}{(\lambda+\delta)/c-\theta_2}e^{\theta_2 x}}{\theta_1 e^{\theta_1 a} - \frac{(\lambda+\delta)/c-\theta_1}{(\lambda+\delta)/c-\theta_2}\theta_2 e^{\theta_2 a}}, & x < a \\ \frac{\alpha c-\lambda-\delta}{\alpha\delta} + x - a, & x \geq a. \end{cases}$$

This proves the theorem. $\square$

4 Concluding remarks

The argument which has been constant throughout this thesis is the general logic behind the optimal control problem. That is, the following logical structure:

1. Formulate the stochastic control problem i.e. the optimal dividend problem;
2. Derive the HJB equation i.e. we want to show the following: **if** V is sufficiently regular and **if** V is the optimal value function **then** V satisfies the HJB equation;
3. Derive the verification theorem, i.e. we want show that **if** f is a solution to the HJB equation **then** $f = V$;
4. Solve the HJB equation, whose solution we then know through the verification theorem is the optimal value function.

Both the restricted and the unrestricted problems we have studied in this thesis follow this logic. In providing evidence for this claim, we have relied on ad hoc regularity assumptions relating to the optimal value function. This might be viewed as a limitation, since these assumptions could be too constricting. However, general results regarding optimal dividends might still be of importance. Moreover, in the case of the Cramér–Lundberg model we can only find explicit solutions in the case of exponentially distributed claim sizes, or at least with sufficiently differentiable claim size distributions. However, this is quite an unrealistic assumption for claim size distributions.

References

[1] Hansjörg Albrecher and Stefan Thonhauser. Optimality results for dividend problems in insurance. *RACSAM-Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A. Matematicas*, 103(2):295–320, 2009.

[2] Søren Asmussen, Bjarne Højgaard, and Michael Taksar. Optimal risk control and dividend distribution policies. example of excess-of loss reinsurance for an insurance corporation. *Finance and Stochastics*, 4(3):299–324, 2000.

[3] Søren Asmussen and Mogens Steffensen. *Risk and Insurance: A Graduate Text*. Springer, 2020.

[4] Søren Asmussen and Michael Taksar. Controlled diffusion models for optimal dividend pay-out. *Insurance: Mathematics and Economics*, 20(1):1–15, 1997.

[5] Tomas Björk. *Arbitrage theory in continuous time*. Oxford university press, 2009.

[6] Tomas Björk. Stochastic optimal control with finance applications. `https://www.math.kth.se/matstat/seminarier/Tomas2015/KTH2015.pdf`, 2017 (accessed April 19, 2021).

[7] Michael Björklund. Existens och entydighet för ordinära differentialekvationer. `https://people.kth.se/~tranberg/5B1200.existens.pdf`, 2017 (accessed April 19, 2021).

[8] Simone Calogero. Stochastic calculus, financial derivatives and pde's. `http://www.math.chalmers.se/~calogero/Lecture_Notes11.pdf`, 2017 (accessed April 19, 2021).

[9] Allan Gut. *Probability: a graduate course*, volume 75. Springer Science & Business Media, 2013.

[10] Monique Jeanblanc-Picqué and Albert Nikolaevich Shiryaev. Optimization of the flow of dividends. *Uspekhi Matematicheskikh Nauk*, 50(2):25–46, 1995.

[11] Fima C Klebaner. *Introduction to stochastic calculus with applications*. World Scientific Publishing Company, 2005.

[12] T Mikosch. Non-life insurance mathematics: An introduction with the poisson processes, springer verlag. 2009.

[13] Bernt Oksendal. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2013.

[14] Hanspeter Schmidli. *Stochastic control in insurance*. Springer Science & Business Media, 2007.

[15] Steven E Shreve. *Stochastic calculus for finance II: Continuous-time models*, volume 11. Springer Science & Business Media, 2004.