# A joint analysis of financial and biometrical risks in life insurance

Marcus C. Christiansen

# A joint analysis of financial and biometrical risks in life insurance

Dissertation

zur

Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. nat.)

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität Rostock

vorgelegt von

Marcus Christian Christiansen

geboren am 13. Oktober 1978 in Michigan (Vereinigte Staaten)

1. Gutachter: Prof. Dr. Hartmut Milbrodt, Universität Rostock
2. Gutachter: Prof. Dr. Holger Drees, Universität Hamburg
3. Gutachter: PD Dr. Ulrich Orbanz, Towers Perrin Tillinghast

Tag der Verteidigung: 17. Januar 2007

# Contents

# Introduction

Though the actuary of the 20th century used to model the lifespan of an insured stochastically, he usually relied on a deterministic prognosis of interest rates and mortality probabilities, denoted as 'actuarial assumptions' or 'technical basis'. The past has shown that these assumptions can vary significantly within a contract period. Especially in recent years financial markets have experienced increased volatility, and life expectancies have risen in many developed countries with an unforeseen rate. As changes of the technical basis can have a crucial effect on profits and losses, the actuary of the 21st century is well advised to pay attention to the financial and the systematic biometrical risk, in particular the longevity risk. This need is also reflected in the International Financial Reporting Standard (IFRS) No. 4 of the International Accounting Standards Board (IASB, 2004).

Already in 1905 Lidstone studied the effect of interest rate and mortality rate changes on premium values. Since then his sensitivity analysis concept has been improved by various authors leading to a number of insights. A younger and very capable approach of sensitivity analysis in life insurance is to study partial derivatives with respect to the parameters of the technical basis. Classically this works for (actuarial) functionals with an infinite-dimensional domain. However, if one wants to employ that sensitivity analysis concept for continuous time models where the technical basis parameters are functions on the real line, some kind of generalized gradient for functionals on function spaces is needed. This paper presents a proper approach, which will turn out to have some similarities to concepts of robust or nonparametric statistics. Applying the generalized sensitivity analysis concept to typical life insurance contracts will show that not only a change of interest rates but also of mortality or disability rates can have a significant influence on profits or losses.

Though a sensitivity analysis is a helpful approach for studying risks, it does not take into account the diversity of the volatilities of the technical basis parameters, neither in respect it different rates (e.g., interest rate, mortality rate, etc.) nor in respect to time (e.g., interest rate at different time points). For an exhaustive risk study it is therefore inevitable to model the technical basis stochastically.

Regarding the financial risk, the thriving development of financial theory during the last decades has inspired many actuaries to model capital profits in life insurances stochastically, too. It proved to be fruitful to adopt techniques and insights of financial mathematics for actuarial tasks such as valuation or pricing. Today the literature offers quite a number of life insurance models with stochastic interest rates.

In contrast, the systematic biometrical risks were for a long time widely ignored. Hoem (1988, p. 192), for example, recommended to cover the systematic mortality risk by a generous security loading on the interest rate. Lately, the unforeseen increase of life expectancies directed the attention to the systematic mortality risk. In the meantime, the literature offers several life insurance models with stochastic mortality rates.

For quantifying and comparing the financial and biometrical risks, the overall uncertainty of actuarial quantities such as present benefit/premium value has to be decomposed with regard to its different sources. Though the literature offers several approaches for studying financial and biometrical risks, it is up to now lacking a concept that allows for (a) quantifying financial (interest rate) risk, unsystematic biometrical risk, systematic mortality risk, systematic disability risk, et cetera simultaneously (b) with risk measures which are comparable to each other (c) for a wide variety of life insurance contract types. This paper wants to fill that gap, presenting some uncertainty analysis concept.

Applying that concept to typical life insurance contracts will show that even though the financial (interest rate) risk is largely predominant over the systematic biometrical risks, the latter are in many cases of significant size.

The structure of this paper is as follows:

After a short overview over the types of risk considered here, chapter 1 introduces the life insurance model of Milbrodt and Helbig (1999), which is one of the most general modeling frameworks of individual contracts in life insurance and includes both discrete time and continuous time approaches. Section 1.3 expands this model on a very general level to a stochastic technical basis, which will be further specified in chapter 4. Section 1.4 addresses the task of decomposing the overall randomness to its different sources.

Modeling the compounding factor and the transition probabilities as functions on the real line, the prospective reserve and the premium level have as mappings of the technical basis a function space as domain. For this reason it is not possible to perform a classical sensitivity analysis on them by just calculating their partial derivatives. Chapter 2 presents a new concept for a sensitivity analysis of functionals on specific function spaces by introducing a kind of generalized gradient vector. The latter has some similarities with concepts in robust statistics and nonparametric statistics; a comparison is given in section 2.3.

Using the tools of chapter 2, chapter 3 performs a sensitivity analysis on the actuarial functionals 'prospective reserve' and 'premium level' as mappings of the technical basis. Several realistic examples in section 3.5 exemplify the capability of the introduced concept. An empirical study of the basic life insurance contract types 'annuity insurance', 'pure endowment insurance', 'temporary life insurance', 'disability insurance', and their combinations yields valuable hints for risk management.

Based on the preliminary work of section 1.3, in chapter 4 the technical basis is modeled stochastically by assuming that the interest rate and the transition intensities are linear combinations of diffusion processes. Section 4.2 enhances the risk decomposition of section 1.4, and allows one to separate the financial risk, the unsystematic biometrical risk, and the systematic biometrical risks such as systematic mortality risk or systematic disability risk. In section 4.3, the approach of section

4.2 is applied to examples of of typical life insurance contracts. An empirical study leads to the following insights: Contrary to the statement of Hoem (1988, p. 192), the systematic mortality risk can be of great importance, especially for temporary life insurances. For annuity insurances it plays a smaller role but is still not negligible. For disability insurances the systematic disability risk is as well of significant size, having about the same dimension as the financial risk. Further on typical combinations of basic insurance contracts are studied with regard to lowering the technical basis risk.

# 1 Financial and biometrical risks in life insurance

This chapter presents the modeling frameworks for the study of financial and biometrical risks. After a short classification of risks in section 1.1, section 1.2 introduces 'classical' life insurance modeling, where the lifespan of an insured is stochastic but interest rates and life tables are deterministic. Section 1.3 extends this classical model in order to allow for a stochastic technical basis. This is done on a very general level, including many models to be found in the literature. It will be further specified in chapter 4. Section 1.4 addresses the task of identifying the contributions different risk sources make to the overall risk.

## 1.1 Classification of risks

The modeling framework of section 1.3 will contain two sources of risk:

(a) A *financial risk* brought into the model by a stochastic compounding factor, which is linked to the general economical situation. The term financial risk is here synonymous to 'interest rate risk'.

(b) A *biometrical risk* brought into the model by a stochastic jump process, which represents the biography of an individual insured.

Despite these two sources there are many other uncertainties an insurer is faced with, for example, the regulatory policy or the administration costs. However, these circumstances are not taken into consideration here.

Following Dahl (2004) and Cairns et al. (2005) the biometrical risk splits into

(b1) the *unsystematic biometrical risk* referring to the randomness of the biography for given transition probabilities (random fluctuations around expected values), and

(b2) the *systematic biometrical risk* referring to the uncertain future development of the underlying transition probabilities (systematic deviations of observed values from expected ones).

The distinction of unsystematic and systematic biometrical risk is convenient, since the former is largely diversifiable under the usual assumption that the biographies of different individuals in a portfolio are independent random variables, whereas the systematic biometrical risk is non-diversifiable. In contrast, a distinction of unsystematic and systematic financial risk is not implemented here as both are non-diversifiable. The financial risk (a) and the systematic biometrical risk (b2) correspond to the uncertainty of the technical basis, hence denote them summarized as *technical basis risk*.

Some authors propose a more subtle risk classification, introducing an *estimation risk* referring to misguided estimations caused by random fluctuations in the data (cf. Farny (1995), section 1553). This kind of risk is disregarded here. Further note that the unsystematic biometrical risk is here only studied for an individual contract and not for a portfolio of contracts.

The aim of this paper is to quantify the risks (a), (b1), and (b2) and to compare them with each other. 'Comparing of risks' is a topic frequently met in actuarial literature, but in many cases the intention differs from the approach presented here: for example, Kaas et al. (2003, chapter 10) or Müller and Stoyan (2002, chapter 8) assume the different risks are available as separate random variables, and the task is to find a stochastic order for comparing them. In contrast, the starting point is here just one random variable – the present value of future payments (see (1.2.5) and (1.3.7)) or functions of it – and the task is first of all to isolate the different risks from this one quantity.

## 1.2 Life insurance model

In this section, 'classical' life insurance modeling for individual contracts is introduced, where classical means that the actuarial assumptions 'compounding factor' and 'transition probabilities' are deterministic.

The model presented here follows the outline of Milbrodt and Helbig (1999) and differs from it mainly in postulating weaker requirements for the compounding factor. In contrast to a wide range of actuarial literature, the approach of Milbrodt and Helbig combines continuous time and discrete time approaches.

Consider a life insurance policy whose random state shall be represented by a Markovian jump process $((X_t, \mathfrak{A}_t))_{t \geq 0}$ on a finite state space $\mathcal{S}$. Denote by $J \subset \mathcal{S} \times \mathcal{S}$ the set of possible direct transitions. With $\mathcal{S}$ being finite, there exists a *transition probability matrix* $p$,

$$p_{yz}(s,t) = P(X_t = z | X_s = y)\,, \qquad (y,z) \in \mathcal{S}^2, s \leq t\,,$$

for which the Chapman-Kolmogorov equations hold (cf. Proposition 4.21 and Definition 4.18 in Milbrodt and Helbig (1999)). According to Milbrodt and Helbig (1999, Definition 4.28), it corresponds to a so-called *cumulative transition intensity matrix* $q$ – but differing from them use a one parameter notation, $q(t) := q(0,t)$. (The equation $q(s,t) := q(t) - q(s)$ translates the results presented here back into the notation of Milbrodt and Helbig.)

*Let the cumulative transition intensity matrix* $q$ *be regular,* (1.2.1) where regularity is defined in accordance with Definition 4.30 in Milbrodt and Helbig (1999). This allows for calculating $p$ out of $q$ via product-integration[1]

$$p(s,t) = \prod_{(s,t]} (\mathbb{I} + \mathrm{d}q) \in [0,1]^{|\mathcal{S}|\times|\mathcal{S}|}, \qquad 0 \le s < t < \infty. \tag{1.2.2}$$

Payments between insurer and policyholder are of two types:

(a) Lump sums are payable upon a transition $(y,z) \in J$ between two states. The amounts falling due are specified by nonnegative functions $D_{yz} \in BVC$ (*B*ounded *V*ariation on *C*ompacta, see section A.2). The actual payment date may differ from the time of transition. The monotone nondecreasing function $DT : [0,\infty) \to [0,\infty)$, $DT(t) \ge t$ for all $t \ge 0$, specifies that difference.

(b) Annuity payments fall due during sojourns in a state, modeled in a cumulative manner via functions $F_z \in BVC_{\leftarrow}$, $z \in \mathcal{S}$. (For the definition of $BVC_{\leftarrow}$ see section A.2.) If the policy stays during the time interval $(s,t]$ in state $z$, then $F_z(t) - F_z(s)$ is the accumulated amount falling due. Benefits paid to the insured get a positive sign, premiums paid by the insured get a negative sign.

Differing from Milbrodt and Helbig (1999) the functions $D_{yz}$ are not only Borel-measurable but also have finite variation on compacts. This restriction does not matter in practice.

The value of a payment depends on the payment date. This interrelation is specified by the *accumulation factor* $K : [0,\infty) \to (0,\infty)$. (A payment of one at time zero has at time $t$ the value $K(t)$.) The function $K$ shall be representable by the product-integral

$$K(t) = K_{\Phi}(t) = \prod_{(0,t]} (1 + \mathrm{d}\Phi), \tag{1.2.3}$$

where $\Phi$ is the so-called *cumulative interest intensity* satisfying

$$\Phi \in BVC_{\leftarrow}, \quad \Delta\Phi(t) := \Phi(t) - \Phi(t-0) \ge C_{\Phi} > -1 \quad \forall\, t \in \mathbb{R}. \tag{1.2.4}$$

**Remark 1.2.1.** Like many other authors Milbrodt and Helbig (1999), let the accumulation factor $K$ be nondecreasing. In contrast, Norberg and Møller (1996, p. 45) argue that a decreasing of the accumulation factor "is credible in a world where insurance companies may suffer losses on their investments". For this reason, and to allow for analyzing variations of the interest in both positive and negative directions, the weaker condition (1.2.4), when compared to Milbrodt and Helbig, is claimed here.

Replacing the second part in (1.2.4) with the even weaker condition $\Delta\Phi(t) > -1$ would be sufficient to ensure that the accumulation factor $K$ is positive, which is essential to the existence of the discounting factor $t \mapsto 1/K(t)$. Nevertheless the stronger condition $\Delta\Phi(t) \ge C_{\Phi} > -1$ is postulated: If there is no lower bound $C_{\Phi} > -1$, the accumulation factor could drop arbitrarily low by just one jump. Then, the sensitivity with respect to interest is unlimited, which conflicts with boundedness required later on.

[1] For some introduction to product-integration see Gill (1994) or Gill and Johansen (1990).

**Prospective reserve**

Let $DB_s$ be the present value at time $s$ of all transition benefits that are triggered (strictly) after $s$, and let $SB_s$ be the present value at time $s$ of all payments falling due during sojourns in a state due at and after time $s$. Write

$$B_s := SB_s + DB_s \tag{1.2.5}$$

for the (total) present value at time $s$ (cf. Milbrodt and Helbig (1999), p. 435). If $DT(t) \equiv t$, the time when a benefit is triggered coincides with the time when a benefit is payable. Otherwise denote by $DB_{s,u}$, $s < u$, the present value at time $s$ of all transition benefits triggered (strictly) after time $s$ and payable at and after time $u$. Further on, let $SB_{s,u}$ be the present value at time $s$ of all payments during sojourns in a state payable at and after time $u$. Write $B_{s,u} := SB_{s,u} + DB_{s,u}$.

Assume that the integrability conditions

$$\mathrm{InCo_{SB}} := \sum_{z \in \mathcal{S}} \int_{[0,\infty)} \frac{1}{K(t)} \, |F_z|(\mathrm{d}t) < \infty \,, \tag{1.2.6}$$

$$\mathrm{InCo_{DB}} := \sum_{(z,\zeta) \in J} \int_{(0,\infty)} \frac{1}{K(DT(t))} \, D_{z\zeta}(t) \, q_{z\zeta}(\mathrm{d}t) < \infty \tag{1.2.7}$$

hold (cf. (10.2.1) and (10.4.2) in Milbrodt and Helbig (1999)). Under these conditions denote

$$V_{y,s} := \mathbb{E}\big(B_s \big| \, X_s = y\big) \tag{1.2.8}$$

as the *prospective reserve at time* $s \geq 0$ *in state* $y \in \mathcal{S}$, and write for $s \leq u$

$$V_{y,s,u} := \mathbb{E}\big(B_{s,u} \big| \, X_s = y\big) \,. \tag{1.2.9}$$

Another useful quantity is the conditional variance of the present value $B_s$ for given initial state $X_s \in \mathcal{S}$. Define

$$\mathcal{V}_{y,s} := \mathrm{Var}\big(B_s \big| \, X_s = y\big) \,, \qquad s \geq 0 \,, y \in \mathcal{S} \,, \tag{1.2.10}$$

provided the conditional variance exists. Following Milbrodt and Helbig (1999, Proposition 10.4) the prospective reserve is representable as follows:

**Proposition 1.2.2.** *Under integrability conditions* (1.2.6) *and* (1.2.7)

$$\begin{aligned} V_{y,s} \overset{a.s.}{=} & \sum_{z \in \mathcal{S}} \int_{[s,\infty)} \frac{K(s)}{K(t)} \, p_{yz}(s,t) \, F_z(\mathrm{d}t) \\ & + \sum_{(z,\zeta) \in J} \int_{(s,\infty)} \frac{K(s)}{K(DT(t))} \, D_{z\zeta}(t) \, p_{yz}(s, t-0) \, q_{z\zeta}(\mathrm{d}t) \,, \end{aligned} \tag{1.2.11}$$

*for all* $s \geq 0$ *and* $y \in \mathcal{S}$*. For any* $y \in \mathcal{S}$ *the right hand side is, as a function of reference time* $s$*, bounded on compacts.*

*Proof.* According to Proposition A.4.1, the accumulation factor $K$ (see (1.2.3)) has finite variation on compacts. Hence, it is representable as a difference of two monotonic functions (see Theorem A.2.1) and thus Borel-measurable. The further proof is analogous to that of Proposition 10.4 in Milbrodt and Helbig (1999). □

From now on let the expression '$V_{y,s}$' always denote the representative (1.2.11). Whenever the specification of the actuarial assumptions $\Phi$ and $q$ plays an important role, write

$$V_{y,s}[\Phi,q] := \text{'representative (1.2.11) of the prospective reserve with actuarial assumptions } \Phi \text{ and } q\text{'}. \tag{1.2.12}$$

Similar to Proposition 1.2.2 one can show that for all $0 \le s < u$ and $y \in \mathcal{S}$

$$\begin{aligned} V_{y,s,u} &\overset{\text{a.s.}}{=} \sum_{z\in\mathcal{S}} \int_{[s,\infty)} \mathbf{1}_{[u,\infty)}(t)\, \frac{K(s)}{K(t)}\, p_{yz}(s,t)\, F_z(\mathrm{d}t) \\ &\quad + \sum_{(z,\zeta)\in J} \int_{(s,\infty)} \mathbf{1}_{\{r\,|\,u\le DT(r)\}}(t)\, \frac{K(s)}{K(DT(t))}\, D_{z\zeta}(t)\, p_{yz}(s,t-0)\, q_{z\zeta}(\mathrm{d}t)\,. \end{aligned} \tag{1.2.13}$$

Applying Corollary 10.39 in Milbrodt and Helbig (1999) yields also some representation for the conditional variance:

**Proposition 1.2.3.** *Suppose the integrability conditions*

$$\sum_{y\in\mathcal{S}} \sum_{(z,\zeta)\in J} \int_{(0,\infty)} \int_{[t,\infty)} \frac{1}{K(\tau)}\, |F_y|(\mathrm{d}\tau)\, q_{z\zeta}(\mathrm{d}t) < \infty\,,$$

$$\sum_{(\zeta,\eta)\in J} \sum_{(y,z)\in J} \int_{(0,\infty)} \int_{(t,\infty)} \frac{1}{K(DT(\tau))}\, D_{yz}(\tau)\, q_{yz}(\mathrm{d}\tau)\, q_{\zeta\eta}(\mathrm{d}t) < \infty$$

*hold. For* $s \ge 0$

$$\begin{aligned} \mathcal{V}_{y,s} &\overset{a.s.}{=} \sum_{(z,\zeta)\in J} \int_{(s,\infty)} \frac{K(s)^2}{K(t)^2} \big(R_{z\zeta}(t)\big)^2\, p_{yz}(s,t-0)\, q_{z\zeta}(\mathrm{d}t) \\ &\quad - \sum_{z\in\mathcal{S}} \sum_{\zeta,\eta\in\mathcal{S}\setminus\{z\}} \int_{(s,\infty)} \frac{K(s)^2}{K(t)^2}\, R_{z\zeta}(t)\, R_{z\eta}(t)\, p_{yz}(s,t-0)\, q_{z\zeta}(\{t\})\, q_{z\eta}(\mathrm{d}t)\,, \end{aligned} \tag{1.2.14}$$

*where* $R_{z\zeta}(t) := V_{\zeta,t} - V_{z,t} + D_{z\zeta}(t)$ *for* $t \in \mathbb{R}$ *is the so-called 'sum at risk'.*

*Proof.* In case of $s = 0$, equation (1.2.14) is a consequence of Corollary 10.39 in Milbrodt and Helbig (1999). (The present value $B_0$ is equal to their loss function at $t = \infty$.) For $s > 0$ the proof is analogous. □

Analogously to the prospective reserve, let from now on the expression '$\mathcal{V}_{y,s}$' always denote the representative (1.2.14). Whenever the specification of the actuarial assumptions $\Phi$ and $q$ plays an important role, write

$\mathcal{V}_{y,s}[\Phi, q] :=$ 'representative (1.2.14) of (1.2.10) with actuarial assumptions $\Phi$ and $q$'.

In case the cumulative transition intensity matrix $q$ is continuous, the right hand side of (1.2.14) reduces to

$$\mathcal{V}_{y,s}[\Phi, q] = \sum_{(z,\zeta)\in J} \int_{(s,\infty)} \frac{K(s)^2}{K(t)^2} \big(R_{z\zeta}(t)\big)^2 \, p_{yz}(s, t-0)\, q_{z\zeta}(\mathrm{d}t)\,. \tag{1.2.15}$$

#### Calculation of premiums

The prospective reserve $V_{y,s}$ is the mean present value of all future benefits (with respect to time $s$) minus all future premiums, assuming that the policy is at present in state $y \in \mathcal{S}$. This is roughly the amount an insurer needs on the average to meet all future obligations. Therefore, the so-called *equivalence principle* states that a fair contract satisfies the condition

$$V_{a,0} \stackrel{!}{=} 0\,, \qquad \text{where } X_0 = a \text{ is the } \textit{initial state},$$

that is, the mean present value of all benefits is equal to the mean present value of all premiums at the beginning $s = 0$ of the contract. In fact, this approach is widely used to calculate *net premiums*. Typically the procedure is as follows:

First, the desired benefits (e.g., death grant, disability pension, etc.) are selected and a premium scheme (e.g., monthly rates, single lump sum, etc.) is chosen, whose mean present values are here denoted by $VB_{a,0}$ and $VE_{a,0}$, respectively. Because of the linearity of the prospective reserve regarding the benefit and premium payments, the overall mean present value is $V_{a,0} = VB_{a,0} + C \cdot VE_{a,0}$. Then the *premium level* $C \in \mathbb{R}$ is determined in such a way that the equivalence requirement

$$V_{a,0} = VB_{a,0} + C \cdot VE_{a,0} \stackrel{!}{=} 0 \tag{1.2.16}$$

is met (note that benefits get a positive sign and premiums get a negative sign), which is equivalent to

$$C = \frac{-VB_{a,0}}{VE_{a,0}}\,. \tag{1.2.17}$$

## 1.3 Expanded life insurance model

Now the actuarial assumptions 'cumulative interest intensity' and 'cumulative transition intensity matrix' are modeled as stochastic processes. The general framework presented here includes various other models of the literature (cf. Example 1.3.5), but it is to be noted that only an individual insurance contract is modeled here, not a portfolio of contracts.

**Condition 1.3.1.** (a) Let the cumulative interest intensity be a stochastic process $(\Phi_t)_{t\geq 0}$ on $(\Omega, \mathfrak{A}, P)$ with paths $t \mapsto \Phi_t(\omega)$ in

$$\mathcal{F} \subset \left\{ F \in BVC_{\leftarrow} \,\middle|\, F \text{ satisfies (1.2.4)} \right\}. \tag{1.3.1}$$

Defining $\mathfrak{F} := \mathcal{F} \cap (\mathfrak{B}(\mathbb{R}))^{[0,\infty)}$, the process $(\Phi_t)$ as a whole is $(\Omega, \mathfrak{A}, P)$-$(\mathcal{F}, \mathfrak{F})$-measurable (cf. Bauer (1992), Theorem 7.4).

(b) Let the cumulative transition intensity matrix be a stochastic process $(q_t)_{t\geq 0}$ on $(\Omega, \mathfrak{A}, P)$ with paths $t \mapsto q_t(\omega)$ in

$$\mathcal{Q} \subset \left\{ Q \in (BVC_{\leftarrow})^{|\mathcal{S}|\times|\mathcal{S}|} \,\middle|\, Q \text{ is a regular cumulative intensity matrix} \right\}. \tag{1.3.2}$$

Defining $\mathfrak{Q} := \mathcal{Q} \cap (\mathfrak{B}(\mathbb{R}))^{[0,\infty)}$, the process $(q_t)$ as a whole is $(\Omega, \mathfrak{A}, P)$-$(\mathcal{Q}, \mathfrak{Q})$-measurable (cf. Bauer (1992), Theorem 7.4).

According to Milbrodt and Helbig (1999, Theorem 4.35), for any regular cumulative transition intensity matrix $Q \in \mathcal{Q}$ and initial distribution $\pi = \mathcal{L}(X_0^Q|P)$ there exists a corresponding Markovian jump process

$$(X_t^Q)_{t\geq 0} : (\Omega, \mathfrak{A}, P) \to (\mathcal{X}, \mathfrak{X}),$$

where $\mathcal{X} \subset \mathcal{S}^{[0,\infty)}$ is a set of right-continuous trajectories with, at most, a finite number of jumps on any compact interval and $\mathfrak{X} := \mathcal{X} \cap (2^{\mathcal{S}})^{[0,\infty)}$ (cf. Milbrodt and Helbig (1999), Definition 4.1). For an arbitrary but fixed initial distribution $\pi$ define

$$K(Q, \cdot) := \mathcal{L}\big((X_t^Q) \,\big|\, P\big), \quad Q \in \mathcal{Q}. \tag{1.3.3}$$

**Proposition 1.3.2.** *The mapping* $K : (\mathcal{Q}, \mathfrak{X}) \mapsto \mathbb{R}$, $(Q, A) \mapsto K(Q, A)$ *is a Markov kernel from* $(\mathcal{Q}, \mathfrak{Q})$ *to* $(\mathcal{X}, \mathfrak{X})$.

(The second half of the following proof is due to a personal comment of F. Liese, University of Rostock, Department of Mathematics.)

*Proof.* At first, it is shown that for each $y, z \in \mathcal{S}$ and $s, t \in \mathbb{R}$, $s \leq t$, the mapping

$$p_{yz}^{(\cdot)}(s,t) : \mathcal{Q} \to \mathbb{R}, \quad Q \mapsto p_{yz}^{Q}(s,t) := \left[ \prod_{(s,t]} \big(\mathbb{I} + \mathrm{d}Q\big) \right]_{yz} \tag{1.3.4}$$

is $(\mathcal{Q}, \mathfrak{Q})$-$(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$-measurable: By means of Theorem 2 in section 2 in Gill (1994)

$$\left[ \prod_{(s,t]} \Big(\mathbb{I} + \mathrm{d}Q\Big) \right]_{yz} = \left[ \lim_{n\to\infty} \prod_{\mathcal{T}_n} \Big(\mathbb{I} + Q(t_i) - Q(t_{i-1})\Big) \right]_{yz}, \qquad \forall\, (y,z) \in \mathcal{S}^2,$$

for any sequence $(\mathcal{T}_n)$ of interval decompositions $s \leq t_1 < \cdots < t_n \leq t$ satisfying $\lim_{n\to\infty} \max_{\mathcal{T}_n} |t_{i+1} - t_i| = 0$. Since the projection functions

$$\mathcal{Q} \ni Q \mapsto Q_{gl}(u) \in \mathbb{R}\,, \qquad u \in (s,t]\,,\ (g,l) \in \mathcal{S}^2,$$

are $(\mathcal{Q},\mathfrak{Q})$-$(\mathbb{R},\mathfrak{B}(\mathbb{R}))$-measurable, the mapping $p_{yz}^{(\cdot)}(s,t)$ is as well. As

$$P(X_{t_1}^Q \in A_1, \ldots, X_{t_n}^Q \in A_n) = \sum_{y_0\in\mathcal{S}} \sum_{y_1\in A_1} \cdots \sum_{y_n\in A_n} \pi(y_0)\, p_{y_0y_1}^Q(0,t_1) \cdot \ldots \cdot p_{y_{n-1}y_n}^Q(t_{n-1},t_n)$$

for $A_1,\ldots,A_n \in 2^{\mathcal{S}}$ and $t_1,\ldots,t_n \in [0,\infty)$, the finite dimensional marginal distributions of $(X_t^Q)$ are $(\mathcal{Q},\mathfrak{Q})$-$(\mathbb{R},\mathfrak{B}(\mathbb{R}))$-measurable, too.

Now the measurability of $K(\cdot,A)$ for each $A \in \mathfrak{X}$ is shown: The set

$$\widetilde{\mathfrak{A}} := \left\{ A \in \mathfrak{X} \,\middle|\, Q \mapsto K(Q,A) = P((X_t^Q) \in A) \text{ is } (\mathcal{Q},\mathfrak{Q})\text{-measurable} \right\}$$

is a Dynkin system, as (i) the mapping $Q \mapsto P\big((X_t^Q) \in \mathcal{X}\big) = 1$ is $(\mathcal{Q},\mathfrak{Q})$-measurable, (ii) the mapping $Q \mapsto P\big((X_t^Q) \in \mathcal{X} \setminus A\big) = 1 - P\big((X_t^Q) \in A\big)$ is $(\mathcal{Q},\mathfrak{Q})$-measurable for all $A \in \widetilde{\mathfrak{A}}$, and (iii) for each pairwise disjoint sequence $(D_n)_{n\in\mathbb{N}} \subset \widetilde{\mathfrak{A}}$ the mapping

$$Q \mapsto P\left((X_t^Q) \in \bigcup_{n=1}^{\infty} D_n\right) = \sum_{n=1}^{\infty} P\big((X_t^Q) \in D_n\big)$$

is $(\mathcal{Q},\mathfrak{Q})$-measurable. The set

$$\mathcal{C} := \bigcup_{n=1}^{\infty} (X_{t_1}^m, .., X_{t_n}^m)^{-1}((2^{\mathcal{S}})^n)$$

is contained in $\widetilde{\mathfrak{A}}$ since the finite dimensional marginal distributions of $(X_t^Q)$ are $(\mathcal{Q},\mathfrak{Q})$-measurable. For each pair $C, D \in \mathcal{C}$ the intersection $C \cap D$ is itself an element of $\mathcal{C}$. Now Theorem 2.4 in Bauer (1992) yields

$$\widetilde{\mathfrak{A}} \supseteq \delta(\mathcal{C}) = \sigma(\mathcal{C}) = \mathfrak{X}\,,$$

where $\delta(\mathcal{C})$ and $\sigma(\mathcal{C})$ denote the smallest Dynkin system and the smallest $\sigma$-algebra containing $\mathcal{C}$, respectively. □

This property of $K$ is the key for a consistent extension of the classical life insurance model to a model with a stochastic technical basis:

**Theorem 1.3.3.** *There exists a probability space $(\tilde{\Omega},\tilde{\mathfrak{A}},\tilde{P})$ and stochastic processes*

$$(\tilde{X}_t)_{t\geq 0} : (\tilde{\Omega},\tilde{\mathfrak{A}}) \to (\mathcal{X},\mathfrak{X})\,,$$
$$(\tilde{\Phi}_t)_{t\geq 0} : (\tilde{\Omega},\tilde{\mathfrak{A}}) \to (\mathcal{F},\mathfrak{F})\,,$$
$$(\tilde{q}_t)_{t\geq 0} : (\tilde{\Omega},\tilde{\mathfrak{A}}) \to (\mathcal{Q},\mathfrak{Q})$$

*such that $\mathcal{L}((\tilde{\Phi}_t)|\tilde{P}) = \mathcal{L}((\Phi_t)|P)$, $\mathcal{L}((\tilde{q}_t)|\tilde{P}) = \mathcal{L}((q_t)|P)$, and*

$$\tilde{P}\big((\tilde{X}_t) \in \cdot \,\big|\, (\tilde{q}_t) = Q\big) \overset{a.s.}{=} \mathcal{L}((X_t^Q)\,|\,P)\,, \qquad \forall\, Q \in \mathcal{Q}\,. \tag{1.3.5}$$

*Proof.* Let $\tilde{\Omega} := \mathcal{X} \times \mathcal{F} \times \mathcal{Q}$ be the domain of the new probability space. The projections $p_X : \tilde{\Omega} \ni (X, F, Q) \mapsto X \in \mathcal{X}$, $p_\Phi : \tilde{\Omega} \ni (X, F, Q) \mapsto F \in \mathcal{F}$, and $p_q : \tilde{\Omega} \ni (X, F, Q) \mapsto Q \in \mathcal{Q}$ are measurable on $\tilde{\mathfrak{A}} := \mathfrak{X} \otimes \mathfrak{F} \otimes \mathfrak{Q} = \sigma(p_X, p_\Phi, p_q)$. Define the new stochastic processes by $(\tilde{X}_t) := p_X$, $(\tilde{\Phi}_t) := p_\Phi$, and $(\tilde{q}_t) := p_q$. Proposition 1.3.2 allows for defining

$$\tilde{P}(A) := \iint \mathbf{1}_{(X,F,Q) \in A}\, K(Q, \mathrm{d}X)\, P_{((\Phi_t),(q_t))}(\mathrm{d}(F, Q))\,, \quad A \in \tilde{\mathfrak{A}}\,. \tag{1.3.6}$$

This mapping is a measure since $0 = \tilde{P}(\varnothing) \leq \tilde{P}(A) \leq \tilde{P}(\tilde{\Omega}) = \tilde{P}(\mathcal{X} \times \mathcal{F} \times \mathcal{Q}) = 1$ for all $A \in \tilde{\mathfrak{A}}$, and by means of the Monotone Convergence Theorem

$$\begin{aligned}
\tilde{P}\left(\bigcup_{n=1}^{\infty} A_n\right) &= \iint \sum_{n=1}^{\infty} \mathbf{1}_{(X,F,Q) \in A_n}\, K(Q, \mathrm{d}X)\, P_{((\Phi_t),(q_t))}(\mathrm{d}(F, Q)) \\
&= \sum_{n=1}^{\infty} \iint \mathbf{1}_{(X,F,Q) \in A_n}\, K(Q, \mathrm{d}X)\, P_{((\Phi_t),(q_t))}(\mathrm{d}(F, Q)) \\
&= \sum_{n=1}^{\infty} \tilde{P}(A_n)
\end{aligned}$$

for each sequence $(A_n)_{n \in \mathbb{N}} \subset \tilde{\mathfrak{A}}$ of pairwise disjoint sets. Because of

$$\begin{aligned}
\tilde{P}\big((\tilde{q}_t) \in A\big) &= \iint \mathbf{1}_{(X,F,Q) \in \mathcal{X} \times \mathcal{F} \times A}\, K(Q, \mathrm{d}X)\, P_{((\Phi_t),(q_t))}(\mathrm{d}(F, Q)) \\
&= \int \mathbf{1}_{(F,Q) \in \mathcal{F} \times A}\, P_{((\Phi_t),(q_t))}(\mathrm{d}(F, Q)) \\
&= P\big((q_t) \in A\big)
\end{aligned}$$

for all $A \in \mathfrak{Q}$ one has $\mathcal{L}((\tilde{q}_t)|\tilde{P}) = \mathcal{L}((q_t)|P)$. In the same way, one gets $\mathcal{L}((\tilde{\Phi}_t)|\tilde{P}) = \mathcal{L}((\Phi_t)|P)$. Equation (1.3.5) holds as

$$\begin{aligned}
\int_C P((X_t^Q) \in A)\, \tilde{P}_{(\tilde{q}_t)}(\mathrm{d}Q) &= \int_C K(Q, A)\, P_{(q_t)}(\mathrm{d}Q) \\
&= \iint \mathbf{1}_{(X,F,Q) \in A \times \mathcal{F} \times C}\, K(Q, \mathrm{d}X)\, P_{((\Phi_t),(q_t))}(\mathrm{d}(F, Q)) \\
&= \tilde{P}((\tilde{X}_t) \in A, (\tilde{q}_t) \in C)
\end{aligned}$$

for all $A \in \mathfrak{X}$ and $C \in \mathfrak{Q}$. □

Instead of $(X_t)_{t \geq 0}$ as defined in section 1.2, regard henceforth $(\tilde{X}_t)_{t \geq 0}$ as the biography of the insured. In case $(\tilde{\Phi}_t)_{t \geq 0}$ and $(\tilde{q}_t)_{t \geq 0}$ are deterministic, this new approach coincides with the classical model because of property (1.3.5). To simplify the notation write from now on $(\Phi_t)_{t \geq 0}$, $(q_t)_{t \geq 0}$, and $P$ instead of $(\tilde{\Phi}_t)_{t \geq 0}$, $(\tilde{q}_t)_{t \geq 0}$, and $\tilde{P}$.

**Remark 1.3.4.** The jump process $(\tilde{X}_t)_{t\geq 0}$ is not necessarily Markovian: Let $\mathcal{Q} = \{Q_a, Q_b\}$ with $(X_t^{Q_a})_{t\geq 0}$ being (almost sure) in the state space $\{0,1\}$ and with $(X_t^{Q_b})_{t\geq 0}$ being (almost sure) in the state space $\{1,2\}$. Assume that $P(X_3^{Q_a} = 0 \,|\, X_2^{Q_a} = 1) > 0$ and $0 < P((q_t) = Q_a) < 1$. Then,

$$\begin{aligned} 0 < P\big(\tilde{X}_3 = 0 \big| \tilde{X}_2 = 1\big) &\neq P\big(\tilde{X}_3 = 0 \big| \tilde{X}_2 = 1, \tilde{X}_1 = 2\big) \\ &= P\big(\tilde{X}_3 = 0, (q_t) = Q_a \big| \tilde{X}_2 = 1, \tilde{X}_1 = 2, (q_t) = Q_b\big) = 0\,, \end{aligned}$$

that is, $(\tilde{X}_t)_{t\geq 0}$ is not Markovian.

**Example 1.3.5.** As suggested by Norberg (1999), let the interest intensity $\varphi$ and the transition intensity matrix $\mu$ be controlled by a Markovian jump process $(Y_t)_{t\geq 0}$ with finite state space $\mathcal{R}$:

$$\varphi(t) = \sum_{e\in\mathcal{R}} \mathbf{1}_{Y_t=e}\, \varphi_e\,, \quad \mu_{jk}(t) = \sum_{e\in\mathcal{R}} \mathbf{1}_{Y_t=e}\, \mu_{jk;e}(t)\,,$$

where the $\varphi_e$ are constants and the $\mu_{jk;e}(t)$ are piecewise continuous intensity functions, all deterministic. Assume the process $(Y_t)_{t\geq 0}$ is homogenous with transition intensities $\lambda_{ef}$ for $e, f \in \mathcal{R}$. This implies that the trajectories have almost sure only a finite number of jumps in any compact interval. The corresponding cumulative intensities

$$\Phi(t) = \int_{(0,t]} \varphi(\tau)\,\mathrm{d}\tau\,, \qquad q_{jk}(t) = \int_{(0,t]} \mu_{jk}(\tau)\,\mathrm{d}\tau$$

satisfy Condition 1.3.1 (cf. Milbrodt and Helbig (1990), Exercise 20, pp. 197-198).

### Conditional expectation and conditional variance of the present value

Analogously to the present value $B_s$ in section 1.2, which has according to Proposition A.1.1 a representation of the form $B_s = b_s((X_t), \Phi)$, denote by $\tilde{B}_s$ the present value corresponding to the stochastic state trajectory $(\tilde{X}_t)_{t\geq 0}$ and the stochastic cumulative interest intensity $(\Phi_t)_{t\geq 0}$. Since the mapping $b_s$ is $(\mathcal{X} \times \mathcal{F}, \mathfrak{X} \otimes \mathfrak{F})$-$(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$-measurable, $\tilde{B}_s$ is representable by

$$\tilde{B}_s = b_s\big((\tilde{X}_t), (\Phi_t)\big)\,. \tag{1.3.7}$$

The two following theorems yield some form of 'insertion rule' for the conditional expectation and the conditional variance of $\tilde{B}_s$ for given actuarial assumptions $(\Phi, q) \in \mathcal{F} \times \mathcal{Q}$ and given initial state $\tilde{X}_s \in \mathcal{S}$.

**Theorem 1.3.6.** *For* $s \in [0,\infty)$, $F \in \mathcal{F}$, $Q \in \mathcal{Q}$, *and* $y \in \mathcal{S}$

$$\mathbb{E}\Big(b_s\big((X_t^Q), F\big) \,\Big|\, X_s^Q = y\Big) \overset{a.s.}{=} \mathbb{E}\Big(b_s\big((\tilde{X}_t), (\Phi_t)\big) \,\Big|\, \tilde{X}_s = y, (\Phi_t) = F, (q_t) = Q\Big)\,, \tag{1.3.8}$$

*provided the conditional expectations exist.*

The theorem will be proven with the help of the two following Propositions:

**Proposition 1.3.7.** *The mapping* $K_s : (\mathcal{S} \times \mathcal{Q}, \mathfrak{X}) \to \mathbb{R}$ ,

$$((y,Q),A) \mapsto \begin{cases} \dfrac{K\big(Q, A \cap \{X_s^Q = y\}\big)}{K(Q,\{X_s^Q = y\})} & : K(Q,\{X_s^Q = y\}) \neq 0\,, \\ K\big(Q, A\big) & : K(Q,\{X_s^Q = y\}) = 0\,, \end{cases} \tag{1.3.9}$$

*is for any fixed* $s \in [0,\infty)$ *a Markov kernel from* $(\mathcal{S}\times\mathcal{Q}, 2^{\mathcal{S}}\otimes\mathfrak{Q})$ *to* $(\mathcal{X},\mathfrak{X})$ *. Particularly*

$$K_s\big((y,Q),A\big) \overset{a.s.}{=} P\big((X_t^Q) \in A \big| X_s^Q = y\big)\,, \qquad \forall\, A \in \mathfrak{X}\,, y \in \mathcal{S}\,, Q \in \mathcal{Q}\,. \tag{1.3.10}$$

*Proof.* Since $\{\omega : X_s^Q(\omega) = y\} \in \mathfrak{X}$, the mapping $K_s$ is well defined. As $K(Q,\cdot)$ is – according to Proposition 1.3.2 – a probability measure for each $Q \in \mathcal{Q}$ , the mapping $K_s((y,Q),\cdot)$ is for each $s \in [0,\infty)$ , $y \in \mathcal{S}$ , and $Q \in \mathcal{Q}$ a probability measure as well.

Since for each $A \in \mathfrak{X}$ and $B \in \mathfrak{B}(\mathbb{R})$ Proposition 1.3.2 yields

$$\Big\{ Q \in \mathcal{Q} \Big| K\big(Q, A \cap \{X_s^Q = y\}\big) \in B \Big\} \in \mathfrak{Q}\,,$$

the set

$$\begin{aligned} &\Big\{ (y,Q) \in \mathcal{S} \times \mathcal{Q} \Big| K\big(Q, A \cap \{X_s^Q = y\}\big) \in B \Big\} \\ &= \bigcup_{y \in \mathcal{S}} \{y\} \times \Big\{ Q \in \mathcal{Q} \Big| K\big(Q, A \cap \{X_s^Q = y\}\big) \in B \Big\} \end{aligned}$$

is an element of $2^{\mathcal{S}} \otimes \mathfrak{Q}$. That means the mapping $(y,Q) \mapsto K\big(Q, A \cap \{X_s^Q = y\}\big)$ is $(S \times \mathcal{Q}, 2^{\mathcal{S}} \otimes \mathfrak{Q})$-$(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$-measurable for each $A \in \mathfrak{X}$, and consequently

$$\mathbf{1}_{K(Q,\{X_s^Q = y\}) \neq 0}(y,Q) \cdot \frac{K\big(Q, A \cap \{X_s^Q = y\}\big)}{K(Q,\{X_s^Q = y\})} + \mathbf{1}_{K(Q,\{X_s^Q = y\}) = 0}(y,Q) \cdot K(Q,A)$$

is $(S \times \mathcal{Q}, 2^{\mathcal{S}} \otimes \mathfrak{Q})$-$(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$-measurable, too. Equation (1.3.10) is equivalent to the Radon-Nikodym equation

$$\begin{aligned} \int_S K_s\big((y,Q),A\big)\, P(X_s^Q = \mathrm{d}y) &= \sum_{y \in S} K_s\big((y,Q),A\big)\, P(X_s^Q = y) \\ &= K(Q, A \cap \{X_s^Q \in S\}) \\ &= P\big((X_t^Q) \in A\,, X_s^Q \in S\big)\,, \qquad \forall\, S \in 2^{\mathcal{S}}\,. \end{aligned}$$

□

**Proposition 1.3.8.** *For each measurable function* $g : (\mathcal{X} \times \mathcal{F}, \mathfrak{X} \otimes \mathfrak{F}) \to \big(\mathbb{R}, \mathfrak{B}(\mathbb{R})\big)$ *and for any* $s \in [0,\infty)$ , $F \in \mathcal{F}$ , $Q \in \mathcal{Q}$ *, and* $y \in \mathcal{S}$

$$\mathbb{E}\Big( g\big((X_t^Q), F\big) \,\Big|\, X_s^Q = y \Big) \overset{a.s.}{=} \mathbb{E}\Big( g\big((\tilde{X}_t), (\Phi_t)\big) \,\Big|\, \tilde{X}_s = y\,,\, (\Phi_t) = F, (q_t) = Q \Big)\,, \tag{1.3.11}$$

*provided the conditional expectations exist.*

*Proof.* Because of Proposition 1.3.7, for any $F \in \mathcal{F}$, $Q \in \mathcal{Q}$, and $y \in \mathcal{S}$

$$\mathbb{E}\Big(g\big((X_t^Q), F\big)\Big|X_s^Q = y\Big) = \int g(X, F)\, P\big((X_t^Q) \in \mathrm{d}X \mid X_s^Q = y\big)$$
$$\overset{\text{a.s.}}{=} \int g(X, F)\, K_s\big((y, Q), \mathrm{d}X\big)\,.$$

With that, and by using (1.3.6) and (1.3.3) combined with (1.3.10) one gets

$$\begin{aligned}
&\int_A \mathbb{E}\Big(g\big((X_t^Q), F\big)\Big|X_s^Q = y\Big)\, P_{(\tilde{X}_s,(\Phi_t),(q_t))}(\mathrm{d}(y, F, Q)) \\
&= \iint \mathbf{1}_{(y,F,Q)\in A} \int g(X, F)\, K_s((y, Q), \mathrm{d}X)\, P_{(\tilde{X}_s,(\Phi_t),(q_t))}(\mathrm{d}(y, F, Q)) \\
&= \iint \mathbf{1}_{(y,F,Q)\in A} \int g(X, F)\, K_s((y, Q), \mathrm{d}X)\, K(Q, X_s^Q = \mathrm{d}y)\, P_{((\Phi_t),(q_t))}(\mathrm{d}(F, Q)) \\
&= \iint \mathbf{1}_{(X_s,F,Q)\in A}\, g(X, F)\, K(Q, \mathrm{d}X)\, P_{((\Phi_t),(q_t))}(\mathrm{d}(F, Q)) \\
&= \iint \mathbf{1}_{(X_s,F,Q)\in A}\, g(X, F)\, P_{((\tilde{X}_t),(\Phi_t),(q_t))}(\mathrm{d}(X, F, Q)) \\
&= \int_{(\tilde{X}_s,(\Phi_t),(q_t))\in A} g \circ \big((\tilde{X}_t), (\Phi_t)\big)\, \mathrm{d}P
\end{aligned}$$

for all $A \in 2^{\mathcal{S}} \otimes \mathfrak{F} \otimes \mathfrak{Q}$, where it is to be noted that $K$ and $K_s$ are Markov kernels. □

*Proof of Theorem* 1.3.6. Equation (1.3.8) is a consequence of Proposition 1.3.8 and Proposition A.1.1. □

**Theorem 1.3.9.** *For* $s \in [0, \infty)$, $F \in \mathcal{F}$, $Q \in \mathcal{Q}$, *and* $y \in \mathcal{S}$

$$\mathrm{Var}\Big(b_s\big((X_t^Q)), F\big)\,\Big|\, X_s^Q = y\Big) \overset{a.s.}{=} \mathrm{Var}\Big(b_s\big((\tilde{X}_t), (\Phi_t)\big)\,\Big|\, \tilde{X}_s = y, (\Phi_t) = F, (q_t) = Q\Big)\,, \tag{1.3.12}$$

*provided the conditional variances exist.*

*Proof.* Because of Proposition A.1.1 and Theorem 1.3.6 the mapping

$$g\big((X_t^Q), F\big) := \Big(b_s\big((X_t^Q), F\big) - \mathbb{E}\Big(b_s\big((X_t^Q), F\big)\,\Big|\, X_s^Q = y\Big)\Big)^2$$
$$\overset{\text{a.s.}}{=} \Big(b_s\big((X_t^Q), F\big) - \mathbb{E}\Big(b_s\big((\tilde{X}_t), (\Phi_t)\big)\,\Big|\, \tilde{X}_s = y, (\Phi_t) = F, (q_t) = Q\Big)\Big)^2$$

is $(\mathcal{X} \times \mathcal{F}, \mathfrak{X} \otimes \mathfrak{F})$-$\big(\mathbb{R}, \mathfrak{B}(\mathbb{R})\big)$-measurable. Now apply Proposition 1.3.8. □

## 1.4 Decomposition of risk

Unfortunately the risk contributions of the financial risk, the unsystematic biometrical risk, and the systematic biometrical risk are not on hand as separate random variables, but are only available in merged form in terms of the present value $\tilde{B}_s$. For studying the different risks separately the present value needs to be decomposed to separate components referring to the different risk contributions.

In a first step, the unsystematic biometrical risk is separated. Therefore, follow the concept in Fischer (2004, section 3.5) to obtain some orthogonal decomposition by means of conditional expectations. (Note that already Bühlmann (1992, 1995) used conditional expectations to decompose financial risks.) Building the conditional expectation

$$\mathbb{E}\Big(\tilde{B}_s \Big| (\Phi_t), (q_t)\Big) = \mathbb{E}\Big(b_s\big((\tilde{X}_t), (\Phi_t)\big) \Big| (\Phi_t), (q_t)\Big) \tag{1.4.1}$$

averages out the randomness of the biography $(\widetilde{X}_t)$, but keeps the uncertainty of the technical basis $((\Phi_t), (q_t))$. That means the unsystematic biometrical risk is eliminated but the financial and the systematic biometrical risk are still at hand. This motivates to decompose the present value $\tilde{B}_s$ to

$$\tilde{B}_s = \Big(\tilde{B}_s - \mathbb{E}\big(\tilde{B}_s \,\big|\, (\Phi_t), (q_t)\big)\Big) + \mathbb{E}\big(\tilde{B}_s \,\big|\, (\Phi_t), (q_t)\big)\,. \tag{1.4.2}$$

With the second addend corresponding to the technical basis risk, the first addend may be interpreted as representative for the unsystematic biometrical risk. (Fischer (2004) did not implement stochastic transition probabilities; that means his technical basis risk is only a financial risk, but apparently his concept also works for a technical basis with stochastic transition probabilities.)

Due to the 'projection property' of conditional expectations, the two addends of decomposition (1.4.2) are uncorrelated (provided their second moments exist). Calculating the variance on both hand sides leads to the following well-known variance decomposition:

**Proposition 1.4.1.** *Assume that Condition* 1.3.1 *holds and* $\mathbb{E}(\tilde{B}_s)^2 < \infty$. *Then,*

$$\mathrm{Var}\big(\tilde{B}_s\big) = \mathbb{E}\Big(\mathrm{Var}\big(\tilde{B}_s \big| (\Phi_t)\,, (q_t)\big)\Big) + \mathrm{Var}\Big(\mathbb{E}\big(\tilde{B}_s \big| (\Phi_t)\,, (q_t)\big)\Big)\,. \tag{1.4.3}$$

*Proof.* As the second moment of $\tilde{B}_s$ is finite, the same holds for the two addends of decomposition (1.4.2). Since they are uncorrelated, one gets

$$\mathrm{Var}\Big(\tilde{B}_s\Big) = \mathrm{Var}\Big(\tilde{B}_s - \mathbb{E}\big(\tilde{B}_s \,\big|\, (\Phi_t), (q_t)\big)\Big) + \mathrm{Var}\Big(\mathbb{E}\big(\tilde{B}_s \big| (\Phi_t)\,, (q_t)\big)\Big)\,. \tag{1.4.4}$$

The first addend on the right hand side is equal to

$$\mathbb{E}\Big(\tilde{B}_s - \mathbb{E}\big(\tilde{B}_s \,\big|\, (\Phi_t), (q_t)\big)\Big)^2 = \mathbb{E}\bigg(\mathbb{E}\bigg(\Big(\tilde{B}_s - \mathbb{E}\big(\tilde{B}_s \,\big|\, (\Phi_t), (q_t)\big)\Big)^2 \bigg| (\Phi_t), (q_t)\bigg)\bigg)\,.$$

The equality is a consequence of the linearity, the 'chain rule', and the 'pull-out property' of conditional expectations (cf. Kallenberg (1997), pp. 81-82). □

This variance decomposition or similar versions have been used by various authors in the actuarial literature to decompose an overall risk to separate components. See, for example, Parker (1997), Olivieri (2001), or Helwich (2003).

Now let $(\tilde{X}_{t,1}), \ldots, (\tilde{X}_{t,n})$ denote independent biographies of a homogenous portfolio of $n \in \mathbb{N}$ insured. The term

$$\frac{1}{n} \sum_{i=1}^{n} b_s\big((\tilde{X}_{t,i}), (\Phi_t)\big)$$

is the arithmetic mean of the corresponding present values. Similar to Parker (1997), calculating its decomposed variance components in terms of (1.4.3) shows the diversifiability of the unsystematic biometrical risk,

$$\begin{aligned} &\mathbb{E}\Bigg(\mathrm{Var}\Big(\frac{1}{n} \sum_{i=1}^{n} b_s\big((\tilde{X}_{t,i}), (\Phi_t)\big) \Big| \,(\Phi_t)\,, (q_t)\Big)\Bigg) \\ &= \frac{1}{n}\mathbb{E}\Bigg(\mathrm{Var}\Big(b_s\big((\tilde{X}_{t,1}), (\Phi_t)\big) \Big| \,(\Phi_t)\,, (q_t)\Big)\Bigg) \longrightarrow 0\,, \quad n \to \infty\,, \end{aligned}$$

and the non-diversifiability of the technical basis risk,

$$\begin{aligned} &\mathrm{Var}\Bigg(\mathbb{E}\Big(\frac{1}{n} \sum_{i=1}^{n} b_s\big((\tilde{X}_{t,i}), (\Phi_t)\big) \Big| \,(\Phi_t)\,, (q_t)\Big)\Bigg) \\ &= \mathrm{Var}\Bigg(\mathbb{E}\Big(b_s\big((\tilde{X}_{t,1}), (\Phi_t)\big) \Big| \,(\Phi_t)\,, (q_t)\Big)\Bigg) = \mathrm{const}\,, \quad \forall\, n \in \mathbb{N}\,. \end{aligned}$$

Hence it is convenient to pay some special attention to the financial and the systematic biometrical risks, which correspond to the uncertainty of (1.4.1). In case of the classical life insurance model as stated in section 1.2, the conditional expectation (1.4.1) is a deterministic functional of the technical basis. Taking Theorem 1.3.6 into account, this motivates to perform a sensitivity analysis on the prospective reserve (1.2.8) in order to study the effect of fluctuations of interest rate and transition probabilities. Such a sensitivity analysis will be performed in chapter 3.

The task of decomposing the present value $\tilde{B}_s$ to its risk components is not finished yet. Although the unsystematic risk was separated, the financial risk and the systematic biometrical risks, such as systematic mortality risk or systematic disability risk, are still cumulated in (1.4.1). The sensitivity analysis tools of chapters 2 and 3 will allow for a linearization of (1.4.1), resulting in a sum whose addends uniquely correspond to the different risk factors, see section 4.2.

# 2 A sensitivity analysis approach for functionals on specific function spaces

Suppose $F$ is a real-valued mapping on the Euclidean linear space $\mathbb{R}^n$, and assume that it is differentiable at $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. If one is interested in the sensitivity of the output to variations of the argument $x$, a common approach of sensitivity analysis is to analyze the gradient of $F$ at $x$,

$$\nabla_x F = \left( \frac{\partial F}{\partial x_1}(x) \, , \; \ldots \; , \; \frac{\partial F}{\partial x_n}(x) \right) \in \mathbb{R}^n \, . \tag{2.0.1}$$

According to Saltelli et al. (2000, section 2.3), the partial derivatives are called 'first-order local sensitivities', which is motivated by the two following characteristic properties of gradients:

(a) *The directional derivatives can be composed as a bilinear mapping of gradient and direction.* That means here

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} F(x + th) = < \nabla_x F, h > = \sum_{i=1}^{n} (\nabla_x F)_i \, h_i \, , \quad \forall \, h \in \mathbb{R}^n \, . \tag{2.0.2}$$

The $i$-th entry of the gradient corresponds to the sensitivity of $F$ to changes of the $i$-th argument.

(b) *If the gradient is not equal to zero, it is the unique orientation in which the directional derivative at $x$ has the largest value.* That means here

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} F\left(x + t \frac{h}{\|h\|}\right) \; < \; \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} F\left(x + t \frac{\nabla_x F}{\|\nabla_x F\|}\right) \tag{2.0.3}$$

for all $h \in \mathbb{R}^n$, which are linearly independent of $\nabla_x F$.

Unfortunately many functionals of interest, including various actuarial quantities, depend on an infinite number of parameters.

*Is it possible to generalize the gradient concept to functionals on infinite-dimensional spaces, while preserving the properties* (a) *and* (b)*?*

For functionals on $L_p(\nu)$-spaces, where $1 < p < \infty$ and $\nu$ is a Borel-measure on $\mathbb{R}$, section 2.1 shows the answer is 'yes' if one replaces (2.0.2) by

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} F(x+th) = \int \nabla_x F \, h \, \mathrm{d}\nu \,, \qquad \forall\, h \in L_p(\nu) \,, \tag{2.0.4}$$

and substitutes some transformation of the gradient for the original one in (2.0.3). In case of $p = 1$ one also finds a solution $\nabla_x F \in L_\infty(\nu)$ for (2.0.4), but has to abandon the maximality property (b). With property (a), the representation of the directional derivatives by some bilinear mapping of gradient and direction, being worthwhile on its own, it will solely be the defining property of generalized gradients.

This conforms with Courant and Hilbert (1968, p. 193), who define generalized gradients for mappings on function spaces by a condition similar to (2.0.4): In their approach, $\nu$ is the Lebesgue-Borel measure, and the functional $F$ needs to be Hadamard (or compact) differentiable, which is a stronger condition than Condition 2.1.1 will be. While Courant and Hilbert do not mention anything about existence or uniqueness, such statements are the objective of section 2.1.

For later purposes it is convenient to free the $L_1(\nu)$-functionals from their dependence on the dominating measure $\nu$. To this effect they will be embedded into the set of functionals on $BV_{\leftarrow}$ by identifying each $x \in L_1(\nu)$ with a measure $\mu = x\nu$, which in turn corresponds uniquely to an element $X \in BV_{\leftarrow}$ via $X(t) \equiv \mu((-\infty, t])$, provided $\nu$ is a Borel-measure concentrated on $[0,\infty)$. Equation (2.0.4) gets

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \widetilde{F}(X+tH) = \int \nabla_X \widetilde{F} \, \mathrm{d}H \,, \qquad \forall\, H \in BV_{\leftarrow} \,. \tag{2.0.5}$$

Section 2.2 will show that a solution $\nabla_X \widetilde{F}$ of (2.0.5) exists for a wide bandwidth of functionals $\widetilde{F} : BV_{\leftarrow} \to \mathbb{R}$. Because the maximality property (b) cannot be maintained for $L_1(\nu)$-functionals, it also will not hold for $BV_{\leftarrow}$-functionals.

If the arguments $X + tH \in BV_{\leftarrow}$ correspond to probability measures, the gradient concept (2.0.5) looks similar to the so-called 'influence function' and to generalized gradients in the field of robust and of nonparametric statistics, respectively. Section 2.2 will go into detail as regards the differences and similarities.

Differing from (2.0.1) the term 'gradient at $x$' is frequently used for the linear and continuous mapping $h \mapsto \, < \nabla_x F, h >$, which is an element of the corresponding dual space. From this perspective, many authors denote the generalized differential $\mathrm{D}_x F$ (cf. section A.3) as generalized gradient of a functional $F$. To distinguish this approach from the gradient concept used here, the term *gradient vector* is used from now on in referring to (2.0.1) or generalizations (2.0.4) and (2.0.5).

## 2.1 A gradient vector for functionals on $L_p(\nu)$

As indicated above, let equation (2.0.4) be the defining property for generalizing the gradient vector to functionals on $L_p(\nu)$. The left hand side of (2.0.4) implies the existence of the directional derivatives. Therefore, a natural condition to claim here is the existence of the Gâteaux derivative (cf. Definition A.3.1), which is a generalization of the directional derivative for mappings on function spaces:

**Condition 2.1.1.** Let the functional $F : L_p(\nu) \to \mathbb{R}$, $1 \le p < \infty$, be Gâteaux-differentiable at $x \in L_p(\nu)$, and let the Gâteaux differential $\mathrm{D}_x F$ be an element of the dual space $(L_p(\nu))'$.

Gâteaux differentials are in general homogenous (cf. (A.3.2)), but not necessarily linear and continuous. As a result of the Riesz Representation Theorem, these two additional conditions are necessary and sufficient for existence and uniqueness of representation (2.0.4):

**Definition 2.1.2** (gradient vector I)**.** Given that Condition 2.1.1 holds, denote the unique solution $\nabla_x F \in L_{p/(p-1)}(\nu)$ of

$$\mathrm{D}_x F(h) = \int \nabla_x F \, h \, \mathrm{d}\nu \,, \qquad \forall\, h \in L_p(\nu)\,, \tag{2.1.1}$$

as *gradient vector of $F$ at $x$*.

The idea to generalize the gradient vector concept to functionals on Hilbert spaces (here $L_2(\nu)$ is a Hilbert space) appears already in Golomb (1934, pp. 66, 67), who implicitly uses Fréchet differentiability. Since then, many authors used this idea, combined with the Riesz Representation Theorem, to obtain generalized gradient vectors, as for example Flett (1980, Exercise 3.5.1). Particularly in the field of nonparametric statistics this idea plays an important role (cf. Bickel et al. (1998, pp. 58, 178)). Parallels of the latter concept to the one presented herein are discussed more in detail in section 2.2.

**Remark 2.1.3.** In case of $1 < p < \infty$, the gradient vector (2.1.1) has an equivalent to property (2.0.3): By applying Hölders Inequality one can show that

$$\operatorname{sign}\bigl(\nabla_x F(\cdot)\bigr)\, |\nabla_x F(\cdot)|^{1/(p-1)} \;\in L_p(\nu)$$

is the unique orientation, in which the Gâteaux differential $\mathrm{D}_x F$ has its largest value with respect to the $L_p(\nu)$-norm. Note that for $p = 2$ the above vector coincides with the gradient vector $\nabla_x F$. For $p = 1$ the maximizing orientation is not unique. This is seen by the following example: Define the functional $F$ by

$$F : L_1(\nu) \ni x \mapsto \int x \, \mathrm{d}\nu \in \mathbb{R}\,.$$

The Gâteaux differential at $x \in L_1(\nu)$ is linear and continuous and is equal to $D_x F(\cdot) = F(\cdot)$. Now any vector $h \in L_1(\nu)$ with $h \ge 0$, $\nu$-almost everywhere, exemplifies a maximizing orientation.

## 2.2 A gradient vector for functionals on $BV_{\leftarrow}$

Section 2.1 provided a gradient vector concept in particular for functionals $F$ on $L_1(\nu)$. Now let $\nu$ be a Borel-measure concentrated on $[0,\infty)$. Sometimes it is convenient to get rid of the dependence on that dominating measure. This can be achieved by embedding the $L_1(\nu)$-functional $F$ into the set of functionals on $BV_{\leftarrow}$:

(1) Identify each $x \in L_1(\nu)$ with its cumulative counterpart $X \in BV_{\leftarrow}$ via the isometric isomorphism

$$I_\nu : L_1(\nu) \to I_\nu\big(L_1(\nu)\big) \subset BV_{\leftarrow}\,, \quad x \mapsto \int_{(-\infty,\cdot]} x \,\mathrm{d}\nu =: X(\cdot)\,. \tag{2.2.1}$$

(2) Let $\widetilde{F} \in \{BV_{\leftarrow} \to \mathbb{R}\}$ be a functional which satisfies $\widetilde{F}(X) = \widetilde{F}(I_\nu(x)) \stackrel{!}{=} F(x)$ for all $x \in L_1(\nu)$.

Since the gradient vector I according to Definition 2.1.2 was implicitly defined by (2.0.4), a consistent gradient vector extension should satisfy equation (2.0.5), at least for all $H \in I_\nu(L_1(\nu))$. Aiming to get rid of the dependence on $\nu$, let (2.0.5) for $H \in BV_{\leftarrow}$ be the defining property for gradient vectors of $\widetilde{F}$:

**Definition 2.2.1** (gradient vector II)**.** Let the functional $\widetilde{F} : BV_{\leftarrow} \to \mathbb{R}$ be Gâteaux differentiable at $X \in BV_{\leftarrow}$ with Gâteaux differential $\mathrm{D}_X\widetilde{F} \in (BV_{\leftarrow})'$. If there exists a representation of the form

$$\mathrm{D}_X\widetilde{F}(H) = \int g \,\mathrm{d}H\,, \qquad \forall H \in BV_{\leftarrow}\,, \tag{2.2.2}$$

the function $g$ is called the *gradient vector of* $\widetilde{F}$ *at* $X$. Write $\nabla_X\widetilde{F} := g$.

As the mapping $I_\nu$ is isometric, that is, $\|X\|_{BV} = \|I_\nu(x)\|_{BV} = \|x\|_{L_1}$ for all $x \in L_1(\nu)$, and $L_1(\nu)$-functionals have no equivalent for the maximality property (2.0.3) (cf. Remark 2.1.3), the gradient vector II has no maximality property either.

**Proposition 2.2.2.** *The gradient vector II of a Gâteaux differentiable functional is unique on* $[0,\infty)$.

*Proof.* The uniqueness is a consequence of $\nabla_X\widetilde{F}(u) = \mathrm{D}_X\widetilde{F}(\mathbf{1}_{[u,\infty)})$ for all $u \geq 0$ (see (2.2.2)). □

The following proposition shows that the gradient vector I of the embedded functional $F$ and the gradient vector II of the corresponding enlarged functional $\widetilde{F}$ are indeed consistent.

**Proposition 2.2.3.** *Let* $\widetilde{F} : BV_{\leftarrow} \to \mathbb{R}$ *be Gâteaux differentiable at* $X = I_\nu(x) \in BV_{\leftarrow}$, $x \in L_1(\nu)$, *with Gâteaux differential* $\mathrm{D}_X\widetilde{F} \in (BV_{\leftarrow})'$ *and existing gradient vector* $\nabla_X\widetilde{F}$ *according to Definition* 2.2.1*. Then,* $F = \widetilde{F} \circ I_\nu$ *is Gâteaux differentiable at* $x$ *with Gâteaux differential* $\mathrm{D}_xF \in (L_1(\nu))'$ *and gradient vector* $\nabla_X\widetilde{F}$.

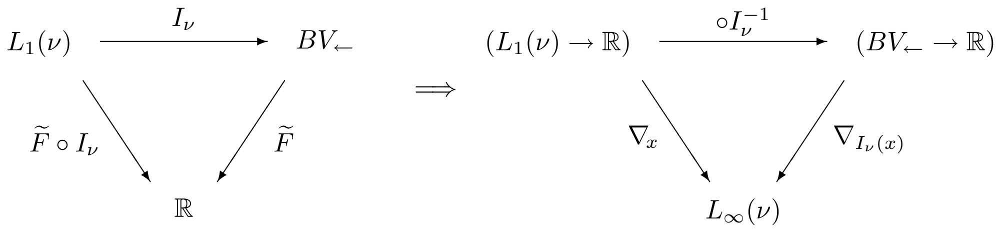


Figure 2.2.1: Illustration of Proposition 2.2.3

*Proof.* Per Definition A.3.1 and the linearity of $I_\nu$ is for any $h \in L_1(\nu)$

$$\mathrm{D}_x F(h) := \frac{\mathrm{d}}{\mathrm{d}t} F(x + th))\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t} \widetilde{F}\big(I_\nu(x) + tI_\nu(h)\big)\Big|_{t=0} =: \mathrm{D}_{I_\nu(x)} \widetilde{F}\big(I_\nu(h)\big)\,.$$

Since the functional $\mathrm{D}_X \widetilde{F}$ was supposed to be linear and continuous, the same holds for $\mathrm{D}_x F$, using the linearity and isometry of $I_\nu$. Hence, the gradient vector $\nabla_x F$ of $F$ at $x$ exists (cf. Condition 2.1.1). Consequently,

$$\int \nabla_x F \, h \,\mathrm{d}\nu = \mathrm{D}_x F(h) = \mathrm{D}_{I_\nu(x)} \widetilde{F}\big(I_\nu(h)\big) = \int \nabla_{I_\nu(x)} \widetilde{F} \,\mathrm{d}I_\nu(h) = \int \nabla_{I_\nu(x)} \widetilde{F} \, h \,\mathrm{d}\nu$$

for all $h \in L_1(\nu)$. The Riesz Representation Theorem states then $\nabla_x F = \nabla_{I_\nu(x)} \widetilde{F}$ $\nu$-almost everywhere. □

While the existence, the linearity, and the continuity of the Gâteaux differential of $L_1(\nu)$-functionals implied already the existence of a unique solution of (2.0.4), the same conditions are not sufficient for the existence of gradient vector II:

**Example 2.2.4.** According to the Lebesgue Decomposition Theorem (see Hewitt and Stromberg (1965), p. 326), each function $H \in BV_{\leftarrow}$ is uniquely decomposable into $H = H^{as} + H^{\perp}$, where $H^{as}$ is absolutely continuous and $H^{\perp}$ is singular with respect to the Lebesgue-Borel measure. Now define $\widetilde{F} \in (BV_{\leftarrow})'$ by

$$\widetilde{F}(H) := \int \mathrm{d}H^{as} + 2 \int \mathrm{d}H^{\perp}\,.$$

Because of the linearity of $\widetilde{F}$, its Gâteaux differential exists at any $X \in BV_{\leftarrow}$ and has the form $\mathrm{D}_X \widetilde{F} = \widetilde{F} \in (BV_{\leftarrow})'$. Assume now $\widetilde{F}$ has a gradient vector $\nabla_X \widetilde{F}$ at $X$ in accordance with Definition 2.2.1, that is,

$$\mathrm{D}_X \widetilde{F}(H) = \int \nabla_X \widetilde{F} \,\mathrm{d}H \qquad \forall H \in BV_{\leftarrow}\,. \tag{2.2.3}$$

Then, $\nabla_X \widetilde{F}(u) = \mathrm{D}_X \widetilde{F}(\mathbf{1}_{[u,\infty)}) = \widetilde{F}(\mathbf{1}_{[u,\infty)}) = 2$ for all $u \geq 0$, which is a contradiction to $\mathrm{D}_X \widetilde{F} = \widetilde{F}$. Hence, a representation of the form (2.2.3) does not exist, although $\mathrm{D}_X \widetilde{F} \in (BV_{\leftarrow})'$.

In constructing the gradient vector I (Definition 2.1.2), it was the Riesz Representation Theorem which allowed for solving (2.0.4). In contrast, no completely satisfactory representation for the dual space of $BV_{\leftarrow}$ seems to be known (cf. Dunford and Schwartz (1988) pp. 374, 378, 392-393).

In the following, a theorem is given that provides representations for at least a comprehensive subset of $(BV_{\leftarrow})'$. This subset is chosen in such a way that it will contain the actuarial functionals 'prospective reserve' and 'premium level'. It is characterized by the following conditions:

**Condition 2.2.5.** Let $G : BV_{\leftarrow} \to \mathbb{R}$ be an element of $(BV_{\leftarrow})'$.

(a) Let the function $t \mapsto G\big(\mathbf{1}_{[t,\infty)}(\cdot)\big)$ be an element of $BVC_b$ (cf. appendix A.2).

(b) For all monotonic nondecreasing sequences $(H_n)_{n\in\mathbb{N}} \subset BV_{\leftarrow}^{+}$ converging uniformly to an $H \in BV_{\leftarrow}^{+}$ let there be a subsequence $(\widetilde{H}_n)_{n\in\mathbb{N}} \subset (H_n)_{n\in\mathbb{N}}$, for which

$$\lim_{n\to\infty} G(\widetilde{H}_n) = G(H)\,.$$

**Theorem 2.2.6.** (a) *Let $y$ be an element of $BVC_b$. Then,*

$$G : BV_{\leftarrow} \to \mathbb{R}\,, \quad G(H) := \int y \,\mathrm{d}H \tag{2.2.4}$$

*is a functional which satisfies Condition* 2.2.5.

(b) *For each functional $G$ satisfying Condition* 2.2.5 *there exists a unique function*

$$y = t \mapsto G\big(\mathbf{1}_{[t,\infty)}(\cdot)\big) \in BVC_b\,,$$

*for which representation* (2.2.4) *holds.*

The proof is based on the ideas of the proof of Theorem 2.31 in Milbrodt and Helbig (1999).

*Proof.* At first, part (a) is proven: The functional $G$ is linear, since integration is a linear operation, and bounded, as with $C := \sup_{t\in\mathbb{R}} |y(t)| < \infty$

$$|G(H)| = \left|\int y \,\mathrm{d}H\right| \leq C\,\|H\|_{BV}\,.$$

Condition 2.2.5(a) holds because of $G\big(\mathbf{1}_{[t,\infty)}(\cdot)\big) = \int y\,\mathrm{d}\mathbf{1}_{[t,\infty)} = y(t)$. Applying Proposition A.2.4 yields Condition 2.2.5(b).

Now, part (b) of the theorem is shown: At first, let $H$ be an element of $BV_{\leftarrow}^{+}$. Subject to Proposition A.2.5(a) there is a monotonic nondecreasing sequence of step functions $(H_n)_{n\in\mathbb{N}} \subset BV_{\leftarrow}^{+}$ converging uniformly to $H$. Using Condition 2.2.5(b) and applying Proposition A.2.5(b) lead to

$$G(H) = \lim_{n\to\infty} G(\widetilde{H}_n) = \lim_{n\to\infty} \int G(\mathbf{1}_{[t,\infty)})\,\mathrm{d}\widetilde{H}_n(t)$$

for a subsequence $(\widetilde{H}_n)_{n\in\mathbb{N}} \subset (H_n)_{n\in\mathbb{N}}$. According to part (a) of Theorem 2.2.6, which has already been proven, the integral on the right hand side is a functional satisfying Condition 2.2.5. Thus, one has

$$\lim_{n\to\infty} \int G(\mathbf{1}_{[t,\infty)})\,\mathrm{d}\widetilde{H}_n(t) = \int G(\mathbf{1}_{[t,\infty)})\,\mathrm{d}H(t)\,.$$

Hence, Theorem 2.2.6(b) holds for functionals on $BV_{\leftarrow}^{+}$. The Jordan-Hahn decomposition (Theorem A.2.1) allows now to extend the result to functionals on $BV_{\leftarrow}$: Using the linearity of $G$ (Condition 2.2.5(a)), for any $H \in BV_{\leftarrow}$ with Jordan-Hahn decomposition $H = H_+ - H_-$

$$G(H_+) - G(H_-) = \int G(\mathbf{1}_{[t,\infty)})\,\mathrm{d}H_+(t) - \int G(\mathbf{1}_{[t,\infty)})\,\mathrm{d}H_-(t) = \int G(\mathbf{1}_{[t,\infty)})\,\mathrm{d}H(t)\,.$$

□

That means the mapping

$$T : BVC_b \to \mathcal{G} \subset (BV_{\leftarrow})'\,, \quad (Ty)(H) := \int y\,\mathrm{d}H$$

is an isomorphism, where $\mathcal{G}$ denotes the set of all functionals in accordance with Condition 2.2.5. Hence, for any Gâteaux differentiable functional $\widetilde{F} : BV_{\leftarrow} \to \mathbb{R}$ whose Gâteaux differential $\mathrm{D}_X\widetilde{F}$ is an element of $\mathcal{G}$ there exists a gradient vector II.

The additional properties (a) and (b) in Condition 2.2.5 are sufficient, but not necessary, for the existence of a solution of (2.2.2). For example, let in definition (2.2.4) the function $y$ be continuous and bounded, but not necessarily of finite variation on compacts. Then, the functional $G$ is still linear and bounded. However, for the intended actuarial applications, continuity of the gradient vector is not on hand. Condition 2.2.5 will turn out to be loose enough to comprehend various actuarial functionals.

## 2.3 Comparison with concepts of the statistical literature

The gradient vector concepts presented are quite similar to approaches used in robust and in nonparametric statistics. In both areas, the functionals considered are mappings of probability measures. They can be embedded into the set of functionals on $BV_{\leftarrow}$ in case the probability measures are Borel-measures concentrated on $[0,\infty)$. Throughout this section, write $X$ for elements of $BV_{\leftarrow}$ as well as for their corresponding signed measures. The actual meaning has to be interpreted according to the context.

In robust statistics, the influence function – originally called 'influence curve' – was introduced by Hampel, who gives a definition with very weak existence conditions:

The *influence function of* $\widetilde{F}$ *at* $X$ is given by

$$IF(u;\widetilde{F},X) := \lim_{t\downarrow 0} \frac{\widetilde{F}\big((1-t)X + t\mathbf{1}_{[u,\infty)}\big) - \widetilde{F}(X)}{t} \tag{2.3.1}$$

in those $X$ where this limit exists (cf. Hampel (1986), p. 84). In Hampels approach, $X$ is the distribution function of a probability measure, which is infinitesimal contaminated at point $u$. Note that $\mathbf{1}_{[u,\infty)}$ is the distribution function of the probability measure which puts mass 1 at point $u$. The existence of the limit in (2.3.1) is an even weaker condition than the existence of $g$ in (2.2.2) (cf. Huber (1996), p. 10).

**Proposition 2.3.1.** *Let* $\widetilde{F}$ *be a functional in accordance with Definition* 2.2.1 *and existing gradient vector* $\nabla_X\widetilde{F}$ *at* $X$. *Then,* (2.3.1) *exists as well and*

$$IF(\cdot;\widetilde{F},X) = \nabla_X\widetilde{F}(\cdot) - \int \nabla_X\widetilde{F}\,\mathrm{d}X\,. \tag{2.3.2}$$

*Particularly*

$$\mathrm{D}_X\widetilde{F}(H-X) = \int IF(\cdot;\widetilde{F},X)\,\mathrm{d}H \tag{2.3.3}$$

*for all* $H \in BV_{\leftarrow}$ *that correspond to a probability measure.*

*Proof.* Applying equation (2.2.2) with $H = \mathbf{1}_{[u,\infty)} - X$ yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\widetilde{F}\big(X + t(\mathbf{1}_{[u,\infty)} - X)\big) - \widetilde{F}(X)}{t} = \int \nabla_X\widetilde{F}\,\mathrm{d}(\mathbf{1}_{[u,\infty)} - X) = \nabla_X\widetilde{F}(u) - \int \nabla_X\widetilde{F}\,\mathrm{d}X$$

for all $u \geq 0$. That means the limit (2.3.1) exists and is equal to (2.3.2). Applying equation (2.2.2) together with (2.3.2),

$$\mathrm{D}_X\widetilde{F}(H-X) = \int \nabla_X\widetilde{F}\,\mathrm{d}H - \int \nabla_X\widetilde{F}\,\mathrm{d}X \int \mathrm{d}H = \int IF(\cdot;\widetilde{F},X)\,\mathrm{d}H\,.$$

Note that $\int \mathrm{d}H = 1$ when $H$ is the distribution function of a probability measure. □

Equation (2.3.3) motivates to locally approximate the functional $\widetilde{F}$ by its first order Taylor expansion

$$\widetilde{F}(H) = \widetilde{F}(X) + \int IF(\cdot;\widetilde{F},X)\,\mathrm{d}H + \mathit{Remainder}\,. \tag{2.3.4}$$

This idea was introduced by R. von Mises in 1947 to obtain asymptotic distribution results (cf. Fernholz (1983), pp. 7, 8). With $\mathrm{D}_X\widetilde{F}$ in (2.3.3) being a Gâteaux differential, the convergence of the remainder to zero is not uniformly. Therefore, oftentimes Fréchet differentiability of $\widetilde{F}$ at $X$ is claimed, which implies a uniform convergence of the remainder with respect to $\|H - X\|$. Some authors vary this approach by claiming the convergence of the remainder with respect to some distance function $\delta(X,H)$

instead (cf. Pfanzagl and Wefelmeyer (1982), p. 66).

In nonparametric statistics, representations of the form (2.3.4) play an important role, running under the name 'gradient'. There, the concept was introduced by Koshevnik and Levit and further developed by Pfanzagl and Van der Vaart.

Instead of convergence of the remainder (to zero) on linear paths – this is the case if $\mathrm{D}_X \widetilde{F}$ represents a Gâteaux differential –, the convergence of the remainder is claimed on $L_r$-differentiable paths. (Following Pfanzagl (1982), various authors use a differing but equivalent approach instead. Compare Exercise 1.50 in Witting (1985).) On the one hand, this is a stronger condition, as not only linear paths, but also paths that are just asymptotic to linear paths (quasi the Hadamard differentiability of $\widetilde{F}$) are considered. On the other hand, only tangential directions with existing Radon-Nikodym derivative with respect to some fixed dominating measure are taken into account, which makes this concept rather similar to Definition 2.1.2 in section 2.1. The gradients in nonparametric statistic are equal to the corresponding influence function, provided they exist. That means their relation to the gradient vector concept presented here is analogous to (2.3.2).

Both in robust statistics and in nonparametric statistics, statements about the existence of influence functions and gradients are rare. An analogous result to Theorem 2.2.6 is Proposition 5.1 in Huber (1981, p. 37), where he claims weak continuity and Fréchet differentiability for the functional $\widetilde{F}$ to ensure a representation of the form (2.2.2). However, the condition of weak continuity allows only for continuous gradients $g$. In contrast, Theorem 2.2.6 allows the gradient vector II to have discontinuities, which will turn out to be inevitable for later purposes.

## 2.4 An extended gradient vector for $L_p^d(\nu)$-functionals

In sections 2.1 and 2.2 the considered functionals are mappings on one-dimensional function spaces. Now the domain shall be extended to $d \in \mathbb{N}$ dimensions. Let $\mathbf{x}$ be an element of the linear space

$$L_p^d(\nu) := \underbrace{L_p(\nu) \times \cdots \times L_p(\nu)}_{d \text{ times}}, \quad d \in \mathbb{N},$$

where multiplication by a scalar and summation is defined componentwise. A combination of $L_p(\nu)$-norm and $l_p$-Norm,

$$\|\mathbf{x}\|_p := \left( \sum_{i=1}^{d} \|x_i\|_{L_p(\nu)}^p \right)^{1/p} = \left\| \left( \|x_i\|_{L_p(\nu)} \right)_{i=1,..,d} \right\|_{l_p}, \qquad \mathbf{x} = (x_1, \ldots, x_d) \in L_p^d(\nu), \tag{2.4.1}$$

yields a norm on $L_p^d(\nu)$ for all $1 \leq p < \infty$. Based on the Riesz Representation Theorem, the linear mapping

$$T^d : L^d_{p/(p-1)}(\nu) \to \big(L_p^d(\nu)\big)', \quad (T^d\mathbf{y})(\mathbf{h}) := \int \mathbf{y}\cdot\mathbf{h}\,\mathrm{d}\nu = \sum_{i=1}^d \int y_i\,h_i\,\mathrm{d}\nu \tag{2.4.2}$$

is an isometric isomorphism for any $1 \leq p < \infty$.

**Definition 2.4.1** (gradient vector I)**.** For each functional $F : L_p^d(\nu) \to \mathbb{R}$ whose Gâteaux differential $\mathrm{D}_{\mathbf{x}}F$ exists and is an element of the dual space $(L_p^d(\nu))'$, denote the unique solution $\nabla_{\mathbf{x}}F \in L^d_{p/(p-1)}(\nu)$ of

$$\mathrm{D}_{\mathbf{x}}F(\mathbf{h}) = \int \nabla_{\mathbf{x}}F\cdot\mathbf{h}\,\mathrm{d}\nu := \sum_{i=1}^d \int (\nabla_{\mathbf{x}}F)_i\,h_i\,\mathrm{d}\nu\,, \qquad \mathbf{h}\in L_p^d(\nu)\,, \tag{2.4.3}$$

as *gradient vector of* $F$ *at* $\mathbf{x}$.

**Remark 2.4.2.** Applying Hölders Inequality for $L_p$-spaces and for $l_p$-spaces consecutively (cf. Hewitt and Stromberg (1965), pp. 190, 194) yields

$$\big\|(x_1y_1,\ldots,x_dy_d)\big\|_1 \leq \|\mathbf{x}\|_p\,\|\mathbf{y}\|_{p/(p-1)}\,, \qquad \forall\,\mathbf{x}\in L_p^d(\nu)\,,\mathbf{y}\in L^d_{p/(p-1)}(\nu)\,,$$

for $1 < p < \infty$. Equality is obtained if and only if there are nonnegative real numbers $\alpha$ and $\beta$, not both zero, in such a manner that

$$\alpha\,(|x_1|^p,\ldots,|x_d|^p) = \beta\,(|y_1|^{p/(p-1)},\ldots,|y_d|^{p/(p-1)})\,.$$

With that, analogously to Remark 2.1.3, one gets a unique maximizing orientation for the differential $\mathrm{D}_{\mathbf{x}}F$,

$$\Big(\mathrm{sign}\big((\nabla_{\mathbf{x}}F)_i\big)\,\big|(\nabla_{\mathbf{x}}F)_i\big|^{1/(p-1)}\Big)_{i=1,\ldots,d}\,.$$

## 2.5 An extended gradient vector for $BV_{\leftarrow}^d$-functionals

Now let $\mathbf{X}$ be an element of the linear space

$$BV_{\leftarrow}^d := \underbrace{BV_{\leftarrow}\times .. \times BV_{\leftarrow}}_{d\text{ times}}\,, \qquad d\in\mathbb{N}\,,$$

where multiplication by a scalar and summation is defined componentwise. Then,

$$\|\mathbf{X}\|_{BV} := \sum_{i=1}^d \|X_i\|_{BV}\,, \qquad \mathbf{X} = (X_1,\ldots,X_d)\in BV_{\leftarrow}^d\,, \tag{2.5.1}$$

defines a norm on $BV_{\leftarrow}^d$. Extend Definition 2.2.1 as follows:

**Definition 2.5.1** (gradient vector II)**.** Let the functional $\widetilde{F} : BV^d_{\leftarrow} \to \mathbb{R}$ be Gâteaux differentiable at $\mathbf{X} \in BV^d_{\leftarrow}$ with Gâteaux differential $\mathrm{D}_{\mathbf{X}}\widetilde{F} \in (BV^d_{\leftarrow})'$. If there exists a representation of the form

$$\mathrm{D}_{\mathbf{X}}\widetilde{F}(\mathbf{H}) = \int \mathbf{g} \cdot \mathrm{d}\mathbf{H} := \sum_{i=1}^{d} \int g_i \, \mathrm{d}H_i \,, \qquad \forall\, \mathbf{H} \in BV^d_{\leftarrow} \,, \tag{2.5.2}$$

$\mathbf{g} = (g_1, \dots, g_d)$ is called *gradient vector of* $\widetilde{F}$ *at* $\mathbf{X}$. Write $\nabla_{\mathbf{X}}\widetilde{F} := \mathbf{g}$.

Similar to Proposition 2.2.2 the extended gradient vector II is still unique:

**Proposition 2.5.2.** *The gradient vector II of a Gâteaux differentiable functional according to Definition* 2.5.1 *is unique on* $[0,\infty)^d$.

*Proof.* The uniqueness is a consequence of

$$\nabla_{\mathbf{X}}\widetilde{F}(\mathbf{u}) = \Big((\nabla_{\mathbf{X}}\widetilde{F})_1(u_1), \dots, (\nabla_{\mathbf{X}}\widetilde{F})_d(u_d)\Big) = \mathrm{D}_{\mathbf{X}}\widetilde{F}\big((\mathbf{1}_{[u_1,\infty)}, \dots, \mathbf{1}_{[u_d,\infty)})\big) \,,$$

for all $\mathbf{u} = (u_1, \dots, u_d) \in [0,\infty)^d$. □

Naturally the question arises, whether the extended gradient vector Definitions 2.4.1 and 2.5.1 still complement one another as Definitions 2.1.2 and 2.2.1 did in Proposition 2.2.3. In fact, by extending (2.2.1) to

$$I_\nu : L^d_1(\nu) \to BV^d_{\leftarrow} \,, \quad I_\nu(\mathbf{x}) := \big(I_\nu(x_1), \dots, I_\nu(x_d)\big) \,,$$

Proposition 2.2.3 adapted to the multidimensional setting holds; in particular

$$\nabla_{\mathbf{x}}(\widetilde{F} \circ I_\nu) = \nabla_{I_\nu(\mathbf{x})}\widetilde{F} \quad \nu\text{-almost everywhere.} \tag{2.5.3}$$

The proof is analogous to that of Proposition 2.2.3.

A beneficial property of the extended gradient vector is that its $d$ entries are equal to the separate gradient vectors of the functional $\widetilde{F}$ restricted to its $d$ one-dimensional function spaces:

**Proposition 2.5.3.** *Let* $\nabla_{\mathbf{X}}\widetilde{F} = \big((\nabla_{\mathbf{X}}\widetilde{F})_1, \dots, (\nabla_{\mathbf{X}}\widetilde{F})_d\big)$ *be the gradient vector of* $\widetilde{F}$ *at* $\mathbf{X} = (X_1, \dots, X_d) \in BV^d_{\leftarrow}$ *in accordance with Definition* 2.5.1, *and define the mappings* $e_i : BV_{\leftarrow} \to BV^d_{\leftarrow}$ *by* $Y \mapsto (X_1, \dots, X_{i-1}, Y, X_{i+1}, \dots, X_d)$. *Then,*

$$(\nabla_{\mathbf{X}}\widetilde{F})_i = \nabla_{X_i}(\widetilde{F} \circ e_i) \,, \qquad \forall i \in \{1, \dots, d\} \,, \tag{2.5.4}$$

*where* $\nabla_{X_i}(\widetilde{F} \circ e_i)$ *is a gradient vector in accordance with Definition* 2.2.1.

*Proof.* For an arbitrary but fixed $i \in \{1, \dots, d\}$, let $\mathbf{H} := (0, \dots, 0, H_i, 0, \dots, 0)$ and $H_i := \mathbf{1}_{[u,\infty)}$. Because of (2.5.2), (A.3.1), and (2.2.2)

$$(\nabla_{\mathbf{X}}\widetilde{F})_i(u) = \int (\nabla_{\mathbf{X}}\widetilde{F})_i \, \mathrm{d}H_i = \mathrm{D}_{\mathbf{X}}\widetilde{F}(\mathbf{H}) = \mathrm{D}_{X_i}(\widetilde{F} \circ e_i)(H_i) = \nabla_{X_i}(\widetilde{F} \circ e_i)(u) \,, \;\; u \in \mathbb{R}.$$

□

Rewriting Condition 2.2.5 to fit the multidimensional case, one gets an analogue to Theorem 2.2.6 for

$$G : BV^d_{\leftarrow} \to \mathbb{R}\,, \quad G(\mathbf{H}) := \int \mathbf{y} \cdot \,\mathrm{d}\mathbf{H} := \sum_{i=1}^{d} \int y_i \,\mathrm{d}H_i\,, \qquad \mathbf{y} \in (BVC_b)^d\,, \quad (2.5.5)$$

and

$$\mathbf{y}(\mathbf{u}) := \big(G(\mathbf{1}_{[u_1,\infty)}, 0, \ldots, 0), \ldots, G(0, \ldots, 0, \mathbf{1}_{[u_d,\infty)})\big)\,, \qquad \mathbf{u} = (u_1, \ldots, u_d) \in \mathbb{R}^d\,.$$

# 3 A sensitivity analysis of life insurance contracts

As pointed out in the introduction, the actuarial assumptions such as interest rate or mortality rate can vary significantly within a contract period. For studying the influence of such changes on profits and losses or premiums a sensitivity analysis can help.

A first attempt was made by Lidstone (1905), who studied in a discrete time setting the effect on reserves of changes in technical basis and contract terms. However, Lidstone dealt only with simple single life policies with payments dependent on survival and death. The same did Norberg (1985), who transferred Lidstones ideas to a continuous case where premiums and annuities are payable continuously. Likewise using Thieles differential equations, Hoem (1988), Ramlau-Hansen (1988), and Linnemann (1993) obtained further results on a more general level. All of these studies have in common that they mainly yield qualitative and less quantitative results. They show which direction the prospective reserve or the premium level are shifted to by a parameter change, but hardly quantify the magnitude of that effect.

Another approach is to calculate different scenarios and to compare them with each other, as for example Olivieri (2001) or Khalaf-Allah et al. (2006), but this idea only works for a small number of parameters.

A third way is to study sensitivities by means of derivatives, which turned out to be a very efficient concept. References using such an approach are Dienst (1995), Bowers et al. (1997), Kalashnikov and Norberg (2003), and Helwich (2003, 2005):

Dienst (1995, pp. 66-68, 147-150) uses a finite number of partial derivatives of the net premium with respect to time-discrete disablement probabilities to approximate the relative change of the net premium caused by altered disablement probabilities.

Bowers et al. (1997, pp. 490, 491) calculate the first order derivative of the expected loss with respect to the interest rate, which they assume to be constant.

Kalashnikov and Norberg (2003) differentiate the prospective reserve and the premium level with respect to one arbitrary real parameter. This also includes parameters such as contract terms. Similar to Thieles differential equation for the prospective reserve they present differential equations for their derivative. In section 5, Kalashnikov and Norberg generalize their approach to a finite number of real parameters.

Helwich (2003, 2005) models the actuarial assumptions as finite dimensional and real-valued vectors, allowing for parameter changes at a finite number of discrete time points. He calculates the gradient (2.0.1) of the expected loss of a portfolio of insurance contracts with respect to yearly constant interest and retirement rates.

All of those studies have in common that they only allow for a finite number of parameters. This chapter presents a sensitivity analysis based on the generalized gradient vector concept of chapter 2. This makes it possible to study sensitivities with respect to an infinite number of parameters, which meets, for example, the more realistic idea of actuarial assumptions (e.g., mortality) being functions on the real line rather than on a discrete time grid. Nonetheless, the present approach is including also discrete time models, and thus it generalizes Helwichs (2003) chapter 5 in case of just a single contract.

Sections 3.1 to 3.4 calculate the gradient vectors of the prospective reserve and the premium value in general as mappings of the technical basis. Interpreting there gradient vectors as sensitivities, section 3.5 performs a sensitivity analysis for several typical examples, yielding some substantial insights.

#### Actuarial functionals of interest

Since the prospective reserve (1.2.8) and the premium level (1.2.17) are largely the crucial quantities for shaping fundamental contract terms, calculating premiums, or evaluating future obligations, they are the actuarial functionals for which a sensitivity analysis shall be performed here. As shown later in section 3.4, the generalized gradient vector of the premium level is easily deducible from that of the prospective reserve. Hence, focus initially only on the latter:

Taking into consideration (1.2.2) and (1.2.3), the prospective reserve $V_{y,s}$ given by (1.2.11) may be seen as a functional of the cumulative interest intensity $\Phi \in BVC_{\leftarrow}$ and the cumulative intensities of transition $q_{yz} \in BVC_{\leftarrow}$ , $(y,z) \in J$ . However, the sensitivity analysis concept of sections 2.2 and 2.5 allows only for arguments with finite total variation! For example, finite total variation for the cumulative interest intensity $\Phi$ would imply that the discounting factor $1/K$ has an upper bound on $[0,\infty)$, which disagrees with established notions of the economical reality. How to handle that discrepancy?

A possible solution is to distinguish between *initial point* and *deviation*: Decompose the cumulative intensities to

$$(\Phi, q) = (\Phi^*, q^*) + (H_\Phi, H_q)\,, \qquad H_\Phi \in BV_{\leftarrow}\,, H_q \in (BV_{\leftarrow})^{|\mathcal{S}|\times|\mathcal{S}|}\,, \tag{3.0.1}$$

where $\Phi^* \in BVC_{\leftarrow}$ and $q^* \in (BVC_{\leftarrow})^{|\mathcal{S}|\times|\mathcal{S}|}$ build an arbitrary but fixed initial point which varies with deviation $(H_\Phi, H_q) \in BV_{\leftarrow} \times (BV_{\leftarrow})^{|\mathcal{S}|\times|\mathcal{S}|}$. Now regard the prospective reserve as a functional of the deviation only. This ambivalent approach allows for lifelike actuarial assumptions on $[0,\infty)$; but note that a sensitivity analysis considers just a shortened set of fluctuations! In most practical cases it suffices to act on the assumption of finite time horizons, where the distinction of initial point and deviation vanishes. (In case of a finite time horizon $T < \infty$, all cumulative intensities may, without loss of generality, be multiplied with $\mathbf{1}_{[0,T]}$ or $\mathbf{1}_{[0,DT(T)]}$, which makes them elements of $BV_{\leftarrow}$.)

**Remark 3.0.4.** Facing the unusual split between 'initial point' and 'deviation', one could ask: why not expanding Definition 2.2.1 to functionals on $BVC_{\leftarrow}$? The two following items run contrary:

(i) The directional derivative of the prospective reserve as a functional on $(BVC_{\leftarrow})^d$ usually does not exist for all directions $H \in (BVC_{\leftarrow})^d$. Moreover the set of directions with existing directional derivative depends highly on the contract terms, which is problematic if one aims to compare the sensitivities of heterogenous insurance contracts.

Look at the following simple example: Let the state space have just a single state $\mathcal{S} = \{z\}$ and define $F_z$ by

$$F_z := \sum_{n=1}^{\infty} \frac{1}{n^2} \mathbf{1}_{[n,\infty)} .$$

Then, the prospective reserve at $\Phi^* \equiv 0$ exists,

$$\sum_{z\in\mathcal{S}} \int_{[0,\infty)} \frac{1}{K_{\Phi^*}(t)} F_z(\mathrm{d}t) = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty .$$

Defining

$$H_\Phi := \sum_{n=1}^{\infty} -\frac{1}{2} \mathbf{1}_{[n,\infty)} ,$$

the cumulative interest intensity $\Phi = \Phi^* + \tau H_\Phi = 0 + \tau H_\Phi$ meets condition (1.2.4) for all $0 < \tau < 1$, but the prospective reserve at $\Phi$ does not exist for any $0 < \tau < 1$,

$$\sum_{z\in\mathcal{S}} \int_{[0,\infty)} \frac{1}{K_{\Phi}(t)} F_z(\mathrm{d}t) = \sum_{n=1}^{\infty} \frac{1}{(1-\tau/2)^n} \frac{1}{n^2} = \infty ,$$

that is, the directional derivative at initial point $\Phi^*$ does not exist in direction $H_\Phi$.

Similar examples can be constructed for deviations of the cumulative transition intensities. Note that the choice of $\Phi$ is not just a theoretical example but a customary model for the financial development. That means directions with no existing directional derivative are not only on the fringes.

(ii) The prospective reserve according to (3.0.1) as a functional of the deviation will prove to be even Fréchet differentiable, that is, the convergence of the directional difference quotients is even uniform. But extending the domain to functions of bounded variation on compacts, the total variation is not a norm anymore, and one loses the Fréchet differentiability, which is only meaningful on normed spaces.

Taking Proposition 2.5.3 into consideration, at first gradient vectors are calculated with respect only to the cumulative interest or one of the cumulative transition intensities.

## 3.1 Gradient vector with respect to interest

In this section, regard the prospective reserve as a functional solely with respect to the cumulative interest intensity $\Phi$. As stated in (3.0.1), let $\Phi$ decompose to $\Phi = \Phi^* + H$, where the starting point $\Phi^*$ satisfies (1.2.4) and the integrability conditions (1.2.6) and (1.2.7), and $H \in BV_{\leftarrow}$ is a deviation from $\Phi^*$. Write

$$V_{y,s}^{\Phi^*}(H) := V_{y,s}[\Phi^* + H, q^*] .$$

Unfortunately, this functional does not generally exist on the whole of $BV_{\leftarrow}$, as $\Phi = \Phi^* + H$ does not necessarily satisfy the lower jump bound condition $\Delta\Phi(t) > -1$, $t \in \mathbb{R}$, in (1.2.4). Aiming to obtain the gradient vector II of $V_{y,s}^{\Phi^*}$ at zero, one needs at least the existence of $V_{y,s}^{\Phi^*}(H)$ for $H$ in a neighborhood of zero. Within this neighborhood, the corresponding cumulative interest intensity $\Phi = \Phi^* + H$ ought to satisfy the lower jump bound in (1.2.4). This motivates the definition

$$\mathcal{E}_{\Phi^*} := \left\{ H \in BV_{\leftarrow} \;\middle|\; \Delta(\Phi^* + H)(t) \geq \widetilde{C}_{\Phi^*} , \quad \forall\, t \in \mathbb{R} \right\} , \tag{3.1.1}$$

where $C_{\Phi^*}$ is the lower jump bound of $\Phi^*$ according to (1.2.4) and $\widetilde{C}_{\Phi^*}$ is an arbitrary but fixed constant satisfying $-1 < \widetilde{C}_{\Phi^*} < C_{\Phi^*} \leq 0$.

**Proposition 3.1.1.** *For each $H \in \mathcal{E}_{\Phi^*}$, $y \in \mathcal{S}$, and $s \geq 0$, one has $|V_{y,s}^{\Phi^*}(H)| < \infty$.*

The proof of that Proposition is based on the following useful Proposition:

**Proposition 3.1.2.** *For all $H \in \mathcal{E}_{\Phi^*}$ with corresponding Jordan-Hahn decomposition $H =: H_+ - H_-$ (cf. Theorem A.2.1)*

$$\prod_{(s,t]} \big(1 + \mathrm{d}(\Phi^* + H)\big) \leq \frac{1}{1 + C_{\Phi^*}} \prod_{(s,t]} \big(1 + \mathrm{d}\Phi^*\big) \prod_{(s,t]} \big(1 + \mathrm{d}H\big) , \tag{3.1.2}$$

$$\prod_{(s,t]} \big(1 + \mathrm{d}(\Phi^* + H)\big) \geq \overline{C} \prod_{(s,t]} \big(1 + \mathrm{d}\Phi^*\big) \prod_{(s,t]} \big(1 + \mathrm{d}(-H_-)\big) > 0 \tag{3.1.3}$$

*for $\overline{C} := (1 + \widetilde{C}_{\Phi^*})/((1 + C_{\Phi^*})(1 + \widetilde{C}_{\Phi^*} - C_{\Phi^*})) \in (0, 1]$.*

*Proof.* Because of $1 + \Delta\Phi^*(t) \geq 1 + C_{\Phi^*}$ for all $t \in \mathbb{R}$ and $0 < 1 + C_{\Phi^*} \leq 1$

$$1 + \Delta(\Phi^* + H)(t) \leq \frac{1 + \Delta\Phi^*(t)}{1 + C_{\Phi^*}} + \frac{1 + \Delta\Phi^*(t)}{1 + C_{\Phi^*}} \Delta H(t) = \frac{\big(1 + \Delta\Phi^*(t)\big)\big(1 + \Delta H(t)\big)}{1 + C_{\Phi^*}}$$

for all $t \in \mathbb{R}$. Combined with Proposition A.4.2 this leads to (3.1.2),

$$\begin{aligned}
&\prod_{(s,t]} \big(1 + \mathrm{d}(\Phi^* + H)\big) \\
&= \exp\left(\Phi^*_c(t) - \Phi^*_c(s) + H_c(t) - H_c(s)\right) \prod_{\tau \in (s,t]} \Big(1 + \Delta(\Phi^* + H)(\tau)\Big) \\
&\leq \exp\Big(\Phi^*_c(t) - \Phi^*_c(s) + H_c(t) - H_c(s)\Big) \frac{1}{1 + C_{\Phi^*}} \prod_{\tau \in (s,t]} \Big(1 + \Delta\Phi^*(\tau)\Big)\Big(1 + \Delta H(\tau)\Big) \\
&= \frac{1}{1 + C_{\Phi^*}} \prod_{(s,t]} \big(1 + \mathrm{d}\Phi^*\big) \prod_{(s,t]} \big(1 + \mathrm{d}H\big) .
\end{aligned}$$

Since for all $a \in [C_{\Phi^*}, \infty)$ and $b \in [\widetilde{C}_{\Phi^*} - C_{\Phi^*}, 0]$

$$\frac{1 + a + b}{(1 + a)(1 + b)} \geq \frac{1 + \widetilde{C}_{\Phi^*}}{(1 + C_{\Phi^*})(1 + \widetilde{C}_{\Phi^*} - C_{\Phi^*})} = \overline{C} \in (0, 1] ,$$

one gets $1 + \Delta(\Phi^* - H_-)(\tau) \geq \overline{C}\,\big(1 + \Delta(\Phi^*)(\tau)\big)\big(1 - \Delta(H_-)(\tau)\big)$ for all $\tau \in \mathbb{R}$. Combined with Proposition A.4.2 this leads to (3.1.3),

$$\begin{aligned}
&\prod_{(s,t]} \big(1 + \mathrm{d}(\Phi^* + H)\big) \\
&\geq \exp\left(\Phi^*_c(t) - \Phi^*_c(s) + H_c(t) - H_c(s)\right) \prod_{\tau \in (s,t]} \Big(1 + \Delta(\Phi^* - H_-)(\tau)\Big) \\
&\geq \exp\Big(\Phi^*_c(t) - \Phi^*_c(s) + H_c(t) - H_c(s)\Big)\, \overline{C} \prod_{\tau \in (s,t]} \Big(1 + \Delta\Phi^*(\tau)\Big)\Big(1 - \Delta H_-(\tau)\Big) \\
&= \overline{C} \prod_{(s,t]} \big(1 + \mathrm{d}\Phi^*\big) \prod_{(s,t]} \big(1 + \mathrm{d}(-H_-)\big) > 0 .
\end{aligned}$$

□

*Proof of Proposition* 3.1.1. As $\Phi^* + H$ has finite variation on compacts, the corresponding accumulation factor $K_{\Phi^*+H}(s)$ is finite for any fixed $s \geq 0$ (cf. Proposition A.4.1). Therefore, it suffices to show

$$\begin{aligned}
&\sum_{z \in \mathcal{S}} \int_{[0,\infty)} \frac{1}{\left|\prod_{(0,t]}(1 + \mathrm{d}(\Phi^* + H))\right|} |F_z|(\mathrm{d}t) < \infty , \\
&\sum_{(z,\zeta) \in J} \int_{(0,\infty)} \frac{1}{\left|\prod_{(0,DT(t)]}(1 + \mathrm{d}(\Phi^* + H))\right|} D_{z\zeta}(t)\, q_{z\zeta}(\mathrm{d}t) < \infty ,
\end{aligned} \tag{3.1.4}$$

since $0 \leq p_{yz}(s, t) \leq 1$ for any $0 \leq s \leq t$ and any $y, z \in \mathcal{S}$. With (3.1.3) and Proposition A.4.2

$$0 < \frac{1}{\prod_{(0,t]}(1 + \mathrm{d}(\Phi^* + H))} \leq \overline{C} \frac{1}{\prod_{(0,t]}(1 + \mathrm{d}\Phi^*)} \exp\left(\frac{\ln(1 + \widetilde{C}_{\Phi^*} - C_{\Phi^*})}{\widetilde{C}_{\Phi^*} - C_{\Phi^*}} \|H_-\|_{BV}\right) .$$

That means

$$\frac{1}{|K_{\Phi^*+H}(t)|} \leq \frac{1}{|K_{\Phi^*}(t)|} \operatorname{const} e^{\operatorname{const} \|H\|_{BV}}, \quad t \geq 0, \tag{3.1.5}$$

and analogously

$$\frac{1}{|K_{\Phi^*+H}(DT(t))|} \leq \frac{1}{|K_{\Phi^*}(DT(t))|} \operatorname{const} e^{\operatorname{const} \|H\|_{BV}}, \quad t \geq 0, \tag{3.1.6}$$

which allows to bound the terms in (3.1.4) with the terms of (1.2.6) and (1.2.7) multiplied with factor $const\, e^{const\, \|H\|_{BV}} < \infty$. □

That means the prospective reserve at $\Phi = \Phi^* + H$ as a functional of the deviation $H$ is well-defined on $\mathcal{E}_{\Phi^*}$,

$$V_{y,s}^{\Phi^*} : \mathcal{E}_{\Phi^*} \to \mathbb{R}, \quad H \mapsto V_{y,s}[\Phi^* + H, q^*]. \tag{3.1.7}$$

Note that the space $\mathcal{E}_{\Phi^*}$ contains all continuous functions of $BV_{\leftarrow}$ and is not, in contrast to $BV_{\leftarrow}$, linear anymore. As $\mathcal{E}_{\Phi^*}$ contains even a whole ball around zero

$$\mathcal{B}_{\Phi^*} := \left\{ H \in BV_{\leftarrow} \,\middle|\, \|H\|_{BV} < C_{\Phi^*} - \widetilde{C}_{\Phi^*} \right\} \subset \mathcal{E}_{\Phi^*} \tag{3.1.8}$$

(note that $C_{\Phi^*} - \widetilde{C}_{\Phi^*} > 0$), not only Gâteaux but also Fréchet differentiability of $V_{y,s}^{\Phi^*}$ at zero may be studied. In fact, even the stronger Fréchet differentiability holds here:

**Theorem 3.1.3.** *The functional $V_{y,s}^{\Phi^*}$ is Fréchet differentiable at zero with differential*

$$\begin{aligned} &\mathrm{D}_0 V_{y,s}^{\Phi^*}(H) = \\ &-\sum_{z\in\mathcal{S}} \int_{[s,\infty)} \left( \int_{(s,t]} \frac{1}{1+\Delta\Phi^*(u)} \,\mathrm{d}H(u) \right) \frac{K_{\Phi^*}(s)}{K_{\Phi^*}(t)}\, p_{yz}(s,t)\, F_z(\mathrm{d}t) \\ &-\sum_{(z,\zeta)\in J} \int_{(s,\infty)} \left( \int_{(s,DT(t)]} \frac{1}{1+\Delta\Phi^*(u)} \,\mathrm{d}H(u) \right) \frac{K_{\Phi^*}(s)}{K_{\Phi^*}(DT(t))}\, D_{z\zeta}(t)\, p_{yz}(s,t-0)\, q_{z\zeta}(\mathrm{d}t) \end{aligned} \tag{3.1.9}$$

*for $H \in BV_{\leftarrow}$.*

The proof of this theorem is split into several pieces – following now – with some of them being formulated as an own proposition:

**Proposition 3.1.4.** *For all $H \in \mathcal{B}_{\Phi^*}$ and $-\infty < s < t < \infty$*

$$\left| \frac{1}{\prod_{(s,t]} \left(1 + \mathrm{d}(\Phi^* + H)\right)} - \frac{1}{\prod_{(s,t]} \left(1 + \mathrm{d}\Phi^*\right)} \right| \leq \frac{\operatorname{const} \|H\|_{BV}}{\prod_{(s,t]} \left(1 + \mathrm{d}\Phi^*\right)}, \tag{3.1.10}$$

$$\left|\frac{1}{\prod_{(s,t]}\big(1+\mathrm{d}(\Phi^*+H)\big)}-\frac{1}{\prod_{(s,t]}\big(1+\mathrm{d}\Phi^*\big)}+\frac{\int_{(s,t]}\frac{1}{1+\Delta\Phi^*(u)}\,\mathrm{d}H(u)}{\prod_{(s,t]}\big(1+\mathrm{d}\Phi^*\big)}\right|\le\frac{\mathrm{const}\,\|H\|_{BV}^2}{\prod_{(s,t]}\big(1+\mathrm{d}\Phi^*\big)}\,. \tag{3.1.11}$$

*For all continuous functions $H \in BV_{\leftarrow}$ and $-\infty < s < t < \infty$*

$$\left|\frac{1}{\prod_{(s,t]}\left(1+\mathrm{d}(\Phi^*+H)\right)}-\frac{1}{\prod_{(s,t]}\left(1+\mathrm{d}\Phi^*\right)}\right|\le\frac{\|H\|_{BV}\left(1+e^{H(s)-H(t)}\right)}{\prod_{(s,t]}\left(1+\mathrm{d}\Phi^*\right)}\,, \tag{3.1.12}$$

$$\begin{aligned}&\left|\frac{1}{\prod_{(s,t]}\left(1+\mathrm{d}(\Phi^*+H)\right)}-\frac{1}{\prod_{(s,t]}\left(1+\mathrm{d}\Phi^*\right)}+\frac{\int_{(s,t]}\frac{1}{1+\Delta\Phi^*(u)}\,\mathrm{d}H(u)}{\prod_{(s,t]}\left(1+\mathrm{d}\Phi^*\right)}\right|\\&\le\frac{\|H\|_{BV}^2\left(1+e^{H(s)-H(t)}\right)}{\prod_{(s,t]}\left(1+\mathrm{d}\Phi^*\right)}\,.\end{aligned} \tag{3.1.13}$$

*Proof.* The proofs of (3.1.10) and (3.1.12) are just simpler forms of the proofs of (3.1.11) and (3.1.13). Thus, only the latter are presented here.

Applying the equation of Duhamel (see equation (14) of section 2 in Gill (1994)) two times and using (3.1.2) and (A.4.3) leads for $\|H\|_{BV} \le C_{\Phi^*} - \widetilde{C}_{\Phi^*}$ to

$$\begin{aligned}&\left|\prod_{(s,t]}\left(1+\mathrm{d}(\Phi^*+H)\right)-\prod_{(s,t]}\left(1+\mathrm{d}\Phi^*\right)-\prod_{(s,t]}\left(1+\mathrm{d}\Phi^*\right)\int_{(s,t]}\frac{1}{1+\Delta\Phi^*(u)}\,\mathrm{d}H(u)\right|\\
&=\left|\prod_{(s,t]}\left(1+\mathrm{d}(\Phi^*+H)\right)-\prod_{(s,t]}\left(1+\mathrm{d}\Phi^*\right)-\int_{(s,t]}\prod_{(s,u)}\left(1+\mathrm{d}\Phi^*\right)\prod_{(u,t]}\left(1+\mathrm{d}\Phi^*\right)\mathrm{d}H(u)\right|\\
&=\left|\int_{(s,t]}\left(\int_{(s,u)}\prod_{(s,v)}\left(1+\mathrm{d}(\Phi^*+H)\right)\prod_{(v,u)}\left(1+\mathrm{d}\Phi^*\right)\mathrm{d}H(v)\right)\prod_{(u,t]}\left(1+\mathrm{d}\Phi^*\right)\mathrm{d}H(u)\right|\\
&\le\int_{(s,t]}\left(\int_{(s,u)}\prod_{(s,v)}\left(1+\mathrm{d}(\Phi^*+H)\right)\prod_{(v,u)}\left(1+\mathrm{d}\Phi^*\right)\mathrm{d}|H|(v)\right)\prod_{(u,t]}\left(1+\mathrm{d}\Phi^*\right)\mathrm{d}|H|(u)\\
&\le\frac{e^{\|H\|_{BV}}}{1+C_{\Phi^*}}\int_{(s,t]}\left(\int_{(s,u)}\prod_{(s,v)}\left(1+\mathrm{d}\Phi^*\right)\prod_{(v,u)}\left(1+\mathrm{d}\Phi^*\right)\mathrm{d}|H|(v)\prod_{(u,t]}\left(1+\mathrm{d}\Phi^*\right)\right)\mathrm{d}|H|(u)\\
&\le\mathrm{const}\int_{(s,t]}\int_{(s,u)}\frac{\prod_{(s,t]}\left(1+\mathrm{d}\Phi^*\right)}{\left(1+\Delta\Phi^*(v)\right)\left(1+\Delta\Phi^*(u)\right)}\,\mathrm{d}|H|(v)\,\mathrm{d}|H|(u)\\
&\le\mathrm{const}\prod_{(s,t]}\left(1+\mathrm{d}\Phi^*\right)\frac{\|H\|_{BV}^2}{(1+C_{\Phi^*})^2}\\
&\le\mathrm{const}\,\|H\|_{BV}^2\prod_{(s,t]}\left(1+\mathrm{d}\Phi^*\right).\end{aligned}$$

Analogously, applying the equation of Duhamel only one time,

$$\left| \prod_{(s,t]} \big(1 + \mathrm{d}(\Phi^* + H)\big) - \prod_{(s,t]} \big(1 + \mathrm{d}\Phi^*\big) \right| \le \mathrm{const}\, \|H\|_{BV} \prod_{(s,t]} \big(1 + \mathrm{d}\Phi^*\big) .$$

These two inequalities combined with the inequality

$$\left| \frac{1}{a} - \frac{1}{b} + \frac{c}{b^2} \right| \le \left| \frac{(a-b)^2}{ab^2} \right| + \left| \frac{a-b-c}{b^2} \right| , \qquad \forall a, b > 0\,, c \in \mathbb{R}\,,$$

property (3.1.5), and $\|H\|_{BV} \le C_{\Phi^*} - \widetilde{C}_{\Phi^*}$ yield

$$\begin{aligned}
&\left| \frac{1}{\prod_{(s,t]} \big(1 + \mathrm{d}(\Phi^* + H)\big)} - \frac{1}{\prod_{(s,t]} \big(1 + \mathrm{d}\Phi^*\big)} + \frac{\int_{(s,t]} \frac{1}{1+\Delta\Phi^*(u)}\, \mathrm{d}H(u)}{\prod_{(s,t]} \big(1 + \mathrm{d}\Phi^*\big)} \right| \\
&\le \frac{\Big(\mathrm{const}\, \|H\|_{BV} \prod_{(s,t]} \big(1 + \mathrm{d}\Phi^*\big)\Big)^2}{\prod_{(s,t]} \big(1 + \mathrm{d}(\Phi^* + H)\big) \Big(\prod_{(s,t]} \big(1 + \mathrm{d}\Phi^*\big)\Big)^2} + \frac{\mathrm{const}\, \|H\|_{BV}^2 \prod_{(s,t]} \big(1 + \mathrm{d}\Phi^*\big)}{\Big(\prod_{(s,t]} \big(1 + \mathrm{d}\Phi^*\big)\Big)^2} \\
&\le \frac{\mathrm{const}\, e^{\mathrm{const}\, \|H\|_{BV}} \|H\|_{BV}^2}{\prod_{(s,t]} \big(1 + \mathrm{d}\Phi^*\big)} + \frac{\mathrm{const}\, \|H\|_{BV}^2}{\prod_{(s,t]} \big(1 + \mathrm{d}\Phi^*\big)} \\
&\le \frac{\mathrm{const}\, \|H\|_{BV}^2}{\prod_{(s,t]} \big(1 + \mathrm{d}\Phi^*\big)} .
\end{aligned}$$

Now assume that $H$ is an arbitrary but continuous function of $BV_{\leftarrow}$. Because of (A.4.1)

$$\prod_{(s,t]} \big(1 + \mathrm{d}(\Phi^* + H)\big) = e^{H(t)-H(s)} \prod_{(s,t]} \big(1 + \mathrm{d}\Phi^*\big) .$$

Further on, the finite variation of $\Phi^*$ on $(s,t]$ implies that the set $\{u \in (s,t] : \Delta\Phi^*(u) \neq 0\}$ has, at most, a countable number of elements and is therefore a null set with respect to $H$. Consequently,

$$\int_{(s,t]} \frac{1}{1+\Delta\Phi^*(u)}\, \mathrm{d}H(u) = \int_{(s,t]} \mathrm{d}H(u) = H(t) - H(s) .$$

Since $|e^x - 1 - x| \le |x|^2\,(1 + e^x)$ for $x \in \mathbb{R}$, one gets

$$\begin{aligned}
&\left| \frac{1}{\prod_{(s,t]} \big(1 + \mathrm{d}(\Phi^* + H)\big)} - \frac{1}{\prod_{(s,t]} \big(1 + \mathrm{d}\Phi^*\big)} + \frac{\int_{(s,t]} \frac{1}{1+\Delta\Phi^*(u)}\, \mathrm{d}H(u)}{\prod_{(s,t]} \big(1 + \mathrm{d}\Phi^*\big)} \right| \\
&= \frac{1}{\prod_{(s,t]} \big(1 + \mathrm{d}\Phi^*\big)} \left| e^{H(s)-H(t)} - 1 - \big(H(s) - H(t)\big) \right| \\
&\le \frac{|H(t) - H(s)|^2 \Big(1 + e^{H(s)-H(t)}\Big)}{\prod_{(s,t]} \big(1 + \mathrm{d}\Phi^*\big)} .
\end{aligned}$$

□

**Proposition 3.1.5.** *Define for an arbitrary but fixed* $s \geq 0$ *the Borel-measures*

$$\begin{aligned} \mu_{SB}(A) &:= \sum_{z\in\mathcal{S}} \int_{A\cap[s,\infty)} p_{yz}(s,t)\, F_z(\mathrm{d}t)\,, \qquad \forall A \in \mathfrak{B}(\mathbb{R})\,, \\ \mu_{DB}(A) &:= \sum_{(z,\zeta)\in J} \int_{A\cap(s,\infty)} D_{z\zeta}(t)\; p_{yz}(s,t-0)\; q_{z\zeta}(\mathrm{d}t)\,, \qquad \forall A \in \mathfrak{B}(\mathbb{R})\,. \end{aligned} \tag{3.1.14}$$

*Then, the mapping*

$$\nu_s : \mathcal{E}_{\Phi^*} \to L_1\Big(|\mu_{SB}| + |\mu_{DB}| \circ DT^{-1}\Big)\,, \quad H \mapsto \left( t \mapsto \frac{1}{\prod_{(s,t]} \big(1 + \mathrm{d}(\Phi^* + H)\big)} \right) \tag{3.1.15}$$

*is Fréchet differentiable at zero with differential*

$$\mathrm{D}_0\nu_s : BV_{\leftarrow} \to L_1\Big(|\mu_{SB}| + DT(|\mu_{DB}|)\Big)\,, \quad H \mapsto \left( t \mapsto -\frac{\displaystyle\int_{(s,t]} \frac{1}{1+\Delta\Phi^*(u)}\,\mathrm{d}H(u)}{\displaystyle\prod_{(s,t]} \big(1 + \mathrm{d}\Phi^*\big)} \right). \tag{3.1.16}$$

*Proof.* The mapping (3.1.15) is well defined as for any $H \in \mathcal{E}_{\Phi^*}$

$$\begin{aligned} &\|\nu_s(H)\|_{L_1\left(|\mu_{SB}|+|\mu_{DB}|\circ DT^{-1}\right)} \\ &= \int \frac{1}{\prod_{(s,t]} \big(1 + \mathrm{d}(\Phi^* + H)\big)}\, |\mu_{SB}|(\mathrm{d}t) + \int \frac{1}{\prod_{(s,t]} \big(1 + \mathrm{d}(\Phi^* + H)\big)}\, |\mu_{DB}| \circ DT^{-1}(\mathrm{d}t) \\ &= \int \frac{K_{\Phi^*+H}(s)}{K_{\Phi^*+H}(t)}\, |\mu_{SB}|(\mathrm{d}t) + \int \frac{K_{\Phi^*+H}(s)}{K_{\Phi^*+H}(DT(t))}\, |\mu_{DB}|(\mathrm{d}t) \\ &\leq \mathrm{const}\, e^{\mathrm{const}\,\|H\|_{BV}}\, K_{\Phi^*+H}(s) \big(\mathrm{InCo_{SB}} + \mathrm{InCo_{SDB}}\big) < \infty\,, \end{aligned}$$

where the inequality is due to (3.1.5) and (3.1.6), and the constants $\mathrm{InCo_{SB}}$ and $\mathrm{InCo_{DB}}$ are defined according to (1.2.6) and (1.2.7). With (3.1.11) and the above estimation

$$\begin{aligned} &\frac{\big\|\nu_s(H) - \nu_s(0) - \mathrm{D}_0\nu_s(H)\big\|_{L_1\left(|\mu_{SB}|+|\mu_{DB}|\circ DT^{-1}\right)}}{\|H\|_{BV}} \\ &\leq \frac{\mathrm{const}\, \|H\|_{BV}^2\, \|\nu_s(0)\|_{L_1\left(|\mu_{SB}|+|\mu_{DB}|\circ DT^{-1}\right)}}{\|H\|_{BV}} \\ &\leq \mathrm{const}\, \|H\|_{BV} \end{aligned} \tag{3.1.17}$$

for all $H \in \mathcal{B}_{\Phi^*}$. The differential $\mathrm{D}_0\nu_s$ is linear and continuous, the latter property is due to

$$\|\mathrm{D}_0\nu_s(H)\|_{L_1\left(|\mu_{SB}|+|\mu_{DB}|\circ DT^{-1}\right)} \leq \frac{\|H\|_{BV}}{1+C_{\Phi^*}}\, \|\nu_s(0)\|_{L_1\left(|\mu_{SB}|+|\mu_{DB}|\circ DT^{-1}\right)} = \mathrm{const}\, \|H\|_{BV}\,.$$

□

*Proof of Theorem* 3.1.3. Let $F$ be the sum of two Fréchet differentiable mappings of the form (A.3.4) with measures $\mu_{SB}$ and $\mu_{DB} \circ DT^{-1}$ according to (3.1.14),

$$F : L_1\Big(|\mu_{SB}| + |\mu_{DB}| \circ DT^{-1}\Big) \to \mathbb{R}\,, \quad v \mapsto \int v \,\mathrm{d}\mu_{SB} + \int v \,\mathrm{d}\mu_{DB} \circ DT^{-1}\,.$$

As stated in Example A.3.3, $F$ is Fréchet differentiable at zero with differential $\mathrm{D}_0 F = F$. As $V^{\Phi^*}_{y,s} = F \circ \nu_s$, the chain rule (A.3.6) yields

$$\mathrm{D}_0 V^{\Phi^*}_{y,s} = \mathrm{D}_0\big(F \circ \nu_s\big) = \mathrm{D}_{\nu_s(0)} F \circ \mathrm{D}_0 \nu_s = F \circ \mathrm{D}_0 \nu_s\,.$$

Now (3.1.16) leads to (3.1.9). □

**Remark 3.1.6.** In case $H$ is an arbitrary but continuous element of $BV_{\leftarrow}$, property (3.1.13) yields analogously to (3.1.17)

$$\begin{aligned} &\Big\|\nu_s(H) - \nu_s(0) - \mathrm{D}_0\nu_s(H)\Big\|_{L_1\left(|\mu_{SB}|+|\mu_{DB}|\circ DT^{-1}\right)} \\ &\le \Big\|\|H\|^2_{BV}\Big(1 + e^{H(s)-H(\cdot)}\Big)\nu_s(0)\Big\|_{L_1\left(|\mu_{SB}|+|\mu_{DB}|\circ DT^{-1}\right)}\,. \end{aligned}$$

With $F$ being defined as in the above proof, one gets

$$\begin{aligned} &\Big|V^{\Phi^*}_{y,s}(H) - V^{\Phi^*}_{y,s}(0) - \mathrm{D}_0 V^{\Phi^*}_{y,s}(H)\Big| \\ &= \Big|F \circ \nu_s(H) - F \circ \nu_s(0) - F \circ \mathrm{D}_0\nu_s(H)\Big| \\ &= \Big|F \circ \big(\nu_s(H) - \nu_s(0) - \mathrm{D}_0\nu_s(H)\big)\Big| \\ &\le \Big\|\|H\|^2_{BV}\Big(1 + e^{H(s)-H(\cdot)}\Big)\nu_s(0)\Big\|_{L_1\left(|\mu_{SB}|+|\mu_{DB}|\circ DT^{-1}\right)}\,, \end{aligned} \tag{3.1.18}$$

which is a useful property for later purposes.

As Fréchet differentiability implies Gâteaux differentiability with coinciding differentials, the task is now to find a representation for $\mathrm{D}_0 V^{\Phi^*}_{y,s}$ in accordance with (2.2.2).

**Theorem 3.1.7.** *The gradient vector II of $V^{\Phi^*}_{y,s}$ exists at zero and has the form*

$$\begin{aligned} \nabla_0 V^{\Phi^*}_{y,s}(u) = &-\sum_{z\in\mathcal{S}} \int_{[s,\infty)} \frac{\mathbf{1}_{(s,t]}(u)}{1+\Delta\Phi^*(u)}\,\frac{K_{\Phi^*}(s)}{K_{\Phi^*}(t)}\,p_{yz}(s,t)\,F_z(\mathrm{d}t) \\ &- \sum_{(z,\zeta)\in J} \int_{(s,\infty)} \frac{\mathbf{1}_{(s,DT(t)]}(u)}{1+\Delta\Phi^*(u)}\,\frac{K_{\Phi^*}(s)}{K_{\Phi^*}(DT(t))}\,D_{z\zeta}(t)\,p_{yz}(s,t-0)\,q_{z\zeta}(\mathrm{d}t) \\ \overset{a.s.}{=} &-\frac{\mathbf{1}_{(s,\infty)}(u)}{1+\Delta\Phi^*(u)}\,V_{y,s,u}\,, \qquad \forall\, u \in \mathbb{R}\,. \end{aligned} \tag{3.1.19}$$

That means the gradient vector at $u > s$ is equal to the negative of the present value at time $s$ of all payments falling due not before $u$ and triggered not before $s$, normalized by the relative height of a possible jump of the accumulation factor $K$ at $u$.

*Proof.* Apply Fubinis Theorem to (3.1.9). Note that the functional $\mathrm{D}_0 V_{y,s}^{\Phi^*}$ is continuous at zero, that is, $|\mathrm{D}_0 V_{y,s}^{\Phi^*}(H)| \le \mathrm{const}\, \|H\|_{BV} < \infty$ for all $H \in BV_{\leftarrow}$. □

If the Fréchet differential $\mathrm{D}_0 V_{y,s}^{\Phi^*}$ satisfied Condition 2.2.5, Theorem 2.2.6 would already imply the existence of the gradient vector $\nabla_0 V_{y,s}^{\Phi^*}$, which could then be easily calculated by $\nabla_0 V_{y,s}^{\Phi^*}(u) = \mathrm{D}_0 V_{y,s}^{\Phi^*}(\mathbf{1}_{[u,\infty)})$. In fact, one can verify Condition 2.2.5, but calculating the gradient vector by Fubinis Theorem turns out to be easier here.

**Remark 3.1.8** (Computing the gradient vector)**.** Though having the explicit formula (3.1.19) for the gradient vector $\nabla_0 V_{y,s}^{\Phi^*}$, it can still be challenging to compute it. Since (3.1.19) looks very similar to (1.2.11), one could aim to find an analogon to the Thiele Integral/Differential Equations (cf. Milbrodt and Helbig (1999), section 10.C). In case payments due to transitions are paid immediately, that is, $DT = Id$, there is no need for it: Using the Chapman-Kolmogorov-equations, the gradient vector has the form

$$\begin{aligned}&\nabla_0 V_{y,s}^{\Phi^*}(u)\\ &= -\frac{\mathbf{1}_{(s,\infty)}(u)}{1+\Delta\Phi^*(u)}\frac{K_{\Phi^*}(u)}{K_{\Phi^*}(s)}\left(\sum_{z\in\mathcal{S}} p_{yz}(s,u)\,V_{z,u} + \sum_{(z,\zeta)\in J} D_{z\zeta}(u)\,p_{yz}(s,u-0)\,\Delta q_{z\zeta}(u)\right).\end{aligned}$$

The only challenge in this formula is to compute the prospective reserves $V_{z,u}$ and the transition probabilities $p_{yz}(s,u-0)$. For the former, use the Thiele Integral/Differential Equations already mentioned. Provided the transition intensity matrix $\mu$ exists, the latter are easily calculable with the Kolmogorov Forward/Backward Equations.

## 3.2 Gradient vector with respect to a single transition

In this section, the prospective reserve is regarded as a functional solely with respect to an arbitrary but fixed transition $(g,l) \in J$. As stated in (3.0.1), let the cumulative intensity decompose to $q_{gl} = q_{gl}^* + H$, where $q_{gl}^* \in BVC_{\leftarrow}^{+}$ is a starting point satisfying the integrability conditions (1.2.6) and (1.2.7), and $H \in BV_{\leftarrow}$ is a deviation from $q_{gl}^*$. Let $q^*$ be a regular transition intensity matrix.

A remark on notation:

- Denote by $[M]_{ij}$ the element of the $i$-th row and the $j$-th column of matrix $M$.
- Let $\mathbb{I}_{ij}$ be a zero matrix except of the element in the $i$-th row and the $j$-th column being $1$.

- Let $|\cdot|$ denote the maximum-row-sum norm when it is applied to a matrix. Note that the same symbol applied to a measure $\mu$ denotes the measure $|\mu| := \mu_+ + \mu_-$, where $\mu = \mu_+ - \mu_-$ is the unique Jordan-Hahn decomposition of $\mu$.

With the cumulative intensity $q^*_{gl}$ being shifted by $H$, the matrix of the cumulative transition intensities gets

$$q(\cdot) = q^*(\cdot) + (\mathbb{I}_{gl} - \mathbb{I}_{gg})\, H(\cdot) =: q(\cdot; H)\,, \tag{3.2.1}$$

since $q_{zz} = -\sum_{z \neq \zeta} q_{z\zeta}$ for all $z \in \mathcal{S}$. Via (1.2.2) it corresponds to

$$p(s,t) = \prod_{(s,t]} \big(\mathbb{I} + \mathrm{d}q(\cdot; H)\big) =: p(s,t;H)\,. \tag{3.2.2}$$

For a shorter notation write $p^*(s,t) := p(s,t;0)$. Since only the influence of changes of $q_{gl}$ shall be studied here, define

$$V^{q^*_{gl}}_{y,s} : BV_{\leftarrow} \to \mathbb{R}\,, \quad H \mapsto V_{y,s}\big[\Phi^*, q^* + (\mathbb{I}_{gl} - \mathbb{I}_{gg})\, H\big]\,. \tag{3.2.3}$$

The following proposition shows that this functional is well-defined.

**Proposition 3.2.1.** *Let $t \mapsto D_{gl}(t)/K(DT(t))$ be bounded on $[0,\infty)$. For each $H \in BV_{\leftarrow}$, $y \in \mathcal{S}$, and $s \geq 0$, one has $|V^{q^*_{gl}}_{y,s}(H)| < \infty$.*

The proof of that Proposition is based on the following useful Proposition:

**Proposition 3.2.2.** *For all $H \in BV_{\leftarrow}$ and $s < t$, $\big|p(s,t;H)\big| \leq e^{2\,\|H\|_{BV}}$.*

*Proof.* Equation (14) in Gill (1994, p. 126), the submultiplicativity of the maximum-row-sum-norm, and the property of the stochastic matrix $p^*$ to have row sums of one, yield

$$\begin{aligned}
\big|p(s,t;H)\big| = \Bigg| &\prod_{(s,t]} \big(\mathbb{I} + \mathrm{d}q^*\big) + \sum_{m=1}^{\infty} \int_{s<u_1<\cdots<u_m\leq t} \prod_{(s,u_1)} (\mathbb{I} + \mathrm{d}q^*)\, (\mathbb{I}_{gl} - \mathbb{I}_{gg})\, \mathrm{d}H(u_1) \cdot \\
&\cdot \prod_{(u_1,u_2)} (\mathbb{I} + \mathrm{d}q^*)\, (\mathbb{I}_{gl} - \mathbb{I}_{gg})\, \mathrm{d}H(u_2) \,\ldots\, (\mathbb{I}_{gl} - \mathbb{I}_{gg})\, \mathrm{d}H(u_m) \prod_{(u_m,t]} (\mathbb{I} + \mathrm{d}q^*) \Bigg| \\
\leq 1 + &\sum_{m=1}^{\infty} \int_{s<u_1<\cdots<u_m\leq t} \big|p^*(0, u_1 - 0)\big|\, 2\,\mathrm{d}|H|(u_1) \cdot \\
&\cdot \big|p^*(u_1, u_2 - 0)\big|\, 2\,\mathrm{d}|H|(u_2) \,\ldots\, 2\,\mathrm{d}|H|(u_m)\, \big|p^*(u_m, t)\big| \\
= 1 + &\sum_{m=1}^{\infty} \int_{s<u_1<\cdots<u_m\leq t} 2\,\mathrm{d}|H|(u_1)\, 2\,\mathrm{d}|H|(u_2) \,\ldots\, 2\,\mathrm{d}|H|(u_m) \\
= 1 + &\sum_{m=1}^{\infty} \frac{1}{m!} \int_{(s,t]^m} 2\,\mathrm{d}|H|(u_1)\, 2\,\mathrm{d}|H|(u_2) \,\ldots\, 2\,\mathrm{d}|H|(u_m) \\
\leq &\sum_{m=0}^{\infty} \frac{\big(2\,\|H\|_{BV}\big)^m}{m!}\,.
\end{aligned}$$

□

*Proof of Proposition* 3.2.1. Applying Proposition 3.2.2 yields

$$\left|\sum_{z\in\mathcal{S}}\int_{(s,\infty)}\frac{1}{K(t)}\,p_{yz}(s,t;H)\,F_z(\mathrm{d}t)\right|\le e^{2\,\|H\|_{BV}}\,\mathrm{InCo}_{\mathrm{SB}}<\infty$$

with constant $\mathrm{InCo}_{\mathrm{SB}}$ defined according to (1.2.6). Further on,

$$\begin{aligned}
&\left|\sum_{(z,\zeta)\in J}\int_{(s,\infty)}\frac{1}{K(DT(t))}\,D_{z\zeta}(t)\,p_{yz}(s,t-0;H)\,q_{z\zeta}(\mathrm{d}t;H)\right|\\
&\le\left|\sum_{(z,\zeta)\in J}\int_{(s,\infty)}\frac{1}{K(DT(t))}\,D_{z\zeta}(t)\,p_{yz}(s,t-0;H)\,q^*_{z\zeta}(\mathrm{d}t)\right|\\
&\quad+\left|\int_{(s,\infty)}\frac{1}{K(DT(t))}\,D_{gl}(t)\,p_{yg}(s,t-0;H)\,\mathrm{d}H(t)\right|\\
&\le e^{2\,\|H\|_{BV}}\mathrm{InCo}_{\mathrm{DB}}+\sup_{t\in\mathbb{R}}\left\{\frac{D_{gl}(t)}{K(DT(t))}\right\}\,\mathrm{e}^{2\,\|H\|_{BV}}\,\|H\|_{BV}<\infty\,,
\end{aligned}$$

with constant $\mathrm{InCo}_{\mathrm{DB}}$ defined according to (1.2.7). □

A necessary condition for $V_{y,s}^{q^*_{gl}}$ having a gradient vector II at zero is its Gâteaux differentiability. In fact, it is even Fréchet differentiable at zero.

**Theorem 3.2.3.** *Let* $t\mapsto D_{gl}(t)/K(DT(t))$ *be bounded on* $[0,\infty)$*. Then, the functional* $V_{y,s}^{q^*_{gl}}$ *is Fréchet differentiable at zero with differential*

$$\begin{aligned}
\mathrm{D}_0V_{y,s}^{q^*_{gl}}(H)=&\sum_{z\in\mathcal{S}}\int_{[s,\infty)}\frac{K(s)}{K(t)}\,\varrho^s_{yz}(t;H)\,F_z(\mathrm{d}t)\\
&+\sum_{(z,\zeta)\in J}\int_{(s,\infty)}\frac{K(s)}{K(DT(t))}\,D_{z\zeta}(t)\,\varrho^s_{yz}(t-0;H)\,q^*_{z\zeta}(\mathrm{d}t)\\
&+\int_{(s,\infty)}\frac{K(s)}{K(DT(t))}\,D_{gl}(t)\,p^*_{yg}(s,t-0)\,\mathrm{d}H(t)\,,
\end{aligned}\tag{3.2.4}$$

*for* $H\in BV_{\leftarrow}$*, where*

$$\varrho^s_{yz}(t;H):=\left[\int_{(s,t]}p^*(s,u-0)\,(\mathbb{I}_{gl}-\mathbb{I}_{gg})\,p^*(u,t)\,\mathrm{d}H(u)\right]_{yz}\,,\qquad (y,z)\in\mathcal{S}\times\mathcal{S}\,,s\le t\,.$$

The proof of this theorem is split into several pieces – following now – with some of them being formulated as an own proposition:

**Proposition 3.2.4.** *For all* $H\in BV_{\leftarrow}$ *with* $\|H\|_{BV}\le c<\infty$ *for an arbitrary but fixed constant* $c$ *and all* $-\infty<s<t<\infty$

$$\left|p(s,t;H)-p^*(s,t)\right|\le\mathrm{const}\,\|H\|_{BV}\,,\tag{3.2.5}$$

$$\left| p(s,t;H) - p^*(s,t) - \int_{(s,t]} p^*(s,u-0)\,(\mathbb{I}_{gl} - \mathbb{I}_{gg})\,p^*(u,t)\,\mathrm{d}H(u) \right| \leq \text{const}\,\|H\|_{BV}^2\,. \tag{3.2.6}$$

*For all $H \in BV_{\leftarrow}$ with $q(\cdot;H)$ being a regular cumulative transition intensity matrix and all $-\infty < s < t < \infty$*

$$\left| p(s,t;H) - p^*(s,t) \right| \leq 2\,\|H\|_{BV}\,, \tag{3.2.7}$$

$$\left| p(s,t;H) - p^*(s,t) - \int_{(s,t]} p^*(s,u-0)\,(\mathbb{I}_{gl} - \mathbb{I}_{gg})\,p^*(u,t)\,\mathrm{d}H(u) \right| \leq 4\,\|H\|_{BV}^2\,. \tag{3.2.8}$$

*Proof.* The proofs of (3.2.5) and (3.2.7) are just simpler forms of the proofs of (3.2.6) and (3.2.8). Thus, only the latter are presented here.

Applying the equation of Duhamel (see equation (14) of section 2 in Gill (1994)) two times, and using Proposition 3.2.2 and the property of $p^*$ being a stochastic matrix (that is the row sums are equal to one) leads for $\|H\|_{BV} \leq c < \infty$ to

$$\begin{aligned}
&\left| p(s,t;H) - p^*(s,t) - \int_{(s,t]} p^*(s,u-0)\,(\mathbb{I}_{gl} - \mathbb{I}_{gg})\,p^*(u,t)\,\mathrm{d}H(u) \right| \\
&= \left| \int_{(s,t]} \left( \int_{(s,u)} p(s,v-0;H)\,(\mathbb{I}_{gl} - \mathbb{I}_{gg})\,p^*(v,u-0)\,\mathrm{d}H(v) \right) (\mathbb{I}_{gl} - \mathbb{I}_{gg})\,p^*(u,t)\,\mathrm{d}H(u) \right| \\
&\leq \int_{(s,t]} \left( \int_{(s,u)} \left| p(s,v-0;H) \right| 2 \left| p^*(v,u-0) \right| \mathrm{d}|H|(v) \right) 2 \left| p^*(u,t) \right| \mathrm{d}|H|(u) \\
&\leq 4\,e^{2c} \int_{(s,t]} \left( \int_{(s,u)} \mathrm{d}|H|(v) \right) \mathrm{d}|H|(u) \\
&\leq 4\,e^{2c}\,\|H\|_{BV}^2\,.
\end{aligned}$$

If $q(\cdot;H)$ is a regular cumulative transition intensity matrix, one has $|p(s,v-0;H)| = 1$, which leads to the upper bound (3.2.8). □

**Proposition 3.2.5.** *The mapping*

$$\begin{aligned}
\rho^s : BV_{\leftarrow} \to B &:= \left\{ f : \mathbb{R} \to \mathbb{R}^{|\mathcal{S}|\times|\mathcal{S}|} \,\middle|\, \|f\|_B := \sup_{t\in\mathbb{R}} |f(t)| < \infty \right\}, \\
H &\mapsto \left( t \mapsto \mathbf{1}_{(s,\infty)}(t)\,p(s,t;H) \right)
\end{aligned} \tag{3.2.9}$$

*is Fréchet differentiable at zero with differential*

$$\mathrm{D}_0\rho^s(H) = \int_{(s,\cdot]} p^*(s,u-0)\,(\mathbb{I}_{gl} - \mathbb{I}_{gg})\,p^*(u,\cdot)\,\mathrm{d}H(u)\,, \qquad \forall H \in BV_{\leftarrow}\,. \tag{3.2.10}$$

*Proof.* The mapping $\rho^s$ is well defined as by means of Proposition 3.2.2 $\left|\rho^s(H)(t)\right| \leq \exp(2\,\|H\|_{BV})$ for all $t \in \mathbb{R}$. Property (3.2.6) yields for $\|H\|_{BV} \leq c < \infty$

$$\frac{\left|\rho^s(H)(t) - \rho^s(0)(t) - \mathrm{D}_0\rho^s(H)(t)\right|}{\|H\|_{BV}} \leq \text{const}\,\|H\|_{BV}\,, \qquad \forall\, t \in \mathbb{R}\,. \tag{3.2.11}$$

According to Definition A.3.2, it is left to show that the functional $H \mapsto \mathrm{D}_0\rho^s(H)$ is linear and continuous. Linearity is obvious. Continuity is here equivalent to boundedness,

$$\left\|\mathrm{D}_0\rho^s(H)\right\|_B \leq \sup_{t\in\mathbb{R}} \int \mathbf{1}_{(s,t]}(u)|p^*(s,u-0)|\,|\mathbb{I}_{gl} - \mathbb{I}_{gg}|\,|p^*(u,t)|\,\mathrm{d}|H|(u) \leq 2\,\|H\|_{BV}\,. \tag{3.2.12}$$

□

One gets an analogous result when altering $\rho^s(H)$ to

$$\overline{\rho}^s(H) := \Big(t \mapsto \mathbf{1}_{(s,\infty)}(t)\, p(s,t-0;H)\Big)\,, \qquad H \in BV_{\leftarrow}\,. \tag{3.2.13}$$

The corresponding Fréchet differential has the form

$$\mathrm{D}_0\overline{\rho}^s(H) = \int_{(s,\cdot-0]} p^*(s,u-0)\,(\mathbb{I}_{gl} - \mathbb{I}_{gg})\,p^*(u,\cdot-0)\,\mathrm{d}H(u)\,, \qquad H \in BV_{\leftarrow}\,. \tag{3.2.14}$$

*Proof of Theorem* 3.2.3. Denoting by $F_{\eta_z}$ and $F_{\eta_{z\zeta}}$ Fréchet differentiable mappings of the form (A.3.4) with measures

$$\begin{aligned}
\eta_z(A) &:= \int_A \frac{K(s)}{K}\,\mathrm{d}F_z\,, \qquad \forall\, A \in \mathfrak{B}(\mathbb{R})\,,\\
\eta_{z\zeta}(A) &:= \int_A \frac{K(s)}{K \circ DT} D_{z\zeta}\,\mathrm{d}q^*_{z\zeta}\,, \qquad \forall\, A \in \mathfrak{B}(\mathbb{R})\,,
\end{aligned}$$

the prospective reserve at $q = q^* + (\mathbb{I}_{gl} - \mathbb{I}_{gg})\,H$ is equal to

$$\begin{aligned}
V^{q^*_{gl}}_{y,s}(H) &= \sum_{z\in\mathcal{S}} \int_{[s,\infty)} p_{yz}(s,t;H)\,\eta_z(\mathrm{d}t) + \sum_{(z,\zeta)\in J} \int_{(s,\infty)} p_{yz}(s,t-0;H)\,\eta_{z\zeta}(\mathrm{d}t)\\
&\quad + \int_{(s,\infty)} \frac{K(s)}{K \circ DT(t)}\,D_{gl}(t)\,p_{yg}(s,t-0;H)\,\mathrm{d}H(t)\\
&= \sum_{z\in\mathcal{S}} F_{\eta_z}\big(p_{yz}(s,\cdot;H)\big) + \sum_{(z,\zeta)\in J} F_{\eta_{z\zeta}}\big(p_{yz}(s,\cdot-0;H)\big)\\
&\quad + \int_{(s,\infty)} \frac{K(s)}{K \circ DT(t)}\,D_{gl}(t)\,p_{yg}(s,t-0;H)\,\mathrm{d}H(t)\,.
\end{aligned} \tag{3.2.15}$$

By means of the linearity of Fréchet differentials, the chain rule (A.3.6), the property $\mathrm{D}_0 F = F$ for mappings in accordance with Example A.3.3, and Proposition 3.2.5, the Fréchet differential of the first two addends has the form

$$\begin{aligned}
&\mathrm{D}_0 \left( \sum_{z\in\mathcal{S}} F_{\eta_z}\big(p_{yz}(s,\cdot;H)\big) + \sum_{(z,\zeta)\in J} F_{\eta_{z\zeta}}\big(p_{yz}(s,\cdot-0;H)\big)\right)(H) \\
&= \sum_{z\in\mathcal{S}} \mathrm{D}_0\Big(F_{\eta_z}\circ[\rho^s]_{yz}\Big)(H) + \sum_{(z,\zeta)\in J} \mathrm{D}_0\Big(F_{\eta_{z\zeta}}\circ[\overline{\rho}^s]_{yz}\Big)(H) \\
&= \sum_{z\in\mathcal{S}} \Big(\mathrm{D}_{[\rho^s(0)]_{yz}} F_{\eta_z}\Big)\circ\Big(\mathrm{D}_0[\rho^s]_{yz}\Big)(H) + \sum_{(z,\zeta)\in J} \Big(\mathrm{D}_{[\overline{\rho}^s(0)]_{yz}} F_{\eta_{z\zeta}}\Big)\circ\Big(\mathrm{D}_0[\overline{\rho}^s]_{yz}\Big)(H) \\
&= \sum_{z\in\mathcal{S}} F_{\eta_z}\circ\Big(\mathrm{D}_0[\rho^s]_{yz}\Big)(H) + \sum_{(z,\zeta)\in J} F_{\eta_{z\zeta}}\circ\Big(\mathrm{D}_0[\overline{\rho}^s]_{yz}\Big)(H) \\
&= \sum_{z\in\mathcal{S}} \int_{[s,\infty)} \frac{K(s)}{K(t)}\,\varrho^s_{yz}(t;H)\,F_z(\mathrm{d}t) + \sum_{(z,\zeta)\in J} \int_{(s,\infty)} \frac{K(s)}{K(DT(t))}\,D_{z\zeta}(t)\,\varrho^s_{yz}(t-0;H)\,q^*_{z\zeta}(\mathrm{d}t)\,,
\end{aligned}$$

which is equal to the first two addends of (3.2.4). For the third addend in (3.2.15) Proposition 3.2.4 yields

$$\begin{aligned}
&\left|\int_{(s,\infty)} \frac{K(s)}{K(DT(t))} D_{gl}(t)\,p_{yg}(s,t-0;H)\,\mathrm{d}H(t) - \int_{(s,\infty)} \frac{K(s)}{K(DT(t))} D_{gl}(t)\,p^*_{yg}(s,t-0)\,\mathrm{d}H(t)\right| \\
&\le \int \frac{K(s)}{K(DT(t))}\,D_{gl}(t)\,\Big|p_{yg}(s,t-0;H) - p^*_{yg}(s,t-0)\Big|\,\mathrm{d}|H|(t) \\
&\le K(s)\,\sup_{t\in\mathbb{R}}\left\{\frac{D_{gl}}{K\circ DT}(t)\right\}\,\mathrm{const}\,\|H\|_{BV}\int \mathrm{d}|H|(t) \\
&\le \mathrm{const}\,\|H\|^2_{BV}
\end{aligned}$$

for all $\|H\|_{BV} \le c < \infty$. It is left to show that the differential is linear and continuous. The linearity is a consequence of the linearity of integration. The continuity is equivalent to boundedness, which holds because of (3.2.12), the boundedness of $D_{gl}(\cdot)/K(DT(\cdot))$, and the integrability conditions (1.2.6) and (1.2.7). □

**Remark 3.2.6.** In case $H \in BV_{\leftarrow}$ is chosen in such a way that $q(\cdot;H)$ is a *regular* cumulative transition intensity matrix, property (3.2.8) allows for replacing the upper bound in (3.2.11) by $4\,\|H\|_{BV}$ without the necessity of an upper bound $\|H\|_{BV} \le c$. With that and (3.2.7) one gets – following the lines of the above proof –

$$\begin{aligned}
&\left|V^{q^*_{gl}}_{y,s}(H) - V^{q^*_{gl}}_{y,s}(0) - \mathrm{D}_0 V^{q^*_{gl}}_{y,s}(H)\right| \\
&\le 4\,\|H\|^2_{BV}\,K(s)\,\big(\mathrm{InCo_{SB}} + \mathrm{InCo_{DB}}\big) + K(s)\,\sup_{t>0}\left\{\frac{D_{gl}}{K\circ DT}(t)\right\}\,2\,\|H\|^2_{BV}\,,
\end{aligned} \tag{3.2.16}$$

which is a useful property for later purposes.

The functional $V_{y,s}^{q_{gl}^*}$ is not only differentiable at zero, but also has a gradient vector II in accordance with Definition 2.2.1:

**Theorem 3.2.7.** *Let $t \mapsto D_{gl}(t)/K(DT(t))$ be bounded on $[0,\infty)$. Then, the gradient vector II of $V_{y,s}^{q_{gl}^*}$ exists at zero and has the form*

$$\begin{aligned}\nabla_0 V_{y,s}^{q_{gl}^*}(u) &= \mathbf{1}_{(s,\infty)}(u)\, p_{yg}^*(s,u-0)\, K(s) \Bigg\{ \sum_{z\in\mathcal{S}} \int_{[u,\infty)} \frac{1}{K(t)} \Big(p_{lz}^*(u,t) - p_{gz}^*(u,t)\Big) F_z(\mathrm{d}t) \\ &\quad + \sum_{(z,\zeta)\in J} \int_{(u,\infty)} \frac{1}{K(DT(t))} D_{z\zeta}(t) \Big(p_{lz}^*(u,t-0) - p_{gz}^*(u,t-0)\Big) q_{z\zeta}^*(\mathrm{d}t) \\ &\quad + \frac{1}{K(DT(u))} D_{gl}(u) \Bigg\} \\ &= \mathbf{1}_{(s,\infty)}(u)\, p_{yg}^*(s,u-0)\, \frac{K(s)}{K(u)} \left( V_{l,u} - V_{g,u} + \frac{K(u)}{K(DT(u))} D_{gl}(u) \right). \end{aligned} \tag{3.2.17}$$

*Proof.* Since for all $(y,z) \in \mathcal{S}\times\mathcal{S}$

$$\begin{aligned}\varrho_{yz}^s(t;\mathbf{1}_{[u,\infty)}) &= \mathbf{1}_{(s,t]}(u) \Big[p^*(s,u-0)\,(\mathbb{I}_{gl} - \mathbb{I}_{gg})\, p^*(u,t)\Big]_{yz} \\ &= \mathbf{1}_{(s,t]}(u)\, p_{yg}^*(s,u-0) \Big(p_{lz}^*(u,t) - p_{gz}^*(u,t)\Big)\end{aligned}$$

and $\mathrm{D}_0 V_{y,s}^{q_{gl}^*}$ is continuous at zero, that is, $|\mathrm{D}_0 V_{y,s}^{q_{gl}^*}(H)| \le \text{const}\, \|H\|_{BV}$ for all $H \in BV_{\leftarrow}$, equation (3.2.17) is obtained by applying Fubinis Theorem to (3.2.4). □

Similar to section 3.1, it is here easier to apply Fubinis Theorem to $\mathrm{D}_0 V_{y,s}^{q_{gl}^*}$ than to verify Condition 2.2.5 and to apply Theorem 2.2.6 in order to obtain the gradient vector $\nabla_0 V_{y,s}^{q_{gl}^*}$. The factor

$$R_{gl}(u) := V_{l,u} - V_{g,u} + \frac{K(u)}{K(DT(u))} D_{gl}(u), \qquad u \ge 0, \tag{3.2.18}$$

is the so-called *sum at risk* associated with a possible transition from state $g$ to state $l$ at time $u$ (cf. Milbrodt and Helbig (1999), p. 470). That means the gradient vector at $u > s$ is a product of (i) the probability to be in state $g$ just before $u$, (ii) a discounting factor, and (iii) the sum at risk.

**Remark 3.2.8** (Computing the gradient vector)**.** Starting from equation (3.2.17), the only challenges for computing the gradient vector $\nabla_0 V_{y,s}^{q_{gl}^*}$ are to get to know the transition probabilities $p_{yg}^*(s,u-0)$ and the prospective reserves $V_{l,u}, V_{g,u}$. For the former. use the Kolmogorov Forward/Backward Equations (provided the transition intensity matrix $\mu$ exists) and, for the latter, use the Thiele Integral/Differential Equations (cf. Milbrodt and Helbig (1999), section 10.C).

## 3.3 Gradient vector with respect to interest and all transitions simultaneously

Now regard the prospective reserve as a mapping of the cumulative interest intensity $\Phi$ and all cumulative transition intensities $q_{z\zeta}$, $(z,\zeta) \in J$, simultaneously. As stated in (3.0.1), let the cumulative intensities decompose to

$$(\Phi, q_J) = (\Phi^*, q_J^*) + (H_\Phi, H_J) := \big(\Phi^*, (q^*_{z\zeta})_{(z,\zeta)\in J}\big) + \big(H_\Phi, (H_{z\zeta})_{(z,\zeta)\in J}\big)\,, \qquad (3.3.1)$$

where $(\Phi^*, q_J^*)$ denotes an arbitrary but fixed starting point and $(H_\Phi, H_J)$ is a deviation of it. Let $\Phi^*$ satisfy (1.2.4) and the integrability conditions (1.2.6) and (1.2.7). Further let $q^*$ be a regular transition intensity matrix. Analogously to (3.2.1) and (3.2.2), write

$$\begin{aligned} q(\cdot; H_J) &:= q^*(\cdot) + \sum_{(z,\zeta)\in J} (\mathbb{I}_{z\zeta} - \mathbb{I}_{zz})\, H_{z\zeta}(\cdot)\,, \\ p(s,t;H_J) &:= \prod_{(s,t]} \big(\mathbb{I} + \mathrm{d}q(\cdot;H_J)\big)\,, \qquad -\infty < s < t < \infty\,. \end{aligned} \qquad (3.3.2)$$

Again, for a shorter notation write $p^*(s,t) := p(s,t;0)$. Regard the prospective reserve as a functional of the deviation $(H_\Phi, H_J)$ and write

$$V_{y,s}^{(\Phi^*,q^*)}(H_\Phi, H_J) := V_{y,s}\big[\Phi^* + H_\Phi, q(\cdot; H_J)\big]\,.$$

Since $V_{y,s}^{\Phi^*}$ is not well-defined on all of $BV_{\leftarrow}$, the functional $V_{y,s}^{(\Phi^*,q^*)}$ is not well-defined on all of $(BV_{\leftarrow})^{1+|J|}$. A suitable domain is obtained by restricting it to the subset $\mathcal{E}_{\Phi^*} \times (BV_{\leftarrow})^{|J|}$,

$$V_{y,s}^{(\Phi^*,q^*)} : \mathcal{E}_{\Phi^*} \times (BV_{\leftarrow})^{|J|} \to \mathbb{R}\,, \quad (H_\Phi, H_J) \mapsto V_{y,s}\big[\Phi^* + H_\Phi, q(\cdot;H_J)\big]\,. \qquad (3.3.3)$$

**Proposition 3.3.1.** *Let $t \mapsto D_{z\zeta}(t)/K_{\Phi^*}(DT(t))$ be bounded on $[0,\infty)$ for each $(z,\zeta) \in J$. For all $(H_\Phi, H_J) \in \mathcal{E}_{\Phi^*} \times (BV_{\leftarrow})^{|J|}$, $y \in \mathcal{S}$, and $s \geq 0$, one has $|V_{y,s}^{(\Phi^*,q^*)}(H_\Phi, H_J)| < \infty$.*

Before presenting the proof, another useful proposition is shown:

**Proposition 3.3.2.** *For all $H_J \in (BV_{\leftarrow})^{|J|}$ and $-\infty < s < t < \infty$*

$$\big|p(s,t;H_J)\big| \leq e^{2\,\|H_J\|_{BV}}\,. \qquad (3.3.4)$$

(For the definition of $\|H_J\|_{BV}$, see (2.5.1).)

*Proof.* Since

$$\Bigg|\sum_{(z,\zeta)\in J} (\mathbb{I}_{z\zeta} - \mathbb{I}_{zz})\, H_{z\zeta}(\cdot)\Bigg| \leq 2 \sum_{(z,\zeta)\in J} \|H_{z\zeta}\|_{BV} = 2\,\|H_J\|_{BV}\,,$$

the proof is analogous to that of Proposition 3.2.2. □

*Proof of Proposition* 3.3.1. Follow the lines of the proof of Proposition 3.2.1, and take into account (3.1.5), (3.1.6), and (3.3.4). □

In the same way Proposition 3.2.2 was generalized to Proposition 3.3.2, one gets analogously to Proposition 3.2.4

$$\big|p(s,t;H_J)-p^*(s,t)\big| \leq \text{const}\,\|H_J\|_{BV}\,, \tag{3.3.5}$$

$$\left|p(s,t;H_J)-p^*(s,t)-\sum_{(z,\zeta)\in J}\int_{(s,t]} p^*(s,u{-}0)\,(\mathbb{I}_{z\zeta}{-}\mathbb{I}_{zz})\,p^*(u,t)\,\mathrm{d}H_{z\zeta}(u)\right| \leq \text{const}\,\|H_J\|_{BV}^2 \tag{3.3.6}$$

for all $H_J \in (BV_{\leftarrow})^{|J|}$ satisfying $\|H_J\|_{BV} \leq c$ for an arbitrary but fixed constant $c<\infty$, and

$$\big|p(s,t;H_J)-p^*(s,t)\big| \leq 2\,\|H_J\|_{BV}\,, \tag{3.3.7}$$

$$\left|p(s,t;H_J)-p^*(s,t)-\sum_{(z,\zeta)\in J}\int_{(s,t]} p^*(s,u-0)\,(\mathbb{I}_{z\zeta}-\mathbb{I}_{zz})\,p^*(u,t)\,\mathrm{d}H_{z\zeta}(u)\right| \leq 4\,\|H_J\|_{BV}^2 \tag{3.3.8}$$

for all $H_J \in (BV_{\leftarrow})^{|J|}$ for which $q(\cdot;H_J)$ is a *regular* cumulative transition intensity matrix. In the same way Proposition 3.2.5 is expandable to the mapping

$$\begin{aligned}\rho^s : (BV_{\leftarrow})^{|J|} \to B &:= \Big\{f:\mathbb{R}\to\mathbb{R}^{|\mathcal{S}|\times|\mathcal{S}|}\;\Big|\;\|f\|_B := \sup_{t\in\mathbb{R}}|f(t)|<\infty\Big\}\,,\\ H_J &\mapsto \Big(t\mapsto \mathbf{1}_{(s,\infty)}(t)\,p(s,t;H_J)\Big)\,,\end{aligned}$$

with Fréchet differential

$$\mathrm{D}_0\rho^s(H) = \sum_{(z,\zeta)\in J}\int_{(s,\cdot]} p^*(s,u-0)\,(\mathbb{I}_{z\zeta}-\mathbb{I}_{zz})\,p^*(u,\cdot)\,\mathrm{d}H_{z\zeta}(u)\,, \qquad \forall\, H_J\in(BV_{\leftarrow})^{|J|}\,. \tag{3.3.9}$$

In the same manner, Theorem 3.2.3 holds for

$$V_{y,s}^{q^*} : (BV_{\leftarrow})^{|J|} \to \mathbb{R}\,,\quad H_J \mapsto V_{y,s}\big[\Phi^*, q(\cdot;H_J)\big]\,,$$

provided the mappings $t\mapsto D_{z\zeta}(t)/K_{\Phi^*}(DT(t))$, $(z,\zeta)\in J$, are bounded on $[0,\infty)$.

The Fréchet differential at zero is

$$\begin{aligned}
\mathrm{D}_0 V_{y,s}^{q^*}(H_J) &= \sum_{z\in\mathcal{S}} \int_{[s,\infty)} \frac{K_{\Phi^*}(s)}{K_{\Phi^*}(t)}\, \varrho_{yz}^{s}(t;H_J)\, F_z(\mathrm{d}t) \\
&\quad + \sum_{(z,\zeta)\in J} \int_{(s,\infty)} \frac{K_{\Phi^*}(s)}{K_{\Phi^*}(DT(t))}\, D_{z\zeta}(t)\, \varrho_{yz}^{s}(t-0;H_J)\, q_{z\zeta}^{*}(\mathrm{d}t) \\
&\quad + \sum_{(z,\zeta)\in J} \int_{(s,\infty)} \frac{K_{\Phi^*}(s)}{K_{\Phi^*}(DT(t))}\, D_{z\zeta}(t)\, p_{yz}^{*}(s,t-0)\, \mathrm{d}H_{z\zeta}(t) \\
&= \sum_{(z,\zeta)\in J} \mathrm{D}_0 V_{y,s}^{q_{z\zeta}^*}(H_{z\zeta})
\end{aligned} \tag{3.3.10}$$

for $H_J \in (BV_{\leftarrow})^{|J|}$, where for $(y,z)\in\mathcal{S}\times\mathcal{S}$ and $t\geq 0$

$$\varrho_{yz}^{s}(t;H_J) := \left[\sum_{(z,\zeta)\in J} \int_{(s,t]} p^*(s,u-0)\,(\mathbb{I}_{z\zeta}-\mathbb{I}_{zz})\, p^*(u,t)\,\mathrm{d}H_{z\zeta}(u)\right]_{yz} .$$

Because of the second equation in (3.3.10), the existence at zero of the gradient vectors of $V_{y,s}^{q_{z\zeta}^*}$, $(z,\zeta)\in J$, (cf. Theorem 3.2.7) implies the existence at zero of a gradient vector according to (2.5.2) for $V_{y,s}^{q^*}$. That means the concept of section 3.2 is expanded to a simultaneous study of all cumulative transition intensities. Now the task is to include the cumulative interest intensity as well.

**Theorem 3.3.3.** *Let* $t\mapsto D_{z\zeta}(t)/K_{\Phi^*}(DT(t))$ *be bounded on* $[0,\infty)$ *for each* $(z,\zeta)\in J$*. Then,* $V_{y,s}^{(\Phi^*,q^*)}$ *is Fréchet differentiable at zero with differential*

$$\mathrm{D}_0 V_{y,s}^{(\Phi^*,q^*)}(H_\Phi,H_J) = \mathrm{D}_0 V_{y,s}^{\Phi^*}(H_\Phi) + \mathrm{D}_0 V_{y,s}^{q^*}(H_J)\,, \qquad \forall\,(H_\Phi,H_J)\in\mathcal{E}_{\Phi^*}\times(BV_{\leftarrow})^{|J|}\,. \tag{3.3.11}$$

*Proof.* At first (A.3.3) is shown. For all $(H_\Phi,H_J)\in\mathcal{E}_{\Phi^*}\times(BV_{\leftarrow})^{|J|}$

$$\begin{aligned}
&\frac{1}{\|(H_\Phi,H_J)\|_{BV}}\left|V_{y,s}^{(\Phi^*,q^*)}(H_\Phi,H_J) - V_{y,s}^{(\Phi^*,q^*)}(0,0) - \mathrm{D}_0 V_{y,s}^{\Phi^*}(H_\Phi) - \mathrm{D}_0 V_{y,s}^{q^*}(H_J)\right| \leq \\
&\frac{1}{\|(H_\Phi,H_J)\|_{BV}}\left|V_{y,s}^{(\Phi^*,q^*)}(H_\Phi,0) - V_{y,s}^{(\Phi^*,q^*)}(0,0) - \mathrm{D}_0 V_{y,s}^{\Phi^*}(H_\Phi)\right| \\
&+ \frac{1}{\|(H_\Phi,H_J)\|_{BV}}\left|V_{y,s}^{(\Phi^*,q^*)}(0,H_J) - V_{y,s}^{(\Phi^*,q^*)}(0,0) - \mathrm{D}_0 V_{y,s}^{q^*}(H_J)\right| \\
&+ \frac{1}{\|(H_\Phi,H_J)\|_{BV}}\left|V_{y,s}^{(\Phi^*,q^*)}(H_\Phi,H_J) - V_{y,s}^{(\Phi^*,q^*)}(0,H_J) - V_{y,s}^{(\Phi^*,q^*)}(H_\Phi,0) + V_{y,s}^{(\Phi^*,q^*)}(0,0)\right|.
\end{aligned} \tag{3.3.12}$$

The first and the second addend converge to zero, because the functionals $V_{y,s}^{(\Phi^*,q^*)}(\cdot,0) = V_{y,s}^{\Phi^*}(\cdot)$ and $V_{y,s}^{(\Phi^*,q^*)}(0,\cdot) = V_{y,s}^{q^*}(\cdot)$ are Fréchet differentiable (cf. Theorem 3.1.3 and (3.3.10)).

Now look at the third addend. Let $\mu_{SB,0}$ and $\mu_{DB,0}$ be defined by (3.1.14), and denote by $\mu_{SB,H_J}$ and $\mu_{DB,H_J}$ analogous measures for which $q^*$ is replaced by $q(\cdot;H_J)$ and $p^*$ is replaced by $p(\cdot,\cdot;H_J)$. Then,

$$\begin{aligned}&\Big|V_{y,s}^{(\Phi^*,q^*)}(H_\Phi,H_J)-V_{y,s}^{(\Phi^*,q^*)}(0,H_J)-V_{y,s}^{(\Phi^*,q^*)}(H_\Phi,0)+V_{y,s}^{(\Phi^*,q^*)}(0,0)\Big|\\&=\Big|\int\Big(\frac{K_{\Phi^*+H_\Phi}(s)}{K_{\Phi^*+H_\Phi}(t)}-\frac{K_{\Phi^*}(s)}{K_{\Phi^*}(t)}\Big)\Big(\mu_{SB,H_J}+\mu_{DB,H_J}\circ DT^{-1}-\mu_{SB,0}-\mu_{DB,0}\circ DT^{-1}\Big)(\mathrm{d}t)\Big|.\end{aligned}$$

Applying (3.1.10) the integrand is bounded by

$$\Big|\frac{K_{\Phi^*+H_\Phi}(s)}{K_{\Phi^*+H_\Phi}(t)}-\frac{K_{\Phi^*}(s)}{K_{\Phi^*}(t)}\Big|\le \mathrm{const}\,\|H_\Phi\|_{BV}\frac{K_{\Phi^*}(s)}{K_{\Phi^*}(t)}\,,\qquad \forall\, H_\Phi\in\mathcal{B}_{\Phi^*}\subset\mathcal{E}_{\Phi^*}\,.$$

Looking at the integrator, for any $H_J\in(BV_{\leftarrow})^{|J|}$ with $\|H_J\|_{BV}\le c<\infty$ for an arbitrary but fixed constant $c$ the properties (3.3.4) and (3.3.5) lead to

$$\begin{aligned}&\Big|\mu_{SB,H_J}+\mu_{DB,H_J}\circ DT^{-1}-\mu_{SB,0}-\mu_{DB,0}\circ DT^{-1}\Big|(A)\\&\le\sum_{z\in\mathcal{S}}\int_{A\cap[s,\infty)}\Big|p_{yz}(s,t;H_J)-p^*_{yz}(s,t)\Big|\,|F_z|(\mathrm{d}t)\\&\quad+\sum_{(z,\zeta)\in J}\int_{DT^{-1}(A)\cap(s,\infty)}D_{z\zeta}(t)\,\Big|p_{yz}(s,t-0;H_J)-p^*_{yz}(s,t-0)\Big|\,q^*_{z\zeta}(\mathrm{d}t)\\&\quad+\sum_{(z,\zeta)\in J}\int_{DT^{-1}(A)\cap(s,\infty)}D_{z\zeta}(t)\big|p_{yz}(s,t-0;H_J)\big|\,\mathrm{d}|H_{z\zeta}|(t)\\&\le\mathrm{const}\,\|H_J\|_{BV}\Big(\sum_{z\in\mathcal{S}}\int_{A\cap[s,\infty)}|F_z|(\mathrm{d}t)+\sum_{(z,\zeta)\in J}\int_{DT^{-1}(A)\cap(s,\infty)}D_{z\zeta}(t)\,q^*_{z\zeta}(\mathrm{d}t)\Big)\\&\qquad+\mathrm{const}\sum_{(z,\zeta)\in J}\int_{DT^{-1}(A)\cap(s,\infty)}D_{z\zeta}(t)\,\mathrm{d}|H_{z\zeta}|(t)\Big)\end{aligned}$$

for all $A\in\mathfrak{B}(\mathbb{R})$. Consequently,

$$\begin{aligned}&\Big|V_{y,s}^{(\Phi^*,q^*)}(H_\Phi,H_J)-V_{y,s}^{(\Phi^*,q^*)}(0,H_J)-V_{y,s}^{(\Phi^*,q^*)}(H_\Phi,0)+V_{y,s}^{(\Phi^*,q^*)}(0,0)\Big|\\&\le\mathrm{const}\,\|H_\Phi\|_{BV}\,\|H_J\|_{BV}\big(\mathrm{InCo_{SB}}+\mathrm{InCo_{DB}}\big)\\&\quad+\mathrm{const}\,\|H_\Phi\|_{BV}\sum_{(z,\zeta)\in J}\int_{(s,\infty)}\frac{K_{\Phi^*}(s)}{K_{\Phi^*}(DT(t))}D_{z\zeta}(t)\,\mathrm{d}|H_{z\zeta}|(t)\\&\le\mathrm{const}\,\|H_\Phi\|_{BV}\,\|H_J\|_{BV}\\&\le\mathrm{const}\,\|(H_\Phi,H_J)\|^2_{BV}\,.\end{aligned}\tag{3.3.13}$$

Hence, the functional $V_{y,s}^{(\Phi^*,q^*)}$ satisfies (A.3.3) with differential (3.3.11). The linearity and the continuity of $\mathrm{D}_0V_{y,s}^{\Phi^*}$ and $\mathrm{D}_0V_{y,s}^{q^*}$ imply the same for $D_0V_{y,s}^{(\Phi^*,q^*)}$. □

**Remark 3.3.4.** In case $H_J \in (BV_{\leftarrow})^{|J|}$ is chosen in such a way that $q(\cdot; H_J)$ is a *regular* cumulative transition intensity matrix, the properties (3.3.7) and (3.3.8) yield analogously to Remark 3.2.6

$$\begin{aligned}&\left|V_{y,s}^{q^*}(H_J) - V_{y,s}^{q^*}(0) - \mathrm{D}_0 V_{y,s}^{q^*}(H_J)\right| \\ &\leq 4\|H_J\|_{BV}^2\, K_{\Phi^*}(s)\,\big(\mathrm{InCo_{SB}} + \mathrm{InCo_{DB}}\big) + \sum_{(z,\zeta)\in J} K_{\Phi^*}(s) \sup_{t\in\mathbb{R}} \left\{\frac{D_{z\zeta}}{K_{\Phi^*}\circ DT}(t)\right\} 2\,\|H_{z\zeta}\|_{BV}^2 .\end{aligned} \tag{3.3.14}$$

If additionally $H_\Phi \in \mathcal{E}_{\Phi^*}$ is continuous, properties (3.1.12) and (3.3.7) allow for replacing the upper bound in (3.3.13) by

$$\begin{aligned}&\left|V_{y,s}^{(\Phi^*,q^*)}(H_\Phi, H_J) - V_{y,s}^{(\Phi^*,q^*)}(0, H_J) - V_{y,s}^{(\Phi^*,q^*)}(H_\Phi, 0) + V_{y,s}^{(\Phi^*,q^*)}(0,0)\right| \\ &\leq 2\,\|H_J\|_{BV}\,\|H_\Phi\|_{BV}\, K_{\Phi^*}(s) \int \Big(1 + e^{H_\Phi(s)-H_\Phi(t)}\Big)\,\gamma(\mathrm{d}t) \\ &\quad + \|H_\Phi\|_{BV}\, K_{\Phi^*}(s) \sum_{(z,\zeta)\in J} \sup_{t>0}\left\{\frac{D_{z\zeta}}{K_{\Phi^*}\circ DT}(t)\right\} \int_{(0,\infty)} \Big(1 + e^{H_\Phi(s)-H_\Phi(DT(t))}\Big)\,\mathrm{d}|H_{z\zeta}|(t),\end{aligned} \tag{3.3.15}$$

where

$$\begin{aligned}\gamma(A) := &\sum_{z\in\mathcal{S}} \int_{A\cap[0,\infty)} \frac{1}{K_{\Phi^*}(t)}\,|F_z|(\mathrm{d}t) \\ &+ \sum_{(z,\zeta)\in J} \int_{DT^{-1}(A)\cap(0,\infty)} \frac{1}{K_{\Phi^*}(DT(t))}\, D_{z\zeta}(t)\, q_{z\zeta}(\mathrm{d}t), \qquad \forall\, A \in \mathfrak{B}(\mathbb{R}).\end{aligned} \tag{3.3.16}$$

Thus, following the line of the proof of Theorem 3.3.3, the properties (3.1.18), (3.3.14), and (3.3.15) yield

$$\begin{aligned}&\left|V_{y,s}^{(\Phi^*,q^*)}(H_\Phi, H_J) - V_{y,s}^{(\Phi^*,q^*)}(0) - \mathrm{D}_0 V_{y,s}^{(\Phi^*,q^*)}(H_\Phi, H_J)\right| \\ &\leq \|H_\Phi\|_{BV}^2\, K_{\Phi^*}(s) \int \Big(1 + e^{H_\Phi(s)-H_\Phi(t)}\Big)\,\gamma(\mathrm{d}t) \\ &\quad + \|H_J\|_{BV}^2\, K_{\Phi^*}(s)\left(4\,(\mathrm{InCo_{SB}} + \mathrm{InCo_{DB}}) + 2 \sum_{(z,\zeta)\in J} \sup_{t\in\mathbb{R}}\left\{\frac{D_{z\zeta}}{K_{\Phi^*}\circ DT}(t)\right\}\right) \\ &\quad + \|H_J\|_{BV}\,\|H_\Phi\|_{BV}\, K_{\Phi^*}(s)\, 2 \int \Big(1 + e^{H_\Phi(s)-H_\Phi(t)}\Big)\,\gamma(\mathrm{d}t) \\ &\quad + \|H_\Phi\|_{BV}\, K_{\Phi^*}(s) \sum_{(z,\zeta)\in J} \sup_{t>0}\left\{\frac{D_{z\zeta}}{K_{\Phi^*}\circ DT}(t)\right\} \int_{(0,\infty)} \Big(1 + e^{H_\Phi(s)-H_\Phi(DT(t))}\Big)\,\mathrm{d}|H_{z\zeta}|(t)\end{aligned} \tag{3.3.17}$$

for all $(H_\Phi, H_J) \in \mathcal{E}_{\Phi^*} \times (BV_{\leftarrow})^{|J|}$ with $H_\Phi$ being continuous and $q(\cdot; H_J)$ being a regular cumulative transition intensity matrix. This property will be needed in chapter 4.

**Theorem 3.3.5.** *Let $t \mapsto D_{z\zeta}(t)/K_{\Phi^*}(DT(t))$ be bounded on $[0,\infty)$ for each $(z,\zeta) \in J$. Then, the gradient vector of $V_{y,s}^{(\Phi^*,q^*)}$ exists at zero and has the form*

$$\nabla_0 V_{y,s}^{(\Phi^*,q^*)} = \Big(\nabla_0 V_{y,s}^{\Phi^*}, \big(\nabla_0 V_{y,s}^{q^*_{gl}}\big)_{(g,l)\in J}\Big). \tag{3.3.18}$$

*Proof.* Because of (3.3.11) and the existence of $\nabla_0 V_{y,s}^{\Phi^*}$ and $\nabla_0 V_{y,s}^{q^*}$, there exists a gradient vector II for $V_{y,s}^{(\Phi^*,q^*)}$ at zero. According to Proposition 2.5.3, it is of the form (3.3.18). □

Occasionally, it can be convenient to assume that some subgroup of the cumulative transition intensities varies synchronously, for example, mortality changes of healthy and of disabled persons evolving similarly.

Let $I := \{(g_1,l),\ldots,(g_n,l)\}$ be an arbitrary but fixed subset of $J$. Without loss of generality, write $H_J = (H_{J\setminus I}, H_{g_1 l},\ldots,H_{g_n l})$. Taking into account (3.3.3), define the functional

$$\begin{aligned} W_{y,s} : \mathcal{E}_{\Phi^*} \times (BV_{\leftarrow})^{|J\setminus I|+1} &\to \mathbb{R}, \\ (H_\Phi, H_{J\setminus I}, G) &\mapsto V_{y,s}^{(\Phi^*,q^*)}\big(H_\Phi, H_{J\setminus I}, \underbrace{G,\ldots,G}_{n \text{ times}}\big). \end{aligned} \tag{3.3.19}$$

From Theorem 3.3.3 it follows that

$$\mathrm{D}_0 W_{y,s}(0,\ldots,0,G) = \mathrm{D}_0 V_{y,s}^{(\Phi^*,q^*)}\big(0,\ldots,0,\underbrace{G,\ldots,G}_{n \text{ times}}\big) = \sum_{i=1}^{n} \mathrm{D}_0 V_{y,s}^{g_i l}(G), \quad \forall G \in BV_{\leftarrow}.$$

The uniqueness properties of Proposition 2.2.2 and Proposition 2.5.3 yield that the gradient vector $\nabla_0 W_{y,s}$ is in its $1+|J\setminus I|$ first dimensions equal to that of $\nabla_0 V_{y,s}^{(\Phi^*,q^*)}$, and its last entry is equal to the sum of the $n$ last entries of $\nabla_0 V_{y,s}^{(\Phi^*,q^*)}$.

## 3.4 Gradient vector of the premium level

In classical life insurance with deterministic actuarial assumptions, the premiums are commonly calculated by using the equivalence principle according to section 1.2. In order to study the influence of changes of actuarial assumptions on premiums, regard the premium level (1.2.17) as a functional of the cumulative interest and transition intensities,

$$C^{(\Phi^*,q^*)} : \mathcal{E}_{\Phi^*} \times (BV_{\leftarrow})^{|J|} \to \mathbb{R}, \quad (H_\Phi, H_J) \mapsto -\frac{VB_{a,0}^{(\Phi^*,q^*)}(H_\Phi,H_J)}{VE_{a,0}^{(\Phi^*,q^*)}(H_\Phi,H_J)}, \tag{3.4.1}$$

where $a$ denotes the initial state of the insured at time zero, and the mappings $VB_{a,0}^{(\Phi^*,q^*)}$ and $VE_{a,0}^{(\Phi^*,q^*)}$ are functionals in accordance with (3.3.3), but solely with benefit and with premium payments, respectively. It is well-defined as long as the denominator is not equal to zero.

**Theorem 3.4.1.** *Let* $VE_{a,0}^{(\Phi^*,q^*)}(0,0) \neq 0$. *Under the assumptions of Theorem* 3.3.3, *the functional* $C^{(\Phi^*,q^*)}$ *is Fréchet-differentiable at zero with differential*

$$\mathrm{D}_0 C^{(\Phi^*,q^*)}(H_\Phi, H_J) = -\frac{\mathrm{D}_0 V_{a,0}^{(\Phi^*,q^*)}(H_\Phi, H_J)}{VE_{a,0}^{(\Phi^*,q^*)}(0,0)}\,, \qquad \forall (H_\Phi, H_J) \in (BV_{\leftarrow})^{1+|J|}\,, \tag{3.4.2}$$

*and has a gradient vector II at zero,*

$$\nabla_0 C^{(\Phi^*,q^*)} = -\frac{1}{VE_{a,0}^{(\Phi^*,q^*)}(0,0)} \nabla_0 V_{a,0}^{(\Phi^*,q^*)}\,. \tag{3.4.3}$$

*Proof.* With $VB_{a,0}^{(\Phi^*,q^*)}$ and $VE_{a,0}^{(\Phi^*,q^*)}$ being Fréchet differentiable at zero, one gets

$$\begin{aligned}
&\mathrm{D}_0 C^{(\Phi^*,q^*)}(H_\Phi, H_J)\\
&= -\frac{\mathrm{D}_0 VB_{a,0}^{(\Phi^*,q^*)}(H_\Phi, H_J) \cdot VE_{a,0}^{(\Phi^*,q^*)}(0,0) - VB_{a,0}^{(\Phi^*,q^*)}(0,0) \cdot \mathrm{D}_0 VE_{a,0}^{(\Phi^*,q^*)}(H_\Phi, H_J)}{\left(VE_{a,0}^{(\Phi^*,q^*)}(0,0)\right)^2}\\
&= -\frac{1}{VE_{a,0}^{(\Phi^*,q^*)}(0,0)}\left(\mathrm{D}_0 VB_{a,0}^{(\Phi^*,q^*)}(H_\Phi, H_J) + C^{(\Phi^*,q^*)}(0,0) \cdot \mathrm{D}_0 VE_{a,0}^{(\Phi^*,q^*)}\right)\\
&= -\frac{1}{VE_{a,0}^{(\Phi^*,q^*)}(0,0)} \mathrm{D}_0 V_{a,0}^{(\Phi^*,q^*)}(H_\Phi, H_J)\,, \qquad \forall\, (H_\Phi, H_J) \in (BV_{\leftarrow})^{1+|J|}\,.
\end{aligned}$$

As $\mathrm{D}_0 C^{(\Phi^*,q^*)}$ differs from $\mathrm{D}_0 V_{a,0}^{(\Phi^*,q^*)}$ only in the constant factor $-1/VE_{a,0}^{(\Phi^*,q^*)}(0,0)$, the same holds for the corresponding gradient vectors. □

For a better comparability it is sometimes convenient to normalize the gradient vector $\nabla_0 C^{(\Phi^*,q^*)}$ by dividing it by $C^{(\Phi^*,q^*)}(0,0)$. Then, (3.4.3) implies

$$\frac{1}{C^{(\Phi^*,q^*)}(0,0)} \nabla_0 C^{(\Phi^*,q^*)}(H) = -\frac{1}{C^{(\Phi^*,q^*)}(0,0) \cdot VE_{a,0}^{(\Phi^*,q^*)}(0,0)} \nabla_0 V_{a,0}^{(\Phi^*,q^*)}\,, \tag{3.4.4}$$

where the denominator on the right hand side is equal to the mean present value of all premiums at time zero. That means the normalized gradient vector of the premium level is – despite a factor of minus one – equal to the gradient vector of the prospective reserve normalized by its mean present value of premiums payments. Hence, in many cases it suffices to concentrate solely on the prospective reserve.

## 3.5 Sensitivities of typical life insurance contracts

Now a sensitivity analysis, based on the gradient vector (3.3.18), is performed for some typical life insurance contracts. The following examples are either two-state or three-state models with state spaces $\mathcal{S} = \{\mathrm{a}, \mathrm{d}\}$ and $\mathcal{S} = \{\mathrm{a}, \mathrm{i}, \mathrm{d}\}$, where the states 'a', 'i', and 'd' stand for 'alive and fit', 'incapable of working', and 'dead', respectively. The corresponding sets of possible direct transitions are $J = \{(\mathrm{a}, \mathrm{d})\}$ and $J = \{(\mathrm{a}, \mathrm{i}), (\mathrm{a}, \mathrm{d}), (\mathrm{i}, \mathrm{a}), (\mathrm{i}, \mathrm{d})\}$.

Starting from state 'a', the annual mortality probabilities and the annual disability probabilities are taken from the life table 2002/2004 of the 'Statistisches Bundesamt Deutschland' and from the disability table DAV 1997 I of the 'Deutsche Aktuarvereinigung'. Both tables provide so-called 'independent probabilities' (cf. Milbrodt and Helbig (1999), Definition 3.25), which are transformed to 'dependent probabilities' as stated in Exercise 3.17(c) in Milbrodt and Helbig (1999, pp. 131-132). The procedure is similar when coming from state 'i'. For the transition $(\mathrm{i}, \mathrm{d})$ the ultimate table of the select life table DAV 1997 TI is used, which is a special life table for disabled. For the transition $(\mathrm{i}, \mathrm{a})$ the ultimate table of the select recovery table DAV 1997 RI is used.

As the tables provide only yearly probabilities, one needs additional assumptions to obtain a real continuous time model. Assume that the integer truncated durations of stay in a state are independent of their remainders, where the latter are uniformly distributed. As a result the transition intensities as derivatives of the cumulative transition intensities exist and are easily computable according to Theorem 6.24 in Milbrodt and Helbig (1999, p. 289). The transition probabilities are now numerically computable using the Kolmogorov forward/backward differential equation.

Further on, suppose the annual interest is at 4% and the accumulation factor has the form $K_{\Phi^*}(t) = (1.04)^t$, $t \geq 0$, that is, the interest intensity (or rate) is constantly at $\ln(1.04)$. For the sake of simplicity, let $DT \equiv Id$, that is, any benefits due to a change of state are paid immediately.

Because of Theorem 3.3.5, the effect of simultaneous interest and transition intensity changes is equal to an aggregation of the individual effects of changes of single intensities. Thus, at first the gradient vectors (3.1.19) and (3.2.17) are analyzed separately before simultaneous parameter changes are studied.

#### Sensitivity with respect to interest

At first, gradient vectors with respect only to interest are studied.

**Example 3.5.1.** Consider a male who contracts an insurance at the age of $30$. At first, look at simple two state models with state space $\mathcal{S} = \{\mathrm{a}, \mathrm{d}\}$.

(a) Consider a *pure endowment insurance* with a sum insured of 1 payable upon termination of the contract – here at the age of 65 – in case of survival. The premium is paid yearly in advance with a constant fee of $0.01123$, which approximately satisfies the equivalence condition $V_{\mathrm{a},0} = 0$.

$$F_{\mathrm{a}}(t) = -\sum_{i=0}^{34} 0.01123 \cdot \mathbf{1}_{[i,\infty)}(t) + \mathbf{1}_{[35,\infty)}(t)\,, \quad F_{\mathrm{d}}(t) = 0\,, \qquad \forall\, t \in \mathbb{R}\,.$$

(b) Instead of paying a yearly premium one could have a big lump sum at the beginning of the contract. The equivalence principle yields

$$F_{\mathrm{a}}(t) = -0.21151 \cdot \mathbf{1}_{[0,\infty)}(t) + \mathbf{1}_{[35,\infty)}(t)\,, \quad F_{\mathrm{d}}(t) = 0\,, \qquad \forall\, t \in \mathbb{R}\,.$$

(c) Now look at a *temporary life insurance* with a sum insured of 1 and premiums paid monthly in advance till death or termination of the contract, here the at age of 65. With the equivalence principle one gets

$$F_{\mathrm{a}}(t) = -\sum_{i=0}^{34} 0.003467 \cdot \mathbf{1}_{[i,\infty)}(t)\,, \quad F_{\mathrm{d}}(t) = 0\,, \quad D_{\mathrm{ad}}(t) = \mathbf{1}_{(0,35]}(t)\,, \qquad \forall t \in \mathbb{R}\,.$$

(d) Consider an *annuity insurance* with premiums paid yearly in advance till death or the age of 65 and a yearly annuity payment of 1 starting at age 65 till death. Taking into account the equivalence principle, one gets

$$F_{\mathrm{a}}(t) = -\sum_{i=0}^{34} 0.13272 \cdot \mathbf{1}_{[i,\infty)}(t) + \sum_{i=35}^{\infty} \mathbf{1}_{[i,\infty)}(t)\,, \quad F_{\mathrm{d}}(t) = 0\,, \qquad \forall\, t \in \mathbb{R}\,.$$

The last example is a three-state model with $\mathcal{S} = \{\mathrm{a}, \mathrm{i}, \mathrm{d}\}$:

(e) Consider a *disability insurance* with a yearly disability annuity of 1 payable up to the age of 65 as long as the insured is in state 'i'. A constant premium has to be paid yearly in advance till the age of 64, as long as the insured is in state 'a'. The equivalence principle yields

$$F_{\mathrm{a}}(t) = -\sum_{i=0}^{33} 0.05410 \cdot \mathbf{1}_{[i,\infty)}(t)\,, \quad F_{\mathrm{i}}(t) = \sum_{i=0}^{34} \mathbf{1}_{[i,\infty)}(t)\,, \qquad F_{\mathrm{d}}(t) = 0\,, \quad \forall t \in \mathbb{R}\,.$$

For the sake of comparability the insurance contracts (a) to (e) should be normalized, which means the payments are scaled in such a way that all contracts have the same mean present premium values $C \cdot VE_{a,0}$ or mean present benefit values $VB_{a,0}$ at time zero in initial state 'a' (cf. (1.2.16)). The reciprocals of these mean present values are proper scaling factors.

| | pure endowment ins. | temporary life ins. | annuity ins. | disability ins. |
|---|---|---|---|---|
| $VB_{a,0}$ | 0.21151 | 0.06531 | 2.50038 | 0.95658 |

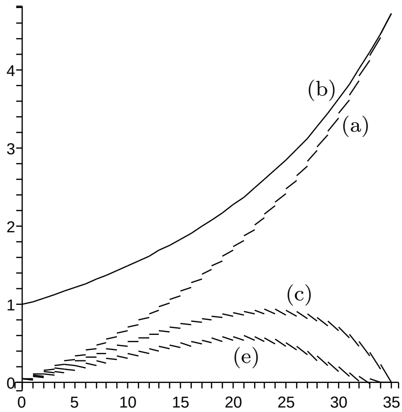


Figure 3.5.1: normalized prospective reserves $V_{\mathrm{a},\cdot}$ of Examples 3.5.1 (a) to (c) and (e)

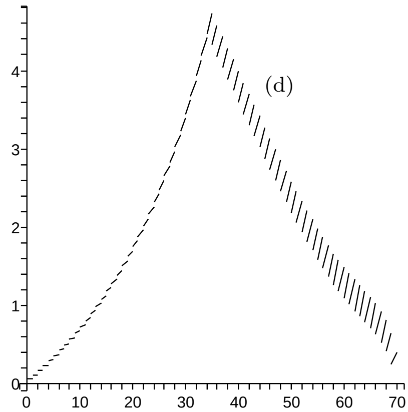


Figure 3.5.2: normalized prospective reserve $V_{\mathrm{a},\cdot}$ of Example 3.5.1(d)

Figures 3.5.1 and 3.5.2 show the normalized prospective reserves $s \mapsto V_{\mathrm{a},s}$ of Examples 3.5.1. Figures 3.5.3 and 3.5.4 illustrate the corresponding gradient vectors $u \mapsto \nabla_0 V_{\mathrm{a},0}^{\Phi^*}(u)$. With the gradient vector II in Definition 2.2.1 being some form of generalization of classical gradient vectors on $\mathbb{R}^n$, the graphs of Examples 3.5.1 (a), (c), and (d) in Figures 3.5.3 and 3.5.4 are similar to the 'Zinsdurationen' of Helwich (2003, p. 113).

Conspicuously, all gradient vectors are nonpositive. In fact, this points to a deeper insight: Contrary to the insured, the insurer has in many cases no legal right to cancel the contract. To avoid arbitrage possibilities in favor of the insured, the insurer aims to keep the prospective reserve nonnegative at any time for any state, that is, $V_{y,s} \geq 0$ for all $y \in \mathcal{S}$ and all $s \geq 0$, which induces the nonpositivity of the gradient vectors:

**Corollary 3.5.2.** *If the prospective reserve $V_{y,s}$ is nonnegative for any state $y \in \mathcal{S}$ at any time $s \geq 0$, the gradient vector of the prospective reserve with respect to the cumulative interest intensity is nonpositive,*

$$\nabla_0 V_{y,s}^{\Phi^*}(u) \leq 0\,, \quad \forall\, y \in \mathcal{S}\,, s \geq 0\,, u \geq 0\,.$$

*The inequality is even strict for all $u > s$ if there exists at least one state $w \in \mathcal{S}$, for which the transition probability $p_{yw}(s,\cdot)$ and the prospective reserve $s \mapsto V_{w,s}$ are always strictly positive.*

*Proof.* Analogously to Remark 3.1.8, the Chapman-Kolmogorov equations

$$p_{yz}^*(s,t) = \sum_{w \in \mathcal{S}} p_{yw}^*(s,u)\, p_{wz}^*(u,t)\,, \qquad s \leq u \leq t\,,\;\; y,z \in \mathcal{S},$$

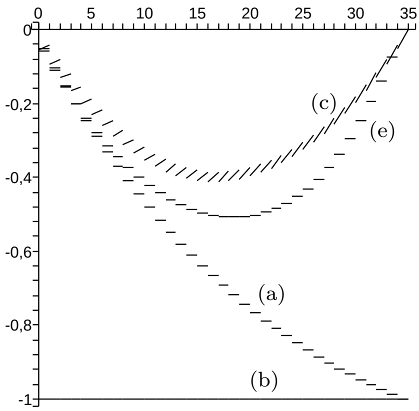


Figure 3.5.3: normalized gradient vectors $\nabla_0 V_{\mathrm{a},0}^{\Phi^*}$ of Examples 3.5.1 (a) to (c) and (e)

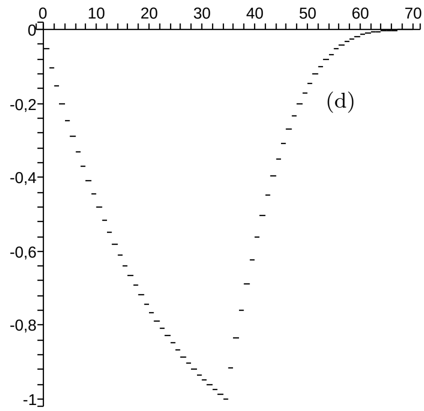


Figure 3.5.4: normalized gradient vector $\nabla_0 V_{\mathrm{a},0}^{\Phi^*}$ of Example 3.5.1 (d)

and Theorem 3.1.7 yield

$$\nabla_0 V_{y,s}^{\Phi^*}(u) = -\frac{\mathbf{1}_{(s,\infty)}(u)}{1+\Delta\Phi^*(u)}\,\frac{K_{\Phi^*}(s)}{K_{\Phi^*}(u)}\Bigg(\sum_{w\in\mathcal{S}} p_{yw}^*(s,u)\,V_{w,u} + \sum_{(z,\zeta)\in J}\int_{(s,u]} \mathbf{1}_{\{DT(t)\geq u\}}(t)\frac{K_{\Phi^*}(u)}{K_{\Phi^*}(DT(t))} D_{z\zeta}(t)\,p_{yz}^*(s,t-0)\,q_{z\zeta}^*(\mathrm{d}t)\Bigg). \tag{3.5.1}$$

Now the nonnegativity of the prospective reserves $V_{w,u}$, the regularity of the cumulative intensities $q_{z\zeta}^*$, and the nonnegativity of $1+\Delta\Phi^*$ and $K_{\Phi^*}$ (cf. (1.2.4) and Proposition A.4.2) lead to the nonpositivity of $\nabla_0 V_{y,s}^{\Phi^*}$. □

*That means higher interest rates normally diminish the prospective reserve and lower interest rates raise it.* This agrees with the existing literature on this subsect, e.g., Hoem (1988, section 8), Linnemann (1993, section 6), or Milbrodt and Helbig (1999, Bemerkungen 9.19(c)).

Comparing Figures 3.5.1 and 3.5.2 with Figure 3.5.3 and 3.5.4 unveils another insight: *As interest is borne by the aggregated reserves, the sensitivity to interest is the lower, the closer the prospective reserve is to zero.*

For example, with the temporary life insurance (c) having a consistently lower prospective reserve than the pure endowment insurances (a)&(b) and the annuity insurance (d), it shows throughout a lower sensitivity to interest rate changes. Comparing the two pure endowment insurances (a) and (b), paying the premium in yearly rates produces a lower reserve than a lump sum premium and by that a lower sensitivity.

The lowest sensitivities are obtained if the 'natural premium' (cf. Milbrodt and Helbig (1999), pp. 376, 380, 381) is charged, which is just the amount needed to cover all benefits for the next short period and by that minimizes the reserves. However, lowering the interest rate sensitivity by altering the premium scheme is not a panacea. In practice, there is often little room to vary the premium scheme, since balancing premium loads within contract periods is a core service of insurers.

Seemingly the disability insurance deviates from the above rule as its prospective reserve is consistently lower than that of the temporary life insurance (c), whereas the gradient vectors show the opposite behavior. Looking at Figures 3.5.1 and 3.5.5 explains the discrepancy: While the normalized prospective reserve starting from state 'a' is comparatively low, the one starting from state 'i' is tremendous. With $DT = Id$ and the continuity of the transition probabilities, the gradient vector (3.5.1) is here equal to

$$\nabla_0 V^{\Phi^*}_{\mathrm{a},s}(u) = -\mathbf{1}_{(s,\infty)}(u)\,\frac{K_{\Phi^*}(s)}{K_{\Phi^*}(u)}\Big(p^*_{\mathrm{aa}}(s,u)\,V_{\mathrm{a},u} + p^*_{\mathrm{ai}}(s,u)\,V_{\mathrm{i},u}\Big)\,, \qquad u \geq 0\,, \quad (3.5.2)$$

which shows the influence of $V_{\mathrm{i},\cdot}$.

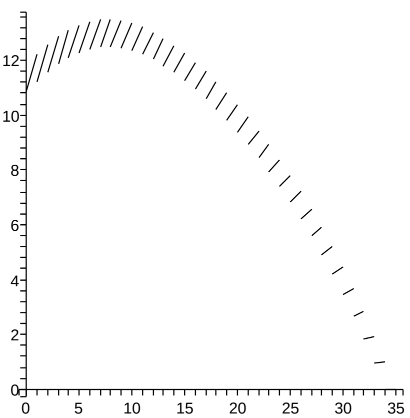


Figure 3.5.5: normalized prospective reserve $V_{\mathrm{i},\cdot}$ of Example 3.5.1(e)

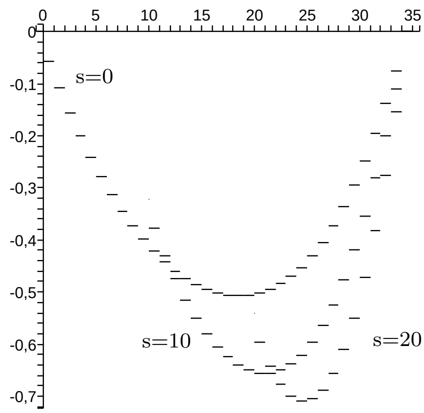


Figure 3.5.6: normalized gradient vector $\nabla_0 V^{\Phi^*}_{\mathrm{a},s}$ of Example 3.5.1(d) with $s = 0, 10, 20$

More interesting insights are gained by studying the progression of the gradient vectors in time. For example, the pure endowment insurance (a) with yearly fees has an increasing sensitivity towards termination of the contract, whereas the pure endowment insurance (b) with a lump sum premium has a constant gradient vector. The temporary life insurance (c) and the disability insurance (e) react most sensitively to interest rate changes after half of the contract period. The annuity insurance (d) has the greatest sensitivity at about the age of retirement.

Another interesting question is which role the reference time $s$ plays in $\nabla_0 V_{y,s}^{\Phi^*}$. For the two-state models of Example 3.5.1 with $DT = Id$ and existing interest and mortality intensity the gradient vector (3.5.1) is of the form

$$\begin{aligned}\nabla_0 V_{\mathrm{a},s}^{\Phi^*}(u) &= -\mathbf{1}_{(s,\infty)}(u)\,\frac{K_{\Phi^*}(s)}{K_{\Phi^*}(u)}\,p_{\mathrm{aa}}^*(s,u)\,V_{\mathrm{a},u}\\ &= -\mathbf{1}_{(s,\infty)}(u)\,\frac{K_{\Phi^*}(s)}{K_{\Phi^*}(\widetilde{s})}\,\frac{K_{\Phi^*}(\widetilde{s})}{K_{\Phi^*}(u)}\,\frac{1}{p_{\mathrm{aa}}^*(\widetilde{s},s)}\,p_{\mathrm{aa}}^*(\widetilde{s},u)\,V_{\mathrm{a},u}\\ &= \mathbf{1}_{(s,\infty)}(u)\,\frac{K_{\Phi^*}(s)}{K_{\Phi^*}(\widetilde{s})}\,\frac{1}{p_{\mathrm{aa}}^*(\widetilde{s},s)}\,\nabla_0 V_{\mathrm{a},\widetilde{s}}^{\Phi^*}(u)\,, \qquad \forall\, 0 \le \widetilde{s} \le s\,, u \ge 0\,.\end{aligned}$$

That means the gradient vectors with respect to reference times $\widetilde{s}$ and $s$ differ on $(s,\infty)$ only in the real number $K_{\Phi^*}(s)/(K_{\Phi^*}(\widetilde{s})\,p_{\mathrm{aa}}^*(\widetilde{s},s))$. *The gradient vectors of the two-state models with continuous transition probabilities and $DT = Id$ do not vary in their fundamental shape if the reference time $s$ is shifted.* Hence, mostly it suffices to solely concentrate on $s = 0$.

Having more than two states complicates the dependence on $s$. If one assumes the prospective reserves are nonnegative, at least the sign of the gradient vector is not changing (cf. Corollary 3.5.2). Figure 3.5.6 shows the gradient vector $\nabla_0 V_{\mathrm{a},s}^{\Phi^*}$ of Example 3.5.1(e) for $s = 0, 10, 20$. The shape of the three plots remains quite similar, but in contrast to the two-state models the location of the minimum moved.

#### Sensitivity with respect to single transitions

Now gradient vectors of Examples 3.5.1 with respect to single cumulative transition intensities are studied.

Figures 3.5.7 and 3.5.8 show the gradient vectors with respect to mortality $\nabla_0 V_{\mathrm{a},0}^{q_{\mathrm{ad}}^*}$ of Examples 3.5.1 (a) to (d). They are similar to the so-called 'Biodurationen' of Helwich (2003, p. 113).

Clearly, the temporary life insurance is most sensitive towards mortality changes, especially at the beginning of the contract. Comparing the two pure endowment insurances, the lump sum premium in (b) produces a constant sensitivity throughout the contract period, whereas the yearly fees in (a) lead to a sensitivity, which starts nearly at zero and then rises monotonously to the level of the former towards termination of the contract. The gradient vector of the annuity insurance (d) is on the interval $(0, 35)$ equal to that of the pure endowment insurance, having its maximum as well at contract time $35$, the time of retirement.

One of the most significant differences between the pure endowment insurances (a) & (b) and the temporary life insurance (c) is the sign of the gradient vectors. The former show a so-called *survival character* (cf. Milbrodt and Helbig (1999), Definition 5.31), that is, a decreasing mortality always raises the prospective reserve, and vice versa.

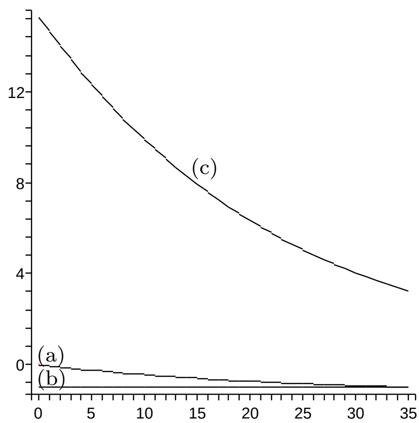


Figure 3.5.7: normalized gradient vectors $\nabla_0 V_{\mathrm{a},0}^{q_{\mathrm{ad}}^*}$ of Examples 3.5.1 (a) to (c)

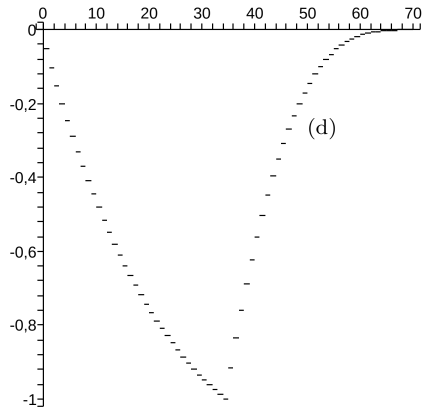


Figure 3.5.8: normalized gradient vector $\nabla_0 V_{\mathrm{a},0}^{q_{\mathrm{ad}}^*}$ of Example 3.5.1(d)

In contrast, the temporary life insurance has a positive gradient vector throughout and hence a so-called *occurrence character* (cf. Milbrodt and Helbig (1999), Definition 5.31), that is, an increasing mortality raises the prospective reserve, and vice versa.

In practice, these characteristics are commonly taken into consideration by using different life tables depending on the contract type. For example, for calculating the premium level of temporary life insurances or pure endowment insurances, life tables are used that overstate or understate the present mortality, respectively (cf. Ramlau-Hansen (1988), pp. 225, 231).

**Corollary 3.5.3.** *With $R_{gl}$ being the 'sum at risk' according to* (3.2.18)*, for any $y \in \mathcal{S}$, $(g,l) \in J$, and $s \geq 0$*

$$sgn\big(R_{gl}(u)\big) = sgn\big(\nabla_0 V_{y,s}^{q_{gl}^*}(u)\big)\,, \qquad \text{for all } u \in (s,\infty) \text{ with } p_{yg}^*(s,u-0) > 0\,.$$

*Proof.* With the second equation in (3.2.17) and with (3.2.18)

$$\nabla_0 V_{y,s}^{q_{gl}^*}(\cdot) = \mathbf{1}_{(s,\infty)}(\cdot)\, p_{yg}^*(s,\cdot-0)\, \frac{K_{\Phi^*}(s)}{K_{\Phi^*}(\cdot)}\, R_{gl}(\cdot)\,. \tag{3.5.3}$$

The accumulation factor $K_{\Phi^*}$ is strictly positive because of (1.2.4) and Proposition A.4.2. □

The sign of the sum at risk at time $u$ determines the direction the prospective reserve $V_{y,s}^{q_{gl}^*}$ is shifted to if the cumulative intensity $q_{gl}^*$ is raised from $u$ on. This result agrees with the existing literature on this subject, e.g., Ramlau-Hansen (1988, p. 231), Linnemann (1993, section 6), or Milbrodt and Helbig (1999, Bemerkungen 9.19(c)).

Note that $R_{gl}(u)$ and with it the sign of $\nabla_0 V_{y,s}^{q_{gl}^*}(u)$ is independent of time $s$ and state $y$! *That means the proper decision between overstating and understating mortalities et cetera in order to reduce the insurers risk is independent of the increasing information about the insured due to progression of time.*

Analogously to section 3.5, for the two-state models not only the signs but also the shapes of the gradient vectors $\nabla_0 V_{\mathrm{a},s}^{q_{\mathrm{ad}}^*}$ are independent of reference time $s$: With (3.5.3)

$$\nabla_0 V_{\mathrm{a},s}^{q_{\mathrm{ad}}^*}(\cdot) = \mathbf{1}_{(s,\infty)}(\cdot)\, \frac{K_{\Phi^*}(s)}{K_{\Phi^*}(\widetilde{s})}\, \frac{1}{p_{\mathrm{aa}}^*(\widetilde{s}, s)}\, \nabla_0 V_{\mathrm{a},\widetilde{s}}^{q_{\mathrm{ad}}^*}(\cdot)\,, \qquad \forall\, 0 \leq \widetilde{s} \leq s\,, \tag{3.5.4}$$

that is, the gradient vectors with respect to reference times $\widetilde{s}$ and $s$ differ on $(s, \infty)$ only in a constant factor. For models with three or more states, the dependency on $s$ is more complex.

While Examples 3.5.1 (a) to (d) are two state models, Example 3.5.1(e) has two ways of reaching state 'd' from initial state 'a': either by directly jumping to 'd' or by detouring via 'i'. Figure 3.5.9 illustrates the gradient vectors with respect to the cumulative transition intensities $q_{\mathrm{ad}}^*$ and $q_{\mathrm{id}}^*$. While in the first half of the contract period a mortality change of active persons has a greater influence on the prospective reserve than a mortality change of disabled persons, the situation is contrary in the second half.

Occasionally, it may be convenient to assume that general mortality changes affect $q_{\mathrm{ad}}$ and $q_{\mathrm{id}}$ likewise. Following the approach in (3.3.19), the sensitivity of the prospective reserve with respect to mortality 'in general' is equal to the totalized gradient vectors of transitions (a,d) and (i,d). Denoting by 'l' the subset of states $\{\mathrm{a}, \mathrm{i}\} \subset \mathcal{S}$ in which the insured is still alive, Figure 3.5.9 shows the effect of 'general' mortality changes, $\nabla_0 V_{\mathrm{a},0}^{q_{\mathrm{ld}}^*} = \nabla_0 V_{\mathrm{a},0}^{q_{\mathrm{ad}}^*} + \nabla_0 V_{\mathrm{a},0}^{q_{\mathrm{id}}^*}$.

Figure 3.5.10 illustrates the influence of reference time $s$. Contrary to the two-state models (cf. (3.5.4)) the gradient vectors with respect to $s = 0, 10, 20$ differ in more than just constant factors, for example, the location of the minimum drifts. However, the principal shape remains here quite similar.

Now look at the transitions (a,i) and (i,a). Figure 3.5.11 shows that the gradient vector with respect to reactivation is very similar to the one with respect to mortality of disabled.

Figure 3.5.12 illustrates the sensitivity with respect to changes of the cumulative disablement intensity. Compared to the other transitions, it has by far the greatest effect on the prospective reserve.

### Combining different insurance contract types

Aiming to reduce the sensitivities to cumulative transition intensity changes, a popular way is to combine different types of insurance contracts. *The gradient vector*

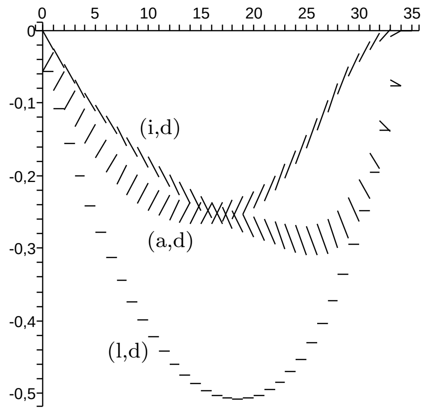


Figure 3.5.9: normalized gradient vectors $\nabla_0 V_{\mathrm{a},0}^{q^*_{z\zeta}}$ of Example 3.5.1(e) with respect to transitions $(z,\zeta) =$ (a,d), (i,d), (l,d)

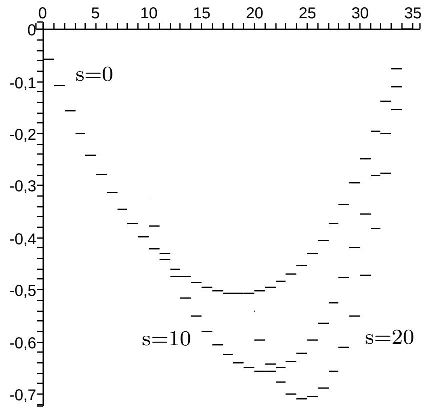


Figure 3.5.10: normalized gradient vector $\nabla_0 V_{\mathrm{a},s}^{q^*_{\mathrm{ld}}}$ of Example 3.5.1(d) with $s = 0, 10, 20$

*calculus is a convenient tool for studying the effect of such combinations as the gradient vectors of the prospective reserve are linear with respect to premiums and benefits.*

**Example 3.5.4.**

(f1) Combining the pure endowment insurance (a) and the temporary life insurance (c) of Examples 3.5.1 leads to an endowment insurance with

$$F_{\mathrm{a}}(t) = -\sum_{i=0}^{34} 0.01469 \cdot \mathbf{1}_{[i,\infty)}(t) + \mathbf{1}_{[35,\infty)}(t)\,, \quad F_{\mathrm{d}}(t) = 0\,, \quad D_{\mathrm{ad}}(t) = \mathbf{1}_{(0,35]}(t)\,.$$

(f2) If one doubles the survival benefit, one has

$$F_{\mathrm{a}}(t) = -\sum_{i=0}^{34} 0.02592 \cdot \mathbf{1}_{[i,\infty)}(t) + 2 \cdot \mathbf{1}_{[35,\infty)}(t)\,, \quad F_{\mathrm{d}}(t) = 0\,, \quad D_{\mathrm{ad}}(t) = \mathbf{1}_{(0,35]}(t)\,.$$

(g) A popular combination is to connect disability insurances with temporary life insurances. Combining (c) and (e) of Examples 3.5.1 leads to

$$F_{\mathrm{a}}(t) = -\sum_{i=0}^{33} 0.05730 \cdot \mathbf{1}_{[i,\infty)}(t)\,, \qquad F_{\mathrm{i}}(t) = \sum_{i=0}^{34} \mathbf{1}_{[i,\infty)}(t)\,,$$
$$F_{\mathrm{d}}(t) = 0\,, \qquad D_{\mathrm{ad}}(t) = \mathbf{1}_{(0,35]}(t)\,.$$

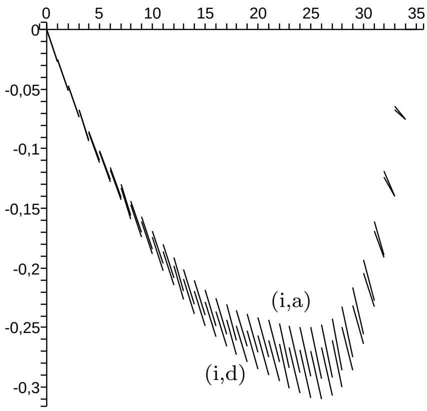


Figure 3.5.11: normalized gradient vectors $\nabla_0 V_{a,0}^{q^*_{z\zeta}}$ of Example 3.5.1(e) with respect to transitions $(z,\zeta) = (\text{i,a}), (\text{i,d})$

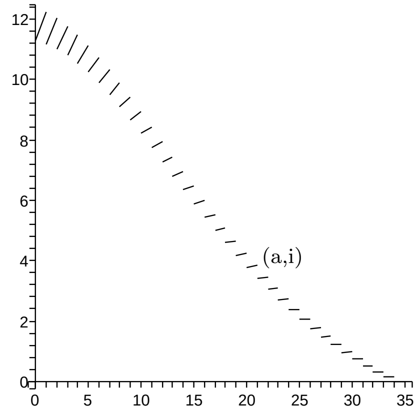


Figure 3.5.12: normalized gradient vector $\nabla_0 V_{a,0}^{q^*_{ai}}$ of Example 3.5.1(e)

(For the sake of simplicity, the premium scheme of the temporary life insurance is adapted to that of the disablement insurance.)

(h) Frequently, disablement insurances are combined with annuity insurances; here let

$$F_{\mathrm{a}}(t) = -\sum_{i=0}^{33} 0.16875 \cdot \mathbf{1}_{[i,\infty)}(t) + \sum_{i=35}^{\infty} \mathbf{1}_{[i,\infty)}(t)\,, \qquad F_{\mathrm{i}}(t) = \sum_{i=0}^{34} \mathbf{1}_{[i,\infty)}(t)\,,$$
$$F_{\mathrm{d}}(t) = 0\,, \qquad D_{\mathrm{ad}}(t) = 0\,.$$

The corresponding mean present benefits, which are needed for normalizing, are:

| | example (f1) | example (f2) | example (g) | example (h) |
|---|---|---|---|---|
| $VB_{a,0}$ | 0.27663 | 0.48824 | 1.01325 | 2.98404 |

Figure 3.5.13 shows the gradient vectors $\nabla_0 V_{a,0}^{q^*_{ad}}$ of examples (a), (c), (f1), and (f2): With the temporary life insurance and the pure endowment insurance having gradient vectors of opposite signs, combining them lets their sensitivities partly cancel out each other.

In example (g), the situation is similar. Figure 3.5.14 shows the gradient vectors with respect to transition (l,d) for examples (c), (e), and their combination (g).

Looking at the gradient vectors in Figure 3.5.15, combining a disability and an annuity insurance as in example (h) seems to be of no advantage. As both insurances

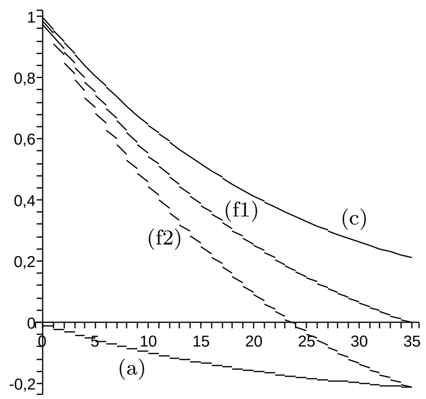


Figure 3.5.13: gradient vectors $\nabla_0 V_{\mathrm{a},0}^{q^*_{\mathrm{ad}}}$ of Examples 3.5.1 (a)&(c) and of Examples 3.5.4 (f1)&(f2)

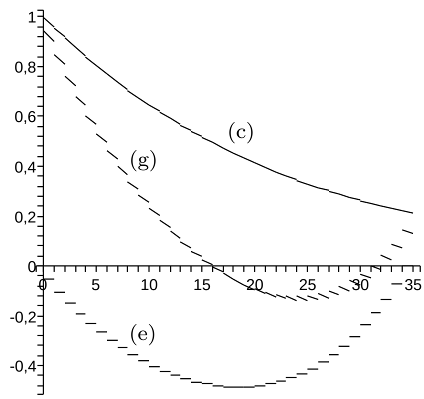


Figure 3.5.14: gradient vectors $\nabla_0 V_{\mathrm{a},0}^{q^*_{\mathrm{ld}}}$ of Examples 3.5.1 (c)&(e) and of Example 3.5.4(g)

have a survival character, there is no cancelation effect of the mortality sensitivities. This arises the question why such combinations are so popular in practice. A possible answer can be found by plotting the normalized gradient vectors with respect to disability, see Figure 3.5.16. While the mean present benefit value $VB_{\mathrm{a},0}$ of the annuity insurance is comparatively large, the sensitivity with respect to disability in example (h) is not absolutely lower than in example (e), but it is relatively lower. That means a proper risk loading has relative to the netto premium less weight.

### Sensitivity with respect to interest and all transitions simultaneously

Now sensitivities with respect to interest and all transitions *simultaneously* are studied. As already mentioned above, because of Theorem 3.3.5 this can be done by just aggregating the results of the above subsections.

Figure 3.5.17 compares the normalized gradient vectors with respect to interest and with respect to mortality of Example 3.5.1(c). The prospective reserve of temporary life insurances is much more sensitive to interest than to mortality changes.

For the pure endowment insurances (a) and (b) and the annuity insurance (d), the gradient vectors with respect to interest and to mortality look similar. In fact, they are even equal:

**Corollary 3.5.5.** *Suppose the transition probabilities are continuous in time. Then, for each* $\zeta \in \mathcal{S}$ *with* $V_{\zeta,\cdot} \equiv 0$ *and* $D_{z\zeta} \equiv 0$, $z \in \mathcal{S} \setminus \{\zeta\}$, *one has*

$$\big(1 + \Delta\Phi^*(\cdot)\big)\, \nabla_0 V_{y,s}^{\Phi^*}(\cdot) = \sum_{z \in \mathcal{S}} \nabla_0 V_{y,s}^{q^*_{z\zeta}}(\cdot)\,, \qquad \forall\, s \geq 0\,, y \in \mathcal{S}\,. \tag{3.5.5}$$

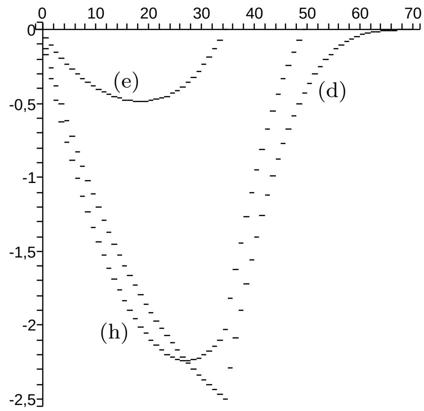


Figure 3.5.15: gradient vectors with respect to transition (l,d) of Examples 3.5.1 (d)&(e) and of Example 3.5.4 (h)

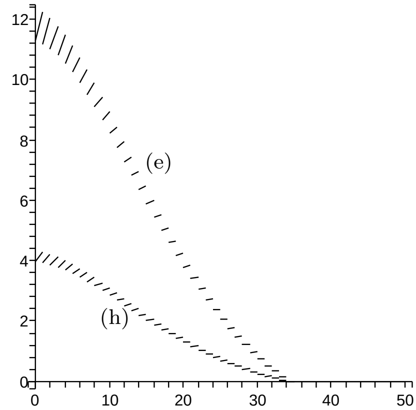


Figure 3.5.16: normalized gradient vectors $\nabla_0 V_{\mathrm{a},0}^{q^*_{\mathrm{ai}}}$ of Example 3.5.1 (e) and of Example 3.5.4 (h)

*Proof.* With (3.5.1) it follows that

$$\nabla_0 V_{y,s}^{\Phi^*}(\cdot) = -\frac{\mathbf{1}_{(s,\infty)}(\cdot)}{1+\Delta\Phi^*(\cdot)}\,\frac{K_{\Phi^*}(s)}{K_{\Phi^*}(\cdot)}\sum_{w\in\mathcal{S}} p^*_{yw}(s,\cdot)\,V_{w,\cdot}\,.$$

On the other hand equation (3.5.3) yields

$$\begin{aligned}\sum_{z\in\mathcal{R}} \nabla_0 V_{y,s}^{q^*_{z\zeta}}(\cdot) &= \mathbf{1}_{(s,\infty)}(\cdot)\,\frac{K_{\Phi^*}(s)}{K_{\Phi^*}(\cdot)}\sum_{z\in\mathcal{R}} p^*_{yz}(s,\cdot-0)\,R_{z\zeta}(\cdot)\\ &= \mathbf{1}_{(s,\infty)}(\cdot)\,\frac{K_{\Phi^*}(s)}{K_{\Phi^*}(\cdot)}\sum_{z\in\mathcal{S}} p^*_{yz}(s,\cdot)\big(-V_{z,\cdot}\big)\,.\end{aligned}$$

□

In the special case of a pure endowment insurance, the equality of the sensitivities with respect to interest and with respect to mortality is already mentioned in Kalashnikov and Norberg (2003, p. 248).

**Remark 3.5.6.** In case of the two state model $\mathcal{S} = \{\mathrm{a,d}\}$, one can abandon the continuity condition for the transition probabilities by replacing (3.5.5) with

$$\big(1+\Delta\Phi^*(\cdot)\big)\,\nabla_0 V_{\mathrm{a},s}^{\Phi^*}(\cdot) = \frac{1}{1-\Delta q^*_{\mathrm{ad}}(\cdot)}\,\nabla_0 V_{\mathrm{a},s}^{q^*_{\mathrm{ad}}}(\cdot)\,, \qquad s\ge 0\,,$$

since $p^*_{\mathrm{aa}}(s,\cdot) = p^*_{\mathrm{aa}}(s,\cdot-0)\,p^*_{\mathrm{aa}}(\cdot-0,\cdot) = p^*_{\mathrm{aa}}(s,\cdot-0)\,(1-\Delta q^*_{\mathrm{ad}}(\cdot))$.

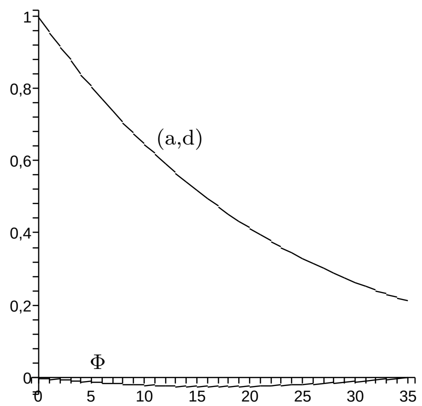


Figure 3.5.17: normalized gradient vectors of Example 3.5.1(c) with respect to interest 'Φ' and mortality '(a,d)'

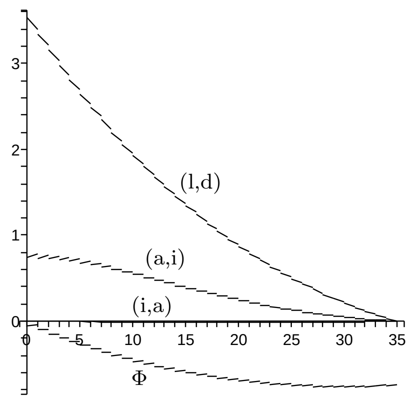


Figure 3.5.18: normalized gradient vectors of Example 3.5.7 (i)

In particular, the corollary implies that for the disability insurance of Example 3.5.1(e)

$$\nabla_0 V_{\mathrm{a},0}^{\Phi^*} = \nabla_0 V_{\mathrm{a},0}^{q^*_{\mathrm{ad}}} + \nabla_0 V_{\mathrm{a},0}^{q^*_{\mathrm{id}}} =: \nabla_0 V_{\mathrm{a},0}^{q^*_{\mathrm{ld}}}\,, \qquad \nabla_0 V_{\mathrm{i},0}^{\Phi^*} = \nabla_0 V_{\mathrm{i},0}^{q^*_{\mathrm{ad}}} + \nabla_0 V_{\mathrm{i},0}^{q^*_{\mathrm{id}}} =: \nabla_0 V_{\mathrm{i},0}^{q^*_{\mathrm{ld}}}\,,$$

that is, the sensitivity of the prospective reserve with respect to interest rate changes is equal to the one with respect to general mortality changes. However, for the disability insurance these sensitivities are clearly outweighed by the sensitivity with respect to the disability probabilities (cf. Figures 3.5.3, 3.5.9, 3.5.11, and 3.5.12).

At last, look at another typical insurance contract type:

**Example 3.5.7.**

(i) Combine a pure endowment insurance, a temporary life insurance, and a disability insurance to an endowment insurance with disability waiver,

$$F_{\mathrm{a}}(t) = -\sum_{i=0}^{33} 0.01570 \cdot \mathbf{1}_{[i,\infty)}(t) + \mathbf{1}_{[35,\infty)}(t)\,, \qquad F_{\mathrm{i}}(t) = \mathbf{1}_{[35,\infty)}(t)\,,$$
$$F_{\mathrm{d}}(t) = 0\,, \qquad D_{\mathrm{ad}}(t) = \mathbf{1}_{(0,35]}(t)\,, \qquad D_{\mathrm{id}}(t) = \mathbf{1}_{(0,35]}(t)\,.$$

Figure 3.5.18 illustrates the gradient vectors with respect to interest '$\Phi$', general mortality '(l,d)', disability '(a,i)', and reactivation '(i,a)' in one coordinate system. The sensitivity as regards the reactivation is negligible. Mortality probability and disability probability changes have greater effects on the prospective reserve towards the beginning of the contract, where the former outweighs the latter. The sensitivities to interest rate changes have the same order of magnitude, but they rise towards expiration of the contract.

### Financial risk versus systematic mortality risk

It is a popular opinion that the financial risk (interest rate risk) is generally much more important than the systematic mortality risk and that the latter is sufficiently covered by choosing a generous safety loading for the former (cf. Hoem (1988), pp. 191-192). This idea emanated from the empirical observation that mortality rates have been smaller and more stable than interest rates. In recent years, several authors questioned this notion in case of term insurances and brought up arguments against it; see, for example, Norberg (2001, section 8.4) and Helwich (2003).

The previous performed sensitivity analysis also strengthens the impression that for term insurances the systematic mortality risk is no less important than the financial risk: For the temporary life insurance of Example 3.5.1(c), the sensitivity to mortality changes by far exceeds the sensitivity to interest rate changes (see Figure 3.5.17). Even if one assumes the mortality rate has a significant smaller volatility than the interest rate, it is unlikely that the systematic mortality risk is negligible compared to the financial risk. Similarly the systematic disability risk of the disability insurance in Example 3.5.1(e) is likely not to be outnumbered by the financial risk (see Figures 3.5.3 and 3.5.12).

In order to decide which of the two contrary opinions meets the reality, one has to quantify the volatility of the interest and transition rates, preferably with respect to time. This is done in the next chapter.

# 4 An uncertainty analysis of life insurance contracts

Based on the preliminary work of section 1.3, the technical basis is here modeled stochastically in order to specify the uncertainty of the actuarial assumptions. The aim is then to quantify the financial risk, the unsystematic biometrical risk, and the systematic biometrical risks such as systematic mortality risk or systematic disability risk of individual insurance contracts.

In financial theory, stochastic interest rate approaches have already been extensively studied, resulting in a great variety of different interest rate models. Some of them found their way into the actuarial literature. A popular idea, which will be taken up here, is to model the interest rate by means of diffusion processes. See, for example, Beekman and Fuelling (1990, 1991), Parker (1994a, 1994b, 1997), Norberg and Møller (1996), Persson (1998), Perry et al. (2001, 2003), and Dahl (2004). Since Persson (1998) argued that under stochastic interest rates the classical principle of equivalence is an improper method for pricing, it became more and more accepted to valuate insurance contracts under risk-adjusted probability measures as it is common in financial mathematics. Such financial market modeling is not implemented here, as the aim is not to calculate market values but to quantify risks, which will be mainly done with the help of second order moments.

Lately, the uncertainty of mortality tables draws attention of actuaries. Some references dealing with stochastic mortality rates are Dahl (2004), Cairns et al. (2005), Biffis (2005), or Dahl and Møller (2005). Many authors just adopt interest rate modeling frameworks, which enables them to use common tools of financial mathematics.

Recalling that the aim here is to quantify and compare the financial risk, the unsystematic biometrical risk, and the systematic biometrical risks such as systematic mortality risk or systematic disability risk, the literature offers the following:

Frees (1990), Beekman and Fuelling (1990, 1991), De Schepper et al. (1992), Parker (1994a, 1997), Møller (1995), Norberg and Møller (1996), Marceau and Gaillardetz (1999), Bruno et al. (2000), Helwich (2003), and Fischer (2004) present moments of second and higher order or approximative probability distributions of the present value of an insurance contract. However, they all use deterministic mortality rates, that means they do not consider systematic biometrical risks. (Though Helwich (2003) studies the systematic mortality risk by using some form of sensitivity analysis, his risk measures for systematic and unsystematic mortality risk are not comparable to each other.)

Norberg (1999) uses a homogenous and Markovian jump process with finite state space to model the interest rate and the mortality rate stochastically. His perspective of view in terms of profits and losses is not that of the insurer but of the insured.

He presents differential equations for calculating moments of the future bonuses an insured is awaiting. However, the risk contributions of the stochastic interest rate and the stochastic mortality rate are not separated. Olivieri (2001) uses a scenario-based method to study the uncertainty of mortality probabilities, but disregards the financial risk. Khalaf-Allah et al. (2006) use a model with deterministic interest rates and stochastic mortality. They quantify the systematic mortality risk of annuities at different ages by simulating the probability distributions of the present values.

Most authors in the actuarial literature who deal with stochastic mortality rates concentrate on valuation and hedging of life insurance liabilities. The quantifying of the various financial and biometrical risks of life insurance contracts in general is still an open field. This chapter offers at least approximative methods for decomposing the overall risk to its different sources and calculating risk measures for its components.

The structure of this chapter is as follows: In a first step, the cumulative interest intensity process and the cumulative transition intensity matrix process of Condition 1.3.1 are further specified.

The task is then to further decompose the technical basis risk (second addend in (1.4.2)) to components referring to the uncertainty of the interest rate, the mortality rate, the disability rate, et cetera. This is done asymptotically in section 4.2, where a further approximation approach is presented in order to study the probability distributions of the elements of the decomposition.

An empirical study in section 4.3 applies the methods of section 4.2 on typical life insurance contract types. It shows that none of the analyzed risks is, in general, negligible and studies the advantage of combining different insurance contract types concerning the technical basis risk.

Throughout this chapter assume that the following condition holds:

**Condition 4.0.8** (finite time horizon)**.** Let $F_z|_{\mathbb{R}\setminus[0,T]} \equiv 0$ for all $z \in \mathcal{S}$, and let $D_{z\zeta}|_{\mathbb{R}\setminus[0,T]} \equiv 0$ for all $(z,\zeta) \in J$, where $T < \infty$ is the *time horizon.*

As a consequence the distinction of 'starting point' and 'deviation' as stated in (3.0.1) vanishes, since for the compact intervals $[0, DT(T)]$ and $[0,T]$

$$V_{y,s}[\Phi, q] = V_{y,s}[\mathbf{1}_{[0,DT(T)]}\,\Phi, \mathbf{1}_{[0,T]}\,q]\,.$$

## 4.1 Probabilistic model for the technical basis

In this section, the cumulative interest intensity process and the cumulative transition intensity matrix process of Condition 1.3.1 are further specified.

### Modeling stochastic interest

The literature offers various approaches for modeling stochastic interest. A comprehensive reference is the book of Brigo and Mercurio (2001). According to Parker (1994b), most approaches in the actuarial literature are of one of the two following forms: The first approach is to let the interest intensity $\varphi$ be a stochastic process, the second approach is to model the *cumulative* interest intensity $\Phi$ as a stochastic process. The first approach is adopted here, because contrary to the second one it will ensure that the cumulative interest intensity $\Phi$ has finite variation on compacts (cf. Condition 1.3.1(a)).

A popular approach of financial mathematics is to let the interest intensity $\varphi$ be a stochastic diffusion process,

$$\mathrm{d}\varphi_t = \alpha(\varphi_t, t)\,\mathrm{d}t + \sigma(\varphi_t, t)\,\mathrm{d}W_t\,, \quad t \geq 0\,, \varphi_0 = \text{const}\,, \tag{4.1.1}$$

where the *drift term* $\alpha(\varphi, t)$ and the *diffusion term* $\sigma(\varphi, t)$ are proper $\mathbb{R} \times [0, \infty) \to \mathbb{R}$ functions and $W_t$ is a standard Wiener process. Commonly the drift or diffusion term are called *linear* if $\alpha(\varphi_t, t) = \alpha_1(t)\,\varphi_t + \alpha_2(t)$ or $\sigma(\varphi_t, t) = \sigma_1(t)\,\varphi_t + \sigma_2(t)$ for all $t \geq 0$, respectively. If additionally $\sigma_1(t) \equiv 0$, it is said that (4.1.1) is a linear stochastic differential equation *with additive noise*.

The following proposition defines the interest rate modeling framework used here.

**Proposition 4.1.1.** *Let the interest intensity be of the form*

$$\varphi_t := e_0(t) + \sum_{i=1}^{n} e_i(t)\,\phi_{i,t}\,, \qquad t \in [0, DT(T)]\,, n \in \mathbb{N}\,, \tag{4.1.2}$$

*where the* $(\phi_{i,t})$ *are stochastically independent and* $t$*-continuous solutions of* (4.1.1) *with linear drift term, and the* $e_i$ *are Lebesgue-measurable and bounded functions on* $[0, DT(T)]$.

*Then, the cumulative interest intensity* $(\Phi_t)$ *is a well-defined stochastic process on* $[0, DT(T)]$ *that meets Condition* 1.3.1(a).

*Proof.* As the processes $(\phi_{i,t})$ are $t$-continuous on $[0, DT(T)]$, the mappings $(t, \omega) \mapsto \phi_{i,t}(\omega)$ are into $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ measurable (cf. Elliot (1982), Theorem 2.32). With the help of Tonellis Theorem, the cumulative interest intensity at arbitrary $t \in [0, DT(T)]$,

$$\Phi_t = \int_{(0,t]} \left( e_0(\tau) + \sum_{i=1}^{n} e_i(\tau)\,\phi_{i,\tau} \right) \mathrm{d}\tau\,, \tag{4.1.3}$$

is measurable as well. Note that for arbitrary but fixed $\omega$ the processes $(\phi_{i,t})(\omega)$ are bounded on $[0, DT(T)]$, that is, the integral (4.1.3) is finite. This also implies that the paths $t \mapsto \Phi_t(\omega)$ have finite variation on compacts. The continuity of the paths implies the validity of the lower jump bound condition in (1.2.4). □

The linearity of the drift term is not necessary here, but will be needed later on. Now look at some examples. First of all, consider linear diffusions with additive noise:

**Example 4.1.2.**

(a) *Vasicek Model:* Vasicek assumed that the interest intensity evolves as an Ornstein-Uhlenbeck process with constant coefficients,

$$\mathrm{d}\varphi_t = \eta\,(\theta - \varphi_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t\,, \qquad \eta, \theta, \sigma > 0\,.$$

(b) *Hull-White extended Vasicek Model:* Hull and White introduced time-varying parameters in the Vasicek Model,

$$\mathrm{d}\varphi_t = \big(\vartheta(t) - \lambda(t)\,\varphi_t\big)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}W_t\,,$$

where $\vartheta$, $\lambda$, and $\sigma$ are deterministic functions of time.

Assuming the initial interest rate $\varphi_0$ is a deterministic constant, the above examples are all Gaussian processes, which is a general property of linear diffusion processes with additive noise (cf. Kloeden and Platen (1992), pp. 110-111, 564). That implies they can get negative with probability greater than zero. Referring to Remark 1.2.1 the opinions in the literature differ in whether negative interest rates meet the reality. In order to avoid this negativity, several non-Gaussian interest rate models have been developed:

**Example 4.1.3.**

(c) *Dothan Model:* Dothan proposed 'multiplicative noise' instead of 'additive noise',

$$\mathrm{d}\varphi_t = \theta\,\varphi_t\,\mathrm{d}t + \sigma\,\varphi_t\,\mathrm{d}W_t\,, \qquad \theta \in \mathbb{R}\,, \sigma > 0\,.$$

(d) *Cox-Ingersoll-Ross Model:* Cox, Ingersoll, and Ross inserted a 'square root' term in the diffusion coefficient of the interest intensity dynamics proposed by Vasicek,

$$\mathrm{d}\varphi_t = \eta\,(\theta - \varphi_t)\,\mathrm{d}t + \sigma\,\sqrt{\varphi_t}\,\mathrm{d}W_t\,, \qquad \eta,\ \theta,\ \sigma > 0\,.$$

Upon starting from a positive initial value $\varphi_0$, this process remains positive if $2\eta\theta > \sigma^2$.

(e) *Hull-White extended Cox-Ingersoll-Ross Model:* As for the Vasicek Model, Hull and White introduced time-varying parameters in the Cox-Ingersoll-Ross Model,

$$\mathrm{d}\varphi_t = \big(\vartheta(t) - \lambda(t)\,\varphi_t\big)\,\mathrm{d}t + \sigma(t)\,\sqrt{\varphi_t}\,\mathrm{d}W_t\,,$$

where $\vartheta$, $\lambda$, and $\sigma$ are deterministic functions of time.

The Examples 4.1.2 and 4.1.3 are all *one-factor* diffusion processes. A frequently used generalization is to form linear combinations of them according to (4.1.2).

### Modeling stochastic transition probabilities

Compared to the modeling of stochastic interest, the modeling of stochastic transition probabilities has been much less studied so far. Recently, an unforseen increase of life expectancies has provoked several approaches to stochastic mortality rates. A reference with a good overview is Cairns et al. (2005). It is a popular concept in the literature to treat the force of mortality in a similar way as short-term rate of interest. With stochastic diffusions being the most frequently used continuous time models for interest rates, various authors employ them to model stochastic mortality rates, e.g., Dahl (2004), Biffis (2005), or Cairns et al. (2005).

Following that idea, let here all transition intensities be linear combinations of stochastically independent diffusion processes. More precisely:

**Proposition 4.1.4.** *Let the transition intensities be of the form*

$$\mu_{z\zeta,t} := e_{z\zeta,0}(t) + \sum_{i=1}^{n} e_{z\zeta,i}(t)\, \pi_{z\zeta,i,t}\,, \qquad t \in [0,T]\,, m \in \mathbb{N}\,, (z,\zeta) \in J\,, \tag{4.1.4}$$

*where the* $(\pi_{z\zeta,i,t})_{t\geq 0}$ *(*$i \in \{1,\ldots,n\}$*,* $(z,\zeta) \in J$*) are stochastically independent and* $t$*-continuous solutions of* (4.1.1) *with linear drift term, and the* $e_{z\zeta,i}$ *are Lebesgue-measurable and bounded functions on* $[0, DT(T)]$*. Further, let the processes* $(\mu_{z\zeta,t})_{t\geq 0}$ *be nonnegative.*

*Then, the cumulative interest intensity* $(q_t)$ *is a well-defined stochastic process on* $[0, DT(T)]$ *that meets Condition* 1.3.1(b).

*Proof.* The proof is analogous to the proof of Proposition 4.1.1. For the regularity see Milbrodt and Helbig (1999, pp. 197-198, Exercise 20). □

The linearity of the drift term is not necessary here, but will be needed later on.

While nonnegativity was a matter of opinion for the interest intensity, it is inevitable for the transition intensities. This disqualifies all linear diffusion processes with additive noise. Because of their positivity, the processes of Example 4.1.3 could be a good choice. For instance, Dahl (2004) suggests to use extended Cox-Ingersoll-Ross models (cf. Example 4.1.3(e)) for modeling the mortality rate, whereas Korn et al. (2006, p. 408) prefer log-normal processes such as Dothans model (cf. Example 4.1.3(c)).

Note that the time parameter $t$ of the transition intensities stands for both proceeding time and increasing age. Considering, for example, the mortality, the stochastic process $(\mu_{\text{ad},t})$ models a cohort life table. Note that the reference time $t = 0$ relates here to the beginning of the contract period and not to the date of birth of the cohort.

## 4.2 Decomposition of risk and approximation of its components

The task is now to further decompose the overall technical basis risk (second addend in (1.4.2)) to components referring to the uncertainty of the interest rate, the mortality rate, the disability rate, et cetera.

Using Theorem 1.3.6 the conditional expectation (1.4.1) is representable by

$$\mathbb{E}\Big(\tilde{B}_s \,\Big|\, (\Phi_t), (q_t)\Big) \overset{a.s.}{=} \sum_{y \in \mathcal{S}} P\big(\tilde{X}_s = y \,\big|\, (\Phi_t), (q_t)\big)\, V_{y,s}[(\Phi_t), (q_t)]\,. \tag{4.2.1}$$

*For the sake of simplicity, from now on let the reference time $s$ be zero.* In case the reference time $s$ is greater than zero and the history up to time $s$ is completely known, the following results hold analogously by means of a time shift. Supposing $\tilde{X}_0$ is deterministic with initial state $y \in \mathcal{S}$, equation (4.2.1) gets

$$\mathbb{E}\Big(\tilde{B}_0 \,\Big|\, (\Phi_t), (q_t)\Big) \overset{a.s.}{=} V_{y,0}[(\Phi_t), (q_t)]\,. \tag{4.2.2}$$

A common concept in 'Uncertainty Analysis' is to approximate functionals of random variables by their first order Taylor expansion (cf. Saltelli et al. (2000), section 5.9.1). Generalizing that idea, the prospective reserve $V_{y,0}$ as a functional of the technical basis is now expanded to a Taylor series of first order by means of the gradient vector (3.3.18) (cf. (2.3.4)): Following (3.0.1) and (3.3.1), decompose the cumulative interest and transition intensities to

$$\begin{aligned}
(\Phi, q) &= (\Phi^*, q^*) + (H_\Phi, H_q) := (\Phi^*, q^*) + (H_\Phi, q(\cdot; H_J) - q^*) \ \in \mathcal{F} \times \mathcal{Q}\,,\\
(\Phi, q_J) &= (\Phi^*, q_J^*) + (H_\Phi, H_J) := (\Phi^*, (q^*_{z\zeta})_{(z,\zeta)\in J}) + (H_\Phi, (H_{z\zeta})_{(z,\zeta)\in J}) \ \in (BVC_{\leftarrow})^{1+|J|}\,,
\end{aligned} \tag{4.2.3}$$

where $(\Phi^*, q^*)$ or $(\Phi^*, q_J^*)$ form some 'starting point' and $(H_\Phi, H_q)$ or $(H_\Phi, H_J)$ are 'deviations' of it. In case the deviation is small, the prospective reserve at $(\Phi, q)$ equals approximately its first order Taylor expansion at $(\Phi^*, q^*)$ (see Theorem 3.3.3):

$$V_{y,0}[\Phi, q] = V_{y,s}^{(\Phi^*,q^*)}(H_\Phi, H_J) = \underbrace{V_{y,0}^{(\Phi^*,q^*)}(0) + \mathrm{D}_0 V_{y,0}^{(\Phi^*,q^*)}(H_\Phi, H_J)}_{\text{first order Taylor expansion}} + R^{(\Phi^*,q^*)}(H_\Phi, H_J)\,, \tag{4.2.4}$$

where the remainder satisfies

$$\frac{R^{(\Phi^*,q^*)}(H_\Phi, H_J)}{\|(H_\Phi, H_J)\|_{BV}} \longrightarrow 0\,, \quad \text{for } \|(H_\Phi, H_J)\|_{BV} \to 0\,.$$

Assume that $H_\Phi$ and $H_J$ have Radon-Nikodym derivatives with respect to some dominating measure $\nu$, write $h_\Phi := \mathrm{d}H_\Phi/\mathrm{d}\nu$ and $h_{z\zeta} := \mathrm{d}H_{z\zeta}/\mathrm{d}\nu$ for $(z,\zeta) \in J$. By applying Theorem 3.3.5, the first order Taylor expansion is equal to

$$\begin{aligned}
&V_{y,0}^{(\Phi^*,q^*)}(0) + \int \nabla_0 V_{y,0}^{(\Phi^*,q^*)} \cdot (h_\Phi, h_J)\,\mathrm{d}\nu\\
&= V_{y,0}^{(\Phi^*,q^*)}(0) + \int \nabla_0 V_{y,0}^{\Phi^*}\, h_\Phi\,\mathrm{d}\nu + \sum_{(z,\zeta)\in J} \int \nabla_0 V_{y,0}^{q^*_{z\zeta}}\, h_{z\zeta}\,\mathrm{d}\nu\,.
\end{aligned} \tag{4.2.5}$$

Suppose now the starting point $(\Phi^*, q^*)$ is deterministic and only $(h_\Phi, h_J)$ accounts for randomness. Then, the conditional expectation (1.4.1) is in (4.2.5) asymptotically decomposed into a sum of random variables, which uniquely correspond to the uncertainties of the different components of the technical basis.

In case the intensities $h_\Phi$ and $h_{z\zeta}$, $(z,\zeta) \in J$, are Gaussian processes and $\nu$ is the Lebesgue measure, the integrals are normally distributed random variables:

**Corollary 4.2.1.** *Let the interest intensity $(\varphi_t)$ be defined in accordance with Proposition 4.1.1, and let the processes $(\phi_{i,t})$ in (4.1.2) be linear diffusions with additive noise. Further, let the transition intensities $(\mu_{z\zeta,t})$ be defined in accordance with Proposition 4.1.4, but here skip the nonnegativity condition, and let the processes $(\pi_{z\zeta,i,t})$ in (4.1.4) also be linear diffusions with additive noise. Perform the decomposition (4.2.3) as follows:*

$$\begin{aligned} \Phi_t = \Phi^*(t) + H_{\Phi,t} &= \int_0^t \mathbb{E}\varphi_\tau \,\mathrm{d}\tau + \int_0^t \big(\varphi_\tau - \mathbb{E}\varphi_\tau\big)\,\mathrm{d}\tau\,, \quad t \in [0, DT(T)]\,, \\ q_{z\zeta,t} = q^*_{z\zeta}(t) + H_{z\zeta,t} &= \int_0^t \mathbb{E}\mu_{z\zeta,\tau}\,\mathrm{d}\tau + \int_0^t \big(\mu_{z\zeta,\tau} - \mathbb{E}\mu_{z\zeta,\tau}\big)\,\mathrm{d}\tau\,, \quad t \in [0,T]\,. \end{aligned} \tag{4.2.6}$$

*Then, under the assumptions of Theorem 3.3.5, the first order Taylor expansion (4.2.5) equals a constant plus a sum of stochastically independent and normally distributed random variables with zero expectation.*

*Proof.* With $\Phi^*$ and $q^*_{z\zeta}$ being deterministic, the prospective reserve $V^{(\Phi^*,q^*)}_{y,0}(0)$ is a constant. Applying Proposition A.5.3, the integrals in (4.2.5) are normally distributed with zero expectations, as the stochastic processes $(h_{\Phi,t}) = (\varphi_t - \mathbb{E}\varphi_t)$ and $(h_{z\zeta,t}) = (\mu_{z\zeta,t} - \mathbb{E}\mu_{z\zeta,t})$ have zero expectations as well. □

In order to calculate the corresponding variances, decompose the intensities $(\varphi_t)$ and $(\mu_{z\zeta,t})$ as stated in (4.1.2) and (4.1.4), and apply Proposition A.5.3 for each of the stochastically independent $(\phi_{i,t})$ and $(\pi_{z\zeta,i,t})$ separately. Note that the supports of the gradient vectors in (4.2.5) are subsets of $[0, DT(T)]$.

**Remark 4.2.2.** The decomposition (4.2.6) is by no means exclusive. As long as the starting point $(\Phi^*, q^*)$ is deterministic, various other decompositions are imaginable. However, the purpose should always be to keep the deviation $(H_\Phi, H_J)$ small (in a proper sense) in order to have a negligible remainder $R$ in (4.2.4). Having this intention in mind, decomposition (4.2.6) is motivated by the forthcoming Proposition 4.2.3.

Unfortunately, letting the intensities be linear combinations of linear diffusions with additive noise implies they can get negative with probability greater than zero. While such a model property is a matter of opinion for the interest intensity, it is definitely unrealistic for the transition intensities. Some authors ignore this fact (cf. Biffis (2005), p. 459), arguing that the parameters of the model can be chosen in such a way that the intensity processes take negative values with negligible low probability. For convenience this approach is adopted here, approximating diffusions with linear drift term and non-additive noise by linear diffusions with additive noise:

**Proposition 4.2.3.** *Let for $0 < \varsigma \leq 1$ the process $(\phi_t)_{t\geq 0}$ be the (unique) solution of*

$$\mathrm{d}\phi_t = \big(\alpha_1(t)\,\phi_t + \alpha_2(t)\big)\,\mathrm{d}t + \varsigma\,\sigma(\phi_t, t)\,\mathrm{d}W_t\,, \quad t \in [0, DT(T)]\,, \phi_0 = \mathrm{const}\,, \tag{4.2.7}$$

*where the drift term $\alpha(x,t) := \alpha_1(t)\,x + \alpha_2(t)$ and and the diffusion term $\sigma(x,t)$ satisfy conditions* (a) *to* (c) *of Proposition* A.5.1 *on* $[0, DT(T)]$. *Further, let $(f_t)_{t\geq 0}$ be the (unique) solution of*

$$\mathrm{d}f_t = \big(\alpha_1(t)\,f_t + \alpha_2(t)\big)\,\mathrm{d}t + \varsigma\,\sigma(\mathbb{E}\phi_t, t)\,\mathrm{d}W_t\,, \quad t \in [0, DT(T)]\,, f_0 = \phi_0 = \mathrm{const}\,. \tag{4.2.8}$$

*Then, for each $(k,l) \in \mathbb{N}_0^2$ there exists a constant $C_{kl} < \infty$ in such a way that*

$$E^{k,l} := \mathbb{E}\Big|(\phi_t - f_t)^k(\phi_t - \mathbb{E}\phi_t)^l\Big| \leq C_{kl}\,\varsigma^{2k+l}\,, \qquad \forall\, t \in [0, DT(T)]\,, 0 < \varsigma \leq 1\,. \tag{4.2.9}$$

That means approximating $(\phi_t)$ by its mean $(\mathbb{E}\phi_t)$ is of order $1$, an approximation by the process $(f_t)$ is of order $2$. Supposing the interest and transition intensities are linear combinations of processes of the form (4.2.7), approximating them by linear combinations of processes of the form (4.2.8) allows for applying Corollary 4.2.1.

*Proof.* For a shorter notation define $\widetilde{\phi}_t := \mathbb{E}(\phi_t)$, which, according to Kloeden and Platen (1992, p. 113), is a deterministic function satisfying $\mathrm{d}\widetilde{\phi}_t = \big(\alpha_1(t)\,\widetilde{\phi}_t + \alpha_2(t)\big)\,\mathrm{d}t$ for $t \in [0, DT(T)]$. The (twice continuously) differentiable function

$$g_\varepsilon : \mathbb{R} \to (0, \infty)\,, \qquad x \mapsto \begin{cases} |x| & : |x| \geq \varepsilon \\ \frac{-1}{8\varepsilon^3}x^4 + \frac{3}{4\varepsilon}x^2 + \frac{3\varepsilon}{8} & : |x| < \varepsilon \end{cases}$$

is for any $\varepsilon > 0$ a majorant of $x \mapsto |x|$ on $\mathbb{R}$. Thus, property (4.2.9) holds if for an $\varepsilon > 0$ there exists a constant $C_{kl,\varepsilon} < \infty$ with

$$E^{k,l} \leq E^{k,l}_\varepsilon := \mathbb{E}g_\varepsilon\big((\phi_t - f_t)^k(\phi_t - \widetilde{\phi}_t)^l\big) \leq C_{kl,\varepsilon}\,\varsigma^{2k+l}\,, \qquad \forall\, t \in [0, DT(T)]\,. \tag{4.2.10}$$

Aiming to apply the Ito Formula, function $g_\varepsilon$ has in contrast to the absolute value function the advantage of being differentiable at zero. Now use the principle of mathematical induction to proof (4.2.10):

*Induction basis:* Let $(k,l) = (0,0)$. Then, $E^{00}_\varepsilon = \mathbb{E}g_\varepsilon(1) =: C_{00,\varepsilon} < \infty$.

*Induction step:* The succession of the induction steps is tricky here. At first, several induction arguments are presented, after that the succession is explained. To avoid case differentiations, define $< z > := \max\{z, 0\}$ and $E^{k,l} := 0$ for all $(k,l) \in \mathbb{Z}^2 \setminus \mathbb{N}_0^2$.

The Ito Formula (cf. Kloeden and Platen (1992), pp. 96, 97) leads to

$$
\begin{aligned}
&g_\varepsilon\big((\phi_t - f_t)^k(\phi_t - \widetilde{\phi}_t)^l\big) \\
&= \int_0^t g'_\varepsilon\big((\phi_s - f_s)^k(\phi_s - \widetilde{\phi}_s)^l\big)\bigg(k\,(\phi_s - f_s)^{<k-1>}(\phi_s - \widetilde{\phi}_s)^l\,\alpha_1(s)\,(\phi_s - f_s) \\
&\qquad\qquad + l\,(\phi_s - f_s)^k(\phi_s - \widetilde{\phi}_s)^{<l-1>}\,\alpha_1(s)\,(\phi_s - \widetilde{\phi}_s)\bigg)\,\mathrm{d}s \\
&\quad + \frac{1}{2}\int_0^t \varsigma^2 g'_\varepsilon\big((\phi_s - f_s)^k(\phi_s - \widetilde{\phi}_s)^l\big)\bigg(l\,(l-1)\,(\phi_s - f_s)^k(\phi_s - \widetilde{\phi}_s)^{<l-2>}\,\big(\sigma(\phi_s, s)\big)^2 \\
&\qquad\qquad + 2\,k\,l\,(\phi_s - f_s)^{<k-1>}(\phi_s - \widetilde{\phi}_s)^{<l-1>}\,\big(\sigma(\phi_s, s) - \sigma(\widetilde{\phi}_s, s)\big)\sigma(\phi_s, s) \\
&\qquad\qquad + k\,(k-1)\,(\phi_s - f_s)^{<k-2>}(\phi_s - \widetilde{\phi}_s)^l\,\big(\sigma(\phi_s, s) - \sigma(\widetilde{\phi}_s, s)\big)^2\bigg)\,\mathrm{d}s \\
&\quad + \frac{1}{2}\int_0^t \varsigma^2 g''_\varepsilon\big((\phi_s - f_s)^k(\phi_s - \widetilde{\phi}_s)^l\big)\bigg(l^2\,(\phi_s - f_s)^{2k}(\phi_s - \widetilde{\phi}_s)^{2<l-1>}\,\big(\sigma(\phi_s, s)\big)^2 \\
&\qquad\qquad + 2\,k\,l\,(\phi_s - f_s)^{<2k-1>}(\phi_s - \widetilde{\phi}_s)^{<2l-1>}\,\big(\sigma(\phi_s, s) - \sigma(\widetilde{\phi}_s, s)\big)\sigma(\phi_s, s) \\
&\qquad\qquad + k^2\,(\phi_s - f_s)^{2<k-1>}(\phi_s - \widetilde{\phi}_s)^{2l}\,\big(\sigma(\phi_s, s) - \sigma(\widetilde{\phi}_s, s)\big)^2\bigg)\,\mathrm{d}s \\
&\quad + \int_0^t \ldots\,\mathrm{d}W_s\,.
\end{aligned}
$$

Taking expectations on both sides makes the last integral zero (cf. Kloeden and Platen (1992), pp. 87, 88). (The quadratic integrability of the integrand is a consequence of Theorem 4.5.4 in Kloeden and Platen (1992, p. 136) and Hölders inequality.) As

- $-1 \le g'_\varepsilon(x) \le 1$ and $0 \le x\, g''_\varepsilon(x) \le 3/2$ for all $x \in \mathbb{R}$,
- $|\alpha_1(s)| \le K$ for all $s \in [0, DT(T)]$,
- $|\sigma(\phi_s, s) - \sigma(\widetilde{\phi}_s, s)| \le K|\phi_s - \widetilde{\phi}_s|$ for all $s \in [0, DT(T)]$,
- $|\sigma(\phi_s, s)| \le |\sigma(\phi_s, s) - \sigma(\widetilde{\phi}_s, s)| + |\sigma(\widetilde{\phi}_s, s)| \le K|\phi_s - \widetilde{\phi}_s| + |\sigma(\widetilde{\phi}_s, s)|$ for all $s \in [0, DT(T)]$, and
- $|\sigma(\phi_s, s)|^2 \le 2K^2|\phi_s - \widetilde{\phi}_s|^2 + 2|\sigma(\widetilde{\phi}_s, s)|^2$ for all $s \in [0, DT(T)]$,

the expectation of the right hand side has for $k \neq 1$ and $l \neq 1$ an upper bound of

$$\begin{aligned}
\mathbb{E} g_\varepsilon\big((\phi_t - f_t)^k(\phi_t - \widetilde{\phi}_t)^l\big) \leq{}& \int_0^t (k+l)\, E^{k,l}\, K \,\mathrm{d}s \\
&+ \frac{\varsigma^2}{2}\int_0^t \Big( l(l-1)\, E^{k,l}\, 2\, K^2 + l(l-1)\, E^{k,l-2}\, 2\, |\sigma(\widetilde{\phi}_s, s)|^2 \\
&\qquad + 2\, k\, l\, E^{k-1,l}\, K^2 + 2\, k\, l\, E^{k-1,l+1}\, K\, |\sigma(\widetilde{\phi}_s, s)| \\
&\qquad + k(k-1)\, E^{k-2,l+2}\, K^2 \Big)\,\mathrm{d}s \\
&+ \frac{\varsigma^2}{2}\int_0^t \frac{3}{2}\Big( l^2\, E^{k,l}\, 2\, K^2 + l^2\, E^{k,l-2}\, 2\, |\sigma(\widetilde{\phi}_s, s)|^2 \\
&\qquad + 2\, k\, l\, E^{k-1,l}\, K^2 + 2\, k\, l\, E^{k-1,l+1}\, K\, |\sigma(\widetilde{\phi}_s, s)| \\
&\qquad + k^2\, E^{k-2,l+2}\, K^2 \Big)\,\mathrm{d}s\,.
\end{aligned}$$

Since the deterministic function $s \mapsto \widetilde{\phi}_s$ is bounded on the compact interval $[0, DT(T)]$, the mapping $s \mapsto |\sigma(\widetilde{\phi}_s, s)|$ is bounded as well, which leads to

$$E^{k,l} \leq E^{k,l}_\varepsilon \leq \int_0^t \Big(\mathrm{const}\, E^{k,l} + \mathrm{const}\, \varsigma^2 \Big(E^{k,l-2} + E^{k-1,l} + E^{k-1,l+1} + E^{k-2,l+2}\Big)\Big)\,\mathrm{d}s\,.$$

In case (4.2.10) holds for $\{(k, l-2), (k-1, l), (k-1, l+1), (k-2, l+2)\}$, one gets

$$E^{k,l} \leq \int_0^t \mathrm{const}\, E^{k,l}\,\mathrm{d}s + \mathrm{const}\, \varsigma^2 \Big(\varsigma^{2k+l-2} + \varsigma^{2k-2+l} + \varsigma^{2k-2+l+1} + \varsigma^{2k-4+l+2}\Big)\,.$$

Then, from Gronwalls Inequality (cf. Kloeden and Platen (1992), Lemma 4.5.1) it follows that

$$E^{k,l} \leq \mathrm{const}\, \varsigma^2 \Big(\varsigma^{2k+l-2} + \varsigma^{2k-2+l} + \varsigma^{2k-2+l+1} + \varsigma^{2k-4+l+2}\Big) \leq \mathrm{const}\, \varsigma^{2k+l} \qquad (4.2.11)$$

for all $(k, l) \in \mathbb{N}_0^2$ with $k \neq 1$ and $l \neq 1$.

Assume now $k = 1$ and $l \neq 1$. Then, the inequality $0 \leq g''_\varepsilon(x) \leq 3/(2\varepsilon)$, $\forall x \in \mathbb{R}$, yields analogously to the above

$$E^{1,l} \leq E^{1,l}_\varepsilon \leq \int_0^t \mathrm{const}\, E^{1,l}\,\mathrm{d}s + \mathrm{const}\, \varsigma^2 \Big(\varsigma^{2+l-2} + \varsigma^{l} + \varsigma^{l+1} + \frac{1}{\varepsilon}\varsigma^{2l+2}\Big)\,,$$

if (4.2.10) holds for $\{(1, l-2), (0, l), (0, l+1), (0, l+2)\}$. Using Gronwalls Inequality again, for arbitrary but fixed $\varepsilon > 0$ and $l \in \mathbb{N}_0 \setminus \{1\}$

$$E^{1,l} \leq \mathrm{const}\, \varsigma^2 \Big(\varsigma^{2+l-2} + \varsigma^{l} + \varsigma^{l+1} + \varsigma^{2l+2}\Big) \leq \mathrm{const}\, \varsigma^{2+l}\,. \qquad (4.2.12)$$

Now suppose that (4.2.10) holds for an arbitrary but fixed $(k,l) \in \mathbb{N}_0 \times \mathbb{N}$ and the corresponding $(k,0) \in \mathbb{N}_0 \times \{0\}$. Define the finite measure $\nu$ by $\mathrm{d}\nu/\mathrm{d}\mathbb{P} = (\phi_t - f_t)^k$. Then, from Hölders Inequality it follows that

$$\begin{aligned}
E^{k,l-1} &= \int |\phi_t - f_t|^k |\phi_t - \widetilde{\phi}_t|^{l-1} \, \mathrm{d}\mathbb{P} \\
&= \int |\phi_t - \widetilde{\phi}_t|^{l-1} \, \mathrm{d}\mu \\
&\leq \left( \int |\phi_t - \widetilde{\phi}_t|^{(l-1)l/(l-1)} \, \mathrm{d}\mu \right)^{(l-1)/l} \left( \int 1^l \, \mathrm{d}\mu \right)^{1/l} \\
&= \left( \int |\phi_t - f_t|^k |\phi_t - \widetilde{\phi}_t|^l \, \mathrm{d}\mathbb{P} \right)^{(l-1)/l} \left( \int |\phi_t - f_t|^k \, \mathrm{d}\mathbb{P} \right)^{1/l} \\
&= \left( E^{k,l} \right)^{(l-1)/l} \left( E^{(k,0)} \right)^{1/l} \\
&\leq \mathrm{const}\, \varsigma^{2k+l-1} .
\end{aligned} \tag{4.2.13}$$

Using (4.2.11), (4.2.12), and (4.2.13), let the induction steps be in the following order:

(1) Starting from the induction basis $(0,0)$, use (4.2.11) to approve (4.2.10) for the pairs $(0,2), (0,4), (0,6), \ldots$ and so forth.
(2) Now apply (4.2.13) to obtain (4.2.10) for $\{(0,1), (0,3), (0,5), \ldots\}$.
(3) As with (1) and (2) the property (4.2.10) holds for all $\{0\} \times \mathbb{N}_0$, use (4.2.12) to approve it for $(1,0), (1,2), (1,4), \ldots$ and so forth.
(4) Again apply (4.2.13) to show (4.2.10) for the odd numbers $\{(1,1), (1,3), (1,5), \ldots\}$.
(5) Analogously to (1), use (4.2.11) to approve the statement for $(2,2), (2,4), (2,6), \ldots$ and so forth.
(6) Analogously to (2), apply (4.2.13) for approving the pairs $\{(2,1), (2,3), (2,5), \ldots\}$.
(7) Proceeding with $k = 3, 4, 5, \ldots$, repeat the steps (5) and (6).

□

The procedure for analyzing (4.2.2) is now as follows:

- Firstly, approximate (4.2.2) by its first order Taylor expansion (4.2.5).
- Secondly, approximate the intensity processes of the technical basis by Gaussian processes in terms of Proposition 4.2.3.

The result is a sum of normally distributed random variables, which correspond uniquely to the different risk sources, and whose distributions are calculable by applying Proposition A.5.3.

For example, if one uses the variance as risk measure and assumes the intensity processes are stochastically independent, the technical basis risk approximately equals

$$\mathrm{Var}\Big( \mathbb{E}\big( \tilde{B}_0 \big| \, (\Phi_t) \,, (q_t) \big) \Big) \approx \mathrm{Var}\left( \int \nabla_0 V_{y,0}^{\Phi^*} \, h_\Phi \, \mathrm{d}\nu \right) + \sum_{(z,\zeta) \in J} \mathrm{Var}\left( \int \nabla_0 V_{y,0}^{q^*_{z\zeta}} \, h_{z\zeta} \, \mathrm{d}\nu \right) . \tag{4.2.14}$$

(The right hand side is easily calculable, not only if $h_\Phi$ and $h_{z\zeta}$, $(z,\zeta)\in J$, are Gaussian processes. With the help of Fubinis Theorem it suffices to know the expectation functions and covariance functions of $(h_{\Phi,t})$ and $(h_{z\zeta,t})$, $(z,\zeta)\in J$, provided $\Phi^*$ and $q^*$ are deterministic.)

Now look at the first addend of decomposition (1.4.2), which refers to the unsystematic biometrical risk. Calculating its probability distribution is very challenging. Again a linearization can help to calculate at least its variance asymptotically: Following the proof of Proposition 1.4.1 and applying Theorem 1.3.9 leads to

$$\begin{aligned}\operatorname{Var}\Big(\tilde{B}_s - \mathbb{E}\big(\tilde{B}_s \,\big|\, (\Phi_t),(q_t)\big)\Big) &= \mathbb{E}\Big(\operatorname{Var}\big(\tilde{B}_s\big|\,(\Phi_t)\,,(q_t)\big)\Big)\\ &= \mathbb{E}\bigg(\sum_{y\in\mathcal{S}} P\big(\tilde{X}_s = y \,\big|\, (\Phi_t),(q_t)\big)\, \mathcal{V}_{y,s}[(\Phi_t),(q_t)]\bigg)\,.\end{aligned} \tag{4.2.15}$$

Similar to the study of (4.2.1), for the sake of simplicity only the special case $s=0$ is studied here. Supposing $\tilde{X}_0$ is deterministic with initial state $y\in\mathcal{S}$, equation (4.2.15) gets

$$\operatorname{Var}\Big(\tilde{B}_0 - \mathbb{E}\big(\tilde{B}_0 \,\big|\, (\Phi_t),(q_t)\big)\Big) = \mathbb{E}\Big(\mathcal{V}_{y,0}[(\Phi_t),(q_t)]\Big)\,. \tag{4.2.16}$$

Since it is still very challenging to calculate the mean of $\mathcal{V}_{y,s}[(\Phi_t),(q_t)]$, the idea is now to approximate $\mathcal{V}_{y,s}[(\Phi_t),(q_t)]$ by its first order Taylor expansion similar to (4.2.4). As the modeling framework of section 4.1 implies that the cumulative transition intensity matrix is continuous, it suffices to study the special case (1.2.15) instead of the general formula (1.2.14).

**Theorem 4.2.4.** *Let $t\mapsto D_{z\zeta}(t)$ be bounded on $[0,T]$ for each $(z,\zeta)\in J$. With the cumulative intensities $(\Phi,q_J)$ being decomposed in accordance with* (4.2.3)*, the conditional variance* (1.2.15) *– regarded as a functional of deviation $(H_\Phi,H_J)$ –*

$$\mathcal{V}_{y,s}: \mathcal{E}_{\Phi^*}\times (BV_{\leftarrow})^{|J|}\to\mathbb{R}\,,\quad (H_\Phi,H_J)\mapsto \mathcal{V}_{y,s}\big[\Phi^*+H_\Phi, q_J^*+H_J\big]\,, \tag{4.2.17}$$

*is Fréchet differentiable at zero.*

*Proof.* The finite time horizon condition allows for substituting $(\Phi^*,q_J^*)$ and $(H_\Phi,H_J)$ by $\mathbf{1}_{[0,DT(T)]}(\Phi^*,q_J^*)$ and $\mathbf{1}_{[0,DT(T)]}(H_\Phi,H_J)$; that is, without loss of generality one may assume finite total variation for the cumulative intensities. Now Proposition 3.1.5, Theorem 3.3.3, and property (3.3.9) imply that for any $t\in[s,T]$ the discounting factor $K(s)/K(t)$, the prospective reserves $V_{y,t}$ $(y\in\mathcal{S})$, and the transition probabilities $p_{yz}(s,t-0)$ $((y,z)\in J)$ are Fréchet differentiable at zero as mappings of the deviation $(H_\Phi,H_J)$. Because of (3.1.11) combined with the boundedness of $1/K(t)$ (cf. (A.4.2)), (3.3.14) combined with the boundedness of $K(t)$ (cf. (A.4.3)), and because of (3.3.17), the convergence (A.3.3) is uniform on $t\in[s,T]$. Using the product rule for Fréchet differentials (cf. Flett (1980), Exercises 3.5 No. 2), the product

$$G\big(t,(H_\Phi,H_J)\big) := \frac{K(s)^2}{K(t)^2}\big(V_{\zeta,t}-V_{z,t}+D_{z\zeta}(t)\big)^2\,p_{yz}(s,t-0)$$

as a mapping of the deviation $(H_\Phi,H_J)$ is for arbitrary $y\in\mathcal{S}$ and $(z,\zeta)\in J$ at zero Fréchet differentiable as well for any $t\in[s,T]$. Again the convergence (A.3.3) is uniform on $t\in[s,T]$, that is,

$$\Big|G\big(t,(H_\Phi,H_J)\big)-G(t,0)-\mathrm{D}_0G\big(t,(H_\Phi,H_J)\big)\Big| \le \mathit{Rem}(H_\Phi,H_J)\,,\qquad \forall\, t\in[s,T]\,,$$

where the remainder satisfies $\lim_{\|(H_\Phi,H_J)\|_{BV}\to 0}\mathit{Rem}(H_\Phi,H_J)/\|(H_\Phi,H_J)\|_{BV}=0$. Consequently, for arbitrary $y\in\mathcal{S}$ and $(z,\zeta)\in J$

$$\begin{aligned}
&\bigg|\int_{(s,T]} G\big(t,(H_\Phi,H_J)\big)\,(q^*_{z\zeta}+H_{z\zeta})(\mathrm{d}t)-\int_{(s,T]} G(t,0)\,q^*_{z\zeta}(\mathrm{d}t)\\
&-\int_{(s,T]}\mathrm{D}_0G\big(t,(H_\Phi,H_J)\big)\,q^*_{z\zeta}(\mathrm{d}t)-\int_{(s,T]} G(t,0)\,H_{z\zeta}(\mathrm{d}t)\bigg|\\
&\le \mathit{Rem}(H_\Phi,H_J)\,\|q^*_{z\zeta}\|_{BV}+\sup_{t\in[0,T]}\Big|G\big(t,(H_\Phi,H_J)\big)-G(t,0)\Big|\,\|H_{z\zeta}\|_{BV}\\
&= o\Big(\|(H_\Phi,H_J)\|_{BV}\Big)\,.
\end{aligned}$$

(Note that the convergence $|G(t,(H_\Phi,H_J))-G(t,0)|\to 0$ is uniform on $t\in[s,T]$, cf. (3.1.10) and (3.3.5).) Thus, the mapping (4.2.17) is Fréchet differentiable at zero with Fréchet differential

$$\sum_{(z,\zeta\in J)}\int_{(0,T]}\mathrm{D}_0G\big(t,(H_\Phi,H_J)\big)\,q^*_{z\zeta}(\mathrm{d}t)+\sum_{(z,\zeta\in J)}\int_{(0,T]} G(t,0)\,H_{z\zeta}(\mathrm{d}t)\,.$$

□

That means $\mathcal{V}_{y,0}[\Phi,q]$ has for continuous $q$ a first order Taylor expansion of the form

$$\mathcal{V}_{y,0}[\Phi,q]=\mathcal{V}_{y,0}(H_\Phi,H_J)=\underbrace{\mathcal{V}_{y,0}(0)+\mathrm{D}_0\mathcal{V}_{y,0}(H_\Phi,H_J)}_{\text{first order Taylor expansion}}+\tilde{R}^{(\Phi^*,q^*)}(H_\Phi,H_J)\,,\quad (4.2.18)$$

where the remainder satisfies

$$\frac{\tilde{R}^{(\Phi^*,q^*)}(H_\Phi,H_J)}{\|(H_\Phi,H_J)\|_{BV}}\longrightarrow 0\,,\quad \|(H_\Phi,H_J)\|_{BV}\to 0\,.$$

Now assume again the deviations $H_\Phi$ and $H_J$ have Radon-Nikodym derivatives $h_\Phi := \mathrm{d}H_\Phi/\mathrm{d}\nu$ and $h_{z\zeta} := \mathrm{d}H_{z\zeta}/\mathrm{d}\nu$ with respect to some dominating measure $\nu$. Since the mapping (2.2.1) is Fréchet differentiable, the variance $\mathcal{V}_{y,0}$ is also Fréchet differentiable at zero as a mapping of its intensities $(h_\Phi, h_J)$ because of the chain rule (A.3.6). According to section 2.1, the first order Taylor expansion in (4.2.18) may now be rewritten to

$$\mathcal{V}_{y,0}(0) + \int \nabla_0 \mathcal{V}_{y,0} \cdot (h_\Phi, h_J)\, \mathrm{d}\nu\,, \tag{4.2.19}$$

where $\nabla_0 \mathcal{V}_{y,s}$ is the gradient vector given by Definition 2.1.2. Supposing $(\Phi^*, q^*)$ is deterministic and only $(h_\Phi, h_J)$ accounts for randomness, it is quite easy to calculate the expectation of (4.2.19) because of the linearity with respect to $(h_\Phi, h_J)$. (With the help of Fubinis Theorem it suffices to know the expectation functions of $(h_{\Phi,t})$ and $(h_{z\zeta,t})$.) In doing so, one gets an approximation of (4.2.16). For deviations $(H_\Phi, H_J)$ in accordance with Corollary 4.2.1, the expectation is always zero,

$$\mathbb{E}\Big( \int \nabla_0 \mathcal{V}_{y,0} \cdot (h_\Phi, h_J)\, \mathrm{d}\nu \Big) = 0\,. \tag{4.2.20}$$

### Convergence of the approximations

The former section yielded approximations (i) by using Taylor expansions of first order and (ii) by applying Proposition 4.2.3. This section studies the convergence rates of these approximations. Again for the sake of simplicity, the following theorems are only formulated for reference time $s = 0$. In case of $s > 0$, one gets similar results.

To begin with, look at the first order Taylor expansions in (4.2.4) and (4.2.18):

**Theorem 4.2.5.** *Let the interest intensity $(\varphi_t)$ and the transition intensities $(\mu_{z\zeta,t})$ be stochastically independent and defined in accordance with Propositions* 4.1.1 *and* 4.1.4*, and let the cumulative interest intensity $(\Phi_t)$ and the cumulative transition intensities $(q_{z\zeta,t})$ be decomposed as stated in* (4.2.6)*. Analogously to* (4.2.7)*, let $0 < \varsigma \leq 1$ be a scaling parameter of the diffusion terms of the processes $(\phi_{i,t})$ and $(\pi_{z\zeta,i,t})$ in* (4.1.2) *and* (4.1.4)*.*

*Then, there exist a constant $C < \infty$ and for each $k \in \mathbb{N}$ constants $C_k < \infty$ with*

$$\begin{aligned} &\mathbb{E}\Big| R^{(\Phi^*,q^*)}(H_\Phi, H_J) \Big|^k = \mathbb{E}\Big| V_{y,0}^{(\Phi^*,q^*)}(H_\Phi, H_J) - V_{y,0}^{(\Phi^*,q^*)}(0) - \mathrm{D}_0 V_{y,0}^{(\Phi^*,q^*)}(H_\Phi, H_J) \Big|^k \\ &\leq C_k\, \varsigma^{2k} \sup_{0 \leq s \leq DT(T)} \mathbb{E} e^{-(1+\varepsilon)k\, H_\Phi(s)}\,, \end{aligned} \tag{4.2.21}$$

$$\begin{aligned} &\Big| \mathbb{E}\, \tilde{R}^{(\Phi^*,q^*)}(H_\Phi, H_J) \Big| = \Big| \mathbb{E}\Big( \mathcal{V}_{y,0}(H_\Phi, H_J) - \mathcal{V}_{y,s}(0) - \mathrm{D}_0 \mathcal{V}_{y,0}(H_\Phi, H_J) \Big) \Big| \\ &\leq C\, \varsigma^2 \sup_{\substack{s_1,\ldots,s_5 \in [0,DT(T)]\,, \\ s_2 > s_3\,,\, s_4 > s_5}} \mathbb{E}\, e^{-(1+\varepsilon)\,(2H_\Phi(s_1) + (H_\Phi(s_2) - H_\Phi(s_3)) + (H_\Phi(s_4) - H_\Phi(s_5)))} \end{aligned} \tag{4.2.22}$$

*for an arbitrary but fixed $\varepsilon > 0$.*

*Proof.* Since $t \mapsto H_{\Phi,t}$ is continuous and the cumulative transition intensity matrices $(q_t)$ and $q^*$ are regular (cf. Exercise 20 on p. 197 in Milbrodt and Helbig (1999)), the remainder $R^{(\Phi^*,q^*)}(H_\Phi, H_J)$ has an upper bound of the form (3.3.17). Because of the finite time horizon (Condition 4.0.8), one may replace $(H_\Phi, H_J)$ by $\mathbf{1}_{[0,DT(T)]}(H_\Phi, H_J)$.

Hölders Inequality implies

$$(a_1 + \cdots + a_n)^k \le n^{k-1}(a_1^k + \cdots + a_n^k)\,, \qquad \forall\,(a_1,\ldots,a_n) \in [0,\infty)^k\,, n \in \mathbb{N}\,, k \ge 1\,. \tag{4.2.23}$$

This inequality, again Hölders Inequality, Fubinis Theorem, Proposition 4.2.3, and the boundedness of the $e_i$ in (4.1.2) lead to

$$\begin{aligned}
\mathbb{E}\|\mathbf{1}_{[0,DT(T)]}H_\Phi\|_{BV}^k &\le \mathbb{E}\Big(\int_0^{DT(T)} \sum_{i=1}^n |e_i(\tau)|\,\big|\phi_{i,\tau} - \mathbb{E}\phi_{i,\tau}\big|\,\mathrm{d}\tau\Big)^k \\
&\le n^{k-1}\sum_{i=1}^n \mathbb{E}\Big(\int_0^{DT(T)} |e_i(\tau)|\,\big|\phi_{i,\tau} - \mathbb{E}\phi_{i,\tau}\big|\,\mathrm{d}\tau\Big)^k \\
&\le n^{k-1}\sum_{i=1}^n \mathbb{E}\Big(\int_0^{DT(T)} |e_i(\tau)|^k\,\big|\phi_{i,\tau} - \mathbb{E}\phi_{i,\tau}\big|^k\,\mathrm{d}\tau\Big)\Big(\int_0^{DT(T)} \mathrm{d}\tau\Big)^{k-1} \\
&= n^{k-1}\sum_{i=1}^n \Big(\int_0^{DT(T)} |e_i(\tau)|^k\,\mathbb{E}\big|\phi_{i,\tau} - \mathbb{E}\phi_{i,\tau}\big|^k\,\mathrm{d}\tau\Big)DT(T)^{k-1} \\
&\le \mathrm{const}_k\,\varsigma^k\,,
\end{aligned}$$

for each $k \in \mathbb{N}$. Analogously, $\mathbb{E}\|\mathbf{1}_{[0,DT(T)]}H_{z\zeta}\|_{BV}^k \le \mathrm{const}_k\,\varsigma^k$, which can be extended to the vector $H_J$,

$$\begin{aligned}
\mathbb{E}\|\mathbf{1}_{[0,DT(T)]}H_J\|_{BV}^k &= \mathbb{E}\Big(\sum_{(z,\zeta)\in J} \|\mathbf{1}_{[0,DT(T)]}H_{z\zeta}\|_{BV}\Big)^k \\
&\le |J|^{k-1}\sum_{(z,\zeta)\in J} \mathbb{E}\|\mathbf{1}_{[0,DT(T)]}H_{z\zeta}\|_{BV}^k \\
&\le \mathrm{const}_k\,\varsigma^k\,.
\end{aligned}$$

As the support of $\gamma$ (cf. definition (3.3.16)) is a subset of $[0, DT(T)]$, using Hölders

Inequality twice and applying Fubinis Theorem yield for $1+\varepsilon = 1+1/m$, $m \in \mathbb{N}$,

$$\begin{aligned}
&\mathbb{E}\Big(\|\mathbf{1}_{[0,DT(T)]}H_\Phi\|_{BV}^2 \int_0^{DT(T)} e^{-H_\Phi(t)}\gamma(\mathrm{d}t)\Big)^k \\
&\leq \Big(\mathbb{E}\|H_\Phi\|_{BV}^{2k(m+1)}\Big)^{1/(m+1)} \bigg(\mathbb{E}\Big(\int_0^{DT(T)} e^{-H_\Phi(t)}\gamma(\mathrm{d}t)\Big)^{k(1+\varepsilon)}\bigg)^{1/(1+\varepsilon)} \\
&\leq \Big(\mathrm{const}_k\, \varsigma^{2k(m+1)}\Big)^{1/(m+1)} \bigg(\mathbb{E}\int_0^{DT(T)} e^{-(1+\varepsilon)k\, H_\Phi(t)}\gamma(\mathrm{d}t)\ \ DT(T)^{k(1+\varepsilon)-1}\bigg)^{1/(1+\varepsilon)} \\
&\leq \mathrm{const}_k\, \varsigma^{2k} \bigg(\int_0^{DT(T)} \sup_{0\leq s\leq DT(T)} \mathbb{E}e^{-(1+\varepsilon)k\, H_\Phi(s)}\, \gamma(\mathrm{d}t)\ \ DT(T)^{k(1+\varepsilon)-1}\bigg)^{1/(1+\varepsilon)} \\
&\leq \mathrm{const}_k\, \varsigma^{2k} \sup_{0\leq s\leq DT(T)} \mathbb{E}e^{-(1+\varepsilon)k\, H_\Phi(s)}\,,
\end{aligned}$$

where the last inequality uses $\sup_{0\leq s\leq DT(T)} \mathbb{E}e^{-(1+\varepsilon)k\, H_\Phi(s)} \geq e^{-(1+\varepsilon)k\, H_\Phi(0)} = 1 \geq 0$. Analogously, with the stochastic independence of interest and transition intensities,

$$\begin{aligned}
&\mathbb{E}\Big(\|\mathbf{1}_{[0,DT(T)]}H_\Phi\|_{BV}\|\mathbf{1}_{[0,DT(T)]}H_J\|_{BV} \int_0^{DT(T)} e^{-H_\Phi(t)}\gamma(\mathrm{d}t)\Big)^k \\
&\leq \mathrm{const}_k\, \varsigma^{2k} \sup_{0\leq s\leq DT(T)} \mathbb{E}e^{-(1+\varepsilon)k\, H_\Phi(s)}\,.
\end{aligned}$$

Similarly one gets

$$\begin{aligned}
&\mathbb{E}\Big(\|\mathbf{1}_{[0,DT(T)]}H_\Phi\|_{BV} \int_0^{DT(T)} e^{-H_\Phi(t)}\,\mathrm{d}|H_{z\zeta}|(t)\Big)^k \\
&\leq \mathrm{const}_k\, \varsigma^k \bigg(\mathrm{const}_k \int_0^{DT(T)} \mathbb{E}e^{-(1+\varepsilon)k\, H_\Phi(t)}\, |b_{z\zeta}(t)|^k\, \mathbb{E}\big|\mu_{z\zeta,t} - \mathbb{E}\mu_{z\zeta,t}\big|^{k(1+\varepsilon)}\,\mathrm{d}t\bigg)^{1/(1+\varepsilon)} \\
&\leq \mathrm{const}_k\, \varsigma^k \sup_{0\leq s\leq DT(T)} \Big(\mathbb{E}e^{-(1+\varepsilon)k\, H_\Phi(s)}\Big)\, \mathrm{const}_k\, \varsigma^k\,.
\end{aligned}$$

Now (4.2.21) follows from applying all these inequalities to (3.3.17) at $\mathbf{1}_{[0,DT(T)]}(H_\Phi, H_J)$.

Taking into consideration Proposition 3.1.4, property (3.3.7), property (3.3.8), and the ideas of Remark 3.3.4, the proof of (4.2.22) is similar to that for (4.2.21). □

This theorem is of help only if the suprema in (4.2.21) and (4.2.22) are finite. At first glance, this seems to be a strong requirement. However, it holds for Examples 4.1.2 and 4.1.3: For the latter, the nonnegativity of the interest intensities lets the mapping $s \mapsto \Phi_s$ be nonnegative, which means

$$\mathbb{E}\, e^{-(1+\varepsilon)k\, H_\Phi(s)} = e^{(1+\varepsilon)k\, \Phi^*(s)}\, \mathbb{E}\, e^{-(1+\varepsilon)k\, \Phi_s} \leq e^{(1+\varepsilon)k\, \Phi^*(s)} \leq \mathrm{const}_k\,, \quad \forall\, s \in [0, DT(T)]\,.$$

For Examples 4.1.2, Proposition A.5.3 yields that for each $s \geq 0$ the random variable $\Phi_s$ is normally distributed with some expectation $\mu_s \in \mathbb{R}$ and some variance $\sigma_s^2 \in [0,\infty)$. Consequently, $e^{-(1+\varepsilon)k\, \Phi_s}$ is log-normal distributed with expectation

$$\mathbb{E}\, e^{-(1+\varepsilon)k\, \Phi_s} = e^{\left(-(1+\varepsilon)\, k\, \mu_s + (1+\varepsilon)^2 k^2 \sigma_s^2/2\right)}\,, \qquad \forall\, k \in \mathbb{N}, s \geq 0\,.$$

The boundedness of the mappings $s \mapsto \mu_s$ and $s \mapsto \sigma_s^2$ on $s \in [0, DT(T)]$ lets the above expectation be bounded on $[0, DT(T)]$ as well. Hence,

$$\mathbb{E}\, e^{-(1+\varepsilon)k\, H_\Phi(s)} = e^{(1+\varepsilon)k\, \Phi^*(s)}\, \mathbb{E}\, e^{-(1+\varepsilon)k\, \Phi_s} \leq \mathrm{const}_k\,, \quad \forall\, s \in [0, DT(T)]\,.$$

Using Hölders Inequality twice leads to

$$\begin{aligned}&\mathbb{E}\, e^{-(1+\varepsilon)\,(2H_\Phi(s_1)+(H_\Phi(s_2)-H_\Phi(s_3))+(H_\Phi(s_4)-H_\Phi(s_5)))}\\ &\quad\leq \mathbb{E}\, e^{-(1+\varepsilon)4\, H_\Phi(s_1)}\, \mathbb{E}\, e^{-(1+\varepsilon)4\,(H_\Phi(s_2)-H_\Phi(s_3))}\, \mathbb{E}\, e^{-(1+\varepsilon)4\,(H_\Phi(s_4)-H_\Phi(s_5))}\,.\end{aligned}$$

Analogous to the above arguments, in case of Examples 4.1.2 and 4.1.3 each of the three factors is bounded on $[0, DT(T)]$, that means the supremum in (4.2.22) is finite, too.

Now look at the error resulting from approximating interest and transition intensities in terms of Proposition 4.2.3.

**Theorem 4.2.6.** *Assume to be in the setting of Theorem* 4.2.5. *Denote by* $(\bar\varphi_t)$ *and* $(\bar\mu_{z\zeta,t})$ *the intensity processes corresponding to* $(\varphi_t)_{t\leq 0}$ *and* $(\mu_{z\zeta,t})_{t\geq 0}$, *for which the diffusions* $(\phi_{i,t})$ *and* $(\pi_{z\zeta,i,t})$ *in* (4.1.2) *and* (4.1.4) *are substituted by diffusions with additive noise in terms of* (4.2.8). *Analogously to* (4.2.6), *define*

$$\begin{aligned}\bar\Phi_t = \bar\Phi^*(t) + \bar H_{\Phi,t} &:= \int_0^t \mathbb{E}\bar\varphi_\tau\,\mathrm{d}\tau + \int_0^t \big(\bar\varphi_\tau - \mathbb{E}\bar\varphi_\tau\big)\,\mathrm{d}\tau\,, \quad t \in [0, DT(T)]\,,\\ \bar q_{z\zeta,t} = \bar q^*_{z\zeta}(t) + \bar H_{z\zeta,t} &:= \int_0^t \mathbb{E}\bar\mu_{z\zeta,\tau}\,\mathrm{d}\tau + \int_0^t \big(\bar\mu_{z\zeta,\tau} - \mathbb{E}\bar\mu_{z\zeta,\tau}\big)\,\mathrm{d}\tau\,,\ \ t \in [0,T]\,, (z,\zeta) \in J\,.\end{aligned} \tag{4.2.24}$$

*Then,* $(\bar\Phi^*, \bar q^*) = (\Phi^*, q^*)$, *and for each* $k \in \mathbb{N}$ *there exists a constant* $C_k < \infty$ *with*

$$\mathbb{E}\Big(\mathrm{D}_0 V_{y,s}^{(\Phi^*,q^*)}(H_\Phi, H_J) - \mathrm{D}_0 V_{y,s}^{(\Phi^*,q^*)}(\bar H_\Phi, \bar H_J)\Big)^k \leq C_k\,\varsigma^2\,, \tag{4.2.25}$$

$$\mathbb{E}\Big(\mathrm{D}_0 \mathcal{V}_{y,s}^{(\Phi^*,q^*)}(H_\Phi, H_J) - \mathrm{D}_0 \mathcal{V}_{y,s}^{(\Phi^*,q^*)}(\bar H_\Phi, \bar H_J)\Big)^k \leq C_k\,\varsigma^2\,. \tag{4.2.26}$$

*Proof.* The equality $(\bar\Phi^*, \bar q^*) = (\Phi^*, q^*)$ is due to $\mathbb{E}\varphi_t \equiv \mathbb{E}\bar\varphi_t$ and $\mathbb{E}\mu_{z\zeta,t} \equiv \mathbb{E}\bar\mu_{z\zeta,t}$ (cf. Kloeden and Platen (1992, p. 113) for the calculation of the expectations). Using the gradient vector calculus, one can write

$$\begin{aligned}&\mathrm{D}_0 V_{y,s}^{(\Phi^*,q^*)}(H_\Phi, H_J) - \mathrm{D}_0 V_{y,s}^{(\Phi^*,q^*)}(\bar H_\Phi, \bar H_J)\\ &= \int_0^{DT(T)} \nabla_0 V_{y,s}^{(\Phi^*,q^*)} \cdot \mathrm{d}(H_\Phi - \bar H_\Phi, H_J - \bar H_J)\\ &= \int_0^{DT(T)} \nabla_0 V_{y,s}^{\Phi^*}(\tau)\,\big(\varphi_\tau - \bar\varphi_\tau\big)\,\mathrm{d}\tau + \sum_{(z,\zeta)\in J} \int_0^{DT(T)} \nabla_0 V_{y,s}^{q^*_{z\zeta}}(\tau)\,\big(\mu_{z\zeta,\tau} - \bar\mu_{z\zeta,\tau}\big)\,\mathrm{d}\tau\,.\end{aligned}$$

By dint of (4.2.23), Hölders Inequality, and Fubinis Theorem, one gets

$$
\begin{aligned}
&\mathbb{E}\Big(\mathrm{D}_0 V_{y,s}^{(\Phi^*,q^*)}(H_\Phi, H_J) - \mathrm{D}_0 V_{y,s}^{(\Phi^*,q^*)}(\bar{H}_\Phi, \bar{H}_J)\Big)^k \\
&\le (1+|J|)^{k-1}\Bigg(\mathbb{E}\bigg|\int_0^{DT(T)} \nabla_0 V_{y,s}^{\Phi^*}(\tau)\ \big(\varphi_\tau - \bar{\varphi}_\tau\big)\,\mathrm{d}\tau\bigg|^k \\
&\qquad\qquad + \sum_{(z,\zeta)\in J} \mathbb{E}\bigg|\int_0^{DT(T)} \nabla_0 V_{y,s}^{q^*_{z\zeta}}(\tau)\,\big(\mu_{z\zeta,\tau} - \bar{\mu}_{z\zeta,\tau}\big)\,\mathrm{d}\tau\bigg|^k\Bigg) \\
&\le (1+|J|)^{k-1}\Bigg(\bigg(\int_0^{DT(T)} \Big|\nabla_0 V_{y,s}^{\Phi^*}(\tau)\Big|^{k/(k-1)}\,\mathrm{d}\tau\bigg)^{k-1} \int_0^{DT(T)} \mathbb{E}\big|\varphi_\tau - \bar{\varphi}_\tau\big|^k\,\mathrm{d}\tau \\
&\qquad + \sum_{(z,\zeta)\in J}\bigg(\int_0^{DT(T)} \Big|\nabla_0 V_{y,s}^{q^*_{z\zeta}}(\tau)\Big|^{k/(k-1)}\,\mathrm{d}\tau\bigg)^{k-1} \int_0^{T} \mathbb{E}\big|\mu_{z\zeta,\tau} - \bar{\mu}_{z\zeta,\tau}\big|^k\,\mathrm{d}\tau\Bigg).
\end{aligned}
$$

Applying (4.2.23) to $|\varphi_\tau - \bar{\varphi}_\tau|^k$ and $|\mu_{z\zeta,\tau} - \bar{\mu}_{z\zeta,\tau}|^k$ combined with (4.2.9) and the boundedness of the gradient vectors on $[0, DT(T)]$ yields (4.2.25). The proof of (4.2.26) is completely analogous. □

**Corollary 4.2.7.** *Under the assumptions of Theorems* 4.2.5 *and* 4.2.6*, there exist a constant* $C < \infty$ *and for any* $k \in \mathbb{N}$ *constants* $C_k$ *in such a manner that*

$$
\begin{aligned}
&\mathbb{E}\Big|V_{y,0}^{(\Phi^*,q^*)}(H_\Phi, H_J) - V_{y,0}^{(\Phi^*,q^*)}(0) - \mathrm{D}_0 V_{y,0}^{(\Phi^*,q^*)}(\bar{H}_\Phi, \bar{H}_J)\Big|^k \\
&\le C_k\,\varsigma^{2k} \sup_{0\le s\le DT(T)} \mathbb{E}e^{-(1+\varepsilon)k\,H_\Phi(s)}\,,
\end{aligned}
\tag{4.2.27}
$$

$$
\begin{aligned}
&\Big|\mathbb{E}\Big(\mathcal{V}_{y,0}^{(\Phi^*,q^*)}(H_\Phi, H_J) - \mathcal{V}_{y,0}^{(\Phi^*,q^*)}(0) - \mathrm{D}_0 \mathcal{V}_{y,0}^{(\Phi^*,q^*)}(\bar{H}_\Phi, \bar{H}_J)\Big)\Big| \\
&\le C\,\varsigma^2 \sup_{\substack{s_1,\dots,s_5\in[0,DT(T)]\,,\\ s_2>s_3\,,\,s_4>s_5}} \mathbb{E}\,e^{-(1+\varepsilon)\,(2H_\Phi(s_1)+(H_\Phi(s_2)-H_\Phi(s_3))+(H_\Phi(s_4)-H_\Phi(s_5)))}\,.
\end{aligned}
\tag{4.2.28}
$$

*Proof.* The corollary is a consequence of Theorems 4.2.5 and 4.2.6. □

**Corollary 4.2.8.** *Suppose the assumptions of Theorems* 4.2.5 *and* 4.2.6 *hold, and let the supremum in* (4.2.27) *be finite. Writing* $D_\varsigma$ *and* $\bar{D}_\varsigma$ *for the probability distribution functions of* $V_{y,0}^{(\Phi^*,q^*)}(H_\Phi, H_J)$ *and* $V_{y,0}^{(\Phi^*,q^*)}(0) + \mathrm{D}_0 V_{y,0}^{(\Phi^*,q^*)}(\bar{H}_\Phi, \bar{H}_J)$*, respectively, one has*

$$
D_\varsigma^{-1}(y) - \bar{D}_\varsigma^{-1}(y) \longrightarrow 0\,, \quad \varsigma \to 0\,, \tag{4.2.29}
$$

*for each* $y \in (0,1)$.

*Proof.* According to Corollary 4.2.1, the term

$$
V_{y,0}^{(\Phi^*,q^*)}(0) + \mathrm{D}_0 V_{y,0}^{(\Phi^*,q^*)}(\bar{H}_\Phi, \bar{H}_J)
$$

is normally distributed with expectation $V_{y,0}^{(\Phi^*,q^*)}(0)$ and has a variance of the form $const \cdot \varsigma^2$. For $\varsigma \to 0$ it converges in quadratic mean to the probability distribution $\varepsilon_c$, which puts mass 1 at point $c := V_{y,0}^{(\Phi^*,q^*)}(0)$. Due to (4.2.27) for $k = 2$, the random variable $V_{y,0}^{(\Phi^*,q^*)}(H_\Phi, H_J)$ also converges in quadratic mean to $\varepsilon_c$. Since convergence in quadratic mean implies weak convergence of the corresponding probability distributions, Theorem 5.67 in Witting and Müller-Funk (1995) yields

$$D_\varsigma^{-1}(y) \longrightarrow \varepsilon_c^{-1}(y)\,,\ \varsigma \to 0\,, \qquad \bar{D}_\varsigma^{-1}(y) \longrightarrow \varepsilon_c^{-1}(y)\,,\ \varsigma \to 0\,,$$

for each $y$ at which $\varepsilon_c^{-1}$ is continuous. □

## 4.3 Uncertainties of typical life insurance contracts

Similar to section 3.5, the following examples are either two-state or three-state models with state spaces $\mathcal{S} = \{\mathrm{a}, \mathrm{d}\}$ and $\mathcal{S} = \{\mathrm{a}, \mathrm{i}, \mathrm{d}\}$, where the states 'a', 'i', and 'd' stand for 'alive and fit', 'incapable of working', and 'dead', respectively. The corresponding sets of possible direct transitions are $J = \{(\mathrm{a}, \mathrm{d})\}$ and $J = \{(\mathrm{a}, \mathrm{i}), (\mathrm{a}, \mathrm{d}), (\mathrm{i}, \mathrm{a}), (\mathrm{i}, \mathrm{d})\}$.

**Scenario 1:** For modeling the interest intensity, follow the statistical investigation of Fischer et al. (2004) of the German bond market. They propose to model the German yield curve by some CIR-2-model, for which they supply parameter estimations based on market data from 1972 to 2002 (see p. 381). The short-rate $(r_t)$ is then a sum of two independent CIR-processes,

$$\begin{aligned} r_t &:= \phi_{1,t} + \phi_{2,t}\,, \quad \forall t \geq 0\,, \\ \mathrm{d}\phi_{1,t} &= (0.0187 - 0.4833\,\phi_{1,t})\,\mathrm{d}t + 0.1156\sqrt{\phi_{1,t}}\,\mathrm{d}W_{1,t}\,, \quad \forall t \geq 0\,, \\ \mathrm{d}\phi_{2,t} &= (0.0010 - 0.0586\,\phi_{2,t})\,\mathrm{d}t + 0.0453\sqrt{\phi_{2,t}}\,\mathrm{d}W_{2,t}\,, \quad \forall t \geq 0\,. \end{aligned}$$

The processes $(\phi_{1,t})$ and $(\phi_{2,t})$ have long-term means of $0.0187/0.4833$ and $0.0010/0.0586$, respectively. Let $\phi_{1,0} = 0.01806$ and $\phi_{2,0} = 0.00797$ be the initial values, which are chosen in such a way that (a) their quotient equals the quotient of their long-term means and (b) their sum – the initial value of the short-rate – equals $\phi_{1,0} + \phi_{2,0} = r_0 \approx \ln(1 + 0.02637)$, where 0.02637 is the 'one-week Euribor' of the 29th March in 2006 (see 'http://www.Euribor.org'). As insurers are commonly able to leverage the higher interest rate returns of long-term bonds, not the short-rate but the 10-year spot-rate is chosen here as the interest intensity $(\varphi_t)$, which according to Fischer et al. (2004) equals

$$\varphi_t = 0.22014\,\phi_{1,t} + 0.98898\,\phi_{2,t} + 0.03756\,, \qquad \forall t \geq 0\,. \tag{4.3.1}$$

To get an idea of the volatility, Figure 4.3.1 illustrates 10 simulations of $(\varphi_t)_{t \geq 0}$ for a period of 35 years.

For modeling the mortality intensity, the concept is to start from a realistic scenario and then to add some fluctuations on it. As a realistic actuarial assumption, use the

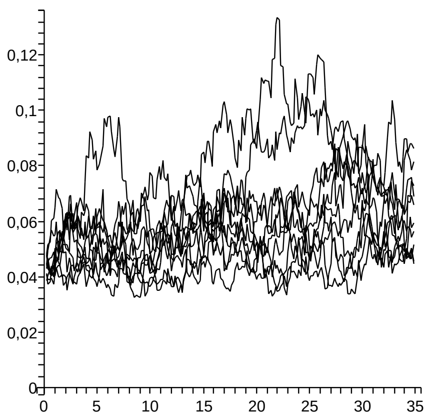


Figure 4.3.1: 10 simulated trajectories of the interest intensity

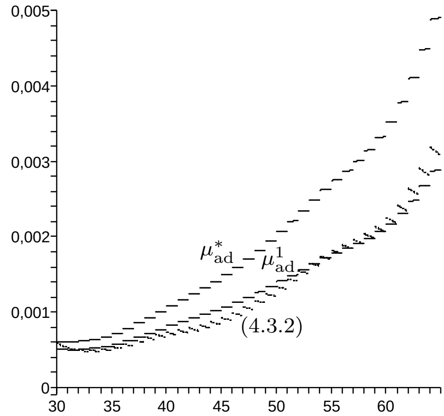


Figure 4.3.2: mortality intensity of first order $\mu^1_{\text{ad}}$ and second order $\mu^*_{\text{ad}}$, and mapping (4.3.2) (dotted)

annual mortalities of the life table DAV 2004 R (aggregated table of second order for men, reference year 2005) for transition (a,d) and the ultimate table of the select life table DAV 1997 TI (for men) for transition (i,d). The probabilities at times in between are calculated by assuming the integer truncated lifetime is independent of its remainder, where the latter is uniformly distributed. This leads to time-continuous mortality probabilities with existing mortality intensities $\mu^*_{\text{ad}}$ and $\mu^*_{\text{id}}$. Now add some random fluctuations,

$$\mu_{\text{ad},t} := \mu^*_{\text{ad}}(t) + m_{\text{ad},t}\,, \quad t \geq 0\,,$$
$$\mu_{\text{id},t} := \mu^*_{\text{id}}(t) + m_{\text{id},t}\,, \quad t \geq 0\,,$$

where $(m_{\text{ad},t})$ and $(m_{\text{id},t})$ are proper diffusion processes. In contrast to the interest intensity, modeling the mortality intensity as stochastic process has not yet been studied very well. Not only the adjusting of some parameters but the fundamental selection of a proper modeling framework is still an open question. Cairns et al. (2005, pp. 10-11) argue that long-run improvements in mortality should not be strongly mean reverting. With $(m_{\text{ad},t})$ and $(m_{\text{id},t})$ being the deviations from the anticipated mortalities $\mu^*_{\text{ad}}$ and $\mu^*_{\text{id}}$, it may be convenient to set the drift terms of $(m_{\text{ad},t})$ an $(m_{\text{id},t})$ to zero. For the diffusion term, intuition says the volatility should rise with increasing time parameter $t$. The suggestion of Dahl (2004) to use Cox-Ingersoll-Ross models means the diffusion term is proportional to the square root of the mortality. This implies a strong increase of the mortality after a medium term and a moderate volatility in the long run. The composers of the life table DAV 2004 R (Deutsche Aktuarvereinigung (2004)) assume a direct proportionality of mortality and its volatility. This leads to a more moderate volatility in mid-terms and a quite high volatility in the long run.

Korn et al. (2006, p. 408) also see Dothans model (cf. Example 4.1.3) without mean reversion as the most qualified one concerning the class of diffusion processes. Hence, let

$$\mathrm{d}m_{\mathrm{ad},t} = 0.08\left(\mu^*_{\mathrm{ad}}(30+t) + m_{\mathrm{ad},t}\right)\mathrm{d}W_{\mathrm{ad},t}\,, \quad \forall t \geq 0\,, \quad m_{\mathrm{ad},0} = 0\,,$$
$$\mathrm{d}m_{\mathrm{id},t} = 0.08\left(\mu^*_{\mathrm{id}}(30+t) + m_{\mathrm{id},t}\right)\mathrm{d}W_{\mathrm{id},t}\,, \quad \forall t \geq 0\,, \quad m_{\mathrm{id},0} = 0\,,$$

where the Wiener processes $(W_{\mathrm{ad},t})$, $(W_{\mathrm{id},t})$, $(W_{1,t})$, and $(W_{2,t})$ are stochastically independent. The constant factor 0.08 was chosen in the following way: The life tables DAV 2004 R also include an aggregated life table of *first* order. According to its authors, it represents a lower bound for the mortality in such a way that some standardized annuity produces no losses in mean at confidence level 95% (cf. Deutsche Aktuarvereinigung (2004), pp. 20-21). Now the factor 0.08 was chosen in such a manner that the mapping

$$\begin{aligned}[0,35] \ni t \;\mapsto\; & \mathbb{E}(\mu_{\mathrm{ad},t}) - 1.645\sqrt{\mathrm{Var}(\mu_{\mathrm{ad},t})} \\ &= \mu^*_{\mathrm{ad}}(t) - 1.645\left(0.08^2\int_0^t \mu^*_{\mathrm{ad}}(30+s)\, e^{0.08^2\,(t-s)}\,\mathrm{d}s\right)^{1/2}\end{aligned} \tag{4.3.2}$$

nearly fits the life table of first order mentioned above, see Figure 4.3.2. If $\mu_{\mathrm{ad},t}$ was normally distributed for each $t \geq 0$, the above mapping would yield its 5% quantiles.

For modeling the disability and the reactivation intensity process there is even less empirical evidence about the proper framework and proper parameters. Therefore, the approach of the mortality intensity processes is simply adopted here. The annual disability probabilities and reactivation probabilities are taken from the disability table DAV 1997 I and the ultimate table of the select reactivation table DAV 1997 RI, respectively. (Note that the 'independent probabilities' of the mentioned tables have to be transformed to 'dependent probabilities'. See section 3.C in Milbrodt and Helbig (1999).) Again calculate the probabilities at non-integer times by assuming the integer truncated lifetime is independent of its remainder, where the latter is uniformly distributed. Let $\mu^*_{\mathrm{ai}}$ and $\mu^*_{\mathrm{ia}}$ be the corresponding disability and reactivation intensities. Random fluctuations are added analogously to the mortality intensities,

$$\mu_{\mathrm{ai},t} := \mu^*_{\mathrm{ai}}(t) + m_{\mathrm{ai},t}\,, \quad \mathrm{d}m_{\mathrm{ai},t} = 0.08\left(\mu^*_{\mathrm{ai}}(30+t) + m_{\mathrm{ai},t}\right)\mathrm{d}W_{\mathrm{ai},t}\,, \quad m_{\mathrm{ai},0} = 0\,,$$
$$\mu_{\mathrm{ia},t} := \mu^*_{\mathrm{ia}}(t) + m_{\mathrm{ia},t}\,, \quad \mathrm{d}m_{\mathrm{ia},t} = 0.08\left(\mu^*_{\mathrm{ia}}(30+t) + m_{\mathrm{ia},t}\right)\mathrm{d}W_{\mathrm{ia},t}\,, \quad m_{\mathrm{ia},0} = 0\,,$$

where the Wiener processes $(W_{\mathrm{ai},t})$ and $(W_{\mathrm{ia},t})$ are stochastically independent of the other Wiener processes and of each other. In contrast to the construction of the interest intensity process and the mortality intensity processes, which was based on empirical evidence, the modeling of the disability and the reactivation intensity processes has to be seen with caution.

Denote by $(X_t)$ and $(\tilde{X}_t)$ the to $\mu^*$ and $\mu$ corresponding biographies of the insured.

Further, let $(\Phi_t)$ be the cumulative version of $(\varphi_t)$. Following Proposition 4.2.3, approximate the short-rate $(r_t)$ by

$$\begin{aligned}
\bar{r}_t &:= \bar{\phi}_{1,t} + \bar{\phi}_{2,t}\,, \quad \forall t \geq 0\,,\\
\mathrm{d}\bar{\phi}_{1,t} &= (0.0187 - 0.4833\,\bar{\phi}_{1,t})\,\mathrm{d}t + 0.1156\sqrt{\mathbb{E}\phi_{1,t}}\,\mathrm{d}W_{1,t}\,, \quad \forall t \geq 0\,,\\
\mathrm{d}\bar{\phi}_{2,t} &= (0.0010 - 0.0586\,\bar{\phi}_{2,t})\,\mathrm{d}t + 0.0453\sqrt{\mathbb{E}\phi_{2,t}}\,\mathrm{d}W_{2,t}\,, \quad \forall t \geq 0\,,
\end{aligned}$$

and denote by $(\bar{\varphi}_t)$ and $(\bar{\Phi}_t)$ the corresponding 10-year spot-rate and cumulative interest intensity. Again with the approach of Proposition 4.2.3, approximate the transition intensities by

$$\begin{aligned}
\bar{\mu}_{\mathrm{ad},t} &:= \mu^*_{\mathrm{ad}}(t) + \bar{m}_{\mathrm{ad},t}\,, \quad \mathrm{d}\bar{m}_{\mathrm{ad},t} = 0.08\,\mu^*_{\mathrm{ad}}(30+t)\,\mathrm{d}W_{\mathrm{ad},t}\,, \quad \bar{m}_{\mathrm{ad},0} = 0\,,\\
\bar{\mu}_{\mathrm{id},t} &:= \mu^*_{\mathrm{id}}(t) + \bar{m}_{\mathrm{id},t}\,, \quad \mathrm{d}\bar{m}_{\mathrm{id},t} = 0.08\,\mu^*_{\mathrm{id}}(30+t)\,\mathrm{d}W_{\mathrm{id},t}\,, \quad \bar{m}_{\mathrm{id},0} = 0\,,\\
\bar{\mu}_{\mathrm{ai},t} &:= \mu^*_{\mathrm{ai}}(t) + \bar{m}_{\mathrm{ai},t}\,, \quad \mathrm{d}\bar{m}_{\mathrm{ai},t} = 0.08\,\mu^*_{\mathrm{ai}}(30+t)\,\mathrm{d}W_{\mathrm{ai},t}\,, \quad \bar{m}_{\mathrm{ai},0} = 0\,,\\
\bar{\mu}_{\mathrm{ia},t} &:= \mu^*_{\mathrm{ia}}(t) + \bar{m}_{\mathrm{ia},t}\,, \quad \mathrm{d}\bar{m}_{\mathrm{ia},t} = 0.08\,\mu^*_{\mathrm{ia}}(30+t)\,\mathrm{d}W_{\mathrm{ia},t}\,, \quad \bar{m}_{\mathrm{ia},0} = 0\,.
\end{aligned}$$

Now, following the ideas of section 4.2 and supposing $\tilde{X}_0 = X_0 = \mathrm{a}$, the variance of the present value $\tilde{B}_0$ is approximately decomposable to

$$\begin{aligned}
&\mathrm{Var}\Big(\tilde{B}_0\big((X_t),(\mathbb{E}\Phi_t)\big)\Big) + \mathrm{Var}\Big(\int \nabla_0 V^{\Phi^*}_{y,s}\,\bar{\varphi}\,\mathrm{d}\lambda\Big) + \mathrm{Var}\Big(\int \nabla_0 V^{q^*_{\mathrm{ad}}}_{y,s}\,\bar{m}_{\mathrm{ad},\cdot}\,\mathrm{d}\lambda\Big)\\
&+ \mathrm{Var}\Big(\int \nabla_0 V^{q^*_{\mathrm{id}}}_{y,s}\,\bar{m}_{\mathrm{id},\cdot}\,\mathrm{d}\lambda\Big) + \mathrm{Var}\Big(\int \nabla_0 V^{q^*_{\mathrm{ai}}}_{y,s}\,\bar{m}_{\mathrm{ai},\cdot}\,\mathrm{d}\lambda\Big) + \mathrm{Var}\Big(\int \nabla_0 V^{q^*_{\mathrm{ia}}}_{y,s}\,\bar{m}_{\mathrm{ia},\cdot}\,\mathrm{d}\lambda\Big)\\
&:= \sigma^2_{(X_t)} + \sigma^2_{(\Phi_t)} + \sigma^2_{(q_{\mathrm{ad},t})} + \sigma^2_{(q_{\mathrm{id},t})} + \sigma^2_{(q_{\mathrm{ai},t})} + \sigma^2_{(q_{\mathrm{ia},t})}\,,
\end{aligned} \tag{4.3.3}$$

which uniquely correspond to (i) the unsystematic biometrical risk, (ii) the financial risk, (iii) the systematic mortality risk in state 'a', (iv) the systematic mortality risk in state 'i', (v) the systematic disability risk, and (vi) the systematic reactivation risk. (To see that, start from decomposition (1.4.4), apply (4.2.15), (4.2.22), and (4.2.20) to the first addend, and apply (4.2.1) and (4.2.27) to the second addend.) All these variances shall now be calculated for Examples 3.5.1 (a), (c), (d), and (e). With the integrals in (4.3.3) being normally distributed, the variance or its square root is a convenient risk measure. The quantity

$$\sigma^2_{((\Phi_t),(q_t))} := \sigma^2_{(\Phi_t)} + \sigma^2_{(q_{\mathrm{ad},t})} + \sigma^2_{(q_{\mathrm{id},t})} + \sigma^2_{(q_{\mathrm{ai},t})} + \sigma^2_{(q_{\mathrm{ia},t})}$$

approximates (4.2.1) at reference time $s = 0$ and corresponds to technical basis risk. The term

$$\sigma^2_{(q_t)} := \sigma^2_{(q_{\mathrm{ad},t})} + \sigma^2_{(q_{\mathrm{id},t})} + \sigma^2_{(q_{\mathrm{ai},t})} + \sigma^2_{(q_{\mathrm{ia},t})}$$

corresponds to the total systematic biometrical risk.

Applying the equivalence principle for interest intensity $(\mathbb{E}\varphi_t)$ and transition intensity $\mu^*$, and scaling the contracts to a present benefit value of 1 at time zero, the yearly premium is $0.06466$ for examples (a)&(c)&(d) and $0.06800$ for example (e). The benefit in case of survival for the pure endowment insurance is $8.4159$, the death grant for the temporary life insurance is $48.4534$, the yearly annuity for the annuity insurance is $0.6766$, and the disability annuity for the disability insurance is $1.5861$.

| SCENARIO 1 | pure endowment ins. | temp. life ins. | annuity ins. | disability ins. |
|---|---|---|---|---|
| $\sigma_{(X_t)}$ | 0.2286 | 4.5165 | 0.3210 | 2.7902 |
| $\sigma_{(\Phi_t)}$ | 0.2485 | 0.0601 | 0.3312 | 0.1177 |
| $\sigma_{(q_t)}$ | 0.0101 | 0.1384 | 0.0232 | 0.1356 |
| $\sigma_{(q_{\mathrm{ad},t})}$ | 0.0101 | 0.1384 | 0.0232 | 0.0014 |
| $\sigma_{(q_{\mathrm{id},t})}$ | – | – | – | 0.0150 |
| $\sigma_{(q_{\mathrm{ai},t})}$ | – | – | – | 0.1097 |
| $\sigma_{(q_{\mathrm{ia},t})}$ | – | – | – | 0.0784 |
| $\sigma_{((\Phi_t),(q_t))}$ | 0.2487 | 0.1509 | 0.3320 | 0.1796 |
| $\sigma_{simulated}$ | 0.2092 | 0.1534 | 0.2762 | 0.1717 |

As these standard deviations are just approximations, the total technical basis risk $\sigma_{((\Phi_t),(q_t))}$ is compared with a simulation of the standard deviation of (4.2.1) at time zero, here denoted as $\sigma_{simulated}$. Therefore, the stochastic processes $(\phi_{1,t})$, $(\phi_{2,t})$, $(\mu_{\mathrm{ad},t})$, $(\mu_{\mathrm{id},t})$, $(\mu_{\mathrm{ai},t})$, and $(\mu_{\mathrm{ia},t})$ were simulated $10^4$ times using the Milstein Scheme (cf. Kloeden and Platen (1992), section 10.3) with equidistant time steps of step size $1/24$. Looking at the results, the approximation errors seem to be acceptable. For the temporary life insurance and the disability insurance, the approximation is even pretty good.

One could argue that in reality the mortality fluctuations in state 'a' and in state 'i' are not really independent. An alternative is to let the Wiener processes $(W_{\mathrm{ad},t})$ and $(W_{\mathrm{id},t})$ be equal, that is, the relative fluctuations of $(\mu_{\mathrm{ad},t})$ and $(\mu_{\mathrm{id},t})$ are similar. The joint standard deviation is then at about $0.0164$ compared to $(\sigma^2_{(q_{\mathrm{ad},t})} + \sigma^2_{(q_{\mathrm{id},t})})^{1/2} = 0.0151$, which means that it does not really make a big difference.

The above table disproves the myth that financial risk (interest rate risk) is generally much more important than systematic biometrical risks and that the latter are sufficiently covered by choosing a generous safety loading for the former. For the disability insurance, the systematic disability risk has the same order of magnitude as the financial risk. For the temporary life insurance, the systematic mortality risk is here even greater than the financial risk. For the latter contract type, Norberg (1999, p. 389) comes to a similar conclusion.

For the pure endowment insurance, the annuity insurance, and the disability insurance, the systematic mortality risk seems to be negligible. At least for the annuity insurance this is astonishing, since the so-called 'longevity risk', which is included in the systematic mortality risk, attracts more and more the attention of the practitioners, and deservedly so. The unforeseen increase of life expectancy especially in recent years had and still has a great impact on the life insurance industry, e.g., the subsequent reserving in Germany due to the new life table DAV 2004 R. Why is the model not reflecting that experience?

Not surprisingly, the above results heavily depend on the choice of the stochastic intensity processes. Therefore, an alternative scenario is studied:

**Scenario 2:** In Scenario 1, the interest intensity is mainly driven by the process $(\phi_{2,t})$ (cf. (4.3.1)), whose relative small mean reversion factor lets the interest intensity have great values with substantial probability. Especially in case of the annuity insurance with a contract period of 90 years, an increase of the interest rate towards extraordinary 20% is quite likely. Hence one could argue that the interest rate risks $\sigma_{(\Phi_t)}$ of Scenario 1 are immoderate. Alternatively, model now the interest intensity $(\varphi_t)$ by a CIR-1-model following Fischer et al. (2004, p. 380, again use the 10-year spot-rate),

$$\varphi_t = 0.62362\,\phi_t + 0.03547\,, \quad \forall t \geq 0\,,$$
$$\mathrm{d}\phi_{0,t} = (0.0097 - 0.1780\,\phi_{0,t})\,\mathrm{d}t + 0.0461\sqrt{\phi_{0,t}}\,\mathrm{d}W_{0,t}\,, \quad \forall t \geq 0\,.$$

The mean reversion factor of 0.1780 keeps the interest intensity process in a narrower bandwidth (cf. Figure 4.3.3). Trying to reconstruct the historical situation associated with the introduction of the life table DAV 2004 R, use as initial value the 10-year spot-rate of the Deutsche Bundesbank from December 2001, $\varphi_0 = 0.0512$ (cf. Fischer et al. (2003), p. 202). One may criticize in this new approach that Fischer et al. (2003) themselves reject their CIR-1-model: In their opinion, the yield curve is too plain and the dynamics of the short-rate is not realistic. Nevertheless, the CIR-1-model is used here, as their criticism does not affect the dynamics of the 10-year spot-rate.

The life table DAV 2004 R suggests a strong future decline of the mortality (though using the second order table). This leads to a comparatively small systematic mortality risk in state active $\sigma_{(q_{\mathrm{ad},t})}$. Alternatively, for the construction of $\mu^*_{\mathrm{ad}}$ use the life table 2002/2004 of the 'Statistisches Bundesamt Deutschland', which represents the actual mortality between 2002 and 2004 in Germany. Figure 4.3.4 shows a comparison of the mortality intensities of Scenario 1 and 2.

Applying the equivalence principle as in Scenario 1 under the new assumptions, and scaling the contracts to a present benefit value of 1 at time zero, leads to a yearly premium of 0.07019 for examples (a)&(c)&(d) and 0.07332 for example (e). The benefit in case of survival for the pure endowment insurance is $12.3024$, the death grant for the temporary life insurance is $26.2620$, the yearly annuity for the annuity insurance is $1.3005$, and the disability annuity for the disability insurance is $1.9076$.

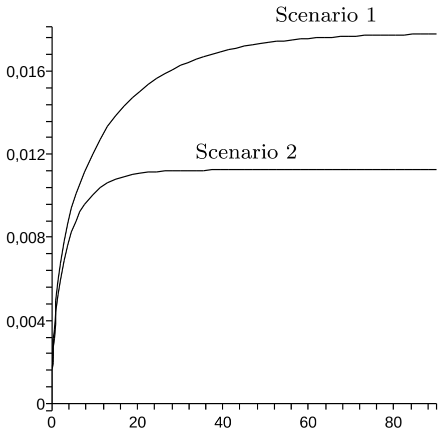


Figure 4.3.3: standard deviation function of $(\varphi_t)$ for a period of 90 years for Scenario 1 and Scenario 2

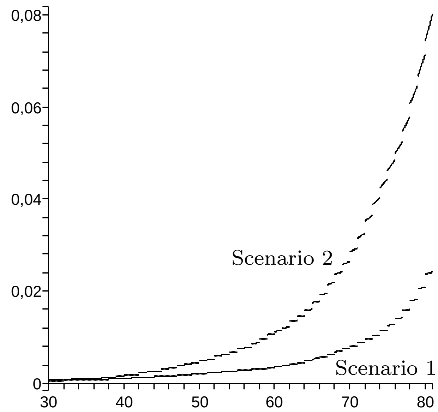


Figure 4.3.4: mortality intensity $\mu^*_{\mathrm{ad}}$ between age 30 and 80 for Scenario 1 and Scenario 2

| SCENARIO 2 | pure endowment ins. | temp. life ins. | annuity ins. | disability ins. |
|---|---|---|---|---|
| $\sigma_{(X_t)}$ | 0.4026 | 2.9342 | 0.5387 | 2.8765 |
| $\sigma_{(\Phi_t)}$ | 0.1462 | 0.0538 | 0.1690 | 0.0713 |
| $\sigma_{(q_t)}$ | 0.0230 | 0.1106 | 0.0487 | 0.1357 |
| $\sigma_{(q_{\mathrm{ad},t})}$ | 0.0230 | 0.1106 | 0.0487 | 0.0025 |
| $\sigma_{(q_{\mathrm{id},t})}$ | – | – | – | 0.0144 |
| $\sigma_{(q_{\mathrm{ai},t})}$ | – | – | – | 0.1104 |
| $\sigma_{(q_{\mathrm{ia},t})}$ | – | – | – | 0.0775 |
| $\sigma_{((\Phi_t),(q_t))}$ | 0.1480 | 0.1230 | 0.1759 | 0.1533 |
| $\sigma_{simulated}$ | 0.1457 | 0.1265 | 0.1749 | 0.1549 |

For all four examples a comparison of $\sigma_{((\Phi_t),(q_t))}$ and $\sigma_{simulated}$ shows the approximation approach works pretty well here. If one assumes that the Wiener processes $(W_{\mathrm{ad},t})$ and $(W_{\mathrm{id},t})$ are equal instead of independent, the joint standard deviation for mortality is 0.0146 compared to $(\sigma^2_{(q_{\mathrm{ad},t})} + \sigma^2_{(q_{\mathrm{id},t})})^{1/2} = 0.0168$.

The systematic disability risk for the disability insurance and the systematic mortality risk for the temporary life insurance are not only of great importance as in Scenario 1, but clearly exceed now the financial risk. For the pure endowment insurance, the annuity insurance, and the disability insurance, the systematic mortality risk still plays a minor role compared to the financial risk. But in contrast to Scenario 1 its importance increased:

- For the annuity insurance the ratio $\sigma_{(q_{\mathrm{ad},t})}/\sigma_{(\Phi_t)}$ rose from about 7% to about

29%. That means although the volatility of the financial markets still contributes a greater risk than the uncertainty of future life expectancies, the systematic mortality risk is definitely not negligible.

- For the pure endowment insurance the ratio $\sigma_{(q_{\mathrm{ad},t})}/\sigma_{(\Phi_t)}$ changed from about 4% to about 16%.

Interestingly, in Scenario 2 the summarized risks $\sigma_{((\Phi_t),(q_t))}$, which correspond to the uncertainty of the technical basis as a whole, have all about the same size. That means if one calculates a proper safety loading by using a standard deviation approach, the risk premiums for the technical basis risks would all be of about the same size.

**Comparison and combination of different insurance contract types:** Studying the tables of Scenarios 1 and 2 suggests to divide the four examples into two groups: (i) the pure endowment insurance and the annuity insurance, and (ii) the temporary life insurance and the disability insurance. The characteristics within these groups are quite similar: Group (i) is more vulnerable to changes of the interest rate than group (ii), whereas group (ii) needs a much greater portfolio than group (i) to diversify its unsystematic biometrical risk.

For both scenarios the annuity has clearly the greatest 'technical basis risk' $\sigma_{((\Phi_t),(q_t))}$. This is not surprising: With the contract not terminating until death, it has mostly a much longer contract period than the other examples.

Among other things, section 3.5 studied the effects combinations of different insurance contract types have on the sensitivities. Now the combinations of Example 3.5.4 shall be reappraised by calculating their decomposed (approximative) standard deviations analogously to the above. Again the premiums are calculated by using the equivalence principle for the technical basis $((\mathbb{E}\varphi_t), \mu^*)$, and the contracts are scaled to a present premium value of 1 at time zero. The (approximative) standard deviations for Example 3.5.4(f1) are:

| pure endowment & temp. life ins. (f1) | $\sigma_{(\Phi_t)}$ | $\sigma_{(q_{\mathrm{ad},t})}$ | $\sigma_{((\Phi_t),(q_t))}$ |
|---|---|---|---|
| Scenario 1 | 0.2201 | 0.0129 | 0.2205 |
| Scenario 2 | 0.1153 | 0.0219 | 0.1174 |

In section 3.5, this kind of combination was motivated by some cancelation effect of the corresponding gradient vectors with respect to mortality (cf. Figures 3.5.13). In fact, in case of Scenario 2 the systematic mortality risk $\sigma_{(q_{\mathrm{ad},t})}$ is lower than that for single contracts. The same holds for the aggregated 'technical basis risk' $\sigma_{((\Phi_t),(q_t))}$; for combination (f1) it is around 0.1174, for the pure endowment insurance and the temporary life insurance it is around 0.1480 and 0.1230. However, in case of Scenario 1 the cancelation effect is not strong enough to let the combined contract have a better risk situation than each of the two single contracts.

The magnitude of the cancelation effect heavily depends on the weight the two basic contracts have. Figure 4.3.5 shows the (approximative) technical basis risks $\sigma_{((\Phi_t),(q_t))}$ of combined pure endowment and temporary life insurances subject to their ratios between survival and death benefit. The abscissa is logarithmic, a x-coordinate of $k$ means the ratio of survival and death benefit is $2^k$. The optimal ratio with minimal (approximative) technical basis risk greatly depends on the chosen scenario. Here the best ratios are around $2^{-5}$ for Scenario 1 and $2^{-2}$ for Scenario 2.

As seen in Figure 3.5.14, Example 3.5.4(g) features also some cancelation effect. Figure 4.3.6 shows a plot analogous to Figure 4.3.5.

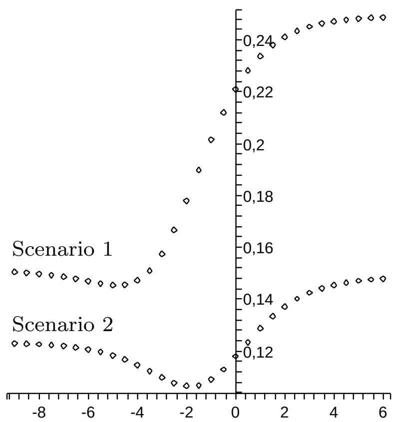


Figure 4.3.5: technical basis risk $\sigma_{((\Phi_t),(q_t))}$ of combined pure endowment & temporary life ins. subject to the ratio of survival and death benefit (logarithmic scale!)

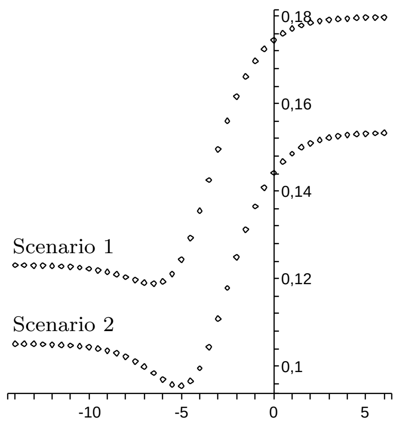


Figure 4.3.6: technical basis risk $\sigma_{((\Phi_t),(q_t))}$ of combined disability & temporary life ins. subject to the ratio of yearly disability annuity and death benefit (logarithmic scale!)

While combination (h) does not feature a cancelation effect, section 3.5 gave another motivation for its existence: The tremendous sensitivity of the disability insurance to changes of the disability probabilities calls for a huge safety loading. A combination with an annuity insurance lowers this sensitivity relative to the overall present premium value (cf. Figure 3.5.16). Figure 4.3.7 shows the technical basis risks for several ratios of disability annuity and pension annuity. For Scenario 1, the optimal combination of disability and annuity insurance is a pure disability insurance. The approach of reducing the relative size of a proper risk loading for the disability insurance by combining it with the annuity insurance does not work here! Interestingly, this is not always the case. In Scenario 2, a risk reduction takes place due to a general property of variances: Since the variance is a risk averse risk measure, a sum of two independent and medium risks produces a lower variance than a comparable independent sum of a huge and a small risk. Although no cancelation effect takes place, a combination of disability and annuity insurance can produce a technical basis risk that is lower than the technical basis risks of each single contract.

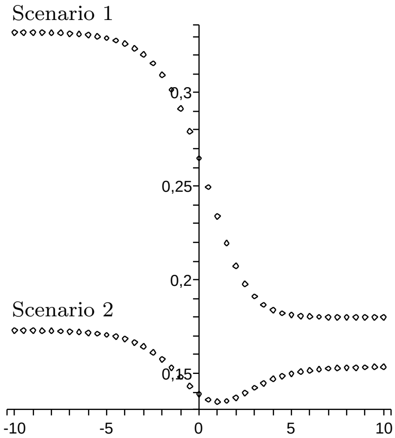


Figure 4.3.7: technical basis risk $\sigma_{((\Phi_t),(q_t))}$ of a combined disability & annuity ins. subject to the ratio of disability annuity and pension annuity (logarithmic scale!)

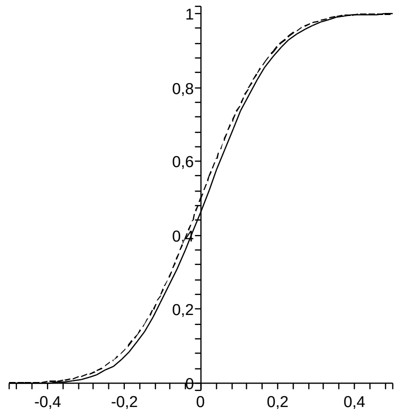


Figure 4.3.8: approximated (dotted) and simulated (solid) distribution function of (4.2.1) for the pure endowment insurance under Scenario 2

**Outlook:** Until now, only variances or standard deviations of the components of decomposition (4.2.5) have been calculated approximately. In fact, Corollary 4.2.7 allows for more: the moments of higher order and the quantiles on the interval $(0, 1)$ converge as well. Given the approximation error is small enough, this enables one to use also other risk measures than the variance, e.g., the 'value-at-risk'.

Unfortunately, the approximation of the quantiles is not satisfying for all of the above examples and scenarios. Positive examples are the two-state contracts in case of Scenario 2, see Figures 4.3.8 to 4.3.10. The approximation in Figures 4.3.8 and 4.3.10 are despite a small bias amazingly well. This enables, for example, to calculate premiums on the safe side by using a percentile principle, which offers a much better risk management than calculating premiums with the equivalence principle for some technical basis of first order.

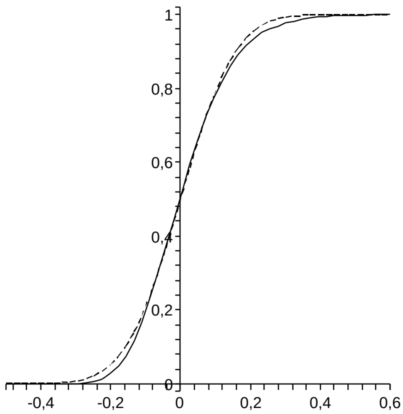


Figure 4.3.9: approximated (dotted) and simulated (solid) distribution function of (4.2.1) for the temporary life insurance under Scenario 2

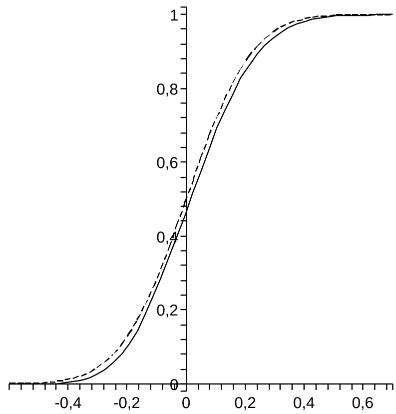


Figure 4.3.10: approximated (dotted) and simulated (solid) distribution function of (4.2.1) for the annuity insurance under Scenario 2

# References


Bauer, H. (1992): Maß- und Integrationstheorie, 2. Auflage. Walter de Gruyter, Berlin.

Bickel, P.J., C.A.J. Klaassen, Y. Ritov, J.A. Wellner (1998): Efficient and Adaptive Estimation for Semiparametric Models. Springer, New York.

Bowers, N.L., H.U. Gerber, J.C. Hickman, D.A. Jones, and C.J. Nesbitt (1997): Actuarial Mathematics, 2nd edition. The Society of Actuaries, Schaumburg, Illinois.

Brigo, D., F. Mercurio (2001): Interest Rate Models – Theory and Practice. Springer, Berlin.

Bruno, M.G., E. Camerini, A. Tomassetti (2000): Financial and demographic risks of a portfolio of life insurance policies with stochastic interest rates: Evaluation methods and applications. North American Actuarial Journal No. 4, pp. 44-55.

Bühlmann, H. (1992): Stochastic discounting. Insurance: Mathematics and Economics No. 11(2), pp. 113-127.

Bühlmann, H. (1995): Life insurance with Stochastic Interest Rates. In: Ottaviani, G. (Ed.) – Financial Risk in Insurance, Springer, Berlin.

Cairns, A.J.G., D. Blake, K. Dowd (2005): Pricing Death: Frameworks for the Valuation and Securitization of Mortality Risk. To appear in: ASTIN Bulletin.

Courant, R. and D. Hilbert (1968): Methoden der Mathematischen Physik I, 3. Auflage. Springer, Berlin.

Dahl, M.H. (2004): Stochastic mortality in life insurance: market reserves and mortality-linked insurance contracts. Insurance: Mathematics and Economics No. 35(1), pp. 113-136.

Dahl, M., T. Møller (2005): Valuation and Hedging of Life Insurance Liabilities with Systematic Mortality Risk. Working paper No. 203, Laboratory of Actuarial Mathematics, Copenhagen.

Deutsche Aktuarvereinigung (2004): Herleitung der DAV Sterbetafel für Rentenversicherungen – Stand 21.05.2004.

De Schepper, A., F. De Vylder, M. Goovaerts, R. Kaas (1992): Interest randomness in annuities certain. Insurance: Mathematics and Economics No. 11, pp. 271-281.

Dienst, H.-R. (1995): Zur aktuariellen Problematik der Invaliditätsversicherung. Verlag Versicherungswirtschaft, Karlsruhe.

Dunford, N., J.T. Schwartz (1988): Linear Operators Part I: General Theory. Wiley, New York.

Elliot, R.J. (1982): Stochastic Calculus and Applications. Springer, New York.

Farny, D. (1995): Versicherungsbetriebslehre, 2. Ausgabe. Verlag Versicherungswirtschaft, Karlsruhe.

Fernholz, L.T. (1983): Von Mises calculus for statistical functionals. Lecture Notes in Statistics No. 19. Springer, New York.

Fischer, T. (2004): Valuation and risk management in life insurance. Dissertation, Department of Mathematics, TU Darmstadt.

Fischer, T., A. May, B. Walther (2003): Anpassung eines CIR-1-Modells zur Simulation der Zinsstrukturkurve. Blätter der DGVFM No. 2, pp. 193-206.

Fischer, T., A. May, B. Walther (2004): Anpassung eines CIR-$k$-Modells zur Simulation der Zinsstrukturkurve. Blätter der DGVFM No. 3, pp. 369-387.

Flett, T.M. (1980): Differential Analysis. Cambridge University Press, Cambridge.

Frees, E. (1990): Stochastic life contingencies with solvency considerations. Transactions of the Society of Actuaries No. 24, pp. 91-130.

Gill, R. (1994): Lectures on Survival Analysis. In: Bakry, D., R. Gill and S. Molchanov – Lectures on Probability Theory. Lecture Notes in Mathematics No. 1581. Springer, Berlin.

Gill, R. and S. Johansen (1990): A survey of product-integration with a view toward application in survival analysis. Annals of Statistics No. 18, pp. 1501-1555.

Golomb, M. (1934): Zur Theorie der nichtlinearen Integralgleichungen, Integralgleichungssystemen und allgemeinen Funktionalgleichungen. Mathematische Zeitschrift No. 39, pp. 45-72. Springer, Berlin.

Hampel, F.R. (1986): Robust Statistics – The Approach Based on Influence Functions. Wiley, New York.

Helwich, M. (2003): Über den Verlgleich des Zinsrisikos mit dem biometrischen Risiko bei Lebensversicherungen. Diplomarbeit, Fachbereich Mathematik, Universität Rostock.

Helwich, M. (2005): Further ideas concerning the interaction between insurance and investment risk. Preprint, Department of Mathematics, University of Rostock.

Hewitt, E. and K. Stromberg (1965): Real and Abstract Analysis. Springer, Berlin.

Hoem, J.M. (1988): The Versatility of the Markov Chain as a Tool in the Mathematics of Life Insurance. Transactions of the 23rd International Congress of Actuaries R, pp. 171-202.

Huber, P.J. (1981): Robust Statistics. Wiley series in probability and mathematical statistics, New York.

Huber, P.J. (1996): Robust Statistical Procedures, 2nd edition. Society for Industrial and Applied Mathematics, Philadelphia.

International Accounting Standards Board (2004): International Financial Reporting Standard No. 4. Available at 'http://www.iasb.org'.

Kaas, R., M. Goovaerts, J. Dhaene, M. Denuit (2003): Modern actuarial risk theory. Kluwer Academic Publishers, Boston.

Kalashnikov, V. and R. Norberg (2003): On the Sensitivity of Premiums and Reserves to Changes in Valuation Elements. Scandinavian Actuarial Journal No. 3, pp. 238-256.

Kallenberg, O. (1997): Foundations of Modern Probability. Springer, New York.

Khalaf-Allah, M., S. Haberman, R. Verrall (2006): Measuring the effect of mortality improvements on the cost of annuities. Insurance: Mathematics and Economics No. 39, pp. 231-249.

Kloeden, P.E. and E. Platen (1992): Numerical Solution of Stochastic Differential Equations. Springer, Berlin.

Korn, R., K. Natcheva, J. Zipperer (2006): Langlebigkeitsbonds – Bewertung, Modellierung und Aspekte für deutsche Daten. Blätter der DGVFM No. XXVII, pp. 397-418.

Lidstone, G.J. (1905): Changes in pure premium values consequent upon variations in the rate of interest or rate of mortality. Journal of the Institute of Actuaries No. 39, pp. 209-252.

Linnemann, P. (1993): On the application of Thieles differential equation in life insurance. Insurance: Mathematics and Economics No. 13, pp. 63-74.

Marceau, E., P. Gaillardetz (1999): On life insurance reserves in a stochastic mortality and interest rates environment. Inusrance: Mathematics and Economics No. 25, pp. 261-280.

Milbrodt, H. and M. Helbig (1999): Mathematische Methoden der Personenversicherung. Walter de Gruyter, Berlin.

Müller, A., D. Stoyan (2002): Comparison Methods for Stochastic Models and Risks. Wiley Series in Probability and Statistics, Chichester.

Norberg, R. (1985): Lidstone in the continuous case. Scandinavian Actuarial Journal, pp. 27-32.

Norberg, R. (1999): A theory of bonus in life insurance. Finance and Stochastics No. 3, pp. 373-390.

Norberg, R. (2001): On bonus and bonus prognoses. Scandinavian Actuarial Journal, pp. 126-147.

Norberg, R. and C.M. Møller (1996): Thieles Differential Equation with Stochastic Interest of Diffusion Type. Scandinavian Actuarial Journal No. 1, pp. 37-49.

Olivieri, A. (2001): Uncertainty in mortality projections: an actuarial perspective. Insurance: Mathematics and Economics No. 29, pp. 231-245.

Parker, G. (1994a): Limiting distribution of the present value of a portfolio. ASTIN Bulletin No. 24(1), pp. 47-60.

Parker, G. (1994b): Two stochastic approaches for discounting actuarial functions. ASTIN Bulletin, No. 24(2), pp. 167-181.

Parker, G. (1997): Stochastic analysis of the interaction between investment and insurance risk. North American Actuarial Journal No. 2, pp. 55-84.

Perry, D., W. Stadje (2001): Function space integration for annuities. Insurance: Mathematics and Economics No. 29, pp. 73-82.

Perry, D., W. Stadje, R. Yosef (2003): Annuities with controlled random interest rates. Insurance: Mathematics and Economics No. 32, pp. 245-253.

Persson, S.-A. (1998): Stochastic Interest Rate in Life Insurance: The Principle of Equivalence Revisited. Scandinavian Actuarial Journal, pp. 97-112.

Pfanzagl, J. and W. Wefelmeyer (1982): Contributions to a general asymptotic statistical theory. Lecture Notes in Statistics No. 13. Springer, New York.

Ramlau-Hansen, H. (1988): The emergence of profit in life insurance. Insurance: Mathematics and Economics No. 7, pp. 225-236.

Riesz, F. and B. Sz.-Nagy (1968): Vorlesungen über Funktionalanalysis, 2. Auflage. Deutscher Verlag der Wissenschaften, Berlin.

Saltelli, A. (ed.), K. Chan (ed.), E.M. Scott (ed.) (2000): Sensitivity analysis. Wiley Series in Probability and Statistics, Chichester.

Statistisches Bundesamt Deutschland: Sterbetafeln für Deutschland 2002/2004. Available at 'http://www.destatis.de'.

Witting, H. (1985): Mathematische Statistik I. Teubner, Stuttgart.

Witting, H., U. Müller-Funk (1995): Mathematische Statistik II. Teubner, Stuttgart.

# A Appendix

This section presents definitions and propositions that are largely common folklore, but for the most part not found in the literature in the form they are needed here.

## A.1 Representation of the present value

**Proposition A.1.1.** *The present value* (1.2.5) *is representable as a* $(\mathcal{X} \times \mathcal{F}, \mathfrak{X} \otimes \mathfrak{F})$*-*$(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$*-measurable mapping of the biography of the insured* $(X_t)$ *and the cumulative interest intensity* $\Phi$*,*

$$B_s = b_s\big((X_t), \Phi\big) .$$

For the definitions of $\mathcal{X}$ and $\mathcal{F}$, see section 1.3.

*Proof.* (i) Because of $\mathfrak{F} = \mathcal{F} \cap (\mathfrak{B}(\mathbb{R}))^{[0,\infty)}$ the mapping

$$F \mapsto K_F(t) = \prod_{(0,t]} (1 + \mathrm{d}F) = \lim_{n\to\infty} \prod_{\mathcal{T}_n} \Big(1 + F(t_i) - F(t_{i-1})\Big)$$

is $(\mathcal{F}, \mathfrak{F})$-$(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$-measurable for each $t \in [0, \infty)$, where $(\mathcal{T}_n)$ is a sequence of interval decompositions $s \le t_1 < .. < t_n \le t$ satisfying $\lim_{n\to\infty} \max_{\mathcal{T}_n} |t_{i+1} - t_i| = 0$. According to Theorem 2.32 in Elliot (1982), it is even $(\mathbb{R}\times\mathcal{F}, \mathfrak{B}(\mathbb{R})\otimes\mathfrak{F})$-$(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$-measurable. The same holds true for $(t, F) \mapsto 1/K_F(t)$, as $K_F$ is strictly positive (cf. Proposition A.4.2).

(ii) Applying Theorem 2.32 in Elliot (1982) again, the right-continuous mapping $t \mapsto \mathbf{1}_{X_t=z}(t)$ is $(\mathbb{R} \times \mathcal{X}, \mathfrak{B}(\mathbb{R}) \otimes \mathfrak{X})$-$(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$-measurable.

(iii) Using (i) and (ii) and Tonellis Theorem, the mapping (cf. (6.15.2) in Milbrodt and Helbig (1999))

$$(X, F) \mapsto sb_s\big(X, F\big) = \sum_{z\in\mathcal{S}} \int_{[s,\infty)} \mathbf{1}_{X_\tau=z}(\tau) \frac{1}{K_F(\tau)}\, F_z(\mathrm{d}\tau)$$

is $(\mathcal{X} \times \mathcal{F}, \mathfrak{X} \otimes \mathfrak{F})$-$(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$-measurable.

(iv) Since the mappings $D_{z\zeta}$ are of bounded variation on compacts, they are $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$-$(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$-measurable. As $|J| < \infty$, the mapping $((z, \zeta), t) \mapsto D_{z\zeta}(t)$ is even $(J \times \mathbb{R}, 2^J \otimes \mathfrak{B}(\mathbb{R}))$-$(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$-measurable.

Following Milbrodt and Helbig (1999, see (4.8.4)), define

$$t_0(X) := 0\,, \qquad t_m(X) := \min\left\{t > t_{m-1}(X) \,|\, X_t \neq X_{t_{m-1}(X)}\right\}, \quad m \in \mathbb{N}\,.$$

According to Theorem 4.12 in Milbrodt and Helbig (1999),

$$t_m : \mathcal{X} \ni X \mapsto t_m(x) \in [0,\infty] \qquad \text{and} \qquad \mathcal{X} \ni X \mapsto X_{t_m(X)} \in \mathcal{S}$$

are $(\mathcal{X},\mathfrak{X})$-$(\mathbb{R},\mathfrak{B}([0,\infty]))$- and $(\mathcal{X},\mathfrak{X})$-$(\mathcal{S},2^{\mathcal{S}})$-measurable, respectively. Thus, $X \mapsto D_{X_{t_{m-1}(X)}X_{t_m(X)}}(t_m(X))$ is $(\mathcal{X},\mathfrak{X})$-$(\mathbb{R},\mathfrak{B}(\mathbb{R}))$-measurable.

Using (i) and the measurability of $DT$ and of the $t_m$, the mapping

$$(X,F) \mapsto \mathbf{1}_{s<t_m(X)<\infty}\,\frac{1}{K_F \circ DT}(t_m(X))$$

is $(\mathcal{X}\times\mathcal{F},\mathfrak{X}\otimes\mathfrak{F})$-$(\mathbb{R},\mathfrak{B}(\mathbb{R}))$-measurable. Thus,

$$(X,F) \mapsto db_s(X,F) = \sum_{m=1}^{\infty} \mathbf{1}_{s<t_m(X)<\infty}\,\frac{D_{X_{t_{m-1}(X)}X_{t_m(X)}}}{K_F \circ DT}(t_m(X))$$

(cf. (6.15.1) in Milbrodt and Helbig (1999)) is $(\mathcal{X}\times\mathcal{F},\mathfrak{X}\otimes\mathfrak{F})$-$(\mathbb{R},\mathfrak{B}(\mathbb{R}))$-measurable as well.

Combining (iii) and (iv), the mapping (cf. (1.2.5))

$$b_s : (X,F) \mapsto sb_s(X,F) + db_s(X,F)$$

is $(\mathcal{X}\times\mathcal{F},\mathfrak{X}\otimes\mathfrak{F})$-$(\mathbb{R},\mathfrak{B}(\mathbb{R}))$-measurable for any $s \geq 0$. By construction $b_s((X_t),\Phi) = B_s$. □

## A.2 Functions of bounded variation

With $\mathcal{T}$ being any finite partition $t_1 < t_2 < \cdots < t_n$ of the interval $I \subset \mathbb{R}$, let

$$\mathcal{V}_I(x) := \sup_{\mathcal{T}\subset I} \sum_{\mathcal{T}} \left|x(t_{i+1}) - x(t_i)\right| \tag{A.2.1}$$

be the *variation of* $x : \mathbb{R} \to \mathbb{R}$ *on* $I$. The functional $\|\cdot\|_{BV} := \mathcal{V}_{\mathbb{R}}(\cdot)$ is denoted as *total variation of* $x$.

Denote by $BVC$ the linear space of functions on $\mathbb{R}$ with finite total variation on compacts and support in $[0,\infty)$. The subset $BV := \{x \in BVC \,|\, \|x\|_{BV} < \infty\}$ is a normed space with the total variation as its norm. The super- and subscripts '$\leftarrow$', '$\rightarrow$', '$b$', and '+' stand for the additional restrictions 'right-continuous', 'left-continuous', 'bounded', and 'monotonic nondecreasing', respectively.

**Theorem A.2.1** (Jordan-Hahn decomposition)**.** *Each element $x$ of $BVC_{\leftarrow}$ or $BVC_{\rightarrow}$ is decomposable to $x = x_+ - x_-$, where $x_+$ and $x_-$ are elements of $BVC^{+}_{\leftarrow}$ or $BVC^{+}_{\rightarrow}$, respectively. If additionally $x \in BV_{\leftarrow}$ or $x \in BV_{\rightarrow}$, then the functions $x_+$ and $x_-$ are elements of $BV^{+}_{\leftarrow}$ or $BV^{+}_{\rightarrow}$, respectively.*

*Proof.* The definitions

$$x_+ := \frac{1}{2}\left(\mathcal{V}_{(-\infty,\cdot\,]}(x) + x\right), \quad x_- := \frac{1}{2}\left(\mathcal{V}_{(-\infty,\cdot\,]}(x) - x\right) \tag{A.2.2}$$

provide a suitable decomposition $x = x_+ - x_-$. Firstly, the functions $x_+$ and $x_-$ are nondecreasing and elements of $BVC$ (cf. Riesz and Sz.-Nagy (1968), pp. 8, 9), secondly, they remain right-(left-)continuous (cf. Riesz and Sz.-Nagy (1968), p. 14), and third $x \in BV$ implies $x_+, x_- \in BV$. □

**Proposition A.2.2.** *Each element $x$ of $BV$ is decomposable into a sum of a right-continuous function $x_\leftarrow$ and a left-continuous function $x_\rightarrow$ with finite total variations.*

*Proof.* According to Riesz and Sz.-Nagy (1968, pp. 11-13), $x$ decomposes into a sum of a continuous function $x_c$ and a step function $x_s$ of the form

$$x_s = \sum_n a_n \mathbf{1}_{[t_n,\infty)} + \sum_n b_n \mathbf{1}_{(s_n,\infty)}$$

with $a_n, b_n \in \mathbb{R}$ as the step heights and $t_n, s_n \in \mathbb{R}$ as the countable locations of jumps. Then, for example,

$$x_\leftarrow := \sum_n a_n \mathbf{1}_{[t_n,\infty)} + \frac{1}{2} x_c \in BV_\leftarrow, \qquad x_\rightarrow := \sum_n b_n \mathbf{1}_{(s_n,\infty)} + \frac{1}{2} x_c \in BV_\rightarrow$$

is a proper decomposition. □

With $\mu_x$ being the Borel-measure corresponding to $x \in BVC_\leftarrow$ (or $x \in BVC_\rightarrow$), the so-called *Lebesgue-Stieltjes* integral is defined by

$$\int y \,\mathrm{d}x := \int y \,\mathrm{d}\mu_x$$

for any Borel-measurable und $\mu_x$-integrable function $y$. Applying Proposition A.2.2 leads to a generalization of Lebesgue-Stieltjes integration: Provided the integrals exist, define for $x \in BV$

$$\int y \,\mathrm{d}x := \int y \,\mathrm{d}x_\leftarrow + \int y \,\mathrm{d}x_\rightarrow$$

for any Borel-measurable function $y$, where $x = x_\leftarrow + x_\rightarrow$ represents a decomposition according to Proposition A.2.2. The integral is well defined as it is independent of the chosen decomposition.

**Proposition A.2.3** (partial integration)**.** *For any $x, y \in BV$, with $y = y_\leftarrow + y_\rightarrow$ being a decomposition according to Proposition* A.2.2*,*

$$\begin{aligned}\int_{[a,b]} y \,\mathrm{d}x = \quad & y(b+0)\,x(b+0) - \int_{[a,b]} x(\cdot - 0)\,\mathrm{d}y_\leftarrow \\ & - \int_{[a,b]} x(\cdot + 0)\,\mathrm{d}y_\rightarrow - y(a-0)\,x(a-0)\end{aligned} \tag{A.2.3}$$

*for all $-\infty < a < b < \infty$.*

*Proof.* Let $\mu_x$, $\mu_{y_\leftarrow}$, and $\mu_{y_\rightarrow}$ be the Borel measures corresponding to $x$, $y_\leftarrow$, and $y_\rightarrow$. Proposition A.2.2 allows for writing

$$\int_{[a,b]} y \,\mathrm{d}x = \int_{[a,b]} y_\leftarrow \,\mathrm{d}x + \int_{[a,b]} y_\rightarrow \,\mathrm{d}x \,.$$

Using Fubinis Theorem, the first addend equals

$$\begin{aligned}\int_{[a,b]} y_\leftarrow \,\mathrm{d}x &= \int \mathbf{1}_{[a,b]}(\tau)\, \mu_{y_\leftarrow}((-\infty,\tau])\, \mu_x(\mathrm{d}\tau) \\ &= \iint \mathbf{1}_{[a,b]}(\tau)\, \mathbf{1}_{(-\infty,\tau]}(t)\, \mu_{y_\leftarrow}(\mathrm{d}t) \\ &= \iint \mathbf{1}_{[a,b]\cap[t,\infty)}(\tau)\, \mu_x(\mathrm{d}\tau)\, \mu_{y_\leftarrow}(\mathrm{d}t) \\ &= \int_{(-\infty,a)} \mu_x([a,b])\, \mu_{y_\leftarrow}(\mathrm{d}t) + \int_{[a,b]} \big(\mu_x((-\infty,b]) - \mu_x((-\infty,t))\big)\, \mu_{y_\leftarrow}(\mathrm{d}t) \\ &= x(b+0)\, y_\leftarrow(b+0) - \int_{[a,b]} x(\cdot - 0)\,\mathrm{d}y_\leftarrow - x(a-0)\, y_\leftarrow(a-0)\,.\end{aligned}$$

Analogously, one gets

$$\int_{[a,b]} y_\rightarrow \,\mathrm{d}x = x(b+0)\, y_\rightarrow(b+0) - \int_{[a,b]} x(\cdot + 0)\,\mathrm{d}y_\rightarrow - x(a-0)\, y_\rightarrow(a-0)\,.$$

□

**Proposition A.2.4.** *Let $(H_n)_{n\in\mathbb{N}} \subset BV^+_{\leftrightarrows}$ be a monotonic nondecreasing sequence converging uniformly to $H \in BV^+_{\leftrightarrows}$. Then, for each $y \in BVC_b$ there exists a subsequence $(\widetilde{H}_n)_{n\in\mathbb{N}} \subset (H_n)_{n\in\mathbb{N}}$ satisfying*

$$\lim_{n\to\infty} \int y \,\mathrm{d}\widetilde{H}_n = \int y \,\mathrm{d}H \,.$$

*Proof.* (See also Proposition 2.32 in Milbrodt and Helbig (1999).) Since for functions of $BV^+$ the total variation norm is equal to the supremum on $\mathbb{R}$ and the sequence $(H_n)_{n\in\mathbb{N}}$ is monotonic nondecreasing, one gets

$$\|H_n\|_{BV} = \sup_{t\in\mathbb{R}} |H_n(t)| \le \sup_{t\in\mathbb{R}} |H(t)| = \|H\|_{BV}\,, \quad \forall\, n \in \mathbb{N}\,.$$

Consequently,

$$\left|\int y \,\mathrm{d}H_n\right| \le \int |y| \,\mathrm{d}H_n \le \sup_{t\in\mathbb{R}} |y(t)| \cdot \|H_n\|_{BV} \le \sup_{t\in\mathbb{R}} |y(t)| \cdot \|H\|_{BV} < \infty\,.$$

Hence, there exists a subsequence $(\widetilde{H}_n)_{n\in\mathbb{N}} \subset (H_n)_{n\in\mathbb{N}}$ for which $\int y \,\mathrm{d}\widetilde{H}_n$ is converging. By defining $y_m(t) := \mathbf{1}_{[-m,m]}(t) \cdot y(t) \in BV$ for any $m \in \mathbb{R}^+$ and by partial integration according to Proposition A.2.3,

$$\begin{aligned}\lim_{n\to\infty}\int_{[-m,m]} y\,\mathrm{d}\widetilde{H}_n &= \lim_{n\to\infty}\int_{[-m,m]} y_m\,\mathrm{d}\widetilde{H}_n \\ &= \lim_{n\to\infty}\Big(y_m(m+0)\,\widetilde{H}_n(m+0) - \int_{[-m,m]}\widetilde{H}_n(\cdot-0)\,\mathrm{d}y_{m,\leftarrow} \\ &\qquad - \int_{[-m,m]}\widetilde{H}_n(\cdot+0)\,\mathrm{d}y_{m,\rightarrow} - y_m(-m-0)\,\widetilde{H}_n(-m-0)\Big),\end{aligned}$$

where $y_{m,\rightarrow}+y_{m,\leftarrow}$ is a decomposition of $y_m$ according to Proposition A.2.2. Applying the Monotone Convergence Theorem and using partial integration in reverse to the above, the limit gets

$$\begin{aligned}&y_m(m+0)\,H(m+0) - \int_{[-m,m]} H(\cdot-0)\,\mathrm{d}y_{m,\leftarrow} \\ &\qquad - \int_{[-m,m]} H(\cdot+0)\,\mathrm{d}y_{m,\rightarrow} - y_m(-m-0)\,H(-m-0) \\ &= \int_{[-m,m]} y_m\,\mathrm{d}H = \int_{[-m,m]} y\,\mathrm{d}H\,.\end{aligned}$$

Consequently,

$$\begin{aligned}\left|\int y\,\mathrm{d}H - \lim_{n\to\infty}\int y\,\mathrm{d}\widetilde{H}_n\right| &= \left|\int_{\mathbb{R}\setminus[-m,m]} y\,\mathrm{d}H - \lim_{n\to\infty}\int_{\mathbb{R}\setminus[-m,m]} y\,\mathrm{d}\widetilde{H}_n\right| \\ &\le \int_{\mathbb{R}\setminus[-m,m]} |y|\,\mathrm{d}|H| + \lim_{n\to\infty}\int_{\mathbb{R}\setminus[-m,m]} |y|\,\mathrm{d}|\widetilde{H}_n| \\ &\le \sup_{t\in\mathbb{R}}|y(t)|\cdot\mathcal{V}_{\mathbb{R}\setminus[-m,m]}(H) + \sup_{t\in\mathbb{R}}|y(t)|\cdot\lim_{n\to\infty}\mathcal{V}_{\mathbb{R}\setminus[-m,m]}(\widetilde{H}_n) \\ &\xrightarrow{m\to\infty} 0\,.\end{aligned}$$

□

**Proposition A.2.5.** (a) *For each $H \in BV_{\leftarrow}^{+}$ there exists a monotonic nondecreasing sequence of step functions converging uniformly to $H$,*

$$H_n = \sum_{i=0}^{\infty} a_i^n\,\mathbf{1}_{[t_i^n,\infty)} \in BV_{\leftarrow}^{+}\,, \qquad a_i^n \ge 0\,,\ t_i^n \ge 0\,.$$

(b) *For all step functions $Z$ as declared in* (a) *and all functionals $G$ satisfying Condition* 2.2.5,

$$G(Z) = \int G\big(\mathbf{1}_{[t,\infty)}(\cdot)\big)\,\mathrm{d}Z(t)\,.$$

*Proof.* (See also Proposition 2.34 in Milbrodt and Helbig (1999).) Ad (a): Without loss of generality, set $H(\infty) = 1$. Define

$$H_n := \sum_{i=0}^{2^n} \frac{i}{2^n} \mathbf{1}_{\{\frac{i}{2^n} \le H < \frac{i+1}{2^n}\}} = \frac{1}{2^n} \sum_{i=1}^{2^n} \mathbf{1}_{[H^{-1}(\frac{i}{2^n}),\infty)} .$$

Consequently,

$$\lim_{n\to\infty} \sup_{t\in\mathbb{R}} |H(t) - H_n(t)| \le \lim_{n\to\infty} \frac{1}{2^n} = 0 ,$$

that is, the sequence converges uniformly. The functions $H_n$ are elements of $BV^+$ because as they are monotonic nondecreasing and bounded.

Ad (b): Let $Z$ be represented by

$$Z = \sum_{i=0}^{\infty} a_i \, \mathbf{1}_{[t_i,\infty)} , \quad a_i \ge 0 , \; t_i \ge 0 .$$

Then, the sequence

$$Z_n := \sum_{i=0}^{n} a_i \, \mathbf{1}_{[t_i,\infty)}$$

is monotonic nondecreasing and converging uniformly to $Z$, since $\sum_{i=0}^{\infty} a_i = Z(\infty) < \infty$. The functions $Z, Z_1, Z_2, \ldots$ are elements of $BV^+_{\leftarrow}$. According to Condition 2.2.5(c), there is a subsequence $(\widetilde{Z}_n)_{n\in\mathbb{N}} \subset (Z_n)_{n\in\mathbb{N}}$ for which

$$G(Z) = G(\lim_{n\to\infty} \widetilde{Z}_n) = \lim_{n\to\infty} G(\widetilde{Z}_n) .$$

Because of the presumed linearity of $G$, this is equal to

$$\lim_{n\to\infty} \sum_{i=0}^{n} a_i \, G(\mathbf{1}_{[a_i,\infty)}) = \sum_{i=0}^{\infty} a_i \, G(\mathbf{1}_{[t_i,\infty)}) = \int G(\mathbf{1}_{[t,\infty)}) \, \mathrm{d}Z(t) .$$

Notice that the series

$$\sum_{i=0}^{\infty} a_i \, |G(\mathbf{1}_{[a_i,\infty)})| = \int |G(\mathbf{1}_{[t,\infty)})| \, \mathrm{d}Z(t) \le \sup_{t\in\mathbb{R}} |G(\mathbf{1}_{[t,\infty)})| \cdot \|Z\|_{BV} < \infty$$

is convergent, as $\sup_{t\in\mathbb{R}} |G(\mathbf{1}_{[t,\infty)})|$ is finite according to Condition 2.2.5(b). □

## A.3 Derivatives

Let $F$ be a mapping from a normed vector space $X$ to another normed vector space $Y$.

**Definition A.3.1** (Gâteaux differential)**.** The mapping $F : X \to Y$ is said to be *Gâteaux differentiable at* $x \in X$ if the limit

$$\lim_{\tau \to 0} \frac{F(x+\tau h) - F(x)}{\tau} = \frac{\mathrm{d}}{\mathrm{d}\tau} F(x+\tau h)\Big|_{\tau=0} =: \mathrm{D}_x F(h) \tag{A.3.1}$$

exists for each $h \in X$. The functional $\mathrm{D}_x F : X \to Y$ is called the *Gâteaux differential at* $x$.

The Gâteaux differential $\mathrm{D}_x F$ is homogenous, that is,

$$\mathrm{D}_x F(\sigma h) = \sigma\, \mathrm{D}_x F(h)\,, \quad \forall \sigma \in \mathbb{R}\,, \tag{A.3.2}$$

but not necessarily linear. Gâteaux differentiability is a generalization of directional derivatives. Note that Definition A.3.1 includes the existence of the directional derivative in *all* directions.

Now let $X$ and $Y$ be normed.

**Definition A.3.2** (Fréchet differential)**.** The mapping $F : X \to Y$ is said to be *Fréchet differentiable at* $x \in X$ if there exists a linear and continuous mapping $\mathrm{D}_x F : X \to Y$, the *Fréchet differential at* $x$, with

$$\lim_{\|h\|_X \to 0} \frac{\|F(x+h) - F(x) - \mathrm{D}_x F(h)\|_Y}{\|h\|_X} = 0\,. \tag{A.3.3}$$

In case the Fréchet differential $D_x F$ exists, the Gâteaux differential does as well, and the two differentials are equal. Hence, the same notation is used.

**Example A.3.3.** Define the mapping

$$F : L_1(|\mu|) \to \mathbb{R}\,, \quad x \mapsto \int x \,\mathrm{d}\mu \tag{A.3.4}$$

for any $\sigma$-finite and signed Borel-measure $\mu$. Then,

$$\lim_{\|h\| \to 0} \frac{|F(x+h) - F(x) - F(h)|}{\|h\|_{L_1(|\mu|)}} = \lim_{\|h\| \to 0} \frac{\left|\int (x+h)\,\mathrm{d}\mu - \int x\,\mathrm{d}\mu - \int h\,\mathrm{d}\mu\right|}{\|h\|_{L_1(|\mu|)}} = 0\,,$$

that is, the Fréchet differential of $F$ at $x$ exists and is equal to $\mathrm{D}_x F = F$.

**Definition A.3.4** (Hadamard differential)**.** The mapping $F : X \to Y$ is said to be *Hadamard (or compactly) differentiable at* $x \in X$ if there exists a linear and continuous mapping $\mathrm{D}_x F : X \to Y$, the *Hadamard differential at* $x$, in such a way that

$$\lim_{n \to 0} \left\| \frac{\|F(x+\tau_n h_n) - F(x)}{\tau_n} - \mathrm{D}_x F(h) \right\|_Y = 0 \tag{A.3.5}$$

for any $\tau_n \to 0$ and any $\{h_n\}_{n \geq 0} \subset X$ with $\|h_n - h_0\|_X \to 0$.

Fréchet differentiability at $x$ implies Hadamard differentiability at $x$, which in turn implies Gâteaux differentiability at $x$. All three differentials coincide, provided they exist. A helpful survey of these derivatives is given in Bickel et al. (1998, Appendix A.5).

**Proposition A.3.5** (chain rule)**.** *Suppose that $F : X \supset A \to B \subset Y$ is Fréchet differentiable at $x \in A$ and $G : Y \supset B \to Z$ is Fréchet differentiable at $F(x) \in B$, where $X$, $Y$, and $Z$ are normed vector spaces. Then, $G \circ F$ is Fréchet differentiable at $x$ and*

$$\mathrm{D}_x(G \circ F)(\cdot) = (\mathrm{D}_{F(x)} G) \circ (\mathrm{D}_x F)\,. \tag{A.3.6}$$

*Proof.* See (3.1.4) in Flett (1980). □

## A.4 Some properties of product integrals

General references for this section are Gill (1994) or Gill and Johansen (1990).

**Proposition A.4.1.** *Let $x \in BVC_{\leftarrow}$ and $\Delta x(t) := x(t) - x(t-0) \geq \mathrm{const} > -1$ for all $t \in \mathbb{R}$. Then,*

$$\left( t \mapsto \prod_{(0,t]} (1 + \mathrm{d}x) \right) \in BVC_{\leftarrow}\,.$$

*Proof.* Denote by $x = x_+ - x_-$ the Jordan-Hahn decomposition of $x$ according to Theorem A.2.1. As $\Delta x(t) \geq \mathrm{const} > -1$ and $x$ is dominated by $|x| := x_+ + x_-$, one has

$$0 \leq \prod_{(s,t]} \big(1 + \mathrm{d}x\big) \leq \prod_{(s,t]} \big(1 + \mathrm{d}|x|\big)\,, \quad \forall\, s \leq t\,.$$

Thus, with the multiplicativity of product-integration and two times the Forward

Integral Equation in Gill (1994, p. 125) one gets

$$\begin{aligned}
\mathcal{V}_{[0,t]}\left(\prod_{(0,\cdot]}(1+\mathrm{d}x)\right) &= \sup_{\mathcal{T}\subset[0,t]}\sum_{\mathcal{T}}\left|\prod_{(0,t_{i+1}]}(1+\mathrm{d}x)-\prod_{(0,t_i]}(1+\mathrm{d}x)\right| \\
&= \sup_{\mathcal{T}\subset[0,t]}\sum_{\mathcal{T}}\prod_{(0,t_i]}(1+\mathrm{d}x)\left|\prod_{(t_i,t_{i+1}]}(1+\mathrm{d}x)-1\right| \\
&= \sup_{\mathcal{T}\subset[0,t]}\sum_{\mathcal{T}}\prod_{(0,t_i]}(1+\mathrm{d}x)\left|\int_{(t_i,t_{i+1}]}\prod_{(t_i,u)}(1+\mathrm{d}x)\,\mathrm{d}x(u)\right| \\
&\le \sup_{\mathcal{T}\subset[0,t]}\sum_{\mathcal{T}}\prod_{(0,t_i]}(1+\mathrm{d}|x|)\int_{(t_i,t_{i+1}]}\prod_{(t_i,u)}\big(1+\mathrm{d}|x|\big)\,\mathrm{d}|x|(u) \\
&= \sup_{\mathcal{T}\subset[0,t]}\sum_{\mathcal{T}}\left|\prod_{(0,t_{i+1}]}(1+\mathrm{d}|x|)-\prod_{(0,t_i]}(1+\mathrm{d}|x|)\right| \\
&= \mathcal{V}_{[0,t]}\left(\prod_{(0,\cdot]}(1+\mathrm{d}|x|)\right).
\end{aligned}$$

The function $\prod_{(0,\cdot]}\big(1+\mathrm{d}|x|\big)$ has finite variation on compacts, since it is monotone nondecreasing and finite for each $t\in\mathbb{R}$. Hence $\prod_{(0,\cdot]}\big(1+\mathrm{d}x\big)$ has finite variation on compacts, too. The right-continuity is due to integrating over the right-closed intervals $(0,t]$. □

**Proposition A.4.2.** *Let $x\in BVC_{\leftarrow}$ and $\Delta x(t):=x(t)-x(t-0)\ge C>-1$ for all $t\in\mathbb{R}$. Denote by $x_c$ the continuous part of $x$ (see the proof of Proposition* A.2.2*). Then,*

$$\prod_{(0,t]}\Big(1+\mathrm{d}x\Big)=\prod_{\tau\in(0,t]}\Big(1+\Delta x(\tau)\Big)\cdot\exp\Big(x_c(t)\Big) \tag{A.4.1}$$

*for all $t>0$ with upper and lower bounds of*

$$\prod_{(0,t]}\Big(1+\mathrm{d}x\Big)\ge\exp\Big(-\frac{\ln(1+C)}{C}\|\mathbf{1}_{[0,t]}x\|_{BV}\Big)>0\,, \tag{A.4.2}$$

$$\prod_{(0,t]}\Big(1+\mathrm{d}x\Big)\le\exp\Big(\|\mathbf{1}_{[0,t]}x\|_{BV}\Big). \tag{A.4.3}$$

The proof follows that of Theorem 2.7 in Milbrodt and Helbig (1999):

*Proof.* According to Gill (1994, pp. 126, 127), the product-integral in the scalar case is just the unique solution of the Volterra integral equations (also denoted as forward and backward integral equations). Therefore, it is shown that

$$w(t):=\prod_{\tau\in(0,t]}\big(1+\Delta x(\tau)\big)\cdot\exp\big(x_c(t)\big)\,,\quad t>0\,,$$

solves the forward integral equation

$$w(t) - 1 = \int_{(0,t]} w(\tau - 0)\,\mathrm{d}x(\tau)\,, \quad t > 0\,. \tag{A.4.4}$$

Decompose $w$ into a product of the functions

$$y : t \mapsto \prod_{\tau \leq t} \big(1 + \Delta x(\tau)\big)\,, \qquad z : t \mapsto \exp\big(x_c(t)\big)\,, \quad t > 0\,,$$

which are well defined, as a function of finite variation on compacts has, at most, a countable number of jumps, and because

$$\ln\Big(\prod_{\tau \leq t} \big(1 + \Delta x(\tau)\big)\Big) \leq \sum_{\tau \leq t} \Delta x(\tau) \leq x(t) < \infty\,.$$

(i) The function $y$ is strictly positive as (without loss of generality let $C < 0$)

$$\begin{aligned}
\ln\Big(\prod_{\tau \leq t} \big(1 + \Delta x(\tau)\big)\Big) &\geq \ln\Big(\prod_{\tau \leq t} \big(1 - \Delta x_-(\tau)\big)\Big) \\
&= \sum_{\tau \leq t} \ln\Big(\big(1 - \Delta x_-(\tau)\big)\Big) \\
&\geq \sum_{\tau \leq t} -\Delta x_-(\tau) \frac{\ln(1 + C)}{C} \\
&\geq -\|\mathbf{1}_{[0,t]} x_-\|_{BV} \frac{\ln(1 + C)}{C} > -\infty\,,
\end{aligned}$$

where $x = x_+ - x_-$ denotes the Jordan-Hahn decomposition according to Theorem A.2.1. The second inequality is due to $0 \leq \Delta x_-(t) \leq -C < 1$ and the concavity of the logarithm.

(ii) With the strict positivity in (i),

$$\Delta y(t) = y(t - 0)\left(\frac{y(t)}{y(t - 0)} - 1\right) = y(t - 0)\,\Delta x(t)\,, \quad t > 0\,. \tag{A.4.5}$$

(iii) Arguing analogously to the proof of Proposition A.4.1 yields that $z$ is an element of $BVC_{\leftarrow}$. The corresponding measure to $z$ is absolutely continuous with respect to that corresponding to $x_c$ with density

$$\frac{\mathrm{d}z}{\mathrm{d}x_c} = z\,. \tag{A.4.6}$$

To proof this, it suffices to show

$$\int_{(a,b]} \mathrm{d}z = \int_{(a,b]} z\,\mathrm{d}x_c\,, \qquad \text{for all } -\infty < a < b < \infty\,. \tag{A.4.7}$$

For any fixed $-\infty < a < b < \infty$ and any finite interval segmentation $\mathcal{T}$ with $a = t_0 < t_1 < \cdots < t_n = b$,

$$
\begin{aligned}
&\left| \int_{(a,b]} \mathrm{d}e^{x_c(t)} - \int_{(a,b]} e^{x_c(t)} \, \mathrm{d}x_c(t) \right| \\
&\leq \sum_{\mathcal{T}} \left| \int_{(t_i,t_{i+1}]} \mathrm{d}e^{x_c(t)} - \int_{(t_i,t_{i+1}]} e^{x_c(t)} \, \mathrm{d}x_c(t) \right| \\
&\leq \sum_{\mathcal{T}} \left( \left| \int_{(t_i,t_{i+1}]} \mathrm{d}e^{x_c(t)} - \int_{(t_i,t_{i+1}]} e^{x_c(t_i)} \, \mathrm{d}x_c(t) \right| + \left| \int_{(t_i,t_{i+1}]} \big( e^{x_c(t_i)} - e^{x_c(t)} \big) \, \mathrm{d}x_c(t) \right| \right) \\
&\leq \sum_{\mathcal{T}} e^{x_c(t_i)} \left| e^{x_c(t_{i+1}) - x_c(t_i)} - 1 - \big( x_c(t_{i+1}) - x_c(t_i) \big) \right| \\
&\quad + \sum_{\mathcal{T}} \sup_{t \in (t_i,t_{i+1}]} \left| e^{x_c(t_i)} - e^{x_c(t)} \right| \big( |x_c|(t_{i+1}) - |x_c|(t_i) \big) \, .
\end{aligned}
$$

As $x_c$ is continuous on the compact interval $[a,b]$, it is bounded and for each $\varepsilon > 0$ there exists a finite interval segmentation $\mathcal{T}$ with $|x_c(t_{i+1}) - x_c(t_i)| < \varepsilon$. The same holds for $|x_c|$ and $e^{x_c(t)}$. Uniting the three corresponding interval segmentations lets $|x_c(t_{i+1}) - x_c(t_i)| < \varepsilon$, $0 \leq |x_c|(t_{i+1}) - |x_c|(t_i) < \varepsilon$, and $|e^{x_c(t_{i+1})} - e^{x_c(t_i)}| < \varepsilon$, simultaneously. Further on, $|e^y - 1 - y| \leq y^2$ for any $|y| \leq 1$. Consequently, the upper estimate continuous to

$$
\begin{aligned}
\ldots &\leq \sup_{t \in (a,b]} e^{x_c(t)} \sum_{\mathcal{T}} \big( x_c(t_{i+1}) - x_c(t_i) \big)^2 \\
&\quad + \sum_{\mathcal{T}} \sup_{t \in (t_i,t_{i+1}]} \left| e^{x_c(t_i)} - e^{x_c(t)} \right| \big( |x_c|(t_{i+1}) - |x_c|(t_i) \big) \\
&\leq \sup_{t \in (a,b]} e^{x_c(t)} \big( |x_c|(b) - |x_c|(a) \big) \, \varepsilon \\
&\quad + \big( |x_c|(b) - |x_c|(a) \big) \, \varepsilon \\
&= \mathrm{const} \cdot \varepsilon
\end{aligned}
$$

for any $0 < \varepsilon \leq 1$ and proper interval segmentations $\mathcal{T}$. This implies (A.4.7).

Partial integration (Proposition A.2.3), equation (A.4.6), and (A.4.5) lead to the forward equation (A.4.4),

$$\begin{aligned} 1 - w(t) &= y(0)\, z(0) - y(t)\, z(t) \\ &= - \int_{(0,t]} y(\tau - 0)\, \mathrm{d}z(\tau) - \int_{(0,t]} z(\tau)\, \mathrm{d}y(\tau) \\ &= - \int_{(0,t]} y(\tau - 0)\, z(\tau)\, \mathrm{d}x_c(\tau) - \sum_{0<\tau\le t} z(\tau)\, y(\tau - 0)\, \Delta x(\tau) \\ &= - \int_{(0,t]} y(\tau - 0)\, z(\tau)\, \mathrm{d}x(\tau) \\ &= - \int_{(0,t]} w(\tau - 0)\, \mathrm{d}x(\tau)\,. \end{aligned}$$

□

## A.5 Integration of stochastic diffusions with additive noise

**Proposition A.5.1.** *The differential equation* (4.1.1) *has a unique and $t$-continuous solution on $[0,T]$ if*

(a) (Measurability) *the mappings $\alpha = \alpha(x,t)$ and $\sigma = \sigma(x,t)$ are jointly square integrable on $(x,t) \in \mathbb{R} \times [0,T]$,*

(b) (Lipschitz condition) *there exists a constant $K > 0$ in such a way that*

$$|\alpha(x,t) - \alpha(y,t)| \le K\,|x-y|\,, \quad |\sigma(x,t) - \sigma(y,t)| \le K\,|x-y| \tag{A.5.1}$$

*for all $t \in [0,T]$ and $x, y \in \mathbb{R}$, and*

(c) (Linear growth bound) *there exists a constant $K > 0$ in such a way that*

$$|\alpha(x,t)|^2 \le K^2 \left(1 + |x|^2\right), \quad |\sigma(x,t)|^2 \le K^2 \left(1 + |x|^2\right) \tag{A.5.2}$$

*for all $t \in [0,T]$ and $x \in \mathbb{R}$.*

*Proof.* See Kloeden and Platen (1992, section 4.5). □

**Proposition A.5.2** (linear stochastic differential equation with additive noise)**.** *Let the functions $\alpha_1$, $\alpha_2$, and $\sigma_2$ be Lebesgue measurable and bounded on $[0,T]$. Then,*

$$\phi_t = \phi_0\, e^{\int_0^t \alpha_1(s)\,\mathrm{d}s} + \int_0^t e^{\int_\tau^t \alpha_1(s)\,\mathrm{d}s} \alpha_2(\tau)\,\mathrm{d}\tau + \int_0^t e^{\int_\tau^t \alpha_1(s)\,\mathrm{d}s} \sigma_2(\tau)\,\mathrm{d}W_\tau \tag{A.5.3}$$

*is for $t \in [0,T]$ a solution of the stochastic differential equation*

$$\mathrm{d}\phi_t = \big(\alpha_1(t)\,\phi_t + \alpha_2(t)\big)\,\mathrm{d}t + \sigma_2(t)\,\mathrm{d}W_t\,, \quad \phi_0 = \text{const}\,. \tag{A.5.4}$$

*Particularly, the stochastic process $(\phi_t)$ has a $t$-continuous version.*

*Proof.* For the explicit solution (A.5.3), see section 4.4 in Kloeden and Platen (1992), for the continuity, see Theorem 3.2.6 in Kloeden and Platen (1992). □

**Proposition A.5.3.** *Let the functions $\alpha_1$, $\alpha_2$, $\sigma_2$, and $f$ be Lebesgue-measurable and bounded on $[0,T]$, and let $(\phi_t)$ be a $t$-continuous solution of* (A.5.4)*. Then, for any fixed $t \in [0,T]$ the term*

$$\int_0^t f(s)\,\phi_s\,\mathrm{d}s$$

*is normally distributed with expectation*

$$\mu_t := \int_0^t f(s)\Big(\phi_0\, e^{\int_0^s \alpha_1(u)\,\mathrm{d}u} + \int_0^s e^{\int_u^s \alpha_1(v)\,\mathrm{d}v}\alpha_2(u)\,\mathrm{d}u\Big)\,\mathrm{d}s$$

*and variance*

$$\sigma_t^2 := \int_0^t \bigg(\Big(\int_s^t f(u)\, e^{\int_0^u \alpha_1(v)\,\mathrm{d}v}\,\mathrm{d}u\Big)\, e^{-\int_0^s \alpha_1(v)\,\mathrm{d}v}\sigma_2(s)\bigg)^2\,\mathrm{d}s\,.$$

*Proof.* With $(\phi_t)$ being $t$-continuous, the integral is well-defined pathwise. Applying Proposition A.5.2 leads to

$$\begin{aligned}\int_0^t f(s)\,\phi_s\,\mathrm{d}s &= \int_0^t f(s)\Big(\phi_0\, e^{\int_0^s \alpha_1(u)\,\mathrm{d}u} + \int_0^s e^{\int_u^s \alpha_1(v)\,\mathrm{d}v}\alpha_2(u)\,\mathrm{d}u\Big)\,\mathrm{d}s \\ &\quad + \int_0^t f(s)\Big(e^{\int_0^s \alpha_1(v)\,\mathrm{d}v}\int_0^s e^{-\int_0^u \alpha_1(v)\,\mathrm{d}v}\sigma_2(u)\,\mathrm{d}W_u\Big)\,\mathrm{d}s\,.\end{aligned}$$

Integration by parts (see Kloeden and Platen (1992), formula (4.10) in Example 3.4.1) yields for the second addend

$$\begin{aligned}&\int_0^t f(s)\, e^{\int_0^s \alpha_1(v)\,\mathrm{d}v}\Big(\int_0^s e^{-\int_0^u \alpha_1(v)\,\mathrm{d}v}\sigma_2(u)\,\mathrm{d}W_u\Big)\,\mathrm{d}s \\ &= \int_0^t f(s)\, e^{\int_0^s \alpha_1(v)\,\mathrm{d}v}\,\mathrm{d}s \int_0^t e^{-\int_0^s \alpha_1(v)\,\mathrm{d}v}\sigma_2(s)\,\mathrm{d}W_s \\ &\quad - \int_0^t \Big(\int_0^s f(u)\, e^{\int_0^u \alpha_1(v)\,\mathrm{d}v}\,\mathrm{d}u\Big)\, e^{-\int_0^s \alpha_1(v)\,\mathrm{d}v}\sigma_2(s)\,\mathrm{d}W_s \\ &= \int_0^t \Big(\int_s^t f(u)\, e^{\int_0^u \alpha_1(v)\,\mathrm{d}v}\,\mathrm{d}u\Big)\, e^{-\int_0^s \alpha_1(v)\,\mathrm{d}v}\sigma_2(s)\,\mathrm{d}W_s\,.\end{aligned}$$

The latter term is normally distributed with expectation zero and variance

$$\int_0^t \bigg(\Big(\int_s^t f(u)\, e^{\int_0^u \alpha_1(v)\,\mathrm{d}v}\,\mathrm{d}u\Big)\, e^{-\int_0^s \alpha_1(v)\,\mathrm{d}v}\sigma_2(s)\bigg)^2\,\mathrm{d}s\,,$$

which is a consequence of the Isometry Property of Ito integration (cf. Kloeden and Platen (1992), Theorem 3.2.3). □

Ich versichere hiermit an Eides statt, dass ich die vorliegende Arbeit selbständig angefertigt und ohne fremde Hilfe verfasst habe, keine außer den von mir angegebenen Hilfsmitteln und Quellen dazu verwendet habe und die den benutzten Werken inhaltlich und wörtlich entnommenen Stellen als solche kenntlich gemacht habe.

Rostock, den 4. Oktober 2006
Marcus C. Christiansen

# Lebenslauf des Autors

| | |
|---|---|
| 13. Oktober 1978 | geboren in Michigan (Vereinigte Staaten) |
| 1985 – 1988 | Grundschule Erich-Kästner (Göttingen) |
| 1988 – 1992 | Grundschule Garbsen (bei Hannover),<br>Orientierungsstufe Garbsen,<br>Gymnasium Garbsen |
| 1992 – 1997 | Gymnasium Burg (bei Magdeburg),<br>abgeschlossen mit der Allgemeinen Hochschulreife |
| Oktober 1997 | Immatrikulation an der Otto-von-Guericke Universität Magdeburg für den Studiengang Mathematik mit Nebenfach Physik |
| 1998 – 1999 | Unterbrechung des Studiums im 3. und 4. Semester für den Zivildienst bei einer Blindenhörbücherei in Darmstadt |
| März 2003 | erfolgreicher Abschluss des Studiums,<br>Diplomarbeit betreut durch Prof. Dr. Gerd Christoph mit dem Titel:<br>*Untersuchungen von Konvergenzgeschwindigkeiten in der Risiko-Theorie für regulär variierende Verteilungen*<br><br>Träger des Fakultätspreises Mathematik 2003 der Otto-von-Guericke-Universität Magdeburg |
| seit Juli 2003 | Wissenschaftlicher Mitarbeiter am Lehrstuhl für Finanz- und Versicherungsmathematik an der Universität Rostock bei Prof. Dr. Hartmut Milbrodt |
| seit 2004 | Mitglied der Fachgruppe Stochastik |

Rostock, den 4. Oktober 2006