On Valuation and Control in Life and Pension Insurance

Mogens Steffensen

Supervisor: Ragnar Norberg
Co-supervisor: Christian Hipp
Thesis submitted for the Ph.D. degree
Laboratory of Actuarial Mathematics
Institute for Mathematical Sciences
Faculty of Science
University of Copenhagen
May 2001

Preface

This thesis has been prepared in partial fulfillment of the requirements for the Ph.D. degree at the Laboratory of Actuarial Mathematics, Institute for Mathematical Sciences, University of Copenhagen, Denmark. The work has been carried out in the period from May 1998 to April 2001 under the supervision of Professor Ragnar Norberg, London School of Economics (University of Copenhagen until April 2000), and Professor Christian Hipp, Universität Karlsruhe.

My interest in the topics dealt with in this thesis was aroused during my graduate studies and the preparation of my master's thesis. I realized a number of open questions and wanted to search for some of the answers. This search started with my master's thesis and continues with the present thesis. Chapter 2 is closely related to parts of my master's thesis. However, the framework and the results are generalized to such an extent that it can be submitted as an integrated part of this thesis.

Each chapter is more or less self-contained and can be read independently from the rest. This prepares a submission for publication of parts of the thesis. Some parts have already been published. However, Chapters 3 and 4 build strongly on the framework developed in Chapter 2. For the sake of independence, they will both contain a brief introduction to this framework and a few motivating examples.

Acknowledgments

I wish to thank my supervisors Ragnar Norberg and Christian Hipp for their cheerful supervision during the last three years. I owe a debt of gratitude to Ragnar Norberg for shaping my understanding of and interest in various involved problems of insurance and financial mathematics and for encouraging me to go for the Ph.D. degree. Christian Hipp sharpened my understanding and I thank him for numerous fruitful discussions, in particular during my six months stay at University of Karlsruhe. A special thank goes to Professor Michael Taksar, State University of New York at Stony Brook, for his hospitality during my three months stay at SUNY at Stony Brook. Despite no supervisory duties, he took his time for many valuable discussions on stochastic control theory.

I also wish to thank my colleagues, fellow students, and friends Sebastian Aschenbrenner, Claus Vorm Christensen, Mikkel Jarbøl, Svend Haastrup, Bjarne Højgaard, Thomas Møller, Bo Normann Rasmussen, and Bo Søndergaard for interesting discussions and all their support. Finally, thanks to Jeppe Ekstrøm who, under my supervision, prepared a master's thesis from which the figures in Chapter 3 are taken.

Mogens Steffensen
Copenhagen, May 2001

Summary

This thesis deals with financial valuation and stochastic control methods and their application to life and pension insurance. Financial valuation of payment streams flowing from one party to another, possibly controlled by one of the parties or both, is important in several areas of insurance mathematics. Insurance companies need theoretically substantiated methods of pricing, accounting, decision making, and optimal design in connection with insurance products. Insurance products like e.g. endowment insurances with guarantees and bonus and surrender options distinguish themselves from traditional so-called plain vanilla financial products like European and American options by their complex nature. This calls for a thorough description of the contingent claims given by an insurance contract including a statement of its financial and legislative conditions. This thesis employs terminology and techniques fetched from financial mathematics and stochastic control theory for such a description and derives results applicable for pricing, accounting, and management of life and pension insurance contracts.

In the first part we give a survey of the theoretical framework within which this thesis is prepared. We explain how both traditional insurance products and exotic linked products can be viewed as contingent claims paid to and from the insurance company in the form of premiums and benefits. Two main principles for valuation, diversification and absence of arbitrage, are briefly described. We give examples of application of stochastic control theory to finance and insurance and relate our work to these applications.

In the second part we focus on the description and the valuation of payment streams generated by life insurance contracts. We introduce a general payment stream with payments released by a counting process and linked to a general Markov process called the index. The dynamics of the index is sufficiently general to include both traditional insurance products and various exotic unit-linked insurance products where the payments depend explicitly on the development of the financial market. An implicit dependence is present in a certain class of insurance products, pension funding and participating life insurance. However, we describe explicit forms which mimic these products, and we study them under the name surplus-linked insurance. We also introduce intervention options like e.g. the surrender and free policy options of a policy holder by allowing him to intervene in the index which determines the payments. We develop deterministic differential equations for the market value of future payments which can be used for construction of fair con-

tracts. In presence of intervention options the corresponding constructive tool takes the form of a variational inequality.

In the third part, we take a closer look at the options, in a wide sense, held by the insurance company in the cases of pension funding and participating life insurance. To these options belong the investment and redistribution of the surplus of an insurance contract or of a portfolio of contracts. The dynamics of the surplus is modelled by diffusion processes. It is relevant for the management and the optimal design of such insurance contracts to search for optimal strategies, and stochastic control theory applies. Out starting point is an optimality criterion based on a quadratic cost function which is frequently used in pension funding and which leads to optimal linear control there. This classical situation is modified in three respects: We introduce a notion of risk-adjusted utility which remedies a general problem of counter-intuitive investment strategies in connection with quadratic object functions; we introduce an absolute cost function leading to singular redistribution of surplus; and we work with a constraint on the control which leads to results which are directly applicable to participating life insurance.

Resumé

Denne afhandling beskæftiger sig med metoder til finansiel værdiansættelse og stokastisk kontrol samt deres anvendelse i livs- og pensionsforsikring. Finansiel værdiansættelse af betalingsstrømme mellem to parter, eventuelt kontrolleret af en af parterne eller begge, er vigtig i adskillige områder inden for forsikringsmatematik. Forsikringsselskaber har behov for teoretisk velfunderede metoder til prisfastsættelse, regnskabsaflæggelse, beslutningstagning og optimalt design i forbindelse med forsikringsprodukter. Forsikringsprodukter som f.eks. oplevelsesforsikringer med garantier og bonus- og genkøbsoptioner adskiller sig fra traditionelle såkaldt plain vanilla finansielle produkter som europæiske og amerikanske optioner ved deres komplekse natur. Dette nødvendiggør en grundig beskrivelse af de betingede krav indeholdt i en forsikringskontrakt, herunder en redegørelse for dens finansielle og lovgivningsmæssige betingelser. Denne afhandling anvender terminologi og teknikker hentet fra finansmatematik og stokastisk kontrolteori til en sådan beskrivelse og udleder resultater som kan anvendes til prisfastsættelse, regnskabsaflæggelse og styring af livs- og pensionsforsikringskontrakter.

I den første del gives en oversigt over den teoretiske ramme indenfor hvilken denne afhandling er lavet. Det forklares hvordan både traditionelle forsikringskontrakter og eksotiske unit link produkter kan opfattes som betingede krav til og fra forsikringsselskabet i form af præmier og ydelser. To hovedprincipper for værdiansættelse, diversifikation og fravær af arbitrage, beskrives kort. Der gives eksempler på anvendelse af stokastisk kontrolteori i finans og forsikring, og vores arbejde relateres til disse anvendelser.

I den anden del fokuseres på beskrivelsen og værdiansættelsen af betalingsstrømme genereret af livsforsikringskontrakter. Der introduceres en generel betalingsstrøm med betalinger udløst af en tælleproces og knyttet til en generel Markov proces kaldet indekset. Indeksets dynamik er tilstrækkeligt generelt til at inkludere både traditionelle forsikringsprodukter og forskellige eksotiske link forsikringsprodukter hvor betalingerne afhænger eksplicit af udviklingen af det finansielle marked. En implicit afhængighed er til stede i en særlig klasse af forsikringsprodukter, pension funding og forsikringer med bonus. Eksplicitte former som efterligner disse produkter beskrives imidlertid, og disse studeres under navnet overskudslink forsikring. Der introduceres også interventionsoptioner som f.eks. forsikringstagerens genkøbs- og fripoliceoption ved at tillade denne at intervenere i det indeks der bestemmer betalingerne. Der udvikles deterministiske differentialligninger for markedsværdien

af fremtidige betalinger som kan bruges til konstruktion af fair kontrakter. Ved tilstedeværelse af interventionsoptioner tager det tilsvarende konstruktive redskab form af en variationsulighed.

I den tredje del kigges nærmere på optionerne, i bred forstand, ejet af forsikringsselskabet i forbindelse med pension funding og livsforsikring med bonus. Til disse optioner hører investering og tilbageføring af overskud på en forsikringskontrakt eller på en portefølje af kontrakter. Dynamikken af overskuddet modelleres ved diffusionsprocesser. Det er relevant for styring og optimalt design af sådanne forsikringskontrakter at søge efter optimale strategier, og stokastisk kontrolteori er her et naturligt redskab. Udgangspunktet er et optimalitetskriterium baseret på en kvadratisk tabsfunktion, som ofte bruges i pension funding og som fører til lineær kontrol der. Denne klassiske situation er modificeret i tre henseender: Der introduceres et begreb kaldet risikojusteret nytte der afhjælper et generelt problem med ikke-intuitive investeringsstrategier som ofte opstår i forbindelse med kvadratiske objektfunktioner; der introduceres en absolut tabsfunktion som fører til singulær tilbageføring af overskud; og der introduceres en begrænsning på kontrollen som fører til resultater der er direkte anvendelige på livsforsikring med bonus.

Contents

Part I

Survey

Chapter 1

A survey of valuation and control in life and pension insurance

This thesis deals with valuation and control problems in life and pension insurance. In this introductory chapter we give a survey of notation, terminology, and methodology used throughout the thesis, and we summarize some of the results obtained. The chapter contains references to literature related to the thesis. In some cases notation and terminology used in the thesis differs from notation and terminology used in the references. We shall already here use the notation and terminology of the thesis for the sake of consistence and such that the chapter can serve to prepare the reader for the remaining chapters. This includes a partial change of notation when going from valuation problems to control problems.

1.1 Introduction

Life and pension insurance contracts are contracts which stipulate an exchange of payments between an insurance company and a policy holder. The payments are contingent on events in the life history of an insured and possibly other contingencies. Though it need not be the case, the policy holder and the insured are often the same person. By connecting payments to the life history of the insured and possibly other contingencies, a contract can be viewed as a bet on the life history and these contingencies.

Section 1.2 deals with the terms of the contract. Those terms are supposed to be comprehensible without any knowledge of probability theory, statistics, or finance. Of course, one cannot expect the policy holder to have proficiency in these areas. Though formulated in mathematical terms, Section 1.2 therefore explains the terms of the insurance contract without use of probability theoretical terminology. Valuation of the contract or the bet, on the other hand, builds on assumptions on probability laws governing the life history and the contingencies of the insurance contract. Various principles of valuation and corresponding probability laws are introduced in Section 1.3. That section also introduces intervention options of the policy holder and discusses briefly their effect on the valuation problem. The intervention options of the policy holder make up an example of a decision problem imbedded in the insurance contract. In general, the payments of an insurance contract may be rather involved and may contain various imbedded options held by both the insurance company and the policy holder. Some of the imbedded decision problems held by the insurance company are brought to the surface in Section 1.4. That section also relates these decision problems to other decision problems previously treated in the fields of finance and insurance.

1.2 Continuous-time life and pension insurance

Classical payment processes

In this section we specify payment processes in classical life and pension insurance contracts. References to the mathematics of classical life and pension insurance contracts are Gerber [25] and Norberg [54].

We let the payments stipulated in an insurance contract be formalized by a payment process $(B_t)_{t\geq 0}$, where B_t represents the accumulated payments from the policy holder to the insurance company over the time period $[0,t]$. Thus, payments that go from the insurance company to the policy holder appear in B as negative payments. We shall specify the payments in a continuous-time framework. In order to formalize the connection between payments and the life history of the insured, we introduce an indicator process $(X_t)_{t\geq 0}$. The process X indicates whether the insured is dead or not in the sense that $X_t = 0$ if the insured is alive at time t and $X_t = 1$ if the insured is dead at time t. The process X is illustrated in Figure 1.1.

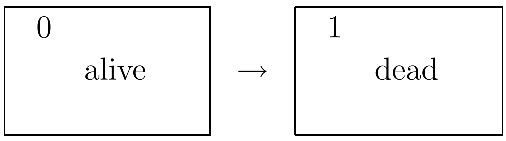

Figure 1.1: A survival model

We also introduce a counting process $(N_t)_{t\geq 0}$ counting the number of deaths of the insured (equals 0 or 1) over $[0,t]$. Note that $N = X$ in this case. Fixing a time horizon T for the insurance contract, most insurance payment processes are given by a payment process B in the form

$$B_t = \int_{0-}^{t} dB_s,\ 0 \leq t \leq T, \tag{1.1}$$

where

$$dB_t = B_0 d1_{(t\geq 0)} + b^c\left(t, X_t\right) dt - b^d\left(t, X_{t-}\right) dN_t - \Delta B_T\left(X_T\right) d1_{(t\geq T)}. \tag{1.2}$$

Here B_0 is a lump sum payment from the policy holder to the insurance company at time 0, b^c are continuous payments from the policy holder to the insurance company,

b^d is a lump sum payment at time of death from the insurance company to the policy holder, and $\Delta B\,(X_T)$ is a lump sum payment at time T from the insurance company to the policy holder. The minus signs in front of b^d and ΔB conform to the typical situation where B_0 and b^c are premiums and b^d and ΔB are benefits, all positive.

We can now specify the elements of some standard forms of benefit payment processes ($B_0 = 0$),

	$b^c\,(t, X_t)$	$b^d\,(t, X_{t-})$	$\Delta B_T\,(X_T)$
pure endowment	0	0	$1_{(X_T=0)}$
term insurance	0	$1_{(t<T, X_{t-}=0)}$	0
endowment insurance	0	$1_{(t<T, X_{t-}=0)}$	$1_{(X_T=0)}$
temporary life annuity	$-1_{(t<T, X_t=0)}$	0	0

and specify the elements of some standard forms of premium payment processes ($b^d\,(t, X_{t-}) = \Delta B_T\,(X_T) = 0$),

	B_0	$b^c\,(t, X_t)$
single premium	1	0
level premium	0	$1_{(t<T, X_t=0)}$

It is clear that the event $X_{t-} = 0$ in the indicator function of $b^d\,(t, X_{t-})$ is redundant since we know that $X_{t-} = 0$ if $dN_t = 1$. Nevertheless, we choose to expose a dependence on X_{t-} to prepare for the generalized payment processes to be introduced below.

Although the payment process in (1.1) formalizes a number of standard forms of insurances and premiums, there are a number of situations which cannot be covered by this process. One example is the situation where the premium is paid as level premium but modified such that no premium is payable during periods of disability. This modification is called premium waiver. Premium waiver and different types of disability insurances can be covered by extending X with a third state, "disabled". In general, we let $(X_t)_{t\geq 0}$ be a process moving around in a finite number of states J. The case with a disability state is illustrated in Figure 1.2.

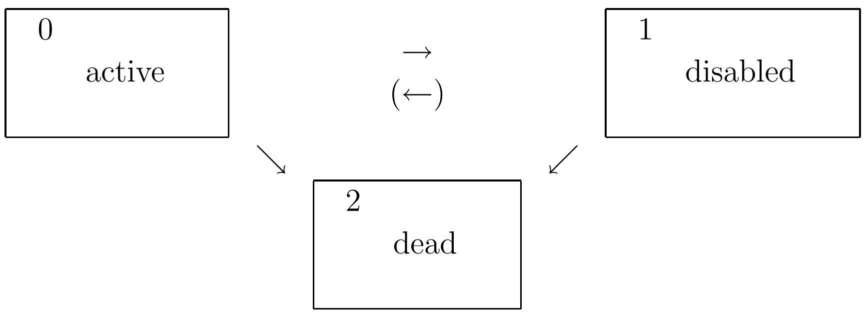

Figure 1.2: A survival model with disability and possibly recovery

Corresponding to the general J state process X, we introduce a generalized counting process N, a J-dimensional column vector where the jth entry, denoted by N^j, counts the number of jumps into state j. Correspondingly, we also generalize $b^d(t, X_{t-})$ to be a J-dimensional row vector where the jth entry, denoted by $b^{dj}(t, X_{t-})$, is the payment due upon a jump from state X_{t-} to state j at time t.

With the generalized jump process and jump payments we can specify a number of generalized insurance and premium forms. In the disability model illustrated by Figure 1.2, we can e.g. specify the elements of some standard forms of disability benefit payment processes ($B_0 = \Delta B_T(X_T) = 0$),

	$b^c(t, X_t)$	$b^d(t, X_{t-})$
disability annuity	$-1_{(t<T, X_t=1)}$	$(0, 0, 0)$
disability insurance	0	$\left(0, 1_{(t<T, X_{t-}=0)}, 0\right)$

and the elements of a premium payment process ($B_0 = b^d(t, X_{t-}) = \Delta B_T(X_T) = 0$),

	$b^c(t, X_t)$
level premium with premium waiver	$1_{(t<T, X_t=0)}$

The disability model is a three state model, i.e. $J = 3$. Models with more states are relevant for other types of insurances e.g. contracts on two lives where either member of a married pair is covered against the death of the other or multiple cause of death where payments depend on the cause of death.

Generalized payment processes

In our construction of the payment process (1.1), we have carefully distinguished between the process X, determining at any point in time the size of possible payments, and the process N, releasing these payments. So far the purpose of this distinction is not very clear since there is a one-to-one correspondence between X and N, in the sense that X determines N uniquely and vice versa. However, with the introduction of e.g. duration dependent payments or unit-linked life insurance this simple situation changes.

Duration dependent payments are payments that depend, not only on the present state of the process X, but also on the time elapsed since this state was entered. Such a construction is relevant in e.g. the disability model if the insurance company works with a so-called qualification period. Then the disability annuity does not start until the insured has qualified through uninterrupted (by activity) disability during a certain amount of time, e.g. three months. Another example is a so-called unit-linked insurance contract which is a type of contract where the payments are linked to some stock index or the value of some more or less specified portfolio.

Both in the case of duration dependent payments and in the case of unit-linked insurance, information beyond the present state of X determines the possible payment. We formalize this by allowing of a general index S to determine the possible payments. Thus, replacing X by S in the payment process (1.1), the generalized payment process becomes

$$dB_t = B_0 d1_{(t \geq 0)} + b^c(t, S_t)\, dt - b^d(t, S_{t-})\, dN_t - \Delta B_T(S_T)\, d1_{(t \geq T)}. \qquad (1.3)$$

A specification of payments is obtained by a recording of the process S and a specification of B_0 and the functions b^c, b^d, and ΔB. A special case is, of course, to let $S = X$, hereby returning to the classical payment process given by (1.2).

In the case of duration dependent payments we put $S = (X, Y)$, where Y_t equals the time elapsed since the present state X_t was entered. Considering the disability model illustrated by figure 1.2, we let Y indicate the duration of disability and see that the dynamics of Y is given by

$$dY_t = 1_{(X_t=1)}dt - Y_{t-}1_{(X_{t-}=1)}dN_t^0 - Y_{t-}1_{(X_{t-}=1)}dN_t^2,\ Y_0 = 0.$$

An example of elements of an insurance coverage with qualification period y is given by $(B_0 = b^d(t, S_{t-}) = \Delta B_T(S_T) = 0)$

	$b^c(t, S_t)$
disability annuity with qualification period	$-1_{(t<T, X_t=1, Y_t>y)}$

A simple unit-linked insurance contract can be constructed by putting $S = (X, Y)$, where Y is some stock index or the value of some portfolio. Letting G denote a guaranteed minimum payment and letting X be the simple two-state life death model illustrated in Figure 1.1, some examples of simple guaranteed unit-linked contracts are given by $(B_0 = b^c(t, S_t) = 0)$

	$b^d(t, S_{t-})$	$\Delta B_T(S_T)$
pure endowment	0	$1_{(X_T=0)} \max(Y(T), G)$
term insurance	$1_{(t<T, X_{t-}=0)} \max(Y(t), G)$	0

Once the insurance company and the policy holder have agreed on a payment process, including the recording of the index S, an insurance contract is specified. Thus, the insurance contract does not specify any assumptions as to the probability laws for the processes driving the payments, the interest rate, and other features of the market. Such assumptions are invoked by the insurer in the valuation of the payments and are needed to answer questions like: How many units of level premium with premium waiver represent a fair price to pay for a simple unit-linked endowment insurance with a guarantee?

1.3 Valuation

Valuation by diversification

This section deals with valuation of the payment streams described in Section 1.2, and we need for that purpose the probabilistic apparatus. We assume that the processes S and N are defined on a probability space $\left(\Omega, \mathcal{F}, \mathbf{F} = \{\mathcal{F}_t\}_{t\geq 0}, P\right)$.

We assume that payments are currently deposited on (or withdrawn from) a bank account that bears interest. If we denote by Z_t^0 the *(present) value* at time t of a unit deposited at time 0, we find that the (present) value at time t of a unit deposited at time s equals the amount $\frac{Z_t^0}{Z_s^0}$. We shall assume there exists a *force of interest* or *(short) rate of interest* r such that

$$dZ_t^0 = r_t Z_t^0 dt, \; Z_0^0 = 1. \tag{1.4}$$

Conforming to actuarial terminology, a present value at time t need not be $\mathcal{F}_t$-measurable. We can now speak of the present value at time t of a payment process by adding up the value of all elements in the payment process, and we get the present value at time t of the payment process B,

$$\int_{0-}^{T} \frac{Z_t^0}{Z_s^0} dB_s.$$

The value at time 0 of a payment process B is the net gain at time 0 which the insurance company faces by issuing the insurance contract. If the time of death and other contingencies determining B are known at time 0, this gain can be calculated at that point in time. To avoid gains one should balance the elements in the payment process such that the net gain equals zero,

$$\int_{0-}^{T} \frac{1}{Z_s^0} dB_s = 0. \tag{1.5}$$

However, the time of death and other contingencies determining B are in general not known at time 0. We consider these contingencies as stochastic variables defined on our probability space such that the left hand side of (1.5) becomes a stochastic variable. The question is how one should balance the elements of the insurance contract in this situation. A particular situation arises if

- the insurance company issues (or can issue) contracts on a "large" number n of insured with identically distributed payment processes $(B^i)_{i=1,\ldots,n}$,
- B^i is independent of B^j for $i \neq j$,
- the interest rate and hereby Z^0 is deterministic.

Then the law of large numbers applies and provides that the gain of the insurance portfolio per insured converges towards the expectation of the gain of an insured as the number of contracts increases, i.e.

$$\frac{1}{n} \sum_{i=1}^{n} \int_{0-}^{T} \frac{1}{Z_s^0} dB_s^i \to E\left[\int_{0-}^{T} \frac{1}{Z_s^0} dB_s\right] \text{ as } n \to \infty.$$

To avoid *systematic* gains, one should balance the elements in the payment process such that the expected net gain equals zero

$$E\left[\int_{0-}^{T} \frac{1}{Z_s^0} dB_s\right] = 0. \tag{1.6}$$

This balance equation formalizes the *principle of equivalence* which is fundamental in classical life insurance mathematics.

If one of the three assumptions above fails, the classical principle of equivalence fails as balancing tool for the payment process: If the insurance company cannot issue a large number of contracts, it makes no sense to draw conclusions from the law of large numbers; if B^i and B^j are dependent for $i \neq j$, the law of large numbers does not apply; if the interest rate is not deterministic, we cannot conclude independence between $\int_{0-}^{T} \frac{1}{Z_s} dB_s^i$ and $\int_{0-}^{T} \frac{1}{Z_s} dB_s^j$ from the independence between B^i and B^j, $i \neq j$. It should be mentioned, however, that the first two assumptions can be weakened such that they are only required to hold in a certain asymptotic sense.

So far, we have not said much about the distribution of S and N. The principle of equivalence is only based on the assumption that payment processes of different insured are identically distributed and independent. We are now going to assume that there exist deterministic piecewise continuous functions $\mu^j(t,s)$ such that N^j admits the $\mathcal{F}_t^S$-intensity process $\mu^j(t, S_t)$. This means that the $\mathcal{F}_t^S$-intensity of N is a function of t and S_t only. In the classical case where the index S is made up by the process X, a consequence of this assumption is that X is a Markov process, i.e. Markov with respect to the filtration generated by the process itself. In the set-up with a general index S this need not be the case. However, a consequence is that X is $\mathcal{F}_t^S$-Markov, i.e. Markov with respect to the filtration generated by the index S.

Consider the classical situation where $S = X$, assume that the life histories of the insured are independent, and assume that the interest rate is deterministic. We can then use the classical principle of equivalence (1.6) to determine fair premiums for the standard forms of insurance introduced in Section 1.2. Consider e.g. the calculation of a fair level premium for an endowment insurance of 1 in the survival model illustrated by Figure 1.1. Putting $\mu_t = \mu(t, 0)$, the principle of equivalence states

$$\begin{aligned} E\left[\int_{0-}^{T} \frac{1}{Z_t^0} dB_t\right] &= E\left[\int_{0}^{T} \frac{1}{Z_t^0} \left(\pi 1_{(X_t=0)} dt - 1_{(X_{t-}=0)} dN_t - 1_{(X_T=0)} d1_{(t \geq T)}\right)\right] \\ &= \pi \int_0^T e^{-\int_0^t r_s + \mu_s ds} dt - \int_0^T e^{-\int_0^t r_s + \mu_s ds} \mu_t dt - e^{-\int_0^T r_s + \mu_s ds} \\ &= 0 \Rightarrow \\ \pi &= \frac{\int_0^T e^{-\int_0^t r_s + \mu_s ds} \mu_t dt + e^{-\int_0^T r_s + \mu_s ds}}{\int_0^T e^{-\int_0^t r_s + \mu_s ds} dt}. \end{aligned}$$

Actuaries have developed a special notation for present values and expected present values of basic payment streams. An actuary would write the premium formula above on coded form (given that the insured has age x at time 0),

$$\pi = \frac{\overline{A}^{1}_{x\overline{T}|} + {}_T E_x}{\overline{a}_{x\overline{T}|}} = \frac{\overline{A}_{x\overline{T}|}}{\overline{a}_{x\overline{T}|}}.$$

The calculations for disability insurances, premium waiver and deferred benefit policies can be carried out in the same way, but they become, obviously, more involved.

Valuation by absence of arbitrage

A crucial assumption underlying the principle of equivalence was the independence between payment processes. For certain payment processes this independence comes from independence between life histories and makes sense. We shall now consider a payment process where this assumption cannot be argued to hold, and we shall reflect on a reasonable valuation principle in this situation. It is clear that if we cannot rely on the law of large of numbers, we have to rely on something else.

Arbitrage pricing theory relies on investment possibilities in a market and introduces a principle of *absence of arbitrage* i.e. avoidance of risk-free capital gains. The theory has been one of the most explosive fields of applied mathematics over the last decades. The breakthrough of this theory was the option pricing problem formulated and solved in Black and Scholes [6] and in Merton [44]. Later, rigorous mathematical content was given to notions like investment strategy, arbitrage, and completeness, and their connection to martingale theory was disclosed in Harrison and Kreps [31] and in Harrison and Pliska [32]. We shall only make a few comments on the basic theory and ask the reader to confer the cornucopia of textbooks for further insight.

A fundamental theorem in arbitrage pricing theory states that a sufficient condition which ensures that no risk-free capital gains are available is that the expected value of gains equals zero,

$$E^Q\left[\int_{0-}^T \frac{1}{Z_s^0} dB_s\right] = 0, \tag{1.7}$$

where the expectation is taken with respect to a so-called martingale measure. A martingale measure is a probability measure Q, equivalent to the measure P, such that discounted prices of traded assets are martingales under Q.

One of the simplest illustrations of (1.7) one can think of, is to find the single premium π of a payment at time T, a so-called T-claim, of a stock index Y_T where Y is included in S, i.e. $\left(b^c(t,S_t) = b^d(t,S_t) = 0\right)$

	B_0	$\Delta B_T(S_T)$
simple claim	π	Y_T

If the stock index is not available as an investment possibility, one has not necessarily enough information on the probability measure Q to say much about the price π. If the stock index is available as an investment possibility, $\frac{Y}{Z}$ is a martingale under the valuation measure Q such that

$$\begin{aligned} E^Q\left[\int_{0-}^T \frac{1}{Z_s^0} dB_s\right] &= \pi - E^Q\left[\frac{Y_T}{Z_T^0}\right] = 0 \Leftrightarrow \\ \pi &= E^Q\left[\frac{Y_T}{Z_T^0}\right] = \frac{Y_0}{Z_0^0} = Y_0. \end{aligned} \tag{1.8}$$

Why is (1.8) a reasonable result in the case where the stock index is available as an investment possibility? The issuer can, instead of investing money in the bank account, invest money in the stock index. If he does so, the gain at time T amounts to

$$\frac{Y_T}{Y_0}\pi - Y_T,$$

and, obviously, in order to avoid risk-free capital gains, we need to put $\pi = Y_0$.

Indeed, Y_T is a particularly simple T-claim, but what about a (European) option $(Y_T - K)^+$? Arbitrage pricing theory deals with general claims pricing and investment strategies which in general need to be dynamical as opposed to the static strategy above. One of the key results is that (1.7) is sufficient for absence of arbitrage.

Although the pricing formulas (1.6) and (1.7) only differ by a topscript indicating the probability measure, one should carefully note that they rely on fundamentally different properties of the risk in the payment process. Whereas (1.6) relies on diversification, (1.7) relies on absence of arbitrage in an underlying market.

The left hand side of the formulas (1.6) and (1.7) value the future payments of the contract at time 0. For various reasons one may be interested in valuing the future payments at any point of time before termination. Obviously, if one wishes to sell these future payments one must set a price. But even if one does not wish to sell the future payments, various institutions may be interested in their value. Owners of the insurance company and other investors are interested in the value of future payments for the purpose of assessing the value of the company; supervisory authorities are interested in ensuring that the payments are payable by the company and set up solvency requirements which are to be met; tax authorities are interested in the current surplus as a basis for taxation. All these parties are interested in the value of outstanding payments or *liabilities*. In a life insurance company these liabilities are called the *reserve*.

Different institutions may be interested in different notions of reserve. Whereas the payment process is (more or less) uniquely specified by (1.3), the valuation formulas (1.6) and (1.7) build on a (more or less) subjective choice of interest rate and valuation probability measure. In particular, if one does not search for information on r and Q on the financial market, values are certainly subjective and possibly not consistent with absence of arbitrage. We call a set of interest rate and Q-dynamics a valuation basis because such a set produces one version of the reserve. In Chapter 3, we introduce various special valuation bases and study the dynamics of the surplus under these.

The actual calculation of reserves, not giving rise to arbitrage possibilities, relies on the probability law of processes driving the payment process and on the underlying investment possibilities. So far we have only specified one probabilistic structure by introduction of the $\mathcal{F}_t^S$-intensities for the counting process N. We need some probabilistic structure on the index S in order to obtain applicable pricing formulas. The relation (1.7) is not worth much if we have no idea of the probabilistic structure of S. A crucial property that one is apt to rely on is the Markov property. Assuming that S is a Markov process and requiring that the reserve is

$\mathcal{F}_t^S$-Markov leads to appealing computational tools in the search for arbitrage free reserves and payment processes. This is due to the close relation between expected values of (functionals of) Markov processes and deterministic differential equations. This relation is often used in applied probability, and it is used (and partly proved) several times in this thesis.

Guaranteed payments and dividends.

In (1.7) the probability measure Q is to some extent determined by the market. However, there may be risk present in S (and N) which is not "priced by the market" and which cannot be diversified by independence of payment processes. The question is what to do with risk which is neither diversifiable nor hedgeable. A nice example is the classical case where the only investment possibility is the bank account. We now, realistically, allow the intensities of N to depend, not only on the life history of the individual insured, but also on demographic, economic, and socio-medical conditions. These conditions are formalized by the index S. Now, the individual payment processes can no longer be said to be independent. Also the assumption of deterministic interest, which is implicit in (1.6), seems unrealistic under time horizons extending to 50 years. In general, the insurance company may be unwilling to face undiversifiable and unhedgeable risk and needs to do something else.

One resolution, developed by life and pension insurance companies, is to add to the (*first order*) payment process an additional payment process of dividends conditioned on a particular performance of a policy or a portfolio of policies. This dividend can be constrained to be to the benefit of the policy holder or not, depending on the type of insurance product. If the dividends are constrained to be to the benefit of the policy holder, the first order payments must represent an overpricing, roughly speaking. In this case the dividends can be seen as a compensation for this overpricing. One way of producing first order payments which represent an overpricing is to use a certain artificial valuation basis consisting of an artificial rate of interest rate $\widehat{r}$ driving an artificial risk-free asset $\widehat{Z}_0$, and an artificial valuation measure $\widehat{Q}$, called a first order basis, to lay down payments at the time of issue. The payment process produced is called the first order payment process $\widehat{B}$, and it is determined subject to the artificial valuation formula,

$$E^{\widehat{Q}}\left[\int_{0-}^{T}\frac{1}{\widehat{Z}_s}d\widehat{B}_s\right]=0.$$

The payment process of dividends is denoted by $\widetilde{B}$. The first order payments and the dividends make up the total payments B stemming from the contract,

$$B=\widehat{B}+\widetilde{B}.$$

Now the problem of setting fair payments is translated to the problem of allotting fair dividends. At the end of the day, the insurance company needs to make up its mind about the assessment of (the value of) non-diversifiable and non-hedgeable risk and balance dividends by the corresponding equivalence relation

$$E^Q\left[\int_{0-}^T \frac{1}{Z_s^0} dB_s\right] = 0. \tag{1.9}$$

However, by the introduction of dividend payments, it is to some extent possible for the insurance company to transfer a part of the risk from the insurance company to the policy holders. Hereby the insurance company is less exposed to risk than in a situation without dividends, of course depending on how these are determined. We shall not go deeper into the interpretation of dividend distribution as a risk management instrument now but content ourselves with a simple illustrative example.

Assume e.g. that dividends are only paid out at time T and that this dividend payment is a function of the performance of the first order payments. Then, by introducing the process $\int_{0-}^t \frac{Z_t}{Z_s} d\widehat{B}_s$ in the index S, we can define for a some function f, $\left(\widetilde{B}_0 = \widetilde{b}^c(t, S_t) = \widetilde{b}^d(t, S_{t-}) = 0\right)$,

	$\Delta\widetilde{B}_T(S_T)$
terminal dividends	$-f\left(\int_{0-}^T \frac{Z_T^0}{Z_s^0} d\widehat{B}_s\right)$

This dividend plan leads to a total gain of

$$\int_{0-}^T \frac{1}{Z_s^0} dB_s = \int_{0-}^T \frac{1}{Z_s^0} d\widehat{B}_s - \frac{1}{Z_T^0} f\left(\int_{0-}^T \frac{Z_T^0}{Z_s^0} d\widehat{B}_s\right).$$

If e.g. the insurance company is allowed to choose as function f the identity function the gain is zero and all risk is transferred to the policy holder. This is, of course, an extreme (and extremely uninteresting) case, but it illustrates what is meant by transferring risk to the policy holder. Another function f, which moreover ensures that dividends are to the benefit of the policy holder, is

$$f\left(\int_{0-}^T \frac{Z_T^0}{Z_s^0} d\widehat{B}_s\right) = q\left(\int_{0-}^T \frac{Z_T^0}{Z_s^0} d\widehat{B}_s\right)^+,$$

where q is a constant. In the case of no constraints on the dividends, we shall speak of pension funding, and in the case where dividends are constrained to be to the policy holder's benefit, we shall speak of participating life insurance.

Chapter 3 deals with valuation bases, surplus, and dividends. The relation between expected values and deterministic differential equations gives a constructive tool for calculation of fair strategies for investment and repayment of surplus through dividends. Numerical results shall illustrate this tool.

Valuation under intervention options

It is implicitly assumed in all valuation formulas above that the insurance company and the policy holder have no influence on the performance of the insurance contract, hereunder the dynamics of the index. In practice, there are a number of intervention options that may (or may not) affect the valuation of payments.

One example of an intervention option is the exercise option of an American option. The exercise option allows the owner of an American option to exercise the contract at any point in time up to the expiration date T. For a life insurance contract the most important intervention option is probably the surrender option of the policy holder. Holding this option, he can at any point in time t up to T close the contract and convert all future payments into an immediate payment of the surrender value. Also the issuer of an insurance contract may hold intervention options. E.g. the bankruptcy option of the owners of the insurance company can be considered as an intervention option held by the insurance company.

It turns out that a very convenient way of modelling these intervention options is to allow the policy holder and/or the insurance company to intervene in the index S in some specified way. This enables us to capture exactly the types of intervention options in which we are interested. Disregarding all intervention options held by the insurance company but taking into consideration intervention options of the policy holder, arbitrage arguments lead to a valuation formula on the form

$$\sup_{I\in\mathcal{I}} E^{Q^I}\left[\int_{0-}^{T}\frac{1}{Z_s^{0I}}dB_s^I\right], \tag{1.10}$$

where topscript I indicates that the quantity is dependent on a certain admissible intervention strategy taken by the policy holder and the supremum is taken over all admissible intervention strategies.

The results building on optimal intervention represent one approach to the problem of valuation, taking into account intervention options. This approach expects the policy holder to behave financially optimal. Whereas this assumption may be reasonable for short-term pure financial contracts, a simple example demonstrates that one should follow this approach with care in connection with long-term insurance contracts.

Consider an insured holding a term insurance and assume the possibility of starting to smoke with an increasing effect on mortality. We model this situation by introducing an index indicating whether the insured is a smoker or not. Before starting to smoke, the mortality is $\mu(t,0)$ and after starting to smoke it increases to $\mu(t,1)$. We disregard the possibility of stopping smoking.

The question is now on which mortality rate should the insurance company base the premium calculation and the reservation if the new customer tells that he is a non-smoker. The insured can advance his death occurrence and hereby maximize his expected benefit payments by starting smoking, and, in fact, the valuation formula (1.10) tells the insurance company to use the high mortality rate assuming that he does so immediately. However, the insured may take other things into consideration than the benefits from the insurance contract and conclude that, after all, it is optimal not to advance death whereafter he chooses not to start smoking.

This is a toy example which, nevertheless, shows that the insurance company should use a valuation formula based on optimization with care. The policy holder may have other objectives than increasing the value of his insurance contract, and when it comes to payments linked to his life history, he probably will. Nevertheless, in Chapter 4 we give mathematical content to intervention options and work with the valuation formula (1.10). In the case of no intervention options, the relation between expected values and deterministic differential equations is demonstrated in Chapter 2. In presence of intervention option, (1.10) relates to a so-called quasi-variational inequality. This is a constructive tool for determining fair contracts under intervention options and is derived in Chapter 4.

1.4 Control

In Section 1.3 we discussed valuation of payment streams. At the end of that section we unveiled one control problem imbedded in the payment process, namely the control by intervention of the policy holder. Furthermore we constructed guaranteed payments and dividends, and we argued that this construction of payments allows the insurance company to transfer risk to the policy holder. In fact, the design of the dividend payment process $\widetilde{B}$ can be considered as a genuine control problem on the part of the insurance company. E.g. one could simply formulate an objective of risk reduction in some sense and then look for an optimal dividend process.

We shall now consider a framework frequently used in finance and insurance decision problems. Within this framework we recall some classical decision problems in finance and non-life insurance, and we consider how the decision problem of the life insurance company also has been approached within this framework in the literature on life and pension insurance. The approach to the life insurance decision problem studied in Chapters 5 and 6 is a modification of this classical framework. At the end of this section we explicate this. We will now partially change notation in order to conform to Chapters 5 and 6.

Consider the *wealth* (*reserve*, *value*, or *surplus*) of an *agent* (*consumer* or *insurance company*) with the following dynamics

$$\begin{aligned} dX(t) &= \alpha(\theta(t), X(t))\,dt + \sigma(\theta(t), X(t))\,dW(t) - dU(t), \qquad (1.11)\\ X(0-) &= x_0, \end{aligned}$$

where x is the initial wealth, and W is a standard Brownian motion defined on a probability space $\left(\Omega, \mathcal{F}, \mathbf{F} = \{\mathcal{F}_t\}_{t\geq 0}, P\right)$. The parameter θ is chosen by the agent to balance drift and diffusion in the wealth process through the functions α and σ. In this section, θ will represent the proportion of wealth invested in a risky asset and/or a parameter indicating the extent of cover by some type of reinsurance. $U(t)$ is the amount withdrawn from the wealth either for *consumption* or as *dividend distribution* until time t.

The agent needs an objective for his decisions, and he wishes to choose (θ, U) so as to maximize

$$E\left[\int_0^\tau \upsilon\left(t, X\left(t\right), dU\left(t\right), dt\right) + \Upsilon\left(\tau, X\left(\tau\right)\right)\right], \tag{1.12}$$

for some utility functions υ and Υ and some stopping time τ. Note that the compact notation for current utility only makes sense for particular functions υ. Examples will be given below. We use the terms utility and disutility functions for general (not necessarily increasing, concave/concave and continuously differentiable) reward and cost functions, respectively.

Control in finance

Suppose there is a market in which two assets are traded continuously. One asset is the bank account introduced in (1.4), Z^0. The other asset is a risky asset denoted by Z^1 with a price process modelled as a geometric Brownian motion with drift, i.e.

$$\begin{aligned} dZ^0\left(t\right) &= rZ^0\left(t\right)dt, \\ Z^0\left(0\right) &= 1, \\ dZ^1\left(t\right) &= \mu Z^1\left(t\right)dt + \sigma Z^1\left(t\right)dW_t, \\ Z^1\left(0\right) &= z^1. \end{aligned} \tag{1.13}$$

Now the wealth process of an agent following the proportional investment strategy θ and the consumption strategy U can be shown to have the dynamics given by (1.11), with

$$\begin{aligned} \alpha\left(\theta\left(t\right), X\left(t\right)\right) &= \left(r + \theta\left(t\right)\left(\mu - r\right)\right)X\left(t\right), \\ \sigma\left(\theta\left(t\right), X\left(t\right)\right) &= \theta\left(t\right)\sigma X\left(t\right). \end{aligned}$$

One class of problems are so-called investment-consumption problems where one requires consumption to be positive and absolutely continuous with respect to the Lebesgue measure such that $u\left(t\right) = \frac{dU(t)}{dt} \geq 0$ exists and where utilities are typically given by

$$\begin{aligned} \upsilon\left(t, X\left(t\right), dU\left(t\right), dt\right) &= e^{-\gamma t}\upsilon\left(u\left(t\right)\right)dt, \\ \Upsilon\left(\tau, X\left(\tau\right)\right) &= e^{-\gamma\tau}\Upsilon\left(X\left(\tau\right)\right). \end{aligned}$$

Merton initiated the study of this problem in Merton [42] and [43] and found explicit solutions for some particular utility functions. Primary examples of utility functions in this case are the logarithmic and the power functions.

Another class of problems are hedging problems connected with a contingent claim $Y\left(\tau\right)$. One approach is to let the consumption be fixed at zero and let the utility of $X\left(\tau\right)$ depend on $Y\left(\tau\right)$ such that for e.g. a quadratic loss function,

$$\begin{aligned} \upsilon\left(t, X\left(t\right), dU\left(t\right), dt\right) &= 0, \\ \Upsilon\left(\tau, X\left(\tau\right)\right) &= -\left(X\left(\tau\right) - Y\left(\tau\right)\right)^2. \end{aligned}$$

In this way deviations from the τ-claim are punished and an optimal hedging strategy can be searched for. In this hedging problem, one may also consider the starting point of our wealth, x_0, as a decision variable and speak of the optimal x_0 as some kind of price of Y. This approach to optimal investment strategies (and prices) is called mean-variance hedging. Of course, the idea of mean-variance hedging is not restricted to the simple market given by 1.13 but can studied for general (non-Markovian) markets as well, see e.g. Schweizer [61].

Control in non-life insurance

Consider a non-life insurance company receiving premiums continuously and paying out claims. The company balances its gains by an extent $\theta(t)$ of cover of some type of reinsurance where the company pays premiums continuously and receives some compensation for claims. Furthermore, the company decides to pay out dividends to share-holders. We emphasize that dividend here is the share of profits paid to share-holders as opposed to the dividend introduced in Section (1.3) which goes to the policy holders. If we denote the rate of premiums (net of the reinsurance premium) by $\pi(\theta(t))$, the number of claims received up to time t by $N(t)$, the time for occurrence of claim number i by τ_i, and the size of the ith claim (net of the reinsurance compensation) by $Y_i(\theta(t))$, then X, the company's reserve (net of reinsurance payments) at time t is given by

$$X(t) = x_0 + \int_0^t \pi(\theta(s))\,ds - \sum_{i=1}^{N(t)} Y_i(\theta(\tau_i)) - U(t).$$

We remark that the notion of reserve in non-life insurance has a different meaning than in life insurance. If $N(t)$ is a Poisson process with intensity λ and the claims are i.i.d., then X can be approximated by the process given by (1.11) with

$$\begin{aligned} \alpha(\theta(t), X(t)) &= \pi(\theta(t)) - \lambda E[Y_i(\theta(t))], \\ \sigma(\theta(t), X(t)) &= \sqrt{\lambda E[Y_i^2(\theta(t))]}. \end{aligned}$$

Usually one lets τ in (1.12) be the time of ruin, i.e. the first time the reserve hits zero, and wishes to maximize discounted dividends,

$$v(t, X(t), dU(t), dt) = e^{-\gamma t} dU(t).$$

One argument for using the identity function as utility function is that the value of the firm may be represented by expected discounted dividends. This argument, however, seems criticizable since it is by no means clear which discount factor and which measure to use. For optimal proportional reinsurance see Højgaard and Taksar [34]. In the non-life insurance model above, we have disregarded capital gains, but optimal control of reinsurance can, of course, be combined with optimal investment.

Control in life and pension insurance

The literature on control in life and pension insurance has until now concentrated primarily on control of pension funds. For references to literature on control of pension funds, see Cairns [12], which is partly a survey article gathering results of several authors. The control parameters are usually a proportion in risky assets and/or the level of premiums/benefits. The institutional conditions for pension funds may be rather involved, and it is by no means clear how the objectives of the fund manager, the employer (pays the premium), and the employed (receives the benefits) should be reflected in the objective function of the control problem. Here, we shall briefly demonstrate how pension funds are modelled and controlled in continuous-time as exposed in Cairns [12].

Assume that the pension fund receives premiums and pays benefits. The infinitesimal net outgo of the fund is normally distributed with expected value $u(t)\,dt$ and variance $\beta^2(t)\,dt$, independently of the financial market. The mean rate of the net outgo is controllable by the fund manager who can adjust this according to the performance of the fund. The employer and the employed, respectively, experience this control by changes in the premium level in the case of defined benefits and in the benefit level in the case of defined contributions, respectively. We assume that the money in the fund is invested in the market described by (1.13). Then the dynamics of the fund is given by (1.11) with

$$\begin{aligned}\alpha(\theta(t),X(t)) &= (r+\theta(t)(\alpha-r))X(t),\\ \sigma(\theta(t),X(t)) &= \sqrt{\theta^2(t)\,\sigma^2X^2(t)+\beta^2(t)},\\ dU(t) &= u(t)\,dt.\end{aligned}$$

So far the problem only differs from the investment-consumption problem of finance by the term β. This indicates a connection between investment-consumption problems and the control problems of the life insurance company. This connection is briefly mentioned in Cairns [12] and will be clarified in Chapters 5 and 6.

Now we introduce an objective which rewards a certain kind of stability of the pension fund and in the payments. This is done by working with a quadratic disutility the expected total of which is now to be minimized,

$$\begin{aligned}&\upsilon(t,X(t),dU(t),dt)\\ =\;& e^{-\gamma t}\left(a(X(t)-\widehat{x})^2+b(u(t)-\widehat{u})^2+c(X(t)-\widehat{x})(u(t)-\widehat{u})\right)dt. \quad (1.14)\end{aligned}$$

This disutility function punishes distance between $X(t)$ and $\widehat{x}$ and distance between $u(t)$ and $\widehat{u}$. The punishments are weighted by (a,b,c).

This disutility function is clearly somewhat connected to the quadratic approaches to hedging of contingent claims in finance like mean-variance hedging, but the one cannot be considered as a special case of the other. However, the quadratic approaches share a counter-intuitive conclusion on investment which is easily explained: If one wishes to have $X(t)$ close to $\widehat{x}$ (or eventually $X(\tau)$ close to $Y(\tau)$), and $X(t)$ exceeds its target, then one is urged to throw away money on the financial market by investing non-efficiently. In the case of hedging one can argue that the quadratic approach is only relevant if $X(t)$ is below its target. This argument calls for further studies in connection with pension funding where $X(t)$ very well in practice may be above its target.

Risk-adjusted utility

In this section we seek to remedy the drawback of the quadratic approaches to hedging and pension fund controlling concerning counter-intuitive investments by introducing a notion of *risk-adjusted utility*. This is an alternative to the traditional objective function given by (1.12). It is based on a certain kind of state-dependence of utilities which, in some special cases, separates the problem of optimal investment from the problem of optimal consumption. The idea is to allow for dependence of a state price deflator Λ, such that we wish to optimize

$$E\left[\int_{0-}^{\tau} \upsilon\left(t, \Lambda(t), X(t), dU(t), dt\right) + \Upsilon\left(\Lambda(\tau), X(\tau)\right)\right].$$

The state price deflator is just the one used to calculate values of payment processes by their expected value, namely (see (1.7))

$$E^Q\left[\int_{0-}^{T} \frac{1}{Z_s^0} dB_s\right] = E\left[\int_{0-}^{T} \Lambda_s dB_s\right].$$

The dependence on Λ can be introduced directly in the control problems explained above in both finance and life insurance. In finance, we introduce dependence of Λ such that the investment-consumption problem reads

$$\begin{aligned} \upsilon\left(t, \Lambda(t), X(t), dU(t), dt\right) &= e^{-\gamma t} \upsilon\left(\Lambda(t) u(t)\right) dt, \\ \Upsilon\left(\Lambda(\tau), X(\tau)\right) &= e^{-\gamma t} \Upsilon\left(\Lambda(\tau) X(\tau)\right). \end{aligned}$$

In pension funding, the dependence on Λ is introduced by $(c = 0)$,

$$\begin{aligned} & \upsilon\left(t, \Lambda(t), X(t), dU(t), dt\right) \\ = \; & e^{-\gamma t}\left(a\left(\Lambda(t) X(t) - \widehat{x}\right)^2 + b\left(\Lambda(t) u(t) - \widehat{u}\right)^2\right) dt. \end{aligned} \tag{1.15}$$

The idea is the same as in the classical formulation, but instead of measuring utility of consumption or wealth by their nominal values we measure it by their deflated values. In Chapter 5 we pursue this idea and study its effect on variations of Merton's problem and its effect on pricing by utility indifference. Although we show that the concept produces nice and intuitively appealing prices and strategies for consumption and investment, we do not claim that it will work well in every area where utility theory is the basis for decision making or pricing. We shall not conceal that it is a pragmatic idea which should be used with care.

Modifications of control in life and pension insurance

The quadratic disutility functions given by (1.14) and (1.15) bring to mind the classical control problem called the linear regulator problem, exposed in just about every textbook on stochastic control, see e.g. Fleming and Rishel [23]. This problem differs from most other stochastic control problems by being easily solved explicitly and by having a simple solution. This carries over to the life insurance company's decision problem in the case of pension funding. Both in Cairns [12] and in our Chapter 6 based on risk-adjusted utility, u is optimally controlled by a linear function of X.

The simple solution to the linear regulator problem relies heavily on absence of constraints on the control variable u. We shall also be interested in the case where $dU(t)$ is constrained to be chosen in $\mathbf{R}_+ \cup \{0\}$. We are interested in this situation because it relates to the situation in life insurance where dividends are constrained to be to the benefit of the policy holder. Thus, we are going to speak of pension funding and participating life insurance as the unconstrained case and the constrained case, respectively.

The constrained case represents one modification of the pension fund control problem with quadratic disutility. Another modification is to punish the distance of consumption to its target by absolute value instead of quadratic value,

$$\upsilon(t, \Lambda(t), X(t), dU(t), dt) = a(\Lambda(t) X(t) - \widehat{x})^2 dt + b|\Lambda(t) dU(t) - \widehat{u}dt|.$$

While quadratic disutility leads to optimal consumption of classical (absolutely continuous with respect to the Lebesgue measure) type, absolute disutility leads to optimal consumption of singular type. The idea of rewarding stability of payments and wealth is the same as in the quadratic case, but by measuring distance in another way one gets an optimal control of completely different nature. It becomes optimal to keep the surplus or fund within a certain area by singular repayments. If U is not constrained to be positive, the area is bounded from above and below; if U is constrained to be positive, the area is bounded from above only. This optimal behavior is well-known in the literature on life insurance mathematics where the boundary is called the bonus barrier, see e.g. Daykin et al. [16, p. 419].

The introduction of risk-adjusted utility is one modification of the traditional pension funding control problem leading to intuitively feasible investment strategies. Quadratic and absolute disutility of payments, constrained or not, lead to intuitively feasible strategies for dividend payments. However, other ways of adjusting the objective of control may lead to different but still intuitively feasible strategies. In Hansen [30], a traditional concave utility of dividends combined with no utility of surplus is studied. This resembles the classical optimization problem in finance but differs by the way in which dividends are paid out; dividends are not immediately turned into payments but currently traded into future payments. This can lead to a class of problems related to habit formation utility specifications also resulting in optimal dividend distribution of singular type. In Taylor [68], the quadratic disutility of dividends is maintained but the quadratic disutility of the fund (ratio) is replaced by a decreasing function such that high funds are not punished in the same way as low funds are.

1.5 Overview and contributions of the thesis

Valuation in life and pension insurance

In Chapter 2 we work with general valuation of payment streams. A classical result in life insurance mathematics, Thiele's differential equation, is generalized to valuation by absence of arbitrage, and examples of insurance contracts where the equation proves to be a constructive tool, are given. A special case of the set-up in Chapter 2 constitutes a part of Steffensen [64].

Chapter 3 takes a closer look at the contingent claims that are actually present in life and pension insurance contracts. Varying from pension funds to participating life insurance contracts, we explain how the payments of these contracts are made up by first order payments and dividends. Linking dividends to the surplus represents one way of explicitly imitating the implicit dependence on the surplus present in these products, and we study the dynamics of various versions of the surplus intensively. We show how Thiele's generalized differential equation in the case of such surplus-linked dividend payments is a constructive tool in the search for fair strategies for investment and redistribution of surplus. A main example illustrates notation, terminology, and results and is also the basis for a few illustrative figures. These are borrowed from Ekstrøm [21].

In Chapter 4 we introduce intervention options held by the policy holder. This leads to a further generalization of Thiele's differential equation such that it takes the form of a so-called quasi-variational inequality. A simple example of an intervention option is the exercise option in an American option in finance, but the free policy and surrender options in life and pension insurance are more important in our context. The framework and the resulting quasi-variational inequality are illustrated by these three examples.

Control in life and pension insurance

Chapter 5 and Chapter 6 extends the study of fair dividends to the study of optimal dividends. In Chapter 5 the idea of risk-adjusted utility is introduced and illustrated in a number of optimization problems in finance. Also the problem of pricing claims by equivalence of utility is considered. Chapter 5 prepares for Chapter 6 and demonstrates no connection to decision problems of the life insurance company, whatsoever.

In Chapter 6 risk-adjusted utility is applied to the problem of optimal investment and distribution of surplus in life and pension insurance. This is done in a general framework of consumption and investment with risky income and risky debt. A range of problems of classical and singular type, constrained and unconstrained, lead to a small collection of problems and solutions: Some solutions are known and cited here; some solutions are, we believe, carried out here for the first time; some solutions are not to be found explicitly but can be illustrated numerically; and in a few cases we do not even get that far.

Part II

Valuation in life and pension insurance

Chapter 2

A no arbitrage approach to Thiele's differential equation

The multistate life insurance contract is reconsidered in a framework of securitization where insurance claims may be priced by the principle of no arbitrage. This way a generalized version of Thiele's differential equation is obtained for insurance contracts linked to indices, possibly marketed securities. The equation is exemplified by a traditional policy, a simple unit-linked policy and a path-dependent unit-linked policy. This chapter is an adapted version of Steffensen [66].

2.1 Introduction

The reserve on an insurance contract is traditionally defined as the expected present value of future contractual payments and is provided by the insurance company to cover these payments. The reserve thus defined can be calculated under various conditions depending e.g. on the choice of discount factor used for calculation of the present value. We shall take a different approach and define the reserve as the market price of future payments. This redefinition of the reserve inspires a reconsideration of the premium calculation principle. Financial mathematics suggest the principle of no arbitrage, and our purpose is to derive the structure of the reserve imposed by this principle. Fortunately, this structure specializes to well-known results in actuarial mathematics like Thiele's differential equation, introduced by Thiele in 1875, and since then generalized in various directions. Thus, the traditionally defined reserve coincides with the price under certain market conditions.

The key to market prices is the notion of securitization of insurance contracts. Securitization of insurance contracts is making progress in various respects these years. At the stock exchanges all over the world attempts are made at securitizing insurance risk as an alternative to traditional exchange of risk by reinsurance contracts. This development on the exchanges is the background for an interest in modelling and pricing a variety of new products (see e.g. Cummins and Geman [13] and Embrechts and Meister [22]). Parallel to this development, securitization has become an important concept in the unification of actuarial mathematics and mathematical finance since it plays an important role in stating actuarial problems in the framework of mathematical finance and vice versa (see e.g. Delbaen and Haezendonck [17] and Sondermann [63]).

Financial theory applies to markets where there exist assets correlated with the claim subject to pricing, and finance is thus particularly apt to analysis of insurance contracts if such a market exists. An obvious example is unit-linked life insurance, at least if the unit is traded, and this subject of actuarial mathematics has been an issue of financial theory since Brennan and Schwartz [10] recognized the option structure of a unit-linked life insurance with a guarantee. Aase and Persson [1] gives an overview of existing literature up to 1994.

Aase and Persson [1] obtained a generalized version of Thiele's differential equation for unit-linked insurance contracts. Our model framework covers their set-up, and we show how the securitization leads to further generalization of Thiele's differential equation by means of arguments fetched from finance exclusively. The fundamental connection between the celebrated Thiele's differential equation and the Black-Scholes differential equation (just as celebrated but in a different forum) is indicated by Aase and Persson [1]. Our derivation brings to the surface more directly this connection by treatment of financial risk and insurance risk on equal terms.

The target group of the chapter is twofold. On one hand, we approach an actuarial problem of evaluating an insurance payment process. The tools are imported from financial mathematics, and the reader with a background in traditional actuarial mathematics will benefit from knowledge of the concept of arbitrage as well as its connection to martingale measures. References are Harrison and Pliska [32] and Delbaen and Schachermayer [18]. On the other hand, the chapter may also form an introduction to life insurance mathematics for financial mathematicians. A statistical model frequently used in life insurance mathematics is presented, an insurance contract is constructed, and our main result is specialized to Thiele's differential equation. The statistical model and the construction of an insurance contract is not motivated here, however, and the reader with a background only in financial mathematics is asked to consult a textbook on basic life insurance mathematics, e.g. Gerber [25].

In Section 2.2 we present the basic stochastic model. In Section 2.3 we define an index and a market based on this model and in Section 2.4 we introduce a payment process and an insurance contract based on the index. In Section 2.5 the price process of an insurance contract is derived, whereas the differential equation imposed by a no arbitrage condition on the market forming this price process is derived in Section 2.6. Section 2.7 contains three examples of which one is the traditional actuarial set-up, whereas two treat unit-linked insurance in a simple and a path-dependent set-up, respectively.

2.2 The basic stochastic environment

We take as given a probability space $(\Omega, \mathcal{F}, \mathbf{F} = \{\mathcal{F}_t\}_{t\geq 0}, P)$. We let $(X_t)_{t\geq 0}$ be a cadlag (i.e. its sample paths are almost surely right continuous with left limits) jump process with finite state space $\mathcal{J} = (1, \ldots, J)$ defined on $(\Omega, \mathcal{F}, \mathbf{F} = \{\mathcal{F}_t\}_{t\geq 0}, P)$ and associate a marked point process (T_n, Φ_n), where T_n denotes the time of the nth jump of X_t, and Φ_n is the state entered at time T_n, i.e. $X_{T_n} = \Phi_n$. We introduce the counting processes

$$N_t^j = \sum_{n=1}^{\infty} 1_{(T_n \leq t, X_{T_n} = j)}, \quad j \in \mathcal{J},$$

and the J-dimensional vector

$$N_t = \begin{bmatrix} N_t^1 \\ \vdots \\ N_t^J \end{bmatrix}.$$

We let $(W_t)_{t\geq 0} = \left(W_t^1, \ldots W_t^K\right)_{t\geq 0}$ be a standard K-dimensional Brownian motion defined on $(\Omega, \mathcal{F}, \mathbf{F} = \{\mathcal{F}_t\}_{t\geq 0}, P)$.

For a matrix A we let A^T denote the transpose of A and let $A^{i\cdot}$ and $A^{\cdot i}$ denote the ith row and the ith column of A, respectively. For a vector a, we let $diag\,(a)$ denote the diagonal matrix with the components of a in the principal diagonal and 0 elsewhere. We shall write $\delta^{1\times J}$ and $\delta^{J\times 1}$ instead of $(\delta, \ldots, \delta)$ and $(\delta, \ldots, \delta)^T$, respectively. For derivatives we shall use the notation $\partial_x = \frac{\partial}{\partial x}$ and $\partial_{xy} = \frac{\partial^2}{\partial x \partial y}$. For a vector a we let $\int a$ and da mean componentwise integration and componentwise differentiation, respectively.

2.3 The index and the market

In the subsequent sections we shall define and study an insurance contract. Instead of letting the payments in the insurance contract be directly driven by the stochastic basis we shall work with an index which is driven by the stochastic basis and which will form the basis for the payments.

We introduce an *index* S, an $(I+1)$-dimensional vector of processes, the dynamics of which is given by

$$dS_t = \alpha_t dt + \beta_{t-} dN_t + \sigma_t dW_t, \; S_0 = s_0,$$

where $\alpha \in \mathbf{R}^{(I+1)}$, $\beta \in \mathbf{R}^{(I+1)\times J}$, and $\sigma \in \mathbf{R}^{(I+1)\times K}$ are functions of (t, S_t) and $s_0 \in \mathbf{R}^{I+1}$ is $\mathcal{F}_0$-measurable. We denote by S^i, α^i, β^{ij}, and σ^{ik} the ith entry of S, the ith entry of α, the (i,j)th entry of β, and (i,k)th entry of σ, respectively. The information generated by S is formalized by the filtration $\mathbf{F}^S = \left\{\mathcal{F}_t^S\right\}_{t\geq 0}$, where

$$\mathcal{F}_t^S = \sigma(S_s, \; 0 \leq s \leq t) \subseteq \mathcal{F}_t.$$

We assume that S is a Markov process and that there exist deterministic piecewise continuous functions $\mu^j(t,s)$, $j \in \mathcal{J}$, $s \in \mathbf{R}^{I+1}$ such that N_t^j admits the $\mathcal{F}_t^S$-intensity process $\mu_t^j = \mu^j(t, S_t)$, informally given by

$$\begin{aligned} \mu_t^j dt &= E\left(dN_t^j \middle| \mathcal{F}_{t-}^S\right) + o(dt) \\ &= E\left(dN_t^j \middle| S_{t-}\right) + o(dt), \end{aligned}$$

where $o(h)/h \to 0$ as $h \to 0$. We introduce the J-dimensional vectors containing the intensity processes and martingales associated with N,

$$\mu_t = \begin{bmatrix} \mu_t^1 \\ \vdots \\ \mu_t^J \end{bmatrix}, \quad M_t = \begin{bmatrix} M_t^1 \\ \vdots \\ M_t^J \end{bmatrix} = \begin{bmatrix} N_t^1 - \int_0^t \mu_s^1 ds \\ \vdots \\ N_t^J - \int_0^t \mu_s^J ds \end{bmatrix}.$$

To help the reader fix ideas, we explain briefly the roles of the introduced processes. Their roles will become more clear when we formalize the payment process below. The process N describes (at least) some specification of the life history of an insured. Whereas the process N will partly determine the points in time where payments fall due, the process S determines the amounts of these payments (and the intensities for the process N). In classical life insurance mathematics, payments are allowed to depend on the state of the policy, X. We can cover this situation by taking S^1 to be equal to X by the coefficients

$$\alpha_t^1 = 0,\ \beta_t^{1j} = j - S_t^1,\ \sigma_t^1 = 0,\ s_0^1 = X_0. \tag{2.1}$$

If e.g. X is included in the index S, $\mu(t, X_t)$ candidates to the intensity process corresponding to the classical situation, see e.g. Hoem [33]. However, in general, the intensity process μ may differ from the intensity process with respect to the natural filtration of N.

However, this classical contract can be extended in various directions. We can e.g. allow for payments (and intensities) to depend on the duration of the sojourn in the current state by letting S^2 be defined by

$$\alpha_t^2 = 1,\ \beta_t^{2j} = -S_t^2,\ \sigma_t^2 = 0,\ s_0^2 = 0, \tag{2.2}$$

and allow for payments (and intensities) to depend on the total number of jumps by letting S^3 be defined by

$$\alpha_t^3 = 0,\ \beta_t^{3j} = 1,\ \sigma_t^3 = 0,\ s_0^3 = 0. \tag{2.3}$$

In Møller [47] and Norberg [52] generalized versions of Thiele's differential equation have been studied where payments depend on the duration of the sojourn in the current state.

We introduce a *market* Z, an $(n+1)$-dimensional vector $(n \leq I)$ of price processes assumed to be positive, and denote by Z^i the ith entry of Z. The market Z consists of exactly those entries of S that are prices of traded assets. We assume that there exists a short rate of interest such that the market contains a price process Z^0 with the dynamics given by

$$dZ_t^0 = r_t Z_t^0 dt, Z_0^0 = 1.$$

This price process can be considered as the value process of a unit deposited on a bank account at time 0, and we shall call this entry for the risk-free asset even though r_t is allowed to depend on (t, S_t). Furthermore, we assume that the set of martingale measures, $\mathcal{Q}$, i.e. the set of probability measures Q equivalent to P such that $\frac{Z^i}{Z^0}$ is a Q-martingale for each i, is non-empty. From fundamental theory of asset pricing this assumption is known to be essentially equivalent to the assumption that no arbitrage possibilities exist on the market Z. The entries of an index S will also be called indices, and the indices appearing in Z will then be called marketed indices or assets. With this formulation the set of marketed indices is a subset of the set of indices and it contains at least one entry, namely Z^0. We let $\alpha^Z \in \mathbf{R}^{(n+1)}$, $\beta^Z \in \mathbf{R}^{(n+1)\times J}$, and $\sigma^Z \in \mathbf{R}^{(n+1)\times K}$ denote the coefficients of the asset prices Z.

2.4 The payment process and the insurance contract

Fixing some time horizon T, we formally take an *insurance contract* to be a *payment process* B which is an $\mathcal{F}_t^S$-adapted, cadlag process of finite variation with dynamics given by

$$dB_t = B_0 d1_{(t\geq 0)} + b_t^c dt - b_{t-}^d dN_t - \Delta B_T d1_{(t\geq T)},$$

where $B_0 \in \mathbf{R}$ is a function of S_0, $b^c \in \mathbf{R}$ and $b^d \in \mathbf{R}^J$ are functions of (t, S_t), and $\Delta B_T \in \mathbf{R}$ is a function of S_T. We denote by b^{dj} the jth entry of b^d. Note that the $\mathcal{F}_t^S$-adaptedness of B places demands on the connection between the coefficients of S and the coefficients of B. Although it need not be the case, the reader should have in mind the case where $\mathcal{F}_0$ and thus also $\mathcal{F}_0^S$ are trivial, i.e. $\mathcal{F}_0 = \mathcal{F}_0^S = \{\emptyset, \Omega\}$. Then B_0 is deterministic.

B_t represents the cumulative payments from the policy holder to the insurance company over $[0, t]$. Both continuous payments and lump sum payments are thus allowed to depend on the present state of the process (t, S_t). The minus signs in front of b^d and ΔB in dB_t conform to the typical situation where B_0 and b^c are premiums and b^d and ΔB are benefits, all positive. To simplify notation, lump sum payments at deterministic times are restricted to time 0 and time T. Thus, an insurance contract is given by a set of functions $\left(B_0, b^c, b^d, \Delta B\right)$ such that a recording of S completely determines the payment stream.

It should be noted that the inclusion of Z in S opens for insurance contracts with payments depending on marketed indices and on indices which are somehow driven by the market. This remark motivates a continuation of our series of examples of entries in S (2.1)-(2.3) showing that the general insurance contract opens for a number of well-known simple and advanced unit-linked life insurance policies. We will still focus on the process S and take as given that this process is Markov and that the $\mathcal{F}_t^S$-intensity process exists in all cases. Given S, interesting insurance contracts are easily devised. The reader will recognize some elements from this illustration in the examples in Section 2.7 below.

We extend the dependence of payments given in (2.1)-(2.3) by letting payments depend on the present state of a marketed index, e.g. a geometric Brownian motion. The geometric Brownian motion and the process Z^0 constitute the Black-Scholes model and is obtained by defining

$$\alpha_t^4 = S_t^4 \alpha^4,\ \beta_t^4 = 0,\ \sigma_t^4 = S_t^4 \sigma^4.$$

A Markovian multidimensional diffusion model is obtained by just adding further processes similar to S^4 to the market. The Black-Scholes case and the multidimensional diffusion case have been studied previously in connection with unit-linked insurance in Aase and Persson [1] and Ekern and Persson [20], respectively, with a particular construction of X, namely a classical survival model. However, they did not bring the processes N and X to the surface.

The pure diffusion price processes are continuous. Price processes involving jumps involve a development of the traditional actuarial idea of the process X, where X describes the state of life of an individual or a group of individuals only. In a jump model for prices, the process X may also partly describe the financial state. This development of the traditional actuarial idea has been studied previously in e.g. Norberg [51] where a stochastic interest rate is driven by a process of the same type as X. We mention that e.g. a price process modelled by a geometric compound Poisson process, where the jump distribution is discrete and finite, is included in the present model.

Just as we opened the possibility of path-dependence on the non-marketed index X through addition of the state variables S^2 and S^3, we can also have dependence on the path of the marketed index S^4 e.g. by defining

$$\alpha_t^5 = g\left(t, S_t^4\right),\ \beta_t^5 = 0,\ \sigma_t^5 = 0,$$

where g is some specified function. Introduction of path-dependence of marketed indices through addition of state variables is well-known in the theory of Asian derivatives, and it opens for quite exotic unit-linked products as will be exemplified in Section 2.7.3. Previously, path-dependent unit-linked insurance has been studied in Bacinello and Ortu [2] and Nielsen and Sandmann [49] in set-ups quite different from ours.

In Aase and Persson [1], Ekern and Persson [20], Bacinello and Ortu [2], and Nielsen and Sandmann [49], standard arbitrage pricing theory is applied to (complete market) financial risk in unit-linked insurance. Working in a framework of securitization, we apply, however, (possibly incomplete market) financial mathematics to both financial risk and insurance risk, and we obtain thereby a generalized version of Thiele's differential equation and a corresponding pricing formula, where financial risk and insurance risk are treated on equal terms.

2.5 The derived price process

The insurance contract forms the basis for introduction of two price processes, F and V:

$$\begin{aligned} F_t &= \text{the price at time } t \text{ of the contractual payments to the} \\ & \text{insurance company over } [0,T]\text{, i.e. premiums less benefits,} \\ V_t &= \text{the price at time } t \text{ of the contractual payments from the} \\ & \text{insurance company over } (t,T]\text{, i.e. benefits less premiums.} \end{aligned}$$

We make some preliminary comments on these processes as a preparation and motivation for a detailed study.

By the price at time t of contractual payments we mean the amount against which the payments stipulated by a contract are taken over by one agent from another. Thus, buying and selling means 'taking over' and 'handing over', respectively, the contractual payments over some specified period of time. This consideration of contractual payments as dynamically marketed objects is called 'securitization' of insurance contracts and plays an important role in the adaptation and application of financial theory to insurance problems. Important contributions are Delbaen and Haezendonck [17] and Sondermann [63].

By the securitization of contractual payments, we have implicitly taken as given the existence of a market, on which these contractual payments are allowed to be traded and, furthermore, that these contractual payments actually are bought and sold by the agents on the market. We shall assume that Z constitutes such a market.

In many countries government regulations appear to prohibit a securitization of insurance contracts. One of the reasons may be that the supervisory authorities are not at all prepared for a free exchange of the kind of financial interests appearing on the insurance market. On the other hand, traditional reinsurance contracts actually represent one allowable way of forwarding risk to a third party. The agents on the insurance market, i.e. the customers, the direct insurance companies, and the reinsurance companies are, of course, the primary investors, but also other parties may consider the contractual payments of insurance contracts as possible investment objects. This statement is substantiated by the fact that F can be interpreted as the surplus of the company stemming from the insurance contract. This surplus is reflected in the equity, which is definitely a relevant investment object for all investors.

We have introduced two price processes, one covering all contractual payments and one covering future payments only. Even though we may be interested in the price of the future payments only, we shall work with the process F since this process forms the asset, in a financial sense, arising from the securitization of the insurance contract. One could consider the introduction of the process F as a preliminary step leading to the definition of and derivation of formulas for the process V. In practice, one should have trading and marketing of V in mind. This also explains why we have not taken the investment strategy (to be introduced below) chosen by the insurance company as an integrated part of the insurance contract. As we shall see, this strategy affects the price of past payments but not the price of future payments.

In actuarial terminology, the outstanding liabilities are called the reserve, and these liabilities can be calculated under various assumptions. Since F and V are price processes arising on a market, it seems natural to call V_t the market reserve at time t. We will, however, suppress the word 'market', and simply speak of V_t as the reserve at time t. One should carefully note that, whereas the reserve is traditionally defined as the expected present value of future payments, we take the reserve to be the market price of future payments.

Our approach to the price process F is the following: Assuming that the market Z is arbitrage free, we require that also the market (Z, F) be arbitrage free. We use the essential equivalence between arbitrage free markets and existence of a so-called martingale measure, i.e. a measure under which discounted asset prices are martingales. If the no arbitrage condition is fulfilled for (Z, F), we shall speak of B as an arbitrage free insurance contract and about V as the corresponding arbitrage free reserve.

Already at this stage we will argue for side conditions on the price process V. These side conditions are due to the no arbitrage condition on the market (Z, F) and the structure of the payment process B. Here it is important to state clearly the problems we actually want to solve: Given S, including Z, we wish to determine a payment process B such that no arbitrage possibilities arise from marketing the insurance contract. Afterwards, given the payment process B we wish to determine arbitrage free prices of the insurance contract.

When determining the payment process, this process is to be considered as a balancing tool and is as such comparable with the delivery price of a future or the price of an option. However, the payment process contains a continuum of balancing elements (premiums and benefits) and in practice all but one of these elements are predetermined by the customer and the last one acts as the balancing tool of the insurance company. Which elements are predetermined and which element is the balancing tool depends on the type of insurance contract (defined benefits, defined contributions etc.). Since the contract can be entered into at time 0 with no past payments, B should be balanced such that the equivalence relation

$$F_{0-} = -V_{0-} = 0 \tag{2.4}$$

is fulfilled in order to prevent the obvious arbitrage possibility that arises if an agent can enter into the insurance contract and immediately sell the same contract on the market at a price different from 0. If $\mathcal{F}_0^S$ is trivial such that B_0 is deterministic, the equivalence relation (2.4) can also be written as

$$V_0 = B_0.$$

If e.g. B_0 is fixed at 0, the remaining elements of B are to be determined subject to $V_0 = 0$. Hereby, the insurance contract is somewhat similar to a future contract.

The side condition at time T is also given by a no arbitrage argument. Since the price at time T of a payment of ΔB_T at time T in an arbitrage free market must be ΔB_T, we have

$$V_{T-} = \Delta B_T. \tag{2.5}$$

So, the side conditions (2.4) and (2.5), imposed by the no arbitrage condition, should be included in the basis for balancing the payment process B. Given B, this payment process is to be considered as an, indeed unusual, contingent claim and achieves as such at least one arbitrage free price at any time in an arbitrage free market. Here again, the insurance contract is somewhat similar to the future contracts which has a price, positive or negative, at any time during the term of the contract.

The insurance company receives payments in accordance with the insurance contract B, and we assume that these are currently deposited on (withdrawn from) an account which is invested in a portfolio with positive value process U, generated by a self-financing investment strategy $\theta \in \mathbf{R}^{n+1}$, i.e.

$$\begin{aligned} U_t &= \theta_t \cdot Z_t = \sum_{i=0}^{n} \theta_t^i Z_t^i > 0, \\ dU_t &= \theta_t \cdot dZ_t. \end{aligned}$$

The strategy is furthermore assumed to comply with whatever institutional requirements there may be. Throughout this chapter one can think of θ as the strategy corresponding to a constant relative portfolio, i.e. a strategy θ such that for a constant $(n+1)$-dimensional vector γ, $\theta_t^i Z_{t-}^i = \gamma^i U_{t-}$, $i = 0, \ldots, n$. This strategy reflects an investment profile possibly restricted by the supervisory authorities, e.g. such that θ^i is non-negative for all i if short-selling is not allowed. We emphasize that θ, in general, is not a strategy aiming at hedging some contingent claim.

Consequently, the present value at time t of the contractual payments over $[0, T]$ becomes $U_t \int_0^T \frac{1}{U_s} dB_s$, where U is the value process corresponding to the chosen trading strategy θ. This present value is composed of an $\mathcal{F}_t^S$-measurable part,

$$L_t = U_t \int_0^t \frac{1}{U_s} dB_s,$$

and a part which is not in general $\mathcal{F}_t^S$-measurable,

$$U_t \int_t^T \frac{1}{U_s} dB_s.$$

If the price operator, denoted by π_t, is assumed to be additive, pricing the contractual payments over $[0, T]$ amounts to replacing $U_t \int_t^T \frac{1}{U_s} dB_s$ by some $\mathcal{F}_t$-measurable process, the price process $-V_t$. Thus,

$$F_t = \pi_t\left(U_t\int_0^T \frac{1}{U_s}dB_s\right) = L_t - V_t.$$

We restrict ourselves to prices allowing V_t to be written in the form $V(t, S_t)$. This restriction seems reasonable since S is Markov and since the payments by B and the intensities of N depend only on time and the current value of S, but it is actually a restrictive assumption on the formation of prices in the market. It corresponds to a restrictive It corresponds to the restrictive structure of the measure transformation in Section 2.6.

If X jumps to state j at time t, S will jump to $S_{t-} + \beta_{t-}^{\cdot j}$, and thus V_t jumps to V_{t-}^j, where $V_t^j \equiv V\left(t, S_t + \beta_t^{\cdot j}\right)$. Each V_{t-}^j is $\mathcal{F}_t^S$-predictable, and we can introduce the J-dimensional $\mathcal{F}_t^S$-predictable row vector

$$V_{t-}^{\mathcal{J}} = \left[V_{t-}^1, \ldots, V_{t-}^J\right].$$

Assume that the partial derivatives $\partial_t V(t,s)$, $\partial_s V(t,s)$, and $\partial_{ss} V(t,s)$ exist and are continuous, abbreviate

$$\partial_s V_t = \partial_s V(t, S_t) = \partial_s V(t,s)|_{s=S_t},$$

and denote $\frac{1}{2}tr\left(\sigma_t^T \partial_{ss} V_t \sigma_t\right)$ by ψ_t. Then Ito's lemma applied to the process V gives the differential form,

$$dV_t = \left(\partial_t V_t + (\partial_s V_t)^T \alpha_t + \psi_t\right)dt + \left(V_{t-}^{\mathcal{J}} - V_{t-}^{1\times J}\right)dN_t + (\partial_s V_t)^T \sigma_t dW_t.$$

Ito's lemma also gives the differential form of the process L,

$$\begin{aligned} dL_t &= b_t^c dt - b_{t-}^d dN_t + \frac{L_{t-}}{U_{t-}}dU_t \\ &= b_t^c dt - b_{t-}^d dN_t + L_t r_t dt - L_t r_t dt + \frac{L_{t-}}{U_{t-}}dU_t \\ &= b_t^c dt - b_{t-}^d dN_t + L_t r_t dt + \frac{L_{t-}Z_t^0}{U_{t-}}d\left(\frac{U_t}{Z_t^0}\right). \end{aligned}$$

It should be noted that we can also write $dL_t = dB_t + \frac{\theta_t L_{t-}}{U_{t-}}dZ_t$ and consider the process L as a value process corresponding to a trading strategy given by $\upsilon_t = \frac{\theta_t L_{t-}}{U_{t-}}$. Because of the payment process B, this strategy is not self-financing, though.

Now, collecting terms gives the differential form of the process F,

$$\begin{aligned} dF_t &= dL_t - dV_t \\ &= r_t F_t dt + \\ &\quad \left(b_t^c + r_t V_t - \partial_t V_t - (\partial_s V_t)^T \alpha_t - \psi_t - \left(b_t^d + V_t^{\mathcal{J}} - V_t^{1\times J}\right)\mu_t\right)dt \\ &\quad - (\partial_s V_t)^T \sigma_t dW_t - \left(b_{t-}^d + V_{t-}^{\mathcal{J}} - V_{t-}^{1\times J}\right)dM_t + \frac{L_{t-}Z_t^0}{U_{t-}}d\left(\frac{U_t}{Z_t^0}\right). \end{aligned}$$

Upon introducing

$$\begin{aligned} R_t &= b_t^d + V_t^{\mathcal{J}} - V_t^{1\times J}, \\ TD\left(\alpha_t^S, \mu_t\right) &= b_t^c + r_t V_t - \partial_t V_t - \left(\partial_s V_t\right)^T \alpha_t - R_t \mu_t - \psi_t, \end{aligned}$$

and abbreviating

$$\begin{aligned} \alpha_t^F &= r_t F_t + TD\left(\alpha_t, \mu_t\right), \\ \beta_t^F &= -R_t, \\ \sigma_t^F &= -\left(\partial_s V_t\right)^T \sigma_t, \\ \rho_t^F &= \frac{L_t Z_t^0}{U_t}, \end{aligned}$$

we arrive at the simple form

$$dF_t = \alpha_t^F dt + \beta_{t-}^F dM_t + \sigma_t^F dW_t + \rho_{t-}^F d\left(\frac{U_t}{Z_t^0}\right). \tag{2.6}$$

The abbreviations R and TD are motivated by the terms sum at **R**isk and **T**hiele's **D**ifferential, respectively. In Section 2.6, we shall see that TD, taken in a point different from (α_t, μ_t), equated to 0 constitutes a generalized version of Thiele's differential equation. A differential equation for the reserve of a life insurance contract was derived by Thiele in 1875, but we shall refer to Hoem [33] for a classical version presented in probabilistic terms.

Note that (2.6) is not the semimartingale form under P, since $\frac{U}{Z^0}$ is not in general a P-martingale. This is, however, a convenient form as the succeeding section will show.

2.6 The set of martingale measures and Thiele's differential equation

In this section we study the consequences of the no arbitrage condition on the markets Z and (Z, F) by studying the conditions for existence of an equivalent martingale measure on these markets.

For construction of a new measure Q, we shall define a likelihood process Λ by

$$\begin{aligned} d\Lambda_t &= \Lambda_{t-}\left(\sum_j g_{t-}^j dM_t^j + \sum_k h_t^k dW_t^k\right) \\ &= \Lambda_{t-}\left(g_{t-}^T dM_t + h_t^T dW_t\right), \\ \Lambda_0 &= 1, \end{aligned}$$

where we have introduced

$$g_t^j = g^j\left(t, S_t\right), \; h_t^k = h^k\left(t, S_t\right), \tag{2.7}$$

and

$$g_t = \begin{bmatrix} g_t^1 \\ \vdots \\ g_t^J \end{bmatrix}, \; h_t = \begin{bmatrix} h_t^1 \\ \vdots \\ h_t^K \end{bmatrix}.$$

Assume that g_t and h_t are chosen such that the conditions

$$\begin{aligned} E^P\left[\Lambda_T\right] = 1, \\ g^j\left(t,s\right) > -1, \quad j \in \mathcal{J}, \end{aligned} \tag{2.8}$$

are fulfilled. Then we can change measure from P to Q on (Ω, F_T) by the definition,

$$\Lambda_T = \frac{dQ}{dP},$$

and it follows from Girsanov's Theorems that W_t^k under Q has the local drift h_t^k and that N_t^j under Q admits the $\mathcal{F}_t^S$-intensity process $(1+g_t^j)\mu_t^j$. In vector notation, W_t has the local drift h_t and N_t admits the $\mathcal{F}_t^S$-intensity process $diag\left(1^{J\times 1} + g_t\right)\mu_t$ under Q. Note that by (2.7) we consider only the part of possible measure transformations that allow g and h to be stochastic processes in a particular form. This restriction on the measure transformation is imposed by the restriction on the price operator leading to $V_t = V\left(t, S_t\right)$. This will be argued at the end of this section. Defining the Q-martingales

$$\begin{aligned} M_t^Q &= N_t - \int_0^t diag\left(1^{J\times 1} + g_s\right)\mu_s ds, \\ W_t^Q &= W_t - \int_0^t h_s ds, \end{aligned}$$

we can write the dynamics of (Z, F) as

$$\begin{aligned} dZ_t &= \alpha_t^{ZQ} dt + \beta_{t-}^Z dM_t^Q + \sigma_t^Z dW_t^Q, \\ dF_t &= \alpha_t^{FQ} dt + \beta_{t-}^F dM_t^Q + \sigma_t^F dW_t^Q + \rho_{t-}^F d\left(\frac{U_t}{Z_t^0}\right), \end{aligned} \tag{2.9}$$

where

$$\begin{aligned} \alpha_t^{ZQ} &= \alpha_t^Z + \sigma_t^Z h_t + \beta_t^Z diag\left(1^{J\times 1} + g_t\right)\mu_t, \\ \alpha_t^{FQ} &= r_t F_t + TD\left(\alpha_t + \sigma_t h_t, diag\left(1^{J\times 1} + g_t\right)\mu_t\right). \end{aligned}$$

We define the market prices of diffusion and jump risk, respectively, by

$$\begin{aligned} \eta_t &= -h_t, \\ \xi_t &= -diag\left(g_t\right)\mu_t, \end{aligned}$$

and we say that the insurance company is risk-neutral with respect to diffusion risk k or jump risk j if $\eta_t^k = 0$ or $\xi_t^j = 0$, respectively.

Now we determine the set of martingale measures $\mathcal{Q}$ in the market Z by requiring $\frac{Z}{Z^0}$ to be a martingale under Q. We see that g_t and h_t should be chosen such that

$$\alpha_t^Z + \sigma_t^Z h_t + \beta_t^Z diag\left(1^{J\times 1} + g_t\right)\mu_t - r_t Z_t = 0.$$

We have that also $\frac{U}{Z^0}$ is a martingale under Q and (2.9) is seen to be written on semimartingale form under Q. Thus, requiring that also $\frac{F}{Z^0}$ is a martingale under Q gives the equation

$$TD\left(\alpha_t + \sigma_t h_t, diag\left(1^{J\times 1} + g_t\right)\mu_t\right) = 0,$$

which constitutes a generalized version of Thiele's differential equation (TDE). In Section 2.7, we recognize the classical version from Hoem [33].

Adding to TDE the side conditions $V_{0-} = 0$ and $V_{T-} = \Delta B_T$, we formulate our result as a theorem:

Theorem 1 *Assume that the partial derivatives $\partial_t V$, $\partial_s V$, and $\partial_{ss} V$ exist and are continuous. Assume that (g, h) can be chosen such that*

$$\alpha_t^Z + \sigma_t^Z h_t + \beta_t^Z diag\left(1^{J\times 1} + g_t\right)\mu_t - r_t Z_t = 0. \tag{2.10}$$

Then, if the arbitrage free reserve on an insurance contract B can be written in the form $V(t, S_t)$, $V(t,s)$ solves for some (g,h) subject to (2.10) the deterministic differential equation (coefficients are (t,s) and (T,s), respectively)

$$\begin{aligned} \partial_t V_t &= b^c + r_t V_t - \left(\partial_s V_t\right)^T\left(\alpha_t + \sigma_t h_t\right) - R_t diag\left(1^{J\times 1} + g_t\right)\mu_t - \psi_t, \\ V_{T-} &= \Delta B_T. \end{aligned}$$

An arbitrage free insurance contract fulfills the equivalence relation

$$V_{0-} = 0. \tag{2.11}$$

Although the semimartingale form of F under P was not needed in our derivation of TDE, it may be interesting for other reasons. After some straightforward calculations one gets

$$\begin{aligned} dZ_t &= \left(r_t Z_t + \beta_t^Z \xi_t + \sigma_t^Z \eta_t\right)dt + \beta_{t-}^Z dM_t + \sigma_t^Z dW_t, \\ dF_t &= \left(r_t F_t + \left(\beta_t^F + \frac{\theta_t L_t}{U_t}\beta_t^Z\right)\xi_t + \left(\sigma_t^F + \frac{\theta_t L_t}{U_t}\sigma_t^Z\right)\eta_t\right)dt \\ &\quad + \left(\beta_{t-}^F + \frac{\theta_t L_{t-}}{U_{t-}}\beta_{t-}^Z\right)dM_t + \left(\sigma_t^F + \frac{\theta_t L_t}{U_t}\sigma_t^Z\right)dW_t. \end{aligned} \tag{2.12}$$

This representation of (Z, F) motivates the term *market price of risk* and shows how the expected return on the marketed indices, now including F, is increased compared to the return on the asset Z^0. From Theorem 1 it is seen that the price process V does not depend on θ, but (2.12) shows that the price process F indeed does. This implies that when it comes to laying down the payment process B, the only marketed indices of importance are those actually appearing as indices in B. Only if we consider the price process F, the remaining entries of Z, i.e. those entries of S that play the role of investment possibilities but do not appear as indices in B, are important. In the succeeding section we consider examples of insurance contracts. Since we focus on the process V, we let the market comprise only those assets on which payments depend. The representation in (2.12) may be an appropriate starting point for choice of an admissible strategy θ, taking into account e.g. the preferences of (the owners of) the insurance company.

Here we finish the general study of the process (V, F) by pinning down its stochastic representation formula. The traditionally educated life insurance actuary may rejoice at recognizing the reserve as an expected value. We have postponed the representation of the reserve as an expected value in order to emphasize that this is rather a fortunate consequence of the no arbitrage condition than a (measure-adjusted) consequence of traditional actuarial reasoning. In order to prevent arbitrage possibilities we have constructed Q such that $\left(\frac{F}{Z^0}, \frac{U}{Z^0}\right)$ under Q is a martingale, and then it follows that

$$\begin{aligned}
\frac{F_t}{Z_t^0} &= E^Q\left(\left.\frac{F_T}{Z_T^0}\right|\mathcal{F}_t^S\right) \\
&= E^Q\left(\left.\frac{1}{Z_T^0}\int_0^t \frac{U_T}{U_s}dB_s\right|\mathcal{F}_t^S\right) + E^Q\left(\left.\frac{1}{Z_T^0}\int_t^T \frac{U_T}{U_s}dB_s\right|\mathcal{F}_t^S\right) \\
&= E^Q\left(\left.\frac{U_T}{Z_T^0}\right|\mathcal{F}_t^S\right)\int_0^t \frac{1}{U_s}dB_s + \int_t^T E^Q\left(\left.\frac{1}{U_s}dB_s E^Q\left(\left.\frac{U_T}{Z_T^0}\right|\mathcal{F}_s^S\right)\right|\mathcal{F}_t^S\right) \\
&= \frac{U_t}{Z_t^0}\int_0^t \frac{1}{U_s}dB_s + \int_t^T E^Q\left(\left.\frac{1}{U_s}dB_s\frac{U_s}{Z_s^0}\right|\mathcal{F}_t^S\right) \\
&= \frac{L_t}{Z_t^0} + E^Q\left(\left.\int_t^T \frac{1}{Z_s^0}dB_s\right|\mathcal{F}_t^S\right).
\end{aligned}$$

Thus,

$$V_t = E^Q\left(\left.Z_t^0\int_t^T \frac{1}{Z_s^0}d\left(-B_s\right)\right|\mathcal{F}_t^S\right), \tag{2.13}$$

and the equivalence relation $V_{0-} = 0$, can be written

$$E^Q\left(\int_{0-}^T \frac{1}{Z_s^0}dB_s\right) = 0.$$

A special case of this constraint is known in actuarial mathematics as the equivalence principle, namely the case of risk-neutrality. In general, market prices of risk are not zero, and here we have taken into account the existence of a market Z which may contain information of these market prices of risk.

A calculation similar to the one leading to (2.13) shows that the equivalence relation $V_{0-} = 0$ corresponds to the alternative equivalence relation

$$E^Q\left(\frac{L_T}{Z_T^0}\right) = 0, \tag{2.14}$$

Thus, instead of checking the condition (2.11) for a payment process, one can just as well check the condition (2.14).

The representation in (2.13) explains why the restricted class of measure transformations corresponds to reserves in the form $V(t, S_t)$. The structure of Λ determined by (2.7) is necessary and sufficient for the following relation to hold,

$$\begin{aligned} V_t &= E^Q\left(Z_t^0 \int_t^T \frac{1}{Z_s^0} d(-B_s) \middle| \mathcal{F}_t^S\right) \\ &= E^P\left(\frac{\Lambda_T}{\Lambda_t} Z_t^0 \int_t^T \frac{1}{Z_s^0} d(-B_s) \middle| \mathcal{F}_t^S\right) \\ &= E^P\left(\frac{\Lambda_T}{\Lambda_t} Z_t^0 \int_t^T \frac{1}{Z_s^0} d(-B_s) \middle| S_t\right) \\ &= V(t, S_t). \end{aligned}$$

2.7 Examples

2.7.1 A classical policy

In this section we consider a model where payments depend on the present state of X. Hoem [33] obtained in this model a version of Thiele's differential equation which has taken a central position in life insurance mathematics and is widely used by practitioners.

Let r be constant, put $K = 0$, define S by

	α_t	β_t^j	σ_t^k	s_0
S^0	rS_t^0	0	0	1
S^1	0	$j - S_t^1$	0	X_0

(2.15)

and let $Z = S^0$. Thus the market Z consists of the risk-free asset only and contains thereby no information on market prices of risk. We fix a martingale measure by assuming that the insurance company is risk-neutral with respect to risk due to the policy state, which is the only risk present in this model, i.e.

$$g = 0^{J \times 1}.$$

Then the reserve function solves the classical TDE

$$\begin{aligned} \partial_t V_t &= b_t^c + r_t V_t - \left(b_t^d + V_t^{\mathcal{J}} - V_t^{1\times J}\right) \mu_t, \\ V_{T-} &= \Delta B_T, \end{aligned}$$

and the payment process should be based on the equivalence relation

$$V_{0-} = 0.$$

2.7.2 A simple unit-linked policy

In this section we consider a model where payments depend on the present state of X and the present state of a marketed index given by a geometric Brownian motion. Aase and Persson [1] obtained a version of Thiele's differential equation in a similar model. We use the word simple since the payments depend on only the present state of the marketed index. Modelling the marketed index by a geometric Brownian motion, we now work with the Black-Scholes model.

Let r be constant, put $K = 1$, define S by adding to (2.15)

	α_t	β_t^j	σ_t^k	s_0
S^2	$\overline{\alpha}S_t^2$	0	$\overline{\sigma}S_t^2$	S_0^2

(2.16)

where $\overline{\alpha}$ and $\overline{\sigma}$ are constant, and let $Z = \begin{bmatrix} S^0 \\ S^2 \end{bmatrix}$.

Theorem 1 states that (g, h) should be chosen subject to

$$\overline{\alpha}S_t^2 + \overline{\sigma}S_t^2 h_t - rS_t^2 = 0,$$

implying that

$$h = \frac{r - \overline{\alpha}}{\overline{\sigma}}.$$

With g determined as in Section 2.7.1, the reserve function solves the TDE

$$\begin{aligned} \partial_t V_t &= b_t^c + r_t V_t - \partial_{s^2} V_t r_t s^2 - \left(b_t^d + V_t^{\mathcal{J}} - V_t^{1\times J}\right)\mu_t - \frac{1}{2}\overline{\sigma}s^2 \partial_{s^2 s^2} V_t \overline{\sigma} s^2, \\ V_{T-} &= \Delta B_T, \end{aligned}$$

and the payment process should be based on the equivalence relation

$$V_{0-} = 0.$$

2.7.3 A path-dependent unit-linked policy

In this section we consider a model where payments depend on the present state of X and the present state of two accounts. One account is defined as the value of as-if investments in a marketed index and the other is defined as the value of as-if investments in a non marketed index. We call it a path-dependent policy since the payments depend on the path of the marketed index through the account defined by the value of as-if investments in this index. Modelling the marketed index by a geometric Brownian motion, the market Z constitutes the Black-Scholes model, and as in Section 2.7.2 the geometric Brownian motion will be described by S^2.

Now we construct the process S. First we introduce an artificial bank account $\widehat{S^0}$ given by

$$d\widehat{S_t^0} = \widehat{r}\widehat{S_t^0}dt, \ \widehat{S_0^0} = 1,$$

where $\widehat{r}$ is an artificial short rate of interest, which may differ from the short rate of interest r.

The insurance contract specifies two artificial payment processes, $B^{\widehat{S^0}}$ and B^{S^2}, defined in the same way as B. The notation $b_t^{\widehat{S^0}c}$, $b_t^{S^2c}$, $b_t^{\widehat{S^0}d}$, $b_t^{S^2d}$ for artificial continuous and jump payments, respectively, is natural. Now we make up two artificial accounts $A^{\widehat{S^0}}$ and A^{S^2} by pretending to invest payments from $B^{\widehat{S^0}}$ and B^{S^2} in $\widehat{S^0}$ and S^2, respectively. Consequently, the two accounts can be written

$$\begin{aligned} A_t^{\widehat{S^0}} &= \widehat{S_t^0} \int_0^t \frac{dB_s^{\widehat{S^0}}}{\widehat{S_s^0}}, \\ A_t^{S^2} &= S_t^2 \int_0^t \frac{dB_s^{S^2}}{S_s^2}. \end{aligned}$$

The payment process B is specified to depend on these two accounts. It is easily seen that if we put $K = 1$ and define S by adding to (2.15) and (2.16)

	α_t	β_t^j	σ_t^k	s_0
S^3	$\frac{b_t^{S^2c}}{S_t^2}$	$\frac{b_t^{S^2dj}}{S_t^2}$	0	0
S^4	$b_t^{\widehat{S^0}c} + \widehat{r} S_t^4$	$b_t^{\widehat{S^0}dj}$	0	0

we actually have that

$$\begin{aligned} A_t^{\widehat{S^0}} &= S_t^4, \\ A_t^{S^2} &= S_t^2 S_t^3. \end{aligned}$$

As an example one could think of a policy specifying that e.g. continuous premiums b_t^c are invested in the pseudo bank account, the payment process $B_t^{S^2} = t$ is invested in the index S^2, and the benefit payment b_t^d is the maximum of the two accounts made up hereby. In that case one could interpret $\widehat{r}$ as the guaranteed rate of interest and the pseudo bank account thus represents one way of introducing an interest rate guarantee. By this example we point out that elements of the payment process B are allowed to appear in the processes B^{S^2} and $B^{\widehat{S^0}}$.

Bacinello and Ortu [2] and Nielsen and Sandmann [49] have studied a variant of this set-up and the occurrence of elements of B in $B^{\widehat{S^0}}$ is exactly what the endogeneity in the title of Bacinello and Ortu [2] refers to. Bacinello and Ortu [2] indicate that the resulting insurance product is an Asian-like derivative, and comparing our construction of state variables with the one known from theory of Asian options shows in which way the derivative is Asian-like. Bacinello and Ortu [2] end up with a delicate fix point problem and discuss conditions for existence of a payment process satisfying the equivalence relation $V_0 = 0$ (no lump sum payment at time 0). These conditions constrain the two accounts and the payments depending on them. Apart from the mathematical conditions, one could deal with conditions of the payments being practically feasible. However, we shall not enter into any of these discussions, but rather allow of general contributions to the two accounts as described above.

With (g, h) determined as in 2.7.2, the TDE can now easily be found by Theorem 1.

Chapter 3

Contingent claims analysis in life and pension insurance

The application of mathematical finance to unit-linked life insurance is unified with the theory of distribution of surplus in life and pension insurance. The unification is based a consideration of distribution of surplus as an integrated part of the insurance contract. We suggest a distinction between the retrospective surplus and the prospective surplus and study these versions of the surplus in detail. The retrospective surplus and the prospective surplus are proposed as indices in a type of index linked insurance which we appropriately call surplus-linked insurance. This chapter is an extended version of Steffensen [65].

3.1 Introduction

The term life and pension insurance is here used for the general type of life insurance where premiums and benefits are calculated on a certain basis at the time of issue of the contract and then revised currently according to the performance of the insurance company. The revision of premiums and benefits can take various forms depending on the type of the contract. Examples are various types of participating life insurance (in some countries called with-profit life insurance) and various types of pension funding.

The revision of premiums and benefits is based on payment of *dividends*, which in general may be positive or negative, from the insurance company to the policy holder. It is important to distinguish between two aspects of the revision, the *dividend plan* and the *bonus plan*. The dividend plan is the plan for recording of dividends. However, often the dividends are not paid out immediately as cash but are converted into a stream of future payments. The bonus plan is the plan for how the dividends eventually are turned into payments.

In Chapter 2 a framework of securitization is developed where reserves are no longer defined as expected present values but as market prices of streams of payments (which, however, happen to be expressible as expected present values under adjusted measures). An insurance contract is defined as a stream of payments linked to dynamical indices, opening for a wide range of insurance contracts including various forms of unit-linked contracts. Securitization is one way of dealing with the dependence between the risk in the insurance policy and the risk in the financial market. It is built on the consideration of the stream of payments stipulated in an insurance contract as a dynamically traded object on the financial market. The insurance company is then considered as a participant in this market and has to adapt prices and strategies to the market conditions.

In the present chapter we construct a general life and pension insurance contract within the framework developed in Chapter 2. Working with general index-linked payments in participating life insurance and pension funding we go beyond the traditional set-up of payments in existing literature on emergence of surplus and dividends. However, index-linked payments open for a number of appealing set-ups. An important one is obtained by linking payments directly to the surplus as it will be defined in this chapter. The study of such surplus-linked insurance has a two-fold motivation: Firstly, it represents a new product combining properties of participating life insurance, pension funding, and unit-linked insurance. Secondly, it seems to represent a good imitation of the behavior of managers. As such it can be used as a management tool as well as a market analysis tool.

Even though payments need not be linked directly to the surplus, the surplus may carry great weight when the insurance company decides on dividends. It is one of the main purposes of this chapter to provide insight in the dynamics of the surplus, and an important step is the classification of surplus into the retrospective surplus and the prospective surplus. The recognition of previous definitions of surplus as either a part of the retrospective surplus or a part of the prospective surplus contributes to the insight in the notion of surplus and constitutes, together with other remarks, a whole section of comparisons with related literature in the field.

Subjugating life and pension insurance to the market conditions, the appropriate tool seems to be mathematical finance or, more specifically, contingent claims analysis. Option pricing theory was introduced as a tool for analysis and management of unit-linked insurance in the seventies (see Brennan and Schwartz [10] and references in Aase and Persson [1]). Also the consideration of schemes in pension funding as options goes back to the seventies (see e.g. Sharpe [62] and references in Blake [7]). In participating life insurance, however, contingent claims analysis as a tool for analysis and management has been long in coming and was, to our knowledge, introduced in Briys and de Varenne [11]. A main reason for this delay may be that the link between the payments and the performance of the company in participating life insurance may be laid down by statute so vaguely that it may seem unreasonable to consider dividends as contractual. Working in a framework of securitization, our main objection to this argument is, of course, that the insurance business and, hereby, the participation in the performance takes place in a competitive market. Thus, the insurance company is forced to adapt e.g. its plans for revision of payments to the market conditions. This objection is at the same time the primary argument for applying contingent claims analysis to life and pension insurance at all.

The chapter is structured as follows. In Section 3.2, we recapitulate the framework developed in Chapter 2. In Section 3.3, we construct the general life and pension insurance contract within that framework and extract some decision problems of a life insurance company. The notions of retrospective and prospective surplus are defined and studied in Section 3.4. In Section 3.5, we consider dividend plans in general and surplus-linked dividends in particular, whereas various bonus plans are studied in Section 3.6. Aspects of our approach is compared to related literature in Section 3.7, and Section 3.8 reviews the balance sheet of a life insurance company in the light of our terminology and definitions. Section 3.9 contains a few numerical illustrations. These are based on a main example which already at the end of Sections 3.3-3.6 serves to illustrate the contents of each section.

3.2 The insurance contract

3.2.1 The basics

In this section we recapitulate the framework developed in Chapter 2 and state the main result obtained there. For motivation, details, and examples the reader is asked to confer Chapter 2.

We take as given a probability space $(\Omega, \mathcal{F}, \mathbf{F} = \{\mathcal{F}_t\}_{t\geq 0}, P)$. We let $(X_t)_{t\geq 0}$ be a cadlag (i.e. its sample paths are almost surely right continuous with left limits) jump process with finite state space $\mathcal{J} = (1, \ldots, J)$ defined on $(\Omega, \mathcal{F}, \mathbf{F} = \{\mathcal{F}_t\}_{t\geq 0}, P)$ and associate a marked point process (T_n, Φ_n), where T_n denotes the time of the nth jump of X_t, and Φ_n is the state entered at time T_n, i.e. $X_{T_n} = \Phi_n$. We introduce the counting processes

$$N_t^j = \sum_{n=1}^{\infty} 1_{(T_n \leq t, X_{T_n} = j)}, \quad j \in \mathcal{J},$$

and the J-dimensional vector

$$N_t = \begin{bmatrix} N_t^1 \\ \vdots \\ N_t^J \end{bmatrix}.$$

We let $(W_t)_{t\geq 0} = \left(W_t^1, \ldots W_t^K\right)_{t\geq 0}$ be a standard K-dimensional Brownian motion defined on $(\Omega, \mathcal{F}, \mathbf{F} = \{\mathcal{F}_t\}_{t\geq 0}, P)$.

For a matrix A we let A^T denote the transpose of A and let $A^{i\cdot}$ and $A^{\cdot i}$ denote the ith row and the ith column of A, respectively. For a vector a, we let $diag\,(a)$ denote the diagonal matrix with the components of a in the principal diagonal and 0 elsewhere. We shall write $\delta^{1\times J}$ and $\delta^{J\times 1}$ instead of $(\delta, \ldots, \delta)$ and $(\delta, \ldots, \delta)^T$, respectively. For derivatives we shall use the notation $\partial_x = \frac{\partial}{\partial x}$ and $\partial_{xy} = \frac{\partial^2}{\partial x \partial y}$. For a vector a we let $\int a$ and da mean componentwise integration and componentwise differentiation, respectively.

We introduce an *index* S, an $(I+1)$-dimensional vector of processes, the dynamics of which is given by

$$dS_t = \alpha_t dt + \beta_{t-} dN_t + \sigma_t dW_t,\ S_0 = s_0,$$

where $\alpha \in \mathbf{R}^{(I+1)}$, $\beta \in \mathbf{R}^{(I+1)\times J}$, and $\sigma \in \mathbf{R}^{(I+1)\times K}$ are functions of (t, S_t) and $s_0 \in \mathbf{R}^{I+1}$ is $\mathcal{F}_0$-measurable. We denote by S^i, α^i, β^{ij}, and σ^{ik} the ith entry of S, the ith entry of α, the (i,j)th entry of β, and (i,k)th entry of σ, respectively. The information generated by S is formalized by the filtration $\mathbf{F}^S = \left\{\mathcal{F}_t^S\right\}_{t\geq 0}$, where

$$\mathcal{F}_t^S = \sigma(S_s,\ 0 \leq s \leq t) \subseteq \mathcal{F}_t.$$

We assume that S is a Markov process and that there exist deterministic piecewise continuous functions $\mu^j(t,s)$, $j \in \mathcal{J}$, $s \in \mathbf{R}^{I+1}$ such that N_t^j admits the $\mathcal{F}_t^S$-intensity process $\mu_t^j = \mu^j(t, S_t)$, informally given by

$$\begin{aligned} \mu_t^j dt &= E\left(dN_t^j \middle| \mathcal{F}_{t-}^S\right) + o(dt) \\ &= E\left(dN_t^j \middle| S_{t-}\right) + o(dt), \end{aligned}$$

where $o(h)/h \to 0$ as $h \to 0$. We introduce the J-dimensional vectors containing the intensity processes and martingales associated with N,

$$\mu_t = \begin{bmatrix} \mu_t^1 \\ \vdots \\ \mu_t^J \end{bmatrix}, \quad M_t = \begin{bmatrix} M_t^1 \\ \vdots \\ M_t^J \end{bmatrix} = \begin{bmatrix} N_t^1 - \int_0^t \mu_s^1 ds \\ \vdots \\ N_t^J - \int_0^t \mu_s^J ds \end{bmatrix}.$$

To help the reader fix ideas, we explain briefly the roles of the introduced processes. Their roles will become more clear when we formalize the payment process below. The process N describes (at least) some specification of the life history of an insured. Whereas the process N will partly determine the points in time where payments fall due, the process S determines the amounts of these payments (and the intensities for the process N). In classical life insurance mathematics, payments are allowed to depend on the state of the policy, X. We can cover this situation by taking S^1 to be equal to X by the coefficients

$$\alpha_t^1 = 0,\ \beta_t^{1j} = j - S_t^1,\ \sigma_t = 0,\ s_0^1 = X_0.$$

If e.g. X is included in the index S, $\mu(t, X_t)$ candidates to the intensity process corresponding to the classical situation, see e.g. Hoem [33]. However, in general, the intensity process μ may differ from the intensity process with respect to the natural filtration of N.

However, this classical contract can be extended in various directions. We can e.g. allow for payments (and intensities) to depend on the duration of the sojourn in the current state by letting S^2 be defined by

$$\alpha_t^2 = 1,\ \beta_t^2 = -S_t^2,\ \sigma_t^2 = 0,\ s_0^2 = 0,$$

and allow for payments (and intensities) to depend on the total number of jumps by letting S^3 be defined by

$$\alpha_t^3 = 0,\, \beta_t^{3j} = 1,\, \sigma_t^3 = 0,\, s_0^3 = 0.$$

In Møller [47] and Norberg [52] generalized versions of Thiele's differential equation have been studied where payments depend on the duration of the sojourn in the current state.

We introduce a *market* Z, an $(n+1)$-dimensional vector $(n \leq I)$ of price processes assumed to be positive, and denote by Z^i the ith entry of Z. The market Z consists of exactly those entries of S that are prices of traded assets. We assume that there exists a short rate of interest such that the market contains a price process Z^0 with the dynamics given by

$$dZ_t^0 = r_t Z_t^0 dt, Z_0^0 = 1.$$

This price process can be considered as the value process of a unit deposited on a bank account at time 0, and we shall call this entry for the risk-free asset even though r_t is allowed to depend on (t, S_t). Furthermore, we assume that the set of martingale measures, $\mathcal{Q}$, i.e. the set of probability measures Q equivalent to P such that $\frac{Z^i}{Z^0}$ is a Q-martingale for each i, is non-empty. From fundamental theory of asset pricing this assumption is known to be essentially equivalent to the assumption that no arbitrage possibilities exist on the market Z. The entries of an index S will also be called indices, and the indices appearing in Z will then be called marketed indices or assets. With this formulation the set of marketed indices is a subset of the set of indices and it contains at least one entry, namely Z^0. We let $\alpha^Z \in \mathbf{R}^{(n+1)}$, $\beta^Z \in \mathbf{R}^{(n+1)\times J}$, and $\sigma^Z \in \mathbf{R}^{(n+1)\times K}$ denote the coefficients of the asset prices Z.

Fixing some time horizon T, we now formally take an *insurance contract* to be a *payment process* B which is an $\mathcal{F}_t^S$-adapted, cadlag process of finite variation with dynamics given by

$$dB_t = B_0 d1_{(t\geq 0)} + b_t^c dt - b_{t-}^d dN_t - \Delta B_T d1_{(t\geq T)},$$

where $B_0 \in \mathbf{R}$ is a function of S_0, $b^c \in \mathbf{R}$ and $b^d \in \mathbf{R}^J$ are functions of (t, S_t), and $\Delta B_T \in \mathbf{R}$ is a function of S_T. We denote by b^{dj} the jth entry of b^d. Note that the $\mathcal{F}_t^S$-adaptedness of B places demands on the connection between the coefficients of S and the coefficients of B. Although it need not be the case, the reader should have in mind the case where $\mathcal{F}_0$ and thus also $\mathcal{F}_0^S$ are trivial, i.e. $\mathcal{F}_0 = \mathcal{F}_0^S = \{\emptyset, \Omega\}$. Then B_0 is deterministic.

B_t represents the cumulative payments from the policy holder to the insurance company over $[0, t]$. Both continuous payments and lump sum payments are thus allowed to depend on the present state of the process (t, S_t). The minus signs in front of b^d and ΔB in dB_t conform to the typical situation where B_0 and b^c are premiums and b^d and ΔB are benefits, all positive. To simplify notation, lump sum payments at deterministic times are restricted to time 0 and time T. Thus, an insurance contract is given by a set of functions $\left(B_0, b^c, b^d, \Delta B\right)$ such that a recording of S completely determines the payment stream.

The insurance contract forms the basis for introduction of two price processes, F and V:

$$\begin{aligned} F_t &= \text{the price at time } t \text{ of the contractual payments to the} \\ &\quad \text{insurance company over } [0,T]\text{, i.e. premiums less benefits,} \\ V_t &= \text{the price at time } t \text{ of the contractual payments from the} \\ &\quad \text{insurance company over } (t,T]\text{, i.e. benefits less premiums.} \end{aligned}$$

3.2.2 The main result

Our approach to the price process F is the following: Assuming that the market Z is arbitrage free, we require that also the market (Z, F) be arbitrage free. We use the essential equivalence between arbitrage free markets and existence of a so-called martingale measure, i.e. a measure under which discounted asset prices are martingales. If the no arbitrage condition is fulfilled for (Z, F), we shall speak of B as an arbitrage free insurance contract and about V as the corresponding arbitrage free reserve.

Since the market may be incomplete, there may be several martingale measures and, correspondingly, several arbitrage free reserves. Thus, when we talk of *the* arbitrage free reserve, we think of having fixed a martingale measure according to some criterion. Alternatively, one could imagine that there exists only one martingale measure reflecting the market participants' attitudes to risk although this measure, in the incomplete market, is not to be identified by looking at asset prices only. In this case *the* martingale measure could, appropriately, be fixed as the unique measure reflecting the attitudes to risk.

We restrict ourselves to prices allowing V_t to be written in the form $V(t, S_t)$. This restriction seems reasonable since S is Markov and since the payments by B and the intensities of N depend only on time and the current value of S, but it is actually a restrictive assumption on the formation of prices in the market. It corresponds to the restrictive structure of the measure transformation that we now enter by defining the likelihood process Λ by

$$\begin{aligned} d\Lambda_t &= \Lambda_{t-}\left(\sum_j g^j_{t-} dM^j_t + \sum_k h^k_t dW^k_t\right) \\ &= \Lambda_{t-}\left(g^T_{t-} dM_t + h^T_t dW_t\right), \\ \Lambda_0 &= 1, \end{aligned}$$

where we have introduced

$$g^j_t = g^j(t, S_t), \; h^k_t = h^k(t, S_t),$$

and

$$g_t = \begin{bmatrix} g_t^1 \\ \vdots \\ g_t^J \end{bmatrix}, h_t = \begin{bmatrix} h_t^1 \\ \vdots \\ h_t^K \end{bmatrix}.$$

With conditions on (g, h) (see Chapter 2) we can now change measure from P to Q on (Ω, F_T) by the definition,

$$\Lambda_T = \frac{dQ}{dP}.$$

Upon introducing

$$\begin{aligned} V_t^j &= V\left(t, S_t + \beta_t^{\cdot j}\right), \\ V_t^{\mathcal{J}} &= \left[V_t^1, \ldots, V_t^J\right], \\ \psi_t &= \frac{1}{2} tr\left(\sigma_t^T \partial_{ss} V_t \sigma_t\right), \\ R_t &= b_t^d + V_t^{\mathcal{J}} - V_t^{1\times J}, \end{aligned}$$

we can state the main result of Chapter 2:

Theorem 2 *Assume that the partial derivatives $\partial_t V$, $\partial_s V$, and $\partial_{ss} V$ exist and are continuous. Assume that (g, h) can be chosen such that*

$$\alpha_t^Z + \sigma_t^Z h_t + \beta_t^Z diag\left(1^{J\times 1} + g_t\right) \mu_t - r_t Z_t = 0. \tag{3.1}$$

If the arbitrage free reserve on an insurance contract B can be written in the form $V(t, S_t)$, then $V(t, s)$ solves for some (g, h) subject to (3.1) the deterministic differential equation (coefficients are (t, s) and (T, s), respectively)

$$\begin{aligned} \partial_t V_t &= b_t^c + r_t V_t - \left(\partial_s V_t\right)^T \left(\alpha_t + \sigma_t h_t\right) - R_t diag\left(1^{J\times 1} + g_t\right) \mu_t - \psi_t, \quad (3.2) \\ V_{T-} &= \Delta B_T, \end{aligned}$$

and has the representation

$$V_t = E^Q\left(\int_t^T \frac{Z_t^0}{Z_s^0} d\left(-B_s\right) \middle| S_t\right). \tag{3.3}$$

An arbitrage free insurance contract fulfills the equivalence relation

$$V_{0-} = 0. \tag{3.4}$$

The relation (3.1) is just the standard condition for existence of a martingale measure on Z and arises solely from the dynamics of Z. The relation (3.3) can be derived directly from existence of a martingale measure on (Z, F). For calculation of V_t, (3.2) yields a constructive tool such that an arbitrage free insurance contract can be constructed subject to the side condition (3.4). The equation (3.2) generalizes the classical deterministic differential equation for a reserve introduced by Thiele in 1875, and we will simply speak of (3.2) as Thiele's differential equation.

3.3 The general life and pension insurance contract

3.3.1 The first order basis and the technical basis

In this section we formulate the structure of the general life and pension insurance contract within the framework recapitulated in Section 3.2.

We introduce a *first order basis*, $(\widehat{r}, \widehat{g}, \widehat{h})$, let the first order short rate of interest $\widehat{r}$ drive a first order risk-free asset $\widehat{Z}$ (not a part of the market Z), and let the first order Girsanov kernel $(\widehat{g}, \widehat{h})$ determine a first order measure $\widehat{Q}$. Here, $\widehat{r}$, $\widehat{g}$, and $\widehat{h}$ are functions of (t, S_t). We define a stream of *first order payments* $\widehat{B}$ in the same way as B is defined in Section 3.2, i.e. linked to (t, S_t), and we define the *first order reserve* by

$$\widehat{V}_t = E^{\widehat{Q}}\left(\int_t^T \frac{\widehat{Z}_t}{\widehat{Z}_s} d\left(-\widehat{B}_s\right) \middle| S_t \right).$$

Then, upon introducing

$$\begin{aligned} \widehat{V}_t^j &= \widehat{V}\left(t, S_{t-} + \beta_{t-}^{\cdot j}\right), \\ \widehat{V}_t^{\mathcal{J}} &= \left[\widehat{V}_t^1, \ldots, \widehat{V}_t^J\right], \\ \widehat{R}_t &= \widehat{b}_t^d + \widehat{V}_t^{\mathcal{J}} - \widehat{V}_t^{1\times J}, \\ \widehat{\psi}_t &= \frac{1}{2} tr\left(\sigma_t^T \partial_{ss} \widehat{V}_t \sigma_t\right), \end{aligned}$$

we have, according to Theorem 1, the first order Thiele's differential equation and the first order terminal condition,

$$\begin{aligned} \partial_t \widehat{V}_t &= \widehat{b}_t^c + \widehat{r}_t \widehat{V}_t - \left(\partial_s \widehat{V}_t\right)^T \left(\alpha_t + \sigma_t \widehat{h}_t\right) - \widehat{R}_t diag\left(1^{J\times 1} + \widehat{g}_t\right) \mu_t - \widehat{\psi}_t, \quad (3.5) \\ \widehat{V}_{T-} &= \Delta \widehat{B}_T. \end{aligned}$$

We let $\widehat{B}$ be constrained by the first order prospective equivalence relation,

$$\widehat{V}_{0-} = 0.$$

Actually, this construction of the first order payments amounts to requiring that $\widehat{B}$ be an arbitrage free contract on an artificial market with only one asset, $\widehat{Z}$.

The first order basis serves solely to determine the first order payments at time 0. However, also during the term of the contract, the insurance company needs to valuate the first order payments for different operations. The appropriate conditions for such a valuation depend on what operation is performed. We shall introduce a *technical basis*, (r^*, g^*, h^*) with functions r^*, g^*, and h^* of (t, S_t) for such a valuation of first order payments and define the *technical reserve* as

$$V_t^* = E^{Q^*}\left(\int_t^T \frac{Z_t^*}{Z_s^*} d\left(-\widehat{B}_s\right) \middle| S_t \right).$$

The technical Thiele's differential equation and the technical terminal condition are now obtained upon replacing $(\widehat{r}, \widehat{g}, \widehat{h}, \widehat{V}, \widehat{R})$ in (3.5) with $(r^*, g^*, h^*, V^*, R^*)$, where $R^* = \widehat{b}_t^d + V_t^{*\mathcal{J}} - V_t^{*1\times J}$. Note that there exists no technical equivalence relationships. Only if the first order basis is used as technical basis, the relation $V_{0-}^* = 0$ holds.

The technical basis plays a role in the operation of reporting to the owners of the company and to the supervisory authorities. The insurance company may draw up a statement of accounts at market value if the owners of the company and/or the supervising authorities want a true picture of the company. For this operation the basis given by (r, g, h) seems to be an obvious choice for technical basis. However, specific conditions for solvency may be formulated under another basis and the supervisory authorities may require a presentation of accounts on such a basis. Such a basis could e.g. be the first order basis.

Thus, the first order basis and the basis (r, g, h) certainly candidate to the technical basis, but other technical bases may apply. This conforms with recent accounting rules in Denmark, where the insurance companies set aside reserves on a basis that differs from the first order basis (and probably also from the real basis) on portfolios where the first order reserves seems not to be adequate in some sense.

3.3.2 The real basis and the dividends

As opposed to the first order basis $(\widehat{r}, \widehat{g}, \widehat{h})$ and the technical basis (r^*, g^*, h^*) we shall speak of (r, g, h) as the real basis. Since the first order basis may differ from the real basis, the first order payments may impose arbitrage possibilities in the real environment. However, the *real payments* B are to be determined such that B constitutes an arbitrage free insurance contract in the real environment. The real payments are composed by the first order payments and an additional payment stream $\widetilde{B}$ called the *dividends*, i.e.

$$B = \widehat{B} + \widetilde{B}. \tag{3.6}$$

Note that both $\widehat{B}$ and $\widetilde{B}$ are payments to the insurance company such that e.g. dividend payments to the policy holder will appear with a minus sign in $\widetilde{B}$. We want to work within the framework of Section 3.2, and we are therefore interested in index-linked dividends. The index to which dividends are linked may be the same as the one to which the first order payments are linked. However, we may also augment this index with further state variables.

The formulas of Theorem 1 then read the real martingale measure constraint, the real Thiele's differential equation, the real expected value representation, and the real equivalence relationships. If the dividends are designed in such a way that the contract B is arbitrage free, i.e. the real equivalence relation holds, we shall simply say that the dividends are arbitrage free. We shall be interested in designing the dividends in such a manner that they are index-linked and arbitrage free.

The dividends rectify a possible imbalance between the first order basis and the real basis in the sense that we get from putting (3.6) into (3.3) and (3.4),

$$E^Q\left(\int_{0-}^{T}\frac{1}{Z_t^0}d\widetilde{B}_t\right) = -E^Q\left(\int_{0-}^{T}\frac{1}{Z_t^0}d\widehat{B}_t\right). \tag{3.7}$$

The sign of $E^Q\left(\int_{0-}^{T}\frac{1}{Z_t^0}d\widehat{B}_t\right)$ decides whether an insurance contract has positive or negative dividends in expectation. In particular, if the real basis is used as first order basis, then the expected dividends become zero. In this case, the dividends given by $\widetilde{B} = 0$ would, obviously, be arbitrage free, and the *unrevised* contract would be the appropriate name for this particular construction.

In participating life insurance the dividends are restricted to be to the policy holder's advantage, i.e. $\widetilde{B}$ must be a non-increasing process with $\widetilde{B}_0 \leq 0$. From (3.7) it is seen that there will exist arbitrage free dividend plans to the policy holder's advantage only if

$$E^Q\left(\int_{0-}^{T}\frac{1}{Z_t^0}d\widehat{B}_t\right) \geq 0. \tag{3.8}$$

On the other hand, if (3.8) is fulfilled, an arbitrage free dividend plan can easily be devised. We conclude that (3.8) is a necessary and sufficient condition on the relation between the first order basis and the real basis for existence of an arbitrage free dividend plan. The interpretation is that the insurance company cannot come up with dividends to the policy holder's advantage arbitrage freely if the first order payments are to the policy holder's advantage in the first place. But if the first order payments are to the policy holder's disadvantage, there will exist a continuum of arbitrage free dividend plans.

3.3.3 A delicate decision problem

When designing a life insurance product we face a delicate decision problem. First of all, we have to decide on a first order basis. Given this first order basis, we need to decide on a dividend plan such that the insurance contract becomes arbitrage free. One can think of many dividend plans, some of them rather obscure. To mention a few, one could e.g. pay out $\widetilde{B}_0 = -E^Q\left(\int_{0-}^{T}\frac{1}{Z_t^0}d\widehat{B}_t\right)$ as a deterministic lump sum payment at time 0 and thereby finish the revision of payments at time of issue. If the policy holder does not find this plan appealing, one could simply toss a coin to see whether the policy holder should receive a deterministic lump sum at time 0 or not. The size of the lump sum would depend on the market's attitude to toss-up. Note that this toss-up example describes a special case where $\mathcal{F}_0$ is not trivial. Usually, however, the policy holder is more interested in gambling on the financial market. We shall see how this can be obtained by letting some indices represent accounts that are invested on the market. Given such a construction, also the underlying investment strategy becomes a part of the decision problem.

We want to design products which in these decision aspects imitate the manager of a life insurance company. The problem is to come up with an appropriate index which, on the one hand, contains the information on which the manager bases the decisions and, on the other hand, is mathematical tractable, i.e. not dependent on "too many" state variables. We shall in the succeeding section study thoroughly the notion of surplus since this seems to be the all-important piece of information on which the manager bases the decisions concerning dividends. The surplus introduced in the succeeding section depends on the technical basis. Hereby determination of dividends is added to the list of operations for which a technical basis must be specified.

The decision is made subject to two basic constraints. Firstly, we have the arbitrage condition

$$V_{0-} = 0,$$

which appropriately could be called the market constraint. Secondly, we have the legislative constraints. They could e.g. simply put bounds on the first order rate of interest. More interesting are possible constraints on the relation between the dividends and the surplus. If such a relation is included in the legislative constraints, it is of course important that the insurance company and the supervisory authorities agree upon what surplus is and, possibly, on which technical basis it should be based.

3.3.4 Main example

We shall work with a simple insurance contract as an illustration. The insurance contract is a single life endowment insurance with a sum insured of 1 and a constant premium π paid as long as the insured is alive. The insurance contract is introduced on the Black-Scholes market defined below by coefficients of S^0 and S^2. Let X be the two-state process defined by $X_t = 0$ if the insured is alive at time t, $X_t = 1$ if the insured is dead at time t, and $X_0 = 0$. Since state 2 is absorbing, we need only one counting process N^1 counting the number of deaths, and for notational convenience we skip the topscript 1. Assume that the intensity of N is given by $\mu_t = \mu(t, X_t)$, let the first order intensity of N be given by $\widehat{\mu}_t = (1 + \widehat{g})\,\mu_t$ for some constant $\widehat{g}$, and let $\widehat{r}$ be a constant first order rate of interest. The first order diffusion kernel $\widehat{h}$ plays no role and can be chosen arbitrarily.

We define S by

	α_t	β_t	σ_t	s_0
S^0	rS_t^0	0	0	1
S^1	0	$1 - S_t^1$	0	0
S^2	$\overline{\alpha}S_t^2$	0	$\overline{\sigma}S_t^2$	S_0^2

where r, $\overline{\alpha}$, and $\overline{\sigma}$ are constant, and let $Z = \begin{bmatrix} S^0 \\ S^2 \end{bmatrix}$.

Theorem 1 states that (g, h) should be chosen subject to

$$\overline{\alpha}S_t^2 + \overline{\sigma}S_t^2 h_t - rS_t^2 = 0,$$

implying that

$$h = \frac{r - \overline{\alpha}}{\overline{\sigma}}.$$

The market contains no information on the price of mortality risk, and we fix a martingale measure by assuming risk neutrality with respect to mortality risk, i.e.

$$g = 0.$$

This example will serve to illustrate the contents of each of the succeeding sections. Throughout the example we will only consider the state-wise quantities (reserve, surplus, contributions, etc. all to be defined below) for $S_t^1 = 0$, i.e. the policy holder being alive. For notational convenience we will then skip the explicit dependence on this state variable in the formulas, i.e. $\widehat{V}_t \equiv \widehat{V}(t,0)$, $\widehat{\mu}_t = \widehat{\mu}(t,0)$. The system of deterministic differential equations for the reserve can then be written

$$\begin{aligned} \partial_t \widehat{V}_t &= \pi + \widehat{r}\widehat{V}_t - \left(1 - \widehat{V}_t\right)\widehat{\mu}_t, \\ \widehat{V}_{T-} &= 1, \end{aligned}$$

and π is determined by

$$\widehat{V}_0 = 0.$$

3.4 The notion of surplus

3.4.1 The investment strategy

The insurance company receives payments in accordance with the insurance contract B, and we assume that these are currently deposited on/drawn from an account which is invested in a portfolio with positive value process U, generated by a self-financing investment strategy $\theta \in \mathbf{R}^{n+1}$, i.e.

$$\begin{aligned} U_t &= \theta_t \cdot Z_t = \sum_{i=0}^{n} \theta_t^i Z_t^i > 0, \\ dU_t &= \theta_t \cdot dZ_t. \end{aligned}$$

The strategy is furthermore assumed to comply with whatever institutional requirements there may be. Throughout this chapter one can think of θ as the strategy corresponding to a constant relative portfolio, i.e. a strategy θ such that for a constant $(n+1)$-dimensional vector γ, $\theta_t^i Z_{t-}^i = \gamma^i U_{t-}$, $i = 0, \ldots, n$. This strategy reflects an investment profile possibly restricted by the supervising authorities, e.g. such that θ^i is non-negative for all i if short-selling is not allowed. We emphasize that θ, in general, is not a strategy aiming at hedging some contingent claim. Introduce α_t^U, β_t^U, and σ_t^U such that

$$dU_t = U_t \alpha_t^U dt + U_{t-} \beta_{t-}^U dM_t + U_t \sigma_t^U dW_t.$$

3.4.2 The retrospective surplus

We define the *retrospective surplus* $\overleftarrow{F}^*_t$ corresponding to a technical basis (r^*, g^*, h^*) by

$$\begin{aligned} \overleftarrow{F}^*_{0-} &= 0, \\ \overleftarrow{F}^*_t &= L_t - V^*_t,\ t \geq 0, \end{aligned}$$

where

$$L_t = U_t \int_{0-}^{t} \frac{1}{U_s} dB_s.$$

The real equivalence principle can be expressed in terms of the discounted retrospective surplus at time T: Using that $\frac{U}{Z^0}$ is a Q-martingale, one can show that (3.4) is equivalent to (see also (2.14))

$$E^Q\left(\frac{\overleftarrow{F}^*_T}{Z^0_T}\right) = 0. \tag{3.9}$$

Using Ito's formula on $\overleftarrow{F}^*_t$, we get

$$\begin{aligned} d\overleftarrow{F}^*_t &= dL_t - dV^*_t \\ &= b^c_t dt - b^d_{t-} dN_t + L_{t-}\frac{dU_t}{U_{t-}} - \left(\partial_t V^*_t + \left(\partial_s V^*_t\right)^T \alpha_t + \psi^*_t\right) dt \\ &\quad - \left(V^{*\mathcal{J}}_{t-} - V^{*1\times J}_{t-}\right) dN_t - \left(\partial_s V^*_t\right)^T \sigma_t dW_t. \end{aligned}$$

Subtraction of the technical Thiele's differential equation and some rearrangements give the form

$$\begin{aligned} d\overleftarrow{F}^*_t &= \overleftarrow{F}^*_{t-}\left(\alpha^U_t dt + \beta^U_{t-} dM_t + \sigma^U_t dW_t\right) \\ &\quad + \left(\left(\alpha^U_t - r^*_t\right) V^*_t + \left(\partial_s V^*_t\right)^T \sigma_t h^*_t + R^*_t diag\left(g^*_t\right) \mu_t\right) dt \\ &\quad + \left(V^*_t \sigma^U_t - \left(\partial_s V^*_t\right)^T \sigma_t\right) dW_t + \left(V^*_{t-}\beta^U_{t-} - R^*_{t-}\right) dM_t + d\widetilde{B}_t. \end{aligned} \tag{3.10}$$

It can now be shown, by Ito's formula, that $\overleftarrow{F}^*_t$ can be written in the form

$$\overleftarrow{F}^*_t = \int_{0-}^{t} \frac{U_t}{U_s}\left(d\left(C^*_s + \widetilde{B}_s\right) - \sigma^*_s \sigma^U_s ds\right), \tag{3.11}$$

where C^* is a process defined by

$$\begin{aligned} C^*_0 &= \widehat{B}_0 - V^*_0, \\ dC^*_t &= \alpha^*_t dt + \beta^*_{t-} dM_t + \sigma^*_t dW_t,\ t > 0, \end{aligned} \tag{3.12}$$

with

$$\begin{aligned} \alpha^*_t &= \left(\alpha^U_t - r^*_t\right) V^*_t + \left(\partial_s V^*_t\right)^T \sigma^S_t h^*_t + R^*_t diag\left(g^*_t\right) \mu_t, \\ \beta^*_t &= V^*_t \beta^U_t - R^*_t, \\ \sigma^*_t &= V^*_t \sigma^U_t - \left(\partial_s V^*_t\right)^T \sigma^S_t. \end{aligned}$$

The term $\sigma_t^* \sigma_t^U dt$ in (3.11) is a correction term stemming from the correlation between the increment of dU_t and the increment of dC_t^*.

Apart from the term $d\widetilde{B}_t$, (3.10) and (3.12) are written in semimartingale form under the measure P. Since (3.9) is a relation under the measure Q, we shall derive a corresponding semimartingale form under Q,

$$\begin{aligned} d\overleftarrow{F}_t^* &= \overleftarrow{F}_{t-}^* \left(r_t dt + \beta_{t-}^U dM_t^Q + \sigma_t^U dW_t^Q \right) \\ &\quad + \left(\left(r_t - r_t^* \right) V_t^* + \left(\partial_s V_t^* \right)^T \sigma_t \left(h_t^* - h_t \right) + R_t^* diag \left(g_t^* - g_t \right) \mu_t \right) dt \\ &\quad + \left(V_t^* \sigma_t^U - \left(\partial_s V_t^* \right)^T \sigma_t \right) dW_t^Q + \left(V_{t-}^* \beta_{t-}^U - R_{t-}^* \right) dM_t^Q + d\widetilde{B}_t, \end{aligned} \tag{3.13}$$

and

$$dC_t^* = \alpha_t^{*Q} dt + \beta_{t-}^* dM_t^Q + \sigma_t^* dW_t^Q,$$

with

$$\alpha_t^{*Q} = \left(r_t - r_t^* \right) V_t^* + \left(\partial_s V_t^* \right)^T \sigma_t \left(h_t^* - h_t \right) + R_t^* diag \left(g_t^* - g_t \right) \mu_t. \tag{3.14}$$

Allowing, for a moment, of diffusion payments, we note that the retrospective surplus can be considered as the retrospective reserve of an insurance contract with payments given by $C^* + \widetilde{B}$ minus the correction term $\int \sigma_s^* \sigma_s^U ds$. An appealing interpretation of this payment process is to consider the process C^* as the premium payments, in general positive or negative, and the process $\widetilde{B}$ as the benefit process, in general positive or negative. The correction term comes from the correlation between payments and investment gains. The payments of this contract start out with a lump sum payment at time 0 of $C_0^* + \widetilde{B}_0 = B_0 - V_0^*$ and develop according to dC^* and $d\widetilde{B}$, including a lump sum payment at time T of $\Delta \widetilde{B}_T$. The relation (3.9) is simply an equivalence relation for this contract.

3.4.3 The prospective surplus

We define the *prospective surplus* $\overrightarrow{F}_t^*$ corresponding to a technical basis (r^*, g^*, h^*) by

$$\begin{aligned} \overrightarrow{F}_t^* &= V_t - V_t^*, \ t \geq 0, \\ \overrightarrow{F}_{0-}^* &= 0, \end{aligned} \tag{3.15}$$

Using Ito's formula on $\overrightarrow{F}_t^*$, we get

$$\begin{aligned} d\overrightarrow{F}_t^* &= \left(\partial_t V_t + \partial_s V_t \alpha_t + \psi_t \right) dt + \left(V_{t-}^{\mathcal{J}} - V_{t-}^{1 \times J} \right) dN_t + \left(\partial_s V_t \right)^T \sigma_t dW_t \\ &\quad - \left(\partial_t V_t^* + \partial_s V_t^* \alpha_t + \psi_t^* \right) dt - \left(V_{t-}^{*\mathcal{J}} - V_{t-}^{*1 \times J} \right) dN_t - \left(\partial_s V_t^* \right)^T \sigma_t dW_t. \end{aligned}$$

Now, a subtraction of the real Thiele's differential equation, an addition of the technical Thiele's differential equation, and some rearrangements give the form

$$\begin{aligned} d\overrightarrow{F}^*_t &= r_t \overrightarrow{F}^*_t dt \\ &+ \left((r_t - r^*_t) V^*_t + \left((\partial_s V^*_t)^T \sigma_t h^*_t - (\partial_s V_t)^T \sigma_t h_t \right) \right) dt \\ &+ (R^*_t diag(g^*_t) - R_t diag(g_t)) \mu_t dt \\ &- (\partial_s V^*_t - \partial_s V_t)^T \sigma_t dW_t - \left(R^*_{t-} - R_{t-} \right) dM_t + d\widetilde{B}_t. \end{aligned} \tag{3.16}$$

Apart from the term $d\widetilde{B}_t$, (3.16) is written on semimartingale form under the measure P. To compare with the dynamics of the retrospective surplus, we shall also derive the semimartingale form under the measure Q,

$$\begin{aligned} d\overrightarrow{F}^*_t &= r_t \overrightarrow{F}^*_t dt \\ &+ \left((r_t - r^*_t) V^*_t + (\partial_s V^*_t)^T \sigma_t (h^*_t - h_t) + R^*_t diag(g^*_t - g_t) \mu_t \right) dt \\ &+ (\partial_s V_t - \partial_s V^*_t)^T \sigma_t dW^Q_t + \left(R_{t-} - R^*_{t-} \right) dM^Q_t + d\widetilde{B}_t. \end{aligned} \tag{3.17}$$

Since $F_t = \overleftarrow{F}^*_t - \overrightarrow{F}^*_t$, we know that $\frac{\overleftarrow{F}^*_t - \overrightarrow{F}^*_t}{Z^0_t}$ is an $\mathcal{F}^S_t$-martingale under the measure Q (see Chapter 2). Writing the retrospective surplus in the form (3.11), this can be used to derive an appealing representation of the prospective surplus. It follows that

$$\begin{aligned} \frac{\overleftarrow{F}^*_t - \overrightarrow{F}^*_t}{Z^0_t} &= E^Q\left(\left. \frac{\overleftarrow{F}^*_T}{Z^0_T} \right| \mathcal{F}^S_t \right) \\ &= E^Q\left(\left. \frac{1}{Z^0_T} \int_{0-}^t \frac{U_T}{U_s} \left(d\left(C^*_s + \widetilde{B}_s \right) - \sigma^*_s \sigma^U_s ds \right) \right| \mathcal{F}^S_t \right) \\ &\quad + E^Q\left(\left. \frac{1}{Z^0_T} \int_t^T \frac{U_T}{U_s} \left(d\left(C^*_s + \widetilde{B}_s \right) - \sigma^*_s \sigma^U_s ds \right) \right| \mathcal{F}^S_t \right) \\ &= E^Q\left(\left. \frac{U_T}{Z^0_T} \right| \mathcal{F}^S_t \right) \int_{0-}^t \frac{1}{U_s} \left(d\left(C^*_s + \widetilde{B}_s \right) - \sigma^*_s \sigma^U_s ds \right) \\ &\quad + \int_t^T E^Q\left(\left. \frac{1}{U_s} \left(d\left(C^*_s + \widetilde{B}_s \right) - \sigma^*_s \sigma^U_s ds \right) E^Q\left(\left. \frac{U_T}{Z^0_T} \right| \mathcal{F}^S_s \right) \right| \mathcal{F}^S_t \right) \\ &= \frac{U_t}{Z^0_t} \int_{0-}^t \frac{1}{U_s} \left(d\left(C^*_s + \widetilde{B}_s \right) - \sigma^*_s \sigma^U_s ds \right) \\ &\quad + \int_t^T E^Q\left(\left. \frac{1}{U_s} d\left(C^*_s + \widetilde{B}_s \right) \frac{U_s}{Z^0_s} \right| \mathcal{F}^S_t \right) \\ &= \frac{\overleftarrow{F}^*_t}{Z^0_t} + E^Q\left(\left. \int_t^T \frac{1}{Z^0_s} d\left(C^*_s + \widetilde{B}_s \right) \right| \mathcal{F}^S_t \right), \end{aligned} \tag{3.18}$$

and we get the following representation of the prospective surplus

$$\overrightarrow{F}^*_t = E^Q\left(\left. Z^0_t \int_t^T \frac{1}{Z^0_s} d\left(-C^*_s - \widetilde{B}_s \right) \right| \mathcal{F}^S_t \right). \tag{3.19}$$

In (3.18) the correction term has disappeared since we work with Ito-integrals which are based on forward increments.

This shows that the prospective surplus can be considered as the prospective reserve of an insurance contract with payments given by $C^*+\widetilde{B}$ (allowing for diffusion payments) and (3.15) is simply the prospective equivalence relation of this contract.

A deterministic differential equation for the state-wise prospective surplus can be obtained in the same way as for the prospective reserve in Chapter 2. The only extension is the allowance for diffusion payments. Upon introducing

$$\begin{aligned} \overrightarrow{F}_t^{*\mathcal{J}} &= V\left(t, S_t + \beta_t^{\cdot j}\right) - V^*\left(t, S_t + \beta_t^{\cdot j}\right), \\ \overrightarrow{R}_t^* &= \widetilde{b}_t^d + \overrightarrow{F}_t^{*\mathcal{J}} - \overrightarrow{F}_t^{*1\times J}, \\ \overrightarrow{\psi}_t^* &= \frac{1}{2} tr\left(\sigma_t^T \partial_{ss} \overrightarrow{F}_t^* \sigma_t\right), \end{aligned}$$

$\overrightarrow{F}_t^*$ solves the differential equation

$$\begin{aligned} \partial_t \overrightarrow{F}_t^* &= \alpha_t^{*Q} + \widetilde{b}_t^c + r_t \overrightarrow{F}_t^* - \left(\partial_s \overrightarrow{F}_t^*\right)^T (\alpha_t + \sigma_t h_t) \\ &\quad - \overrightarrow{R}_t^* diag\left(1^{J\times 1} + g_t\right)\mu_t - \overrightarrow{\psi}_t^*, \\ F_{T-} &= \Delta \widetilde{B}_T. \end{aligned} \tag{3.20}$$

As an alternative to the real Thiele's differential equation and the real equivalence relation, (3.20) may be used together with the equivalence relation (3.15) to determine an arbitrage free dividend plan.

3.4.4 Two important cases

We now specify the dynamics of the retrospective and the prospective surplus in the cases where the technical basis coincides with the first order basis and the real basis, respectively.

If the **technical basis equals the first order basis**, i.e. $(r^*, g^*, h^*) = (\widehat{r}, \widehat{g}, \widehat{h})$, we get from (3.13), (3.14), and (3.17),

$$\begin{aligned} d\overleftarrow{F}_t^* &= r_t \overleftarrow{F}_t^* dt + \left(\frac{dU_t}{U_{t-}} - r_t dt\right)\left(\overleftarrow{F}_{t-}^* + \widehat{V}_{t-}\right) \\ &\quad + \widehat{\alpha}_t^Q dt - \left(\partial_s V_t^*\right)^T \sigma_t dW_t^Q - R_{t-}^* dM_t^Q + d\widetilde{B}_t, \\ d\overrightarrow{F}_t^* &= r_t \overrightarrow{F}_t^* dt \\ &\quad + \widehat{\alpha}_t^Q dt + \left(\partial_s V_t - \partial_s \widehat{V}_t\right)^T \sigma_t dW_t^Q + \left(R_{t-} - \widehat{R}_{t-}\right) dM_t^Q + d\widetilde{B}_t, \\ \widehat{\alpha}_t^Q &= (r_t - \widehat{r}_t)\widehat{V}_t + \left(\partial_s \widehat{V}_t\right)^T \sigma_t \left(\widehat{h}_t - h_t\right) + \widehat{R}_t diag\left(\widehat{g}_t - g_t\right)\mu_t. \end{aligned} \tag{3.21}$$

We have that $C_0^* = \widehat{B}_0 - \widehat{V}_0$, i.e. $E\left[C_0^*\right] = 0$. Interpreting C^* as a premium process corresponding to the benefit process $\widetilde{B}$, this means that there is no systematic single premium paid at time 0. All systematic premiums of the insurance contract $C^* + \widetilde{B}$ fall due during $(0, T)$.

If the **technical basis equals the real basis**, i.e. $(r^*, g^*, h^*) = (r, g, h)$, we get from (3.13), (3.14), and (3.17),

$$\begin{aligned} d\overleftarrow{F}^*_t &= r_t \overleftarrow{F}^*_t dt + \left(\frac{dU_t}{U_{t-}} - r_t dt\right)\left(\overleftarrow{F}^*_{t-} + V^*_{t-}\right) \\ &\quad - (\partial_s V^*_t)^T \sigma_t dW^Q_t - R^*_{t-} dM^Q_t + d\widetilde{B}_t, \\ d\overrightarrow{F}^*_t &= r_t \overrightarrow{F}^*_t dt \\ &\quad + (\partial_s V_t - \partial_s V^*_t)^T \sigma_t dW^Q_t + \left(R_{t-} - R^*_{t-}\right) dM^Q_t + d\widetilde{B}_t. \end{aligned}$$

We have that $C^*_0 = \widehat{B}_0 - V^*_0$, i.e. $E[C^*_0] = E^Q\left(\int_{0-}^T \frac{1}{Z^0_t} d\widehat{B}_t\right)$. Interpreting C^* as a premium process corresponding to the benefit process $\widetilde{B}$, this means that there is a systematic single premium of $E^Q\left(\int_{0-}^T \frac{1}{Z^0_t} d\widehat{B}_t\right)$ at time 0. On the other hand, no systematic premiums fall due during $(0, T)$.

By subtraction of (3.16) from (3.10) and subtraction of (3.17) from (3.13), we get the P and Q dynamics of F, derived also in Chapter 2,

$$\begin{aligned} dF_t &= r_t F_t dt + \left(\frac{dU_t}{U_{t-}} - r_t dt\right)(F_{t-} + V_{t-}) \\ &\quad + (\partial_s V_t)^T \sigma_t h_t dt + R_t diag(g_t)\mu_t dt \\ &\quad - (\partial_s V_t)^T \sigma_t dW_t - R_{t-} dM_t \\ &= r_t F_t dt + \left(\sigma^U_t (F_t + V_t) - (\partial_s V_t)^T \sigma_t\right) dW^Q_t \\ &\quad + \left((F_{t-} + V_{t-})\beta^U_{t-} - R_{t-}\right) dM^Q_t. \end{aligned}$$

The dynamics of F is, of course, independent of the technical basis. However, it is interesting to see how the increments of F for a given technical basis split into increments of the retrospective and the negative prospective surplus (the first column of F increments is the sum of the second column of $\overleftarrow{F}^*$ increments and the third column of minus $\overrightarrow{F}^*$ increments):

$r_t F_t dt$	$r_t \overleftarrow{F}^*_t dt$	$-r_t \overrightarrow{F}^*_t dt$
$+\left(\frac{dU_t}{U_{t-}} - r_t dt\right)(F_{t-} + V_{t-})$	$+\left(\frac{dU_t}{U_{t-}} - r_t dt\right)\left(\overleftarrow{F}^*_{t-} + V^*_{t-}\right)$	$+0$
$-(\partial_s V_t)^T \sigma_t dW_t$	$-(\partial_s V^*_t)^T \sigma_t dW_t$	$-(\partial_s V_t - \partial_s V^*_t)^T \sigma_t dW_t$
$-R_{t-} dM_t$	$-R^*_{t-} dM_t$	$-\left(R_{t-} - R^*_{t-}\right) dM_t$
$+0$	$+\alpha^*_t + d\widetilde{B}_t$	$-\left(\alpha^*_t + d\widetilde{B}_t\right)$
dF_t	$d\overleftarrow{F}^*_t$	$d\overrightarrow{F}^*_t$

Note, in particular, from the second line that possible risky investments influence only the retrospective surplus and not the prospective surplus.

3.4.5 Main example continued

Let γ be the constant relative portfolio of risky assets corresponding to the investment strategy θ, i.e. $\theta^2_t S^2_t = \gamma U_t$ and $\theta^0_t S^0_t = (1 - \gamma) U_t$. Then the dynamics of U

becomes

$$dU_t = (r + \gamma(\overline{\alpha} - r)) U_t dt + \gamma\overline{\sigma} U_t dW_t.$$

In the rest of our example we consider the case where the first order basis is chosen as technical basis and, for notational convenience, we skip the topscript "$*$" on $\overleftarrow{F}_t^*$ and $\overrightarrow{F}_t^*$. We get from (3.21),

$$\begin{aligned} d\overleftarrow{F}_t &= r\overleftarrow{F}_t dt + (\gamma(\overline{\alpha} - r) dt + \gamma\overline{\sigma} dW)\left(\overleftarrow{F}_t + \widehat{V}_t\right) \\ &\quad + \widehat{\alpha}_t dt - \left(1 - \widehat{V}_{t-}\right) dM_t + d\widetilde{B}_t, \\ d\overrightarrow{F}_t &= r\overrightarrow{F}_t dt + \widehat{\alpha}_t dt + \left(\widehat{V}_{t-} - \widetilde{b}_{t-}^d - V_{t-}\right) dM_t + d\widetilde{B}_t, \\ \widehat{\alpha}_t &= (r - \widehat{r})\widehat{V}_t + \left(1 - \widehat{V}_t\right)(\widehat{\mu}_t - \mu_t). \end{aligned}$$

3.5 Dividends

3.5.1 The contribution plan and the second order basis

We start this section by considering a simple dividend plan, called the *contribution plan.* It plays an important role in the definition of a notion from participating life insurance, the *second order basis.* The contribution plan amounts to arranging $\widetilde{B}$ such that the discounted retrospective surplus $\frac{\overleftarrow{F}_t^*}{Z_t^0}$ becomes a zero mean Q-martingale, i.e.

$$\begin{aligned} \widetilde{B}_0 &= V_{0-}^*, \\ \widetilde{b}_t^c - \widetilde{b}_t^d diag\left(1^{J\times 1} + g_t\right)\mu_t &= -\alpha_t^{*Q}. \end{aligned} \tag{3.22}$$

Since all terms of α_t^{*Q} are functions of (t, S_t), the contribution plan is index-linked. Furthermore, since $E^Q\left(\frac{\overleftarrow{F}_T^*}{Z_T^0}\right) = 0$, the contribution plan is arbitrage free by construction. Note that also the discounted prospective surplus $\frac{\overrightarrow{F}_t^*}{Z_t^0}$ is a zero mean Q-martingale under the contribution plan.

The equation (3.22) is actually one equation with two free parameters to be chosen, $\widetilde{b}_t^c$ and $\widetilde{b}_t^d$, and we have a whole set of contribution plans. Usually, *the* contribution plan is considered as the special case where $\widetilde{b}_t^d = 0$, i.e.

$$\widetilde{b}_t^c = -\left((r_t - r_t^*) V_t^* + (\partial_s V_t^*)^T \sigma_t (h_t^* - h_t) + R_t^* diag(g_t^* - g_t)\mu_t\right). \tag{3.23}$$

If, for a given dividend plan $\widetilde{B}$ (not necessarily the contribution plan) there exists a vector $(\widetilde{r}, \widetilde{g}, \widetilde{h})$ such that (3.23) holds with (r, g, h) replaced by $(\widetilde{r}, \widetilde{g}, \widetilde{h})$ we call $(\widetilde{r}, \widetilde{g}, \widetilde{h})$ *the second order basis.* This means that the second order basis for a given dividend plan is the basis which, playing the role as real basis, turns the dividend plan into the contribution plan. It is seen that the basis (r, g, h) candidates to the second order basis only if the insurance contract actually follows the contribution plan. If the contract does not follow the contribution plan, i.e. if (3.23) is not fulfilled, then the real basis (r, g, h) does not candidate to the second order basis. The word candidate is appropriate here, since even if the insurance company actually follows the contribution plan, $\widetilde{b}_t$ can be obtained from (3.23) with other bases $(\widetilde{r}, \widetilde{g}, \widetilde{h})$ than (r, g, h) ((3.23) is in this connection one equation with three free parameters, $\widetilde{r}$, $\widetilde{g}$, and $\widetilde{h}$). This recognition of the second order basis as a decision variable has to our knowledge not, previously, been described in the literature, although it is well-known in practice. Sometimes the triplet of bases is completed by naming the real basis *the third order basis*.

In Section 3.3 it was realized that there will exist arbitrage free dividend plans to the policy holder's advantage if and only if

$$E^Q\left(\int_{0-}^{T}\frac{1}{Z_t^0}d\widehat{B}_t\right) \geq 0.$$

In the actuarial literature on participating life insurance one normally works with the much stricter requirement that, in particular, the contribution plan has to be arbitrage free and to the policy holder's advantage. This can be obtained if

$$\begin{aligned} -V_{0-}^* &\geq 0, \\ \alpha_t^{*Q} &\geq 0. \end{aligned}$$

One way of achieving $\alpha_t^{*Q} \geq 0$ is by having all the components of α_t^{*Q} greater or equal to zero, i.e.

$$\begin{aligned} (r_t - r_t^*)\, V_t^* &\geq 0, \\ R_t^* diag\,(g_t^* - g_t)\,\mu_t &\geq 0, \\ (\partial_s V_t^*)^T \sigma_t\,(h_t^* - h_t) &\geq 0. \end{aligned}$$

These are well-known relations (perhaps except for the last one which is a consequence of our diffusion type of indices), although they take different forms and are established differently in other expositions (e.g. Norberg [53] and Sverdrup [67]).

3.5.2 Surplus-linked insurance

The apparently appealing contribution plan has considerable drawbacks in our framework. Here we refer to the fact that, under the contribution plan, the policy holder does not participate in the game of investment. The contribution plan leads to the same dividends independently of the investment strategy underlying U, namely the dividends corresponding to pure investment in the risk-free asset. In some actuarial literature on surplus (e.g. Norberg [53] and Ramlau-Hansen [59]) only the possibility of investment in the risk-free asset has been taken into consideration. Consequently, the drawbacks of the contribution plan are not brought to the surface there. In practice the insurance companies hold large positions in risky assets, and an important issue of this chapter is to integrate this circumstance.

We propose another plan which we shall call the surplus-linked dividend plan. We add the retrospective surplus and the prospective surplus to the index and let the dividend plan be linked to this augmented index. It is clear from Section 3.4 that the augmented index is really an index since both the retrospective and the prospective surplus possess the dynamics of an index and the coefficients appearing in them are functions of (t, S_t).

We now proceed to specify the functional dependence of the dividends on the retrospective surplus and the prospective surplus. One can think of various constructions, but we shall go directly to a continuous affine form given by

$$\begin{aligned} \widetilde{B}_0 &= 0, \\ \widetilde{b}_t^d &= 0, \\ \widetilde{b}_t^c &= -\left(p_t + q_t \overleftarrow{F}_t^*\right) \text{ or } \widetilde{b}_t^c = -\left(p_t + q_t \overrightarrow{F}_t^*\right), \\ \Delta \widetilde{B}_T &= 0, \end{aligned} \tag{3.24}$$

where p_t and q_t are specified functions of (t, S_t). In the case of participating life insurance (3.24) should be modified such that $\widetilde{b}_t^c \leq 0$, e.g. by

$$\widetilde{b}_t^c = -\left(p_t + q_t \overleftarrow{F}_t^*\right)^+ \text{ or } \widetilde{b}_t^c = -\left(p_t + q_t \overrightarrow{F}_t^*\right)^+. \tag{3.25}$$

Here the option structure of products in participating life insurance can be recognized. We shall refer to the form in (3.24) as the pension funding form and the form in (3.25) as the participating life insurance form.

There is a variety of candidates for the functions p_t and q_t. A simple form would be to let p_t and q_t be deterministic functions. Other examples are $q_t = \frac{q_t'}{V_t^*}$ or $q_t = \frac{q_t'}{R_t^* \mu_t}$, q_t' being a deterministic function. Hereby we measure the surplus per technical reserve or per sum at risk. Such formulations could be motivated by solvency regulations.

As mentioned in Section 3.4, only the retrospective surplus (not the prospective surplus) is affected by risky investments. Thus, only dividends linked to the retrospective surplus amend the drawback of the contribution plan as mentioned above. Nevertheless, we propose also dividends linked to the prospective reserve for two reasons. Firstly, it makes the account more coherent. Secondly, it has some nice features from a mathematical point of view as we shall see below.

3.5.3 Main example continued

First we consider **dividends linked to the retrospective surplus**. Assume that the dividend plan specifies that $\left(p + q \overleftarrow{F}_t\right)^+$, for constants p and q, is paid out as long as the insured is alive, i.e.

$$\widetilde{b}_t = -\left(p + q \overleftarrow{F}_t\right)^+ 1_{\left(S_t^1 = 0\right)}. \tag{3.26}$$

The dynamics of the retrospective surplus becomes

$$\begin{aligned} d\overleftarrow{F}_t &= r\overleftarrow{F}_t dt + \left(\gamma\left(\overline{\alpha}-r\right)dt + \gamma\overline{\sigma}dW\right)\left(\overleftarrow{F}_t + \widehat{V}_t\right) + \widehat{\alpha}_t dt \\ &\quad -\left(1-\widehat{V}_{t-}\right)dM_t - \left(p + q\overleftarrow{F}_t\right)^+ 1_{\left(S_t^1=0\right)}dt. \end{aligned}$$

The deterministic differential equation and terminal condition for the reserve is given by ($V_t = V\left(t,0,f\right)$)

$$\begin{aligned} \partial_t V_t &= \pi - \left(p+qf\right)^+ + rV_t - \left(1-V_t\right)\mu_t \\ &\quad -d_f V_t\left(rf - \left(p+qf\right)^+ + \widehat{\alpha}_t + \left(1-\widehat{V}_t\right)\mu_t\right) \\ &\quad -\frac{1}{2}\gamma^2\overline{\sigma}^2\left(f+\widehat{V}_t\right)^2 d_{ff}V_t, \\ V_{T-} &= 1. \end{aligned} \tag{3.27}$$

In order to have arbitrage free dividends, (p,q) are to be determined subject to the equivalence relation

$$V_0 = 0.$$

We have no general analytical solution for the differential equation (3.27) and must resort to a numerical solution. However, an analytical solution can be obtained in the pension funding case where the topscript + in (3.26) is skipped. We get the differential equation ($V_t = V\left(t,0,f\right)$),

$$\begin{aligned} \partial_t V_t &= \pi - p - qf + rV_t - \left(1-V_t\right)\mu_t \\ &\quad -d_f V_t\left(rf - p - qf + \widehat{\alpha}_t + \left(1-\widehat{V}_t\right)\mu_t\right) \\ &\quad -\frac{1}{2}\gamma^2\overline{\sigma}^2\left(f+\widehat{V}_t\right)^2 d_{ff}V_t, \\ V_{T-} &= 1. \end{aligned}$$

We guess a solution in the form $V_t = \mathcal{X}_t + f\mathcal{Y}_t$, and find that the functions $\mathcal{X}$ and $\mathcal{Y}$ solve

$$\begin{aligned} \partial_t\mathcal{X}_t &= \pi - p + r\mathcal{X}_t - \mu_t\left(1-\mathcal{X}_t\right) + \left(p - \widehat{\alpha}_t - \left(1-\widehat{V}_t\right)\mu_t\right)\mathcal{Y}_t, \\ \mathcal{X}_{T-} &= 1, \\ \partial_t\mathcal{Y}_t &= -q + \left(\mu_t + q\right)\mathcal{Y}_t, \\ \mathcal{Y}_{T-} &= 0, \end{aligned}$$

leading to

$$\begin{aligned} \mathcal{Y}_t &= q\int_t^T e^{-\int_t^\tau(\mu_s+q)ds}d\tau, \\ \mathcal{X}_t &= \int_t^T e^{-\int_t^\tau(r+\mu_s)ds}H_\tau d\tau + e^{-\int_t^T(r+\mu_s)ds}, \end{aligned}$$

where

$$H_t = -\pi + p + \mu_t - \left(p - \widehat{\alpha}_t - \left(1-\widehat{V}_t\right)\mu_t\right)\mathcal{Y}_t.$$

Now we consider **dividends linked to the prospective surplus**. Assume that the dividend plan specifies that $\left(p+q\overrightarrow{F}_t\right)^+$, for constants p and q, is paid out as long as the insured is alive, i.e.

$$\widetilde{b}_t = -\left(p+q\overrightarrow{F}_t\right)^+ 1_{\left(S_t^1=0\right)}. \tag{3.28}$$

The dynamics of the prospective surplus becomes

$$d\overrightarrow{F}_t = r\overrightarrow{F}_t dt + \widehat{\alpha}_t dt + \left(\widehat{V}_{t-} - V_{t-}\right) dM_t - \left(p+q\overrightarrow{F}_t\right)^+ 1_{\left(S_t^1=0\right)} dt,$$

One tool is the deterministic differential equation and terminal condition for the reserve given by Theorem 1, $(V_t = V(t,0))$,

$$\begin{aligned} \partial_t V_t &= \pi - \left(p+q\left(V_t-\widehat{V}_t\right)\right)^+ + rV_t - (1-V_t)\mu_t, \\ V_{T-} &= 1. \end{aligned} \tag{3.29}$$

In order to have arbitrage free dividends, (p,q) are to be determined subject to the equivalence relation

$$V_0 = 0.$$

In the pension funding case without topscript + in (3.28) we get the differential equation,

$$\begin{aligned} \partial_t V_t &= \pi - p - q\left(V_t-\widehat{V}_t\right) + rV_t - (1-V_t)\mu_t, \\ V_{T-} &= 1. \end{aligned}$$

leading to the solution

$$V_t = \int_t^T e^{-\int_t^\tau (r+\mu_s-q)ds}\left(-\pi+p-q\widehat{V}_\tau+\mu_\tau\right) d\tau. \tag{3.30}$$

Another tool is the deterministic differential equation and terminal condition for the prospective surplus, (3.20), $\left(\overrightarrow{F}_t = \overrightarrow{F}(t,0)\right)$,

$$\begin{aligned} \partial_t \overrightarrow{F}_t &= r\overrightarrow{F}_t + \widehat{\alpha}_t - \left(p+q\overrightarrow{F}_t\right)^+ + \overrightarrow{F}_t\mu_t, \\ \overrightarrow{F}(T-,1) &= 0. \end{aligned} \tag{3.31}$$

In order to have arbitrage free dividends, (p,q) are to be determined subject to the equivalence relation

$$\overrightarrow{F}_0 = 0.$$

In the pension funding case without topscript + in (3.28) we get the differential equation,

$$\begin{aligned} \partial_t \overrightarrow{F}_t &= (r+\mu_t-q)\overrightarrow{F}_t + \widehat{\alpha}_t - p, \\ \overrightarrow{F}_{T-} &= 0. \end{aligned}$$

leading to

$$\overrightarrow{F}_t = \int_t^T e^{-\int_t^\tau (r+\mu_s - q)ds} \left(-\widehat{\alpha}_\tau + p\right) d\tau. \tag{3.32}$$

The two methods building on the deterministic differential equations for the reserve and for the prospective surplus, respectively, should lead to the same set of arbitrage free dividends. It is left to the reader to verify that for V and $\overrightarrow{F}$ given by (3.30) and (3.32) we get $V_0 = \overrightarrow{F}_0$ (use integration by parts and the first order Thiele's differential equation).

3.6 Bonus

3.6.1 Cash bonus versus additional insurance

Once payments of dividends have been determined, these are credited to the policy holder's account. This happens in basically two ways. Dividends are either paid out immediately or used as single premiums in insurance contracts and thus converted into a stream of future payments. If dividends are paid out immediately we speak of cash bonus and if dividends are used as single premiums we speak of additional insurance. Additional insurance can take various forms depending on the insurance contracts added during the term of the original contract. In Sections 3.6.2 and 3.6.3 we shall pay attention to two schemes of additional insurance called terminal bonus without guarantee and additional first order payments.

We endow all quantities related to the insurance contract initiated at time s with a topscript s, denoting e.g. by V_t^0 and V_t^s the reserves at time t for the insurance contracts initiated at time 0 and time s, respectively. A quantity without superscript covers a sum of the corresponding quantities related to insurance contracts initiated in the past, e.g.

$$V_t = \int_{0-}^t V_t^{ds}.$$

The notation is apt for working with general dividend plans not necessarily absolutely continuous. The circumstance that the dividend paid at time t is used as single premium of an insurance contract initiated at time t can be written

$$d\widetilde{B}_t = -\widehat{B}_t^{dt}, \tag{3.33}$$

from which we get

$$\begin{aligned}
dB_t &= d\widehat{B}_t + d\widetilde{B}_t \\
&= d\left(\int_{0-}^t \widehat{B}_t^{ds}\right) + d\widetilde{B}_t \\
&= \widehat{B}_t^{dt} + \int_{0-}^{t-} d\widehat{B}_t^{ds} + d\widetilde{B}_t \\
&= \int_{0-}^{t-} d\widehat{B}_t^{ds}.
\end{aligned}$$

This relation states that the payments at time t paid from the policy holder consists of all first order payments related to contracts initiated over $[0,t)$, including the original set of first order payments initiated at time 0.

3.6.2 Terminal bonus without guarantee

We consider a bonus plan which is usually not classified as additional insurance but rather as opposed to additional insurance. However, terminal bonus without guarantee can actually be considered as additional insurance with the benefits linked to the retrospective reserves. Payments linked to retrospective reserves can be considered as index-linked, since the retrospective reserve has index dynamics and all coefficients appearing in the retrospective reserve are dependent on (t, S_t). This makes it possible to formulate such pure saving contracts as special cases of insurance contracts.

For $s < T$ the dividends are used as single premiums on insurance contracts paying out the retrospective reserve at time T, i.e.

$$d\widehat{B}_t^{ds} = -L_{t-}^{ds} d1_{(t\geq T)}.$$

Then

$$\begin{aligned} dB_t &= \int_{0-}^{t-} d\widehat{B}_t^{ds} \\ &= d\widehat{B}_t^0 + \int_0^{t-} d\widehat{B}_t^{ds} \\ &= d\widehat{B}_t^0 - \int_0^{t-} L_{t-}^{ds} d1_{(t\geq T)} \\ &= d\widehat{B}_t^0 - d1_{(t\geq T)} \int_0^{t-} L_{t-}^{ds} \\ &= d\widehat{B}_t^0 - d1_{(t\geq T)} \int_0^{t-} \frac{U_{t-}}{U_s} \widehat{B}_s^{ds} \\ &= d\widehat{B}_t^0 + d1_{(t\geq T)} \int_0^{t-} \frac{U_{t-}}{U_s} d\widetilde{B}_s. \end{aligned}$$

Thus in addition to the first order payments, the policy holder simply receives at time T the value of past dividends including capital gains.

3.6.3 Additional first order payments

We consider a bonus plan increasing/decreasing the future payments proportionally to the first order payments contracted at time 0.

For $s < T$ the dividends are used as single premiums on insurance contracts paying out future payments of the original contract times a proportionality factor dK_s, i.e.

$$d\widehat{B}_t^{ds} = dK_s d\widehat{B}_t^0.$$

Variants of such a dividend plan are obtained by linking the payments of the additional insurance contracts to positive or negative payments (premiums or benefits) of the original contract only. The following results are easily carried over to such constructions. The distinction between additional premiums and additional benefits relates to the distinction between defined benefits or defined contributions, respectively, in pension funding. In the particular case of the contribution plan, the construction with additional benefits was studied in Norberg [53] and Ramlau-Hansen [59]. Then $\widetilde{B}$ and K are absolutely continuous. We work with general dividend plans and obtain thereby generalized results. Firstly,

$$\begin{aligned} dB_t &= \int_{0-}^{t-} d\widehat{B}_t^{ds} \\ &= d\widehat{B}_t^0 + \int_0^{t-} d\widehat{B}_t^{ds} \\ &= d\widehat{B}_t^0 + \int_0^{t-} dK_s d\widehat{B}_t^0 \\ &= (1 + K_{t-})\, d\widehat{B}_t^0. \end{aligned}$$

We conclude from this that payments are index-linked if K can be added to the index. We shall now derive the dynamics of K and see that in case of e.g. surplus-linked insurance this is in fact so.

We have

$$\begin{aligned} V_t^{*ds} &= E^{Q^*}\left(\int_t^T \frac{Z_t^*}{Z_u^*} d\left(-\widehat{B}_u^{ds}\right) \middle| \mathcal{F}_t\right) \\ &= E^{Q^*}\left(\int_t^T \frac{Z_t^*}{Z_u^*} d\left(-dK_s\widehat{B}_u^0\right) \middle| \mathcal{F}_t\right) \\ &= dK_s E^{Q^*}\left(\int_t^T \frac{Z_t^*}{Z_u^*} d\left(-\widehat{B}_u^0\right) \middle| \mathcal{F}_t\right) \\ &= dK_s V_t^{*0}, \end{aligned} \tag{3.34}$$

and in particular,

$$\widehat{V}_t^{dt} = dK_t \widehat{V}_t^0. \tag{3.35}$$

According to the first order equivalence relation exercised on the contract initiated at time t we have

$$\widehat{V}_t^{dt} = \widehat{B}_t^{dt}, \tag{3.36}$$

and from (3.35), (3.36) and (3.33) we find

$$dK_t = \frac{\widehat{V}_t^{dt}}{\widehat{V}_t^0} = \frac{-d\widetilde{B}_t}{\widehat{V}_t^0}. \tag{3.37}$$

From (3.34) we have that $V_t^* = (1 + K_t)\, V_t^{*0}$, and it can then be shown that the dynamics of the retrospective surplus are given by

$$\begin{aligned} d\overleftarrow{F}^*_t &= dC^*_t + d\widetilde{B}_t + \overleftarrow{F}^*_{t-}\frac{dU_t}{U_{t-}} \\ &= (1+K_{t-})\, dC^{*0}_t + d\widetilde{B}_t + \overleftarrow{F}^*_{t-}\frac{dU_t}{U_{t-}}. \end{aligned} \tag{3.38}$$

We know that all terms of dC^{*0}_t are functions of (t, S_t). It is now clear from (3.37) and (3.38) that we, in the case of retrospective surplus-linked dividends, can work with the augmented index $\left(S, K, \overleftarrow{F}^*\right)$. Correspondingly, if we consider prospective surplus-linked dividends, we can work with the augmented index (S, K).

3.6.4 Main example continued

First we consider **dividends linked to the retrospective surplus**. Assume that the dividend plan specifies that $\left(p + q\overleftarrow{F}_t\right)^+$, for constants p and q, is paid out as long as the insured is alive, i.e.

$$\widetilde{b}_t = -\left(p + q\overleftarrow{F}_t\right)^+ 1_{(S^1_t=1)}. \tag{3.39}$$

The dynamics of the retrospective surplus becomes

$$\begin{aligned} d\overleftarrow{F}_t &= r\overleftarrow{F}_t dt + (\gamma(\overline{\alpha} - r)\, dt + \gamma_t \overline{\sigma} dW)\left(\overleftarrow{F}_t + \widehat{V}_t\right) + (1+K_t)\,\widehat{\alpha}^0_t dt \\ &\quad - (1+K_{t-})\left(1 - \widehat{V}^0_{t-}\right) dM_t - \left(p + q\overleftarrow{F}_t\right)^+ 1_{(S^1_t=1)} dt. \end{aligned}$$

The deterministic differential equation and terminal condition for the reserve is then $(V_t = V(t, 0, f, k))$

$$\begin{aligned} \partial_t V_t &= \pi(1+k) + rV_t - (1+k-V_t)\,\mu_t \\ &\quad - d_f V_t\left(rf - (p+qf)^+ + (1+k)\left(\widehat{\alpha}^0_t + \left(1-\widehat{V}^0_t\right)\mu_t\right)\right) \\ &\quad - d_k V_t\left(\frac{(p+qf)^+}{\widehat{V}^0_t}\right) - \frac{1}{2}\gamma^2\overline{\sigma}^2\left(f + (1+k)\,\widehat{V}^0_t\right)^2 d_{ff} V_t, \\ V_{T-} &= 1+k. \end{aligned} \tag{3.40}$$

In order to have arbitrage free dividends, (p, q) are to be determined subject to the equivalence relation

$$V_0 = 0.$$

We have no general analytical solution for the differential equation (3.40) and must resort to a numerical solution. However, an analytical solution can be obtained in the pension funding case where the topscript + in (3.39) is skipped. We get the differential equation ($V_t = V(t,0,f,k)$),

$$\begin{aligned}
\partial_t V_t &= \pi(1+k) + rV_t - (1+k-V_t)\mu_t \\
&\quad -d_f V_t\left(rf - p - qf + (1+k)\left(\widehat{\alpha}_t^0 + \left(1-\widehat{V}_t^0\right)\mu_t\right)\right) \\
&\quad -d_k V_t\left(\frac{p+qf}{\widehat{V}_t^0}\right) - \frac{1}{2}\gamma^2\sigma^2\left(f+(1+k)\widehat{V}_t^0\right)^2 d_{ff}V_t, \\
V_{T-} &= 1+k.
\end{aligned}$$

We guess a solution in the form $V_t = \mathcal{X}_t + f\mathcal{Y}_t + k\mathcal{Z}_t$, and find that the functions $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ solve

$$\begin{aligned}
\partial_t \mathcal{X}_t &= \pi + r\mathcal{X}_t - \mu_t(1-\mathcal{X}_t) \\
&\quad + \left(p - \widehat{\alpha}_t^0 - \left(1-\widehat{V}_t^0\right)\mu_t\right)\mathcal{Y}_t - \frac{p}{\widehat{V}_t^0}\mathcal{Z}_t, \\
\mathcal{X}_{T-} &= 1, \\
\partial_t \mathcal{Y}_t &= (\mu_t + q)\mathcal{Y}_t - \frac{q}{\widehat{V}_t^0}\mathcal{Z}_t, \\
\mathcal{Y}_{T-} &= 0, \\
\partial_t \mathcal{Z}_t &= \pi + r\mathcal{Z}_t - \mu_t(1-\mathcal{Z}_t) - \left(\widehat{\alpha}_t^0 + \left(1-\widehat{V}_t^0\right)\mu_t\right)\mathcal{Y}_t, \\
\mathcal{Z}_{T-} &= 1,
\end{aligned}$$

leading to

$$\begin{aligned}
\mathcal{X}_t &= \int_t^T e^{-\int_t^\tau (r+\mu_s)ds} H_\tau d\tau + e^{-\int_t^T (r+\mu_s)ds}, \\
\begin{bmatrix} \mathcal{Y}_t \\ \mathcal{Z}_t \end{bmatrix} &= \int_t^T e^{-\int_t^\tau A_s ds}\begin{bmatrix} 0 \\ -\pi+\mu_\tau \end{bmatrix} d\tau + e^{-\int_t^T A_s ds}\begin{bmatrix} 0 \\ 1 \end{bmatrix},
\end{aligned}$$

where

$$\begin{aligned}
H_t &= -\left(p - a_t^0 - \left(1-\widehat{V}_t^0\right)\mu_t\right)\mathcal{Y}_t - \pi + \mu_t + \frac{p}{\widehat{V}_t^0}\mathcal{Z}_t, \\
A_t &= \begin{bmatrix} \mu_t + q & -\frac{q}{\widehat{V}_t^0} \\ -\left(\widehat{\alpha}_t^0 + \left(1-\widehat{V}_t^0\right)\mu_t\right) & r+\mu_t \end{bmatrix}.
\end{aligned}$$

Now we consider **dividends linked to the prospective reserve**. Assume that the dividend plan specifies that $\left(p+q\overrightarrow{F}_t\right)^+$, for constants p and q, is paid out as long as the insured is alive, i.e.

$$\widetilde{b}_t = -\left(p+q\overrightarrow{F}_t\right)^+ 1_{(S_t^1=1)}. \tag{3.41}$$

The dynamics of the retrospective surplus are given by

$$d\overrightarrow{F}_t = r\overrightarrow{F}_t dt + (1+K_t)\widehat{\alpha}_t^0 dt + \left(\widehat{V}_{t-} - V_{t-}\right)dM_t - \left(p+q\overrightarrow{F}_t\right)^+ 1_{(S_t^1=1)}dt,$$

The deterministic differential equation and terminal condition for the prospective surplus becomes $\left(\overrightarrow{F}_t = \overrightarrow{F}(t,0,k)\right)$

$$\begin{aligned}
\partial_t \overrightarrow{F}_t &= r\overrightarrow{F}_t + (1+k)\,\widehat{\alpha}_t^0 - \left(p + q\overrightarrow{F}_t\right)^+ \\
&\quad + \overrightarrow{F}_t \mu_t - d_k \overrightarrow{F}_t \left(\frac{\left(p + q\overrightarrow{F}_t\right)^+}{\widehat{V}_t^0} \right), \\
\overrightarrow{F}_{T-} &= 0.
\end{aligned}$$

In order to have arbitrage free dividends, (p,q) are to be determined subject to the equivalence relation

$$\overrightarrow{F}_0 = 0.$$

In the pension funding case without topscript $+$ in (3.41) we get the differential equation $\left(\overrightarrow{F}_t = \overrightarrow{F}(t,0,k)\right)$,

$$\begin{aligned}
\partial_t \overrightarrow{F}_t &= (r + \mu_t - q)\,\overrightarrow{F}_t + (1+k)\,\widehat{\alpha}_t^0 - p - d_k \overrightarrow{F}_t \left(\frac{\left(p + q\overrightarrow{F}_t\right)}{\widehat{V}_t^0} \right), \\
\overrightarrow{F}_{T-} &= 0.
\end{aligned}$$

We guess a solution in the form $\overrightarrow{F}_t = \mathcal{X}_t + k\mathcal{Z}_t$, and find that the functions $\mathcal{X}$ and $\mathcal{Z}$ solve

$$\begin{aligned}
\partial_t \mathcal{X}_t &= \widehat{\alpha}_t^0 - p\left(1 + \frac{\mathcal{Z}_t}{\widehat{V}_t^0}\right) + \left(r + \mu_t - q\left(1 + \frac{\mathcal{Z}_t}{\widehat{V}_t^0}\right)\right)\mathcal{X}_t, \\
\mathcal{X}_{T-} &= 0, \\
\partial_t \mathcal{Z}_t &= \widehat{\alpha}_t^0 + (r + \mu_t - q)\,\mathcal{Z}_t - \frac{q}{\widehat{V}_t^0}\mathcal{Z}_t^2, \qquad (3.42) \\
\mathcal{Z}_{T-} &= 0.
\end{aligned}$$

Now we can write the solution

$$\mathcal{X}_t = \int_t^T e^{-\int_t^\tau \left(r + \mu_s - q\left(1 + \frac{\mathcal{Z}_s}{\widehat{V}_s^0}\right)\right)ds} \left(p\left(1 + \frac{\mathcal{Z}_\tau}{\widehat{V}_\tau^0}\right) - \widehat{\alpha}_\tau^0\right) d\tau,$$

where $\mathcal{Z}_t$ is the solution to the ordinary nonlinear differential equation (3.42).

3.7 A comparison with related literature

3.7.1 The set-up of payments and the financial market

In the classical actuarial set-up of payments (see Hoem [33]), $S = (Z^0, X)$ with r being deterministic. In this set-up, the surplus was studied in Ramlau-Hansen [58].

Mathematically, it is only a small step to let r depend on X, still linking payments and intensities to X only (not Z^0). However, conceptually, it is a large step since it is actually the step from working with deterministic financial market modelling to stochastic financial market modelling. This was done in Norberg [51]. A stochastic rate of interest was introduced in the analysis of surplus in Norberg [53].

Generalizing the financial market and linking the payments to this financial market is a natural step. Conceptually, this is another large step because it opens for an explicit interaction between the insurance contract and the financial market introducing financial mathematics as an integrated part of life insurance mathematics. One way of approaching this interaction is, simply, to assume that the insurance contract is integrated completely in the financial market and consider it as a dynamically traded security itself. This was done in Chapter 2 and is the starting point for the reserves and surplus as introduced in this chapter.

3.7.2 Prospective versus retrospective

We have here introduced two notions of surplus, the retrospective surplus and the prospective surplus. This terminology is inherited from the theory of retrospective and prospective reserves as introduced in Norberg [50]. There the prospective reserve at time t is defined as the expected present value of benefits minus premiums, i.e. $-B$, over $(t, T]$ conditioned on information formalized by a sigma-algebra $\mathcal{G}_t \subseteq \mathcal{F}_t$, where $\mathcal{G}$ is not necessarily a filtration. The retrospective reserve is defined as the expected present value of premiums minus benefits, i.e. B, over $[0, t]$ conditioned on $\mathcal{G}_t$. Thus, the two reserves differ in the sign on B and by the time interval over which B is considered.

Taking $\mathcal{G}_t = \mathcal{F}_t^S$, we see that L_t is a retrospective reserve and V_t is a prospective reserve. Thus, the retrospective and the prospective surplus equal the retrospective and the prospective reserve, respectively, with subtraction of the technical prospective reserve. We can, in fact, represent the retrospective surplus and the prospective surplus as expected values themselves,

$$\overleftarrow{F}_t^* = E^Q \left(L_t - V_t^* | \, \mathcal{F}_t^S \right), \tag{3.43}$$

$$\overrightarrow{F}_t^* = E^Q \left(V_t - V_t^* | \, \mathcal{F}_t^S \right). \tag{3.44}$$

Furthermore, as was shown in (3.11) and (3.19) the retrospective surplus and the prospective surplus are, in fact, reserves themselves. The retrospective and prospective surplus are the retrospective and prospective reserves, respectively, corresponding to the payment process $C^* + \widetilde{B}$.

3.7.3 Surplus

In the traditional approach to emergence of surplus (see Sverdrup [67]), one defines the contribution to the surplus C as the payment stream in the form, $dC_t = c_t dt$, fulfilling

$$V_t' = V_t^* + E^Q \left(\int_t^T \frac{Z_t^0}{Z_u^0} d\left(-C_u\right) \middle| S_t \right). \tag{3.45}$$

From the corresponding differential equations one easily gets $c_t = \alpha_t^*$. Using that

$$V_t' = V_t - E^Q \left(\int_t^T \frac{Z_t^*}{Z_u^*} d\left(-\widetilde{B}_u\right) \middle| S_t \right),$$

(3.45) can be written as

$$V_t - V_t^* = E^Q \left(\int_t^T \frac{Z_t}{Z_u} d\left(-C_u - \widetilde{B}_u\right) \middle| S_t \right),$$

and Sverdrup [67] can be said to be based on prospective reasoning. Note that our prospective surplus does not show up here since we have enforced $dC_t = c_t dt$.

In Ramlau-Hansen [58] a realized profit, which corresponds to our retrospective surplus with $\widetilde{B} = 0$ and $(r^*, g^*, h^*) = \left(\widehat{r}, \widehat{g}, \widehat{h}\right)$, is introduced and studied intensively. Ramlau-Hansen [58] remarks that the realized profit only differs from the difference between the "second-order retrospective premium reserve" and the "first-order prospective reserve" by a martingale term. Since the second-order retrospective reserve in Ramlau-Hansen [58] is rather a retrospectively calculated prospective reserve $\widehat{V}_t$ (as pointed out in Norberg [50]) and the first-order prospective equals V^*, this difference corresponds to our prospective surplus, still with $\widetilde{B} = 0$. Thus, the remark by Ramlau-Hansen [58] can be seen as consequence of the fact that the retrospective surplus and the prospective surplus only differ by a martingale term. This follows immediately from the fact that $\frac{F}{Z^0}$ is a martingale. This is true even with $\widetilde{B} = 0$, although $\frac{F}{Z^0}$ is no zero-mean martingale then.

In Norberg [53] the surplus is defined as the realized profit in Ramlau-Hansen [58]. Norberg [53] sets out by deriving c_t in (3.45) in the traditional way and obtains that this equals the systematical contribution to the individual surplus. Again, this is just a consequence of the fact that F is a martingale and the retrospective and prospective surplus therefore only differ by a martingale term, even with $\widetilde{B} = 0$.

Finally, consider the special case of our surplus appearing from the set-up of payments and the financial market considered in Norberg [53]: Invest in Z^0 only, let the payments depend on X only and price, referring to diversifiability, risk to zero ($g = 0$). Let the technical basis coincide with the first order basis. Then (3.10) coincides with (3.13) and reduces to

$$d\overleftarrow{F}_t^* = r_t \overleftarrow{F}_t^* dt + \left((r_t - \widehat{r}_t) \widehat{V}_t + \widehat{R}_t \widehat{g}_t \mu_t \right) dt - \widehat{R}_t dM_t + d\widetilde{B}_t,$$

whereas α_t^* coincides with α_t^{*Q} and reduces to

$$\alpha_t^* = (r_t - \widehat{r}_t) \widehat{V}_t + \widehat{R}_t \widehat{g}_t \mu_t.$$

Letting $\widetilde{B} = 0$ and defining, appropriately, the first order intensity by $(1 + \widehat{g}) \mu$, these quantities equals the corresponding quantities in Norberg [53].

3.7.4 Information

In (3.43) and (3.44) we have represented the retrospective and the prospective surplus as expected values conditioned on $\mathcal{F}_t^S$. In Norberg [50] and Norberg [53], of such expected values conditioned on the full information are spoken of as individual quantities. It is one of the main ideas in Norberg [50] to relax this information and work with different versions of the reserve and the surplus corresponding to different sub-sigma-algebras $\mathcal{G}_t \subseteq \mathcal{F}_t^S$, not necessarily filtrations. This idea also plays a role in Norberg [53]. We shall not go into a study of the surplus under such relaxed information, but just remark that different sub-sigma-algebras $\mathcal{G}_t \subseteq \mathcal{F}_t^S$ replacing $\mathcal{F}_t^S$ in (3.43) and (3.44), correspondingly, define different versions of the retrospective and prospective surplus, respectively. In Norberg [53] the relaxation is restricted to concern information on diversifiable risk. We have no problem with relaxation of any kind of risk since we have, appropriately, adjusted for a possible non-zero price of risk by taking expectation under the measure Q.

3.7.5 The arbitrage condition

Last but not least, we compare the two apparently different requirements on the dividends showing up in Norberg [53] and in our framework. Our requirement is that dividends should be arbitrage free, leading to

$$E^Q\left(\int_{0-}^T \frac{1}{Z_t^0} dB_t\right) = 0, \tag{3.46}$$

whereas in Norberg [53] the requirement is

$$E\left(\int_{0-}^T \frac{1}{Z_t^0} dB_t \middle| \mathcal{H}_T' \vee \mathcal{G}_T\right) = 0, \tag{3.47}$$

where $\mathcal{H}_T'$ contains a part of the information over $(0, T]$ on diversifiable risk in the environment and $\mathcal{G}_T$ contains all information over $(0, T]$ on non-diversifiable risk in the environment. It can easily be shown that (3.46) follows from (3.47) if diversifiable risk is priced to zero,

$$\begin{aligned} E^Q\left(\int_{0-}^T \frac{1}{Z_t^0} dB_t\right) &= E^Q E^Q\left(\int_{0-}^T \frac{1}{Z_t^0} dB_t \middle| \mathcal{H}_T' \vee \mathcal{G}_T\right) \\ &= E^Q E\left(\int_{0-}^T \frac{1}{Z_t^0} dB_t \middle| \mathcal{H}_T' \vee \mathcal{G}_T\right) = 0. \end{aligned}$$

By (3.47), the company is allowed to carry a part of the diversifiable risk only. This may be an unfulfillable hard requirement. Taking instead expectation under the risk adjusted measure this requirement is not necessary. The company is then properly paid for the risk they are left with when all accounts have been settled at time T. However, in our set-up, the owners of the insurance company are, of course, welcome on the market if they want to get rid of some of this risk. This would be a matter of hedging, to some extent, the contingent claim L_T.

We hereby indicate that our set-up imposes a very important separation of the investment strategy underlying the investment of payments of the insurance contract, θ, which should be based on the objectives of the policy holder, and the investment strategy of the owners of the insurance company, which should be based on the objectives of these owners and could include an extent of hedging of the claim L_T.

3.8 Reserves, surplus, and accounting principles

In fact, many quantities treated in this chapter are reflected in the balance sheet of the insurance company. We shall in this section work with the following layout of a balance sheet,

Assets	Liabilities
Assets	*Market reserve* *+Hidden reserve* *Technical reserve* *Market dividend reserve* *−Hidden reserve* *Technical dividend reserve* *Equity*
Balance sum	*Balance sum*

Consider an insurance company with a value (i.e. an equity) of $\mathcal{E}_0$ at time 0 and at that time issuing (i.e. investing in) an insurance contract B. Apart from this insurance contract the insurance company invests in the market Z according to an investment strategy ξ. The balance sheet at time $0 \leq t \leq T$ becomes

Assets	Liabilities
$\mathcal{E}_0 + L_t + \int_0^t \xi_s dZ_s$	V_t' $+\left(V_t^* - V_t'\right)$ V_t^* $V_t - V_t'$ $-\left(V_t^* - V_t'\right)$ $V_t - V_t^*$ $\mathcal{E}_0 + L_t + \int_0^t \xi_s dZ_s - V_t$
$\mathcal{E}_0 + L_t + \int_0^t \xi_s dZ_s$	$\mathcal{E}_0 + L_t + \int_0^t \xi_s dZ_s$

where

$$V_t' = E^Q\left(\int_t^T \frac{Z_t}{Z_s} d\left(-\widehat{B}_s\right) \middle| S_t\right)$$

We have assumed throughout the chapter that $\mathcal{E}_0$ is sufficiently large to fulfill the inequality

$$\mathcal{E}_0 + L_t + \int_0^t \xi_s dZ_s - V_t \geq 0, \ t \geq 0.$$

If this inequality does not hold, the owners of the insurance company will, presumably, declare the company bankrupt and leave the remaining assets up to the policy holders to share. Thus, the inequality can be interpreted as some sort of non-ruin condition.

The balance sheet shows that a given technical basis results in a corresponding hidden reserve. The hidden reserve is a value transferred from the market dividend reserve to the market reserve and resulting in the technical dividend reserve and the technical reserve. Different technical bases can be interpreted as different accounting principles. Accountancy at the market principle amounts to evaluating future payments on the real basis. We see that the hidden reserve then becomes 0, and the balance sheet reflects a true picture of the company. The market reserve is a natural piece of information for the owners of the company, the policy holders and the supervisory authorities. However, as mentioned, other technical bases may be relevant for other operations.

Accountancy at the market principle is complicated by the fact that an appropriate real basis (r, g, h) may be difficult to identify. In the incomplete market the measure Q reflecting attitudes to risk can not be identified through the prices in the market. So we need another way to fix a measure Q. Should the real basis be determined by the market participants, separately, on the basis of possibly different criteria and methods? One objection to this is, that although every participant does its best to give a true picture, the direct comparability of balance sheets disappears. Should the market participants agree on a common real basis on which all participants should make up the accounts? Then the balance sheets would be directly comparable, but possibly some participants not agreeing on the real basis would not consider the result as a true picture. We shall not go further into this discussion here.

In this chapter we have focused on the decision variable $\widetilde{B}$ and the constraints on it. If the investment strategy θ underlying U is affecting $\widetilde{B}$ (e.g. through $\overleftarrow{F}$), another important decision variable in the design of the contract is of course θ. In fact $\left(\theta, \widetilde{B}\right)$ can be considered as a delicate constrained investment-consumption problem, where the technical basis decides when income is realized. In the third part of this thesis, we shall formulate this as an optimization problem in a diffusion framework and try to answer questions like: Can the solution sketched throughout the main example, i.e. constant relative investment and the forms of (3.24) and (3.25) in any sense be said to be optimal?

3.9 Numerical illustrations

In this section we want to illustrate results of the formula (3.27) with and without the topscript $+$, i.e. both for participating life and for pension funding. For a given choice of parameters in the first order basis and in the market, we fix $p = 0$ and want to solve for combinations of (q, γ) such that the side condition at time 0 is fulfilled.

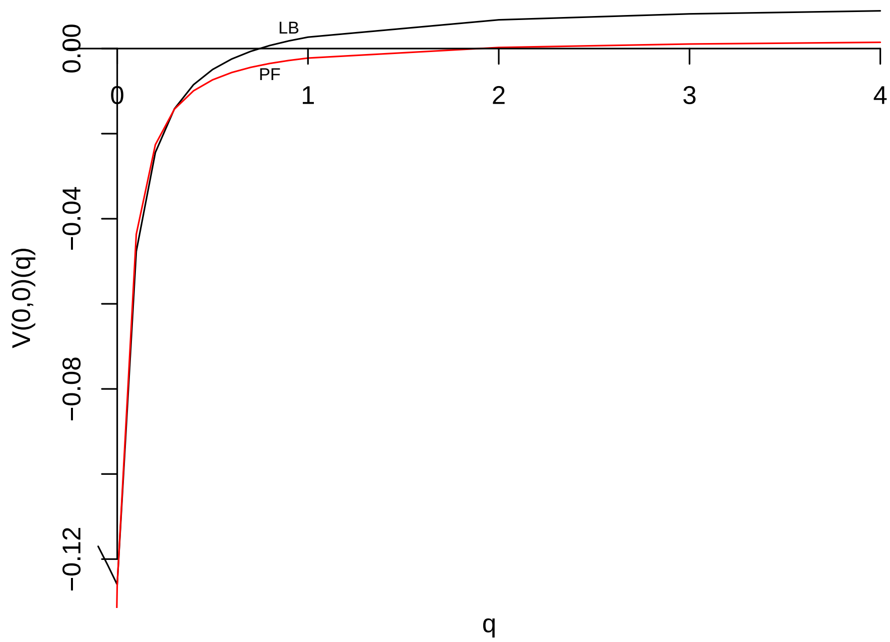

Figure 3.1: The reserve as a function of q, $V(q)$.

This will produce a fair strategy for investment and distribution of surplus.

We shall consider an insured at age 30 for a period of $T = 30$ years. As first order basis we use the Danish technical mortality intensity G82M for males and an (infinitesimal) interest rate 1.5%,

$$\begin{aligned} \widehat{\mu}_t &= 0.0005 + 10^{0.038(30+t)+5.88-10}, \\ \widehat{r} &= 0.015. \end{aligned}$$

As real basis we use

$$\begin{aligned} \mu_t &= (0.02 + 0.001t)\,\widehat{\mu}_t, \\ r &= 0.025, \\ \sigma &= 0.1. \end{aligned}$$

The form of the real mortality is inspired by a report from the Danish Society for Assessment of Personal Insurance Risk [4].

The reserve as a function of q is illustrated in Figure 3.1. Figure 3.2 is a zoom of Figure 3.1. We have fixed γ at 0.5, i.e. the insurance company places 50% of its assets in stocks and 50% in bonds. We can vary the parameter q in order to obtain arbitrage free contracts. There are two graphs, one for participating life (thick line called LB in the figure) and one for pension funding (thin line called PF in the figure).

For $q = 0$ the two reserves coincide in the point -0.125. This is the value of first order payments, because if $q = 0$ no dividends are paid out in any case.

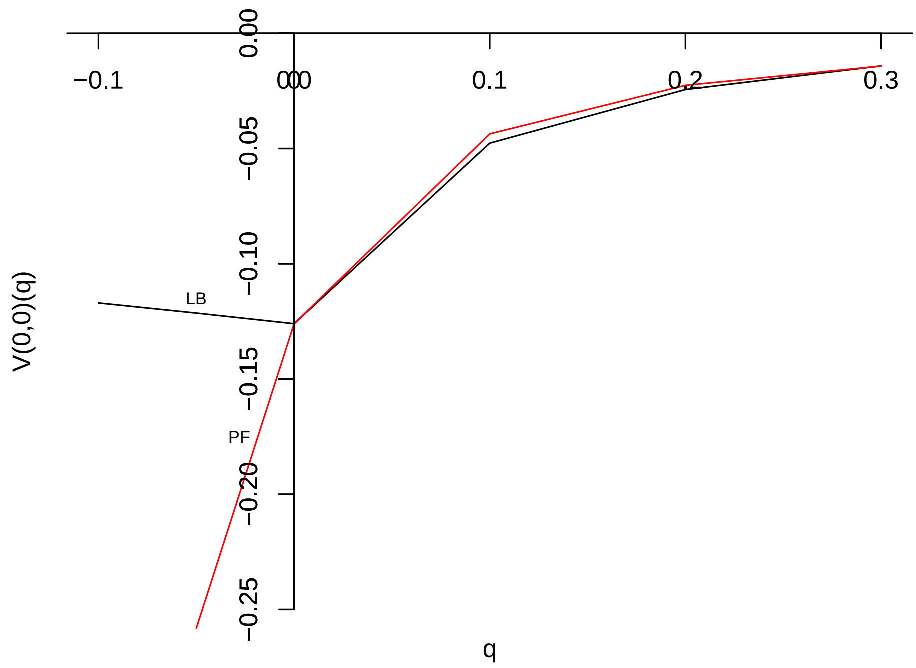

Figure 3.2: The reserve as a function of q, $V(q)$, zoom.

When q increases, the insurance company needs an additional liability for dividend payments. In the point where the reserves are zero the contracts are arbitrage free, i.e. for participating life the arbitrage free dividend parameter q is 0.74 whereas for pension funding the arbitrage free dividend parameter q is 1.88.

The reserves seem to have asymptotes for q tending to infinity. In case of participating life, this corresponds to a situation where positive surplus is paid out immediately after its emergence. The strategy will in the limit correspond to a dividend barrier below which the surplus is kept by singular dividend payments as long as the insured is alive. In case of pension funding, both positive and negative surplus is paid out immediately after its emergence. The strategy will in the limit give a surplus which is kept equal to zero by diffusion dividend payments as long as the insured is alive. The fact that the reserve for pension funding exceeds the reserve for participating life for $q \in (0, 0.3)$ is imputed to uncertainties in the numerical procedure.

For the sake of curiosity we have also calculated a reserve for a negative q. This corresponds to the rather hypothetical situation where the insurance company pays out positive dividends if the surplus is negative and pays out nothing and negative dividends for participating life and pension funding, respectively, if the surplus is positive. From a practical point of view this construction seems senseless because it undermines the idea of dividend payments from the insurance company point of view, namely to transfer (undiversifiable) risk to the policy holder. However, mathematically the construction makes sense and one could look for negative q for which the contract is arbitrage free. In pension funding positive q corresponds, in

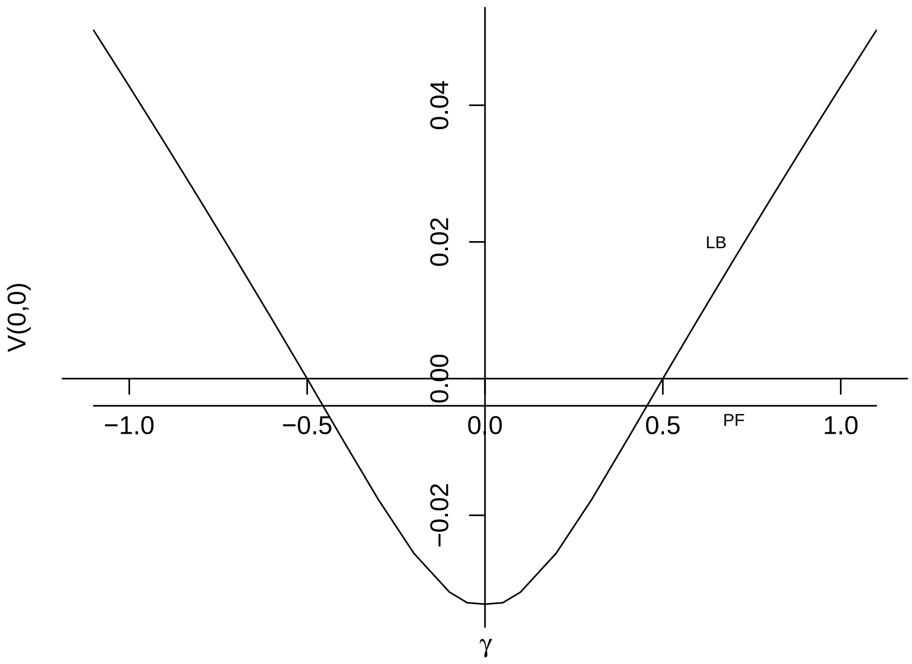

Figure 3.3: The reserve as a function of γ, $V(\gamma)$

some sense, to go long in the surplus whereas negative q corresponds to go short in the surplus. In participating life positive q corresponds, in the same sense, to buy call options on the surplus whereas negative q corresponds to buy put options on the surplus.

The reserve as a function of γ is illustrated in Figure 3.3. We have fixed q at 0.74 (where it produces arbitrage free dividends with $\gamma = 0.5$, see above) and can now vary γ in order to obtain arbitrage free contracts. There are two graphs, one for participating life (called LB in the figure) and one for pension funding (called PF in the figure).

One sees that the insurance contract is arbitrage free ($V(0,0) = 0$) if $\gamma = 0.5$. However, it is seen that another arbitrage free contract is obtained by letting $\gamma = -0.5$, i.e. going short in stocks. This is, however, an uninteresting observation from a practical point of view. In pension funding the reserve is independent of the proportional investments in stocks. The contract cannot be arbitrage free then, since q equals 0.74 and not 1.88 as it should be to obtain an arbitrage free pension funding contract (see above).

Not surprisingly, the shape of the participating life graph resembles the shape of a Black-Scholes option price as a function of the volatility of the underlying stock.

Arbitrage free points $(\mathbf{q}, \boldsymbol{\gamma})$ for participating life insurance is illustrated in Figure 3.4. We let the parameter q vary and find for each q the parameter γ which produces an arbitrage free dividend plan. In pension funding this makes no sense

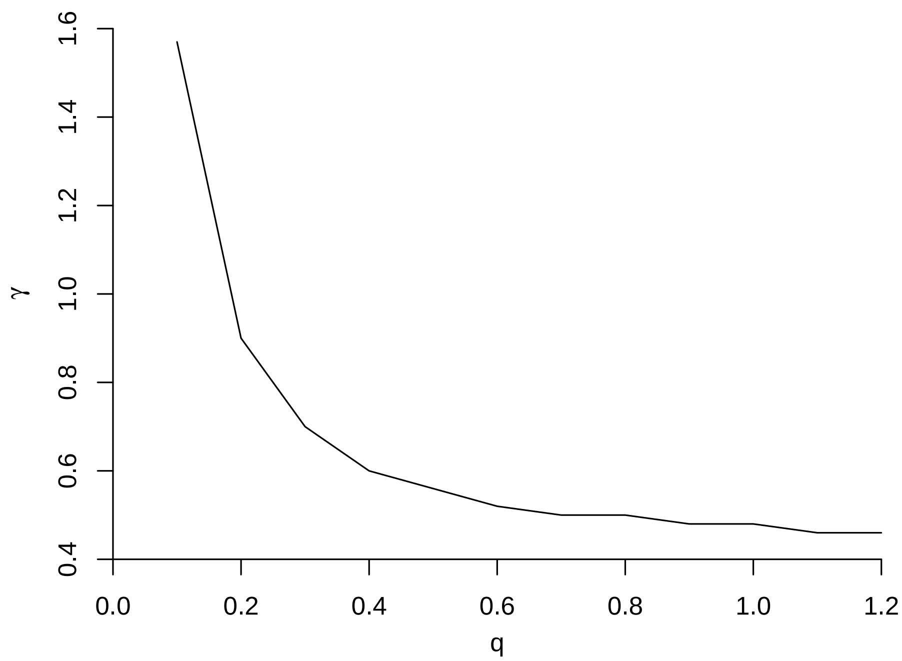

Figure 3.4: Arbitrage free points (q, γ)

since the reserve is independent of γ (see above). In participating life the graph forms a set of combinations of investment strategies and dividend strategies which produce arbitrage free contracts.

Figure 3.4 shows that there is a trade off between investments and dividends. It is not surprising that if one invests aggressively (high γ) then it should be combined with defensive dividend payments (low q) and the other way around. Given a set of realistic parameters, figures similar to Figure 3.4 can be a guideline both for the insurance company and for the supervisory authorities in their management and supervision, respectively.

Chapter 4

Control by intervention options

This chapter deals with the intervention options of the policy holder in life and pension insurance. To these options belong the surrender and the free policy (paid-up policy) options. Our approach is to let payments be driven by processes in which the policy holder is allowed to intervene. The main result is a quasi-variational inequality describing the market reserve on an insurance contract taking into account intervention options. The quasi-variational inequality generalizes Thiele's differential equation used for calculation of reserves on a policy without intervention options. It also generalizes the classical variational inequality used for calculation of the price of an American option.

4.1 Introduction

The market reserve on an insurance contract is in Chapter 2 defined as the market price of future contractual payments. Since this is the primary reserve in this chapter we will suppress the word market and simply speak of the reserve. It is a difficult task to describe in detail the payments stipulated in an insurance contract including the various options that may be held by both the policy holder, the insurance company, and the supervisory authorities. In several articles, published during the last decades, the authors bring some of these options to the surface and deal with their impact on the pricing and the reservation problems.

Starting with unit-linked life insurance, Brennan and Schwartz [10] recognized the option structure of a unit-linked life insurance contract with a guarantee. Brennan and Schwartz [10] integrate the mathematics of finance inevitably as a part of the mathematics of insurance. Going to participating life insurance, the application of mathematical finance has been long in coming probably due to the complex nature of these products. Briys and de Varenne [11] made the first attempt at dealing with the bonus option of the policy holder and the bankruptcy option of the (owners of the) insurance company in terms of contingent claims analysis. Since then, the idea has been developed in various respects. Miltersen and Persson [46] deal with the bonus option, whereas Grosen and Jørgensen [28] in addition take into consideration the surrender option. Grosen and Jørgensen [29] formalize the bankruptcy option of the insurance company and the intervention option of the supervising authorities. Important references for the point of view taken in this chapter are Grosen and Jørgensen [27] and [28], connecting early exercise with the surrender option in life insurance.

This chapter deals with the type of options that can be described by control by intervention. With control by intervention is meant that the controller is allowed to intervene in the evolution of some specified controlled process by introducing a jump at some controller-specified stopping time in the sense that he can actively move the process to some new point in the state space. The simplest form of intervention is optimal stopping, and as in Grosen and Jørgensen [27] we shall connect theory of optimal stopping with the surrender option in life insurance. The surrender option is the primary intervention option, often held by the policy holder, but there may be others. An example is the free policy option where the policy holder can, at any point in time, stop the payment of premiums but continue the contract with subscribed benefits. The free policy option cannot be described completely by means of theory of optimal stopping, at least if it is combined with the surrender option, and shows the need of introducing general theory of control by intervention.

The intervention options described above belong exclusively to the policy holder. However, also the insurance company may hold intervention options, and the bankruptcy option is the primary example. Allowing for intervention options of both the policy holder and the insurance company, we can regard the insurance contract as a so-called game option, see Kifer [39]. However, we shall not pursue this idea but leave it to the reader to generalize the results in this chapter to the situation with intervention options of the insurance company.

Due to the fact that most of the derivatives traded on the markets in practice are of American type, i.e. derivatives including an early exercise option, theory of optimal stopping plays an important role in derivative pricing. Many textbooks on mathematical finance contain an introduction, see e.g. Lamberton and Lapeyre [41].

A main result in optimal stopping is the expression of the optimal value function in terms of a solution to a so-called variational inequality. Description of the solution in the case of the American option and procedures for calculation of the solution were studied in e.g. Jaillet et al. [35]. The main result of the present chapter is a quasi-variational inequality, the solution of which expresses the reserve on an insurance contract with intervention options held by the policy holder. Since the American option is a special case of the setup of payments in this chapter, we choose to repeat, in a separate section, the well-known result for this case. On one hand, this will give the reader with a background in insurance but without particular knowledge to American option pricing an introduction to this well-studied object in the literature on mathematical finance. On the other hand, this will help the reader with a background in finance but without particular knowledge to insurance options to comprehend the setup introduced here.

In Section 4.2, we present the stochastic environment and introduce the payment process that makes up our insurance contract. In Section 4.3, the main results are presented in three theorems. In Sections 4.4, 4.5, and 4.6, we illustrate the main results in the cases of the American option in finance, the surrender option in life insurance, and the free policy option in life insurance, respectively.

4.2 The environment

In this section we recapitulate and extend the framework developed in Chapter 2. The special case of no intervention or no impulse control is exactly the framework of Chapter 2. Therefore, we ask the reader to confer Chapter 2 for motivation, details, and examples in this special framework, and we give here only interpretations and comments on the inclusion of intervention options.

We take as given a probability space $(\Omega, \mathcal{F}, \mathbf{F} = \{\mathcal{F}_t\}_{t\geq 0}, P)$. We let $(X_t)_{t\geq 0}$ be a cadlag (i.e. its sample paths are almost surely right continuous with left limits) jump process with finite state space $\mathcal{J} = (1, \ldots, J)$ defined on $(\Omega, \mathcal{F}, \mathbf{F} = \{\mathcal{F}_t\}_{t\geq 0}, P)$ and associate a marked point process (T_n, Φ_n), where T_n denotes the time of the nth jump of X_t, and Φ_n is the state entered at time T_n, i.e. $X_{T_n} = \Phi_n$. We introduce the counting processes

$$N_t^j = \sum_{n=1}^{\infty} 1_{(T_n \leq t, X_{T_n} = j)}, \quad j \in \mathcal{J},$$

and the J-dimensional vector

$$N_t = \begin{bmatrix} N_t^1 \\ \vdots \\ N_t^J \end{bmatrix}.$$

We let $(W_t)_{t\geq 0} = \left(W_t^1, \ldots W_t^K\right)_{t\geq 0}$ be a standard K-dimensional Brownian motion defined on $(\Omega, \mathcal{F}, \mathbf{F} = \{\mathcal{F}_t\}_{t\geq 0}, P)$.

For a matrix A we let A^T denote the transpose of A and let $A^{i\cdot}$ and $A^{\cdot i}$ denote the ith row and the ith column of A, respectively. For a vector a, we let $diag\,(a)$ denote the diagonal matrix with the components of a in the principal diagonal and 0 elsewhere. We shall write $\delta^{1\times J}$ and $\delta^{J\times 1}$ instead of $(\delta, \ldots, \delta)$ and $(\delta, \ldots, \delta)^T$, respectively. For derivatives we shall use the notation $\partial_x = \frac{\partial}{\partial x}$ and $\partial_{xy} = \frac{\partial^2}{\partial x \partial y}$. For a vector a we let $\int a$ and da mean componentwise integration and componentwise differentiation, respectively.

We introduce an *index* S, an $(I+1)$-dimensional vector of processes, the dynamics of which is given by

$$dS_t = \alpha_t dt + \beta_{t-} dN_t + \sigma_t dW_t, \;\; S_0 = s_0,$$

where $\alpha \in \mathbf{R}^{(I+1)}$, $\beta \in \mathbf{R}^{(I+1)\times J}$, and $\sigma \in \mathbf{R}^{(I+1)\times K}$ are functions of (t, S_t) and $s_0 \in \mathbf{R}^{I+1}$ is $\mathcal{F}_0$-measurable. We denote by S_t^i, α_t^i, β_t^{ij}, and σ_t^{ik} the ith entry of S_t, the ith entry of α_t, the (i,j)th entry of β_t, and (i,k)th entry of σ_t, respectively. The information generated by S is formalized by the filtration $\mathbf{F}^S = \{\mathcal{F}_t^S\}_{t\geq 0}$, where

$$\mathcal{F}_t^S = \sigma(S_s,\ 0 \leq s \leq t) \subseteq \mathcal{F}_t.$$

We assume that S is a Markov process and that there exist deterministic piecewise continuous functions $\mu^j(t,s)$, $j \in \mathcal{J}$, $s \in \mathbf{R}^{I+1}$ such that N_t^j admits the $\mathcal{F}_t^S$-intensity process $\mu_t^j = \mu^j(t,S_t)$, informally given by

$$\begin{aligned} \mu_t^j dt &= E\left(dN_t^j \middle| \mathcal{F}_{t-}^S\right) + o(dt) \\ &= E\left(dN_t^j \middle| S_{t-}\right) + o(dt), \end{aligned}$$

where $o(h)/h \to 0$ as $h \to 0$. We introduce the J-dimensional vectors containing the intensity processes and martingales associated with N,

$$\mu_t = \begin{bmatrix} \mu_t^1 \\ \vdots \\ \mu_t^J \end{bmatrix}, \quad M_t = \begin{bmatrix} M_t^1 \\ \vdots \\ M_t^J \end{bmatrix} = \begin{bmatrix} N_t^1 - \int_0^t \mu_s^1 ds \\ \vdots \\ N_t^J - \int_0^t \mu_s^J ds \end{bmatrix}.$$

To help the reader fix ideas, we explain briefly the roles of the introduced processes. Their roles will become more clear when we formalize the payment process below. The process N describes (at least) some specification of the life history of an insured. Whereas the process N will partly determine the points in time where payments fall due, the process S determines the amounts of these payments (and the intensities for the process N). In classical life insurance mathematics, payments are allowed to depend on the state of the policy, X. We can cover this situation by taking S^1 to be equal to X by the coefficients

$$\alpha_t^1 = 0,\ \beta_t^{1j} = j - S_t^1,\ \sigma_t^1 = 0,\ s_0^1 = X_0.$$

If e.g. X is included in the index S, $\mu(t,X_t)$ candidates to the intensity process corresponding to the classical situation, see e.g. Hoem [33]. However, in general, the intensity process μ may differ from the intensity process with respect to the natural filtration of N.

However, this classical contract can be extended in various directions. We can e.g. allow for payments (and intensities) to depend on the duration of the sojourn in the current state by letting S^2 be defined by

$$\alpha_t^2 = 1,\ \beta_t^2 = -S_t^2,\ \sigma_t^2 = 0,\ s_0^2 = 0,$$

and allow for payments (and intensities) to depend on the total number of jumps by letting S^3 be defined by

$$\alpha_t^3,\ \beta_t^{3j} = 1,\ \sigma_t^3 = 0,\ s_0^3 = 0.$$

In Møller [47] and Norberg [52] generalized versions of Thiele's differential equation have been studied where payments depend on the duration of the sojourn in the current state.

We introduce a *market* Z, an $(n+1)$-dimensional vector $(n \leq I)$ of price processes assumed to be positive, and denote by Z^i the ith entry of Z. The market Z consists of exactly those entries of S that are prices of traded assets. We assume that there exists a short rate of interest such that the market contains a price process Z^0 with the dynamics given by

$$dZ_t^0 = r_t Z_t^0 dt, Z_0^0 = 1.$$

This price process can be considered as the value process of a unit deposited on a bank account at time 0, and we shall call this entry for the risk-free asset even though r_t is allowed to depend on (t, S_t). Furthermore, we assume that the set of martingale measures, $\mathcal{Q}$, i.e. the set of probability measures Q equivalent to P such that $\frac{Z^i}{Z^0}$ is a Q-martingale for each i, is non-empty. From fundamental theory of asset pricing this assumption is known to be essentially equivalent to the assumption that no arbitrage possibilities exist on the market Z. The entries of an index S will also be called indices, and the indices appearing in Z will then be called marketed indices or assets. With this formulation the set of marketed indices is a subset of the set of indices and it contains at least one entry, namely Z^0. We let $\alpha^Z \in \mathbf{R}^{(n+1)}$, $\beta^Z \in \mathbf{R}^{(n+1)\times J}$, and $\sigma^Z \in \mathbf{R}^{(n+1)\times K}$ denote the coefficients of the asset prices Z.

So far we have followed the notation and the framework introduced in Chapter 2. We shall now go an important step further by allowing for intervention in the index S. Generally, for any point $(t,s) \in \mathbf{R}_+ \times \mathbf{R}^{(I+1)}$ there will be a set of points $K_t = K(t, S_t)$ to which S_t can be moved. The formal probabilistic apparatus necessary to describe this situation precisely is unfortunately rather cumbersome. One reason is that we need to work with an indexed set of filtrations $\left(\mathbf{F}^I\right)_{I\in\mathcal{I}}$ where $\mathbf{F}^I$ is the filtration belonging to the intervention strategy I (to be defined below). For the case where S is a piecewise deterministic process this apparatus is given in Davis [15]. Here, we skip the technical details and concentrate on the ideas and potentials of such a construction.

An *intervention strategy* is a marked point process $I = \left(T_n^I, \Phi_n^I\right)_{n=1,2,\ldots}$, where T_n^I denotes the time of the nth intervention, and Φ_n^I is the state entered at time T_n^I, i.e. $S_{T_n^I} = \Phi_n^I$. An intervention strategy is said to be *admissible* if for all ω and n that $T_n^I(\omega)$ is an $\mathcal{F}^{S^I}$-stopping time and $\Phi_n^I(\omega) \in K\left(T_n^I(\omega)-, S_{T_n^I(\omega)-}\right)$. This ensures that intervention at time t is based on only the information available at that point in time and takes place to a state that actually can be reached by intervention. We denote by $\mathcal{I}$ the set of admissible intervention strategies. Throughout the rest of the chapter, we will not need a notation for the dimension of the index, and the letter I will be used for specification of intervention strategies only. Introduce the counting processes

$$N_t^I = \sum_{n=1}^{\infty} 1_{(T_n^I \leq t)},$$

counting the number of interventions. Now, for all ω an intervention strategy $I \in \mathcal{I}$

will give rise to a realization of the index given by

$$dS_t^I = \alpha_t dt + \beta_{t-} dN_t + \sigma_t dW_t + \left(S_t^I - S_{t-}^I\right) dN_t^I,\ S_0^I = s_0^I.$$

Here, the term $\left(S_t^I - S_{t-}^I\right) dN_t^I$ will, if intervention happens at time t, move the index from S_{t-} to S_t.

Fixing some time horizon T, we now formally take an *insurance contract* to be a set of admissible intervention strategies $\mathcal{I}$ and an indexed *payment process* $\left(B^I\right)_{I\in\mathcal{I}}$, where B^I is an $\mathcal{F}_t^{S^I}$-adapted, cadlag process of finite variation with dynamics given by

$$dB_t^I = B_0 d1_{(t\geq 0)} + b_t^c dt - b_{t-}^d dN_t - b_{t-}^{S_t^I} dN_t^I - \Delta B_T d1_{(t\geq T)}, \tag{4.1}$$

where $B_0 \in \mathbf{R}$ is a function of S_0^I, $b^c \in \mathbf{R}$ and $b^d \in \mathbf{R}^J$ are functions of $\left(t, S_t^I\right)$, and $\Delta B_T \in \mathbf{R}$ is a function of S_T^I. $b_{t-}^{S_t^I}$ is the intervention payment due if intervention at time t moves S^I from S_{t-}^I to S_t. Thus, for all $k \in K\left(t, S_t^I\right)$, $b_t^k \in \mathbf{R}$ is a function of $\left(t, S_t^I\right)$. We denote by b^{dj} the jth entry of b^d. Note that the $\mathcal{F}_t^{S^I}$-adaptedness of B^I places demands on the connection between the coefficients of S^I and the coefficients of B^I.

B_t^I represents for a given intervention strategy $I \in \mathcal{I}$ the cumulative payments from the policy holder to the insurance company over $[0,t]$ following this strategy. Both continuous payments and lump sum payments are thus allowed to depend on the present state of the process $\left(t, S_t^I\right)$. The minus signs in front of b^d, b^k, and ΔB in dB_t conform to the typical situation where B_0 and b^c are premiums and b^d, $b^{S_t^I}$, and ΔB are benefits, all positive. To simplify notation, lump sum payments at deterministic times are restricted to time 0 and time T. Thus, an insurance contract is given by a set of functions $\left(B_0, b^c, b^d, \Delta B, b^k, k \in K\left(t,s\right)\right)$ such that a recording of S^I completely determines the payment stream.

By (4.1) the roles of the introduced processes become clear. It is the introduction of the intervention option that generalizes the payment process in Chapter 2. The reader could think of letting an entry of S indicate whether the insurance contract is surrendered or not. Then it is clear that this entry is moved by intervention of the policy holder. It is also clear that the payments depend on this entry, because if the policy is surrendered no future payments are to be paid. Finally, it is clear that we should open for a payment of the surrender value upon surrender (by the term $b_{t-}^{S_t^I} dN_t^I$). In Section 4.5, we illustrate the main results in the case of a surrender option in life insurance.

The insurance contract forms the basis for introduction of two price processes, F and V:

$$\begin{aligned} F_t &= \text{the price at time } t \text{ of the contractual payments to the} \\ &\quad\ \text{insurance company over } [0,T]\text{, i.e. premiums less benefits,} \\ V_t &= \text{the price at time } t \text{ of the contractual payments from the} \\ &\quad\ \text{insurance company over } (t,T]\text{, i.e. benefits less premiums.} \end{aligned}$$

Our approach to the price process F is the following: Assuming that the market Z is arbitrage free, we require that also the market (Z, F) be arbitrage free. We use the essential equivalence between arbitrage free markets and existence of a so-called martingale measure, i.e. a measure under which discounted asset prices are martingales. If the no arbitrage condition is fulfilled for (Z, F), we shall speak of $\left(B^I\right)_{I\in\mathcal{I}}$ as an arbitrage free insurance contract and about V as the corresponding arbitrage free reserve.

Since the market may be incomplete, there may be several martingale measures and, correspondingly, several arbitrage free reserves. Thus, when we talk of *the* arbitrage free reserve, we think of having fixed a martingale measure according to some criterion. Alternatively, one could imagine that there exists only one martingale measure reflecting the market participants' attitudes to risk although this measure, in the incomplete market, is not to be identified by looking at asset prices only. In this case *the* martingale measure could, appropriately, be fixed as the unique measure reflecting the attitudes to risk.

We restrict ourselves to prices allowing V_t to be written in the form $V\left(t, S_t^I\right)$. This restriction seems reasonable since S is Markov and since the payments by B and the intensities of N depend only on time and the current value of S^I, but it is actually a restrictive assumption on the formation of prices in the market. It corresponds to the restrictive structure of the measure transformation that we now enter by defining the likelihood process Λ^I by

$$\begin{aligned} d\Lambda_t^I &= \Lambda_{t-}^I \left(\sum_j g_{t-}^j dM_t^j + \sum_k h_t^k dW_t^k\right) \\ &= \Lambda_{t-}^I \left(g_{t-}^T dM_t + h_t^T dW_t\right), \\ \Lambda_0^I &= 1, \end{aligned}$$

where we have introduced

$$g_t^j = g^j\left(t, S_t^I\right),\ h_t^k = h^k\left(t, S_t^I\right),$$

and

$$g_t = \begin{bmatrix} g_t^1 \\ \vdots \\ g_t^J \end{bmatrix},\ h_t = \begin{bmatrix} h_t^1 \\ \vdots \\ h_t^K \end{bmatrix}.$$

With conditions on (g, h) (see Chapter 2) we can now change measure from P to Q^I on (Ω,F_T) by the definition,

$$\Lambda_T^I = \frac{dQ^I}{dP}. \tag{4.2}$$

In the case where the coefficients of Λ^I depend on the index S (no interventions), we simply denote the measure defined in (4.2) by Q.

Before starting to describe the price process V_t, we consider the market Z. Assuming that Z is arbitrage free, we can directly conclude that

$$\left(Z_t^I - Z_{t-}^I\right) dN_t^I = 0. \tag{4.3}$$

If we are allowed to trade in an asset and at the same time can move the asset price by intervention, it is easy to devise an arbitrage strategy. Given (4.3), we obtain absence of arbitrage by assuming existence of a martingale measure, i.e. existence of a solution (g, h) to

$$\alpha_t^{Z^I} + \sigma_t^{Z^I} h_t + \beta_t^{Z^I} diag\left(1^{J\times 1} + g_t\right)\mu_t - r_t Z_t^I = 0. \tag{4.4}$$

Note carefully that, although the asset prices are not allowed to be affected by intervention through the term $\left(Z_t^I - Z_{t-}^I\right) dN_t^I$, intervention may actually affect prices through the coefficients that are generally dependent on the index S^I. Also the coefficients of Λ are generally dependent on the index S^I. These circumstances imply that both each martingale measure and the set of possible martingale measures, in general, depend on the intervention strategy. Thus, in addition to the probabilistic admissibility condition on the intervention strategy we should add an absence of arbitrage condition stating that an intervention strategy is admissible if for all ω there exists a solution to (4.4). Then, a martingale measure exists for all admissible intervention strategies. We shall not discuss the interpretation of intervention in martingale measures but content ourselves with the fact that the situation where intervention does not affect the market at all, is, of course, just a special case.

4.3 The main results

We warm up with the situation where no intervention is possible. This is the special case treated in Chapter 2. We slightly reformulate the main result obtained there such that we have a version which is directly comparable with the result obtained in the situation with intervention. Note that both Theorem 3 and Theorem 4 presented in this section basically contain two results. The first is a financial result on the price process V. The second is a purely mathematical result which gives a tool for calculating the price. Theorem 5 gives conditions for simple optimal intervention strategies which are fulfilled in some simple cases.

Introduce the notation for a function $U(t, S_t)$,

$$\begin{aligned}
U_t^j &= U\left(t, S_t + \beta_t^{S\cdot j}\right), \\
U_t^{\mathcal{J}} &= \left[U_t^1, \ldots, U_t^J\right], \\
\psi_t &= \frac{1}{2} tr\left(\left(\sigma_t^S\right)^T \partial_{ss} U_t \sigma_t^S\right), \\
R_t &= b_t^d + U_t^{\mathcal{J}} - U_t^{1\times J}, \\
A_t U_t &= \left(\partial_s U_t\right)^T \left(\alpha_t^S + \sigma_t^S h_t\right) + R_t diag\left(1^{J\times 1} + g_t\right)\mu_t + \psi_t.
\end{aligned}$$

Theorem 3 *1. Assume that there are no intervention options. Assume existence of a solution to (4.4) and that the arbitrage free reserve V_t can be written in the form $V(t, S_t)$. Then*

$$V(t, S_t) = E^Q\left[\int_t^T \frac{Z_t^0}{Z_s^0} d\left(-B_s\right) \middle| S_t\right] \tag{4.5}$$

for some martingale measure $Q \in \mathcal{Q}$, and

$$V(0-) = E^{Q}\left[\int_{0-}^{T} \frac{1}{Z_s^0} d(-B_s)\right] = 0.$$

2. Assume that $U(t,s)$ is a (sufficiently) regular solution to the following differential equation

$$\begin{aligned} \partial_t U_t &= b_t^c - b_t^d diag\left(1^{J\times 1} + g_t\right)\mu_t + r_t U_t - A_t U_t, && (4.6)\\ U_{T-} &= \Delta B_T. && (4.7) \end{aligned}$$

Then

$$U(t,S_t) = V(t,S_t)$$

Sketch of proof. 1. This follows if $\left(\frac{Z}{Z^0},\frac{F}{Z^0}\right)$ is a Q-martingale (see Chapter 2).
2. If U is sufficiently regular, an Ito calculation shows that for $U_t = U(t,S_t)$,

$$m_t = \frac{U_t}{Z_t^0} - \int_0^t \frac{1}{Z_s^0}\left(\partial_s U_s - r_s U_s + A_s U_s\right) ds$$

is a Q-martingale. Then, the optional sampling theorem gives

$$E^{Q}\left[m_T \middle| \mathcal{F}_t\right] = m_t. \qquad (4.8)$$

By applying (4.8), (4.6), and (4.7), we have that

$$\begin{aligned} U_t &= E^{Q}\left[-\int_t^T \frac{Z_t^0}{Z_s^0}\left(\partial_s U_s - r_s U_s + A_s U_s\right) ds + \frac{Z_t^0}{Z_T^0} U_{T-}\middle| \mathcal{F}_t\right]\\ &= E^{Q}\left[\int_t^T \frac{Z_t^0}{Z_s^0}\left(-b_s^c + b_s^d diag\left(1^{J\times 1} + g_s\right)\mu_s\right) ds + \frac{Z_t^0}{Z_T^0}\Delta B_T\middle| \mathcal{F}_t\right]\\ &= E^{Q}\left[\int_t^T \frac{Z_t^0}{Z_s^0} d(-B_s)\middle| \mathcal{F}_t\right]\\ &= V(t,S_t).\square \end{aligned}$$

Theorem 4 *1. Assume existence of a solution to (4.4) and that the arbitrage free reserve V_t can be written in the form $V\left(t,S_t^I\right)$. Then*

$$V\left(t,S_t^I\right) = \sup_{I\in\mathcal{I}} E^{Q^I}\left[\int_t^T \frac{Z_t^{0I}}{Z_s^{0I}} d\left(-B_s^I\right)\middle| S_t^I\right] \qquad (4.9)$$

for some martingale measure $Q^I \in \mathcal{Q}(I)$, and

$$V(0-) = \sup_{I\in\mathcal{I}} E^{Q^I}\left[\int_{0-}^{T} \frac{1}{Z_s^{0I}} d\left(-B_s^I\right)\right] = 0.$$

2. Assume that $U(t,s)$ is a (sufficiently) regular solution to the following system of partial differential inequalities (quasi-variational inequality)

$$\begin{aligned} \partial_t U_t &\leq b_t^c - b_t^d diag\left(1^{J\times 1} + g_t\right)\mu_t + r_t U_t - A_t U_t, & (4.10)\\ U(t,s) &\geq \max_{k\in K(t,s)}\left(U(t,k) + b_t^k\right), & (4.11)\\ 0 &= \left(-\partial_t U_t + b_t^c - b_t^d diag\left(1^{J\times 1} + g_t\right)\mu_t + r_t U_t - A_t U_t\right)\\ &\quad \times\left(\max_{k\in K(t,s)}\left(U(t,k) + b_t^k\right) - U(t,s)\right), & (4.12)\\ U_{T-} &= \Delta B_T. & (4.13) \end{aligned}$$

Then

$$U\left(t, S_t^I\right) = V\left(t, S_t^I\right).$$

and the optimal intervention I is given by $\left(T_n^I, \Phi_n^I\right)_{n=1,2,\ldots}$, where $\left(T_n^I, \Phi_n^I\right)_{n=1,2,\ldots}$ are found successively by $\left(T_0^I = 0\right)$

$$\begin{aligned} T_n^I &= \inf_{T_{n-1}^I < \tau \leq T}\left\{U\left(\tau, S_\tau^I\right) = \max_{k\in K\left(\tau-, S_{\tau-}^I\right)}\left(U(\tau,k) + b_{\tau-}^k\right)\right\},\\ \Phi_n^I &= \arg\max_{k\in K\left(T_n^I-, S_{T_n^I-}^I\right)}\left(U\left(T_n^I, k\right) + b_{T_n^I-}^k\right). \end{aligned}$$

Sketch of proof. *1.* An arbitrage argument, left to the reader, shows that no one is willing to buy for more and no one is willing to sell for less than V given by (4.9).

2. The first part of the proof is a classical optimal stopping argument. An Ito calculation shows that for $U_t = U(t, S_t)$,

$$m_t = \frac{U_t}{Z_t^0} - \int_0^t \frac{1}{Z_s^0}\left(\partial_s U_s - r_s U_s + A_s U_s\right) ds$$

is a Q-martingale. Then, for a stopping time τ, $t \leq \tau \leq T$, the optional sampling theorem gives

$$E^Q\left[m_\tau \middle| \mathcal{F}_t\right] = m_t. \tag{4.14}$$

On one hand, by applying (4.14), (4.10), (4.11), and (4.13) we have that

$$\begin{aligned} U_t &= E^Q\left[-\int_t^\tau \frac{Z_t^0}{Z_s^0}\left(\partial_s U_s - r_s U_s + A_s U_s\right) ds + \frac{Z_t^0}{Z_\tau^0} U_{\tau-}\middle| \mathcal{F}_t\right]\\ &\geq E^Q\left[\int_t^\tau \frac{Z_t^0}{Z_s^0}\left(-b_s^c + b_s^d diag\left(1^{J\times 1} + g_s\right)\mu_s\right) ds\middle| \mathcal{F}_t\right]\\ &\quad + E^Q\left[\frac{Z_t^0}{Z_\tau^0}\left(\max_{k\in K(\tau-, S_{\tau-})}\left(U(\tau,k) + b_{\tau-}^k\right) 1_{(\tau<T)} + \Delta B_T 1_{(\tau=T)}\right)\middle| \mathcal{F}_t\right]\\ &= E^Q\left[\int_t^\tau \frac{Z_t^0}{Z_s^0} d\left(-B_s\right) + \frac{Z_t^0}{Z_\tau^0}\max_{k\in K(\tau-, S_{\tau-})}\left(U(\tau,k) + b_{\tau-}^k\right) 1_{(\tau<T)}\middle| \mathcal{F}_t\right]. \end{aligned}$$

Taking supremum over all stopping times $t \leq \tau \leq T$, this proves

$$U_t \geq \sup_{t\leq\tau\leq T} E^Q\left[\int_t^\tau \frac{Z_t^0}{Z_s^0} d\left(-B_s\right) + \frac{Z_t^0}{Z_\tau^0} \max_{k\in K(\tau-,S_{\tau-})} \left(U\left(\tau,k\right) + b_{\tau-}^k\right) 1_{(\tau<T)} \middle| \mathcal{F}_t\right]. \tag{4.15}$$

On the other hand, by defining

$$\begin{aligned}\tau_t &= \min\left(\inf_{t\leq\tau<T}\left\{U\left(\tau,S_\tau\right) = \max_{k\in K(\tau-,S_{\tau-})}\left(U\left(\tau,k\right)+b_{\tau-}^k\right)\right\},T\right),\\ k_t &= \arg\max_{k\in K\left(\tau_t-,S_{\tau_t-}\right)}\left(U\left(\tau_t,k\right)+b_{\tau_t-}^k\right),\end{aligned}$$

we get by (4.12) and (4.13) for $t \leq s < \tau_t \leq T$,

$$\partial_s U_s = b_s^c - b_s^d diag\left(1^{J\times 1} + g_s\right)\mu_s + r_s U_s - A_s U_s$$

and

$$U\left(\tau_t, S_{\tau_t}\right) = \max_{k\in K\left(\tau_t-,S_{\tau_t-}\right)}\left(U\left(\tau_t,k\right)+b_{\tau_t-}^k\right)1_{(\tau_t<T)} + \Delta B_T 1_{(\tau_t=T)},$$

such that the optional sampling theorem yields

$$\begin{aligned}U_t &= E^Q\left[-\int_t^{\tau_t}\frac{Z_t^0}{Z_s^0}\left(\partial_s U_s - r_s U_s + A_s U_s\right)ds + \frac{Z_t^0}{Z_{\tau_t}^0}U_{\tau_t}\middle|\mathcal{F}_t\right]\\ &= E^Q\left[\int_t^{\tau_t}\frac{Z_t^0}{Z_s^0}\left(-b_s^c + b_s^d diag\left(1^{J\times1}+g_s\right)\mu_s\right)ds\middle|\mathcal{F}_t\right]\\ &\quad + E^Q\left[\frac{Z_t^0}{Z_{\tau_t}^0}\left(\max_{k\in K\left(\tau_t-,S_{\tau_t-}\right)}\left(U\left(\tau_t,k\right)+b_{\tau_t-}^k\right)1_{(\tau_t<T)}+\Delta B_T 1_{(\tau_t=T)}\right)\middle|\mathcal{F}_t\right]\\ &= E^Q\left[\int_t^{\tau_t}\frac{Z_t^0}{Z_s^0}d\left(-B_s\right)+\frac{Z_t^0}{Z_{\tau_t}^0}\max_{k\in K\left(\tau_t-,S_{\tau_t-}\right)}\left(U\left(\tau_t,k\right)+b_{\tau_t-}^k\right)1_{(\tau_t<T)}\middle|\mathcal{F}_t\right].\end{aligned}$$

This proves that

$$U_t \leq \sup_{t\leq\tau\leq T} E^Q\left[\int_t^\tau \frac{Z_t^0}{Z_s^0} d\left(-B_s\right) + \frac{Z_t^0}{Z_\tau^0} \max_{k\in K(\tau-,S_{\tau-})} \left(U\left(\tau,k\right) + b_{\tau-}^k\right) 1_{(\tau<T)} \middle| \mathcal{F}_t\right]. \tag{4.16}$$

By (4.15) and (4.16) we have

$$U_t = \sup_{t\leq\tau\leq T} E^Q\left[\int_t^\tau \frac{Z_t^0}{Z_s^0} d\left(-B_s\right) + \frac{Z_t^0}{Z_\tau^0} \max_{k\in K(\tau-,S_{\tau-})} \left(U\left(\tau,k\right) + b_{\tau-}^k\right) 1_{(\tau<T)} \middle| \mathcal{F}_t\right].$$

We now have to show that the quantity characterized by this relation actually is

$$V_t = \sup_{I\in\mathcal{I}} E^{Q^I}\left(\int_t^T \frac{Z_t^{0I}}{Z_s^{0I}} d\left(-B_s^I\right)\middle| S_t^I\right).$$

Define

$$V_t^n = \sup_{I\in\mathcal{I}^n} E^{Q^I}\left(\int_t^T \frac{Z_t^{0I}}{Z_s^{0I}} d\left(-B_s^I\right)\middle| S_t^I\right),$$

where $\mathcal{I}^n$ is the set of admissible intervention strategies involving at most n interventions. Then

$$V_t^{n+1} = \sup_{t \leq \tau \leq T} E^Q \left[\int_t^\tau \frac{Z_t^0}{Z_s^0} d(-B_s) + \frac{Z_t^0}{Z_\tau^0} \max_{k \in K(\tau, S_\tau)} \left(V^n(\tau, k) + b_{\tau-}^k \right) 1_{(\tau<T)} \middle| \mathcal{F}_t \right]. \tag{4.17}$$

With appropriate regularity conditions on $\mathcal{I}^n$ and $\mathcal{I}$, $\lim_{n\to\infty} V_t^n$ will exist and $\lim_{n\to\infty} V_t^n = V_t$. Taking the limit on both sides of (4.17), bounded convergence gives that

$$V_t = \sup_{t \leq \tau \leq T} E^Q \left[\int_t^\tau \frac{Z_t^0}{Z_s^0} d(-B_s) + \frac{Z_t^0}{Z_\tau^0} \max_{k \in K(\tau-, S_{\tau-})} \left(V(\tau, k) + b_{\tau-}^k \right) 1_{(\tau<T)} \middle| \mathcal{F}_t \right]. \square$$

Theorem 5 *Fix a time ρ. Let Y_t, $\rho \leq t \leq T$, be given by*

$$Y_t = \int_\rho^t \frac{Z_\rho^0}{Z_s^0} d(-B_s) + \frac{Z_\rho^0}{Z_t^0} \max_{k \in K(t-, S_{t-})} \left(V(t, k) + b_{t-}^k \right) 1_{(t<T)}. \tag{4.18}$$

If Y is a Q-submartingale, then the optimal intervention is not to intervene. If Y is a Q-supermartingale, then the optimal intervention is immediately to transmit S to

$$k = \arg \max_{k \in K(\rho-, S_{\rho-})} \left(V(\rho, k) + b_{\rho-}^k \right).$$

Sketch of proof. If Y_t is a Q-submartingale, optional sampling theorem gives for any stopping time $\rho \leq \tau \leq T$,

$$Y_\tau \leq E^Q \left[Y_T \middle| \mathcal{F}_\tau \right].$$

Taking conditional expected value yields

$$\begin{aligned} & E^Q \left[\int_\rho^\tau \frac{Z_\rho^0}{Z_s^0} d(-B_s) + \frac{Z_\rho^0}{Z_\tau^0} \max_{k \in K(\tau-, S_{\tau-})} \left(V(\tau, k) + b_{\tau-}^k \right) 1_{(\tau<T)} \right] \\ \leq \;& E^Q \left[\int_\rho^T \frac{Z_\rho^0}{Z_s^0} d(-B_s) \right]. \end{aligned}$$

Supremum of the left hand side over all stopping times $\rho \leq \tau \leq T$ is obtained for $\tau = T$, since then an equality holds.

If Y_t is a Q-supermartingale, optional sampling theorem gives for any stopping time $\rho \leq \tau \leq T$,

$$Y_\rho \geq E^Q \left[Y_\tau \middle| \mathcal{F}_\rho \right].$$

This reads

$$\begin{aligned} & \max_{k \in K(\rho-, S_{\rho-})} \left(V(\rho, k) + b_{\rho-}^k \right) \\ \geq \;& E^Q \left[\int_\rho^\tau \frac{Z_\rho^0}{Z_s^0} d(-B_s) + \frac{Z_\rho^0}{Z_\tau^0} \max_{k \in K(\tau-, S_{\tau-})} \left(V(\tau, k) + b_{\tau-}^k \right) 1_{(\tau<T)} \middle| \mathcal{F}_\rho \right]. \end{aligned}$$

Supremum of the right hand side over all stopping times $\rho \leq \tau \leq T$ is obtained for $\tau = \rho$, since then an equality holds.

Given intervention at time ρ, the optimal intervention is to move S to

$$k = \arg \max_{k \in K(\rho-, S_{\rho-})} \left(V(\rho, k) + b_{\rho-}^k \right). \square$$

4.4 The American option in finance

In this section we consider the European and American options in finance. They are very well-studied objects in mathematical finance, and the results obtained here can also be found in almost every text book on mathematical finance, see e.g.- Lamberton and Lapeyre [41]. Nevertheless, the section serves well as an illustration of the framework introduced in Section 4.2 and of the results obtained in Section 4.3.

Consider the Black-Scholes market where

$$\begin{aligned} dS_t^0 &= rS_t^0 dt, \\ dS_t^1 &= \overline{\alpha} S_t^1 dt + \overline{\sigma} S_t^1 dW_t. \end{aligned}$$

These processes make up the two processes contained in Z. Introduce a European call option stating that a payment of $(S_T^1 - K)^+$ is paid out at time T, i.e.

$$\Delta B_T = \left(S_T^1 - K\right)^+.$$

From (4.4) it is seen that the set of martingale measures contains only one point,

$$h = \frac{r - \overline{\alpha}}{\overline{\sigma}}.$$

From Theorem 3 we get that for an arbitrage free contract, the reserve, and hereby the option price B_0, is determined by a regular solution to the differential equation (known as the Black-Scholes differential equation),

$$\begin{aligned} \partial_t V_t &= rV_t - rs\partial_s V_t - \frac{1}{2}\sigma^2 s^2 \partial_{ss} V_t, \\ V_{T-} &= (s - K)^+. \end{aligned}$$

Now we assume that the owner of the option is allowed to exercise prematurely which means that he can close the contract at any point in time $t < T$ and convert the future payment of $(S_T^1 - K)^+$ at time T into an immediate payment of $(S_t^1 - K)^+$ at time t. To handle this situation we introduce a third index S^2 (in addition to S^0 and S^1) which is not traded on the market, i.e. an index present in S but not in Z. This index is not affected by the underlying stochastic process W, but serves only to indicate whether the option is exercised or not. We assign to S_t^2 the value 0 if the option is not exercised at time t and the value 1 if the option is exercised at time t. Then

$$dS_t^2 = dN_t^I, \quad S_0^2 = 0,$$

where N_t^I counts the number of exercises until time t (equals 0 or 1). The payment due on transition of S^2 from 0 to 1 at time t is

$$b^1\left(t, S_t^0, S_t^1, 0\right) = \left(S_t^1 - K\right)^+.$$

From Theorem 4 we get that for an arbitrage free contract, the reserve, and hereby the option price B_0, is determined by a regular solution to the variational inequality,

$$\begin{aligned} \partial_t V(t,s) &\leq rV(t,s) - rs\partial_s V(t,s) - \frac{1}{2}\sigma^2 s^2 \partial_{ss} V(t,s), \\ V(t,s) &\geq (s-K)^+, \\ 0 &= \left(-\partial_t V(t,s) + rV(t,s) - rs\partial_s V(t,s) - \frac{1}{2}\sigma^2 s^2 \partial_{ss} V(t,s)\right) \\ &\quad \times \left((s-K)^+ - V(t,s)\right), \\ V(T-,s) &= (s-K)^+. \end{aligned} \tag{4.19}$$

The variational inequality (4.19) is a constructive tool for determination of the reserve V_t.

For the special case of a call option treated in this section, we also can draw conclusions from Theorem 5. Fix a time ρ. Now, Y_t in (4.18) is given by

$$Y_t = \left(Z_\rho^0 \left(\frac{S_t}{Z_t^0} - \frac{K}{Z_t^0}\right)\right)^+, \ \rho \leq t \leq T.$$

If $r \geq 0$, then $\frac{K}{Z_t^0}$ is a Q-supermartingale. Since $\frac{S_u}{Z_u^0}$ is a Q-martingale, $Z_\rho^0 \left(\frac{S_t}{Z_t^0} - \frac{K}{Z_t^0}\right)$ is then a Q-submartingale. Since the function $(\cdot)^+$ is convex, Jensen's inequality gives that Y_t is a Q-submartingale. Theorem 2 states that it is never optimal to exercise, and we conclude the well-known result that the reserve/price of an American call option equals the price of a European call option (in the Black-Scholes model).

4.5 The surrender option in life insurance

In this section, we consider a model where payments depend on the present state of X. Hoem obtained in [33] in this model a version of Thiele's differential equation which has taken a central position in life insurance mathematics and is widely used by practitioners. See also Chapter 2, example 7.1.

Let r be deterministic, put $K = 0$, define S by

$$\alpha_t^S = \begin{bmatrix} r_t S_t^0 \\ 0 \end{bmatrix}, \ \beta_t^S = \begin{bmatrix} 0^{1\times J} \\ 1 - S_t^1 \cdots J - S_t^1 \end{bmatrix}, \ s_0 = \begin{bmatrix} 1 \\ X_0 \end{bmatrix}, \tag{4.20}$$

and let $Z = S^0$. Thus, the market consists of the risk-free asset only and contains thereby no information on market prices of risk. We fix a martingale measure by assuming that the insurance company is risk-neutral with respect to risk due to the policy state, which is the only risk present in this model, i.e.

$$g_t = 0.$$

From Theorem 3, we see that for an arbitrage free contract the reserve, and hereby the single premium B_0 (or another balancing element of B), is determined by a regular solution to the differential equation (known as Thiele's differential equation),

$$\begin{aligned} \partial_t V_t &= b_t^c + r_t V_t - \left(b_t^d + V_t^{\mathcal{J}} - V_t^{1\times J}\right)\mu_t, \\ V_{T-} &= \Delta B_T. \end{aligned} \tag{4.21}$$

Now we assume that the policy holder is allowed to surrender the contract. This means that he can close the contract at any time $t < T$ and convert the future payments into an immediate payment of a surrender value which we shall denote by V_t^*. We assume that V_t^* only depends on S^1. To handle this situation we introduce a third index S^2 (in addition to S^0 and S^1) which is not traded on the market, i.e. a second index present in S but not in Z (in addition to S^1). This index is not affected by the underlying stochastic process N, but serves only to indicate whether the contract is surrendered or not. We assign to S_t^2 the value 0 if the contract is not surrendered at time t and the value 1 if the contract is surrendered at time t. Then

$$dS_t^2 = dN_t^I, \quad S_0^2 = 0,$$

where N_t^I counts the number of surrenders until time t (equals 0 or 1). The payment falling due upon transition of S^2 from 0 to 1 at time t is the surrender value, V_t^*, i.e.

$$b^1\left(t, S_t^0, S_t^1, 0\right) = V_t^*.$$

From Theorem 4 we get that for an arbitrage free contract the reserve, and hereby the single premium B_0 (or another balancing element of B), is determined by a regular solution to the variational inequality,

$$\begin{aligned} \partial_t V_t &\leq b_t^c + r_t V_t - \left(b_t^d + V_t^{\mathcal{J}} - V_t^{1\times J}\right)\mu_t, \\ V_t &\geq V_t^*, \\ 0 &= \left(-\partial_t V_t + b_t^c + r_t V_t - \left(b_t^d + V_t^{\mathcal{J}} - V_t^{1\times J}\right)\mu_t\right)\left(V_t^* - V_t\right), \\ V_{T-} &= \Delta B_T. \end{aligned} \tag{4.22}$$

We need to specify a surrender value V_t^*. One possibility could be to define as surrender value a so-called technical reserve defined as V_t in (4.5) but with Z^0 replaced by a technical risk-free asset Z^{0*} with interest rate r^* and the measure Q replaced by a technical measure Q^* given through a vector of technical transition intensities μ^*, i.e.

$$V_t^* = E^{Q^*}\left[\int_t^T \frac{Z_t^{0*}}{Z_s^{0*}} d\left(-B_s\right) \middle| S_t\right]. \tag{4.23}$$

To the reader without particular knowledge to life insurance mathematics, such a surrender value may seem rather odd, but this is actually how surrender values are calculated in practice. It is closely related to general pricing and accounting in life insurance, see Chapter 3. The surrender value V_t^* can, of course, be determined by the differential equation (4.21) with r and μ replaced by r^* and μ^*. By the variational inequality (4.22), we now have a constructive tool for determination of the reserve V_t.

Assuming that the surrender value is given by (4.23), we can draw conclusions from Theorem 5. Fix a time ρ. Now, Y_t in (4.18) is given by

$$Y_t = \int_\rho^t \frac{Z_\rho^0}{Z_s^0} d(-B_s) + \frac{Z_\rho^0}{Z_t^0} V_t^*,$$

such that Ito's lemma and (4.6) with (r, μ) replaced by (r^*, μ^*) gives

$$\begin{aligned}
\frac{Z_t^0}{Z_\rho^0} dY_t &= d(-B_t) - rV_t^* + dV_t^* \\
&= d(-B_t) - rV_t^* \\
&\quad + \left(b_t^c + r_t^* V_t^* - \left(b_t^d + V_t^{*\mathcal{J}} - V_t^{*1\times J}\right) \mu_t^*\right) dt \\
&\quad + \left(V_t^{*\mathcal{J}} - V_t^{*1\times J}\right) dN_t \\
&= (r^* - r) V_t^* - \left(b_t^d + V_t^{*\mathcal{J}} - V_t^{*1\times J}\right) \mu_t^* dt \\
&\quad + \left(b_{t-}^d + V_{t-}^{*\mathcal{J}} - V_{t-}^{*1\times J}\right) dN_t \\
&= (r^* - r) V_t^* + R_t^* (\mu_t - \mu_t^*) + R_{t-}^* dM_t^Q,
\end{aligned}$$

with

$$R_t^* = b_t^d + V_t^{*\mathcal{J}} - V_t^{*1\times J}.$$

We see that a sufficient condition for Y to be a Q-supermartingale is that the technical elements (r^*, μ^*) are chosen such that

$$\begin{aligned}
(r - r^*) V_t^* &\geq 0, \\
(\mu_t^* - \mu_t) R_t^* &\geq 0.
\end{aligned} \tag{4.24}$$

If these inequalities are fulfilled, it is then, according to Theorem 5, optimal to surrender immediately and the insurance company should set aside a reserve simply given by the technical reserve,

$$V_\rho = V_\rho^*.$$

4.6 The free policy option in life insurance

The exercise of an American option and the surrender of a life insurance contract result both in no payments beyond the date of exercise or surrender, respectively. In these cases the theory of optimal stopping would be adequate, and we would not have to introduce general theory of impulse control. Now we replace the surrender option of a life insurance contract introduced in Section 4.5 by a free policy option. A free policy option is an option to subscribe all future premium payments to zero against a corresponding subscription of all future benefits. What could be meant by a corresponding subscription will be explained below. After the time of conversion to free policy, the contract is not stopped but continues with converted payments and therefore general theory of impulse control applies. The free policy option could be considered in combination with the surrender option. Then one could allow for surrender of a policy before or after conversion into free policy or both. However, in order to keep things relatively simple we choose to disregard the surrender option in this section and take into account the free policy option exclusively.

We consider once more the classical multi-state life insurance policy described in Section 4.5, now with the free policy option. To handle this situation we introduce, as in Section 4.5, a third non-marketed index S^2, now indicating whether the contract is converted or not. We let

$$dS_t^2 = dN_t^I, \quad S_0^2 = 0.$$

where N_t^I counts the number of conversions until time t (equals 0 or 1). We assume that no payments fall due on transition of S^2 from 0 to 1, but now a free policy reserve has to be set aside. This will be a reserve for subscribed benefits. Since these subscribed payments, to be defined below, will depend on the time elapsed since conversion, we need a fourth index measuring this duration,

$$dS_t^3 = dt - S_{t-}^3 dN_t^I, \quad S_0^3 = 0.$$

Note that as long as $S_t^2 = 0$, we know that $S_t^3 = t$, the time elapsed since the issue of the contract. For all processes below, we abbreviate the argument $(t, S_t^0, S_t^1, S_t^2, S_t^3)$ by a subscript t and an argument (S_t^2, S_t^3), such that we e.g. can write $V_t(S_t^2, S_t^3)$ instead of $V(t, S_t^0, S_t^1, S_t^2, S_t^3)$. Then the fact that no payments fall due on transition of S^2 from 0 to 1 is written

$$b_t^1(0, t) = 0.$$

Let $\widehat{B}$ be the payment process given that the insurance contract is not yet converted, i.e. $dB_t(0, t) = d\widehat{B}_t$. We emphasize that this notation has no direct connection with first order payments introduced in Chapter 3. Now, we assume that the benefits are subscribed proportionally with a proportionality factor given by the ratio between the technical reserve and the technical free policy reserve at the conversion time. Then

$$dB_t\left(1, S_t^3\right) = d\widehat{B}_t^- \frac{V_{t-S_t^3}^*}{V_{t-S_t^3}^{*-}}, \tag{4.25}$$

where

$$\begin{aligned} d\widehat{B}_t^- &= \left(d\widehat{B}_t\right)^-, \\ V_{t-S_t^3}^* &= E^{Q^*}\left[\int_{t-S_t^3}^T \frac{Z_{t-S_t^3}^{0*}}{Z_s^{0*}} d\left(-\widehat{B}_s\right) \middle| S_{t-S_t^3}^1\right], \\ V_{t-S_t^3}^{*-} &= E^{Q^*}\left[\int_{t-S_t^3}^T \frac{Z_{t-S_t^3}^{0*}}{Z_s^{0*}} d\left(-\widehat{B}_s^-\right) \middle| S_{t-S_t^3}^1\right], \end{aligned}$$

and, accordingly,

$$\begin{aligned}V_t\left(1,S_t^3\right) &= E^Q\left[\left.\int_t^T \frac{Z_t^0}{Z_s^0} d\left(-B_s\left(1,S_t^3\right)\right)\right| S_t\right] \\ &= \frac{V^*_{t-S_t^3}}{V^{*-}_{t-S_t^3}} V_t^-,\end{aligned}$$

where

$$V_t^- = E^Q\left(\left.\int_t^T \frac{Z_t^0}{Z_s^0} d\left(-\widehat{B}_s^-\right)\right| S_t\right).$$

From Theorem 4 we get that for an arbitrage free contract the reserve, and hereby the single premium B_0 (or another balancing element of B), is given by a regular solution to the variational inequality (4.22) with V_t^* replaced by $V_t\left(1,0\right)$, the reserve for free policy benefits at the time of conversion.

Now we only need a differential system for calculation of the free policy reserve $V_t\left(1,S_t^3\right)$. Since no intervention is allowed for in a free policy, we see from Theorem 3 that for an arbitrage free contract, the free policy reserve $V_t\left(1,S_t^3\right)$ is determined by a regular solution to the differential equation ($V_t = V\left(t,s^0,s^1,1,s^3\right)$),

$$\begin{aligned}\partial_t V_t &= b_t^c + r_t V_t - \left(b_t^d + V_t^{\mathcal{J}} - V_t^{1\times J}\right)\mu_t - \partial_{s^3} V_t, \\ V_{T-} &= \Delta B_T.\end{aligned} \tag{4.26}$$

By (4.22) with V_t^* replaced by $V_t\left(1,0\right)$ and $V_t\left(1,S_t^3\right)$ given by (4.26), we now have a constructive tool for determination of the reserve V_t.

If the free policy payments are given by (4.25), we can draw conclusions from Theorem 5. Now, Y_t in (4.18) is given by

$$Y_t = \int_\rho^t \frac{Z_\rho^0}{Z_s^0} d\left(-B_s\left(0,s\right)\right) + \frac{Z_\rho^0}{Z_t^0} V_t\left(1,0\right),$$

such that Ito's lemma and (4.6) with (r,μ) replaced by (r^*,μ^*) gives

$$\begin{aligned}\frac{Z_t^0}{Z_\rho^0} dY_t &= d\left(-B_t\left(0,t\right)\right) - r_t\left(1,0\right)V_t\left(1,0\right) + dV_t\left(1,0\right) \\ &= d\left(-B_t\left(0,t\right)\right) - r_t\left(1,0\right)V_t\left(1,0\right) + \left(V_t^{*\mathcal{J}}\left(1,0\right) - V_t^{*1\times J}\left(1,0\right)\right)dN_t \\ &\quad + \left(b_t^c\left(1,0\right) + r_t\left(1,0\right)V_t\left(1,0\right)\right)dt \\ &\quad - \left(\left(b_t^d\left(1,0\right) + V_t^{\mathcal{J}}\left(1,0\right) - V_t^{1\times J}\left(1,0\right)\right)\mu_t\left(1,0\right) - \partial_{s^3}V_t\left(1,0\right)\right)dt \\ &= d\left(B_t\left(1,0\right) - B_t\left(0,t\right)\right) - \partial_{s^3}V_t\left(1,0\right)dt + R_t\left(1,0\right)dM_t^Q,\end{aligned} \tag{4.27}$$

with

$$R_t\left(1,0\right) = b_t^d\left(1,0\right) + V_t^{\mathcal{J}}\left(1,0\right) - V_t^{1\times J}\left(1,0\right).$$

For the first term of (4.27), we have

$$\begin{aligned} d\left(B_t(1,0) - B_t(0,t)\right) &= \frac{V_t^*}{V_t^{*-}} d\widehat{B}_t^- - d\widehat{B}_t \\ &= \left(\frac{V_t^*}{V_t^{*-}} \left(\widehat{b}_t^{c-} - \widehat{b}_t^{d-} \mu_t \right) - \widehat{b}_t^c + \widehat{b}_{t-}^d \mu_t \right) dt \\ &\quad + \left(-\frac{V_{t-}^*}{V_{t-}^{*-}} \widehat{b}_{t-}^{d-} + \widehat{b}_{t-}^d \right) dM_t^Q, \end{aligned} \tag{4.28}$$

and, introducing

$$\begin{aligned} R_t^* &= \widehat{b}_t^d + V_t^{*\mathcal{J}} - V_t^{*1\times J}, \\ R_t^- &= \widehat{b}_t^{d-} + V_t^{-\mathcal{J}} - V_t^{-1\times J}, \\ R_t^{*-} &= \widehat{b}_t^{d-} + V_t^{*-\mathcal{J}} - V_t^{*-1\times J}, \end{aligned}$$

we have for the second term of (4.27),

$$\begin{aligned} \partial_{s^3} V_t(1,0) &= \frac{V_t^-}{V_t^{*-}} \partial_{s^3} V_t^* - \frac{V_t^- V_t^*}{\left(V_t^{*-}\right)^2} \partial_{s^3} V_t^{*-} \\ &= -\frac{V_t^-}{V_t^{*-}} \partial_t V_t^* + \frac{V_t^- V_t^*}{\left(V_t^{*-}\right)^2} \partial_{s^3} V_t^{*-} \\ &= -\frac{V_t^-}{V_t^{*-}} \left(\widehat{b}_t^c + r_t^* V_t^* - R_t^* \mu_t^* \right) dt \\ &\quad + \frac{V_t^- V_t^*}{\left(V_t^{*-}\right)^2} \left(\widehat{b}_t^{c-} + r_t^* V_t^{*-} - R_t^{*-} \mu_t \right) dt \\ &= \left(-\frac{V_t^-}{V_t^{*-}} \left(\widehat{b}_t^c - R_t^* \mu_t^* \right) + \frac{V_t^- V_t^*}{\left(V_t^{*-}\right)^2} \left(\widehat{b}_t^{c-} - R_t^{*-} \mu_t \right) \right) dt. \end{aligned} \tag{4.29}$$

Then, by (4.27), (4.28), and (4.29), we get

$$\begin{aligned} \frac{Z_t^0}{Z_\rho^0} dY_t &= \left[\frac{V_t^-}{V_t^{*-}} \left(\widehat{b}_t^c - R_t^* \mu_t^* \right) - \frac{V_t^- V_t^*}{\left(V_t^{*-}\right)^2} \left(\widehat{b}_t^{c-} - R_t^{*-} \mu_t^* \right) \right. \\ &\quad \left. + \frac{V_t^*}{V_t^{*-}} \left(\widehat{b}_t^{c-} - \widehat{b}_t^{d-} \mu_t \right) - \widehat{b}_t^c + \widehat{b}_t^d \mu_t \right] dt \\ &\quad + \left(R_{t-}(1,0) - \frac{V_{t-}^*}{V_{t-}^{*-}} \widehat{b}_{t-}^{d-} + \widehat{b}_{t-}^d \right) dM_t^Q \\ &= \left[\left(\frac{V_t^-}{V_t^{*-}} - 1 \right) \left(\widehat{b}_t^c - \frac{V_t^*}{V_t^{*-}} \widehat{b}_t^{c-} - \left(\widehat{b}_t^d - \frac{V_t^*}{V_t^{*-}} \widehat{b}_t^{d-} \right) \mu_t \right) \right. \\ &\quad - \frac{V_t^-}{V_t^{*-}} \left(V_t^{*\mathcal{J}} - \frac{V_t^*}{V_t^{*-}} V_t^{*-\mathcal{J}} \right) \mu_t \\ &\quad \left. + \frac{V_t^-}{V_t^{*-}} \left(R_t^* - \frac{V_t^*}{V_t^{*-}} R_t^{*-} \right) \left(\mu_t - \mu_t^* \right) \right] dt \\ &\quad + \left(R_{t-}(1,0) - \frac{V_{t-}^*}{V_{t-}^{*-}} \widehat{b}_{t-}^{d-} + \widehat{b}_{t-}^d \right) dM_t^Q \end{aligned} \tag{4.30}$$

We are interested in the drift term of $\frac{Z_t^0}{Z_\rho^0}dY_t$ in order to draw conclusions from Theorem 5. It seems difficult to come up with general conditions for r^* and μ^* but we shall take a closer look at the drift term in a special case.

An example

We shall work with a simple insurance contract as an illustration of the drift term of (4.30). The insurance contract is a single life endowment insurance with a sum insured of 1 and a constant premium π as long as the insured is alive. Let X be the two-state process defined by $X_t = 0$ if the insured is alive at time t, $X_t = 1$ if the insured is dead at time t, and $X_0 = 0$. Then

$$\alpha_t^S = \begin{bmatrix} r_t S_t^0 \\ 0 \end{bmatrix},\ \beta_t^S = \begin{bmatrix} 0 \\ 1 \end{bmatrix},\ s_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Given that $S_t^1 = 0$, we have

$$\begin{aligned} \widehat{b}_t^c &= \pi, \\ \widehat{b}_t^{c-} &= 0, \\ \widehat{b}_t^d &= \widehat{b}_t^{d-} = 1, \end{aligned}$$

such that the drift term of $\frac{Z_t^0}{Z_\rho^0}dY_t$, (4.30), given $S_t^1 = 0$, equals

$$\left(\frac{V_t^-}{V_t^{*-}} - 1\right)\left(\pi - \left(1 - \frac{V_t^*}{V_t^{*-}}\right)\mu_t\right) + \frac{V_t^-}{V_t^{*-}}\left(1 - \frac{V_t^*}{V_t^{*-}}\right)(\mu_t - \mu_t^*).$$

Introducing traditional actuarial notation, one can write this drift term as

$$\begin{aligned} &\left(\frac{\overline{A}_{x+t\,\overline{T-t}|}}{\overline{A}^*_{x+t\,\overline{T-t}|}} - 1\right)\left(\pi - \left(1 - \frac{\overline{A}^*_{x+t\,\overline{T-t}|} - \pi\overline{a}^*_{x+t\,\overline{T-t}|}}{\overline{A}^*_{x+t\,\overline{T-t}|}}\right)\mu_t\right) \\ &+\frac{\overline{A}_{x+t\,\overline{T-t}|}}{\overline{A}^*_{x+t\,\overline{T-t}|}}\left(1 - \frac{\overline{A}^*_{x+t\,\overline{T-t}|} - \pi\overline{a}^*_{x+t\,\overline{T-t}|}}{\overline{A}^*_{x+t\overline{T-t}|}}\right)(\mu_t - \mu_t^*) \\ = \;&\left(\frac{\overline{A}_{x+t\,\overline{T-t}|}}{\overline{A}^*_{x+t\,\overline{T-t}|}} - 1\right)\frac{\pi}{\overline{A}^*_{x+t\,\overline{T-t}|}}\left({}_{T-t}E^*_{x+t} + \int_t^T \overline{a}^*_{\overline{s-t}|}\,{}_{s-t}p^*_{x+t}\left(\mu^*_{x+s} - \mu_{x+t}\right)ds\right) \\ &+\frac{\overline{A}_{x+t\,\overline{T-t}|}\pi\overline{a}^*_{x+t\,\overline{T-t}|}}{\left(\overline{A}^*_{x+t\,\overline{T-t}|}\right)^2}(\mu_t - \mu_t^*). \end{aligned}$$

We are interested in sufficient conditions for this quantity to be non-positive such that Y is a Q-supermartingale. Term by term, it can be seen that this will be the case if $\mu^* > \mu$, $r > r^*$, and μ^* is increasing. According to Theorem 5, it is now optimal to convert the into a free policy immediately and the insurance company should set aside the free policy reserve, $V_t(1,0)$.

It is worth noting that these conditions on μ^* and r^* also would impose optimality of immediate surrender in the case of the surrender option: If μ^* is increasing, V_t^* and R_t^* are non-negative for all $t \le T$. Now, this conclusion follows from (4.24).

Part III

Control in life and pension insurance

Chapter 5

Risk-adjusted utility

This chapter introduces an idea of risk-adjusted utility. Instead of measuring moral value or utility of nominal value, we suggest to measure utility of deflated value. As deflator is chosen the same deflator which is used for determination of price or financial value. Using this concept we study the problem of optimal investment-consumption and the problem of utility indifference pricing in an incomplete market.

5.1 Introduction

In this chapter, we study two applications of a very particular state-dependence of utility which leads to what we choose to call risk-adjusted utility. The two applications are in the fields of optimal investment-consumption and pricing in incomplete markets.

We start out by the problem of optimal consumption and portfolio selection of an agent playing the role as both investor and consumer. The work of Robert Merton around 1970 is usually considered as the starting point of the continuous-time formulation of the problem, see references in Merton [45]. In the work by Merton, the preferences of an agent over consumption and wealth are given by time-additive utility functions, and he solves for some specific markets and utility functions explicitly the problem of choosing consumption and investment in order to maximize expected total utility. Also in this chapter, we shall work in a continuous-time framework and base decisions about consumption and investment on the same fundamental idea. However, on various points, the objective of our decisions differs from that of Merton.

Whereas Merton applied control theory to his problem, several authors around 1990 approached the problem with martingale methods from finance. The idea is to split the dynamic decision problem into a static decision problem and a hedging problem. Both methods have, compared to their original formulations, been refined in various directions in order to obtain results that are more applicable. Important are constraints on trading strategies, consumption, and wealth. See Korn [40] for the martingale method, references on constrained problems, and for refinements connected with the introduction of transaction costs.

The preference ordering in Merton's problem is rather special. A generalized ordering of preferences is obtained by considering state-dependent utilities. However, generalizations in this direction in general increases the number of state variables and complicates the problem accordingly. In this chapter, we shall work with a very particular state-dependence of utility which does not add difficulties to the problem. On the contrary, it turns out that our state-dependent utility simplifies the problem considerably since it separates the investment problem and the consumption problem into two problems which can be solved independently of each other. The idea is that the agent in his preferences takes into consideration the attitudes to risk to which the market has already made up its mind. This is based on the usual assumption that the decisions of the agent do not affect the market's attitudes to risk.

Utility plays an important role in pricing in incomplete markets. We shall also study the impacts of risk-adjusted utility on pricing by utility indifference. In particular, we shall consider the exponential principle, the variance principle, and the standard deviation principle studied in actuarial literature. The variance and the standard deviation principles are connected with rather special utility functions, as explained in Schweizer [60]. There, financial versions of these principles, consistent with absence of arbitrage, are introduced. Working with risk-adjusted utilities, we come up with a variance principle and a standard deviation principle different from those obtained in Schweizer [60] but still consistent with absence of arbitrage. For a survey over utility functions and applications to finance and insurance, see Gerber and Pafumi [26].

The reader will realize a partial change of notation compared to Chapters 2-4. We choose, throughout the remaining of the thesis, to conform to a notation which is more conventional in finance and control. This serves to emphasize that the problems and results go beyond optimization in life insurance. In fact, the present chapter demonstrates no connection to decision problems in life and pension insurance, whatsoever, but prepares the reader to Chapter 6. In that chapter insurance is our main application, and Section 6.3 helps the reader to interpret the results of Chapter 6 in terms of life and pension insurance.

In Section 5.2 we introduce the market, the wealth process, and the objective in the traditional optimization problem. In Section 5.3 we introduce the concept of risk-adjusted utility and formulate a new optimization problem arising from these state-dependent utilities. In Section 5.4 we solve in general the problem of optimal investment. An optimal consumption plan is obtained in the case where utility of consumption and terminal wealth coincide and in the case where there is no utility of terminal wealth but the constraint that the financial value of consumption equals initial wealth. Finally, in Section 5.5 we study the impacts of risk-adjusted utility on utility indifference pricing in incomplete markets.

5.2 The traditional optimization problem

We consider the problem of optimal consumption and portfolio selection of an agent where the investment possibilities are provided by a Black-Scholes market. In this market, the dynamics of the price processes of the bond, $(B(t))_{t\geq 0}$, and the stock, $(S(t))_{t\geq 0}$, are given by

$$\begin{aligned} dB(t) &= B(t)\, r dt,\ B(0) = 1, \\ dS(t) &= S(t)\left(\mu dt + \beta dW(t)\right),\ S(0) = s, \end{aligned} \tag{5.1}$$

respectively, where r, μ, and β are constants and W is a Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \mathbf{F} = \{\mathcal{F}_t\}_{t\geq 0}, P)$. Let $\{\mathcal{F}_t^W\}_{t\geq 0}$ denote the filtration generated by W.

Fix a time horizon T. A dynamic portfolio strategy is a predictable process, $(\eta^B, \eta^S) = (\eta^B(t), \eta^S(t))_{0\leq t\leq T}$, where $\eta^B(t)$ and $\eta^S(t)$ denote the number of units held at time t of B and S, respectively. The wealth at time t of the agent holding the portfolio $(\eta^S(t), \eta^B(t))$ is given by

$$X'(t) = \eta^S(t)\, S(t) + \eta^B(t)\, B(t), \tag{5.2}$$

and the relative portfolio in the risky asset is defined by

$$\theta(t) = \frac{\eta^S(t)\, S(t)}{X'(t)}. \tag{5.3}$$

We assume that the cumulative consumption of the agent up to time t is given by $U'(t) = \int_0^t u'(s)\, ds$, such that the wealth of the agent following a self-financing investment-consumption strategy is given by

$$X'(t) = x_0 + \int_0^t \left(-u'(s)\, ds + \eta^S(s)\, dS(s) + \eta^B(s)\, dB(s)\right), \tag{5.4}$$

where x_0 is the initial wealth. Then, by (5.1), (5.2), (5.3), and (5.4),

$$\begin{aligned} dX'(t) &= -u'(t)\, dt + \eta^S(t)\, dS(t) + \eta^B(t)\, dB(t) \\ &= -u'(t)\, dt + X'(t)\left(r + \theta(t)(\mu - r)\right) dt \\ &\quad + X'(t)\,\theta(t)\,\beta dW(t). \end{aligned} \tag{5.5}$$

The agent's preferences over consumption and wealth profiles in a traditional problem of optimal investment-consumption are given by time-additive utility functions υ for consumption and Υ for the terminal wealth. Given an endowment at time t of x, an agent wishes to choose a consumption profile and an investment policy so as to maximize his total expected utility of consumption over $(t, T]$ and utility of wealth at time T using feasible policies. The agent wishes to maximize

$$J(t, x, \theta, u') = E\left[\int_t^T \upsilon(s, u'(s))\, ds + \Upsilon(X'(T)) \,\middle|\, X'(t) = x\right],$$

and the value function of the agent is defined by

$$V(t,x) = \max_{\theta,u'} J(t,x,\theta,u').$$

The utility functions υ and Υ are assumed to be concave, and one typically works with the power function, the logarithmic function, or the negative exponential function. We use the terms utility and disutility functions for general (not necessarily increasing, concave/convex and continuously differentiable) reward and cost functions, respectively. For derivatives of V we shall use the notation $V_t = \frac{\partial}{\partial t}V(t,x)$, $V_x = \frac{\partial}{\partial x}V(t,x)$, and $V_{xx} = \frac{\partial^2}{\partial x^2}V(t,x)$.

This traditional formulation of the problem can be criticized in various respects. The ordering of preferences is rather special and, in particular, the preferences of the agent are independent of the attitudes to risk taken up by the market. It seems natural to let the utility of wealth and consumption at time t depend on the state of the world at that point in time. In general, state-dependent utilities complicate the optimization problem, but we shall take a stand point which turns out to simplify the problem considerably.

5.3 Risk-adjusted utility

Instead of measuring and comparing payments by utility or moral value we shall make an excursion to derivative pricing and compare payments by price or financial value. The market described above is complete and takes a unique position on the price of W-risk. From arbitrage pricing theory we know that the price or financial value at time 0 of an $\mathcal{F}_t^W$-measurable payment of $u'(t)$ at time t is given by $E(\Lambda(t)u'(t))$, where the dynamics of the deflator Λ is given by

$$d\Lambda(t) = -\Lambda(t)\left(rdt + \frac{\mu - r}{\beta}dW(t)\right), \Lambda(0) = 1. \tag{5.6}$$

Thus, taking stand at time 0, the deflator Λ plays a crucial role in measuring and comparing prices or financial values of payments at different points in time. Arbitrage pricing theory tells us that Λ is the appropriate yardstick to use. Moreover, for financial valuation, the yardstick Λ makes payments time-additive in the sense that the price of the payment $u'(s)$ at time s plus the payment $u'(t)$ at time t equals $E[\Lambda(s)u'(s) + \Lambda(t)u'(t)]$.

When it comes to utility or moral value of money, one usually measures moral value of nominal payments. Instead, we suggest using the yardstick from financial valuation. More precisely, we suggest to measure at time 0 the moral value of a payment of $u'(t)$ at time t by the quantity

$$E[\upsilon(\Lambda(t)u'(t))], \tag{5.7}$$

for some utility function υ. Thereby, the agent takes into account the market's attitudes to risk before making decisions based on his own individual utility function.

This seems to be a controversial suggestion, but we believe that it makes sense to use the same deflator for calculating financial and moral value of money. Furthermore, it simplifies our problem considerably and the inclusion of a certain terminal constraint in particular. Note, that our approach has no direct connection with the criticizable approach to deflate by Λ the utility of nominal value,

$$E\left[\Lambda(t)\,\upsilon\left(t,u'(t)\right)\right].$$

In (5.7) we have skipped the time-dependence, assuming that the deflator Λ not only adjusts for utility in different states at time t, but also adjusts for utility at different points in time. This assumption could easily be weakened, at least in the finite time problem studied in this chapter.

One way of thinking of the state-dependent utility suggested above is that prices are deflated by Λ. If a consumer has preferences over purchasing power given by a utility function υ, the quantity to be measured is exactly purchasing power rather than the nominal amount of money. Using Λ as price deflator relates to the assertion that prices go up when the market goes up – or the other way around. Though this seems natural from a macro-economic point of view, we admit that our choice of deflator is not underpinned by basic economic principles, but is rather a matter of mathematical convenience.

After having argued for deflation of money, we admit that there are still critical remarks to be made on the time-additivity. In the classical situation the marginal utility of consumption at time t is independent of the amount of consumption at time s, no matter how close s is to t. This seems to be criticizable. However, using the picture in the previous paragraph, the same criticism goes for purchasing power in our approach.

Abbreviating

$$\begin{aligned} X(t) &= \Lambda(t)\,X'(t)\,, \\ u(t) &= \Lambda(t)\,u'(t)\,, \end{aligned} \tag{5.8}$$

we see, by using Ito's formula and (5.1)-(5.6), that the dynamics of X can be written in various forms,

$$\begin{aligned} dX(t) &= d\left(\eta^S(t)\,\Lambda(t)\,S(t)+\eta^B(t)\,\Lambda(t)\,B(t)\right) \\ &= d\left(\Lambda(t)\,x+\Lambda(t)\int_0^t\left(-u'(s)\,ds+\eta^S(s)\,dS(s)+\eta^B(s)\,dB(s)\right)\right) \\ &= -u(t)\,dt+\eta^B(t)\,d\left(\Lambda(t)\,B(t)\right)+\eta^S(t)\,d\left(\Lambda(t)\,S(t)\right) && (5.9) \\ &= -u(t)\,dt+X(t)\left(\theta(t)\,\beta-\frac{\mu-r}{\beta}\right)dW(t)\,, && (5.10) \end{aligned}$$

and that (5.7) can be written as

$$E\left[\upsilon\left(u(t)\right)\right].$$

5.4 Optimal investment and consumption

In this section, we show that decisions concerning investment and decisions concerning consumption can be separated in the sense that the optimal investment policy and the optimal consumption process can be found independently of each other. We find the optimal investment policy and for particular cases also the optimal consumption plan. Let

$$J(t,x,\theta,u) = E\left[\int_t^T \upsilon(u(s))\,ds + \Upsilon(X(T))\middle| X(t)=x\right],$$

and let the value function of the agent be defined by

$$V(t,x) = \max_{\theta,u} J(t,x,\theta,u).$$

The Hamilton-Jacobi-Bellman equation for the optimization problem is

$$-V_t = \max_{u,\theta}\left[-uV_x + \frac{1}{2}x^2\left(\theta\beta + \frac{r-\mu}{\beta}\right)^2 V_{xx} + \upsilon(u)\right]. \tag{5.11}$$

Differentiation of (5.11) with respect to θ and equating the right hand side to zero results in the optimizer,

$$\theta(t) = \frac{\mu - r}{\beta^2}, \tag{5.12}$$

and differentiation of (5.11) once more shows that θ is maximizing if V is concave.

Proposition 6 *If υ and Υ are concave, then V is concave.*

Sketch of proof. Consider two initial points x_1 and x_2 and strategies (u_1,θ_1) and (u_2,θ_2) and let $\lambda \in (0,1)$. We can now define a new strategy, $(u,\theta) = (\lambda u_1 + (1-\lambda)u_2, \lambda\theta_1 + (1-\lambda)\theta_2)$ for the initial point $x = \lambda x_1 + (1-\lambda)x_2$, and we get by linearity of X, $X(t) = \lambda X_1(t) + (1-\lambda)X_2(t)$. Since υ and Υ are concave, we know that

$$\begin{aligned}\upsilon(u(t)) &\geq \lambda\upsilon(u_1(t)) + (1-\lambda)\upsilon(u_2(t)),\\ \Upsilon(X(T)) &\geq \lambda\Upsilon(X_1(T)) + (1-\lambda)\Upsilon(X_2(T)),\end{aligned}$$

Now, it follows that

$$J(t,x,\theta,u) \geq \lambda J(t,x_1,\theta_1,u_1) + (1-\lambda)J(t,x_2,\theta_2,u_2).$$

For any $\varepsilon > 0$, we can choose (u_1,θ_1) such that $J(t,x_1,\theta_1,u_1) \geq V(t,x_1) - \varepsilon$ and (u_2,θ_2) such that $J(t,x_2,\theta_2,u_2) \geq V(t,x_2) - \varepsilon$. Since (u,θ) is suboptimal we have

$$V(t,x) \geq \lambda V(t,x_1) + (1-\lambda)V(t,x_2) - \varepsilon,$$

and since ε was arbitrary, concavity follows.□

It follows from Proposition 6 and the comment following (5.12) that the optimal investment policy is given by

$$\theta(t) = \frac{\mu - r}{\beta^2}.$$

Inserting this investment policy in (5.11), we get

$$-V_t = \max_u \left[-uV_x + \upsilon(u)\right]. \tag{5.13}$$

This is optimized by

$$u(t) = \upsilon_u^{-1}(V_x), \tag{5.14}$$

and even maximized at this point if υ is concave. Plugging (5.14) into the partial differential equation (5.13) gives the equation

$$V_t - \upsilon_u^{-1}(V_x) V_x + \upsilon\left(\upsilon_u^{-1}(V_x)\right) = 0, \tag{5.15}$$

which is to be solved subject to the terminal condition, $V(T, x) = \Upsilon(x)$.

5.4.1 Terminal utility

We consider the situation where utility of consumption coincides with utility of terminal wealth,

$$\Upsilon = \upsilon.$$

We guess a solution to (5.15) in the form

$$V(t, x) = (T + 1 - t)\, \upsilon\left(\frac{x}{T + 1 - t}\right),$$

and find

$$\begin{aligned} V_t &= -\upsilon\left(\frac{x}{T+1-t}\right) + \frac{x}{T+1-t} \upsilon_u\left(\frac{x}{T+1-t}\right), \\ V_x &= \upsilon_u\left(\frac{x}{T+1-t}\right), \end{aligned}$$

such that (5.15) is fulfilled and $V(T, x) = \upsilon(x) = \Upsilon(x)$.

The optimal consumption is given by

$$u(t) = \frac{X(t)}{T + 1 - t},$$

or, by (5.8), in nominal terms,

$$u'(t) = \frac{X'(t)}{T + 1 - t}.$$

This is the same solution as Merton's solution for the special case of logarithmic utility. In our formulation with risk-adjusted utility, this result is not dependent on a particular form of υ but on the concavity only.

5.4.2 Terminal constraint

We consider the problem with a terminal constraint,

$$E\left[X\left(T\right)\right]=0. \tag{5.16}$$

From (5.10) and (5.16) we get that

$$x_0=E\left[\int_0^T \Lambda\left(t\right)u'\left(t\right)dt\right], \tag{5.17}$$

and we see that (5.16) amounts to requiring that the financial value of the consumption over $(0,T]$ is equal to initial wealth.

Remark 7 *For readers familiar with optimization problems in finance, the constraint (5.16) is recognized from the martingale approach to optimal consumption. In a complete market with positive consumption, (5.16) replaces the admissibility condition, $X\left(t\right)\geq 0$, $0\leq t\leq T$. This condition will consequently be fulfilled in our solution below. However, this is not our motivation for (5.16). In Chapter 6 we shall work with a constraint similar to (5.16) in an incomplete framework where there are no requirements, in general, concerning positive wealth and positive consumption. The terminal constraint can there be motivated by a no arbitrage condition in a life and pension insurance portfolio, as we shall see in Chapter 6.*

We guess a solution to (5.15) in the form

$$V\left(t,x\right)=\left(T-t\right)\upsilon\left(\frac{x}{T-t}\right),$$

and find

$$\begin{aligned} V_t &= -\upsilon\left(\frac{x}{T-t}\right)+\frac{x}{T-t}\upsilon_u\left(\frac{x}{T-t}\right), \\ V_x &= \upsilon_u\left(\frac{x}{T-t}\right), \end{aligned}$$

such that (5.15) is fulfilled. Now we need to find the optimal control and check the constraint.

The optimal consumption is given by

$$u\left(t\right)=\frac{X\left(t\right)}{T-t},$$

or, by (5.8), in nominal terms,

$$u'\left(t\right)=\frac{X'\left(t\right)}{T-t},$$

such that the dynamics of the optimally controlled process is

$$dX\left(t\right)=-\frac{X\left(t\right)}{T-t}dt,\; X\left(0\right)=x_0.$$

This is just an ordinary differential equation with the solution

$$X(t) = \frac{(T-t)\, x_0}{T},$$

such that

$$X(T) = 0. \tag{5.18}$$

Thus, the constraint (5.16) is fulfilled. The reason why we even arrive at (5.18) is that the optimal consumption problem with Hamilton-Jacobi-Bellman equation given by (5.13) is actually a deterministic control problem, although the nominal process X' is still stochastic.

5.5 Pricing by risk-adjusted utility

In this section, we consider the problem of pricing claims in an incomplete market by the principle of equivalent utility. We put consumption equal to 0 throughout the section and denote by $X'_x(t)$ the wealth process of an agent with initial wealth x, defined by (5.4) with $u' = 0$. Consider pricing of γ units of the T-claim H', and let H be a T-claim defined by

$$H = \Lambda(T)\, H'.$$

Let $\pi(x,\gamma)$ denote the price of γ units of H' for an investor with initial wealth x. The market is incomplete in the sense that H' is not necessarily $\mathcal{F}_T^W$-measurable. In Schweizer [60], a general utility indifference pricing principle is suggested,

$$\sup_\theta E\left[\upsilon\left(X'_{x+\pi(x,\gamma)}(T) - \gamma H'\right)\right] = \sup_\theta E\left[\upsilon\left(X'_x(T)\right)\right]. \tag{5.19}$$

However, following the idea of risk-adjusted utility, we shall instead propose to derive the price from

$$\sup_\theta E\left[\upsilon\left(\Lambda(T)\left(X'_{x+\pi(x,\gamma)}(T) - \gamma H'\right)\right)\right] = \sup_\theta E\left[\upsilon\left(\Lambda(T)\, X'_x(T)\right)\right],$$

equivalent to

$$\sup_\theta E\left[\upsilon\left(X_{x+\pi(x,\gamma)}(T) - \gamma V^H(T)\right)\right] = \sup_\theta E\left[\upsilon\left(X_x(T)\right)\right], \tag{5.20}$$

where

$$V^H(t) = E\left[H \mid \mathcal{F}_t\right].$$

Since the market is incomplete, it is no longer clear that the risk-adjustment factor should be Λ. In fact, there will be one risk-adjustment factor for each martingale measure, and the question is which factor to use. The idea with risk-adjustment of utility is to take into consideration the market's attitude to risk. Adding to this that this consideration can only be taken into account to the extent that the market really takes a position, we suggest to use the *minimal risk-adjustment* corresponding to discounting and a change of measure into the *minimal martingale measure* (see e.g. Schweizer [61] for an account on the notion of minimal martingale measure). The factor Λ is exactly such a minimal risk-adjustment.

Consider the right hand side of (5.20). If υ is concave, we have by Jensen's inequality

$$E\left[\upsilon\left(X_x\left(T\right)\right)\right] \leq \upsilon\left(E\left[X_x\left(T\right)\right]\right) = \upsilon\left(x\right).$$

Since equality is obtained by the strategy

$$\theta\left(t\right) = \frac{\mu - r}{\beta^2},$$

the supremum is obtained by this strategy, and the right hand side of (5.20) becomes $\upsilon\left(x\right)$.

In order to represent the optimizing strategy for the left hand side of (5.20), we shall write the martingale V^H, according to the *Galtchouk-Kunita-Watanabe decomposition*, uniquely as

$$V^H\left(T\right) = V^H\left(0\right) + \int_0^T \xi^H\left(t\right) dW\left(t\right) + L^H, \tag{5.21}$$

such that

$$E\left[L^H\right] = 0 \tag{5.22}$$

and

$$E\left[L^H \int_0^T \xi\left(t\right) dW\left(t\right)\right] = 0 \tag{5.23}$$

for any $\left(\xi\left(t\right)\right)_{t\geq 0}$. From (5.21), (5.9), and (5.10) with $u = 0$ we get that

$$\begin{aligned}
& X_{x+\pi(x,\gamma)}\left(T\right) - \gamma V^H\left(T\right) \\
= \; & x + \pi\left(x,\gamma\right) + \int_0^T \eta^B\left(t\right) d\left(\Lambda\left(t\right) B\left(t\right)\right) + \int_0^T \eta^S\left(t\right) d\left(\Lambda\left(t\right) S\left(t\right)\right) \\
& -\gamma V^H\left(0\right) - \gamma \int_0^T \xi^H\left(t\right) dW\left(t\right) - \gamma L^H \\
= \; & x + \pi\left(x,\gamma\right) - \gamma V^H\left(0\right) - \gamma L^H + \int_0^T h\left(t\right) dW\left(t\right),
\end{aligned} \tag{5.24}$$

where

$$h\left(t\right) = X_{x+\pi(x,\gamma)}\left(t\right)\left(\theta\left(t\right)\beta - \frac{\mu - r}{\beta}\right) - \gamma\xi^H\left(t\right). \tag{5.25}$$

Consider the left hand side of (5.20). If υ is concave, we have by Jensen's inequality, (5.24), and (5.22),

$$\begin{aligned}
E\left[\upsilon\left(X_{x+\pi(x,\gamma)}\left(T\right) - \gamma V^H\left(T\right)\right)\right] & \leq \upsilon\left(E\left[X_{x+\pi(x,\gamma)}\left(T\right) - \gamma V^H\left(T\right)\right]\right) \\
& = \upsilon\left(x + \pi\left(x,\gamma\right) - \gamma V^H\left(0\right)\right).
\end{aligned}$$

In general we cannot come up with a strategy such that equality holds. If H is $\mathcal{F}_T^W$-measurable, however, such that $L^H = 0$, we get an equality by putting $h(t) = 0$, i.e.

$$\theta(t) = \frac{\mu - r}{\beta^2} + \frac{\gamma \xi^H(t)}{\beta X_{x+\pi(x,\gamma)}(t)}.$$

Then (5.20) becomes

$$\upsilon\left(x + \pi(x,\gamma) - \gamma V^H(0)\right) = \upsilon(x),$$

and the equivalence price is given by

$$\pi(x,\gamma) = \gamma V^H(0) = \gamma E\left[\Lambda(T) H'\right].$$

From this we conclude that our pricing approach is consistent with absence of arbitrage.

5.5.1 Exponential utility

Consider the case

$$\upsilon(x) = \frac{1}{a}\left(1 - e^{-ax}\right).$$

Since exponential utility is concave, we get by (5.20) and the remarks above

$$\pi(x,\gamma) = \frac{1}{a}\log \inf_\theta E\left[e^{a\left(\gamma V^H(0) + \gamma L^H - \int_0^T h(t) dW(t)\right)}\right],$$

with $h(t)$ given by (5.25).

In general, we cannot say much about this expression. However, if H is independent of S we have that

$$\begin{aligned} L^H &= H - E[H], \\ \xi^H &= 0, \end{aligned}$$

such that

$$\begin{aligned} \pi(x,\gamma) &= \frac{1}{a}\log\left(E\left[e^{a\gamma H}\right] \inf_\theta E\left[e^{a\int_0^T \left(X(t)_{x+\pi(x,\gamma)}\left(\theta(t)\beta - \frac{\mu-r}{\beta}\right)\right) dW(t)}\right]\right) \\ &= \frac{1}{a}\log E\left[e^{a\gamma H}\right], \end{aligned} \tag{5.26}$$

where infimum is obtained by

$$\theta(t) = \frac{\mu - r}{\beta^2}.$$

The pricing formula (5.26) is almost equivalent to the traditional actuarial exponential principle. However, it is important to notice that it is the claim H (not H') for which we have assumed independence of S, and which appears in (5.26). A reference to pricing and hedging by indifferent exponential utility (not risk-adjusted) with applications to insurance is Becherer [3].

5.5.2 Mean-variance utility

We shall now consider a rather special utility, $\upsilon : L^2 \to \mathbf{R}$, given by

$$\upsilon(X) = E[X] - a(Var[X])^c, \; X \in L^2. \tag{5.27}$$

This utility function (see Schweizer [60] for further comments on this expression) covers both the variance principle ($c = 1$) and the standard deviation principle $\left(c = \frac{1}{2}\right)$, and we call a utility function in the form (5.27) for mean-variance utility. In Dana [14] utility is said to be mean-variance if there exists a function $f : \mathbf{R} \times \mathbf{R}_+ \to \mathbf{R}$ such that $\upsilon(X) = f(E[X], Var[X])$, $X \in L^2$. Note that in case of mean-variance utility, the expectation taken on both sides of (5.20) and (5.19) is redundant. Now (5.20) becomes

$$\begin{aligned} & \sup_\theta \left(E\left[X_{x+\pi(x,\gamma)}(T) - \gamma V^H(T)\right] - a\left(Var\left[X_{x+\pi(x,\gamma)}(T) - \gamma V^H(T)\right]\right)^c\right) \\ = & \sup_\theta \left(E[X_x(T)] - a(Var[X_x(T)])^c\right) \end{aligned}$$

or, equivalently,

$$\begin{aligned} & x + \pi(x,\gamma) - \gamma V^H(0) - a \inf_\theta \left(Var\left[X_{x+\pi(x,\gamma)}(T) - \gamma V^H(T)\right]\right)^c \\ = & \; x - a \inf_\theta (Var[X_x(T)])^c. \end{aligned} \tag{5.28}$$

Consider the right hand side of (5.28). Here we have that

$$(Var[X_x(T)])^c \geq 0,$$

and since equality is obtained by the strategy

$$\theta(t) = \frac{\mu - r}{\beta^2},$$

the infimum is obtained by this strategy. Thus, the right hand side of (5.28) equals x.

Consider the left hand side of (5.28). Here, by monotonicity of the power function of positive numbers, (5.23), and (5.24), we have that

$$\begin{aligned} & \inf_\theta \left(Var\left[X_{x+\pi(x,\gamma)}(T) - \gamma V^H(T)\right]\right)^c \\ = & \inf_\theta \left(Var\left[\gamma L^H + \int_0^T h(t)\, dW(t)\right]\right)^c \\ = & \left(Var\left[\gamma L^H\right] + \inf_\theta Var\left[\int_0^T h(t)\, dW(t)\right]\right)^c, \end{aligned}$$

with h given by (5.25). We know that

$$Var\left[\int_0^T h(t)\, dW(t)\right] \geq 0,$$

and since equality is obtained by the strategy given by $h(t) = 0$, i.e.

$$\theta(t) = \frac{\mu - r}{\beta^2} + \frac{\gamma \xi^H(t)}{\beta X_{x+\pi(x,\gamma)}(t)},$$

we get

$$\begin{aligned}\pi(x,\gamma) &= \gamma V^H(0) + a\gamma^{2c}\left(Var\left[L^H\right]\right)^c \\ &= \gamma E\left[\Lambda(T) H'\right] + a\gamma^{2c}\left(Var\left[L^H\right]\right)^c.\end{aligned}$$

Thus, in the case of mean-variance utility, the absence of correlation given by (5.23) is sufficient for obtaining an explicit investment strategy and an explicit pricing formula.

If H is attainable, we have that $L^H = 0$, such that

$$\pi(x,\gamma) = \gamma E\left[\Lambda(T) H'\right].$$

Thus, the pricing formula is consistent with absence of arbitrage.

If H is independent of S, we have that

$$L^H = H - E[H],$$

such that

$$\pi(x,\gamma) = \gamma E[H] + a\gamma^{2c}(Var[H])^c \tag{5.29}$$

and

$$\theta(t) = \frac{\mu - r}{\beta^2}.$$

As in the case of exponential utility, we get by (5.29) a formula almost equivalent to the traditional actuarial pricing formula. Again, one should notice that it is the claim H (not H') for which we have assumed independence of S, and which appears in (5.29). A reference to pricing and hedging by indifferent mean-variance utility (not risk-adjusted) with applications to insurance is Møller [48].

Chapter 6

Optimal investment and consumption in life and pension insurance

In this chapter, we study the problem of optimal consumption and portfolio selection of an agent. The decision problem takes into consideration risky income and risky debt. As opposed to the traditional formulation of the problem we shall formulate the preferences of the agent in terms of so-called risk-adjusted utility. This leads to intuitively reasonable strategies for consumption and investment under quadratic and absolute disutility. We give an interpretation of the set-up in terms of investment and distribution of surplus in life and pension insurance such that the results can be applied there. As a matter of precaution we remind the reader about the partial change in notation, compared to Chapters 2-4, which started in Chapter 5 and continues in the present chapter.

6.1 Introduction

In this chapter, we study the problem of optimal consumption and portfolio selection of an agent playing the role as both investor and consumer. The work of Robert Merton around 1970 is usually considered as the starting point of the continuous-time formulation of the problem, see references in Merton [45]. In the work by Merton, the preferences of an agent over consumption and wealth are given by time-additive utility functions, and he solves for some specific markets and utility functions explicitly the problem of choosing consumption and investment in order to maximize expected total utility. In this chapter, we shall also work in a continuous-time framework and base decisions about consumption and investment on the same fundamental idea. However, in various respects the objective of our decisions differs from that of Merton.

Whereas Merton applied control theory to his problem, several authors around 1990 approached the problem with martingale methods from finance. The idea is to separate the dynamic decision problem into a static decision problem and a hedging problem. Both methods have, compared to their original formulations, been refined in various directions in order to obtain results that are more applicable. Important are constraints on trading strategies, consumption and wealth. See Korn [40] for the martingale method, references on constrained problems, and for refinements connected with the introduction of transaction costs.

We shall take into account a possible income stream increasing the wealth. There are several articles on the issue. In Duffie and Zariphopoulou [19], which is also based on control theory, the income stream is absolutely continuous with respect to the Lebesgue measure. We call such an income risk-free even though the rate of income is allowed to be stochastic. In this chapter, we shall work with an additional diffusion part to the income stream and therefore speak of risky income. We shall also take into account a possible debt. Debt is modelled as a stochastic process and is subtracted from the value process consisting of income and gains on the financial market (gross wealth) in order to determine the surplus or (net) wealth of the agent. Our debt process contains a diffusion part and we shall speak of risky debt.

The introduction of risky income and risky debt gives us the opportunity to interpret all our results in terms of distribution and investment of surplus in life and pension insurance contracts. The link between investment-consumption problems and the problem of redistribution of surplus in life and pension insurance that is brought to the surface here, was also noted in Cairns [12]. We explain in Section 6.3 how an insurance contract or a portfolio of insurance contracts could be covered by our model. The construction of the insurance contract in that section is based on ideas and terminology fetched from Chapter 3. The main difference compared to Chapter 3 is that payments are driven by diffusion processes instead of jump processes.

Risky income and risky debt will be introduced as rather general stochastic processes. However, when we come to solving our control problems, we will specialize to a very simple structure. The relevance of the simple structure for practical problems can be discussed. However, we believe that solving the distribution/investment problem in a simple case is a first step in the understanding of the problem and the nature of its solution in a more involved case.

The preference ordering in Merton's problem is rather special. A generalized ordering of preferences is obtained by considering state-dependent utilities. However, generalizations in this direction, in general, increases the number of state variables and complicates the problem accordingly. In Chapter 5, a notion of risk-adjusted utility is introduced which simplifies the problem of optimal consumption and investment. The investment and the consumption problems are separated into two problems which can be solved independently of each other. We shall follow this idea and formulate the ordering of preferences of the agent in terms of risk-adjusted utility.

After separation of the investment problem and the consumption problem and solution of the investment problem, the consumption problem reduces to a traditional controlled diffusion problem. In our formulation with quadratic disutility of current wealth and quadratic or absolute disutility of consumption, this controlled diffusion problem has been studied intensively in the literature on stochastic control. Sections 6.8, 6.9, and 6.10 of this chapter deal with these problems, and we shall make extensive use of results from existing literature. In the case of absolute disutility of consumption, we work in the framework of singular control (Section 6.8). In the case of quadratic disutility of consumption, we work in the framework of the linear regulator problem known from just about every textbook on stochastic control, see e.g. Fleming and Soner [24] (Section 6.9). Finally, we consider a class of suboptimal control problems (Section 6.10).

The dynamic programming principle states, in non-mathematical words, that, in a certain sense, locally optimal behavior is as good as globally optimal behavior. Under certain regularity conditions, this principle leads to the dynamic programming equation, henceforth abbreviated the DPE, which is a certain kind of differential equation. The particular form of the DPE varies with the underlying problem. We expect the reader to be familiar with these concepts such that the DPE for each case can be put up directly. We work rather systematically with the consumption problems of various kinds. In some cases we find the explicit solutions, in some cases we are content with a description of a solution, and in some cases we do not even get that far. These sections can be seen as a concentrated collection of solutions to a certain class of control problems, some of which, to our knowledge, have not been studied before.

6.2 The general model

In this section, we model the wealth of an agent with risky income and risky debt. Let the stochastic basis be a standard two-dimensional Brownian motion $\left(\overline{W}, W\right)$ defined on a probability space $(\Omega, \mathcal{F}, \mathbf{F} = \{\mathcal{F}_t\}_{t\geq 0}, P)$. Let the market be a Black-Scholes market, where the dynamics of the price processes of the bond, $(B(t))_{t\geq 0}$, and the stock, $(S(t))_{t\geq 0}$, are given by

$$\begin{aligned} dB(t) &= B(t)\, r dt,\ B(0) = 1, \\ dS(t) &= S(t)\left(\mu dt + \beta d\overline{W}(t)\right),\ S(0) = s, \end{aligned}$$

respectively, where r, μ, and β are constants.

The total payment process is the sum of two semimartingales, A' and $-U'$, of which A' is a *gross income* process and U' is a *consumption* process. It is assumed that gross income A' is uncontrollable whereas consumption U' is controllable. We let A' be a continuous semimartingale with dynamics given by

$$dA'(t) = \alpha'(t)\, dt + \sigma'(t)\, dW(t) + \overline{\sigma}'(t)\, d\overline{W}(t)\,, \tag{6.1}$$

where α', σ', $\overline{\sigma}'$ are adapted processes. Lump sum income at deterministic points in time could be taken into account but is omitted for notational convenience. Note that we allow negative income.

Assuming that, at time t, the (controllable) relative portfolio $\theta'(t)$ is invested in asset S and the relative portfolio $(1-\theta'(t))$ is invested in asset B, the *gross wealth* process L' consisting of payments and gains on the financial market is given by the dynamics

$$\begin{aligned} dL'(t) &= dA'(t) - dU'(t) \\ &\quad + \frac{\theta'(t) L'(t)}{S(t)} dS(t) + \frac{(1-\theta'(t)) L'(t)}{B(t)} dB(t), \\ L'(0-) &= l_0, \end{aligned} \tag{6.2}$$

where l_0 is an initial gross wealth. Note that initial gross wealth is specified at time $0-$, since there may be a lump sum consumption at time 0.

The agent takes debt into consideration, and we assume that the *debt* process is given by a continuous semimartingale, V', which can be written on the form

$$\begin{aligned} dV'(t) &= rV'(t)\,dt + \delta'(t)\,dt + \psi'(t)\,dW(t) + \overline{\psi}'(t)\,d\overline{W}(t), \\ V'(0) &= v_0, \end{aligned} \tag{6.3}$$

where, δ', ψ', $\overline{\psi}'$ are adapted processes and v_0 is the initial debt. The term $rV'(t)\,dt$ could, of course, be included in $\delta' dt$, but we choose not to do this for later notational convenience.

The introduction of a debt process gives us the opportunity to interpret all our results in terms of distribution and investment of surplus in life and pension insurance. In the succeeding section, we provide the reader with the concepts needed for such an interpretation. That section is, at the same time, an example of a possible connection between the coefficients in the income process A' and the debt process V', respectively.

The *net wealth*, henceforth just called *wealth* or *surplus* of the agent is now defined by

$$X' = L' - V',$$

and by (6.2) and (6.3), it is given by the stochastic differential equation

$$\begin{aligned} dX'(t) &= (\alpha'(t) - \delta'(t))\,dt - dU'(t) \\ &\quad + (\sigma'(t) - \psi'(t))\,dW(t) + \left(\overline{\sigma}'(t) - \overline{\psi}'(t)\right) d\overline{W}(t) \\ &\quad + X'(t)\,r dt + L'(t)\,\theta'(t)\,(\mu - r)\,dt + L'(t)\,\theta'(t)\,\beta d\overline{W}(t), \quad (6.4) \\ X'(0-) &= x = l_0 - v_0. \end{aligned}$$

In case U' is continuous, we have that $X'(0) = x$. Since the objectives in our optimization problems below are formulated in terms of X' (not L'), we now, without loss of generality, let $l_0 = 0$.

We are going to work with risk-adjusted utility and need for that purpose an appropriate deflator, see Chapter 5. A deflator Λ is a process in the form

$$\begin{aligned} d\Lambda(t) &= -\Lambda(t)\left(rdt - h(t)\,d\overline{W}(t) - g(t)\,dW(t)\right), \\ \Lambda(0) &= 1. \end{aligned} \tag{6.5}$$

With conditions on (g, h), we can define Q by

$$\Lambda(t) = e^{-rt} E\left[\left. \frac{dQ}{dP} \right| \mathcal{F}_t \right],$$

such that Λ is the product of a discount factor and the likelihood function of a measure Q with respect to the measure P.

Abbreviating

$$\begin{aligned} X(t) &= \Lambda(t) X'(t), \\ dU(t) &= \Lambda(t) dU'(t), \\ \sigma(t) &= \Lambda(t) (\sigma'(t) - \psi'(t)), \\ \overline{\sigma}(t) &= \Lambda(t) \left(\overline{\sigma}'(t) - \overline{\psi}'(t) \right), \\ \alpha(t) &= \Lambda(t) (\alpha'(t) - \delta'(t)) + h(t) \overline{\sigma}(t) + g(t) \sigma(t), \\ \theta(t) &= \Lambda(t) L'(t) \theta'(t), \end{aligned} \tag{6.6}$$

Ito's lemma, (6.4), and (6.5) give the dynamics of the *deflated wealth process*, $X = \Lambda X'$,

$$\begin{aligned} dX(t) &= \Lambda(t) dX'(t) + X'(t) d\Lambda(t) + d\Lambda(t) dX'(t) \\ &= \alpha(t) dt - dU(t) + (\sigma(t) + g(t) X(t)) dW(t) \\ &\quad + (\theta(t) \beta + h(t) X(t) + \overline{\sigma}(t)) d\overline{W}(t), \\ X(0-) &= x. \end{aligned} \tag{6.7}$$

In Chapter 5, it is argued that the relevant deflator to use in connection with risk-adjusted utility is the *minimal deflator* corresponding to choosing as measure Q the *minimal martingale measure* (see Schweizer [61]). Following this idea we get that

$$h(t) = \frac{r - \mu}{\beta}, \tag{6.8}$$

$$g(t) = 0, \tag{6.9}$$

such that (6.7) becomes

$$\begin{aligned} dX(t) &= \alpha(t) dt - dU(t) + \sigma(t) dW(t) \\ &\quad + \left(\theta(t) \beta + \frac{r - \mu}{\beta} X(t) + \overline{\sigma}(t) \right) d\overline{W}(t). \end{aligned} \tag{6.10}$$

6.3 A diffusion life and pension insurance contract

In this section we show how a general life and pension insurance contract with payments of diffusion type is covered by the model described above in the sense that *an agent* now should be understood as *an insurance company having issued an insurance contract (portfolio).* It is, at the same time, an example of a possible connection between the coefficients in the income process A' and the debt process V', respectively. The section reconstructs the general life and pension insurance contract along the lines of Section 3.3 in Chapter 3 with payments of diffusion type, and the reader will recognize the structure of payments from there.

We introduce a *first order deflator* $\widehat{\Lambda}$ given by

$$\begin{aligned} d\widehat{\Lambda}(t) &= -\widehat{\Lambda}(t)\left(\widehat{r}(t)\,dt - \widehat{h}(t)\,d\overline{W}(t) - \widehat{g}(t)\,dW(t)\right), \\ \widehat{\Lambda}(0) &= 1, \end{aligned}$$

where $(\widehat{r}, \widehat{g}, \widehat{h})$ is a function of $\left(t, W(t), \overline{W}(t)\right)$. Define a stream of *first order payments* A' by (6.1) where $(\alpha'(t), \sigma'(t), \overline{\sigma}'(t))$ is now a function of $\left(t, W(t), \overline{W}(t)\right)$. The *first order reserve* $\widehat{V}'$ is defined by

$$\begin{aligned} \widehat{V}'(t) &= E\left[\int_t^T \frac{\widehat{\Lambda}(s)}{\widehat{\Lambda}(t)}\left(d(-A'(s)) - \left(\sigma'(s)\,\widehat{g}(s) + \overline{\sigma}'(s)\,\widehat{h}(s)\right)ds\right)\middle|\,\mathcal{F}_t\right] \\ &= E^{\widehat{Q}}\left[\int_t^T e^{-\int_t^s \widehat{r}(\tau)dt}\,d(-A'(s))\middle|\,\mathcal{F}_t\right], \end{aligned}$$

Let A' be constrained by the first order equivalence relation

$$\widehat{V}'(0-) = 0.$$

The first order deflator serves solely as a tool for laying down the first order payments at time 0. During the term of the contract we work with a *technical deflator* Λ^*, defined by

$$\begin{aligned} d\Lambda^*(t) &= -\Lambda^*(t)\left(r^*(t)\,dt - h^*(t)\,d\overline{W}(t) - g^*(t)\,dW(t)\right), \\ \Lambda^*(0) &= 1, \end{aligned}$$

where (r^*, g^*, h^*) is a function of $\left(t, W(t), \overline{W}(t)\right)$. The *technical reserve* V' is defined by

$$\begin{aligned} V'(t) &= E\left[\int_t^T \frac{\Lambda^*(s)}{\Lambda^*(t)}\left(d(-A'(s)) - (\sigma'(s)\,g^*(s) + \overline{\sigma}'(s)\,h^*(s))\,ds\right)\middle|\,\mathcal{F}_t\right] \\ &= E^{Q^*}\left[\int_t^T e^{-\int_t^s r^*(\tau)dt}\,d(-A'(s))\middle|\,\mathcal{F}_t\right] \end{aligned}$$

As opposed to $\widehat{\Lambda}$ and Λ^* we shall speak of Λ as the *real deflator.* The real payments are composed by the first order payments A' and an additional payment stream of bounded variation $-U'$, called the *dividends.* The total payment process $A' - U'$ is required to fulfill the real equivalence relation

$$\begin{aligned} 0 &= E\left[\int_{0-}^T \Lambda(t)\left(dA'(t) - dU'(t) + (\sigma'(t)\,g(t) + \overline{\sigma}'(t)\,h(t))\,dt\right)\right] \qquad (6.11) \\ &= E^Q\left[\int_{0-}^T e^{-\int_t^s r(\tau)dt}\left(dA'(t) - dU'(t)\right)\right] \end{aligned}$$

With Chapter 3 on hand, it is now a (more or less) trivial exercise in Ito Calculus to derive the dynamics of various versions of the surplus, hereunder the *retrospective surplus* defined by $X' = L' - V'$. A study includes a statement of a diffusion version of Thiele's differential equation. In this exposition, however, we shall content ourselves with identifying, without proof, the coefficients in the technical reserve (see (6.3)),

$$\begin{aligned} \psi'(t) &= V'_w(t), \\ \overline{\psi}'(t) &= V'_{\overline{w}}(t), \\ \delta'(t) &= (r^*(t) - r) V'(t) + \alpha'(t) \\ &\quad + (\sigma'(t) - \psi'(t)) g^*(t) + \left(\overline{\sigma}'(t) - \overline{\psi}'(t)\right) h^*(t), \end{aligned}$$

such that

$$\begin{aligned} \frac{\alpha}{\Lambda(t)} &= (r - r^*(t)) V'(t) + (h(t) - h^*(t)) (\overline{\sigma}'(t) - V'_{\overline{w}}(t)) \\ &\quad + (g(t) - g^*(t)) (\sigma'(t) - V'_w(t)), \\ \frac{\sigma}{\Lambda(t)} &= \sigma'(t) - V'_w(t), \\ \frac{\overline{\sigma}}{\Lambda(t)} &= \overline{\sigma}'(t) - V'_{\overline{w}}(t). \end{aligned}$$

The quantities $\frac{\alpha}{\Lambda(t)}$, $\frac{\sigma}{\Lambda(t)}$, and $\frac{\overline{\sigma}}{\Lambda(t)}$ are the coefficients of the contributions to the surplus and can be compared to the coefficients in the process C^* in Section 3.3. The reader should note the systematic contribution to the surplus, $\frac{\alpha}{\Lambda(t)}$, which appears a couple of times in our optimal dividend processes below.

We shall now assume that the agent is risk-neutral with respect to W-risk, and one can think of W as being some kind of diversifiable risk e.g. stemming from mortality risk in the insurance portfolio. Comparing to (3.9), it can then be shown that the equivalence relation (6.11) also can be written as the relation

$$E[X(T)] = 0,$$

which shall play the important role as terminal condition in our (finite time) optimization problems.

For the sake of generality we will, throughout the chapter, use the general terminology introduced in Section 6.2. However, the reader should have in mind the application to a life and pension insurance contract (portfolio). We end this section with a small dictionary relating terms from our general set-up to the terms from the general life and pension insurance interpretation. We write in brackets the corresponding symbols used in Chapter 3.

Symbol	General set-up	General life and pension insurance
A'	gross income	first order payments $\left(\widehat{B}\right)$
V'	debt	technical reserve (V^*)
U'	consumption	dividends $\left(-\widetilde{B}\right)$
X'	wealth	(retrospective) surplus $\left(\overleftarrow{F}^*\right)$

6.4 Objectives

Given the model described in Section 6.2, we now formulate our control problem. The control variables are the relative portfolio invested in the risky asset θ, and the consumption U. The process to be controlled is the deflated wealth process with dynamics given by (6.10). We introduce cost functions υ^u and υ^x and introduce, for each control (θ, U), a cost functional. In this chapter, we are going to work with two different cost functionals:

Finite time consumption:

$$J(t, x, \theta, U) = E\left[\int_t^T \upsilon^x(X(s))\,ds + a\upsilon^u(dU(s), ds)\middle| X(t) = x\right]. \tag{6.12}$$

Stationary consumption:

$$J(x, \theta, U) = \lim_{T\to\infty} \frac{1}{T} E\left[\int_0^T \upsilon^x(X(s))\,ds + a\upsilon^u(dU(s), ds)\middle| X(0) = x\right]. \tag{6.13}$$

Note that the compact notation for current utility of consumption only makes sense for particular functions υ^u. The cost functions introduced below serve as examples. We use the terms utility and disutility functions for general (not necessarily increasing, concave and continuously differentiable) reward and cost functions, respectively.

The objective is to find the corresponding optimal cost functions

$$V(t, x) = \min_{\theta, U} J(t, x, \theta, U) \text{ and } \gamma = \min_{\theta, U} J(x, \theta, U),$$

respectively, and to find the optimal control (θ^*, U^*) such that

$$V(t, x) = J(t, x, \theta^*, U^*) \text{ and } \gamma = J(x, \theta^*, U^*),$$

respectively. Note that the letter V is used for the optimal cost function or value function, not to be confused with the deflated debt process $\Lambda V'$. For derivatives of V we shall use the notation $V_t = \frac{\partial}{\partial t}V(t, x)$, $V_x = \frac{\partial}{\partial x}V(t, x)$, and $V_{xx} = \frac{\partial^2}{\partial x^2}V(t, x)$.

These cost functions show that we are interested in controlling the deflated wealth X. We have specified an additional structure of the cost of wealth and the cost of consumption. In order to be able to vary weight on cost of wealth against weight on cost of consumption, we have introduced the weight factor a. In the case of finite time consumption one could let a be time-dependent, but for simplicity we let a be a constant which we take to be positive.

Usually, when working with optimization in investment-consumption problems, one wishes to maximize J to reach a maximal total utility. Instead, we minimize J to reach a minimal total disutility. Thus, we need to specify disutility functions υ^x and υ^u. We are going to work with disutility of distance to a target. The underlying idea is utility of stability, where stability is measured in relation to a target process. Deviations from the target are punished by a cost of distance. In the case of υ^x, we shall measure distance by the quadratic function. In the case of υ^u, we shall measure distance both by the absolute function and by the quadratic function. Sections 6.4.1 and 6.4.2 below substantiate this.

The quadratic cost function is well-known in both finance and insurance. In finance it is the basis for quadratic hedging approaches to incomplete markets (see Schweizer [61] for a survey). In insurance one has, since decades, been interested in controlling pension funds in order to obtain some kind of stability, see Cairns [12] and references therein. Important references in relation to our work are O'Brien [55] and O'Brien [56], where the stability criterion and minimization of quadratic cost are introduced for the first time in connection with pension funding in a continuous-time framework similar to ours.

A survey of results in continuous-time quadratic control of pension funds is given in Cairns [12] which in many respects relates to our work. However, on important points our cost function differs from the cost function used in Cairns [12] and most other approaches to quadratic control. Firstly, by introducing risk-adjusted utility, we work with cost of X instead of cost of X'. As it will turn out, this has a most appealing effect on the optimal investment policy, which in connection with quadratic utility usually leads to counter-intuitive investment strategies. We obtain intuitively reasonable investment strategies. Secondly, we wish to control the surplus itself instead of controlling the so-called funding level, $\frac{X'}{V'}$, which measures the surplus relative to debt.

6.4.1 Cost of wealth

In this section we consider disutility of current deflated wealth. We assume that the agent's preferences over wealth are given by the quadratic disutility function,

$$\upsilon^x \left(X \left(t \right) \right) = \left(X \left(t \right) - \widehat{x} \left(t \right) \right)^2 ,$$

where, in general, $\widehat{x}(t)$ is an adapted process representing a target for our deflated wealth at time t. Define

$$\widehat{x}' \left(t \right) = \frac{\widehat{x} \left(t \right)}{\Lambda \left(t \right)}.$$

6.4.2 Cost of consumption

In Section 6.4.1, we introduced the quadratic disutility of deflated wealth and we shall now discuss disutility of deflated consumption. This section introduces two cost functions leading to singular control and classical control, respectively, where "classical" refers to a consumption plan which is absolute continuous with respect to the Lebesgue measure.

In singular consumption problems, we assume that the agent's preferences over consumption are given by the absolute disutility function

$$\upsilon^u (dU(t), dt) = |dU(t) - \widehat{u}(t)\, dt|, \tag{6.14}$$

where, in general, $\widehat{u}(t)$ is an adapted process. We call problems of this kind singular consumption problems simply because the resulting optimal consumption contains a singular part. Define

$$\widehat{u}'(t) = \frac{\widehat{u}(t)}{\Lambda(t)}.$$

Denoting by $\mathcal{U}^{bv}$ the set of processes of bounded variation, we require that $U \in \mathcal{U}^{bv}$. A process $U \in \mathcal{U}^{bv}$ can be decomposed uniquely in a pure jump part, U^j, and a continuous part, U^c, which further can be decomposed uniquely into an absolutely continuous part, U^{ac}, and a singular continuous part, U^{sc}. By defining $\Delta U(t) \equiv U(t) - U(t-)$ and, existing almost everywhere, $u(t) \equiv \frac{dU^c(t)}{dt}$, we have

$$\begin{aligned} U^j(t) &\equiv \sum_{0 \le s \le t} \Delta U(s), \\ U^c(t) &\equiv U(t) - U^j(t), \\ U^{ac}(t) &\equiv \int_0^t u(s)\, ds, \\ U^{sc}(t) &\equiv U^c(t) - U^{ac}(t). \end{aligned}$$

In control problems of singular type the state space splits into two regions, a "push region" and a "low action region". If the state process starts from the "push region", the optimal state process moves immediately ($dU^j(0) \neq 0$) into the "low action" region, where it is controlled by absolutely continuous control ($dU^{ac}(t) \neq 0$). Exit from the "low action" region is prevented by reflection ($dU^{sc}(t) \neq 0$) at the boundary in an appropriate direction.

In the literature, authors dealing with singular control usually formulate the problem such that the optimal control in the "low action" is to do nothing, i.e. $u(t) = 0$, and they speak instead of a "no action" region. We call such a problem a pure singular problem. In order to apply directly results from the literature, we transform our singular problems into pure singular problems. Assume that we start out by a process with dynamics given by

$$\begin{aligned} dX(t) &= \widetilde{\alpha}(t) - d\widetilde{U}(t) + \sigma(t)\, dW(t) \\ &\quad + \left(\theta(t)\beta - \frac{\mu - r}{\beta} X(t) + \overline{\sigma}(t)\right) d\overline{W}(t), \end{aligned}$$

and a disutility function

$$\upsilon^u \left(d\widetilde{U}(t), dt\right) = \left|d\widetilde{U}(t) - \widehat{u}(t)\, dt\right|.$$

Now, introducing controls

$$U(t) = \widetilde{U}(t) - \int_0^t \widehat{u}(s)\, ds,$$

and the process

$$\alpha(t) = \widetilde{\alpha}(t) - \widehat{u}(t),$$

we can write the dynamics of X as

$$\begin{aligned} dX(t) &= \alpha(t) - dU(t) + \sigma(t)\, dW(t) \\ &+ \left(\theta(t)\beta - \frac{\mu - r}{\beta} X(t) + \overline{\sigma}(t) \right) d\overline{W}(t), \end{aligned}$$

and the disutility function as

$$\upsilon^u(dU(t), dt) = |dU(t)|.$$

This is a pure singular problem with a "push" region and a "no-action" region. Given an optimal solution to this pure singular problem, we return to an optimal solution to our original problem by the relation $\widetilde{U}(t) = U(t) + \int_0^t \widehat{u}(s)\, ds$. This means that the "no action" region is replaced by a "low action" region in which the absolutely continuous control $\widehat{u}(t)\, dt$ is performed.

In classical consumption problems, we assume that the agent's preferences over consumption are given by the quadratic disutility function

$$\upsilon^u(dU(t), dt) = \left(\frac{dU(t)}{dt} - \widehat{u}(t) \right)^2 dt. \tag{6.15}$$

Here, we require that $U \in \mathcal{U}^{ac}$. We call these problems classical consumption problems simply because the resulting optimal consumption, of course, turns out to be on classical (absolutely continuous) form.

6.5 Constraints

In addition to objectives, one needs to specify possible constraints under which an optimal control is to be chosen. One can work with constraints on both the state process X and the control (θ, U).

The process X may be constrained by upper or lower barriers set by the agent or the legislative environment in which the agent makes his decisions. We shall not consider constraints of this type although the optimally controlled process, in the case of singular control, actually will be bounded by such barriers. However, in the finite time consumption problem we shall work with a terminal constraint on the process X in the form

$$E[X(T)] = 0. \tag{6.16}$$

In Section 6.5.1 we motivate our constraint and explain how it is dealt with in our optimization problem.

With respect to the investment strategy θ, a typical constraint is restriction on short-selling of risky assets. For references on investment restrictions in finance, see

Korn [40]. For investment restrictions in insurance, see Cairns [12], where restriction on short-selling of risky assets is an important issue. We shall not work with constraints of this type.

With respect to U, constraints shall play a very important role. All problems will be approached both without constraints on U and with U constrained to be non-decreasing. In Boulier et al. [9] a constraint is introduced which can be compared to our constraint on U being non-decreasing. However, as it will turn out, such a constraint will typically complicate the problem considerably. The motivation for so much emphasis on this particular constraint is to be found in Chapter 3. In terms of life and pension insurance, and using the terminology introduced in Chapter 3, the unconstrained problem corresponds to the problem of redistribution of surplus in pension funding whereas the constrained problem corresponds to the problem of redistribution of surplus in participating life insurance.

This terminology refers to the fact that redistribution in pension funding is basically assumed not to be constrained. In the literature on the subject, only Boulier et al. [9] goes beyond this assumption and introduces a constraint somewhat similar to ours. As opposed to pension funding, a crucial characterization of participating life insurance is that redistribution is required to be positive. As in Chapter 3, we are interested in both pension funding and participating life situation. We shall, therefore, attempt to find optimal controls in both cases.

If consumption is constrained to be positive, we endow $\mathcal{U}^{bv}$ and $\mathcal{U}^{ac}$ with a subscript '+', i.e. $\mathcal{U}^{bv}_{+}$ and $\mathcal{U}^{ac}_{+}$.

In addition to optimal control, we shall consider suboptimal consumption where consumption is required to be in a certain parametric form. Suboptimal consumption can be considered as optimal consumption in the constrained problem, where the constraint is exactly the particular parametric form of the consumption. We denote by $\mathcal{U}^{l}$ the class of linear controls in the form

$$u(t) = w + vX(t),$$

where w and v are constants. We shall also consider problems where the rate of consumption in addition to the certain parametric form is constrained to be non-negative. We denote by $\mathcal{U}^{l}_{+}$ the class of piecewise linear non-negative controls that can be written in the form

$$u(t) = (w + v(X(t) - x_0))\,1_{(X(t)>x_0)},$$

where w, v, and x_0 are constant and

$$w \geq 0.$$

6.5.1 A terminal constraint

In the finite time problems where the agent optimizes over a finite fixed period of time, we can specify a disutility of terminal wealth. This shall play a special role in our problem. Whereas $\widehat{x}(t)$ plays the role as current target for $X(t)$, we shall introduce 0 as a target for $X(T)$, in the sense that our optimization problem is formulated with a terminal constraint,

$$E[X(T)] = 0. \tag{6.17}$$

The motivation for this constraint is to be found in Section 6.3 where (6.17) is obtained as a no arbitrage relation for the payment process $A' - U'$.

Remark 8 *For readers familiar with optimization problems in finance, the constraint (6.17) is recognized from the martingale approach to optimal consumption. In a complete market with positive consumption, (5.16) replaces the admissibility condition, $X(t) \geq 0$, $0 \leq t \leq T$. Here, however, the constraint (6.17) is motivated by a no arbitrage condition on a life and pension insurance portfolio in an incomplete framework where there are no requirements, in general, concerning positive wealth and positive consumption.*

The terminally constrained problem is solved by application of a Lagrange multiplier. The idea is to introduce a related, but unconstrained, problem with disutility of terminal deflated wealth,

$$\Upsilon(X(T)) = \lambda X(T).$$

If there exists a λ_0 and a solution to the related optimization problem such that

$$E[X(T)] = 0,$$

then this solution will be the solution to the terminally constrained problem. For application of a Lagrange multiplier to terminally constrained problems in control theory, see Øksendal [57].

In the stationary problem, it makes no sense to work with a terminal condition. Actually, it is easy to realize that for quadratic disutility of current wealth, $v^x(X(t)) = (X(t) - \widehat{x})^2$, where $\widehat{x}$ is now constant, we have that if X is optimally controlled, then

$$E[X(T)] \rightarrow \widehat{x} \text{ for } T \rightarrow \infty,$$

independently of the form of utility of consumption.

This ends our introduction and discussion of the type of problems the solutions of which we are interested in. We have introduced the coefficients in the payment process α', σ' and $\overline{\sigma}'$ as general adapted processes leading to general adapted processes α, σ, and $\overline{\sigma}$. Also the targets $\widehat{x}$ and $\widehat{u}$ have been introduced as general adapted processes. In the rest of the chapter, we are going to look for solutions in particular cases. And, in particular, concerning income coefficients and targets we shall restrict ourselves to a very simple structure. We shall assume that α, σ, $\widehat{u}$, and $\widehat{x}$ are constant and that $\overline{\sigma}$ is a function of $(t, X(t))$. This is evidently a very special structure and its relevance for practical problems can be discussed. However, we believe that solving the investment-consumption problem in a simple case is a first step in the understanding of the problem and the nature of its solution in a more involved case.

The restrictive structure assumed above means, broadly speaking, that coefficients of nominal income and nominal targets grow with the market. If the market is high, then Λ is low and, consequently, coefficients and targets are high. In the life insurance portfolio this means that contributions to the surplus and the targets for surplus and redistribution are high when the market is high.

6.6 The dynamic programming equations

In this section we briefly explain the structure of the dynamic programming equation, the DPE, appearing from the various control problems described in the previous sections. We also comment on the DPE as a constructive tool for solution of the control problems and on the verification. Finally, we prepare the reader to a number of remarks in Section 6.8.

The DPE is a system of differential equations (inequalities). In the finite time problems (6.12), the system concerns the value function V, and in the stationary problem (6.13), the system concerns the derivative of a so-called potential function which we also denote by V. An important difference between the DPEs appearing in the finite time problems and the stationary problems, respectively, is the time-dependence. In the finite time problem, V is time-dependent, and the DPE is a system of partial differential equations (inequalities), whereas in the stationary problem, V is not time-dependent, and the DPE is a system of ordinary differential equations (inequalities).

Another important characteristic of the DPEs comes from our distinction between singular problems (6.14) and classical problems (6.15). In the singular problem, the DPE is a system of (partial) differential inequalities, or a so-called variational inequality, which also can be formulated as a free boundary value problem. In the classical problem, the DPE is one (partial) differential equation, the so-called Hamilton-Jacobi-Bellman equation.

The DPE plays two different roles. One has the result that **if** the function V is sufficiently regular, **then** $V(t,x)$ follows the DPE. This result is difficult to use because it is difficult to establish the sufficient differentiability. The so-called verification theorem is often more useful. This result states that **if** a function $\widetilde{V}(t,x)$ follows the DPE, meets some special conditions, and accords with an admissible control, **then** $\widetilde{V}(t,x) = V(t,x)$ and the according control is optimal. In the finite time problems the special conditions boil down to a side condition at time T. In the stationary problems the special conditions boil down to a condition on polynomial growth in x.

In our work we shall use the verification conditions as constructive tools in our search for candidates for the function $\widetilde{V}$ and an according control. Thus, these conditions will be fulfilled by construction, and if we succeed in finding a candidate this candidate will consequently coincide with V. Therefore there will be no explicit verification theorems in Chapters 6.8 and 6.9. In some problems we do not obtain an explicit candidate, and in these cases we take one of three resorts: Either we refer to results in the literature; or we search for a solution by numerical methods; or we give up. However, in all cases we expose our search for a candidate as far as we get.

There are deep relations between problems of singular control and problems of optimal stopping, and several authors work on linking singular control problems to optimal stopping problems in order to apply e.g. existence theorems established in one field to the other, see Karatzas and Shreve [36], [37], and Boetius and Kohlmann [8]. The main idea is to verify that the derivative with respect to x of the value function of the singular control problem is nothing but the value function of the stopping problem. There are basically two ways of realizing this connection. The analytic way, where the DPEs of each problem are derived and related to each other, and the probabilistic way, where the value functions of the two problems are directly related. The former method assumes regularities in order to conclude relations between underlying control problems from relations between the DPEs. In Section 6.8 on singular control problems, we shall for each singular problem, in a remark, specify the corresponding optimal stopping problem, which in some cases is a generalized optimal stopping problem, namely an optimal stopping game. We leave it to readers familiar with optimal stopping time problems and optimal stopping games to realize these connections analytically.

6.7 Optimal investment

In this section we show that the decisions concerning investment and the decisions concerning consumptions can be separated in the sense that the optimal investment policy and the optimal consumption process can be found independently of each other. We find the optimal investment policy.

For the singular finite time problem, the DPE in the "low/no action" region is

$$\begin{aligned} -V_t &= \min_{\theta}[(\alpha-\widehat{u})\,V_x+v^x(x) \\ &\quad +\frac{1}{2}\left(\sigma^2+\left(\theta\beta+\frac{r-\mu}{\beta}x+\overline{\sigma}(t,x)\right)^2\right)V_{xx}]. \end{aligned} \tag{6.18}$$

For the classical finite time problem, the DPE is

$$\begin{aligned} -V_t &= \min_{u,\theta}[(\alpha-u)\,V_x+v^x(x)+av^u(u) \\ &\quad +\frac{1}{2}\left(\sigma^2+\left(\theta\beta+\frac{r-\mu}{\beta}x+\overline{\sigma}(t,x)\right)^2\right)V_{xx}]. \end{aligned} \tag{6.19}$$

In the corresponding stationary problems, $-V_t$ in the equations above is replaced by a constant γ.

Differentiation of (6.18) and (6.19) with respect to θ and equating the right hand side of (6.19) to zero results in an optimizer which for both the singular and the classical problem becomes

$$\theta(t,x) = -\frac{\overline{\sigma}(t,x)}{\beta} + \frac{\mu - r}{\beta^2} x. \tag{6.20}$$

Differentiation once more shows that θ is minimizing if V is convex.

Proposition 9 *If υ^x, υ^u, and Υ are convex, then V is convex.*

Sketch of proof. Consider two initial points x_1 and x_2 and admissible strategies (U_1, θ_1) and (U_2, θ_2) and let $\lambda \in (0,1)$. Defining the strategy

$$(U,\theta) \equiv (\lambda U_1 + (1-\lambda) U_2, \lambda\theta_1 + (1-\lambda)\theta_2),$$

for the initial point $x = \lambda x_1 + (1-\lambda) x_2$, we get by linearity of X, $X(t) = \lambda X_1(t) + (1-\lambda) X_2(t)$. Since υ^x, υ^u, and Υ are convex, we know that

$$\begin{aligned} \upsilon^x(X(t)) &\leq \lambda \upsilon^x(X_1(t)) + (1-\lambda)\upsilon^x(X_2(t)), \\ \upsilon^u(dU(t)) &\leq \lambda \upsilon^u(dU_1(t)) + (1-\lambda)\upsilon^u(dU_2(t)), \\ \Upsilon(X(T)) &\leq \lambda \Upsilon(X_1(T)) + (1-\lambda)\Upsilon(X_2(T)). \end{aligned}$$

Now, it follows that

$$J(t,x,\theta,U) \leq \lambda J(t,x_1,\theta_1,U_1) + (1-\lambda) J(t,x_2,\theta_2,U_2).$$

We can choose (U_1,θ_1) such that $J(t,x_1,\theta_1,U_1) \leq V(t,x_1) + \varepsilon$ and (U_2,θ_2) such that $J(t,x_2,\theta_2,U_2) \leq V(t,x_2) + \varepsilon$. Since (U,θ) is suboptimal we have

$$V(t,x) \leq \lambda V(t,x_1) + (1-\lambda) V(t,x_1) + \varepsilon,$$

and since ε was arbitrary, convexity follows.□

Since the square, absolute, and linear functions are convex, the reasoning leading to (6.20) and Proposition 9 gives us directly

Corollary 10 *For $\upsilon^x(X(t)) = (X(t) - \widehat{x})^2$, $\upsilon^u(dU(t), dt) = \left(\frac{dU(t)}{dt} - \widehat{u}(t)\right)^2 dt$ or $\upsilon^u(dU(t), dt) = |dU - \widehat{u}dt|$, and $\Upsilon(X(T)) = \lambda X(T)$, the optimal investment policy is given by*

$$\theta(t) = -\frac{\overline{\sigma}(t,X(t))}{\beta} + \frac{\mu - r}{\beta^2} X(t), \tag{6.21}$$

or, equivalently,

$$\theta'(t) = -\frac{\overline{\sigma}'(t,X(t)) - \overline{\psi}'(t,X(t))}{\beta L'(t)} + \frac{\mu - r}{\beta^2}\frac{X'(t)}{L'(t)}.$$

Inserting the investment policy (6.21) in (6.18) and (6.19), we get

$$\begin{aligned} V_t + (\alpha - \widehat{u}) V_x + \frac{1}{2}\sigma^2 V_{xx} + v^x(x) &= 0, \\ \min_u \left[V_t + (\alpha - u) V_x + \frac{1}{2}\sigma^2 V_{xx} + v^x(x) + a v^u(u) \right] &= 0, \end{aligned}$$

respectively. These are recognized as the DPEs appearing if we start out with a state process with the dynamics

$$dX(t) = \alpha dt - dU(t) + \sigma dW(t). \tag{6.22}$$

This is not surprising since this is the state process for the optimally invested wealth process.

By (6.21) the solution of the optimal investment problem is given. The rest of the chapter deals with finding the optimal consumption for the deflated wealth process with dynamics given by (6.22). In every problem we search for optimal consumption U, and in order to return to optimal nominal consumption, we simply need to divide by Λ, $dU'(t) = \frac{dU(t)}{\Lambda(t)}$.

6.8 Optimal singular consumption

6.8.1 Finite time unconstrained consumption

The singular finite time unconstrained problem has been studied in the special case of $\alpha = 0$ by Karatzas and Shreve [37]. The singular unconstrained problem is there called the bounded variation follower problem. By symmetry, this problem coincides with the reflected follower problem where X is reflected in $\widehat{x}$. The reflected follower problem is the actual object of study in Karatzas and Shreve [37]. We shall here see how far we can get in a search for an explicit solution. However, at the end we are left with a system which must be approached by numerical methods.

The life and pension insurance interpretation of the problem studied in this section is that of finding optimal time-dependent upper and lower surplus barriers for a pension fund working with a finite time horizon.

We deal with the problem to minimize over $\mathcal{U}^{bv}$, subject to the terminal condition $E[X(T)] = 0$,

$$E\left[\int_0^T \left((X(t) - \widehat{x})^2 dt + a|dU(t)|\right)\right].$$

We introduce the related, but unconstrained, problem to minimize over $\mathcal{U}^{bv}$,

$$E\left[\int_0^T \left((X(t) - \widehat{x})^2 dt + a|dU(t)|\right) + \lambda X(T)\right].$$

We let

$$V(t,x) = \inf_{U \in \mathcal{U}^{bv}} E\left[\int_t^T \left((X(s) - \widehat{x})^2 dt + a|dU(s)|\right) + \lambda X(T) \middle| X(t) = x\right],$$

and we want to find $V(0,x)$ with λ determined such that $E[X(T)] = 0$. As described in Section 6.4.2, it is optimal to consume $dU(t) = \widehat{u}dt$, or $dU'(t) = \widehat{u}'(t)\,dt$, in a "low-action" region and to keep the state process within this region by singular consumption on its boundary. The task is to find the boundary.

The DPE connected with this problem is given by the variational inequality

$$\begin{aligned} -V_t &\leq \alpha V_x + \frac{1}{2}\sigma^2 V_{xx} + (x - \widehat{x})^2, \; V(T,x) = \lambda x, \\ V_x &\leq a, \; V_x \geq -a, \\ 0 &= \left(-V_t + \alpha V_x + \frac{1}{2}\sigma^2 V_{xx} + (x-\widehat{x})^2\right)(a - V_x)(a + V_x), \end{aligned}$$

or by the free-boundary problem

$$\begin{array}{cc} V_x = -a, & x \leq x_0(t), \\ -V_t = \alpha V_x + \frac{1}{2}\sigma^2 V_{xx} + (x-\widehat{x})^2, & x_0(t) \leq x \leq x_1(t), \\ V_x = a, & x_1(t) \leq x, \\ V(T,x) = \lambda x. & \end{array}$$

By subtracting a particular polynomial solution, one finds that

$$V(t,x) = \begin{cases} -ax + S_0(t), & x \leq x_0(t), \\ P(t)x^2 + Q(t)x + R(t) + U(t,x), & x_0(t) \leq x \leq x_1(t), \\ ax + S_1(t), & x_1(t) \leq x, \end{cases}$$

with $(P(t), Q(t), R(t))$ given by

$$\begin{aligned} P(t) &= T - t, \\ Q(t) &= \alpha(T-t)^2 - 2\widehat{x}(T-t) + \lambda, \\ R(t) &= \int_t^T \left(\sigma^2 P(s) + \alpha Q(s)\right) ds + \widehat{x}(T-t), \end{aligned} \tag{6.23}$$

and U solving the PDE,

$$-U_t = U_x \alpha + \frac{1}{2}U_{xx}\sigma^2, \; U(T,x) = 0. \tag{6.24}$$

The principle of smooth fit is now supposed to give boundary conditions for a unique solution U to this free boundary heat equation which then also determines (S_0, S_1).

Remark 11 *The differential stopping game problem corresponding to the problem studied in this section is that of finding*

$$\sup_{\rho} \inf_{\tau} E_{0,x}\left[\int_0^{\rho\wedge\tau\wedge T} 2(X(t) - \widehat{x})\,dt + a1_{(\tau<\rho\wedge T)} - a1_{(\rho<\tau\wedge T)} + \lambda 1_{(T<\tau\wedge\rho)}\right].$$

In the case $\alpha = 0$, the differential stopping game reduces to a stopping problem with absorption of finding

$$\min_{\tau} E_{0,x}\left[\int_0^{T\wedge\eta\wedge\tau} 2(X(t)-\widehat{x})\,dt + a1_{(\tau<T\wedge\eta)} + \lambda 1_{(T<\tau\vee\eta)}\right],$$

where

$$\eta = \inf\{t \geq 0 : X(t) = \widehat{x}\},$$

the first time X hits $\widehat{x}$.

6.8.2 Finite time constrained consumption

The singular finite time constrained problem has been studied in the special case of $\alpha = 0$ by Karatzas and Shreve [36]. The singular constrained problem is there called the monotone follower problem. We shall here see how far we can get in a search for an explicit solution. However, at the end we are left with a system which must be approached by numerical methods. The special case with $\alpha = 0$, $a = 0$, and no terminal condition was solved explicitly in Beneš et al. [5]. Since this special case is of minor importance to us, we shall not repeat this solution here but just give the reference as a rare example of an explicitly solvable singular problem.

The life and pension insurance interpretation of the problem studied in this section is that of finding an optimal time-dependent upper surplus barrier for a participating life insurance company working with a finite time horizon.

We deal with the problem to minimize over $\mathcal{U}_+^{bv}$, subject to the terminal condition $E[X(T)] = 0$,

$$E\left[\int_0^T \left((X(t) - \widehat{x})^2\, dt + a\,|dU(t)|\right)\right].$$

We introduce the related, but unconstrained, problem to minimize over $\mathcal{U}_+^{bv}$,

$$E\left[\int_0^T \left((X(t) - \widehat{x})^2\, dt + a\,|dU(t)|\right) + \lambda X(T)\right].$$

We let

$$V(t,x) = \inf_{U \in \mathcal{U}_+^{bv}} E\left[\int_t^T \left((X(s) - \widehat{x})^2\, dt + a\,|dU(s)|\right) + \lambda X(T) \middle|\, X(t) = x\right],$$

and we want to find $V(0,x)$ with λ determined such that $E[X(T)] = 0$. As described in Section 6.4.2, it is optimal to consume $dU(t) = \widehat{u}dt$, or $dU'(t) = \widehat{u}'(t)\,dt$, in a "low-action" region and to keep the state process within this region by singular consumption on its boundary. The task is to find the boundary.

The DPE connected with this problem is given by the variational inequality

$$\begin{aligned}
-V_t &\leq \alpha V_x + \frac{1}{2}\sigma^2 V_{xx} + (x - \widehat{x})^2,\ V(T,x) = \lambda x, \\
V_x &\leq a, \\
0 &= \left(-V_t + \alpha V_x + \frac{1}{2}\sigma^2 V_{xx} + (x - \widehat{x})^2\right)(a - V_x),
\end{aligned}$$

or by the free-boundary problem

$$\begin{aligned}
-V_t = \alpha V_x + \tfrac{1}{2}\sigma^2 V_{xx} + (x - \widehat{x})^2, &\quad x \leq x_1(t), \\
V_x = a, &\quad x_1(t) \leq x, \\
V(T,x) = \lambda x. &
\end{aligned}$$

By subtracting a particular polynomial solution, one finds that

$$V(t,x) = \begin{cases} P(t)\,x^2 + Q(t)\,x + R(t) + U(t,x), & x \leq x_1(t), \\ ax + S(t), & x_1(t) \leq x, \end{cases}$$

with $(P(t), Q(t), R(t))$ given by (6.23) and U solving (6.24). The principle of smooth fit and the polynomial growth condition are now supposed to give boundary conditions for a unique solution U to this free boundary heat equation which then also determines S.

Remark 12 *The stopping time problem corresponding to the problem studied in this section is that of finding*

$$\min_{\tau} E_{0,x} \int_0^{\tau \wedge T} 2\left(X(t) - \widehat{x}\right) dt + a 1_{(\tau < T)} + \lambda 1_{(T < \tau)}.$$

6.8.3 Stationary unconstrained consumption

The singular stationary unconstrained problem will be solved almost explicitly below. We are left with a numerical solution to a non-linear system of two equations with two unknowns. Fleming and Soner illustrate in [24] singular stochastic control by explicitly solving a similar infinite time problem, but we have not found the explicit solution for the stationary problem in the literature. However, one may argue that this is no more than a nice exercise in smooth fit.

In the case $\alpha = 0$, by symmetry, the control problem is equivalent to the same optimization problem for X reflected in $\widehat{x}$. This follows from the reasoning by Karatzas and Shreve [37] in the finite time case.

The life and pension insurance interpretation of the problem studied in this section is that of finding optimal upper and lower surplus barriers in a pension fund working with an infinite time horizon (the stationary view).

We deal with the problem to minimize over $\mathcal{U}^{bv}$,

$$\lim_{T \to \infty} E\left[\frac{1}{T} \int_0^T \left((X(t) - \widehat{x})^2 dt + a \left| dU(t) \right| \right)\right].$$

We let

$$\gamma = \min_{U \in \mathcal{U}^{bv}} \lim_{T \to \infty} E\left[\frac{1}{T} \int_0^T \left((X(t) - \widehat{x})^2 dt + a \left| dU(t) \right| \right)\right],$$

denote by $\int V$ the potential function, and want to find γ and V. As described in Section 6.4.2, it is optimal to consume $dU(t) = \widehat{u} dt$, or $dU'(t) = \widehat{u}(t) dt$, in a "low-action" region and to keep the state process within this region by singular consumption on its boundary. The task is to find the boundary.

The DPE connected with this problem is given by the variational inequality

$$\begin{aligned} \gamma &\leq \alpha V + \frac{1}{2}\sigma^2 V' + (x - \widehat{x})^2, \\ V &\leq a, \; V \geq -a, \\ 0 &= \left(-\gamma + \alpha V + \frac{1}{2}\sigma^2 V' + (x - \widehat{x})^2 \right) (a - V)(a + V), \end{aligned}$$

or the free-boundary problem

$$\begin{array}{cc} V=-a, & x\leq x_0, \\ \gamma=\alpha V+\frac{1}{2}\sigma^2 V'+(x-\widehat{x})^2, & x_0\leq x\leq x_1, \\ V=a, & x_1\leq x. \end{array} \tag{6.25}$$

Explicit solution in case $\boldsymbol{\alpha}=\mathbf{0}$ By subtracting a particular polynomial solution, one finds that

$$V(x)=\begin{cases} -a, & x\leq x_0, \\ Px^3+Qx^2+Rx+S, & x_0\leq x\leq x_1, \\ a, & x_1\leq x, \end{cases}$$

with

$$\begin{aligned} P &= -\frac{2}{3\sigma^2}, \\ Q &= \frac{2\widehat{x}}{\sigma^2}, \\ R &= \frac{2\left(\gamma-\widehat{x}^2\right)}{\sigma^2}. \end{aligned}$$

The principle of smooth fit gives

$$\begin{aligned} Px_0^3+Qx_0^2+Rx_0+S &= -a, \\ Px_1^3+Qx_1^2+Rx_1+S &= a, \\ 3Px_0^2+2Qx_0+R &= 0, \\ 3Px_1^2+2Qx_1+R &= 0, \end{aligned}$$

from which we can determine (x_0,x_1,S,γ),

$$\begin{aligned} S &= \frac{-2\widehat{x}\left(3\left(\frac{\left(6a\sigma^2\right)^{1/3}}{2}\right)^2-\widehat{x}^2\right)}{3\sigma^2}, \\ \gamma &= \left(\frac{\left(6a\sigma^2\right)^{1/3}}{2}\right)^2, \\ x_0 &= \widehat{x}-\frac{\left(6a\sigma^2\right)^{1/3}}{2}, \\ x_1 &= \widehat{x}+\frac{\left(6a\sigma^2\right)^{1/3}}{2}. \end{aligned}$$

The region within which the state process is optimally kept by singular control, is determined by the barriers (x_0,x_1). The corresponding region for the nominal wealth is determined by $\left(\frac{x_0}{\Lambda},\frac{x_1}{\Lambda}\right)$.

Explicit solution in case $\alpha \neq 0$ By subtracting a particular polynomial solution, one finds that

$$V(x) = \begin{cases} -a, & x \le x_0, \\ Px^2 + Qx + R + U(x), & x_0 \le x \le x_1, \\ a, & x_1 \le x, \end{cases}$$

with

$$\begin{aligned} P &= -\frac{1}{\alpha}, \\ Q &= \frac{\sigma^2}{\alpha^2} + \frac{2\widehat{x}}{\alpha}, \\ R &= \frac{\gamma - \widehat{x}^2}{\alpha} - \frac{\sigma^2}{\alpha^2}\widehat{x} - \frac{\sigma^4}{2\alpha^3}, \end{aligned} \tag{6.26}$$

and U solving the differential equation

$$0 = \alpha U + \frac{1}{2}\sigma^2 U_x, \text{ i.e. } U = Ce^{-\frac{2\alpha}{\sigma^2}x}.$$

The principle of smooth fit gives

$$\begin{aligned} Px_0^2 + Qx_0 + R + Ce^{-\frac{2\alpha}{\sigma^2}x_0} &= -a, \\ Px_1^2 + Qx_1 + R + Ce^{-\frac{2\alpha}{\sigma^2}x_1} &= a, \\ 2Px_0 + Q - \frac{2\alpha}{\sigma^2}Ce^{-\frac{2\alpha}{\sigma^2}x_0} &= 0, \\ 2Px_1 + Q - \frac{2\alpha}{\sigma^2}Ce^{-\frac{2\alpha}{\sigma^2}x_1} &= 0, \end{aligned}$$

from which we can determine (x_0, x_1, C, γ) numerically. The system of 4 unknowns can be reduced to a system of 2 unknowns (x_0, x_1),

$$\begin{aligned} -\alpha a + (x_0 - \widehat{x})^2 &= \alpha a + (x_1 - \widehat{x})^2, \\ \left(\widehat{x} - x_0 + \frac{\sigma^2}{2\alpha}\right) e^{\frac{2\alpha}{\sigma^2}x_0} &= \left(\widehat{x} - x_1 + \frac{\sigma^2}{2\alpha}\right) e^{\frac{2\alpha}{\sigma^2}x_1}. \end{aligned}$$

The region within which the state process is optimally kept by singular control, is determined by the barriers (x_0, x_1). The corresponding region for the nominal wealth is determined by $\left(\frac{x_0}{\Lambda}, \frac{x_1}{\Lambda}\right)$.

In Figure 6.1 we have illustrated how the barriers within which X is to be kept, depend on α. We have fixed the parameters at

$$a = 1,\ \sigma = \sqrt{2},\ \widehat{x} = 0,\ \widehat{u} = 0. \tag{6.27}$$

The solid lines (PF stands for pension funding) represent the upper and the lower barrier as a function of α. The interpretation is clear: If we increase the drift α, we need more pressure on X from above and the upper barrier tends to 0 whereas the lower barrier tends to minus infinity.

Remark 13 *The differential stopping game problem corresponding to the problem studied in this section is that of finding*

$$\max_{\rho} \min_{\tau} E\left[\int_0^{\rho \wedge \tau} 2\left(X\left(t\right)-\widehat{x}\right) dt + a1_{(\tau<\rho)} - a1_{(\rho<\tau)}\right].$$

In the case $\alpha = 0$, the corresponding differential stopping game reduces to a stopping time problem with absorption of finding

$$\min_{\tau} E\left[\int_0^{\eta \wedge \tau} 2\left(X\left(t\right)-\widehat{x}\right) dt + a1_{(\tau<\eta)}\right],$$

where

$$\eta = \inf\left\{t \geq 0 : X\left(t\right) = \widehat{x}\right\},$$

the first time X hits $\widehat{x}$.

6.8.4 Stationary constrained consumption

The singular stationary constrained problem will be solved explicitly below. We have not found the explicit solution for the stationary problem in the literature. However, one can argue that this is no more than a nice exercise in smooth fit.

The life and pension insurance interpretation of the problem studied in this section is that of finding an optimal upper surplus barrier for a participating life insurance company working with an infinite time horizon (the stationary view).

We deal with the problem to minimize over $\mathcal{U}_+^{bv}$,

$$\lim_{T\to\infty} E\left[\frac{1}{T}\int_0^T \left(\left(X\left(t\right)-\widehat{x}\right)^2 dt + a\left|dU\left(t\right)\right|\right)\right].$$

We let

$$\gamma = \min_{U\in\mathcal{U}_+^{bv}} \lim_{T\to\infty} E\left[\frac{1}{T}\int_0^T \left(\left(X\left(t\right)-\widehat{x}\right)^2 dt + a\left|dU\left(t\right)\right|\right)\right],$$

denote by $\int V$ the potential function, and want to find γ and V. As described in Section 6.4.2, it is optimal to consume $dU\left(t\right) = \widehat{u}dt$, or $dU'\left(t\right) = \frac{\widehat{u}}{\Lambda(t)}dt$, in a "low-action" region and to keep the state process within this region by singular consumption on its boundary. The task is to find the boundary.

The DPE connected with this problem is given by the variational inequality

$$\begin{aligned}
\gamma &\leq \alpha V + \frac{1}{2}\sigma^2 V_x + \left(x-\widehat{x}\right)^2, \\
V &\leq a, \\
0 &= \left(-\gamma + \alpha V + \frac{1}{2}\sigma^2 V' + \left(x-\widehat{x}\right)^2\right)\left(a - V\right),
\end{aligned}$$

or the free-boundary problem

$$\begin{array}{rl}
\gamma = \alpha V + \frac{1}{2}\sigma^2 V' + \left(x-\widehat{x}\right)^2, & x \leq x_1, \\
V = a, & x_1 \leq x.
\end{array}$$

For $\alpha > 0$, by subtracting a particular polynomial solution and requiring polynomial growth, one finds that

$$V(x) = \begin{cases} Px^2 + Qx + R, & x \leq x_1, \\ a, & x_1 \leq x, \end{cases}$$

with (P, Q, R) given by (6.26). The principle of smooth fit gives

$$\begin{aligned} Px_1^2 + Qx_1 + R &= a, \\ 2Px_1 + Q &= 0, \end{aligned}$$

from which we can determine (x_1, γ),

$$\begin{aligned} x_1 &= \widehat{x} + \frac{\sigma^2}{2\alpha}, \\ \gamma &= \frac{\sigma^4}{4\alpha^2} + a\alpha. \end{aligned}$$

The region within which the state process is optimally kept by singular control, is determined by the barrier x_1. The corresponding region for the nominal wealth is determined by $\frac{x_1}{\Lambda}$.

In Figure 6.1 we have illustrated how the barrier below which X is to be kept depends on α. We have fixed the parameters in accordance with (6.27). The broken line (PL stands for participating life) for $\alpha > 0$ represent the upper barrier as a function of α. For $\alpha < 0$ the broken line is a lower barrier that works for the constrained case where the control is required to be chosen negative. For $\alpha \to 0^+$, the (upper) barrier tends to infinity. I.e. for small α we should not press much from above in order to keep the process "close" to 0. For $\alpha = 0$, no stationary distribution exists.

Remark 14 *The stopping time problem corresponding to the problem studied in this section is that of finding*

$$\min_{\tau} E\left[\int_0^{\tau} 2\left(X(t) - \widehat{x}\right) dt\right].$$

6.9 Optimal classical consumption

6.9.1 Finite time unconstrained consumption

The classical finite time unconstrained problem will be solved explicitly below. This problem is the finite time linear regulator problem known from just about every textbook on stochastic control, see e.g. Fleming and Soner [24]. However, we have not found the solution to the terminally constrained problem anywhere in the literature. This is certainly an interesting variation of the linear regulator problem, and we have placed the calculations in Appendix A to this problem. However, one may

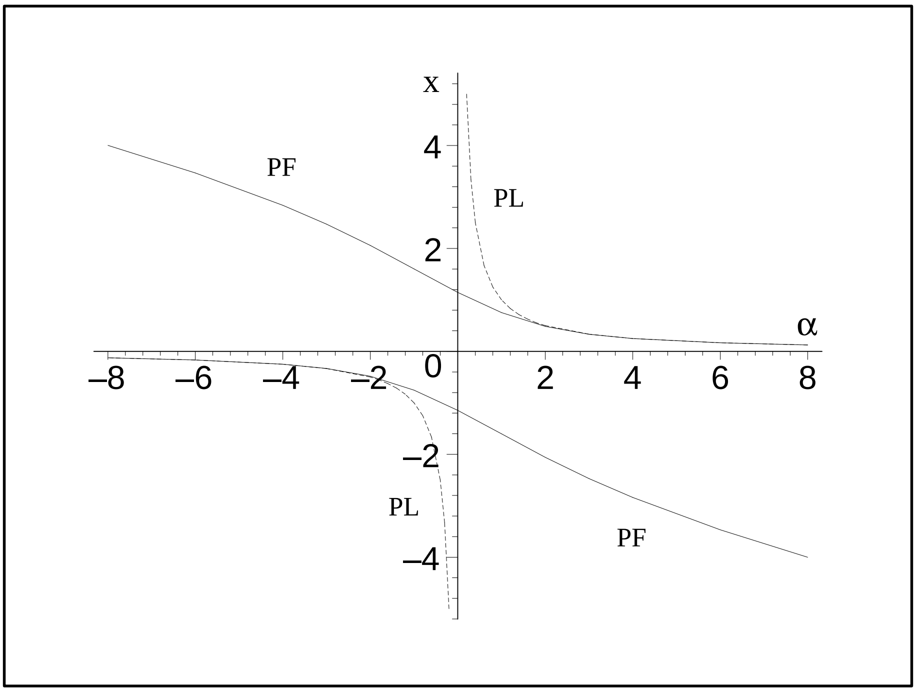

Figure 6.1: Surplus barriers

argue that this is no more than a nice exercise in the application of the Lagrange multiplier for a terminal constraint.

The life and pension insurance interpretation of the problem studied in this section is that of finding an optimal time-dependent absolutely continuous dividend strategy for a pension fund working with a finite time horizon.

We deal with the problem to minimize over $\mathcal{U}^{ac}$, subject to the terminal condition $E[X(T)] = 0$,

$$E\left[\int_0^T \left((X(t) - \widehat{x})^2 dt + a(u - \widehat{u})^2 dt\right)\right].$$

We introduce the related, but unconstrained, problem to minimize over $\mathcal{U}^{ac}$,

$$E\left[\int_0^T \left((X(t) - \widehat{x})^2 dt + a(u - \widehat{u})^2 dt\right) + \lambda X(T)\right].$$

We let

$$V(t,x) = \min_{U \in \mathcal{U}^{ac}} E\left[\int_t^T \left((X(t) - \widehat{x})^2 dt + a(u - \widehat{u})^2 dt\right) + \lambda X(T) \middle| X(t) = x\right],$$

and we want to find $V(0,x)$ with λ determined such that $E[X(T)] = 0$ and the corresponding optimal consumption plan.

In Appendix A, we have solved the linear regulator problem with terminal constraint. We get the solution to our problem by replacing u and $\widehat{u}$ by $\alpha - \widehat{u}$ and $\alpha - u$, respectively. The optimal control is then given by

$$u(t) = \alpha + \frac{P(t)}{a}(X(t) - \widehat{x}) + h(t),$$

where

$$
\begin{aligned}
P(t) &= \sqrt{a}\tanh\left(\sqrt{\frac{1}{a}}(T-t)\right),\\
h(t) &= \frac{\frac{2\sqrt{a}(x-\widehat{x})}{\sinh\left(\sqrt{\frac{1}{a}}T\right)}+\frac{2\sqrt{a}\widehat{x}}{\tanh\left(\sqrt{\frac{1}{a}}T\right)}}{2a\cosh\left(\sqrt{\frac{1}{a}}(T-t)\right)},
\end{aligned}
$$

and the terminally constrained optimally controlled process is an Ornstein-Uhlenbeck process with time-dependent coefficients. We remark that the term $h(t)$ stems from the terminal condition, and the optimal control for the problem with no terminal constraints is as above with $h(t)=0$.

We return to nominal consumption by

$$
u'(t)=\frac{\alpha}{\Lambda(t)}+\frac{P(t)}{a}\left(X'(t)-\widehat{x}'(t)\right)+h'(t),
$$

where

$$
h'(t)=\frac{\frac{2\sqrt{a}(x'(t)-\widehat{x}'(t))}{\sinh\left(\sqrt{\frac{1}{a}}T\right)}+\frac{2\sqrt{a}\widehat{x}'(t)}{\tanh\left(\sqrt{\frac{1}{a}}T\right)}}{\cosh\left(\sqrt{\frac{1}{a}}(T-t)\right)}
$$

and

$$
\frac{\alpha}{\Lambda(t)}=\alpha'(t)-\delta'(t)+h(t)\left(\overline{\sigma}'(t)-\overline{\psi}'(t)\right)+g(t)\left(\sigma'(t)-\psi'(t)\right).
$$

In Figure 6.2 we have illustrated the optimal control of the linear regulator problem with terminal condition by displaying $E[X(t)]$ for an optimally controlled process X as a function of t. In Appendix A this is the quantity $m(t)$. We have fixed the parameters at

$$
\alpha=0,\ \sigma=\sqrt{2},\ x=0,\ \widehat{x}=1,\ \widehat{u}=0,\ T=10.
$$

This means that we start out with zero wealth and expect zero wealth at time T, but in between we wish our wealth to be close to 1. The figure shows that we, on the average, should save money at the beginning by negative consumption in order to build up a wealth and then, towards the end of the period, consume the reserved wealth. The expected wealth is drawn for three different choices of a. For a large a stability of consumption (relative to 0) is more important than stability of wealth (relative to 1), and we should not build up as large a wealth as in the case of a small a.

6.9.2 Finite time constrained consumption

The classical finite time constrained problem will be approached below. We shall here see how far we can get in a search for an explicit solution. However, at the end

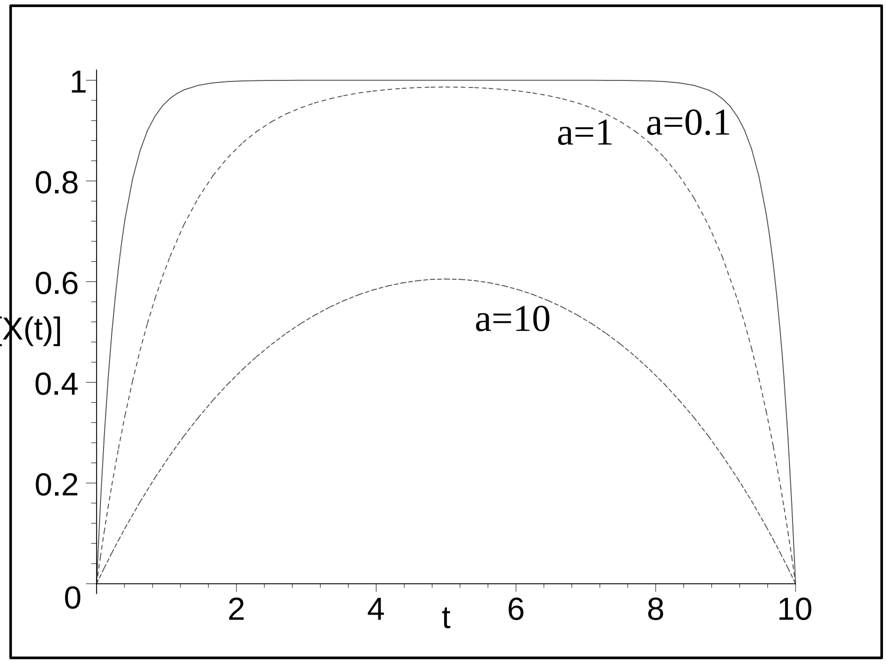

Figure 6.2: $E[X(t)]$ as a function of t

we are left with a rather nasty PDE which needs to be approached by numerical methods. Unfortunately, this section shows that constrained consumption complicates substantially the linear regulator problem. This is probably the explanation why we have not found in the literature any attempts at solving any constrained versions of the otherwise very well-studied linear regulator problem.

The life and pension insurance interpretation of the problem studied in this section is that of finding an optimal time-dependent absolutely continuous dividend strategy for a participating life insurance company working with a finite time horizon.

We deal with the problem to minimize over $\mathcal{U}_+^{ac}$, subject to the terminal condition $E[X_T] = 0$,

$$E\left[\int_0^T \left((X(t) - \widehat{x})^2 dt + a(u - \widehat{u})^2 dt\right)\right].$$

We introduce the related, but unconstrained, problem to minimize over $\mathcal{U}_+^{ac}$

$$E\left[\int_0^T \left((X(t) - \widehat{x})^2 dt + a(u - \widehat{u})^2 dt\right) + \lambda X_T\right].$$

We let

$$V(t,x) = \min_{U \in \mathcal{U}_+^{ac}} E\left[\int_t^T \left((X(t) - \widehat{x})^2 dt + a(u - \widehat{u})^2 dt\right) + \lambda X_T \middle| X(t) = x\right],$$

and we want to find $V(0,x)$ with λ determined such that $E[X(T)] = 0$ and the corresponding optimal consumption plan.

The DPE connected with this problem is given by

$$\begin{aligned} -V_t &= \min_{U \in \mathcal{U}_+^{ac}} \left[(\alpha - u) V_x + \frac{1}{2}\sigma^2 V_{xx} + (x - \widehat{x})^2 + a (u - \widehat{u})^2 \right], \\ V(T, x) &= \lambda x. \end{aligned}$$

Differentiating the DPE with respect to u and equating the right hand side to 0 suggests the control

$$u = \begin{cases} 0, & x \le x_0(t), \\ \widehat{u} + \frac{V_x}{2a}, & x \ge x_0(t), \end{cases}$$

and plugging this control into the DPE gives the differential equation,

$$-V_t = \begin{cases} V_x \alpha + \frac{1}{2} V_{xx} \sigma^2 + (x - \widehat{x})^2 + a \widehat{u}^2, & x \le x_0(t), \\ V_x (\alpha - \widehat{u}) - \frac{(V_x)^2}{4a} + \frac{1}{2} V_{xx} \sigma^2 + (x - \widehat{x})^2, & x \ge x_0(t). \end{cases} \tag{6.28}$$

Upon subtracting a particular polynomial solution with time-dependent coefficients, one finds that

$$V(t, x) = \begin{cases} P^-(t) x^2 + Q^-(t) x + R^-(t) + U(t, x), & x \le x_0(t), \\ P^+(t) x^2 + Q^+(t) x + R^+(t) + U(t, x), & x \ge x_0(t), \end{cases}$$

with (P^-, Q^-, R^-) given by (6.23) with $\alpha \widehat{u} (T - t)$ added to R^- and (P^+, Q^+, R^+) given by (A.1) in Appendix A. For $x \le x_0(t)$, U solves (6.24), and for $x \ge x_0(t)$, U solves

$$-U_t = \left(\alpha - \widehat{u} - \frac{2P^+(t) x + Q^+(t)}{2a} - \frac{U_x}{4a} \right) U_x + \frac{1}{2} U_{xx} \sigma^2. \tag{6.29}$$

The suggested optimal control is then given by

$$\begin{aligned} u &= \left(\widehat{u} + \frac{2P^+(t) X(t) + Q^+(t) + U_x(t, X(t))}{2a} \right)^+ \\ &= \left(\alpha + \frac{P^+(t)}{a} (X(t) - \widehat{x}) + \frac{-2(\alpha - \widehat{u}) a + \lambda}{2a \cosh\left(\sqrt{\frac{1}{a}} (T - t) \right)} + \frac{U_x(t, X(t))}{2a} \right)^+. \end{aligned}$$

The principle of smooth fit and the polynomial growth condition are now supposed to give boundary conditions for a unique solution U to this problem which we shall not pursue further, though.

6.9.3 Stationary unconstrained consumption

The classical stationary unconstrained problem will be solved explicitly below. This problem is the stationary version of the linear regulator problem known from just about every textbook on stochastic control, see e.g. Fleming and Soner [24]. Although the linear regulator problem has been studied intensively in the literature, we have not found the explicit solution to the stationary version of the problem there. However, one may argue that it is no more than a nice exercise in smooth fit.

The life and pension insurance interpretation of the problem studied in this section is that of finding an optimal absolutely continuous dividend strategy for a pension fund working with an infinite time horizon (stationary view).

We deal with the problem to minimize over $\mathcal{U}^{ac}$,

$$\lim_{T\to\infty} E\left[\frac{1}{T}\int_0^T \left((X(t)-\widehat{x})^2\,dt + a\,(u-\widehat{u})^2\,dt\right)\right].$$

We introduce

$$\gamma = \min_{U\in\mathcal{U}^{ac}} \lim_{T\to\infty} E\left[\frac{1}{T}\int_0^T \left((X(t)-\widehat{x})^2\,dt + a\,(u-\widehat{u})^2\,dt\right)\right],$$

denote by $\int V$ the potential function, and want to find γ, V, and the corresponding optimal consumption plan.

The DPE connected with this problem is given by

$$\gamma = \min_{U\in\mathcal{U}^{ac}} \left[(\alpha - u)\,V + \frac{1}{2}\sigma^2 V' + (x-\widehat{x})^2 + a\,(u-\widehat{u})^2\right].$$

Differentiating the DPE with respect to u and equating the right hand side to 0 results in the optimizer

$$u = \widehat{u} + \frac{V}{2a}, \tag{6.30}$$

and plugging this control into the DPE gives the differential equation,

$$\gamma = (\alpha-\widehat{u})\,V - \frac{V^2}{4a} + \frac{1}{2}\sigma^2 V' + (x-\widehat{x})^2.$$

Differentiating the right hand side of the DPE once more with respect to u shows that u given by (6.30) is minimizing.

We have a solution in the form

$$V(x) = Px + Q,$$

with (P, Q) given by

$$\begin{aligned} P &= 2\sqrt{a}, \\ Q &= 2a\,(\alpha-\widehat{u}) - 2\sqrt{a}\widehat{x}. \end{aligned} \tag{6.31}$$

We find that

$$\gamma = a\,(\alpha-\widehat{u})^2 + \sqrt{a}\sigma^2.$$

The optimal control reads

$$u(t) = \alpha + \sqrt{\frac{1}{a}}\,(X(t)-\widehat{x}). \tag{6.32}$$

The optimal nominal consumption is now given by

$$u'(t) = \frac{\alpha}{\Lambda(t)} + \sqrt{\frac{1}{a}}\,(X'(t)-\widehat{x}'(t)), \tag{6.33}$$

where

$$\frac{\alpha}{\Lambda(t)} = \alpha'(t) - \delta'(t) + h(t)\left(\overline{\sigma}'(t) - \overline{\psi}'(t)\right) + g(t)\,(\sigma'(t)-\psi'(t)).$$

6.9.4 Stationary constrained consumption

The classical stationary constrained problem will be approached by numerical methods below. This section is very important since we more or less gave up on the infinite time version of the classical constrained problem. We reach at a certain Riccati problem with growth condition. This problem is, also from a purely mathematical point of view, interesting, and a vain attempt at finding a (more or less) explicit solution is exposed in Appendix B. However, in the present section we illustrate the solution by approaching the Riccati equation directly with numerical methods.

The life and pension insurance interpretation of the problem studied in this section is that of finding an optimal absolutely continuous dividend strategy for a participating life insurance company working with an infinite time horizon (stationary view).

We deal with the problem to minimize over $\mathcal{U}_{+}^{ac}$,

$$\lim_{T\to\infty} E\left[\frac{1}{T}\int_0^T \left(\left(X\left(t\right)-\widehat{x}\right)^2 dt + a\left(u-\widehat{u}\right)^2 dt\right)\right].$$

We let

$$\gamma = \min_{U\in\mathcal{U}_{+}^{ac}} \lim_{T\to\infty} E\left[\frac{1}{T}\int_0^T \left(\left(X\left(t\right)-\widehat{x}\right)^2 dt + a\left(u-\widehat{u}\right)^2 dt\right)\right],$$

denote by $\int V$ the potential function, and want to find γ, V, and the corresponding optimal consumption plan.

The DPE connected with this problem is given by

$$\gamma = \min_{U\in\mathcal{U}_{+}^{ac}} \left[\left(\alpha-u\right)V + \frac{1}{2}\sigma^2 V' + \left(x-\widehat{x}\right)^2 + a\left(u-\widehat{u}\right)^2\right].$$

Differentiating the DPE with respect to u and equating the right hand side to 0 suggests existence of an x_0 such that

$$u = \begin{cases} \widehat{u} + \frac{V}{2a}, & x \geq x_0, \\ 0, & x \leq x_0. \end{cases}$$

Plugging this control into the DPE gives the differential equation,

$$\gamma = \begin{cases} V\alpha + \frac{1}{2}V'\sigma^2 + \left(x-\widehat{x}\right)^2 + a\widehat{u}^2 = 0, & x \leq x_0, \\ V\left(\alpha-\widehat{u}\right) - \frac{V^2}{4a} + \frac{1}{2}V'\sigma^2 + \left(x-\widehat{x}\right)^2 = 0, & x \geq x_0. \end{cases} \tag{6.34}$$

By subtraction of quadratic functions with constant coefficients, one finds that

$$V\left(x\right) = \begin{cases} P^- x^2 + Q^- x + R^-, & x \leq x_0, \\ P^+ x + Q^+ + U\left(x\right), & x \geq x_0, \end{cases}$$

with (P^-, Q^-, R^-) given by (6.26) and (P^+, Q^+) given by (6.31). U solves

$$U' = \left(\frac{2}{\sigma^2\sqrt{a}}\left(x-\widehat{x}\right) + \frac{1}{2\sigma^2 a}U\right)U + \frac{2\left(\gamma - a\left(\alpha-\widehat{u}\right)^2 - \sqrt{a}\sigma^2\right)}{\sigma^2}. \tag{6.35}$$

The optimal control is given by

$$u(t)=\left(\alpha+\sqrt{\frac{1}{a}}\left(X(t)-\widehat{x}\right)+\frac{U\left(X(t)\right)}{2a}\right)^{+}. \tag{6.36}$$

The principle of smooth fit gives directly from (6.34) γ expressed as a function of x_0,

$$\gamma(x_0)=\left(\widehat{x}-x_0\right)^2+\frac{\sigma^2}{\alpha}\left(\widehat{x}-x_0\right)+\frac{\sigma^4}{2\alpha^2}+a\widehat{u}^2-2\alpha a\widehat{u}, \tag{6.37}$$

and the following initial conditions for U,

$$\begin{aligned} U(x_0) &= -2a\alpha+2\sqrt{a}\left(\widehat{x}-x_0\right), \\ U'(x_0) &= \frac{\sigma^2}{\alpha^2}-2\sqrt{a}-\frac{2}{\alpha}\left(x_0-\widehat{x}\right). \end{aligned} \tag{6.38}$$

Now, our consumption problem is reduced to the problem of finding an initial point x_0 such that the Riccati equation (6.35) with γ given by (6.37) and initial values given by (6.38) obeys a polynomial growth condition. Since we have subtracted a particular polynomial solution, the polynomial growth condition reads $U\to 0$ as $x\to\infty$. A problem of this kind is in itself an interesting mathematical problem. A vain attempt at finding a solution to this problem can be found in Appendix B.

By numerical/graphical experimentation we can solve the problem of finding x_0 and then illustrate the optimal consumption by means of the differential equation (6.35). Figures 6.3 and 6.4 show the optimal control in the unconstrained case and in the constrained case for two different choices of α. The parameters are fixed at

$$a=0.5,\ \widehat{u}=0,\ \widehat{x}=0,\ \sigma=\sqrt{2},$$

and Figure 6.3 and Figure 6.4 show optimal control u as a function of X in the cases of $\alpha=1$ and $\alpha=0.5$, respectively. The broken line (PF stands for pension funding) represents the linear control given by (6.32) and the solid line (PL stands for participating life) represents (6.36). The solid line shows how one, in the constrained case, should not control until x exceeds a point x_0, from which one should perform a control which is asymptotically linear. Control in participating life is, in a sense, more defensive than control in pension funding in the sense that x_0 exceeds the point of neutral control ($u=0$) in pension funding. However, for $X>x_0$, it is worth noticing that the marginal control of PL actually exceeds the marginal control of PF. At the end of Section 6.10.4 we compare the cost of optimal control of pension funding and participating life, respectively, and we shall there also compare with two different suboptimal controls.

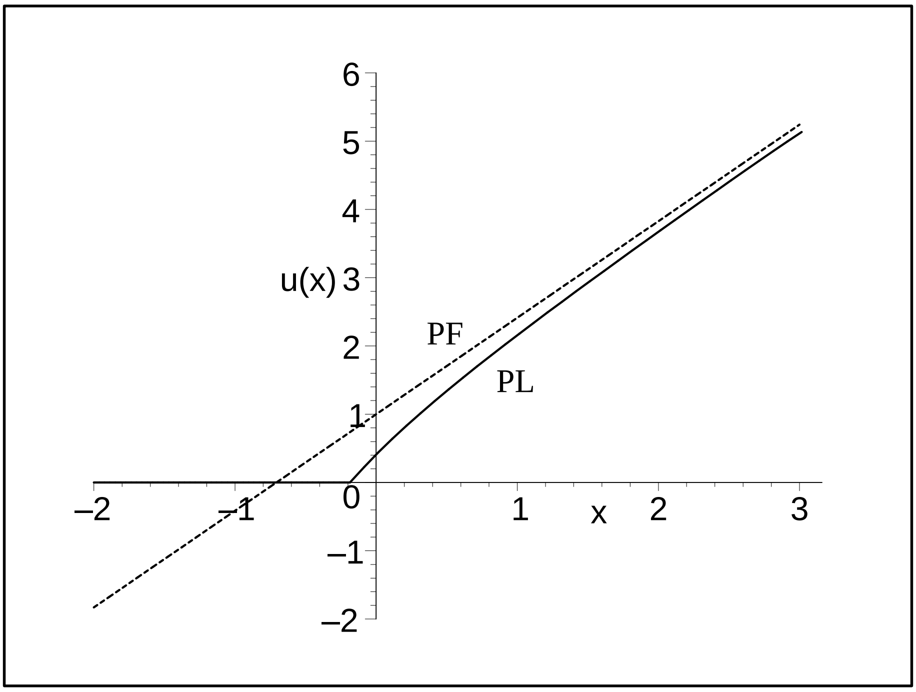

Figure 6.3: Optimal consumption, $\alpha = 1$

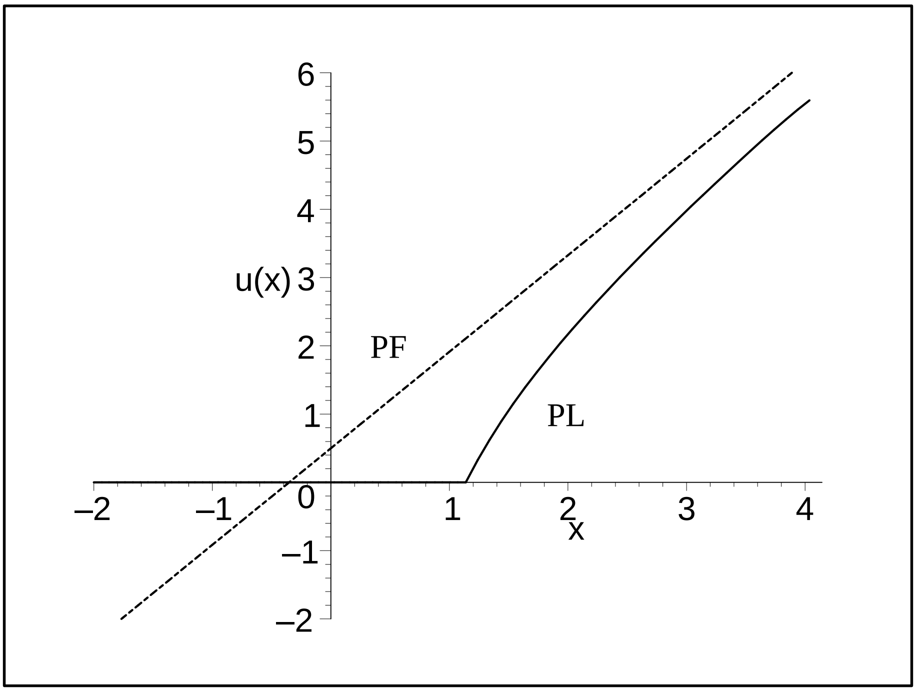

Figure 6.4: Optimal consumption, $\alpha = 0.5$

6.10 Suboptimal consumption

6.10.1 Finite time unconstrained consumption

A suboptimal finite time unconstrained problem will be solved almost explicitly below. We are left with a numerical solution to a non-linear equation. Though we found the explicit solution to the optimal problem in 6.9.1, one may be interested in a simplified procedure of control where the coefficients in the linear regulation do not depend on time.

The life and pension insurance interpretation of the problem studied in this section is that of finding an optimal time-independent linear dividend strategy for a pension fund working with a finite time horizon.

We deal with the problem to minimize over $\mathcal{U}^l$, subject to the terminal condition $E[X(T)] = 0$,

$$E\left[\int_0^T (X(t)-\widehat{x})^2\,dt + a(u-\widehat{u})^2\,dt\right].$$

Since for $U \in \mathcal{U}^l$ we have $u(t) = w + vX(t)$, the problem is to minimize over $(w,v) \in \mathbf{R}^2$, subject to $E[X(T)] = 0$, the quantity

$$g(w,v) \equiv E\left[\int_0^T (X(t)-\widehat{x})^2\,dt + a(w+vX(t)-\widehat{u})^2\,dt\right].$$

Note that given a solution, we simply obtain suboptimal nominal consumption by

$$u'(t) = \frac{w}{\Lambda(t)} + vX'(t).$$

The controlled process $(X(t))_{t\geq 0}$ is an Ornstein-Uhlenbeck process the marginal distributions of which are known to be normal. More specifically, we know that

$$\mathcal{L}(X(t)|\,X(0)=x) = N\left(\frac{\alpha-w}{v} + e^{-vt}\left(x - \frac{\alpha-w}{v}\right), \sigma^2\frac{1-e^{-2vt}}{2v}\right).$$

Now we simply use the moments of the normal distribution to get

$$\begin{aligned}
g(w,v) &= E\int_0^T \left((X(t)-\widehat{x})^2 + a(vX(t)+w-\widehat{u})^2\right)dt \\
&= E\int_0^T \left(\left(1+av^2\right)X(t)^2 + 2\left(awv - a\widehat{u}v - \widehat{x}\right)X(t)\right)dt \\
&\quad + \left(\widehat{x}^2 + a(w-\widehat{u})^2\right)T \\
&= \left(1+av^2\right)\left(\left(x-\frac{\alpha-w}{v}\right)^2 - \frac{\sigma^2}{2v}\right)\frac{1}{2v}\left(1-e^{-2vT}\right) \\
&\quad +2\left(1+av^2\right)\frac{\alpha-w}{v}\left(x-\frac{\alpha-w}{v}\right)\frac{1}{v}\left(1-e^{-vT}\right) \\
&\quad +\left(1+av^2\right)\left(\frac{\sigma^2}{2v}+\left(\frac{\alpha-w}{v}\right)^2\right)T \\
&\quad +\left(awv-a\widehat{u}v-\widehat{x}\right)\left(\frac{\alpha-w}{v}T+\left(x-\frac{\alpha-w}{v}\right)\frac{1}{v}\left(1-e^{-vT}\right)\right) \\
&\quad +\left(\widehat{x}^2+a(w-\widehat{u})^2\right)T, \qquad (6.39)
\end{aligned}$$

which we minimize subject to

$$\begin{aligned}
E[X(T)] &= \frac{\alpha-w}{v} + e^{-vT}\left(x-\frac{\alpha-w}{v}\right) = 0 \\
&\Leftrightarrow\ w = \alpha + v\frac{e^{-vT}}{1-e^{-vT}}x. \qquad (6.40)
\end{aligned}$$

Upon inserting (6.40) in (6.39), the problem reduces to that of minimizing with respect to v,

$$\begin{aligned}
g(v) &= \left(1+av^2\right)\left(\left(\frac{x}{1-e^{-vT}}\right)^2 - \frac{\sigma^2}{2v}\right)\frac{1}{2v}\left(1-e^{-2vT}\right) \\
&\quad -2\left(1+av^2\right)\frac{x^2}{1-e^{-vT}}\frac{1}{v}e^{-vT} \\
&\quad +\left(1+av^2\right)\left(\frac{\sigma^2}{2v}+\left(\frac{xe^{-vT}}{(1-e^{-vT})}\right)^2\right)T \\
&\quad +x\left(a\left(\alpha-\widehat{u}+vx\left(\frac{e^{-vT}}{(1-e^{-vT})}\right)\right)v-\widehat{x}\right)\left(\frac{-e^{-vT}}{(1-e^{-vT})}T+\frac{1}{v}\right) \\
&\quad +\left(\widehat{x}^2+a\left(\alpha-\widehat{u}+vx\frac{e^{-vT}}{(1-e^{-vT})}\right)^2\right)T.
\end{aligned}$$

For $x=0$, this reduces further to

$$g(v) = \left(1+av^2\right)\left(-\frac{\sigma^2}{4v^2}\left(1-e^{-2vT}\right)+\frac{\sigma^2}{2v}T\right)+\left(\widehat{x}^2+a(\alpha-\widehat{u})^2\right)T.$$

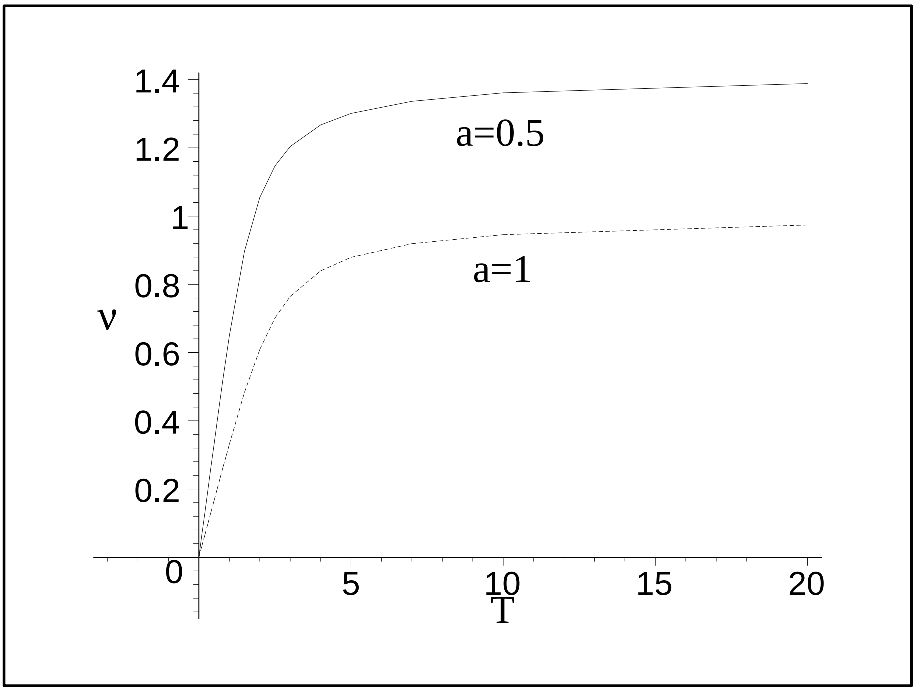

Figure 6.5: Optimal ν as a function of T

Figure 6.5 shows the optimal choice of v as a function of T for two different choices of a. We have fixed parameters at

$$\alpha = 0,\ \widehat{u} = 0,\ x = \widehat{x} = 0,\ \sigma = \sqrt{2}.$$

For T close to 0 one should not control aggressively since the control is expensive compared to what is gained from stability of wealth. As T goes to infinity the optimal v tends to the optimal coefficient $\sqrt{\frac{1}{a}}$ from optimal stationary control, (6.32), which for $a = 1$ equals 1 and for $a = 0.5$ equals $\sqrt{2}$.

6.10.2 Finite time constrained consumption

A suboptimal finite time unconstrained problem will be approached below. However, the marginal distributions of a process regulated linearly in only one direction seems hard to obtain. Therefore, we do not get far.

The life and pension insurance interpretation of the problem studied in this section is that of finding an optimal time-independent piecewise linear dividend strategy for a participating life insurance company working with a finite time horizon.

We deal with the problem to minimize over $\mathcal{U}_+^l$, subject to the terminal condition $E\left[X\left(T\right)\right] = 0$,

$$E\left[\int_0^T \left(\left(X\left(t\right) - \widehat{x}\right)^2 dt + a\left(u - \widehat{u}\right)^2 dt\right)\right].$$

Since for $U \in \mathcal{U}_+^l$ we have $u\left(t\right) = \left(w + v\left(X\left(t\right) - x_0\right)\right) 1_{\left(X\left(t\right) \geq x_0\right)}$, $w \geq 0$, the problem is to minimize over $(w, v, x_0) \in \mathbf{R}_+ \times \mathbf{R}^2$, subject to $E[X_T] = 0$,

$$E\left[\int_0^T \left((X(t) - \widehat{x})^2 dt + a\left((w + v(X(t) - x_0))1_{(X(t) \geq x_0)} - \widehat{u}\right)^2 dt\right)\right].$$

The controlled process $(X(t))_{t \geq 0}$ has linear regulation in one direction and constant regulation in the other direction. We shall not pursue its marginal distributions but take a close look at its stationary distribution below.

6.10.3 Stationary unconstrained consumption

A suboptimal stationary constrained problem will be solved explicitly below. Since the optimal solution obtained in 6.9.3 is actually in the desired linear form, we know the answer on beforehand: If a solution to a suboptimal problem is required to be in a certain parametric form, and the solution to the optimal problem is, in fact, in this form, then of course the optimal solution also solves the suboptimal problem. Nevertheless, we carry out the calculations as a warm-up for the succeeding section.

The life and pension insurance interpretation of the problem studied in this section is that of finding an optimal linear dividend strategy for a pension fund working with an infinite time horizon (stationary view).

We deal with the problem to minimize over $\mathcal{U}^l$,

$$\lim_{T \to \infty} E\left[\frac{1}{T}\int_0^T \left((X(t) - \widehat{x})^2 dt + a(u - \widehat{u})^2 dt\right)\right].$$

Since for $U \in \mathcal{U}^l$ we have $u(t) = w + vX(t)$, the problem is to minimize over $(w, v) \in \mathbf{R}^2$, the quantity

$$g(w, v) \equiv \lim_{T \to \infty} E\left[\frac{1}{T}\int_0^T \left((X(t) - \widehat{x})^2 dt + a(w + vX(t) - \widehat{u})^2 dt\right)\right].$$

We let

$$\gamma = \min_{(v,w) \in \mathbf{R}^2} \lim_{T \to \infty} E\left[\frac{1}{T}\int_0^T \left((X(t) - \widehat{x})^2 dt + a(w + vX_t - \widehat{u})^2 dt\right)\right],$$

and we want to find γ and the corresponding optimal (w, v).

The controlled process $(X(t))_{t \geq 0}$ is an Ornstein-Uhlenbeck process, the stationary distribution of which is known to be normal. More specifically, we know that

$$\mathcal{L}(X_t) \to N\left(\frac{\alpha - w}{v}, \frac{\sigma^2}{2v}\right) \text{ as } t \to \infty.$$

Now we simply use the moments of the normal distribution to get

$$\begin{aligned} g(w,v) &= \frac{1}{\sqrt{2\pi\frac{\sigma^2}{2v}}}\int_{-\infty}^{\infty}\left((x-\widehat{x})^2+a(vx+w-\widehat{u})^2\right)e^{-\frac{\left(x-\frac{\alpha-w}{v}\right)^2}{\frac{\sigma^2}{v}}}dx \\ &= \left(1+av^2\right)\left(\frac{\sigma^2}{2v}+\left(\frac{\alpha-w}{v}\right)^2\right) \\ &\quad +2\frac{\alpha-w}{v}\left(awv-a\widehat{u}v-\widehat{x}\right)+\widehat{x}^2+a\left(w^2+\widehat{u}^2-2w\widehat{u}\right) \\ &= a(\alpha-\widehat{u})^2+\frac{\sigma^2}{2}\left(\frac{1}{v}+av\right)+\left(\frac{\alpha-w}{v}-\widehat{x}\right)^2, \end{aligned}$$

which we minimize.

The derivative with respect to (w,v) is

$$\begin{aligned} l'_w &= -2\left(\frac{\alpha-w}{v}-\widehat{x}\right)\frac{1}{v}, \\ l'_v &= \frac{\sigma^2}{2}\left(-\frac{1}{v^2}+a\right)+2\left(\frac{\alpha-w}{v}-\widehat{x}\right)\frac{w}{v^2}, \end{aligned}$$

and equating these to 0 gives

$$\begin{aligned} v &= \sqrt{\frac{1}{a}}, \\ w &= \alpha-v\widehat{x}=\alpha-\sqrt{\frac{1}{a}}\widehat{x}. \end{aligned}$$

Thus, γ and the optimal control become

$$\begin{aligned} \gamma &= \sqrt{a}\sigma^2+a(\alpha-\widehat{u})^2, \\ u &= \alpha+\sqrt{\frac{1}{a}}(x-\widehat{x}). \end{aligned}$$

This was also the result which we found in Section 6.9.3.

6.10.4 Stationary constrained consumption

Two suboptimal stationary unconstrained problems will be solved almost explicitly below. We are left with numerical solutions to non-linear systems of equations. Since the solution to the optimal problem was approached in Section 6.9.4, we can compare the solutions to the optimal and suboptimal problems, respectively. The suboptimal problem and its solution is relevant if one is interested in a simplified procedure of control compared to the relatively involved solution to the optimal problem approached in Section 6.9.4.

The search for an explicit solution involves a study of the stationary distribution of what we choose to call the defective Ornstein-Uhlenbeck process. This process is regulated linearly in one direction (here from above) but only constantly in the other (here from below). We have not found any studies in the literature on the defective Ornstein-Uhlenbeck process. The process is certainly interesting in itself, and Appendix C contains a derivation of its stationary distribution and the functional needed in our optimization problem.

The life and pension insurance interpretation of the problem studied in this section is that of finding an optimal piecewise linear dividend strategy for a participating life insurance company working with an infinite time horizon (stationary view).

We deal with the problem to minimize over $\mathcal{U}_+^l$,

$$\lim_{T\to\infty} E\left[\frac{1}{T}\int_0^T \left((X(t)-\widehat{x})^2\,dt + a\,(u-\widehat{u})^2\,dt\right)\right]. \tag{6.41}$$

Since for $U \in \mathcal{U}_+^l$ we have $u(t) = (w + v(X(t)-x_0))\,1_{(X(t)\geq x_0)}$, $w \geq 0$, the problem is to minimize over $(w,v,x_0) \in \mathbf{R}_+ \times \mathbf{R}^2$, the quantity

$$\begin{aligned} g(x_0,w,v) &\equiv \lim_{T\to\infty} E[\frac{1}{T}\int_0^T ((X(t)-\widehat{x})^2\,dt \\ &\quad + a\left((w+v(X(t)-x_0))\,1_{(X(t)\geq x_0)}-\widehat{u}\right)^2 dt)]. \end{aligned}$$

We let

$$\begin{aligned} \gamma &= \min_{(x_0,w,v)\in\mathbf{R}_+\times\mathbf{R}^2} \lim_{T\to\infty} E[\frac{1}{T}\int_0^T ((X(t)-\widehat{x})^2\,dt + \\ &\quad a\left((w+v(X(t)-x_0))\,1_{(X(t)\geq x_0)}-\widehat{u}\right)^2 dt)], \end{aligned}$$

and want to find γ and the corresponding optimal (w,v,x_0). Note that given a solution, we simply obtain suboptimal nominal consumption by

$$u'(t) = \left(\frac{w}{\Lambda(t)} + v\left(X'(t) - \frac{x_0}{\Lambda(t)}\right)\right) 1_{\left(X'(t)\geq \frac{x_0}{\Lambda(t)}\right)}.$$

The defective Ornstein-Uhlenbeck process is treated in Appendix C, and from there we have

$$\begin{aligned} g(w,v,x_0) &= \frac{\sigma^2}{2}\left(\frac{1}{v}+av\right) + \left(\frac{\alpha-w}{v}+x_0-\widehat{x}\right)^2 \\ &\quad + a(\alpha-\widehat{u})^2 + \frac{\sigma^2}{2\alpha}C(w,v)\,N(w,v,x_0), \end{aligned}$$

where

$$\begin{aligned} C(w,v) &= \left(\frac{\sigma^2}{2\alpha} + \sqrt{\frac{\sigma^2}{2v}}\frac{\left(1-\Phi\left(-\sqrt{\frac{2}{v}}\frac{\alpha-w}{\sigma}\right)\right)}{\Phi'\left(-\sqrt{\frac{2}{v}}\frac{\alpha-w}{\sigma}\right)}\right)^{-1}, \\ N(w,v,x_0) &= \frac{\sigma^4}{2\alpha^2} - \frac{w^2}{v^2} + \left(\frac{\alpha}{v} + \alpha a v + 2(x_0-\widehat{x})\right)\left(\frac{w}{v} - \frac{\sigma^2}{2\alpha}\right). \end{aligned}$$

Instead of minimizing this quantity as a function of three variables, we can make use of the relation (which is easily realized)

$$\int_{-\infty}^{\infty} x\psi(x)\,dx = \widehat{x},$$

i.e.

$$\begin{aligned}\widehat{x} &= C(w,v)\,e^{-\frac{2\alpha}{\sigma^2}x_0}\int_{-\infty}^{x_0} x e^{\frac{2\alpha}{\sigma^2}x}dx \\ &\quad + C(w,v)\int_{x_0}^{\infty} x e^{\frac{2(\alpha-w+vx_0)}{\sigma^2}x-\frac{v}{\sigma^2}x^2-\frac{2(\alpha-w)}{\sigma^2}x_0-\frac{v}{\sigma^2}x_0^2}.\end{aligned}$$

By this constraint, $g(w,v,x_0)$ can be written as a function of (w,v),

$$\begin{aligned}g(w,v) &= \frac{\sigma^2}{2}\left(\frac{1}{v}+av\right)+\left(\frac{\alpha-w}{v}+x_0(w,v)-\widehat{x}\right)^2 \\ &\quad +a(\alpha-\widehat{u})^2+\frac{\sigma^2}{2\alpha}C(w,v)\,N(w,v),\end{aligned} \tag{6.42}$$

where

$$\begin{aligned}x_0(w,v) &= \widehat{x}-\frac{\alpha-w}{v}+C(w,v)\left(\frac{\sigma^4}{4\alpha^2}-\frac{\sigma^2}{2\alpha}\frac{w}{v}\right), \\ N(w,v) &= \frac{\sigma^4}{2\alpha^2}-\frac{w^2}{v^2}+\left(\frac{\alpha}{v}+\alpha av+2(x_0(w,v)-\widehat{x})\right)\left(\frac{w}{v}-\frac{\sigma^2}{2\alpha}\right),\end{aligned}$$

and $g(w,v)$ is to be minimized over $(w,v)\in\mathbf{R}_+\times\mathbf{R}^2$.

In the special case where we look for continuous control, we put $w=0$ in (6.42), and need to minimize over $v\in\mathbf{R}$ the quantity

$$\begin{aligned}g(v) &= \frac{\sigma^2}{2}\left(\frac{1}{v}+av\right)+\left(\frac{\alpha}{v}+x_0(0,v)-\widehat{x}\right)^2 \\ &\quad +a(\alpha-\widehat{u})^2+\frac{\sigma^2}{2\alpha}C(0,v)\,N(0,v).\end{aligned}$$

Note that given a solution, we simply obtain suboptimal continuous nominal consumption by

$$u'(t)=\left(vX'(t)-\frac{vx_0}{\Lambda(t)}\right)^+. \tag{6.43}$$

Figure 6.6 compare the optimal control; suboptimal piecewise linear control; and suboptimal piecewise linear continuous control. We have fixed parameters at

$$a=0.5,\ \alpha=0.5,\ \widehat{u}=0,\ \widehat{x}=0,\ \sigma=\sqrt{2}.$$

The solid line for optimal control ('O' stands for optimal) is the same as in Figure 6.4. The dotted line for suboptimal piece-wise linear control ('L' stands for linear) shows that one should start consuming at a point $x_0(w,v)$ larger than x_0 for the optimal control and jump up to start a linear control. The broken line for suboptimal

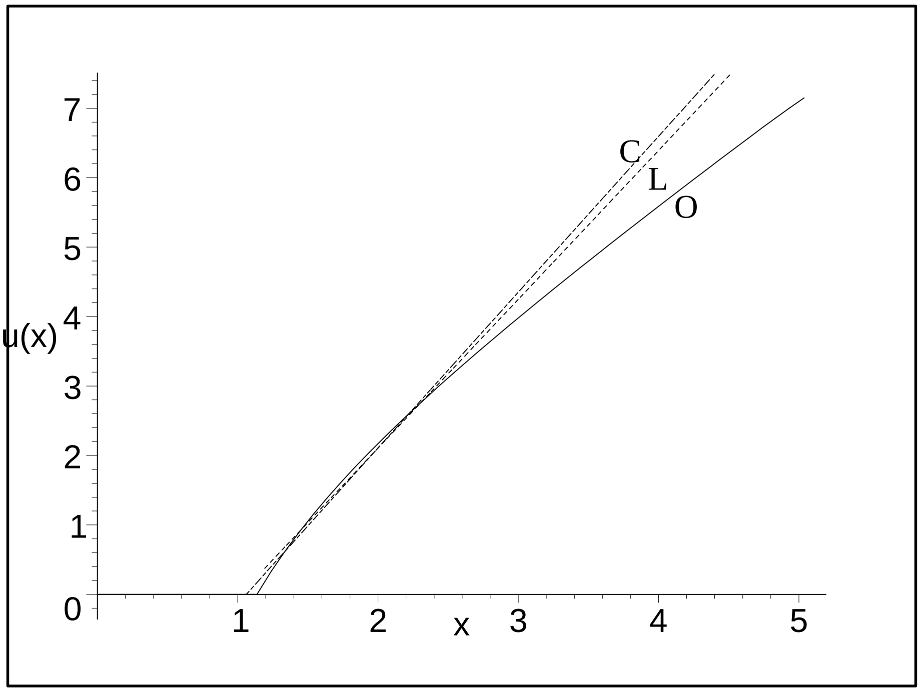

Figure 6.6: Suboptimal control

piecewise linear continuous control ('C' stands for continuous) shows that one should start consuming at a point $x_0(v)$ smaller that x_0 for the optimal control and then consume linearly from there. Figure 6.6 shows that the suboptimal control lies very close to the optimal control for small x. We can also compare the performance of the optimal control for the unconstrained case with the performance of the optimal and suboptimal controls in the constrained case.

	PF	PL, O	PL, L	PL, C
$\alpha = 1$	1.539	2.404	2.405	2.406
$\alpha = 0.5$	1.914	4.474	4.475	4.476

Table 6.1: Optimal cost of control

Table 6.1 shows the optimal cost of control γ for the various control problems for two different choices of α. We see, by comparing the first column with the second, that there is a substantial cost in constraining the problem and that this cost increases relatively when the drift α decreases. An important lesson to learn from Table 6.1 is that there is not much to gain from working with optimal control compared with suboptimal control, at least for the parameters chosen here. Actually, one needs three decimals to detect a difference in the performance of these controls. Depending on the parameters, one can conclude that the continuous control ('C') is so close to being optimal that it could be recommended due to simplicity.

The continuous suboptimal control for the constrained problem given by (6.43) and the optimal control for the unconstrained problem given by (6.33) make up relatively simple rules of consumption. We conclude by directing the reader's attention to the resemblance between these forms and the forms of retrospective surplus-linked dividends proposed in (3.24) and (3.25) in Chapter 3. Of course, the forms are proposed and derived, respectively, in rather different set-ups, and we shall not jump to conclusions on optimality of the surplus-linked form in Chapter 3. Nevertheless, the results in the present chapter throw light on the type of problems where surplus-linked payments actually can be considered as (close to) optimal, and contain suggestions, at the very least, for the parameters of the dividend strategy.

Appendix A

The linear regulator problem with terminal condition

Let X be given by

$$dX(t) = udt + \sigma dW(t),\; X(0) = x.$$

The objective is to minimize

$$E\left[\int_0^T \left((X(s) - \widehat{x})^2 + a(u - \widehat{u})^2\right) ds\right],$$

subject to $E[X(T)] = 0$. We introduce the related unconstrained problem to minimize

$$E\left[\int_0^T \left((X(t) - \widehat{x})^2 + a(u - \widehat{u})^2\right) dt + \lambda X(T)\right].$$

We let

$$V(t,x) = \min_{U \in \mathcal{U}^{AC}} E\left[\int_t^T \left((X(s) - \widehat{x})^2 + a(u - \widehat{u})^2\right) ds + \lambda X(T) \middle| X(t) = x\right],$$

and we want to find $V(0,x)$ with λ determined such that $E[X(T)] = 0$ and the corresponding optimal consumption plan.

The DPE equation connected with this problem is given by

$$\begin{aligned} -V_t &= \min_{U \in \mathcal{U}^{AC}} \left[uV_x + \frac{1}{2}\sigma^2 V_{xx} + (x - \widehat{x})^2 + a(u - \widehat{u})^2\right], \\ V(T,x) &= \lambda x. \end{aligned}$$

Differentiating the DPE with respect to u and equating the right hand side to 0, gives the control

$$u = \widehat{u} - \frac{V_x}{2a},$$

and plugging this control into the DPE gives the differential equation,

$$V_t + \widehat{u}V_x - \frac{V_x^2}{4a} + \frac{1}{2}\sigma^2 V_{xx} + (x - \widehat{x})^2 = 0.$$

Guessing a solution in the form

$$V(t,x) = P(t)\,x^2 + Q(t)\,x + R(t),$$

leads to an optimal control in the form

$$u = \widehat{u} - \frac{P(t)}{a}x - \frac{Q(t)}{2a},$$

and a Riccati system of differential equations for $(P(t), Q(t), R(t))$,

$$\begin{aligned}
P_t &= \frac{(P)^2}{a} - 1,\\
Q_t &= \frac{P}{a}Q - 2\widehat{u}P + 2\widehat{x},\\
R_t &= \frac{1}{4a}Q^2 - \widehat{u}Q - \sigma^2 P - \widehat{x}^2.
\end{aligned}$$

The side conditions $P(T) = 0$, $Q(T) = \lambda$, and $R(T) = 0$ lead, by some efforts, to the solutions

$$\begin{aligned}
P(t) &= \sqrt{a}\tanh\left(\sqrt{\frac{1}{a}}(T-t)\right),\\
Q(t) &= 2\widehat{u}a\left(1 - \frac{1}{\cosh\left(\sqrt{\frac{1}{a}}(T-t)\right)}\right)\\
&\quad -2\widehat{x}P(t) + \frac{\lambda}{\cosh\left(\sqrt{\frac{1}{a}}(T-t)\right)},\\
R(t) &= \int_t^T \left(-\frac{1}{4a}Q^2(s) + \widehat{u}Q(s) + \sigma^2 P(s) + \widehat{x}^2\right) ds.
\end{aligned} \tag{A.1}$$

where we for calculation of $Q(t)$ have made use of the relation

$$\begin{aligned}
e^{-\int_s^t \frac{P(\tau)}{a}d\tau} &= e^{\sqrt{\frac{1}{a}}\int_{T-s}^{T-t}\tanh\left(\sqrt{\frac{1}{a}}y\right)dy}\\
&= \frac{\cosh\left(\sqrt{\frac{1}{a}}(T-t)\right)}{\cosh\left(\sqrt{\frac{1}{a}}(T-s)\right)}.
\end{aligned}$$

The optimally controlled process is an Ornstein-Uhlenbeck process with time-dependent coefficients,

$$dX(t) = \left(\widehat{u} - \frac{P(t)}{a}X(t) - \frac{Q(t)}{2a}\right)dt + \sigma dW(t).$$

Now we need to determine λ such that $E[X(T)] = 0$, and we denote this λ by λ_0. Letting $m(t) = E[X(t)]$, we have that

$$m(t) = x + \int_0^t \left(\widehat{u} - \frac{P(s)}{a}m(s) - \frac{Q(s)}{2a}\right) ds,$$

or in differential form

$$\frac{dm(t)}{dt} = -\frac{P(t)}{a}m(t) + \widehat{u} - \frac{Q(t)}{2a}, \;\; m(0) = x.$$

This leads to

$$\begin{aligned} m(t) &= \cosh\left(\sqrt{\frac{1}{a}}(T-t)\right)\left[\frac{x}{\cosh\left(\sqrt{\frac{1}{a}}T\right)}\right. \\ &\quad + \left(\widehat{u} - \frac{\lambda}{2a}\right)\sqrt{a}\left(\tanh\left(\sqrt{\frac{1}{a}}T\right) - \tanh\left(\sqrt{\frac{1}{a}}(T-t)\right)\right) \\ &\quad \left. + \widehat{x}\left(\frac{1}{\cosh\left(\sqrt{\frac{1}{a}}(T-t)\right)} - \frac{1}{\cosh\left(\sqrt{\frac{1}{a}}T\right)}\right)\right]. \end{aligned} \tag{A.2}$$

Now the condition $m(T) = 0$ determines λ_0 by

$$\lambda_0 = \frac{2\sqrt{a}(x-\widehat{x})}{\sinh\left(\sqrt{\frac{1}{a}}T\right)} + \frac{2\sqrt{a}\widehat{x}}{\tanh\left(\sqrt{\frac{1}{a}}T\right)} + 2a(\alpha - \widehat{u}).$$

Finally, we find $m(t)$ and $Q(t)$ by setting $\lambda = \lambda_0$ in (A.1) and (A.2),

$$\begin{aligned} m(t) &= (x-\widehat{x})\frac{\sinh\left(\sqrt{\frac{1}{a}}(T-t)\right)}{\sinh\left(\sqrt{\frac{1}{a}}T\right)} - \widehat{x}\frac{\sinh\left(\sqrt{\frac{1}{a}}t\right)}{\sinh\left(\sqrt{\frac{1}{a}}T\right)} + \widehat{x}, \\ Q(t) &= 2a\widehat{u} - 2\widehat{x}P(t) + \frac{\frac{2\sqrt{a}(x-\widehat{x})}{\sinh\left(\sqrt{\frac{1}{a}}T\right)} + \frac{2\sqrt{a}\widehat{x}}{\tanh\left(\sqrt{\frac{1}{a}}T\right)}}{\cosh\left(\sqrt{\frac{1}{a}}(T-t)\right)}, \end{aligned}$$

such that the optimally controlled process having the right expectation at termination is the Ornstein-Uhlenbeck process with time-dependent coefficients,

$$dX(t) = \left(-\frac{P(t)}{a}(X(t)-\widehat{x}) - h(t)\right)dt + \sigma dW(t),$$

where

$$h(t) = \frac{\frac{(x-\widehat{x})}{\sinh\left(\sqrt{\frac{1}{a}}T\right)} + \frac{\widehat{x}}{\tanh\left(\sqrt{\frac{1}{a}}T\right)}}{\sqrt{a}\cosh\left(\sqrt{\frac{1}{a}}(T-t)\right)}.$$

Appendix B

Riccati equation with growth condition

Consider the Riccati equation

$$\begin{aligned} U' &= (a + bx + cU)\, U + f(x_0), \\ U(x_0) &= g(x_0), \\ U'(x_0) &= h(x_0), \end{aligned}$$

and the problem of finding x_0 such that $U \to 0$ for $x \to \infty$.

$$\begin{aligned} U &= -\frac{1}{c}\frac{Y'}{Y}, \\ U' &= -\frac{1}{c}\frac{Y''}{Y} + \frac{1}{c}\frac{Y'^2}{Y^2}, \end{aligned}$$

$$Y'' = (a + bx)\, Y' - cf(x_0)\, Y.$$

Now, the transformation

$$\begin{aligned} z &= \frac{1}{2b}(a + bx)^2, \\ Z(z) &= Y(x), \end{aligned}$$

gives

$$zZ'' = \left(z - \frac{1}{2}\right) Z' - \frac{cf(x_0)}{2b} Y,$$

which is the confluent hypergeometric differential equation the solution of which can be represented by the hypergeometric function

$$Z = C_1 F\left(-\frac{cf(x_0)}{2b}, \frac{1}{2}, z\right) + \sqrt{z} C_2 F\left(\frac{1}{2} - \frac{cf(x_0)}{2b}, \frac{3}{2}, z\right),$$

where

$$F(a, c, z) = \sum_{k=0}^{\infty} \frac{a(a+1)\cdots(a+k-1)}{c(c+1)\cdots(c+k-1)} \frac{1}{k!} z^k.$$

The problem here is to find the solution of Z which leads to non-exponential growth of U.

Appendix C

The defective Ornstein-Uhlenbeck process

Consider a defective Ornstein-Uhlenbeck process X with linear regulation in one direction and constant regulation in the other direction. The dynamics of X is given by

$$dX(t) = \alpha - (w + v(X(t) - x_0)) 1_{(X(t) \geq x_0)} dt + \sigma dW(t), \ X(0) = x,$$

for $v > 0$.

The stationary distribution We want to calculate the stationary density ψ and, following Karlin and Taylor [38], page 221, we get

$$\frac{1}{2}\sigma^2 \frac{\partial^2}{\partial x^2}\psi(x) - \frac{\partial}{\partial x}\left[\left(\alpha - (v(x - x_0) + w) 1_{(x > x_0)}\right)\psi(x)\right] = 0.$$

For $x < x_0$:

$$\begin{aligned}
\psi(x) &= e^{\int_{x_0}^{x} \frac{2\alpha}{\sigma^2} dz}\left[C_0 \int_{x_0}^{x} e^{-\int_{x_0}^{y} \frac{2\alpha}{\sigma^2} dz} dy + C_1\right] \\
&= C_1 e^{\int_{x_0}^{x} \frac{2\alpha}{\sigma^2} dz} \\
&= C_1 e^{\frac{2\alpha}{\sigma^2}(x - x_0)}.
\end{aligned}$$

For $x > x_0$:

$$\begin{aligned}
\psi(x) &= e^{\int_{x_0}^{x} \frac{2(\alpha - v(z - x_0) - w)}{\sigma^2} dz}\left[C_0 \int_{x_0}^{x} e^{-\int_{x_0}^{y} \frac{2(\alpha - v(z - x_0) - w)}{\sigma^2} dz} dy + C_2\right] \\
&= C_2 e^{\int_{x_0}^{x} \frac{2(\alpha - v(z - x_0) - w)}{\sigma^2} dz} \\
&= C_2 e^{\frac{2(\alpha - w + v x_0)}{\sigma^2} x - \frac{v}{\sigma^2} x^2 - \frac{2(\alpha - w)}{\sigma^2} x_0 - \frac{v}{\sigma^2} x_0^2}.
\end{aligned}$$

From the absolute continuity of $X - \sigma W$ we can conclude continuity of the density function, i.e.

$$\psi(x_0-) = \psi(x_0+),$$

which gives

$$C_1 = C_2 \equiv C.$$

Now, C must be determined such that

$$\int_{-\infty}^{\infty} \psi(x)\, dx = 1,$$

which gives

$$C = \left(\int_{-\infty}^{x_0} e^{\frac{2\alpha}{\sigma^2}(x-x_0)} + e^{-\frac{2(\alpha-w)}{\sigma^2}x_0 - \frac{v}{\sigma^2}x_0^2} \int_{x_0}^{\infty} e^{\frac{2(\alpha-w+vx_0)}{\sigma^2}x - \frac{v}{\sigma^2}x^2} dx \right)^{-1}.$$

Substituting

$$\begin{aligned} y &= -\sqrt{\frac{2}{v}}\frac{\alpha-w}{\sigma} + \sqrt{\frac{2v}{\sigma^2}}(x-x_0), \\ y_0 &= -\sqrt{\frac{2}{v}}\frac{\alpha-w}{\sigma}, \end{aligned}$$

we can write C as

$$C(w,v,x_0) = \left(\frac{\sigma^2}{2\alpha} + \sqrt{\frac{\sigma^2}{2v}} \frac{(1-\Phi(y_0))}{\Phi'(y_0)} \right)^{-1}.$$

The quantity g We want to calculate the quantity

$$g(w,v,x_0) = E\left[(X-\widehat{x})^2 + a\left((v(X-x_0)+w) 1_{(X>x_0)} - \widehat{u} \right)^2 \right],$$

where the distribution of X is the stationary distribution of the defective Ornstein-Uhlenbeck process. The stationary distribution above gives

$$\begin{aligned} g(w,v,x_0) &= C\int_{-\infty}^{x_0} \left((x-\widehat{x})^2 + a\widehat{u}^2 \right) e^{\frac{2\alpha}{\sigma^2}(x-x_0)} dx \\ &\quad + C\int_{x_0}^{\infty} k(x)\, e^{\frac{2(\alpha-w+vx_0)}{\sigma^2}x - \frac{v}{\sigma^2}x^2 - \frac{2(\alpha-w)}{\sigma^2}x_0 - \frac{v}{\sigma^2}x_0^2} dx, \end{aligned} \tag{C.1}$$

where

$$k(x) = (x-\widehat{x})^2 + a(v(x-x_0)+w-\widehat{u})^2.$$

Firstly, the substitution

$$\begin{aligned} z &= -\frac{2\alpha}{\sigma^2}x,\ z_0 = -\frac{2\alpha}{\sigma^2}x_0, \\ x &= -\frac{\sigma^2}{2\alpha}z, x_0 = -\frac{\sigma^2}{2\alpha}z_0, \end{aligned}$$

gives

$$\int_{-\infty}^{x_0} \left((x-\widehat{x})^2 + a\widehat{u}^2\right) e^{\frac{2\alpha}{\sigma^2}(x-x_0)} dx \tag{C.2}$$

$$= \frac{\sigma^2}{2\alpha} e^{z_0} \int_{z_0}^{\infty} \left(\frac{\sigma^4}{4\alpha^2} z^2 + \frac{\sigma^2}{\alpha}\widehat{x} z + \widehat{x}^2 + a\widehat{u}^2\right) e^{-z} dz$$

$$= \frac{\sigma^2}{2\alpha} \left(\frac{\sigma^4}{4\alpha^2}\left(z_0^2 + 2z_0 + 2\right) + \frac{\sigma^2}{\alpha}\widehat{x}\left(z_0+1\right) + \widehat{x}^2 + a\widehat{u}^2\right),$$

where we have used the relation

$$\begin{aligned} \int_x^{\infty} e^{-z} dz &= \left[-e^{-z}\right]_x^{\infty} = e^{-x}, \\ \int_x^{\infty} z e^{-z} dz &= \left[-e^{-z}\left(1+z\right)\right]_x^{\infty} = e^{-x}\left(x+1\right), \\ \int_x^{\infty} z^2 e^{-z} dz &= \left[-e^{-z}\left(2+2z+z^2\right)\right]_x^{\infty} = e^{-x}\left(x^2+2x+2\right). \end{aligned}$$

Secondly, the substitution

$$\begin{aligned} y &= -\sqrt{\frac{\sigma^2}{2v}} \frac{2\left(\alpha - w\right)}{\sigma^2} + \sqrt{\frac{2v}{\sigma^2}}\left(x - x_0\right), \\ y_0 &= -\sqrt{\frac{\sigma^2}{2v}} \frac{2\left(\alpha - w\right)}{\sigma^2}, \end{aligned}$$

gives

$$\int_{x_0}^{\infty} k\left(x\right) e^{\frac{2(\alpha - w + v x_0)}{\sigma^2}x - \frac{v}{\sigma^2}x^2 - \frac{2(\alpha-w)}{\sigma^2}x_0 - \frac{v}{\sigma^2}x_0^2} dx \tag{C.3}$$

$$= \sqrt{\frac{\sigma^2}{2v}} e^{\frac{1}{v}\left(\frac{\alpha-w}{\sigma}\right)^2} \int_{y_0}^{\infty} \widetilde{k}\left(y\right) e^{-\frac{1}{2}z^2} dy$$

$$\begin{aligned} = {} & \sqrt{\frac{\sigma^2}{2v}} \frac{\sigma^2}{2}\left(\frac{1}{v} + av\right)\left(y_0 + \frac{1-\Phi\left(y_0\right)}{\Phi'\left(y_0\right)}\right) \\ & + \frac{\sigma^2}{v}\left(\frac{\left(\alpha - w + v x_0\right)}{v} - \widehat{x} + av\left(\alpha - \widehat{u}\right)\right) \\ & + \sqrt{\frac{\sigma^2}{2v}}\left(\left(\frac{\left(\alpha - w + v x_0\right)}{v} - \widehat{x}\right)^2 + a\left(\alpha - \widehat{u}\right)^2\right)\frac{1-\Phi\left(y_0\right)}{\Phi'\left(y_0\right)} \end{aligned}$$

where

$$\begin{aligned} \widetilde{k}\left(y\right) = {} & \frac{\sigma^2}{2}\left(\frac{1}{v} + av\right) y^2 \\ & + 2\sqrt{\frac{\sigma^2}{2v}}\left(\frac{\left(\alpha - w + v x_0\right)}{v} - \widehat{x} + av\left(\alpha - \widehat{u}\right)\right) y \\ & + \left(\frac{\left(\alpha - w + v x_0\right)}{v} - \widehat{x}\right)^2 + a\left(\alpha - \widehat{u}\right)^2, \end{aligned}$$

and where we have used the relations

$$\begin{aligned}
\int_x^{\infty} e^{-\frac{z^2}{2}} dz &= \sqrt{2\pi}\left(1-\Phi(x)\right), \\
\int_x^{\infty} z e^{-\frac{z^2}{2}} dz &= -\left[e^{-\frac{z^2}{2}}\right]_x^{\infty} = e^{-\frac{x^2}{2}}, \\
\int_x^{\infty} z^2 e^{-\frac{z^2}{2}} dz &= -\left[z e^{-\frac{z^2}{2}}\right]_x^{\infty} + \int_x^{\infty} e^{-\frac{z^2}{2}} dz \\
&= x e^{-\frac{x^2}{2}} + \sqrt{2\pi}\left(1-\Phi(x)\right).
\end{aligned}$$

By (C.1), (C.2), and (C.1) we reach at

$$\begin{aligned}
g(w,v,x_0) &= \frac{\sigma^2}{2}\left(\frac{1}{v}+av\right)+\left(\frac{\alpha-w}{v}+x_0-\widehat{x}\right)^2 \\
&\quad +a\left(\alpha-\widehat{u}\right)^2+\frac{\sigma^2}{2\alpha}CN(w,v,x_0)
\end{aligned}$$

with

$$N(w,v,x_0)=\frac{\sigma^4}{2\alpha^2}-\frac{w^2}{v^2}+\left(\frac{\alpha}{v}+\alpha a v+2\left(x_0-\widehat{x}\right)\right)\left(\frac{w}{v}-\frac{\sigma^2}{2\alpha}\right).$$

Bibliography

[1] K. K. Aase and S.-A. Persson. Pricing of unit-linked life insurance policies. *Scandinavian Actuarial Journal*, pages 26–52, 1994.

[2] A. R. Bacinello and F. Ortu. Pricing equity-linked life insurance with endogeneous minimum guarantees. *Insurance: Mathematics and Economics*, 12:245–257, 1993.

[3] D. Becherer. Rational pricing and hedging under constant absolute risk aversion. Technical report, Technische Universität Berlin, 2001.

[4] Bedømmelsesforeningen. Dødelighedsundersøgelsen 1987-91. Foreningen til bedømmelse af personforsikringsrisiko, 1993.

[5] V. E. Beneš, L. A. Shepp, and H. S. Witsenhausen. Some solvable stochastic control problems. *Stochastics*, 4:39–83, 1980.

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

[7] D. Blake. Pension schemes as options on pension fund assets: implications for pension fund management. *Insurance: Mathematics and Economics*, 23(3):263–286, 1998.

[8] F. Boetius and M. Kohlmann. Connections between optimal stopping and singular stochastic control. *Stochastic Processes and their Applications*, 77:253–281, 1998.

[9] J.-F. Boulier, S. Michel, and V. Wisnia. Optimizing investment and contribution policies of a defined benefit pension fund. *Proceedings of the 6th AFIR International Colloquium*, 1:593–607, 1996.

[10] M. J. Brennan and E. S. Schwartz. The pricing of equity-linked life insurance policies with an asset value guarantee. *Journal of Financial Economics*, 3:195–213, 1976.

[11] E. Briys and F. de Varenne. Life insurance in a contingent claim framework: pricing and regulatory implications. *The Geneva Papers on Risk and Insurance Theory*, 19(1):53–72, 1994.

[12] A. J. G. Cairns. Some notes on the dynamics and optimal control of stochastic pension fund models in continuous time. *ASTIN Bulletin*, 30(1):19–55, 2000.

[13] J. D. Cummins and H. Geman. Pricing catastrophe futures and call spreads: An arbitrage approach. *Journal of Fixed Income*, pages 46–57, 1995.

[14] R.-A. Dana. Existence, uniqueness and determinacy equilibrium in CAPM with a riskless asset. *Journal of Mathematical Economics*, pages 167–175, 1999.

[15] M. H. A. Davis. *Markov Models and Optimization*. Chapman and Hall, 1993.

[16] C. D. Daykin, T. Pentikäinen, and M. Pesonen. *Practical Risk Theory for Actuaries*. Chapman and Hall, 1994.

[17] F. Delbaen and J. M. Haezendock. A martingale approach to premium calculation principles in an arbitrage free market. *Insurance: Mathematics and Economics*, (8):269–277, 1989.

[18] F. Delbaen and W. Schachermayer. A general version of the fundamental theorem of asset pricing. *Math. Ann.*, pages 463–520, 1994.

[19] D. Duffie and T. Zariphopoulou. Optimal investment with undiversifiable income risk. *Mathematical Finance*, 3(2):135–148, 1993.

[20] S. Ekern and S.-A. Persson. Exotic unit-linked life insurance contracts. *The Geneva Papers on Risk and Insurance Theory*, 21:35–63, 1996.

[21] J. Ekstrøm. Numerisk beregning af fair strategier for investering og udbetaling af overskud. Master's thesis, Laboratory of Actuarial Mathematics, University of Copenhagen, 2001.

[22] P. Embrechts and S. Meister. Pricing insurance derivatives, the case of CAT-futures. *Proceedings of the Bowles Symposium on Securitization of Risk*, 1995.

[23] W. H. Fleming and R. W. Rishel. *Deterministic and Stochastic Optimal Control*. Springer-Verlag, 1975.

[24] W. H. Fleming and H. M. Soner. *Controlled Markov Processes and Viscosity Solutions*. Springer-Verlag, 1993.

[25] H. U. Gerber. *Life Insurance Mathematics*. Springer-Verlag, 1990.

[26] H. U. Gerber and G. Pafumi. Utility functions: from risk theory to finance. *North American Actuarial Journal*, 2(3):74–100, 1998.

[27] A. Grosen and P. L. Jørgensen. Valuation of early exercisable interest rate guarantees. *Journal of Risk and Insurance*, 64(3):481–503, 1997.

[28] A. Grosen and P. L. Jørgensen. Fair valuation of life insurance liabilities: The impact of interest rate guarantees, surrender options, and bonus policies. *Insurance: Mathematics and Economics*, 26:37–57, 2000.

[29] A. Grosen and P. L. Jørgensen. Life insurance liabilities at market values: An analysis of insolvency risk, bonus policy, and regulatory intervention rules in a barrier option framework. Technical report, Department of Management, University of Aarhus, 2000.

[30] M. S. Hansen. Optimal portfolio policies and the bonus option in a pension fund. Technical report, Department of Finance, University of Odense, 2001.

[31] J. M. Harrison and D. M. Kreps. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20:381–408, 1979.

[32] J. M. Harrison and S. R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, 11:215–260, 1981.

[33] J. M. Hoem. Markov chain models in life insurance. *Blätter der Deutschen Gesellschaft für Versicherungsmathematik*, 9:91–107, 1969.

[34] B. Højgaard and M. Taksar. Controlling risk exposure and dividend pay-out schemes: Insurance company example. *Mathematical Finance*, 9(2), 1999.

[35] P. Jaillet, D. Lamberton, and B. Lapeyre. Variational inequalities and the pricing of American options. *Acta Applicandae Mathematicae*, 21:263–289, 1990.

[36] I. Karatzas and S. E. Shreve. Connections between optimal stopping and singular stochastc control I. Monotone follower problems. *SIAM Journal of Control and Optimization*, 22(6):856–877, 1984.

[37] I. Karatzas and S. E. Shreve. Connections between optimal stopping and singular stochastc control II. Reflected follower problems. *SIAM Journal of Control and Optimization*, 23(3):433–451, 1985.

[38] S. Karlin and H. M. Taylor. *A Second Course in Stochastic Processes*. Academic Press, 1981.

[39] Y. Kifer. Game options. *Finance and Stochastics*, 4(4):443–463, 2000.

[40] R. Korn. *Optimal Portfolios*. World Scientific, 1997.

[41] D. Lamberton and B. Lapeyre. *Introduction to Stochastic Calculus Applied to Finance*. Chapman and Hall, 1996.

[42] R. C. Merton. Lifetime portfolio selection under uncertainty: The continuous time case. *Review of Economics and Statistics*, 51:247–257, 1969.

[43] R. C. Merton. Optimum consumption and portfolio rules in a continuous time model. *Journal of Economic Theory*, 3:373–413; Erratum 6 (1973); 213–214, 1971.

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

[45] R. C. Merton. *Continuous-time Finance*. Blackwell, 1990.

[46] K. Miltersen and S.-A. Persson. Guaranteed investment contracts: Distributed and undistributed excess return. Technical report, Department of Finance, University of Odense, 1998.

[47] C. M. Møller. Aspects of prospective mean values in risk theory. *Insurance: Mathematics and Economics*, 18:173–181, 1996.

[48] T. Møller. Quadratic hedging approaches and indifference pricing in insurance. Ph.D. thesis, Laboratory of Actuarial Mathematics, University of Copenhagen, 2000.

[49] J. A. Nielsen and K. Sandmann. Equity-linked life insurance: A model with stochastic interest rates. *Insurance: Mathematics and Economics*, 16:225–253, 1995.

[50] R. Norberg. Reserves in life and pension insurance. *Scandinavian Actuarial Journal*, pages 3–24, 1991.

[51] R. Norberg. A time-continuous Markov chain interest model with applications to insurance. *Journal of Applied Stochastic Models and Data Analysis*, 11:245–256, 1995.

[52] R. Norberg. Addendum to Hattendorff's theorem and Thiele's differential equation generalized, SAJ 1992, 2-14. *Scandinavian Actuarial Journal*, pages 50–53, 1996.

[53] R. Norberg. A theory of bonus in life insurance. *Finance and Stochastics*, 3(4):373–390, 1999.

[54] R. Norberg. Basic life insurance mathematics. Lecture notes, Laboratory of Actuarial Mathematics, University of Copenhagen, 2000.

[55] T. O'Brien. A stochastic-dynamic approach to pension funding. *Insurance: Mathematics and Economics*, 5:141–146, 1986.

[56] T. O'Brien. A two-parameter family of pension contribution functions and stochastic optimization. *Insurance: Mathematics and Economics*, 6:129–134, 1987.

[57] B. Øksendal. *Stochastic Differential Equations*. Springer-Verlag, 1995.

[58] H. Ramlau-Hansen. The emergence of profit in life insurance. *Insurance: Mathematics and Economics*, 7:225–236, 1988.

[59] H. Ramlau-Hansen. Distribution of surplus in life insurance. *ASTIN Bulletin*, 21:57–71, 1991.

[60] M. Schweizer. From actuarial to financial pricing principles. Technical report, Technische Universität Berlin, 1998.

[61] M. Schweizer. A guided tour through quadratic hedging approaches. Technical report, Technische Universität Berlin, 1999.

[62] W. F. Sharpe. Corporate pension funding policy. *Journal of Financial Economics*, 3:183–193, 1976.

[63] D. Sondermann. Reinsurance in arbitrage free markets. *Insurance: Mathematics and Economics*, (10):191–202, 1991.

[64] M. Steffensen. An approach to unit-linked insurance risk. Master's thesis, Laboratory of Actuarial Mathematics, University of Copenhagen, 1997.

[65] M. Steffensen. Contingent claims analysis in life and pension insurance. *Proceedings AFIR 2000*, pages 587–603, 2000.

[66] M. Steffensen. A no arbitrage approach to Thiele's differential equation. *Insurance: Mathematics and Economics*, (27):201–214, 2000.

[67] E. Sverdrup. Noen forsikringsmatematiske emner. Statistical Memoirs no. 1, Institute of Mathematics, University of Oslo, 1969.

[68] G. Taylor. Stochastic control of funding systems. Technical report, Centre for Actuarial Studies, University of Melbourne, 2000.