# Contributions to Optimal Stopping and Long-Term Average Impulse Control

**Dissertation**

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von
Tobias Sohr

Kiel, 2020

| | |
|---|---|
| Erstgutachter: | Prof. Dr. Sören Christensen |
| Zweitgutachter: | Prof. Dr. Paavo Heikki Salminen |

| | |
|---|---|
| Tag der mündlichen Prüfung: | 21.09.2020 |

## Danksagung

Zuallererst gilt mein Dank Prof. Dr. Sören Christensen für seine exzellente Betreuung und dafür immer Zeit, Geduld und hilfreiche Ratschläge, außerdem oft sogar eine Tasse guten Kaffees für mich übrig zu haben, wenn ich mit Fragen oder Ideen zu ihm kam.

Weiter möchte ich all meinen Kolleginnen und Kollegen an den Universitäten Göteborg, Hamburg und Kiel für ihre Hilfe bei all den kleinen Alltagsproblemen, für jede Mege fruchtbare Diskussionen und für eine schöne Zeit an allen drei Universitäten danken.

Ich danke meiner Familie für die Unterstützung und Ermutigung dabei, vor fast zehn Jahren den Weg des Mathematikstudiums zu beginnen und bis zur Promotion weiterzugehen. Ich danke all meinen Freunden, die mich auf diesem Weg begleitet und mich zur richtigen Zeit angespornt und im richtigen Moment manchmal auch abgelenkt haben. Und ich danke Kerstin dafür, mir in den letzten Jahren Rückhalt, Hilfe und so viel mehr gewesen zu sein.

## Zusammenfassung

Gegenstand dieser Arbeit sind undiskontierte optimale Stoppprobleme mit unendlichem Zeithorizont und verallgemeinert-linearen Kosten sowie ergodische Impulskontrollprobleme. In beiden Problemtypen ist das Hauptanliegen dieser Arbeit das Finden (semi-)expliziter Lösungen im Falle, dass der zugrundeliegende Prozess Sprünge aufweist.

Zum Lösen der Stoppprobleme machen wir uns dem Ausgangsproblem innewohnende monotone Probleme zu Nutzen und finden gut handhabbare hinreichende Bedingungen dafür, dass Erstübertrittszeiten Optimierer sind. Darüber hinaus charakterisieren wir die optimale Stoppgrenze im Falle, dass der zugrundeliegende Prozess ein eindimensionaler Markovprozess in stetiger oder diskreter Zeit ist. Während in diskreter Zeit das Konzept der Leiterzeiten angewendet werden kann, um die innewohnenden monotonen Strukturen zu nutzen, entwickeln wir in stetiger Zeit eine Maximumsdarstellung von Integraltyp, um eine vergleichbare Argumentationsweise zu ermöglichen.

Betrachtet werden zudem Impulskontrollprobleme. Die Resultate bezüglich dieser Probleme gliedern sich in zwei Hauptbereiche. Einerseits charakterisieren wir für zugrundegelegte allgemeine eindimensionale Markovprozesse den Wert des Impulskontrollproblems sowie mögliche optimale Strategien durch ein assoziiertes Stoppproblem. Andererseits entwickeln wir eine Schritt für Schritt durchführbare Lösungstechnik für den Fall, dass der zugrundeliegende Prozess ein Lévyprozess ist. Die Nützlichkeit der Technik wird dadurch veranschaulicht, dass wir sie auf mehrere Beispiele anwenden, darunter Fragestellungen aus dem Gebiet der Lagerhaltung und des Ressourcenmanagements. Neben diesen klassischen Anwendungen benutzen wir unsere theoretischen Resultate dazu, den Einfluss des Fixkostentermes auf das Kontrollproblem zu erörtern, untersuchen ein Impulskontrollproblem mit einer Einschränkung bezüglich der Kontrollfrequenz und behandeln Mean-Field-Spiele und -Probleme der Impulskontrolle.

# Abstract

In this thesis we consider undiscounted, infinite time horizon optimal stopping problems with generalized linear costs and long-term average impulse control problems. The main goal is to find (semi-)explicit solutions in case the underlying process contains jumps.

In order to solve the stopping problems, we utilize embedded monotone problems to find sufficient conditions, that are easy to handle, for a threshold time to be optimal. Further, we characterize the threshold for one-dimensional Markov processes in both discrete and continuous time. While in the discrete time case the concept of ladder times can be used to exploit inherent monotone structures, in continuous time we develop an integral type maximum representation to enable a comparable line of argument.

The findings on long-term average impulse control problems are structured in two main areas. First, for a general one-dimensional Markov process we characterize the problem's value and possible optimal strategies by an associated stopping problem. Then, we develop a step-by-step solution technique in case the process is a Lévy process and demonstrate its usefulness by applying it to relevant examples, among others problems from inventory control and optimal harvesting. Apart from these direct applications we use our theoretical findings to investigate the influence of varying fixed costs on the impulse control problem, study a control problem with a restriction to the impulse frequency and treat mean field games and problems of impulse control.

# Contents

# Chapter 1

# Introduction

Two mathematical problems and their inherent connection form the basis of this thesis. The first one is optimal stopping with general linear costs, the second one is long-term average (sometimes also called ergodic) impulse control.
Optimal stopping problems heuristically ask for the best time to perform some kind of an action in a randomly evolving environment. A classical, yet in some generalization still discussed, example is the parking problem. Here, the question is whether one should take the next free parking spot or keep on searching if one wants to reduce the average walking distance after parking. Another prominent example from the field of finance is the question how to find the best time to exercise an American option. Further, common applications are the optimal time to end a clinical study and decide on the approval of a new drug or the question when to stop playing some game in a casino in order to maximize the expected profit (admittedly most of the times to answer the last question with 'as soon as possible' no elaborate mathematical theory is needed). Apart from these externally motivated questions plenty of stopping problems originate inside the world of mathematics. Some examples are explicit stopping games like the yet to be fully solved Chow-Robins-Problem or questions from neighbouring fields like hypotheses testing in sequential analysis.
While these examples already outline the broad range of applicability of optimal stopping in fields like economics, biology, engineering, computer science and finance, they also point to the next obvious question: what if we are faced with the task to perpetually act optimal instead of only having to make one optimal decision?
This is where impulse control theory comes into play. In this field we may choose an infinite sequence of (random) times and may also determine actions at all these chosen times to influence a stochastic process to our advantage. This process can model a stock price or even the evolution of a whole market and our controls reflect the ongoing management of a portfolio. Another prominent example is the management of a natural resource, like a forest or the population of fish in the sea. Here, economical gains have to be balanced against environmental aspects, therefore the modelling requires special care and attention. The so-called 'long-term average' pay-off structure we investigate in this thesis aims for solutions that – in terms of these resource management examples – do not deplete the resources and therefore can be viewed as sustainable solutions.
In the following, we define the two problems, optimal stopping and impulse control, mathematically, briefly sketch some main ideas that greatly influenced the study of these problems and give a short overview of recent results in these fields. Afterwards, we illustrate how this thesis contributes to the two fields of optimal stopping and impulse control as well as some neighbouring fields and give an overview over its contents. We close this chapter with an structural outline of the rest of the thesis.

## 1.1 Optimal Stopping Problems

The origin of stopping problems lie in Bellman's work in the field of dynamic programming (see [Bel57] for an overview) as well as in Wald's and Wolfowitz's work regarding sequential statistics ([WW48], [WW50]), and the work of Snell [Sne51]. Nowadays the problem is usually stated as follows: given some process $(Y_t)_{t\in T}$ adapted to a filtration $\mathcal{F} = (\mathcal{F}_t)_{t\in T}$ where $T \subseteq [0,\infty)$ (most prominent examples are $[0,\infty)$, $\mathbb{N}_0$, $[0,T')$ or $\{0,...,n\}$ for some fixed $T' \in \mathbb{R}$ and $n \in \mathbb{N}$), we want to find a stopping time $\tau^*$ such that

$$\mathbb{E}(Y_{\tau^*}) = \sup_{\tau} \mathbb{E}(Y_\tau),$$

where the supremum is taken over a set of stopping times. Multifarious manifestations of this type of problem with diverse motivations ranging from engineering to economics and clinical studies have been investigated and still are of great interest. These differ fundamentally not only in their origin but also in their mathematical structure. As a result, the developed solution techniques as well as the structure of the stopping problem's value (function) and the optimal stopping time are equally as diverse. For example, it plays a crucial role whether the time set $T$ is discrete or continuous, bounded or unbounded. Additionally, discounting, (in)homogeneity or more general the class of processes $Y$ belongs to bring along both difficulties and possible approaches for the solution. Nevertheless, there are two main approaches to tackle optimal stopping problems that are widely used in various applications. The first is the use of martingale methods. This approach goes along with the game theoretic interpretation of optimal stopping problems. The underlying heuristic can basically be phrased as: 'Continue as long as you don't expect to lose anything on average in the future, stop as soon as you do'. This leads to the characterization of the stopping problem's value over time as the smallest super-martingale dominating $Y$, that is defined by $S_t := \operatorname{esssup}_{\tau\geq t} \mathbb{E}(Y_\tau \mid \mathcal{F}_t)$ for all $t \in T$ and named Snell envelope after Laurie James Snell. While in the easiest case, the case that $T = \{0,...,N\}$ for some $N \in \mathbb{N}$, $S$ can be obtained explicitly via backwards induction, i.e., by $S_N = Y_N$ and $S_n = \max\{Y_n, \mathbb{E}(S_{n+1} \mid \mathcal{F}_n)\}$ for all $n < N$, explicit calculations are usually not possible if the time set is infinite. Nevertheless,

$$\tau^* := \inf\{n \mid S_n = Y_n\}$$

is the almost surely smallest optimal stopping time in all of these cases, provided $\tau^*$ is attained almost surely in finite time. Chapter 1 of [Shi78] provides a comprehensive overview over these super-martingale techniques. To get more explicit results in the more complicated cases for $T$, some more structure in $Y$ is useful, or to state it more precisely: If $Y = \gamma(X)$, where $X$ is a strong Markov process, $\gamma$ some function and $T = \mathbb{N}_0$ or $T = [0,\infty)$ (these problems are usually described as problems with an infinite time horizon), the idea of the Snell envelope leads to a useful spatial characterization of the optimal stopping time that utilizes the so-called pay-off function $\gamma$. Simplified as much as possible, the approach can be paraphrased as: 'If stopping or continuing once was right at some state $x$, due to the homogeneity this action will also be right at any other time at $x$.' This idea leads to a separation of the state space into one stopping region and one continuation region and leads to 'stop as soon as the process hits the stopping region' as an optimal stopping time. To formalize this, we work with the usual associated family of measures $\mathbb{P}_x(\cdot) := \mathbb{P}(\cdot \mid X = x)$. Then, we get a stopping problem for each measure $\mathbb{P}_x$ and therefore may define the value function by

$$V(x) := \sup_{\tau} \mathbb{E}_x \left( \gamma(X_\tau) \right). \tag{1.1}$$

Depending on the area of application often additional time dependence is added, like, for example, a finite end time or exponential discounting or running costs. The additional dependence of the value on the starting point first looks like a major complication of the problem, because instead of just solving one stopping problem, now one has to solve a whole bunch of them. But actually it is a major simplification. So due to the homogeneity under minor technical assumptions the first entry time in the set $S := \{x \mid V(x) = \gamma(x)\}$, called stopping region, is optimal regardless of the starting point. Snell's description of the value as the smallest super-martingale dominating the pay-off process carries over to the Markovian case as follows: the value function is characterized as the smallest super-harmonic majorant of $\gamma$. In the discrete case $V$ additionally satisfies the Wald Bellman equation

$$V(x) = \max\{AV(x), \gamma(x)\}$$

for all $x$, where $A$ is the transition operator of $X$. Under some technical conditions, this property becomes a characterization, meaning each solution of this equation is the value function, see [Shi78, Chapter 2] for proofs and details. A direct continuous time analogue does not exist. In this case, of course under some technical conditions, for example a smooth fit condition on the boundaries of $S$, $V$ satisfies the system of equations

$$\begin{aligned} \mathcal{A}V(x) &= 0 && \forall x \in S^c, \\ V(x) &= \gamma(x) && \forall x \in S\ , \end{aligned}$$

where $\mathcal{A}$ is the characteristic operator of $X$. This is often used as an approach to find the value function: one first solves the system of equations to obtain a candidate, then tries to verify that the obtained solution indeed is the value function. While this approach heavily relies on PDE methods, another idea, which stays in the area of stochastics, is the notion of monotone stopping problems. This idea was developed back in the early days of sequential decision making, see, e.g., [CR61] for one of the first works in this field and [CRS71] for an overview over the field. In the discrete time case, where the definition is most intuitive, a stopping problem with value function (1.1) is called monotone if for all $n \in \mathbb{N}$ holds

$$X_n \geq \mathbb{E}(X_{n+1} \mid \mathcal{F}_n) \Rightarrow X_{n+1} \geq \mathbb{E}(X_{n+2} \mid \mathcal{F}_{n+1}),$$

see [CI19, Section 4]. This definition can not be directly carried over to the continuous time case. Instead one essentially has to model an infinitesimal look into the future with a limit, see [Irl79]. For a monotone problem, it is easily seen that the myopic stopping time

$$\tau' := \inf\{n \in \mathbb{N}_0 \mid X_n > \mathbb{E}(X_{n+1} \mid \mathcal{F}_n)\}$$

is optimal. While not too many stopping problems are monotone ones by default, it has proven itself a fruitful approach to transform given problems to monotone ones or search for embedded monotone ones that help to find the optimizer for the initial, non-monotone problem. More recently, the power of this easy line of argument was rediscovered for more advanced problems, see [Chr17, CI19], or [CS19a] in the context of impulse control problems.
In the case of a discounted pay-off functional the problem is very well investigated. Especially if the underlying process is a one-dimensional diffusion, the problem is very well understood, see [Sal85]. But more recently some advances towards processes with jumps were made: for the case that the underlying process is a Lévy process see, e.g., [Mor02], [NS07], works covering more general classes of Markov processes are, for example, [MS07] and [CST13]. Most of these works utilize in some way the running maximum process $\overline{X} := \left(\sup_{s\leq t} X_s\right)_{t\in T}$ of $X$ evaluated at an independent exponentially distributed time. For undiscounted stopping problems with linear or general running costs, problems that occur, for example, in sequential decision making, most results are also obtained in the diffusion setting, see, e.g., [IP04], [Pau00] or [CI16]. For Lévy processes, [Bei98] is one of the few examples for undiscounted, infinite time problems. Here, the pay-off function is assumed to be bounded from above, concave and unimodal, no running cost term is present and a discretization technique is utilized to find a solution.

## 1.2 Impulse Control Problems

Impulse control problems may either be seen as a problem of repeated stopping or they can be interpreted as a special case or rather a restricted version of continuous control problems.
While we already have introduced the first ones, the latter ones usually start with a stochastic differential equation that contains an additional adapted process as input. This process models the controller's continuously performed action and the goal is to optimize a functional of the SDE's solution over all these controls (see [ØS05, Chapter 3] for a formal definition in a jump diffusion setting). The mathematical origins of these problems lie in dynamic programming. Also, the most common starting point for the search for solutions, the Hamilton Jacobi Bellman equation, roots in the Bellman principle and its non-random counterpart. Probably one of the most commonly known examples even beyond the world of mathematics is the work of Merton regarding portfolio management (see [Mer69] and [Mer75]). A comprehensive overview over this field may be found in [ØS05]. However, continuous control also shows a major drawback.

Usually the optimal strategy requires an infinite number of actions in a finite time interval and is therefore not realizable in practice. Additionally, the continuous activity of the controller would lead to instant bankruptcy if there were fixed costs for each control.
To circumvent this issue, in the seventies a new class of control problems was introduced by Lions and Bensoussan: impulse control problems. These only allow strategies consisting of countably many actions. Therefore, impulse control models are the natural choice when the underlying problems entail some fixed costs for each action or one aims for realizable optimizers. Here, for a given continuous time Markov process $(X_t)_{t\in[0,\infty)}$ adapted to a filtration $(\mathcal{F}_t)_{t\in[0,\infty)}$, the allowed strategies are sequences $(\tau_i,\zeta_i)_{i\in\mathbb{N}}$ of stopping times $\tau_i$ that indicate when the controller acts and $\mathcal{F}_{\tau_i}$-measurable random variables $\zeta_i$ that model, whereto the controller shifts the process. The sequence of stopping times is required to converge to infinity almost surely. Depending on the specific situation, usually there are some restrictions on how the controller may chose the $\zeta_i$. For example, often only downwards (or upwards) shifts of the process are allowed or the values of the $\zeta_i$ may be required to lie in a certain set or can depend on the values $X_{\tau_i,-}$ right before the shift. The goal is to find the value function that is given by

$$v(x) := \sup_{S\in\mathcal{S}_B} J_x(S),$$

where $\mathcal{S}_B$ is a set of admissible strategies. Of course, yet another goal is to find optimal strategies. The functional $J$ usually depends on a pay-off function $\gamma$ that models the controller's gain at each action, a fixed cost term $K \geq 0$ modelling, e.g., fixed transaction costs, and a running cost function $h$ that is used to model ongoing cost occurring regardless of the controller's action. The three most common choices for $J$ are the finite time functional

$$J_x^{fin}(S) = \mathbb{E}_x\left(\sum_{n:\tau_n\leq T'}\left(\gamma\left(X^S_{\tau_n,-}\right)-\gamma(\zeta_n)-K\right)-\int_0^{T'} h\left(X^S_s\right)ds\right)$$

for some end time $T' \in (0,\infty)$, the discounted functional

$$J_x^{disc}(S) = \mathbb{E}_x\left(\sum_{n\in\mathbb{N}} e^{-r\tau_n}\left(\gamma\left(X^S_{\tau_n,-}\right)-\gamma(\zeta_n)-K\right)-\int_0^{\infty} e^{-rs}h\left(X^S_s\right)ds\right),$$

where $r > 0$ models the discount rate, and the long-term average functional

$$J_x^{lta}(S) = \liminf_{T\to\infty}\frac{1}{T}\mathbb{E}_x\left(\sum_{n:\tau_n\leq T}\left(\gamma\left(X^S_{\tau_n,-}\right)-\gamma(\zeta_n)-K\right)-\int_0^{T} h\left(X^S_s\right)ds\right).$$

These impulse control problems were intensively studied over the last decades. The foundation of the theoretical framework was laid by Bensoussan and Lions (see [BL84]) by connecting impulse control problems to quasi variational

inequalities (QVIs). The overwhelmingly large field of applications ranges from economics and finance ([BC19], [Kor99]), control of the exchange rate ([MØ98]), over optimal harvesting ([Alv04]) to inventory control ([HSZ15]). Furthermore, [ØS05] provides a broad range of additional applications. Variants of the discounted pay-off functional $J^{disc}$ are probably used most. Partly, because in finance the discount factor possesses the reasonable interpretation as interest rate. Partly, because the discounting enables the use of resolvents and can be interpreted as exponential killing, two properties that are useful on the methodical side. Still, in various fields of application there are at least reasonable doubts on the justification of a discount factor, amongst them forest management (see [AH20]), the control of exchange rates (see [JZ06]) and inventory control (see [HSZ18], [HSZ17]). Especially a long-term average criterion is of interest when one aims for a sustainable and long-term nature of the problem.
Regarding solution approaches, a traditionally popular way is to use the characterization of the value function by a system of QVIs. This has the disadvantage that the occurring QVIs are often tremendously difficult to solve and therefore are rather used as a verification theorem than as a provider of explicit candidates for value function and optimal strategies. Therefore, especially in the discounted case, progress was recently made to characterize the value function in a more accessible and less technical way, similar to the characterization of the value function of stopping problems. In the case of $J^{disc}$ and $h = 0$ [Chr14] characterized the value function under quite general conditions as the smallest super-harmonic function $v$, such that

$$v(x) \geq Mv(x) := \sup_{a \in B(x)} (v(a) - (\gamma(x) - \gamma(a)))$$

for all $x$, see [Chr14, Proposition 2.3]; $B(x)$ is the set, whereto the process is allowed to be shifted to from $x$. [Chr14, Proposition 3.1], a result that itself relies on [Ega08, Proposition 3.1], states that this characterization of the value function by super-harmonicity and the maximum operator $M$ in the case that $X$ is a diffusion yields a characterization of $v$ by an implicit stopping problem. Namely, $v$ is the smallest non-negative function such that

$$v(x) = \sup_{\tau} \mathbb{E}_x \left( e^{-r\tau} Mv(X_\tau) \right) = \sup_{\tau} \mathbb{E}_x \left( e^{-r\tau} \sup_{a \in B(X_\tau)} \gamma(X_\tau) - \gamma(a) + v(a) \right),$$

where the supremum is taken over all stopping times fulfilling some technical conditions. In [CS17] the characterization by super-harmonicity and the maximum operator is used to obtain (semi-)explicit solutions under the assumption that $\gamma$ is representable by a function $f$ via

$$\gamma(x) = \mathbb{E}_x \left( f(\overline{X}_e) \right)$$

for all $x$, where $\overline{X}$ again is the running supremum of $X$ and $e$ an exponentially distributed random variable independent of $X$. Therein it is shown that, given the function $f$ in the representation above is shaped appropriately, a threshold strategy is optimal, where the threshold $x^*$ is given as the solution to $f(x) = c$ for a certain value $c$. The existence of such a function $f$ under quite general assumptions as well as a way to obtain it in specific cases including Lévy processes is worked out in [CST13]. Apart from these results there are few explicit or semi-explicit results that do not require continuity of the sample path and do not rely on characteristic properties of diffusions. One of the few examples in the branch of inventory control is [Yam17], where the underlying process is assumed to be a spectrally one-sided Lévy process and the pay-off function is assumed to be linear. Herein the theory of scale functions for spectrally one-sided Lévy processes, which was developed in the last two decades (see [Kyp14] for an overview), is utilized to obtain results comparable to those for diffusions. In contrast to these advances in the discounted case, apart from the connection to QVIs ([LP86]) and some results with strong ergodicity assumptions on the underlying process ([Ste86], [PS17]), the vast majority of results for the long-term average functional $J^{lta}$ stays in the case of a one-dimensional diffusion as underlying process, see, e.g., [HSZ17] and [HSZ18] for examples in the field of inventory control or [JZ06] for a model with the control of exchange rates or inflation rates as potential applications. Especially when the goal is a (semi-)explicit characterization of optimal control strategies, continuity of sample paths is needed almost always and often only special types of diffusions are studied.

## 1.3 The Contributions of this Thesis

One of the main results of this thesis is devoted to long-term average control problem for underlying processes with jumps. While in the discounted case the results in [Chr13] and [CS17] for general underlying Markov processes give conditions under that the control problem is one-sided, similar results for long-term average problems are only present for diffusions ([HSZ17], [HSZ18]) or require quite strong ergodicity assumptions and just yield existence results but no explicit solution methods ([Ste86], [PS17]). This thesis contributes to filling that gap: In the setting of one-dimensional Markov processes assuming an integral type representation of the pay-off function in terms of the running maximum under quite general conditions we elaborate, when $(s, S)$ strategies are optimal and characterize the boundaries $s$ and $S$ in terms of the function occurring in the maximum representation. So-called $(s, S)$ strategies for two real numbers $s, S$ are a particularly easy type of control strategies: wait till the process hits or exceeds the boundary $S$ the first time and then shift it back to $s$. In the case of Lévy processes we develop such a representation with help of the ladder height process. This leads to a solution technique for Lévy processes yielding (semi-)explicit characterizations of the boundaries.
A deep theoretical result, as well as a substantial ingredient in our line of proof of the mentioned results, is a characterization of the value of the long-term average impulse control problem by a stopping problem with generalized linear costs that we develop under minimal conditions on the underlying Markov process. The second key topic of this thesis is to show under which condition these type of stopping problems have a threshold time as an optimizer, again under the assumption that an integral type maximum representation exists. This is not only an auxiliary result for the control problems, these undiscounted stopping problems with generalized linear costs are also of interest on their own. So these problems have been investigated for quite a long time, nevertheless general solution techniques are only known for underlying diffusion processes ([IP04, CPT12]) or certain subclasses of problems ([WLK94, Bei98]). Here, we are able to show under what conditions those kind of problems are one sided and have a threshold time as an optimizer, even if the underlying process contains jumps. In case of Lévy processes we even characterize the stopping boundary in terms of the Lévy triple and the ladder height process. The line of argument again strongly relies on the maximum representation. Apart from the continuous time stopping problems we also treat the discrete time equivalent. In the discrete time stopping problem the analogue to the maximum representation can straightforwardly be given by a functional utilizing ladder times. We also investigate in how far the discrete and continuous time problems are connected and how the representation used in the discrete time problem converges to the maximum representation in the continuous time case. To show the applicability of our results, we give a variety of examples, including many long-term average equivalents to discounted one-sided control problems of interest that were collected in [ØS05] as well as applications in forest management and inventory control. Here, the focus lies on giving as explicit results as possible in case the

underlying process is a Lévy process, because there are yet very few of such results when the underlying process contains jumps. Further, we investigate the dependence of both value and optimal strategy on the fixed cost term by use of the maximum representation. Here, the behavior of the control problem's value and the optimal strategies suggest interesting connections to singular control when the fixed costs converge to zero.
Finally, we treat a topic that exceeds the status of a mere application, although we utilize our previous technique for long-term average impulse control problems. We study mean field impulse control games and show that both for a diffusion model and a Lévy driven model there exist mean field equilibria in threshold strategies in the context of resource management under economically reasonable assumptions. Additionally, we solve associated mean field problems that – contrary to the competitive mean field games – model markets where the participants cooperate. Then, we compare the solutions for game and problem.

### 1.3.1 Structure of the thesis

Chapter 2 briefly collects the necessary results from the present literature. Additionally the end of Chapter 2, in Section 2.6, marks the beginning of the authors own contributions. Here, a key ingredient to many of the most important later results, an integral type maximum representation, is developed and its structure is specified in the case of Lévy processes.
Chapter 3 contains the solution of three undiscounted stopping problems: a discrete time problem in Section 3.1, its continuous time analogue in Section 3.2 and in Section 3.4 one, that will serve as an auxiliary tool for the control problems later on. Further, in Section 3.3 the connection between the discrete time problem and the continuous time problem is examined.
Chapter 4 deals with long-term average impulse control problems. Here, the theoretical foundation for the later chapters is laid by developing a super-martingale type verification theorem and characterizing value of the problem and optimal strategies by an associated stopping problem in Section 3.3. This connection is then condensed in the theoretical main theorem of this thesis in Section 4.3. Section 4.4 deals with the most important special case, Lévy processes. Therein the theoretical results are used to develop a step-by-step solution technique for long-term average impulse control problems in case, the underlying process is a Lévy process.
Chapter 5 contains four areas of application. First, in Section 5.1 the solution technique developed in Section 4.4 is applied to questions in the area of inventory control. Then, in Section 5.2, optimal harvesting problems, a special branch of impulse control problems, are studied. Both fields have in common that there are yet very few examples in the literature where (semi-) explicit solutions are obtained for underlying processes with jumps. Section 5.3 shows that the impulse control problem continuously depends on the fixed cost term and further examines the behavior of the control problem's value and the optimal threshold strategies if the fixed costs converge to zero and hereby unveils connections to singular control. The last field of application in Section 5.4 form

impulse control problems with a restriction to the average amount of controls per time unit.
Chapter 6 contains the study of mean field control problems and mean field games with an underlying long-term average impulse control problem.
The last chapter, Chapter 7, briefly discusses possible further questions this thesis has given rise to.

### 1.3.2 Connected scientific articles

Parts of this thesis' work was submitted to journals in form of scientific articles. Some of Chapter 3's results on stopping problems may be found in [CS20]. The results on long-term average impulse control theory from Chapter 4 as well as the tailor-made stopping problem from Section 3.4 and the development of the maximum representation, that can be found in Section 2.6, is presented in [CS19a] in a Lévy process centred fashion. Additionally, the results on mean field theory in Chapter 6 have their origins in the collaborative and yet to be published work [CNS20] of Sören Christensen, Berenice Neumann and myself. Here, the results in Sections 6.3, 6.5, 6.6 and 6.7 were developed in direct cooperation. The rest of Chapter 6 as well as the results in the whole Chapters 3, 4, 5 and the results in Section 2.6 are my own work carried out under the consulting and supporting supervision of Sören Christensen.

# Chapter 2

# Toolbox

This chapter serves as preparation for the following ones. In as much brevity as possible, we collect the necessary results from the present literature in a broad range of topics. Further, we give the needed basic definitions. Since this really is an application centred collection, it is abstained from giving broader context and proofs, for that we rely on a range of well known standard references in probability theory as well as a variety of specialized articles. When the needed results are stated, it is always referred to said sources to provide context and proofs. An exception is Section 2.6. After an overview of the relevant literature regarding maximum representations we develop an explicit way to construct a maximum representation in the case that the underlying process is a Lévy process. This was originally developed in [CS19a] by Christensen and Sohr.

## 2.1 Markov Processes

Throughout the whole thesis we work with a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Markov chains in discrete time or Markov processes in continuous time can be heuristically understood as processes whose progression only depends on the present state, but not on the path in the past. Almost all results in this thesis are formulated for Markov processes. Because this process class is very broad and, therefore, often per se provides not enough structure to admit explicit results, usually additional assumptions are needed. Hereby sometimes we straightforwardly seclude ourself on well known subclasses like Lévy processes or diffusions, sometimes we try to stay as general as possible and only state precisely the needed additional assumptions, depending on whether we aim for as general results as possible or prefer explicity and accessibility.
The intuitive explanation of Markov processes, which is also called memorylessness, can be directly translated to a formal definition in the discrete time case.

**Definition 2.1.1.** *A discrete time process $(X_n)_{n\in\mathbb{N}}$ that is adapted to a filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ is called Markov chain if for all $n \in \mathbb{N}$ we have*

$$\mathbb{P}(X_n \in \cdot \mid \mathcal{F}_{n-1}) = \mathbb{P}(X_n \in \cdot \mid X_{n-1}).$$

To define the continuous time analogous, a bit more preparation is needed. First of all we need the notion of a stopping time.

**Definition 2.1.2.** *Let $T \subseteq [0, \infty)$ be a set and $(\mathcal{F})_{t\in T}$ be a filtration. A random variable $\tau : \Omega \to T$ is called stopping time if for all $t \in T$*

$$\{\tau \leq t\} \in \mathcal{F}_t.$$

Now we have everything at hand to define a continuous time Markov process on the real line; we follow the definition given in [CW05].

**Definition 2.1.3.** *Let $E \subseteq \mathbb{R}$ be an interval that we will call state space, $\mathcal{B}$ its Borel $\sigma$-field and $(\mathcal{F}_t)_{t\geq 0}$ a filtration that fulfils the usual conditions as,* *for example, stated and discussed in [Bic02, Warning 1.3.39], implicating it is a right continuous, complete filtration. Let* $(X_t)_{t\geq 0}$ *be a process adapted to* $(\mathcal{F}_t)_{t\geq 0}$ *with values in* $(E, \mathcal{B})$*. Let* $(\mathbb{P}_x)_{x\in E}$ *be a family of probability measures on* $(\Omega, \mathcal{F})$*. The process* $(X_t)_{t\geq 0}$ *with values in* $E$ *is called Markov process (on the real line) on* $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0}, (\mathbb{P}_x)_{x\in E})$ *if*

1. *for each* $t \geq 0$ *and each* $B \in \mathcal{B}$ *the mapping* $x \mapsto \mathbb{P}_x(X_t \in B)$ *is measurable,*
2. *for each* $\omega \in \Omega$ *and each* $h > 0$ *there is a* $\hat{\omega} \in \Omega$*, such that*
$$X_{t+h}(\omega) = X_t(\hat{\omega})$$
*for all* $t \geq 0$*,*
3. *for each* $x \in E$ *holds* $X_0 = x$ *a.s. under* $\mathbb{P}_x$*,*
4. *for each* $s, t \geq 0$*,* $x \in E$ *and* $B \in \mathcal{B}$ *holds*
$$\mathbb{P}_x(X_{t+s} \in B \mid \mathcal{F}_t) = \mathbb{P}_{X_t}(X_s \in B) \text{ a.s. under } \mathbb{P}_x.$$

*Further, we assume a Markov process to have càdlàg paths. If additionally holds that*
$$\mathbb{P}_x(X_{\tau+s} \in B \mid \mathcal{F}_\tau) = \mathbb{P}_{X_\tau}(X_s \in B) \ \mathbb{P}_x \text{ a.s.}$$
*for each* $s \geq 0$*,* $B \in \mathcal{B}(E)$ *and each stopping time* $\tau$*, then the process* $(X_t)_{t\geq 0}$ *is called a strong Markov process. Without loss of generality, whenever working with a continuous time process, we will assume* $\Omega$ *to be the canonical function-space. Furthermore, we assume the time shift operator that is defined via*
$$\theta_t((\omega_s)_{s\geq 0}) = (\omega_{t+s})_{s\geq 0}$$
*for every* $t \geq 0$*, to be measurable.*

Later on most of the times we will just write 'let $X$ be a strong Markov process' and by doing so implicitly define the associated objects $\Omega$, $\mathcal{F}$, $\mathbb{P}$, $(\mathcal{F}_t)_{t\geq 0}$, $E$, $(\mathbb{P}_x)_{x\in E}$ as in the definition above. Only when we deviate from this exact definition and, for example, specify the state space we will explicitly mention this.
Note that in contrast to this technical way of defining Markov processes, which can be found, e.g., in [CW05], often definitions of Markov processes directly translate the heuristic interpretation that the progression of the process after time $t$ may only depend on the value $X_t$, but not on the 'pre-$t$-past', into a profound mathematical definition that, of course, is equivalent to the one given here (see, e.g., [Kal02, Chapter 8]). We refer to [CW05] or [Kal02] for comprehensive treatises of Markov processes. Here, we will, in as much brevity as possible, define and introduce the later needed tools and objects, always in a manner targeted towards their later purpose. As the name suggests, one of the most important objects in optimal stopping, but also in impulse control theory, are stopping times, hence in addition to the mere definition above, now we establish a nomenclature for some of the most used stopping times.

**Definition 2.1.4.** *For a stochastic process* $(X)_{t\in T}$ *on* $E \subseteq \mathbb{R}$, $T \subseteq [0,\infty)$ *adapted to a filtration* $(\mathcal{F})_{t\in T}$ *set for all* $x \in E$

$$\tau_x := \inf\{t \geq 0 \mid X_t \geq x\},$$

$$\mathring{\tau}_x := \inf\{t \geq 0 \mid X_t > x\},$$

$$\tau_{=x} := \inf\{t \geq 0 \mid X_t = x\}$$

*and for sets* $B \in \mathcal{B}(E)$ *(where* $\mathcal{B}(E)$ *the Borel* $\sigma$*-algebra) define*

$$\tau_B := \inf\{t \geq 0 \mid X_t \in B\}.$$

Note that provided they are finite all of these are stopping times, if $T = \mathbb{N}_0$, see [Kal02, Lemma 7.6, (i)]. In the case $T = [0,\infty)$, if the filtration fulfils the usual conditions and $X$ is a right continuous Markov process, the first three are also stopping times, see [Kal02, Lemma 7.6, (iii)] combined with [Kal02, Lemma 7.3]. Under fairly general conditions on the set $B$ the last one is also a stopping time, see [Kal02, Theorem 7.7]. Later on the stopping times that occur in the optimization problems have to be compatible with pay-off and running cost functions. This leads to the following definition.

**Definition 2.1.5.**
1. *For a Markov process* $X$ *on* $E$*, a measurable function* $\gamma : E \to \mathbb{R}$ *and a continuous function* $h : E \to \mathbb{R}$ *let* $\mathcal{T}(X,\gamma,h)$ *be the set of all stopping times* $\tau$ *such that* $\mathbb{E}_x(\gamma(X_\tau))$ *exists and we have* $\mathbb{E}_x(\tau) < \infty$ *and* $\mathbb{E}_x\left(\int_0^\tau \mid h(X_s) \mid ds\right) < \infty$ *for all* $x \in E$.
2. *For each* $x \in E$ *set*

$$\mathcal{T}_x(X,\gamma,h) := \{\tau \in \mathcal{T}(X,\gamma,h) \mid X_\tau \geq x \text{ a.s. under } \mathbb{P}_x\}.$$

*If it is clear by the context, which functions and which process are meant, we also shortly write* $\mathcal{T}$ *for* $\mathcal{T}(X,\gamma,h)$ *and* $\mathcal{T}_x$ *for* $\mathcal{T}_x(X,\gamma,h)$.

The next important object for our later analysis is the generator of a Markov process that can be interpreted as a derivative of the processes semi-group, which is an essential tool to find martingales associated to Markov processes.

**Definition 2.1.6.** *Let* $X$ *be a Markov process on* $E$ *and for all* $f \in C_0$, $x \in E$ *set*

$$A_X f(x) := \lim_{t\to 0} \frac{\mathbb{E}_x f(X_t) - f(x)}{t}$$

*if this limit exits in* $(C_0, \; ||\cdot||_\infty)$*. The set of all functions* $f$ *for which this limit exist is denoted by* $\mathcal{D}(A_X)$ *and* $A$ *is called the generator of* $X$.

There are many useful properties of the generator, that we discuss in more detail in Section 2.5, when we have introduced Lévy processes and diffusions. Another process that will be of importance later on is the running maximum process of a Markov process.

**Definition 2.1.7.** *Let $X$ be a Markov process. The process $\overline{X}$ defined by*

$$\overline{X}_t := \sup_{s \leq t} X_s$$

*for all $t \in [0, \infty)]$ is called the running maximum process of $X$.*

**Remark 2.1.8.** *Note that $\overline{X}$ is not a Markov process with respect to the filtration $(\mathcal{F}_t)_{t \geq 0}$, but the two dimensional process $(\overline{X}, X)$ is. Whenever the running maximum occurs later on, we tacitly assume $\mathbb{P}_x$ to be $\mathbb{P}_{(x,x)}$, the measure corresponding to the two dimensional Markov process $\left(\overline{X}_t, X_t\right)_{t \geq 0}$ started in $(x, x)$. So we are still able to exploit the Markovian structure.*

## 2.2 Lévy Processes

There are two ways to heuristically describe Lévy processes. Either as stochastic analogues to linear functions or as continuous time extensions of random walks. Both interpretations are directly visible in the following definition.

**Definition 2.2.1.** *A process $(Y_t)_{t \in [0,\infty)}$ on a interval $E \subseteq \mathbb{R}$, that is adapted to a filtration $(\mathcal{F}_t)_{t \in [0,\infty)}$ fulfilling the usual conditions, is called a Lévy process if*

1. *$Y_0 = 0$ a.s.*
2. *for all $s, t \in [0, \infty)$ with $s \leq t$ the random variable $Y_t - Y_s$ is independent of $\mathcal{F}_s$,*
3. *for all $s, t \in [0, \infty)$ with $s \leq t$ the random variable $Y_t - Y_s$ has the same distribution as $Y_{t-s}$,*
4. *$Y$ has càdlàg paths.*

For a more detailed description as well as the proofs to the results we present here see, e.g., [Kyp14], [Sat13], [Kal02] or [App09].

**Remark 2.2.2.** *By $\mathbb{P}_0 := \mathbb{P}$ and $\mathbb{P}_x(Y_t \in \cdot) = \mathbb{P}_0(Y_t + x \in \cdot)$ for all $x \in E$ a Lévy process $Y$ becomes a strong Markov process in the sense of Definition 2.1.3.*

A common tool to analyse Lévy processes is the characteristic exponent that for each Lévy process $Y$ is defined by

$$\psi_Y : \mathbb{R} \to \mathbb{C}; a \mapsto -\ln(\mathbb{E}(e^{aiY_1})).$$

One of the most important results on Lévy processes is the Lévy-Kintchin (often also spelled Khinchine) formula. It connects the characteristic exponent with the Lévy triple that consists of two real numbers and a Lévy measure.

**Proposition 2.2.3.** *For each Lévy process $Y$ there are unique $\mu_Y \in \mathbb{R}$, $\sigma_Y \in [0,\infty)$, and a measure $\Pi_Y$ that is concentrated on $\mathbb{R}\setminus\{0\}$ and fulfils $\int_{\mathbb{R}}(1\wedge x^2)\Pi_Y(dx) < \infty$ such that for all $a \in \mathbb{R}$*

$$\psi_Y(a) = \mu_Y a + \frac{\sigma_Y^2}{2}a^2 + \int \left(1 - e^{iax} + iax\mathbb{1}_{\{|x|<1\}}\right)\Pi_Y(dx). \tag{2.1}$$

*Further, for each such triple $(\mu_Y, \sigma_Y, \Pi_Y)$ there exists a Lévy process $Y$ such that its characteristic exponent is given by* (2.1).

This result can be found in [Sat13, Chapter 1, Theorem 8.1] or in [Kyp14, Theorem 1.3]. The measure $\Pi_Y$ is called the Lévy or jump measure of $Y$. The value $\mu_Y$ is sometimes called drift and $\sigma_Y$ diffusion coefficient, although the nomenclature of these two values is not entirely consistent. The next results connects the so called characteristic triple $(\mu_Y, \sigma_Y, \Pi_Y)$ with the generator of a Lévy process and is proven in [App09, Theorem 3.3.3] and also in a slightly different formulation in [Kal02, Theorem 19.10].

**Proposition 2.2.4.** *Let $Y$ be a Lévy process. Let $g$ be a continuous, infinitely often differentiable function, such that for all natural numbers $n,k \in \mathbb{N}$ we have $\sup_{x\in E} \mid x^k g^{(n)}(x) \mid < \infty$ (These functions are often called rapidly decreasing functions or Schwartz functions). Then $g$ lies in the range of $Y$'s generator $A_Y$ and it holds*

$$\begin{aligned}A_Y g(x) =& \mu_Y \frac{d}{dx}g(x) + \frac{1}{2}\sigma_Y^2 \frac{d^2}{dx^2}g(x)\\ &+ \int (g(x+a) - g(a) + \mathbb{1}_{\{a\le 1\}} a\frac{d}{dx}g(x))\Pi_Y(da). \end{aligned}\tag{2.2}$$

**Definition 2.2.5.** *A special type of Lévy processes are subordinators. Subordinators are defined as Lévy processes with a.s. non-decreasing paths. For a subordinator $S$ instead of the characteristic exponent usually the Laplace exponent*

$$\phi_S(a) := \psi_S(ia)$$

*is used and furthermore, instead of the Lévy triple, a slightly different parametrization is used: set $\tilde{\mu}_S := \mu_S - \int_0^1 t\Pi(dt)$ then*

$$\phi_S(a) = \tilde{\mu}_S a + \int_0^\infty \left(1 - e^{-at}\right)\Pi_S(dt).$$

*We call $\tilde{S}$ a killed subordinator if there is a subordinator $S$ and an Exp($\eta$)-distributed random variable $e_\eta$ for some $\eta \ge 0$ (for $\eta > 0$ the distribution function is $\mathbb{R} \to [0,1]$; $x \mapsto (1-e^{-\eta x})\mathbb{1}_{[0,\infty)}(x)$ and for $\eta = 0$ we use the convention $e_0 = \infty$) such that*

$$\tilde{S}_t := \begin{cases} S_t; & \text{if } t \le e_\eta,\\ \delta\ ; & \text{if } t > e_\eta,\end{cases}$$

*where $\delta$ denotes a cemetery state outside of $E$. With $\eta_{\tilde{S}} := \eta$, $\tilde{\mu}_{\tilde{S}} := \tilde{\mu}_S$ and $\Pi_{\tilde{S}} := \Pi_S$ the Laplace exponent of $\tilde{S}$ then is given by*

$$\phi_{\tilde{S}}(a) = \eta_{\tilde{S}} + \tilde{\mu}_{\tilde{S}} a + \int_0^\infty (1 - e^{-at}) \Pi_{\tilde{S}}(dt).$$

**Remark 2.2.6.** *For a subordinator $S$ the representation of the generator in Proposition 2.2.4 simplifies to*

$$A_S g(x) = \tilde{\mu}_S \frac{d}{dx} g(x) + \int (g(x+a) - g(a)) \Pi_S(da)$$

*for all $x \in E$ and $g$ as in Proposition 2.2.4.*

Especially for subordinators (but sometimes also for general Lévy processes) often $\tilde{\mu}_S$ is called drift instead of $\mu_S$.

**Definition 2.2.7.** *For each killed subordinator $S$ we define its potential measure $U_S$ by*

$$U_S(dx) := \mathbb{E}\left( \int_0^\infty \mathbb{1}_{\{S_t \in dx\}} dt \right).$$

**Lemma 2.2.8** (Wald's equation, continuous version)**.** *Let $Y$ be a Lévy process such that $\mathbb{E}(Y_1)$ exists and $0 < \mathbb{E}(Y_1) \leq \infty$. Let $\tau$ be a stopping time. Then,*

$$\mathbb{E}(Y_\tau) = \mathbb{E}(Y_1) \mathbb{E}(\tau).$$

*Proof.* This result supposedly goes back to Doob in 1957 and an even more general version can be found in [Hal70, Corollary 1]. □

**Lemma 2.2.9.** *Let $Y$ be a Lévy process such that $\mathbb{E}(Y_1)$ exists and $0 < \mathbb{E}(Y_1) < \infty$. Then for all $a \geq 0$ holds $E(\tau_a) < \infty$ and $\mathbb{E}(Y_{\tau_a}) < \infty$.*

*Proof.* The first part of the claim is a direct consequence from the analogous result for random walks that is proven in [Gut74, Theorem 2.1]. The second part then follows with Lemma 2.2.8. □

**Lemma 2.2.10.** *Let $Y$ be a Lévy process. If $Y$ either is not a compound Poisson process or if $Y$ is a compound Poisson process whose Lévy measure has no atoms, then for all $x \in E$ holds $\mathring{\tau}_x = \tau_x$ a.s. under all $\mathbb{P}_y$, $y \in E \setminus \{x\}$.*

*Proof.* This is proven in [PR69, Lemma 2] in case that $Y$ is not a compound Poisson process. The case that $Y$ is a compound Poisson process whose Lévy measure has no atoms follows with elementary arguments. □

**Lemma 2.2.11.** *Let $Y$ be a Lévy process with $0 < \mathbb{E}(Y_1) < \infty$. Assume $Y$ is not a compound Poisson process or a compound Poisson process whose jump measure has no atoms. Define the mapping $\xi$ by*

$$\xi(x, y) := \mathbb{E}_x(\tau_y)$$

*for all $x, y \in E$ with $x < y$. Then, $\xi$ is a continuous real valued mapping that is non-decreasing in the second and non-increasing in the first argument.*

*Proof.* Lemma 2.2.9 yields that for all $x, y \in E$ with $x \leq y$ holds $\mathbb{E}_x(Y_{\tau_y}) < \infty$. Further, Lemma 2.2.10 implies that for all $y \in E$ holds $\lim_{a \nearrow y} \tau_a = \tau_y$ a.s. under all $\mathbb{P}_z$, $z \in E \setminus \{y\}$. With dominated convergence we hence get continuity of $\xi$ in the second argument. Further, for all $x, y \in E$ with $x < y$ we can use the homogeneity of $Y$ to write $\mathbb{E}_x(\tau_y) = \mathbb{E}(\tau_{y-x})$, which yields the claim. □

**Lemma 2.2.12.** *Let $Y$ be a Lévy process. Let $\psi : E \to \mathbb{R}$ be a continuous non-decreasing mapping, let $y_0 \in E$ and assume that for all $y \in E$ holds $\mathbb{E}_{y_0}(\psi(Y_{\tau_y})) < \infty$. If $Y$ either is not a compound Poisson process or if $Y$ is a compound Poisson process whose Lévy measure has no atoms, then the mapping*

$$\Psi : E \cap (y_0, \infty) \to \mathbb{R};\ y \mapsto \mathbb{E}_{y_0}\left(\psi(Y_{\tau_y})\right)$$

*is continuous.*

*Proof.* Let $y \in E \cap (y_0, \infty)$. Let $\delta > 0$. Then, we have

$$\begin{aligned}\mathbb{E}_{y_0}\left(\psi(Y_{\tau_{y+\delta}})\right) \leq &\mathbb{E}_{y_0}\left(\psi(Y_{\tau_y})\right) + \mathbb{P}_{y_0}(Y_{\tau_y} < y+\delta) \sup_{y \leq z \leq y+\delta} \mathbb{E}_z(\psi(Y_{\tau_{y+\delta}})) \\ \leq &\mathbb{E}_{y_0}\left(\psi(Y_{\tau_y})\right) + \mathbb{P}_{y_0}(X_{\tau_y} < y+\delta)\mathbb{E}_{y+\delta}(\psi(Y_{\tau_{y+2\delta+1}})).\end{aligned}$$

Now when $\delta \to 0$ we have due to the assumptions on $Y$ that

$$\mathbb{P}_{y_0}(Y_{\tau_y} < y+\delta) \to 0,$$

hence $\Psi$ is continuous. □

**Lemma 2.2.13.** *Let $f : E \to \mathbb{R}$ be a continuous function, $\overline{x} \in E$ a root of $f$ and let $Y$ be a subordinator that is either no compound Poisson process, or a compound Poisson process whose Lévy measure has no atoms. Further, assume that $\mathbb{E}(Y_1) < \infty$. Then:*

1. *The function*

$$\Xi : E \to \mathbb{R}; x \mapsto \mathbb{E}_x\left(\int_0^{\tau_{\overline{x}}} f(Y_s)ds\right)$$

   *is continuous,*

2. *The function*

$$\tilde{\Xi} : \{(a,b) \in E \times E \mid a < b\} \to \mathbb{R}; (x,y) \mapsto \mathbb{E}_x\left(\int_0^{\tau_y} f(Y_s)ds\right)$$

   *is continuous.*

*Proof.* If $\overline{x}$ is the left boundary of $E$ nothing has to be shown. Hence, assume there is a $z \in E$ with $z < \overline{x}$. Continuity in all points $y \in E$ with $y < \overline{x}$ follows from approximation with simple functions and Lemma 2.2.11. To see that $\Xi$ is continuous on the whole set $E$, note that $\Xi(y) = 0$ for all $y \in E$ with $y \geq \overline{x}$ hence it remains to be shown that $\lim\limits_{y \nearrow \overline{x}} \Xi(y) = 0$. Let $\epsilon > 0$. Then, for all $y \in E$ with $\overline{x} - y < \overline{x} - z$ such that $\max_{x \in [y, \overline{x}]} |f(x)| < \mathbb{E}_z(\tau_{\overline{x}})\epsilon$ we have

$$\Xi(y) < \epsilon$$

and since $f$ is continuous and $\overline{x}$ a root of $f$, the set of such values $y$ is a non-empty interval. Hence, 1. holds. 2. then follows with analogous arguments by the homogeneity of Lévy processes, since we restricted the domain of $\tilde{\Xi}$ to elements $a, b \in E$ with $a < b$. □

### 2.2.1 Ladder height process

For the whole subsection let $X$ be a Lévy process. Following [Kyp14, Definition 6.1] we define a local time at the maximum of $X$ as a continuous, non-decreasing, adapted process $(L)_{t \in [0,\infty)}$ on $[0, \infty)$ with the following properties:

1. The support of $dL$ is $\overline{\{t \in [0, \infty) \mid \overline{X}_t = X_t\}}$.
2. For each stopping time $\tau$ with the property that a.s. $X_\tau = \overline{X}_\tau$ the process
$$(L_{\tau+t} - L_\tau)_{t \in [0,\infty)}$$
is independent of $\mathcal{F}_\tau$ and is distributed as $(L_t)_{t \in [0,\infty)}$ under $\mathbb{P}$.

Such a local time exists for a wide class of Lévy processes. Nevertheless, if (and only if) 0 is not regular for $[0, \infty)$ (this means that $\mathring{\tau}_0 \neq 0$ a.s. under $\mathbb{P}_0$) such a continuous local time fails to exist. Then, it is possible to construct a right continuous alternative we will tacitly work with instead, see [Kyp14, Theorem 6.6], the references thereafter for the proof of existence and [Kyp14, Section 6.1] for more details and explicit constructions of local times in several cases. For a local time $L$ we set $L_\infty := \lim\limits_{t \to \infty} L_t$ and define the inverse local time process $L^{-1}$ by

$$L_t^{-1} := \begin{cases} \inf\{s > 0 \mid L_s > t\}; & \text{if } t < L_\infty \\ \infty; & \text{else} \end{cases} \qquad \forall t \in [0, \infty).$$

It can be seen in the definition that a local time can only be unique up to a multiplicative factor which we chose conveniently for our purposes in the next definition.

**Definition 2.2.14.** *Let $L$ be a local time at the maximum and $H$ defined by*

$$H_t := X_{L_t^{-1}}$$

*for all $t \geq 0$. The process $H$ is called the ladder height process of $X$. $\left(H, L^{-1}\right)$ is a two-dimensional Lévy process, even a bivariate subordinator. Since $L$ is only defined up to a multiplicative constant, w.l.o.g. whenever we have $\mathbb{E}(\tau_x) < \infty$ for all $x \in E$ we choose $L$ such that $\mathbb{E}\left(L_1^{-1}\right) = 1$ and, hence, $\mathbb{E}\left(L_{\tau_x}^{-1}\right) = \mathbb{E}\left(\tau_x\right)$ for all $x \in \mathbb{R}$. Further, we set*

$$\hat{\tau}_x := L_{\tau_x}^{-1} = \inf\{t \geq 0 \mid H_t \geq x\}$$

*for all $x \in \mathbb{R}$.*
*In the same way we define the descending ladder height process $H^{\downarrow}$ as the ladder height process of $-X$.*

**Remark 2.2.15.** *With Wald's equation (see Lemma 2.2.8) we have*

$$\mathbb{E}\left(L_{\tau_x}^{-1}\right) = \mathbb{E}\left(L_1^{-1}\right)^{-1} \mathbb{E}\left(\tau_x\right).$$

*This explains that in the definition above $\mathbb{E}\left(L_1^{-1}\right) = 1$ implies*

$$\mathbb{E}\left(L_{\tau_x}^{-1}\right) = \mathbb{E}\left(\tau_x\right)$$

*for all $x \in \mathbb{R}$.*

### 2.2.2 Special subordinators

The following definitions and results can be found in [SSV12], also [Kyp14, Section 5.6] provides an overview over Bernstein functions that is rather Lévy process centred.

**Definition 2.2.16.** *Let $S$ be a killed subordinator. Then, $S$ is called a special subordinator if $\phi_S$ is a special Bernstein function, i.e., $\tilde{\phi} := \frac{\mathrm{id}}{\phi}$ is also the Laplace exponent of a subordinator.*

**Lemma 2.2.17.** *Let $S$ be a subordinator with potential measure $U_S$. Then, $S$ is special if and only if $U_S \mid_{(0,\infty)}$ has a non-increasing density $u$ with*

$$\int_0^1 u\left(t\right) dt < \infty.$$

*Proof.* This is [Kyp14, Theorem 5.19]. □

**Remark 2.2.18.** *Many common examples of subordinators are special, including:*

1. *All stable subordinators.*
2. *Each subordinator whose jump measure has a log convex density.*
3. *Each subordinator $S$ whose jump measure has a completely monotone density.*

4. *Each subordinator whose Lévy measure* $\Pi_S$ *has the property that* $\mathbb{R} \to \mathbb{R};\ x \mapsto \log \Pi_S(x, \infty)$ *is a convex function.*

**Remark 2.2.19.** *Since later on the most favourable case for us will be that the ladder height process* $H$ *is a special subordinator, the question arises how one can make sure that* $H$ *falls in the class of special subordinators by looking at characteristics of* $X$*. [Kyp14, Theorem 7.8] yields that for each* $y > 0$

$$\Pi_H(y, \infty) = \int_{[0,\infty)} \Pi_X(z + y, \infty) U_{H^\downarrow}(dz)$$

*where* $\Pi_X$ *is the Lévy measure of* $X$*,* $\Pi_H$ *the one of* $H$ *and* $U_{H^\downarrow}(dz)$ *is the potential measure of the descending ladder height process.*
*Now this formula may help to verify one of the necessary conditions for* $H$ *to be a special subordinator stated in Remark 2.2.18 by using our knowledge of* $\Pi_X$*. Especially the condition 2.2.18, 3 turns out to be a handy one since, if* $\Pi_X$ *has a completely monotone density, so has* $\Pi_H$*. And the former applies to many Lévy processes of interest, like, for example, gamma processes, inverse Gaussian processes or Lévy processes with phase type jumps.*

## 2.3 Diffusions

When speaking of diffusions, usually a strong Markov process with continuous paths and some additional regularity properties is meant.

**Definition 2.3.1.** *We call a time homogeneous strong Markov process* $X$ *that takes values in the interval* $E \subseteq \mathbb{R}$ *a (linear one-dimensional) diffusion if it has a.s. continuous sample paths. If for all* $x, y \in E$ *with* $x \in int(E)$ *holds* $\mathbb{P}_x(\tau_{=y} < \infty) > 0$*,* $X$ *is called regular.*

A special subclass of diffusions, that is often used to model phenomena, e.g., in finance, economics or biology by its infinitesimal behaviour, is the class of Itô diffusions.

**Definition 2.3.2.** *Let* $E \subseteq \mathbb{R}$ *be an interval, let* $W$ *be a standard Brownian motion on* $\mathbb{R}$*,* $\sigma : E \to [0, \infty)$ *and* $\mu_X : E \to \mathbb{R}$ *continuous functions and* $x \in E$*. Call a process* $(X_t)_{t\geq 0}$ *a (linear time homogeneous) Itô-diffusion with drift* $\mu_X$ *and diffusion coefficient* $\sigma_X$ *if it is a unique strong solution to the stochastic differential equation* $X_0 = x$*,*

$$dX_t = \mu_X(X_t)dt + \sigma_X(X_t)dW_t.$$

It is well established that if $\mu$ and $\sigma$ are sufficiently smooth (e.g. Lipschitz continuity is a sufficient condition), this SDE has a unique solution for each given starting point, that, moreover, indeed is a diffusion in sense of Definition 2.3.1. see, e.g., [Øks03, Theorem 5.4.1 & Theorem 7.2.4.].

**Lemma 2.3.3.** *Let $X$ be an Itô diffusion as in Definition 2.3.2. Then, its generator $A_X$ is given by*

$$A_X f(x) = \frac{1}{2}\sigma_X^2(x) f''(x) + \mu_X(x) f'(x) \tag{2.3}$$

*for all functions $f$ that lie in the range of $A_X$. Further, the range of $A_X$ contains all infinitely often differentiable functions $f$ with compact support.*

This result may, for example, be found in [Kal02, Theorem 19.24]. Note that we do not allow killing of the diffusion, otherwise an additional term would occur in the generator.

### 2.3.1 Speed measure and scale function

Now let $X$ be a diffusion as in Definition 2.3.1, furthermore assume, that $X$ is regular. There are three basic characteristics of a diffusion: speed measure, scale function and killing measure. Since we will not study killed diffusions explicitly later on, we will only define the first two here.

**Definition 2.3.4.** *A strictly increasing function $S : E \to \mathbb{R}$ is called scale function of $X$ if for each $a, b \in E$ with $a < b$*

$$\mathbb{P}_x(\tau_{=b} < \tau_{=a}) = \frac{S(x) - S(a)}{S(b) - S(a)} \quad \forall x \in [a, b].$$

Each regular linear diffusion possesses a scale function and a scale function is unique up to affine linear transformations, for further details see [Kal02, Theorem 23.7] or [BS15, Section II.4.].

**Definition 2.3.5.** *The speed measure $M$ of $X$ is defined as the unique measure $M$ on the interior of $E$ such that for all $a, b \in E$ with $a < b$ and all $x \in (a, b)$ holds*

$$\mathbb{E}_x\left(\tau_{=a} \wedge \tau_{=b}\right) = \int_a^b G_{a,b}(x, y) M(dy), \tag{2.4}$$

*where for all $x, y \in [a, b]$*

$$G_{a,b}(x, y) := \begin{cases} \frac{(S(x)-S(a))(S(b)-S(y))}{S(b)-S(a)}, & x \le y \\ \frac{(S(y)-S(a))(S(b)-S(x))}{S(b)-S(a)}, & y \le x. \end{cases}$$

That such a measure indeed exists and is unique can be found in [RY05, page 304]. A definition also can be found in [BS15, Section II.4]. Immediately from the definition we may deduce a first result we need in the following chapters.

**Remark 2.3.6.** *A function that later on will be of needed in Chapter 6 on several occasions is*

$$\xi : \{(x, y) \in E \times E \mid x < y\} \to (0, \infty]; \ (x, y) \mapsto \mathbb{E}_x(\tau_{=y}).$$

*Provided for $x, y \in E$ with $x < y$ the function $\xi$ is finite on a neighbourhood of $(x, y)$, the formula (2.4) together with dominated convergence yields that $\xi$ is continuous at $(x, y)$. Similar arguments also may be used to find sufficient conditions to ensure that for a continuous function $f : E \to \mathbb{R}$ the function*

$$\Xi : E \times E \to \mathbb{R};\ (x, y) \mapsto \mathbb{E}_x\left(\int_0^{\tau_y} f(X_s)ds\right)$$

*is continuous, for the details we refer to [KT81, Section 15.3].*

While the scale function can be interpreted as an indicator for the average drift of the diffusion, the speed measure can be understood as a change of measure transforming a diffusion that is natural in scale (meaning the identity is a scale function) to a time changed Brownian motion, see [Kal02, Theorem 23.9] for more details. Therefore, it is not surprising that for Itô diffusions scale function and speed measure are connected with $\mu$ and $\sigma$. In order to state this connection, assume, $X$ is an Itô-diffusion as defined in Definition 2.3.2 with values $\mu$ and $\sigma$ defined as in Definition 2.3.2, additionally assume $\sigma > 0$, that $S$ is continuous with smooth derivative $s := S'$ and $M$ is absolute continuous with respect to the Lebesgue measure with a smooth Lebesgue density $m := \frac{dM}{d\lambda}$. Then, (see [BS15, II.9.] and also [KT81]) speed measure and scale function are connected to drift $\mu$ and diffusion coefficient $\sigma$ via

$$m(x) = \frac{2}{\sigma^2(x)} e^{B(x)}$$

and

$$s(x) = e^{-B(x)}$$

for all $x \in E$, where $B(x) := \int^x \frac{2\mu(y)}{\sigma^2(y)} dy$ for all $x \in E$.

### 2.3.2 Recurrence and stable distribution

**Definition 2.3.7.** *A diffusion is called recurrent if for all $x, y \in E$ we have $\mathbb{P}_x(\tau_{=y} < \infty) = 1$ and transient otherwise. A recurrent diffusion is called positively recurrent if for all $x, y \in E$ we have $\mathbb{E}_x(\tau_{=y}) < \infty$ and null recurrent otherwise.*

Recurrence is connected to the speed measure in the following way, see, e.g., [BS15, II.12] or [KT81, p.234].

**Lemma 2.3.8.** *$X$ is positively recurrent if and only if $M(E) < \infty$.*

If $X$ is positively recurrent, the speed measure $M$ and, hence, also the probability measure $\tilde{M} := \frac{M}{M(E)}$ is an invariant measure in the sense that for all $A \in \mathcal{B}(E)$, $t \geq 0$ holds

$$X_0 \sim \tilde{M} \Rightarrow X_t \sim \tilde{M}.$$

Also general theory about positively recurrent Markov processes yields that $\tilde{M}$ is a limiting distribution in the sense that

$$\lim_{t\to\infty}\frac{f(X_t)}{t}=\int f d\tilde{M}$$

a.s. and in mean, see [Kal02, Theorem 23.15].

## 2.4 Renewal Theory

The long-term average criterion in the formulation of the control problem, at least notationally, bears a resemblance to some expressions in the topic of renewal theory. And indeed, later on at least for some impulse control strategies that are in a sense stationary, renewal theory will prove itself a versatile tool to compute values of the control problem. Hence, here we will collect the needed tools from that topic. For a detailed treatise of renewal theory and the proofs of the lemmas originating in that field stated below, we refer to [Asm03] and [GS01], also [Als91] is a reference in German.

**Lemma 2.4.1** (Wald's equation, discrete version)**.** *Let $Y_0 := 0$ and $(Y_i)_{i\in\mathbb{N}}$ be a sequence of i.i.d. random variables. Let $(\mathcal{A}_n)_{n\in\mathbb{N}_0}$ be a filtration, such that $(Y_i)_{\in\mathbb{N}_0}$ becomes an adapted process and let $\tau$ be a stopping time with $\mathbb{E}(\tau)<\infty$. Then we have*

$$\mathbb{E}\left(\sum_{i=0}^{\tau}Y_i\right)=E(Y_1)E(\tau)$$

*(with the convention $\pm\infty * 0 = 0$).*

*Proof.* This is [Als91, Satz 1.4.7]. □

**Lemma 2.4.2** (Renewal Reward Theorem)**.** *Assume $(Z_i,R_i)_{i\in\mathbb{N}}$ is a sequence of i.i.d. random variables, with $Z_i>0$ a.s. for all $i\in\mathbb{N}$. Set $T_n:=\sum_{i\leq n}Z_i$ and $N(t):=\sup\{n\in\mathbb{N}\mid T_n\leq t\}$. Assume $\mathbb{E}(Z_1)<\infty$ and $\mathbb{E}(\mid R_1\mid)<\infty$. Then we have*

$$\frac{\sum_{i=1}^{N(t)}R_i}{t}\overset{a.s.}{\to}\frac{\mathbb{E}(R_1)}{\mathbb{E}(Z_1)}$$

*and*

$$\frac{\mathbb{E}\left(\sum_{i=1}^{N(t)}R_i\right)}{t}\to\frac{\mathbb{E}(R_1)}{\mathbb{E}(Z_1)}.$$

*Proof.* This is [GS01, Section 10.5, Theorem 1]. □

These renewal processes are not only a powerful tool on their own, they also help to provide versatile results on regenerative processes that lead to helpful existence and representation results regarding stable or limiting distributions.

**Definition 2.4.3.** *Following [Asm03, p. 169] we call a process* $(Z_t)_{t\in[0,\infty)}$ *on* $\mathbb{R}$ *regenerative if is has right continuous paths, there is a (possibly delayed) renewal process* $R$ *given by* $R_n = \sum_{i=0}^n Y_i$ *for all* $n \in \mathbb{N}$*, such that for all* $n \in \mathbb{N}$ *the pair of random variables* $\left(\left(((R_{n+i})_{i\in\mathbb{N}}), (Z_{R_n+t})_{t\in[0,\infty)}\right), (R_0, ..., R_n)\right)$ *are independent.* $R$ *is called embedded renewal process of* $Z$*.*

**Lemma 2.4.4.** *Assume* $Z$ *is a regenerative process and the process* $(R_n)_{n\in\mathbb{N}} = (\sum_{i=0}^n Y_i)_{n\in\mathbb{N}}$ *is an embedded renewal process as defined in Definition 2.4.3. Assume the* $Y_i$*,* $i \in \mathbb{N}$*, have finite mean and non-lattice distribution. Then, the limiting distribution* $\nu$ *of* $Z$ *exists and is given by*

$$\int f d\nu = \frac{1}{\mathbb{E}(Y_1)} \mathbb{E}\left(\int_0^{Y_1} f(Z_s) ds\right).$$

*Proof.* This is proven in [Asm03, Chapter VI, Theorem 1.2.]. □

## 2.5 An Add-On Concerning the Generator

After having introduced the most important process classes for our purposes, we will come back to the generator of a Markov process. Here, we will discuss some of its useful properties and illustrate how to work around some difficulties one has to face when working with the generator. One of the most useful properties for our purposes is that the generator provides a way to construct martingales.

**Proposition 2.5.1.** *Let* $X$ *be a Markov process and* $g \in \mathcal{D}(A)$*. Then the process given by*

$$g(X_t) - g(X_0) - \int_0^t A_X g(X_s) ds$$

*for all* $t \in [0, \infty)$ *is a martingale.*

*Proof.* This in [RY05, Chapter VII, Proposition (1.6)]. □

As a consequence by applying the optional sampling theorem we get

**Lemma 2.5.2.** *[Dynkin's formula] Let* $X$ *be a Markov process,* $x \in E$*,* $\tau$ *a stopping time with* $\mathbb{E}_x(\tau) < \infty$ *and* $g \in \mathcal{D}(A)$*. Then*

$$\mathbb{E}_x\left(g(X_\tau)\right) = g(x) + \mathbb{E}_x\left(\int_0^\tau A_X g(X_s) ds\right).$$

Now as beautiful and helpful these two results are, there is an obstacle when it comes to applying them: for a general Markov process $X$ it is difficult to find the range of its generator $\mathcal{D}(A_X) \subseteq C_0$. And even if this is possible, being restricted to $C_0$, the class of continuous functions vanishing in infinity, is not what one hopes for when one wants to construct martingales, e.g., out of common payoff functions. For Feller processes, a process class including both diffusions and Lévy processes, the range of the generator lies dense in $C_0$ and the characteristic operator is an extension of the generator.

**Definition 2.5.3.** *For all continuous functions $f$ and all $x \in E$ we set, if well defined,*

$$\mathcal{A}f(x) := \lim_{\epsilon \to 0} \frac{\mathbb{E}_x f(X_{\tau_{B_x(\epsilon)^c}}) - f(x)}{\mathbb{E}_x(\tau_{B_x(\epsilon)^c})}$$

*where $B_x(\epsilon) := \{y \in E \mid |y - x| < \epsilon\}$ and $\tau_{B_x(\epsilon)^c}$ is the first entry time in the set $B_x(\epsilon)^c$ as defined in Definition 2.1.4. The operator $\mathcal{A}$ is called characteristic operator of $X$.*

Nevertheless, since we basically only need the properties of Proposition 2.5.1 and Lemma 2.5.2 to hold in the case that the stopping time is of the form $\tau_x$ we will use these properties to create our own definition of an extended generator, following the approach in [RY05, Chapter VII]. To that end, first note that the generator in a sense is the maximal operator with the property in Proposition 2.5.1.

**Lemma 2.5.4.** *Let $X$ be a Markov process and $g \in C_0$ such that a function $h \in C_0$ exists with the property that the process given by*

$$g(X_t) - g(X_0) - \int_0^t h(X_s) ds$$

*for all $t \in [0, \infty)$ is a martingale under all $\mathbb{P}_x$, $x \in E$. Then $g \in \mathcal{D}(A_X)$ and $A_X g = h$.*

*Proof.* This is [RY05, ChapterVII, Proposition (1.7)]. □

Revuz and Yor in [RY05, ChapterVII, Definition (1.8)] then proceed to define the extended generator as follows:

**Definition 2.5.5.** *Let $X$ be a Markov process and $\mathcal{B}$ the set of measurable functions from $E$ to $\mathbb{R}$. We define the relation $\mathfrak{A}_X \subseteq \mathcal{B}^2$ by defining $(g, h) \in \mathfrak{A}_X$ if and only if*

- *for all $t \geq 0$ holds $\int_0^t |h(X_s)| ds < \infty$ a.s,*
- *$\left(g(X_t) - g(X_0) - \int_0^t h(X_s) ds\right)_{t \in [0,\infty)}$ is a martingale under all $\mathbb{P}_x,\ x \in E$.*

*The relation $\mathfrak{A}_X$ is called extended generator of $X$. Furthermore, we call the set*

$$\mathbb{D}_A := \{g \in \mathcal{B} | \exists h \in \mathcal{B} : (g, h) \in \mathfrak{A}_X\}$$

*range of the extended generator and for each $g \in \mathbb{D}_A$ we set*

$$\mathfrak{A}_X(g) := \{h \in \mathcal{B} \mid (g, h) \in \mathfrak{A}_X\}.$$

As discussed in [RY05, ChapterVII], the extended generator is no real mapping, but what Revuz and Yor call a 'multi-valued almost linear mapping'. The reason is, that, if $(g, h) \in \mathfrak{A}_X$ for some Markov process $X$ then $(g, \tilde{h}) \in \mathfrak{A}$ for each $\tilde{h}$, that equals $h$ on a set, whose complement has potential zero (meaning, $X$ a.s does not occupy it for a positive amount of time). For our purpose of constructing maximum representations, we add two nuances to this definition. Therefore we remind that for Lévy processes, the generator is explicitly given in terms of the characteristic triple by a pseudo differential operator, see Proposition 2.2.4.

**Definition 2.5.6.** *Let $X$ be a Markov process and $g \in \mathbb{D}(\mathfrak{A}_X)$. Then we call a function $h \in \mathfrak{A}_X(g)$ useful (for $g$) if and only if for all $x, y \in E$ with $x \leq y$ holds*

$$\mathbb{E}_x\left(g(X_{\tau_y})\right) = g(x) + \mathbb{E}_x\left(\int_0^{\tau_y} h(X_s)ds\right).$$

*If $X$ is a Lévy process, we call a function $h \in \mathfrak{A}_X(g)$ constructive (for $g$) on a set $I \subseteq E$ if for all $x \in I$ holds*

$$\begin{aligned} h(x) =& \mu_X \frac{d}{dx} g(x) + \sigma_X \frac{d^2}{dx^2} g(x) \\ &+ \int (g(x+a) - g(a) + \mathbb{1}_{\{a \leq 1\}} a \frac{d}{dx} g(x)) \Pi_X(da). \end{aligned}$$

*and constructive, if $h$ is constructive on $E$.*

Later on, when developing maximum representations, we will aim for useful elements of the domain of the extended generator to guarantee existence of a maximum representation. When it comes to the explicit construction of the maximum representation for underlying Lévy processes, the useful constructive elements of the extended generator of the ladder height process are what we need. So let $X$ be a Lévy process and $H$ its ladder height process. The usefulness of the somehow lengthy definitions above hinges on the question if there is a large amount of functions $g \in \mathbb{D}(\mathfrak{A}_H)$ such that there are useful and (on a sufficiently large set) constructive $h \in \mathfrak{A}_H(g)$ for $g$. We will see later in full detail, that this is indeed the case, because we need to develop the precise requirements for the maximum representation. But here we already want to remark, that $H$ being a subordinator and therefore having a.s. monotone sample paths comes in handy for approximation purposes, on the other hand Remark 2.2.6 yields that for all $g \in \mathcal{D}(A_H)$ and all $x \in E$ holds

$$A_H g(x) = \tilde{\mu}_H g'(x) + \int_0^\infty \left(g(a+x) - g(x)\right) da,$$

so $A_H$ has a comparably easy structure.

## 2.6 Maximum Representations

In this section we develop the main ingredient to exploit monotone structures embedded in stopping problems as well as in impulse control problems later on, namely integral type maximum representations of the pay-off function. While in discrete time such representations simply boil down to telescoping sums, the development of such representations in continuous time bears some technical difficulties. In this section we first present the state of research depending that topic, discuss applicability to the setting herein and shortly point out the connection of maximum representations and super-harmonic functions. Then, we develop an own, (semi-)explicit way to generate the needed maximum representations for Lévy processes by utilizing the ladder height process. This construction has been developed in [CS19a] and was further adapted in [CS20]. Lastly, we discuss in how far this approach can be adapted for general Markov processes.

### 2.6.1 General remarks

The type of maximum representation for our particular setting looks as follows: for a given continuous function $\gamma$ and a continuous function $h$ we need a function $f$ such that for all $x, \overline{y} \in E$ with $x \leq \overline{y}$

$$\gamma(x) = -\mathbb{E}_x\left[\int_0^{\tau_{\overline{y}}} f\left(\sup_{r\leq t} X_r\right) dt\right] + \mathbb{E}_x\left[\gamma\left(X_{\tau_{\overline{y}}}\right) - \int_0^{\tau_{\overline{y}}} h(X_s)\, ds\right]. \tag{2.5}$$

Specifically, such functions $f$ are needed in Assumptions 3.2.2 and 3.4.2, both crucial ingredients to our solution approaches for stopping and control problem later on. Nice properties of these functions $f$ therein, as, for example, monotonicity or the right number of roots, ensure threshold times to be optimizers of stopping problems or threshold strategies to be optimizers of control problems. But before developing this particular maximum representation, we give a brief overview on literature regarding maximum representations for general Markov processes and briefly discuss to what extend maximum representations are present in control and stopping. The comprehensive intuition for the use of maximum representations in optimal stopping arises from the fact that under quite general assumptions super-harmonic functions can be characterized as functions of the form

$$x \mapsto \mathbb{E}_x\left(\sup_{t\geq 0} f(X_t)\right)$$

for some function $f$, see [FK07] and the references therein for a potential theoretic perspective on the topic. That the value function of a stopping problem under general conditions is basically the smallest super-harmonic majorant of the pay-off function, suggests the assumption that, when any maximum representation of the pay-off function in terms of a function $f$ can be found, a candidate for the smallest super-harmonic majorant is the maximum representation in terms of $f^+$, the positive part of aforementioned $f$. And, indeed, in the discounted case in [CST13], it is shown that, if the pay-off function has some kind of maximum representation and a suitable representing function $f$ is shaped nicely, indeed $f^+$ is the representing function of the value function's maximum representation. Further, in [CST13] a suggestion can be found to get such a representation: if somehow there is a terminal time $\zeta$ of $X$ such that neither $X_\zeta$ nor the running maximum $M_\zeta$ are degenerate, then one can first find a representation of $\gamma$ in terms of $X_\zeta$ (for example, with Dynkin's formula or using resolvents) and then condition on $M_\zeta$. This can be used to show existence of a maximum representation. This approach works particularly well in the discounted setting since the discounting can be interpreted as killing at an exponentially distributed time independent of the underlying process. For the problems in this thesis, the approach cannot be applied since both stopping and impulse control problems are undiscounted problems.

### 2.6.2 General Markov processes

While without restricting the process further not much can be said regarding the explicit obtainability of a maximum representation as in (2.5), with the notions developed in Section 2.5 we can at least give sufficient conditions regarding existence of such a representation. To that end let $X$ be a Markov process and take a measurable function $\gamma : E \to \mathbb{R}$ and a non-negative, measurable function $h : E \to \mathbb{R}$. Now we look at the two dimensional Markov process $(\overline{X}, X)$ consisting of $X$ and its running maximum $\overline{X}$ introduced in Definition 2.1.7 (and abuse the notation of Section 2.5 a tiny bit in a hopefully understandable way). For each function $\varphi : E \to \mathbb{R}$ we establish the notation

$$\tilde{\varphi} : E \times E \to \mathbb{R}; (x, y) \mapsto \varphi(x). \tag{2.6}$$

Assume that $\tilde{\gamma} \in \mathbb{D}(\mathfrak{A}_{(\overline{X},X)})$ and assume there is a useful $f_1 \in \mathfrak{A}_{(\overline{X},X)}(\tilde{\gamma})$ for $\tilde{\gamma}$ . Further assume, there is some $g \in \mathbb{D}(\mathfrak{A}_X)$ such that $h \in \mathfrak{A}_X(g)$ and $h$ is useful for $g$. Assume $\tilde{g} \in \mathbb{D}(\mathfrak{A}_{(\overline{X},X)})$ such that there is a useful $f_2 \in \mathfrak{A}_{(\overline{X},X)}(\tilde{g})$ for $\tilde{g}$. Then we have for all $x, \overline{y} \in E$ with $x \;\le\; \overline{y}$

$$\begin{aligned}
\mathbb{E}_x(\gamma(X_{\tau_{\overline{y}}})) &= \mathbb{E}_x(\tilde{\gamma}(\overline{X}_{\tau_{\overline{y}}}, X_{\tau_{\overline{y}}})) \\
&= \tilde{\gamma}(x, x) + \mathbb{E}_{(x,x)}\left(\int_0^{\tau_{\overline{y}}} f_1(\overline{X}_s, X_s) ds\right) \\
&= \gamma(x) + \mathbb{E}_{(x,x)}\left(\int_0^{\tau_{\overline{y}}} f_1(\overline{X}_s, X_s) ds\right)
\end{aligned}$$

and furthermore

$$\begin{aligned}
\mathbb{E}_x\left(\int_0^{\tau_{\overline{y}}} h(X_s) ds\right) &= \mathbb{E}_x(g(X_{\tau_{\overline{y}}})) - g(x) \\
&= \mathbb{E}_{(x,x)}(\tilde{g}(\overline{X}_{\tau_{\overline{y}}}, X_{\tau_{\overline{y}}})) - \tilde{g}(x, x) \\
&= \mathbb{E}_{(x,x)}\left(\int_0^{\tau_{\overline{y}}} f_2(\overline{X}_s, X_s) ds\right).
\end{aligned}$$

Putting these two equations together and setting $f := f_1 - f_2$ yields

$$\mathbb{E}_x\left(\gamma\left(X_{\tau_{\overline{y}}}\right) - \int_0^{\tau_{\overline{y}}} h(X_s) ds\right) = \gamma(x) + \mathbb{E}_{(x,x)}\left(\int_0^{\tau_{\overline{y}}} f(\overline{X}_s, X_s) ds\right)$$

If now $f$ is constant in the second argument, we have found a maximum representation as in (2.5). Although this is in general difficult to see, this approach at least provides starting points for showing existence of the needed object and in some cases even to explicitly obtain a maximum representation. One of these possible starting points is to use, that under some technical assumptions the generator of a general $d$-dimensional Markov process $(X_1, ..., X_d)$ on a given relatively compact subset of the state space $\mathbb{R}^d$ has the form

$$\begin{aligned} A_X g(x) =& \lambda(x)g(x) + \sum_{i=1}^{d} \mu_i(x) \frac{\partial g}{\partial x_i}(x) + \sum_{i,j=1}^{d} \sigma_{i,j} \frac{\partial^2 g}{\partial x_i \partial x_j} \\ &+ \int_{\mathbb{R}^d \setminus \{0\}} \left( g(y) - g(x) - \sum_{i=1}^{d} (y_i - x_i) \frac{\partial g}{\partial x_i} \right) \nu(x, dy) \end{aligned}$$

where the real valued function $\lambda$ indicates killing or creation, the real valued functions $\mu_i$ and $\sigma_{i,j}$ are drift and diffusion coefficient and the measure $\nu$ corresponds to jumps. This representation goes back to Dynkin (see [Dyn65]) and may be found in [PS06, page 129]. Now the special structure of the two dimensional Markov Process $(\overline{X}, X)$ together with approximation procedures may be used to create the objects we need for the maximum representation. We will proceed to put this general idea to use for underlying Lévy processes. Since to get a grasp on the generator of the maximum process of a Lévy process directly is quite tedious, the approach requires a little workaround using the ladder height process.

### 2.6.3 Lévy processes

In this subsection we develop an own approach for an explicit obtainability of a maximum representation of integral type as in (2.5) under the assumption that $X$ is a Lévy process, such that $\mathbb{E}(X_1)$ exists and $0 < \mathbb{E}(X_1) < \infty$. First, we will give sufficient conditions for a function $f$ as in (2.5) to exist and thereafter take some steps to the (semi-)explicit obtainability in interesting cases.
The approach heavily utilizes the ascending as well as the descending ladder height process of $X$. The definition and the needed properties can be found in Section 2.2.1. We fix a $\overline{y} \in \mathbb{R}$, a continuous non-decreasing function $\gamma : E \to \mathbb{R}$ and a continuous function $h : E \to [0, \infty)$ throughout the section and let $H$ denote the ascending ladder height process of $X$ and $H^{\downarrow}$ denote the descending ladder height process of $X$ as defined in Definition 2.2.14. Our aim is to find a function $f$ such that for all $x, \overline{y} \in \mathbb{R}$ with $x \leq \overline{y}$

$$\gamma(x) = -\mathbb{E}_x \left[ \int_0^{\tau_{\overline{y}}} f \left( \sup_{r \leq t} X_r \right) dt \right] + \mathbb{E}_x \left[ \gamma \left( X_{\tau_{\overline{y}}} \right) - \int_0^{\tau_{\overline{y}}} h(X_s)\, ds \right].$$

The applicability of such a representation is not only inseparably intertwined with the existence of such a function $f$, but also relies on the explicit obtainability. Thus we will give sufficient conditions for such an $f$ to exist and thereafter take some steps to the (semi-)explicit obtainability of $f$ in interesting cases.

**Lemma 2.6.1.** *For each non-negative function $g$ define for all $y \in E$*

$$\hat{g}(y) := \mathbb{E}_y \left( \int\limits_0^\infty g\left(H_t^{\downarrow}\right) \, dt \right).$$

*Then, for all $x, y \in E$ with $x \leq y$*

$$\mathbb{E}_x \left( \int_0^{\tau_y} g(X_t) \, dt \right) = \mathbb{E}_x \left( \int_0^{\hat{\tau}_y} \hat{g}(H_t) \, dt \right).$$

*Proof.* This result is a reformulation of [Kyp14, Exercise 7.10] and originates in [Sil80]. □

**Remark 2.6.2.** *The process $H^{\downarrow}$ acts in law like a killed subordinator, see [Kyp14, Theorem 6.9].*

Essentially, now we want to exploit that for all $y \in E$ holds $\gamma(Y_{\tau_y}) = \gamma(H_{\hat{\tau}_y})$ (where $\hat{\tau}$ is the first entry time of the ladder hight process defined in Definition 2.2.14) and then apply Dynkin's formula to $\gamma(H_{\hat{\tau}_y})$. Hence, we will make the following assumption.

**Assumption 2.6.3.** *Assume $\gamma$ is in the range of the extended generator $\mathfrak{A}_H$ of $H$ and for each compact interval $I \subseteq E$ there is a $\psi \in \mathfrak{A}_H(\gamma)$ that is useful and constructive on $I$.*

Now our candidate for $f$ looks as follows:

**Definition 2.6.4.** *Define $A_H\gamma(x) := \tilde{\mu}_H \gamma'(x) + \int_0^\infty (\gamma(a+x) - \gamma(x))\Pi_H(da)$ (Note that hereby we extend the generator to functions not lying in $C_0$). Further define*

$$f := \left(A_H\gamma - \hat{h}\right).$$

**Lemma 2.6.5.** *For all $x, y \in E$ with $x, y < \overline{y}$ holds*

$$\mathbb{E}_x \left( \gamma\left(X_{\tau_y}\right) - \int_0^{\tau_y} h(X_s) \, ds \right) = \mathbb{E}_x \left( \int_0^{\hat{\tau}_y} f(H_s) \, ds \right) + \gamma(x).$$

*Proof.* For all $x, y \in E$ with $x, y < \overline{y}$ we have, using Dynkin's formula and Lemma 2.6.1:

$$\begin{aligned}
\mathbb{E}_x \left( \gamma\left(X_{\tau_y}\right) - \int_0^{\tau_y} h(X_s) \, ds \right) &= \mathbb{E}_x \left( \gamma\left(H_{\hat{\tau}_y}\right) - \int_0^{\hat{\tau}_y} \hat{h}(H_s) \, ds \right) \\
&= \mathbb{E}_x \left( \int_0^{\hat{\tau}_y} \left(A_H\gamma - \hat{h}\right)(H_s) \, ds \right) + \gamma(x) \\
&\overset{2.6.4}{=} \mathbb{E}_x \left( \int_0^{\hat{\tau}_y} f(H_s) \, ds \right) + \gamma(x).
\end{aligned}$$

□

**Lemma 2.6.6.** *For all $x, y \in E$ with $x < y$ and all measurable functions $\varphi$ such that the following expressions exist holds*

$$\mathbb{E}_x\left[\int_0^{\hat{\tau}_y} \varphi\left(H_s\right)\ ds\right] = \mathbb{E}_x\left[\int_0^{\tau_y} \varphi\left(\sup_{r\leq s} X_r\right)\ ds\right].$$

*Proof.* This can be proven by algebraic induction (or in other words, the monotone class theorem): Wald's identity shows

$$\mathbb{E}_x\left(\hat{\tau}_y\right) = \mathbb{E}\left(L_1^{-1}\right)\mathbb{E}_x\left(\tau_y\right) = \mathbb{E}_x\left(\tau_y\right),$$

hence, the claim holds for indicator functions of the form $\mathbb{1}_{[x,y]}$ and with the Markov property this extends to indicator functions of general intervals. This carries over to simple positive functions due to linearity and with Fatou's lemma to general positive functions. Decomposition in a positive and a negative part yields the claim for general regular functions. □

As an easy consequence we get

**Lemma 2.6.7.** *For all $x \leq \overline{y}$*

$$\gamma\left(x\right) = \mathbb{E}_x\left[\int_0^{\tau_{\overline{y}}} -f\left(\sup_{r\leq t} X_r\right) dt\right] + \mathbb{E}_x\left[\gamma\left(X_{\tau_{\overline{y}}}\right) - \int_0^{\tau_{\overline{y}}} h\left(X_s\right) ds\right].$$

*Proof.* Lemma 2.6.5 with $y = \overline{y}$ combined with Lemma 2.6.6 and Lemma 2.6.1 yields

$$\begin{aligned}\gamma\left(x\right) &\overset{2.6.5}{=} \mathbb{E}_x\left[\int_0^{\hat{\tau}_{\overline{y}}} -f\left(H_s\right) ds\right] + \mathbb{E}_x\left[\gamma\left(X_{\tau_{\overline{y}}}\right) - \int_0^{\hat{\tau}_{\overline{y}}} \hat{h}\left(H_s\right) ds\right] \\ &\overset{\substack{2.6.6\\2.6.1}}{=} \mathbb{E}_x\left[\int_0^{\tau_{\overline{y}}} -f\left(\sup_{r\leq t} X_r\right) dt\right] + \mathbb{E}_x\left[\gamma\left(X_{\tau_{\overline{y}}}\right) - \int_0^{\tau_{\overline{y}}} h\left(X_s\right) ds\right].\end{aligned}$$

□

**Remark 2.6.8.** *Assume, $X$ is spectrally positive (that means $X$ has only upward jumps). Then, its descending ladder hight process $H^{\downarrow}$ acts in law as the non-random process $(t)_{t\in[0,\infty)}$ killed exponentially with a positive rate $q > 0$ where $q := \phi(0)$, $\phi$ being the right inverse of the Laplace exponent of $-X$, see [Kyp14, Subsection 6.6.2.]. Hence, the function $\hat{h}$ can be obtained via*

$$\hat{h}\left(x\right) = \int_0^{\infty} e^{-qt} h\left(t+x\right)\ dt.$$

*Further, the Lévy measure $\Pi_H$ can be expressed in terms of $q$ and the Lévy measure $\Pi_X$ of $X$ via the formula*

$$\Pi_H\left(x,\infty\right) = e^{qx}\int_x^{\infty} e^{-qy}\Pi_X\left(y,\infty\right)\ dy,$$

*see [Kyp14, Corollary 7.9].*

**Discussion of Assumption 2.6.3**

Now the crucial question is of course, how restrictive Assumption 2.6.3 is. Phrased differently, the question is, to what extend functions $\gamma$ may be approximated with functions in the domain of the generator in a suitable way for our purposes. Here, an answer is given in [ØS05, Theorem 1.24]. Assume, that $\gamma$ is twice continuously differentiable. Assume $\tau$ is an a.s. finite stopping time and assume that for all $x \in E$

$$A_H\gamma(x) = \tilde{\mu}_H\gamma'(x) + \int_0^\infty \gamma(x+y) - \gamma(x)\Pi_H(dy) \in \mathbb{R} \tag{2.7}$$

(Note that this definition of $A_H$ in Definition 2.6.4 is a generalization of the generator of $H$ to possibly unbounded functions). Then [ØS05, Theorem 1.24] states that for all $x \in E$ if both

$$\mathbb{E}_x|\gamma(H_\tau)| < \infty$$

and

$$\mathbb{E}_x\left(\int_0^\tau |A_H\gamma(H_s)|ds\right) < \infty,$$

then it holds that

$$\mathbb{E}_x\left(\gamma(X_\tau)\right) = \gamma(x) + \mathbb{E}_x\left(\int_0^\tau A_H\gamma(H_s)ds\right).$$

Therefore, only the finiteness of the above integrals in the case $\tau = \tau_y$ is needed to ensure that Assumption 2.6.3 is fulfilled.

# Chapter 3

# Stopping With Generalized Linear Costs

Optimal stopping problems with linear running costs, that often are interpreted as costs of observation, occur in several fields. One of the most prominent fields of application is sequential decision making, see [IP04] for an overview of many examples ranging from sequential statistics to finance. Although these problems have been investigated for quite a long time, general solution techniques are only known for underlying diffusion processes ([IP04, CPT12]) or certain subclasses of problems ([WLK94, Bei98]). The first two parts of this chapter are dedicated to fill in the blank. The main tool here will be the concept of monotone problems that we discussed in the introduction. First in discrete time, then in continuous time we solve one-sided stopping problems. In the discrete time case we utilize an approach similar to the one in [CI19]: we embed monotone problems into a priori non-monotone stopping settings to make use of the handy monotone structure. This technique enables us to tackle (undiscounted) problems with generalized linear costs and characterize under which conditions the problems have a threshold time as an optimizer. The second part of this chapter tackles the continuous time analogous to the discrete time problem. The right tool to carry over the idea to utilize monotonicity turns out to be the maximum representation of the pay-off function that we developed in Section 2.6. The root of the function $f$ occurring in said maximum representation yields the optimal threshold. We proceed to discuss to what extent the two cases are related and show that under some integrability conditions the continuous problem may be approximated by discrete time problems and the optimal thresholds of the discrete problems converge to the one for the continuous problem. Lastly, we tackle a non-standard continuous time stopping problem that later one is needed as an auxiliary tool to solve impulse control problems.

### Structure of the chapter

Section 3.1 entails the solution of the discrete time stopping problem. In Subsection 3.1.3 we take a closer look on the important special case that the underlying process is a random walk. Section 3.2 tackles the continuous time stopping problem. Section 3.3 compares the two problems and discusses in how far an approximation of the continuous problem with discrete problems is possible. Section 3.4 treats a stopping problem tailor-made to be applied in the solution of impulse control problems later on.

## 3.1 Discrete Time

This section entails a discrete time stopping problem for general Markov chains without discounting and general linear costs. First, we give sufficient conditions for a threshold time to be optimal in terms of what can be understood as the discrete time equivalent to the maximum representation of Section 2.6 or a spatial infinitesimal look ahead rule. After that, in Subsection 3.1.3 we show how our results simplify and essentially become a characterization in case the underlying process is a random walk.

### 3.1.1 Notation and prerequisites

Let $(Y_n)_{n\in\mathbb{N}}$ be a discrete time Markov chain on the real line. We aim to study the stopping problem for

$$\gamma\left(Y_k\right)-\sum_{i=1}^{k} h\left(Y_i\right), \quad k\in\mathbb{N}_0, \tag{3.1}$$

where $\gamma$ is a non-decreasing function and $h$ a non-negative non-decreasing function (we define the empty sum $\sum_{i=1}^{0} a_i := 0$ for all $a_i\in\mathbb{R}$).
Namely, we want to find the value function

$$V\left(y\right) := \sup_{\tau\in\mathcal{T}} \mathbb{E}_y\left(\gamma\left(Y_\tau\right)-\sum_{i=1}^{\tau} h\left(Y_i\right)\right)$$

where $\mathcal{T}$ is the set of all (real valued) stopping times. In order to make sure that waiting infinitely long is not optimal and hence there is no need to allow $\infty$ as possible value for the admissible stopping times, we make the following assumption that is quite standard for stopping problems.

**Assumption 3.1.1.**

$$\mathbb{E}_y\left(\sup_n\left(\gamma\left(Y_n\right)-\sum_{i=1}^{n} h\left(Y_i\right)\right)\right)<\infty \quad \forall y\in\mathbb{R},$$

$$\gamma\left(Y_n\right)-\sum_{i=1}^{n} h\left(Y_i\right)\to-\infty \quad a.s.$$

Additionally, to finding the value function, we want to characterize the cases in which a first entry time in an interval is an optimal stopping time, even the almost surely smallest.

**Remark 3.1.2.** *A standard result in optimal stopping without running costs is that under assumptions similar to Assumption 3.1.1 the stopping time*

$$\tau^* := \inf\{n \mid Y_n\in S^*\}$$

*for $S^* := \{y \in \mathbb{R} \mid V(y) = \gamma(y)\}$ is the a.s. smallest optimal stopping time. For example, in [Shi78, Chapter 2, Theorem 8] a similar result without running costs can be found. Many works in optimal stopping start from this characterization and then, for example, proceed to show that $S^*$ is an interval in order to verify optimality of threshold times. The approach here is different and we do not rely on knowing that $\tau^*$ is optimal beforehand.*

A main ingredient in our line of argument will be the function $f$ defined by

$$f(y) = \frac{\phi(y) - \gamma(y)}{\mathbb{E}_y(\tau^+)}, \tag{3.2}$$

for all $y \in \mathbb{R}$, where $\tau^+ := \inf\{n \in \mathbb{N} \mid Y_t > Y_0\}$ and

$$\phi(y) := \mathbb{E}_y\left(\gamma(Y_{\tau^+}) - \sum_{i=1}^{\tau^+} h(Y_i)\right)$$

for all $y \in \mathbb{R}$. Note that with the convention $\frac{-\infty}{\infty} := -\infty$ $f$ becomes a well defined function $\mathbb{R} \to \mathbb{R} \cup \{-\infty\}$. Heuristically, $f$ being positive means that the gain one would get by waiting for the process to rise above the present level exceeds the possible pay-off at the present level. The following assumption on the shape of $f$ will be essential to verify optimality of a threshold time.

**Assumption 3.1.3.**

- *There is exactly one $\overline{x} \in \mathbb{R}$ such that $f(x) > 0$ for all $x \in (-\infty, \overline{x})$, and $f(x) < 0$ for all $x \in (\overline{x}, \infty)$.*
- *On $[\overline{x}, \infty)$ the function $f$ is non-increasing.*

From now on we will always assume Assumption 3.1.1 and Assumption 3.1.3 to be true for the rest of the section. Under these assumptions we are able to show that the first entrance time into $(\overline{x}, \infty)$ or $[\overline{x}, \infty)$ is the optimizer for $V(y)$ for all $y \in \mathbb{R}$. Which type of interval is the right choice, depends on the value $f(\overline{x})$. In the following, we assume $f(\overline{x}) > 0$ and show that the first entry time in $(\overline{x}, \infty)$ is optimal. With the same line of argument, one can show that the first entry time into $[\overline{x}, \infty)$ is optimal if $f(\overline{x}) \leq 0$.

**Remark 3.1.4.** *While the numerator has an obvious interpretation, the necessity of dividing by the expected waiting time for the process to exceed the present level is not entirely obvious (although, especially if we compare the discrete time case with the continuous time case, it forges links to the generator of the ladder height process occurring therein). And, indeed, if the function $f$ as in Assumption 3.1.3 fails to be monotone after the root $\overline{x}$, one can instead work with a function*

$$\tilde{f}(y) := \frac{\phi(y) - \gamma(y)}{\mathbb{E}_y\left(\sum_{i=1}^{\tau^+} g(Y_i)\right)} = f(y) \frac{\mathbb{E}_y(\tau^+)}{\mathbb{E}_y\left(\sum_{i=1}^{\tau^+} g(Y_i)\right)}$$

*for some suitable positive function $g$. All the later proofs work with such a $\tilde{f}$ that fulfils Assumption 3.1.3 as well, nevertheless, for the sake of brevity and clarity, we will just use the 'standard' $f$.*

### 3.1.2 Proof that a threshold time is optimal

For all $x \in \mathbb{R}$ we remind of the notation $\mathring{\tau}_x = \inf\{n \in \mathbb{N}_0 \mid Y_n > x\}$ established in Definition 2.1.4. The first step towards the proof of the optimality of $\mathring{\tau}_{\overline{x}}$ is to show that $\mathring{\tau}_{\overline{x}}$ is optimal in the class of threshold times. Here, the next result will be helpful, that also connects this function $f$ with the maximum representations that we developed in Section 2.6.

**Lemma 3.1.5.** *Let*

$$\kappa_0 := 0$$

*and for all $n > 0$ let*

$$\kappa_n := \inf\{k \geq \kappa_{n-1} \mid Y_k > Y_{\kappa_{n-1}}\}$$

*be the $n$-th ladder time. Then, for all $x, y \in \mathbb{R}$ with $x \leq y$ we have*

$$\mathbb{E}_x \left( \gamma(Y_{\mathring{\tau}_y}) - \sum_{i=1}^{\mathring{\tau}_y} h(Y_i) \right) = \gamma(x) + \mathbb{E}_x \left( \sum_{\kappa_n \leq \mathring{\tau}_y} f(Y_{\kappa_n}) \mathbb{E}_{Y_{\kappa_n}} \left(\tau^+\right) \right).$$

*Proof.* Expand $\mathbb{E}_x \left(\gamma(Y_{\mathring{\tau}_y})\right) - \gamma(x)$ in a telescoping series. □

**Corollary 3.1.6.** *The stopping time $\mathring{\tau}_{\overline{x}}$ is an optimizer for the original stopping problem amongst all threshold times.*

*Proof.* Lemma 3.1.5 yields that in order to maximize $\mathbb{E}_x \left(\gamma(Y_{\mathring{\tau}_y}) - \sum_{i=1}^{\mathring{\tau}_y} h(Y_i)\right)$ in $y$, one has to sum as many positive summands of the form $f(Y_{\kappa_n})\mathbb{E}_{Y_{\kappa_n}}(\tau^+)$ on the right-hand side as possible. Assumption 3.1.3 ensures that $\mathring{\tau}_{\overline{x}}$ indeed yields the maximum, because $f$ only changes sign once. □

**Theorem 3.1.7.** *The stopping time*

$$\mathring{\tau}_{\overline{x}} = \inf\{n \geq 0 \mid Y_n > \overline{x}\}$$

*is optimal for $V(y)$ for all $y \in \mathbb{R}$.*

*Proof.* As Corollary 3.1.6 indicates, the value function when only threshold times are admissible is given by

$$\tilde{V}(y) = \begin{cases} \gamma(y) & ;\ y > \overline{x} \\ \mathbb{E}_y \left(\gamma\left(Y_{\mathring{\tau}_{\overline{x}}}\right) - \sum_{i=1}^{\mathring{\tau}_{\overline{x}}} h(Y_i)\right) & ;\ y \leq \overline{x}. \end{cases}$$

Let $y \in \mathbb{R}$. We have

$$\tilde{V}(y) \geq \gamma(y).$$

Hence, it remains to show $h$-exessivity, meaning

$$\tilde{V}(y) \geq \mathbb{E}_y \left(\tilde{V}(Y_1) - h(Y_1)\right).$$

First look at the case $y \leq \overline{x}$: here we have

$$\tilde{V}(y) = \mathbb{E}_y \left( \gamma \left( Y_{\mathring{\tau}_{\overline{x}}} \right) - \sum_{i=1}^{\mathring{\tau}_{\overline{x}}} h(Y_i) \right).$$

Further, by the strong Markov property, we have By the strong Markov property we have since $y \leq \overline{x}$

$$\mathbb{E}_y \left( \tilde{V}(Y_1) - h(Y_1) \right) = \mathbb{E}_y \left( \gamma \left( Y_{\mathring{\tau}_{\overline{x}}} \right) - \sum_{i=1}^{\mathring{\tau}_{\overline{x}}} h(Y_i) \right). \tag{3.3}$$

and (3.3) yields

$$\mathbb{E}_y \left( \tilde{V}(Y_1) - h(Y_1) \right) = \mathbb{E}_y \left( \gamma \left( Y_{\mathring{\tau}_{\overline{x}}} \right) - \sum_{i=1}^{\mathring{\tau}_{\overline{x}}} h(Y_i) \right) = \tilde{V}(y).$$

Now assume $y > \overline{x}$. We again use a type of adjusted ladder times, which we define by $\tilde{\kappa}_0 := 0$, $\tilde{\kappa}_1 := \inf\{n > 0 \mid Y_n > \overline{x}\}$ and for each $n > 1$ we set $\tilde{\kappa}_n := \inf\{n > \tilde{\kappa}_{n-1} \mid Y_n > Y_{\tilde{\kappa}_{n-1}}\}$. Again we use the notation $\tau^+ = \inf\{t \geq 0 \mid Y_t > Y_0\}$. Since this expression will occur later, we notice that

$$\begin{aligned}
&\sum_{n=0}^{\infty} \mathbb{E}_y \left( \mathbb{E}_{Y_{\tilde{\kappa}_n}} \left( \sum_{i=1}^{\tau^+} h(Y_i) \right) 1_{\{\tilde{\kappa}_n < \mathring{\tau}_y\}} \right) \\
&= \sum_{n=0}^{\infty} \mathbb{E}_y \left( \mathbb{E}_y \left( \sum_{i=\tilde{\kappa}_n+1}^{\tilde{\kappa}_{n+1}} h(Y_i) \,|\mathcal{F}_{\tilde{\kappa}_n} \right) 1_{\{\tilde{\kappa}_n < \mathring{\tau}_y\}} \right) \\
&= \sum_{n=0}^{\infty} \mathbb{E}_y \left( \mathbb{E}_y \left( \left( \sum_{i=\tilde{\kappa}_n+1}^{\tilde{\kappa}_{n+1}} h(Y_i) \right) 1_{\{\tilde{\kappa}_n < \mathring{\tau}_y\}} |\mathcal{F}_{\tilde{\kappa}_n} \right) \right) \\
&= \sum_{n=0}^{\infty} \mathbb{E}_y \left( \left( \sum_{i=\tilde{\kappa}_n+1}^{\tilde{\kappa}_{n+1}} h(Y_i) \right) 1_{\{\tilde{\kappa}_n < \mathring{\tau}_y\}} \right) \\
&= \mathbb{E}_y \left( \sum_{i=1}^{\mathring{\tau}_y} h(Y_i) \right).
\end{aligned} \tag{3.4}$$

We obtain

$$
\begin{aligned}
\mathbb{E}_y\left(\tilde{V}\left(Y_1\right)-h\left(Y_1\right)\right)-\phi\left(y\right) &= \mathbb{E}_y\left(\gamma\left(Y_{\mathring{\tau}_{\overline{x}}}\right)-\sum_{i=1}^{\mathring{\tau}_{\overline{x}}} h\left(Y_i\right)\right)-\phi\left(y\right)\\
&= \mathbb{E}_y\left(\gamma\left(Y_{\mathring{\tau}_{\overline{x}}}\right)-\gamma\left(Y_{\mathring{\tau}_y}\right)+\sum_{i=\mathring{\tau}_{\overline{x}}+1}^{\mathring{\tau}_y} h\left(Y_i\right)\right)\\
&= \sum_{n=0}^{\infty}\mathbb{E}_y\left(\left(\gamma\left(Y_{\tilde{\kappa}_n}\right)-\gamma\left(Y_{\tilde{\kappa}_{n+1}}\right)+\sum_{i=\tilde{\kappa}_n+1}^{\tilde{\kappa}_{n+1}} h\left(Y_i\right)\right)1_{\{\tilde{\kappa}_n<\mathring{\tau}_y\}}\right)\\
&= \sum_{n=0}^{\infty}\mathbb{E}_y\left(\left(\gamma\left(Y_{\tilde{\kappa}_n}\right)-\left(\gamma\left(Y_{\tilde{\kappa}_{n+1}}\right)-\sum_{i=\tilde{\kappa}_n+1}^{\tilde{\kappa}_{n+1}} h\left(Y_i\right)\right)\right)1_{\{\tilde{\kappa}_n<\mathring{\tau}_y\}}\right)\\
&= \sum_{n=0}^{\infty}\mathbb{E}_y\left(\left(\gamma\left(Y_{\tilde{\kappa}_n}\right)-\mathbb{E}_y\left(\left(\gamma\left(Y_{\tilde{\kappa}_{n+1}}\right)-\sum_{i=\tilde{\kappa}_n+1}^{\tilde{\kappa}_{n+1}} h\left(Y_i\right)\right)|\mathcal{F}_{\tilde{\kappa}_n}\right)\right)1_{\{\tilde{\kappa}_n<\mathring{\tau}_y\}}\right)\\
&= \sum_{n=0}^{\infty}\mathbb{E}_y\left(\left(\gamma\left(Y_{\tilde{\kappa}_n}\right)-\mathbb{E}_{Y_{\tilde{\kappa}_n}}\left(\gamma\left(Y_{\tau^+}\right)-\sum_{i=1}^{\tau^+} h\left(Y_i\right)\right)\right)1_{\{\tilde{\kappa}_n<\mathring{\tau}_y\}}\right)\\
&= \sum_{n=0}^{\infty}\mathbb{E}_y\left(\left(\gamma\left(Y_{\tilde{\kappa}_n}\right)-\phi\left(Y_{\tilde{\kappa}_n}\right)\right)1_{\tilde{\kappa}_n<\mathring{\tau}_y}\right)\\
&= \sum_{n=0}^{\infty}\mathbb{E}_y\left(\left(-f\left(Y_{\tilde{\kappa}_n}\right)\mathbb{E}_{Y_{\tilde{\kappa}_n}}\left(\sum_{i=1}^{\tau^+} h\left(Y_i\right)\right)\right)1_{\{\tilde{\kappa}_n<\mathring{\tau}_y\}}\right)\\
&\overset{f\searrow}{\leq} \sum_{n=0}^{\infty}\mathbb{E}_y\left(\left(-f\left(y\right)\mathbb{E}_{Y_{\tilde{\kappa}_n}}\left(\sum_{i=1}^{\tau^+} h\left(Y_i\right)\right)\right)1_{\{\tilde{\kappa}_n<\mathring{\tau}_y\}}\right)\\
&= \sum_{n=0}^{\infty}\mathbb{E}_y\left(\left(\left(\gamma\left(y\right)-\phi\left(y\right)\right)\frac{\mathbb{E}_{Y_{\tilde{\kappa}_n}}\left(\sum_{i=1}^{\tau^+} h\left(Y_i\right)\right)}{\mathbb{E}_y\left(\sum_{i=1}^{\mathring{\tau}_y} h\left(Y_i\right)\right)}\right)1_{\{\tilde{\kappa}_n<\mathring{\tau}_y\}}\right)\\
&\overset{(3.4)}{=} \gamma\left(y\right)-\phi\left(y\right)\\
&= \tilde{V}\left(y\right)-\phi\left(y\right).
\end{aligned}
$$

□

**Remark 3.1.8.** *Since we are aiming for the one-sided case, it, of course, seems natural to use the ascending ladder times to construct $f$. We consider it worth mentioning that when one instead uses 'skew' ladder times of the form $\tilde{\tau}^+ := \inf\{n \in \mathbb{N} \mid l(Y_n) > l(y)\}$ for some suitable (possibly non-monotonic) function $l$, in principle the same proofs as above still work and can lead to a characterization of the stopping sets by roots of an analogue function $f$ even in more complicated cases. However, the applicability of this more general result in* *concrete examples requires to 'guess' the right function $l$, which we only managed to do in trivial cases or situations which can be reduced to the one-sided case anyway. So we decided to stick to the one-sided situation in the proofs.*

### 3.1.3 Special case: random walk

A nice application of our theory is the stopping problem for a random walk with linear costs. Assume $(X_i)_{i\in\mathbb{N}}$ is a sequence of independent, identically distributed random variables with $\mathbb{P}(X_1 > 0) > 0$. For all $n \in \mathbb{N}_0$ let

$$S_n := \sum_{i=1}^{n} X_i.$$

Then with the usual family of measures given by

$$\mathbb{P}_y(S_n \in dx) := \mathbb{P}(S_n + y \in dx)$$

for all $y \in \mathbb{R}, n \in \mathbb{N}_0$ the process $S$ becomes a discrete time Markov process and the previous results are applicable. Using the notation of the previous sections, we have

$$\phi(y) = \mathbb{E}\left(\gamma(y + S_{\tau^+}) - \sum_{i=1}^{\tau^+} h(y + S_i)\right).$$

Further,

$$\gamma(y) - \phi(y) = \mathbb{E}\left(\gamma(y) - \gamma\left(y + S_{\tau_+}\right)\right) + \mathbb{E}\left(\sum_{i=1}^{\tau_+} h(y + S_i)\right)$$

for all $y \in \mathbb{R}$. Now $y \mapsto \mathbb{E}\left(\gamma(y) - \gamma\left(y + S_{\tau_+}\right)\right)$ is non-decreasing if

$$y \mapsto \gamma(y) - \gamma(y + s)$$

is non-decreasing for all $s > 0$ or equivalently if $\gamma$ is concave. With the same argument we get that

$$y \mapsto \mathbb{E}\left(\sum_{i=1}^{\tau_+} h(y + S_i)\right)$$

is non-decreasing if $h$ is non-decreasing. Hence, in the random walk case our findings read as follows:

**Theorem 3.1.9.** *Assume that $Y$ is a random walk and define*

$$\overline{x} := \inf\left\{ z \in \mathbb{R} \;\middle|\; \mathbb{E}\left(\gamma(z) - \gamma\left(z + S_{\tau_+}\right)\right) \geq -\mathbb{E}\left(\sum_{i=1}^{\tau_+} h(z + S_i)\right)\right\}.$$

*Further, let Assumption 3.1.1 hold true and assume that $\gamma$ is concave on $[\overline{x}, \infty)$ and that $h$ in non-decreasing and non-negative.*

- *If* $f(\overline{x}) \leq 0$, *then*
$$\tau^* := \inf\{n \geq 0 \mid S_n \geq \overline{x}\}$$
*is an optimal stopping time.*

- *If* $f(\overline{x}) > 0$, *then*
$$\mathring{\tau}^* := \inf\{n \geq 0 \mid S_n > \overline{x}\}$$
*is an optimal stopping time.*

**Corollary 3.1.10.** *Assume that* $Y$ *is a random walk. Further, let Assumption 3.1.1 hold true, assume that* $\gamma$ *is concave on* $[\overline{x}, \infty)$, *and that* $h$ *is constant, i.e.,* $h(x) = c$ *for some* $c > 0$. *Define*

$$\overline{x} := \inf\{z \in \mathbb{R} \mid \mathbb{E}\left(\gamma\left(z\right) - \gamma\left(z + S_{\tau_+}\right)\right) \geq -c\mathbb{E}\left(\tau_+\right)\}.$$

*Then, we get:*

- *If* $f(\overline{x}) \leq 0$, *then*
$$\tau^* := \inf\{n \geq 0 \mid S_n \geq \overline{x}\}$$
*is an optimal stopping time.*

- *If* $f(\overline{x}) > 0$, *then*
$$\mathring{\tau}^* := \inf\{n \geq 0 \mid S_n > \overline{x}\}$$
*is an optimal stopping time.*

Here we want to point out the connection to [WLK94] who study traditional parking problems for random walks without running costs. If $h$ is constant, Wald's identity can be used to transform the problem to a stopping problem with no running costs. So these results can be viewed as a generalization of the results in the mentioned article. Also the line of argument here is inspired by the considerations there, but we avoid the use of the Wiener-Hopf factorization, see also the discussion in [CI19].

## 3.2 Continuous Time

This section translates the ideas from the discrete time case to the continuous time case. As discussed before, a more elaborate approach to define an analogous function $f$ is needed. Nevertheless, the underlying idea stays roughly the same. Once again looking at one of the key ingredients for the discrete time case, Lemma 3.1.5, we see that for $x, y \in \mathbb{R}$ we have with the notations of the previous section

$$\mathbb{E}_x \left( \gamma(Y_{\mathring{\tau}_y}) - \sum_{i=1}^{\mathring{\tau}_y} h(Y_i) \right) = \gamma(x) + \mathbb{E}_x \left( \sum_{\kappa_n \leq \mathring{\tau}_y} f(Y_{\kappa_n}) \mathbb{E}_{Y_{\kappa_n}} \left( \tau^+ \right) \right).$$

The right-hand side herein can be viewed as the discrete time analogue to an expected integral of $Y$'s maximum process plugged in to $f$. This integral representation together with the monotonicity properties of $f$ postulated in Assumption 3.1.3 first ensured that $\mathring{\tau}_{\overline{x}}$ was the optimizer amongst all threshold times and later on also helped to ensure maximality in the initial problem. In the continuous time case we aim to take the same road utilizing the idea of maximum representations from Section 2.6. So in this section later we will assume that there is a function $f$ as in Assumption 2.5 that enables an integral type maximum representation and furthermore assume, this function $f$ has the right shape. Given that, we are able to follow an analogue line of argument as in the discrete time case. The approach here was developed in [CS19a] and [CS20] by Christensen and Sohr in the contexts of impulse control and stopping.

### 3.2.1 Notation and prerequisites

Let $X$ be a strong Markov process on $\mathbb{R}$ as defined in Definition 2.1.3. Let $\gamma : E \to \mathbb{R}$ and $h : E \to \mathbb{R}$ be real functions. From now on we will always assume:

**Assumption 3.2.1.** *1. $\gamma$ is non-decreasing and continuous.*

*2. $h$ is continuous and for all $x, y \in \mathbb{R}$ with $x < y$, we have*

$$\mathbb{E}_x \left( \int_0^{\tau_y} |h(X_s)| \; ds \right) < \infty.$$

Define the set $\mathcal{T}$ as the set of all stopping times $\tau$ with $\mathbb{E}_x(\tau) < \infty$ and $\mathbb{E}_x \left( \int_0^{\tau} |h(X_s)| \; ds \right) < \infty$ for all $x \in \mathbb{R}$ (note that in our main case of interest, that is $h \geq 0$, this only excludes stopping times that would yield a pay-off of the form '$\infty - \infty$'). We look at the stopping problem

$$\mathcal{V}(x) := \sup_{\tau \in \mathcal{T}} \mathbb{E}_x \left( \gamma(X_\tau) - \int_0^{\tau} h(X_t) \; dt \right)$$

for all $x \in \mathbb{R}$. With similar steps as in the discrete time case we develop sufficient conditions under that a threshold time whose threshold is given as the root of a function $f$ is optimal. A key issue of these stopping problems is that there occurs no killing and/or discounting, which excludes the possibility to use resolvents to obtain maximum representations, what is one of the major approaches in settings with discounting, see, among others, [MS07], [NS07], [Sur07] and [CST13]. Instead, the right type of maximum representation is an integral type one, as (2.5) in Section 2.6. To make this approach work, we then proceed to justify that it suffices to only work with stopping times bounded by first entry times to intervals of the form $[\overline{y}, \infty)$ for large enough $\overline{y}$.

**Assumption 3.2.2.** *We assume that there is a function $f$ such that:*

1. *For all $x, \overline{y} \in \mathbb{R}$ with $x \leq \overline{y}$ holds*

$$\gamma(x) = -\mathbb{E}_x\left[\int_0^{\tau_{\overline{y}}} f\left(\overline{X}_t\right) dt\right] + \mathbb{E}_x\left[\gamma\left(X_{\tau_{\overline{y}}}\right) - \int_0^{\tau_{\overline{y}}} h\left(X_s\right) ds\right].$$

2. *The function $f$ has a unique root $\overline{x} \in \mathbb{R}$, is strictly positive on $(-\infty, \overline{x})$ and is strictly decreasing on $[\overline{x}, \infty)$.*

For the remainder of the section we fix such a function $f$ and denote its unique root with $\overline{x}$. Note that this maximum representation may be seen as the analogue to the discrete telescoping sum in Lemma 3.1.5. As discussed in Remark 2.1.8, in a slight abuse of notations, the measure $\mathbb{P}_x$ is assumed to be $\mathbb{P}_{(x,x)}$, the measure corresponding to the two dimensional Markov process $\left(\overline{X}_t, X_t\right)_{t\geq 0}$ started in $(x, x)$. This enables us to still use the Markov property. To be able to make use of the maximum representation, we need to show that the value of the stopping problem does not change if we only maximize over a subset of stopping times. This subset will be called the set of upper regular stopping times.

**Definition 3.2.3.** *Call a stopping time $\tau$ upper regular if there is a value $\overline{y} \in \mathbb{R}$ such that $\tau$ is under all $\mathbb{P}_y$ a.s. bounded by the first entry time of $X$ into $[\overline{y}, \infty)$. Define $\mathcal{U} := \{\tau \in \mathcal{T} \mid \tau$ is upper regular$\}$.*

The name of as well as the idea to utilize upper regular stopping times is inspired by [IP04], where in a two sided optimal stopping problem for diffusions the set of stopping times to maximize over is restricted to regular stopping times.

## 3.2.2 Proof that a threshold time is optimal

Now we have everything at hand to show that under the assumptions made above a threshold time is optimal. The first step is to show that it suffices to consider upper regular stopping times.

**Lemma 3.2.4.** *For all $x \in \mathbb{R}$ we have*

$$\mathcal{V}(x) = \sup_{\tau \in \mathcal{U}} \mathbb{E}_x\left(\gamma\left(X_\tau\right) - \int_0^\tau h\left(X_s\right)\ ds\right).$$

*Proof.* Let $\epsilon > 0$. Take $\tau \in \mathcal{T}$ that is $\epsilon$-optimal. Set for all $n \in \mathbb{N}$

$$\sigma_n := \tau \wedge \inf\{t \geq 0 \mid X_t \geq n\}.$$

The stopping time $\tau$ is a.s. finite, thus we have $\sigma_n \to \tau$ a.s. under all $\mathbb{P}_z$ and since $\int_0^\tau |h(X_s)| \, ds$ works as an integrable majorant, we get with dominated convergence

$$\mathbb{E}_x \left( \int_0^\tau h(X_s) \, ds \right) = \lim_{n\to\infty} \mathbb{E}_x \left( \int_0^{\sigma_n} h(X_s) \, ds \right).$$

Again dominated convergence yields

$$\begin{aligned} \mathbb{E}_x (\gamma(X_\tau)) &= \mathbb{E}_x \left( \lim_{n\to\infty} \gamma(X_\tau) \wedge \gamma(X_{\sigma_n}) \right) \\ &= \lim_{n\to\infty} \mathbb{E}_x (\gamma(X_\tau) \wedge \gamma(X_{\sigma_n})) \\ &\leq \lim_{n\to\infty} \mathbb{E}_x (\gamma(X_{\sigma_n})). \end{aligned}$$

Altogether we get

$$\mathbb{E}_x \left( \gamma(X_\tau) - \int_0^\tau h(X_s) \, ds \right) \leq \lim_{n\to\infty} \mathbb{E}_x \left( \gamma(X_{\sigma_n}) - \int_0^{\sigma_n} h(X_s) \, ds \right).$$

□

The next step towards the solution is the analogue to the monotone problem in the discrete time case.

**Lemma 3.2.5.** *Let $x, \overline{y} \in \mathbb{R}$ with $x \leq \overline{y}$. For $a_{\overline{y}} := \overline{y} \wedge \overline{x}$ holds*

$$\sup_{\tau \leq \tau_{\overline{y}}} \mathbb{E}_x \left[ \int_0^\tau f\left(\overline{X}_t\right) dt \right] = \mathbb{E}_x \left[ \int_0^{\tau_{a_{\overline{y}}}} f\left(\overline{X}_t\right) dt \right].$$

*Proof.* This is a direct consequence of the properties of $f$ posed upon it in Assumption 3.2.2 and the monotonicity of $\overline{X}$. □

Now we have everything at hand to prove the main theorem of this section: showing that the threshold time $\tau_{\overline{x}}$ is optimal provided we have a maximum representation as in Assumption 3.2.2. Note that this maximum representation will be used in a similar way as the telescoping sum over the ladder times in the proof of Theorem 3.1.7.

**Theorem 3.2.6.** *Let Assumption 3.2.2 hold. Then, for all $x \in \mathbb{R}$ holds*

$$\mathcal{V}(x) = \mathbb{E}_x \left( \gamma\left(X_{\tau_{\overline{x}}}\right) - \int_0^{\tau_{\overline{x}}} h(X_s) \, ds \right).$$

*Proof.* We define for all $x \in \mathbb{R}$

$$\tilde{g}(x) := \mathbb{E}_x\left(\gamma\left(X_{\tau_{\overline{x}}}\right) - \int_0^{\tau_{\overline{x}}} h(X_s)\ ds\right).$$

One immediately sees that $\mathcal{V} \geq \tilde{g} \geq \gamma$. Let $x \in \mathbb{R}$. Lemma 3.2.4 tells us that it suffices to show

$$\tilde{g}(x) \geq \sup_{\tau \in \mathcal{U}} \mathbb{E}_x\left(\gamma(X_\tau) - \int_0^{\tau} h(X_s)\ ds\right)$$

in order to prove $\mathcal{V} = \tilde{g}$.
Let $\tau \in \mathcal{U}$ be an upper regular stopping time and fix a $\overline{y} > \overline{x}, x$ such that $\tau \leq \tau_{\overline{y}}$ a.s. under $\mathbb{P}_x$. Then we have, using Assumption 3.2.2 four times,

$$\begin{aligned}
&\mathbb{E}_x\left[\gamma(X_\tau) - \int_0^{\tau} h(X_s)\ ds\right] \\
=&\mathbb{E}_x\left\{-\mathbb{E}_{X_\tau}\left[\int_0^{\tau_{\overline{y}}} f\left(\overline{X}_t\right) dt\right]\right. \\
&\left.+\mathbb{E}_{X_\tau}\left[\gamma\left(X_{\tau_{\overline{y}}}\right) - \int_0^{\tau_{\overline{y}}} h(X_s)\ ds\right] - \left[\int_0^{\tau} h(X_s)\ ds\right]\right\} \\
=&-\mathbb{E}_x\left\{\mathbb{E}_{X_\tau}\left[\int_0^{\tau_{\overline{y}}} f\left(\overline{X}_t\right) dt\right]\right\} + \mathbb{E}_x\left[\gamma\left(X_{\tau_{\overline{y}}}\right) - \int_0^{\tau_{\overline{y}}} h(X_s)\ ds\right] \\
=&-\mathbb{E}_x\left\{\mathbb{E}_x\left[\int_\tau^{\tau_{\overline{y}}} f\left(\overline{X}_t\right) dt \mid \mathcal{F}_\tau\right]\right\} + \mathbb{E}_x\left[\gamma\left(X_{\tau_{\overline{y}}}\right) - \int_0^{\tau_{\overline{y}}} h(X_s)\ ds\right] \\
\leq&-\mathbb{E}_x\left[\int_{\tau_{\overline{x}}}^{\tau_{\overline{y}}} f\left(\overline{X}_t\right) dt\right] + \mathbb{E}_x\left[\gamma\left(X_{\tau_{\overline{y}}}\right) - \int_0^{\tau_{\overline{y}}} h(X_s)\ ds\right] \\
=&-\mathbb{E}_x\left[\int_{\tau_{\overline{x}}}^{\tau_{\overline{y}}} f\left(\overline{X}_t\right) dt\right] + \gamma(x) + \mathbb{E}_x\left[\int_0^{\tau_{\overline{y}}} f\left(\overline{X}_t\right)\ dt\right] \\
=&\gamma(x) + \mathbb{E}_x\left[\int_0^{\tau_{\overline{x}}} f\left(\overline{X}_t\right) dt\right] \\
=&\mathbb{E}_x\left[\gamma\left(X_{\tau_{\overline{x}}}\right) - \int_0^{\tau_{\overline{x}}} h(X_s)\ ds\right].
\end{aligned}$$

□

**Remark 3.2.7.** *The line of proof in Theorem 3.2.6 also directly yields, what happens if the function $f$ in the maximum representation is negative. Then stopping immediately is the optimal stopping time regardless of the starting point. This can be seen by just inserting* 0 *in the calculations for $\tau_{\overline{x}}$.*

## 3.3 Connection of the Problems

As already mentioned when discussing the lines of argument in Sections 3.1 and 3.2, there are deep inherent similarities between the discrete time and the continuous time stopping problem. Especially the similarity in the functions $f$ that determine the optimal threshold suggests that the solution of the continuous problem can be found via discretization. Hence, this section aims to give conditions under that the solution of suitably embedded discrete problems converges to the solution of the continuous problem.
Again, let $X$ be a strong Markov process on $\mathbb{R}$ as defined in Definition 2.1.3. As before, we look at the continuous time stopping problem

$$\mathcal{V}(x) = \sup_{\tau\in\mathcal{T}} \mathbb{E}_x \left( \gamma\left(X_\tau\right) - \int_0^\tau h\left(X_t\right)\, dt \right)$$

for all $x \in \mathbb{R}$, where $\gamma$ and $h$ are function that fulfil Assumption 3.2.1. Further for the sake of easier lines of argument assume that $h \geq 0$.

**Definition 3.3.1.** *We call a sequence of ascending sequences of stopping times* $((\tau_k^n)_{k\in\mathbb{N}})_{n\in\mathbb{N}}$ *a suitable discretization if*

- $\{\tau_k^n \mid k \in \mathbb{N}\} \subseteq \{\tau_k^{n+1} \mid k \in \mathbb{N}\}$ *for all* $n \in \mathbb{N}$,
- *for each* $n \in \mathbb{N}$ *the process* $Y^n$ *defined by* $Y_k^n := X_{\tau_k^n}$ *for all* $k \in \mathbb{N}$ *is a Markov process,*
- *for each* $X$*-stopping time* $\sigma$, *for each* $n \in \mathbb{N}$ *there is a* $Y^n$*-stopping time* $\tilde{\sigma}^n$ *such that with the definition* $\lceil\sigma\rceil^n := \tau_{\tilde{\sigma}^n}^n$ *we have for all* $n \in \mathbb{N}$ *that* $\sigma \leq \lceil\sigma\rceil^{n+1} \leq \lceil\sigma\rceil^n$ *and a.s.*

$$\lim_{n\to\infty} \lceil\sigma\rceil^n = \sigma.$$

We say a suitable discretization $((\tau_k^n)_{k\in\mathbb{N}})_{n\in\mathbb{N}}$ harmonizes with $\gamma$, if for all $x \in \mathbb{R}$ and all $X$-stopping times $\sigma$

$$\lim_{n\to\infty} \mathbb{E}_x \left(\gamma(X_{\lceil\sigma\rceil^n})\right) = \mathbb{E}_x\left(\gamma(X_\sigma)\right).$$

We say a suitable discretization $((\tau_k^n)_{k\in\mathbb{N}})_{n\in\mathbb{N}}$ harmonizes with $h$, if for each $n \in \mathbb{N}$ there is a function $\lceil h\rceil^n$ such that for all $x \in \mathbb{R}$ and all $X$-stopping times $\sigma$ we have

$$\mathbb{E}_x \left( \sum_{i=1}^{\tilde{\sigma}^n} \lceil h\rceil^n(Y_i^n) \right) \geq \mathbb{E}_x \left( \sum_{i=1}^{\tilde{\sigma}^{n+1}} \lceil h\rceil^{n+1}(Y_i^{n+1}) \right)$$

and

$$\lim_{n\to\infty} \mathbb{E}_x \left( \sum_{i=1}^{\tilde{\sigma}^n} \lceil h\rceil^n(Y_i^n) \right) = \mathbb{E}_x \left( \int_0^\sigma h(X_s)ds \right).$$

With these definitions at hand we immediately get the following result:

**Theorem 3.3.2.** *Assume $((\tau_k^n)_{k\in\mathbb{N}})_{n\in\mathbb{N}}$ is a suitable discretization. Using the notations from Definition 3.3.1, define for all $n \in \mathbb{N}$*

$$\mathcal{V}_n(x) := \sup_\sigma \mathbb{E}_x \left( \gamma(Y_\sigma^n) - \sum_{i=1}^{\sigma} \lceil h \rceil^n (Y_i^n) \right),$$

*where the supremum is taken over all $Y^n$-stopping times $\sigma$ such that $\tau_\sigma^n \in \mathcal{T}$. Assume each $\mathcal{V}_n$ has an optimal stopping time and let*

$$\mathcal{S}_n := \{x \in \mathbb{R} \mid \mathcal{V}_n(x) = \gamma(x)\}$$

*be the stopping set of $\mathcal{V}_n$ and*

$$\mathcal{S} := \{x \in \mathbb{R} \mid \mathcal{V}(x) = \gamma(x)\}$$

*the one for $\mathcal{V}$. Then we have point-wise*

$$\mathcal{V}_n \nearrow \mathcal{V}$$

*and*

$$\bigcap_{n\in\mathbb{N}} \mathcal{S}_n = \mathcal{S}.$$

*Proof.* Observe that

$$\gamma \le \mathcal{V}_n \le \mathcal{V}_{n+1} \le \mathcal{V}$$

for all $n \in \mathbb{N}_0$. Hence, $\lim\limits_{n\to\infty} \mathcal{V}_n$ exists and $\lim\limits_{n\to\infty} \mathcal{V}_n \le \mathcal{V}$. Let $x \in \mathbb{R}$. We only treat the case that $\mathcal{V}(x) < \infty$, the case $\mathcal{V}(x) = \infty$ works analogously with the obvious alterations. For each $x \in \mathbb{R}$ and each $\epsilon > 0$ there is an $X$-stopping time $\sigma$ such that

$$\mathcal{V}(x) \le \mathbb{E}_x \left( \gamma(X_\sigma) - \int_0^\sigma h(X_s)\, ds \right) + \frac{\epsilon}{2}$$

and an $N \in \mathbb{N}$ such that

$$\mathbb{E}_x \left( \gamma(X_\sigma) - \int_0^\sigma h(X_s)\, ds \right) \le \mathbb{E}_x \left( \gamma\left(X_{\lceil\sigma\rceil^N}\right) - \sum_{i=1}^{\tilde{\sigma}^N} \lceil h \rceil^N \left(Y_i^N\right) \right) + \frac{\epsilon}{2}.$$

Hence, we get

$$\begin{aligned} \mathcal{V}(x) &\le \mathbb{E}_x \left( \gamma(X_\sigma) - \int_0^\sigma h(X_s)\, ds \right) + \frac{\epsilon}{2} \\ &\le \mathbb{E}_x \left( \gamma\left(X_{\lceil\sigma\rceil^N}\right) - \sum_{i=1}^{\tilde{\sigma}^N} \lceil h \rceil^N (Y_i) \right) + \epsilon \\ &\le \mathcal{V}_N(x) + \epsilon \\ &\le \lim_{n\to\infty} \mathcal{V}_n(x) + \epsilon. \end{aligned}$$

Assume $x \in \mathcal{S}_n$ for all $n \in \mathbb{N}$. Then, $\gamma(x) = \lim_{n\to\infty} \mathcal{V}_n(x) = \mathcal{V}(x)$. □

**Remark 3.3.3.** *Note that these proofs don't even depend on one-dimensionality. The only requirement needed to have suitable discretizations that harmonizes with $\gamma$ and $h$ is strong enough combined continuity properties of the functions and the process.*

**Remark 3.3.4.** *While the theory above is of a very general nature, the most interesting case for the setting here is the one-dimensional one-sided case. Above results imply that, if the discretized problems are one-sided, the continuous one also is. Additionally, then the threshold of the continuous problem is the monotone limit of the discretized problems' thresholds.*

Now that we have given conditions under that suitable discrete problems approximate the continuous one in the right way, depending on the process and the functions $h$ and $\gamma$, we have to find suitable discretizations.

### 3.3.1 Non-random discretization for Lévy processes

As an example for a discretization, let $X$ be a Lévy process with $0 < \mathbb{E}(X_1) < \infty$, nevertheless, we want to mention that under the right assumptions the results remain to hold for more general processes. One of the simplest imaginable discretizations is $((\tau_k^n)_{k\in\mathbb{N}})_{n\in\mathbb{N}} = \left((\frac{k}{2^n})_{k\in\mathbb{N}}\right)_{n\in\mathbb{N}}$. We assume $\gamma$ to be differentiable and that there is an $M \in \mathbb{R}$ such that $|\gamma|, |h| < M$. For all $x \in \mathbb{R}$ and all $n \in \mathbb{N}$ we set

$$\lceil h \rceil^n(x) := \mathbb{E}_x\left(\int_0^{\frac{1}{2^n}} h(X_s)ds\right)$$

and for all $X$-stopping times $\sigma$ and all $n \in \mathbb{N}$

$$\lceil \sigma \rceil^n := \frac{\lfloor 1 + \sigma 2^n \rfloor}{2^n}$$

and with that implicitly $\tilde{\sigma}^n$ for all $n \in \mathbb{N}$. Now $\mathbb{E}_x|X_1| < \infty$ implies

$$\mathbb{E}_x\left(\sup_{t\le 1}|X_t|\right) < \infty,$$

see [Gut75], and hence we have for all $\sigma \in \mathcal{T}$

$$\left|\mathbb{E}_x\left(\gamma\left(X_{\lceil\sigma\rceil^n}\right) - \gamma\left(X_\sigma\right)\right)\right| \le M\mathbb{E}_x\left|X_\sigma - X_{\lceil\sigma\rceil^n}\right| \le M\mathbb{E}_0(\sup_{t\le\frac{1}{2^n}}|X_t|) \overset{n\to\infty}{\to} 0$$

as well as

$$\left|\mathbb{E}_x\left(\sum_{i=1}^{\tilde{\sigma}^n}\lceil h\rceil^n\left(Y_i^n\right) - \int_0^\sigma h\left(X_s\right)ds\right)\right| \le M\frac{1}{2^n} \overset{n\to\infty}{\to} 0.$$

This yields that the non-random discretization is a suitable discretization that harmonizes with $\gamma$ and $h$. We want to remark two things: First, one can weaken the boundedness assumptions to $\gamma$ and $h$, as shown in [Bei98] by approximation with bounded functions. Second, the trade-off for the relatively strong restrictions on the functions is a pretty high compatibility of discrete and continuous problems. When one assumes concavity of $\gamma$ and monotonicity of $h$, the functions $\lceil h \rceil^n$ are also monotone, the $Y^n$ are random walks and hence this immediately yields that the $\mathcal{S}_n$ are one-sided intervals as discussed in Subsection 3.1.3.

### 3.3.2 Spatial discretization

Another approach, which nicely stresses out the connection of the representing functions $f$ of the discrete and continuous problem, is to use a separation of the state space instead of the time axis. Again let $X$ be a Lévy process that fulfils $0 < \mathbb{E}_0(X_1) < \infty$ and assume $X$ to be no compound Poisson process. For each $k, n \in \mathbb{N}$, define $\tau_0^n := 0$ and

$$\tau_k^n := \inf \left\{ t \geq \tau_{k-1}^n \middle| X_t \in \mathbb{R} \setminus \left[ \frac{\lfloor 2^n X_{\tau_{k-1}^n} \rfloor}{2^n}, \frac{\lfloor 1 + 2^n X_{\tau_{k-1}^n} \rfloor}{2^n} \right) \right\}.$$

Again we use the obvious choice

$$\lceil h \rceil^n (x) := \mathbb{E}_x \left( \int_0^{\tau_1^n} h(X_s) ds \right)$$

and define

$$\lceil \sigma \rceil^n := \inf \left\{ t \geq \sigma \middle| X_t \in \mathbb{R} \setminus \left[ \frac{\lfloor 2^n X_\sigma \rfloor}{2^n}, \frac{\lfloor 1 + 2^n X_\sigma \rfloor}{2^n} \right) \right\}$$

for each $\sigma \in \mathcal{T}$ and each $n \in \mathbb{N}$ (and with that also implicitly $\tilde{\sigma}^n$). we see that $((\tau_k^n)_{k\in\mathbb{N}})_{n\in\mathbb{N}}$ is a suitable discretization and if we assume $h$ to be continuous, $\gamma$ to be smooth enough to be in the range of the generator $A_X$ (understood in the extended case as defined in Definition 2.6.4) of $X$ and that for all $n \in \mathbb{N}$ Dynkin's formula is applicable to $\tau_1^n$, we get for each $\sigma \in \mathcal{T}$

$$\begin{aligned} \left| \mathbb{E}_x \left( \gamma \left( X_{\lceil \sigma \rceil^n} \right) - \gamma \left( X_\sigma \right) \right) \right| &= \mathbb{E}_x \left( \int_\sigma^{\lceil \sigma \rceil^n} A_X \gamma(X_s) ds \right) \\ &\leq \mathbb{E}_x \left( \mathbb{E}_{X_\sigma} \left( \int_0^{\tau_1^n} A_X \gamma(X_s) ds \right) \right) \end{aligned}$$

and also

$$\left| \mathbb{E}_x \left( \sum_{i=1}^{\tilde{\sigma}^n} \lceil h \rceil^n \left( Y_i^n \right) - \int_0^\sigma h \left( X_s \right) ds \right) \right| \leq \mathbb{E}_x \left( \mathbb{E}_{X_\sigma} \left( \int_0^{\tau_1^n} h(X_s) ds \right) \right).$$

Now only some boundedness assumptions are needed to ensure that both these terms converge to zero. For example, known properties of the stopping region, like boundedness of the continuation region and monotonicity of $\gamma$, or if one can justify only to consider bounded stopping times may help here. Instead of going into more detail regarding this, we want to emphasize that with this discretization one can nicely see how the representing functions of the discrete and the continuous time problems are connected. If we denote the representing function of the discrete time problem of stopping

$$\gamma(Y_k^n) - \sum_{i=1}^{\tau_k^n} \lceil h \rceil^n (Y_i^n)$$

as defined in (3.2) in Section 3.1 with $f_n$, we see that if we first assume $h = 0$ and for the sake of notational simplicity also $x \in \frac{1}{2^n}\mathbb{N}$, we have

$$\begin{aligned} f_n(x) &= \frac{\mathbb{E}_x\left(\gamma(Y^n_{\tau^+})\right) - \gamma(x)}{\mathbb{E}(\tau_1^n)} \\ &= \frac{\mathbb{E}_x\left(\gamma(H_{\hat{\tau}_{x+\frac{1}{2^n}}})\right) - \gamma(x)}{\mathbb{E}(\hat{\tau}_{\frac{1}{2^n}})} \\ &\to A_H\gamma(x), \end{aligned}$$

provided $\gamma$ is in the range of the extended generator of the ladder height process $H$, that was defined in Definition 2.2.14. To treat the case of arbitrary $h$ we use Lemma 2.6.1. With the notations used therein we see that, again for $x \in \frac{1}{2^n}\mathbb{N}$,

$$\begin{aligned} \frac{\mathbb{E}_x(\lceil h \rceil(Y^n_{\tau^+}))}{\mathbb{E}_x(\tau_1^n)} &= \frac{\mathbb{E}_x(\int_0^{\tau_{x+\frac{1}{2^n}}} h(X_s)ds)}{\mathbb{E}_x(\tau_{x+\frac{1}{2^n}})} \\ &= \frac{\mathbb{E}_x(\int_0^{\hat{\tau}_{x+\frac{1}{2^n}}} \hat{h}(H_s)ds)}{\mathbb{E}_x(\hat{\tau}_{x+\frac{1}{2^n}})} \\ &\to \hat{h}(x), \end{aligned}$$

provided $h$ is smooth enough.

## 3.4 A Tailor-Made Problem for Impulse Control

Later on, in the chapter about impulse control problems when it comes to solve control problems motivated by inventory control, the need arises to solve a non-standard stopping problem. Despite not being of the usual form of a stopping problem, it is nevertheless in some sense similar to the stopping problem we tackled in Section 3.2. Therefore, in this section adapt the approach developed therein to solve the mentioned non-standard problem.

### 3.4.1 Notations and prerequisites

Throughout the section let $X$ be a strong Markov process on a possibly unbounded interval $E \subseteq \mathbb{R}$. Let $\gamma : E \to \mathbb{R}$ and $h : E \to \mathbb{R}$ be real functions and assume:

**Assumption 3.4.1.** *1. For all $x, y \in E$ with $x < y$ we have $\mathbb{E}_x(\tau_y) < \infty$.*

*2. $\gamma$ is non-decreasing and differentiable.*

*3. $h$ is continuous and for all $x, y \in E$ with $x < y$, we have*

$$\mathbb{E}_x \left( \int_0^{\tau_y} \mid h\left(X_s\right) \mid ds \right) < \infty.$$

Later in the chapter on impulse control problems, Corollary 4.2.9 stresses the importance of finding an optimal stopping time for stopping problems of the form

$$\mathfrak{g}_\rho^{\mathcal{T}_x}\left(x\right) := \sup_{\tau \in \mathcal{T}_x} \mathbb{E}_x \left( \gamma\left(X_\tau\right) - \gamma\left(x\right) - K - \int_0^\tau \left(h\left(X_t\right) + \rho\right)\ dt \right).$$

for all $x \in E$ for some $K \in [0, \infty)$ and $\rho \in \mathbb{R}$, where the optimal stopping time is required to be in the set

$$\mathcal{T}_x = \{\tau \in \mathcal{T} \mid X_\tau \geq x \ \ \mathbb{P}_x \text{ a.s. }\}.$$

This type of problem deviates from standard ones twofold. First, only stopping above the starting point is allowed, which makes it impossible to straightforwardly exploit the Markovian structure. Second, the pay-off depends on the starting point. Despite this special structure of the problem, the techniques we developed in the previous section can be modified and adapted to this scenario. A first step herein is to instead show that a threshold time is optimal for the stopping problem given by

$$g_\rho^{\mathcal{T}_x}\left(x\right) := \sup_{\tau \in \mathcal{T}_x} \mathbb{E}_x \left( \gamma\left(X_\tau\right) - \int_0^\tau \left(h\left(X_t\right) + \rho\right)\ dt \right)$$

for all $x \in E$ for a large enough set of starting points. Then, we will show that in order to find

$$\mathfrak{G}(\rho) := \sup_{x \in E} \mathfrak{g}_\rho^{\mathcal{T}_x}\left(x\right)$$

– and that value is what we actually have to get a grip on later on in Chapter 4 – it suffices to know that a threshold time is optimal in the mentioned set of starting points by making sure the supremum in $\mathfrak{G}(\rho)$ is attained on that set. In the following, we fix a $\rho \in \mathbb{R}$ and establish sufficient conditions for a threshold time $\tau_a$ to be an optimizer for $g_\rho^{\mathcal{T}_x}(x)$ and partition the state space $\mathbb{R}$ in a set where this threshold time is the optimal one and a set that does not contribute to the supremum in $\mathfrak{G}$. The main tool to characterize the pay-off functions, or more precisely the pairs $(\gamma, h)$ of pay-off function and running costs, for this to hold is again a representation of $\gamma$ in terms of expected running suprema of integral type as in (2.5). As in the previous section, we use the maximum representation to find a solution on $(-\infty, \overline{y}]$ for given $\overline{y}$ and later on by again utilizing upper regularity show that our obtained optimizers in fact already are general optimizers if $\overline{y}$ is chosen large enough.
Now, step-by-step, we modify the arguments of the previous chapter to work in this particular setting. The adapted maximum representation we need here reads as follows:

**Assumption 3.4.2.** *We assume that there is a function $f : E \to \mathbb{R}$ such that:*

1. *For all $x, \overline{y} \in E$ with $x \leq \overline{y}$*

$$\gamma(x) = -\mathbb{E}_x\left[\int_0^{\tau_{\overline{y}}} f\left(\overline{X}_t\right) dt\right] + \mathbb{E}_x\left[\gamma\left(X_{\tau_{\overline{y}}}\right) - \int_0^{\tau_{\overline{y}}} h\left(X_s\right) ds\right].$$

2. *The function $f$ has a unique maximum $a \in E$, is strictly increasing on $(-\infty, a] \cap E$ and strictly decreasing on $[a, \infty) \cap E$.*

We again silently identify $\mathbb{P}_x$ and $\mathbb{P}_{(x,x)}$ with each other, whenever the running maximum occurs as described in Remark 2.1.8. While for the unrestricted stopping problem a unique root of the function $f$ therein was needed, our aim to apply this stopping problem in the context of impulse control later on requires some different structure of the function $f$. There are main two reasons. First, to tackle the impulse control problem, we need the solution of a stopping problem not just for one particular $\rho$ but for a range of $\rho$. Second, the representing function $f$ will also be used to determine the optimal restarting point for the control problem. These reasons lead to the following assumption in addition to Assumption 3.4.2 regarding the roots of $f$. We introduce the notation

$$f_\rho := f - \rho$$

for each $\rho \in \mathbb{R}$ and for this section fix a $\rho \in \mathbb{R}$ such that the following assumption holds:

**Assumption 3.4.3.** *The function $f_\rho$ has exactly two roots $\underline{x}, \overline{x} \in E$ with $\underline{x} < \overline{x}$, is positive on $(\underline{x}, \overline{x})$ and negative on $(\underline{x}, \overline{x})$.*

Note that Assumption 3.4.2 implies that, if Assumption 3.4.3 holds, $f$ is negative on $[\underline{x}, \overline{x}]^c$. Again we will work with upper regular stopping times as defined in Definition 3.2.3. Additionally, define for all $y \in E$ the set

$$\mathcal{U}_y := \{\tau \in \mathcal{T}_y \mid \tau \text{ is upper regular}\}.$$

### 3.4.2 Proof that in the relevant cases a threshold time is optimal

With all definitions and assumptions at hand, we are now able to prove a refinement of Lemma 3.2.4.

**Lemma 3.4.4.** *For all $x \in E$ we have*

$$g_\rho^{\mathcal{T}_x}(x) = \sup_{\tau \in \mathcal{U}_x} \mathbb{E}_x \left( \gamma(X_\tau) - \int_0^\tau (h(X_s) + \rho) \ ds \right).$$

*Proof.* This proof works just as the one for Lemma 3.2.4, with the small alteration of the definition

$$\sigma_n := \tau \wedge \inf\{t \geq 0 \mid X_t \geq n + y\}$$

to ensure that $\sigma_n \in \mathcal{U}_n$ for all $n \in \mathbb{N}$.

□

The analogue to the embedded monotone problem here reads as follows.

**Lemma 3.4.5.** *Let $x, \overline{y} \in E$ with $x \leq \overline{y}$. For $a_{\overline{y}} := \overline{y} \wedge \overline{x}$ holds*

$$\sup_{\tau \leq \tau_{\overline{y}}} \mathbb{E}_x \left[ \int_0^\tau f_\rho(\overline{X}_t) \, dt \right] = \mathbb{E}_x \left[ \int_0^{\tau_{a_{\overline{y}}}} f_\rho(\overline{X}_t) \, dt \right].$$

*Proof.* This is a direct consequence of the properties of $f$ posed upon it in Assumptions 3.4.2 and 3.4.3 as well as the monotonicity of $\overline{X}$. □

**Remark 3.4.6.** *The claim of the Lemma above even holds pathwise.*

In Section 3.2 the structure of the function $f$ divides the real line into stopping and continuation region in a very beautiful manner since one is the support of $f^- := \max\{0, -f\}$, the other the support of $f^+ := \max\{0, f\}$. The control problem we want to solve later with this stopping problem here needs the function $f$ also to determine where to restart the process again. But that requires $f_\rho$ to have another root, as postulated in Assumption 3.4.3, and therefore bears a somehow conflicting nature to the nice relation of $f$ towards the one-sidedness of the problem. To work around that obstacle, intuitively our approach can be interpreted as follows: we show that right from the second root of $f_\rho$ the monotonicity arguments of the previous section still work and the rest of the starting points are not important in a way that will clarify when we treat the impulse control problem later in its full extent. The next ingredients, a simplifying assumption and a lemma, need to be read in that light. Also the version of the main result, Theorem 3.4.9, entails this distinction in a statement about the 'important' part of the real line and a part that ensures that the rest of the real line indeed is unimportant.

**Lemma 3.4.7.** *For all $x \in E$ holds*

$$\mathbb{E}_x\left[\int_0^{\tau_{\overline{x}}} f_\rho^+\left(\overline{X}_t\right)dt\right] \le \sup_{x^*\in[\underline{x},\overline{x}]} \mathbb{E}_{x^*}\left[\int_0^{\tau_{\overline{x}}} f_\rho^+\left(\overline{X}_t\right)dt\right].$$

*Proof.* Let $\epsilon > 0$ and take a $x^* \in [\overline{x}, \underline{x}]$ such that

$$\mathbb{E}_{x^*}\left[\int_0^{\tau_{\overline{x}}} f_\rho^+\left(\overline{X}_t\right)dt\right] + \epsilon \ge \sup_{x'\in[\underline{x},\overline{x}]} \mathbb{E}_{x'}\left[\int_0^{\tau_{\overline{x}}} f_\rho^+\left(\overline{X}_t\right)dt\right].$$

Now for each $x \in E$ with $x \ge \overline{x}$ we have

$$\begin{aligned}\mathbb{E}_x\left[\int_0^{\tau_{\overline{x}}} f_\rho^+\left(\overline{X}_t\right)dt\right] = 0&\\ &=\mathbb{E}_{\overline{x}}\left[\int_0^{\tau_{\overline{x}}} f_\rho^+\left(\overline{X}_t\right)dt\right]\\ &\le\mathbb{E}_{x^*}\left[\int_0^{\tau_{\overline{x}}} f_\rho^+\left(\overline{X}_t\right)dt\right] + \epsilon.\end{aligned}$$

And for each $x \le \underline{x}$ we get

$$\begin{aligned}\mathbb{E}_x\left[\int_0^{\tau_{\overline{x}}} f_\rho^+\left(\overline{X}_t\right)dt\right] &= \mathbb{E}_x\left[\int_{\tau_{\underline{x}}}^{\tau_{\overline{x}}} f_\rho^+\left(\overline{X}_t\right)dt\right]\\ &= \mathbb{E}_x\left[\mathbb{E}_x\left[\int_{\tau_{\underline{x}}}^{\tau_{\overline{x}}} f_\rho^+\left(\overline{X}_t\right)dt\middle|\mathcal{F}_{\tau_{\underline{x}}}\right]\right]\\ &= \mathbb{E}_x\left[\mathbb{E}_{X_{\tau_{\underline{x}}}}\left[\int_0^{\tau_{\overline{x}}} f_\rho^+\left(\overline{X}_t\right)dt\right]\right]\\ &\le \mathbb{E}_x\left[\mathbb{E}_{x^*}\left[\int_0^{\tau_{\overline{x}}} f_\rho^+\left(\overline{X}_t\right)dt\right] + \epsilon\right]\\ &= \mathbb{E}_{x^*}\left[\int_0^{\tau_{\overline{x}}} f_\rho^+\left(\overline{X}_t\right)dt\right] + \epsilon.\end{aligned}$$

□

Now to avoid the need to introduce such an $\epsilon$ as in the previous proof each time, we use Lemma 3.4.7 and by that further complicate the notation, we will assume, an optimizer for the supremum used in Lemma 3.4.7 exists. Later we will see that in many cases of interest under fairly general conditions such an optimizer indeed exists.

**Assumption 3.4.8.** *There is an $x^* \in [\underline{x}, \overline{x}]$ such that for all $x \in E$:*

$$\mathbb{E}_x\left[\int_0^{\tau_{\overline{x}}} f_\rho^+\left(\overline{X}_t\right)dt\right] \le \mathbb{E}_{x^*}\left[\int_0^{\tau_{\overline{x}}} f_\rho^+\left(\overline{X}_t\right)dt\right].$$

Now we have everything at hand to state and prove the main theorem of this subsection.

**Theorem 3.4.9.** *1. For all $x \in E \cap [\underline{x}, \infty)$ holds*

$$g_\rho^{\mathcal{T}_x}(x) = \mathbb{E}_x \left( \gamma\left(X_{\tau_{\overline{x}}}\right) - \int_0^{\tau_{\overline{x}}} \left(h\left(X_s\right) + \rho\right) \, ds \right)$$

*and*

$$\mathfrak{g}_\rho^{\mathcal{T}_x}(x) = \mathbb{E}_x \left( \gamma\left(X_{\tau_{\overline{x}}}\right) - K - \gamma(x) - \int_0^{\tau_{\overline{x}}} \left(h\left(X_s\right) + \rho\right) \, ds \right).$$

*2. For all $x \in E \cap (\infty, \underline{x})$ holds*

$$\mathfrak{g}_\rho^{\mathcal{T}_x}(x) \leq \mathfrak{g}_\rho^{\mathcal{T}_{x^*}}(x^*).$$

*Proof.* We define our candidate for the value function by setting for all $x \in \mathbb{R}$

$$\tilde{g}(x) := \mathbb{E}_x \left[ \gamma\left(X_{\tau_{\overline{x}}}\right) - \int_0^{\tau_{\overline{x}}} h - \rho\left(X_s\right) ds \right].$$

Lemma 3.4.4 tells us that it suffices to show

$$\tilde{g}(x) \geq \sup_{\tau \in \mathcal{U}_x} \mathbb{E}_x \left( \gamma\left(X_\tau\right) - \int_0^{\tau} \left(h\left(X_s\right) + \rho\right) \, ds \right)$$

for all $x \in E \cap [\underline{x}, \infty)$ in order to prove $g_\rho^{\mathcal{T}_x}(x) = \tilde{g}(x)$ for all $x \in E \cap [\underline{x}, \infty)$. Let $x \in E$. Let $\tau \in \mathcal{U}_x$ be an upper regular stopping time and fix a $\overline{y} > x$ such that $\tau \leq \tau_{\overline{y}}$ $\mathbb{P}_x$ a.s. Then, we have

$$
\begin{aligned}
&\mathbb{E}_x\left[\gamma\left(X_\tau\right)-\int_0^\tau\left(h\left(X_s\right)+\rho\right)\ ds\right] \\
&\overset{3.4.2}{=}\mathbb{E}_x\left\{\mathbb{E}_{X_\tau}\left[\int_0^{\tau_{\overline{y}}}-f_\rho\left(\overline{X}_t\right)dt\right]\right. \\
&\qquad+\mathbb{E}_{X_\tau}\left[\gamma\left(X_{\tau_{\overline{y}}}\right)-\int_0^{\tau_{\overline{y}}}\left(h\left(X_s\right)+\rho\right)\ ds\right] \\
&\qquad\left.-\left[\int_0^\tau\left(h\left(X_s\right)+\rho\right)\ ds\right]\right\} \\
&=\mathbb{E}_x\left\{\mathbb{E}_{X_\tau}\left[\int_0^{\tau_{\overline{y}}}-f_\rho\left(\overline{X}_t\right)dt\right]\right\} \\
&\qquad+\mathbb{E}_x\left[\gamma\left(X_{\tau_{\overline{y}}}\right)-\int_0^{\tau_{\overline{y}}}\left(h\left(X_s\right)+\rho\right)\ ds\right] \\
&=\mathbb{E}_x\left\{\mathbb{E}_x\left[\int_\tau^{\tau_{\overline{y}}}-f_\rho\left(\overline{X}_t\right)dt\middle|\mathcal{F}_\tau\right]\right\} \\
&\qquad+\mathbb{E}_x\left[\gamma\left(X_{\tau_{\overline{y}}}\right)-\int_0^{\tau_{\overline{y}}}\left(h\left(X_s\right)+\rho\right)\ ds\right] \\
&=\mathbb{E}_x\left\{\mathbb{1}_{\{\tau\leq\tau_{\underline{x}}\}}\mathbb{E}_x\left[\int_\tau^{\tau_{\overline{y}}}-f_\rho\left(\overline{X}_t\right)dt\middle|\mathcal{F}_\tau\right]\right\} \\
&\qquad+\mathbb{E}_x\left\{\mathbb{1}_{\{\tau>\tau_{\underline{x}}\}}\mathbb{E}_x\left[\int_\tau^{\tau_{\overline{y}}}-f_\rho\left(\overline{X}_t\right)dt\middle|\mathcal{F}_\tau\right]\right\} \\
&\qquad+\mathbb{E}_x\left[\gamma\left(X_{\tau_{\overline{y}}}\right)-\int_0^{\tau_{\overline{y}}}\left(h\left(X_s\right)+\rho\right)\ ds\right] \\
&\qquad=\ldots
\end{aligned}
$$

$$\begin{aligned}
\ldots =&\mathbb{E}_x\left\{\mathbb{E}_x\left[\mathbb{1}_{\{\tau\leq\tau_{\underline{x}}\}}\int_\tau^{\tau_{\overline{y}}} -f_\rho\left(\overline{X}_t\right)dt\middle|\mathcal{F}_\tau\right]\right\}\\
&+\mathbb{E}_x\left\{\mathbb{E}_x\left[\mathbb{1}_{\{\tau>\tau_{\underline{x}}\}}\int_\tau^{\tau_{\overline{y}}} -f_\rho\left(\overline{X}_t\right)dt\middle|\mathcal{F}_\tau\right]\right\}\\
&+\mathbb{E}_x\left[\gamma\left(X_{\tau_{\overline{y}}}\right)-\int_0^{\tau_{\overline{y}}}\left(h\left(X_s\right)+\rho\right)\,ds\right]\\
\overset{3.4.3}{\leq}&\mathbb{E}_x\left[\mathbb{1}_{\{\tau\leq\tau_{\underline{x}}\}}\int_0^{\tau_{\overline{y}}} -f_\rho\left(\overline{X}_t\right)dt\right]\\
&+\mathbb{E}_x\left[\mathbb{1}_{\{\tau>\tau_{\underline{x}}\}}\int_{\tau_{\overline{x}}}^{\tau_{\overline{y}}} -f_\rho\left(\overline{X}_t\right)dt\right]\\
&+\mathbb{E}_x\left[\gamma\left(X_{\tau_{\overline{y}}}\right)-\int_0^{\tau_{\overline{y}}}\left(h\left(X_s\right)+\rho\right)\,ds\right]\\
\overset{3.4.2}{=}&\mathbb{E}_x\left[\mathbb{1}_{\{\tau\leq\tau_{\underline{x}}\}}\int_0^{\tau_{\overline{y}}} -f_\rho\left(\overline{X}_t\right)dt\right]\\
&+\mathbb{E}_x\left[\mathbb{1}_{\{\tau>\tau_{\underline{x}}\}}\int_{\tau_{\overline{x}}}^{\tau_{\overline{y}}} -f_\rho\left(\overline{X}_t\right)dt\right]\\
&+\gamma\left(x\right)+\mathbb{E}_x\left[\int_0^{\tau_{\overline{y}}} f_\rho\left(\overline{X}_t\right)\,dt\right]\\
=&\gamma\left(x\right)+\mathbb{E}_x\left[\mathbb{1}_{\{\tau>\tau_{\underline{x}}\}}\int_0^{\tau_{\overline{x}}} f_\rho\left(\overline{X}_t\right)dt\right]\\
=&:\star.
\end{aligned}$$

To prove the claim, we have to distinguish the cases $x\leq\underline{x}$ and $x>\underline{x}$. If $x\geq\underline{x}$, applying the identity from Assumption 3.4.2 yet another time yields

$$\begin{aligned}
\star&\leq\gamma\left(x\right)+\mathbb{E}_x\left[\int_0^{\tau_{\overline{x}}} f_\rho^+\left(\overline{X}_t\right)dt\right]\\
&=\mathbb{E}_x\left[\gamma\left(X_{\tau_{\overline{x}}}\right)-\int_0^{\tau_{\overline{x}}}\left(h\left(X_s\right)+\rho\right)\,ds\right]\\
&=\tilde{g}(x).
\end{aligned}$$

This shows

$$g_\rho^{\mathcal{T}_x}\left(x\right)=\mathbb{E}_x\left(\gamma\left(X_{\tau_{\overline{x}}}\right)-\int_0^{\tau_{\overline{x}}}\left(h\left(X_s\right)+\rho\right)\,ds\right)$$

and

$$\mathfrak{g}_\rho^{\mathcal{T}_x}\left(x\right)=\mathbb{E}_x\left(\gamma\left(X_{\tau_{\overline{x}}}\right)-K-\gamma(x)-\int_0^{\tau_{\overline{x}}}\left(h\left(X_s\right)+\rho\right)\,ds\right).$$

If $x < \underline{x}$, we get by using Lemma 3.4.7 and an $x^* \in [\underline{x}, \overline{x}]$ as defined in Assumption 3.4.8

$$\begin{aligned}
\star - \gamma(x) - K &= \mathbb{E}_x \left[ \mathbb{1}_{\{\tau > \tau_{\underline{x}}\}} \int_0^{\tau_{\overline{x}}} f_\rho \left( \overline{X}_t \right) dt \right] - K \\
&< \mathbb{E}_{x^*} \left[ \int_0^{\tau_{\overline{x}}} f_\rho^+ \left( \overline{X}_t \right) dt \right] - K \\
&= \mathbb{E}_{x^*} \left[ \gamma \left( X_{\tau_{\overline{x}}} \right) - \int_0^{\overline{x}} h \left( X_s \right) + \rho \; ds \right] - \gamma(x^*) - K \\
&\leq \mathfrak{g}^{\tau_{x^*}} (x^*),
\end{aligned}$$

this yields

$$\mathfrak{g}^{\tau_x} (x) \leq \mathfrak{g}^{\tau_{x^*}} (x^*).$$

□

# Chapter 4

# Long-Term Average Impulse Control

This chapter treats long-term average impulse control problems for Markov processes with generalized linear running costs. Under minimal conditions we characterize the long-term average impulse control problem with super-martingales and from that deduce a general verification theorem. Further, also under minimal conditions on the underlying process we characterize the impulse control problem by an associated stopping problem. The connection to super-martingales in Theorem 4.2.1 and the connection to a stopping problem in and after Theorem 4.2.7 can be viewed as a type of long-term average analogous to the results for discounted problems from [Chr14] we described in Section 1.2. The characterization by a stopping problem of the type solved in Section 3.4 leads to one of the main results of this chapter, namely that given an integral type maximum representation as in Assumption 3.4.2 exists an $(s, S)$-strategy is optimal and $S$ is given in easy terms of the maximum representation. We further give conditions under that $s$ is also given in a similar way. This theoretical result will be the basis for a step-by-step solution technique for Lévy processes that leads to nice (semi-)explicit solutions in many interesting special cases.

### Structure of the chapter

In Section 4.1 we introduce the necessary notations for the following proofs, define and motivate the problem and collect a bunch of necessary general results as well as some first insights on the structure of the problem that will be needed later.

In Section 4.2 we first prove a verification theorem utilizing super-martingale techniques. Then, we characterize the value of the control problem by a value of a 'tailor-made' stopping problem, that we treated in Section 3.4. We show that, provided this stopping problem has an optimizer and an optimal restarting point can be found, this optimizer can be used to construct an optimal strategy for the control problem.

In Section 4.3 we discuss, under what conditions such an optimal restarting point exists and then condense our findings to what can be viewed as the theoretical main result of this chapter, maybe even the whole thesis, Theorem 4.3.2. From Section 4.4 on, we focus on Lévy processes and under some conditions in Subsection 4.4.1 derive a characterization of the optimal restarting point by using the maximum representation. In Subsection 4.4.2 we condense our findings to an explicit step-by-step solution technique and illustrate the applicability of our technique by demonstrating how to show existence of optimal $(s, S)$-strategies in quite general settings that cover many examples of interest.

Subsection 4.4.4 is devoted to the proof of the validity of the solution technique.

## 4.1 Detailled Setup and General Results

First, we formally state the problem, give detailed definitions and introduce the necessary notations. Then, we collect some elementary results and discuss degenerate cases.

### 4.1.1 Notation and prerequisites

Let $X$ be a strong Markov process on a possibly unbounded interval $E \subseteq \mathbb{R}$ as defined in Definition 2.1.3. Let $\gamma : E \to \mathbb{R}$ and $h : E \to \mathbb{R}$ be real functions, such that Assumption 3.4.1 is fulfilled. Note that these assumptions and prerequisites are the same as posed upon process and functions in Section 3.4 what makes the results there applicable.

#### Model of the controlled process

To model the controlled process for underlying general Markov processes requires a bit of technical work and is usually done by constructing a new probability measure that resembles the distribution of the controlled process. Most works on impulse control problems refer to the dissertation [Rob81] where a construction is given. [Ste83] provides a reference in English, that itself refers to [Rob81]. [Chr14] provides a more recent source for the construction that, despite in principle being the same construction as in the previously mentioned works, outlines some other aspects of the construction and especially mentions the importance to be able to distinguish the values of the process at the exact time of the control right before and after the control is exercised (note that this discussion may be found in [Chr14, Section 2] and not in [Chr14, Appendix], where the model itself is constructed). Here, we will use the same model as [Rob81, Ste83, Chr14], albeit go into a bit more detail than the latter two works, that seem to be the only references in English for the model.
Axiomatically the model should fulfil the following:

- A control strategy $S = (\tau_n, \zeta_n)_{n\in\mathbb{N}}$ should consist of a sequence of increasing stopping times $(\tau_n)_{n\in\mathbb{N}}$ and $\mathcal{F}_{\tau_n}$-measurable random variables $\zeta_n$, that indicate whereto the process is shifted.
- Between the controls the controlled process evolves following the initial Markovian dynamics of $X$.
- At the time of a control $\tau_n$ we have to be able to work with the value of the controlled process 'right before the control', or to be more precise, at time $\tau_n$ but with the control not having taken place yet, that we call $X^S_{\tau_n,-}$, which in general (for processes with jumps) may deviate from both the value $X^S_{\tau_n} = \zeta_n$ at time $\tau_n$ after the control has taken place and the left limit $X^S_{\tau_n-}$.

To model the controlled process, we work on the new state space $\tilde{\Omega} = \Omega^{\mathbb{N}}$. A control strategy $S = (\tau_n, \zeta_n)_{n\in\mathbb{N}}$ is defined as a sequence consisting of a sequence of increasing stopping times $(\tau_n)_{n\in\mathbb{N}}$ in $\mathcal{T}$ (as defined in Definition 2.1.5, 1. with the obvious adaptation of $\mathcal{T}$ to the new space). When calling the $\tau_i$ stopping times, it is not clear with respect to which filtration. For all $n \in \mathbb{N}$

$$\tau_n \text{ shall be a } \left( \bigotimes_{i\leq n} \mathcal{F}_t \otimes \bigotimes_{i\in\mathbb{N}, i>n} \{\emptyset, \Omega\} \right)_{t\in[0,\infty)} \text{ stopping time}$$

and

$$\zeta_n \text{ shall be } \bigotimes_{i\leq n} \mathcal{F}_{\tau_n} \otimes \bigotimes_{i\in\mathbb{N}, i>n} \{\emptyset, \Omega\} - \text{ measurable,}$$

such that $\gamma(\zeta_n)$ is integrable (as the new 'default' probability measure we take $\tilde{\mathbb{P}} := \bigotimes_{n\in\mathbb{N}} \mathbb{P}$). Now define for each $n \in \mathbb{N}$ and each $(\omega_i)_{i\in\mathbb{N}} \in \tilde{\Omega}$

$$X^n \left( (\omega_i)_{i\in\mathbb{N}} \right) := X(\omega_n).$$

We construct a new probability measure by setting for all $n \in \mathbb{N}$, all $x \in E$, all $t \in [0, \infty)$ and all measurable sets $A_1, ..., A_n, A_{n+1}$

$$\tilde{\mathcal{F}}^n := \bigotimes_{i\leq n} \mathcal{F}_{\tau_n} \otimes \bigotimes_{i\in\mathbb{N}, i>n} \{\emptyset, \Omega\}$$

and

$$\mathbb{P}_x^S \left( X^i_{\tau_n+t} \in A_i, \forall i \leq n+1 \mid \tilde{\mathcal{F}}^n \right) := \prod_{i\leq n} \delta_{X^i_{\tau_i}}(A_i)\, \mathbb{P}_{\zeta_n}(X_t \in A_{n+1})$$

on the set $\{\tau_n + t < \tau_{n+1}\}$, where $\delta$ denotes the Dirac measure. Now we define the trajectories of the controlled process $\tilde{X}$ by

$$\tilde{X}_t((\omega_i)_{i\in\mathbb{N}}) := X^n_t(\omega_n)$$

for all $(\omega_i)_{i\in\mathbb{N}} \in \tilde{\Omega}$, $t \in [0, \infty)$, where $n \in \mathbb{N}$ is chosen such that $t \in [\tau_{n-1}, \tau_n)$ (with $\tau_0 := 0$). Note that this implies

$$\tilde{X}_{\tau_n} = \zeta_n$$

and to take care of the third point from the requirements on the controlled process listed above we set

$$\tilde{X}_{\tau_n,-} := X^n_{\tau_n}$$

to describe the value of the process at time $\tau_n$, but right before the $n$-th control is exercised. When working with general Markov processes in the following we will always tacitly assume this change of underlying probability space and this construction of controlled processes and suppressing the '~' and just write $X$, $\mathbb{P}$, etc. as it is common in the literature. Further, we slightly abuse the notation by writing

$$\mathbb{P}(X^S \in \cdot) := \mathbb{P}^S(\tilde{X} \in \cdot)$$

although of course we do not have absolute continuity. Also, while this model serves its purpose well for most theoretical aspects, when it comes to applications and with that specifications of the underlying process, not to be able to compare paths of the process controlled with different control strategies might raise some issues, as, for example, may be seen in Chapter 6. This amplifies the need for suitable couplings or path-wise constructions of the controlled process. Thus, whenever we work with diffusions or our main case of interest in this treatise, Lévy processes, we work with path-wise construction that are given in the following.

**Model for diffusions** In the case that $X$ is a Itô diffusion as defined in Definition 2.3.2 with diffusion coefficients $\mu$ and $\sigma$ as defined ibd. we may stay on the initial filtered probability space and define a control strategy by a sequence $S = (\tau_n, \zeta_n)_{n\in\mathbb{N}}$ where $(\tau_n)_{n\in\mathbb{N}} \in \mathcal{T}^{\mathbb{N}}$ (where $\mathcal{T}$ is defined in Definition 2.1.5, Part 1.) and each $\zeta_n$ is a $\mathcal{F}_{\tau_n}$-measurable random variable. Then, we define the controlled process $X^S$ by

$$X_t^S = X_0^S + \int_0^t \mu(X^S(s))\, ds + \int_0^t \sigma(X_s^S)\, dW_s - \sum_{n;\ \tau_n \leq t} (X_{\tau_n -}^S - \zeta_n),$$

where $W$ is the same Brownian motion used to construct the uncontrolled process. Further, due to the sample paths being continuous we may set

$$X_{\tau_n,-}^S := X_{\tau_n -}^S$$

as the left limit for all $n \in \mathbb{N}$.

**Model for Lévy processes** If $X$ is a Lévy process as defined in Definition 2.2.1 with $\mathbb{E}(X_1) \in (0, \infty)$ we again may work on the original filtered probability space, again model a control strategy $S = (\tau_n, \zeta_n)_{n\in\mathbb{N}}$ by a sequence of stopping times $(\tau_n)_{n\in\mathbb{N}} \in \mathcal{T}^{\mathbb{N}}$ (where $\mathcal{T}$ is defined in Definition 2.1.5, Part 1.) and each $\zeta_n$ is a $\mathcal{F}_{\tau_n}$-measurable random variable. Then, we set

$$X_t^S := X_t - \sum_{n;\tau_n \leq t} \left(X_{\tau_n,-}^S - \zeta_n\right) \tag{4.1}$$

for each strategy $S = (\tau_n, \zeta_n)_{n\in\mathbb{N}}$. Herein we use

$$X_{\tau_n,-}^S := X_{\tau_n} - \sum_{i=1}^{n-1} \left(X_{\tau_i,-}^S - \zeta_i\right)$$

for the value right before the $n$-th shift (Note that here indeed due to $X$ not being continuous this value may deviate from both $X_{\tau_n}^S$ and $X_{\tau_n -}^S$).

**Remark 4.1.1.** *Note that in full generality the explicit constructions of the model for Lévy processes are not directly couplings for the family of the controlled* *processes constructed by the general model, since in the general model there might be a richer filtration. But this may be addressed by just enlarging the filtrations in the explicit models; as it is commonly known for Markov processes, this does not change the optimal value of the control problem.*

**Value function and admissible strategies**

We define the long-term average value of the process controlled by a strategy $S = (\tau_n, \zeta_n)$ by

$$J_x(S) := \liminf_{T \to \infty} \frac{1}{T} \mathbb{E}_x \left( \sum_{n:\tau_n \le T} \left( \gamma\left(X^S_{\tau_n,-}\right) - \gamma(\zeta_n) - K \right) - \int_0^T h\left(X^S_s\right) ds \right)$$

for all $x \in E$ where $\gamma$ is the so called pay-off function, $K \ge 0$ models fixed costs and $h$ is called running costs. Further, for each $B \subseteq E$ we define $\mathcal{S}_B$ as the set of all control strategies $S = (\tau_n, \zeta_n)_{n \in \mathbb{N}}$ such that $X^S_{\tau_n,-} \ge \zeta_n \in B$ a.s. under all $\mathbb{P}_x$ for all $n \in \mathbb{N}$ and call all elements of $\mathcal{S}_B$ admissible strategies. We always assume that $\mathcal{S}_B \neq \emptyset$. Fix a $B \subseteq E$ throughout the following sections and let

$$v(x) := \sup_{S \in \mathcal{S}_B} J_x(S) \tag{4.2}$$

define the value function for all $x \in E$.
Before defining further necessary objects, let us make one remark on some of the assumptions. The Assumption 3.4.1, 3 is a quite natural one to make. Without this condition, the state space is basically divided in two or more regions such that we cannot let the process go from one region to the another by itself, otherwise we have to pay an infinite amount of costs. So in this case we would basically end up with several disjoint control problems, depending on the starting point. On the other hand, with the later developed tools and notations it will be clear that the Assumption 3.4.1, 3 is not too restrictive and holds in all interesting examples for $h$. Especially in the case of an underlying Lévy process the finiteness of the integral depends mainly on the length and amplitude of excursions from the maximum of $X$ and since these are for Lévy processes not dependent on the starting point, these integrals for most functions $h$ will be finite either for all or for no pairs of points $x, y \in E$ with $x < y$.

**Some important stopping times and strategies**

Since these strategies occur frequently later on, we set $\mathcal{T}_x := \mathcal{T}_x(\gamma, h)$ for all $x \in E$ as defined in Definition 2.1.5, 2. For all $\tau \in \mathcal{T}_x$ we set

$$\tau_1 := \tau,$$

$$\tau_n := \tau \circ \theta_{\tau_{n-1}} + \tau_{n-1} \text{ for all } n > 1$$

and set $R(\tau,x) := (\tau_n, x)_{n\in\mathbb{N}}$. Note that $R(\tau,x) = (\tau_n, x)_{n\in\mathbb{N}}$ is an admissible strategy whenever $\mathbb{E}_x(\tau) > 0$ and $x \in B$.
We call these strategies admissible stationary strategies. Note that this name is not only justified by the stationarity in the controller's action. The controlled process also possesses a stationary distribution.

**Lemma 4.1.2.** *For each admissible stationary strategy $R(\tau,x)$, $\tau \in \mathcal{T}_x$, a limiting distribution for the process $X^{R(\tau,x)}$ (or more precisely for the process $X$ under the measures $\mathbb{P}_y^{R(\tau,x)}$, $y \in E$), denoted by $\Pi^{R(\tau,x)}$, exists and is given by*

$$\int f(x)\Pi^{R(\tau,x)}(dx) = \frac{1}{\mathbb{E}_x\tau}\mathbb{E}_x \int_0^\tau f(X_s)ds.$$

*Proof.* It is immediately seen that $X^R$ is a regenerative processes in the sense of Section 2.4, therefore the result follows by Lemma 2.4.4. □

We call all stationary strategies $R(\tau_y, x)$ given by a threshold time $\tau_y$ for some $y > x$ threshold strategies and shortly write $R(y,x) := R(\tau_y, x)$.

### 4.1.2 Degenerate case

As last precaution, we take a look at the degenerate case that an infinite gain per period is possible.

**Lemma 4.1.3.** *If there is $x \in B$ and $\tau \in \mathcal{T}_x$ such that*

$$\mathbb{E}_x\left[\gamma(X_\tau) - \int_0^\tau h(X_s)\ ds\right] = \infty,$$

*then $v(y) = \infty$ for all $y \in E$.*

*Proof.* We take the strategy $S := (\tau_i, x)_{i\in\mathbb{N}} := R(\tau,x)$, write $\Delta\tau_i := \tau_i - \tau_{i-1}$ for all $i \in \mathbb{N}$ and, making use of Lemma 2.4.2, we first show that $J_x(S) = \infty$. To that end we split $h$ up in positive and negative part. Define for all $i \in \mathbb{N}$

$$R_i := \gamma\left(X^S_{\tau_{i,-}}\right) - \gamma(x) - K.$$

Now $(\Delta\tau_i, R_i)_{i\in\mathbb{N}}$ is a sequence of i.i.d. random variables under $\mathbb{P}_x$, but the $R_i$ violate the integrability requirements of Lemma 2.4.2. To circumvent that issue, we note that since for all $i \in \mathbb{N}$

$$R_i \overset{d}{=} \gamma(X_\tau) - \gamma(x) - K,$$

under $\mathbb{P}_x$ and since $\tau \in \mathcal{T}_x$ for the negative part $R_1^-$ of $R_1$ we have $\mathbb{E}_x\left(R_1^-\right) < \infty$. Hence, for all $a > 0$ the random variable $R_i \wedge a$ is integrable, and $(\Delta\tau_i, R_i \wedge a)_{i\in\mathbb{N}}$ fulfils the requirements of Lemma 2.4.2. This yields

$$\frac{1}{T}\mathbb{E}_x\left(\sum_{i=1}^{N(T)} R_i \wedge a\right) \overset{T\to\infty}{\to} \frac{\mathbb{E}_x(R_1 \wedge a)}{\mathbb{E}_x(\tau)}$$

where $N(t) := \sup\{n \in \mathbb{N}_0 \mid \tau_n \leq t\}$. We tackle the running cost term also with Lemma 2.4.2. To this end write $h^+ := \max\{h, 0\}$ and $h^- := \max\{-h, 0\}$. We define

$$Q_i^+ := \int_{\tau_{i-1}}^{\tau_i} h^+\left(X_s^S\right)\ ds$$

and

$$Q_i^- := \int_{\tau_{i-1}}^{\tau_i} h^-\left(X_s^S\right)\ ds$$

and note that for each $* \in \{+,-\}$ the process $\left(\Delta\tau_i, Q_i^*\right)_{i\in\mathbb{N}}$ also fulfils the requirements of Lemma 2.4.2 because $\tau \in \mathcal{T}_x$ implies that $\int_0^\tau \mid h(X_s) \mid\ ds < \infty$ and hence we get for each $* \in \{+,-\}$

$$\begin{aligned}
&\frac{1}{T}\mathbb{E}_x\left(\int_0^T h^*\left(X_s^S\right) ds\right)\\
&\leq \frac{1}{T}\mathbb{E}_x\left(\int_0^{\tau_{N(T)+1}} h^*\left(X_s^S\right) ds\right)\\
&= \frac{1}{T}\mathbb{E}_x\left(\sum_{i=1}^{N(T)+1} Q_i^*\right)\\
&\overset{T\to\infty}{\to} \frac{\mathbb{E}_x\left(Q_1^*\right)}{\mathbb{E}_x(\tau)}\\
&= \frac{\mathbb{E}_x\left(\int_0^\tau h^*\left(X_s\right)\ ds\right)}{\mathbb{E}_x(\tau)}\\
&=: C^*.
\end{aligned}$$

Now for each $T \geq 0$ and each $a \geq 0$:

$$\begin{aligned}
&\frac{1}{T}\mathbb{E}_x\left(\sum_{n:\tau_n\leq T}\left(\gamma\left(X_{\tau_n,-}^S\right) - \gamma(x) - K\right) - \int_0^T h\left(X_s^S\right) ds\right)\\
&\geq \frac{1}{T}\mathbb{E}_x\left(\sum_{n=1}^{N(T)}\left(R_n \wedge a\right) - \int_0^T h\left(X_s^S\right) ds\right)\\
&\overset{T\to\infty}{\to} \frac{\mathbb{E}_x\left(R_1 \wedge a\right)}{\mathbb{E}_x(\tau)} - C^+ + C^-.
\end{aligned}$$

Now the monotone convergence theorem yields

$$\lim_{a\to\infty} \mathbb{E}_x\left(R_1 \wedge a\right) = \mathbb{E}_x\left(R_1\right) = \infty$$

and we finally get

$$\frac{1}{T}\mathbb{E}_x\left(\sum_{n:\tau_n\leq T}\left(\gamma\left(X^S_{\tau_n,-}\right)-\gamma\left(x\right)-K\right)-\int\limits_0^T h\left(X^S_s\right)ds\right)\overset{T\to\infty}{\to}\infty.$$

It remains to show that also for all $y \in \mathbb{R}$ with $y \neq x$ holds $J_y\left(S\right) = \infty$. This can be easily done by adding a new first control to the strategy constructed above, where we shift the process back to $x$ as soon it exceeds $x$ for the first time. Assumption 3.4.1, 3. ensures that this is still an admissible control strategy and the renewal processes we worked with above then are delayed renewal processes, hence the renewal theoretic results we used still hold, as is worked out, e.g., in [Asm03]. □

## 4.2 Connection to Martingales and Optimal Stopping

This section aims to characterize the control problem in a general manner in order to unveil inherent structures and forge links to neighbouring fields. When looking at the more extensively studied control problems with discounting, in [Chr14] the value function was characterized as the smallest super-harmonic majorant $h$ of the pay-off function that sufficed $h \geq Mh$ for the maximum operator $M$, see Section 1.2 or directly in [Chr14, Proposition 2.2]. Further, therein was a connection made between the value function of the control problem and a connected value function of a stopping problem. The two main results of this section can – in a very broad sense – be understood as the analogue for the long-term average control problems. Because the value function of the long-term average control problem often is constant, no direct connection to superharmonic functions can be made, but, nevertheless, the idea behind that characterization still can be used. So, in Theorem 4.2.1 we use an underlying super-martingale to describe value and optimal strategy by a function occuring in the super-martingale. The second part, which is also one of the main ingredients of the solution technique we will present in Section 4.4.2, connects the control problem to an associated optimal stopping problem, even more directly than in the discounted case. In Theorem 4.2.7 and the corollaries thereafter it is shown that the value of the long-term average control problem can directly be derived from the value of a stopping problem. It is also shown that an optimal stopping time for said stopping problem and an optimal starting point for the problem can be merged to an optimal control strategy.

### 4.2.1 Super-martingale type verification theorem

The main theorem of this subsection, Theorem 4.2.1, establishes sufficient conditions for a strategy to be an optimizer and a real number to be the value of the control problem by super-martingale techniques. While it does not explicitly provide optimal strategies or candidates for the value function, it may serve as a verification theorem.

**Theorem 4.2.1.** *Let $g : E \to \mathbb{R}$ be a measurable function, let $u$ be defined by*

$$u(x,y) = \gamma(x) - \gamma(y) - K - g(x) + g(y)$$

*for all $x, y \in E$ with $y \leq x$, let $\rho \in \mathbb{R}$ and define*

$$M := \left( g(X_t) - \int_0^t \left( h(X_s) + \rho \right) \, ds \right)_{t \geq 0}$$

*and for each $S \in \mathcal{S}_B$ set*

$$M^S := \left( g\left(X_t^S\right) - \int_0^t \left( h\left(X_s^S\right) + \rho \right) \, ds \right)_{t \geq 0}.$$

*(i) Assume*

*(a) $M$ is a super-martingale under $\mathbb{P}_x$ for all $x \in E$,*

*(b)*

$$\limsup_{T \to \infty} \frac{\mathbb{E}_x g\left(X_T^S\right)}{T} \geq 0 \text{ for all } S \in \mathcal{S}_B, \ x \in E,$$

*(c)*

$$u(x,y) \leq 0 \text{ for all } x \in E, y \in B \text{ with } \ y \leq x.$$

*Then*

$$v(x) \leq \rho \text{ for all } x \in E.$$

*(ii) If there is a strategy $S^\uparrow = \left(\tau_n^\uparrow, \zeta_n^\uparrow\right)_{n \in \mathbb{N}} \in \mathcal{S}_B$ such that*

*(a)*

$$\mathbb{E}_x \left( M_{\tau_n^\uparrow \wedge T, -} - M_{\tau_{n-1}^\uparrow \wedge T} \right) \geq 0 \text{ for all } n \in \mathbb{N}, \ x \in E, \ T \geq 0,$$

*(b)*

$$\lim_{T \to \infty} \frac{\mathbb{E}_x g\left(X_T^{S^\uparrow}\right)}{T} \leq 0 \text{ for all } x \in E,$$

*(c)*

$$u\left(X_{\tau_n^\uparrow, -}, \zeta_n^\uparrow\right) \geq 0 \quad \mathbb{P}_x^{S^\uparrow}\text{-a.s. for all } x \in E, \ n \in \mathbb{N}.$$

*Then,*

$$v(x) \geq J_x\left(S^\uparrow\right) \geq \rho, \text{ for all } x \in E.$$

*(iii) If (i) holds and a strategy $S^* = (\tau_n^*, \zeta_n^*)_{n\in\mathbb{N}} \in \mathcal{S}_B$ as in* (ii) *exists, then*

$$v(x) = \rho \quad \textit{for all } x \in E$$

*and $S^*$ is optimal in $\mathcal{S}_B$ in the sense that $v(x) = J_x(S^*)$ for all $x \in E$.*

*Proof.* We first fix $S = (\tau_n, \zeta_n)_{n\in\mathbb{N}} \in \mathcal{S}_B$ and $T > 0$ and assume the conditions (a), (b), (c) in (i) to hold. Since the process $X^S$ uncontrolledly follows the initial dynamic of $X$ on each interval $[\tau_{k-1}, \tau_k)$, the optional sampling theorem yields that $\mathbb{E}_x\left(M^S_{\tau_k\wedge T,-} - M^S_{\tau_{k-1}\wedge T}\right) \leq 0$ for each $k \in \mathbb{N}$, $x \in \mathbb{R}$. Hence,

$$\begin{aligned}
&\mathbb{E}_x\left[\sum_{n\in\mathbb{N}:\tau_n\leq T}\left(\gamma\left(X^S_{\tau_n,-}\right) - \gamma(\zeta_n) - K\right) - \int_0^T h\left(X^S_s\right)\,ds\right] \\
&\leq \mathbb{E}_x\left[\sum_{n\in\mathbb{N}:\tau_n\leq T}\left(\gamma\left(X^S_{\tau_n,-}\right) - \gamma(\zeta_n) - K\right)\right. \\
&\qquad\left. - \sum_{k=1}^{\infty}\left(M^S_{\tau_k\wedge T,-} - M^S_{\tau_{k-1}\wedge T}\right) - \int_0^T h\left(X^S_s\right)\,ds\right] \\
&= \mathbb{E}_x\left[\sum_{n\in\mathbb{N}:\tau_n\leq T}\left(\gamma\left(X^S_{\tau_n,-}\right) - \gamma(\zeta_n) - K\right) - \sum_{k=1}^{\infty}\left(g\left(X^S_{\tau_k\wedge T,-}\right) - g\left(X^S_{\tau_{k-1}\wedge T}\right)\right.\right. \\
&\qquad\left.\left. - \int_{\tau_{k-1}\wedge T}^{\tau_k\wedge T} h\left(X^S_s\right)\,ds - \rho(\tau_k\wedge T) + \rho(\tau_{k-1}\wedge T)\right) - \int_0^T h\left(X^S_s\right)\,ds\right] \\
&= \mathbb{E}_x\left[\sum_{1\leq n:\tau_n\leq T}\left(\gamma\left(X^S_{\tau_n,-}\right) - \gamma(\zeta_n)\right.\right. \\
&\qquad\left.\left. - K - g\left(X^S_{\tau_n,-}\right) + g(\zeta_n)\right) - g\left(X^S_T\right) + g\left(X^S_0\right) + \rho T\right] \\
&= \mathbb{E}_x\left[\sum_{1\leq n:\tau_n\leq T} u\left(X^S_{\tau_n,-}, \zeta_n\right)\right] - \mathbb{E}_x g\left(X^S_T\right) + g(x) + \rho T \\
&\leq -\mathbb{E}_x g\left(X^S_T\right) + g(x) + \rho T.
\end{aligned}$$

Dividing by $T$ and taking the limit $T \to \infty$, we obtain the first assertion.
To prove (ii), we see that similar calculations for a strategy $S^{\uparrow}$ as defined in (ii) yield

$$\begin{aligned}
&\mathbb{E}_x \left[ \sum_{n\in\mathbb{N}:\tau_n^\uparrow \le T} \left(\gamma\left(X^{S^\uparrow}_{\tau_n^\uparrow,-}\right) - \gamma\left(\zeta_n^\uparrow\right) - K\right) - \int_0^T h\left(X_s^{S^\uparrow}\right)\, ds\right] \\
&\ge \mathbb{E}_x \left[ \sum_{n\in\mathbb{N}:\tau_n^\uparrow \le T} \left(\gamma\left(X^{S^\uparrow}_{\tau_n^\uparrow,-}\right) - \gamma\left(\zeta_n^\uparrow\right) - K\right) \right.\\
&\qquad \left. - \sum_{k=1}^{\infty} \left( M^{S^\uparrow}_{\tau_k^\uparrow \wedge T,-} - M^{S^\uparrow}_{\tau_{k-1}^\uparrow \wedge T}\right) - \int_0^T h\left(X_s^{S^\uparrow}\right)\, ds\right] \\
&= \mathbb{E}_x \left[ \sum_{n\in\mathbb{N}:\tau_n^\uparrow \le T} \left(\gamma\left(X^{S^\uparrow}_{\tau_n^\uparrow,-}\right) - \gamma\left(\zeta_n^\uparrow\right) - K\right) \right.\\
&\qquad - \sum_{k=1}^{\infty} \left( g\left(X^{S^\uparrow}_{\tau_k^\uparrow \wedge T,-}\right) - g\left(X^{S^\uparrow}_{\tau_{k-1}^\uparrow \wedge T}\right)\right.\\
&\qquad \left.\left. - \int_{\tau_{k-1}^\uparrow \wedge T}^{\tau_k^\uparrow \wedge T} h\left(X_s^{S^\uparrow}\right)\, ds - \rho\tau_k^\uparrow \wedge T + \rho\tau_{k-1}^\uparrow \wedge T\right) - \int_0^T h\left(X_s^{S^\uparrow}\right)\, ds\right] \\
&= \mathbb{E}_x \left[ \sum_{1\le n:\tau_n^\uparrow \le T} \left(\gamma\left(X^{S^\uparrow}_{\tau_n^\uparrow,-}\right) - \gamma\left(\zeta_n^\uparrow\right)\right.\right.\\
&\qquad \left.\left. - K - g\left(X^{S^\uparrow}_{\tau_n^\uparrow,-}\right) + g\left(\zeta_n^\uparrow\right)\right) - g\left(X_T^{S^\uparrow}\right) + g\left(X_0^{S^\uparrow}\right) + \rho T\right] \\
&= \mathbb{E}_x \left[ \sum_{1\le n:\tau_n^\uparrow \le T} u\left(X^{S^\uparrow}_{\tau_n^\uparrow,-}, \zeta_n^\uparrow\right)\right] - \mathbb{E}_x g\left(X_T^{S^\uparrow}\right) + g\left(x\right) + \rho T \\
&\ge -\mathbb{E}_x g\left(X_T^{S^\uparrow}\right) + g\left(x\right) + \rho T.
\end{aligned}$$

Again $T \to \infty$ yields the claim.
Lastly, (iii) is a direct consequence of (i) and (ii). □

### 4.2.2 Reduction to stopping

In this subsection we characterize the value of the impulse control problem by the value of a stopping problem that in a way resembles the maximal gain with one control. In that process we also show that optimal stopping times for that stopping problem, when repeatedly used, form an optimal control strategy, provided an optimal restarting point exists. For all $x, y \in E$ and all $\rho \in \mathbb{R}$ we set

$$g_\rho^{\mathcal{T}_y}\left(x\right) := \sup_{\tau\in\mathcal{T}_y} \mathbb{E}_x \left(\gamma\left(X_\tau\right) - \int_0^\tau \left(h\left(X_t\right) + \rho\right)\, dt\right)$$

and

$$\mathfrak{g}_{\rho}^{\mathcal{T}_y}(x) := \sup_{\tau\in\mathcal{T}_y} \mathbb{E}_x\left(\gamma(X_\tau) - \gamma(x) - K - \int_0^\tau (h(X_t) + \rho)\ dt\right).$$

Recall that herein $\mathcal{T}_y$ was defined as the set of all stopping times $\tau$ with $\mathbb{E}_y(\tau) < \infty$ as well as $\mathbb{E}_y\left(\int_0^\tau | h(X_s) | ds\right) < \infty$ and $X_\tau \geq y$ under $\mathbb{P}_y$ for all $y \in E$, see Definition 2.1.5.

**Remark 4.2.2.** *Looking at the definition of $\mathcal{T}$ and $\mathcal{T}_y$, we see that for all $x, y \in E$ the expressions $g_{\rho}^{\mathcal{T}_y}(x)$ and $\mathfrak{g}_{\rho}^{\mathcal{T}_y}(x)$ are well defined and further $-K \leq \mathfrak{g}_{\rho}^{\mathcal{T}_x}(x)$ since immediate stopping is allowed is the case $y = x$.*

**Definition 4.2.3.** *We define*

$$\mathfrak{G} : E \to [-K, \infty];\ \rho \mapsto \sup_{x\in B} \mathfrak{g}_{\rho}^{\mathcal{T}_x}(x).$$

**Lemma 4.2.4.** *$\mathfrak{G}$ is decreasing, on $\mathfrak{G}^{-1}((-K,\infty))$ even strictly decreasing, convex and continuous on $\mathbb{R}\setminus\{\beta\}$ where $\beta := \sup\{\rho \in \mathbb{R} \mid \mathfrak{G}(\rho) = \infty\}$ (with the convention $\inf\emptyset = \infty = -\sup\emptyset$).*

*Proof.* The monotonicity is clear, $\mathfrak{G}$ is convex as supremum over affine functions, hence also continuous on $\mathbb{R}\setminus\{\beta\}$. □

**Lemma 4.2.5.** *There is a $\rho^+ \in \mathbb{R}$ such that $\mathfrak{G}(\rho^+) > 0$.*

*Proof.* Here Assumption 3.4.1 comes into play. We take $x \in B$ and $y \in E$ with $x < y$ and $\mathbb{E}_x(\tau_y) \neq 0$. Note that this is possible since otherwise the only element in $B$ would be the supremum of $E$ and hence $\mathcal{S}_B = \emptyset$ which by assumption must not hold. Assumption 3.4.1 ensures $\mathbb{E}_x(\tau_y) < \infty$ and $\mathbb{E}_x(\int_0^{\tau_y} h(X_s)ds) < \infty$. Hence, if we set

$$\rho^+ := \frac{\mathbb{E}_x(\int_0^{\tau_y} h(X_t)\,dt) + K + 1}{\mathbb{E}_x(\tau_y)},$$

then we have, since $\gamma$ is non-decreasing,

$$\begin{aligned}
&\mathbb{E}_x\left(\gamma\left(X_{\tau_y}\right) - \int_0^{\tau_y} \left(h(X_t) + \rho^+\right)\ dt\right) - \gamma(x) - K \\
\geq& -\mathbb{E}_x\left(\int_0^{\tau_y} h(X_t)\ dt\right) - K + \rho^+\mathbb{E}_x(\tau_y) \\
>& \ 0
\end{aligned}$$

This yields $\mathfrak{g}_{\rho^+}^{\mathcal{T}_x}(x) > 0$ and hence also $\mathfrak{G}(\rho^+) = \sup_{x\in B} g_{\rho^+}^{\mathcal{T}_x}(x) > 0$. □

**Definition 4.2.6.** *We define*

$$\rho^* := \sup\{\rho \in \mathbb{R} \mid \mathfrak{G}(\rho) > 0\}.$$

Note that due to the monotonicity

$$\rho^* = \inf\{\rho \in \mathbb{R} \mid \mathfrak{G}(\rho) \leq 0\}$$

and if $\rho^* \neq \beta$, $\rho^*$ is the only root of $\mathfrak{G}$ (In the case $\rho^* = \beta$ the function $\mathfrak{G}$ may jump from infinity to a negative value). Now, in the following, we show that $v(y) = \rho^*$ for all $y \in E$.

**Theorem 4.2.7.** *For all $\rho \in \mathbb{R}$ with $\mathfrak{G}(\rho) \in \mathbb{R}$ holds*

$$\mathfrak{G}(\rho) > 0 \Leftrightarrow \forall x \in E : v(x) > \rho.$$

*Proof.* If there is an $x \in B$ and a stopping time $\tau \in \mathcal{T}_x$, such that $\mathbb{E}_x(\gamma(X_\tau)) = \infty$. Then, Lemma 4.1.3 yields the equivalence, therefore, in the following, we assume that no such stopping time exists. First, let $\rho \in \mathbb{R}$ such that for all $x \in E$ holds $v(x) > \rho$. Then, there is an admissible strategy $S = (\tau_n, \zeta_n) \in \mathcal{S}_B$ such that

$$\liminf_{T\to\infty} \frac{1}{T}\mathbb{E}_x \left\{ \sum_{1\leq n:\tau_n \leq T} \left(\gamma\left(X^S_{\tau_n,-}\right) - \gamma(\zeta_n) - K\right) - \int_0^T h\left(X^S_s\right)\,ds \right\} > \rho$$

and due to excluding strategies with an infinite gain in one period or infinite costs in one period, we also have

$$\liminf_{T\to\infty} \frac{1}{T}\mathbb{E}_x \left\{ \sum_{1\leq n:\tau_{n-1} \leq T} \left(\gamma\left(X^S_{\tau_n,-}\right) - \gamma(\zeta_{n-1}) - K - \int_{\tau_{n-1}}^{\tau_n} h\left(X^S_s\right)\,ds\right) \right\} > \rho,$$

because we sum over one more control and hence add one summand with finite expectation. For notational convenience set $\tau_0 := 0$ and $\zeta_0 := x$.
The equation above implies that there is a $\tilde{T} > 0$ such that for all $T \geq \tilde{T}$ we have

$$\frac{1}{T}\mathbb{E}_x \left\{ \sum_{1\leq n:\tau_{n-1} \leq T} \left(\gamma\left(X_{\tau_n,-}\right) - \gamma(\zeta_{n-1}) - K - \int_{\tau_{n-1}}^{\tau_n} h\left(X_s\right)\,ds\right) \right\} > \rho$$

and hence

$$\frac{1}{T}\mathbb{E}_x \left\{ \sum_{1\leq n:\tau_{n-1} \leq T} \left(\gamma\left(X_{\tau_n,-}\right) - \gamma(\zeta_{n-1}) - K - \int_{\tau_{n-1}}^{\tau_n} \left(h\left(X_s\right) + \rho\right)\,ds\right) \right\} > 0.$$

For all $n \in \mathbb{N}$ set $\tilde{\tau}^n_i = \begin{cases} \tau_i; & i \leq n \\ \infty; & i > n \end{cases}$ and $\tilde{S}_n := (\tilde{\tau}^n_i, \zeta_i)_{i\in\mathbb{N}}$. Although this is not an admissible impulse control strategy, we still use the established notations for these strategy.
Fix an $n \in \mathbb{N}$. The process $Y$ given by $Y_t := X^{\tilde{S}_{n-1}}_{\tau_{n-1}+t}$ is still a Markov process

(started in $X_{\tau_{n-1}}^{\tilde{S}_{n-1}}$) both under its natural filtration $\mathcal{F}^Y$ and under the filtration $\tilde{\mathcal{F}}$ given by $\tilde{\mathcal{F}}_t := \mathcal{F}_{\tau_{n-1}+t}$ for all $t \in [0,\infty)$. It is well established that in optimal stopping problems for Markov processes the value of the problem does not change, when one only considers optimization over first entry times, which are in the natural filtration of the process. We define

$$\mathcal{S}_x(\mathcal{G}) := \{\tau \mid \tau \text{ is } \mathcal{G} \text{ st. time}, Y_\tau \geq x, \mathbb{E}_x\left(\int_0^\tau (\mid h(Y_t) \mid + \rho)\ dt\right),\ \mathbb{E}_x(\tau) < \infty\}$$

for each $x \in E$, and each $\mathcal{G} \in \{\tilde{\mathcal{F}}, \mathcal{F}^Y\}$. We set

$$\sigma := \mathbb{1}_{\{\tau_{n-1} \leq T\}}(\tau_n - \tau_{n-1}) \in \mathcal{S}_x\left(\tilde{\mathcal{F}}\right)$$

and we have due to the aforementioned reason

$$\begin{aligned}
&\mathbb{E}_x\left\{\mathbb{1}_{\{\tau_{n-1}\leq T\}}\left(\gamma(X_{\tau_n,-}) - \gamma(\zeta_{n-1}) - K - \int_{\tau_{n-1}}^{\tau_n} h(X_s) + \rho\ ds\right)\right\}\\
&\quad= \mathbb{E}_x\left\{\mathbb{E}_x\left[\mathbb{1}_{\{\tau_{n-1}\leq T\}}\right.\right.\\
&\qquad\left.\left.\left(\gamma(X_{\tau_n,-}) - \gamma(\zeta_{n-1}) - K - \int_{\tau_{n-1}}^{\tau_n} h(X_s) + \rho\ ds\right)\middle|\mathcal{F}_{\tau_{n-1}}\right]\right\}\\
&\quad= \mathbb{E}_x\left\{\mathbb{1}_{\{\tau_{n-1}\leq T\}}\right.\\
&\quad\left.\mathbb{E}_x\left[\left(\gamma\left(Y_{(\tau_n,-)-\tau_{n-1}}\right) - \gamma(Y_0) - K - \int_0^{(\tau_n,-)-\tau_{n-1}} h(Y_s) + \rho\ ds\right)\middle|\mathcal{F}_{\tau_{n-1}}\right]\right\}\\
&\quad= \mathbb{E}_x\left\{\mathbb{1}_{\{\tau_{n-1}\leq T\}}\mathbb{E}_{X_{\tau_{n-1}}}\left(\gamma(Y_\sigma) - \gamma(Y_0) - K - \int_0^\sigma h(Y_s) + \rho\ ds\right)\right\}\\
&\quad\leq \mathbb{E}_x\left\{\mathbb{1}_{\{\tau_{n-1}\leq T\}}\right.\\
&\qquad\left.\sup_{\tau\in S_{X_{\tau_{n-1}}}(\tilde{\mathcal{F}})}\mathbb{E}_{X_{\tau_{n-1}}}\left(\gamma(Y_\tau) - \gamma(Y_0) - K - \int_0^\tau (h(Y_t) + \rho)\ dt\right)\right\}\\
&\quad= \mathbb{E}_x\left\{\mathbb{1}_{\{\tau_{n-1}\leq T\}}\right.\\
&\qquad\left.\sup_{\tau\in \mathcal{S}_{X_{\tau_{n-1}}}(\mathcal{F}^Y)}\mathbb{E}_{X_{\tau_{n-1}}}\left(\gamma(Y_\tau) - \gamma(Y_0) - K - \int_0^\tau (h(Y_t) + \rho)\ dt\right)\right\}\\
&\quad\leq \mathbb{E}_x\left(\mathbb{1}_{\{\tau_{n-1}\leq T\}}\mathfrak{G}(\rho)\right)\\
&\quad= \mathbb{P}_x(\tau_{n-1} \leq T)\,\mathfrak{G}(\rho),
\end{aligned}$$

hence for all $T \geq \tilde{T}$

$$\begin{aligned} 0 <& \frac{1}{T}\mathbb{E}_x \left\{ \sum_{1 \leq n : \tau_{n-1} \leq T} \left( \gamma\left(X_{\tau_n,-}\right) - \gamma\left(\zeta_{n-1}\right) - K - \int_{\tau_{n-1}}^{\tau_n} h\left(X_s\right) + \rho \; ds \right) \right\} \\ =& \frac{1}{T} \sum_{n \in \mathbb{N}} \mathbb{P}_x\left(\tau_{n-1} < T\right) \mathfrak{G}\left(\rho\right) \end{aligned}$$

and we get

$$0 < \mathfrak{G}\left(\rho\right).$$

Now we show the reverse inequality.
Let $\rho \in \mathbb{R}$ such that $\mathfrak{G}\left(\rho\right) > 0$. Then, there is an $y \in B$ and a $\tau \in \mathcal{T}_y$ with

$$\mathbb{E}_y \left( \gamma\left(X_\tau\right) - \gamma\left(y\right) - K - \int_0^\tau \left(h\left(X_s\right) + \rho\right) \; ds \right) > 0. \tag{4.3}$$

We set

$$S^{\uparrow} := \left(\tau_n, y\right)_{n \in \mathbb{N}} := R\left(\tau, y\right),$$

and we define

$$R_i := \gamma\left(X_{\tau_i}\right) - \gamma\left(y\right) - K - \int_{\tau_{i-1}}^{\tau_i} \left(h\left(X_s\right)\right) \; ds.$$

Then,

$$\mathbb{E}_y\left(R_i\right) > \rho \mathbb{E}_y\left(\tau\right)$$

and Lemma 2.4.2 yields

$$\begin{aligned} \liminf_{T \to \infty} \frac{1}{T}\mathbb{E}_x &\left\{ \sum_{1 \leq n : \tau_n \leq T} \left( \gamma\left(X^{S^{\uparrow}}_{\tau_n,-}\right) - \gamma\left(\zeta_n\right) - K \right) - \int_0^T h\left(X^{S^{\uparrow}}_s\right) \; ds \right\} \\ &= \lim_{T \to \infty} \frac{\mathbb{E}_y\left(\sum_{\tau_n \leq T} R_i\right)}{T} \\ &= \frac{\mathbb{E}_y\left(R_1\right)}{\mathbb{E}_y\left(\tau\right)} \\ &> \rho. \end{aligned}$$

□

**Corollary 4.2.8.** *It holds $v(y) = \rho^*$ for all $y \in E$.*

*Proof.* By the definition of $\rho^*$ in Definition 4.2.6 we have that $\rho^*$ is the smallest upper bound of the set $L := \{\rho \in \mathbb{R} \mid \mathfrak{G}(\rho) > 0\}$. Theorem 4.2.7 yields that for all $x \in E$ the value $v(x)$ also is an upper bound of $L$. This yields for all $x \in E$

$$v(x) \geq \rho^*.$$

On the other hand the characterization of $\rho^*$ right after Definition 4.2.6 yields that $\rho^*$ is the largest lower bound of $\mathbb{R} \setminus L$. Again Theorem 4.2.7 yields that for all $x \in E$ the value $v(x)$ is a lower bound of $\mathbb{R} \setminus L$. This yields for all $x \in E$

$$\rho^* \geq v(x).$$

□

**Corollary 4.2.9.** *If $\mathfrak{G}(\rho^*) = 0$ and there are $y \in B$ and $\tau \in \mathcal{T}_y$ such that*

$$\mathfrak{G}(\rho^*) = \mathbb{E}_y \left( \gamma(X_\tau) - \int_0^\tau (h(X_t) + \rho^*) \, dt \right) - \gamma(y) - K,$$

*then the strategy $R(\tau, y)$ is optimal for $v$. Further,*

$$v = \frac{\mathbb{E}_y \left( \gamma(X_\tau) - \int_0^\tau h(X_t) \, dt \right) - \gamma(y) - K}{\mathbb{E}_y(\tau)} \tag{4.4}$$

*Proof.* For each $\rho \in \mathbb{R}$ with $\rho < \rho^*$ holds $\mathfrak{G}(\rho) > 0$ and $y$ and $\tau$ fulfil (4.3) in the previous proof of Theorem 4.2.7. Hence, the calculations therein show that $J_y(R(\tau, y)) > \rho$ for all $\rho < \rho^*$, This is why $J_y(R(\tau, y)) \geq \rho^* = v$, which means that $R(\tau, y)$ is an optimizer for $v$. □

## 4.3 Main Result

After we have characterized the impulse control problem by use of a stopping problem, the purpose of this section is twofold. First, we will unravel the dense and technical structure of the previous section and summarize the results of that section into clear, relatively short statements. Further, we connect these results with our findings on the tailor-made stopping problem from Section 3.4. This connection will unveil a nice existence result of optimal threshold strategies, given a maximum representation as in Assumption 3.4.2 exists. But before we are able to state, what is the main theorem of this chapter, maybe even the main theoretical result of this whole thesis, we have to make a short observation on possible optimal restarting points.

**Lemma 4.3.1.** *Assume $B = E$, a maximum representation in terms of a function $f$ as in Assumption 3.4.2 exists and let $\rho \in \mathbb{R}$ such that Assumption 3.4.3 holds. Then we have $\mathfrak{G}(\rho) = \sup_{x\in[\underline{x},\overline{x}]} \mathfrak{g}^{\mathcal{T}_x}(x)$.*

*Proof.* This is a direct consequence of Theorem 3.4.9, Part 2. □

Now we have gathered everything to state and prove the main theorem.

**Theorem 4.3.2.** *Let Assumption 3.4.1 hold and assume that $K > 0$.*

1. *The value $v$ defined in (4.2) is constant and given by*

$$v = \sup_{y\in B} \sup_{\tau\in\mathcal{T}_y} \frac{\mathbb{E}_y\left(\gamma\left(X_\tau\right) - \int_0^\tau h\left(X_t\right)\ dt\right) - \gamma\left(y\right) - K}{\mathbb{E}_y(\tau)}. \tag{4.5}$$

2. *If a pair of maximizers $y \in B$, $\tau \in \mathcal{T}_y$ for the term in (4.5) exists, then the strategy $R(\tau, y)$ is optimal for the impulse control problem.*

3. *If $B = E$ and a maximum representation $f$ as in Assumption 3.4.2 exists and fulfils Assumption 3.4.3 and furthermore Assumption 3.4.8 holds, then an optimal threshold strategy $R(\tau_{\overline{x}}, x^*)$ exists, where $\overline{x}$ is given as the larger of the two roots of $f - v$, $x^*$ lies between the two roots of $f - v$. Furthermore $x^*$ equals the smaller root, if $X$ has no upward jumps.*

4. *If $B = E$ and a maximum representation $f$ as in Assumption 3.4.2 exists and fulfils Assumption 3.4.3, then*

$$v = \sup_{y,z\in E; y<z} \frac{\mathbb{E}_y\left(\gamma\left(X_{\tau_z}\right) - \int_0^{\tau_z} h\left(X_t\right)\ dt\right) - \gamma\left(y\right) - K}{\mathbb{E}_y(\tau_z)}. \tag{4.6}$$

*Proof.* Corollary 4.2.8 yields that the value is characterized as the, due to Lemma 4.2.4 unique, value $\rho^*$ where the function

$$\mathfrak{G} : \mathbb{R} \to [-K, \infty];\ \rho \mapsto \sup_{x\in B} \mathfrak{g}_\rho^{\mathcal{T}_x}\left(x\right),$$

defined in Definition 4.2.3, changes sign and moves from a positive to a negative value. Now we split the proof into two cases.
Case 1: $\rho^*$ is the root of $\mathfrak{G}$.
Then, for each $x \in B$, $\tau \in \mathcal{T}_x$ we have

$$0 \geq \mathbb{E}_x \left( \gamma(X_\tau) - \int_0^\tau (h(X_t) + \rho^*) \ dt \right) - \gamma(x) - K \tag{4.7}$$

and hence, if not $\tau = 0$ a.s. under $\mathbb{P}_x$, then

$$\rho^* \geq \frac{\mathbb{E}_x \left( \gamma(X_\tau) - \int_0^\tau h(X_t) \ dt \right) - \gamma(x) - K}{\mathbb{E}_x(\tau)}.$$

On the other hand, with appropriate pairs $x, \tau$ one can get arbitrary close to 0 in (4.7) due to $\mathfrak{G}$'s definition as supremum. Further, these appropriate pairs can without loss of generality be chosen such that not $\tau = 0$ a.s. under $\mathbb{P}_x$ since such a $\tau$ would regardless of $x$ yield a value of $-K$. Hence, if a root $\rho^*$ of $\mathfrak{G}$ exists, then Part 1 of the claim holds.
Case 2: $\rho^*$ is not the root of $\mathfrak{G}$.
If no such root of $\mathfrak{G}$ exists, the only case left to analyze due to $\mathfrak{G}$ being convex, decreasing (see Lemma 4.2.4) and not constant (see Lemma 4.2.5) is that the function $\mathfrak{G}$ jumps from $\infty$ to a negative value. Then, with the same argument as above we still see that

$$\rho^* \geq \frac{\mathbb{E}_x \left( \gamma(X_\tau) - \int_0^\tau h(X_t) \ dt \right) - \gamma(x) - K}{\mathbb{E}_x(\tau)}$$

for all $x \in B$, $\tau \in \mathcal{T}_x$. Additionally, for all $\epsilon > 0$, there is an $x \in B$ and a $\tau \in \mathcal{T}_x$ such that

$$0 < \mathbb{E}_x \left( \gamma(X_\tau) - \int_0^\tau (h(X_t) + \rho^* - \epsilon) \ dt \right) - \gamma(x) - K,$$

hence,

$$\rho^* < \frac{\mathbb{E}_x \left( \gamma(X_\tau) - \int_0^\tau h(X_t) \ dt \right) - \gamma(x) - K}{\mathbb{E}_x(\tau)} + \epsilon$$

which also yields the claim.
Now Corollary 4.2.9 directly yields Part 2 of the claim. Regarding part 3, now let Assumption 3.4.2 hold true and let Assumption 3.4.3 hold true for $\rho^*$. Fix a function $f$ as in Assumption 3.4.3 and denote the two solutions of $f(x) = \rho^*$ with $\underline{x}$ and $\overline{x}$ with the usual convention $\underline{x} < \overline{x}$. Theorem 3.4.9 yields that for each $x \in E$ with $x \geq \underline{x}$ the threshold time $\tau_{\overline{x}}$ is optimal for $\mathfrak{g}^{\mathcal{T}_x}(x)$. Additionally, Lemma 4.3.1 yields that there is an $x^* \in [\underline{x}, \overline{x}]$ such that

$$\mathfrak{G}(\rho^*) = \mathfrak{g}^{\mathcal{T}_{x^*}}(x^*) = \mathbb{E}_{x^*} \left( \gamma\left(X_{\tau_{\overline{x}}}\right) - \int_0^{\tau_{\overline{x}}} (h(X_t) + \rho^*) \ dt \right) - \gamma(x^*) - K.$$

Now with Part 2 of this theorem here and Corollary 4.2.9 follows that $\mathbb{R}(\tau_{\overline{x}}, x^*)$ is optimal for the impulse control problem. Lastly Part 4 directly follows from Part 3 and 1. □

### 4.3.1 Discussion of Assumption 3.4.3 and further results if a maximum representation exists

While Assumption 3.4.2 certainly is essential in order for Theorem 4.3.2, Part 3 to hold, since it was one of the main tools to solve the tailor-made stopping problem, the question is if Assumption 3.4.3 is also essential and furthermore, if Assumption 3.4.3 is a restrictive one. The short answer to both question is no. In the following, we will justify that answer and pick up some helpful results on the way. Throughout this subsection we assume a maximum representation as in Assumption 3.4.2 exists and fix a function $f$ as in Assumption 3.4.2. We will only assume $K \geq 0$ and thus in contrast to Theorem 4.3.2 include the case $K = 0$. Note that this is a slightly weaker assumption as in Theorem 4.3.2. Now step-by-step we will look at all of the cases that may occur, when Assumption 3.4.3 fails, and hence the function $f_{\rho^*} = f - \rho^*$ has less than two roots. Now in Assumption 3.4.2 $f$ is assumed to have one global maximum at a value $a \in E$ and is strictly increasing, right of $a$ and strictly decreasing left of $a$.
Hence the first reason for $f_{\rho^*}$ to violate Assumption 3.4.3 could be that it is discontinuous and jumps from a positive to a negative value. But then the place of the jump could be used to replace the 'missing' root and all of the proofs would work just the same, since we only used that $f_{\rho^*}$ is positive between the roots and negative elsewhere. When the right of the two roots is replaced by a value $\overline{x}'_{\rho^*}$ with $f_{\rho^*}(\overline{x}'_{\rho^*}) \neq 0$ one has to distinguish the cases $f_{\rho^*}(\overline{x}'_{\rho^*}) > 0$ and $f_{\rho^*}(\overline{x}'_{\rho^*}) < 0$ and as in the discrete time stopping problem in Section 3.1, depending on the case, instead of $\tau_{\overline{x}'_{\rho^*}}$ also $\mathring{\tau}_{\overline{x}'_{\rho^*}}$ could be an optimizer for the tailor-made stopping problem from Section 3.4 and hence a strategy $R(\mathring{\tau}_{\overline{x}'_{\rho^*}}, x^*)$ for some $x^* \in B$ could be optimal for the long-term average control problem.
The second reason could be that $\rho^*$ lies strictly beneath all values $f(x)$, $x \in E$ and is bounded away from the graph of $f$. Under minimal assumptions on expected hitting times, we are able to show that this is not possible.

**Lemma 4.3.3.** *Assume that the sets* $\{\mathbb{E}_x(\tau_y) \mid (x, y) \in E^2, x < y \leq a\}$ *and* $\{\mathbb{E}_x(\tau_y) \mid (x, y) \in E^2, a \leq x < y\}$ *are not bounded from above. Then*

$$\rho^* \geq \max\{\inf\{f(x)|x \in E \cap (-\infty, a)\}, \inf\{f(x)|x \in E \cap (a, \infty)\}\}.$$

*Proof.* Assume

$$\rho^* < \max\{\inf\{f(x)|x \in E \cap (-\infty, a)\}, \inf\{f(x)|x \in E \cap (a, \infty)\}\}.$$

Due to all the cases needing similar arguments, we will only discuss the case

$$\alpha := \inf\{f(x)|x \in E \cap (-\infty, a)\} > \rho^* > \inf\{f(x)|x \in E \cap (a, \infty)\}.$$

Define $\epsilon := \frac{\alpha - \rho^*}{2}$. Now we use the function $\mathfrak{G}$ defined in Definition 4.2.3 and that

$$\rho^* = \inf\{\rho \mid \mathfrak{G}(\rho) \leq 0\} = \sup\{\rho \mid \mathfrak{G}(\rho) > 0\} \tag{4.8}$$

as Definition 4.2.6 and the discussion right after that definition yields. Let $y \in E \cap (a, \infty)$ be the value where $f_{\rho^*+\epsilon}$ changes sign. Now take a $x \in E$ such that $\mathbb{E}_x(\tau_y) > \frac{K}{\epsilon}$ (due to our boundedness assumptions this is possible). Then, we have utilizing the maximum representation and due to the fact that our choice of $\epsilon$ and $y$ yields $f(z) - (\rho^* + \epsilon) \geq \epsilon$ for all $z \in [x, y]$

$$\begin{aligned}
\mathfrak{G}(\rho^* + \epsilon) &\geq \mathbb{E}_x \left( \gamma \left( X_{\tau_y} \right) - \gamma(x) - K - \int_0^{\tau_y} \left( h\left( X_t \right) + \rho^* + \epsilon \right) \, dt \right) \\
&= \mathbb{E}_x \left( \int_0^{\tau_y} \left( f(\overline{X}_s) - (\rho^* + \epsilon) \right) ds \right) - K \\
&\geq \mathbb{E}_x \left( \int_0^{\tau_y} \epsilon \, ds \right) - K \\
&= \epsilon \cdot \mathbb{E}_x \left( \tau_y \right) - K \\
&> 0.
\end{aligned}$$

But this is a contradiction to (4.8).

□

The third reason why the assumption could be violated is that $\rho^*$ strictly could lie above $f$, but that case can be shown to not occur even without additional assumptions. That $\rho^*$ is also bounded from above by the maximum of the function $f$, is needed later on, hence we formulate this result as a lemma.

**Lemma 4.3.4.** *Let $a \in E$ be the due to Assumption 3.4.2 unique maximum of $f$. Then $v \leq f(a)$.*

*Proof.* That everywhere stopping immediately is optimal when the function in the maximum representation is negative, was discussed in Remark 3.2.7. Transferred to our situation this yields that if we would have $\rho^* > f(a)$, for $\epsilon := \frac{\rho^* - f(a)}{2} > 0$ this would yield $\mathfrak{G}(\rho^* - \epsilon) = -K \leq 0$, but this is again a contradiction to (4.8). □

Since for Assumption 3.4.3 to hold we need $v < f(a)$ instead of $v \leq f(a)$, a refinement of the previous result is needed, that requires the additional assumption that we have non-zero fixed costs.

**Lemma 4.3.5.** *Assume $K > 0$. Assume that there is an open interval $I \subseteq E$ with $a \in I$ such that the set $\{\mathbb{E}_x(\tau_y) \mid (x, y) \in I^2, x < y\}$ is bounded from above. Then $\rho^* < f(a)$.*

*Proof.* Again we utilize the function $\mathfrak{G}$ and remind that in Lemma 4.2.4 we have shown that $\mathfrak{G}$ is strictly decreasing on $\mathfrak{G}^{-1}(-K, \infty)$. Now for all $\delta \in (0, \infty)$ that are small enough the discussion above yields that

$$\mathfrak{G}(f(a) - \delta) \leq (f(a) + \delta - f(a)) \sup_{x, y \in I, x < y} \mathbb{E}_x(\tau_y) - K$$

and this term converges to $-K$ if $\delta \to 0$. □

To summarize our results, there are two reasons, why Assumption 3.4.3 may fail: First, $f$ may jump over $\rho^*$, but then one can replace the needed root of $f_{\rho^*}$ with the point of the jump. The second reason could be, that the assumption of Lemma 4.3.3 fails. Heuristically this may be interpreted as '$X$ gets too large too fast'. This could be circumvented by including the boundaries of $E$ to the domain of $f$ (if they are not included already) and setting $f$ to $-\infty$ at these boundaries. But we will not dive into the technical details required for that here, since our main case of interest is a Lévy process on $\mathbb{R}$ and such a process fulfils the Assumption of Lemma 4.3.3 anyway.
Our last aim is to show that even if $K = 0$ the formula (4.6) still holds.

**Theorem 4.3.6.** *Assume a maximum representation as in Assumption 3.4.2 exists, fix a function $f$ as in Assumption 3.4.2 and denote its unique maximum with $a$. Assume $f$ is continuous in $a$. Further, we assume $K \geq 0$. Then holds*

$$v = \sup_{y,z\in E; y<z} \frac{\mathbb{E}_y\left(\gamma\left(X_{\tau_z}\right) - \int_0^{\tau_z} h\left(X_t\right)\ dt\right) - \gamma\left(y\right) - K}{\mathbb{E}_y(\tau_z)}.$$

*Proof.* Since $v \leq f(a)$ due to Lemma 4.3.4 there is $\epsilon > 0$ small enough such that Assumption 3.4.3 holds for $\rho^* - \epsilon$ (with the roots postulated to exists in Assumption 3.4.3 if needed replaced with jumping points of as discussed before). Let $\underline{x}_\epsilon, \overline{x}_\epsilon$ be the roots of $f_{\rho^*-\epsilon}$. Then, Theorem 3.4.9 yields that there is $x^*_\epsilon$ such that

$$0 < \frac{1}{2}\mathfrak{G}(\rho^* - \epsilon) \leq \mathbb{E}_{x^*_\epsilon}\left(\gamma\left(X_{\tau_{\overline{x}_\epsilon}}\right) - \gamma\left(x^*_\epsilon\right) - K - \int_0^{\tau_{\overline{x}_\epsilon}} \left(h\left(X_t\right) + \rho^* - \epsilon\right)\ dt\right)$$

and therefore

$$\rho^* \leq \frac{\mathbb{E}_{x^*_\epsilon}\left(\gamma\left(X_{\tau_{\overline{x}_\epsilon}}\right) - \gamma\left(x^*_\epsilon\right) - K - \int_0^{\tau_{\overline{x}_\epsilon}} \left(h\left(X_t\right)\right)\ dt\right)}{\mathbb{E}_{x^*_\epsilon}(\tau_{\overline{x}_\epsilon})} + \epsilon$$

□

Especially the case that $K = 0$ will be further investigated in Section 5.3. For example there will be shown that for $K = 0$ we get $\epsilon$ optimal strategies with boundaries that are arbitrary close to $a$.

## 4.4 Specification for Lévy Processes

After we have proven the theoretical main results of this chapter, we now concentrate our attention on how to explicitly find optimal threshold strategies for control problems in case the underlying process is a Lévy process. In contrast to results for diffusions (see [HSZ18] and [HSZ17]) especially regarding explicit calculation procedures for such strategies little is known if the underlying process has jumps. One of the few present works on this topic is [Yam17], where in the discounted case for an underlying spectrally one-sided Lévy process the relatively new theory of scale functions for spectrally one-sided Lévy processes is utilized. Hence, our aim here is to connect our results on maximum representation for Lévy processes from Section 2.6 and our theoretical findings on solving impulse control problems from Sections 4.2 and 4.3 in order to develop a step-by-step solution technique for impulse control problems for underlying Lévy processes. By applying our technique to interesting special cases, amongst them the long-term average analogues to a variety discounted problems discussed in [ØS05], we will emphasize the versatility of our technique. Hence, now we assume $X$ to be a Lévy process on $E := \mathbb{R}$ as defined in Definition 2.2.1 and, furthermore, write $\mathbb{P} := \mathbb{P}_0$ and $\mathbb{E} := \mathbb{E}_0$. Note that as discussed in Subsection 4.1.1 in this situation it is more convenient to work with an explicit construction of the controlled process, namely we model the controlled process as in the explicit model for Lévy processes described in Subsection 4.1.1 recursively by

$$X_t^S := X_t - \sum_{n;\tau_n \le t} \left( X^S_{\tau_n,-} - \zeta_n \right) \tag{4.9}$$

for each strategy $S = (\tau_n, \zeta_n)_{n\in\mathbb{N}}$. Herein we use

$$X^S_{\tau_n,-} := X_{\tau_n} - \sum_{i=1}^{n-1} \left( X^S_{\tau_i,-} - \zeta_i \right)$$

for the value right before the $n$-th shift (Note that still due to $X$ not being continuous this value may deviate from both $X^S_{\tau_n}$ and $X^S_{\tau_n-}$). Assume:

**Assumption 4.4.1.** *1.* $0 < \mathbb{E}(X_1) < \infty$.

*2.* $\gamma$ *is non-decreasing and differentiable.*

*3.* $h$ *is continuous and for all* $x, y \in \mathbb{R}$ *with* $x < y$*, we have*

$$\mathbb{E}_x \left( \int_0^{\tau_y} \mid h(X_s) \mid \ ds \right) < \infty.$$

### 4.4.1 The optimal restarting point

So far, in Corollary 4.2.9 we characterized the (random) optimal times to exercise controls by optimal stopping times of an associated stopping problem. Assuming a supremum representation of the pay-off function $\gamma$ that involves a sufficiently favourably shaped function $f$, the characterization boils down to exercise a control, whenever the process exceeds the rightmost root of the function $f_{\rho^*}$ defined in and after Assumption 3.4.2. The optimal restarting point so far only is characterized as an optimizer of

$$\sup_{y\in B} \mathfrak{g}_{\rho^*}^{\mathcal{T}_y}(y)$$

and Lemma 4.3.1 tells us if $B = E$ we may take the supremum only over all $y$ lying between the roots of $f - \rho^*$.
While in full generality we do not see much hope to characterize the optimal restarting point any further, we now assume $B = \mathbb{R}$ since this is the most interesting case for the applications later. In that case we proceed to show that provided the ladder height process $H$ of $X$ is a special subordinator as defined in Definition 2.2.16, the lower root of $f_\rho$, $\underline{x}$, is indeed the maximizer of

$$\mathfrak{G}(\rho) = \sup_{y\in B} \mathfrak{g}_{\rho}^{\mathcal{T}_y}(y).$$

If the Assumption 3.4.3 also particularly holds for $\rho^*$, it immediately follows that the $(s, S)$-strategy with $s = \underline{x}$ and $S = \overline{x}$ is optimal for the control problem.
Again we fix a $\rho \in \mathbb{R}$ throughout the section and assume $f_\rho$ has precisely two roots $\underline{x} < \overline{x}$.

**Theorem 4.4.2.** *Let* $x^* := \inf \arg\max_{y\in\mathbb{R}} \ \mathfrak{g}_{\rho}^{\mathcal{T}_y}(y)$. *Then holds*

$$x^* \geq \underline{x}.$$

*If $X$ is not a compound Poisson process and $H$ is a special subordinator then furthermore*

$$\underline{x} \in \arg\max_{y\in\mathbb{R}} \mathfrak{g}_{\rho}^{\mathcal{T}_y}(y)$$

*and hence*

$$x^* = \underline{x}.$$

*Proof.* First, we show $x^* \geq \underline{x}$.
Assume $x < \underline{x}$. Then,

$$\mathbb{E}_x\left(\int_0^{\hat{\tau}_{\underline{x}}} f_\rho(H_s)\ ds\right) < 0$$

and hence we obtain by use of Assumption 3.4.2 in combination with Lemma 2.6.6

$$\begin{aligned}
\mathfrak{g}_\rho^{\mathcal{T}_x}(x) + K &= \mathbb{E}_x \left( \gamma\left(X_{\tau_{\overline{x}}}\right) - \int_0^{\tau_{\overline{x}}} \left(h\left(X_s\right) + \rho\right)\ ds \right) - \gamma(x) \\
&\overset{3.4.2}{=} \mathbb{E}_x \left( \int_0^{\tau_{\overline{x}}} f_\rho\left(\overline{X}_s\right)\ ds \right) \\
&\overset{2.6.6}{=} \mathbb{E}_x \left( \int_0^{\hat{\tau}_{\overline{x}}} f_\rho\left(H_s\right)\ ds \right) \\
&= \mathbb{E}_x \left( \int_0^{\hat{\tau}_{\underline{x}}} f_\rho\left(H_s\right)\ ds + \int_{\hat{\tau}_{\underline{x}}}^{\hat{\tau}_{\overline{x}}} f_\rho\left(H_s\right)\ ds \right) \\
&= \mathbb{E}_x \left( \int_0^{\hat{\tau}_{\underline{x}}} f_\rho\left(H_s\right)\ ds \right) + \mathbb{E}_x \left\{ \mathbb{E}_{H_{\hat{\tau}_{\underline{x}}}} \left( \int_{\hat{\tau}_{\underline{x}}}^{\hat{\tau}_{\overline{x}}} f_\rho\left(H_s\right)\ ds \right) \right\} \\
&< \mathbb{E}_x \left\{ \mathbb{E}_{H_{\hat{\tau}_{\underline{x}}}} \left( \int_{\hat{\tau}_{\underline{x}}}^{\hat{\tau}_{\overline{x}}} f_\rho\left(H_s\right)\ ds \right) \right\} \\
&\le \mathbb{E}_a \left\{ \int_{\hat{\tau}_{\underline{x}}}^{\hat{\tau}_{\overline{x}}} f_\rho\left(H_s\right)\ ds \right\}
\end{aligned}$$

for some $a \in [\underline{x}, \bar{x}]$.
Now we show $x^* \le \underline{x}$ under the assumption that $H$ is a special subordinator and not a compound Poisson process.
Let $U$ be the potential measure of $H$. Since $H$ is a special subordinator, according to Lemma 2.2.17 $U|_{(0,\infty)}$ has a non-increasing density $u$. Since $X$ is not a compound Poisson process, $U$ has, furthermore, no point mass at 0.

Let $x \in [\underline{x}, \bar{x}]$. Then,

$$\begin{aligned}
\mathfrak{g}_\rho^{\mathcal{T}_x}(x) + K &= \mathbb{E}_x \left( \gamma\left(X_{\tau_{\overline{x}}}\right) - \int_0^{\tau_{\overline{x}}} \left(h\left(X_s\right) + \rho\right)\ ds \right) - \gamma(x) \\
&\overset{3.4.2}{=} E_x \left( \int_0^{\tau_{\overline{x}}} f_\rho\left(\sup_{r \le s} X_s\right)\ ds \right) \\
&\overset{2.6.6}{=} \mathbb{E}_x \left( \int_0^{\hat{\tau}_{\overline{x}}} f_\rho\left(H_s\right)\ ds \right) \\
&= \int_x^{\bar{x}} f_\rho(y)\ U(dy - x) \\
&= \int_x^{\bar{x}} f_\rho(y)\, u(y - x)\ dy \\
&= \ldots
\end{aligned}$$

$$\begin{aligned}
... &\leq \int_{x}^{\bar{x}} f_\rho(y)\, u(y-\bar{x})\ dy \\
&\leq \int_{\underline{x}}^{\bar{x}} f_\rho(y)\, u(y-\bar{x})\ dy \\
&= \mathbb{E}_{\underline{x}}\left[\gamma\left(X_{\tau_{\bar{x}}}\right) - \int_0^{\tau_{\bar{x}}} \left(h\left(X_s\right)+\rho\right)\ ds\right].
\end{aligned}$$

These calculations yield

$$\underline{x} \in \arg\max_{y\in\mathbb{R}} \mathfrak{g}_\rho^{\mathcal{T}_y}(y)$$

and hence also $x^* = \underline{x}$. □

### 4.4.2 Step-by-step solution technique

The core of our results is a step-by-step solution technique for long-term average impulse control problems of the form

$$v = \sup_{S=(\tau_n,\zeta_n)_{n\in\mathbb{N}}} \liminf_{T\to\infty} \frac{1}{T}\mathbb{E}_x\left(\sum_{n:\tau_n\leq T}\left(\gamma\left(X^S_{\tau_n,-}\right)-\gamma\left(\zeta_n\right)-K\right)-\int\limits_0^T h\left(X^S_s\right)ds\right)$$

where $X$ is a Lévy process with $0<\mathbb{E}\left(X_1\right)<\infty$ and arbitrary downshifts are allowed. The approach not only provides a sufficient criterion to verify existence of an optimal $(s,S)$-strategy, in many cases it leads to a semi-explicit or explicit characterization of the boundaries. Our step-by-step solution technique reads as follows:

1. Find a function $f$ such that for all $x,\bar{y}\in\mathbb{R}$ with $x<\bar{y}$

   $$\gamma(x) = -\mathbb{E}_x\left[\int_0^{\tau_{\bar{y}}} f\left(\sup_{r\leq t} X_r\right)dt\right] + \mathbb{E}_x\left[\gamma\left(X_{\tau_{\bar{y}}}\right) - \int_0^{\tau_{\bar{y}}} h\left(X_s\right)ds\right].$$

   One way to obtain such a function, provided the occurring objects exist, is the choice

   $$f = \left(A_H\gamma - \hat{h}\right)$$

   as discussed in Definition 2.6.4 and thereafter. Here, $A_H$ is the generator of the ascending ladder hight process $H$ of $X$, extended to non-bounded function as defined in Definition 2.6.4, normed appropriately, see Definition 2.2.14 and for all $y\in\mathbb{R}$

   $$\hat{h}(y) = \mathbb{E}_y\left(\int\limits_0^\infty h\left(H_t^{\downarrow}\right)\ dt\right) = \int_0^\infty h(y+x)\, dU^{\downarrow}(dx) \tag{4.10}$$

   where $H^{\downarrow}$ is the descending ladder height process of $X$ and $U^{\downarrow}$ the occupation measure of $H^{\downarrow}$.

2. Find a $\rho^* \in \mathbb{R}$ such that the equation $f(x) = \rho^*$ has exactly two solutions $\underline{x}_{\rho^*} < \overline{x}_{\rho^*}$ and

$$0 = \sup_{x \in [\underline{x}_{\rho^*}, \overline{x}_{\rho^*}]} \mathbb{E}_x \left( \gamma \left( X_{\tau_{\overline{x}_{\rho^*}}} \right) - \int_0^{\tau_{\overline{x}_{\rho^*}}} \left( h\left(X_s\right) + \rho^* \right) \; ds \right)$$
$$= \sup_{x \in [\underline{x}_{\rho^*}, \overline{x}_{\rho^*}]} \mathbb{E}_x \left( \int_0^{\hat{\tau}_{\overline{x}_\rho}} f\left( \sup_{r \le t} X_r \right) - \rho^* \; dt \right).$$

If such a $\rho^*$ exists, we have

$$v = \rho^*,$$

and the $(s, S)$-strategy with

$$S := \overline{x}_{\rho^*}$$

and

$$s := \arg \max_{x \in [\underline{x}_{\rho^*}, \overline{x}_{\rho^*}]} \mathbb{E}_x \left( \gamma \left( X_{\tau_{\overline{x}_{\rho^*}}} \right) - \int_0^{\tau_{\overline{x}_{\rho^*}}} \left( h\left(X_s\right) + \rho^* \right) \; ds \right)$$

is optimal.

3. If $X$ is not a compound Poisson process and $H$ is a special subordinator (as defined in Definition 2.2.16), we also have

$$s = \underline{x}_{\rho^*}.$$

There are two things we want to remark: first, our results exceed this condensed solution technique and, for example, deliver $\epsilon$-optimal $(s, S)$-strategies, even if no optimizers exist. Moreover technically the continuity of $f$ is not required, it only makes live easier, because if $f$ is not continuous lengthy case differentiations would be required, see Subsection 4.3.1.
Second, the ascending and descending ladder height processes are in general difficult to handle. However, in many special cases there are numerous helpful results known about these processes that often enable us to handle the objects occurring in our solution technique quite well. Andreas Kyprianou's book ([Kyp14]) is an excellent source for theoretical results about these processes, for some short Remarks focussed on the situation here, see also Subsection 2.2.1 and Section 2.6. Also for a dense class of Lévy processes, namely the ones whose positive jumps are of phase type distribution, in [Pis06] an iterative method to explicitly calculate the law of the ascending ladder height process was developed. Furthermore, for many examples of interest, not the full distribution of the ladder height processes needs to be known, but only certain moments. To illustrate the utility of our solution technique, in the following we look at some interesting examples and special cases, that cover the long-term average equivalents to almost all examples of impulse control problems in [ØS05]. We don't pose any restrictions on the process $X$ apart from sufficient integrability to make the problem well defined and non-degenerate.

**Remark 4.4.3.** *So far for underlying Lévy processes we only have studied the case that $B = E$ meaning that arbitrary downshifts are allowed in the control problem. While this is one of the most two common cases, the other is, that for some $y_0 \in E$ we have $B = \{y_0\}$ and hence $B$ is one elementary. Then, according to our main theorem, Theorem 4.3.2, the value is given by*

$$\begin{aligned} v = &\sup_{y\in B}\sup_{\tau\in\mathcal{T}_y} \frac{\mathbb{E}_y\left(\gamma\left(X_\tau\right) - \int_0^\tau h\left(X_t\right)\ dt\right) - \gamma\left(y\right) - K}{\mathbb{E}_y(\tau)} \\ = &\frac{\mathbb{E}_{y_0}\left(\gamma\left(X_\tau\right) - \int_0^\tau h\left(X_t\right)\ dt\right) - \gamma\left(y_0\right) - K}{\mathbb{E}_{y_0}(\tau)}. \end{aligned}$$

*Thus there is one optimization step less to do. Also in this case our solution technique remains unchanged, only the last part of Step 2., namely the search of the optimal restarting point, vanishes. The proof, that this shortened technique again yields the optimal value and the optimal strategy works just the same. There is only one potential obstacle: since we cannot adapt the restarting point, it might be the case that for the optimal $\rho^*$ the value $y_0$ lies left of the lower root $\underline{x}_{\rho^*}$ of $f - \rho^*$. Then, additional arguments are needed to verify that still the threshold time $\tau_{\overline{x}}$ is optimal and not stopping immediately. Nevertheless, in all applications below, we have that $\underline{x}_{\rho^*} \leq y_0$ hence we can follow the solution techniques without worries.*

### 4.4.3 Interesting special cases

#### Linear $\gamma$ and convex $h$

The first special case we want to focus on is the case when for all $x \in \mathbb{R}$

$$\gamma\left(x\right) = Cx$$

for some $C \in (0,\infty)$ and $h$ is positive, convex, continuous and we have $\lim_{x\to\infty} h\left(x\right) = \lim_{x\to-\infty} h\left(x\right) = \infty$. This type of pay-off and cost functions occur in inventory control, see, e.g., [HSZ17] if the underlying process is a diffusion, but for Lévy processes there are yet very few results. In [Yam17] in the setting of discounted pay-offs Yamazaki proves existence of an optimal $(s,S)$-strategy only under the additional assumption that $X$ is spectrally positive. Also the long-term average equivalent to the dividend problem presented in [ØS05, Example 6.4], and the prominent exchange rate control problem ([ØS05, Example 6.5]) is covered by this special case.
With our solution technique one can show existence of an optimal $(s,S)$-strategy quite easy.

1. To obtain $f$ we first observe that $H$ is a subordinator and therefore

$$A_H\gamma = C\tilde{\mu}_H + \int_0^\infty y\Pi_H\left(dy\right)$$

for the drift term $\tilde{\mu}_H$ and the jump measure $\Pi_H$ of $H$ as defined in Remark 2.2.5. Hence $A_H\gamma(\cdot)$ is constant. Further, for all $y \in \mathbb{R}$

$$\hat{h}(y) = \int_0^\infty h(y+x)\, dU^{\downarrow}(dx),$$

thus $\hat{h}$ is still convex with $\lim_{x\to\infty} \hat{h}(x) = \lim_{x\to-\infty} \hat{h}(x) = \infty$. Further, it follows that the equation $f(x) = \rho$ always has exactly two solutions if we choose $\rho$ large enough.

2. The function given by

$$\rho \mapsto \sup_{x\in[\underline{x}_\rho, \overline{x}_\rho]} \mathbb{E}_x \left( \int_0^{\tau_{\overline{x}_\rho}} f\left(\sup_{r\le t} X_r\right) - \rho \; dt \right)$$

is monotone and continuous, thus either the intermediate value theorem provides the desired root $\rho^*$, or we are in a degenerate case and the value is either $\infty$ or $-\infty$.

As consequence, our solution technique verifies the existence of an $(s, S)$-strategy. Furthermore, concerning the explicit obtainability, we want to remark that the extended generator of the ladder height process in general is difficult to obtain. But to obtain the function

$$\begin{aligned} f(x) &= -\left(A_H\gamma(x) + \hat{h}(x)\right) \\ &= \left\{ C\tilde{\mu}_H + \int_0^\infty y\Pi_H(dy) + \mathbb{E}_x\left( \int_0^\infty h\left(H_t^{\downarrow}\right) \; dt \right) \right\} \end{aligned}$$

one does not need full knowledge of $A_H$. Only the drift term $\tilde{\mu}_H$ and the expected jump size of the ladder height process are needed. These parameters in principal can be expressed in parameters of $X$ which in some cases leads to good characterizations, as we already have discussed in Section 2.6. But more importantly they are accessible by path-wise simulation techniques, e.g., Monte Carlo methods since by simulating paths of the initial process $X$, one can often directly derive the ladder height processes path and therefore also its jumps and drift parts. The same arguments hold for the obtainability of $\hat{h}$. In most treated examples, $h$ are relatively simple functions, like (piecewise) linear ones, the square function, or even just a constant function. Hence, in that cases $\hat{h}$ is just an integral moment of the descending ladder height process.

### Polynomial $\gamma$ and $h$

To show that basically the same arguments for the obtainability of $f$ hold in more general cases we now turn our attention to the case where $\gamma$ and $h$ are polynomials. This example is certainly one of great interest since on one hand polynomials are interesting special cases of pay-off and running cost functions

on its own. On the other hand, polynomials may serve as a tool for approximating more general functions. In the following, we will see that the necessary transformations we have to apply on $\gamma$ and $h$ have the very compelling property to transform polynomials to polynomials of the same degree. This makes our solution technique boil down to an analysis of a polynomial of known degree whose coefficients can be expressed in terms of $\gamma$, $h$ and parameters of the process.
This setting includes the long-term average analogous to [ØS05, Example 7.8]. So now we assume $\gamma(x) = \sum_{i=0}^{l} a_i x^i$ and $h(x) = \sum_{i=0}^{k} c_i x^i$ for some $l, k \in \mathbb{N}$ and $a_i, c_i \in \mathbb{R}$ for all $i \leq l \vee k$. For the sake of simplicity and brevity we simply assume the occurring moments and integrals to exist and be finite. Then,

$$A_H \gamma(x) = \sum_{i=0}^{l} b_i x^i$$

with

$$b_i = (i+1)\tilde{\mu}_H a_{i+1} + \sum_{j=i}^{l} a_j \binom{j}{i} \int_0^\infty y^{j-i} d\Pi_H(y)$$

for all $i \in \{0, ..., l\}$ where $a_{l+1} = 0$ and $\Pi_H$ is the Lévy measure of $H$. Furthermore, using Fubini's theorem we obtain

$$\hat{h}(x) = \sum_{j=0}^{k} d_j x^j$$

where

$$d_j := \sum_{i=1}^{k} c_i \binom{i}{j} \int_0^\infty \mathbb{E}\left(\left(H_t^{\downarrow}\right)^{i-j}\right) dt.$$

Again we remark that some useful formulas to obtain $\Pi_H$ can be found in Section 2.6, further [Kyp14, Section 7.4] entails a comprehensive discussion. Further, $H^{\downarrow}$ acts in law like an exponentially killed subordinator, hence the occurring moments can be obtained via the cumulant function of said subordinator. Also, we don't need to know the full distribution of ascending and descending ladder height processes. To explicitly get the occurring coefficients one only has to calculate drift rate, moments of the jump measure of the ascending ladder height process, and cumulative moments of the descending ladder height process, both up to a previously known given degree. So again this is a good starting point for simulations.
The observation that $f = \left(A_H \gamma - \hat{h}\right)$ is also a polynomial with degree $\max\{l, k\}$ on top of that yields a starting point for a different procedure to (computationally) find $f$. Our calculations above show that $f$ is a polynomial of known degree. Hence, it is possible to just start with the desired equation

$$\gamma(x) = -\mathbb{E}_x\left[\int_0^{\tau_{\bar{y}}} f\left(\sup_{r \leq t} X_r\right) dt\right] + \mathbb{E}_x\left[\gamma\left(X_{\tau_{\bar{y}}}\right) - \int_0^{\tau_{\bar{y}}} h(X_s)\, ds\right],$$

plug in a general polynomial of the right degree for $f$, compute the occurring moments and integrals either explicitly or approximate them numerically and then compare the coefficients. For example, the aforementioned work [Pis06] also gives an iterative method to explicitly obtain the Laplace transform of the law of the running supremum of $X$, in case the upward jumps of $X$ are of phase type.

**Exponential Lévy processes**

A class of processes which is of great interest in mathematical finance is the class of exponential Lévy processes, for example, [ØS05, Exercises 6.2 and 7.2] present examples where this setting is used to determine the optimal stream of dividends.
Set for all $x \in \mathbb{R}$

$$\gamma(x) := e^x$$

and for all $x \in \mathbb{R}$

$$h(x) := a_1 e^{a_2 x} + b_1 e^{-b_2 x}$$

for $a_1, a_2, b_1, b_2 \in (0, \infty)$. Again apart from $\mathbb{E}(X_1) > 0$, we only assume the occurring moments and integrals to exist and be finite.
To obtain $f$ similar to the polynomial case we see that

$$\begin{aligned} A_H \gamma(x) &= \tilde{\mu}_H e^x + e^x \int_0^\infty e^y \Pi_H(dy) \\ &= \tilde{\mu}_H e^x + e^x \int_0^\infty e^y \int_0^\infty \Pi(z + y, \infty) U^{\downarrow}(dz)\ dy. \end{aligned}$$

Further,

$$\begin{aligned} \hat{h}(x) &= \int_0^\infty a_1 e^{a_2(x+y)} - b_1 e^{b_2(x+y)} dU^{\downarrow}(dy) \\ &= a_1 e^{a_2 x} \int_0^\infty e^{a_2 y} dU^{\downarrow}(dy) + b_1 e^{-b_2 x} \int_0^\infty e^{-b_2 y} dU^{\downarrow}(dy) \end{aligned}$$

for all $x \in \mathbb{R}$. Hence, in this case finding the optimal value and an optimal strategy boils down to the analysis of exponential functions.

### 4.4.4 Proof of the validity of the solution technique

The scope of this subsection is to briefly connect the dots and use our findings in order to show that the step-by-step solution technique presented Section 4.4.2 indeed is valid.

1. Lemma 2.6.7 shows that for

$$f = \left(A_H \gamma - \hat{h}\right)$$

we have

$$\gamma(x) = -\mathbb{E}_x\left[\int_0^{\tau_{\bar{y}}} f\left(\overline{X}_t\right) dt\right] + \mathbb{E}_x\left[\gamma\left(X_{\tau_{\bar{y}}}\right) - \int_0^{\tau_{\bar{y}}} h\left(X_s\right) ds\right]$$

for all $x, \bar{y} \in \mathbb{R}$ with $x \leq \bar{y}$, hence the desired maximum representation of $\gamma$.

2. The second step of the solution technique is to find a $\rho^* \in \mathbb{R}$ such that $f(x) = \rho^*$ has exactly two solutions $\underline{x}_{\rho^*} < \overline{x}_{\rho^*}$ and

$$\begin{aligned} 0 &= \sup_{x \in [\underline{x}_{\rho^*}, \overline{x}_{\rho^*}]} \mathbb{E}_x\left(\gamma\left(X_{\tau_{\overline{x}_{\rho^*}}}\right) - \int_0^{\tau_{\overline{x}_{\rho^*}}} \left(h\left(X_s\right) + \rho^*\right)\ ds\right) - \gamma(x) - K \\ &= \sup_{x \in [\underline{x}_{\rho^*}, \overline{x}_{\rho^*}]} \mathbb{E}_x\left(\int_0^{\hat{\tau}_{\overline{x}}} f\left(\overline{X}_t\right) - \rho^*\ dt\right) - K. \end{aligned}$$

Assume we have found such elements. Then, Theorem 3.4.9 yields that the threshold time $\tau_{\overline{x}_{\rho^*}}$ is an optimizer for the stopping problem with value function $g_\rho^{\mathcal{T}_\cdot}$. Further, the first part of Theorem 4.4.2 ensures that there is an

$$s \in \arg\max_{x \in [\underline{x}_{\rho^*}, \overline{x}_{\rho^*}]} \mathbb{E}_x\left(\gamma\left(X_{\tau_{\overline{x}_{\rho^*}}}\right) - \int_0^{\tau_{\overline{x}_{\rho^*}}} \left(h\left(X_s\right) + \rho^*\right)\ ds\right).$$

For this $s$ by definition we have

$$\mathfrak{G}(\rho^*) = \mathfrak{g}_{\rho^*}^{\mathcal{T}_s}(s)$$

where $\mathfrak{G}$ is defined in Definition 4.2.3 and $\mathfrak{g}$ right before. Now Corollary 4.2.8 yields

$$v = \rho^*$$

and Corollary 4.2.9 shows that the strategy $R(\tau_{\overline{x}_{\rho^*}}, s)$ as defined in Subsection 4.1.1 is optimal.

3. The second part in Theorem 4.4.2 shows that in case that $H$ is a special subordinator and not a compound Poisson process, $\underline{x}$ is a valid choice for $s$. Further, Lemma 2.2.17 and Lemma 2.2.18 provided conditions in term of properties of $X$ under that $H$ is a special subordinator and give some examples.

# Chapter 5

# Applications

After we have proven the validity of our solution technique and already have seen some interesting special cases, where we got existence results and good starting points for numerical analysis, we now demonstrate how our solution technique yields (semi-)explicit characterizations of the control boundaries by applying it to two important fields of application. These are inventory control and optimal harvesting. In both fields we will demonstrate that, especially when it comes to models with jumps, our technique yields results that exceed the present results in the literature. After these two applications in a narrower sense of the word, we turn our attention to a topic that sheds some light on the inherent structure of the impulse control problem: we investigate the dependence on the fixed cost term and analyse how value and strategy behave when this term goes to infinity or zero. By that we unveil connections to singular control. A last application in the broader sense is the study of long-term average impulse control problems with a restriction to the average impulse frequency. Therein we show by a Lagrange type ansatz that the restriction to the impulse frequency is in some sense equivalent to adding artificial fixed costs.

### Structure of the chapter

The first section, Section 5.1, deals with inventory control problems for spectrally positive Lévy processes. Section 5.2 contains applications of our technique to a harvesting problem with an underlying spectrally negative Lévy process. Section 5.3 entails the analysis of the problem's dependence on the fixed cost term. Section 5.4 treats the long-term average impulse control problem with a restriction to the average impulse frequency.

## 5.1 Inventory Control for Spectrally One-Sided Lévy Processes

The first example we treat is inventory control. We want to remark that in inventory control one usually seeks to minimize the costs of ordering supplies and maintaining a stock depending on a draining inventory level modelled by a process $X$. Hence here we have to turn the usual setting of inventory control (see, e.g., [Yam17] or [HSZ17]) 'upside down' to translate it to the maximization problem, that this thesis deals with. Although the majority of inventory control problems uses discounting cost and pay-off functionals, we want to emphasize that in inventory control the long-term average reward is of no less interest compared to the discounted one. For instance in [HSZ17] the authors show optimality for $(s, S)$ strategies (this means for two values $s, S \in E$ with $s < S$ shifting the process down to $s$ whenever it exceeds $S$) in the long-term average problem under roughly similar conditions as ours here, provided the underlying process $X$ is a diffusion, after they obtained comparable results in the case with discounted pay-off in [HSZ15]. However, the lack of existence of a 0-resolvent and the often constant value function makes it impossible to directly adapt techniques from the discounted case in the long-term average one. Usual results in inventory control prove existence of optimal $(s, S)$ strategies and sometimes even characterize the optimal boundaries as maximizers of some functionals given by parameters of the process. In the long-term average setting such results currently are only present for diffusions, see [HSZ17] and [HSZ18]. Therein the optimal values are given as optimizers of a functional that consists of integrals over speed measure and scale function of the underlying diffusion, see [HSZ17, Proposition 3.5]. For similar results in the discounted case, see, e.g., [HSZ15]. Although there are yet very few comparable results for Lévy processes, over the course of the last decade the theory of scale functions for spectrally one-sided Lévy processes gave rise to new advances in control theory of these processes. In inventory control in [Yam17] Yamazaki applied these techniques to show optimality of an $(s, S)$ strategy when the reward functional is a discounted one. Additionally, he characterized the boundaries as optimizers of a certain functional by use of the scale functions under roughly the following conditions:

1. The process $X$ is a Lévy process, drifts upwards, is spectrally positive and $\mathbb{E}(\exp(\beta X_1)) < \infty$ for some $\beta > 0$.

2. The pay-off function $\gamma$ is linear.

3. The running costs function $h$ is unimodal, convex on the right of its minimum, grows at least polynomially, $h'(x) > c > 0$ for all $x < x_0$ for some $x_0$ and $c > 0$ and fulfils some more smoothness and integrability conditions (which in their full extend can be seen in [Yam17, Assumption 1]).

Under these conditions in [Yam17, Theorem 1] it is shown that an optimal $(s, S)$ strategy exists. [Yam17, Proposition 1] furthermore states that the value function can be expressed in terms of integral identities that comprise the running cost function, the scale functions and the Lévy exponent of $X$, as well as the right inverse of the Lévy exponent of $X$. The optimal pair of values $(s^*, S^*)$ is in Proposition 3 therein characterized as an optimizer of $\min_s \max_S \mathcal{G}(s, S)$, where $\mathcal{G}$ is also a function comprising all the objects that occur in the representation of the value function.

In this subsection we use our solution technique to first prove existence of an optimal $(s, S)$ strategy in the long-term average case under less restrictive assumptions than the ones used in [Yam17] for the discounted case. Also we characterize the optimal value and the optimal boundaries using only the Lévy triplet of $X$ and the root of the right inverse of its Lévy exponent. Namely we assume

**Assumption 5.1.1.** *1. The process $X$ is a Lévy process that fulfils $0 < \mathbb{E}(X_1) < \infty$, is spectrally positive and all later occurring integrals exist.*

*2. For the pay-off function $\gamma$ we have $\gamma(x) = Cx$ for a $C \in [0, \infty)$.*

*3. $h$ is positive and unimodal with unique minimum $a \in \mathbb{R}$, it only grows polynomially and we have $\lim_{x\to\infty} h(x) = \infty = \lim_{x\to-\infty} h(x)$.*

Now we follow the steps laid out in the solution technique in Subsection 4.4.2. To tackle 1. we first have to get a grip on $f$. Note that since $X$ is spectrally positive its descending ladder high process $H^{\downarrow}$ acts in law as the non-random process $(t)_{t\geq 0}$ killed exponentially with a positive rate $q > 0$ where $q = -\phi(0)$, $\phi$ being the right inverse of the Laplace exponent of $-X$, see Remark 2.6.8. Hence, the function $\hat{h}$ can be obtained via

$$\hat{h}(x) = \int_0^\infty e^{-qt} h(t + x)\ dt.$$

Further,

$$A_H \gamma(x) = C\tilde{\mu}_H + C \int_0^\infty y\ \Pi_H(dy),$$

where $A_H$ is the generator of $H$ understood in a generalized sense as discussed in Section 2.6 and defined in Definition 2.6.4, $\Pi_H$ is the Lévy measure of the ladder height process $H$ and $\tilde{\mu}_H$ is the drift term of $H$ as defined in Definition 2.2.5, so $A_H\gamma$ does not depend on $x$. The Lévy measure $\Pi_H$ can be expressed in terms of $q$ and the Lévy measure $\Pi_X$ of $X$ via the formula

$$\Pi_H(x, \infty) = e^{qx} \int_x^\infty e^{-qy} \Pi_X(y, \infty)\ dy,$$

see Remark 2.6.8, hence

$$\begin{aligned}A_H\gamma(x) &= C\tilde{\mu}_H + C\int_0^\infty y\Pi_H(dy)\\ &= C\tilde{\mu}_H + C\int_0^\infty \Pi_H(z,\infty)\ dz\\ &= C\tilde{\mu}_H + C\int_0^\infty e^{qz}\int_z^\infty e^{-qy}\Pi_X(y,\infty)\ dy\ dz\\ &= C\tilde{\mu}_H + C\int_0^\infty\int_0^\infty e^{-qy}\Pi_X(y+z,\infty)\ dz\ dy.\end{aligned}$$

This yields for all $x \in \mathbb{R}$

$$\begin{aligned}f(x) &= A_H\gamma(x) - \hat{h}(x)\\ &= C\tilde{\mu}_H + C\int_0^\infty\int_0^\infty e^{-qy}\Pi_X(y+z,\infty)\ dz\ dy - \int_0^\infty e^{-qt}h(t+x)\ dt.\end{aligned}$$

Now we have for each $\rho$ such that $\underline{x}_\rho$ and $\overline{x}_\rho$ exist

$$\begin{aligned}&\mathbb{E}_{\underline{x}_\rho}\left(\gamma\left(X_{\tau_{\overline{x}_\rho}}\right) - \int_0^{\tau_{\overline{x}_\rho}}(h(X_s)+\rho)\ ds\right) - \gamma(\underline{x}_\rho) - K\\ &= \mathbb{E}_{\underline{x}_\rho}\left(\int_0^{\hat{\tau}_{\overline{x}_\rho}} f_\rho(H_s)\ ds\right) - K\\ &= (A_H\gamma+\rho)\,\mathbb{E}_{\underline{x}_\rho}\left(\tau_{\overline{x}_\rho}\right) - \mathbb{E}_{\underline{x}_\rho}\left(\int_0^{\hat{\tau}_{\overline{x}_\rho}}\hat{h}(H_s)\ ds\right) - K.\end{aligned}$$

Since $\hat{h}$ is a Laplace transform of a continuous function and $A_H\gamma$ is constant, this term is a continuous function of $\rho$. The growth condition Assumption 5.1.1, 3. therefore enables us to use the mean value theorem and we obtain that there is a $\rho^*$ such that

$$0 = \mathbb{E}_{\underline{x}_{\rho^*}}\left(\gamma\left(X_{\tau_{\overline{x}_{\rho^*}}}\right) - \int_0^{\tau_{\overline{x}_{\rho^*}}}(h(X_s)+\rho^*)\ ds\right) - \gamma(\underline{x}_{\rho^*}) - K.$$

Thus our solution technique not only yields existence of an optimal threshold (or (s,S)) strategy, it also characterises the optimal boundaries.

## 5.2 Optimal Harvesting

Another field of application for our solution technique is optimal harvesting and forest management. This problem originates in the work of Martin Faustmann starting with his seminal paper [Fau49] from 1849. Until now advancements and derivatives of this approach are used and usually called Faustmann's formula, see [Bra01] for an overview. In this branch of impulse control the underlying process models the growth of a forest, or more general: a natural resource, and impulse control theory is used to determine the optimal strategy to repeatedly harvest the resource. The question how to optimally exploit a natural resource whose dynamics involve randomness goes back several decades, see [MBHS78] for one of the earlier works. Nowadays there is a vast amount of literature present ranging from very applied to rather theoretical treatises (see, e.g., [BS88, Wil98, Alv04, AK06, SS10]). While both modeling and solution approaches differ varying by the specific field of application, most of these works have in common that they describe the dynamics of the natural resource by a logistic diffusion. [AS98] provides a solution to the impulse control problem with an underlying logistic diffusion in the discounted case. Although in this fields the discounted pay-off functional is the most common choice, recently more and more works point out that on one hand it is difficult to determine the right discounting factor in practice and on the other hand the discounted model has the flaw to favor the present compared to the future and therefore might not be the right choice when one aims for sustainable solutions. The recent article [AH20] provides an example for the application of the long-term average criterion to find a 'sustainable' harvesting strategy and a discussion of the model, see also [CS19b]. Here, we take a look at a typical Faustmann-type forest management problem as presented for instance in [AL08] or [AK06], but we deviate from modelling the forest growth by a logistic diffusion. Instead we introduce downward jumps to the process since sudden events like storms, bushfire or diseases of the trees could abruptly destroy large quantities of the forest stand or make it worthless. First, we tackle the traditional Faustmann problem with $|B| = 1$ and hence only a single possible restarting point. Then, we will investigate the same model, only with arbitrary downward shifts allowed and compare both cases. After these two quickly solved harvesting problems here, we will come back to these kind of questions later in Chapter 6. To make our solution technique applicable we make the following assumption:

**Assumption 5.2.1.** *The process $X$ is a Lévy process with $0 < \mathbb{E}(X_1) < \infty$, spectrally negative and all later occurring integrals exist and are finite.*

Since the homogeneity of a Lévy process doesn't match the growth of a forest very well, we will later use the pay-off function to shape the model in a realistic way.

### 5.2.1 Fixed restarting point

As discussed, first set $B = \{0\}$, hence only one fixed restarting point is allowed. We will use the solution technique presented in Subsection 4.4.2 with the alteration for a fixed restarting point discussed in Remark 4.4.3. We set $h = 0$ and

$$\gamma : \mathbb{R} \to \mathbb{R};\ x \mapsto \frac{L}{1 + e^{-sx}}$$

for $L, s > 0$. The function $\gamma$ is called logistic function and is used in many (mostly non-random) contexts. It was supposedly first introduced over 170 years ago by Pierre Francois Verhulst to model population growth. And it is still a popular choice when it comes to modelling natural growth in a bounded environment. The choice of $\gamma$ is motivated as follows: in the aforementioned works on optimal harvesting a logistic diffusion is used to model the tree stand. Since, contrary to a diffusion model, we can't model different growth rates dependent on the current population with our Lévy process directly, we interpret $\gamma(X_t)$ as the forest stand at time $t$ and since $sX$ is still a Lévy process, we don't lose generality by assuming $s = 1$.
Since there is no choice in the restarting point, no running costs and no upward jumps, the procedure to find the optimal strategy breaks down to:

1. For arbitrary $\rho \in \mathbb{R}$ find the rightmost root $\overline{x}_\rho$ of $\mathbb{E}(X_1)\gamma' - \rho$.
2. Find $\rho^*$ such that

$$\gamma\left(\overline{x}^*_\rho\right) - \gamma(0) - K - \rho^* \mathbb{E}_0\left(\tau_{\overline{x}_\rho}\right) = 0.$$

It is well known (or easily obtained) that

$$\gamma'(x) = L\frac{e^x}{(1 + e^x)^2}$$

for all $x \in \mathbb{R}$. Hence, we only have to solve a system of two equations. While for general $L$ no closed form of the solution may be written down, this is a problem that easily may be solved numerically for given parameters.
There are some interesting observations to make on this model. First of all, by looking at the non-random case $X_t = t$ for all $t$, this model indeed becomes the standard non-random model for natural growth as described before. When one then adds randomness by, say, a compound Poisson process with only downward jumps to take unforeseeable unfavourable events into account, this only influences the equations above via the $\mathbb{E}(X_1)$ term (note that the $\mathbb{E}_0(\tau_{\overline{x}_\rho})$ may be transformed with Wald's equation (Lemma 2.2.8)). Hence these added downward jumps would make the forester harvest already at a lower level and therefore more often and also let the problem's value decrease. But since only the expected value of the process influences value and strategy, it does not play any role whether there are many small or few large downward jumps. This is due to our choice of $h = 0$; if we would have a non-zero running cost term, we would have a more complex influence of $X$'s path structure to value and strategies.

### 5.2.2 Arbitrary downshifts allowed

Now we consider the same setting as in the previous example with the only difference that we allow arbitrary downshifts. Then, the values $\underline{x}_{\rho^*}, \overline{x}_{\rho^*}, \rho^*$ can be found as follows:

1. For arbitrary $\rho \in \mathbb{R}$ find the two roots $,\underline{x}_\rho, \overline{x}_\rho$ of $\gamma' - \rho$ (since $\gamma'$ is symmetric, we have $\overline{x}_\rho = -\underline{x}_\rho$).

2. Find $\rho^*$ such that

$$\gamma\left(\overline{x}_{\rho^*}\right) - \gamma\left(\underline{x}_{\rho^*}\right) - K - \rho^* \mathbb{E}_{\underline{x}_{\rho^*}}\left(\tau_{\overline{x}_\rho}\right) = 0.$$

Due to the symmetry these equations are as easy to solve as the ones with only one allowed restarting point. Of course these two applications mostly serve the purpose of easy examples to illustrate our findings nicely on a not too abstract level. Nevertheless, even this easy examples stress out some noteworthy observations:

- The only thing we have to know about the underlying Lévy process (apart from the absence of upward jumps) is $\mathbb{E}(X_1)$. This opens the door to easy estimation and calculation procedures of the optimal boundaries.

- More freedom in the choice of the restarting point, of course, yields a higher value for the control problem as well as more frequent trading. Heuristically this makes sense since we may lay our thresholds in the area where $\gamma$ has the highest gradient, so it does not take as long to compensate the fixed cost term.

## 5.3 The Influence of the Fixed Cost Term

Our next aim is to study how the fixed cost term influences the control problem. Particularly we show here that the value continuously depends on the fixed cost term $K$ and is non-increasing in $K$. Further, we investigate the behaviour of the value and the optimal control boundaries if the fixed cost term converges to zero and infinity. We see that the former reveals connection to singular control whereas the latter one (not surprisingly) calls for 'do nothing' as optimal solution. Because both singular controls and to do nothing are technically excluded in our setting, both for too large $K$ ad $K = 0$ we hence usually only get $\epsilon$-optimal strategies.

Throughout this section we again work with the notations and assumptions given in Subsection 4.1.1, especially $X$ is assumed to be an general Markov process as defined in Definition 2.1.3 and the general model for controlled processes explained in Subsection 4.1.1 is used. Let Assumption 3.4.1 regarding the functions $\gamma$ and $h$ hold and let Assumption 3.4.2 regarding existence of a maximum representation hold, further assume that a continuous representing function $f$ as in Assumption 3.4.2 exist. Lastly assume $E = B = \mathbb{R}$. Since here we want to study effects of changing fixed cost term, we also put more emphasis on the variable $K$ and write for each $K \geq 0$ and each $S = (\tau_n, \zeta_n) \in \mathcal{S}_\mathbb{R}$

$$J_{K,x}(S) := \liminf_{T\to\infty} \frac{1}{T}\mathbb{E}_x \left( \sum_{n:\tau_n \leq T} \left(\gamma\left(X^S_{\tau_n,-}\right) - \gamma(\zeta_n) - K\right) - \int\limits_0^T h\left(X^S_s\right) ds \right)$$

for all $x \in \mathbb{R}$. Further, we define

$$\begin{aligned} v(K)(x) &:= \sup_{S\in\mathcal{S}_\mathbb{R}} J_{K,x}(S) \\ &= \sup_{S\in\mathcal{S}_\mathbb{R}} \liminf_{T\to\infty} \frac{1}{T}\mathbb{E}_x \left( \sum_{n:\tau_n \leq T} \left(\gamma\left(X^S_{\tau_n,-}\right) - \gamma(\zeta_n) - K\right) - \int\limits_0^T h\left(X^S_s\right) ds \right) \end{aligned} \quad (5.1)$$

for all $x \in \mathbb{R}$ and to avoid having to discuss several degenerate cases, we assume for all $x \in \mathbb{R}$ and all $K \geq 0$ that $v(K)(x) < \infty$. This implies that for each $K \geq 0$ the function $v(K)$ is constant (This may be seen elementary: since arbitrary downward shifts are allowed, for $x, y \in \mathbb{R}$ with $x < y$ we have $v(K)(x) \leq v(K)(y)$ but when the process is started in $x$ waiting until is exceeds $y$ for the first time due to Assumption 3.4.1 only yields finite costs which are negligible in the limit, hence $v(K)(x) = v(K)(y)$). In accordance with Assumption 3.4.2 and the assumptions above let $f$ be a function such that:

1. For all $x, \overline{y} \in \mathbb{R}$ with $x \leq \overline{y}$ holds
$$\gamma(x) = -\mathbb{E}_x\left[\int_0^{\tau_{\overline{y}}} f\left(\overline{X}_t\right) dt\right] + \mathbb{E}_x\left[\gamma\left(X_{\tau_{\overline{y}}}\right) - \int_0^{\tau_{\overline{y}}} h(X_s)\, ds\right].$$
2. The function $f$ has a unique maximum $a \in \mathbb{R}$, is strictly increasing on $(-\infty, a]$ and strictly decreasing on $[a, \infty)$.

3. $f$ is continuous.

Since here we need actual optimizers for the control problems and not just $\epsilon$-optimal strategies we make the following assumption:

**Assumption 5.3.1.** *The functions*

$$\Xi : \{(x, y \in \mathbb{R}^2 | x < y)\} \to \mathbb{R}; (a, b) \mapsto \mathbb{E}_a \left( \int_0^{\tau_b} f(\overline{X}_s) ds \right)$$

*and*

$$\xi : \{(x, y \in \mathbb{R}^2 | x < y)\} \to \mathbb{R}; (a, b) \mapsto \mathbb{E}_a(\tau_b)$$

*are continuous.*

That this assumptions hold in many cases of interest is discussed in Lemma 2.2.13, Lemma 2.2.11 and Remark 2.3.6 for Lévy processes and diffusions. To analyse $v$'s dependence on the fixed cost term $K$, let us shortly review the key results on long-term average impulse control problems again. Theorem 4.3.6 yields that

$$v(K) = \sup_{y^*, \overline{y} \in \mathbb{R}; y^* < \overline{y}} \frac{\mathbb{E}_{y^*} \left( \gamma \left( X_{\tau_{\overline{y}}} \right) - \int_0^{\tau_{\overline{y}}} h\left(X_t\right) \, dt \right) - \gamma\left(y^*\right) - K}{\mathbb{E}_{y^*}\left(\tau_{\overline{y}}\right)}. \tag{5.2}$$

Set $\alpha := \max\{\inf\{f(x) \mid x \le a\}, \inf\{f(x) \mid x \ge a\}\}$.

**Remark 5.3.2.** *The continuity of the function $f$ ensures that for all $K$ with $\alpha < v(K) < f(a)$ the function $f_{v(K)} = f - v(K)$ has two roots and hence Assumption 3.4.3 holds. Assumption 3.4.8 holds, because the function $\Xi$ is continuous, as assumed in Assumption 5.3.1. Theorem 4.3.2 therefore yields that for $K > 0$ with $v(K) > \alpha$ an optimal threshold strategy $R(\tau_{\overline{x}_K}, x_K^*)$ is given by values $\underline{x}_K$, $x_K^*$ and $\overline{x}_K$ that are defined as follows: $\underline{x}_K$ and $\overline{x}_K$ are due to our assumption the unique two roots of the equation*

$$f(x) = v(K)$$

*with $\underline{x}_K \le \overline{x}_K$. The value $x_K^* \in [\underline{x}_K, \overline{x}_K]$ in general is an element of the non-empty set*

$$\arg \max_{y \in [\underline{x}_K, \overline{x}_K]} \frac{\mathbb{E}_y \left( \gamma \left( X_{\tau_{\overline{x}_K}} \right) - \int_0^{\tau_{\overline{x}_K}} h\left(X_t\right) \, dt \right) - \gamma\left(y\right) - K}{\mathbb{E}_y(\tau_{\overline{x}_K})}.$$

*In special cases we also have $x_K^* = \underline{x}_K$. This was discussed in Lemma 4.4.1 and Lemma 4.3.1 and covers many cases of interest.*

The obvious next question is, whether for all $K \in [0, \infty)$ we have

$$\alpha < v(K) < f(a).$$

Here, Lemma 4.3.3 yields that if the sets $\{\mathbb{E}_x(\tau_y) \mid (x,y) \in \mathbb{R}^2, x < y \leq a\}$ and $\{\mathbb{E}_x(\tau_y) \mid (x,y) \in \mathbb{R}^2, a \leq x < y\}$ are not bounded from above, then for all $K \in [0,\infty)$ holds $v(K) \geq \alpha$. Further, Lemma 4.3.4 yields

$$v(K) \leq f(a)$$

for all $K \in [0,\infty)$ and Lemma 4.3.5 additionally yields that under our assumption that $\xi$ is continuous we have

$$v(K) < f(a)$$

for all $K \in (0,\infty)$. We will see later that $\alpha$ is not always the lower boundary, but before we will prove this, we have to analyse the function $v$ further.

**Theorem 5.3.3.** *The function $v : [0,\infty) \to [-\infty, f(a)]; K \mapsto v(K)$ is non-increasing, even strictly decreasing on $\{K \mid K \geq 0, v(K) > \alpha\}$, convex, continuous and almost everywhere differentiable.*

*Proof.* To check that $v$ is well defined, remind that for all $K > 0$ holds $v(K) \leq f(a)$ as was shown in Lemma 4.3.4. Further, by just looking at the definition of $v$ in (5.1) or the characterization (5.2), it is immediately clear that $v$ is strictly decreasing until it reaches the (possible infinite) value $\alpha$. The definition by (5.1) yields convexity since $v$ is defined as the supremum over linear functions. This also yields differentiability and continuity outside zero. To see that $v$ is continuous in 0 we first note that $v(0) \leq f(a)$ as was shown in Lemma 4.3.4. Now under our assumptions here we have for all $y \in \mathbb{R}$

$$\begin{aligned}
&\lim_{x \searrow y} \frac{\mathbb{E}_y\left(\gamma\left(X_{\tau_x}\right) - \int_0^{\tau_x} h\left(X_t\right)\ dt\right) - \gamma\left(y\right)}{\mathbb{E}_y(\tau_x)} \\
=&\lim_{x \searrow y} \frac{\mathbb{E}_y\left(\int_0^{\tau_x} f\left(X_t\right)\ dt\right) - \gamma\left(y\right)}{\mathbb{E}_y(\tau_x)} \\
=&f(y).
\end{aligned}$$

Hence

$$v(0) = \sup_{x,y \in \mathbb{R}, x > y} \frac{\mathbb{E}_y\left(\gamma\left(X_{\tau_x}\right) - \int_0^{\tau_x} h\left(X_t\right)\ dt\right) - \gamma\left(y\right)}{\mathbb{E}_y(\tau_x)} = f(a).$$

Further, the thresholds $\underline{x}_K$ are non-increasing in $K$ and the thresholds $\overline{x}_K$ are non-decreasing in $K$, since they are the roots of $f - v(K)$ and $f$ is unimodal and $v$ non-increasing. Because $\xi$ is continuous (see Assumption 5.3.1), there is a $K_1 \in (0,\infty)$ such that the set $\{\mathbb{E}_x(\tau_y) \mid x,y \in [\underline{x}_{K_1}, \overline{x}_{K_1}], x < y\}$ is bounded by some $C \in (0,\infty)$. Now let $\epsilon > 0$ and take $x', y' \in [\underline{x}_{K_1}, \overline{x}_{K_1}]$ such that

$$\begin{aligned}
v(0) &= \sup_{x,y \in \mathbb{R}, x > y} \frac{\mathbb{E}_y\left(\gamma\left(X_{\tau_x}\right) - \int_0^{\tau_x} h\left(X_t\right)\ dt\right) - \gamma\left(y\right)}{\mathbb{E}_y(\tau_x)} \\
&< \frac{\mathbb{E}_{y'}\left(\gamma\left(X_{\tau_{x'}}\right) - \int_0^{\tau_{x'}} h\left(X_t\right)\ dt\right) - \gamma\left(y'\right)}{\mathbb{E}_{y'}(\tau_{x'})} + \frac{\epsilon}{2}
\end{aligned}$$

Then, for all $K < \min\{\frac{\epsilon}{2C}, K_1\}$ holds

$$\begin{aligned}
v(K) &= \sup_{x,y\in\mathbb{R},x>y} \frac{\mathbb{E}_y\left(\gamma\left(X_{\tau_x}\right) - \int_0^{\tau_x} h\left(X_t\right)\ dt\right) - \gamma\left(y\right) - K}{\mathbb{E}_y(\tau_x)} \\
&\geq \frac{\mathbb{E}_{y'}\left(\gamma\left(X_{\tau_{x'}}\right) - \int_0^{\tau_{x'}} h\left(X_t\right)\ dt\right) - \gamma\left(y'\right) - K}{\mathbb{E}_{y'}(\tau_{x'})} \\
&\geq \frac{\mathbb{E}_{y'}\left(\gamma\left(X_{\tau_{x'}}\right) - \int_0^{\tau_{x'}} h\left(X_t\right)\ dt\right) - \gamma\left(y'\right)}{\mathbb{E}_{y'}(\tau_{x'})} - \frac{K}{C} \\
&\geq \frac{\mathbb{E}_{y'}\left(\gamma\left(X_{\tau_{x'}}\right) - \int_0^{\tau_{x'}} h\left(X_t\right)\ dt\right) - \gamma\left(y'\right)}{\mathbb{E}_{y'}(\tau_{x'})} - \frac{\epsilon}{2} \\
&> v(0) - \epsilon.
\end{aligned}$$

This shows continuity of $v$ in 0. □

Since it is needed later, we will write down a result we picked up along the way in the proof of Theorem 5.3.3 as an own lemma.

**Lemma 5.3.4.** *The threshold $\overline{x}_K$ is non-increasing in $K$ and $\underline{x}_K$ is non-decreasing in $K$.*

*Proof.* This was already shown in the proof of Theorem 5.3.3. □

Naturally, the next question that arises is if for $K \to 0$ we have $\underline{x}_K - \overline{x}_K \to 0$.

**Theorem 5.3.5.** *It holds that $\underline{x}_K \overset{K\to 0}{\to} a$ and $\overline{x}_K \overset{K\to 0}{\to} a$ for the maximizer $a$ of $f$. Hence in particular*

$$\underline{x}_K - \overline{x}_K \overset{K\to 0}{\to} 0.$$

*Proof.* We have shown that $v$ is continuous and for small enough $K \in (0,\infty)$ there are indeed two distinct roots $\underline{x}_K, \overline{x}_K$ of $f(x) = v(K)$. Further, we know that $v(0) = f(a)$. Now the continuity of $f$ and the continuity of $v$ directly yields the result. □

**Theorem 5.3.6.** *Assume the function $f$ in the maximum representation in Assumption 3.4.2 fulfils $\lim_{x\to-\infty} f(x) = \lim_{x\to\infty} f(x) = -\infty$. Then holds*

$$\lim_{K\to\infty} v(K) = -\infty.$$

*Proof.* Let $R \in \mathbb{R}$. Further, as usual, let $a$ be the maximizer of $f$. Then, the support of $(f - R)^+$, that we will denote with $L$, is compact. Hence, $\underline{b} := \inf L$, and $\overline{b} := \sup L$ are real numbers and since $\xi$ is continuous, some $(a^*, b^*) \in \arg \max_{(c,b)\in[\underline{b},\overline{b}]^2} \mathbb{E}_c(\tau_b)$ exist. Now set

$$K_1 := \mathbb{E}_{a^*}(\tau_{b^*})\left(f(a) - R\right) + 1.$$

Assume $v(K_1) \geq R$. For $v(K_1)$ there are roots $\underline{x}_{K_1}$, and $\overline{x}_{K_1}$ of $f - v(K_1)$ and an $x^*_{K_1} \in [\underline{x}_{K_1}, \overline{x}_{K_1}]$ such that

$$v_{K_1} = \frac{\mathbb{E}_{x^*_{K_1}}\left(\gamma\left(X_{\tau_{\overline{x}_{K_1}}}\right) - \int_0^{\tau_{\overline{x}_{K_1}}} h\left(X_t\right)\ dt\right) - \gamma\left(\underline{x}_{K_1}\right) - K_1}{\mathbb{E}_{x^*_{K_1}}(\tau_{\overline{x}_{K_1}})}$$

and since we assumed $v_{K_1} \geq R$ we have $x^*_{K_1}, \underline{x}_{K_1}, \overline{x}_{K_1} \in [\underline{b}, \overline{b}]$ and hence

$$\begin{aligned}
v(K_1) &= \frac{\mathbb{E}_{x^*_{K_1}}\left(\gamma\left(X_{\tau_{\overline{x}_{K_1}}}\right) - \int_0^{\tau_{\overline{x}_{K_1}}} h\left(X_t\right)\ dt\right) - \gamma\left(x^*_{K_1}\right) - K_1}{\mathbb{E}_{x^*_{K_1}}(\tau_{\overline{x}_{K_1}})} \\
&< \frac{\mathbb{E}_{x^*_{K_1}}\left(\gamma\left(X_{\tau_{\overline{x}_{K_1}}}\right) - \int_0^{\tau_{\overline{x}_{K_1}}} h\left(X_t\right)\ dt\right) - \gamma\left(x^*_{K_1}\right) - (f(a) - R)\mathbb{E}_{x^*_{K_1}}(\tau_{\overline{x}_{K_1}})}{\mathbb{E}_{x^*_{K_1}}(\tau_{\overline{x}_{K_1}})} \\
&= \frac{\mathbb{E}_{x^*_{K_1}}\left(\gamma\left(X_{\tau_{\overline{x}_{K_1}}}\right) - \int_0^{\tau_{\overline{x}_{K_1}}} h\left(X_t\right)\ dt\right) - \gamma\left(x^*_{K_1}\right)}{\mathbb{E}_{x^*_{K_1}}(\tau_{\overline{x}_{K_1}})} - (f(a) - R) \\
&= \frac{\mathbb{E}_{x^*_{K_1}}\left(\int_0^{\tau_{\overline{x}_{K_1}}} f(\overline{X}_s) ds\right)}{\mathbb{E}_{x^*_{K_1}}(\tau_{\overline{x}_{K_1}})} - (f(a) - R) \\
&\leq \frac{\mathbb{E}_{x^*_{K_1}}\left(\int_0^{\tau_{\overline{x}_{K_1}}} f(a) ds\right)}{\mathbb{E}_{x^*_{K_1}}(\tau_{\overline{x}_{K_1}})} - (f(a) - R) \\
&= R,
\end{aligned}$$

which is a contradiction to the assumption that $v(K_1) \geq R$. □

**Theorem 5.3.7.** *Assume $\alpha \neq -\infty$. Then,*

$$\lim_{K \to \infty} v(K) \leq \alpha$$

*and we have equality if with the notations $\beta_r := \lim_{x \to \infty} f(x)$ and $\beta_l := \lim_{x \to -\infty} f(x)$ at least one of the following conditions is fulfilled:*

- *$\beta_l \geq \beta_r$ and the set $\{\mathbb{E}_x(\tau_y) \mid (x, y) \in \mathbb{R}^2, x < y \leq a\}$ is unbounded.*
- *$\beta_l \leq \beta_r$ and the set $\{\mathbb{E}_x(\tau_y) \mid (x, y) \in \mathbb{R}^2, a \leq x < y\}$ is unbounded.*

*Proof.* The condition $\alpha \neq -\infty$ means that the function $f$ fulfils $\beta_r > -\infty$ or $\beta_l > -\infty$. For each $R \in \mathbb{R}$ with $R > \max\{\beta_l, \beta_r\} = \alpha$ we can use the same compactness argument as in the proof of Theorem 5.3.6 that for large enough $K$ we have $v(K) < R$.
Let $K > 0$. Now we proceed to show that $v(K) \geq \max\{\beta_l, \beta_r\} = \alpha$, if the right one of the two sets mentioned in the theorem is not bounded. We assume $\beta_l < \beta_r$ and unboundedness of the set $\{\mathbb{E}_x(\tau_y) \mid (x, y) \in \mathbb{R}^2, a \leq x < y\}$; the

other cases work in a very similar manner. Let $x^*$ be the only root of $f - \alpha$. Then, for all $y > x^*$

$$\begin{aligned}&\frac{\mathbb{E}_{x^*}\left(\gamma\left(X_{\tau_y}\right) - \int_0^{\tau_y} h\left(X_t\right)\ dt\right) - \gamma\left(x^*\right) - K}{\mathbb{E}_{x^*}\left(\tau_y\right)}\\ =&\frac{\mathbb{E}_{x^*}\left(\int_0^{\tau_y} f\left(\overline{X}_t\right)\ dt\right) - K}{\mathbb{E}_{x^*}\left(\tau_y\right)}\\ \geq&\frac{\mathbb{E}_{x^*}\left(\int_0^{\tau_y} \alpha\ dt\right) - K}{\mathbb{E}_{x^*}\left(\tau_y\right)}\\ =&\alpha - \frac{K}{\mathbb{E}_{x^*}\left(\tau_y\right)}.\end{aligned}$$

But now, due to $\{\mathbb{E}_x(\tau_y) \mid (x, y) \in \mathbb{R}^2, a \leq x < y\}$ being unbounded, for a large enough value $y$ the value $\frac{K}{\mathbb{E}_{x^*}(\tau_y)}$ gets arbitrary small, hence $v(K) \geq \alpha$. □

**Corollary 5.3.8.** *Under the set of assumptions made in the beginning of this section we have*

$$\begin{aligned}x_K^* &\overset{K\to\infty}{\to} -\infty,\\ \underline{x}_K &\overset{K\to\infty}{\to} -\infty,\\ \overline{x}_K &\overset{K\to\infty}{\to} \ \ \infty\end{aligned}$$

*and hence also*

$$x_K^* - \overline{x}_K \overset{K\to\infty}{\to} \infty$$

*and*

$$\underline{x}_K - \overline{x}_K \overset{K\to\infty}{\to} \infty,$$

*where we use for all $K > 0$ the convention, that $x_K^* := -\infty$, $\underline{x}_K := -\infty$ and $\overline{x}_K := \infty$ if these values otherwise do not exist.*

*Proof.* This is a direct consequence of the previous two theorems. □

**Remark 5.3.9.** *We do not want to go into too much detail regarding singular control here, since even the definition of singular control problems requires some technical work. To describe it very quickly: Usually monotone processes are used as allowed control strategies and in some examples of interest impulse control strategies lie dense in the set of allowed strategies. Thus, we want to remark that the previous results at least heuristically hint towards the fact, that by letting the cost converge to zero the impulse control problems for $K > 0$ may approximate the optimal singular strategy for the associated singular problem with $K = 0$ and therefore could be a tool to show that the reflection strategy at the point $a$ could be optimal for the singular problem. For a substantial analysis of such a question (albeit in the setting of portfolio optimization in a Black-Scholes-market) we refer to [CIL17], where further references on this question may be found.*

## 5.4 A Restriction to the Impulse Frequency

Throughout this section we again work with the notations and assumptions given in Subsection 4.1.1, especially let Assumption 3.4.1 regarding the functions $\gamma$ and $h$ hold let Assumption 3.4.2 regarding existence of a maximum representation hold and additionally assume that a continuous representing function as in Assumption 3.4.2 exists, assume $E = B = \mathbb{R}$ and, furthermore, let Assumption 5.3.1 hold, meaning that the functions

$$\Xi : \{(x, y \in \mathbb{R}^2 | x < y)\} \to \mathbb{R}; (a, b) \mapsto \mathbb{E}_a \left( \int_0^{\tau_b} f(\overline{X}_s) ds \right)$$

and

$$\xi : \{(x, y \in \mathbb{R}^2 | x < y)\} \to \mathbb{R}; (a, b) \mapsto \mathbb{E}_a(\tau_b)$$

are continuous. Now we fix a $c > 0$ and only allow impulse control strategies $S = (\tau_n, \zeta_n)_{n\in\mathbb{N}} \in \mathcal{S}_{\mathbb{R}}$ such that

$$\lim_{T\to\infty} \frac{1}{T} \mathbb{E}_x \left(|\{n \in \mathbb{N} \mid \tau_n \leq T\}|\right) \leq \frac{1}{c}. \tag{5.3}$$

We call the set of all such strategies $\mathcal{S}^c_{\mathbb{R}}$. Now define the value of the restricted problem by

$$v^c(x) := \sup_{S\in\mathcal{S}^c_{\mathbb{R}}} \liminf_{T\to\infty} \frac{1}{T} \mathbb{E}_x \left( \sum_{n:\tau_n \leq T} \left( \gamma\left(X^S_{\tau_n,-}\right) - \gamma\left(\zeta_n\right) - K \right) - \int_0^T h\left(X^S_s\right) ds \right). \tag{5.4}$$

For each stationary strategy $R(\tau, x)$, $x \in \mathbb{R}$, $\tau \in \mathcal{T}_x$, by the renewal reward theorem, Lemma 2.4.2, we see that

$$\begin{aligned} \frac{1}{T} \mathbb{E}_x \left(|\{n \in \mathbb{N} \mid \tau_n \leq T\}|\right) &= \frac{1}{T} \mathbb{E}_x \left( \sum_{n;\tau_n \leq T} 1 \right) \\ &\overset{T\to\infty}{\to} \frac{1}{\mathbb{E}_x(\tau)}. \end{aligned}$$

Hence, for these strategies the adaptation to the restriction can be done straightforwardly. But since it is a non-standard problem, due to the restriction of the strategies it is by no means obvious that still a threshold strategy is optimal. Now to check that indeed threshold times are optimal, we use a Lagrange type ansatz. To that end let $\lambda \geq 0$. Define the auxiliary problem

$$\begin{aligned} \tilde{v}_\lambda(x) := \sup_{S\in\mathcal{S}_{\mathbb{R}}} \liminf_{T\to\infty} \frac{1}{T} \mathbb{E}_x & \left( \sum_{n:\tau_n \leq T} \left( \gamma\left(X^S_{\tau_n,-}\right) - \gamma\left(\zeta_n\right) - K \right) - \int_0^T h\left(X^S_s\right) ds \right) \\ & - \lambda \left( \frac{1}{T} \mathbb{E}_x \left(|\{n \in \mathbb{N} \mid \tau_n \leq T\}|\right) - \frac{1}{c} \right). \end{aligned} \tag{5.5}$$

To avoid trivialities, we assume that provided there is an optimal strategy $S_0$ for $v_0$, it holds that $S_0 \notin \mathcal{S}^c_{\mathbb{R}}$. Now for each $S := (\tau_n, \zeta_n) \in \mathcal{S}_{\mathbb{R}}$ we have

$$\frac{1}{T}\mathbb{E}_x\left(|\{n \in \mathbb{N} \mid \tau_n \leq T\}|\right) = \frac{1}{T}\mathbb{E}_x\left(\sum_{n;\tau_n \leq T} 1\right)$$

and hence can write

$$\begin{aligned}&\tilde{v}_\lambda(x) - \frac{\lambda}{c}\\ &= \sup_{S \in \mathcal{S}_{\mathbb{R}}} \liminf_{T \to \infty} \frac{1}{T}\mathbb{E}_x\left(\sum_{n:\tau_n \leq T}\left(\gamma\left(X^S_{\tau_n,-}\right) - \gamma(\zeta_n) - (K+\lambda)\right) - \int_0^T h\left(X^S_s\right) ds\right).\end{aligned}$$

Since the term $\frac{\lambda}{c}$ does not depend on the strategy, we see that by Theorem 4.3.2 a threshold time is optimal for $\tilde{v}_\lambda$ if $\lambda > 0$. But on the other hand by the standard Lagrange argument we have for all $S := (\tau_n, \zeta_n) \in \mathcal{S}^c_{\mathbb{R}}$ that

$$\lim_{T \to \infty} \lambda\left(\frac{1}{T}\mathbb{E}_x\left(|\{n \in \mathbb{N} \mid \tau_n \leq T\}|\right) - \frac{1}{c}\right) \leq 0.$$

It follows that

$$\tilde{v}_\lambda \geq v^c.$$

Hence, if we can verify that an optimal threshold strategy of the restricted problem $\tilde{v}_\lambda$ lies in $\mathcal{S}^c$ and hence fulfils the restriction, then it is also optimal for $v^c$.

**Theorem 5.4.1.** *Assume the mapping given by $\lambda \mapsto x^*_\lambda$ is continuous. Then there is a threshold time that is optimal for $v^c$ given by an optimal threshold time for the auxiliary problem $\tilde{v}_{\lambda^*}$ for some $\lambda^* > 0$.*

*Proof.* By the definition of $\tilde{v}_\lambda$ we see that $\tilde{v}_\lambda - \frac{\lambda}{c}$ is just a value function of a standard ergodic impulse control problem with fixed costs $K+\lambda$. Hence for each $\lambda > 0$ there is an optimal threshold strategy $R(\tau_{\overline{x}_\lambda}, x^*_\lambda)$ (even for each $\lambda \geq 0$ if $K > 0$). Now our analysis of the dependence of the value on the cost term yields that the difference $\overline{x}_\lambda - x^*_\lambda$ continuously converges to $\infty$ if $\lambda \to \infty$, see Corollary 5.3.8. Further, we have under our assumptions due to Theorem 5.3.3 and Theorem 5.3.5 that

$$x^*_\lambda \overset{\lambda \to 0}{\to} \underline{x}_0$$

and

$$\overline{x}_\lambda \overset{\lambda \to 0}{\to} \overline{x}_0$$

where in the case $K > 0$, the values $\underline{x}_0, \overline{x}_0$ as usual denote the optimal thresholds and if $K = 0$ denote the unique maximizer of $f$. Since we assumed $R(\tau_{\overline{x}_0}, x^*_0) \notin \mathcal{S}^c$, for all small enough $\lambda > 0$ we have that $\frac{1}{\mathbb{E}_{x^*_\lambda}(\tau_{\overline{x}_\lambda})} > \frac{1}{c}$. Further, for all large enough $\lambda > 0$ we have that $\frac{1}{\mathbb{E}_{x^*_\lambda}(\tau_{\overline{x}_\lambda})} < \frac{1}{c}$. Hence the mean value theorem yields that there is a $\lambda^* > 0$ such that $\frac{1}{\mathbb{E}_{x^*_{\lambda^*}}\left(\tau_{\overline{x}_{\lambda^*}}\right)} = \frac{1}{c}$ and hence

$$\begin{aligned} J(R(\tau_{\overline{x}_0}, x_0^*)) &= J(R(\tau_{\overline{x}_0}, x_0^*)) - \lambda \left( \frac{1}{T} \mathbb{E}_x \left( |\{n \in \mathbb{N} \mid \tau_n \leq T\}| \right) - \frac{1}{c} \right) \\ &= v_\lambda \\ &\geq v^c \\ &\geq J(R(\tau_{\overline{x}_0}, x_0^*)). \end{aligned}$$

□

**Remark 5.4.2.** *As usual when applying a Lagrange type ansatz, the value $\lambda^*$ in Theorem 5.4.1 exceeds the role of a mere auxiliary object. It rather may be interpreted as some kind of artificial cost or 'shadow cost' of the restriction. So again interpreting this result in the exemplary setting of forest management this means that the restriction to the harvesting frequency from the foresters point of view is equivalent to additional fixed cost of $\lambda^*$. This is especially interesting in the case of $K = 0$. Here, usually no impulse control strategy may expected to be optimal, but with the restriction in place actually an impulse control strategy becomes optimal. This supports the (usually) heuristic arguments that a benefit of impulse control problems as a model is that they provide more practically realizable optimizers than other kind of models.*

## Chapter 6

# Mean Field Control

Mean field theory is a modelling or approximation approach with origin in physics. Here, complicated interactions in many-body-systems are replaced with some kind of an averaging influence. As justification for the reasonableness of this procedure the law of large numbers is usually adduced. As the problem how to simplify influences of a large number of sources is widespread amongst several scientific fields, it is not surprising that mean field theory is used in various branches, ranging from physics and engineering to neuroscience, artificial intelligence, statistics and game theory, to name just a few. [OS01] provides an overview over the areas of application. The introduction of mean field games into stochastic control theory goes back to [LL07] and [HMC06] and the range of applications entails many economic applications, like growth models, resource management and dynamics of opinions, see, e.g., [GLL11]. The underlying heuristic idea of this theory is that on markets with a large number of similar participants each individual's point of view may be modelled as a game against all the other participants. Now, the mean field approach is to replace the other participants' influences by an influence of one 'average' participant and therefore make the otherwise not solvable problem solvable. The solution approach then consists of the search for an equilibrium with the justification that the participant whose standpoint we take while modelling the game is similar to the other participants one and therefore no participant benefits from deviating from a strategy that is optimal, given all others (and therefore also the artificial 'average' participant) use the same strategy.
A natural question in these economic applications is how the problem changes, when instead of being in competition with each other the market participants cooperate and agree on a common strategy everyone uses. Price rigging or agreeing on certain production or sale-quotas are economic examples of this. This so called mean field control problem is also of recent interest (see, e.g., [AD11]) and often apart from the separate study of both mean field problem and game, also the comparison is of interest, see [CDL13] or [AD18].
While most mean field games and problems in control theory work with underlying continuous control problems, here, we will study both a mean field game and a mean field problem in the setting of long-term average impulse control theory.

### Structure of the section

After stating the model in Section 6.1, we first define the mean field game in Section 6.2 and then solve it both for diffusions (Section 6.3) and for Lévy processes (Section 6.4). Then, in Section 6.5 we study the mean field control problem. Section 6.6 compares the solutions of game and problem and in Section 6.7 we study two examples. The results of Section 6.3, 6.5, 6.6 and 6.7 originate in the collaborative work of Christensen, Neumann and Sohr ([CNS20]) on mean field games for diffusions.

## 6.1 Model

Again we will first work with a general strong Markov process $X$ as defined in Definition 2.1.3 the notations from Section 4.1 and the model of the controlled process given in Subsection 4.1.1 (later when working with Lévy processes and diffusions we will switch to the explicit models for these kind of processes given in the end of Subsection 4.1.1), but since we want to add an influence of an 'average market participant', the pay-off function $\gamma : E \times E \to \mathbb{R}$ will be two dimensional, twice differentiable and strictly increasing in the first component, differentiable and strictly decreasing in the second component. The first component models the pay-off's dependence on the level of $X$, the second component resembles the average level of resources of the other market participants. For the sake of simplicity we fix a $y_0 \in E$ which will be the only allowed restarting point, hence, we stay in the classical Faustmann setting that was discussed and motivated in the beginning of Section 5.2. Since a core of the mean field approach is to model the influence of a large number of similar agents by something like an average, we define the set $\mathcal{Q} \subseteq \mathcal{S}_{\{y_0\}}$ as the set of all admissible impulse strategies $Q \in S_{\{y_0\}}$, such that the controlled process $X^Q$ has a stable distribution $\Pi^Q$. We call the elements of $\mathcal{Q}$ admissible invariant strategies. To be in accordance with the interpretation of $\Pi^Q$ being a limiting distribution of the controlled process, we will write

$$\mathbb{E}[X_\infty^Q] := \int x \Pi^Q(dx).$$

Now, for each admissible strategy $R := (\tau_i, y_0)_{i \in \mathbb{N}} \in \mathcal{S}_{\{y_0\}}$, each admissible invariant strategy $Q \in \mathcal{Q}$ and each $x \in E$ we define

$$J_x(R,Q) := \liminf_{T \to \infty} \frac{1}{T} \mathbb{E}_x \left( \sum_{n : \tau_n \leq T} (\gamma(X_{\tau_n,-}^R, \mathbb{E}[X_\infty^Q]) - K) \right),$$

as the value of the strategy $R$, if the other market participants use $Q$.

## 6.2 Mean Field Game

In order to formally define a mean field equilibrium, first define the relation

$$\mathfrak{A} \subseteq \mathcal{Q} \times \mathcal{S}_{\{y_0\}}$$

by

$$Q \mathfrak{A} R \;:\Leftrightarrow R \in \operatorname*{arg\,max}_{S \in \mathcal{S}_{\{y_0\}}} J_{y_0}(S,Q).$$

We call each strategy $Q \in \mathcal{Q}$ a mean field equilibrium if

$$Q \mathfrak{A} Q.$$

Note that if all $J_{y_0}(\cdot, Q)$ had a unique optimizer, $\mathfrak{A}$ would become a mapping and mean field equilibria would be its fixed points. Hence, to be in line with the usual literature in this branch of game theory, we will also call the elements $Q \in \mathcal{Q}$ with $Q\mathfrak{A}Q$ fixed points of $\mathfrak{A}$. There are several ways to approach the search for a fixed point of $\mathfrak{A}$. Since there will be no uniqueness even in canonical examples (as we will see later), a promising approach is to shape $\mathfrak{A}$ into a mapping and then use Brouwer's fixed point theorem (or related fixed point theorems). Here, we have to find a topology $\mathcal{O}$ such that $\mathfrak{A}$ becomes a continuous mapping and maps a compact, convex set onto itself. Now one way is to directly tackle the problem, another is to first restrict the space of strategies in order to simplify the application of fixed point theorems. Here, we use the latter one. Hence, for some cases of interest, namely diffusions and Lévy processes, we will show that there is a mean field equilibrium in a threshold strategy $R(\tau_{y^g}, y_0)$ for some $y^g \in E$ with $y^g > y_0$, provided threshold strategies are optimal. This means

$$R(\tau_{y^g}, y_0)\mathfrak{A}R(\tau_{y^g}, y_0)$$

or in other words

$$R(\tau_{y^g}, y_0) \in \arg\max{}_{S\in\mathcal{S}_{\{y_0\}}} J(S, R(\tau_{y^g}, y_0))$$

where $R(\tau_{y^g}, y_0)$ is defined as the stationary threshold strategy for $y_g$, see Subsection 4.1.1. Because the restarting point $y_0$ is assumed to be fixed for this section, we will write

$$R(x) := R(\tau_x, y_0)$$

for short for all $x > y_0$ throughout the section. Here, we will find an equilibrium amongst all threshold strategies and rely on general theory or the previous chapters to justify that this is indeed a general equilibrium. To this end, we define

$$\mathfrak{a} : E\cap(y_0,\infty) \to E\cup\{\sup E\};\ y \mapsto \inf\left(\underset{\tilde{y}\in(y_0,\infty)}{\arg\max} \frac{\mathbb{E}_{y_0}\left(\gamma(X_{\tau_{\tilde{y}}}, \mathbb{E}[X_\infty^{R(y)}])\right) - K}{\mathbb{E}_{y_0}\left(\tau_{\tilde{y}}\right)}\right) \tag{6.1}$$

where we set $\mathfrak{a}(y) = \sup E$ if the arg max set in the above definition is empty. Provided the infimum is always attained, mean field equilibria in threshold strategies are precisely the ones whose thresholds are given by the fixed points of $\mathfrak{a}$, hence, we have to show that $\mathfrak{a}$ indeed has a fixed point and if so, the next interesting question is whether it is unique. To that end we will use Brouwer's fixed point theorem, hence, we have to show that $\mathfrak{a}$ is continuous and there is a non-empty closed interval $I \subseteq E$ such that $\mathfrak{a}(E) \subseteq I$. Since the analysis of $\mathfrak{a}$ heavily relies on continuity arguments, we need stronger assumptions on $X$ than just $X$ being a Markov process. Hence, in the following we will find a mean field equilibrium in threshold strategies under two sets of additional assumptions on $X$. The first one is $X$ being a diffusion, the second one $X$ being a Lévy process.

## 6.3 Mean Field Game for Diffusions

Here we show that under reasonable assumptions for diffusions, there is a mean field equilibrium in threshold strategies (or in other words a fixed point of $\mathfrak{a}$). To prove that this equilibrium is also a equilibrium in general strategies, results, e.g., from [HSZ17] may be used. Note, however, that the ordinary control problem for a diffusion $X$ with fixed restarted point $y_0$ according to Theorem 4.3.2 (under the assumptions given there) has the value

$$v = \sup_{\tau \in \mathcal{T}_{y_0}} \frac{\mathbb{E}_{y_0}\left(\gamma\left(X_\tau\right) - \int_0^\tau h\left(X_t\right)\ dt\right) - \gamma\left(y\right) - K}{\mathbb{E}_{y_0}(\tau)}. \tag{6.2}$$

Hence, all we need to have an optimal threshold strategy is that there is an optimizer $\tau$ for (6.2) (which according to Section 4.2.7 is equivalent to being the optimizer of a stopping problem with general linear costs), that is a first entry time in some set $S$, that is bounded away form $y_0$. Because then due to the continuity of $X$ the control strategy $R(\tau, y_0)$ yields the same value as the threshold strategy $R(\tau_{\bar{y}}, y_0)$ with $\bar{y} := \inf S \cap [y_0, \infty)$.

### Additional model assumptions

In this subsection, we assume $X$ to be an Itô diffusion on $\mathbb{R}_+ := (0, \infty)$ as defined in Definition 2.3.2, meaning $X$ is given by

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t$$

for a standard Brownian motion $W$ and continuous functions $\mu : \mathbb{R}_+ \to \mathbb{R}$ and $\sigma : \mathbb{R}_+ \to \mathbb{R}$ that are smooth enough to guarantee a unique strong solution $X$ to the SDE above. Further, following the explicit model for controlled diffusions given in Subsection 4.1.1, for each admissible strategy $R = (\tau_n, y_0)_{n\in\mathbb{N}}$ we define the controlled process $X^R$ by

$$X_t^R = X_0^R + \int_0^t \mu(X^R(s))\ ds + \int_0^t \sigma(X_s^R)\ dW_s - \sum_{n;\ \tau_n \le t} (X_{\tau_n-}^R - y_0).$$

Following Subsection 2.3.1, we denote the speed measure of $X$ by $M$ and its scale function by $S$. We assume existence of their (Lebesgue-)densities $\frac{dM}{d\lambda} =: m$ and $S' =: s$, that are given by

$$m(x) = \frac{2}{\sigma^2(x)} \exp\left(\int^x \frac{2\mu(y)}{\sigma^2(y)} dy\right), \quad s(x) = \exp\left(-\int^x \frac{2\mu(y)}{\sigma^2(y)} dy\right),$$

see Subsection 2.3.1 for more details.
Further, we assume $X$ to be positively recurrent with integrable stationary distribution (meaning the stationary distribution has a mean) denoted by $\mathbb{P}(X_\infty \in dx)$ (with a slight abuse of notation). General diffusion theory hence yields that $M(\mathbb{R}_+) < \infty$ and the invariant distribution is given by

$$\mathbb{P}(X_\infty \in dx) = \frac{m(x)}{M(\mathbb{R}_+)} dx$$

and

$$\mathbb{E}X_\infty := \int_{\mathbb{R}_+} x \frac{m(x)}{M(\mathbb{R}_+)} dx < \infty,$$

see Section 2.3. Also, our assumptions imply that for all $x, y \in \mathbb{R}_+$ we have $\mathbb{E}_x(\tau_y) < \infty$. Hence, the function

$$\xi : \mathbb{R}_+ \to [0, \infty); \ x \mapsto \mathbb{E}_{y_0}(\tau_x)$$

is a well defined real valued function, is even differentiable and has a representation in terms of $m$ and $s$, see Subsection 2.3.2. The assumption that $M(\mathbb{R}_+) < \infty$ implies that the boundary $\infty$ is natural and 0 is either entrance or natural (see [KT81, p.234]). Note that $X$ being a diffusion and thus having continuous sample paths and the existence of a representation of $\xi$ simplifies the structure of $\mathfrak{a}$ considerably. So under the assumptions made in this subsection we have for all $y > y_0$ that

$$\mathfrak{a}(y) = \inf \left( \underset{\tilde{y} \in (y_0, \infty)}{\arg\max} \frac{\mathbb{E}_{y_0}\left(\gamma\left(\tilde{y}, \mathbb{E}[X_\infty^{R(y)}]\right)\right) - K}{\xi(\tilde{y})} \right). \tag{6.3}$$

Now the feasibility of the search for $\mathfrak{a}$'s possible fixed points, of course, heavily depends on the properties of $\gamma$ and $\xi$. Here, we aim for reasonably simple assumptions that cover the main examples of interest while making the proofs not too theory-heavy. So, in addition to the assumptions on the process made above, we will also work with the following assumptions:

**Assumption 6.3.1.** *1. $\gamma$ is of the form $\gamma(y, z) = (y - y_0)\varphi(z)$ for a continuously differentiable and strictly decreasing function $\varphi : [z_1, z_2] \to \mathbb{R}_+$.*

*2. There is a $y_1$ such that the drift function $\mu$ is non-increasing on $(y_1, \infty)$ and strictly increasing on $[y_0, y_1]$.*

Note that it is possible to generalize these assumptions as it is done in [CNS20]. But the assumptions here already cover a wide class of applications. On one hand, Itô's lemma can be used to work with more general pay-off functions, on the other hand, the assumptions posed on $X$ include the main candidates of $X$ that are used to model growth of natural resources. Hence, the logistic diffusion that is given by

$$dX_t = aX_t(1 - bX_t)dt + \sigma(X_t)dW_t,$$

for some $a, b \in (0, \infty)$ and maybe the most common model for stochastic growth is included here. Also the connection to non-random modelling of natural growth, that was discussed in Section 5.2, is apparent. So our model may be seen as the random analogous to the Richards curve, which is the standard model for non-random natural growth in biology and is given as the solution of the ODE

$$R'(x) = a(R(x))^m - bR(x)$$

(see [Ric59, Pre19]) for some $a, b \in (0, \infty)$ and $m \in (0, 1)$.

**Finding a fixed point**

Now to find a fixed point for $\mathfrak{a}$ under our assumptions, we define the functions

$$b : (y_0, \infty) \to (0, \infty);\ y \mapsto \mathbb{E}_{y_0}[X_\infty^{R(y)}]$$

and

$$c : (0, \infty) \to (y_0, \infty);\ z \mapsto \inf \left( \underset{\tilde{y} \in (y_0, \infty)}{\arg\max} \frac{\mathbb{E}_{y_0} \left( \gamma \left( \tilde{y}, z \right) \right) - K}{\xi(\tilde{y})} \right)$$

and see that

$$\mathfrak{a} = c \circ b.$$

In the following, we will show that both $b$ and $c$ are continuous and $\overline{b((y_0, \infty))}$ is a compact set.

First we tackle $b$. For each $y > y_0$ an explicit formula for the limiting distribution of $X^{R(y)}$ is known. By standard diffusion theory, see, e.g., [HSZ17, Proposition 3.1], we have that $X_\infty^{R(y)}$ has the density

$$\pi_{y_0,y}(x) = \begin{cases} 0, & x > y \\ \kappa m(x) S[x, y], & x \in [y_0, y] \\ \kappa m(x) S[y_0, y] & x \le y_0 \end{cases}, \tag{6.4}$$

where

$$\kappa = \left( \int_{y_0}^{y} S[x, y] dM(x) + S[y_0, y] M[a, y_0] \right)^{-1}.$$

**Lemma 6.3.2.** *The function $b$ is continuously differentiable and strictly increasing.*

*Proof.* The differentiability directly follows from the definition of $\pi_{y_0,y}$. To tackle the monotonicity, we show that given an arbitrary pair $y_1, y_2 \in \mathbb{R}_+$ with $y_1 < y_2$ there is a switching point $z \in [y_0, y_1]$ such that for the corresponding densities it holds that

$$\pi_{y_0,y_1}(x) > \pi_{y_0,y_2}(x), \text{ for all } x < z,$$

and

$$\pi_{y_0,y_1}(x) \le \pi_{y_0,y_2}(x), \text{ for all } x \ge z.$$

This immediately yields the statement. To this end, we first prove that for fixed $y_0$ and $x \le y_0$ the density $\pi_{y_0,y}(x)$ is strictly decreasing in $y$: We have

$$\pi_{y_0,y}(x) = m(x) \frac{g_1(y)}{f_1(y) + f_2(y)}$$

with

$$f_1(y) = \int_{y_0}^{y} S[w, y] dM(w), \quad f_2(y) = S[y_0, y] M[0, y_0], \quad g_1(y) = S[y_0, y]$$

(since $S$ is continuous and $M$ is absolute continuous w.r.t. the Lebesgue measure, it doesn't matter if the intervals are open or closed). Using

$$f_1'(y) = S'(y)M[y_0, y], \quad f_2'(y) = S'(y)M[0, y_0], \quad g_1'(y) = S'(y)$$

we obtain

$$\begin{aligned}\frac{\partial}{\partial y}\pi_{y_0,y}(x) &= m(x)\frac{g_1'(y)(f_1(y)+f_2(y)) - (f_1'(y)+f_2'(y))g_1(y)}{(f_1(y)+f_2(y))^2} \\ &= \frac{m(x)}{(f_1(y)+f_2(y))^2}S'(y)\Bigg[\int_{y_0}^{y} S[w,y]dM(w) + S[y_0,y]M[0,y_0] \\ &\quad - S[y_0,y]M[y_0,y] - S[y_0,y]M[0,y_0]\Bigg] \\ &= \frac{m(x)}{(f_1(y)+f_2(y))^2}S'(y)\left[\int_{y_0}^{y} \underbrace{(S[w,y]-S[y_0,y])}_{=-S[y_0,w]} dM(w)\right] < 0.\end{aligned}$$

This yields $\pi_{y_0,y_1}(x) > \pi_{y_0,y_2}(x)$ for all $x \leq y_0$. It remains to consider the case $x > y_0$. We first show that for $y \in [y_0, y_1]$ and $x \in [y_0, y]$ the derivative $\frac{\partial}{\partial y}\pi_{y_0,y}(x)$ may be decomposed as

$$\frac{\partial}{\partial y}\pi_{y_0,y}(x) = m(x)h(x,y)$$

where $h(x,y)$ is non-decreasing in $x$. Indeed, for all $x \in [y_0, y]$ using the notation $g_2(y) = S[x,y]$ we get

$$\begin{aligned}\frac{\partial}{\partial y}\pi_{y_0,y}(x) &= m(x)\frac{g_2'(y)(f_1(y)+f_2(y)) - g_2(y)(f_1'(y)+f_2'(y))}{(f_1(y)+f_2(y))^2} \\ &= \frac{m(x)S'(y)}{(f_1(y)+f_2(y))^2}\Bigg[\int_{y_0}^{y} S[w,y]dM(w) + S[y_0,y]M[0,y_0] \\ &\quad - S[x,y]M[y_0,y] - S[x,y]M[0,y_0]\Bigg] \\ &= m(x)\left\{\frac{1}{(f_1(y)+f_2(y))^2}\left(\int_{y_0}^{y} S[w,x]dM(w) + S[y_0,x]M[0,y_0]\right)\right\} \\ &=: m(x)h(x,y)\end{aligned}$$

where $h(x,y)$ is indeed obviously strictly increasing in $x$. This decomposition is sufficient as it yields that

$$\pi_{y_0,y_2}(x) - \pi_{y_0,y_1}(x) = m(x)\int_{y_1}^{y_2} h(x,y)\mathrm{d}y$$

changes sign just once. Hence, using $\pi_{y_0,y_1}(x) > \pi_{y_0,y_2}(x)$ for $x < y_0$ and $\pi_{y_0,y_1}(x) = 0 < \pi_{y_0,y_2}(x)$ for $x \in (y_1, y_2)$, there exists some $z \in [y_0, y_1]$ satisfying above conditions. □

Next we show that $\overline{b((y_0,\infty))}$ is a compact set.

**Lemma 6.3.3.** *Let $R$ be an admissible impulse control strategy. Then*

$$z_1 \leq \liminf_{T\to\infty} \mathbb{E}X_T^R \leq \limsup_{T\to\infty} \mathbb{E}X_T^R \leq z_2,$$

*where $z_1 := \mathbb{E}X_\infty^r \in \mathbb{R}_+$ denotes the mean of the limiting distribution of the diffusion process $X^r$ reflected downwards in $y_0$ and $z_2 := \mathbb{E}X_\infty \in \mathbb{R}_+$ denotes the mean of the limiting distribution of the uncontrolled diffusion process $X$.*

*Proof.* First note that the expectations $z_1, z_2$ exist since we assumed $X$ to be ergodic and $X^r$ therefore is positively recurrent. The inequalities can be proved using an easy (partial) coupling argument:
We construct $X^r$ by letting it run coupled with $X^R$ until a state $\geq y_0$ is reached. Then, we reflect $X^r$ in $y_0$ downwards and let both processes run independently following their dynamics until the first time the two paths meet again. Then, we couple the paths and follow this rule. Consequently, for each $t$ and each $\omega$ we have $X_t^r(\omega) \leq X_t^R(\omega)$, proving the first inequality.
Similarly, we construct a version of the uncontrolled diffusion $X$ by running coupled with $X^R$ until the first impulse time. Then, we let both processes run independently following their dynamics until we couple them the next time the two paths meet and so on. Again, for each $t$ and each $\omega$ we have $X_t^R(\omega) \leq X_t(\omega)$. □

Next, we tackle $c$. Here, we first want to make sure that for each $z \in [z_1, z_2]$ the set

$$\operatorname*{arg\,max}_{\tilde{y}\in(y_0,\infty)} \frac{\gamma(\tilde{y}, z) - K}{\xi(\tilde{y})}$$

in the definition of $c(z)$ only has one element. To that end, we first analyse $\xi$.

**Lemma 6.3.4.** *The function $\xi(\cdot)$ is convex on $(y_1,\infty)$ and concave on $[y_0, y_1]$. If $\mu$ is strictly increasing (strictly decreasing) on one of the intervals, $\xi$ is strictly concave (strictly convex).*

*Proof.* We first show the convexity on $(y_1,\infty)$. For $r > 0$, we introduce the function $\psi_r$ via

$$\psi_r(x) = \begin{cases} \mathbb{E}_x(e^{-r\tau_{y_0}}), & x \leq y_0, \\ [\mathbb{E}_{y_0}(e^{-r\tau_x})]^{-1}, & x > y_0, \end{cases}$$

where $\tau_z$ denotes the first hitting time of level $z$. As is well known, see [BS15, II.10], $\psi_r$ is the (up to a multiplicative factor unique) strictly increasing fundamental solution to $A_X f = rf$ where $A_X$ as usual denotes the generator of $X$ extended to possibly non-bounded functions. Due to standard results for the Laplace transform, it holds that for all $x > y_0$

$$\begin{aligned} \xi(x) &= -\frac{\partial}{\partial r}\mathbb{E}_{y_0}(e^{-r\tau_x})|_{r=0} = -\frac{\partial}{\partial r}\frac{1}{\psi_r(x)}|_{r=0} = \frac{\frac{\partial}{\partial r}\psi_r(x)}{\psi_r(x)^2}|_{r=0} \\ &= \lim_{r\to 0}\frac{\psi_r(x) - 1}{r}. \end{aligned}$$

We see that it is enough to prove convexity of $\psi_r$. By [Alv03, Theorem 1] we have that with $\theta_r(x) := rx - \mu(x)$ it holds that

$$\sigma^2(x)\frac{\psi_r''(x)}{S'(x)} = 2r\int_x^\infty \psi_r(x)(\theta_r(y)-\theta_r(x))m'(y)\mathrm{d}y,$$

thus, whenever $\theta_r(y) \geq \theta_r(x)$ for all $y \geq x$, we obtain that $\psi_r''(x) > 0$. Since $\mu$ is strictly decreasing on $(y_1, \infty)$, it is obvious that $\theta_r(x)$ is strictly increasing on $(y_1, \infty)$, which yields $\psi_r''(x) > 0$, implying convexity of $\psi_r(x)$ for all $x \in (y_1, \infty)$, as desired. To show concavity on $[y_0, y_1]$ basically works just the same. Only when utilizing the equality

$$\sigma^2(x)\frac{\psi_r''(x)}{S'(x)} = 2r\int_x^\infty \psi_r(x)(\theta_r(y)-\theta_r(x))m'(y)dy,$$

we have to note that concavity is a local property and we can therefore modify $\mu$ outside $[y_0, y_1]$ to be strictly increasing on $[y_0, \infty)$. □

From this particular shape of $\xi$ we can deduce that for each $z \in [z_1, z_2]$ there is a unique critical point of

$$F_z : (y_2, \infty) \to (0, \infty); y \mapsto \frac{\gamma(y,z) - K}{\xi(y)} = \frac{(y-y_0)\phi(z) - K}{\xi(y)}$$

where $y_2 := \frac{K}{\phi(z)} + y_0$ and said critical point is a global maximum, hence equals $c(z)$. In order to do that, we fix a $z \in [z_1, z_2]$. The critical points are the roots of the function $\tilde{F}_z$ given by

$$\tilde{F}_z(y) := \varphi(z)\xi(y) - ((y-y_0)\varphi(z) - K)\xi'(y).$$

We use

$$\frac{\partial}{\partial y}\tilde{F}_z(y) = -((y-y_0)\varphi(z) - K)\xi''(y).$$

We see that

(i) $\tilde{F}_z(y_2) > 0$,

(ii) $\frac{\partial}{\partial y}\tilde{F}_z(y) \geq 0$ on the (possibly empty) interval $(y_2, y_1)$,

(iii) $\frac{\partial}{\partial y}\tilde{F}_z(y) < 0$ for all $y \geq y_1$,

(iv) $\lim_{y\to\infty}\tilde{F}_z(y) < 0$ since $\xi$ is strictly convex on $[y_1, \infty)$.

This yields that $\tilde{F}_z$ has exactly one root $y^* = y^*(z)$ on $(y_2, \infty)$. Since $\tilde{F}_z$ has the same sign as the derivative of $F_z$, this root is the only maximum of $F_z$ in $(y_2, \infty)$, hence $c(z) = y^*(z)$. This enables us to obtain $c(z)$ by help of the implicit function theorem. To that end, we have to differentiate $\tilde{F}_z$ with respect to $z$.

**Lemma 6.3.5.** *It holds* $\frac{\partial}{\partial z}\tilde{F}(y^*, z) > 0$ *for the unique root* $y^*$ *of* $\tilde{F}_z$ *in* $(y_2, \infty)$.

*Proof.* We first see that for this critical point $y^*$, since it is a root of $\tilde{F}_z$ and we have $\varphi(z) > 0$, it holds that

$$\begin{aligned} 0 =& \varphi(z)\xi(y^*) - ((y^* - y_0)\varphi(z) - K)\xi'(y^*) \\ =& \varphi(z)\left(\xi(y^*) - (y^* - y_0)\xi'(y^*) + \xi'(y^*)\frac{K}{\varphi(z)}\right), \end{aligned}$$

hence,

$$(y^* - y_0)\xi'(y^*) - \xi(y^*) = \xi'(y^*)\frac{K}{\varphi(z)}.$$

This yields

$$\begin{aligned} \frac{\partial}{\partial z}\tilde{F}_z(y^*) &= \varphi'(z)\xi(y^*) - \varphi'(z)\left((y^* - y_0)\xi'(y^*)\right) \\ &= -\varphi'(z)\xi'(y^*)\frac{K}{\varphi(z)} \\ &> 0 \end{aligned}$$

where we used that $\xi$ is strictly increasing and $\varphi$ positive and strictly decreasing. □

Now, with Lemma 6.3.5 we are able to apply the implicit function theorem on $\tilde{F}_z$ and get that the function $c$ is continuously differentiable. This additionally yields that

$$c([z_1, z_2]) \subseteq [\underline{y}, \overline{y}]$$

for some $\underline{y}, \overline{y} \in (y_0, \infty)$ with $\underline{y} < \overline{y}$.
Puzzling all the results together, we see that $\mathfrak{a} = c \circ b$ is continuous and

$$\mathfrak{a}([\underline{y}, \overline{y}]) \subseteq [\underline{y}, \overline{y}].$$

Brouwer's fixed point theorem hence yields:

**Theorem 6.3.6.** *There is a mean field equilibrium in threshold strategies.*

*Proof.* As discussed, $\mathfrak{a}$ is a continuous mapping with

$$\mathfrak{a}\left([\underline{y}, \overline{y}]\right) \subseteq [\underline{y}, \overline{y}].$$

Brouwer's fixed point theorem thus guarantees existence of a fixed point. □

## 6.4 Mean Field Game for Lévy Processes

While it was suprisingly difficult to verify existence of mean field equilibria for diffusions, it is considerably easier for Lévy processes. Now, instead of $X$ being a diffusion, we assume $X$ to be a Lévy process with $0 < \mathbb{E}(X_1) < \infty$. We again will work with the explicit model for controlled Lévy processes presented in Subsection 4.1.1 and used in Subsection 4.4.2. Further, we assume

**Assumption 6.4.1.** *$\gamma$ is for all $y, z \in \mathbb{R}$ of the form*

$$\gamma(y,z) = (\psi(y) - \psi(y_0))\varphi(z)$$

*for a continuously differentiable and strictly decreasing function $\varphi : \mathbb{R} \to \mathbb{R}_+$ and a twice differentiable, unbounded, concave and strictly increasing function $\psi : \mathbb{R} \to [\epsilon, \infty)$ for some $\epsilon > 0$. Further, assume that $\lim\limits_{x\to\infty} \frac{\psi(x)}{x} < 1$.*

Now, to make sure it suffices to search a fixed point of $\mathfrak{a}$ instead of one of $\mathfrak{A}$, we verify that a threshold strategy is optimal for the auxiliary control problems with value $v_z$ given by

$$v_z := \sup_R \liminf_{T\to\infty} \frac{1}{T}\mathbb{E}_x\left(\sum_{n:\tau_n \le T} (\gamma(X^R_{\tau_n,-}, z) - \gamma(y_0, z) - K)\right)$$

for all $z \in (0,\infty)$. This can be done by applying the solution technique developed in Subsection 4.4.2 with the alterations for a fixed restarting point discussed in Remark 4.4.3. Here, we use the usual notations for ladder height process $H$, its generator $A_H$, its drift $\tilde{\mu}_H$ and its Lévy measure $\Pi_H$ as defined in Definition 2.2.14. With that at hand, notice that we have for all $x \ge y_0$ and $z \in \mathbb{R}$

$$\begin{aligned} A_H\gamma(x,z) &= A_H(\psi(x) - \psi(y_0))\varphi(z) \\ &= \varphi(z)\left(\tilde{\mu}_H\psi'(x) + \int_0^\infty (\psi(x+y) - \psi(x))\Pi_H(dy)\right). \end{aligned}$$

Because $\psi$ is concave and strictly increasing, both $\psi'$ and $\psi(\cdot + y) - \psi(\cdot)$ are positive and strictly decreasing functions. Thus, the same is true for $A_H\gamma(\cdot, z)$ for all $z \in (0,\infty)$ and for all $\rho \in (\lim_{x\to\infty} A_H\gamma(x,z), A_H\gamma(y_0,z))$ there is exactly one solution to

$$A_H\gamma(x,z) = \rho.$$

Here, the generator is again understood to be defined for general twice continuously differentiable functions as discussed in Subsection 2.2.4. Thus, step two of the solution technique yields that for fixed $z \in \mathbb{R}$, provided this function is well defined, is given via the function

$$c : \mathbb{R} \to (y_0, \infty);\ z \mapsto \inf\left(\operatorname*{arg\,max}_{\tilde{y}\in(y_0,\infty)} \frac{\mathbb{E}_{y_0}\left(\left(\psi(X_{\tau_{\tilde{y}}}) - \psi(y_0)\right)\varphi(z)\right) - K}{\xi(\tilde{y})}\right).$$

Here, again we define $\xi : [y_0, \infty) \to (0,\infty);\ y \mapsto \mathbb{E}_{y_0}(\tau_y)$ and with Wald's equation we get for all $y \ge y_0$

$$\xi(y) = \frac{1}{\mathbb{E}(X_1)}\mathbb{E}_{y_0}(X_{\tau_y} - y_0)$$

and since $\psi$ is strictly concave this yields that for fixed $z \in (0,\infty)$ the set

$$\operatorname*{arg\,max}_{y\in(y_0,\infty)} \frac{\mathbb{E}_{y_0}\left(\left(\psi(X_{\tau_y}) - \psi(y_0)\right)\varphi(z)\right) - K}{\xi(y)}$$

has exactly one element that, since $\psi$ is unbounded, yields a positive value. Therefore, $c$ is well defined. Further, that $\psi$ is strictly concave and $\phi$ strictly decreasing, yields that $c$ is strictly decreasing as well.

**Lemma 6.4.2.** *The mapping $c$ is continuous.*

*Proof.* The mapping

$$d : \mathbb{R} \to \mathbb{R};\ z \mapsto \max_{\tilde{y} \in (y_0,\infty)} \frac{\mathbb{E}_{y_0}\left(\left(\psi(X_{\tau_{\tilde{y}}}) - \psi(y_0)\right)\varphi(z)\right) - K}{\xi(\tilde{y})}$$

is well defined due to the discussion before this lemma. Further, it is continuous, since it is strictly increasing and as a supremum over affine linear functions strictly convex. By Lemma 2.2.12 the mapping

$$\Psi : (y_0, \infty) \to \mathbb{R}; y \mapsto \mathbb{E}_{y_0}(\psi(X_{\tau_{\tilde{y}}}))$$

is continuous and by Lemma 2.2.11 the mapping $\xi$ is continuous on $(y_0, \infty)$. Hence, the function

$$D : (y_0, \infty) \times \mathbb{R};\ (\tilde{y}, z) \mapsto \frac{\mathbb{E}_{y_0}\left(\left(\psi(X_{\tau_{\tilde{y}}}) - \psi(y_0)\right)\varphi(z)\right) - K}{\xi(\tilde{y})} - d(z)$$

is continuous and strictly decreasing in the second component. Therefore, the implicit function theorem (in the version for monotone functions) yields that the function $c$ is continuous, because for each $z \in \mathbb{R}$ the value $c(z)$ is characterized as the unique solution to

$$D(c(z), z) = 0.$$

□

Now, we want to show that the function

$$b : (y_0, \infty) \to \mathbb{R}, y \mapsto \mathbb{E}(X_\infty^{R(y)})$$

is continuous and the set $\overline{\psi(b((y_0, \infty)))}$ is compact and bounded away from zero.

**Lemma 6.4.3.** *There is $C \in \mathbb{R}$ such that $\psi(b((y_0, \infty))) \subseteq [\epsilon, C]$.*

*Proof.* Fix a strategy $S = (\tau_n, y_0)_{n \in \mathbb{N}} \in \mathcal{Q}$. Now, since we use the explicit model for the controlled process given by (4.1), we have for each $\omega \in \Omega$ and for all $t \geq 0$

$$\max\{\sup_{s \leq t} X_s(\omega) - y_0, 0\} \geq \sum_{\tau_n(\omega) \leq t} (X_{\tau_n(\omega)}(\omega) - y_0)$$

and hence for the with $S$ controlled process $X^S$ that

$$X_t^S(\omega) = X_t(\omega) - \sum_{\tau_n(\omega) \leq t} (X_{\tau_n(\omega)}(\omega) - y_0) \tag{6.5}$$

$$\geq X_t(\omega) - \max\{\sup_{s \leq t} X_s(\omega) - y_0, y_0\} =: X_t^r(\omega) \tag{6.6}$$

for all $t \geq 0$. This process $X^r$ is often called the reflection of $X$ at $y_0$ and is pretty well investigated. It is obviously positive recurrent and hence possesses a stable distribution, whose mean we denote with $\mathbb{E}(X^r_\infty)$. Now since the function $\varphi$ is strictly increasing and bounded below by $\epsilon > 0$ we have

$$\varphi(\mathbb{E}(X^r_\infty)) \geq \varphi(\mathbb{E}(X^r_\infty)) \geq \epsilon.$$

□

**Lemma 6.4.4.** *The function $b$ is continuous.*

*Proof.* Let $y \in (y_0, \infty)$. The process $X^{R(y)}$ possesses a stationary distribution as shown in Lemma 4.1.2 and thus

$$\mathbb{E}(X^{R(y)}_\infty) = \frac{\mathbb{E}_{y_0}(\int_0^{\tau_y} X^{R(y)}_s ds)}{\mathbb{E}_{y_0}(\tau_y)} = \frac{\mathbb{E}_{y_0}\left(\int_0^{\tau_y} X_s ds\right)}{\xi(y)}.$$

Now with similar arguments as in Lemma 6.4.2 we see that both numerator and denominator are continuous functions in $y$, hence $b$ is continuous. □

**Theorem 6.4.5.** *There is a mean field equilibrium that is given by a threshold strategy.*

*Proof.* Lemma 6.4.4 and Lemma 6.4.2 yield that $\mathfrak{a} = c \circ b$ is continuous, Lemma 6.4.3 yields that $\mathfrak{a}$ maps a compact interval to itself. Hence, we may apply Brouwer's fixed point theorem and get that there is a fixed point of $\mathfrak{a}$. The application of our solution technique in the beginning of this subsection already proved that fixed points of $\mathfrak{a}$ also are fixed points of $\mathfrak{A}$, hence mean field equilibria. □

## 6.5 The Mean Field Control Problem

We now consider the so called mean field control problem. The economic interpretation is that the market participants (if we again imagine, say, a market of many 'similar' foresters selling their forest stand) cooperate in the sense that they agree to choose an invariant strategy $Q = (\tau_n, y_0)_{n\in\mathbb{N}} \in \mathcal{Q}$. We again work with the general model, hence with a general underlying Markov process $X$, and the notations introduced in Section 6.1. The mean-field control problem consists of maximizing

$$J_x(Q, Q) = \liminf_{T\to\infty} \frac{1}{T}\mathbb{E}_x\left(\sum_{n:\tau_n\leq T}(\gamma(X^R_{\tau_n,-}, \mathbb{E}[X^R_\infty]) - K)\right) \tag{6.7}$$

over all invariant admissible strategies $Q$. Here, we will show that by a Lagrange type approach this problem can be transformed into an ergodic impulse control problem with generalized linear cost of the type that was extensively studied in Chapter 4. This circumvents the issue that the problem (6.7) is a non-standard

stochastic control problem due to the expectation term.
Instead of restricting the process and the function $\gamma$ further as we did in Section 6.3 and Section 6.4, we make the following more general assumptions regarding the restricted problems defined by

$$v_{z,\lambda}(x) = \sup_{Q\in\mathcal{Q}} \liminf_{T\to\infty} \frac{1}{T}\mathbb{E}_x \left( \sum_{n:\tau_n\le T} (\gamma(X^Q_{\tau_n,-}, z) - K) - \int_0^T \lambda X^Q_t dt \right), \quad (6.8)$$

for all $\lambda \geq 0$, $x \in E$, $z \in \mathbb{R}$.

**Assumption 6.5.1.** *1. Assume $G := \{\mathbb{E}(X^Q_\infty) \mid Q \in \mathcal{Q}\}$ is a non-empty interval.*

2. *Assume for fixed $z \in G$ and $\lambda \in [0,\infty)$ the function $v_{z,\lambda}$ is constant, finite and a threshold strategy $R(y(\lambda, z)) \in \mathcal{Q}$ for some $y(\lambda, z) > y_0$ is optimal for $v_z$.*

3. *Assume for each $z \in E$ the mapping $L_z : [0,\infty) \to (y_0,\infty), \lambda \mapsto y(\lambda, z)$ is continuous.*

4. *Assume the mapping $b : E \to \mathbb{R}; y \mapsto \mathbb{E}(X^{R(y)}_\infty)$ is non-decreasing and continuous.*

5. *It holds $b(E) = G$.*

Note that to verify that this assumption's Part 2 holds for particular processes $X$, we may use the theory of Chapter 4. Especially for Lévy processes, the main example of this chapter, this may be done by utilizing maximum representations, also [HSZ17] and [HSZ18] may be used for diffusions. The assumption that for all pairs $(\lambda, z) \in [0,\infty) \times G$ an optimal threshold exists for $v_{z,\lambda}$ is quite restrictive. However as it can easily seen in the line of argument later on, it is not essential but only needed to apply the mean value theorem and therefore may be weakened. Regarding Part 4 of the assumption, we know from the previous section that both the diffusion and the Lévy process therein fulfil this assumption. Part 3 may be tackled with Theorem 4.3.2 and the representation

$$v_{z,\lambda} = \frac{\mathbb{E}_{y_0}\left(\gamma\left(X_{\tau_{y(\lambda,z)}}\right) - \int_0^{\tau_{y(\lambda,z)}} \lambda\left(X_t\right)\ dt\right) - \gamma\left(y_0\right) - K}{\mathbb{E}_{y_0}(\tau y(\lambda, z))} \quad (6.9)$$

obtained from (4.5). Lastly, Part 5 also holds in the two cases of the previous sections.
Now, the first step to reduce the mean field problem to a standard impulse control problem is to write the problem (6.7) as

$$\sup_{Q\in\mathcal{Q}} J_x(Q,Q) = \sup_{z\in E}\ \sup_{R \text{ with } \mathbb{E}[X^R_\infty]=z}\ \liminf_{T\to\infty} \frac{1}{T}\mathbb{E}_x \left( \sum_{n:\tau_n\le T} \gamma(X^R_{\tau_n,-}, z) \right). \quad (6.10)$$

This enables us to use a Lagrange-type approach to transform the restricted problem and thus show that there is a threshold type strategy that optimizes the mean field problem.

**Theorem 6.5.2.** *The value of the problem given by* (6.7) *is*

$$\sup_{Q\in\mathcal{Q}} J_{y_0}(Q,Q) = \sup_{y\in E, y>y_0} H(y)$$

*where for all* $y \in (y_0,\infty)\cap E$

$$H(y) := \frac{\mathbb{E}_{y_0}\gamma(X_{\tau_y}, \mathbb{E}[X_\infty^{R(y)}]) - K}{\xi(y)}$$

*and for each maximizer* $y^*$ *of* $H$*, the threshold strategy* $R(y^*)$ *is optimal for the problem* (6.7)*.*

*Proof.* By the previous discussion it suffices to fix an arbitrary $z \in G$ and to show that there is a threshold strategy as an optimizer of the problem

$$\sup_{R\in\mathcal{Q} \text{ with } \mathbb{E}[X_\infty^R]=z} \liminf_{T\to\infty} \frac{1}{T}\mathbb{E}_x\left(\sum_{n:\tau_n\le T}(\gamma(X^R_{\tau_n,-},z)-K)\right). \tag{6.11}$$

Now, with a standard Lagrange approach we consider for fixed $\lambda$ the associated unconstrained problem

$$\sup_{R\in\mathcal{Q}} \liminf_{T\to\infty} \frac{1}{T}\mathbb{E}_x\left(\sum_{n:\tau_n\le T}(\gamma(X^R_{\tau_n-},z)-K)\right) - \lambda\left(\mathbb{E}[X_\infty^R]-z\right).$$

By standard calculus it holds that for all $R\in\mathcal{Q}$

$$\mathbb{E}[X_\infty^R] = \lim_{T\to\infty}\mathbb{E}X_T^R = \lim_{T\to\infty}\mathbb{E}_x X_T^R = \lim_{T\to\infty}\frac{1}{T}\mathbb{E}_x\int_0^T X_t^R dt,$$

hence, problem (6.11) may be rewritten as

$$\sup_{R\in\mathcal{Q}} \liminf_{T\to\infty} \frac{1}{T}\mathbb{E}_x\left(\sum_{n:\tau_n\le T}(\gamma(X^R_{\tau_n,-},z)-K) - \int_0^T \lambda X_t^R dt\right) + \lambda z.$$

As $\lambda z$ is just a constant, this problem is a standard long-term average impulse control problem as (6.8) and by Assumption 6.5.1 there exists a threshold $y = y(\lambda,z)$ such that $R(y)$ is optimal. Denote by $y(z)$ a threshold value satisfying $z = \mathbb{E}[X_\infty^{R(y(z))}]$. Then, we obtain that only those values $z$ can achieve the supremum in (6.10) that satisfy $y(z)\le y(0,z)$. Indeed, assume that $y(z) > y(0,z)$. Since $\gamma(y,z)$ is strictly decreasing in the second component, we obtain

$$J_x(R(y(z)),R(y(z))) \le J_x(R(y(z)),R(y(0,z))) \le J_x(R(y(0,z)),R(y(0,z))),$$

which is a contradiction. Since

$$\lim_{\lambda\to\infty} y(\lambda,z) < y(z),$$

and $\lambda \mapsto y(\lambda, z)$ is continuous, we obtain from

$$\lim_{\lambda\to\infty} \int w\pi^{R(y(\lambda,z))}(dw) \le z \le \int w\pi^{R(y(0,z))}(dw)$$

by Assumption 6.5.1, Part 2. that there is a $\lambda_z$ such that

$$z = \int w\pi^{R(y(\lambda_z,z))}(dw).$$

As $R(y(\lambda_z, z))$ is an (unconstrained) maximizer for

$$\sup_R \liminf_{T\to\infty} \frac{1}{T}\mathbb{E}_x \left( \sum_{n:\tau_n\le T} \gamma(X^R_{\tau_n -}, z) \right) - \lambda_z \left( \mathbb{E}[X^R_\infty] - z \right)$$

and fulfils

$$z = \mathbb{E}[X_\infty^{R(y(\lambda_z,z))}],$$

it is a maximizer for (6.11) as well, proving the result. □

## 6.6 Comparison of the Solutions

In the last two sections we have discussed conditions under which both mean field game and mean field problem have threshold strategies as equilibria. Now, we will shortly discuss the relationship of the two equilibria. Hence, we assume both game and problem to have an equilibrium in threshold strategies. Further we assume $\gamma : E \times E \to \mathbb{R}$ to be continuous, strictly increasing in the first and strictly decreasing in the second argument and we assume all later occurring integrals to exist. In both cases the threshold can be obtained by maximizing the function

$$G : (y_0, \infty) \times \mathbb{R}_+; \quad (x_1, x_2) \mapsto \frac{\mathbb{E}_{y_0}\left(\gamma(X_{\tau_{x_1}}, \mathbb{E}[X_\infty^{R(x_2)}])\right) - K}{\xi(x_1)}$$

in a certain way.

Our assumptions, in particular, the assumption that $\gamma(x, z)$ is strictly decreasing in $z$ for all $x$, yields the following comparison result stating that the threshold under competition is larger than in the cooperative regime:

**Theorem 6.6.1.** *Let $\gamma(x, z)$ be strictly decreasing in $z$ for all $x$ and strictly increasing in $x$ for all $z$, further assume that the mapping*

$$b : E \to \mathbb{R}; y \mapsto \mathbb{E}(X_\infty^{R(y)})$$

*is strictly increasing. Then, for each threshold value $y^p$ being optimizer of the mean field problem and each threshold value $y^g$ being optimizer of the mean field game holds $y^p \le y^g$.*

*Proof.* Since $\gamma(x,z)$ is strictly decreasing in $z$ and $b$ is strictly increasing, we obtain that $G(x,z)$ is also strictly decreasing in $z$. Assume that there is an equilibrium threshold $y^g$ and a threshold for the mean field control problem $y^p$ such that $y^g < y^p$. Since $y^g$ is an equilibrium, we obtain that

$$G(y^p, y^g) \leq G(y^g, y^g).$$

Since $y^p$ is a solution of the mean field control problem, we have

$$G(y^g, y^g) \leq G(y^p, y^p).$$

Together with the fact that $G$ is strictly decreasing in the second argument we obtain

$$G(y^g, y^g) \leq G(y^p, y^p) < G(y^p, y^g) \leq G(y^g, y^g),$$

which is a contradiction. □

## 6.7 Examples

### 6.7.1 Classical example

To illustrate our results, we first consider a standard example of a stochastic growth model, that is fitted in the mean field model in an easy, natural way. Namely, we let the underlying process be a (Verlhust-Pearl) logistic diffusion and hence stay in a classical logistic stochastic growth model. Thus, let the uncontrolled process follow the dynamics

$$dX_t = X_t(a - bX_t)dt + \beta X_t dW_t,$$

where $a, b, \beta$ are positive constants. This diffusion is well-studied and we refer to [GVWM15] for the results we use here and further references. We set $q := 1/2 - a\rho^{-2}$ and $\rho := 2b\beta^{-2}$, and with this notations it is well-known that $X$ converges towards a unique stationary distribution if $q \leq 0$; otherwise $X$ converges to 0 a.s. We use the notation developed in Subsection 2.3.1 for speed measure and scale function. As outlined in Section 2.3 speed measure and scale function are given by the densities

$$s(x) = x^{2q-1} \exp\left(\frac{2}{\rho^2} b(x-1)\right)$$

and

$$m(x) = \frac{2}{\rho^2} x^{-2q-1} \exp\left(-\frac{2}{\rho^2} b(x-1)\right).$$

Thus, for each $y \in$ the expectation

$$\mathbb{E}[X_\infty^{R(y)}] = \int_{-\infty}^{y} x\pi_{y_0,y}(x)\mathrm{d}x$$

can be calculated as in (6.4). Assume $q < 0$ in the following. In this case, the function $\xi$ is known to be give (semi-)explicitly by:

$$\xi(y) = \frac{1}{\rho^2|q|}\left(\log\left(\frac{y}{y_0}\right) + \sum_{n=1}^{\infty}\frac{1}{(1-2q)_n}\frac{(\rho y)^n}{n} - \sum_{n=1}^{\infty}\frac{1}{(1-2q)_n}\frac{(\rho y_0)^n}{n}\right),$$

where $(u)_n = u(u+1)\cdots(u+n-1)$ denotes the Pochhammer symbol. Now for the mean field game we have to find a threshold $y^g$ such that

$$y^g = \arg\max_y \frac{\gamma\left(y, \mathbb{E}X_\infty^{R(y^g)}\right) - K}{\xi(y)}.$$

As all expressions are known explicitly, this task can be carried out straightforwardly. For illustrative purposes we use the parameters

$$q := -1,\ \beta := 1,\ y_0 := 1,\ b := 1/2,\ \rho := 2b/\beta, K := 1$$

and as pay-off function we use the function

$$\gamma : [y_0, \infty) \times (0, \infty);\ (y, z) \mapsto y/z.$$

Numerically we obtain $y^g \approx 5.20$ and the corresponding value of the game is approximately 0.2256. For the mean field problem the function

$$[y_0, \infty) \to [y_0, \infty);\ y \mapsto \arg\max_y \frac{\gamma\left(y, \mathbb{E}X_\infty^{R(y)}\right) - K}{\xi(y)} \tag{6.12}$$

has to be optimized. With the same set of parameters as above, the optimizer is $y_p \approx 4.23$ with approximated value 0.2446. As discussed in Section 6.6, the threshold $y_p$ is lower than $y^g$. The corresponding expected present volumes of wood per market participant are $\mathbb{E}X_\infty^{R(y_p)} = 1.624$ and $\mathbb{E}X_\infty^{R(y^g)} = 1.78$. Here, the value in the game is again higher than the value in the problem.

### 6.7.2 An example with no unique equilibrium

Always a question of interest in mean field theory is whether the equilibria of mean field games are unique. In our situation this is not the case in general and modifying the example above only at one point yields a case with several distinct equilibria. Thus, we work with the same process $X$ as above and also take the same set of parameters for our numerical analysis. Compared to the example above, we change the pay-off function and now use

$$\gamma : [y_0, \infty) \times (0, \infty) \to \mathbb{R}_+;\ (y, z) \mapsto \frac{y - y_0}{1 + \exp(10(z - 1.9))}.$$

Modelling the dependence on $z$ by this 'mirrored logistic function' indeed yields three equilibria which are approximately at the points $y_1^g \approx 4.6$, $y_2^g \approx 6.8$ and

$y_3^g \approx 55.5$. While the first and the last one of them are stable in the sense that when starting with a value $y_1$ in an interval around the equilibrium point the iteration used to numerically determine the equilibrium points defined by

$$y_{n+1} = \arg\max_{\tilde{y}} \frac{\gamma\left(\tilde{y}, \mathbb{E}X_\infty^{R(y_n)}\right) - K}{\xi(y)}$$

for all $n \in \mathbb{N}$ will converge to $y_1^g$ and $y_3^g$ respectively, this is not the case for $y_2^g$.

# Chapter 7

# Wrap up and Further Questions

This thesis connects long-term average impulse control problems and stopping problems with general linear costs for quite general Markov processes. A solution to the stopping problems is given and utilized to also show, under which assumptions the impulse control problem has threshold strategies as an optimizer. These theoretical findings are condensed to a quite constructive step-by-step solution technique in the case, the underlying process is a Lévy process. The main ingredient for the theoretical results is an integral type maximum representation that in the Lévy process case is developed and made explicit by use of the ladder height process. While the usefulness of this solution technique is demonstrated by solving control problems for processes with jumps in a way that exceeds the results present in the literature, the power of the theoretical findings unveils itself by enabling a study on how the control problem depends on the cost term that itself is used to show that a restriction to the average allowed number of controls per time unit in a way is equivalent to an added artificial fixed cost term. Lastly, mean field impulse control games and problems are formulated and studied.

Thus, this thesis' topics – although versatile in its origins and diverse in its applications – form a self contained unit with deep inherent connections and therefore nicely wrap up over four years of fruitful research. And nevertheless, or maybe rather precisely because of that, it raises a plethora of further questions. One of the obvious ones is the question, whether the (semi-)explicit results that are proven here for Lévy processes and are already known for diffusions may be unified to theorems on jump diffusions. The difficulty here lies in finding an accessible maximum representation.

Then, of course, there is the ever-present question whether the techniques used here to tackle one-sided problems both in control and stopping theory may be adapted to the two-sided case or maybe even more complicated types of stopping regions. First this does not seem to be the case because the application of the running maximum seems pretty tailor-made (and therefore limited to) problems with some kind of one-sided structure. On the second view, however, it could be possible to add a dimension and use representations, that in addition to the maximum also contain the running minimum. Also the connection to potential theory might be used to replace the integral containing the maximum by some other kind of additive process. The discussion of generalizations of the maximum representations in the discrete stopping problem might be a hint that approaches in this direction might be promising. However these possible generalizations of the discrete stopping problem also point out one of the major obstacles of such generalizations: In order to find the right additive process/potential (or, however, the possible generalization of the running maximum might be called) one basically has to already know the shape of the stopping region. That this phenomenon is not limited to the discrete time case may be seen in the works of Föllmer, Knispel and El Karoui (see, e.g., [FK07]) that show how super-harmonic functions may be seen as expected suprema. There, in order to find the maximum representation of a super-harmonic function, a stopping problem is solved. Thus, using these representations to solve stopping problems bears the serious risk of circular arguments, especially when one aims for con-

structive, explicit solutions.
The same difficulties, of course, also are to be expected when considering to look at multidimensional problems. However, here the so far in this wrap up a bit overlooked connection of discrete and continuous stopping problem bears some hope. Establishing this connection never utilized one dimensionality or the one-sided structure. And quite recently in [CI20] Christensen and Irle showed that (albeit in the discounted setting) by utilizing embedded monotone problems at least some explicit examples of truly multidimensional stopping problems may be solved. Now establishing the connection of control and stopping problems was not entirely free of arguments that used the one sided structure, but at least did not crucially rely on them. Therefore, examining first easy examples of multidimensional problems might be a next question worth looking at.
A last possible subsequent question worth mentioning is the connection of impulse control problems with statistics. Although the repetitive structure of long-term average problems seemingly is beneficial for statistical analysis, apart from some recent work of Claudia Strauch and Sören Christensen (see [CS19b]) there is not much literature present. And especially the characterization of control problems by stopping problems might prove itself useful when one wants to develop statistical procedures.

# Bibliography


[AD11] ANDERSSON, Daniel ; DJEHICHE, Boualem: A maximum principle for SDEs of mean-field type. In: *Applied Mathematics & Optimization* 63 (2011), Nr. 3, S. 341–356

[AD18] AURELL, Alexander ; DJEHICHE, Boualem: Mean-Field Type Modeling of Nonlocal Crowd Aversion in Pedestrian Crowd Dynamics. In: *SIAM Journal on Control and Optimization* 56 (2018), Nr. 1, 434–455. http://dx.doi.org/10.1137/17M1119196. – DOI 10.1137/17M1119196

[AH20] ALVAREZ, Luis H. R. ; HENING, Alexandru: Optimal sustainable harvesting of populations in random environments. In: *Stochastic Processes and their Applications* to appear (2020+)

[AK06] ALVAREZ, Luis H. R. ; KOSKELA, Erkki: Does risk aversion accelerate optimal forest rotation under uncertainty? In: *Journal of Forest Economics* 12 (2006), Nr. 3, S. 171–184

[AL08] ALVAREZ, Luis H. R. ; LEMPA, Jukka: On the Optimal Stochastic Impulse Control of Linear Diffusions. In: *SIAM J. Control Optim.* 47 (2008), Nr. 2, S. 703–732

[Als91] ALSMEYER, Gerold: *Erneuerungstheorie*. Stuttgart : B. G. Teuber, 1991 (Teubner-Skripten zur Mathematischen Stochastik). – ISBN 3–519–02730–5

[Alv03] ALVAREZ, Luis H. R.: On the properties of $r$-excessive mappings for a class of diffusions. In: *Ann. Appl. Probab.* 13 (2003), Nr. 4, 1517–1533. http://dx.doi.org/10.1214/aoap/1069786509. – DOI 10.1214/aoap/1069786509. – ISSN 1050–5164

[Alv04] ALVAREZ, Luis H. R.: Stochastic Forest Stand Value and Optimal Timber Harvesting. In: *SIAM J. Control Optim.* 42 (2004), Nr. 6, S. 1972–1993 (electronic)

[App09] APPLEBAUM, David: *Levy Processes and Stochastic Calculus*. 2nd. Cambridge : Cambridge University Press, 2009

[AS98] Alvarez, Luis H. R. ; Shepp, Larry A.: Optimal harvesting of stochastically fluctuating populations. In: *Journal of Mathematical Biology* 37 (1998), S. 155–177

[Asm03] Asmussen, Søren: *Applied probability and queues*. Bd. 51. Springer Science & Business Media, 2003

[BC19] Belak, Christoph ; Christensen, Sören: Utility maximisation in a factor model with constant and proportional transaction costs. In: *Finance and Stochastics* 23 (2019), 29–96. http://dx.doi.org/10.1007/s00780-018-00380-1. – DOI 10.1007/s00780–018–00380–1. – ISSN 1432–1122

[Bei98] Beibel, M.: Generalized parking problems for Lévy processes. In: *Sequential Anal.* 17 (1998), Nr. 2, 151–171. http://dx.doi.org/10.1080/07474949808836404. – DOI 10.1080/07474949808836404. – ISSN 0747–4946

[Bel57] Bellman, Richard: *Dynamic programming*. Princeton : Princeton University Press, 1957

[Bic02] Bichteler, K.: *Stochastic Integration with Jumps*. Second. Cambridge : Cambridge University Press, 2002

[BL84] Bensoussan, Alain ; Lions, Jacques-Louis: *Impulse control and quasi-variational inequalities*. Montrouge : Gauthier-Villars, 1984. – Translated from the French by J. M. Cole

[Bra01] Brazee, Richard J.: The Faustmann Formula: Fundamental to forest economics 150 years after publication. In: *Journal of Forest Science* 47 (2001), Nr. 4, S. 441–442

[BS88] Bhattacharyya, Rabindra N. ; Snyder, Donald L.: Stumpage Price Uncertainty and the Optimal Rotation of a Forest: An Application of the Sandmo Model. In: *Journal of Environmental Systems* 17 (1988), Nr. 4, S. 305–313

[BS15] Borodin, Andrei N. ; Salminen, Paavo: *Handbook of Brownian Motion – Facts and Formulae, 2nd edition, corrected printing*. Basel, Boston, Berlin : Birkhäuser, 2015

[CDL13] Carmona, René ; Delarue, François ; Lachapelle, Aimé: Control of McKean–Vlasov dynamics versus mean field games. In: *Mathematics and Financial Economics* 7 (2013), 131–166. http://dx.doi.org/10.1007/s11579-012-0089-y. – DOI 10.1007/s11579–012–0089–y. – ISSN 1862–9660

[Chr13] Christensen, Sören: Optimal decision under ambiguity for diffusion processes. In: *Mathematical Methods of Operations Research* 77 (2013), S. 207–226

[Chr14] CHRISTENSEN, Sören: On the solution of general impulse control problems using superharmonic functions. In: *Stochastic Processes and their Applications* 124 (2014), Nr. 1, S. 709–729

[Chr17] CHRISTENSEN, Sören: An effective method for the explicit solution of sequential problems on the real line. In: *Sequential Anal.* 36 (2017), Nr. 1, 2–18. http://dx.doi.org/10.1080/07474946.2016.1275314. – DOI 10.1080/07474946.2016.1275314. – ISSN 0747–4946

[CI16] CHRISTENSEN, Sören ; IRLE, Albecht: On Sequential Decision Problems with Constant Costs of Observation. Version: 2016. http://dx.doi.org/10.4018/978-1-5225-0044-5. In: ANBAZHAGAN, Neelamegam (Hrsg.): *Stochastic Processes and Models in Operations Research*. IGI Global, 2016. – DOI 10.4018/978–1–5225–0044–5, Kapitel 13, S. 218–229

[CI19] CHRISTENSEN, Sören ; IRLE, Albrecht: A general method for finding the optimal threshold in discrete time. In: *Stochastics: An International Journal of Probability and Stochastic Processes* 91 (2019), Nr. 5, S. 728–753

[CI20] CHRISTENSEN, Sören ; IRLE, Albrecht: The monotone case approach for the solution of certain multidimensional optimal stopping problems. In: *Stochastic Processes and Their Applications* 130 (2020), Nr. 4, S. 1972–1993

[CIL17] CHRISTENSEN, Sören ; IRLE, Albrecht ; LUDWIG, Andreas: Optimal portfolio selection under vanishing fixed transaction costs. In: *Advances of Applied Probability* 49 (2017), Nr. 4, S. 1116 – 1143. http://dx.doi.org/10.1017/apr.2017.36. – DOI 10.1017/apr.2017.36

[CNS20] CHRISTENSEN, Sören ; NEUMANN, Berenice A. ; SOHR, Tobias: *Competition versus Cooperation: A class of solvable mean-field impulse control problem*. 2020. – to appear

[CPT12] CISSÉ, Mamadou ; PATIE, Pierre ; TANRÉ, Etienne: Optimal stopping problems for some Markov processes. In: *The Annals of applied probability* 22 (2012), Nr. 3, S. 1243–1265

[CR61] CHOW, Y. S. ; ROBBINS, H.: A martingale system theorem and applications. In: *Proc.4th Berkeley Sympos. Math. Statist. and Prob.* 1 (1961), S. 93–104

[CRS71] CHOW, Yuan S. ; ROBBINS, Herbert ; SIEGMUND, David: *Great expectations: the theory of optimal stopping*. Boston : Houghton Mifflin, 1971

[CS17] Christensen, Sören ; Salminen, Paavo: Impulse Control and Expected Suprema. In: *Advances in Applied Probability* 49 (2017), Nr. 1, S. 238–257. http://dx.doi.org/10.1017/apr.2016.86. – DOI 10.1017/apr.2016.86

[CS19a] Christensen, Sören ; Sohr, Tobias: *A Solution Technique for Lévy Driven Long Term Average Impulse Control Problems.* 2019. – arXiv preprint arXiv:1909.10182

[CS19b] Christensen, Sören ; Strauch, Claudia: Nonparametric learning for impulse control problems. In: *arXiv preprint arXiv:1909.09528* (2019)

[CS20] Christensen, Sören ; Sohr, Tobias: *General Optimal Stopping with Linear Costs.* 2020. – arXiv preprint arXiv:2001.09470

[CST13] Christensen, Sören ; Salminen, Paavo ; Ta, Bao Q.: Optimal stopping of strong Markov processes. In: *Stochastic Processes and their Applications* 123 (2013), Nr. 3, S. 1138–1159

[CW05] Chung, Kai L. ; Walsh, John B.: *Markov Processes, Brownian Motion, and Time Symmetry.* 2nd. New York : Springer, 2005 (Grundlehren der mathematischen Wissenschaften Bd. 249)

[Dyn65] Dynkin, E. B.: *Markov Processes.* Bd. I + II. New York : Academic Press, 1965

[Ega08] Egami, Masahiko: A direct solution method for stochastic impulse control problems of one-dimensional diffusions. In: *SIAM J. Control Optim.* 47 (2008), Nr. 3, S. 1191–1218

[Fau49] Faustmann, Martin: Berechnung des Werthes, welchen Waldboden, sowie noch nicht haubare Holzbestande fur die Waldwirthschaft besitzen [Calculation of the value which forest land and immature stands possess for forestry]. In: *Allgemeine Forst- und Jagd-Zeitung* 25 (1849), S. 441–455

[FK07] Föllmer, Hans ; Knispel, Thomas: Potentials of a Markov process are expected suprema. In: *ESAIM: Probability and Statistics* 11 (2007), S. 89–101

[GLL11] Guéant, Olivier ; Lasry, Jean-Michel ; Lions, Pierre-Louis: Mean Field Games and Applications. In: *Paris-Princeton Lectures on Mathematical Finance 2010* Bd. 2003. Berlin, Heidelberg : Springer-Verlag, 2011. – ISBN 978–3–642–14660–2, S. 205–266

[GS01] Grimmett, Geoffrey R. ; Stirzaker, David R.: *Probability and Random Processes 3. Edition.* Oxford University Press, 2001

[Gut74] GUT, Allan: On the Moments and Limit Distributions of Some First Passage Times. In: *The Annals of Probability* 2 (1974), Nr. 2, S. 227–308

[Gut75] GUT, A.: On a.s. and t-mean convergence of random processes with an application to first passage times. In: *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 31 (1975), S. 333–341

[GVWM15] GIET, Jean-Sébastien ; VALLOIS, Pierre ; WANTZ-MÉZIÈRES, Sophie: The logistic S.D.E. In: *Theory Stoch. Process.* 20 (2015), Nr. 1, S. 28–62. – ISSN 0321–3900

[Hal70] HALL, W. J.: On Wald's Equations in Continuous Time. In: *Journal of Applied Probability* 7 (1970), Nr. 1, 59–68. http://www.jstor.org/stable/3212148

[HMC06] HUANG, Minyi ; MALHAMÉ, Roland P. ; CAINES, Peter E.: Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. In: *Commun. Inf. Syst.* 6 (2006), Nr. 3, S. 221–252. http://dx.doi.org/10.4310/CIS.2006.v6.n3.a5. – DOI 10.4310/CIS.2006.v6.n3.a5

[HSZ15] HELMES, Kurt L. ; STOCKBRIDGE, Richard H. ; ZHU, Chao: A Measure Approach for Continuous Inventory Models: Discounted Cost Criterion. In: *SIAM Journal on Control and Optimization* 53 (2015), Nr. 4, S. 2100–2140

[HSZ17] HELMES, Kurt L. ; STOCKBRIDGE, Richard H. ; ZHU, Chao: Continuous inventory models of diffusion type: Long-term average cost criterion. In: *The Annals of Applied Probability* 27 (2017), Nr. 3, S. 1831–1885

[HSZ18] HELMES, Kurt L. ; STOCKBRIDGE, Richard H. ; ZHU, Chao: A weak convergence approach to inventory control using a long-term average criterion. In: *Advances in Applied Probability* 50 (2018), Nr. 4, S. 1032–1074

[IP04] IRLE, A. ; PAULSEN, V.: Solving problems of optimal stopping with linear costs of observations. In: *Sequential Anal.* 23 (2004), Nr. 3, 297–316. http://dx.doi.org/10.1081/SQA-200027048. – DOI 10.1081/SQA–200027048. – ISSN 0747–4946

[Irl79] IRLE, Albrecht: Monotone Stopping Problems and Continuous Time Processes. In: *Z. Wahrscheinlichkeitstheorie verw. Gebiete* 48 (1979), S. 49–56

[JZ06] Jack, Andrew ; Zervos, Mihail: Impulse Control of One-Dimensional Itô Diffusions with an Expected and a Pathwise Ergodic Criterion. In: *Applied Mathematics and Optimization* 54 (2006), S. 71–93. http://dx.doi.org/10.1007/s00245-005-0853-y. – DOI 10.1007/s00245–005–0853–y

[Kal02] Kallenberg, Olav: *Foundations of Modern Probability*. Second. New York : Springer, 2002 (Probability and its Applications)

[Kor99] Korn, Ralf: Some applications of impulse control in mathematical finance. In: *Math. Meth. Oper. Res.* 50 (1999), S. 493–518

[KT81] Karlin, Samuel ; Taylor, Howard M.: *A second course in stochastic processes*. New York : Academic Press, Inc., 1981

[Kyp14] Kyprianou, Andreas: *Fluctuations of Lévy Processes with Applications, 2nd edition*. Berlin, Heidelberg : Springer, 2014

[LL07] Lasry, Jean-Michel ; Lions, Pierre-Louis: Mean field games. In: *Jp. J. Math.* 2 (2007), S. 229–260. http://dx.doi.org/10.1007/s11537-007-0657-8. – DOI 10.1007/s11537–007–0657–8

[LP86] Lions, P. L. ; Perthame, B.: Quasi Variational Inequalities and Ergodic Impulse Control. In: *SIAM Journal of Control and Optimization* 24 (1986), Nr. 4, S. 604–615

[MBHS78] May, R. M. ; Beddington, J. R. ; Horwood, J. W. ; Shepherd, J. G.: Exploiting natural populations in an uncertain world. In: *Mathematical Biosciences* 42 (1978), Nr. 3-4, S. 219–252

[Mer69] Merton, Robert C.: Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case. In: *The Review of Economics and Statistics* 51 (1969), Nr. 3, S. 247–257

[Mer75] Merton, Robert C.: Optimum Consumption and Portfolio Rules in a Continuous-Time Model. In: *Stochastic Optimization Models in Finance*. 1975, S. 621–661

[MØ98] Mundaca, Gabriela ; Øksendal, Bernt: Optimal stochastic intervention control with application to the exchange rate. In: *J. Math. Econom.* 29 (1998), Nr. 2, S. 225–243

[Mor02] Mordecki, Ernesto: Optimal stopping and perpetual options for Lévy processes. In: *Finance Stoch.* 6 (2002), S. 473–493

[MS07] Mordecki, Ernesto ; Salminen, Paavo: Optimal stopping of Hunt and Lévy processes. In: *An International Journal of Probability and Stochastic Processes* 79 (2007), Nr. 3-4, 233–251. https://doi.org/10.1080/17442500601100232

[NS07] Novikov, Alexander ; Shiryaev, Alexander: On a solution of the optimal stopping problem for processes with independent increments. In: *Stochastics: An International Journal of Probability and Stochastic Processes* 79 (2007), Nr. 3-4, S. 393–406

[Øks03] Øksendal, Bernt: *Stochastic Differential Equations*. Sixth edition. Heidelberg : Springer, 2003 (Universitext)

[OS01] Opper, Manfred (Hrsg.) ; Saad, David (Hrsg.): *Advanced Mean Field Methods: Theory and Practice*. Cambridge, London : The MIT Press, 2001 (Neural Information Processing Series)

[ØS05] Øksendal, Bernt ; Sulem, Agnès: *Applied stochastic control of jump diffusions*. Springer Berlin, 2005 (Universitext)

[Pau00] Paulsen, Volker: *On optimal stopping of one-dimensional symmetric diffusions with nonlinear costs of observations, Habilitation treatise*. 2000

[Pis06] Pistorius, Martijn: On Maxima and Ladder Processes for a Dense Class of Lévy Processes. In: *J. Appl. Prob.* 43 (2006), Nr. 1, S. 208–220

[PR69] Pecherskii, E. A. ; Rogozin, B. A.: On Joint Distriburions od Random Variables Associated with Fluctuations of a Process with Independent Increments. In: *Theory of Probability and its Applications* 14 (1969), Nr. 3, S. 333–341

[Pre19] Pretsch, Hans: *Grundlagen der Waldwachstumsforschung*. Berlin, Heidelberg : Springer Spektrum, 2019. – ISBN 978–3–662–58155–1

[PS06] Peškir, Goran ; Shiryaev, Albert N.: *Optimal stopping and free-boundary problems*. Basel : Birkhäuser Verlag, 2006 (Lectures in Mathematics ETH Zürich). – ISBN 978–3–7643–2419–3; 3–7643–2419–8

[PS17] Palczewski, Jan ; Stettner, Lukasz: Impulse Control Maximizing Average Cost per Unit Time: A Nonuniformly Ergodic Case. In: *SIAM Journal on Control and Optimization* 55 (2017), Nr. 2, 936–960. https://doi.org/10.1137/16M1085991

[Ric59] Richards, F. J.: A flexible growth function for empirical use. In: *Journal of Experimental Botany* (1959), Nr. 10, S. 290–301

[Rob81] Robin, Maurice: On Some Impulse Control Problems with Long Run Average Cost. In: *SIAM Journal on Control and Optimization* 19 (1981), Nr. 3, S. 333–358. http://dx.doi.org/10.1137/0319020. – DOI 10.1137/0319020

[RY05] Revuz, Daniel ; Yor, Marc: *Continuous Martingales and Brownian Motion*. 3rd. Berlin, Heidelberg, New York : Springer, 2005 (Grundlehren der mathematischen Wissenschaften Bd. 293)

[Sal85] Salminen, Paavo: Optimal Stopping of One-Dimensional Diffusions. In: *Mathematische Nachrichten* 124 (1985), Nr. 1, S. 85–101

[Sat13] Sato, Ken-Iti: *Lévy Processes and Infinitely Divisible Distributions: Revised Edition*. 2nd. Cambridge : Cambridge University Press, 2013

[Shi78] Shiryayev, Albert N.: *Optimal Stopping Rules*. New York : Springer, 1978 (Applications of Mathematics)

[Sil80] Silverstein, M. L.: Classification of coharmonic and coinvariant functions for a Lévy process. In: *The Annals of Probability* 8 (1980), Nr. 3, S. 539–575

[Sne51] Snell, Laurie J.: *Applications of martingale system theorems (Abstract)*, University of Illinois, Diss., 1951

[SS10] Shackleton, Mark B. ; Sødal, Sigbjørn: Harvesting and recovery decisions under uncertainty. In: *Journal of Economic Dynamics and Control* 34 (2010), Nr. 12, S. 2533–2546

[SSV12] Schilling, René L. ; Song, Renming ; Vondracek, Zoran: *Bernstein Functions Theory and Applications, 2nd ed.* De Gruyter, 2012

[Ste83] Stettner, Łukasz: On impulsive control with long run average cost criterion. In: *Studia Math.* 76 (1983), Nr. 3, S. 279–298

[Ste86] Stettner, Ł: On continuous time adaptive impulsive control. In: *System Modelling and Optimization: Proceedings of 12th IFIP Conference, Budapest, Hungary, September 2–6, 1985* Springer, 1986, S. 913–922

[Sur07] Surja, B. A.: An approach for solving perpetual optimal stopping problems driven by Lévy processes. In: *Stochastics: An International Journal of Probability and Stochastic Processes* 79 (2007), Nr. 3-4, S. 337–361

[Wil98] Willassen, Yngve: The stochastic rotation problem: a generalization of Faustmann's formula to stochastic forest growth. In: *Journal of Economic Dynamics and Control* 22 (1998), Nr. 4, S. 573–596

[WLK94] Woodroofe, Michael ; Lerche, Hans R. ; Keener, Robert: A generalized parking problem. In: *Statistical decision theory and related topics, V (West Lafayette, IN, 1992)*. Springer, New York, 1994, S. 523–532

[WW48] WALD, Abraham ; WOLFOWITZ, Jacob: Optimum character of the sequential probability ratio test. In: *The Annals of Mathematical Statistics* 19 (1948), Nr. 3, S. 326–339

[WW50] WALD, Abraham ; WOLFOWITZ, Jacob: Bayes solutions of sequential decision problems. In: *The Annals of Mathematical Statistics* 21 (1950), Nr. 1, S. 82–99

[Yam17] YAMAZAKI, Kazutoshi: Inventory Control for Spectrally Positive Lévy Demand Processes. In: *Math. Oper. Res.* 42 (2017), Nr. 1, 212–237. https://doi.org/10.1287/moor.2016.0801

# Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit – abgesehen von der Beratung durch meinen Betreuer Herrn Prof. Dr. Sören Christensen – nach Inhalt und Form eigenständig angefertigt habe und nur die angegebenen Hilfsmittel benutzt habe. Dabei habe ich die Regeln guter wissenschaftlicher Praxis der *Deutschen Forschungsgemeinschaft* eingehalten. Die Arbeit hat weder ganz noch in Teilen an einer anderen Stelle im Rahmen eines Prüfungsverfahrens vorgelegen. Die Arbeit ist nicht als Ganzes zur Veröffentlichung eingereicht oder veröffentlicht worden. Teile der Arbeit sind in den folgenden Artikeln zur Veröffentlichung eingereicht:

- Christensen, Sören; Sohr, Tobias: *A Solution Technique for Lévy Driven Long Term Average Impulse Control Problems.*
- Christensen, Sören; Sohr, Tobias: *General Optimal Stopping with Linear Costs.*
- Christensen, Sören; Neumann, Berenice Anne; Sohr, Tobias: *Competition versus Cooperation: A class of solvable mean field impulse control problems*

Weiter ist mir kein akademischer Grad entzogen worden.

Kiel, den 19.10.2020