TECHNISCHE UNIVERSITÄT WIEN

D I P L O M A R B E I T

Optimal Dividend Strategies of Two Collaborating Businesses

zur Erlangung des akademischen Grades

Diplom-Ingenieurin

im Rahmen des Studiums

Finanz- und Versicherungsmathematik

eingereicht von

Verena Köck
Matrikelnummer 01226270

ausgeführt am Institut für Wirtschaftsmathematik
der Fakultät für Mathematik und Geoinformation der Technischen Universität Wien

Betreuer:
Ao.Univ.Prof. Dr. Peter Grandits

Wien, 07.11.2018

Unterschrift Verfasserin

Unterschrift Betreuer

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Diplomarbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe.

Wien, am 07.11.2018.

Verena Köck

Einverständniserklärung zur Plagiatsprüfung

Ich nehme zur Kenntnis, dass die vorgelegte Arbeit mit geeigneten und dem derzeitigen Stand der Technik entsprechenden Mitteln (Plagiat-Erkennungssoftware) elektronisch- technisch überprüft wird. Dies stellt einerseits sicher, dass bei der Erstellung der vorgelegten Arbeit die hohen Qualitätsvorgaben im Rahmen der ausgegebenen der an der TU Wien geltenden Regeln zur Sicherung guter wissenschaftlicher Praxis - "Code of Conduct" (Mitteilungsblatt 2007, 26. Stück, Nr. 257 idgF.) an der TU Wien eingehalten wurden. Zum anderen werden durch einen Abgleich mit anderen studentischen Abschlussarbeiten Verletzungen meines persönlichen Urheberrechts vermieden.

Wien, am 07.11.2018.

Verena Köck

Abstract

In this diploma thesis the optimal dividend payment strategy problem is analysed. It is assumed that there are two collaborating insurance companies aiming to maximize their joint dividends to satisfy their shareholders. The surplus processes of both insurances are modelled as diffusions. The companies are allowed to transfer money freely without any costs. In the considered model a collaboration contract obligates both to cover losses the other one eventually experiences. As a consequence only simultaneous ruin is possible. In this work the optimal value function for the corresponding stochastic control problem is computed and an optimal strategy is constructed. Finally, a simulation study is implemented to compare the performance of this optimal strategy with other strategies.

Zusammenfassung

In dieser Diplomarbeit wird das optimale Dividendenproblem analysiert. Im betrachteten Modell maximieren zwei zusammenarbeitende Versicherungsunternehmen ihre gemeinsame Dividende, um ihre Aktionäre zufriedenzustellen. Der Überschussprozesse beider Versicherungen werden als Diffusion Prozesse modelliert. Beide Unternehmen können Gelder ohne Transaktionskosten an den jeweils anderen transferieren. Der Kooperationsvertrag beeinhaltet die Verpflichtung eventuelle Verluste des jeweils anderen auszugleichen. Als Konzequenz ist nur gemeinsamer Ruin möglich. In dieser Arbeit wird die optimale Wertfunktion des zugehörigen stochastischen Kontrollproblem hergeleitet und eine optimale Strategie konstruiert. Abschließend wird eine Simulationsstudie durchgeführt, um diese optimalen Strategie mit anderen Strategien zu vergleichen.

Contents

List of Figures

List of Tables

1 Introduction

This diploma thesis is concerned with optimal dividend payout, an optimization problem of control theory, which has been discussed from different aspects in the past. The question of interest is always the same: How to maximize dividend payments of a company to be paid to shareholders?

The classical optimal dividend problem goes back to Bruno De Finetti (1957) [8], who argued that calculating and minimizing the ruin probability is not sufficient. It is not realistic to assume that the surplus will grow indefinitely if ruin doesn't occur. Instead, a company should seek to maximize the expected present value of all dividends before ruin. The motivation for an insurance company is the following. Apart from the policyholders, who pay premiums to receive insurance coverage, shareholders are especially interested in a high dividend. As shareholders are inpatient, i.e. they prefer immediate payments over payments in the future, dividends shouldn't be held back longer than necessary. On the other hand high payments could result in a fast ruin and no dividends can be paid out at all. Thus, the insurance company has to find a compromise to optimize its dividend policy. Finetti showed, under the assumption that the surplus process follows a very simple model, that transferring all surplus above a certain barrier is optimal in this sense.

In 1969, Gerber [9] argued that the free surplus process of an insurance company can be modelled by a compound Poisson process. Asmussen and Taksar (1997) [3] modified the problem and computed the optimal dividend strategy in a diffusion approximation model. They considered two cases: First, when the dividend rate is restricted and the second, when the dividend rate is unrestricted. Gerber and Shiu (2006) [10] examined the analogous questions in the compound Poisson model. The book of Azcue and Muler (2014) [4] provides a summary of these stochastic control problems.

The above mentioned literature studied optimal dividend strategies only in the one-dimensional case, i.e. only one insurance company is involved and operates on its own. In 2017, Albrecher, Azcue and Muler [1] extended the one-dimensional problem to a two-dimensional setup of two collaborating companies under a compound Poisson model framework. Here, both aim to maximize their joint dividend and the collaboration involves balancing out the losses of the partner company with the own surplus.

Gu, Steffensen and Zheng (2017) [11] considered the two-dimensional model, when surplus processes are modelled as diffusion processes. The collaboration of two companies allows transferring money freely between the partners without transaction costs. If the surplus process of one company becomes negative, the other one is obligated to help and transfer money to cover the loss. In comparison to the two-dimensional compound Poisson model, here only simultaneous ruin is possible. In the Poisson model it can happen that one company cannot afford helping the other one, but continues alone while the other one goes ruin. As the Brownian motion model is continuous, this cannot happen here.

This diploma thesis mainly analyses the two-dimensional model in a diffusion framework as introduced by [11]. The structure is as follows. First, the Cramér Lundberg model is presented and the diffusion approximation is motivated. Second, an overview about control theory is given and some verification theorems are stated. The one-dimensional problem of [3] is explained for restricted as well as for unrestricted dividends. Then, the two-dimensional problem [11] is analysed in more detail. A short introduction in local time of Brownian motion and the reflected Brownian motion, which is related with the optimal dividend problem, is given. Finally, a simulation study is implemented to compare the optimal barrier strategy of the two-dimensional model with alternative strategies in numerical examples. Especially, it is analysed how the collaboration model performs in comparison to the case when both companies operate separately or if the present value of all dividends could be increased if a company decides to terminate collaboration.

2 Ruin theory

In this chapter the risk model, as described for example in Asmussen and Albrecher (2010) [2] or Azcue and Muler (2014) [4], is introduced and general properties are derived.

2.1 The Cramér Lundberg model

The *free surplus process* (or risk reserve process) $(X_t)_{t\geq 0}$ models the time evolution of the reserve of an insurance company. It is assumed that the portfolio of clients is fixed and that premiums are received at a constant rate p. The surplus is used to pay claims, which are of random sizes and occur at random times. The time and size of the ith claim is denoted by (τ_i, U_i) and N_t is defined as the number of claims up to time t, which is

$$N_t = \max\{i : \tau_i \leq t\}.$$

With x denoting the initial reserve, the risk process or surplus process of the insurance company at time t is defined as

$$X_t = x + pt - \sum_{i=1}^{N_t} U_i.$$

The *ruin time* is the time when the surplus of the company drops below zero the first time, which is

$$\tau = \inf\{t : X_t < 0\}.$$

The probability that ruin occurs given an initial surplus of x is the total ruin probability

$$\psi(x) = P(\tau < \infty | X_0 = x).$$

Then, the survival probability is

$$\delta(x) = 1 - \psi(x) = P(\tau = \infty | X_0 = x).$$

The following assumptions for the claim size distributions and the claim arrival times are made:

1. The first claim cannot occur at time zero, two claims cannot occur at the same time and the number of claims in any finite time interval is also finite.

2. The claim sizes are positive, mutually independent and identically distributed. They are also independent of the claim-arrival times and the number of claims N_t.

3. For the number of claims holds

$$P(N_{t+h} - N_t = k) = P(N_{s+h} - N_s = k) \qquad \text{for any } t, s \geq 0.$$

4. The number of claims in disjoint intervals are independent.

A model, which fulfils these assumptions, is called *Cramér-Lundberg model* or *classical collective risk model.* By Bühlmann (1970) [5, theorem 2] the last two assumptions imply that N_t is a Poisson process with intensity $\beta = E[N_1]$. This means that

$$P(N_{t+h} - N_t = k) = \frac{(\beta h)^k}{k!} e^{-\beta h}.$$

This implies that the arrival times τ_n can be expressed as a sum of exponentially distributed random variables Y_i with parameter β, i.e. $\tau_n = \sum_{i=1}^{n} Y_i$. The process $\sum_{i=1}^{N_t} U_i$ is a *compound Poisson process* and describes the total amount of claims up to time t. This model was introduced and developed by Lundberg (1903) [12] and Cramér (1930) [7]. It can be fully described by the premium rate p, the intensity β and the claim size distribution function $F(x) = P(U_i \leq x)$.

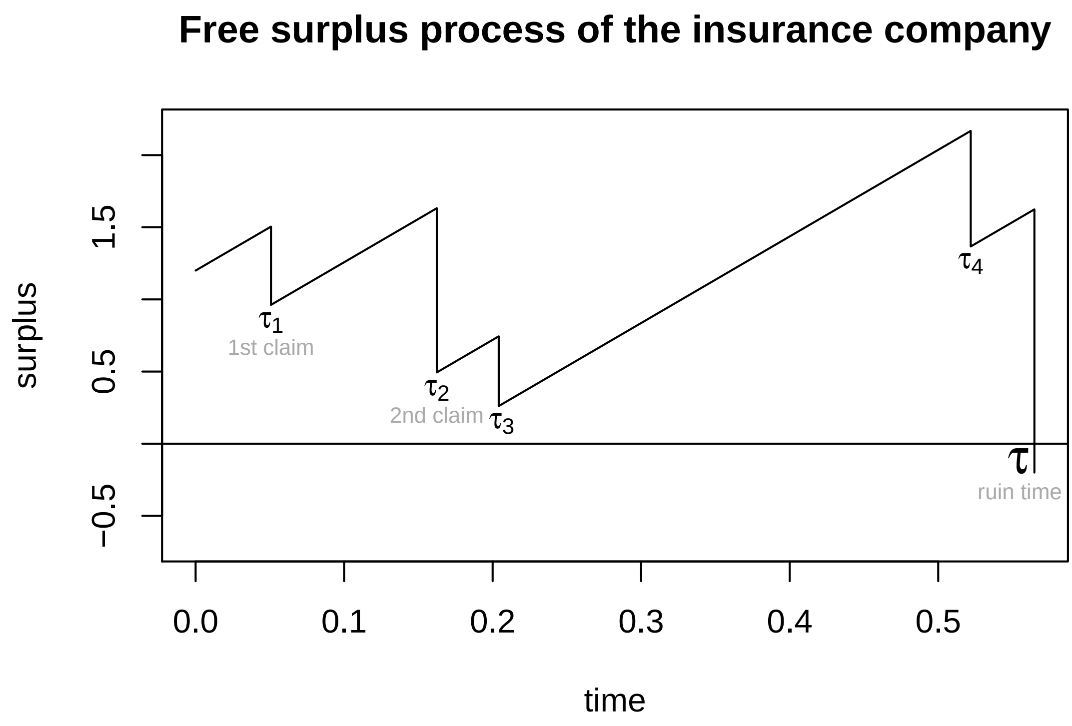

Figure 1: Free surplus process

The graph illustrates the free surplus process. At times τ_i claims occur and the surplus jumps down. These claim arrival times and the claim size (height of the jump) are random. If there is no claim the free surplus increases as premium is earned. After the 5th claim the surplus process drops below zero and therefore τ_5 is the ruin time.

2.2 Diffusion approximation of the Cramér Lundberg model

Under the assumption that a big portfolio is considered, the claim sizes are relatively small and arrive at high frequency. In this case the surplus process can be approximated with a Brownian motion with drift. The following plot gives a visual motivation of this idea. Here, a Cramér Lundberg risk model with relatively small claims arriving at high frequency is contrasted with a Brownian motion trajectory. One notices that both look very similar and that distinction is hardly possible.

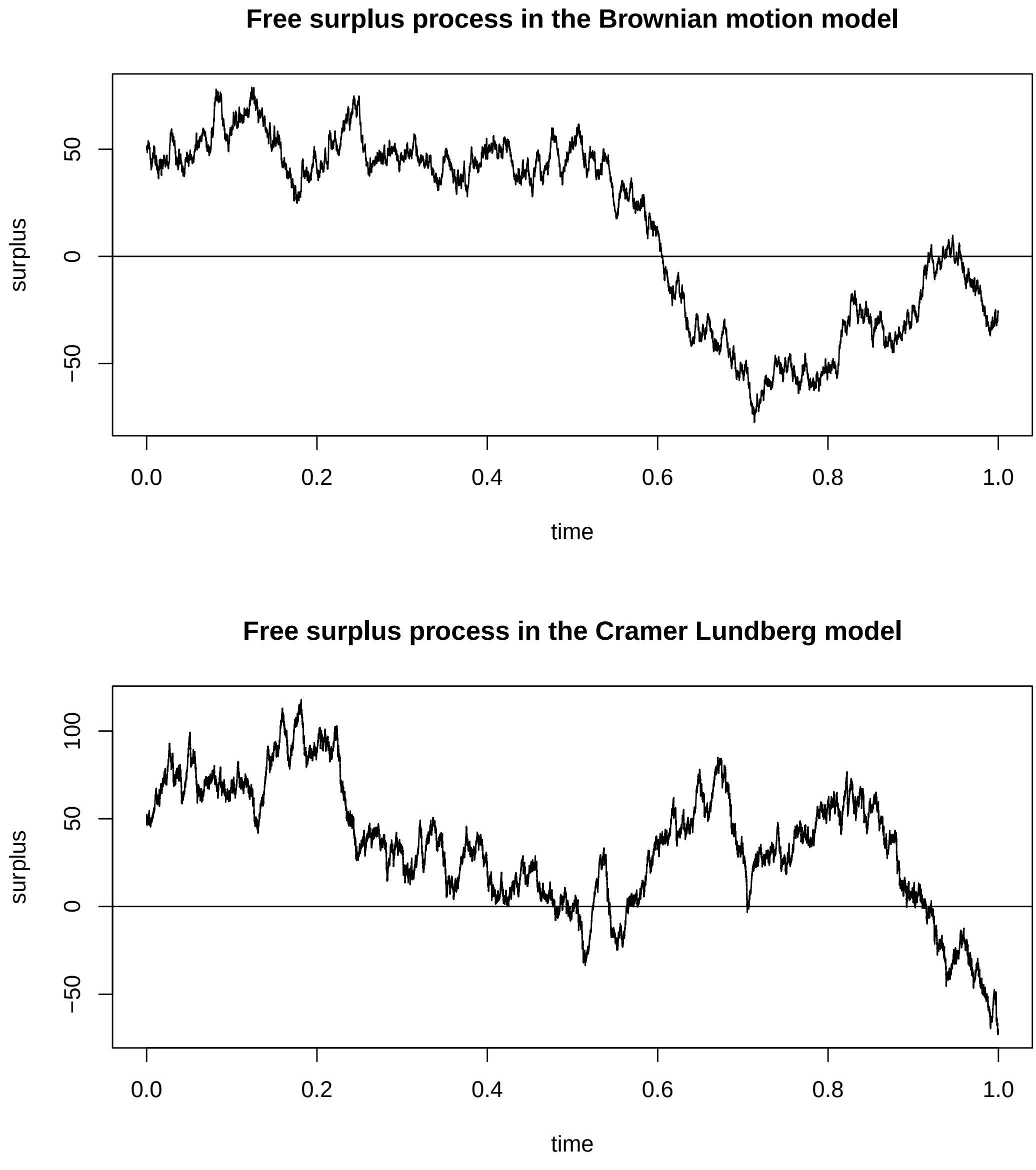

Figure 2: Cramér Lundberg versus Brownian motion

This approximation can be done by matching the first two moments, as shown in the book of Asmussen and Albrecher (2010) [2].

The approximation is based on Donsker's theorem (related to the classical central limit theorem) for a random walk $(S_n^*)_{n=0,1,\dots}$ in discrete time. If the drift $\mu = E[S_1^*]$ and variance $\sigma^2 = \mathrm{Var}[S_1^*]$, then

$$\left(\frac{1}{\sigma\sqrt{c}}(S^*_{\lfloor tc\rfloor} - tc\mu)\right)_{t\geq 0} \xrightarrow{\mathcal{D}} (W_0(t))_{t\geq 0}\,, \qquad c\to\infty, \tag{1}$$

where $(W_\zeta(t))$ describes a Brownian motion with drift ζ and variance 1 and $\xrightarrow{\mathcal{D}}$ means weak convergence.

Definition 2.1. A sequence $X_1, X_2, \dots$ of real-valued random variables converge weakly (denoted by $X_n \xrightarrow{\mathcal{D}} X$) if

$$\lim_{n\to\infty} F_n(x) = F(x),$$

for every number $x \in \mathbb{R}$ at which F is continuous. Here F_n and F are the cumulative distribution functions of random variables X_n and X, respectively.

A compound Poisson model is of the form

$$X_t = x + pt - \sum_{i=1}^{N_t} U_i$$

with the claim surplus process

$$S_t^{(p)} = \sum_{i=1}^{N_t} U_i - pt,$$

where p is the premium rate is considered. In the following $\rho > 0$ is the premium rate without safety loading (critical premium rate)

$$\rho t = E\left[\sum_{i=1}^{N_t} U_i\right] = \beta t E[U_1].$$

Theorem 2.1. *As $p \to \rho$,*

$$\left(\frac{|\mu|}{\sigma^2} S^{(p)}_{t\sigma^2/\mu^2}\right)_{t\geq 0} \xrightarrow{\mathcal{D}} (W_{-1}(t))_{t\geq 0}$$

where $\mu = \mu_p = \rho - p, \sigma^2 = \beta E[U_1^2]$.

Proof. First one notes

$$\left(\frac{1}{\sigma\sqrt{c}}(S_{tc}^{(p)} - tc\mu_p)\right)_{t\geq 0} = \left(\frac{1}{\sigma\sqrt{c}}(\sum_{i=1}^{N_{tc}} U_i - ptc - tc(\rho - p))\right)_{t\geq 0} \tag{2}$$

$$= \left(\frac{1}{\sigma\sqrt{c}}(\sum_{i=1}^{N_{tc}} U_i - tc\rho)\right)_{t\geq 0} \tag{3}$$

$$= \left(\frac{1}{\sigma\sqrt{c}} S_{tc}^{(\rho)}\right)_{t\geq 0} \xrightarrow{\mathcal{D}} (W_0(t))_{t\geq 0} \tag{4}$$

for $c = c_p \to \infty$ for $p \to \rho$.

This is a consequence of (1) with $S_n^* = S_n^{(\rho)}$ and the inequalities [2, IV Lemma 1.3]

$$S_{n/c}^{(\rho)} - \rho/c \leq S_t^{(\rho)} \leq S_{(n+1)/c}^{(\rho)} + \rho/c, \qquad n/c \leq t \leq (n+1)/c.$$

Defining $c = \sigma^2/\mu_p^2 = \sigma^2/(\rho - p)^2$, then

$$\left(\frac{|\mu|}{\sigma^2} S_{t\sigma^2/\mu^2}^{(\rho)} + t\right)_{t\geq 0} = \left(\frac{|\mu|}{\sigma^2}\big(S_{t\sigma^2/\mu^2}^{(\rho)} + \frac{\sigma^2}{|\mu|} t\big)\right)_{t\geq 0} = \left(\frac{1}{\sigma\sqrt{\frac{\sigma^2}{|\mu|^2}}}\big(S_{t\sigma^2/\mu^2}^{(p)} - t\frac{\sigma^2}{\mu^2}(\rho - p)\big)\right)_{t\geq 0}$$

$$\xrightarrow{\mathcal{D}} \quad \big(W_0(t)\big)_{t\geq 0} \qquad \text{by (2).}$$

Therefore,

$$\left(\frac{|\mu|}{\sigma^2} S_{t\sigma^2/\mu^2}^{(\rho)}\right)_{t\geq 0} \quad \xrightarrow{\mathcal{D}} \quad \big(W_0(t) - t\big)_{t\geq 0} = \big(W_{-1}(t)\big)_{t\geq 0}$$

□

It was shown that there is a relation of the Cramér Lundberg model and a Brownian motion with drift for a downwards scaling (decreasing claim sizes) and long time horizon (more claims). Next, moments have to be matched to approximate the model appropriately.

Define the process $B_t = x + \mu t + \sigma W_t$, where W is a standard Brownian motion. In the next step the parameter μ and σ of B_t are determined, such that its moments match the moments of the original process $X_t = x + pt - \sum_{i=1}^{N_t} U_i$.

$$E[X_t] = E\left[x + pt - \sum_{i=1}^{N_t} U_i\right] = x + pt - \beta t E[U_1],$$
$$E[B_t] = E\left[x + \mu t + \sigma W_t\right] = x + \mu t.$$

Setting $E[X_t] = E[B_t]$, leads to

$$\mu = p - \beta E[U_1].$$

The variances are

$$V[X_t] = \beta t E[U_1^2],$$
$$V[B_t] = \sigma^2 t.$$

Matching them, gives

$$\sigma^2 = \beta E[U_1^2].$$

In the following chapters this approximation is used for the free surplus process. For this representation control methods are introduced and optimal dividend problems are solved.

3 Control theory

In the following chapter the general theory of stochastic control is introduced based on [14].

3.1 Stochastic control problem

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $T > 0$ the maturity, $\mathcal{F}$ a filtration satisfying the usual conditions (completeness and right-continuity) and $(W_t)_{t\in[0,T]}$ a m-dimensional Brownian motion w.r.t $\mathcal{F}$.

$$W_t = \begin{pmatrix} W_t^1 \\ \vdots \\ W_t^m \end{pmatrix}.$$

- The *control process (the action)* is a $\mathcal{F}$-progressively measurable process $u = (u_t)_{t\in[0,T]}$ with values in a set $\mathcal{U} \in \mathbb{R}^p$.
- The *controlled process (state of the system)* is a n-dimensional process $(X_t)_{t\in[0,T]}$ described by

  $$dX_t = b(t, X_t, u_t)\ dt + \sigma(t, X_t, u_t)\ dW_t, \qquad X_0 = x_0, \tag{5}$$

  where the coefficients

  $$b : [0,T] \times \mathbb{R}^n \times \mathcal{U} \longrightarrow \mathbb{R}^n, \quad \sigma : [0,T] \times \mathbb{R}^n \times \mathcal{U} \longrightarrow \mathbb{R}^{n\times m}$$

  are measurable. To guarantee the existence of X further conditions have to be imposed. Sometimes the state process X_t is denoted by X_t^u as it is dependent on u.
- The *optimization/performance criterion* is

  $$J(t,x,u) = E\Big[\int_t^T \psi(s, X_s^u, u_s)\ ds + \Psi(T, X_T^u)\Big| X_t^u = x\Big].$$

- An *admissible control* is a control $(u_s)_{s\in[t,T]}$ for which a unique strong solution of (5) exists on $[t,T]$ for $X_t = x$ and for which the performance measure is well defined. The set of all admissible controls is denoted by $\mathcal{A}(t,x)$.
- The *value function* of the control problem is

  $$V(t,x) = \sup_{u\in\mathcal{A}(t,x)} J(t,x,u).$$

The goal is to find the value function V and an optimal control strategy u^* for which the maximum is attained, i.e. $V(0,x_0) = J(0,x_0,u^*)$.

3.2 Dynamic programming

The aim of this chapter is to find a partial differential equation for the value function V. Therefore, the dynamic programming principle has to be introduced.

3.2.1 Itô diffusions and their generators

Again, the n-dimensional SDE

$$dX_t = b(t, X_t, u_t)\ dt + \sigma(t, X_t, u_t)\ dW_t$$

where W is a m-dimensional Brownian motion is considered and the measurable *drift* and *diffusion* coefficients

$$b : [0, T] \times \mathbb{R}^n \times \mathcal{U} \longrightarrow \mathbb{R}^n, \quad \sigma : [0, T] \times \mathbb{R}^n \times \mathcal{U} \longrightarrow \mathbb{R}^{n\times m}$$

satisfy for $K > 0$, for all $x, y \in \mathbb{R}^n$, $s, t \geq 0$ the Lipschitz and linear growth conditions

$$\begin{aligned} &||b(s, x) - b(t, y)|| + ||\sigma(s, x) - \sigma(t, y)|| \leq K(||y - x|| + |t - s|), \\ &||b(t, x)||^2 + ||\sigma(t, x)||^2 \leq K^2(1 + ||x||^2). \end{aligned}$$

These conditions imply that there is a unique and strong solution X which is called *Itô diffusion.* The matrix

$$a(t, x) := \sigma(t, x)\sigma(t, x)^t$$

is called the *diffusion matrix* of X. For a random variable Y the notation

$$\begin{aligned} &E_{t,x}[Y] = E[Y|X_t = x], \\ &E_x[Y|X_0 = x] \end{aligned}$$

is used.

Theorem 3.1. *Suppose that X is a time-homogeneous Itô diffusion and $f : \mathbb{R}^n \to \mathbb{R}$ bounded and measurable.*

1. *(Markov property): For all $\omega \in \Omega$*

$$E_x[f(X_{t+s})|\mathcal{F}_t] = E_{X_t}[f(X_s)], \quad t, s \geq 0$$

2. *(Strong Markov property): If τ is a stopping time with $\tau < \infty$, then*

$$E_x[f(X_{\tau+s})|\mathcal{F}_t] = E_{X_\tau}[f(X_s)], \quad s \geq 0$$

Definition 3.1. The *infinitesimal generator* L of X is defined as

$$Lf(s,x) = \lim_{s \to t} \frac{E_{s,x}[f(t,X_t)] - f(s,x)}{t-s}$$

for all $s \geq 0, x \in \mathbb{R}^n$ and f in the domain $\mathcal{D}_L$ of L which is the class of functions $f : [0,\infty) \times \mathbb{R}^n \to \mathbb{R}$ for which the limit exists for all s, x.

Definition 3.2. The *partial differential operator* $\mathcal{L}$ is defined as

$$\mathcal{L} := \frac{\partial}{\partial t} + \sum_{i=1}^{n} b_i \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} a_{ij} \frac{\partial^2}{\partial x_i \partial x_j}.$$

The partial differential operator can by applied to functions f in

$$\mathcal{C}^{1,2} := \{g(t,x) : [0,\infty) \times \mathbb{R}^n \to \mathbb{R}, g \text{ cont. diff. in } t \text{ and twice cont. diff. in } x\}$$

which leads to

$$\begin{aligned} \mathcal{L}f(t,x) &= f_t(t,x) + \sum_{i=1}^{n} f_{x_i}(t,x)^t b_i(t,x) + \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(t,x) f_{x_i,x_j}(t,x) \\ &= f_t(t,x) + (D_x f(t,x))^t b(t,x) + \frac{1}{2} tr((D_{xx} f(t,x)) a(t,x)) \end{aligned}$$

where $D_x f$ is the gradient of f, $D_{xx} f$ the Hessian of f (i.e. $(D_{xx} f)_{ij} = f_{x_i,x_j}$) and $tr(A)$ the trace of a matrix A, which is the sum of the diagonal elements of A.

Theorem 3.2. *Suppose that $f \in \mathcal{C}^{1,2}$ and for all $t \geq s \geq 0, x \in \mathbb{R}$*

$$E_{s,x}\Big[\int_s^t |\mathcal{L}f(u,X_u)| \, du\Big] < \infty, \qquad E_{t,x}\Big[\int_s^t |(D_x f(u,X_u))^t \sigma(u,X_u)^2| \, du\Big] < \infty.$$

Then $f \in \mathcal{D}_L$ and $Lf(s,x) = \mathcal{L}f(s,x), t \geq 0, x \in \mathbb{R}^n$.

Proof. Using Itô's formula for $f(t,X_t)$ yields

$$\begin{aligned} Lf(s,x) &= \lim_{t \searrow s} \frac{E_{s,x}[f(t,X_t)] - f(s,x)}{t-s} \\ &= \lim_{t \searrow s} \frac{E_{s,x}[f(s,x) + \int_s^t \mathcal{L}f(u,X_u) \, du + \int_s^t (D_x f(u,X_u))^t \sigma(u,X_u) \, dW_u] - f(s,x)}{t-s} \\ &= \lim_{t \searrow s} \frac{E_{s,x}[\int_s^t \mathcal{L}f(u,X_u) \, du] + E[\int_s^t (D_x f(u,X_u))^t \sigma(u,X_u) \, dW_u]}{t-s}. \end{aligned}$$

The stochastic integral $M_t := \int_s^t (D_x f(u,X_u))^t \sigma(u,X_u) \, dW_u$ is a local martingale and as it holds $E_{s,x}\Big[\int_s^t |(D_x f(u,X_u))^t \sigma(u,X_u)|^2 \, ds\Big] < \infty$ it is a true martingale. Therefore,

$$E[M_t | \mathcal{F}_s] = M_s = 0.$$

Moreover, as $E_{s,x}\Big[\int_s^t |\mathcal{L}f(u,X_u)|\ du\Big] < \infty$ one can interchange the limit and the expectation. This leads to

$$Lf(s,x) = E_{s,x}\Big[\lim_{t\searrow s} \frac{\int_s^t \mathcal{L}f(u,X_u)\ du}{t-s}\Big] = E_{s,x}[\mathcal{L}f(s,x)] = \mathcal{L}f(s,x).$$

□

The operator $\mathcal{L}$ is also called the *generator* of X. The conditions of this theorem are met for $f \in \mathcal{C}^{1,2}$ with compact support, because in this case the appearing derivatives are bounded. This motivates to the following theorem.

Theorem 3.3 (Dynkin's Formula). *Let $f \in \mathcal{C}^{1,2}$ be a function with compact support and τ a stopping time with $E_x[\tau] < \infty$. Then*

$$E_x[f(\tau, X_\tau)] = f(0,x) + E_x\Big[\int_0^\tau \mathcal{L}f(s,X_s)\ ds\Big].$$

3.2.2 The idea of dynamic programming

To solve the control problem the Hamilton-Jacobi-Bellman equation is derived. It is proceeded as follows.

1. The *Bellman Principle* is used

$$V(t,x) = \sup_{u\in\mathcal{A}(t,x)} E_{t,x}\Big[\int_t^{t_1} \psi(s,X_s^u,u_s)\ ds + V(t_1,X_{t_1}^u)\Big].$$

   The principle states that an optimal control on $[t,t_1]$ and being optimal on $[t_1,T]$ afterwards leads to a global optima, i.e. an optimal control on $[t,T]$.

2. After applying the Bellman Principle one uses the Itô formula for $V(t_1,X_{t_1}^u)$ under the assumption that V is smooth enough.

$$\begin{aligned} V(t,x) &= \sup_{u\in\mathcal{A}(t,x)} E_{t,x}\Big[\int_t^{t_1} \psi(s,X_s^u,u_s)\ ds + V(t_1,X_{t_1}^u)\Big] \\ &= \sup_{u\in\mathcal{A}(t,x)} E_{t,x}\Big[\int_t^{t_1} \psi(s,X_s^u,u_s)\ ds + V(t,X_t^u) \\ &+ \int_t^{t_1} V_t(s,x) + (D_xV(s,X_s^u))^t b(s,X_s^u,u_s)\ ds \\ &+ \frac{1}{2}\int_t^{t_1} tr((D_{xx}V(s,X_s^u))a(s,X_s^u,u_s))\ ds \\ &+ \int_t^{t_1} (D_xV(s,X_s^u))^t \sigma(s,X_s^u,u_s)\ dW_s\Big], \end{aligned}$$

   where $a(s,X_s,u_s) = \sigma(s,X_s,u_s)\sigma(s,X_s,u_s)^t$ is the diffusion matrix.

Assuming that $\int_t^{t_1} (D_x V(s, X_s^u))^t \sigma(s, X_s^u, u_s)\ dW_s,\ t_1 > t$ is a martingale, its expectation is zero and one gets

$$\begin{aligned} V(t,x) = \sup_{u \in \mathcal{A}(t,x)} E_{t,x}\Big[&\int_t^{t_1} \psi(s, X_s^u, u_s)\ ds + V(t, X_t^u) \\ &+ \int_t^{t_1} V_t(s,x) + (D_x V(s, X_s^u))^t b(s, X_s^u, u_s)\ ds \\ &+ \frac{1}{2} \int_t^{t_1} tr((D_{xx} V(s, X_s^u)) a(s, X_s^u, u_s))\ ds \Big]. \end{aligned}$$

3. Note that because of the conditional expectation one can set $X_t^u = x$. Subtracting $V(t,x)$ on both sides, dividing by $t_1 - t$ and taking the limit $t_1 \searrow t$ yields

$$\begin{aligned} 0 = \lim_{t_1 \searrow t} \sup_{u \in \mathcal{A}(t,x)} E_{t,x}\Big[&\frac{\int_t^{t_1} \psi(s, X_s^u, u_s)\ ds}{t_1 - t} \\ &+ \frac{\int_{t_1}^{t} \big(V_t(s, X_s^u) + (D_x V(s, X_s^u))^t b(s, X_s^u, u_s)\big)\ ds}{t_1 - t} \\ &+ \frac{\int_t^{t_1} \frac{1}{2} tr\big((D_{xx} V(s, X_s^u)) a(s, X_s^u, u_s)\big)\ ds}{t_1 - t} \Big] \end{aligned}$$

Under the assumption that 'sup' and 'lim' and expectation and 'lim' can be interchanged it follows

$$\begin{aligned} 0 = \sup_{u \in \mathcal{A}(t,x)} E_{t,x}\Big[&\lim_{t_1 \searrow t} \frac{\int_t^{t_1} \psi(s, X_s^u, u_s)\ ds}{t_1 - t} \\ &+ \lim_{t_1 \searrow t} \frac{\int_{t_1}^{t} \big(V_t(s, X_s^u) + (D_x V(s, X_s^u))^t b(s, X_s^u, u_s)\big)\ ds}{t_1 - t} \\ &+ \lim_{t_1 \searrow t} \frac{\int_t^{t_1} \frac{1}{2} tr\big((D_{xx} V(s, X_s^u)) a(s, X_s^u, u_s)\big)\ ds}{t_1 - t} \Big] \end{aligned}$$

In the limit case $t_1 \searrow t$ finding the optimal control on the interval $[t, t_1]$ is reduced to finding the optima at time t, i.e. $u_t = u \in \mathcal{U}$. Using that $X_t^u = x$ yields

$$0 = \sup_{u \in \mathcal{U}} \Big\{ \psi(t,x,u) + V_t(t,x) + (D_x V(t,x))^t b(t,x,u) + \frac{1}{2} tr\big((D_{xx} V(t,x)) a(t,x,u)\big) \Big\} \quad (6)$$

The equation can be written as

$$0 = \sup_{u \in \mathcal{U}} \Big\{ \psi(t,x,u) + \mathcal{L}^u V(t,x) \Big\} \quad (7)$$

where the on the optimal control u dependent operator $\mathcal{L}^u$ is defined as

$$\mathcal{L}^u f(t,x) := f_t(t,x) + (D_x f(t,x))^t b(t,x,u) + \frac{1}{2} tr\big((D_{xx} f(t,x)) a(t,x,u)\big).$$

The equation (6) resp. (7) is called *Hamilton Jacobi Bellman equation* or *HJB equation*. As seen above, under certain conditions the optimal value function V solves the HJB equation. This means it provides a necessary condition. On the other hand the question arises if a solution of the HJB equation is the value function of the corresponding problem. Therefore, sufficient conditions are needed to solve the optimal control problem. The so-called *verification theorems* are used.

3.3 Verification theorems

Definition 3.3. A control strategy u is *admissible*, i.e. $u \in \mathcal{A}(t,x)$ if

1. $u = (u_s)_{s\in[t,T]}$ is progressively measurable, has values in $\mathcal{U}$, and $E\big[\int_t^T ||u_s||^2\, ds\big] < \infty$
2. The SDE (5) has a strong solution $(X_s)_{s\in[t,T]}$ with $X_t = x$ and $E_{t,x}\big[\sup \lim_{t\leq s\leq T} ||X_s||^2\big] < \infty$
3. $J(t,x,u)$ is well defined.

Condition (3.) is guaranteed by (1.) and the assumptions of Theorem 3.4 below.

3.3.1 Finite time horizon

In this section verification theorems for the stochastic control problem of section 3.1, which guarantee that a solution to the HJB equation (7) is the optimal value function, are presented.

Theorem 3.4 (Verification Theorem). *Suppose that $||\sigma(t,x,u)||^2 < C_\sigma(1+||x||^2+||u||^2)$ and that ψ is continuous with $||\psi(t,x,u)||^2 \leq C_\psi(1+||x||^2+||u||^2)$ for some $C_\sigma, C_\psi > 0$ and all $t \geq 0, x \in \mathbb{R}^n, u \in \mathcal{U}$.*

1. *Suppose that Φ lies in $\mathcal{C}^{1,2}([0,T)\times\mathbb{R}^n)$, is continuous on $[0,T]\times\mathbb{R}^n$ with $||\Phi(t,x)|| \leq C_\Phi(1+||x||^2)$, and satisfies the HJB equation and the boundary condition, i.e.*

$$\sup_{u\in\mathcal{U}} \big\{\psi(t,x,u) + \mathcal{L}^u\Phi(t,x)\big\} = 0, \qquad t \in [0,T), x \in \mathbb{R}^n \tag{8}$$
$$\Phi(T,x) = \Psi(T,x), \qquad x \in \mathbb{R}^n.$$

   Then for all $t \in [0,T], x \in \mathbb{R}^n$

$$\Phi(t,x) \geq V(t,x).$$

2. *If a maximizer $\hat{u}(t,x)$ of $u \mapsto \psi(t,x,u) + \mathcal{L}^u\Phi(t,x)$ exists such that $u^* = (u_t^*)_{t\in[0,T]}, u_t^* = \hat{u}(t,X_t^*)$ is admissible, then $\Phi(t,x) = V(t,x)$ for all $t \in [0,T], x \in \mathbb{R}^n$ and u^* is an optimal control strategy, i.e. $V(t,x) = J(t,x,u^{t,x})$ where $u^{t,x} = (u_s^*)_{s\in[t,T]} \in \mathcal{A}(t,x)$. Here X_t^* is the solution of (5) using control u_s^* on $[0,t)$.*

Proof. Let $t \in [0,T], x \in \mathbb{R}^n$ be fixed and define the stopping time

$$\tau_n = T \wedge \inf\{s > t : ||X_s - X_t|| \geq n\}$$

Using Itôs formula for $X_t = x$ and an admissible u gives

$$\Phi(\tau_n, X_{\tau_n}) = \Phi(t,x) + \int_t^{\tau_n} \mathcal{L}^{u_s}\Phi(s, X_s)\ ds + \int_t^{\tau_n} \Phi_x(s,X_s)^t \sigma(s, X_s, u_s)\ dW_s.$$

A stochastic integral w.r.t. Brownian motion W is a local martingale. A stopped local martingale is again a local martingale. The function Φ is continuous and X bounded on $[t, \tau_n]$ due to the construction of the stopping time τ_n with

$$||X_s|| = ||X_t + X_s - X_t|| \leq ||X_t|| + ||X_s - X_t|| \leq ||x|| + n, \qquad s \in [t, \tau_n].$$

By assumption one has $||\sigma(t,x,u)||^2 < C_\sigma(1+||x||^2+||u||^2)$ and $E\big[\int_t^T ||u_s||^2\ ds\big] < \infty$. Therefore, it holds

$$E_{t,x}\Big[\int_t^{\tau_n} ||\Phi_x(s,X_s)^t\sigma(X_s,u_s)||^2\ ds\Big] < \infty$$

and thus the stochastic integral w.r.t. Brownian motion is a true martingale. As τ_n is bounded it follows with Doobs optimal stopping theorem that

$$E\Big[\int_t^{\tau_n} \Phi_x(s,X_s)^t\sigma(s,X_s,u_s)\ dW_s\Big] = 0.$$

Therefore,

$$\begin{aligned}
&E_{t,x}\Big[\int_t^{\tau_n} \psi(s,X_s,u_s)\ ds + \Phi(\tau_n, X_{\tau_n})\Big] \\
&= E_{t,x}\Big[\int_t^{\tau_n} \psi(s,X_s,u_s)\ ds + \Phi(t,x) + \int_t^{\tau_n} \mathcal{L}^{u_s}\Phi(s,X_s)\ ds\Big] \\
&= \Phi(t,x) + E_{t,x}\Big[\int_t^{\tau_n} \underbrace{\psi(s,X_s,u_s) + \mathcal{L}^{u_s}\Phi(s,X_s)}_{\leq 0 \ \ \text{(because (8))}}\ ds\Big] \\
&\leq \Phi(t,x)
\end{aligned}$$

It is left to show that

$$\lim_{n\to\infty} E_{t,x}\Big[\int_t^{\tau_n} \psi(s,X_s,u_s)\ ds + \Phi(\tau_n, X_{\tau_n})\Big] = J(t,x,u).$$

As

$$\begin{aligned}
\Big| \int_t^{\tau_n} \psi(s, X_s, u_s)\ ds + \Phi(\tau_n, X_{\tau_n}) \Big| &\leq \int_t^{\tau_n} ||\psi(s, X_s, u_s)||\ ds + ||\Phi(\tau_n, X_{\tau_n})|| \\
&\leq \int_t^{\tau_n} C_\psi (1 + ||X_s||^2 + ||u_s||^2)\ ds + C_\Phi(1 + ||X_{\tau_n}||^2) \\
&\leq C_\psi \int_t^{T} (1 + ||X_s||^2 + ||u_s||^2)\ ds + C_\Phi(1 + ||X_{\tau_n}||^2) \\
&\leq C_\psi \Big((T-t) + \int_t^T \sup_{t \leq s \leq T} ||X_s||^2\ ds + \int_t^T ||u_s||^2\ ds \Big) + C_\Phi (1 + \sup_{t \leq s \leq T} ||X_s||^2)
\end{aligned}$$

and with $u \in \mathcal{A}(t,x)$

$$E_{t,x}\Big[C_\psi \Big((T-t) + \int_t^T \sup_{t \leq s \leq T} ||X_s||^2\ ds + \int_t^T ||u_s||^2\ ds \Big) + C_\Phi (1 + \sup_{t \leq s \leq T} ||X_s||^2) \Big] < \infty$$

the dominated convergence theorem allows to interchange 'lim' and expectation

$$\lim_{n\to\infty} E_{t,x}\Big[\int_t^{\tau_n} \psi(s, X_s, u_s)\ ds + \Phi(\tau_n, X_{\tau_n}) \Big] = E_{t,x}\Big[\lim_{n\to\infty} \int_t^{\tau_n} \psi(s, X_s, u_s)\ ds + \Phi(\tau_n, X_{\tau_n}) \Big].$$

Furthermore, $\lim_{n\to\infty} \tau_n = T$ because $\forall \omega \in \Omega\ \exists \tilde{n} > 0 : \ ||X_t - X_s|| < \tilde{n}$. Such a $\tilde{n}$ exists by assumption $E_{t,x}\Big[\sup_{t < s \leq T} ||X_s||^2 \Big] < \infty$ which implies $\sup_{t < s \leq T} ||X_s|| < \infty$. With continuity of Φ it follows

$$\begin{aligned}
\lim_{n\to\infty} E_{t,x}\Big[\int_t^{\tau_n} \psi(s, X_s, u_s)\ ds + \Phi(\tau_n, X_{\tau_n}) \Big] &= E_{t,x}\Big[\int_t^{T} \psi(s, X_s, u_s)\ ds + \Phi(T, X_T) \Big] \\
= E_{t,x}\Big[\int_t^{T} \psi(s, X_s, u_s)\ ds + \Psi(T, X_T) \Big] &= J(t,x,u).
\end{aligned}$$

Now, claim (1.) is proven with

$$V(t,x) = \sup_{u \in \mathcal{A}(t,x)} J(t,x,u) \leq \Phi(t,x).$$

For claim (2.) one uses that

$$\psi(s, X_s, u_s^*) + \mathcal{L}^{u_s^*} \Phi(s, X_s) = 0$$

and the same limit arguments as before to get

$$\begin{aligned}
E_{t,x}\Big[\int_t^{T} \psi(s, X_s, u_s^*)\ ds + \Psi(T, X_T) \Big] &= J(t,x,u^*) = \Phi(t,x) \\
&\overset{\text{Claim (1.)}}{\Longrightarrow} \quad V(t,x) = \Phi(t,x).
\end{aligned}$$

□

3.3.2 Infinite time horizon

In this section a verification theorem for a more specific stochastic control problem with $T = \infty$ is given. Here, it is assumed that neither the coefficients $b(x,u)$ and $\sigma(x,u)$ nor $\psi(x,u)$ depend explicitly on time t. Therefore,

$$dX_t = b(X_t, u_t)\ dt + \sigma(X_t, u_t)\ dW_t, \quad X_0 = x_0, \tag{9}$$

with the corresponding differential operator

$$\mathcal{L}^u f(x) := (D_x f(x))^t b(x,u) + \frac{1}{2} tr\big((D_{xx} f(x)) a(x,u)\big). \tag{10}$$

The performance criterion is

$$J(x,u) = E_x\Big[\int_0^\infty e^{-\beta s}\psi(X_s, u_s)\ ds\Big]$$

with discount factor $\beta > 0$. The value function is

$$V(x) = \sup_{u \in \mathcal{A}(x)} J(x,u).$$

Theorem 3.5 (Verification Theorem)**.** *Suppose that $||\sigma(x,u)||^2 < C_\sigma(1 + ||x||^2 + ||u||^2)$ and that ψ is continuous with $||\psi(x,u)||^2 \leq C_\psi(1 + ||x||^2 + ||u||^2)$ for some $C_\sigma, C_\psi > 0$ and all $x \in \mathbb{R}^n, u \in \mathcal{U}$.*

1. *Suppose that Φ lies in $\mathcal{C}^2(\mathbb{R}^n)$ with $||\Phi(t,x)|| \leq C_\Phi(1+||x||^2)$, and satisfies the HJB equation*

$$\sup_{u \in \mathcal{U}} \big\{\psi(x,u) + \mathcal{L}^u\Phi(x) - \beta\Phi(x)\big\} = 0, \qquad x \in \mathbb{R}^n \tag{11}$$

   Then for all $x \in \mathbb{R}^n$

$$\Phi(x) \geq V(x).$$

2. *If a maximizer $\hat{u}(x)$ of $u \mapsto \psi(x,u) + \mathcal{L}^u\Phi(x) - \beta\Phi(x)$ exists such that $u^* = (u_t^*)_{t\geq 0}$, $u_t^* = \hat{u}(X_t^*)$ is admissible, then $\Phi(x) = V(x)$ for all $x \in \mathbb{R}^n$ and u^* is an optimal control strategy, i.e. $V(x) = J(x, u^*)$. Here X_t^* is the solution of (9) using control u_s^* on $[0,t)$.*

The idea of this proof is to use Theorem 3.4 for the case of a finite T and then use limit arguments for $T \to \infty$ to get the result.

3.3.3 Stopped state process

If the free surplus process of an insurance company is considered as the state process and the expected discounted dividends as performance measure, it makes sense to consider the stochastic control problem only until the ruin time, i.e. when the free surplus process is negative the first time. Therefore, the time horizon is not necessarily infinity because it is stopped when the insurance company goes ruin. The following verification theorem is for the case of a finite time horizon. Together with the theorem for the infinite time horizon case of the last section one gets a verification theorem for a stopped state process with a infinite time horizon.

Let u and X be the control and state process considered in section 3.1. Now, the state process should be constrained to a certain set. Let Q be an open set in $[0,T] \times \mathbb{R}^n$ and ∂Q the boundary of Q. The process should be stopped when leaving the set Q. This is described by the stopping time

$$\tau := \inf\{t > 0 : (t, X_t) \notin Q\}.$$

The set $(\{T\} \times \mathbb{R}^n) \cap \bar{Q}$, where $\bar{Q} = Q \cup \partial Q$, is part of the boundary, so $P_{t,x}(\tau \leq T) = 1$ for all $(t,x) \in Q$. Let $\partial^* Q$ be a subset of the boundary which satisfies

$$P_{t,x}((\tau, X_\tau) \in \partial^* Q) = 1 \quad \text{for all} \quad (t,x) \in Q.$$

The considered performance criterion is

$$J(t,x,u) = E_{t,x}\Big[\int_t^\tau \psi(s, X_s, u_s)\; ds + \Psi(\tau, X_\tau)\Big]$$

and the value function is

$$V(t,x) = \sup_{u \in \mathcal{A}(t,x)} J(t,x,u).$$

Theorem 3.6 (Verfication Theorem). *Suppose that $||\sigma(t,x,u)||^2 < C_\sigma(1 + ||x||^2 + ||u||^2)$ and that ψ is continuous with $||\psi(t,x,u)||^2 \leq C_\psi(1 + ||x||^2 + ||u||^2)$ for some $C_\sigma, C_\psi > 0$ and all $t \geq 0, x \in \mathbb{R}^n, u \in \mathcal{U}$.*

1. *Suppose that $\Phi \in \mathcal{C}^{1,2}(Q) \cap \mathcal{C}(\bar{Q})$, with $||\Phi(t,x)|| \leq C_\Phi(1 + ||x||^2)$, and satisfies the HJB equation and the boundary condition, i.e.*

$$\begin{aligned} &\sup_{u \in \mathcal{U}} \big\{\psi(t,x,u) + \mathcal{L}^u \Phi(t,x)\big\} = 0, \qquad (t,x) \in Q \\ &\Phi(t,x) = \Psi(t,x), \qquad (t,x) \in \partial^* Q. \end{aligned} \tag{12}$$

   Then for all $(t,x) \in Q$

$$\Phi(t,x) \geq V(t,x).$$

2. *If a maximizer $\hat{u}(t,x)$ of $u \mapsto \psi(t,x,u) + \mathcal{L}^u\Phi(t,x)$ exists on Q such that $u^* = (u_t^*)_{t \in [0,T]}, u_t^* = \hat{u}(t, X_t^*)$ is admissible, then $\Phi(t,x) = V(t,x)$ for all $(t,x) \in Q$ and u^* is an optimal control strategy, i.e. $V(t,x) = J(t,x,u^{t,x})$ where $u^{t,x} = (u_s^*)_{s \in [t,T]} \in \mathcal{A}(t,x)$.*

The proof is analogous to the proof of Theorem 3.4. Instead of τ_n the stopping times $\tau_n \wedge \tau$ are used.

3.4 HJB algorithm

With the above results it is possible to find a solution of the stochastic control problem by proceeding according to the following algorithm.

1. Find an optimal $u = \hat{u}(t,x)$ in (6).

2. If such an $\hat{u}$ exists, then it depends on the derivatives $V_t, D_xV, D_{xx}V$, i.e.

$$\hat{u}(t,x) = \tilde{u}\big(t, x, V_t(t,x), D_xV(t,x), D_{xx}V(t,x)\big).$$

   Substituting $\hat{u}$ in (6) gives a partial differential equation for V which has to be solved with boundary condition $\Phi(T,x) = \Psi(T,x)$ to find a candidate V^* for the optimal value function.

3. If V^* satisfies the conditions of a verification theorem and $u_t^* = \hat{u}(t, X_t^*),\ t \in [0,T]$ is an admissible control strategy, then V^* is indeed the value function and u_t^* defines an optimal control strategy. Here, X_t^* is the solution of (5) using the optimal control strategy u^* in $[0,t)$.

Example: Optimal dividend problem in the Cramér Lundberg model

The considered insurance company is assumed to pay out dividends. As defined by Azcue and Muler (2014) [4], the corresponding dividend strategy is a process $\bar{L} = (L_t)_{t\geq 0}$ where L_t is the cumulative dividends paid out until time t. The associated *controlled surplus process* X_t^L is defined as

$$X_T^L = x + pt - \sum_{i=1}^{N_t} U_i - L_t$$

and the corresponding ruin time is

$$\tau^L = \inf\{t \geq 0 : X_t^L < 0\}.$$

The dividend strategy L is called *admissible* if it is

- non-decreasing, as dividends cannot be negative,
- cáglád (left continuous with right limits) and predictable with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$,
- $L_0 = 0$ and $L_t \leq X_t = x + pt - \sum_{i=1}^{N_t} U_i$ for $0 \leq t < \tau^L$, which means that the company cannot pay dividends exceeding the current surplus.

For $t \geq \tau^{\bar{L}}$ the admissible dividend process is $L_t = L_{\tau^L}$ and Π_x^L denotes the set of all admissible dividend strategies with initial surplus x. Given a $L \in \Pi_x^L$ it follows that X_t^L is adapted and τ^L is a stopping time with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$. An upwards jump of the cumulative dividend process L_t at time s means that the company pays out a dividend of $L_{s+} - L_s$. As a compound Poisson process is of finite variation, the controlled risk process $X_t^{\bar{L}}$ is also of finite variation. It holds that $X_{t-}^L \geq X_t^L \geq X_{t+}^L$, where

$$
\begin{aligned}
&X_{t-}^L > X_t^L, \text{ only when a claim arrives,}\\
&X_t^L > X_{t+}^L, \text{ only when dividends are paid out.}
\end{aligned}
$$

Given an initial surplus $x \geq 0$ and a $L \in \Pi_x^L$, the performance measure (cumulative expected discounted dividends) $J(x, L)$ is defined as

$$J(x, L) = E_x \left[\int_0^{\tau^L} e^{-\beta s} \, dL_s \right].$$

The parameter $\beta > 0$ can be interpreted as the impatience rate of the shareholders. As L is a finite variation process the integral can be interpreted pathwise as a Riemann-Stieltjes integral. The optimal value function (optimal dividend function) is defined as

$$V(x) = \sup_{L \in \Pi_x^L} J(x, L) \quad \text{for } x \geq 0.$$

The goal is to look for an *optimal control* $L^* \in \Pi_x^L$, such that $V(x) = J(x, L^*)$.

4 1-dimensional optimal dividend problem

In this section the problem dealing with the optimal pay-out of dividends in a framework of controlled diffusion models solved by Asmussen and Taksar (1997) [3] is introduced. Here, optimal strategies exist and have a simple structure - they are barrier strategies.

4.1 Restricted dividends

The free surplus process without dividend is assumed to evolve like a Brownian motion with drift μ and variance σ^2.

$$d\tilde{X}_t = \mu\ dt + \sigma\ dW_t,$$

where $W = (W_t)_{t\geq 0}$ is a standard Brownian motion. The controlled process X_t is constructed by a dynamical choice of the dividend rate $u = (u_t)_{t\geq 0}$, by allowing it to be depended on the past up to time t. In mathematical terms this is expressed by a filtration $(\mathcal{F}_t)_{t\geq 0}$ with $\mathcal{F}_t = \sigma\{W_s : 0 \leq s \leq t\}$ on a probability space $(\Omega, \mathcal{F}, P)$.

The controlled surplus process is defined as

$$dX_t = (\mu - u_t)\ dt + \sigma\ dW_t, \qquad X_0 = x. \tag{13}$$

The stopping time τ denotes the ruin time $\tau = \inf\{t : X_t < 0\}$. The purpose of the dividend optimization problem is to maximize the discounted expected value of the dividend payments. The performance criterion is

$$J(x,u) = E_x\Big[\int_0^\tau e^{-\beta t} u_t\ dt\Big].$$

The process u is the control variable of the problem with respect to which the performance criterion is maximized over all admissible controls as defined in Section 3.1. The aim is to find an *optimal control* u^* such that $V(x) = J(x, u^*)$. In the following the case, where there is an upper bound u_0 on the rate according to which dividends can be paid out, is considered. The optimal control u_t^* at time t should only depend on the past through X_t only, which is a consequence of the Markov property of the free surplus process.

The optimal value function is

$$V(x) = \sup_{u\in\mathcal{A}(x)} J(x,u).$$

4.1.1 Derivation of the HJB equation

To derive the HJB equation the differential operator $\mathcal{L}^u$ is needed

$$\mathcal{L}^u f(x) = \big(\mu - u_t\big)\frac{\partial}{\partial x} f(x) + \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial x^2} f(x).$$

Then, the HJB equation for the case of an infinite time horizon (Section 3.3.2) is used

$$\sup_{u\in\mathcal{U}}\{\psi(x,u)+\mathcal{L}^u V(x)-\beta V(x)\}=0$$

The function ψ in this setting can be identified as $\psi(x,u)=u$. As the dividend payments are restricted by an upper bound u_0 the HJB equation is

$$\begin{aligned}&\sup_{0\le u\le u_0}\left\{u+(\mu-u)V'(x)+\frac{1}{2}\sigma^2V''(x)-\beta V(x)\right\}=0, \qquad (14)\\ &V(0)=0.\end{aligned}$$

The natural boundary condition $V(0)=0$ holds, because if the initial capital is zero dividends cannot be paid out.

4.1.2 Optimal value function and optimal strategy

Suppose $f(x)$ is a solution of the HJB system. As the optimization problem is linear dependent on u it is maximized by the optimal control $u^*(x)$, which is either 0 or u_0 at each point x, i.e.

$$u^*(x)=\begin{cases}0 & f'(x)>1,\\ u_0 & f'(x)\le 1.\end{cases}$$

Under the assumption that the solution f is concave, which will be shown later, there exists a point $m\ge 0$ such that $f'(x)>1$ as $x<m$ and $f'(x)\le 1$ as $x\ge m$. Thus the HJB equation can be rewritten as

$$\begin{aligned}&\frac{\sigma^2}{2}f''(x)+f'(x)\mu-\beta f(x)=0, && 0\le x\le m \qquad (15)\\ &\frac{\sigma^2}{2}f''(x)+f'(x)(\mu-u_0)-\beta f(x)+u_0=0, && x\ge m. \qquad (16)\end{aligned}$$

The characteristic equation

$$\frac{\sigma^2}{2}\theta^2+\lambda\,\theta-\beta=0$$

has roots

$$\theta_1(\lambda)=\frac{-\lambda+\sqrt{\lambda^2+2\beta\sigma^2}}{\sigma^2}>0,\qquad \theta_2(\lambda)=\frac{-\lambda-\sqrt{\lambda^2+2\beta\sigma^2}}{\sigma^2}<0. \qquad (17)$$

Then the generalsolution of (15) has the form

$$C_1e^{\theta_1(\mu)x}+C_2e^{\theta_2(\mu)x}.$$

For the solution of the inhomogeneous differential equation (16) one observes that the particular solution is $\frac{u_0}{\beta}$. The general solution is the sum of the general solution of the homogeneous equation and the particular solution and therefore has the form

$$\frac{u_0}{\beta} + C_3 e^{\theta_1(\mu - u_0)x} + C_4 e^{\theta_2(\mu - u_0)x}.$$

As for any control $u \in [0, u_0]$ it holds that

$$J(x,u) = E_x\Big[\int_0^\tau e^{-\beta t} u_t \; dt\Big] \leq E_x\Big[\int_0^\infty e^{-\beta t} u_0 \; dt\Big] = \frac{u_0}{\beta}$$

for all admissible u. Hence, as $V(x) = \sup_{u \in \mathcal{U}} J(x,u)$ only f with $f(x) \leq \frac{u_0}{\beta}$ are considered. As $\theta_1 > 0$ it follows that $C_3 = 0$, as exponential growth is not possible. Moreover $C_4 \equiv -d < 0$, such that $f(x) \leq \frac{u_0}{\beta}$ is not violated. As $f(0) = 0$ it holds $C_1 = -C_2 \equiv C$ and as $f(x) > 0$ it holds $C > 0$. Let $\theta_1 = \theta_1(\mu),\ -\theta_2 = \theta_2(\mu),\ -\theta_3 = \theta_2(\mu - u_0)$. The function f should be twice continuously differentiable. To identify the parameter d, C and the unknown boundary m *the principle of smooth fit* or *the smooth pasting condition* is used to get

$$\begin{aligned} f(m+) &= f(m-), \\ f'(m+) &= 1 \\ f'(m-) &= 1. \end{aligned}$$

This is equivalent to

$$C(e^{\theta_1 m} - e^{-\theta_2 m}) = \frac{u_0}{\beta} - d e^{-\theta_3 m}, \tag{18}$$

$$C(\theta_1 e^{\theta_1 m} + \theta_2 e^{-\theta_2 m}) = 1 \tag{19}$$

$$d\theta_3 e^{-\theta_3 m} = 1. \tag{20}$$

The system leads to

$$C(e^{\theta_1 m} - e^{-\theta_2 m}) = \frac{u_0}{\beta} - \frac{1}{\theta_3} =: \alpha$$

As $C > 0$ and $e^{\theta_1 m} > e^{-\theta_2 m}$, a solution can only exist if $\alpha > 0$. In the following it will be shown that this is a sufficient condition to solve (18)-(20).

Dividing (18) by (19) gives

$$\begin{aligned} \frac{e^{\theta_1 m} - e^{-\theta_2 m}}{\theta_1 e^{\theta_1 m} + \theta_2 e^{-\theta_2 m}} &= \alpha \qquad \Leftrightarrow \\ (1 - \alpha\theta_1) e^{\theta_1 m} &= (1 + \alpha\theta_2) e^{-\theta_2 m} \end{aligned}$$

Under the assumption that $\alpha\theta_1 < 1$ it holds

$$m = \frac{1}{\theta_1 + \theta_2} \log \frac{1 + \alpha\theta_2}{1 - \alpha\theta_1} > 0.$$

It follows the proof that $\alpha\theta_1 < 1$, which is equivalent to $\frac{u_0}{\beta} < \frac{1}{\theta_1} + \frac{1}{\theta_3}$. From the inequality $\sqrt{a^2+b} - a < \frac{b}{2a}$ follows that

$$\theta_1 = \frac{-\mu + \sqrt{\mu^2 + 2\beta\sigma^2}}{\sigma^2} < \frac{\frac{2\beta\sigma^2}{2\mu}}{\sigma^2} = \frac{\beta}{\mu}.$$

If $u_0 \leq \mu$

$$\frac{u_0}{\beta} \leq \frac{\mu}{\beta} \leq \frac{\mu}{\beta} + \frac{1}{\theta_3}$$

and the claim follows. If $u_0 \geq \mu$, then use

$$\theta_3 = \frac{(\mu - u_0) + \sqrt{(\mu - u_0)^2 + 2\sigma^2\beta}}{\sigma^2} = \frac{-(u_0 - \mu) + \sqrt{(u_0 - \mu)^2 + 2\sigma^2\beta}}{\sigma^2} < \frac{\beta}{u_0 - \mu}$$

to see that

$$\frac{1}{\theta_1} + \frac{1}{\theta_3} > \frac{1}{\theta_1} + \frac{u_0 - \mu}{\beta} > \frac{u_0 - \mu}{\beta} > \frac{u_0}{\beta}.$$

It was shown that $\alpha\theta_1 < 1$ has to hold. Therefore, there exists a $m > 0$ if and only if $\alpha > 0$ as $\theta_1, \theta_2 > 0$. The coefficients C, d are determined through the equations (19) and (20)

$$C = \frac{1}{\theta_1 e^{\theta_1 m} + \theta_2 e^{-\theta_2 m}},$$
$$d = \frac{e^{\theta_3 m}}{\theta_3}.$$

The next theorem characterizes the solution of the HJB equation, where one has to distinguish the cases $\alpha > 0$ and $\alpha \leq 0$.

Theorem 4.1. *There exists a twice continuously differentiable concave solution to (14). If $\alpha \leq 0$, then the solution is*

$$f(x) = \frac{u_0}{\beta}(1 - e^{-\theta_3 x}). \tag{21}$$

If $\alpha > 0$, then

$$f(x) = \begin{cases} C(e^{\theta_1 x} - e^{-\theta_2 x}), & 0 \leq x \leq m, \\ \frac{u_0}{\beta} - de^{-\theta_3 x}, & x > m, \end{cases} \tag{22}$$

where C, d, m are the unique solutions of (18) – (20).

Proof. The function (21) is concave, satisfies $f(0) = 0$ and $f'(0) = \theta_3 u_0/\beta \leq 1$. Therefore $f'(x) \leq 1$ for all $x > 0$ and

$$(u_0 - u)(f'(x) - 1) \leq 0, \qquad u \in [0, u_0].$$

Adding this inequality to the equality

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

satisfied by f, leads to (14).

Now, concavity for the second case is shown. It is assumed that $\alpha > 0$. The function f satisfies $f(0) = 0$. By construction f is continuous and $f'(m-) = f(m+)$. As f satisfies (15) for $x \in [0, m]$ and (16) for $x \in (m, \infty)$ it holds

$$\begin{aligned} f''(m-) &= \frac{2}{\sigma^2}(\beta f(m-) - \mu f'(m-)), \\ f''(m+) &= \frac{2}{\sigma^2}(\beta f(m+) - (\mu - u_0) f'(m+) - u_0). \end{aligned}$$

As $f'(m) = f'(m-) = f'(m+) = 1$ and $f(m+) = f(m-)$ by construction, it holds

$$\begin{aligned} f''(m-) &= \frac{2}{\sigma^2}(\beta f(m) - \mu), \\ f''(m+) &= \frac{2}{\sigma^2}(\beta f(m) - (\mu - u_0) - u_0). \end{aligned}$$

Therefore, $f''(m) = f''(m-) = f''(m+)$ and f is twice continuously differentiable. Concavity on (m, ∞) is obvious

$$f''(x) = -d\theta_3^2 e^{-\theta_3 x} < 0 \text{ for } x \in (m, \infty).$$

To see that f is concave on $[0, m]$ note that $f'''(x) = C(\theta_1^3 e^{\theta_1 x} + \theta_2^3 e^{-\theta_2 x}) > 0$ for $x \in [0, m]$, which means that f'' is monotone increasing. As $f''(0) < 0$ and $f''(m+) = f''(m-) = -\theta_3 < 0$ it follows $f''(x) < 0$ for $x \in [0, m]$ and f is concave on $[0, m]$. Thus, f is concave on $[0, \infty)$.

Now it is shown that function f for the case $\alpha > 0$ satisfies the HJB equation (14). If $x \leq m$, then $f'(x) \geq 1$ and adding the inequality $-u(f'(x) - 1) \leq 0$ to (15) leads to (14). If $x > m$, then $f'(x) = d\theta_3 e^{-\theta_3 x} = e^{\theta_3(m-x)} < 1$ and (14) follows by adding $(u_0 - u)(f'(x) - 1) \leq 0$ to (16).

Uniqueness of the coefficients C, d, m was already shown before. □

Proposition 4.1. *The function f in Theorem 4.1 is greater than $J(x, u)$ for any admissible control u and any initial capital x, i.e. $f(x) \geq V(x)$.*

Proof. Using Itôs product formula gives

$$\begin{aligned}
&e^{-\beta(\tau\wedge T)}f(X_{\tau\wedge T}) - f(x) \\
&= e^{-\beta 0}f(X_0) + \int_0^{\tau\wedge T} e^{-\beta s}\ df(X_s) - \int_0^{\tau\wedge T} f(X_s)\beta e^{-\beta s}\ ds - f(x) \\
&= \int_0^{\tau\wedge T} e^{-\beta s}\ \Big(f'(X_s)\ dX_s + \frac{\sigma^2}{2}f''(X_s)\ ds\Big) - \int_0^{\tau\wedge T} f(X_s)\beta e^{-\beta s}\ ds \\
&= \int_0^{\tau\wedge T} e^{-\beta s}\Big(f'(X_s)(\mu - u_s) + \frac{\sigma^2}{2}f''(X_s) - f(X_s)\beta\Big)\ ds + \int_0^{\tau\wedge T} e^{-\beta s}f'(X_s)\sigma\ dW_s. \qquad (23)
\end{aligned}$$

The last term is a stopped local martingale (therefore a local martingale itself) and as f is concave ($f'(x) \leq f'(0)$ for any x) it holds

$$\begin{aligned}
E\Big[\Big[\int_0^{\tau\wedge\cdot} e^{-\beta s}f'(X_s)\sigma\ dW_s\Big]_T\Big] &= E\Big[\int_0^{\tau\wedge T} e^{-2\beta s}f'(X_s)^2\sigma^2\ ds\Big] \\
&\leq E\Big[\int_0^{\tau\wedge T} e^{-2\beta s}f'(0)^2\sigma^2\ ds\Big] < \infty.
\end{aligned}$$

Hence, it is a square integrable zero-expectation martingale. From (14) one can see that the bracket in the first integrand of (23) doesn't exceed $-u_s$. Rearranging terms and taking expectations lead to

$$f(x) \geq E_x\Big[\int_0^{\tau\wedge T} e^{-\beta s}u_s\ ds\Big] + E_x\big[e^{-\beta(\tau\wedge T)}f(X_{\tau\wedge T})\big]$$

Taking limits $T \to \infty$, using the Monotone Convergence Theorem and that $f(0) = 0$ and $X_\tau = 0$, yields

$$f(x) \geq E_x\Big[\int_0^{\tau} e^{-\beta s}u_s\ ds\Big] = J(x, u).$$

The claim $f(x) \geq V(x)$ follows trivially. $\square$

Proposition 4.2. *Define*

$$u^*(x) = \begin{cases} 0, & 0 \leq x \leq m \\ u_0, & x > m \end{cases}$$

and let X_t^* be the solution of the stochastic differential equation (13) with $u_t = u^*(X_t^*)$. Then $J(x, u^*) = f(x)$.

Proof. Substitution $u^*(x) = u_0 1_{x>m}$ in the expression for $e^{-\beta(\tau\wedge T)}f(X_{\tau\wedge T}) - f(x)$ derived in the proof for Proposition 4.1 gives

$$\begin{aligned}
&e^{-\beta(\tau\wedge T)}f(X_{\tau\wedge T})-f(x)\\
&=\int_0^{\tau\wedge T}e^{-\beta s}\Big(f'(X_s)(\mu-u_0 1_{X_s>m})+\frac{\sigma^2}{2}f''(X_s)-f(X_s)\beta\Big)\,ds+\int_0^{\tau\wedge T}e^{-\beta s}f'(X_s)\sigma\,dW_s\\
&=\int_0^{\tau\wedge T}e^{-\beta s}f'(X_s)\sigma\,dW_s.
\end{aligned}$$

If $X_s\le m$ the expression in the bracket is zero with (15), if $X_s>m$ the expression is equal to u_0. Taking expectations yields

$$f(x)=E_x\Big[\int_0^{\tau\wedge T}e^{-\beta s}u_0\,ds\Big]+E_x\big[e^{-\beta(\tau\wedge T)}f(X_{\tau\wedge T})\big].$$

Since $f(0)=0$ and $X_\tau=0$,

$$\begin{aligned}
E_x\big[e^{-\beta(\tau\wedge T)}f(X_{\tau\wedge T})\big]&=E_x\big[e^{-\beta(\tau\wedge T)}f(X_{\tau\wedge T})1_{\tau>T}\big]\\
&=E_x\big[e^{-\beta T}f(X_T)1_{\tau>T}\big]\\
&\le e^{-\beta T}\sup_{x\ge 0}f(x)\le e^{-\beta T}u_0/\beta,
\end{aligned}$$

which converges to zero for $T\to\infty$. Passing to the limit with Monotone Convergence yields

$$f(x)=E_x\Big[\int_0^{\tau}e^{-\beta s}u_0\,ds\Big].$$

□

Corollary 4.1. $f(x)=V(x)=J(x,u^*)$

4.2 Unrestricted dividends

In this section the case, where the dividend is not restricted and the situation $u_t=\infty$ is possible, is considered following again the approach of [3]. The cumulative dividend process up time t is defined as

$$L_t=\int_0^t u_s\,ds.$$

As the process $L=(L_t)_{t\ge 0}$ describes the summed up amount of dividends paid out up to the time t, it is called *admissible* if

- $L_t\in\mathcal{F}_t$,
- L is non-decreasing and non-negative.

In comparison to the last section dealing with restricted dividends, instead of u the *dividend strategy* L is the control process. The imposed conditions for L do not imply continuity, but they do not exclude it. It is only assumed, that the process is left continuous with right limits. The controlled surplus process is defined as

$$dX_t = \mu\ dt + \sigma\ dW_t - dL_t.$$

As $X_0 = x - L_0$, it is conventionally assumed that $L_{0-} = 0$ and therefore $X_{0-} = x$. The performance index is

$$J(x, L) = E_x\Big[\int_0^\tau e^{-\beta t}\ dL_t\Big],$$

interpreted as Riemann-Stieltjes integration, with the left endpoint of integration included. Define the optimal return function as the supremum over all admissible L

$$V(x) = \sup_{L\ adm.} J(x, L).$$

The objective is to find an *optimal control* L^* such that

$$V(x) = J(x, L^*).$$

Compared with the last section, when dividends were bounded and the optimal dividend was to pay 0 or the maximal rate, in this case the maximal rate is unrestricted.

4.2.1 Derivation of the HJB equation

The HJB-equation derivation follows the approach of [4]. Here, it is assumed that dividends are paid at a constant rate $l \geq 0$, i.e. $(L_t)_{t\geq 0} = (lt)_{t\geq 0}$ and $dL_t = l\ dt$. To derive the HJB equation the differential operator $\mathcal{L}^l$ is needed

$$\mathcal{L}^l f(x) = \big(\mu - l\big)\frac{\partial}{\partial x} f(x) + \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial x^2} f(x).$$

Then, the HJB equation for the case of an infinite time horizon (Section 3.3.2) is used

$$\sup_{l\geq 0}\{\psi(x, l) + \mathcal{L}^l V(x) - \beta V(x)\} = 0$$

The function ψ in this setting can be identified as $\psi(x, l) = l$. Putting everything together yields

$$\sup_{l\geq 0}\Big\{l + (\mu - l)V'(x) + \frac{1}{2}\sigma^2 V''(x) - \beta V(x)\Big\} = 0, \tag{24}$$
$$V(0) = 0,$$

The natural boundary condition $V(0) = 0$ holds, because if the initial capital is zero dividends cannot be paid out. Rearranging terms gives

$$l\big(1 - V'(x)\big) + \mu V'(x) + \frac{1}{2}\sigma^2 V''(x) - \beta V(x) \le 0 \qquad \text{for all } l \ge 0.$$

As this inequality holds for all $l \ge 0$ it has to hold $V'(x) \ge 1$ because otherwise it will be violated for l large enough, i.e. $l \to \infty$. This means the left side of the inequality attains its maximum at $l = 0$ for $V'(x) > 1$. If $V'(x) = 1$, then it remains

$$\mu V'(x) + \frac{1}{2}\sigma^2 V''(x) - \beta V(x) = 0.$$

Theorem 4.2. *The optimal return function V satisfies the following Hamilton-Jacobi-Bellmann equation:*

$$\begin{aligned} &\max\left\{\frac{1}{2}\sigma^2 V''(x) + \mu V'(x) - \beta V(x), 1 - V'(x)\right\} = 0, \\ &V(0) = 0. \end{aligned} \tag{25}$$

4.2.2 Optimal value function and optimal strategy

This equation is solved in the following way. First, one constructs a twice continuously differentiable function f, which satisfies (25). Second, it has to be shown that this solution is the optimal return function, i.e. $V = f$. Third, an optimal control L^* has to be found such that the corresponding performance criterion is V.

The solution is constructed as follows. It is assumed that f is concave, so that $f'(x)$ is non-increasing. Define $m = \sup\{x : f'(x) > 1\}$. As it has to hold $f'(x) \ge 1$ for a solution, one has

$$f'(x) = \begin{cases} > 1, & x < m, \\ = 1, & x \ge m. \end{cases}$$

Moreover,

$$\frac{1}{2}\sigma^2 f''(x) + \mu f'(x) - \beta f(x) = 0, \quad x \le m. \tag{26}$$

The solution f should therefore consist of two pices: one satisfies (26) on $[0, m]$ and on $f'(x) = 1$ on $[m, \infty]$. To find the boundary m again the *principle of smooth fit* is used. As f is twice continuously differentiable it has to hold that

$$f(m) = f(m-) = f(m+), \tag{27}$$

$$f'(m) = f'(m+) = f'(m-) = 1, \tag{28}$$

$$f''(m) = f''(m+) = f''(m-) = 0. \tag{29}$$

As seen in Section 4.1, a general solution of (26) is $f(x) = C_1 e^{\theta_1 x} + C_2 e^{\theta_2 x}$, where θ_1, θ_2 are given in (17), with $C := C_1 = -C_2 > 0$ because $f(0) = 0$. Then

$$f(x) = C(e^{\theta_1 x} - e^{-\theta_2 x}), \tag{30}$$
$$f'(x) = C(\theta_1 e^{\theta_1 x} + \theta_2 e^{-\theta_2 x}), \tag{31}$$
$$f''(x) = C(\theta_1^2 e^{\theta_1 x} - \theta_2^2 e^{-\theta_2 x}). \tag{32}$$

From (28) and (31) follows

$$C = \frac{1}{\theta_1 e^{\theta_1 x} + \theta_2 e^{-\theta_2 x}} \tag{33}$$

and from (29) and (32)

$$m = \frac{2}{\theta_1 + \theta_2} \log\left(\frac{\theta_2}{\theta_1}\right). \tag{34}$$

As it holds $f'(x) = 1$ on $[m, \infty)$ it follows that $f(x) = x + \mathit{constant}$ on $[m, \infty)$. As the function f is assumed to be twice continuously differentiable it has to satisfy $f(m) = f(m-) = f(m+)$ which leads to the equation $C(e^{\theta_1 m} - e^{-\theta_2 m}) = m + \mathit{const}$. Therefore the constant is computed as $C(e^{\theta_1 m} - e^{-\theta_2 m}) - m$.

Theorem 4.3. *Define*

$$f(x) = \begin{cases} C(e^{\theta_1 x} - e^{-\theta_2 x}), & x \leq m, \\ C(e^{\theta_1 m} - e^{-\theta_2 m}) + x - m, & x \geq m. \end{cases}$$

where the parameter C and m are given by (33) and (34) and the parameter θ_1, θ_2 by (17). Then $f(x)$ is a solution to the Hamilton-Jacobi-Bellman equation (25).

Proof. What remains to be shown is that it holds

$$\begin{aligned} &f'(x) \geq 1, && x \leq m, \\ &\frac{1}{2}\sigma^2 f''(x) + \mu f'(x) - \beta f(x) \leq 0, && x \geq m. \end{aligned}$$

Are this conditions fulfilled, the function f satisfies the HJB equation of Theorem 4.2. Considering the third derivative of f

$$f'''(x) = \begin{cases} C\left(\theta_1^3 e^{\theta_1 x} + \theta_2^3 e^{-\theta_2 x}\right) & x \leq m, \\ 0 & x > m, \end{cases} \quad \geq 0, \tag{35}$$

as θ_1, θ_2 are positive. This means that $f''(x)$ is non-decreasing. Furthermore, $f''(0) < 0$ as $-\theta_2 > \theta_1$. Since $f''(m) = 0$ it holds $f''(x) \leq 0$ for $x \leq m$. This proves that $f(x)$ is concave on $[0, m]$. Hence, $f'(x) \geq f'(m) = 1$ for $x \leq m$.

For $x \geq m$

$$\frac{1}{2}\sigma^2 \underbrace{f''(x)}_{=0} + \mu \underbrace{f'(x)}_{=1} - \beta f(x) = \mu - \beta f(x) \leq \mu - \beta f(m) = \frac{1}{2}\sigma^2 f''(m) + \mu f'(m) - \beta f(m) = 0,$$

as f is monotone increasing because of $f'(x) = 1$ for $x \geq m$. $\square$

Proposition 4.3. *If L is any control, then*

$$f(x) \geq J(x, L) = E\Big[\int_0^\tau e^{-\beta t}\ dL_t\Big]$$

Moreover, $f(x) \geq V(x)$.

Proof. To use Itôs formula the non-decreasing dividend process L is decomposed into its continuous L^C and discontinuous part L^D

$$L_t = L_t^C + L_t^D = \int_0^t dL_s^C + \sum_{L_{s-} \neq L_s, s \leq t} (L_s - L_{s-}).$$

Using Itôs product formula

$$e^{-\beta t} f(X_t) = f(x) + \int_0^t e^{-\beta s} df(X_s) + \int_0^t (-\beta) e^{-\beta s} f(X_s)\ ds$$

where with Itôs formula for semimartingales [16, Proposition 8.19]

$$\begin{aligned}
f(X_t) &= f(X_0) + \int_0^t f'(X_s) dX_s + \int_0^t \frac{1}{2}\sigma^2 f''(X_s)\ d[X]_s^C \\
&+ \sum_{X_s \neq X_{s-}, s \leq t} \big(f(X_s) - f(X_{s-}) - f'(X_s)(X_s - X_{s-})\big) \\
&= \int_0^t f'(X_s)\Big(\mu\ ds + \sigma\ dW_s - dL_s^C\Big) - \sum_{L_{s-} \neq L_s, s \leq t} f'(X_s)(L_s - L_{s-}) \\
&+ \int_0^t \frac{1}{2}\sigma^2 f''(X_s)\ ds + \sum_{X_s \neq X_{s-}, s \leq t} \big(f(X_s) - f(X_{s-}) - f'(X_s)(\underbrace{X_s - X_{s-}}_{-(L_s - L_{s-})})\big) \\
&= \int_0^t \Big(f'(X_s)\mu + \frac{1}{2}\sigma^2 f''(X_s)\Big)\ ds + \int_0^t f'(X_s)\sigma\ dW_s \\
&- \int_0^t f'(X_s)\ dL_s^C + \sum_{X_s \neq X_{s-}, s \leq t} \big(f(X_s) - f(X_{s-})\big)
\end{aligned}$$

as $\{s : X_s \neq X_{s-}\} = \{s : L_s \neq L_{s-}\}$ leads to

$$\begin{aligned} e^{-\beta(t\wedge\tau)} f(X_{t\wedge\tau}) = f(x) + \int_0^{t\wedge\tau} e^{-\beta s} \underbrace{\Big(f'(X_s)\mu + \frac{1}{2}\sigma^2 f''(X_s) - \beta f(X_s) \Big)}_{\leq 0} \, ds \\ + \int_0^{t\wedge\tau} e^{-\beta s} f'(X_s)\sigma \; dW_s - \int_0^{t\wedge\tau} e^{-\beta s} f'(X_s) \; dL_s^C \\ + \sum_{X_s \neq X_{s-}, s \leq t\wedge\tau} e^{-\beta s} \big(f(X_s) - f(X_{s-}) \big) \end{aligned}$$

Since f is concave $f'(x) < f'(0) < \infty$. The stochastic integral $\int_0^{t\wedge\tau} e^{-\beta s} f'(X_s)\sigma \; dW_s$ is a zero expectation martingale because of its square integrable integrands. As f satisfies the HJB equation the following inequality holds

$$\begin{aligned} E\Big[e^{-\beta(t\wedge\tau)} f(X_{t\wedge\tau})\Big] \leq f(x) - E\Big[\int_0^{t\wedge\tau} e^{-\beta s} f'(X_s) \; dL_s^C \Big] \\ + E\Big[\sum_{X_s \neq X_{s-}, s \leq t\wedge\tau} e^{-\beta s} \big(f(X_s) - f(X_{s-}) \big) \Big]. \end{aligned}$$

It holds $f(X_\tau) = f(0) = 0$. Therefore,

$$E\Big[e^{-\beta(t\wedge\tau)} f(X_{t\wedge\tau})\Big] = E\Big[e^{-\beta t} f(X_t) 1_{t<\tau}\Big] = e^{-\beta t} E\Big[f(X_t) 1_{t<\tau}\Big].$$

It holds that $X_t = x + \mu t + \sigma W_t - L_t \leq |x + \mu t + \sigma W_t|$ and as f is concave, there exist $a, b > 0$ such that $f(x) \leq a + bx$. Therefore,

$$e^{-\beta t} E\Big[f(X_t) 1_{t<\tau}\Big] \leq e^{-\beta t} E\Big[(a + bX_t) 1_{t<\tau}\Big] \leq e^{-\beta t} (a + bE[|x + \mu t + \sigma W_t|])$$

As

$$E[|W_t|] = \int_0^\infty 2x \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{1}{2t} x^2 \right) \mathrm{d}x = \sqrt{\frac{2t}{\pi}}$$

it follows that

$$e^{-\beta t} E\Big[f(X_t) 1_{t<\tau}\Big] \leq e^{-\beta t} \left(a + bx + b\mu t + b\sqrt{\frac{2t}{\pi}} \right) \overset{t\to\infty}{\longrightarrow} 0.$$

Since $f'(x) \geq 1$

$$-E\Big[\int_0^{t\wedge\tau} e^{-\beta s} f'(X_s) \; dL_s^C \Big] \leq -E\Big[\int_0^{t\wedge\tau} e^{-\beta s} \; dL_s^C \Big]$$

and since $f(X_s) - f(X_{s-}) = f'(\zeta)(X_s - X_{s-}) = -f'(\zeta)(L_s - L_{s-}) \leq -(L_s - L_{s-})$ for $\zeta \in [X_s, X_{s-}]$

$$E\Big[\sum_{X_s \neq X_{s-}, s \leq t\wedge\tau} e^{-\beta s} \big(f(X_s) - f(X_{s-}) \big) \Big] \leq -E\Big[\sum_{X_s \neq X_{s-}, s \leq t\wedge\tau} e^{-\beta s} \big(L_s - L_{s-} \big) \Big].$$

Combining these inequalities and taking limits $t \to \infty$ gives

$$0 \le f(x) - E\Big[\int_0^\tau e^{-\beta s}\, dL_s^C\Big] - E\Big[\sum_{X_s \neq X_{s-}, s \le \tau} e^{-\beta s}\big(L_s - L_{s-})\big)\Big] = f(x) - \underbrace{E\Big[\int_0^\tau e^{-\beta s} dL_s\Big]}_{=J(x,L)}.$$

Since $f(x) \ge J(x, L)$ for any admissible strategy L it holds

$$f(x) \ge V(x) = \sup_L J(x, L)$$

□

Now, the functional L^* is constructed, such that $J(x, L^*) = f(x)$. Define

$$\begin{aligned} L_t^* &= \max_{s \le t} \big[x + \mu s + \sigma W_s - m\big]^+, \\ X_t^* &= x + \mu t + \sigma W_t - L_t^*. \end{aligned} \tag{36}$$

The process L^* is a continuous non-decreasing process with $L^*(0) > 0$ when $x > m$. In this case L^* has a jump of size $x - m$ at $t = 0$ and $X_0^* = m$. The process X^* is a Brownian motion with upper reflection boundary m. The optimal dividend strategy L^* increases at the same time, when $X^* = m$. This means

$$\begin{aligned} X_t^* &\le m \quad \text{for all } t \ge 0, \\ \int_0^\infty & 1_{X_t^* < m}\, dL_t^* = 0. \end{aligned}$$

This dividend strategy pays out the amount the controlled surplus process X_t^* exceeds the boundary m. If the controlled surplus process is below m, no dividend is paid.

Proposition 4.4. *It holds $J(x, L^*) = f(x) = V(x)$. This means L^* is the optimal control.*

Proof. The same approach as in the proof of Theorem 4.3 is used. First the case $x \le m$ is considered. The stopping time is definied as $\tau^* = \inf\{t : X_t^* \le 0\}$. Now, one notes that

$$f'(x)\mu + \frac{1}{2}\sigma^2 f''(x) - \beta f(x) = 0 \qquad \text{for } x \le m$$

to see that

$$\begin{aligned} E\Big[e^{-\beta(t\wedge\tau^*)} f(X_{t\wedge\tau^*}^*)\Big] = E\Big[e^{-\beta t} f(X_t^*) 1_{t<\tau^*}\Big] &= f(x) - E\Big[\int_0^{t\wedge\tau^*} e^{-\beta s} f'(X_s^*)\, dL_s^{*C}\Big] \\ &+ E\Big[\sum_{X_s^* \neq X_{s-}^*, s \le t\wedge\tau^*} e^{-\beta s}\big(f(X_s^*) - f(X_{s-}^*)\big)\Big] \\ &= f(x) - E\Big[\int_0^{t\wedge\tau^*} e^{-\beta s} f'(X_s^*)\, dL_s^*\Big] \end{aligned}$$

because L is a continuous process for $x \leq m$.

$$\begin{aligned} E\Big[\int_0^{t\wedge\tau^*} e^{-\beta s} f'(X_s^*)\ dL_s^*\Big] &= E\Big[\int_0^{t\wedge\tau^*} e^{-\beta s} f'(X_s^*) 1_{X_s^*=m}\ dL_s^*\Big] \\ &= E\Big[\int_0^{t\wedge\tau^*} e^{-\beta s}\ \underbrace{f'(m)}_{=1}\ dL_s^*\Big] \\ &= E\Big[\int_0^{t\wedge\tau^*} e^{-\beta s}\ dL_s^*\Big]. \end{aligned}$$

This gives

$$f(x) = E\Big[e^{-\beta t} f(X_t^*) 1_{t<\tau^*}\Big] + E\Big[\int_0^{t\wedge\tau^*} e^{-\beta s}\ dL_s^*\Big]$$

Letting t tend to ∞ and using the Bounded Convergence Theorem (f bounded on $[0, m]$) and the Monotone Convergence Theorem, one has

$$f(x) = E\Big[\int_0^t e^{-\beta s} dL_s^*\Big] = J(x, L^*).$$

$\square$

5 2-dimensional optimal dividend problem

In this chapter the optimal dividend strategy is analysed for the case when two insurance companies (or business lines) are involved. The problem was solved by Gu, Steffensen and Zheng (2016) [11] in a framework where the free surplus processes are modelled as diffusion processes. In this model it is not possible that only one company survives and the other one goes ruin, i.e. ruin is defined as simultaneous ruin. The collaboration allows money to be transferred from one company to another one, but must be transferred in order to avoid ruin of one company. Both companies aim to maximize their joint paid out dividends.

5.1 The model

The considered model includes two insurance companies: Company 1 and Company 2. The free surplus process of Company 1 is denoted by $\bar{X}_t^1$ and the free surplus process of Company 2 by $\bar{X}_t^2$. Both free surpluses are diffusion processes

$$\bar{X}_t^1 = x_1 + \mu_1 t + \sigma_1 W_t^1,$$
$$\bar{X}_t^2 = x_2 + \mu_2 t + \sigma_2 W_t^2,$$

were x_1, x_2 are the initial surplus levels and W_1 and W_2 are independent standard Brownian motions. The collaboration works as follows. If the current surplus of one company hits zero, the other one is obligated to transfer positive cash to avoid ruin. Moreover, both companies are allowed to transfer money to each other at any time without costs.

If there is a positive surplus, both companies are can to pay out dividends to their shareholders. The dividend strategy $L_t = (L_t^1, L_t^2)$ is the cumulative dividend paid out up to time t. The process C_t^{21} corresponds to the total amount transferred from Company 2 to Company 1 and C_t^{12} to the total amount transferred from Company 1 to Company 2 up to time t. Then the associated controlled surplus processes are

$$X_t^1 = \bar{X}_t^1 + C_t^{21} - C_t^{12} - L_t^1,$$
$$X_t^2 = \bar{X}_t^2 + C_t^{12} - C_t^{21} - L_t^2.$$

Ruin can only occur simultaneously at the moment both processes hit zero. The ruin time of the companies is

$$\tau := \inf\{t > 0 : X_t^1, X_t^2 < 0\}.$$

A dividend and transferring strategy $(L, C) = (L^1, L^2, C^{12}, C^{21})$ is called *admissible*, denoted as $(L, C) \in \pi_{(x_1, x_2)}$, if

1. L^1, L^2 are left continuous with right limits and $\mathcal{F}_t$-predictable, where $\mathcal{F}_t$ is the natural filtration generated by X^1 and X^2. This corresponds to the fact, that an insurance company

plans its dividend policy itself and the amount to be paid out at time t is known before time t.

2. C^{12}, C^{21} are right continuous with left limits and $\mathcal{F}_t$-adapted. In comparison to the dividends payments costs can arise unexpected but are observable right after they are paid.

3. L^1, L^2, C^{12}, C^{21} are non-negative and non-decreasing. These processes describe the accumulated dividends resp. costs of Company 1 resp. Company 2 up to time t. Neither dividends nor costs can be negative. These properties imply finite variation.

4. $L_t^1 \leq X_t^1 + C_t^{21} - C_t^{12}$ and $L_t^2 \leq X_t^2 + C_t^{12} - C_t^{21}$. It is not possible to be pay out more dividends than the available surplus after deducting costs.

Let $\mathcal{R}_+^2$ denote the set $\{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0\}$. For any initial surplus level $(x_1, x_2) \in \mathcal{R}_+^2$, the optimal value function is

$$V(x_1, x_2) = \sup_{(L,C)\in\pi_{(x_1,x_2)}} J_{L,C}(x_1, x_2),$$

where the performance criterion J is

$$J_{L,C}(x_1, x_2) = E_{x_1,x_2}\Big[a \int_0^\tau e^{-\beta s} dL_s^1 + (1-a)\int_0^\tau e^{-\beta s} dL_s^2\Big].$$

The weights a and $1-a$ represent possible different proportional costs when drawing money out of the two companies.

5.2 Derivation of the HJB equation

For fixed $l_1, l_2, \Delta x_1, \Delta x_2 > 0$ and $B_1 > x_1, B_2 > x_2$, the following strategy (L, C) is considered. The processes X^1 and X^2 are the controlled surplus processes associated with (L, C). For a given initial surplus (x_1, x_2), Company 1 pays dividends at rate l_1 and transfers money to Company 2 with rate Δx_1 until time $\bar{\tau}$, where $\bar{\tau} = \inf\{t \geq 0 : X_t^1 = 0 \text{ or } X_t^2 = 0 \text{ or } X_t^1 \geq B_1 \text{ or } X_t^2 \geq B_2\}$. Analogously, Company 2 pays dividends at rate l_2 and transfers money at rate Δx_2 up to time $\bar{\tau}$.

$$\begin{aligned} X_t^1 &= x_1 + (\mu_1 + l_1 + \Delta x_2 - \Delta x_1)t + \sigma_1 W_t^1 \\ X_t^2 &= x_2 + (\mu_2 + l_2 + \Delta x_1 - \Delta x_2)t + \sigma_2 W_t^2 \end{aligned}$$

Using the dynamic programming principle

$$\begin{aligned} V(x_1, x_2) = \sup_{(L,C)\in\pi_{(x_1,x_2)}} & E_{x_1,x_2}\Big[a \int_0^{\bar{\tau}\wedge t} e^{-\beta s} l_1 \, ds + (1-a)\int_0^{\bar{\tau}\wedge t} e^{-\beta s} l_2 \, ds \\ & + e^{-\beta(\bar{\tau}\wedge t)} V(X_{\bar{\tau}\wedge t}^1, X_{\bar{\tau}\wedge t}^2)\Big] \end{aligned} \tag{37}$$

it holds

$$\begin{aligned} V(x_1,x_2) \ \geq \ & E_{x_1,x_2}\Big[a\int_0^{\bar\tau\wedge t} e^{-\beta s} l_1 \ ds + (1-a)\int_0^{\bar\tau\wedge t} e^{-\beta s} l_2 \ ds\Big] \\ & + E_{x_1,x_2}\Big[e^{-\beta(\bar\tau\wedge t)} V(X^1_{\bar\tau\wedge t}, X^2_{\bar\tau\wedge t})\Big]. \end{aligned} \tag{38}$$

Using Itô's formula for $e^{-\beta t}V(X_t^1, X_t^2)$ gives

$$\begin{aligned} d\big(e^{-\beta t}V\big) &= e^{-\beta t}\ dV + V(-\beta)e^{-\beta t}\ dt + d[V, e^{-\beta \cdot}]_t = e^{-\beta t}\ dV - V\beta e^{-\beta t}\ dt, \\ dV(X_t^1, X_t^2) &= \frac{\partial V}{\partial x_1} dX_t^1 + \frac{\partial V}{\partial x_2} dX_t^2 + \frac{1}{2}\sigma_1^2 \frac{\partial^2 V}{\partial x_1^2} dt + \frac{1}{2}\sigma_2^2 \frac{\partial^2 V}{\partial x_2^2} dt + \frac{\partial^2 V}{\partial x_1 x_2} d[X^1, X^2]_t \\ &= \Big(\frac{\partial V}{\partial x_1}(\mu_1 + \Delta x_2 - \Delta x_1 - l_1) + \frac{\partial V}{\partial x_2}(\mu_2 + \Delta x_1 - \Delta x_2 - l_2) + \frac{1}{2}\sigma_1^2\frac{\partial^2 V}{\partial x_1^2} dt + \frac{1}{2}\sigma_2^2 \frac{\partial^2 V}{\partial x_2^2}\Big)\ dt \\ &\quad + \frac{\partial V}{\partial x_1}\sigma_1\ dW_t^1 + \frac{\partial V}{\partial x_2}\sigma_2\ dW_t^2. \end{aligned}$$

Taking expectations and using the optional sampling theorem, under the assumption that the integrands are such that the stochastic integrals resp. W^1 and W^2 are martingales, leads to

$$\begin{aligned} & E_{x_1,x_2}\Big[e^{-\beta(\bar\tau\wedge t)} V(X^1_{\bar\tau\wedge t}, X^2_{\bar\tau\wedge t})\Big] - V(x_1,x_2) \\ & \quad = E_{x_1,x_2}\Big[\int_0^{\bar\tau\wedge t} e^{-\beta s}\mathcal{L}V\ ds\Big] \\ & \quad + E_{x_1,x_2}\Big[\int_0^{\bar\tau\wedge t} e^{-\beta s}\big[\frac{\partial V}{\partial x_1}(\Delta x_2 - \Delta x_1 - l_1) + \frac{\partial V}{\partial x_2}(\Delta x_1 - \Delta x_2 - l_2)\big]\ ds\Big], \end{aligned} \tag{39}$$

where

$$\mathcal{L}V := \mu_1 \frac{\partial V}{\partial x_1} + \mu_2 \frac{\partial V}{\partial x_2} + \frac{1}{2}\sigma_1^2\frac{\partial^2 V}{\partial x_1^2} + \frac{1}{2}\sigma_2^2\frac{\partial^2 V}{\partial x_2^2} - \beta V.$$

Substitution (39) into (38), subtracting $V(x_1,x_2)$ on both sides and taking the limes for $t \searrow 0$ gives

$$\begin{aligned} 0 \geq & \lim_{t\searrow 0} \frac{E_{x_1,x_2}\big[\int_0^{\bar\tau\wedge t} e^{-\beta s}\mathcal{L}V\ ds\big]}{t} \\ & + \lim_{t\searrow 0}\frac{E_{x_1,x_2}\Big[\int_0^{\bar\tau\wedge t} e^{-\beta s}\Big(\frac{\partial V}{\partial x_1}(\Delta x_2 - \Delta x_1 - l_1) + \frac{\partial V}{\partial x_2}(\Delta x_1 - \Delta x_2 - l_2)\Big)\ ds\Big]}{t} \\ & + \lim_{t\searrow 0}\frac{E_{x_1,x_2}\Big[\int_0^{\bar\tau\wedge t} e^{-\beta s}(a l_1 + (1-a) l_2)\ ds\Big]}{t} \end{aligned}$$

and after rearranging terms and using that $(X_0^1, X_0^2) = (x_1, x_2)$

$$\begin{aligned} 0 \ \geq\ & \mathcal{L}V(x_1,x_2) + l_1\left(a - \frac{\partial V}{\partial x_1}(x_1,x_2)\right) + l_2\left(1 - a - \frac{\partial V}{\partial x_2}(x_1,x_2)\right) \\ & + \Delta x_2\left(\frac{\partial V}{\partial x_1} - \frac{\partial V}{\partial x_2}\right)(x_1,x_2) + \Delta x_1\left(\frac{\partial V}{\partial x_2} - \frac{\partial V}{\partial x_1}\right)(x_1,x_2) \end{aligned}$$

By setting $l_1 = l_2 = \Delta x_1 = \Delta x_2 = 0$, one has $0 \geq \mathcal{L}V(x_1,x_2)$. By setting $l_2 = \Delta x_1 = \Delta x_2 = 0$ and assuming that $l_1 \to \infty$ gives that is has to hold that $0 \geq (a - \frac{\partial V}{\partial x_1})(x_1,x_2)$. And letting $l_2 \to \infty, l_1 = \Delta x_1 = \Delta x_2 = 0$, $\Delta x_1 \to \infty, l_1 = l_2 = \Delta x_2 = 0$ and $\Delta x_2 \to \infty, l_1 = l_2 = \Delta x_1 = 0$ gives a similar result. Therefore, under the assumption that 'sup' and 'lim' in (37) can be interchanged one has

$$\begin{aligned} 0 \ =\ & \sup_{l_1,l_2,\Delta x_1,\Delta x_2 \geq 0} \mathcal{L}V(x_1,x_2) + l_1\left(a - \frac{\partial V}{\partial x_1}(x_1,x_2)\right) + l_2\left(1 - a - \frac{\partial V}{\partial x_2}(x_1,x_2)\right) \\ & + \Delta x_2\left(\frac{\partial V}{\partial x_1} - \frac{\partial V}{\partial x_2}\right)(x_1,x_2) + \Delta x_1\left(\frac{\partial V}{\partial x_2} - \frac{\partial V}{\partial x_1}\right)(x_1,x_2) \end{aligned}$$

and the corresponding HJB equation for this problem is given by

$$\begin{aligned} 0 = \max\Bigg\{ & \mathcal{L}V(x_1,x_2), a - \frac{\partial V}{\partial x_1}(x_1,x_2), 1 - a - \frac{\partial V}{\partial x_2}(x_1,x_2), \\ & \frac{\partial V}{\partial x_2}(x_1,x_2) - \frac{\partial V}{\partial x_1}(x_1,x_2), \frac{\partial V}{\partial x_2}(x_1,x_2) - \frac{\partial V}{\partial x_1}(x_1,x_2)\Bigg\}, \qquad (40) \\ 0 = V&(0,0). \end{aligned}$$

Due to symmetry, without loss of generality, it is assumed that $a \geq \frac{1}{2}$.

This optimization problem is solved by following the steps below.

1. A function f, that solves the HJB equation (40), is constructed.
2. It is shown that the solution f is greater than the optimal value function V.
3. A dividend strategy L^* and a transferring strategy C^* are constructed, such that $f(x_1,x_2) = J_{L^*,C^*}(x_1,x_2)$ for any initial capital x_1, x_2. This proves that (L^*, C^*) gives the optimal strategy.

5.3 Optimal value function and optimal strategy

Proposition 5.1. *Let X^1, X^2 be the controlled surplus processes with control (L, C) and initial values x_1, x_2. For any twice continuously differentiable function ψ on $\mathcal{R}_+^2$ and a finite stopping time $\tau^* \leq \tau$, if one of the following two conditions holds,*

1. *X^1, X^2 are bounded,*
2. *ψ has bounded first derivatives,*

then, one has

$$
\begin{aligned}
& e^{-\beta\tau^*}\psi(X^1_{\tau^*}, X^2_{\tau^*}) - \psi(x_1, x_2) \\
& = \int_0^{\tau *} e^{-\beta s}\mathcal{L}\psi(X^1_{s-}, X^2_{s-})\ ds + M(\tau^*) \\
& + \int_0^{\tau *} e^{-\beta s}[\psi_{x_1}(X^1_{s-}, X^2_{s-}) - \psi_{x_2}(X^1_{s-}, X^2_{s-})]\ dC^{21C}_s \\
& + \int_0^{\tau *} e^{-\beta s}[\psi_{x_2}(X^1_{s-}, X^2_{s-}) - \psi_{x_1}(X^1_{s-}, X^2_{s-})]\ dC^{12C}_s \\
& \sum_{X^1_{s-}\neq X^1_s, X^2_{s-}\neq X^2_s, s\leq\tau^*} e^{-\beta s}[\psi(X^1_s, X^2_s) - \psi(X^1_{s-}, X^2_{s-})] \\
& - \int_0^{\tau^*} ae^{-\beta s} dL^1_s - \int_0^{\tau^*}(1-a)e^{-\beta s} dL^2_s \\
& + \int_0^{\tau^*}(a - \psi_{x_1}(X^1_{s-}, X^2_{s-}))e^{-\beta s} dL^{1C}_s \\
& + \sum_{L^1_{s+}\neq L^1_s, s<\tau^*} \int_0^{L^1_{s+}-L^1_s}(a - \psi_{x_1}(X^1_s - \alpha, X^2_s))e^{-\beta s} d\alpha \\
& + \int_0^{\tau^*}(1 - a - \psi_{x_2}(X^1_{s-}, X^2_{s-}))e^{-\beta s} dL^{2C}_s \\
& + \sum_{L^2_{s+}\neq L^2_s, s<\tau^*} \int_0^{L^2_{s+}-L^2_s}(1 - a - \psi_{x_2}(X^1_{s+}, X^2_s - \alpha))e^{-\beta s} d\alpha,
\end{aligned}
$$

where M is a martingale, L^{iC}, C^{ijC} are the continuous parts of L^i, C^{ij} and $\psi_{x_i} := \frac{\partial\psi}{\partial x_i}$, $\psi_{x_ix_j} := \frac{\partial^2\psi}{\partial x_i x_j}$.

Proof. Let $X = (X^1, X^2)$. The dividend process L^i is non-decreasing and left continuous, therefore finite variation and the sum of the jumps converges and it can be written as

$$L^i_t = \int_0^t dL^{iC}_s + \sum_{X^i_{s+}\neq X^i_s, s<t}(L^i_{s+} - L^i_s). \tag{41}$$

The same holds for the right continuous cost process

$$C^{ij}_t = \int_0^t dC^{ijC}_s + \sum_{X_{s-}\neq X_s, s\leq t}(C^{ij}_s - C^{ij}_{s-}). \tag{42}$$

Also the controlled surplus processes are of finite variation and can be expressed as

$$
\begin{aligned}
X^i_t &= \int_0^t dX^{iC}_s + \sum_{\Delta X^i_s\neq 0, s<t}\Delta X^i_s & (43)\\
&= \int_0^t dX^{iC}_s + \sum_{X^i_s\neq X^i_{s-}, s\leq t}(X^i_s - X^i_{s-}) + \sum_{X^i_{s+}\neq X^i_s, s<t}(X^i_{s+} - X^i_s). & (44)
\end{aligned}
$$

Itôs product formula gives

$$
\begin{aligned}
&e^{-\beta\tau^*}\psi(X^1_{\tau^*}, X^2_{\tau^*}) - \psi(x_1, x_2) \\
&\quad = \int_0^{\tau^*} e^{-\beta s}\; d\psi(X^1_s, X^2_s) + \int_0^{\tau*} \psi(X^1_s, X^2_s)\; d\left(e^{-\beta s}\right) + \int_0^{\tau*} d\,\underbrace{[\psi(X^1, X^2), e^{-\beta\cdot}]_s}_{=0} \\
&\quad = \int_0^{\tau^*} e^{-\beta s}\; d\psi(X^1_s, X^2_s) - \beta\int_0^{\tau^*} e^{-\beta s}\psi(X^1_s, X^2_s)\; ds. \qquad (45)
\end{aligned}
$$

The surplus processes X_1 and X_2 are rewritten as (43). The Itô formula for semimartinales [16, Proposition 8.19] leads to

$$
\begin{aligned}
\psi(X_t) &= \int_0^t \psi_{x_1}(X_{s-})\; dX^1_s + \int_0^t \psi_{x_2}(X_{s-})\; dX^2_s + \frac{1}{2}\int_0^t \psi_{x_1x_2}(X_s)\; d\,\underbrace{[X^1, X^2]^C_s}_{=0} \\
&+ \frac{1}{2}\int_0^t \psi_{x_1x_2}(X_s)\; d\,\underbrace{[X^2, X^1]^C_s}_{=0} + \frac{1}{2}\int_0^t \psi_{x_1x_1}(X_s)\; d[X^1, X^1]^C_s + \frac{1}{2}\int_0^t \psi_{x_2x_2}(X_s)\; d[X^2, X^2]^C_s \\
&+ \sum_{\substack{X_s \neq X_{s-} \\ s\leq t}} (\psi(X_s) - \psi(X_{s-}) - \psi_{x_1}(X_{s-})(X^1_s - X^1_{s-}) - \psi_{x_2}(X_{s-})(X^2_s - X^2_{s-})) \\
&+ \sum_{\substack{X_{s+} \neq X_s \\ s<t}} (\psi(X_{s+}) - \psi(X_s) - \psi_{x_1}(X_{s-})(X^1_{s+} - X^1_s) - \psi_{x_2}(X_{s-})(X^2_{s+} - X^2_s)) \\
&= \int_0^t \psi_{x_1}(X_{s-})\; dX^{1C}_s + \int_0^t \psi_{x_2}(X_{s-})\; dX^{2C}_s + \frac{1}{2}\int_0^t \psi_{x_1x_1}(X_s)\; d[X^1]^C_s \\
&+ \frac{1}{2}\int_0^t \psi_{x_2x_2}(X_s)\; d[X^2]^C_s + \sum_{\substack{X_s \neq X_{s-} \\ s\leq t}} (\psi(X_s) - \psi(X_{s-})) + \sum_{\substack{X_{s+} \neq X_s \\ s<t}} (\psi(X_{s+}) - \psi(X_s))
\end{aligned}
$$

The processes L^1, L^2, C^{12} and C^{21} are rewritten as (41) resp. (42). As one company transfers money to the other company, the right continuous processes C^{21} and C^{12} jump at the same time. As the dividend strategy is left continuous and the Brownian motion and drift parts are continuous it holds that $\{s : X^1_{s-} \neq X^1_s\} = \{s : X^2_{s-} \neq X^2_s\} = \{s : C^{12}_{s-} \neq C^{12}_s\} = \{s : C^{21}_{s-} \neq C^{21}_s\}$. Also note that continuous finite variation processes have quadratic variation zero.

Therefore,

$$
\begin{aligned}
\psi(X_t) = {} & \underbrace{\int_0^t \psi_{x_1}(X_{s-})\ dX_s^{1C}}_{\int_0^t \psi_{x_1}(X_{s-})\ d(\mu_1 s + \sigma_1 W_s^1 + C_s^{21C} - C_s^{12C} - L_s^{1C})} + \underbrace{\int_0^t \psi_{x_2}(X_{s-})\ dX_s^{2C}}_{\int_0^t \psi_{x_2}(X_{s-})\ d(\mu_2 s + \sigma_2 W_s^2 + C_s^{12C} - C_s^{21C} - L_s^{2C})} \\
& + \frac{1}{2}\int_0^t \psi_{x_1 x_1}(X_s)\ \underbrace{d[X^1]_s^C}_{=\sigma_1^2\ ds} + \frac{1}{2}\int_0^t \psi_{x_2 x_2}(X_s)\ \underbrace{d[X^2]_s^C}_{=\sigma_2^2\ ds} \\
& + \sum_{\substack{X_s^1 \neq X_{s-}^1, X_s^2 \neq X_{s-}^2 \\ s \leq t}} (\psi(X_s^1, X_s^2) - \psi(X_{s-}^1, X_{s-}^2)) \\
& + \sum_{\substack{L_{s+}^1 \neq L_s^1, L_{s+}^2 = L_s^2 \\ s<t}} (\psi(X_{s+}^1, X_s^2) - \psi(X_s^1, X_s^2)) + \sum_{\substack{L_{s+}^2 \neq L_s^2, L_{s+}^1 = L_s^1 \\ s<t}} (\psi(X_s^1, X_{s+}^2) - \psi(X_s^1, X_s^2)) \\
& + \sum_{\substack{L_{s+}^2 \neq L_s^2, L_{s+}^1 \neq L_s^1 \\ s<t}} (\psi(X_{s+}^1, X_{s+}^2) - \psi(X_s^1, X_s^2)).
\end{aligned}
$$

Rearranging terms yields

$$
\begin{aligned}
\psi(X_t^1, X_t^2) = {} & \int_0^t \big(\psi_{x_1}\mu_1 + \psi_{x_2}\mu_2 + \frac{1}{2}\sigma_1^2 \psi_{x_1 x_1} + \frac{1}{2}\sigma_2^2 \psi_{x_2 x_2}\big)(X_{s-}^1, X_{s-}^2)\ ds \\
& + \int_0^t \psi_{x_1}(X_{s-}^1, X_{s-}^2)\sigma_1\ dW_s^1 + \int_0^t \psi_{x_2}(X_{s-}^1, X_{s-}^2)\sigma_2\ dW_s^2 \\
& + \int_0^t \big(\psi_{x_1} - \psi_{x_2}\big)(X_{s-}^1, X_{s-}^2)\ dC_s^{21C} + \int_0^t \big(\psi_{x_2} - \psi_{x_1}\big)(X_{s-}^1, X_{s-}^2)\ dC_s^{12C} \\
& - \int_0^t \psi_{x_1}(X_{s-}^1, X_{s-}^2)\ dL_s^{1C} - \int_0^t \psi_{x_2}(X_{s-}^1, X_{s-}^2)\ dL_s^{2C} \\
& + \sum_{X_s^1 \neq X_{s-}^1, X_s^2 \neq X_{s-}^2, s \leq t} [\psi(X_s^1, X_s^2) - \psi(X_{s-}^1, X_{s-}^2)] \\
& + \sum_{L_{s+}^1 \neq L_s^1, s<t} (\psi(X_{s+}^1, X_s^2) - \psi(X_s^1, X_s^2)) + \sum_{L_{s+}^2 \neq L_s^2, s<t} (\psi(X_{s+}^1, X_{s+}^2) - \psi(X_{s+}^1, X_s^2)).
\end{aligned}
$$

Using this result to compute the first term of equation (45) gives

$$\begin{aligned}
\int_0^{\tau^*} e^{-\beta s}\, d\psi(X_s^1, X_s^2) &= \int_0^{\tau^*} e^{-\beta s}\big(\psi_{x_1}\mu_1 + \psi_{x_2}\mu_2 + \frac{1}{2}\sigma_1^2\psi_{x_1x_1} + \frac{1}{2}\sigma_2^2\psi_{x_2x_2}\big)(X_{s-}^1, X_{s-}^2)\, ds \\
&+ \int_0^{\tau^*} e^{-\beta s}\psi_{x_1}(X_{s-}^1, X_{s-}^2)\sigma_1\, dW_s^1 + \int_0^{\tau^*} e^{-\beta s}\psi_{x_2}(X_{s-}^1, X_{s-}^2)\sigma_2\, dW_s^2 \\
&+ \int_0^{\tau^*} e^{-\beta s}\big(\psi_{x_1} - \psi_{x_2}\big)(X_{s-}^1, X_{s-}^2)\, dC_s^{21C} + \int_0^{\tau^*} e^{-\beta s}\big(\psi_{x_2} - \psi_{x_1}\big)(X_{s-}^1, X_{s-}^2)\, dC_s^{12C} \\
&- \int_0^{\tau^*} e^{-\beta s}\psi_{x_1}(X_{s-}^1, X_{s-}^2)\, dL_s^{1C} - \int_0^{\tau^*} e^{-\beta s}\psi_{x_2}(X_{s-}^1, X_{s-}^2)\, dL_s^{2C} \\
&+ \sum_{X_{s-}^1 \neq X_s^1, X_{s-}^2 \neq X_s^2, s \leq \tau^*} e^{-\beta s}(\psi(X_s^1, X_s^2) - \psi(X_{s-}^1, X_{s-}^2)) \\
&+ \sum_{L_{s+}^1 \neq L_s^1, s < \tau^*} e^{-\beta s}(\psi(X_{s+}^1, X_s^2) - \psi(X_s^1, X_s^2)) \\
&+ \sum_{L_{s+}^2 \neq L_s^2, s < \tau^*} e^{-\beta s}(\psi(X_{s+}^1, X_{s+}^2) - \psi(X_{s+}^1, X_s^2)).
\end{aligned}$$

As the processes W^1 and W^2 are standard Brownian motions, they are (local) martingales. Stochastic integrals with respective to a local martingale are again local martingales. Therefore the process

$$M_t := \int_0^t e^{-\beta s}\psi_{x_1}(X_{s-}^1, X_{s-}^2)\sigma_1\, dW_s^1 + \int_0^t e^{-\beta s}\psi_{x_2}(X_{s-}^1, X_{s-}^2)\sigma_2\, dW_s^2$$

is a local martingale. If one of the conditions hold, i.e. first derivatives of ψ are bounded or X^1, X^2 are bounded (together with ψ continuous differentiable) it follows that the integrands are bounded resp. bounded on $[0, t]$ (by Weierstrass' Theorem). Therefore, it holds $E\Big[[M(\cdot)]_t\Big] < \infty$ and M_t is a martingale. With the optional sampling theorem it follows that the stopped martingale is again a martingale. Since $X_{s+}^1 - X_s^1 = -(L_{s+}^1 - L_s^1)$,

$$\begin{aligned}
&- \int_0^{\tau^*} e^{-\beta s}\psi_{x_1}(X_{s-}^1, X_{s-}^2)\, dL_s^{1C} \\
&\quad + \sum_{L_{s+}^1 \neq L_s^1, s < \tau^*} e^{-\beta s}(\psi(X_{s+}^1, X_s^2) - \psi(X_s^1, X_s^2)) \\
&= - \int_0^{\tau^*} e^{-\beta s}\psi_{x_1}(X_{s-}^1, X_{s-}^2)\, dL_s^{1C} \\
&\quad - \sum_{L_{s+}^1 \neq L_s^1, s < \tau^*} e^{-\beta s}\int_0^{L_{s+}^1 - L_s^1} \psi_{x_1}(X_s^1 - \alpha, X_s^2)\, d\alpha \\
&= - \int_0^{\tau^*} a e^{-\beta s}\, dL_s^1 + \int_0^{\tau^*} e^{-\beta s}(a - \psi_{x_1}(X_{s-}^1, X_{s-}^2))\, dL_s^{1C} \\
&\quad + \sum_{L_{s+}^1 \neq L_s^1, s < \tau^*} e^{-\beta s}\int_0^{L_{s+}^1 - L_s^1} (a - \psi_{x_1}(X_s^1 - \alpha, X_s^2))\, d\alpha.
\end{aligned}$$

Analogously,

$$
\begin{aligned}
&-\int_0^{\tau^*} e^{-\beta s}\psi_{x_2}(X^1_{s-},X^2_{s-})\ dL^{2C}_s \\
&\quad+\sum_{L^2_{s+}\neq L^2_s, s<\tau^*} e^{-\beta s}(\psi(X^1_{s+},X^2_{s+})-\psi(X^1_{s+},X^2_s)) \\
&=-\int_0^{\tau^*} e^{-\beta s}\psi_{x_2}(X^1_{s-},X^2_{s-})\ dL^{2C}_s \\
&\quad-\sum_{L^2_{s+}\neq L^2_s, s<\tau^*} e^{-\beta s}\int_0^{L^2_{s+}-L^2_s}\psi_{x_2}(X^1_{s+},X^2_s-\alpha)\ d\alpha \\
&=-\int_0^{\tau^*}(1-a)e^{-\beta s}\ dL^2_s+\int_0^{\tau^*} e^{-\beta s}((1-a)-\psi_{x_2}(X^1_{s-},X^2_{s-}))\ dL^{2C}_s \\
&\quad+\sum_{L^2_{s+}\neq L^2_s, s<\tau^*} e^{-\beta s}\int_0^{L^2_{s+}-L^2_s}((1-a)-\psi_{x_2}(X^1_{s+},X^2_s-\alpha))\ d\alpha.
\end{aligned}
$$

Combining all equalities shown above, the proof is complete. □

Now, the goal is to find a twice continuously differentiable function in $\mathcal{R}^2_+$ that solves problem (40). Define the following function f on $\mathcal{R}^2_+$:

$$
f(x_1,x_2)=\begin{cases} aC(e^{\theta_1(x_1+x_2)}-e^{-\theta_2(x_1+x_2)}), & x_1+x_2<m, \\ a[C(e^{\theta_1 m}-e^{-\theta_2 m})+x_1+x_2-m], & x_1+x_2\geq m, \end{cases} \tag{46}
$$

where θ_1,θ_2,C,m are positive constants, which have to be determined. The intuition to find this function is the same as in Section 4.2 for $x=x_1+x_2$. It is required that $\mathcal{L}(f)=0$ when $x_1+x_2\leq m$. The characteristic equation is

$$
\Big(\frac{1}{2}\sigma_1^2+\frac{1}{2}\sigma_2^2\Big)\theta+(\mu_1+\mu_2)\theta-\beta=0,
$$

with roots $\theta_1,-\theta_2$, where $\theta_2>\theta_1>0$

$$
\theta_1=\frac{-(\mu_1+\mu_2)+\sqrt{(\mu_1+\mu_2)^2+2(\sigma_1^2+\sigma_2^2)\beta}}{(\sigma_1^2+\sigma_2^2)}, \tag{47}
$$

$$
-\theta_2=\frac{-(\mu_1+\mu_2)-\sqrt{(\mu_1+\mu_2)^2+2(\sigma_1^2+\sigma_2^2)\beta}}{(\sigma_1^2+\sigma_2^2)}. \tag{48}
$$

As f has to be twice differentiable, analogously to Section 4.2 the *principle of smooth fit* has to hold:

$$
\begin{aligned}
&\frac{\partial f}{\partial x_1}\Big|_{x_1+x_2=m}=\frac{\partial f}{\partial x_2}\Big|_{x_1+x_2=m}=a, \\
&\frac{\partial^2 f}{\partial x_1^2}\Big|_{x_1+x_2=m}=\frac{\partial^2 f}{\partial x_2^2}\Big|_{x_1+x_2=m}=\frac{\partial^2 f}{\partial x_1 x_2}\Big|_{x_1+x_2=m}=0.
\end{aligned}
$$

Again, these conditions lead to

$$C(\theta_1 e^{\theta_1 m} + \theta_2 e^{-\theta_2 m}) = 1,$$
$$\theta_1^2 e^{\theta_1 m} - \theta_2^2 e^{-\theta_2 m} = 0,$$

which gives the same coefficients as in the 1-dimensional problem with unrestricted dividends for $\mu = \mu_1 + \mu_2$ and $\sigma^2 = \sigma_1^2 + \sigma_2^2$:

$$C = \frac{1}{\theta_1 e^{\theta_1 m} + \theta_2 e^{-\theta_2 m}}, \quad m = \frac{2}{\theta_1 + \theta_2} \ln\left(\frac{\theta_2}{\theta_1}\right). \tag{49}$$

Proposition 5.2. *The function f defined in (46), with coefficients given in (47) and (49) is twice continuously differentiable and has bounded first derivatives and is a solution to the HJB equation (40).*

Proof. The function f is twice continuously differentiable, has bounded first derivatives by construction and $\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2}$.

- Case $x_1 + x_2 < m$: By construction, it holds $\mathcal{L}f = 0$. It remains to show that $a - \frac{\partial f}{\partial x_1} \leq 0$ and $(1-a) - \frac{\partial f}{\partial x_2} \leq 0$. One has $\frac{\partial^3 f}{\partial x_1^3} = aC(\theta_1^3 e^{\theta_1(x_1+x_2)} + \theta_2^3 e^{-\theta_2(x_1+x_2)}) > 0$. The derivative $\frac{\partial^2 f}{\partial x_1^2}$ is increasing and $\frac{\partial^2 f}{\partial x_1^2}\Big|_{x_1+x_2=m} = 0$, thus $\frac{\partial^2 f}{\partial x_1^2} \leq 0$. The derivative $\frac{\partial f}{\partial x_1}$ is decreasing and $\frac{\partial f}{\partial x_1}\Big|_{x_1+x_2=m} = a$, thus $a - \frac{\partial f}{\partial x_1} \leq 0$ for $x_1 + x_2 < m$. Similarly, $(1-a) - \frac{\partial f}{\partial x_2} \leq a - \frac{\partial f}{\partial x_2} \leq 0$.
- Case $x_1 + x_2 \geq m$: Here, it holds $(1-a) - \frac{\partial f}{\partial x_2} \leq 0$ and $a = \frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2}$. The cases $x_2 \leq m$ and $x_2 > m$ are considered to prove that $\mathcal{L}f \leq 0$. Note that f is monotone increasing in both arguments. If $x_2 \leq m$, $\mathcal{L}f(x_1, x_2) = \mu_1 a + \mu_2 a - \beta f(x_1, x_2) \leq \mu_1 a + \mu_2 a - \beta f(m - x_2, x_2) = \mathcal{L}f(m - x_2, x_2) = 0$. If $x_2 > m$, $\mathcal{L}f(x_1, x_2) = \mu_1 a + \mu_2 a - \beta f(x_1, x_2) \leq \mu_1 a + \mu_2 a - \beta f(0, m) = \mathcal{L}f(0, m) = 0$.

$\square$

Now, the optimal dividend and transferring strategy (L^*, C^*) is constructed. For this purpose, the domain $\mathcal{R}_+^2$ is divided into three parts:

- $A = \{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \leq m\}$,
- $B = \{(x_1, x_2) : x_1 > 0, x_2 \in [0, m], x_1 + x_2 > m\}$,
- $C = \{(x_1, x_2) : x_1 \geq 0, x_2 > m\}$.

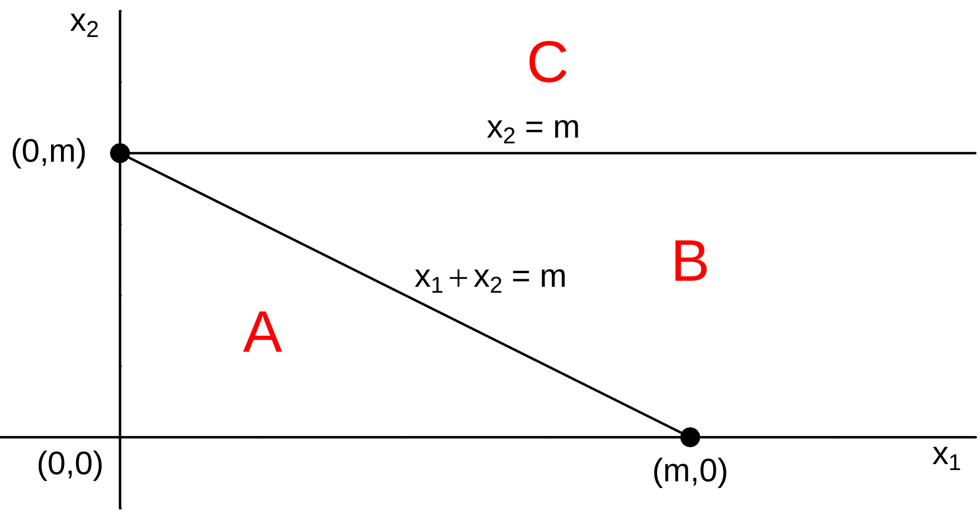

Figure 3: Areas for the optimal strategy

The dividend strategy is L^* and the transferring strategy is C^* with initial surplus level (x_1, x_2). The strategy (L^*, C^*) is constructed as follows (see Figure 4-6):

1. $(x_1, x_2) \in C$: Company 2 transfers $x_2 - m$ to Company 1. Continue with 2.
2. $(x_1, x_2) \in B$: Company 1 pays $x_1 + x_2 - m$ as dividend. Continue with 3.
3. $(x_1, x_2) \in A$: Company 1 pays accumulated $\max_{s \leq t}[X_s^{1*} + X_s^{2*} - m]^+$ up to time t until the process reaches $(0, 0)$. If X^{1*} hits zero, money is transferred from Company 2 to Company 1. Analogously, if X^{2*} hits zero, money is transferred from Company 1 to Company 2.

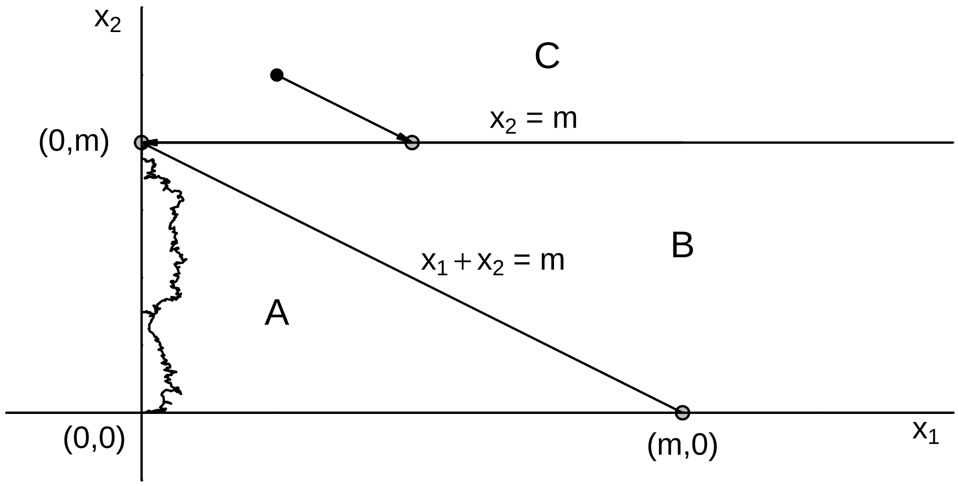

Figure 4: Optimal Strategy when initial value in area C

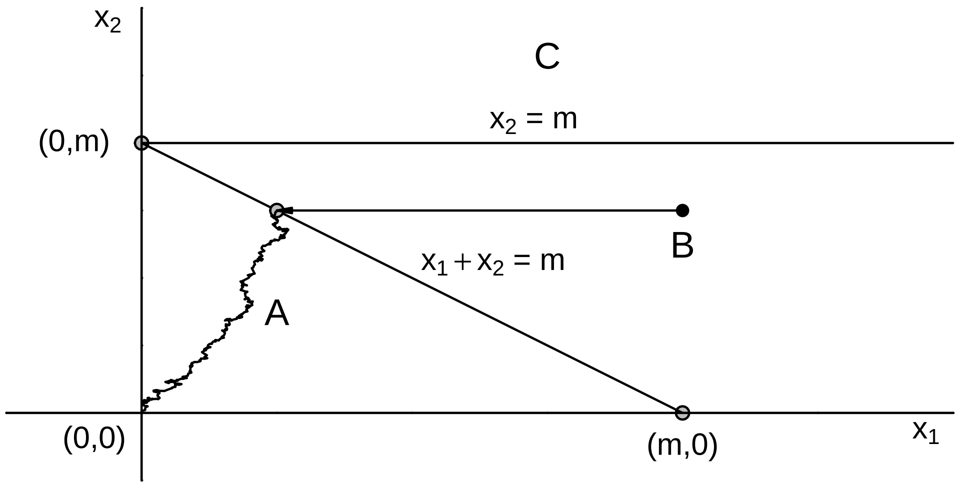

Figure 5: Optimal Strategy when initial value in area B

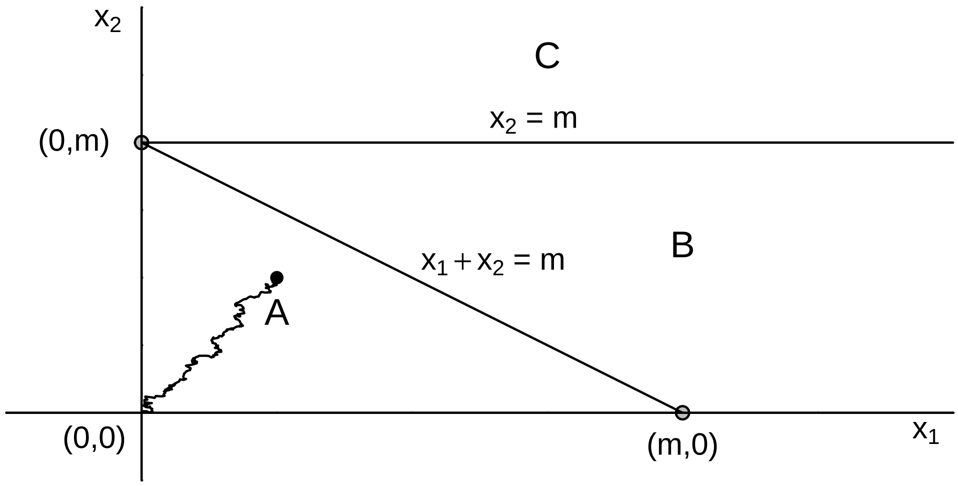

Figure 6: Optimal Strategy when initial value in area A

Here, X^{1*} and X^{2*} denote the controlled surplus processes using the strategies (L^*, C^*). The optimal strategy implies that dividends are only paid if the sum of the two surplus processes exceeds or reaches the barrier m. Company 2 never pays any dividends, but if its surplus is higher than the barrier, it transfers the money exceeding m to Company 1. Only Company 1 pays dividends when the sum of both surplus processes of the companies $X_s^{1*} + X_s^{2*}$ hits the barrier m.

The next theorem shows that the above constructed strategy gives the optimal value function, which means that this barrier strategy is optimal.

Theorem 5.1. *The strategy (L^*, C^*) gives the optimal strategy and the function f gives the optimal value function, i.e. $f(x_1, x_2) \geq J_{L,C}(x_1, x_2)$ and $f(x_1, x_2) = J_{L^*,C^*}(x_1, x_2)$.*

Proof. The first step is to show that for any initial capital $(x_1, x_2) \in \mathcal{R}_+^2$ and any admissible control (L, C),

$$f(x_1, x_2) \geq J_{L,C}(x_1, x_2).$$

Let X^1, X^2 denote the controlled surplus processes of Company 1 resp. Company 2 with L and C. Then, as L is left-continuous, $X_s^1 \neq X_{s-}^1, X_s^2 \neq X_{s-}^2$ only when money is transferred. As the money is only exchanged between the two companies it holds

$$\begin{aligned} X_s^1 - X_{s-}^1 &= X_s^2 - X_{s-}^2 \quad \Leftrightarrow \\ X_s^1 + X_s^2 &= X_{s-}^1 + X_{s-}^2 \end{aligned}$$

and hence,

$$f(X_s^1, X_s^2) = f(X_{s-}^1, X_{s-}^2).$$

As f satisfies the HJB equation (40), by Proposition 5.1 with finite stopping time $\tau \wedge t$,

$$\begin{aligned}
& e^{-\beta(\tau\wedge t)} f(X^1(\tau\wedge t), X^2(\tau\wedge t)) - f(x_1, x_2) \\
&= \int_0^{\tau\wedge t} e^{-\beta\tau\wedge t} \underbrace{\mathcal{L}f(X_{s-}^1, X_{s-}^2)}_{\leq 0} \, ds + M_{\tau\wedge t} \\
&\quad + \int_0^{\tau *} e^{-\beta s} \underbrace{(f_{x_1}(X_{s-}^1, X_{s-}^2) - f_{x_2}(X_{s-}^1, X_{s-}^2))}_{=0} \, dC_s^{21C} \\
&\quad + \int_0^{\tau *} e^{-\beta s} \underbrace{(f_{x_2}(X_{s-}^1, X_{s-}^2) - f_{x_1}(X_{s-}^1, X_{s-}^2))}_{=0} \, dC_s^{12C} \\
&\quad \sum_{X_{s-}^1 \neq X_s^1, X_{s-}^2 \neq X_s^2, s \leq \tau\wedge t} e^{-\beta s} \underbrace{(f(X_s^1, X_s^2) - f(X_{s-}^1, X_{s-}^2))}_{=0} \\
&\quad - \int_0^{\tau\wedge t} a e^{-\beta s} dL_s^1 - \int_0^{\tau\wedge t} (1-a) e^{-\beta s} dL_s^2 \\
&\quad + \int_0^{\tau\wedge t} \underbrace{(a - f_{x_1}(X_{s-}^1, X_{s-}^2))}_{\leq 0} e^{-\beta s} dL_s^{1C} \\
&\quad + \sum_{L_{s+}^1 \neq L_s^1, s < \tau\wedge t} \int_0^{L_{s+}^1 - L_s^1} \underbrace{(a - f_{x_1}(X_s^1 - \alpha, X_s^2))}_{\leq 0} e^{-\beta s} d\alpha \\
&\quad + \int_0^{\tau\wedge t} \underbrace{(1 - a - f_{x_2}(X_{s-}^1, X_{s-}^2))}_{\leq 0} e^{-\beta s} dL_s^{2C} \\
&\quad + \sum_{L_{s+}^2 \neq L_s^2, s < \tau\wedge t} \int_0^{L_{s+}^2 - L_s^2} \underbrace{(1 - a - f_{x_2}(X_{s+}^1, X_s^2 - \alpha))}_{\leq 0} e^{-\beta s} d\alpha, \\
&\leq M_{\tau\wedge t} - \int_0^{\tau\wedge t} a e^{-\beta s} dL_s^1 - \int_0^{\tau\wedge t} (1-a) e^{-\beta s} dL_s^2
\end{aligned}$$

with Proposition 5.1 and Proposition 5.2. The stopped martingale M has expectation zero by the optional stopping theorem. Hence,

$$f(x_1, x_2) \geq E_{x_1, x_2}\Big[\int_0^{\tau\wedge t} a e^{-\beta s} dL_s^1 + \int_0^{\tau\wedge t} (1-a) e^{-\beta s} dL_s^2\Big]$$

and by Monotone Convergence Theorem,

$$\begin{aligned}
f(x_1, x_2) &\geq \lim_{t\to\infty} E_{x_1, x_2}\Big[\int_0^{\tau\wedge t} a e^{-\beta s} \, dL_s^1 + \int_0^{\tau\wedge t} (1-a) e^{-\beta s} \, dL_s^2\Big] \\
&= E_{x_1, x_2}\Big[\int_0^{\tau} a e^{-\beta s} dL_s^1 + \int_0^{\tau} (1-a) e^{-\beta s} \, dL_s^2\Big] = J_{L,C}(x_1, x_2).
\end{aligned}$$

This means f dominates $J_{L,C}$ for all L, C and therefore it follows $f(x_1, x_2) \geq V(x_1, x_2) = V(x_1, x_2)$.

The next step is to show that for the constructed control (L^*, C^*), $f(x_1, x_2) = J_{L^*,C^*}(x_1, x_2)$. This means (L^*, C^*) is optimal. Let X_1^*, X_2^* denote the controlled surplus processes of Company 1 resp. Company 2 with L^* and C^*. For $(x_1, x_2) \in A$, by Proposition 5.1 and Proposition 5.2 it holds $\mathcal{L}f = 0$. Moreover $L^{2*} = 0$, $L^{1*}_{s+} - L^{1*}_s = 0$, therefore

$$
\begin{aligned}
&e^{-\beta(\tau\wedge t)} f(X^{1*}_{\tau\wedge t}, X^{2*}_{\tau\wedge t}) - f(x_1, x_2) \\
&= \int_0^{\tau\wedge t} e^{-\beta\tau\wedge t} \underbrace{\mathcal{L}^{(L^*,C^*)} f(X^1_{s-}, X^2_{s-})}_{=0}\ ds + M_{\tau\wedge t} \\
&\quad + \int_0^{\tau*} e^{-\beta s} \underbrace{(f_{x_1}(X^1_{s-}, X^2_{s-}) - f_{x_2}(X^1_{s-}, X^2_{s-}))}_{=0}\ dC^{21*C}_s \\
&\quad + \int_0^{\tau*} e^{-\beta s} \underbrace{(f_{x_2}(X^1_{s-}, X^2_{s-}) - f_{x_1}(X^1_{s-}, X^2_{s-}))}_{=0}\ dC^{12*C}_s \\
&\quad \sum_{X^1_{s-} \neq X^1_s, X^2_{s-} \neq X^2_s, s \leq \tau\wedge t} e^{-\beta s} \underbrace{(f(X^1_s, X^2_s) - f(X^1_{s-}, X^2_{s-}))}_{=0} \\
&\quad - \int_0^{\tau\wedge t} a e^{-\beta s} dL^{1*}_s - \int_0^{\tau\wedge t} (1-a) e^{-\beta s} \underbrace{dL^{*2}_s}_{=0} \\
&\quad + \int_0^{\tau\wedge t} (a - f_{x_1}(X^1_{s-}, X^2_{s-})) e^{-\beta s} dL^{1*C}_s \\
&\quad + \sum_{L^1_{s+} \neq L^1_s, s < \tau\wedge t} \underbrace{\int_0^{\overbrace{L^{1*}_{s+} - L^{1*}_s}^{=0}} (a - f_{x_1}(X^1_s - \alpha, X^2_s)) e^{-\beta s} d\alpha}_{=0} \\
&\quad + \int_0^{\tau\wedge t} (1 - a - f_{x_1}(X^1_{s-}, X^2_{s-})) e^{-\beta s} \underbrace{dL^{2*C}_s}_{=0} \\
&\quad + \sum_{L^2_{s+} \neq L^2_s, s < \tau\wedge t} \underbrace{\int_0^{\overbrace{L^2_{s+} - L^2_s}^{=0}} (1 - a - f_{x_1}(X^1_{s+} - \alpha, X^2_s)) e^{-\beta s} d\alpha}_{=0} \\
&= M_{\tau\wedge t} - \int_0^{\tau\wedge t} a e^{-\beta s} dL^{1*}_s + \int_0^{\tau\wedge t} (a - f_{x_1}(X^{1*}_{s-}, X^{2*}_{s-})) e^{-\beta s} dL^{1*C}_s,
\end{aligned}
$$

where

$$\begin{aligned}
&\int_0^{\tau\wedge t}(a - f_{x_1}(X^{1*}_{s-}, X^{2*}_{s-}))e^{-\beta s}dL^{1*C}_s\\
&\int_0^{\tau\wedge t}(a - f_{x_1}(X^{1*}_{s}, X^{2*}_{s}))e^{-\beta s}dL^{1*C}_s\\
&\int_0^{\tau\wedge t}(a - f_{x_1}(X^{1*}_{s}, X^{2*}_{s}))e^{-\beta s}1_{X^{1*}_s+X^{2*}_s=m}\ dL^{1*C}_s\\
&= 0
\end{aligned}$$

as $f_{x_1}(x_1,x_2)\big|_{x_1+x_2=m} = a$. Since $(X^{1*}_\tau, X^{2*}_\tau) = (0,0)$

$$\begin{aligned}
&f(x_1,x_2)\\
&= E_{x_1,x_2}\big[e^{-\beta(\tau\wedge t)}f(X^{1*}_{\tau\wedge t}, X^{2*}_{\tau\wedge t})\big] + E_{x_1,x_2}\Big[\int_0^{\tau\wedge t} ae^{-\beta s}dL^{1*}_s\Big]\\
&= E_{x_1,x_2}\big[e^{-\beta t}f(X^{1*}_{t}, X^{2*}_{t})1_{t<\tau}\big] + E_{x_1,x_2}\Big[\int_0^{\tau\wedge t} ae^{-\beta s}dL^{1*}_s\Big].
\end{aligned}$$

Using the Bounded Convergence Theorem and the Monotone Convergence Theorem, it follows

$$\begin{aligned}
f(x_1,x_2) &= E_{x_1,x_2}\big[\lim_{t\to\infty} e^{-\beta t}f(X^{1*}_{t}, X^{2*}_{t})1_{t<\tau}\big] + E_{x_1,x_2}\Big[\lim_{t\to\infty}\int_0^{\tau\wedge t} ae^{-\beta s}dL^{1*}_s\Big]\\
&= E_{x_1,x_2}\Big[\int_0^{\tau} ae^{-\beta s}dL^{1*}_s\Big]
\end{aligned}$$

and

$$f(x_1,x_2) = J_{L^*,C^*}(x_1,x_2).$$

For the case $x_1+x_2 > m$ the strategy L^* pays immediately dividend x_1+x_2-m and $f(x_1,x_2) = x_1+x_2-m+f(\tilde{x}_1,\tilde{x}_2)|_{\tilde{x}_1+\tilde{x}_2=m}$. For $\tilde{x}_1, \tilde{x}_2 \in A$ the previous computation holds and the result follows. □

The optimal value function $V^{2-dim}(X^1, X^2)$ of 2-dimensional optimal dividend problem for two collaborating companies is similar to the optimal value function for the 1-dimensional problem $V^{1-dim}(X)$. Specifically, the relation is $V^{2-dim}(X^1,X^2) = aV^{1-dim}(X^1+X^2)$. Like in the 1-dimensional problem, dividends are only paid out when the sum of the two companies exceeds an optimal barrier. If the sum is below this barrier, no dividend is paid.

6 Local time and Tanaka's formula

This chapter is based on the book of Chung and Williams [6], where more details and the missing proofs can be found. Let W_t denote a Brownian motion. The goal of this chapter to derive a decomposition, known as Tanaka's formula, of the process $|W_t|$ as the sum of another Brownian motion $\hat{W}_t$ and a continuous increasing process L_t

$$|W_t| = \hat{W}_t + L_t.$$

As $|W_t|$ is a submartingale,

$$E\big[|W_t|\big|\mathcal{F}_s\big] \geq \big|E\big[W_t\big|\mathcal{F}_s\big]\big| = \big|W_s\big| \qquad \text{for } \forall s \leq t,$$

it is intuitive that L_t has to be increasing. A more general formula is given for $x \in \mathbb{R}$ with

$$|W_t - x| = \hat{W}_t + L_t(x).$$

The process $L_t(x)$ is called *local time* of Brownian motion and can be expressed as

$$L_t(x) = \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_0^t 1_{(x-\epsilon, x+\epsilon)}(W_s)\, ds = \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \lambda \big\{ s \in [0,t] : B_s \in (x-\epsilon, x+\epsilon) \big\} \tag{50}$$

where λ is the Lebesgue measure. Thus, the local time measures the time the Brownian motion stays in a neighbourhood of x. As the set $\big\{t \in \mathbb{R}_+ : B_t = x\big\}$ has Lebesgue measure zero, it is not obvious that the process $L_t(x)$ doesn't vanish.

6.1 Existence of local time

Theorem 6.1. *For each $t \in \mathbb{R}_+$ and $x \in \mathbb{R}$, it holds almost surely, that*

$$(W_t - x)^+ - (W_0 - x)^+ = \int_0^t 1_{[x,\infty)}(W_s)\, dW_s + \frac{1}{2} L_t(x)$$

where L is defined in (50) with the limit in L^2.

Proof. For $x \in \mathbb{R}$, the function $f_x(y) = (y-x)^+$ is defined. This function is not differentiable but its first two derivatives exist in the sense of generalized functions (Schwartz distributions)

$$\begin{aligned} f_x'(y) &= 1_{[0,\infty)}(y), \\ f_x''(y) &= \delta_x(y) = \lim_{\epsilon \to 0} \frac{1}{2\epsilon} 1_{(x-\epsilon, x+\epsilon)}(y). \end{aligned}$$

A formal application of Itô's formula yields:

$$(W_t - x)^+ - (W_0 - x)^+ = \int_0^t 1_{[x,\infty)}(W_s)\, dW_s + \frac{1}{2} \int_0^t \delta_x(W_s)\, ds.$$

If the Itô formula holds in this case, the last integral exists and coincides with the definition of

$L_t(x)$, then the claim is proven. Now, a function $f_{x\epsilon}$ is defined as

$$f_{x\epsilon}(y) = \begin{cases} 0, & \text{for } y \leq x - \epsilon, \\ \frac{(y-x+\epsilon)^2}{4\epsilon}, & \text{for } x - \epsilon < y < x + \epsilon, \\ y - x, & \text{for } y \geq x + \epsilon \end{cases}$$

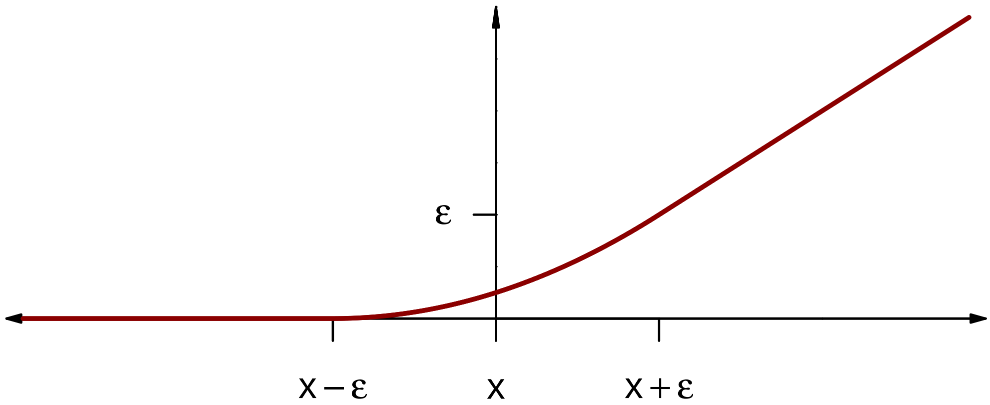

Figure 7: $f_{x\epsilon}$

This construction is a continuous differentiable function, which converges uniformly to $f_x(y)$ for $\epsilon \to 0$. Its first and second derivatives are

$$f'_{x\epsilon}(y) = \begin{cases} 0, & \text{for } y \leq x - \epsilon, \\ \frac{y-x+\epsilon}{2\epsilon}, & \text{for } x - \epsilon < y < x + \epsilon, \\ 1, & \text{for } y \geq x + \epsilon \end{cases}$$

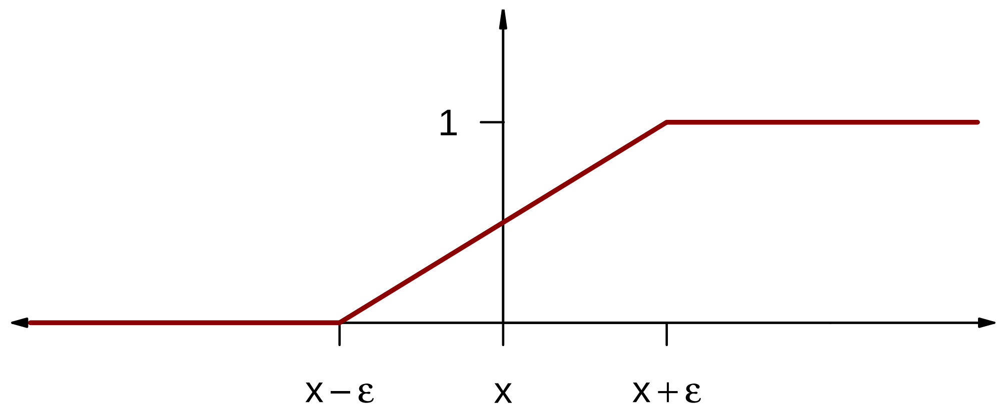

Figure 8: $f'_{x\epsilon}$

and

$$f''_{x\epsilon}(y) = \begin{cases} 0, & \text{for } y \leq x - \epsilon, \\ \frac{1}{2\epsilon}, & \text{for } x - \epsilon < y < x + \epsilon, \\ 0, & \text{for } y \geq x + \epsilon. \end{cases}$$

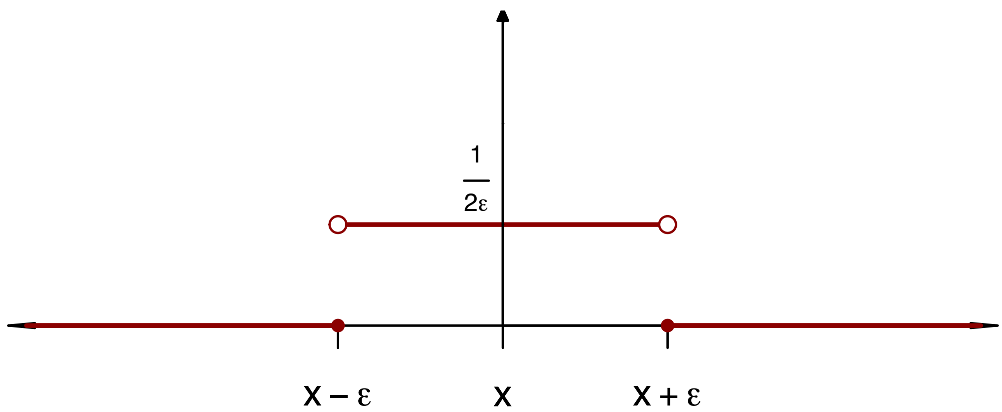

Figure 9: $f''_{x\epsilon}$

The second derivative $f''_{x\epsilon}$ is not defined at $x \pm \epsilon$, therefore it is set to zero. The graphs show $f_{x\epsilon}$ and its derivative

To apply Itôs formula to $f_{x\epsilon}$ a so called *mollifier* is used create sequences of smooth functions approximating this nonsmooth function via convolution. There is a sequence $\phi_n \in \mathcal{C}^\infty$ (mollifier) with compact supports shrinking to $\{0\}$, i.e. support$(\phi_n) \to 0$ for $n \to \infty$, such that

$$g_n = \phi_n * f_{x\epsilon} = \int_{\mathbb{R}} f_{x\epsilon}(y-z)\phi_n(z)\ dz,$$

then

- $g_n \in \mathcal{C}^\infty$,
- $g_n \to f_{x\epsilon}$ uniformly in $\mathbb{R}$,
- $g_n' \to f_{x\epsilon}'$ uniformly in $\mathbb{R}$,
- $g_n'' \to f_{x\epsilon}''$ pointwise except at $x \pm \epsilon$.

The sequence ϕ_n can be defined as follows: $\phi_n(y) = n\phi(ny)$, where $\phi(y) = c\exp\big(-(1-y^2)^{-1}\big)$ for $|y| < 1$ and $\phi(y) = 0$ for $|y| \geq 1$. The constant c is such that $c\int_{-1}^{1}\phi(y)\ dy = 1$. As g_n is smooth enough, Itô's formula can be applied

$$g_n(W_t) - g_n(W_0) = \int_0^t g_n'(W_s)\ dW_s + \frac{1}{2}\int_0^t g_n''(W_s)\ ds. \tag{51}$$

Now, it holds

$$\lim_{n\to\infty} 1_{[0,t]} g_n'(W_s) = 1_{[0,t]} f_{x\epsilon}'(W_s)$$

uniformly on on $\mathbb{R}_+ \times \Omega$ and hence in $L^2(\lambda \times P)$, i.e.

$$\lim_{n\to\infty} E\Big[\int_0^t \big(g_n'(W_s) - f_{x\epsilon}'(W_s)\big)^2\ ds\Big] = 0.$$

With Itôs isometry it holds

$$\lim_{n\to\infty} E\Big[\int_0^t \big(g_n'(W_s) - f_{x\epsilon}'(W_s)\ dW_s\big)^2\Big] = \lim_{n\to\infty} E\Big[\int_0^t \big(g_n'(W_s) - f_{x\epsilon}'(W_s)\big)^2\ ds\Big] = 0.$$

and therefore

$$\lim_{n\to\infty}\int_0^t g_n'(W_s)\ dW_s = \int_0^t f_{x\epsilon}'(W_s)\ dW_s \qquad \text{in } L^2(P).$$

As $P[W_t = x \pm \epsilon] = 0$

$$\lim_{n\to\infty} g_n''(W_t) = f_{x\epsilon}''(W_t) \qquad a.s.$$

and this relation also holds for λ-almost all $s \in \mathbb{R}_+$ a.s. As $|g_n''| \leq \frac{1}{2\epsilon}$ it follows by bounded convergence that

$$\lim_{n\to\infty}\int_0^t g_n''(W_s)\ ds = \int_0^t f_{x\epsilon}''(W_s)\ ds$$

a.s. and in L^2. Thus, by letting $n \to \infty$ in (51), as g_n converges uniformly to $f_{x\epsilon}$ it holds for each x and t, almost surely

$$f_{x\epsilon}(W_t) - f_{x\epsilon}(W_0) = \int_0^t f'_{x\epsilon}(Ws)\, dW_s + \frac{1}{2}\int_0^t \underbrace{f''_{x\epsilon}(W_s)}_{\frac{1}{2\epsilon}1_{(x-\epsilon,x+\epsilon)}(W_s)}\, ds$$

Furthermore,

$$\lim_{\epsilon \to 0} f_{x\epsilon}(W_t) - f_{x\epsilon}(W_0) = f_x(W_t) - f_x(W_0) = (W_t - x)^+ - (W_0 - x)^+$$

in L^2 since $|f_{x\epsilon}(W_t) - f_{x\epsilon}(W_0)| \leq |W_t - W_0|$. Also

$$E\Big[\int_0^t (f'_{x\epsilon}(W_s) - 1_{[x,\infty)}(W_s))^2\, ds\Big] \leq E\Big[\int_0^t 1_{(x-\epsilon,x+\epsilon)}(W_s)\, ds\Big] \leq \int_0^t \frac{2\epsilon}{\sqrt{2\pi s}}\, ds \overset{\epsilon\to 0}{\longrightarrow} 0.$$

where the upper bound in last inequality is the maximal density value $\frac{1}{\sqrt{2\pi s}}$ in the interval $(x-\epsilon, x+\epsilon)$ times the interval length 2ϵ. An illustration of this inequality is given in the graph below. With Itôs isometry

$$\lim_{\epsilon \to 0} \int_0^t f'_{x\epsilon}(W_s)\, dW_s = \int_0^t 1_{[x,\infty)}(W_s)\, dW_s \qquad \text{in } L^2(P).$$

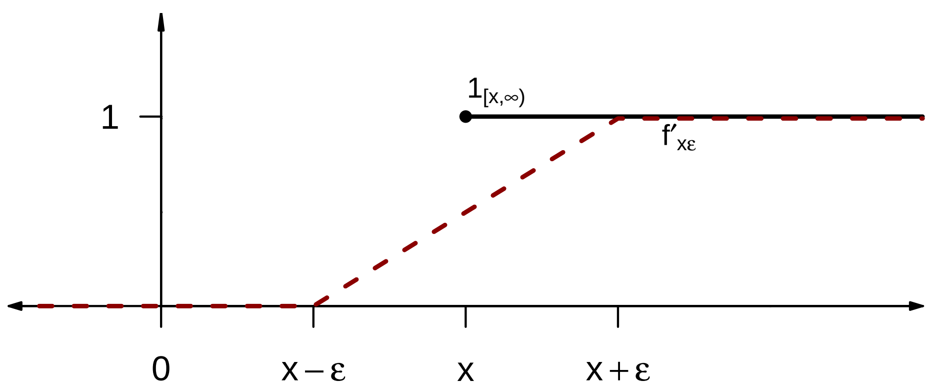

Figure 10: Illustration of the proof idea

It was shown that $\frac{1}{2\epsilon}\int_0^t 1_{(x-\epsilon,x+\epsilon)}(W_s)\, ds$ can be represented as sum of $L(P)^2$-convergent terms and is therefore also $L^2(P)$-convergent. Thus, the local time exists and the claim is proven. □

The local time $L_t(x)$ can be interpreted as the length of time the Brownian motion stays in an interval until time t because

$$\int_0^t 1_{[a,b]}(W_s)\ ds = \int_a^b L_t(x)\ dx.$$

6.2 Tanaka's formula

Theorem 6.2. *For each* (t, x)*, it holds a.s.*

$$|W_t - x| - |W_0 - x| = \int_0^t \operatorname{sgn}(W_s - x)\ dW_s + L_t(x).$$

Here sgn(y) *is* $1, 0,$ *or* -1*, as* y *is greater, equal, or less than zero, respectively.*

Proof. Note that $-W_t$ is again a Brownian motion with local time at $-x$ denoted by $L_t^-(-x)$. Applying Theorem 6.2 to it yields

$$(-W_t + x)^+ - (-W_0 + x)^+ = \int_0^t 1_{[x,\infty)}(-W_s)\ d(-W_s) + \frac{1}{2}L_t^-(-x).$$

As $(-W_t + x)^+ = (W_t - x)^-$ and $L_t^-(-x) = L_t(x)$, because

$$\begin{aligned} L_t^-(-x) &= \lim_{\epsilon\to 0}\frac{1}{2\epsilon}\int_0^t 1_{(-x-\epsilon,-x+\epsilon)}(-W_s)\ ds \\ &= \lim_{\epsilon\to 0}\frac{1}{2\epsilon}\int_0^t 1_{(x-\epsilon,x+\epsilon)}(W_s)\ ds = L_t(x), \end{aligned}$$

it holds

$$(W_t - x)^- - (W_0 - x)^- = -\int_0^t 1_{(-\infty,x]}(W_s)\ dW_s + \frac{1}{2}L_t(x).$$

Adding this to

$$(W_t - x)^+ - (W_0 - x)^+ = \int_0^t 1_{[x,\infty)}(W_s)\ dW_s + \frac{1}{2}L_t(x)$$

yields

$$|W_t - x| + |W_0 - x| = \int_0^t \underbrace{\left(1_{[x,\infty)}(W_s) - 1_{(-\infty,x]}(W_s)\right)}_{=\operatorname{sgn}(W_s - x)}\ dW_s + L_t(x).$$

In this case the definition of *sgn* at zero is not important as $\int_0^t 1_{\{x\}}(W_s)\ dW_s = 0$ a.s. because

$$E\Big[\Big(\int_0^t 1_{\{x\}}(W_s)\ dW_s\Big)^2\Big] = E\Big[\int_0^t 1_{\{x\}}(W_s)^2\ ds\Big] = E\Big[\int_0^t 1_{\{x\}}(W_s)\ ds\Big] = 0.$$

$\square$

Define

$$\hat{W}_t(x) = |W_0 - x| + \int_0^t \operatorname{sgn}(W_s - x)\ dW_s.$$

According to the next Theorem Tanaka's formula holds also pathwise. For the proof see [6].

Theorem 6.3. *For each x, we have a.s.*

$$|W - x| = \hat{W}(x) + L(x)$$

where $\hat{W}(x)$ is a Brownian motion and $L(x)$ is a continuous increasing process with initial value zero. Moreover, almost surely, $L(x)$ can increase only when $|W - x|$ is at zero, i.e.,

$$\int_0^\infty 1_{\{t: W_t \neq x\}} dL_t(x) = 0 \quad a.s.$$

6.3 Reflected Brownian motion at zero

Definition 6.1 (Skorokhod: Problem of reflection)**.** Let $\mathcal{C}$ denote the class of continuous functions from $\mathbb{R}_+$ to $\mathbb{R}$. Given $x \in \mathcal{C}$, a pair (z, y) is called a solution of the problem of reflection for x, denoted by $PR(x)$, if $z \in \mathcal{C}, y \in \mathcal{C}$, and the following three conditions are satisfied

1. $z = x + y$
2. $z \geq 0$
3. $y(0) = 0$, y is increasing on $\mathbb{R}_+$, and $\int_0^\infty z(t)\ dy(t) = 0$.

Lemma 6.1. *Let $x \in \mathcal{C}$ with $x(0) \geq 0$. Then $PR(x)$ has a unique solution given by (z, y) where*

$$z = x + y, \quad y(t) = \max_{0 \leq s \leq t} x^-(s) \text{ for each } t \in \mathbb{R}_+.$$

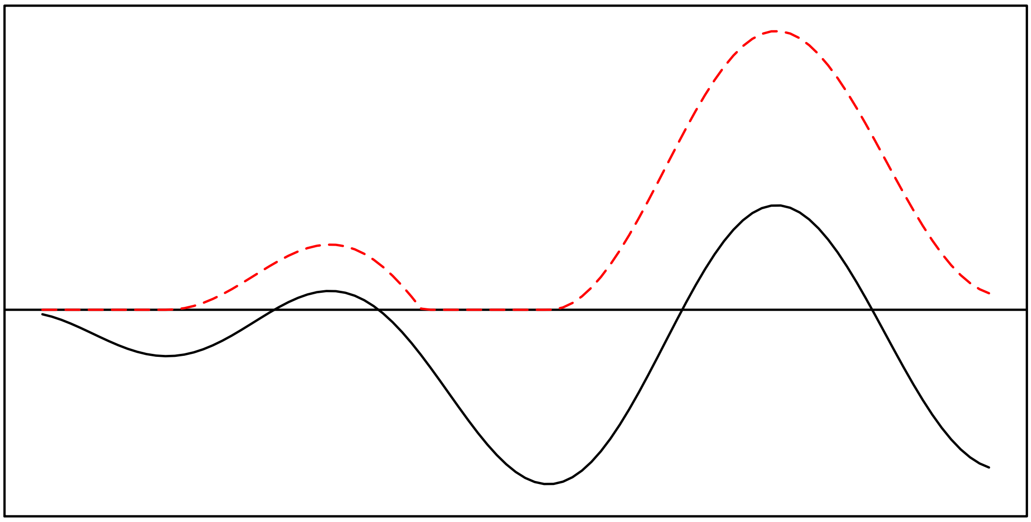

Figure 11: A path (plain line) and its reflection (red dashed line)

The above graph shows the reflection z (red-dashed line) of a path of x.

Tanaka's formula from the previous section is only a special case of the Skorokhod reflection problem. As $(|W|, L)$ is a solution of $PR(\hat{W})$ and $\hat{W}(0) = |B_0| \geq 0$, it follows with the previous lemma that the local time of Brownian motion is

$$L_t = \max_{0 \leq s \leq t} \hat{W}^-(s) \text{for each } t \in \mathbb{R}_+.$$

6.3.1 Optimal dividend problem as reflection problem

The optimal dividend strategy and the optimal value function can be interpreted as the solution of a reflection problem.

The controlled free surplus process was given as

$$X_t = x + \mu\, t + \sigma\, W_t - D_t,$$

where D_t denotes the dividend strategy. The optimal dividend strategy is

$$D_t^* = \sup_{s \leq t} (x + \mu\, t + \sigma\, W_t - m)^+$$

for some barrier m.

According to Peskir (2006) [13] for a Brownian motion with drift

$$Y_t^x = x - \mu \int_0^t \operatorname{sgn}(Y_s^x)\, ds + \sigma W_t,$$

Tanaka's formula gives

$$|Y_t^x| = |x| + \int_0^t \operatorname{sgn}(Y_s^x)\ dY_s^x + L_t^{Y^x}(0) = |x| - \mu t + \sigma \underbrace{\int_0^t \operatorname{sgn}(Y_s^x)\ dW_s}_{=:-\tilde{W}_t} + L_t^{Y^x}(0),$$

where L^{Y^x} denotes the local time for the process Y^x and $-\tilde{W}$ is a standard Brownian motion. Skorokhod's Lemma 6.1 can be applied and provides the unique solution

$$L_t^{Y^x}(0) = \sup_{0\le s\le t}\left(|x| - \mu s - \sigma\tilde{W}_s\right)^- = \sup_{0\le s\le t}\left(-|x| + \mu s + \sigma\tilde{W}_s\right)^+.$$

Thus, assuming that $m \geq x$, the reflection of Y_t^x at barrier m is

$$\begin{aligned} m - |Y_t^x - m| = m - |Y_t^{x-m}| &= m - \left(m - x - \mu t - \sigma\tilde{W}_t + \sup_{0\le s\le t}\left(x - m + \mu s + \sigma\tilde{W}_s\right)^+\right) \\ &= x + \mu t + \sigma\tilde{W}_t - \sup_{0\le s\le t}\left(x + \mu s + \sigma\tilde{W}_s - m\right)^+. \end{aligned}$$

Here, $x + \mu t + \sigma\tilde{W}_t$ can be interpreted as the free surplus process with parameter μ and σ. The process $\sup_{0\le s\le t}\left(x + \mu s + \sigma\tilde{W}_s - m\right)^+$ describes the optimal dividend payments. As the Skorokhod reflection problem provides an unique solution, the local time of the process Y^{x-m} can be interpreted as the optimal dividend strategy.

7 Simulation study

In this chapter the theoretical optimal barrier computed in the previous section is compared with other strategies for numerical examples. This is done by a simulation study using Monte Carlo methods. For more information on numerical methods refer the book of Seydel (2000) [15].

For simplicity the optimal value function is

$$V(x_1, x_2) = \sup_{(L,C)\in\pi_{(x_1,x_2)}} J_{L,C}(x_1, x_2),$$

with the performance criterion

$$J_{L,C}(x_1, x_2) = E_{x_1,x_2}\Big[\int_0^{\tau} e^{-\beta s}\ dL_s^1 + \int_0^{\tau} e^{-\beta s}\ dL_s^2\Big].$$

7.1 Simulation 1: Stable market

In this chapter the optimal strategy is compared in a model with the following parameter for the surplus processes of the two companies:

Company 1	x_1	μ_1	σ_1
surplus X_1	0.05	0.03	0.02
Company 2	x_2	μ_2	σ_2
surplus X_2	0.05	0.02	0.03

Table 1: Simulation 1: Model parameter the 2-dim model

Hence, the two free surplus processes are

$$X_t^1 = 0.05 + 0.03\ t + 0.02\ W_t^1,$$
$$X_t^2 = 0.05 + 0.02\ t + 0.03\ W_t^2.$$

Both companies generate profit in average and the volatility is moderate. The discount factor is $\beta = \log 1.1$, which corresponds to an annual interest rate of 10%. The initial capital for both companies is 0.05. The dividend is paid out until both companies go ruin simultaneously or to infinity otherwise. As an infinite time interval is not possible for numerical computations, the time needs to be cut. Therefore the time interval $[0, T]$ for $T = 1000$ is chosen. This time horizon seems to be a good idea, because of the discount factor the dividend after 1000 years is not significant anymore. Furthermore, a discretization of time is needed. In this simulation the interval $[0, 1000]$ is divided into $N = 100000$ parts, which corresponds to time steps of size $\Delta t = 0.01$.

7.1.1 Comparison of different barriers

Now, the optimal barrier strategy presented by [11] is computed and then compared with different barriers. The optimal barrier for the model is

$$m = \frac{\sigma_1^2+\sigma_2^2}{\sqrt{(\mu_1+\mu_2)^2+2\beta(\sigma_1^2+\sigma_2^2)}} \log\left(\frac{-(\mu_1+\mu_2)-\sqrt{(\mu_1+\mu_2)^2+2\beta(\sigma_1^2+\sigma_2^2)}}{-(\mu_1+\mu_2)+\sqrt{(\mu_1+\mu_2)^2+2\beta(\sigma_1^2+\sigma_2^2)}}\right) = 0.092888$$

and the theoretical optimal value function given in (46) is

$$V(0.05, 0.05) = 0.5317147.$$

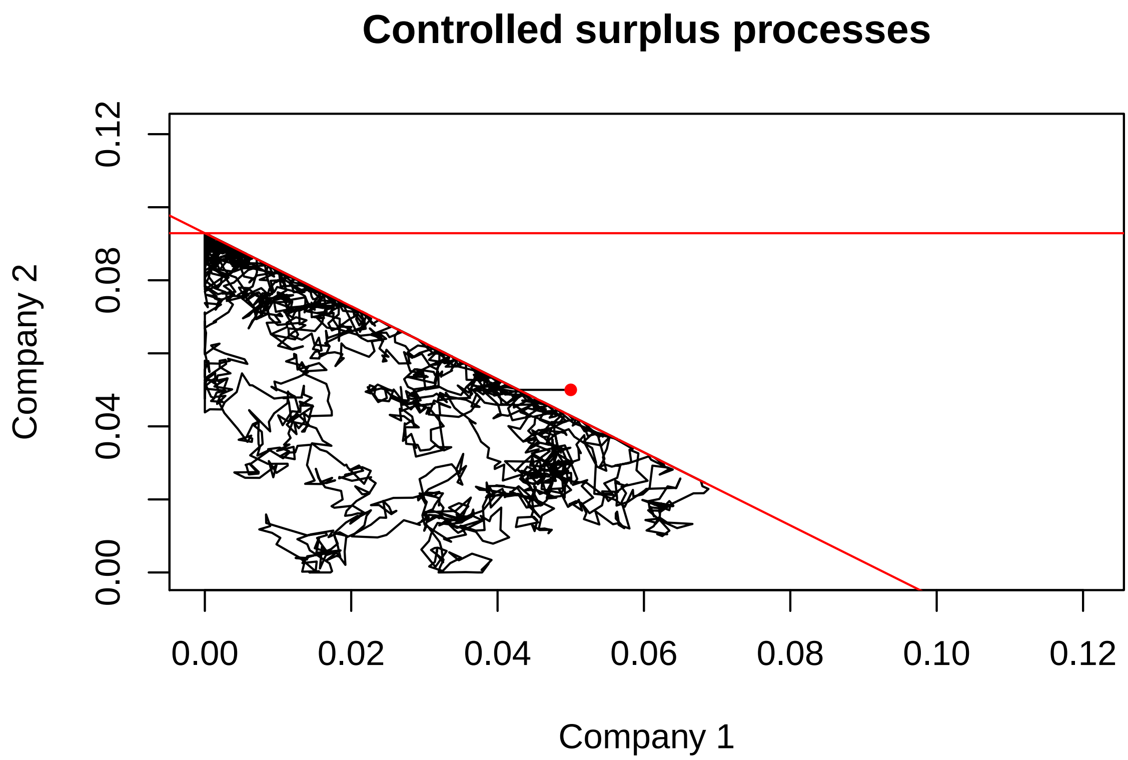

Figure 12: 2-dim controlled surplus simulation

In the previous illustration one trajectory of both controlled surplus processes is simulated for 10 years following this model. Here, the optimal barrier strategy is used to pay out dividends. One can see, that if $x_1 = 0$ or $x_2 = 0$ the other company transfers money. When the diagonal red line, where $x_1 + x_2 = m$, is hit dividends are paid out. The red point indicates the initial capital.

The corresponding paid out accumulated dividends are given in the plot below.

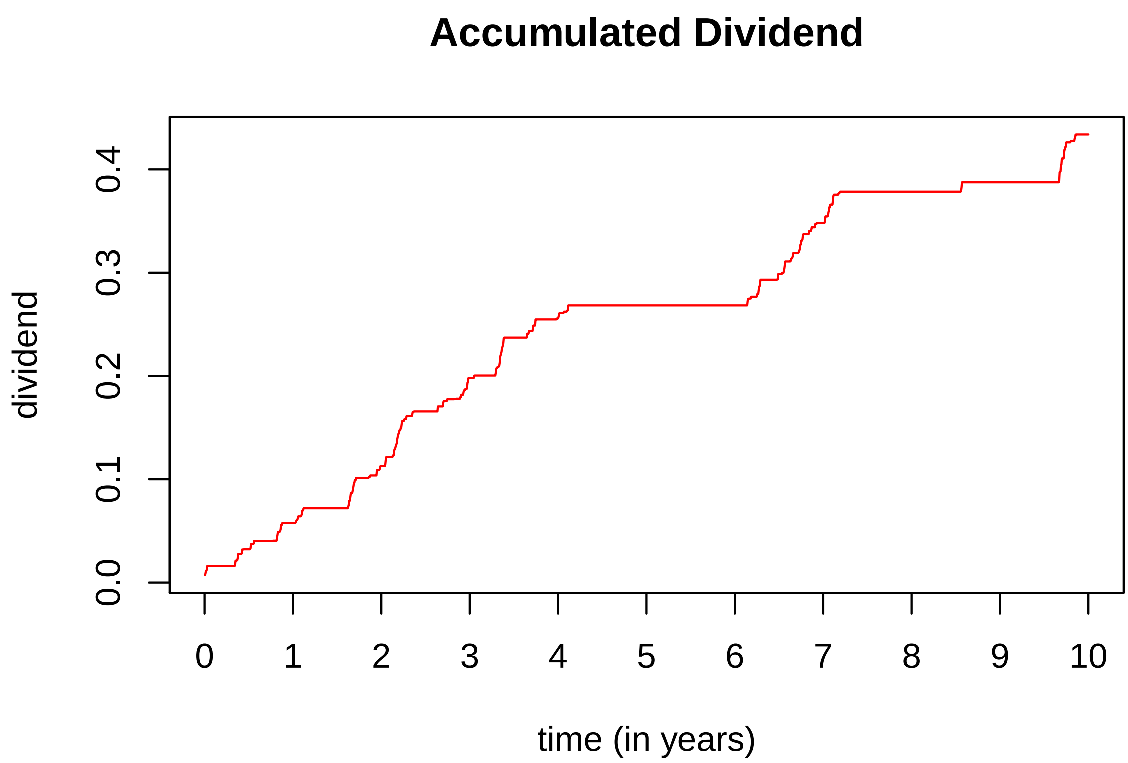

Figure 13: Paid out dividend

A Monte Carlo simulation is implemented to compare strategies, described by Figure 4-6, with different barriers. In each of $n = 10000$ simulations the discounted dividends are summed up, then the mean over all simulations is computed. This method also approximates the value function (when using the optimal barrier). In the next graph, the black points show the value of the expected discounted dividend for the barrier given on the x-axis. For the simulation the different barriers $m + i \cdot 0.01$ for $i \in \{-9, -8, \ldots, 39, 30\}$ are used. The red point indicates the theoretical value function

$$V(0.05, 0.05) = 0.5317147.$$

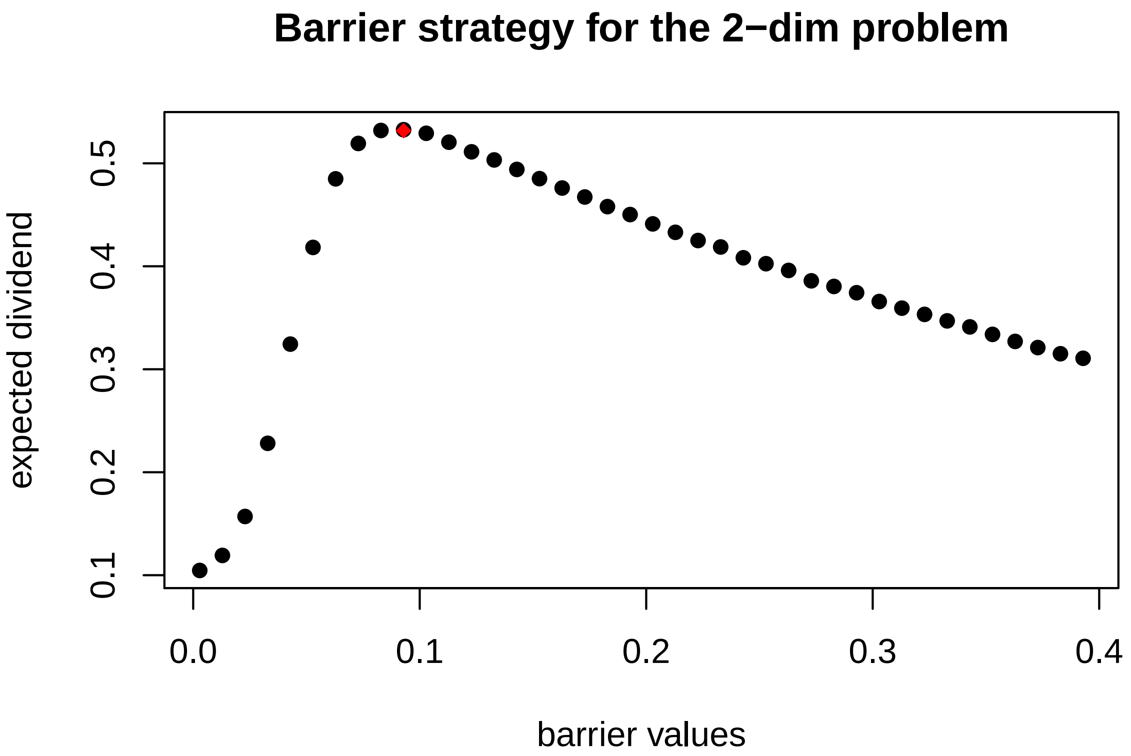

Figure 14: Performance comparison for different barriers

One can observe that the calculated barrier is also optimal compared to other barriers in a simulation. The table below lists the simulated values for all barrier, where the values for the optimal barrier are marked in bold.

The relative deviation of the simulated optimal value V_{MC} and the real theoretical optimal value V is

$$\left|\frac{V_{MC}(0.05, 0.05) - V(0.05, 0.05)}{V(0.05, 0.05)}\right| = \frac{0.5326389 - 0.5317147}{0.5317147} = 0.173815\%$$

This shows that the performance of Monte Carlo simulation is quite satisfactory.

	barrier	exp. dividend
1	0.002888	0.104600
2	0.012888	0.119218
3	0.022888	0.157026
4	0.032888	0.228076
5	0.042888	0.324411
6	0.052888	0.418333
7	0.062888	0.484921
8	0.072888	0.519276
9	0.082888	0.531880
10	**0.092888**	**0.532639**
11	0.102888	0.529158
12	0.112888	0.520501
13	0.122888	0.511166
14	0.132888	0.503281
15	0.142888	0.494051
16	0.152888	0.485113
17	0.162888	0.476000
18	0.172888	0.467278
19	0.182888	0.457930
20	0.192888	0.450215
21	0.202888	0.441159
22	0.212888	0.432998
23	0.222888	0.425021
24	0.232888	0.418743
25	0.242888	0.408267
26	0.252888	0.402513
27	0.262888	0.395969
28	0.272888	0.385891
29	0.282888	0.380418
30	0.292888	0.374326
31	0.302888	0.365788
32	0.312888	0.359335
33	0.322888	0.353267
34	0.332888	0.347004
35	0.342888	0.341147
36	0.352888	0.333863
37	0.362888	0.327036
38	0.372888	0.321062
39	0.382888	0.315071
40	0.392888	0.310655

Table 2: Simulation 1: Results

7.1.2 Optimal strategy II

The value function of the optimal dividend problem is unique as shown in Chapter 5, whereas the optimal strategy is not. Another possibility would be to transfer continuously money to or from Company 1 to keep the surplus of Company 2 at zero. This procedure corresponds to combining two companies into one surplus process.

$$\begin{aligned} 0 &\overset{(!)}{=} X_t^2 = \bar{X}_t^2 + C_t^{12} - C_t^{21} - \underbrace{L_t^2}_{=0} \\ &\Leftrightarrow C_t^{21} - C_t^{12} = \bar{X}_t^2 \end{aligned}$$

Company 2 never pays dividends as its surplus process is always zero. Therefore, the controlled surplus process for Company 1 is

$$X_t^1 = \bar{X}_t^1 + \bar{X}_t^2 - L_t^1 = x_1 + \mu_1 t + \sigma_1 W_t^1 + x_2 + \mu_2 t + \sigma_2 W_t^2 - L_t^1.$$

As W^1 and W^2 are two independent Brownian motions, it holds

$$\sigma_1 W_t^1 + \sigma_2 W_t^2 = \sqrt{\sigma_1^2 + \sigma_2^2}\, W_t$$

where W is another Brownian motion. Therefore,

$$X_t^1 = (x_1 + x_2) + (\mu_1 + \mu_2)t + \sqrt{\sigma_1^2 + \sigma_2^2}\, W_t - L_t^1.$$

The 2-dimensional problem is now transformed into a 1-dimensional problem as described in Section 4.2 for

$$\begin{aligned} x &:= x_1 + x_2 = 0.1 \\ \mu &:= \mu_1 + \mu_2 = 0.05 \\ \sigma &:= \sqrt{\sigma_1^2 + \sigma_2^2} = 0.0361 \end{aligned}$$

and the corresponding results can be used. The 1-dimensional optimal dividend problem for a company with free surplus process

$$X_t = 0.1 + 0.05\, t + 0.0361\, W_t$$

is considered. It follows that optimal barrier is again

$$m = \frac{\sigma^2}{\sqrt{\mu^2 + 2\beta\sigma^2}} \log\left(\frac{-\mu - \sqrt{\mu^2 + 2\beta\sigma^2}}{-\mu + \sqrt{\mu^2 + 2\beta\sigma^2}}\right) = 0.092888$$

and the theoretical optimal value function given in Theorem 4.3 is

$$V(0.05, 0.05) = 0.5317147.$$

In the following plot one trajectory of Company 1's controlled surplus process simulated for 10 years is illustrated. The controlled surplus process of Company 2 is zero.

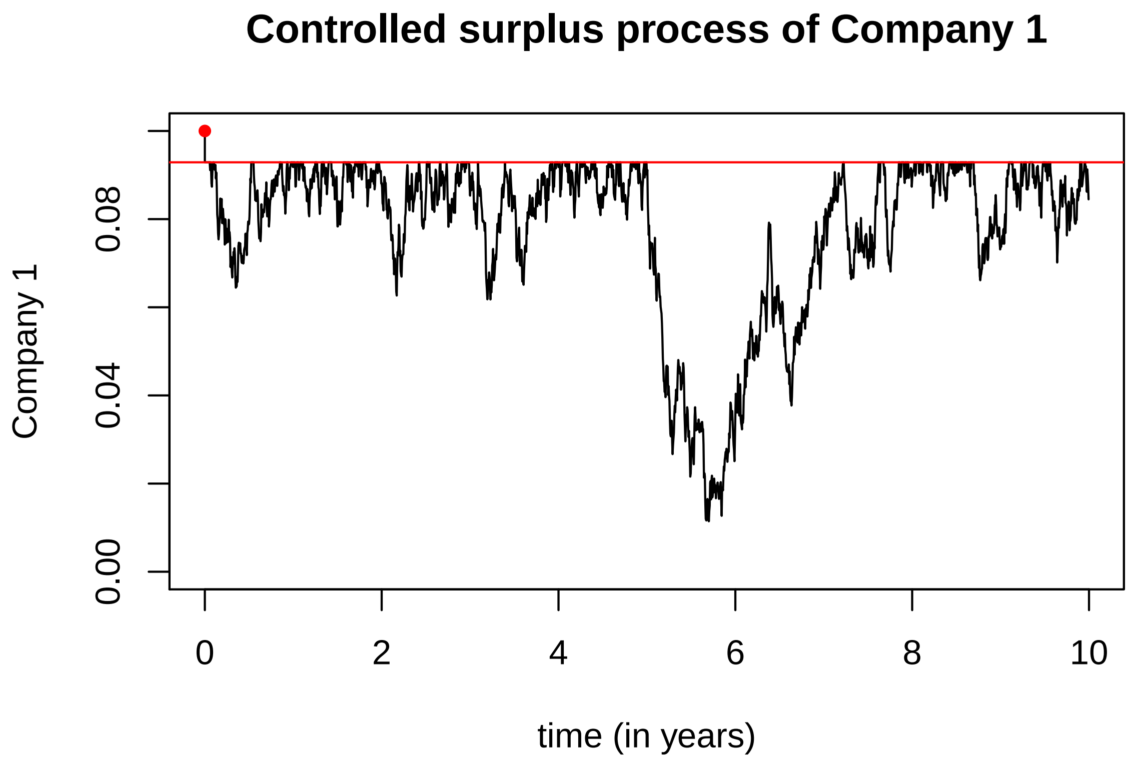

Figure 15: Controlled surplus simulation of Company 1

The red point indicates the initial capital $x_1 + x_2$. Dividends are only paid out when free surplus processes with drift $\mu_1 + \mu_2$, volatility $\sigma_1^2 + \sigma_2^2$ and initial capital $x_1 + x_2$ is above the optimal barrier, which is indicated by the red horizontal line.

7.1.3 No collaboration agreement

It is assumed that there is no collaboration contract between the two companies. This means that they are not liable for each others losses and only maximize their own expected dividend. It is possible that one company goes ruin while the other one survives. The ruin time of Company 1 is denoted by τ_1 and the ruin time of Company 2 by τ_2. Therefore, both optimal dividend problems are considered separately and the optimal value function of Company 1 V^{Co1} and Company 2 V^{Co2} given by Theorem 4.3 are added in the end. This means the following problem is considered

$$V^{sep}(x_1, x_2) = \sup_{L^1} E_{x_1}\Big[\int_0^{\tau_1} e^{-\beta s} dL_s^1\Big] + \sup_{L^2} E_{x_2}\Big[\int_0^{\tau_2} e^{-\beta s} dL_s^2\Big] = V^{Co1}(x_1) + V^{Co2}(x_2).$$

Using Theorem 4.3 for both companies one gets

$$V^{sep}(0.05, 0.05) = V^{Co1}(0.05) + V^{Co2}(0.05) = 0.3148879 + 0.164547 = 0.4794349$$

Comparing this value with the optimal value of the collaborating case $V(0.05, 0.05) = 0.5317147$ one comes to the conclusion that it is advantageous to help each other out.

The question arises if collaboration is reasonable in any case or if there are cases where the companies should only be responsible for themselves. If one company is performing badly compared to the other one (low drift combined with high volatility), then it is better not to collaborate. The bad company would impair the good company as it has to come up for its losses. In this case it is better to let the bad company go ruin, such that the good company can pay out more dividends instead of helping the other one to survive. In the following graph the same model as before is considered for different volatility parameter for Company 1 σ_1. The other parameters are fixed.

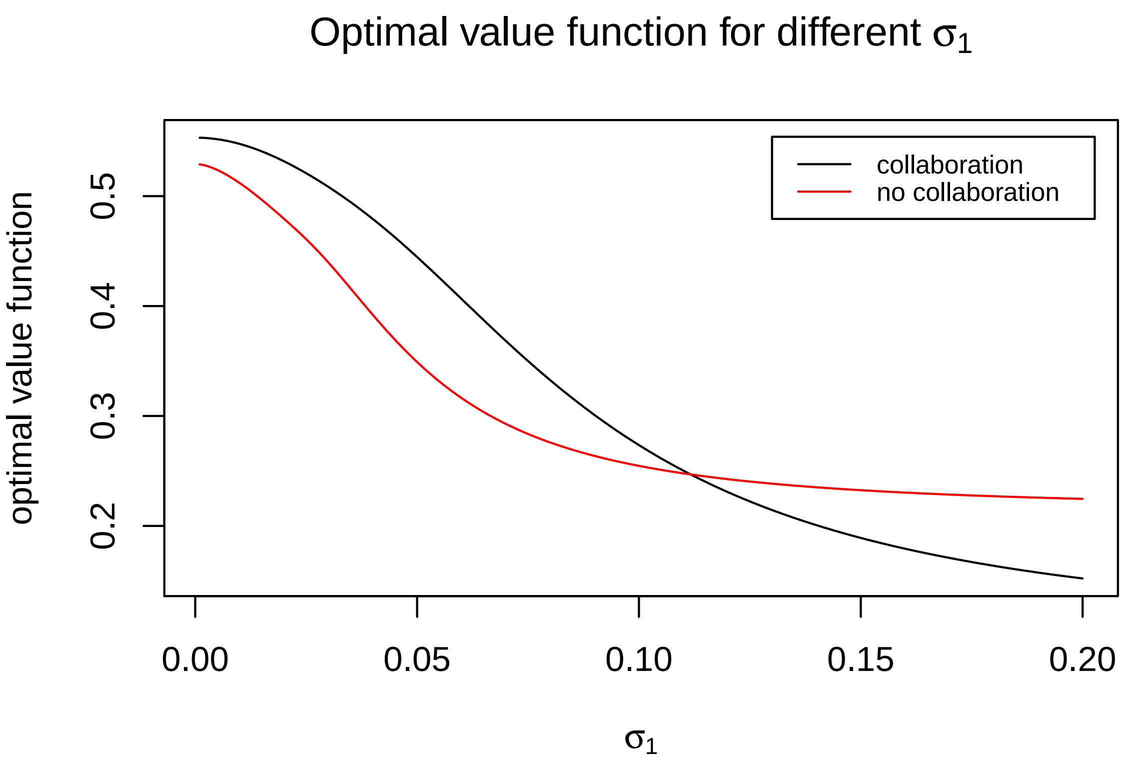

Figure 16: Collaboration versus no collaboration for σ_1

The more insecure (volatile) the surplus of Company 1 is, the better it is to operate separately.

In the next graph the same model as before is considered for different drift parameter for Company 1 μ_1. The other parameters are fixed.

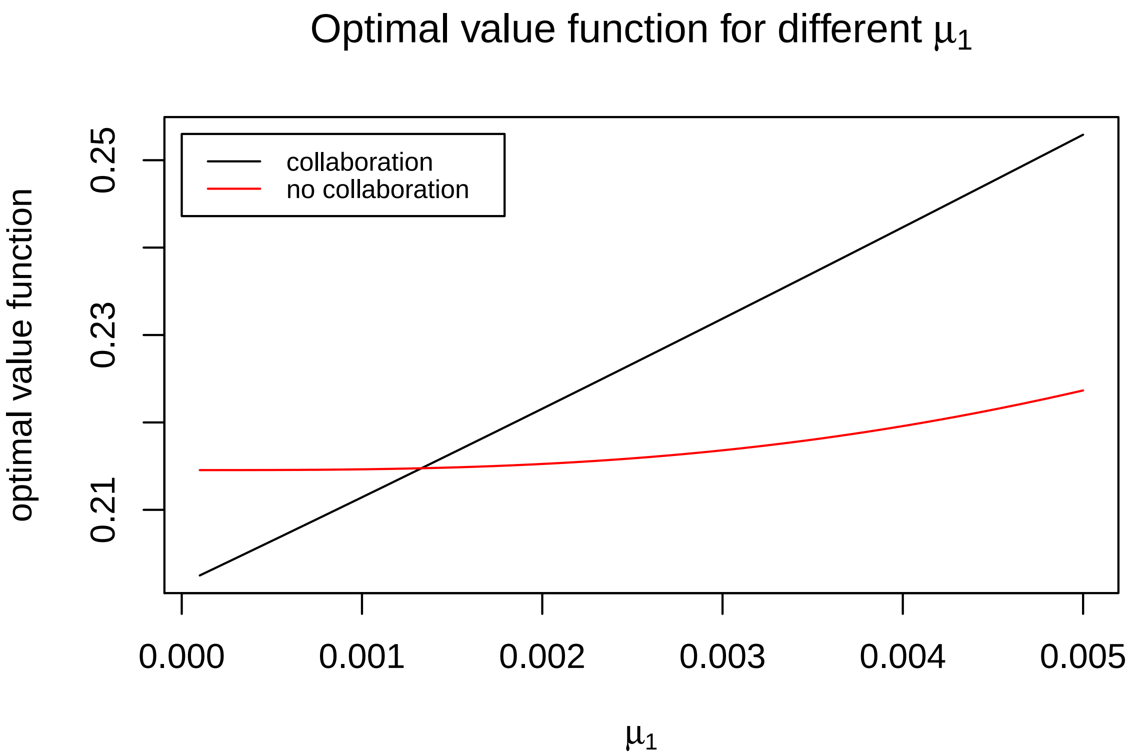

Figure 17: Collaboration versus no collaboration for μ_1

The principle is the same as before. If the drift parameter is low, it is more likely that losses occur and the other company has to balance them out instead of paying dividends to their shareholders.

7.2 Simulation 2: Stressed market

In this simulation a more stressed model is considered. Company 1 performs as before, but Company 2 generates a loss in average. Moreover, its surplus is more volatile than before. As both companies are liable for each others losses, Company 1 has to help out more often. It is analysed how the optimal strategy performs in this setting, when transferring money is a bigger issue. To avoid early ruin a higher initial capital is chosen: for Company 1 $x_1 = 0.1$ and for Company 2 $x_2 = 0.1$. To sum up, the following parameters are used:

Company 1	x_1	μ_1	σ_1
surplus X_1	0.1	0.03	0.02
Company 2	x_2	μ_2	σ_2
surplus X_2	0.1	-0.01	0.05

Table 3: Simulation 2: Model parameter the 2-dim model

Hence, the two free surplus processes are

$$\begin{aligned} X_t^1 &= 0.1 + 0.03\ t + 0.02\ W_t^1, \\ X_t^2 &= 0.1 - 0.01\ t + 0.05\ W_t^2. \end{aligned}$$

As in the first simulation, the discount factor is $\beta = \log 1.1$ and the time interval $[0, T]$ for $T = 1000$ is chosen. In this simulation the interval $[0, 1000]$ is divided into $N = 100000$ parts, which corresponds to time steps of size $\Delta t = 0.01$.

7.2.1 Comparison of different barriers

The optimal barrier for this model is

$$m = \frac{\sigma_1^2 + \sigma_2^2}{\sqrt{(\mu_1 + \mu_2)^2 + 2\beta(\sigma_1^2 + \sigma_2^2)}} \log \left(\frac{-(\mu_1 + \mu_2) - \sqrt{(\mu_1 + \mu_2)^2 + 2\beta(\sigma_1^2 + \sigma_2^2)}}{-(\mu_1 + \mu_2) + \sqrt{(\mu_1 + \mu_2)^2 + 2\beta(\sigma_1^2 + \sigma_2^2)}} \right) = 0.1450075.$$

It is notable that the optimal barrier is now higher than in the previous model. As Company 2 is performing worse than in the last chapter, the barrier needs to be higher to maintain a buffer. Otherwise both companies could go ruin in the next moment as the risk of a loss is higher. The next illustration shows one trajectory of both controlled surplus processes simulated for 10 years.

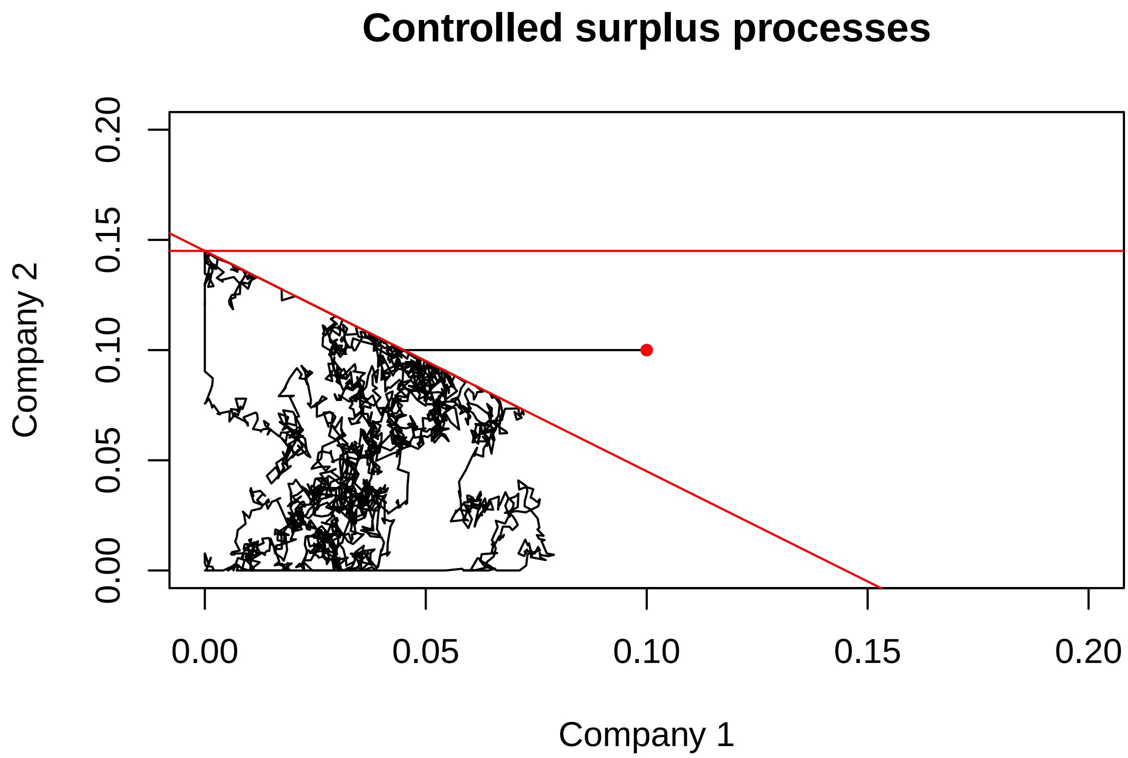

Figure 18: 2-dim controlled surplus simulation

In this case one can observe that the surplus Company 2 is zero more often, which means that Company 1 had to transfer money to compensate the losses of Company 2. Nevertheless, in the end both companies go ruin, as the trajectory reaches the point $(0, 0)$. The corresponding paid out accumulated dividends are visible in the next plot. Here, the accumulated dividend is constant in the end because no dividends are paid out anymore. This is either because both companies went ruin or because their surplus is low and hence below the barrier.

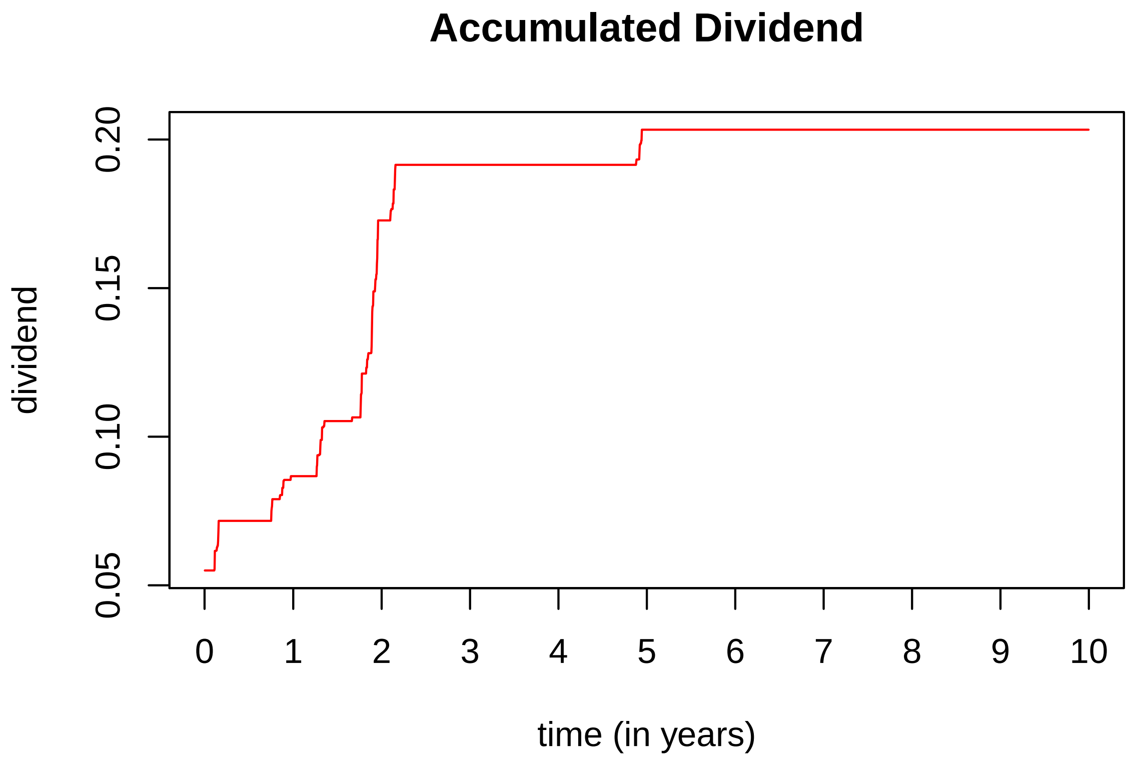

Figure 19: Paid out dividend

Again, a Monte Carlo simulation is implemented to compare strategies, described by Figure 4-6, for different barriers. In the next graph, the black points show the value of the expected discounted dividend for the barrier given on the x-axis. For the simulation the different barriers $m + i \cdot 0.01$ for $i \in \{-14, -8, \ldots, 39, 25\}$ are used. The red point indicates the theoretical value function given in (46), calculated as

$$V(0.1, 0.1) = 0.2648337.$$

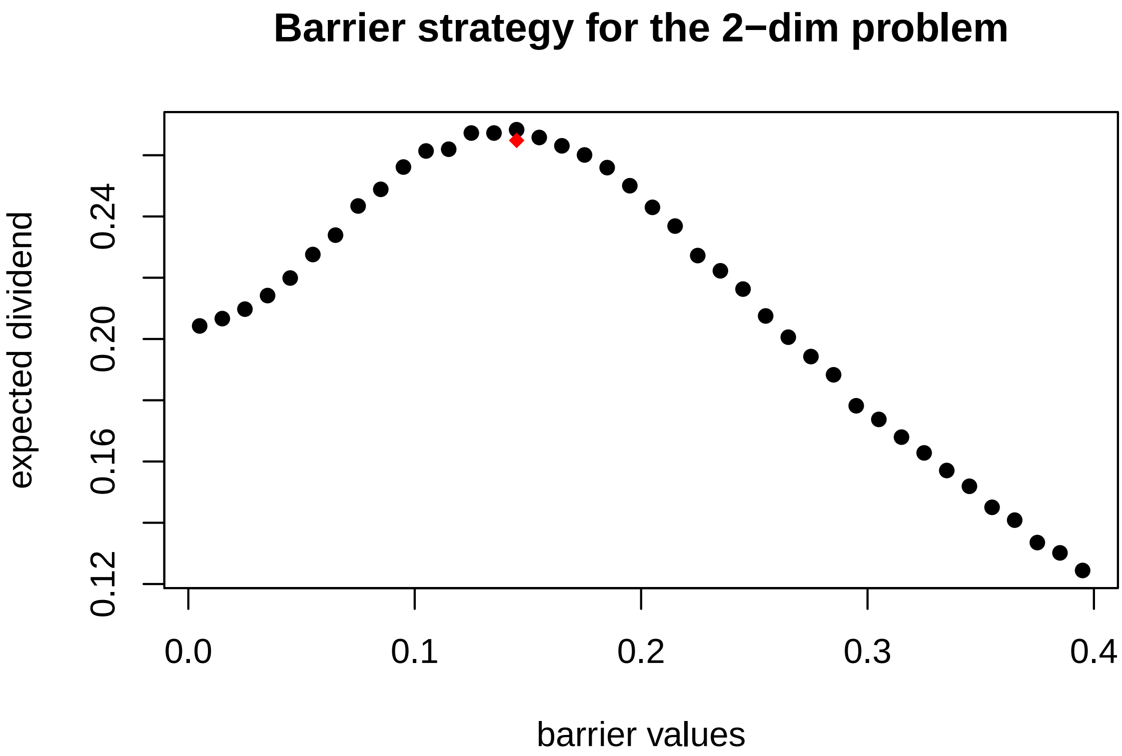

Figure 20: Performance comparison for different barriers

One can observe that the calculated barrier is also optimal compared to other barriers in a simulation. The table below lists the simulated values for all barrier, where the values for the optimal barrier are marked in bold.

The relative deviation of the simulated optimal value V_{MC} and the real theoretical optimal value V is

$$\left| \frac{V_{MC}(0.1, 0.1) - V(0.1, 0.1)}{V(0.1, 0.1)} \right| = \frac{0.2683334 - 0.2648337}{0.2648337} = 1.321471\%$$

The performance of Monte Carlo simulation is worse than in the previous model. The approximation is still satisfactory and the accuracy could be improved by increasing the number of simulations.

	barrier	exp. dividend
1	0.005007	0.204265
2	0.015007	0.206657
3	0.025007	0.209746
4	0.035007	0.214177
5	0.045007	0.219907
6	0.055007	0.227580
7	0.065007	0.233908
8	0.075007	0.243422
9	0.085007	0.248868
10	0.095007	0.256139
11	0.105007	0.261388
12	0.115007	0.261954
13	0.125007	0.267251
14	0.135007	0.267242
15	**0.145007**	**0.268333**
16	0.155007	0.265797
17	0.165007	0.263074
18	0.175007	0.260086
19	0.185007	0.255955
20	0.195007	0.250046
21	0.205007	0.242980
22	0.215007	0.236861
23	0.225007	0.227245
24	0.235007	0.222263
25	0.245007	0.216294
26	0.255007	0.207520
27	0.265007	0.200595
28	0.275007	0.194272
29	0.285007	0.188308
30	0.295007	0.178197
31	0.305007	0.173750
32	0.315007	0.167956
33	0.325007	0.162817
34	0.335007	0.157062
35	0.345007	0.151917
36	0.355007	0.145047
37	0.365007	0.140861
38	0.375007	0.133538
39	0.385007	0.130173
40	0.395007	0.124427

Table 4: Simulation 2: Results

7.2.2 Optimal strategy II

Analogously to the first simulation example, it is possible to transfer the 2-dimensional problem into a 1-dimensional problem analogously to the derivation in Section 7.1.2. The free surplus process

$$X_t = 0.2 + 0.02\ t + 0.05385165\ W_t$$

in the 1-dimensional optimal dividend problem is considered. The optimal barrier is $m = 0.145007$. The following plot shows a surplus trajectory of Company 1 after absorbing the free surplus process of Company 2.

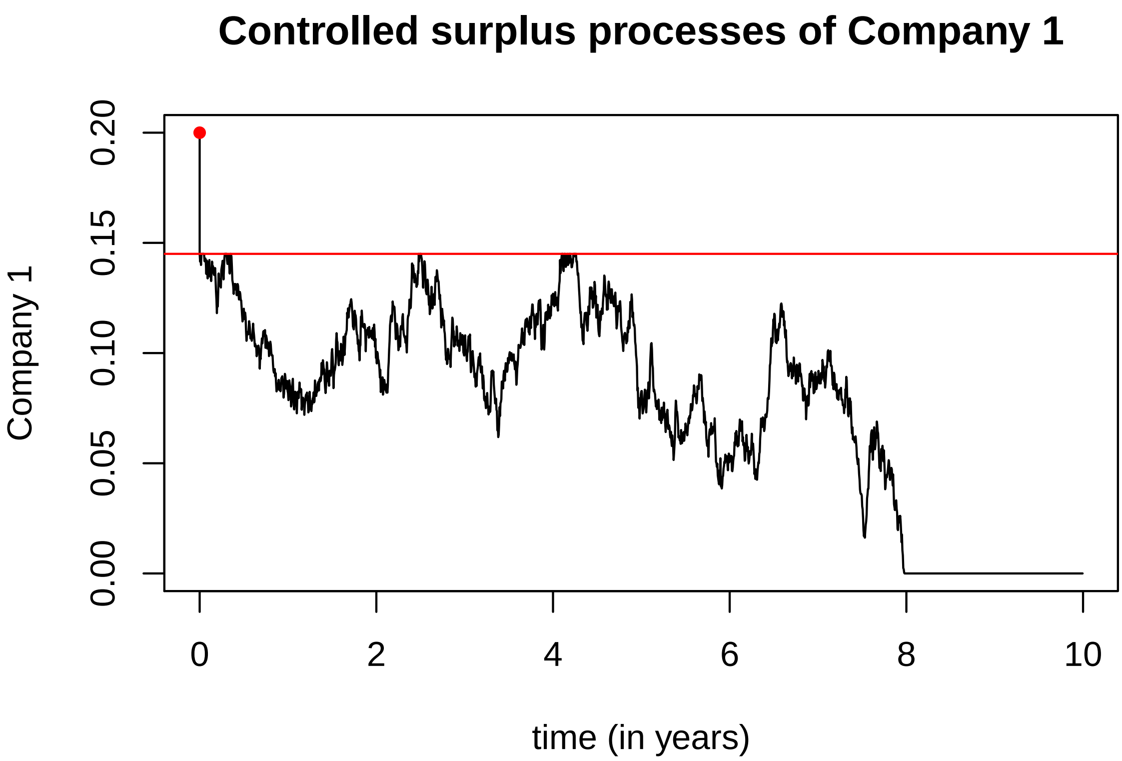

Figure 21: Controlled surplus simulation of Company 1

7.2.3 No collaboration contract

As Company 2 has a negative drift parameter, and will therefore produce a loss in average it is better not to collaborate. In this case Company 1 doesn't have to compensate Company 2's losses and can pay out more dividends.

7.3 Simulation 3: Switching from collaboration to no collaboration

The following parameter set is considered

Company 1	x_1	μ_1	σ_1
surplus X_1	0	0.03	0.02
Company 2	x_2	μ_2	σ_2
surplus X_2	0.05	0.002	0.03

Table 5: Simulation 3: Model parameter of the 2-dim model

The two free surplus processes are

$$X_t^1 = 0.03\ t + 0.02\ W_t^1,$$
$$X_t^2 = 0.05 + 0.002\ t + 0.03\ W_t^2.$$

Again, the discount factor is $\beta = \log 1.1$ and the time interval $[0, T]$ for $T = 500$ is chosen. In this simulation the interval $[0, 500]$ is again divided into $N = 100000$ parts, which corresponds to time steps of size $\Delta t = 0.005$.

Company 1 is performing very well compared to Company 2, which only has a small expected surplus. But Company 1 on the other hand has no initial capital, i.e. ruin would be immediate. Therefore, both companies agree on a collaboration following the strategy illustrated by Figure 4-6, such that Company 1 can start its business and then balance each others losses. Now, the scenario is considered where the companies are allowed to switch to "no collaboration" at some point. This can also be interpreted as disregarding the agreement and "betraying" the other one.

The following plot shows the optimal value function for collaboration and the sum of the optimal value functions for no collaboration dependent on the initial capital of Company 1 x_1.

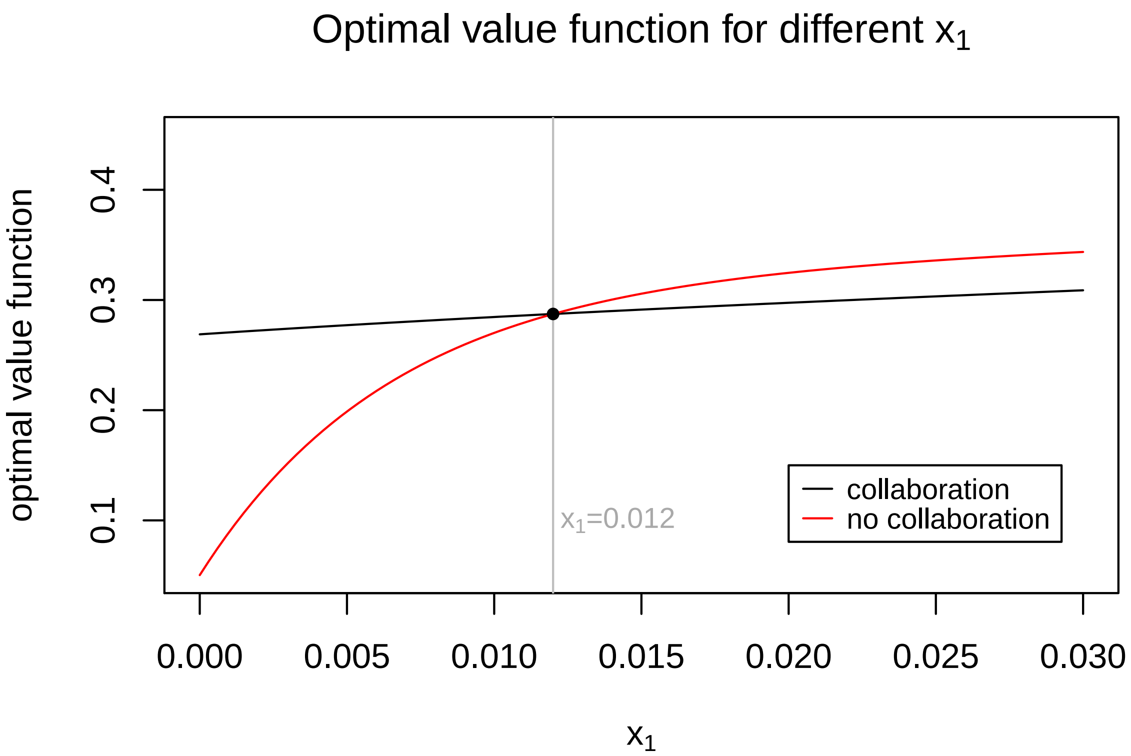

Figure 22: Optimal value functions dependent on x_1

As one can see in the graph the expected discounted dividend for $x_1 = 0$ and $x_2 = 0.05$, if there is collaboration with optimal barrier $m = 0.1063$, is

$$V(0, 0.05) = 0.2688.$$

Without cooperation the expected discounted dividend is

$$V^{Co1}(0) + V^{Co2}(0.05) = V^{Co2}(0.05) = 0.001.$$

Therefore, it is recommended to collaborate in this scenario. According to this illustration one can see that Company 1 won't collaborate if it's initial surplus is more than 0.012

$$\begin{aligned} V(x_1, 0.05) = V^{Co1}(x_1) + V^{Co2}(0.05) \qquad &\Leftrightarrow \\ x_1 = 0.012. \end{aligned}$$

This brings up the question, if there is an incentive to terminate the cooperation after Company 1 reaches a certain surplus level. After hitting this level, both companies will operate optimally on their own. The idea is that the weaker Company 2 won't get any help from its partner anymore and will go ruin sooner. Company 1 on the other hand doesn't have to "waste" its positive surplus to eventually save Company 2. The value function for this problem will be denoted by

V^{switch}. The ruin time of Company 1 and Company 2 are denoted by τ_1 and τ_2, respectively. Simultaneous ruin time within a collaboration contract is τ. The time when the required capital for Company 1 is reached to switch to no collaboration is $\bar{\tau}$. The optimal dividend strategy for collaboration described by Figure 4-6 is denoted by (L^1, L^2). Without collaboration the optimal strategies for Company 1 and Company 2 is $\tilde{L}^1$ resp. $\tilde{L}^2$ given by formula (36).

$$\begin{aligned} V^{switch}(0, 0.05) = & \sup_{\bar{\tau}} E_{x_1,x_2}\Bigg[\int_0^{\tau\wedge\bar{\tau}} e^{-\beta s}\ dL_s^1 \\ & + \int_0^{\tau\wedge\bar{\tau}} e^{-\beta s}\ dL_s^2 + e^{-\beta\bar{\tau}}\left(\int_{\bar{\tau}}^{\tau_1} e^{-\beta s}\ d\tilde{L}_s^1 + \int_{\bar{\tau}}^{\tau_2} e^{-\beta s}\ d\tilde{L}_s^2\right) 1_{\bar{\tau}<\tau}\Bigg] \end{aligned}$$

Monte Carlo simulations with 50000 trajectories are implemented for different surplus levels, which have to be reached such that Company 1 terminates the cooperation, to find an estimation for $V^{switch}(0, 0.05)$. If the free surplus level of Company 1 hits the switching levels $i \cdot 0.001$ for $i \in \{0, 1, 2, \ldots, 100\}$ they continue operating without each other. In the following graph the black points show the value of the expected discounted dividend for the switching level given on the x-axis.

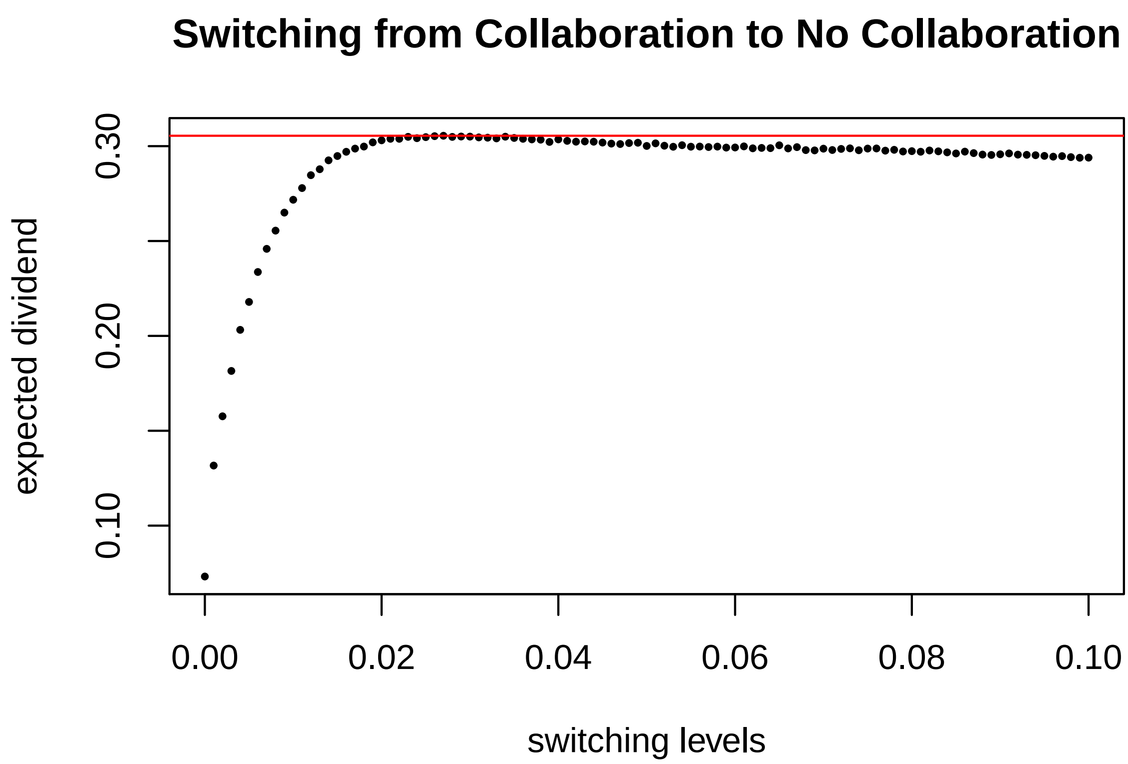

Figure 23: Switching from collaboration to no collaboration

As one can see in the plot, it is best to work together until Company 1 has reached the free surplus level of approximately 0.027 and then it should abandon its partner to achieve a value function of approximately 0.3055.

Obviously, there are other transferring and dividend strategies, which additionally allow switching from collaboration to no collaboration, such that the expected dividend strategy is higher. Another strategy, for example, would be if Company 2 agrees to transfer its initial capital to Company 1 at time 0 and goes ruin itself as a consequence. Then Company 1 operates alone until its own ruin time. The discounted expected dividends of this strategy, is the optimal value function for Company 1

$$V^{Co1}(0.05) = 0.3149.$$

8 Conclusion

In this diploma thesis the problem of finding an optimal dividend strategy, in the sense that the expected discounted dividend pay-out is maximal for the shareholders, is considered. It is assumed that there are two companies involved, whose free surplus processes are are modelled as diffusions. In comparison to the work of Asmussen and Taksar (1997) [3] it is assumed that there are two collaborating companies (or business lines), which are obligated to balance out each others losses to avoid ruin of one company. As a consequence only simultaneous ruin is possible.

In Section 5 the Hamilton-Jacobi-Bellman equation (HJB) for this 2-dimensional model is derived. The corresponding (optimal) partial differential equation allows a unique classical solution, which can be computed explicitly. This solution is the value function of the control problem. Then, an optimal strategy, which has to be executed to achieve this value function, is constructed. It turns out that the optimal strategy has to be a barrier strategy. As the companies are allowed to transfer money freely, the optimal strategy is not unique and there are several possibilities to get maximal pay-out in expectation.

In the last section a simulation study was implemented to compare different alternative strategies with the theoretically optimal strategy. Three models are considered: In the first model both companies are performing well in average. In the second model the second company is performing badly and it is very likely that the other company has to bail out a lot. In the third model a setting is considered, where one company would perform pretty well, but has no initial capital. The other company on the other hand, is performing badly but has initial capital. It is assumed that both agree on a collaboration, but the well-performing company terminated the contract as soon it has enough capital to act alone.

Comparing different barriers showed that the calculated barrier is indeed optimal. The question arises if it is maybe better for the shareholders if the businesses don't work together. In this case the weak one goes ruin while the other one can pay out more dividends instead of rescuing the weak one first. As expected it turns out that it depends on the drift and volatility if collaboration is recommended or if a selfish approach will lead to a possible better result. In the third model is better to start with collaboration and then end the cooperation at a certain (optimal) point.

9 Appendix: Code

```
##########FUNCTIONS##########
barrierstrategy<-function(barrier,start1,start2,mu1,mu2,sigma1,sigma2,beta,N,T){
  theta1<-(-(mu1+mu2)+sqrt((mu1+mu2)^2+2*(sigma1^2+sigma2^2)*beta))/(sigma1^2+sigma2^2)
  theta2<-(+(mu1+mu2)+sqrt((mu1+mu2)^2+2*(sigma1^2+sigma2^2)*beta))/(sigma1^2+sigma2^2)
  m_optimal<-2/(theta1+theta2)*log(theta2/theta1)
  C_optimal<-1/(theta1*exp(theta1*m_optimal)+theta2*exp(-theta2*m_optimal))
  line1<-numeric(N)
  line2<-numeric(N)
  dividend<-numeric(N)
  x1<-start1
  x2<-start2
  BM1<-c(0,rnorm(N-1,mean=mu1*T/N,sd=sigma1*sqrt(T/N))) #mu1*t+sigma*W_t
  BM2<-c(0,rnorm(N-1,mean=mu2*T/N,sd=sigma2*sqrt(T/N)))

  if (barrier!="optimal"){
    m<-barrier
  }
  else m<-m_optimal

  for (i in 1:N) {
    x1<-x1+BM1[i]
    x2<-x2+BM2[i]

    if ((x1+x2)<=0) break

    if (x1<0 & x2>0) {
      x2<-x2+x1
      x1<-0
    }
    if (x2<0 & x1>0) {
      x1<-x2+x1
      x2<-0
    }
    if ((x1+x2>m) && (x2>m)){
      x1<-x1+x2-m
      x2<-m
    }
    if ((x1+x2>m) & (x2 <= m)) {
      dividend[i]<-x1+x2-m
      x1<-m-x2
    }
    line1[i]<-x1
    line2[i]<-x2
    dividend[i]
  }
  return(list("dividend"=cumsum(dividend),"line1"=c(start1,line1),"line2"=c(start2,line2),
              "m_optimal"=m))}
```

```
divmax<-function(sims,barrier,start1,start2,mu1,mu2,sigma1,sigma2,beta,N,T){

  theta1<- (-(mu1+mu2)+sqrt((mu1+mu2)^2+2*(sigma1^2+sigma2^2)*beta))/(sigma1^2+sigma2^2)
  theta2<- (+(mu1+mu2)+sqrt((mu1+mu2)^2+2*(sigma1^2+sigma2^2)*beta))/(sigma1^2+sigma2^2)
  m_optimal<-2/(theta1+theta2)*log(theta2/theta1)
  C_optimal<-1/(theta1*exp(theta1*m_optimal)+theta2*exp(-theta2*m_optimal))
  enddiv<-numeric(sims)
  count<-0
  for (k in 1:sims){
    x1<-start1
    x2<-start2
    dividend<-numeric(N)
    discounted_dividend<-numeric(N)
    BM1<-c(0,rnorm(N-1,mean=mu1*(T/N),sd=sigma1*sqrt(T/N))) #mu1*t+sigma*W_t
    BM2<-c(0,rnorm(N-1,mean=mu2*(T/N),sd=sigma2*sqrt(T/N)))

    if (barrier!="optimal"){
      if (barrier>=0) m<-barrier
      else m<-0
    }
    else if (m_optimal<0){
      m<-0
      m_optimal<-0
    }
    else m<-m_optimal

    for (i in 1:N) {

      x1<-x1+BM1[i]
      x2<-x2+BM2[i]

      if ((x1+x2)<=0) break

      if (x1<0 & x2>0) {
        x2<-x2+x1
        x1<-0
      }

      if (x2<0 & x1>0) {
        x1<-x2+x1
        x2<-0
      }

      if ((x1+x2>m) && (x2>m)){
        x1<-x1+x2-m
        x2<-m
      }


      if ((x1+x2>m) & (x2 <= m)) {
```

```
        dividend[i]<-x1+x2-m
        x1<-m-x2
      }

    }
   discounted_dividend<-exp((-beta)*(T/N)*(0:(N-1)))*dividend
  enddiv[k]<-sum(discounted_dividend)
   }
  if ((start1+start2)<m_optimal) valuefunction<-C_optimal*(exp(theta1*(start1+start2))
                                                        -exp(-theta2*(start1+start2)))
  else valuefunction<-C_optimal*(exp(theta1*(m_optimal))-exp(-theta2*(m_optimal)))
                                                           +start1+start2-m_optimal

  return(c("simulation_value"=mean(enddiv),"optim value"=valuefunction,
           "survival"=sims-count,"m optim"=m_optimal))
}

m_opti_fun<-function(mu1,mu2,sigma1,sigma2,beta){
  theta1<- (-(mu1+mu2)+sqrt((mu1+mu2)^2+2*(sigma1^2+sigma2^2)*beta))/(sigma1^2+sigma2^2)
  theta2<- (+(mu1+mu2)+sqrt((mu1+mu2)^2+2*(sigma1^2+sigma2^2)*beta))/(sigma1^2+sigma2^2)
  m_optimal<-2/(theta1+theta2)*log(theta2/theta1)
  C_optimal<-1/(theta1*exp(theta1*m_optimal)+theta2*exp(-theta2*m_optimal))
  return(m_optimal)
}

m_opti_value<-function(mu1,mu2,sigma1,sigma2,beta,start1,start2){
  theta1<- (-(mu1+mu2)+sqrt((mu1+mu2)^2+2*(sigma1^2+sigma2^2)*beta))/(sigma1^2+sigma2^2)
  theta2<- (+(mu1+mu2)+sqrt((mu1+mu2)^2+2*(sigma1^2+sigma2^2)*beta))/(sigma1^2+sigma2^2)
  m_optimal<-2/(theta1+theta2)*log(theta2/theta1)
  C_optimal<-1/(theta1*exp(theta1*m_optimal)+theta2*exp(-theta2*m_optimal))

  if(start1+start2>m_optimal) valuefunction<-C_optimal*(exp(theta1*(m_optimal))
                                          -exp(-theta2*(m_optimal)))+start1+start2-m_optimal
  else valuefunction<-C_optimal*(exp(theta1*(start1+start2))-exp(-theta2*(start2+start1)))

  return(valuefunction)
}

divmax_switch<-function(sims,grenze,start1,start2,mu1,mu2,sigma1,sigma2,beta,N,T){

  theta1<- (-(mu1+mu2)+sqrt((mu1+mu2)^2+2*(sigma1^2+sigma2^2)*beta))/(sigma1^2+sigma2^2)
  theta2<- (+(mu1+mu2)+sqrt((mu1+mu2)^2+2*(sigma1^2+sigma2^2)*beta))/(sigma1^2+sigma2^2)
  m_optimal<-2/(theta1+theta2)*log(theta2/theta1)
  enddiv<-numeric(sims)

  theta1_1<- (-(mu1)+sqrt((mu1)^2+2*(sigma1^2)*beta))/(sigma1^2)
  theta2_1<- (+(mu1)+sqrt((mu1)^2+2*(sigma1^2)*beta))/(sigma1^2)
  m_optimal_1<-2/(theta1_1+theta2_1)*log(theta2_1/theta1_1)

  theta1_2<- (-(mu2)+sqrt((mu2)^2+2*(sigma2^2)*beta))/(sigma2^2)
```

```
theta2_2<- (+(mu2)+sqrt((mu2)^2+2*(sigma2^2)*beta))/(sigma2^2)
m_optimal_2<-2/(theta1_2+theta2_2)*log(theta2_2/theta1_2)

m1<-m_optimal_1
m2<-m_optimal_2
m<-m_optimal

count<-0
for (k in 1:sims){
  x1<-start1
  x2<-start2
  dividend<-numeric(N)
  line1<-numeric(N)
  line2<-numeric(N)
  discounted_dividend<-numeric(N)
  BM1<-c(0,rnorm(N-1,mean=mu1*(T/N),sd=sigma1*sqrt(T/N))) #mu1*t+sigma*W_t
  BM2<-c(0,rnorm(N-1,mean=mu2*(T/N),sd=sigma2*sqrt(T/N)))
  check1<-0
  check2<-0
  for (i in 1:N) {

    if (check1 != 2) x1<-x1+BM1[i]
    if (check2 != 2) x2<-x2+BM2[i]
    if(x1>=grenze) check1<-1 #keine collaboration

    if ((check1==1) && (x2<0)){
      check2<-2
      x2<-0
    }#no collaboration, Comp2 ruin
    if ((check1==1) && (x1<0)){
      check1<-2
      x1<-0
    }#no collaboration, Comp1 ruin

    if ((check1==2) && (check2==2)) break #both ruin without collaboration

    if (check1==0){ #collaboration
      if ((x1+x2)<=0) break
      if (x1<0 & x2>0) {
        x2<-x2+x1
        x1<-0
      }
      if (x2<0 & x1>0) {
        x1<-x2+x1
        x2<-0
      }
      if ((x1+x2>m) && (x2>m)){
        x1<-x1+x2-m
        x2<-m
      }
```

```
        if ((x1+x2>m) & (x2 <= m)) {
          dividend[i]<-x1+x2-m
          x1<-m-x2
        }
      }

      if(check1==1){
        if ((x1>m1)) {
          dividend[i]<-x1-m1
          x1<-m1
        }
      }
      if((check1 != 0) && (check2==0)){
        if (x2>m2) {
          dividend[i]<-dividend[i]+x2-m2
          x2<-m2
        }
      }
    }
    discounted_dividend<-exp((-beta)*(T/N)*(0:(N-1)))*dividend
    enddiv[k]<-sum(discounted_dividend)
  }
  return(c("simulation_value"=mean(enddiv)))
}

premium<-function(t) return(6*t)

PP_Simulation<-function(lambda,T){
  #simulate homogeneous poisson process
  rpp1<-function(T,lambda) sort(runif(rpois(1,T*lambda), max=T))
  arrivals<-c(0,rpp1(T,lambda))

  #number of claims
  claimnumber<-length(arrivals)

  #size of claims (assume, that claimsize at t=0 is zero)
  claimsize<-c(0,rgamma(claimnumber-1,2,2))

  #cumulated claimes
  cumclaims<-cumsum(claimsize)

  return(list("arrivals"=arrivals,"cumclaims"=cumclaims))
}


##########Figure 1##########
lambda <- 7
T<-1
x<-1.2
```

```
simulation<-PP_Simulation(lambda,T)
arrivals<-simulation$arrivals
cumclaims<-simulation$cumclaims

#wealth before each claim
upper_w<-premium(arrivals)-c(0,cumclaims[-length(cumclaims)])+x
#wealth after each claim
lower_w<-premium(arrivals)-cumclaims+x

#adjust data for plot
M<-rbind(cbind(arrivals,lower_w),cbind(arrivals,upper_w))
C<-M[order(M[,1],-M[,2]), ]
#1. Arrival time are sorted in ascending order
#2. Corresponding wealth is sorted in descending order

plot(rbind(C[1:11,],c(0.56476834,-0.2)),ylim=c(-0.7,2.2),type="l",
     main="Free surplus process of the insurance company",
     xlab="time",ylab="surplus")
abline(c(0,0),c(0,1))

text(0.554,-0.12, labels = expression(paste(tau)),cex=1.5)
text(0.557,-0.33, labels = expression(paste("ruin time ")),cex=0.7,col="darkgrey")
text(0.05075726,0.84, labels = expression(paste(tau[1])),cex=0.9)
text(0.05075726,0.67, labels = expression(paste(1,"st claim")),cex=0.7,col="darkgrey")
text(0.16239057,0.84-0.47, labels = expression(paste(tau[2])),cex=0.9)
text(0.16239057,0.67-0.47, labels = expression(paste(2,"nd claim")),cex=0.7,col="darkgrey")
text(0.20608430,0.84-0.7, labels = expression(paste(tau[3])),cex=0.9)
text(0.52189696,0.84+0.41, labels = expression(paste(tau[4])),cex=0.9)

##########Figure 2##########
PP_Simulation<-function(lambda,T){
  #simulate homogeneous poisson process
  rpp1<-function(T,lambda) sort(runif(rpois(1,T*lambda), max=T))
  arrivals<-c(0,rpp1(T,lambda))

  #number of claims
  claimnumber<-length(arrivals)

  #size of claims (assume, that claimsize at t=0 is zero)
  claimsize<-c(0,rexp(claimnumber-1,1))

  #cumulated claimes
  cumclaims<-cumsum(claimsize)

  return(list("arrivals"=arrivals,"cumclaims"=cumclaims))
}

lambda <- 10000
T<-1
x<-50
```

```
simulation<-PP_Simulation(lambda,T)
arrivals<-simulation$arrivals
cumclaims<-simulation$cumclaims

#wealth before each claim
upper_w<-premium(arrivals)-c(0,cumclaims[-length(cumclaims)])+x
#wealth after each claim
lower_w<-premium(arrivals)-cumclaims+x

#adjust data for plot
M<-rbind(cbind(arrivals,lower_w),cbind(arrivals,upper_w))
C<-M[order(M[,1],-M[,2]), ]
#1. Arrival time are sorted in ascending order
#2. Corresponding wealth is sorted in descending order


par(mfrow=c(2,1))
plot(seq(0,1,by=1/10000),cumsum(c(50,rnorm(10000,mean=-50/10000,sd=sqrt(1/10000)*100))),
     type='l',main="Free surplus process in the Brownian motion model",
     xlab="time",ylab="surplus")
abline(c(0,0),c(0,1))

plot(C,type="l",main="Free surplus process in the Cramer Lundberg model",
     xlab="time",ylab="surplus")
abline(c(0,0),c(0,1))

par(mfrow=c(1,1))

##########Figure 3##########
plot(c(0,4),c(4,0), xlim=c(-1,6), axes=FALSE, pch=19, ylim=c(-1,6),xlab="",ylab="")
axis(1, pos=0,labels=FALSE,at=NULL,tck=0)
axis(2, pos=0,labels=FALSE,tck=0)

segments(0, 4, x1 = 4, y1 = 0)
segments(0, 4, x1 = 6, y1 = 4)

text(c(-0.5,4,-0.35),c(4,-0.5,-0.4), labels = c("(0,m)","(m,0)","(0,0)"),cex=0.8)
text(c(1,4,3),c(1.5,2.5,5.3), labels = c("A","B","C"),col="red",cex=1.5)
text(2.6,2.2, labels = expression(paste(x[1]+x[2]," = ",m)),cex=0.8)
text(2.9,4.3, labels = expression(paste(x[2]," = ",m)),cex=0.8)
text(c(5.8, -0.25),c(-0.3,5.8), labels = c(expression(x[1]),expression(x[2])),cex=0.8)

##########Figure 4##########
plot(c(0,4),c(4,0), xlim=c(-1,6), axes=FALSE, pch=21,bg="grey", cex=0.8,ylim=c(-1,6),xlab=""
     ,ylab="")
axis(1, pos=0,labels=FALSE,at=NULL,tck=0)
axis(2, pos=0,labels=FALSE,tck=0)

segments(0, 4, x1 = 4, y1 = 0)
```

```
segments(0, 4, x1 = 6, y1 = 4)

text(c(-0.5,4,-0.35),c(4,-0.5,-0.4), labels = c("(0,m)","(m,0)","(0,0)"),cex=0.8)
text(c(1,4,3),c(1.5,2.5,5.3), labels = c("A","B","C"),col="black",cex=1)
text(2.6,2.2, labels = expression(paste(x[1]+x[2]," = ",m)),cex=0.8)
text(2.9,4.3, labels = expression(paste(x[2]," = ",m)),cex=0.8)
text(c(5.8, -0.25),c(-0.3,5.8), labels = c(expression(x[1]),expression(x[2])),cex=0.8)

library(shape)
points(1,5,pch=16,cex=0.8)
points(2,4,pch=21,bg="grey",cex=0.8)
Arrows(1,5,2,4,lwd=1,arr.adj = 1,arr.type="triangle",arr.length=0.2,arr.width=0.08)
Arrows(2,4,0,4,lwd=1,arr.adj = 1,arr.type="triangle",arr.length=0.2,arr.width=0.08)
plot_result<-barrierstrategy(4,4,4,0.05,-0.4,0.1,0.1,1.1,1000,20)
lines(plot_result$line1,plot_result$line2)

##########Figure 5##########
plot(c(0,4),c(4,0), xlim=c(-1,6), axes=FALSE, pch=21,bg="grey", cex=0.8,ylim=c(-1,6),xlab=""
     ,ylab="")
axis(1, pos=0,labels=FALSE,at=NULL,tck=0)
axis(2, pos=0,labels=FALSE,tck=0)

segments(0, 4, x1 = 4, y1 = 0)
segments(0, 4, x1 = 6, y1 = 4)

text(c(-0.5,4,-0.35),c(4,-0.5,-0.4), labels = c("(0,m)","(m,0)","(0,0)"),cex=0.8)
text(c(1,4,3),c(1.5,2.5,5.3), labels = c("A","B","C"),col="black",cex=1)
text(2.6,2.2, labels = expression(paste(x[1]+x[2]," = ",m)),cex=0.8)
text(2.9,4.3, labels = expression(paste(x[2]," = ",m)),cex=0.8)
text(c(5.8, -0.25),c(-0.3,5.8), labels = c(expression(x[1]),expression(x[2])),cex=0.8)

library(shape)
points(4,3,pch=16,cex=0.8)
points(1,3,pch=21,bg="grey",cex=0.8)
Arrows(4,3,1,3,lwd=1,arr.adj = 1,arr.type="triangle",arr.length=0.2,arr.width=0.08)
plot_result<-barrierstrategy(4,1,3,-0.1,-0.4,0.1,0.1,1.1,1000,20)
lines(plot_result$line1,plot_result$line2)


##########Figure 6##########
plot(c(0,4),c(4,0), xlim=c(-1,6), axes=FALSE, pch=21,bg="grey", cex=0.8,ylim=c(-1,6),xlab=""
     ,ylab="")
axis(1, pos=0,labels=FALSE,at=NULL,tck=0)
axis(2, pos=0,labels=FALSE,tck=0)

segments(0, 4, x1 = 4, y1 = 0)
segments(0, 4, x1 = 6, y1 = 4)

text(c(-0.5,4,-0.35),c(4,-0.5,-0.4), labels = c("(0,m)","(m,0)","(0,0)"),cex=0.8)
text(c(1,4,3),c(1.5,2.5,5.3), labels = c("A","B","C"),col="black",cex=1)
```

```
text(2.6,2.2, labels = expression(paste(x[1]+x[2]," = ",m)),cex=0.8)
text(2.9,4.3, labels = expression(paste(x[2]," = ",m)),cex=0.8)
text(c(5.8, -0.25),c(-0.3,5.8), labels = c(expression(x[1]),expression(x[2])),cex=0.8)

library(shape)
points(1,2,pch=16,cex=0.8)
plot_result<-barrierstrategy(4,1,2,-0.2,-0.4,0.1,0.1,1.1,1000,20)
lines(plot_result$line1,plot_result$line2)

###########Figure 7##########
y<-seq(-2.89,2.89,by=0.01)
x<-0
epsilon<-1
f<-(x-epsilon<y)*(x+epsilon>y)*(y-x+epsilon)^2/(4*epsilon)+(y-x)*(y>=x+epsilon)
f1<-(y-x+epsilon)/(2*epsilon)*(x-epsilon<y)*(x+epsilon>y)+(y >= x+epsilon)
f2<-1/(2*epsilon)*(x-epsilon<y)*(x+epsilon>y)

plot(y,f,type='l',axes=FALSE,xlab="",ylab="",lwd=2,col="darkred")
axis(1, pos=0,labels=FALSE,at=NULL,tck=0)
axis(1, pos=0,labels=c(expression(x-epsilon),expression(x),expression(x+epsilon)),
     at=c(x-epsilon,x,x+epsilon))
axis(2, pos=0,labels=FALSE,tck=0)
axis(2, pos=0,labels=expression(epsilon),at=epsilon,las=1)
Arrows(0,0,0,3,lwd=1,arr.adj = 1,arr.type="triangle",arr.length=0.2,arr.width=0.08)
Arrows(0,0,3,0,lwd=1,arr.adj = 1,arr.type="triangle",arr.length=0.2,arr.width=0.08)
Arrows(-2.9,0,-3,0,lwd=1,arr.adj = 1,arr.type="triangle",arr.length=0.2,arr.width=0.08)
lines(y,f,lwd=2,col="darkred")

###########Figure 8##########
plot(y,f1,type='l',axes=FALSE,xlab="",ylab="",lwd=2,col="darkred",ylim=c(0,1.5))
axis(1, pos=0,labels=FALSE,at=NULL,tck=0)
axis(1, pos=0,labels=c(expression(x-epsilon),expression(x),expression(x+epsilon)),
     at=c(x-epsilon,x,x+epsilon))
axis(2, pos=0,labels=FALSE,tck=0)
axis(2, pos=0,labels=expression(1),at=epsilon,las=1)
Arrows(0,0,0,1.57,lwd=1,arr.adj = 1,arr.type="triangle",arr.length=0.2,arr.width=0.08)
Arrows(0,0,3,0,lwd=1,arr.adj = 1,arr.type="triangle",arr.length=0.2,arr.width=0.08)
Arrows(-2.9,0,-3,0,lwd=1,arr.adj = 1,arr.type="triangle",arr.length=0.2,arr.width=0.08)
lines(y,f1,lwd=2,col="darkred")

###########Figure 9##########
plot(y,f2,type='l',axes=FALSE,xlab="",ylab="",lwd=2,col="darkred",ylim=c(0,1.5))
lines(c(x-epsilon,x-epsilon),c(0,1/(2*epsilon)),col="white",lwd=4)
lines(c(x+epsilon,x+epsilon),c(0,1/(2*epsilon)),col="white",lwd=4)
axis(1, pos=0,labels=FALSE,at=NULL,tck=0)
axis(1, pos=0,labels=c(expression(x-epsilon),expression(x),expression(x+epsilon)),
     at=c(x-epsilon,x,x+epsilon))
axis(2, pos=0,labels=FALSE,tck=0)
text(-0.16,0.73,expression(frac(1,2*epsilon)),cex=0.7)
Arrows(0,0,0,1.57,lwd=1,arr.adj = 1,arr.type="triangle",arr.length=0.2,arr.width=0.08)
```

```
Arrows(0,0,3,0,lwd=1,arr.adj = 1,arr.type="triangle",arr.length=0.2,arr.width=0.08)
Arrows(-2.9,0,-3,0,lwd=1,arr.adj = 1,arr.type="triangle",arr.length=0.2,arr.width=0.08)
points(c(x-epsilon,x+epsilon),c(0,0),pch=20,col="darkred")
points(c(x-epsilon,x+epsilon),c(1/(2*epsilon),1/(2*epsilon)),pch=21,bg="white",col="darkred")
lines(c(-2.89,x-epsilon),c(0,0),col="darkred",lwd=2)
lines(c(x+epsilon,2.9),c(0,0),col="darkred",lwd=2)

##########Figure 10##########
y<-seq(-0.4,2.5,by=0.01)
x<-1
epsilon<-0.5
f<-(x-epsilon<y)*(x+epsilon>y)*(y-x+epsilon)^2/(4*epsilon)+(y-x)*(y>=x+epsilon)
f1<-(y-x+epsilon)/(2*epsilon)*(x-epsilon<y)*(x+epsilon>y)+(y >= x+epsilon)
f2<-1/(2*epsilon)*(x-epsilon<y)*(x+epsilon>y)
plot(y,f1,type='l',axes=FALSE,xlab="",ylab="",lwd=2,col="darkred",ylim=c(0,1.5),lty=2)
axis(1, pos=0,labels=FALSE,at=NULL,tck=0)
axis(1, pos=0,labels=c(0,expression(x-epsilon),expression(x),expression(x+epsilon)),
     at=c(0,x-epsilon,x,x+epsilon))
axis(2, pos=0,labels=FALSE,tck=0)
axis(2, pos=0,labels=expression(1),at=1.01,las=1)
lines(c(x,2.5),c(1.01,1.01),lwd=2)
points(x,1.01,pch=20,col="black")
text(x+0.1,1.14,expression(1[paste("[x,",infinity,")")]),cex=0.8)
text(x+epsilon+0.2,0.9,expression(f*minute[paste("x",epsilon)]) ,cex=0.8)
Arrows(0,0,0,1.57,lwd=1,arr.adj = 1,arr.type="triangle",arr.length=0.2,arr.width=0.08)
Arrows(0,0,2.5,0,lwd=1,arr.adj = 1,arr.type="triangle",arr.length=0.2,arr.width=0.08)
Arrows(-0.4,0,-0.5,0,lwd=1,arr.adj = 1,arr.type="triangle",arr.length=0.2,arr.width=0.08)
lines(y,f1,type='l',lwd=2,col="darkred",lty=2)

##########Figure 11##########
x<-seq(0,10,by=0.1)
y<-2.9+sin(x)+cos(1.5*x+1)
supi<-NULL
for(i in 1:length(x)){
  supi[i]<-max(-min(y[1:i])+3.5,0)
}
plot(x,y,type='l',ylim=c(1,7.3),xlab="",ylab="",xaxt="n",yaxt="n")
abline(h=3.5)
lines(x,y+supi,lty=5,col="red")

##########Figure 12 und Figure 13##########
mu1<- 0.03
mu2<- 0.02
sigma1<- 0.02
sigma2<- 0.03
start1<- 0.05
start2<- 0.05
beta <- log(1.1)

plot_result<-barrierstrategy("optimal",start1,start2,mu1,mu2,sigma1,sigma2,beta,2000,10)
```

```
m_opt<-plot_result$m_optimal
plot(plot_result$line1,plot_result$line2,type='l',xlim=c(0,m_opt)*1.3,ylim=c(0,m_opt)*1.3,
     xlab="Company 1",ylab="Company 2",main="Controlled surplus processes")
abline(a=m_opt,b=-1,col="red")
abline(h=m_opt,col="red")
points(start1,start2,col="red",pch=20)
plot(c(0,seq(0,10-2*0.005,by=0.005)),plot_result$dividend,type="l",main="Accumulated Dividend",
     col="red",xlab="time (in years)",ylab="dividend")

##########Figure 14##########
T<-1000
N<-100000
sims<-10000

theta1<- (-(mu1+mu2)+sqrt((mu1+mu2)^2+2*(sigma1^2+sigma2^2)*beta))/(sigma1^2+sigma2^2)
theta2<- (+(mu1+mu2)+sqrt((mu1+mu2)^2+2*(sigma1^2+sigma2^2)*beta))/(sigma1^2+sigma2^2)
m_optimal<-2/(theta1+theta2)*log(theta2/theta1)
C_optimal<-1/(theta1*exp(theta1*m_optimal)+theta2*exp(-theta2*m_optimal))

require(foreach)
start_time <- Sys.time()
values<-m_opt+0.01*(-14:25)
n<-length(values)
simvalues<-numeric(n)

require(doSNOW)
cores <- parallel::detectCores()
cl <- makeSOCKcluster(cores)
registerDoSNOW(cl)
pb <- txtProgressBar(min=1, max=n, style=3)
progress <- function(n) setTxtProgressBar(pb, n)
opts <- list(progress=progress)

list<-foreach (a=1:n, .options.snow=opts) %dopar% {
  optimalvalue_est<-divmax(sims,values[a],start1,start2,mu1,mu2,sigma1,sigma2,beta,N,T)
  simvalues[a]<-optimalvalue_est[1]
}

close(pb)
stopCluster(cl)
end_time <- Sys.time()
end_time - start_time

plot(values,as.numeric(list),pch=16,xlab="barrier values",ylab="expected dividend",
     main="Barrier strategy for the 2-dim problem")
optimal_value<-C_optimal*(exp(theta1*(m_optimal))-exp(-theta2*(m_optimal)))+start1+start2
                                          -m_optimal
points(m_optimal,optimal_value,col="red",pch=18)

##########Figure 15##########
```

```
mu_new<-mu1+mu2
sigma_new<-sqrt(sigma1^2+sigma2^2)
start_new<-start1+start2

plot_result<-barrierstrategy("optimal",start_new,0,mu_new,0,sigma_new,0,beta,2000,10)
m_opt<-plot_result$m_optimal
plot_result$dividend[N]
plot(c(0,seq(0,10-2*0.005,by=0.005)),plot_result$dividend,type="l",main="Accumulated Dividend",
     col="red",xlab="time (in years)",ylab="dividend")
plot(c(0,seq(0,10-0.005,by=0.005)),plot_result$line1,type='l',ylim=c(0,start_new),
     xlab="time (in years)",ylab="Company 1",main="Controlled surplus process of Company 1")
abline(h=m_opt,col="red")
points(0,start_new,col="red",pch=20)

###########Figure 16##########
mu1<- 0.03
mu2<- 0.02
sigma1<- 0.02
sigma2<- 0.03
start1<- 0.05
start2<- 0.05
beta <- log(1.1)

sigma1<-seq(0.0001,0.2,by=0.2/100)
zusammen <-numeric(length(sigma1))
einzeln <-numeric(length(sigma1))
for (i in 1:length(sigma1)){
  zusammen[i]<-m_opti_value(mu1,mu2,sigma1[i],sigma2,beta,start1,start2)
  einzeln[i]<-m_opti_value(mu1,0,sqrt(sigma1[i]^2),0,beta,start1,0)
  +m_opti_value(mu2,0,sqrt(sigma2^2),0,beta,start2,0) #einzeln
}
plot(sigma1,y=zusammen,type='l',xlab=expression(sigma[1]),ylab="optimal value function",
     main=expression(paste("Optimal value function for different ",mu[1])))
lines(sigma1,einzeln,col="red")
legend(0.13,0.55,  c("collaboration","no collaboration"),cex = 0.75,lty=c(1,1),
       lwd=c(1,1),col=c("black","red"))

###########Figure 17##########
mu1<- 0.03
mu2<- 0.02
sigma1<- 0.02
sigma2<- 0.03
start1<- 0.05
start2<- 0.05
beta <- log(1.1)

mu1<-seq(0.0001,0.005,by=0.002/100)
zusammen <-numeric(length(mu1))
einzeln <-numeric(length(mu1))
for (i in 1:length(mu1)){
```

```
  zusammen[i]<-m_opti_value(mu1[i],mu2,sigma1,sigma2,beta,start1,start2)
  einzeln[i]<-m_opti_value(mu1[i],0,sqrt(sigma1^2),0,beta,start1,0)
  +m_opti_value(mu2,0,sqrt(sigma2^2),0,beta,start2,0) #einzeln
}
plot(mu1,y=zusammen,type='l',xlab=expression(mu[1]),ylab="optimal value function",
     main=expression(paste("Optimal value function for different ",mu[1])))
lines(mu1,einzeln,col="red")
legend(0,0.253,  c("collaboration","no collaboration"),cex = 0.75,lty=c(1,1),
       lwd=c(1,1),col=c("black","red"))

###########Figure 18-Figure 21##########
#Plots analogously to previous plots with different parameter
mu1<- 0.03
mu2<- -0.01
sigma1<- 0.02
sigma2<- 0.05
start1<- 0.1
start2<- 0.1
beta <- log(1.1)

###########Figure 22##########
beta <- log(1.1)
mu1<- 0.03
mu2<- 0.002
sigma1<- 0.02
sigma2<- 0.03
start1<-0.05
start2<-0.05

start1<-seq(0,0.1,by=0.1/100)
zusammen <-numeric(length(start1))
einzeln <-numeric(length(start1))
for (i in 1:length(start1)){
  zusammen[i]<-m_opti_value(mu1,mu2,sigma1,sigma2,beta,start1[i],start2)
  einzeln[i]<-m_opti_value(mu1,0,sqrt(sigma1^2),0,beta,start1[i],0)
  +m_opti_value(mu2,0,sqrt(sigma2^2),0,beta,start2,0) #einzeln
}
plot(start1,y=zusammen,type='l',xlab=expression(x[1]),ylim=c(0.05,0.45),
     ylab="optimal value function",
     main=expression(paste("Optimal value function for different ",x[1])))
lines(start1,einzeln,col="red")
legend(0.02,0.15,y.intersp=1.7,text.width = strwidth("no collaboration")[1]*1.5,
       c("collaboration","no collaboration"),cex = 0.8,lty=c(1,1),lwd=c(1,1),
       col=c("black","red"))
abline(v=0.012,col="grey")
points(0.012, 0.2873112,pch=20)
text(0.0142,0.1,expression(paste(x[1],"=0.012")), col="darkgrey",cex=0.85)

###########Figure 23##########
mu1<- 0.03
```

```
mu2<- 0.002
sigma1<- 0.02
sigma2<- 0.03
start1<- 0
start2<- 0.05
beta <- log(1.1)

T<-500
N<-100000
sims<-50000

require(foreach)
start_time <- Sys.time()
grenze<-seq(0,0.1,by=0.001)
n<-length(grenze)
simvalues<-numeric(n)

require(doSNOW)
cores <- parallel::detectCores()
cl <- makeSOCKcluster(cores-1)
registerDoSNOW(cl)

pb <- txtProgressBar(min=1, max=n, style=3)
progress <- function(n) setTxtProgressBar(pb, n)
opts <- list(progress=progress)

#running time: 48h (with parallelization on 3 cores)
list<-foreach (a=1:n, .options.snow=opts) %dopar% {
  optimalvalue_est<-divmax_switch(sims,grenze[a],start1,start2,mu1,mu2,sigma1,sigma2,beta,N,T)
  simvalues[a]<-optimalvalue_est[1]
}

close(pb)
stopCluster(cl)
end_time <- Sys.time()
end_time - start_time

plot(grenze,as.numeric(list),pch=16,cex=2.7,xlab="switching levels",ylab="expected dividend",
     main="Switching from Collaboration to No Collaboration")
abline(h=max(as.numeric(list)),col="red")
max(as.numeric(list))
grenze[which(as.numeric(list)==max(as.numeric(list)))]
```

References

[1] H. Albrecher, P. Azcue, and N. Muler, *Optimal dividend strategies for two collaborating insurance companies*, Advances in Applied Probability, 49 (2017), pp. 515–548.

[2] S. Asmussen and H. Albrecher, *Ruin probabilities*, World Scientific Publishing Co Pte Ltd, 2010.

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

[4] P. Azcue and N. Muler, *Stochastic optimization in insurance: a dynamic programming approach*, Springer, 2014.

[5] H. Bühlmann, *Mathematical methods in risk theory*, vol. 172, Springer Science & Business Media, 1970.

[6] K. L. Chung, R. J. Williams, and R. Williams, *Introduction to stochastic integration*, vol. 2, Springer, 1990.

[7] H. Cramér, *On the mathematical theory of risk*, Centraltryckeriet, 1930.

[8] B. De Finetti, *Su un impostazione alternativa della teoria collettiva del rischio*, in Transactions of the XVth international congress of Actuaries, vol. 2, New York, 1957, pp. 433–443.

[9] H. U. Gerber, *Entscheidungskriterien für den zusammengesetzten Poisson-Prozess*, PhD thesis, ETH Zurich, 1969.

[10] H. U. Gerber and E. S. Shiu, *On optimal dividend strategies in the compound poisson model*, North American Actuarial Journal, 10 (2006), pp. 76–93.

[11] J.-W. Gu, M. Steffensen, and H. Zheng, *Optimal dividend strategies of two collaborating businesses in the diffusion approximation model*, Mathematics of Operations Research, 43 (2017), pp. 377–398.

[12] F. Lundberg, *I. Approximerad framställning af sannolikhetsfunktionen. II. Återförsäkring af kollektivrisker.*, Almqvist och Wiksells boktryckeri, 1903.

[13] G. Peskir, *On reflecting brownian motion with drift*, In: Proc. Symp. Stoch. Syst., (Osaka, 2005), ISCIE Kyoto, (2006), pp. 1–5.

[14] J. Sass, *Stochastic control: With applications to financial mathematics*, Appunti delle lezioni, RICAM JKU Linz, 2006, 5 (2006), p. 15.

[15] R. Seydel, *Einführung in die numerische Berechnung von Finanz-Derivaten*, Springer, 2000.

[16] P. Tankov, *Financial modelling with jump processes*, vol. 2, CRC press, 2003.