Durational effects and non-smooth semi-Markov models in life insurance

Dissertation

zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Rostock

vorgelegt von

Dipl.-Math. Marko Helwich,
geboren am 09. Dezember 1976 in Schwerin,
aus Rostock

Rostock, im Dezember 2007

urn:nbn:de:gbv:28-diss2008-0056-4

Gutachter:
Prof. Dr. Hartmut Milbrodt, Institut für Mathematik, Universität Rostock
Prof. Dr. Friedrich Liese, Institut für Mathematik, Universität Rostock
Prof. Dr. Klaus D. Schmidt, Institut für Mathematische Stochastik, TU Dresden

Tag der Verteidigung:
27. Mai 2008

Great are the works of the LORD;
they are pondered by all who delight in them.
Psalm 111, verse 2

Contents

Chapter 1

Introduction

The aim of this work is to model and investigate so-called durational effects in life insurance. This means, in the context of a multiple state model for a single risk, that probabilities of transitions between certain states as well as actuarial payments are not completely determined by the current state of the policy, but may also depend on the time elapsed since entering this state.

The development of an insured risk is usually described by non-homogeneous Markovian pure jump processes with a finite state space containing all possible states of the policy. The Markov property assures that the future development of such a process only depends on the state which is occupied at a certain time. Further, the actuarial payments, which are often separately considered as sojourn payments and payments due to transitions, are constructed such that they also solely depend on the current state of the policy. Here, a generalized model is presented that allows an appropriate implementation of durational effects, such that both transition probabilities and actuarial payments are additionally allowed to depend on the time elapsed since entering the current state. In order to achieve this, the development of an insured risk is modelled by non-homogeneous semi-Markovian pure jump processes. Note that each Markov process is also a semi-Markov process.

A pure jump process $(X_t)_{t\geq 0}$ with finite state space $\mathcal{S}$ can also be described by using a marked point process with space of marks $\mathcal{S}$. The marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$ that appertains to a pure jump process - which is either Markovian or semi-Markovian - is a Markov chain. Hence, in order to develop a model that covers both the Markov as well as the semi-Markov approach, our theory mostly relies on the Markov property of the appertaining marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$. For reasons to be explained later, this process is further assumed to be homogeneous. The Markov property of $((T_m, Z_m))_{m\in\mathbb{N}_0}$ implies the Markov property of the bivariate process $((X_t, U_t))_{t\geq 0}$. For each $t \geq 0$, the process $((X_t, U_t))_{t\geq 0}$ records both the current state of the policy, X_t, and the time spent in state X_t up to time t since the latest transition to that state, U_t. Thus, for our approach, the pure jump process $(X_t)_{t\geq 0}$ is not necessarily Markovian, but the bivariate process $((X_t, U_t))_{t\geq 0}$ is a Markov process. In contrast to the pure jump process $(X_t)_{t\geq 0}$, however, this process does not have a finite state space. The state space of $((X_t, U_t))_{t\geq 0}$ is given by $\mathcal{S} \times [0, \infty)$.

Before giving a survey of the literature that is concerned with durational effects in life insurance, and afterwards presenting the outline of this thesis, three examples from life insurance are discussed for which the probabilities of certain transitions actually depend on the time elapsed since the current state was entered. Doing so, the importance of durational effects with respect to the development of a single policy shall be pointed out. The insurance products being considered are German private health insurance (PKV: **P**rivate **K**ranken**v**ersicherung), permanent disability insurance (also referred to as permanent health insurance (PHI)), and long-term care insurance (LTC).

German private health insurance forms the capital funded part of the health insurance system in Germany. Hence, the following issues are a distinctively German matter of interest. Since policyholders currently lose their ageing provision (i.e. the prospective reserve) if they switch insurance companies, there is, on the one hand, a lack of competition. This is being discussed in German politics at the present time. On the other hand, an interesting actuarial issue is raised, namely that withdrawal probabilities decrease with the time of being insured. The reason for this is that the prospective reserve - and with that the loss due to withdrawal - increases over the first years. Hence, the withdrawal probabilities do not only depend on the attained age of an insured, but also on the previous duration of the contract. Figure 1 gives an impression of this effect. For a real existing PKV portfolio with annual withdrawal rates structured by technical age and time elapsed since entry into the portfolio, mere age-depending withdrawal rates $w_x, x \in \{x_{MIN}, ..., x_{MAX}\}$ are compared with age- and duration-depending withdrawal rates $w_{x,u}, x \in \{x_{MIN}, ..., x_{MAX}\}, u \in \{0, ..., x - x_{MIN}\}$, in the case of 36-year old insured of both gender and previous contract durations $u = 0, ..., 15$.

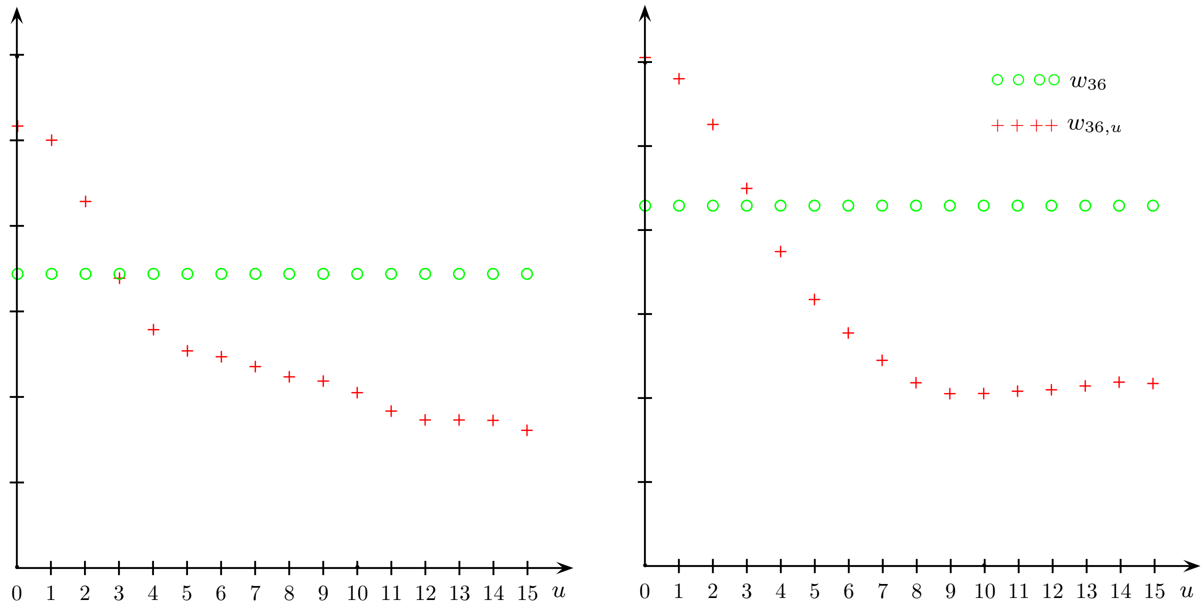

Figure 1: Age- and duration-depending annual withdrawal rates compared with mere age-depending withdrawal rates with respect to the previous contract duration u, for men (left) and women (right) at the age of $x = y = 36$

Currently, the decrement of a PKV portfolio is modelled by using a Markov approach with a mere age-depending actuarial basis. However, figure 1 shows that in doing so, the actual withdrawal rates of insured with a short previous contract duration are underestimated and the withdrawal rates of policyholders being insured for a longer time are overestimated. The latter is a very crucial point, because the withdrawal rates are used to calculate the reserve that is left by the withdrawing insured and put to increase the reserve of the remaining insured. In a certain sense, this is like giving financial support to the remaining insured, resulting in lower net premiums for them. Consequently, if the actual withdrawal rates are lower than the ones used for actuarial calculations, the resulting net premiums might be too low. Hence, the risk of the portfolio is not really covered. For this reason, a calculation relying on an age- and duration-depending actuarial basis could help to avoid losses for the insurance company, or at least to prevent from unintended shifting of risks. Yet, according to the obligatory regulations concerning the actuarial modelling for German private health insurance - i.e. the act concerning the health insurance calculation (KalV: **Kal**kulations**v**erordnung [1996]) and the German

Insurance Supervisory Law (VAG: **V**ersicherungs**a**ufsichts**g**esetz [2004]) - it is not permitted to base the calculation on an age- and duration-depending actuarial basis. For if this were the case, premiums would also depend on previous contract durations, basically in the following way: The longer the previous contract duration, the higher the premium. This, however, is forbidden in order to prevent premiums for long-term insured becoming more expensive than premiums for new entries at the same age. Incidentally, in this case, insured could withdraw their contracts and afterwards enter a new contract at the same company with lower premiums, provided that they pass the health examination. An appropriate solution of this dilemma - on the one hand, the actual withdrawal rates depend on both age and previous contract duration, and on the other hand, it is not permitted to use a model that takes this into account - will be introduced later on (see example 4. 13). Further, we will sketch the change of the present situation for PKV modelling caused by the regulations for improving the competition for statutory health insurance (GKV-WSG: **G**esetzliche **K**ranken**v**ersicherung-**W**ettbewerbs**s**tärkungs**g**esetz [2007]), and we will clarify whether or not our approach remains appropriate.

Figure 2 presents the durational effects in the situation of the PKV portfolio on the level of probabilities of remaining in the portfolio. The pure jump process used to model the decrement of a PKV portfolio allows the states *active*, *withdrawal* and *dead*. Thus, at discrete times, the probabilities of remaining $._{}p_x$ and $._{}p_{x,u}$ can, under certain assumptions, be calculated with the aid of both annual withdrawal rates and annual mortality rates (cf. example 2. 38).

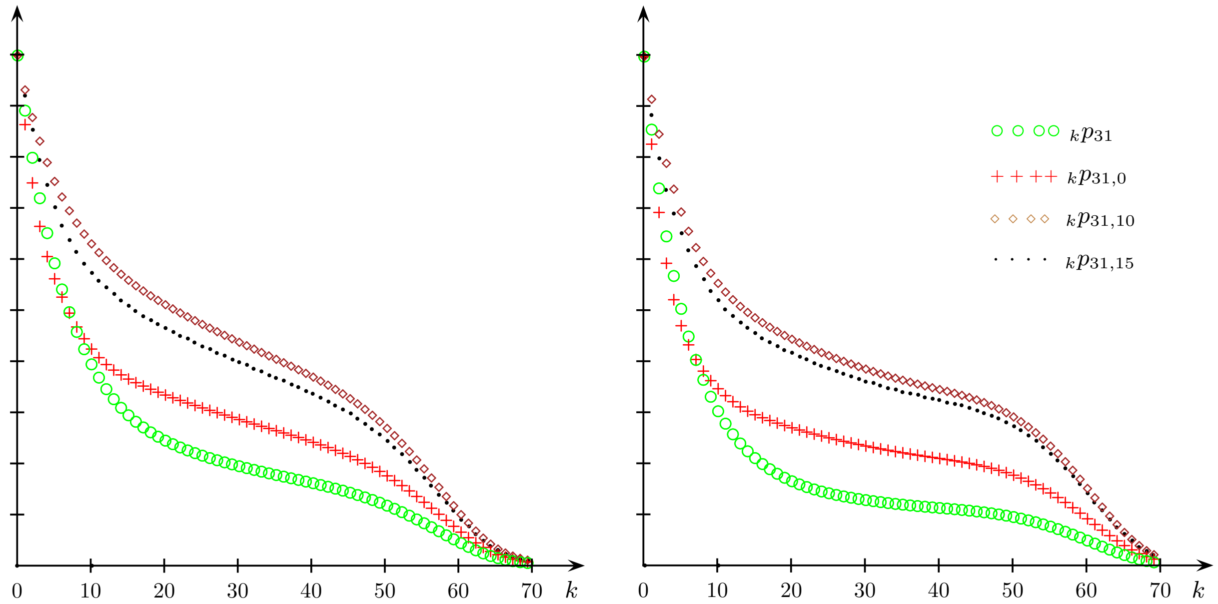

Figure 2: Comparison of probabilities of remaining k-years in the portfolio for a model with a mere age-depending actuarial basis with corresponding probabilities for a model with an age- and duration-depending actuarial basis and previous contract durations $u = 0, 10, 15$, for men (left) and women (right) with attained age $x = y = 31$

We turn to the second example for which durational effects play an important role, the permanent disability insurance. Since the pressure on existing social welfare systems increases for different reasons, the care for disabled and elderly persons must be funded more privately. Therefore, insurance products to cover the risk of financial losses or even the risk of a financial ruin due to disability become more important. We refer to insurance products of this kind as permanent health insurance (PHI). This usage originates from the British one. A permanent health insurance policy should not be mistaken for a private health insurance policy. While the intention of the latter is basically the absorbtion of costs due to health services, a PHI contract provides an insured with an income if the insured is prevented from working by disability due to sickness or injury.

PHI policies are usually also modelled by multiple state models with state space $\mathcal{S} := \{a, i, d\}$. The three possible states are referred to as a ∼*active*, i ∼*invalid* and d ∼*dead*. In cases where recovery is implemented in the model, the set of possible transitions - generally being a subset of $\{(y,z) \in \mathcal{S}^2, y \neq z\}$ - is given by $\mathcal{J} := \{(a,i), (a,d), (i,a), (i,d)\}$. Figure 3 illustrates the set of possible states and the corresponding transitions for the model of a PHI policy.

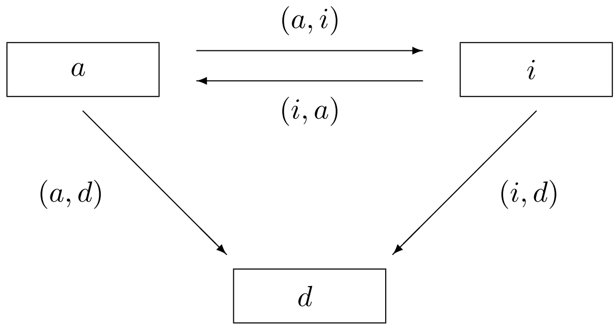

Figure 3: Set of states and set of transitions for the PHI model

One can easily imagine that especially for transitions from the state *invalid* the corresponding probabilities depend not only on the information that a person is disabled at a certain time or at a certain age (both of which are equivalent), but also on the time elapsed since disablement. It is a widespread opinion that for disabled insured both the probability of recovering and the probability of dying decrease with increasing duration of disability. Segerer [1993], for example, investigated mortality and recovery rates relying on a data base coming from reinsurance portfolios of several German life insurance companies. He came to the conclusion that mortality during the first years of disability is significantly higher than the mortality used for the life insurance premium calculation. For higher ages and longer durations of disability, the mortality of disabled persons approaches the normal mortality of insured persons. Regarding the recovery rates, Segerer stated that with increasing age and duration of disability, recovery becomes less probable.

Recovery and mortality rates for disabled insured that depend on both age and time elapsed since disablement are provided by so-called select-and-ultimate tables. Such tables generally contain annual rates that depend on two variables (cf. Bowers et al. [1997], section 3.8). One variable, $[x]$, is the age at selection (e.g. onset of disability), and the second variable, t, is the duration since selection. Thus, a two-dimensional array is generated. The dependence on age is recorded along the columns, and the dependence on time since selection is recorded along the rows. For example, $q^{ii}_{[x-t]+t}$ is understood as the annual mortality rate of an disabled insured with attained age x who became invalid t years ago at the age of $[x-t]$. The impact of the time since selection on the annual rates often diminishes following selection, such that beyond a certain period, the dependence on the time since selection can be neglected. Consequently, it is economical to construct select-and-ultimate tables by truncation of the two-dimensional array after the first $(r+1)$ columns, for example by means of

$$q^{ii}_{[x]},\ q^{ii}_{[x-1]+1}, \ldots,\ q^{ii}_{[x-r]+r} =: q^{ii}_x = q^{ii}_{[x-t]+t},\ \ t \geq r.$$

The number r is referred to as select period. Figure 4 sketches the structure of a select-and-ultimate table for recovery rates r_* with a select period of 5 years.

To provide numerical examples for our results, the actuarial basis for calculations concerning PHI contracts is formed by the German select-and-ultimate tables DAV-SRT 1997 RI M (for the recovery of disabled insured) and DAV-SST TI 1997 M (for the mortality of disabled insured).

Both tables contain annual rates for male insured and have a select period of $r = 5$. Even though these tables are about a decade old, they are still in use for corresponding calculations in practise.

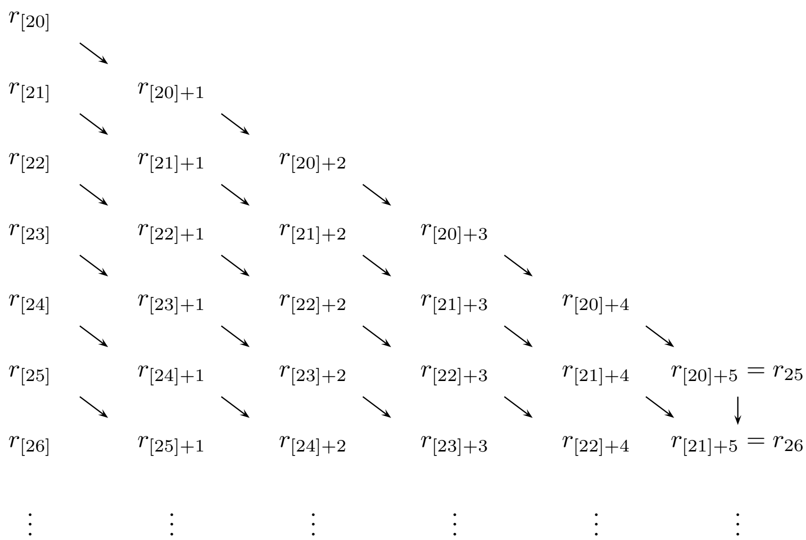

Figure 4: The Structure of a select-and-ultimate table for recovery with a 5-year select period

Figure 5 exhibits the dependence of the annual rates on the time elapsed since onset of disability for three different ages. It can be observed that the mortality rates actually decrease with the time elapsed since disablement. However, the recovery rates increase, especially for young ages, in the first years of invalidity and reach their highest value between two and four years after disablement. Summarizing, the annual rates considered differ substantially with the time elapsed since onset of disability. Hence, the durational effects cannot be easily neglected. This is especially important for the consideration of policies for which the select period carries significant weight when compared with the policy term.

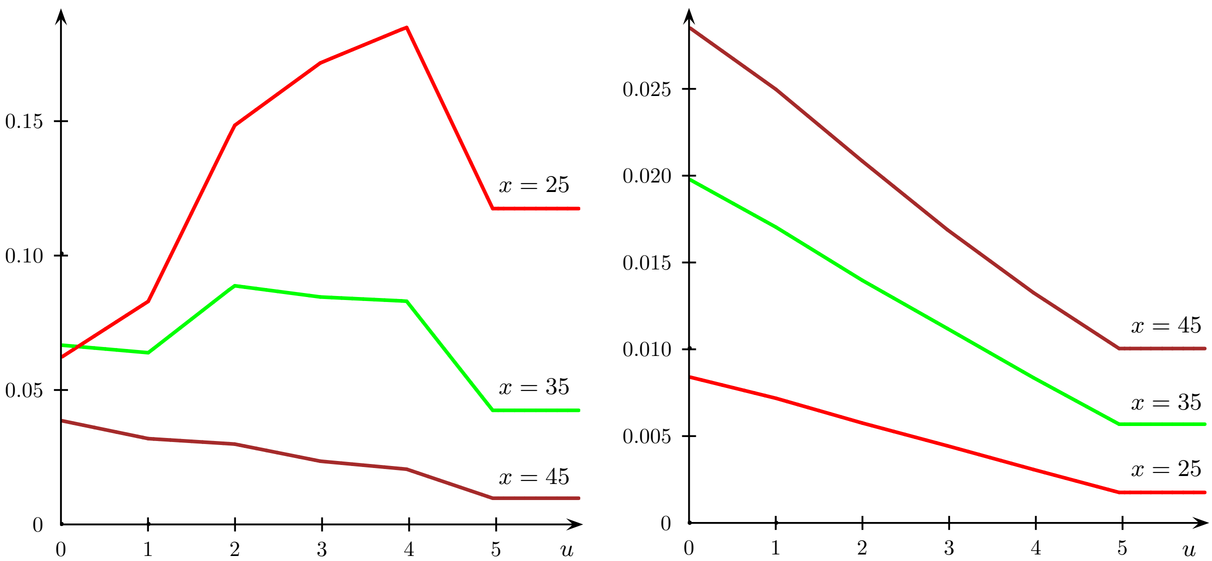

Figure 5: Linear interpolated annual recovery rates (left) and mortality rates (right) for disabled insured at the age of $x = 25, 35, 45$ with respect to the time elapsed since onset of disability, taken from the German tables DAV-SRT 1997 RI M and DAV-SST TI 1997 M for men

Conditional probabilities of remaining in state *invalid* that appertain to the tables DAV-SRT 1997 RI M as well as DAV-SST TI 1997 M are illustrated in figure 6. According to definition 2.24, $\bar{p}_i(s,t,u)$ is the conditional probability of a disabled insured, at a certain age at time s, to remain in this state between the times s and t, given that the onset of disability took place at time $s-u$. Figure 6 shows that this conditional probability has smaller values for insured whose disablement dates back less than five years. Note that for PHI contracts, the major part of benefits is usually due to the state *invalid*. Hence, lower probabilities of remaining in this state lead to lower premiums. Thus, it can be expected that the implementation of durational effects results in lower premiums. Recall that for German private health insurance, the implementation of durational effects seemed to result in higher premiums.

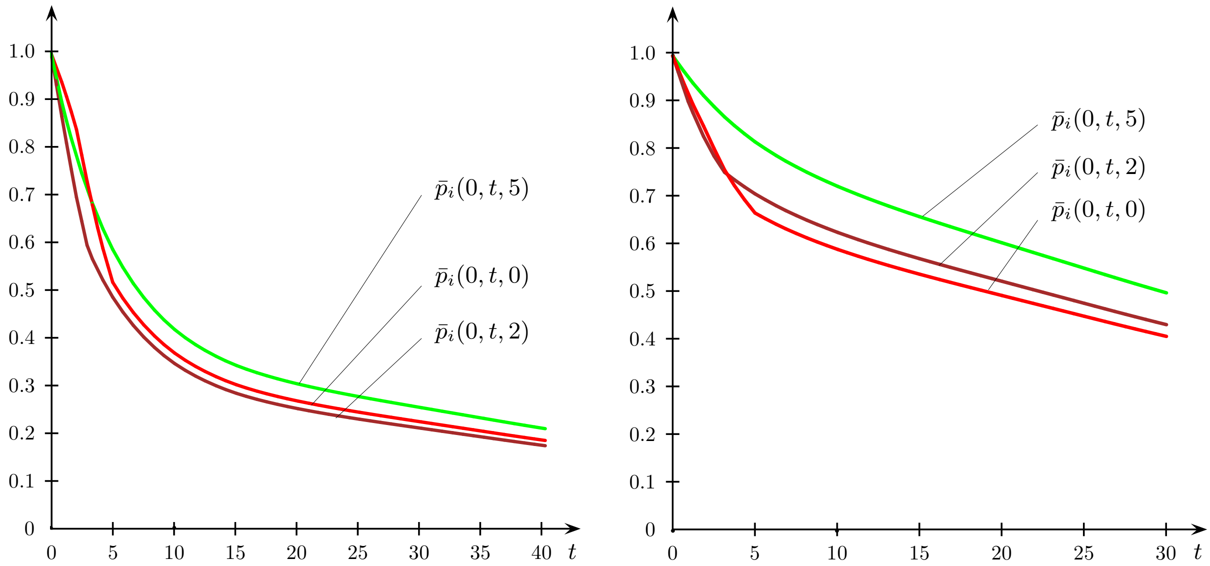

Figure 6: Linear interpolated conditional probabilities of remaining in state invalid according to the German tables DAV-SRT 1997 RI M and DAV-SST TI 1997 M, for male insured with attained age 25 at time $s=0$ (left) and with attained age 35 at time $s=0$ (right), as well as previous durations of invalidity of $0, 2,$ and 5 years

The third insurance product, for which the implementation of age- and duration-depending transition probabilities might be an interesting issue, is long-term care insurance (LTC). Products of this kind provide financial support for insured who are in need of nursing or medical care. The need for care due to the frailty of an insured is classified according to the individuals ability to take care of himself by performing activities of daily living (ADL). Some ADLs are: eating, bathing, moving around, going to the toilet, and dressing (cf. Haberman and Pitacco [1999], section 6.1). In Germany, one distinguishes two different categories of care needs, namely *ambulatory care* (or out-patient care) and *in-patient care*. Further, three different levels of frailty are defined (Pflegestufe I - III).

LTC policies are also modelled by multiple state models. The state space usually consists of the states *active*, *dead*, and the corresponding levels of frailty. For LTC models, the dependence of the transition probabilities on the time spent in a certain state of frailty seems to be even more important than the dependence on the age of an insured. Rudolph [2000], for example, used a proportional hazard model to estimate transition intensities which depend on the time since the beginning of any nursing or medical care. In his model, the age is only a covariable, but nonetheless has a significant impact. Consequently, transition probabilities that depend on both the age of an insured and the time spent in the current state - instead of the time since the beginning of any care - could lead to a more realistic modelling.

Seger et al. [2007] were concerned with providing an actuarial basis for the calculation of

LTC products. From a cohort of almost one hundred thousand long-term care patients of the statutory health insurance “Deutsche BKK”, the authors investigated the probabilities for a change or loss of care level and death with respect to age, gender and the duration of care in each level. Unfortunately, applicable results are not yet publicly available. The provision of such an age- and duration-depending actuarial basis would allow us to apply our model to LTC calculations.

We turn to the literature that is concerned with durational effects in life insurance. At first, we introduce two approaches of taking durational effects into account within a Markov set-up, meaning that the pure jump process describing the development of a single risk is assumed to be Markovian. Both approaches were exemplified by Haberman and Pitacco ([1999], section 1.7 and section 3.2). The first approach is given by the so-called splitting of states. Instead of modelling a single risk by employing a semi-Markovian pure jump process $(X_t)_{t\geq 0}$ along with the bivariate Markov process $((X_t, U_t))_{t\geq 0}$, the set of possible states of the policy will be redefined, in such a way that the states (or some states) also take into account information concerning the duration of presence (see Haberman and Pitacco [1999], section 1.7). Regarding PHI models, for example, the distinction of both recovery rates and mortality rates with respect to the time since disablement can also be achieved by replacing the state *invalid* by the $r+1$ states $i_0, i_1, ..., i_r$, $r \in \mathbb{N}$. Thus, the state i_0 stands for insured who have just recently become disabled and the appertaining transition rates correspond to the first column of an appropriate select-and-ultimate table. The states $i_1, ..., i_{r-1}$ must be interpreted in a similar way. In state i_r, the disablement dates back r or more years.

The disadvantages of the splitting-of-states method are, one the one hand, that it can only be used for discrete durational effects. On the other hand, the model becomes extremely complex and not very clearly arranged. Consider the previously mentioned PHI model with set of states and set of transitions according to figure 3. Assuming a select period of five years, the state space of the corresponding model according to the splitting-of-states method would consist of eight different states instead of three. The appertaining set of transitions would be given as

$$\begin{aligned}\mathcal{J} \;:=\; &\{(a,i_0),(a,d),(i_0,i_1),(i_0,d),(i_0,a),(i_1,i_2),(i_1,d),(i_1,a),(i_2,i_3),(i_2,d),(i_2,a),\\ &(i_3,i_4),(i_3,d),(i_3,a),(i_4,i_5),(i_4,d),(i_4,a),(i_5,a),(i_5,d)\}\,.\end{aligned}$$

Figure 7 displays the set of states and the corresponding transitions for the Markov model according to the splitting-of-states method (cf. figure 3). Our model, to be introduced later on, principally contains this special case of a Markov model with a finite state space.

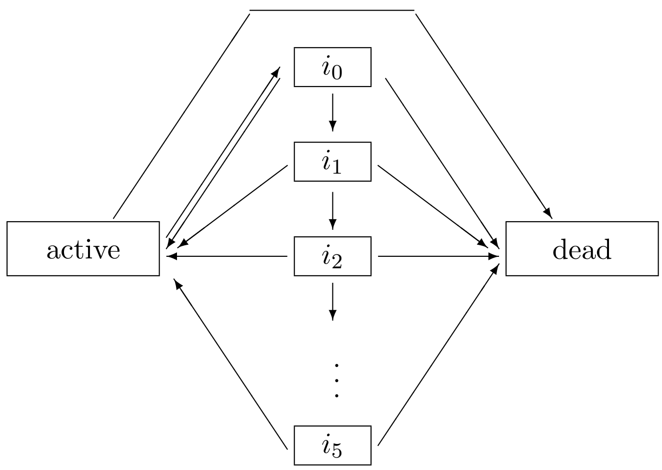

Figure 7: States and transitions for a PHI policy according to the splitting-of-states method

The second approach of taking durational effects into account within a Markov set-up is concerned with actuarial payments. So far, we have only discussed durational effects with respect to the development of a single policy. But actuarial payments can also be designed such that they depend not only on the current state of a policy, but also on the (random) time since entering this state. For example, a PHI policy often includes a deferred period so that the benefit will not be payable until the disability has lasted a certain minimum period. Other policies provide increasing benefits in order to protect the policyholder from the effects of inflation. Another feature in policy design is a benefit level that goes down with duration of sickness to encourage the insured to return to work. For a short introduction of the implementation of duration-depending actuarial payments due to certain policy specifications (e.g. deferred periods) in a Markov model, we follow Haberman and Pitacco ([1999], section 3.2) (see also Pitacco [1995]). This approach relies on a combination of transition probabilities with information concerning whether or not a benefit is intended. We will see that this way of implementing certain policy specifications leads to very complex transition probabilities, especially if different policy specifications are simultaneously involved. Yet, the major drawback of this approach is that it focuses on transition probabilities. For actuarial modelling, however, the natural starting point is formed by (cumulative) transition intensities. Obtaining corresponding transition probabilities from given (cumulative) transition intensities is usually a complicated task, which can be avoided by employing Thiele's integral or differential equations for the prospective reserve. Taking this into consideration, the method introduced below seems not to have a wide range of applications. But it should be briefly explained here, because this concept has also been dealt with by several other authors (e.g. Möller and Zwiesler [1996], Wetzel [2002], Wetzel and Zwiesler [2003]). Some of them have partially adapted it to smooth semi-Markov models.

Let $(X_t)_{t\geq 0}$ be a Markovian pure jump process modelling a single policy with set of states and set of transitions according to figure 3. Further, let

$$\phi(x,t) := P(X_{x+t} = i |\, X_x = a), \;\; x,t \geq 0$$

be the conditional probability of an insured being in state *invalid* (i) at the age $x+t$, given that the insured is *active* (a) at age x. $\phi^{\Gamma}(x,t)$ is the probability that the insured (*active* at age x) is disabled and the benefits are payable at the attained age $x+t$. $\Gamma = [n_1, n_2, f, m, r]$ formally represents the policy specifications by means of five parameters:

- (n_1, n_2) denotes the insured period, meaning that benefits are payable if disability inception time belongs to this interval,
- f denotes the deferred period (from disability inception),
- m is the maximum number of years of annuity payment (from disability inception),
- and r is the stopping time of annuity payment (from policy issue).

Note that $\phi^{\Gamma}(x,t) \leq \phi(x,t) = \phi^{[0,\infty,0,\infty,\infty]}(x,t)$. In addition, the authors defined the probability

$${}_tp_x^{ai}(\tau) := P(X_{x+r} = i \;\text{ for all }\; r \in [t-\tau, t] |\, X_x = a), \;\; 0 \leq \tau \leq t.$$

Using this approach, a number of realistic policy specifications can be modelled. For instance, a deferred period of length $f \geq 0$ can be implemented by means of

$$\phi^{[0,\infty,f,\infty,\infty]}(x,t) \;\; = \;\; \begin{cases} 0, & t \leq f \\ {}_tp_x^{ai}(f), & t > f \end{cases} .$$

To model an n-year PHI policy that provides an annuity for a maximum of m years, the transition probability ${}_tp_x^{ai} = \phi(x,t)$ must be replaced by

$$\phi^{[0,n,0,m,n]}(x,t) \;\; = \;\; \begin{cases} {}_tp_x^{ai} - {}_tp_x^{ai}(m), & t < n \\ 0, & t \geq n \end{cases} .$$

As previously mentioned, actuarial values of both premiums and benefits must be derived for this approach by using transition probabilities. But these probabilities cannot be easily provided. Beside the fact that they must be derived from given transition intensities, the additional consideration of policy specifications concerning actuarial payments makes the calculation of corresponding transition probabilities cumbersome. In contrast, a model that allows actuarial payments to depend on the time elapsed since entering the current state provides a well-arranged opportunity to implement policy specifications according to $\Gamma = [n_1, n_2, f, m, r]$, regardless of whether the process $(X_t)_{t\geq 0}$ is Markovian or not. Additionally, a wide range of payment dynamics can be modelled which cannot be formalized with the aid of Γ.

We now come to semi-Markov approaches. A life insurance model relying on semi-Markov processes was introduced by Hoem [1972]. Three years after expanding on the Markov model for life insurance (cf. Hoem [1969]), Hoem outlined basic characteristics of the application of inhomogeneous semi-Markov processes to that field. In doing so, the bivariate Markov process $((X_t, U_t))_{t\geq 0}$ was investigated. Among other things, forward and backward equations for the appertaining transition probabilities were sketched. They relate the duration-depending intensities to the corresponding transition probabilities of the process $((X_t, U_t))_{t\geq 0}$. In addition, Hoem [1972] provided basic formulas to derive actuarial values of future payments which are also allowed to depend on the previous state duration.

Møller [1993] established stochastic versions of Thiele's differential equations for the prospective reserve in a semi-Markov model. These equations also hold in a framework for which interest is likewise modelled by random processes. All processes, however, were assumed to be of locally bounded variation. Similarly to Hoem [1972], actuarial payments were allowed to depend on the time previously spent in the current state. Further, Møller [1993] provided an integral representation of the prospective loss, established by using tools from the theory of point processes. This issue will be raised again in the sequel.

Several other authors were concerned with semi-Markov models, particularly for PHI. Most of them have already been mentioned. Pitacco [1995] as well as Haberman and Pitacco [1999] also investigated models for LTC and dread disease (DD) insurance. As explained above, they introduced a general multiple state model allowing the implementation of various policy specifications. Further, they gave an overview of different PHI models and practical applications in different countries. Thus, Pitacco [1995] introduced the Norwegian model, the Manchester-Unity approach, and the Swedish model. An introduction of the Danish model, the so-called CMIB model (CMIB: **C**ontinuous **M**ortality **I**nvestigation **B**ureau), and the Finnish model can be found in Haberman and Pitacco [1999]. However, only the CMIB model, the Swedish model, and the Finnish model actually allow duration-depending probabilities for certain transitions. Segerer [1993] presented the corresponding models for Austria, Germany and Switzerland. While most of the models mentioned above are smooth models relying on analytical laws of mortality, disability, or recovery, the models introduced by Segerer are purely discrete models, since in the corresponding countries, it is usual to use tables of annual rates of mortality, etc.

Möller and Zwiesler [1996], Wetzel [2002], and Wetzel and Zwiesler [2003] advocated continuous modelling with subsequent discretization of the formulas. Particularly, Möller and Zwiesler [1996] adapted the ideas of Pitacco [1995] to German needs, in such a way that the model described by Segerer [1993] is principally contained. Special attention was paid to deferred periods. Wetzel [2002] as well as Wetzel and Zwiesler [2003] were basically concerned with establishing equations for specified PHI products, e.g. a deferred annuity with a return-of-premium guarantee over the deferral period combined with an occupational disability rider. In addition to the use of Pitacco's [1995] concept of implementing certain policy specifications, some actuarial payments were also allowed to depend on the previous duration in the current state. Doing so, Wetzel and Zwiesler [2003] defined the disability annuity rate depending on both the time of disability inception and the time elapsed since then. However, in order to discretize the resulting formulas (cf. Wetzel and Zwiesler [2003], section 5), the annuity rate was assumed to be constant.

Another author that should be mentioned here, since he provided results which were used in some of the previously mentioned papers, is Waters [1989] (see also Waters [1984]). He investigated the CMIB model in particular.

With exception of Segerer [1993], all authors mentioned above dealt with smooth models. This means that they assumed the existence of intensities for each of the three usual model components, i.e. transition probabilities, actuarial payments and discounting function. Discretization of results leads to purely discrete models. A more realistic model, however, should comprise both the continuous method as well as the discrete method, because events due to biometrical effects (mortality, disability, etc.) normally appear in continuous time. Events due to administrational effects, in turn, appear in discrete time (actuarial payments, retirement, etc.).

Stracke [1997] as well as Milbrodt and Stracke [1997] (see also Milbrodt and Helbig [1999], sections 4B, 10A, and 10C) introduced a generalized Markov model for life insurance that covers both the discrete method and the continuous method. Concerning a single policy with set of states $\mathcal{S}$, they employed Markovian pure jump processes with cumulative transition intensities q_{yz}, $(y, z) \in \mathcal{S}^2$. Further, accumulated actuarial payments and, regarding interest, cumulative interest intensities were considered. Two systems of integral equations for the prospective reserve were established, referred to as Thiele's integral equations of type 1 and of type 2. These systems of integral equations turned out to be equivalent to the corresponding types of backward integral equations for the transition probabilities of the Markov process $(X_t)_{t\geq 0}$. Further, Stracke ([1997], chapter 5) considered the possibility of relaxing the Markov property of the pure jump process $(X_t)_{t\geq 0}$. By requiring only that the appertaining marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$ is a homogeneous Markov chain, she pointed out that Thiele's integral equations of type 1 as well as the backward integral equations of type 1 remain valid in this generalized set-up (see Stracke [1997], Satz 5.6, Korollar 5.7, and Bemerkung 5.8). In contrast, Thiele's integral equations of type 2 and the backward integral equations of type 2 do not hold within this framework. These systems of integral equations require the existence of a regular cumulative transition intensity matrix $q = (q_{yz})_{(y,z)\in\mathcal{S}^2}$ for the pure jump process $(X_t)_{t\geq 0}$ (cf. Stracke [1997], Beispiel 5.9). For this, however, the Markov property of $(X_t)_{t\geq 0}$ is essential.

Driven by the ideas of Stracke [1997] as well as Milbrodt and Stracke [1997], a generalized semi-Markov model for life insurance will be presented here. The concept of regular cumulative transition intensity matrices is adapted to the bivariate Markov process $((X_t, U_t))_{t\geq 0}$. Further, smoothness assumptions for all other model parameters can be avoided and, in principle, all transition probabilities and actuarial payments are allowed to depend on both the current state of a policy and the previous duration in that state. In order to focus on durational effects, interest is assumed to be non-random. Due to the generality of our approach, corresponding Markov models - including the non-smooth Markov model provided by Milbrodt and Stracke - are contained. Further, models that incorporate durational effects - as the above mentioned smooth semi-Markov models and Markov models according to the splitting-of-states method - are also covered.

This thesis is organized as follows. In chapter 2, classes of processes used to model a single risk are considered. Thus, the pure jump process $(X_t)_{t\geq 0}$, the marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$, and the bivariate process $((X_t, U_t))_{t\geq 0}$ are investigated in more detail. Relying on the Markov property of the marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$, it is stated under which circumstances the appertaining pure jump process is Markovian. Further, implications of the Markov properties of both the marked point process and the bivariate process $((X_t, U_t))_{t\geq 0}$ are demonstrated. The investigation of the bivariate Markov process $((X_t, U_t))_{t\geq 0}$ is structured in a manner similar to the investigation of the Markovian pure jump process in Milbrodt and Helbig ([1999], section 4B and 4C). Starting from a multiple decrement approach, cumulative transition intensities are defined (definition 2. 24). Further, systems of backward and forward integral equations are established that relate the cumulative transition intensities to the transition probabilities of the Markov process $((X_t, U_t))_{t\geq 0}$ (section 2.D.1).

Regarding the bivariate process $((X_t, U_t))_{t\geq 0}$, actuarial payments due to transitions as well as sojourn payments are defined in chapter 3.

Chapter 4 is concerned with the prospective reserve. The prospective reserve at a certain time is specified as conditional expectation of future payments given the history of the process. Due to the Markov property of $((X_t, U_t))_{t\geq 0}$, the prospective reserve depends on the current state of the policy and the time elapsed since entering this state. Note that within a Markov set-up, the prospective reserve at a certain time would only depend on the current state of the policy. In order to derive prospective reserves, Thiele's integral equations of type 1 are established (theorem 4. 8). Solving this system of integral equations within a semi-Markov framework can be split up into two steps. Firstly, the duration-dependence can be disregarded and a system of integral equations must be solved which basically corresponds to Thiele's integral equations of type 1 in a non-smooth Markov set-up. Hence, establishing the uniqueness of solutions of this system of integral equations can be done in almost the same manner as shown by Milbrodt and Stracke [1997]. Secondly, for each current state of the policy and each previous duration in that state, the prospective reserve can be derived by computing an integral that contains the solution provided by the first step. Hence, the mathematical methods to obtain prospective reserves in a semi-Markov framework are principally the same as for the generalized Markov model. Regarding the derivation of prospective reserves, a numerical example dealing with PHI contracts is given by example 4. 14.

Similar to the Markov model by Milbrodt and Stracke [1997], there are two different systems of integral equations for the prospective reserve. Further, these systems of integral equations likewise turn out to be equivalent to the corresponding types of backward integral equations for the underlying Markov process. Milbrodt and Stracke [1997] also pointed out that Thiele's integral equations of type 1 imply Thiele's integral equations of type 2 and vice versa. This relationship also remains valid in a semi-Markov set-up.

Thiele's integral equations of type 2 are also equivalent to a certain integral representation of the prospective loss. This integral representation is used to derive the variance of the prospective loss according to Hattendorff's theorem, namely by calculating the expectations of the predictable variation of losses for certain time periods and certain states. Norberg [1992] established a version of Hattendorff's theorem stating fairly generally that the variance of the prospective loss can be derived this way. However, to compute the corresponding expected values, sufficient structure must be added to the model. This is usually done by assuming the pure jump process $(X_t)_{t\geq 0}$ to be Markovian. Relaxing the Markov assumption makes the computation of variances cumbersome. According to Norberg [1992], the calculation of variances merits attention particularly in a semi-Markov set-up:

> "An interesting issue is to study computational problems under non-Markov assumptions, e.g. when the transition intensities are allowed to depend on the duration of the period that has elapsed since the policy entered the current state. This would complicate matters immensely since integrations would have to be performed over the times of transitions." (Norberg [1992], section 3F)

As pointed out later on, these problems can be avoided by applying the same ideas yielding Thiele's integral equations of type 1 for the prospective reserve. Doing so, integral equations for conditional expectations of the predictable variation of prospective losses are established (theorem 4. 31). They can be solved in almost the same manner as Thiele's integral equations of type 1.

It must be mentioned here that the use of variances as measures of risk is often criticized, mainly for two reasons. The first reason is that the variance disregards the opportunities of financial markets to hedge investment risks. In the present investigation, however, we focus on durational effects and abstain from modelling interest randomly. Thus, this argument can be countered. In doing further research, however, the presented model should be embedded in a framework that does not ignore the surrounding financial markets. The second reason is that the variance is a symmetric measure of risk, meaning that positive and negative deviations from the mean are taken into account equally. Regarding the loss due to a policy, however, only a positive loss actually includes a risk for the insurer. But in spite of these objections, the variance has the advantage of being analytically computable, without knowledge of the distribution of the prospective loss. In contrast, more modern risk measures as the 'positive semi-variance', the 'value-at-risk', or the 'expected shortfall', etc. are often only available on a basis of simulations which are not enforced here. Hence, in order to compare the risk situation of policies within a semi-Markov set-up with the situation of corresponding policies within a Markov set-up, variance-based risk measures are used. Example 4. 32 illustrates this for the PHI policies considered in example 4. 14.

In chapter 5, retrospective reserves are considered. As mentioned above, the integral equations for the prospective reserve generalize the backward integral equations for the transition probabilities of the underlying Markov process. These transition probabilities also satisfy, under certain assumptions, two systems of forward integral equations. This raises the question of whether there are generalizations with a meaningful interpretation in view of a single policy. For a smooth Markov model, Norberg [1991] answered that question positively and illustrated that the forward differential equations for the transition probabilities correspond for a certain concept of the retrospective reserve to the differential equations for this reserve. Due to the lack of a non-smooth theory of retrospective reserves in a Markov set-up as well as in a semi-Markov set-up, we adapt Norberg's definition and outline some characteristics of the retrospective reserve. Doing so, we derive retrospective reserves in a Markov and in a semi-Markov set-up. Within the former, systems of integral equations will be presented that correspond to either Thiele's integral equations of type 1 or to Thiele's integral equations of type 2 for the prospective reserve. These systems of integral equations generalize both types of forward integral equations. In a semi-Markov set-up, such systems of integral equations can also be established. Yet in contrast to the integral equations for the prospective reserve in a semi-Markov framework, the additional integration over the times of transitions cannot be avoided. This is already the case when comparing the backward and forward integral equations for the transition probabilities of the bivariate Markov process $((X_t, U_t))_{t \geq 0}$ (cf. lemma 2. 32 and lemma 2. 36).

Before turning to classes of processes modelling single risks, some remarks concerning the notation of variables with actuarial meaning should be given. On the one hand, there is the actuarial standard, the so-called *Hamza notation* (cf. Haberman and E. Pitacco [1999], section 3.2.1, or Milbrodt and Helbig [1999], section 1C), which is often used by authors who are concerned with more practical actuarial mathematics. On the other hand, for more theoretical actuarial issues, the notation is often originated from probability theory. Here, the actuarial standard notation will mainly be used for numerical examples based on decrement tables. In all other respects, it seems to be more convenient to omit this standard. Nevertheless, our results can be translated into the *Hamza notation* by replacing, for example, the following expressions (cf. (2. 21. 6) and definition 2. 24):

$$p_{yz}(s, s+t, u, v) \leftrightarrow {}_{v,t}p^{yz}_{x+s,u}, \tag{1.0.1}$$

$$p_{yz}(s, r, u, v) \leftrightarrow {}_{v,r-s}p^{yz}_{x+s,u}, \tag{1.0.2}$$

$$\bar{p}_y(s, r, u) \leftrightarrow {}_{r-s}p^{\underline{yy}}_{x+s,u}, \tag{1.0.3}$$

and,

$$\mu_{yz}(t, u) \leftrightarrow \mu^{yz}_{x+t,u}. \tag{1.0.4}$$

Chapter 2

Modelling a single risk

A Classes of processes

Our starting point is a pure jump process $(\Omega, \mathfrak{F}, P, (X_t)_{t\geq 0})$ with finite state space $\mathcal{S}$ and right continuous paths with left-hand limits (càdlàg: **c**ontinue **à d**roite avec des **l**imites **à g**auche). This process describes the development of a single policy in continuous time. For each time $t \geq 0$ after policy issue, X_t is interpreted as the current state of the policy. Following Milbrodt and Stracke [1997] (see also Milbrodt and Helbig [1999], section 4A), the paths of $(X_t)_{t\geq 0}$ are given by pattern of states.

2. 1 Definition. A *pattern of states* is a right continuous map $[0,\infty) \ni t \mapsto x_t \in \mathcal{S}$ with at least one jump and at most finitely many jumps on bounded intervals. Further, let $\mathcal{X} \subset \mathcal{S}^{[0,\infty)}$ denote the set of all patterns of states and $\mathcal{V} \subset \mathcal{X}$ the set of all possible patterns of states. The state space $\mathcal{S}$ will be equipped with the σ-field $2^{\mathcal{S}}$. $\mathcal{X}$ will be equiped with the σ-field generated by the coordinate projections $pr_t : \mathcal{X} \ni x \mapsto x_t \in \mathcal{S}, t \geq 0$:

$$\mathfrak{X} := \mathcal{X} \cap (2^{\mathcal{S}})^{[0,\infty)} = a_\sigma(pr_t |\, t \geq 0).$$

Hence, a filtration on $\mathcal{X}$ is given by

$$\mathfrak{X}_t := a_\sigma(pr_s |\, 0 \leq s \leq t), \;\; t \geq 0.$$

To define a σ-field on $\mathcal{V}$, we set $\mathfrak{V} := \mathcal{V} \cap \mathfrak{X}$. $\mathfrak{V}_t := \mathcal{V} \cap \mathfrak{X}_t, t \geq 0$ is the corresponding filtration.

The set of possible patterns of states $\mathcal{V}$ is defined by imposing the initial state, possible transitions, possible final states, etc. on the policy development. An alternative description of the development of a single policy is given by random patterns of transitions.

2. 2 Definition. Let $\mathcal{J} := \{(y,z) \in \mathcal{S}^2 |\, y \neq z\}$ denote the *transition space*. A *pattern of transitions* is a non-vanishing right continuous map $[0,\infty) \ni t \mapsto (n_{yz,t})_{(y,z)\in\mathcal{J}} \in \mathbb{N}_0^{\mathcal{J}}$ with:

- $n_{yz,0} = 0$ for $(y,z) \in \mathcal{J}$.
- Each component $t \mapsto n_{yz,t}$ has at most finitely many jumps on bounded intervals.
- For $t \geq 0$ and $(y_1,z_1) \neq (y_2,z_2)$ the following holds:

$$\Delta\, n_{y_1z_1,t} \in \{0,1\} \quad \text{and} \quad \Delta\, n_{y_1z_1,t} \cdot \Delta\, n_{y_2z_2,t} = 0.$$

- Let $s < t$ and $(y_i,z_i) \in \mathcal{J}$ with $z_1 \neq y_2$ as well as $\Delta\, n_{y_1z_1,s} \cdot \Delta\, n_{y_2z_2,t} = 1$. Then there is an $r \in (s,t)$ and a pair $(y_3,z_3) \in \mathcal{J}$ satisfying $\Delta\, n_{y_3z_3,r} = 1$.

$\mathcal{N} \subset (\mathbb{N}_0^{\mathcal{J}})^{[0,\infty)}$ denotes the set of all patterns of transitions, equipped with the σ-field

$$\mathfrak{N} := \mathcal{N} \cap (2^{(\mathbb{N}_0^{\mathcal{J}})})^{[0,\infty)} = a_\sigma(n \mapsto n_t |\, t \geq 0).$$

A Filtration on $\mathcal{N}$ is given by $\mathfrak{N}_t := a_\sigma(n \mapsto n_s |\, 0 \leq s \leq t)$, $t \geq 0$.

A pattern of states as well as the corresponding pattern of transitions can be reproduced by recording the sequence of jump times and jump marks, with the latter corresponding to the attained states.

2. 3 Definition. A *chain of jumps* is a sequence $((t_m, z_m))_{m \in \mathbb{N}_0} \subset [0, \infty] \times \mathcal{S}$ such that

- $t_m < \infty \Longleftrightarrow z_m \neq z_{m-1}$,
- $0 = t_0 \leq t_1 \leq \ldots \nearrow \infty;\ t_1 < \infty$,
- $t_m < t_{m+1}$ if $t_m < \infty$.

Let $\mathcal{K}$ denote the set of all chains of jumps, equipped with the σ-field

$$\mathfrak{K} := \mathcal{K} \cap (\mathfrak{B}([0, \infty]) \otimes 2^{\mathcal{S}})^{\mathbb{N}_0}.$$

A filtration on $\mathcal{K}$ is given by the coordinate projections, i.e. $\mathfrak{K}_m := a_\sigma(pr_0, pr_1, \ldots, pr_m), m \in \mathbb{N}_0$.

According to Milbrodt and Helbig ([1999], Satz 4.8 and Satz 4.12), the measurable spaces $(\mathcal{X}, \mathfrak{X})$, $(\mathcal{N}, \mathfrak{N})$, and $(\mathcal{K}, \mathfrak{K})$ are isomorph, meaning that there exist isomorphisms

$$H = (H_{yz})_{(y,z) \in \mathcal{J}} : \mathcal{X} \to \mathcal{N} \quad \text{with} \quad H_{yz,t}(x) := \sum_{s \leq t} \mathbf{1}_{\{x_{s-0}=y,\, x_s = z\}}, \quad t \geq 0, \qquad (2.3.1)$$

and

$$G = (G_m)_{m \in \mathbb{N}_0} : \mathcal{X} \to \mathcal{K} \quad \text{with} \quad G_m(x) := (t_m(x), x_{t_m(x)}), \quad m \in \mathbb{N}_0. \qquad (2.3.2)$$

$t_m(x)$ denotes the time of the m-th jump of the pattern of states $x \in \mathcal{X}$, and $x_{t_m(x)}$ is the state entered by that jump. For a given pattern of states $x \in \mathcal{X}$, the jump times are defined recursively by

$$t_0(x) := 0 \quad \text{and} \quad t_m(x) := \min\{t > t_{m-1}(x) |\, x_t \neq x_{t_{m-1}(x)}\}, \ m \in \mathbb{N}. \qquad (2.3.3)$$

Note that they are stopping times with respect to $(\mathfrak{X}_t)_{t \geq 0}$. Further, the mapping $x \mapsto x_{t_m(x)}$ is $\mathfrak{X}_{t_m}$-measurable. The inverse function of G is given by

$$G^{-1}\big(((t_m, z_m))_{m \in \mathbb{N}_0}\big)_t = z_m, \quad t \in [t_m, t_{m+1}), \quad ((t_m, z_m))_{m \in \mathbb{N}_0} \in \mathcal{K}. \qquad (2.3.4)$$

For a given set $\mathcal{V} \subset \mathcal{X}$ of possible patterns of states, there is a set $\mathcal{M} := H(\mathcal{V})$ of possible patterns of transitions with σ-field $\mathfrak{M} := \mathcal{M} \cap \mathfrak{N}$ and filtration $\mathfrak{M}_t := \mathcal{M} \cap \mathfrak{N}_t, t \geq 0$. Further, $\mathcal{T} := G(\mathcal{V})$ denotes the set of possible chains of jumps with σ-field $\mathfrak{T} := \mathcal{T} \cap \mathfrak{K}$.

In order to implement durational effects in our modelling, we will often additionally consider the pattern of previous durations that correspond to a given pattern of states. Let $x \in \mathcal{X} \subset \mathcal{S}^{[0,\infty)}$ be a pattern of states. The appertaining pattern of previous durations $u(x)$ is then specified by

$$u_t(x) = t - \min\{s \leq t |\, x_s = x_t\} = t - t_m(x), \quad t \in [t_m(x), t_{m+1}(x)). \qquad (2.3.5)$$

The mapping $t \mapsto u_t(x)$ is a real-valued right continuous piecewise linear function with $u_t(x) = 0$ for $t = t_m(x), m \in \mathbb{N}$. Let $\mathcal{U} \subset \mathbb{R}_{\geq 0}{}^{[0,\infty)}$ denote the set of all such patterns of previous durations, equipped with the σ-field $\mathfrak{U} := \mathcal{U} \cap (\mathfrak{B}([0,\infty)))^{[0,\infty)}$. Now consider the set

$$\mathcal{U}_\mathcal{X} := \{(x, u(x)) | \, x \in \mathcal{X}\}$$

containing all pairs of pattern of states and the appertaining pattern of previous durations, equipped with the σ-field $\mathfrak{U}_\mathcal{X} := \mathcal{U}_\mathcal{X} \cap \mathfrak{X} \otimes \mathfrak{U}$. In view of the set of all possible pattern of states, we analogously define $\mathcal{U}_\mathcal{V} := \{(x, u(x)) | \, x \in \mathcal{V}\}$ with σ-field $\mathfrak{U}_\mathcal{V} := \mathcal{U}_\mathcal{V} \cap \mathfrak{V} \otimes \mathfrak{U}$.

In section 3C, we will employ the fact that the measurable spaces $(\mathcal{U}_\mathcal{X}, \mathfrak{U}_\mathcal{X})$ and $(\mathcal{K}, \mathfrak{K})$ are also isomorph. In order to demonstrate this, we define similarly to (2. 3. 2)

$$\bar{G} = (\bar{G}_m)_{m \in \mathbb{N}_0} : \mathcal{U}_\mathcal{X} \to \mathcal{K} \text{ with } \bar{G}_m((x, u(x))) := (t_m(x), x_{t_m(x)}), \quad m \in \mathbb{N}_0. \tag{2. 3. 6}$$

The inverse of $\bar{G}$ is given by

$$\bar{G}^{-1}\big(((t_m, z_m))_{m \in \mathbb{N}_0}\big)_t = (z_m, t - t_m), \quad t \in [t_m, t_{m+1}), \quad ((t_m, z_m))_{m \in \mathbb{N}_0} \in \mathcal{K}. \tag{2. 3. 7}$$

2. 4 Lemma. *The mapping $\bar{G}$ according to* (2. 3. 6) *defines an isomorphism between the spaces* $(\mathcal{U}_\mathcal{X}, \mathfrak{U}_\mathcal{X})$ *and* $(\mathcal{K}, \mathfrak{K})$.

Proof. That $\bar{G}$ is a bijective mapping can be verified by realizing that $\bar{G}^{-1} \circ \bar{G} = Id_{\mathcal{U}_\mathcal{X}}$ as well as $\bar{G} \circ \bar{G}^{-1} = Id_\mathcal{K}$. In order to verify the measurability of both $\bar{G}$ and $\bar{G}^{-1}$, consider the mapping $\bar{G}$ as composition $\bar{G} = G \circ p_x$ of the projection mapping $p_x : \mathcal{U}_\mathcal{X} \to \mathcal{X}$ with $p_x((x, u(x))) = x$, and the isomorphism $G : \mathcal{X} \to \mathcal{K}$. Since the σ-field $\mathfrak{X} \otimes \mathfrak{U}$ is generated by the corresponding projection mappings, the projection p_x is $\mathfrak{U}_\mathcal{X} - \mathfrak{X}$-measurable. Hence, by employing the measurability of G, the composition $\bar{G} = G \circ p_x$ is $\mathfrak{U}_\mathcal{X} - \mathfrak{K}$-measurable. That the inverse mapping $\bar{G}^{-1} = p_x^{-1} \circ G^{-1}$ is $\mathfrak{K} - \mathfrak{U}_\mathcal{X}$-measurable likewise follows from the measurability of both components. □

According to the above definitions, to every single policy (p) corresponds a pure jump process $(\Omega, \mathfrak{F}, P, (X_t)_{t \geq 0})$ with state space $\mathcal{S}$ and paths belonging to $\mathcal{V} \subset \mathcal{X}$. This process can also be described by the associated marked point process $(\Omega, \mathfrak{F}, P, ((T_m, Z_m))_{m \in \mathbb{N}_0})$ with paths $m \mapsto (T_m(\omega), Z_m(\omega)), \omega \in \Omega$ in $\mathcal{T}$. This process can be obtained as $(T, Z) := G \circ X$. Note that according to (2. 3. 4), for all $\omega \in \{T_m \leq t < T_{m+1}\} \subseteq \Omega$

$$X_t(\omega) = G^{-1}\big((T_m(\omega), Z_m(\omega))_{m \in \mathbb{N}_0}\big)_t = Z_m(\omega), \tag{2. 4. 1}$$

and hence, for all $y \in \mathcal{S}$ and $t \geq 0$

$$\begin{aligned} \{X_t = y\} &= \cup_{m \in \mathbb{N}_0}(\{T_m \leq t < T_{m+1}\} \cap \{Z_m = y\}) \\ &= \{\exists m \in \mathbb{N}_0 : T_m \leq t < T_{m+1}, Z_m = y\}. \end{aligned} \tag{2. 4. 2}$$

For a marked point process $((T_m, Z_m))_{m \in \mathbb{N}_0}$, the point of explosion is defined as $T_\infty := \sup_{m \in \mathbb{N}_0} T_m$. The property of having paths in $\mathcal{K}$, however, implies that $T_\infty = \infty$ P-a.s. Such marked point processes are called nonexplosive.

Figure 8 illustrates a possible realization of the process $(X_t)_{t \geq 0}$ with a state space consisting of three states: $a \sim$ *active*, $i \sim$ *invalid*, and $d \sim$ *dead.* It can be observed that in order to reproduce the pattern of states, it is sufficient to record the jump times T_m and the corresponding attained states Z_m.

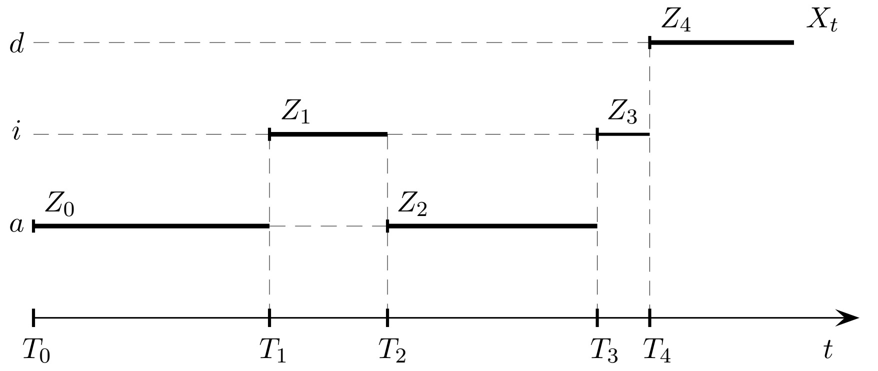

Figure 8: A POSSIBLE REALIZATION OF THE RANDOM PATTERN OF STATES $(X_t)_{t\geq 0}$

The third associated process is the multivariate counting process $(\Omega, \mathfrak{F}, P, (\mathbf{N}_t)_{t\geq 0})$ with paths in $\mathcal{M}$. This process is given as $\mathbf{N} := H \circ X$. Alternatively, it can be obtained with the aid of the marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$, namely by specifying each component of $(\mathbf{N}_t)_{t\geq 0}$ as

$$N_{yz,t} := \sum_{m\in\mathbb{N}_0} \mathbf{1}_{\{T_{m+1}\leq t,\, Z_m=y,\, Z_{m+1}=z\}}, \quad (y,z) \in \mathcal{J}, t \geq 0. \tag{2.4.3}$$

For the pattern of states displayed in figure 8, the component of the corresponding multivariate counting process $(N_{ai,t})_{t\geq 0}$ for the transitions from *active* to *invalid* is sketched in figure 9.

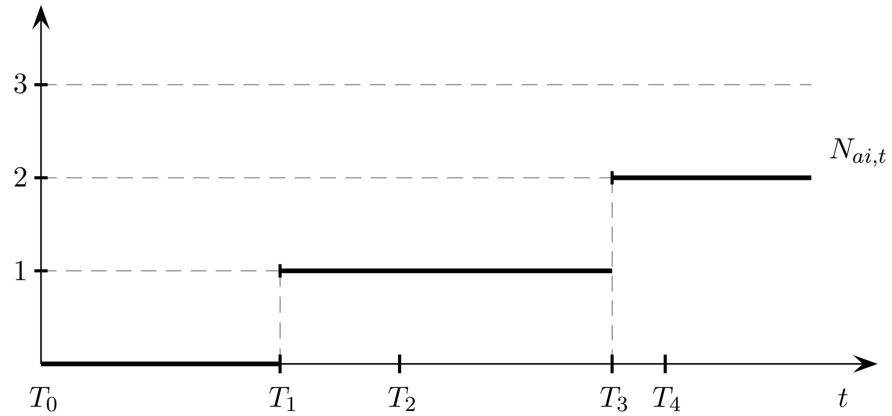

Figure 9: A POSSIBLE REALIZATION OF THE RANDOM PATTERN OF TRANSITIONS $(N_{ai,t})_{t\geq 0}$

According to (2.4.3), the random number of jumps up to time $t \geq 0$ is given by

$$N_t := \sum_{(y,z)\in\mathcal{J}} N_{yz,t} = \sum_{m\in\mathbb{N}_0} \mathbf{1}_{\{T_{m+1}\leq t\}}. \tag{2.4.4}$$

Using this, one gets by (2.4.2) for $t \geq 0$

$$X_t(\omega) = Z_{N_t(\omega)}(\omega), \quad \omega \in \Omega \tag{2.4.5}$$

and

$$\{X_t = y, N_t = k\} = \{T_k \leq t < T_{k+1}, Z_k = y\}. \tag{2.4.6}$$

Now let $(\mathfrak{F}_t)_{t\geq 0}$ denote the so-called point process filtration, i.e.

$$\mathfrak{F}_t := \mathfrak{F}_0 \vee a_\sigma((N_{yz,s})_{(y,z)\in\mathcal{J}} |\, 0 \leq s \leq t), \quad t \geq 0, \tag{2.4.7}$$

where $\mathfrak{F}_0$ contains some initial information. The following theorem gathers some results concerning measurability, which can be found e.g. in Brémaud ([1981], appendix A2).

2. 5 Theorem. *Let $(\Omega, \mathfrak{F}, P, ((T_m, Z_m))_{m\in\mathbb{N}_0})$ be a marked point process, $(\mathbf{N}_t)_{t\geq 0}$ the associated multivariate counting process, and $(\mathfrak{F}_t)_{t\geq 0}$ a filtration according to (2. 4. 7). Then:*

- *For all $m \in \mathbb{N}_0$, T_m is a stopping time with respect to $(\mathfrak{F}_t)_{t\geq 0}$.*
- *For all $m \in \mathbb{N}_0$, the σ-field $a_\sigma((T_i, Z_i)|\, 0 \leq i \leq m)$ coincides with the σ-field related to the past at time T_m, i.e. $\mathfrak{F}_{T_m} := \{A \in \mathfrak{F}|\, A \cap \{T_m \leq t\} \in \mathfrak{F}_t$ for all $t \geq 0\}$.*
- *For another $(\mathfrak{F}_t)_{t\geq 0}$-stopping time S, which is assumed to be finite, the following holds:*

$$\mathfrak{F}_S \cap \{T_m \leq S < T_{m+1}\} = \mathfrak{F}_{T_m} \cap \{T_m \leq S < T_{m+1}\}, \quad m \in \mathbb{N}_0, \tag{2. 5. 1}$$

  where $\mathfrak{F}_S$ and $\mathfrak{F}_{T_m}$ are the σ-fields related to the past at time S and the past at time T_m, respectively.
- *For a process $(R_t)_{t\geq 0}$ to be predictable with respect to $(\mathfrak{F}_t)_{t\geq 0}$, it is necessary and sufficient that it admits the representation*

$$R_t(\omega) \quad = \quad \sum_{m\in\mathbb{N}_0} f^{(m)}(t,\omega)\,\mathbf{1}_{\{T_m(\omega)<t\leq T_{m+1}(\omega)\}} + f^{(\infty)}(t,\omega)\,\mathbf{1}_{\{T_\infty(\omega)<t<\infty\}}, \tag{2. 5. 2}$$

  where for each $m \in \mathbb{N}_0$ the mapping $\Omega \times [0,\infty) \ni (\omega, t) \mapsto f^{(m)}(t,\omega)$ is $\mathfrak{F}_{T_m} \otimes \mathfrak{B}([0,\infty))$-measurable.

Concerning the relationship between the natural filtration of the appertaining pure jump process $(X_t)_{t\geq 0}$,

$$\mathfrak{F}_t^X := a_\sigma(X_s|\, 0 \leq s \leq t), \quad t \geq 0, \tag{2. 5. 3}$$

and the filtration $(\mathfrak{F}_t)_{t\geq 0}$ given by (2. 4. 7) with $\mathfrak{F}_0 := a_\sigma(X_0)$, it is provided by Last and Brandt ([1995], lemma 2.5.5) that for all $t \geq 0$,

$$\mathfrak{F}_t \subseteq \mathfrak{F}_t^X \quad \text{and} \quad \mathfrak{F}_t \cap \{t < T_\infty\} = \mathfrak{F}_t^X \cap \{t < T_\infty\}. \tag{2. 5. 4}$$

Hence, the assertions of the above theorem remain valid for the filtration $\mathfrak{F}_t^X$. Further, in cases where explosion is excluded, both natural filtrations coincide.

In view of a pure jump process $(X_t)_{t\geq 0}$ recording the present state of a policy (p) at every time after policy issue let

$$T(s) := \min\{t > s|\, X_t \neq X_s\} \tag{2. 5. 5}$$

be the time of the next jump after time $s \geq 0$, and

$$R(s) := \min\{t \leq s|\, X_t = X_s\} \tag{2. 5. 6}$$

the time of the last jump before s. Then,

$$W_s := T(s) - s \quad \text{and} \quad U_s := s - R(s) \tag{2. 5. 7}$$

are the sojourn times of $(X_t)_{t\geq 0}$ in the current state after time s and before time s, respectively. For the latter confer (2. 3. 5). In the case of $s \in [T_m, T_{m+1})$, $m \in \mathbb{N}_0$, one obtains

$$V_m := T_{m+1} - T_m = U_s + W_s. \tag{2. 5. 8}$$

According to (2. 5. 7) and (2. 4. 4), the time elapsed since entering the current state $U_t,\ t \geq 0$, can also be represented as

$$U_t = t - T_{N_t} = t - \sum_{m \in N_0} T_m \, \mathbf{1}_{\{T_m \leq t < T_{m+1}\}}. \tag{2. 5. 9}$$

Let $(\mathfrak{F}_t)_{t\geq 0}$ be a filtration with $(X_t)_{t\geq 0}$ being adapted to it. Due to (2. 5. 9), which can be written as

$$U_t = t - \lim_{n\to\infty} \sum_{m=0}^{n} T_m \, \mathbf{1}_{\{T_m \leq t\}} \, \mathbf{1}_{\{T_{m+1} > t\}},$$

and the fact that $T_m, m \in \mathbb{N}_0$, are stopping times with respect to $(\mathfrak{F}_t)_{t\geq 0}$ (cf. theorem 2. 5), it follows that for each $t \geq 0$ the time previously spent in the current state U_t is measurable with respect to $\mathfrak{F}_t$. Therefore, the process $(U_t)_{t\geq 0}$ is also adapted to $(\mathfrak{F}_t)_{t\geq 0}$ and - by definition - right continuous with left-hand limits (càdlàg).

We will later investigate the bivariate process $((X_t, U_t))_{t\geq 0}$ with state space $(\mathcal{S} \times [0, \infty), 2^{\mathcal{S}} \otimes \mathfrak{B}([0, \infty)))$. Figure 10 outlines for the realization of $(X_t)_{t\geq 0}$ illustrated in figure 8 the corresponding path of the bivariate process $((X_t, U_t))_{t\geq 0}$.

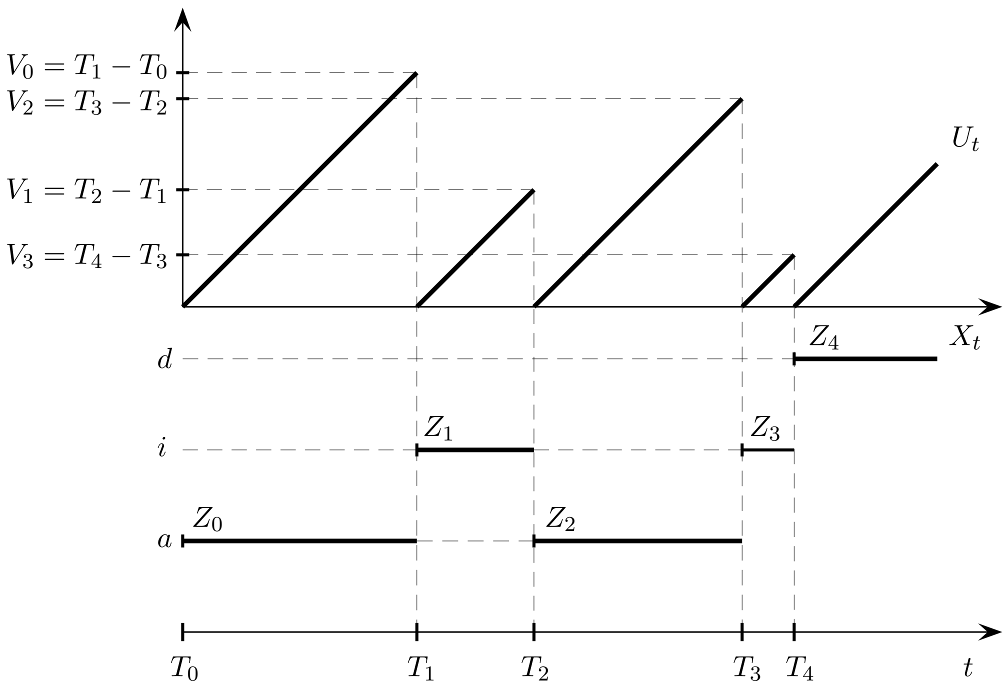

Figure 10: A POSSIBLE REALIZATION OF THE BIVARIATE PROCESS $((X_t, U_t))_{t\geq 0}$

The next section recalls some fundamentals of the theory of Markov processes, providing us with useful tools for future calculations. Afterwards we turn to the investigation of non-homogeneous marked point processes.

B Some general remarks on Markov processes

Without specifying the state space, some general properties of Markov processes will be outlined here. These issues can be found in almost every textbook dealing with random processes. For the most part, we follow the presentation by Milbrodt and Helbig ([1999], section 4B). Proofs are omitted.

Let $(\Omega, \mathfrak{F}, P)$ be a probability space, $(\mathfrak{F}_t)_{t\geq 0}$ a filtration, and $(X_t)_{t\geq 0}$ a random process with state space E, adapted to $(\mathfrak{F}_t)_{t\geq 0}$. The state space E is assumed to be polish and it is equipped with the Borel σ-field $\mathfrak{B}(E)$. Then, $(X_t, \mathfrak{F}_t)_{t\geq 0}$ is called a Markov process if for $0 \leq s \leq t$ and $B \in \mathfrak{B}(E)$

$$P(X_t \in B|\, \mathfrak{F}_s) = P(X_t \in B|\, X_s) \;\; P - a.s. \tag{2.5.10}$$

In the case of $(\mathfrak{F}_t)_{t\geq 0}$ being the natural filtration, it is often not explicitly indicated. The property (2.5.10) is usually called the elementary Markov property. It can be extended to a general statement about the future: If $(X_t, \mathfrak{F}_t)_{t\geq 0}$ is a Markov process and $s \geq 0$, then for each $A \in a_\sigma(X_t | t \geq s)$

$$P(A|\, \mathfrak{F}_s) = P(A|\, X_s) \;\; P - a.s. \tag{2.5.11}$$

Since the state space E is assumed to be polish, there always exists a version of the conditional distribution $P(X_t \in B|\, X_s = e)$, $0 \leq s \leq t, e \in E, B \in \mathfrak{B}(E)$ which is a kernel from E to $\mathfrak{B}(E)$. A family of such conditional distributions

$$p(s,t) : E \times \mathfrak{B}(E) \ni (e, B) \mapsto p_{eB}(s,t) \overset{\text{a.s.}}{=} P(X_t \in B|\, X_s = e), \quad 0 \leq s \leq t \tag{2.5.12}$$

is called a set of transition probabilities for the process $(X_t)_{t\geq 0}$. For all $0 \leq s \leq t$ and $B \in \mathfrak{B}(E)$, the transition probability $p_{\cdot B}(s,t)$ is uniquely determined up to $\mathcal{L}(X_s|\, P)$-exceptional sets.

A set of transition probabilities $(p(s,t))_{0\leq s\leq t<\infty}$, along with an initial distribution $\mathcal{L}(X_0|P)$, uniquely determines the distribution of the process: For $n \in \mathbb{N}_0$, $B_i \in \mathfrak{B}(E)$, $i = 0, \ldots, n$, and $0 = t_0 < \ldots < t_n < \infty$

$$\begin{aligned} &P(X_{t_0} \in B_0, \ldots, X_{t_n} \in B_n) \\ &= \int_{B_0}\int_{B_1} \cdots \int_{B_{n-1}} p_{e_{n-1}\, B_n}(t_{n-1}, t_n)\, p_{e_{n-2}\, de_{n-1}}(t_{n-2}, t_{n-1}) \ldots p_{e_0\, de_1}(t_0, t_1)\, \mathcal{L}(X_0|P)(de_0). \end{aligned}$$

The so-called Chapman-Kolmogorov equations form a special case of the above formula. They are given as follows: For $0 \leq s \leq r \leq t$ and $B \in \mathfrak{B}(E)$, the equation

$$p_{eB}(s,t) \;\; = \;\; \int_E p_{xB}(r,t)\, p_{e\,dx}(s,r) \tag{2.5.13}$$

is satisfied for $\mathcal{L}(X_s|P)-$a.e. $e \in E$. Further, it holds for $\mathcal{L}(X_s|P)-$a.e. $e \in E$

$$p_{eB}(s,s) = \varepsilon_e(B), \;\; s \geq 0, B \in \mathfrak{B}(E). \tag{2.5.14}$$

$\varepsilon_e(\cdot)$ denotes the Dirac measure at e, meaning for a set $B \in \mathfrak{B}(E)$: $\varepsilon_e(B) = 1$ if $e \in B$, and otherwise $\varepsilon_e(B) = 0$.

To ensure the existence of a distribution $\mathcal{L}((X_t)_{t\geq 0})$ - and with that the existence of a Markov process $(X_t, \mathfrak{F}_t)_{t\geq 0}$ - for a given initial distribution and a set of transition probabilities, a regular set of transition probabilities must exist. This is a set of transition probabilities which satisfy (2.5.13) and (2.5.14) identically. If the existence of such a set of transition probabilities is granted, the following theorem ensures the existence of a corresponding Markov process.

2.6 Theorem. *Let $(p(s,t))_{0\leq s\leq t<\infty}$ be a regular set of transition probabilities and $\pi|_{\mathfrak{B}(E)}$ a probability measure. Then, there is a Markov process $(\Omega, \mathfrak{F}, P, (X_t)_{t\geq 0})$ with initial distribution π and transition probabilities $(p(s,t))_{0\leq s\leq t<\infty}$. This process can be chosen as coordinate representation process, which is constructed in the following way: $(\Omega, \mathfrak{F}) := (E^{[0,\infty)}, \mathfrak{B}(E)^{[0,\infty)})$, $X_t := pr_t$, $t \geq 0$, and $P := \mathcal{L}((X_t)_{t\geq 0})$.*

A regular set of transition probabilities is not only used to get the distribution of the process, but also to manufacture versions of all important conditional probabilities, e.g. the distribution of the process $(X_t)_{t\geq s}$ starting out at the point $y \in E$ at time $t_0 = s$. Such a distribution will be denoted by P_{ys}. It can be constructed as follows: For $n \in \mathbb{N}_0$, $B_i \in \mathfrak{B}(E)$, $i = 0, \ldots, n$, and $s = t_0 < \ldots < t_n < \infty$

$$P_{ys}\left(X_s \in B_s, X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n, \prod_{t \notin \{s,t_1,\ldots t_n\}} X_t \in E\right)$$
$$= \int_{B_1} \cdots \int_{B_{n-1}} p_{e_{n-1}\,B_n}(t_{n-1}, t_n)\, p_{e_{n-2}\,de_{n-1}}(t_{n-2}, t_{n-1}) \ldots p_{y\,de_1}(s, t_1) \cdot \varepsilon_y(B_s).$$

Due to the use of a regular set of transition probabilities, P_{ys} is a kernel

$$E \times \mathfrak{B}(E)^{[0,\infty)} \ni (y, C) \mapsto P_{ys}(C) \in [0,1],$$

which forms a version of the conditional distribution of $(X_t)_{t\geq s}$ given $X_s = y$, uniquely determined up to $\mathcal{L}(X_s|P)$-exceptional sets. When considering such a conditional distribution, we will always choose the version provided by the corresponding kernel P_{ys}. Thus, for all measurable mappings

$$f : ([0,\infty) \times E^{[0,\infty)}, \mathfrak{B}([0,\infty)) \otimes \mathfrak{B}(E)^{[0,\infty)}) \to ([0,\infty), \mathfrak{B}([0,\infty))),$$

the conditional expectation is for $\mathcal{L}(X_s|P)-$a.e. $y \in E$ given as

$$\mathbf{E}\left[f(s,(X_t)_{t\geq s})|\, X_s = y\right] = \int_{E^{[0,\infty)}} f(s,\cdot)\, dP_{ys}. \tag{2.6.1}$$

Further, by means of the kernel P_{ys}, the Markov property (2. 5. 10) can for $t, \tau \geq 0$ and $B \in \mathfrak{B}(E)$ be rewritten as

$$P(X_{\tau+t} \in B|\, \mathfrak{F}_t) = p_{yB}(s, \tau + s)|_{s=t,y=X_t} = P_{ys}(X_{\tau+s} \in B)|_{s=t,y=X_t} \quad P - a.s. \tag{2.6.2}$$

In view of (2. 6. 1), this is equivalent to

$$\mathbf{E}\left[f(t,(X_\tau)_{\tau\geq t})|\, \mathfrak{F}_t\right] = \mathbf{E}\left[f(t,(X_\tau)_{\tau\geq t})|\, X_t\right] = \int_{E^{[0,\infty)}} f(s,\cdot)\, dP_{ys}|_{s=t,y=X_t} \quad P - a.s. \tag{2.6.3}$$

Now let $(\mathcal{W}, \mathfrak{W})$ be another measurable space and

$$I : (E^{[0,\infty)}, \mathfrak{B}(E)^{[0,\infty)}) \to (\mathcal{W}, \mathfrak{W})$$

a measurable mapping. Employing the theorem on integration with respect to an image measure, we state the following lemma that will be used in the sequel.

2. 7 Lemma. *In the situation which is described in this section, let the function* f *be given as* $f(s,x) := \hat{f}(s, I \circ x)$, $x \in E^{[0,\infty)}$, *where* $\hat{f}$ *is a mapping*

$$\hat{f} : ([0,\infty) \times \mathcal{W}, \mathfrak{B}([0,\infty)) \otimes \mathfrak{W}) \to ([0,\infty), \mathfrak{B}([0,\infty))).$$

Then, we obtain for the right-hand side of (2. 6. 1)

$$\int_{E^{[0,\infty)}} f(s,\cdot)\, dP_{ys} = \int_{E^{[0,\infty)}} \hat{f}(s, I(\cdot))\, dP_{ys} = \int_{\mathcal{W}} \hat{f}(s,\cdot)\, d\hat{P}_{ys} \tag{2.7.1}$$

with $\hat{P}_{ys}(B) := P_{ys}(I^{-1}(B))$, $B \in \mathfrak{W}$. *Further,* $\hat{P}_{ys}$ *is a kernel*

$$E \times \mathfrak{W} \ni (y, B) \mapsto \hat{P}_{ys}(B) \in [0,1].$$

Hence, regarding the left-hand side of (2.6.1)*, the following holds for* $\mathcal{L}(X_s|P)-$*a.e.* $y \in E$*:*

$$\mathbf{E}\left[f(s,(X_t)_{t\geq s})\right| X_s = y] = \int_{\mathcal{W}} \hat{f}(s,\cdot)\, d\hat{P}_{ys}. \tag{2.7.2}$$

We finish this section by stating some remarks concerning the strong Markov property of a Markov process $(X_t, \mathfrak{F}_t)_{t\geq 0}$. For this, we also refer to Milbrodt and Helbig ([1999], section 4B) and additionally to Gihman and Skorohod ([1975], chapter I, § 1). In principle, the strong Markov property means that the Markov property (cf. (2.6.2)) remains valid when a fixed time is replaced by a random one. Let T be such a random time, meaning that $T : \Omega \to [0,\infty]$ is a stopping time with respect to $(\mathfrak{F}_t)_{t\geq 0}$. Then the process $(X_t, \mathfrak{F}_t)_{t\geq 0}$ is called strong Markovian if

- it is progressively measurable,
- the transition probabilities $p_{eB}(s,t)$, $e \in B, 0 \leq s \leq t < \infty$ are - considered as functions of (e,s,t) - measurable with respect to $\mathfrak{B}(E) \otimes \mathfrak{B}([0,\infty)) \otimes \mathfrak{B}([0,\infty))$ for any $B \in \mathfrak{B}(E)$, and
- for $t, \tau \geq 0$ and $B \in \mathfrak{B}(E)$ the following equality is satisfied

$$P(X_{\tau+T} \in B|\, \mathfrak{F}_T) = p_{yB}(s,\tau+s)|_{s=T,y=X_T} \quad P-a.s., \tag{2.7.3}$$

  where $\mathfrak{F}_T$ is the σ-field related to the past at time T.

The first condition ensures that the random variable $X_T : \omega \mapsto X_{T(\omega)}(\omega)$ is $\mathfrak{F}_T - \mathfrak{B}(E)$-measurable. Together with the second condition, this yields that $p_{X_T B}(T, \tau+T)$ is $\mathfrak{F}_T$-measurable for any $B \in \mathfrak{B}(E)$. According to the third assertion, the conditional probabilities $P(X_{\tau+T} \in B|\, \mathfrak{F}_T)$ and $p_{X_T B}(T, \tau+T)$ coincide almost surely. Hence, $p_{X_T B}(T,\tau+T)$ is a version of the conditional probability $P(X_{\tau+T} \in B|\, \mathfrak{F}_T)$, since it is $\mathfrak{F}_T$-measurable and satisfies $\forall A \in \mathfrak{F}_T$ the Radon-Nikodym equation

$$P(\{X_{\tau+T} \in B\} \cap A) = \int_A P(X_{\tau+T} \in B|\, \mathfrak{F}_T)\, dP = \int_A p_{yB}(s,\tau+s)|_{s=T,y=X_T}\, dP. \tag{2.7.4}$$

As a generalization of the elementary Markov property which is equivalent to (2.6.2) and (2.6.3), the strong Markov property is equivalent to

$$P(X_{\tau+T} \in B|\, \mathfrak{F}_T) = p_{yB}(s,\tau+s)|_{s=T,y=X_T} = P_{ys}(X_{\tau+s} \in B)|_{s=T,y=X_T} \quad P-a.s., \tag{2.7.5}$$

and to

$$\mathbf{E}\left[f(T,(X_\tau)_{\tau\geq T})|\, \mathfrak{F}_T\right] = \int_{E^{[0,\infty)}} f(s,\cdot)\, dP_{ys}|_{s=T,y=X_T} \quad P-a.s. \tag{2.7.6}$$

Note that a process $(X_t, \mathfrak{F}_t)_{t\geq 0}$ having right continuous paths is progressively measurable. Further, a Markov process $(X_t, \mathfrak{F}_t)_{t\geq 0}$ is always strong Markovian for stopping times T with countable range, i.e. $P(T \in \{\infty, t_1, t_2, ...\}) = 1$. The latter can be shown by verifying the Radon-Nikodym equation (2.7.4): For $A \in \mathfrak{F}_T$ one obtains due to $A \cap \{T = t_i\} \in \mathfrak{F}_{t_i}$

$$\begin{aligned}
&P(\{X_{\tau+T} \in B\} \cap A \cap \{T < \infty\}) \\
&= \sum_{i\in\mathbb{N}} P(\{X_{\tau+T} \in B\} \cap A \cap \{T = t_i\}) = \sum_{i\in\mathbb{N}} \int_{A\cap\{T=t_i\}} P(X_{\tau+t_i} \in B|\, \mathfrak{F}_{t_i})\, dP \\
&= \sum_{i\in\mathbb{N}} \int_{A\cap\{T=t_i\}} p_{yB}(s,\tau+s)|_{s=t_i,y=X_{t_i}}\, dP = \sum_{i\in\mathbb{N}} \int_{A\cap\{T=t_i\}} p_{yB}(s,\tau+s)|_{s=T,y=X_T}\, dP \\
&= \int_A p_{yB}(s,\tau+s)|_{s=T,y=X_T}\, dP.
\end{aligned} \tag{2.7.7}$$

The strong Markov property will be used to clarify the implications of the Markov properties of both the marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$ and the appertaining bivariate process $((X_t, U_t))_{t\geq 0}$. For this we refer to section 2D. The following section is concerned with homogeneous marked point processes.

C Homogeneous Markovian marked point processes

Let $(\Omega, \mathfrak{F}, P, ((T_m, Z_m))_{m\in\mathbb{N}_0})$ be a marked point process with state space $([0,\infty]\times\mathcal{S}, \mathfrak{B}([0,\infty])\otimes 2^{\mathcal{S}})$. Since this state space is polish, there exists for each $m \in \mathbb{N}_0$ a regular version of the conditional distribution

$$P((T_{m+1}, Z_{m+1}) \in C|\,(T_i, Z_i)_{i\in\mathbb{N}_0, i\leq m}),\ \ C \in \mathfrak{B}([0,\infty])\otimes 2^{\mathcal{S}}.$$

When considering such a conditional distribution, we will always choose a regular version. Further, we assume for our entire investigation that the marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$ is a homogeneous Markov chain, meaning that for each $m \in \mathbb{N}_0$ and $C \in \mathfrak{B}([0,\infty])\otimes 2^{\mathcal{S}}$

$$P((T_{m+1}, Z_{m+1}) \in C|\,(T_i, Z_i)_{i\in\mathbb{N}_0, i\leq m}) = P((T_{m+1}, Z_{m+1}) \in C|\,T_m, Z_m)\ \ P-a.s., \qquad (2.7.8)$$

where the right-hand side can additionally be chosen as independent of m. In doing so, we make the following definition.

2.8 Definition. For $(y,z) \in \mathcal{J}$ and $s \geq 0$ we define

$$\hat{Q}_{yz}(s,\cdot):\ [s,\infty) \ni t \mapsto P(T_{m+1} \leq t, Z_{m+1} = z|\,T_m = s,\ Z_m = y), \qquad (2.8.1)$$

$$\hat{Q}_{y}(s,\cdot):\ [s,\infty) \ni t \mapsto \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \hat{Q}_{yz}(s,t) = P(T_{m+1} \leq t|\,T_m = s,\ Z_m = y), \qquad (2.8.2)$$

and, with the convention $0/0 := 0$,

$$\hat{q}_{yz}(s,\cdot):\ [s,\infty) \ni t \mapsto \int_{(s,t]} \frac{\hat{Q}_{yz}(s,d\tau)}{1-\hat{Q}_y(s,\tau-0)} \in [0,\infty], \qquad (2.8.3)$$

$$\hat{q}_{yy}(s,\cdot):\ [s,\infty) \ni t \mapsto -\int_{(s,t]} \frac{\hat{Q}_{y}(s,d\tau)}{1-\hat{Q}_y(s,\tau-0)} \in [-\infty,0]. \qquad (2.8.4)$$

$\hat{q}_{yz}$ is referred to as the *cumulative transition intensity from y to z* and $-\hat{q}_{yy}$ is the *cumulative intensity of decrement for state y*. The matrix

$$\hat{q}(s,t) := (\hat{q}_{yz}(s,t))_{(y,z)\in\mathcal{S}^2},\ \ 0 \leq s \leq t < \infty \qquad (2.8.5)$$

is called *cumulative transition intensity matrix for the marked point process* $((T_m, Z_m))_{m\in\mathbb{N}_0}$.

All quantities in definition 2.8 are uniquely determined up to $\mathcal{L}(T_m, Z_m|P)$-null sets. $\hat{Q}_{yz}(s,\cdot)$ and $\hat{Q}_y(s,\cdot)$ define for all $(s,y) \in [0,\infty)\times\mathcal{S}$ distribution functions of measures concentrated on (s,∞). $\hat{q}_{yz}(s,\cdot)$ and $-\hat{q}_{yy}(s,\cdot)$ define the appertaining hazard measures on $\mathfrak{B}((s,\infty))$.

2.9 Lemma. *Let $((T_m, Z_m))_{m\in\mathbb{N}_0}$ be a homogeneous Markovian marked point process with transition probabilities according to definition 2.8. Then, one obtains for $\mathcal{L}(T_m, Z_m|\,P)$-a.e. $(s,y) \in [0,\infty)\times\mathcal{S}$*

$$\frac{d\hat{Q}_{yz}}{d\hat{q}_{yz}}(s,\cdot) = (1-\hat{Q}_y(s,\cdot-0)),\ \ z\in\mathcal{S}, z\neq y, \qquad (2.9.1)$$

$$\frac{d\hat{Q}_y}{d\hat{q}_{yy}}(s,\cdot) = -(1-\hat{Q}_y(s,\cdot-0)), \tag{2.9.2}$$

and

$$\frac{d\hat{Q}_{yz}}{d\hat{Q}_y}(s,\cdot) = P(Z_{m+1}=z|\,T_{m+1}=\cdot, T_m=s,\,Z_m=y),\;\; z\in\mathcal{S}, z\neq y, \tag{2.9.3}$$

as well as

$$\frac{d\hat{q}_{yz}}{d\hat{q}_{yy}}(s,\cdot) = -P(Z_{m+1}=z|\,T_{m+1}=\cdot, T_m=s,\,Z_m=y),\;\; z\in\mathcal{S}, z\neq y. \tag{2.9.4}$$

Proof. (2.9.1) and (2.9.2) follow immediately from (2.8.3) and (2.8.4), respectively. In order to prove (2.9.3), consider for $B\in\mathfrak{B}([0,\infty))$ and $(y,z)\in\mathcal{J}$ the transition probability $\hat{Q}_{yz}(s,B)$. Upon successive conditioning and afterwards inserting (2.8.2), (2.9.3) follows by means of

$$\begin{aligned}\hat{Q}_{yz}(s,B) &= P(T_{m+1}\in B, Z_{m+1}=z|\,T_m=s,\,Z_m=y)\\ &= \int_B P(Z_{m+1}=z|\,T_{m+1}=v,T_m=s,\,Z_m=y)\,P(T_{m+1}\in dv|\,T_m=s,\,Z_m=y)\\ &= \int_B P(Z_{m+1}=z|\,T_{m+1}=v,T_m=s,\,Z_m=y)\,\hat{Q}_y(s,dv). \end{aligned}\tag{2.9.5}$$

With the aid of (2.9.1) and (2.9.2), (2.9.4) can be proved analogously. □

The following lemma gathers some properties of cumulative transition intensities according to definition 2.8. Most of them are properties of cumulative hazard rates in general.

2.10 Lemma. *For $s\geq 0$ and $(y,z)\in\mathcal{J}$ let $\hat{q}_{yz}(s,\cdot)$ and $\hat{q}_{yy}(s,\cdot)$ be cumulative intensities according to definition 2.8. Then for $\mathcal{L}(T_m,Z_m|\,P)$-a.e. $(s,y)\in[0,\infty)\times\mathcal{S}$*

$$\hat{q}_{yz}(s,r)\leq\hat{q}_{yz}(s,t),\;\; s\leq r\leq t, \tag{2.10.1}$$

$$\hat{q}_{yz}(s,s)=0, \tag{2.10.2}$$

$$\lim_{h\searrow 0}\hat{q}_{yz}(s,t+h)=\hat{q}_{yz}(s,t),\;\; s\leq t, \tag{2.10.3}$$

$$\hat{q}_{yz}(s,t)\geq 0,\;\; \textit{and}\;\; \hat{q}_{yy}(s,t)\leq 0,\;\; s\leq t, \tag{2.10.4}$$

$$-\hat{q}_{yy}(s,t)=\sum_{\substack{z\in\mathcal{S}\\ z\neq y}}\hat{q}_{yz}(s,t),\;\; s\leq t, \tag{2.10.5}$$

$$-\hat{q}_{yy}(s,\{t\})\leq 1,\; t\geq s,\;\; \textit{and}\;\; -\hat{q}_{yy}(s,\{t\})=1 \Longrightarrow -\hat{q}_{yy}(s,\tau)=-\hat{q}_{yy}(s,t),\;\; \tau\geq t\geq s. \tag{2.10.6}$$

Further, let $\omega_y := \inf\{t\geq s:\, \hat{Q}_y(s,(s,t])=1)\}$. Then

- *if $\omega_y=\infty$ and $\lim_{t\to\infty}\hat{Q}_y(s,t)=1$ then $\lim_{t\to\infty}-\hat{q}_{yy}(s,t)=\infty$,*
- *if $\omega_y<\infty$ and $\hat{Q}_y(s,\{\omega_y\})=0$ then $-\hat{q}_{yy}(s,(s,\omega_y))=-\hat{q}_{yy}(s,(s,\omega_y])=\infty$,*
- *if $\omega_y<\infty$ and $\hat{Q}_y(s,\{\omega_y\})>0$ then $-\hat{q}_{yy}(s,(s,\omega_y])<\infty$ and $-\hat{q}_{yy}(s,\{\omega_y\})=1$.*

Finally, for a given cumulative intensity of decrement $-\hat{q}_{yy}(s,\cdot)$, the so-called exponential formula determines the conditional distribution $\hat{Q}_y(s,\cdot)$ by means of

$$1-\hat{Q}_y(s,t)=\exp\{\hat{q}^{(c)}_{yy}(s,t)\}\prod_{s<\tau\leq t}(1+\Delta\hat{q}_{yy}(s,\tau)),\;\; t\geq s, \tag{2.10.7}$$

where the operators $^{(c)}$ *and* Δ *refer to the second variable and denote the continuous part and the discrete part of* $\hat{q}_{yy}(s,\cdot)$*, respectively.*

PROOF. (2.10.1), (2.10.2), (2.10.4), and (2.10.5) follow immediately from definition 2.8. (2.10.3) is a consequence of (2.10.5) and lemma A.4 in the appendix. The other assertions are consequences of general properties of hazard measures. For more details, Last and Brandt ([1995], appendix 5) or Milbrodt and Helbig ([1999], Folgerung 3.3) can be consulted. □

To avoid difficulties with null sets, Milbrodt and Stracke [1997] (see also Milbrodt and Helbig [1999], section 4B) have introduced the concept of regular transition intensity matrices. In the present situation, a substantial part of this would be that the assertions (2.10.1) - (2.10.6) are satisfied without exceptional sets. This, however, is granted by the choice of regular versions of the conditional distributions $\hat{Q}_*$, the existence of which follows from the polish state space of the marked point process. Thus, the functions $\hat{Q}_{yz}(s,\cdot)$, $z \in \mathcal{S}$ define for each $(s,y) \in [0,\infty) \times \mathcal{S}$ measures on (s,∞), and the cumulative transition intensities $\hat{q}_{yz}(s,\cdot)$ are the appertaining hazard measures. Hence, the properties (2.10.1) - (2.10.6) are satisfied for each $(s,y) \in [0,\infty) \times \mathcal{S}$. Further, Milbrodt and Stracke [1997] require a regular cumulative transition intensity matrix to be finite. Partially following them, we state the following definition.

2.11 Definition. A matrix-valued map

$$\hat{q} = (\hat{q}_{yz})_{(y,z)\in\mathcal{S}^2} : \{(s,t) \mid 0 \le s \le t < \infty\} \to \mathbb{R}^{\mathcal{S}^2} \tag{2.11.1}$$

possessing the properties (2.10.1) - (2.10.6) without exceptional sets is called a *regular cumulative transition intensity matrix for a marked point process* $((T_m, Z_m))_{m\in\mathbb{N}_0}$. By means of

$$\hat{q}_{yz}(s,(r,t]) := \hat{q}_{yz}(s,t) - \hat{q}_{yz}(s,r),\ s \le r \le t,$$

a regular cumulative transition intensity matrix defines for each $s \ge 0$ and $(y,z) \in \mathcal{S}^2$ a Borel measure concentrated on (s,∞).

Summarizing the above arguments, for a homogenous marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$ with cumulative transition intensities $(\hat{q}_{yz})_{(y,z)\in\mathcal{S}^2}$ being finite on bounded intervals, the existence of a regular cumulative transition intensity matrix according to definition 2.11 is granted.

Note that our understanding of regular cumulative transition intensity matrices is more general than the corresponding concept introduced by Milbrodt and Stracke [1997]. They require regular cumulative transition intensities to be additive, i.e.

$$\hat{q}_{yz}(s,r) + \hat{q}_{yz}(r,t) = \hat{q}_{yz}(s,t),\ \ 0 \le s \le r \le t, (y,z) \in \mathcal{J}. \tag{2.11.2}$$

In view of the general property of measures according to definition 2.11,

$$\hat{q}(s,(s,r]) + \hat{q}_{yz}(s,(r,t]) = \hat{q}_{yz}(s,(s,t]),$$

this means that for $s \le r$ and $(y,z) \in \mathcal{J}$ the measures defined by the distribution functions $\hat{q}_{yz}(s,\cdot)$ and $\hat{q}_{yz}(r,\cdot)$ coincide on (r,∞) (cf. the proof of lemma 2.19). As we will see later, the property (2.11.2) ensures that the appertaining pure jump process $X = G^{-1}(T,Z)$ is Markovian.

A regular cumulative transition intensity matrix $\hat{q}$ along with an initial distribution determines the distribution of a marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$. Moreover, for a given regular cumulative transition intensity matrix and an initial distribution, the existence of a corresponding homogeneous Markovian marked point process is granted.

2.12 Theorem. *Let* π *be a probability measure on* $2^{\mathcal{S}}$*,* ε_a *the Dirac measure at a, and* $\hat{q} = (\hat{q}_{yz})_{(y,z)\in\mathcal{S}^2}$ *a regular cumulative transition intensity matrix according to definition* 2.11*. Then:*

- *There is a homogeneous Markovian marked point process* $(\Omega, \mathfrak{F}, P, ((T_m, Z_m))_{m\in\mathbb{N}_0})$ *with initial distribution* $\mathcal{L}(T_0, Z_0|P) = \varepsilon_0 \otimes \pi$ *and cumulative transition intensity matrix* $\hat{q}$.
- *The distribution of a homogeneous Markovian marked point process* $((T_m, Z_m))_{m\in\mathbb{N}_0}$ *is uniquely determined by* π *and* $\hat{q}$.

Proof. For the first assertion, we refer to the proof of Satz 35 in Milbrodt and Helbig [1999]. Starting with an initial distribution and a regular cumulative transition intensity matrix, the authors verified the existence of a marked point process satisfying the requirements. Though the regular cumulative transition intensity matrix in the sense of Milbrodt and Helbig [1999] forms a special case of our concept according to definition 2. 11, the proof can be reproduced almost literally. The second assertion can likewise be obtained by reproducing the corresponding result in Milbrodt and Helbig. For this, Satz 4.34 must be consulted. □

After answering the question of existence, we now clarify whether or not a marked point process possesses paths in $\mathcal{K}$ (cf. definition 2. 3). To ensure this, the transition probabilities must satisfy some requirements. On the one hand, for each initial distribution $\mathcal{L}(T_0, Z_0|P) = \varepsilon_0 \otimes \pi$, the condition $P(T_1 < \infty) = 1$ is equivalent to

$$\hat{Q}_y(0, \infty) = \lim_{t\to\infty} \hat{Q}_y(0, t) = 1, \ \forall y \in \mathcal{S} \quad \text{satisfying} \quad \pi(y) > 0. \tag{2.12.1}$$

This makes sure that for almost every path of $(X_t)_{t\geq 0}$ at least one jump occurs. One the other hand, the marked point process must be nonexplosive, i.e. $P(T_\infty = \infty) = 1$. This is obviously equivalent to the fact that the number of jumps on bounded intervals is P-a.s. finite. In order to ensure this by verifying that the expected number of jumps up to time $t \geq 0$ is finite,

$$\mathbf{E}\Big[\sum_{(y,z)\in\mathcal{J}} N_{yz,t}\Big] < \infty, \tag{2.12.2}$$

the compensators of $(N_{yz,t})_{t\geq 0}$, $(y, z) \in \mathcal{J}$ are useful tools.

2. 13 Lemma. *Let* $((T_m, Z_m))_{m\in\mathbb{N}_0}$ *be a homogeneous Markovian marked point process with transition probabilities and transition intensities according to definition* 2. 8. *Then, each component of the appertaining multivariate counting process* $(\mathbf{N}_t)_{t\geq 0}$ *possesses with respect to a filtration* $(\mathfrak{F}_t)_{t\geq 0}$ *according to* (2. 4. 7) *the compensator*

$$A_{yz,t} \;=\; \sum_{m\in\mathbb{N}_0} \int_{(T_m\wedge t, T_{m+1}\wedge t]} \mathbf{1}_{\{Z_m=y\}}\, \hat{q}_{yz}(T_m, ds), \ \ t \geq 0, \tag{2.13.1}$$

which is P*-almost surely unique.*

Proof. Consider $N_{yz,t} = \sum_{m\in\mathbb{N}_0} \mathbf{1}_{\{T_{m+1}\leq t,\, Z_m=y, Z_{m+1}=z\}}$, $(y, z) \in \mathcal{J}, t \geq 0$ and let $G^{(m)}, m \in \mathbb{N}_0$ be a sequence of regular conditional distributions of (T_{m+1}, Z_m, Z_{m+1}) given $a_\sigma((T_i, Z_i)|\, 0 \leq i \leq m)$. According to Jacod and Shiryaev ([1987], theorem III 1.33), there is a P-a.s. uniquely determined $\mathfrak{F}_t$-compensator of the form

$$A_{yz,t} = \sum_{m\in\mathbb{N}_0} A^{(m)}_{yz,t}, \ \ t \geq 0, \tag{2.13.2}$$

with

$$A^{(m)}_{yz,t} := \int_{(T_m\wedge t, T_{m+1}\wedge t]} \frac{G^{(m)}(\cdot, ds, \{y, z\})}{G^{(m)}(\cdot, [s, \infty], \mathcal{J})}\,. \tag{2.13.3}$$

Since $((T_m, Z_m))_{m\in\mathbb{N}_0}$ is a homogeneous Markov chain with transition probabilities according to definition 2. 8, one obtains P-a.s.

$$G^{(m)}(\cdot, ds, \{y, z\}) \;=\; P(T_{m+1} \in ds, Z_m = y, Z_{m+1} = z|\, T_m, Z_m) = \mathbf{1}_{\{Z_m=y\}}\, \hat{Q}_{Z_m z}(T_m, ds),$$

and

$$G^{(m)}(\cdot,[s,\infty],\mathcal{J}) \;=\; 1-\hat{Q}_{Z_m}(T_m,s-0).$$

Inserting this into (2. 13. 3), and afterwards applying (2. 8. 3), it follows

$$A^{(m)}_{yz,t} \;=\; \int_{(T_m\wedge t,T_{m+1}\wedge t]} \mathbf{1}_{\{Z_m=y\}}\,\hat{q}_{yz}(T_m,ds)\,.$$

Finally, employing (2. 13. 2) yields the assertion (2. 13. 1). □

The compensators $(A_{yz,t})_{t\geq 0}$, $(y,z)\in\mathcal{J}$ are predictable processes satisfying $\mathbf{E}[N_{yz,t}] = \mathbf{E}[A_{yz,t}]$ for each $t\geq 0$. Hence, in order to ensure (2. 12. 2), we stipulate

2. 14 Assumption.

$$\mathbf{A1}:\quad \mathbf{E}\Big[\sum_{(y,z)\in\mathcal{J}}\sum_{m\in\mathbb{N}_0}\int_{(T_m\wedge t,T_{m+1}\wedge t]}\hat{q}_{yz}(T_m,ds)\Big]<\infty,\quad t\geq 0. \tag{2. 14. 1}$$

In the framework of Milbrodt and Helbig [1999], the finiteness of $\mathbf{E}[A_{yz,t}]$, $t\geq 0$, follows from the regularity of the cumulative transition intensity matrix (see Milbrodt and Helbig [1999], Folgerung 4.38). According to this, the local boundedness of the cumulative transition intensities is sufficient to ensure the number of jumps on bounded intervals being P-a.s. finite. In the present framework, however, the regularity of the cumulative transition intensity matrix is not sufficient to ensure (2. 14. 1). The following counterexample points this out.

2. 15 Example. Let $((T_m,Z_m))_{m\in\mathbb{N}_0}$ be a homogeneous Markovian marked point process with state space $([0,\infty]\times\mathcal{S},\mathfrak{B}([0,\,\infty])\otimes 2^{\mathcal{S}})$, initial distribution $\mathcal{L}(T_0,Z_0|P)=\varepsilon_0\otimes\pi$, and transition probabilities which are for $0\leq s<r$ and $(y,z)\in\mathcal{J}$ given by

$$\hat{Q}_{yz}(s,t) := \frac{1}{|\mathcal{S}|-1}\,\mathbf{1}_{[s+\frac{r-s}{2},\infty)}(t),\;\; t\geq s. \tag{2. 15. 1}$$

Hence

$$\hat{Q}_y(s,t) = \sum_{\substack{z\in\mathcal{S}\\ z\neq y}}\frac{1}{|\mathcal{S}|-1}\,\mathbf{1}_{[s+\frac{r-s}{2},\infty)}(t) = \mathbf{1}_{[s+\frac{r-s}{2},\infty)}(t),\;\; t\geq s. \tag{2. 15. 2}$$

Thus, $\hat{Q}_*$ define measures concentrated on (s,∞). Further, (2. 12. 1) is satisfied. According to (2. 8. 3), one obtains for the appertaining hazard measures (cumulative transition intensities)

$$\hat{q}_{yz}(s,dt) = \frac{1}{|\mathcal{S}|-1}\,\varepsilon_{s+\frac{r-s}{2}}(dt).$$

They are obviously finite on bounded intervals. But nevertheless, all jumps occur P-a.s. before r and hence, the paths of $((T_m,Z_m))_{m\in\mathbb{N}_0}$ are not contained in $\mathcal{K}$. The compensator of the number of jumps $(N_t)_{t\geq 0}$ is given by $A_t := \sum_{(y,z)\in\mathcal{J}} A_{yz,t}, t\geq 0$. For this, one obtains

$$A_t \;=\; \sum_{m\in\mathbb{N}_0}\int_{(T_m\wedge t,T_{m+1}\wedge t]}\varepsilon_{T_m+\frac{r-T_m}{2}}(ds) = \sum_{m\in\mathbb{N}_0}\mathbf{1}_{\{T_m\wedge t<T_m+\frac{r-T_m}{2}\leq T_{m+1}\wedge t\}},$$

which is infinite for $t\geq r$. △

Taking into consideration that in our framework the finiteness of a regular cumulative transition intensity matrix is not sufficient to ensure the marked point process being nonexplosive, more restrictive conditions must be imposed. In several scenarios of practical interest, there are Borel measures - not necessarily the Lebesgue measure - that dominate the cumulative transition intensities, with the appertaining densities being bounded. In a situation like this, (2. 14. 1) can be granted: Let $\mathbf{\Lambda}_{yz}$ be a Borel measure on $\mathfrak{B}([0,\infty))$ such that for $s \geq 0$ and $(y,z) \in \mathcal{J}$

$$\hat{q}_{yz}(s, d\tau) \ll \mathbf{\Lambda}_{yz}(d\tau) \quad \text{with} \quad \frac{d\hat{q}_{yz}(s,\cdot)}{d\mathbf{\Lambda}_{yz}} := \lambda_{yz}(s,\cdot).$$

Then for $t \geq 0$

$$\begin{aligned} A_t &= \sum_{(y,z)\in\mathcal{J}} \sum_{m\in\mathbb{N}_0} \int_{(T_m\wedge t, T_{m+1}\wedge t]} \mathbf{1}_{\{Z_m=y\}}\, \lambda_{yz}(T_m,\tau)\, \mathbf{\Lambda}_{yz}(d\tau) \\ &\leq \sum_{(y,z)\in\mathcal{J}} \sum_{m\in\mathbb{N}_0} \int_{(T_m\wedge t, T_{m+1}\wedge t]} \sup_{r,s\in[0,t]} \lambda_{yz}(r,s)\, \mathbf{\Lambda}_{yz}(d\tau) \\ &= \sum_{(y,z)\in\mathcal{J}} \int_{(0,T_\infty]\cap(0,t]} \sup_{r,s\in[0,t]} \lambda_{yz}(r,s)\, \mathbf{\Lambda}_{yz}(d\tau) \\ &\leq \sum_{(y,z)\in\mathcal{J}} \sup_{r,s\in[0,t]} \lambda_{yz}(r,s)\, \mathbf{\Lambda}_{yz}([0,t]), \end{aligned} \tag{2. 15. 3}$$

which is finite according to the above assumptions.

We now turn to a pair of important tools for future calculations stated in the following two remarks. The first is a chain of equations that will be used in almost the same manner for the derivation of integral equations for both the transition probabilities of the bivariate Markov process $((X_t, U_t))_{t\geq 0}$ and the prospective reserves. The second tool is the adaption of (2. 6. 1) to the present situation. It allows one to calculate conditional expectations by means of integration with respect to a corresponding conditional distribution.

2. 16 Remark. Let $((T_m, Z_m))_{m\in\mathbb{N}_0}$ be a homogeneous Markovian marked point process with state space $([0,\infty]\times\mathcal{S}, \mathfrak{B}([0,\infty])\otimes 2^{\mathcal{S}})$. Further, let $s \geq 0, B \in \mathfrak{B}([0,\infty))$, and $(y,z) \in \mathcal{J}$. For $l \in \mathbb{N}_0$, we define the l-step transition probabilities of the marked point process by

$$\hat{Q}^{(l)}_{yz}(s,B) \;:=\; P(T_{m+l}\in B, Z_{m+l}=z \,|\, T_m = s,\, Z_m = y). \tag{2. 16. 1}$$

Since (T,Z) is homogeneous, these transition probabilities can also be chosen as independent of $m \in \mathbb{N}$. Upon conditioning on (T_{m+1}, Z_{m+1}), applying the Markov property of (T,Z), inserting (2. 8. 1), as well as employing (2. 9. 1), one obtains for $l > 0$ the following backward equations:

$$\begin{aligned} \hat{Q}^{(l)}_{yz}(s,B) &= P(T_{m+l}\in B, Z_{m+l}=z \,|\, T_m = s,\, Z_m = y) \\ &= \int_{(s,\infty)\times\mathcal{S}} P(T_{m+l}\in B, Z_{m+l}=z, T_{m+1}\in d\tau, Z_{m+1}\in d\xi \,|\, T_m=s, Z_m=y) \\ &= \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \int_{(s,\infty)} P(T_{m+l}\in B, Z_{m+l}=z \,|\, T_{m+1}=\tau, Z_{m+1}=\xi) \\ &\qquad\qquad \cdot P(T_{m+1}\in d\tau, Z_{m+1}=\xi \,|\, T_m = s,\, Z_m = y) \\ &= \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \int_{(s,\infty)} \hat{Q}^{(l-1)}_{\xi z}(\tau,B)\, \hat{Q}_{y\xi}(s,d\tau) \\ &= \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \int_{(s,\infty)} \hat{Q}^{(l-1)}_{\xi z}(\tau,B)\, (1-\hat{Q}_y(s,\tau-0))\, \hat{q}_{y\xi}(s,d\tau). \end{aligned} \tag{2. 16. 2}$$

Upon conditioning on (T_{m+l-1}, Z_{m+l-1}), the following forward equations can be obtained analogously:

$$\begin{aligned}
\hat{Q}^{(l)}_{yz}(s,B) &= P(T_{m+l} \in B, Z_{m+l} = z \,|\, T_m = s,\ Z_m = y) \\
&= \int_{[s,\infty)\times\mathcal{S}} P(T_{m+l} \in B, Z_{m+l} = z, T_{m+l-1} \in d\tau, Z_{m+l-1} \in d\xi \,|\, T_m = s, Z_m = y) \\
&= \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{[s,\infty)} \hat{Q}_{\xi z}(\tau, B)\, \hat{Q}^{(m-1)}_{y\xi}(s, d\tau) \\
&= \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{[s,\infty)} \int_B (1 - \hat{Q}_\xi(\tau, x-0))\, \hat{q}_{\xi z}(\tau, dx)\, \hat{Q}^{(l-1)}_{y\xi}(s, d\tau). \qquad (2.16.3)
\end{aligned}$$

2. 17 Remark. Let $\mathcal{L}(((T_l, Z_l))_{l\geq m} \,|\, T_m, Z_m)$, $m \in \mathbb{N}_0$ denote a regular version of the conditional distribution of $((T_l, Z_l))_{l\geq m}$ given (T_m, Z_m). Then for every $\mathfrak{K}$-measurable mapping $f : \mathcal{K} \to [0,\infty]$, the following holds for $\mathcal{L}(T_m, Z_m \,|\, P)$-a.e. $(s,y) \in [0,\infty)\times\mathcal{S}$:

$$\mathbf{E}\left[f(((T_l, Z_l))_{l\geq m}) \,|\, T_m = s, Z_m = y\right] = \int_{\mathcal{K}} f\, d\mathcal{L}(((T_l, Z_l))_{l\geq m} \,|\, T_m = s, Z_m = y), \qquad (2.17.1)$$

provided the integral is well defined. Due to the marked point process being homogeneous, the conditional distribution of $((T_l, Z_l))_{l\geq m}$ given (T_m, Z_m) as well as the right-hand side of (2. 17. 1) can likewise be chosen as independent of m. For this reason, we will in the sequel often write $\mathcal{L}(((T_l, Z_l))_{l\geq 0} \,|\, T_0, Z_0)$.

In order to later derive the variance of the prospective loss, we provide here some more facts concerning counting processes and their compensators. Consider for $(y,z) \in \mathcal{J}$ the corresponding component of the multivariate counting process $(N_{yz,t})_{t\geq 0}$, and the appertaining compensator $(A_{yz,t})_{t\geq 0}$ given by (2. 13. 1). According to Milbrodt and Helbig ([1999], Satz 12.27), it follows that under assumption (2. 14. 1) the innovation process

$$M_{yz,t} := N_{yz,t} - A_{yz,t}, \;\; t \geq 0, \qquad (2.17.2)$$

is a square integrable martingale. Further, the optional covariation processes are of the following form (cf. Andersen et al. [1993], section II.3.2):

$$\begin{aligned}
[M_{yz}, M_{\eta\xi}]_t &= \sum_{s\leq t} \Delta M_{yz,s}\, \Delta M_{\eta\xi,s}, \;\; t \geq 0,\; (y,z), (\eta,\xi) \in \mathcal{J} \\
&= M_{yz,t}\, M_{\eta\xi,t} - \int_{(0,t]} M_{\eta\xi,s-0}\, M_{yz,ds} - \int_{(0,t]} M_{yz,s-0}\, M_{\eta\xi,ds}. \qquad (2.17.3)
\end{aligned}$$

Hence, the optional variation processes are given by

$$[M_{yz}]_t = M^2_{yz,t} - 2\int_{(0,t]} M_{yz,s-0}\, M_{yz,ds}, \;\; t \geq 0,\; (y,z) \in \mathcal{J}. \qquad (2.17.4)$$

The corresponding predictable covariations $\langle M_{yz}, M_{\eta\xi}\rangle_t, t \geq 0$, as well as the predictable variations $\langle M_{yz}\rangle_t, t \geq 0$, can be obtained by deriving the compensators of (2. 17. 3) and (2. 17. 4), respectively. This is the issue of the next lemma.

2. 18 Lemma. *Assume* (2. 14. 1) *and let* $(M_{yz,t})_{t\geq 0}$, $(y,z) \in \mathcal{J}$, *be innovation processes according to* (2. 17. 2). *The corresponding predictable covariations are, up to P-indistinguishability,* *given by*

$$\begin{aligned}\langle M_{yz}, M_{\eta\xi}\rangle_t &= \delta_{y\eta}\,\delta_{z\xi} \sum_{m\in\mathbb{N}_0} \int_{(T_m\wedge t, T_{m+1}\wedge t]} \mathbf{1}_{\{Z_m=y\}}\, \hat{q}_{yz}(T_m, ds) \\ &\quad -\delta_{y\eta} \sum_{m\in\mathbb{N}_0} \int_{(T_m\wedge t, T_{m+1}\wedge t]} \mathbf{1}_{\{Z_m=y\}}\, \hat{q}_{yz}(T_m, \{s\})\, \hat{q}_{y\xi}(T_m, ds), \quad t \geq 0. \quad (2.18.1)\end{aligned}$$

PROOF. To prove (2.18.1), we argue in almost the same manner as in Andersen et al. ([1993], chapter II.4.1). Under the assumptions stipulated above, the optional covariation processes of the innovations according to (2.17.2) are given by (2.17.3). They can be represented in the following way: For $t \geq 0$

$$\begin{aligned}&[M_{yz}, M_{\eta\xi}]_t \\ &= \int_{(0,t]} \Delta M_{yz,s}\, M_{\eta\xi,ds} \\ &= \int_{(0,t]} (\Delta N_{yz,s} - \Delta A_{yz,s})\,(N_{\eta\xi,ds} - A_{\eta\xi,ds}) \\ &= \sum_{s\leq t} \Delta N_{yz,s}\, \Delta N_{\eta\xi,s} - \int_{(0,t]} \Delta N_{yz,s}\, A_{\eta\xi,ds} - \int_{(0,t]} \Delta A_{yz,s}\, N_{\eta\xi,ds} + \int_{(0,t]} \Delta A_{yz,s}\, A_{\eta\xi,ds} \\ &= \delta_{y\eta}\,\delta_{z\xi} \int_{(0,t]} N_{yz,ds} - \int_{(0,t]} \Delta A_{\eta\xi,s}\, N_{yz,ds} - \int_{(0,t]} \Delta A_{yz,s}\, N_{\eta\xi,ds} + \int_{(0,t]} \Delta A_{yz,s}\, A_{\eta\xi,ds}. \quad (2.18.2)\end{aligned}$$

The predictable covariation $\langle M_{yz}, M_{\eta\xi}\rangle_t, t \geq 0$ is the compensator of the optional covariation. Hence, the compensators of the four addends in the last line of (2.18.2) must be derived. Obviously, the compensator of the first addend is $\delta_{y\eta}\,\delta_{z\xi}\, A_{yz,t},\ t \geq 0$. The last addend is predictable and hence, it is its own compensator. Concerning the second addend, note that $\Delta A_{\eta\xi,s}, s \geq 0$ is bounded and predictable, and since $M_{yz,t}, t \geq 0$ is a martingale, the stochastic integral

$$\int_{(0,t]} \Delta A_{\eta\xi,s}\, M_{yz,ds} = \int_{(0,t]} \Delta A_{\eta\xi,s}\, N_{yz,ds} - \int_{(0,t]} \Delta A_{\eta\xi,s}\, A_{yz,ds}$$

is also a martingale. Consequently, the predictable process $\int_{(0,t]} \Delta A_{\eta\xi,s}\, A_{yz,ds}$ is the compensator of the second addend in the last line of (2.18.2). Similarly, it can be verified that $\int_{(0,t]} \Delta A_{yz,s}\, A_{\eta\xi,ds}$ is the compensator of the third addend. Summarizing the above, we obtain for $(y,z), (\eta,\xi) \in \mathcal{J}$

$$\begin{aligned}\langle M_{yz}, M_{\eta\xi}\rangle_t &= \delta_{y\eta}\,\delta_{z\xi}\, A_{yz,t} - \int_{(0,t]} \Delta A_{\eta\xi,s}\, A_{yz,ds} - \int_{(0,t]} \Delta A_{yz,s}\, A_{\eta\xi,ds} + \int_{(0,t]} \Delta A_{yz,s}\, A_{\eta\xi,ds} \\ &= \delta_{y\eta}\,\delta_{z\xi}\, A_{yz,t} - 2\int_{(0,t]} \Delta A_{yz,s}\, A_{\eta\xi,ds} + \int_{(0,t]} \Delta A_{yz,s}\, A_{\eta\xi,ds} \\ &= \delta_{y\eta}\,\delta_{z\xi}\, A_{yz,t} - \int_{(0,t]} \Delta A_{yz,s}\, A_{\eta\xi,ds}, \quad t \geq 0.\end{aligned}$$

Finally, inserting (2.13.1) yields the assertion (2.18.1). □

We close this section by having a look at the appertaining pure jump process $X = G^{-1}(T,Z)$. Since $G^{-1} : \mathcal{K} \to \mathcal{X}$ is a measurable mapping, the distribution of X is uniquely determined by the distribution of (T,Z). But the Markov property of the marked point process (T,Z) does not imply the Markov property of the pure jump process X. For this, we refer to a counterexample given by Stracke ([1997], Beispiel 5.1). Stracke constructed a Markovian marked point process for which the appertaining pure jump process does not possess the Markov property. Hence, the Markov property of a marked point process (T, Z) does not imply the Markov property of the appertaining pure jump process X. Conversely, according to Milbrodt and Helbig ([1999], Hilfssatz 4.33), the Markov property of a pure jump process X implies the Markov property of the appertaining marked point process (T, Z):

$$\boxed{(T,Z)\text{ homogeneous Markov chain}} \quad \substack{\not\Longrightarrow \\ \Longleftarrow} \quad \boxed{X\text{ Markovian pure jump process}}$$

This raises the question of conditions for the distribution of a marked point process (T, Z) such that the Markov property of this marked point process goes together with the Markov property of the appertaining pure jump process X. A sufficient condition for this is that the cumulative transition intensities of the marked point process (T, Z) possess the additivity property (2. 11. 2). That this condition is sufficient will be stated in theorem 2. 42. The following lemma gathers some properties of the transition probabilities and the cumulative transition intensities following from the property (2. 11. 2).

2. 19 Lemma. *Let $((T_m, Z_m))_{m\in\mathbb{N}_0}$ be a homogeneous Markovian marked point process with transition probabilities according to definition 2. 8, and regular cumulative transition intensities $\hat{q}_*$ satisfying (2. 11. 2). Further, let $s \geq 0$ and $(y, z) \in \mathcal{J}$. Then:*

- *There is a Borel measure $q_{yz}(dr)$ such that $\hat{q}_{yz}(s, dr) \ll q_{yz}(dr)$ and*

$$\hat{q}_{yz}(s, dr) = \mathbf{1}_{(s,\infty)}(r)\, q_{yz}(dr). \tag{2. 19. 1}$$

- *The densities*

$$\frac{d\hat{Q}_{yz}}{d\hat{Q}_y}(s, \cdot) = -\frac{d\hat{q}_{yz}}{d\hat{q}_{yy}}(s, \cdot) \tag{2. 19. 2}$$

do not depend on s, and hence, the transition probabilities are of the form

$$\hat{Q}_{yz}(s, B) \;=\; \int_B -\frac{d\hat{q}_{yz}}{d\hat{q}_{yy}}(\tau)\, \hat{Q}_y(s, d\tau), \;\; B \in \mathfrak{B}([0, \infty)). \tag{2. 19. 3}$$

- *The conditional distributions $\hat{Q}_y$ satisfy*

$$1 - \hat{Q}_y(s, t) = (1 - \hat{Q}_y(s, r))\, (1 - \hat{Q}_y(r, t)), \;\; s \leq r \leq t. \tag{2. 19. 4}$$

- *In case of $\hat{Q}_y(s, t) < 1$, $s \leq t$, there is an appropriate version of the conditional distribution that satisfies*

$$P(((T_l, Z_l))_{l\in\mathbb{N}} | \, T_1 > t, T_0 = s, Z_0 = y) = P(((T_l, Z_l))_{l\in\mathbb{N}} | \, T_0 = t, Z_0 = y). \tag{2. 19. 5}$$

PROOF. Let $s \leq r \leq t$ and $(y, z) \in \mathcal{J}$. Since the cumulative transition intensities define measures on $\mathfrak{B}([0, \infty))$, we have on the one hand (cf. (2. 10. 1))

$$\begin{aligned} \hat{q}_{yz}(s, t) &= \hat{q}_{yz}(s, (s, t]) \\ &= \hat{q}_{yz}(s, (s, r]) + \hat{q}_{yz}(s, (r, t]). \end{aligned} \tag{2. 19. 6}$$

On the other hand, we obtain by employing (2. 11. 2)

$$\begin{aligned} \hat{q}_{yz}(s, t) &= \hat{q}_{yz}(s, r) + \hat{q}_{yz}(r, t) \\ &= \hat{q}_{yz}(s, (s, r]) + \hat{q}_{yz}(r, (r, t]). \end{aligned} \tag{2. 19. 7}$$

Hence, $\hat{q}_{yz}(s,(r,t]) = \hat{q}_{yz}(r,(r,t])$. Hence, the measures $\hat{q}_{yz}(s,\cdot)$ and $\hat{q}_{yz}(r,\cdot)$ coincide on a $\cap$-stable generator of $\mathfrak{B}((r,\infty))$. Then, according to uniqueness of measures, they coincide on $\mathfrak{B}((r,\infty))$. To prove the first assertion, we define $q_{yz}(dr) := \hat{q}_{yz}(0,dr)$. Then, as previously argued, for each $s \geq 0$, the measures $q_{yz}(dr)$ and $\hat{q}_{yz}(s,dr)$ coincide on $\mathfrak{B}((s,\infty))$. Hence, on $\mathfrak{B}((s,\infty))$, they are dominated by each other and, particularly, $\hat{q}_{yz}(s,dr) \ll q_{yz}(dr)$. Since further, $\hat{q}_{yz}(s,dr)$ is concentrated on (s,∞), it follows that $\hat{q}_{yz}(s,B) = 0$ for all $B \in \mathfrak{B}([0,s])$. This also holds for sets $B \in \mathfrak{B}([0,s])$ for which $q_{yz}(B) = 0$. Consequently, the measure $\hat{q}_{yz}(s,dr)$ is on $\mathfrak{B}([0,\infty))$ dominated by the Borel measure $q_{yz}(dr)$, and the following holds:

$$\hat{q}_{yz}(s,B) = \int_B \mathbf{1}_{(s,\infty)}(r)\, q_{yz}(dr), \quad s \geq 0, B \in \mathfrak{B}([0,\infty)).$$

The third assertion is a consequence of (2. 11. 2) and the exponential formula (2. 10. 7). One obtains

$$\begin{aligned} 1 - \hat{Q}_y(s,t) &= \exp\{\hat{q}^{(c)}_{yy}(s,t)\} \prod_{s<\tau\leq t} (1 + \Delta\hat{q}_{yy}(s,\tau)) \\ &= \exp\{\hat{q}^{(c)}_{yy}(s,r) + \hat{q}^{(c)}_{yy}(r,t)\} \prod_{s<\tau\leq r} (1 + \Delta\hat{q}_{yy}(s,\tau)) \prod_{r<\tau\leq t} (1 + \Delta\hat{q}_{yy}(r,\tau)) \\ &= (1 - \hat{Q}_y(s,r))\,(1 - \hat{Q}_y(r,t)). \end{aligned}$$

Concerning the remaining assertions, we refer to Milbrodt and Helbig ([1999], Hilfssatz 4.43 and Hilfssatz 4.39). Note that a regular cumulative transition intensity matrix $(\hat{q}_{yz})_{(y,z)\in\mathcal{S}^2}$ that additionally satisfies the property (2. 11. 2) fulfills the requirements of a regular cumulative transition intensity matrix in the sense of the authors. □

The aim of the next section is to introduce semi-Markovian pure jump processes $(X_t)_{t\geq 0}$ modelling the development of a single policy (p). Further, the bivariate process $((X_t,U_t))_{t\geq 0}$ is investigated. Though the pure jump process $(X_t)_{t\geq 0}$ is - in contrast to the classical model assumption - not a Markov process, the bivariate process $((X_t,U_t))_{t\geq 0}$ is Markovian. In a semi-Markov set-up, this process forms the basic quantity of interest. For the investigation of the process $((X_t,U_t))_{t\geq 0}$, the previously discussed marked point process $((T_m,Z_m))_{m\in\mathbb{N}_0}$ will be of considerable help.

D Non-smooth semi-Markovian pure jump processes

Let $(\Omega,\mathfrak{F},P,((T_m,Z_m))_{m\in\mathbb{N}_0})$ be a homogeneous Markovian marked point process with paths in $(\mathcal{K},\mathfrak{K})$. Recall that for this, the assumptions (2. 12. 1) and (2. 14. 1) must be satisfied. Further, let $(X_t)_{t\geq 0}$ be the appertaining pure jump process. $\mathcal{S}$ denotes the finite state space of $(X_t)_{t\geq 0}$, and $\mathcal{J} = \{(y,z) \in \mathcal{S}^2 |\, y \neq z\}$ is the transition space. As in the preceding sections, $(X_t)_{t\geq 0}$ is interpreted as the development of a single policy (p).

Defining a pure jump process $(X_t)_{t\geq 0}$ to be semi-Markovian, there are basically two approaches. The first approach starts from a two-dimensional discrete process. This approach goes back to Lévy, Smith, and Takács, who in 1954-1955 independently and almost simultaneously introduced the concept of semi-Markov processes (for a survey of literature see Korolyuk et al. [1974]), but it is often related with Janssen and De Dominicis [1984]. The second approach starts from the bivariate process $((X_t,U_t))_{t\geq 0}$. It goes back to Pyke and Schaufele [1964] and was adapted to insurance mathematics by Hoem [1972].

We start by considering the first approach. According to Janssen and De Dominicis [1984], a pure jump process $(X_t)_{t\geq 0}$ is called semi-Markovian if it appertains to a Markovian marked point process $((T_m,Z_m))_{m\in\mathbb{N}_0}$. In cases where the Markov chain $((T_m,Z_m))_{m\in\mathbb{N}_0}$ is non-homogeneous, meaning that the transition probabilities (cf. definition 2. 8) also depend on the number of previous jumps, the sequence

$$Q^{(m)}(s,t) := (\hat{Q}^{(m)}_{yz}(s,t))_{(y,z)\in\mathcal{S}^2}, \quad m\in\mathbb{N}, 0\le s\le t,$$

is called a *completely* non-homogeneous semi-Markov kernel. Consequently, the appertaining pure jump process $(X_t)_{t\geq 0}$ is referred to as *completely* non-homogeneous semi-Markov process. If this were the case, the transition probabilities of $(X_t)_{t\geq 0}$ would also depend on the number of previous transitions. This concept, however, seems not to have many useful applications in life insurance. Further, the provision of an appropriate actuarial basis would be enormously complicated. For these reasons, completely non-homogeneous semi-Markov processes are not considered here. We will call a process $(X_t)_{t\geq 0}$ semi-Markovian if it appertains to a homogeneous marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$ (definition 2. 20).

In the notation of Janssen and De Dominicis [1984], a marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$ is referred to as additive process, where the word *additive* indicates the following:

$$T_m = \sum_{k=0}^{m-1} V_k, \quad m\in\mathbb{N} \quad \text{with} \quad V_m = T_{m+1} - T_m, \quad m\in\mathbb{N}_0.$$

V_m corresponds to the sojourn time in the state entered by the m-th jump (see (2. 5. 8)). Instead of starting from a marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$, the first approach of defining semi-Markov processes sometimes starts from the process $((V_m, Z_m))_{m\in\mathbb{N}}$, which is also a two-dimensional discrete process. This process likewise possesses the state space $[0,\infty]\times\mathcal{S}$. The transition probabilities can be defined as

$$\hat{Q}_{yz}(s, s+\cdot) : [0,\infty) \ni r \mapsto P(V_m \le r, Z_{m+1} = z |\, T_m = s,\, Z_m = y), \;\; s\ge 0, (y,z)\in\mathcal{J}, m\in\mathbb{N}_0.$$

We turn to the second approach. According to Hoem [1972], the Markov property of the bivariate process $((X_t, U_t))_{t\geq 0}$ defines a pure jump process $(X_t)_{t\geq 0}$ to be semi-Markovian. Pyke and Schaufele [1964], however, actually start from a Markov process $((X_t, U_t))_{t\geq 0}$ being additionally strong Markovian. We will see that by requiring the bivariate process $((X_t, U_t))_{t\geq 0}$ to be strong Markovian both approaches to define semi-Markov processes are equivalent. Figure 11 sketches the corresponding relationships. Employing the strong Markov property of a homogeneous Markovian marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$ - which is due to the countable time set a consequence of the ordinary Markov property (see (2. 7. 7)) - it can be verified that the appertaining bivariate process $((X_t, U_t))_{t\geq 0}$ is a Markov process (lemma 2. 21). Conversely, starting from a given bivariate Markov process $((X_t, U_t))_{t\geq 0}$ being additionally strong Markovian, it can be shown that the appertaining marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$ is a homogeneous Markov chain (lemma 2. 31). Here, however, the strong Markov property is not generally be given. For this, some regularity conditions must additionally be satisfied. Lemma 2. 30 states a condition for the transition probabilities of the bivariate Markov process $((X_t, U_t))_{t\geq 0}$ ensuring the strong Markov property of this process. This condition is satisfied for the transition probabilities of the bivariate process $((X_t, U_t))_{t\geq 0}$ appertaining to a homogeneous Markovian marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$ (corollary 2. 33). Hence, $((X_t, U_t))_{t\geq 0}$ possesses the strong Markov property iff the appertaining marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$ is a homogeneous Markov chain.

Here, we follow the first approach by defining a semi-Markov process as the pure jump process appertaining to a homogeneous Markovian marked point process. Afterwards, the bivariate process $((X_t, U_t))_{t\geq 0}$ will be investigated in more detail.

2. 20 Definition. Let $(\Omega, \mathfrak{F}, P, ((T_m, Z_m))_{m\in\mathbb{N}_0})$ be a homogeneous Markovian marked point process with transition probabilities according to definition 2. 8. Further, let $(X_t)_{t\geq 0}$ be the appertaining pure jump process. Then, the matrix

$$Q(s,t) := (\hat{Q}_{yz}(s,t))_{(y,z)\in\mathcal{S}^2}, \quad 0\le s\le t,$$

is called a *semi-Markov kernel* and the process $(X_t)_{t\geq 0}$ is referred to as *semi-Markov process.*

As previously mentioned, *completely* non-homogeneous semi-Markov processes are not considered here. Another subclass of semi-Markov processes, which are also not considered here, was investigated in detail by Nollau [1980]. This is the class of homogeneous semi-Markov processes which are characterized by a semi-Markov kernel Q that only depends on the difference $t-s, 0 \leq s \leq t$. Homogeneous semi-Markov processes were also discussed by Hoem [1969].

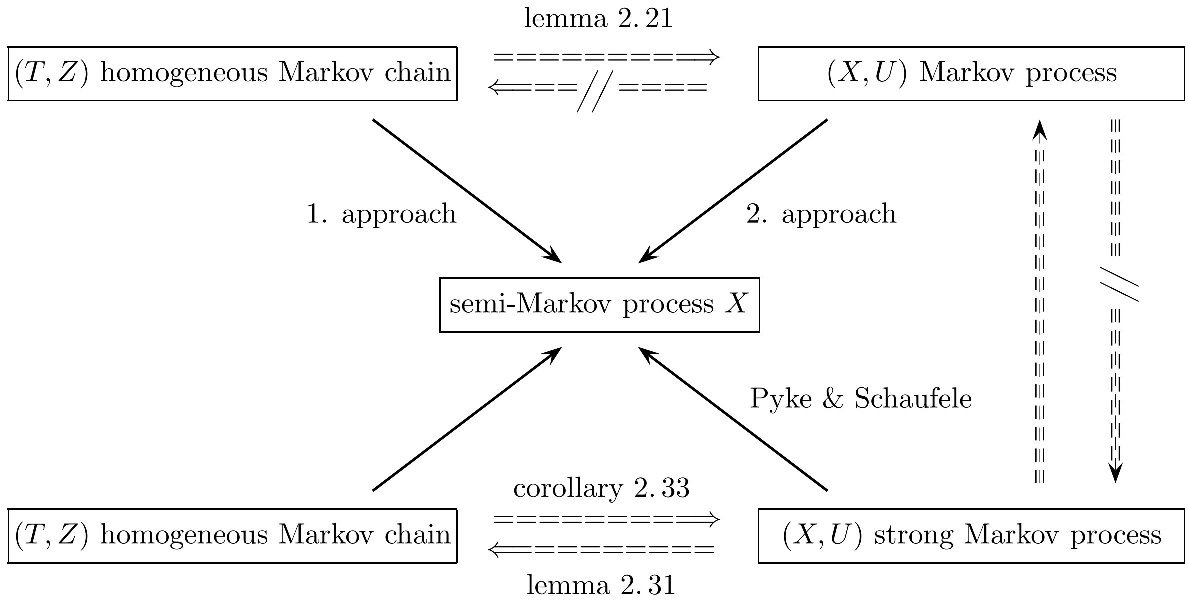

Figure 11: Approaches to define semi-Markov processes and deduced relations

Note that, for a given semi-Markov kernel Q which corresponds to the matrix of transition probabilities of a marked point process (T, Z) (cf. definition 2. 8), the existence of a semi-Markov process $(X_t)_{t\geq 0}$ is granted due to the existence of (T, Z) and the relationship $X = G^{-1}(T, Z)$.

Recall that the pure jump process $(X_t)_{t\geq 0}$ that appertains to a homogeneous Markovian marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$ is not necessarily Markovian. However, lemma 2. 21 states that the bivariate process $((X_t, U_t))_{t\geq 0}$ is always a Markov process. Incidentally, it is a well-known method to convert non-Markovian processes into Markov processes by including supplementary variables (cf. Cox [1955]). Doing so, the semi-Markov process $(X_t)_{t\geq 0}$ can be investigated by applying the theory of Markov processes to $((X_t, U_t))_{t\geq 0}$.

2. 21 Lemma. *Let* $(\Omega, \mathfrak{F}, P, ((T_m, Z_m))_{m\in\mathbb{N}_0})$ *be a homogeneous Markovian marked point process, and* $(X_t)_{t\geq 0}$ *the appertaining pure jump process. Then, the appertaining bivariate process* $((X_t, U_t))_{t\geq 0}$ *is a Markov process: For* $0 \leq s \leq t$ *and* $C \in 2^{\mathcal{S}} \otimes \mathfrak{B}([0,\infty))$

$$P((X_t, U_t) \in C|\, (X_\tau, U_\tau)_{\tau\leq s}) = P((X_t, U_t) \in C|\, X_s, U_s) \quad P-a.s. \tag{2. 21. 1}$$

Proof. To verify (2. 21. 1), it is sufficient to show that for $t \geq t_n \geq t_{n-1} \geq ... \geq t_0 \geq 0,\ n \in \mathbb{N}$, and $C \in 2^{\mathcal{S}} \otimes \mathfrak{B}([0,\infty))$

$$P((X_t, U_t) \in C|\, X_{t_n}, U_{t_n}, ..., X_{t_0}, U_{t_0}) = P((X_t, U_t) \in C|\, X_{t_n}, U_{t_n}) \quad P-a.s. \tag{2. 21. 2}$$

Since sets of the form $\{z\} \times [0, v]$, $z \in \mathcal{S}, v \geq 0$ form, together with the empty set, a $\cap$-stable generator of $2^{\mathcal{S}} \otimes \mathfrak{B}([0,\infty))$, and since further,

$$\mathfrak{D} := \{C \in 2^{\mathcal{S}} \otimes \mathfrak{B}([0,\infty)) |\, (2.21.2) \text{ is satisfied for } C\}$$

is a Dynkin system, (2. 21. 2) only needs to be verified for elements of

$$\mathfrak{D}_0 := \{\{z\} \times [0,v],\ z \in \mathcal{S}, v \geq 0\} \cup \{\emptyset\}. \tag{2.21.3}$$

Thus, $\mathfrak{D}_0 \subset \mathfrak{D}$, and due to $a_\sigma(\mathfrak{D}_0) = 2^{\mathcal{S}} \otimes \mathfrak{B}([0,\infty)) \subset \mathfrak{D}$, (2. 21. 2) holds for all $C \in 2^{\mathcal{S}} \otimes \mathfrak{B}([0,\infty))$. Now consider (2. 21. 2) for $C = \{z\} \times [0,v] \in \mathfrak{D}_0$. By employing (2. 4. 5) and (2. 5. 9), we obtain for the left-hand side of (2. 21. 2)

$$\begin{aligned}
&P(X_t = z, U_t \in [0,v] |\, X_{t_n}, U_{t_n}, ..., X_{t_0}, U_{t_0}) \\
&= P(Z_{N_t} = z, t - T_{N_t} \in [0,v] |\, Z_{N_{t_n}}, t_n - T_{N_{t_n}}, ..., Z_{N_{t_0}}, t_0 - T_{N_{t_0}}) \\
&= P(Z_{N_t} = z, T_{N_t} \in [t-v,t] |\, Z_{N_{t_n}}, t_n - T_{N_{t_n}}, ..., Z_{N_{t_0}}, t_0 - T_{N_{t_0}}) \\
&= P(Z_{N_t} = z, T_{N_t} \in [t-v,t] |\, Z_{N_{t_n}}, t_n - T_{N_{t_n}}) \quad P - a.s.,
\end{aligned} \tag{2.21.4}$$

where the last equality is a consequence of the strong Markov property of the associated marked point process (T, Z), along with the fact that for $t \geq 0$ the number of jumps N_t is a stopping time with respect to the filtration $(\mathfrak{F}_{T_m})_{m \in \mathbb{N}_0}$. The strong Markov property of (T, Z) is a consequence of the ordinary Markov property, since the time set is countable (see (2. 7. 7) or Chung [2001], theorem 9.2.5). The right-hand side of (2. 21. 2) can then be obtained by means of

$$P(Z_{N_t} = z, T_{N_t} \in [t-v,t] |\, Z_{N_{t_n}}, t_n - T_{N_{t_n}}) = P(X_t = z, U_t \in [0,v] |\, X_{t_n}, U_{t_n}).$$

Hence, (2. 21. 2) holds for all elements of $\mathfrak{D}_0$, and thus, also for all $C \in 2^{\mathcal{S}} \otimes \mathfrak{B}([0,\infty))$. □

According to Nollau ([1980], Satz 1.5), the bivariate process $(X_t, W_t)_{t\geq 0}$ (cf. (2. 5. 7)), which records the current state of $(X_t)_{t\geq 0}$ and the future sojourn time in that state, is also a Markov process. In the sequel, however, we only consider the bivariate Markov process $((X_t, U_t))_{t\geq 0}$. Due to the polish state space $\mathcal{S} \times [0,\infty)$, there exists a set of transition probabilities

$$p(s,t) : \big((\mathcal{S} \times [0,\infty)) \times (2^{\mathcal{S}} \otimes \mathfrak{B}([0,\infty)))\big) \ni ((y,u),C) \mapsto p_{(y,u)\,C}(s,t),\ \ 0 \leq s \leq t, \tag{2.21.5}$$

with

$$p_{(y,u)\,C}(s,t) = P\big((X_t, U_t) \in C |\, X_s = y, U_s = u\big)$$

for $\mathcal{L}(X_s, U_s |\, P)$-a.e. $(y,u) \in \mathcal{S} \times [0,\infty)$. Since the system $\mathfrak{D}_0$ according to (2. 21. 3) forms a $\cap$-stable generator of $2^{\mathcal{S}} \otimes \mathfrak{B}([0,\infty))$, it is sufficient to specify the transition probabilities by means of

$$p_{yz}(s,t,u,v) := p_{(y,u)\,\{z\}\times[0,v]}(s,t) \overset{\text{a.s.}}{=} P\big(X_t = z, U_t \leq v |\, X_s = y, U_s = u\big) \tag{2.21.6}$$

for $u, v \geq 0$, and $(y,z) \in \mathcal{S}^2$. Up to $\mathcal{L}(X_s, U_s |\, P)$-exceptional sets, they are also uniquely determined by the right-hand side of (2. 21. 6). The corresponding transition probability matrix can then be defined as

$$p(s,t,u,v) := \big(p_{yz}(s,t,u,v)\big)_{(y,z)\in\mathcal{S}^2}, \quad 0 \leq u \leq s \leq t < \infty, v \geq 0. \tag{2.21.7}$$

Further, we define

$$\begin{aligned}
p_{yz}(s,t,u) &:= p_{yz}(s,t,u,\infty) \\
&\overset{\text{a.s.}}{=} P\big(X_t = z |\, X_s = y, U_s = u\big),\ \ 0 \leq u \leq s \leq t.
\end{aligned} \tag{2.21.8}$$

Thus, the conditional probability of $\{X_t = z\}, t \geq 0, z \in \mathcal{S}$, is allowed to depend not only on the state of the pure jump process at time s, but also on the time elapsed since entering this state. In doing so, durational effects can be implemented in the model.

We now turn to the Chapman-Kolmogorov equations. These equations must be generally satisfied by a set of transition probabilities for a Markov process. For the bivariate process $((X_t, U_t))_{t\geq 0}$, the Chapman-Kolmogorov equations can already be found in Hoem [1972]. In a smooth framework, Hoem outlined basic principles of the theory of semi-Markov processes and their applications in actuarial mathematics.

2. 22 Lemma. [Chapman-Kolmogorov equations] *Let $((X_t, U_t))_{t\geq 0}$ be a bivariate Markov process with transition probability matrix p according to $(2.21.7)$. Further, let $0 \leq u \leq s \leq t < \infty, v \geq 0$, and $(y, z) \in \mathcal{S}^2$. Then, the components of p satisfy for $\mathcal{L}(X_s, U_s|\, P)$-a.e. (y, u) the Chapman-Kolmogorov equations*

$$p_{yz}(s,t,u,v) = \sum_{\xi\in\mathcal{S}} \int_{[0,\infty)} p_{\xi z}(r,t,l,v)\, p_{y\xi}(s,r,u,dl), \quad s \leq r \leq t. \tag{2.22.1}$$

Further, for $\mathcal{L}(X_s, U_s|\, P)$-a.e. (y, u)

$$p_{yz}(s,s,u,v) = \delta_{yz}\, \mathbf{1}_{(v\geq u)}. \tag{2.22.2}$$

Proof. Let $0 \leq u \leq s \leq t < \infty, v \geq 0$, and $(y, z) \in \mathcal{S}^2$. It must be verified that the right-hand side of $(2.22.1)$ is a version of $p_{yz}(s,t,u,v)$, meaning that it is $2^{\mathcal{S}} \otimes \mathfrak{B}([0,\infty))$-measurable and satisfies the corresponding Radon-Nikodym equation. Due to the use of transition probabilities, the $2^{\mathcal{S}} \otimes \mathfrak{B}([0,\infty))$-measurability of the right-hand side of $(2.22.1)$ is granted. Regarding the corresponding Radon-Nikodym equation, an argumentation similar as in the proof of lemma 2.21 yields that it is sufficient to verify this equation for $\{\eta\} \times [0,w] \in \mathfrak{D}_0$. Doing so, one obtains

$$\begin{aligned} &P(X_t = z, U_t \leq v, X_s = \eta, U_s \leq w) \\ &= \int_{\{\eta\}\times[0,w]} P(X_t = z, U_t \leq v|\, X_s = y, U_s = u)\, \mathcal{L}(X_s, U_s|P)(dy, du). \end{aligned} \tag{2.22.3}$$

For $s \leq r \leq t$, we get by conditioning on (X_r, U_r), employing the Markov property of $(X_t, U_t)_{t\geq 0}$, and afterwards inserting $(2.21.6)$

$$\begin{aligned} &P(X_t = z, U_t \leq v, X_s = \eta, U_s \leq w) \\ &= \int_{\{\eta\}\times[0,w]} P(X_t = z, U_t \leq v|\, X_s = y, U_s = u\big)\, \mathcal{L}(X_s, U_s|P)(dy, du) \\ &= \int_{\{\eta\}\times[0,w]} \int_{\mathcal{S}\times[0,\infty)} P(X_t = z, U_t \leq v, X_r \in d\xi, U_r \in dl|\, X_s = y, U_s = u\big)\, \mathcal{L}(X_s, U_s|P)(dy, du) \\ &= \int_{\{\eta\}\times[0,w]} \sum_{\xi\in\mathcal{S}} \int_{[0,\infty)} P(X_t = z, U_t \leq v|\, X_r = \xi, U_r = l\big) \\ &\qquad P(X_r = \xi, U_r \in dl|\, X_s = y, U_s = u\big)\, \mathcal{L}(X_s, U_s|P)(dy, du) \\ &= \int_{\{\eta\}\times[0,w]} \sum_{\xi\in\mathcal{S}} \int_{[0,\infty)} p_{\xi z}(r,t,l,v)\, p_{y\xi}(s,r,u,dl)\, \mathcal{L}(X_s, U_s|P)(dy, du). \end{aligned}$$

Hence, the right-hand side of $(2.22.1)$ satisfies $(2.22.3)$. Thus, $(2.22.1)$ is verified. $(2.22.2)$ follows immediately from $(2.21.6)$ by setting $t = s$. ☐

As explained in section 2B, regular sets of transition probabilities for a Markov process play an important role. Here, regularity means that (2. 22. 1) and (2. 22. 2) must be satisfied identically (i.e. without exceptional sets). For the bivariate Markov process $((X_t, U_t))_{t\geq 0}$, a regular set of transition probabilities can be manufactured by using a regular version of the conditional distribution $\mathcal{L}(((T_l, Z_l))_{l\geq 0}|\, T_0, Z_0)$ for the appertaining marked point process.

2. 23 Theorem. *Let $(\Omega, \mathfrak{F}, P, ((T_m, Z_m))_{m\in\mathbb{N}_0})$ be a homogeneous Markovian marked point process, and $(X_t)_{t\geq 0}$ the appertaining pure jump process. Then, there is a regular set of transition probabilities $p_{yz}(s, t, u, v)$, $0 \leq u \leq s \leq t < \infty, v \geq 0, (y, z) \in \mathcal{S}^2$ for the bivariate Markov process $((X_t, U_t))_{t\geq 0}$, which is given by*

$$p_{yz}(s, t, u, v) := \delta_{yz}\, \mathbf{1}_{(v\geq u)} \tag{2. 23. 1}$$

in case $t = s$, and otherwise

$$p_{yz}(s, t, u, v) := P(\exists l \in \mathbb{N}_0 : T_{l+1} > t, T_l \in [t-v, t], Z_l = z|\, T_1 > s, T_0 = s-u, Z_0 = y). \tag{2. 23. 2}$$

Proof. It must be verified that the right-hand side of (2. 23. 2) is a version of $p_{yz}(s, t, u, v)$ that additionally satisfies (2. 22. 1) identically. Here, we only demonstrate that it is a version of $p_{yz}(s, t, u, v)$. That this version satisfies (2. 22. 1) identically is verified in A.6 in the appendix. Let $0 \leq u \leq s < t < \infty, v \geq 0$, and $(y, z) \in \mathcal{S}^2$. Regarding (2. 21. 6), we define for $l \in \mathbb{N}_0$ the appertaining l-step transition probabilities by

$$p^{(l)}_{yz}(s, t, u, v) := P\big(X_t = z, U_t \leq v, N_t - N_s = l|\, X_s = y, U_s = u\big), \tag{2. 23. 3}$$

with $N_t, t \geq 0$, being the number of jumps according to (2. 4. 4). Hence, for $\mathcal{L}(X_s, U_s|\, P)$-a.e. (y, u)

$$p_{yz}(s, t, u, v) = \sum_{l=0}^{\infty} p^{(l)}_{yz}(s, t, u, v). \tag{2. 23. 4}$$

We will demonstrate that

$$P(T_{l+1} > t, T_l \in [t-v, t], Z_l = z|\, T_1 > s, T_0 = s-u, Z_0 = y) \tag{2. 23. 5}$$

is a version of $p^{(l)}_{yz}(s, t, u, v)$. By taking (2. 23. 4) into account, this gives

$$\begin{aligned} p_{yz}(s, t, u, v) &= \sum_{l=0}^{\infty} P(T_{l+1} > t, T_l \in [t-v, t], Z_l = z|\, T_1 > s, T_0 = s-u, Z_0 = y) \\ &= P(\cup_{l\in\mathbb{N}_0}\{T_{l+1} > t, T_l \in [t-v, t], Z_l = z\}|\, T_1 > s, T_0 = s-u, Z_0 = y) \\ &= P(\exists l \in \mathbb{N}_0 : T_{l+1} > t, T_l \in [t-v, t], Z_l = z|\, T_1 > s, T_0 = s-u, Z_0 = y) \end{aligned} \tag{2. 23. 6}$$

for $\mathcal{L}(X_s, U_s|\, P)$-a.e. (y, u), which verifies that the right-hand side of (2. 23. 2) forms a version of $p_{yz}(s, t, u, v)$.
Similarly to (2. 22. 3), the corresponding Radon-Nikodym equation for $\{\eta\} \times [0, w] \in \mathfrak{D}_0$, which must be satisfied by (2. 23. 5), is given by

$$\begin{aligned} &P(X_t = z, U_t \leq v, N_t - N_s = l, X_s = \eta, U_s \leq w) \\ &= \int_{\{\eta\}\times[0,w]} P(X_t = z, U_t \leq v, N_t - N_s = l|\, X_s = y, U_s = u)\, \mathcal{L}(X_s, U_s|P)(dy, du). \end{aligned} \tag{2. 23. 7}$$

Starting from the left-hand side of this equation, we get the following by using (2. 4. 5), (2. 5. 9), and (2.4.6):

$$
\begin{aligned}
&P(X_t = z, U_t \leq v, N_t - N_s = l, X_s = \eta, U_s \leq w) \\
&= P(Z_{N_s+l} = z, t - T_{N_s+l} \leq v, N_t = N_s + l, Z_{N_s} = \eta, s - T_{N_s} \leq w) \\
&= \sum_{k\in\mathbb{N}_0} P(\{N_s = k\} \cap \{Z_{k+l} = z, t - T_{k+l} \leq v, N_t = k + l, Z_k = \eta, s - T_k \leq w\}) \\
&= \sum_{k\in\mathbb{N}_0} P(T_k \leq s < T_{k+1}, Z_{k+l} = z, T_{k+l} \geq t - v, T_{k+l} \leq t < T_{k+l+1}, Z_k = \eta, s - T_k \leq w) \\
&= \sum_{k\in\mathbb{N}_0} P(T_{k+l+1} > t, T_{k+l} \in [t - v, t], Z_{k+l} = z, T_{k+1} > s, Z_k = \eta, s - T_k \in [0, w]) \\
&= \sum_{k\in\mathbb{N}_0} \int_{\{\eta\}\times[0,w]} P(T_{k+l+1} > t, T_{k+l} \in [t - v, t], Z_{k+l} = z, T_{k+1} > s|\, Z_k = y, s - T_k = u) \\
&\qquad \cdot\mathcal{L}(Z_k, s - T_k|\, P)(dy, du) \\
&= \sum_{k\in\mathbb{N}_0} \int_{\{\eta\}\times[0,w]} P(T_{k+l+1} > t, T_{k+l} \in [t - v, t], Z_{k+l} = z, T_{k+1} > s|\, T_k = r, Z_k = \xi)|_{r=s-u,\xi=y} \\
&\qquad \cdot\mathcal{L}(Z_k, s - T_k|\, P)(dy, du) \\
&= \sum_{k\in\mathbb{N}_0} \int_{\{\eta\}\times[0,w]} P(T_{l+1} > t, T_l \in [t - v, t], Z_l = z, T_1 > s|\, T_0 = r, Z_0 = \xi)|_{r=s-u,\xi=y} \\
&\qquad \cdot\mathcal{L}(Z_k, s - T_k|\, P)(dy, du),
\end{aligned}
$$

where the last equation is due to the homogeneity of the marked point process (T, Z). Applying the theorem on integration with respect to an image measure as well as (2.4.5) and (2.5.9) again, the above chain of equations can with $A := \{Z_k = \eta, s - T_k \in [0, w]\} \subset \Omega$ be continued as

$$
\begin{aligned}
&= \sum_{k\in\mathbb{N}_0} \int_A P(T_{l+1} > t, T_l \in [t - v, t], Z_l = z, T_1 > s|\, T_0 = r, Z_0 = \xi)|_{r=s-s+T_k,\xi=Z_k}\, dP \\
&= \sum_{k\in\mathbb{N}_0} \int_{A\cap\{N_s=k\}} P(T_{l+1} > t, T_l \in [t - v, t], Z_l = z, T_1 > s|\, T_0 = r, Z_0 = \xi)|_{r=s-s+T_{N_s},\xi=Z_{N_s}}\, dP \\
&= \sum_{k\in\mathbb{N}_0} \int_{\{X_s=\eta,U_s\in[0,w]\}\cap\{N_s=k\}} P(T_{l+1} > t, T_l \in [t - v, t], Z_l = z, T_1 > s|\, T_0 = r, Z_0 = \xi)|_{r=s-U_s,\xi=X_s}\, dP \\
&= \int_{\{X_s=\eta,U_s\in[0,w]\}} P(T_{l+1} > t, T_l \in [t - v, t], Z_l = z, T_1 > s|\, T_0 = r, Z_0 = \xi)|_{r=s-U_s,\xi=X_s}\, dP \\
&= \int_{\{\eta\}\times[0,w]} P(T_{l+1} > t, T_l \in [t - v, t], Z_l = z, T_1 > s|\, T_0 = r, Z_0 = \xi)|_{r=s-u,\xi=y}\, \mathcal{L}(X_s, U_s|\, P)(dy, du).
\end{aligned}
\tag{2.23.8}
$$

According to (2.23.8),

$$
P(T_{l+1} > t, T_l \in [t - v, t], Z_l = z, T_1 > s|\, T_0 = s - u, Z_0 = y) \tag{2.23.9}
$$

satisfies the Radon-Nikodym equation (2.23.7). Further, it is $2^{\mathcal{S}} \otimes \mathfrak{B}([0,\infty))$-measurable. The latter can be confirmed by realizing that (2.23.9) is for $s \geq 0$ the composition of the $2^{\mathcal{S}} \otimes \mathfrak{B}([0,\infty)) - \mathfrak{B}([0,\infty]) \otimes 2^{\mathcal{S}}$-measurable mapping

$$
g_s : \mathcal{S} \times [0,\infty) \ni (y, u) \mapsto g_s(y, u) = (s - u, y) \in [0,\infty] \times \mathcal{S}, \;\; u \leq s
$$

and the $\mathfrak{B}([0,\infty]) \otimes 2^{\mathcal{S}}$-measurable mapping

$$
P(T_{l+1} > t, T_l \in [t - v, t], Z_l = z, T_1 > s|\, T_0 = \cdot, Z_0 = \cdot).
$$

Hence, (2.23.9) is a version of $p^{(l)}_{yz}(s,t,u,v)$. Next, we will verify that for $\mathcal{L}(X_s,U_s|\,P)$-a.e. (y,u)

$$\begin{aligned}&P(T_{l+1}>t,T_l\in[t-v,t],Z_l=z,T_1>s|\,T_0=s-u,Z_0=y)\\&=P(T_{l+1}>t,T_l\in[t-v,t],Z_l=z|\,T_1>s,T_0=s-u,Z_0=y).\end{aligned}\tag{2.23.10}$$

In general, it holds for $\mathcal{L}(T_0,Z_0|\,P)$-a.e. $(r,\xi)\in[0,\infty)\times\mathcal{S}$ that

$$\begin{aligned}&P(T_{l+1}>t,T_l\in[t-v,t],Z_l=z,T_1>s|\,T_0=r,Z_0=\xi)\\&=P(T_{l+1}>t,T_l\in[t-v,t],Z_l=z|\,T_1>s,T_0=r,Z_0=\xi)\,P(T_1>s|\,T_0=r,Z_0=\xi).\end{aligned}\tag{2.23.11}$$

For $C=\{1\}$, this is a consequence the following equation: With $C\in 2^{\{0,1\}}$, one obtains by successive conditioning for $\mathcal{L}(T_0,Z_0|\,P)$-a.e. $(r,\xi)\in[0,\infty)\times\mathcal{S}$

$$\begin{aligned}&\int_C P(T_{l+1}>t,T_l\in[t-v,t],Z_l=z,\mathbf{1}_{\{T_1>s\}}\in d\delta|\,T_0=r,Z_0=\xi)\\&=\int_C P(T_{l+1}>t,T_l\in[t-v,t],Z_l=z|\,\mathbf{1}_{\{T_1>s\}}=\delta,T_0=r,Z_0=\xi)\\&\qquad\cdot P(\mathbf{1}_{\{T_1>s\}}\in d\delta|\,T_0=r,Z_0=\xi).\end{aligned}\tag{2.23.12}$$

Inserting (2.23.11) into (2.23.8) yields

$$\begin{aligned}&P(X_t=z,U_t\le v,N_t-N_s=l,X_s=\eta,U_s\le w)\\&=\int_{\{\eta\}\times[0,w]}P(T_{l+1}>t,T_l\in[t-v,t],Z_l=z,T_1>s|\,T_0=r,Z_0=\xi)|_{r=s-u,\xi=y}\,\mathcal{L}(X_s,U_s|\,P)(dy,du)\\&=\int_{\{\eta\}\times[0,w]}P(T_{l+1}>t,T_l\in[t-v,t],Z_l=z|\,T_1>s,T_0=r,Z_0=\xi)|_{r=s-u,\xi=y}\\&\qquad\cdot P(T_1>s|\,T_0=r,Z_0=\xi)|_{r=s-u,\xi=y}\,\mathcal{L}(X_s,U_s|\,P)(dy,du)\\&=\int_{\{\eta\}\times[0,w]}P(T_{l+1}>t,T_l\in[t-v,t],Z_l=z|\,T_1>s,T_0=s-u,Z_0=y)\\&\qquad\cdot P(T_1>s|\,T_0=s-u,Z_0=y)\,\mathcal{L}(X_s,U_s|\,P)(dy,du).\end{aligned}\tag{2.23.13}$$

As explained below, we have $P(T_1>s|\,T_0=s-u,Z_0=y)=1$ for $\mathcal{L}(X_s,U_s|\,P)$-a.e. (y,u). Taking this into account, we get

$$\begin{aligned}&P(X_t=z,U_t\le v,N_t-N_s=l,X_s=\eta,U_s\le w)\\&=\int_{\{\eta\}\times[0,w]}P(T_{l+1}>t,T_l\in[t-v,t],Z_l=z|\,T_1>s,T_0=s-u,Z_0=y)\,\mathcal{L}(X_s,U_s|\,P)(dy,du).\end{aligned}$$

Hence, $P(T_{l+1}>t,T_l\in[t-v,t],Z_l=z|\,T_1>s,T_0=s-u,Z_0=y)$ also satisfies the Radon-Nikodym equation (2.23.7), and therefore, it is a version of $p^{(l)}_{yz}(s,t,u,v)$.
To complete the proof, we must show that $P(T_1>s|\,T_0=s-u,Z_0=y)=1$ for $\mathcal{L}(X_s,U_s|\,P)$-a.e. (y,u). For this, one must realize that for $t\ge s$, $P(T_1>t|\,T_0=s-u,Z_0=y)$ is a version of $P(T(s)>t|\,X_s=y,\,U_s=u)$ with $T(s)$ being the first jump after s. The verifying Radon-Nikodym equation,

$$\begin{aligned}&P(T(s)>t,X_s=\eta,U_s\le w)\\&=\int_{\{\eta\}\times[0,w]}P(T(s)>t|\,X_s=y,U_s=u)\,\mathcal{L}(X_s,U_s|\,P)(dy,du)\\&=\int_{\{\eta\}\times[0,w]}P(T_1>t|\,T_0=s-u,Z_0=y)\,\mathcal{L}(X_s,U_s|\,P)(dy,du),\end{aligned}\tag{2.23.14}$$

can be proved analogously to (2. 23. 8). Hence, for $\mathcal{L}(X_s, U_s|\, P)$-a.e. (y, u)

$$P(T_1 > s|\, T_0 = s - u, Z_0 = y) = P(T(s) > s|\, X_s = y,\, U_s = u) = 1,$$

where the second equation is due to the definition of $T(s)$ according to (2. 5. 5). □

In case of the pure jump process $(X_t)_{t\geq 0}$ being Markovian - which is equivalent to the situation of lemma 2. 19, meaning that the regular cumulative transition intensities of the marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$ satisfy the additivity property (2. 11. 2) - the transition probabilities

$$p_{yz}(s,t,u,v) \overset{\text{a.s.}}{=} P(\exists l \in \mathbb{N}_0 : T_{l+1} > t, T_l \in [t-v,t], Z_l = z|\, T_1 > s, T_0 = s-u, Z_0 = y)$$

can according to the last assertion of lemma 2. 19 be chosen as independent of u, namely by means of

$$p_{yz}(s,t,u,v) \overset{\text{a.s.}}{=} P(\exists l \in \mathbb{N}_0 : T_{l+1} > t, T_l \in [t-v,t], Z_l = z|\, T_0 = s, Z_0 = y).$$

Further, if $v = \infty$, one gets

$$p_{yz}(s,t,u,\infty) \overset{\text{a.s.}}{=} P(\exists l \in \mathbb{N}_0 : T_l \leq t < T_{l+1}, Z_l = z|\, T_0 = s, Z_0 = y), \qquad (2.\,23.\,15)$$

where the right-hand side forms a regular version of the transition probabilities $p_{yz}(s,t)$ of the Markovian pure jump process $(X_t)_{t\geq 0}$ (cf. theorem 2. 42 or Milbrodt and Helbig [1999], Folgerung 4.41).

In order to adapt the multiple decrement model for life insurance to a semi-Markov framework, we define (regular) cumulative transition intensities q_* for the bivariate Markov process $((X_t, U_t))_{t\geq 0}$. It will be stated that, for a given regular cumulative transition intensities matrix q, there exists a semi-Markov process $(X_t)_{t\geq 0}$ with the appertaining bivariate Markov process $((X_t, U_t))_{t\geq 0}$ possessing the cumulative transition intensity matrix q (theorem 2. 29). Note that a regular cumulative transition intensity matrix can, for example, be obtained from given select-and-ultimate tables (cf. figure 4). Further, a regular cumulative transition intensities matrix q along with an initial distribution fully determines the distribution of a semi-Markovian pure jump process $(X_t)_{t\geq 0}$ as well as the distribution of the appertaining bivariate Markov process $((X_t, U_t))_{t\geq 0}$. With the aid of cumulative intensities q_*, backward and forward integral equations for the transition probabilities of $((X_t, U_t))_{t\geq 0}$ will be established. These equations turn out to be generalizations of the corresponding formulas in the non-smooth Markov set-up by Milbrodt and Helbig ([1999], section 4C). In order to point out the differences, the following is organized in a manner similar to the corresponding sections in Milbrodt and Helbig ([1999], sections 4B and 4C).

Recall that for $s \geq 0$, the time of the first jump after s is denoted by $T(s)$. $X_{T(s)}$ denotes the corresponding destination state. Next, we define the conditional distribution of $T(s)$ and $X_{T(s)}$, given the current state X_s as well as the time elapsed since entering this state U_s. Further, the appertaining cumulative intensities are defined (cf. definition 2. 8, and Milbrodt and Helbig [1999], definition 4.28).

2. 24 Definition. For $(y, z) \in \mathcal{J}$ and $0 \leq u \leq s$ we define

$$Q_{yz}(s,\cdot,u) :\ [s,\infty) \ni t \mapsto P(T(s) \leq t, X_{T(s)} = z|\, X_s = y,\, U_s = u), \qquad (2.\,24.\,1)$$

$$Q_y(s,\cdot,u) :\ [s,\infty) \ni t \mapsto \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} Q_{yz}(s,t,u) = P(T(s) \leq t|\, X_s = y,\, U_s = u), \qquad (2.\,24.\,2)$$

$$\bar{p}_y(s,t,u) :\ = 1 - Q_y(s,t,u) = P(X_\tau = y, \tau \in [s,t]|\, X_s = y, U_s = u),\ t \geq s, \quad (2.\,24.\,3)$$

and, with the convention $0/0 := 0$,

$$q_{yz}(s,\cdot,u):\ [s,\infty)\ni t\mapsto \int_{(s,t]}\frac{Q_{yz}(s,d\tau,u)}{1-Q_y(s,\tau-0,u)}\in[0,\infty], \tag{2.24.4}$$

$$q_{yy}(s,\cdot,u):\ [s,\infty)\ni t\mapsto -\int_{(s,t]}\frac{Q_{y}(s,d\tau,u)}{1-Q_y(s,\tau-0,u)}\in[-\infty,0]. \tag{2.24.5}$$

Analogously to the cumulative intensities for the marked point process, q_{yz} is referred to as the *cumulative transition intensity from y to z* and $-q_{yy}$ is the *cumulative intensity of decrement for state y*. The matrix

$$q(s,t,u) := \big(q_{yz}(s,t,u)\big)_{(y,z)\in\mathcal{S}^2},\quad 0\le u\le s\le t<\infty \tag{2.24.6}$$

is referred to as *cumulative transition intensity matrix for the bivariate process* $((X_t,U_t))_{t\ge0}$.

All quantities in the above definition are uniquely determined up to $\mathcal{L}(X_s,U_s|\,P)$-null sets. $\bar p_y(s,t,u)$ is the conditional probability of staying uninterruptedly in state y from time s to time t, given that this state was entered at time $s-u$. For this conditional probability, an exponential formula similar to (2.10.7) is given by

$$\bar p_y(s,t,u) = \exp(q^{(c)}_{yy}(s,t,u))\prod_{s<\tau\le t}(1+\Delta\, q_{yy}(s,\tau,u)),\ \ 0\le u\le s\le t. \tag{2.24.7}$$

Before we turn to a lemma that gathers - as counterpart to Hilfssatz 4.29 in Milbrodt and Helbig [1999] - properties of the cumulative transition intensities for the process $((X_t,U_t))_{t\ge0}$, the relationships between the quantities specified by definition 2.24 and the corresponding quantities with respect to the marked point process (T,Z) (cf. definition 2.8) are investigated. It turns out that the conditional probabilities Q_* can be expressed in terms of the transition probabilities of the marked point process $\hat Q_*$, and vice versa. In almost the same manner, the cumulative transition intensities q_* can be expressed in terms of $\hat q_*$. Here, the converse also holds.

2.25 Lemma. *Let $(\Omega,\mathfrak{F},P,((T_m,Z_m))_{m\in\mathbb{N}_0})$ be a homogeneous Markovian marked point process with transition probabilities $\hat Q_*$ as well as cumulative transition intensities $\hat q_*$ according to definition 2.8. $((X_t,U_t))_{t\ge0}$ denotes the appertaining bivariate Markov process with Q_* and q_* according to definition 2.24. Further, let $0\le u\le s\le t<\infty$, and $(y,z)\in\mathcal{J}$. Then, one obtains for $\mathcal{L}(X_s,U_s|\,P)$-a.e. (y,u)*

$$Q_{yz}(s,t,u) = \hat Q_{yz}(s-u,(s,t]), \tag{2.25.1}$$

and

$$q_{yz}(s,t,u) = \hat q_{yz}(s-u,(s,t]). \tag{2.25.2}$$

Conversely, for each $m\in\mathbb{N}_0$ and $\mathcal{L}(T_m,Z_m|\,P)$-a.e. (s,y)

$$\hat Q_{yz}(s,t) = Q_{yz}(s,t,0), \tag{2.25.3}$$

and

$$\hat q_{yz}(s,t) = q_{yz}(s,t,0). \tag{2.25.4}$$

Proof. Firstly, we verify (2.25.1). Let $0\le u\le s\le t<\infty$, and $(y,z)\in\mathcal{J}$. According to definition 2.8 and the homogeneity of the marked point process (T,Z), we have

$$\begin{aligned}\hat Q_{yz}(s-u,(s,t]) &\overset{\text{a.s.}}{=} P(T_{m+1}\in(s,t],Z_{m+1}=z|\,T_m=s-u,\,Z_m=y)\\ &= P(T_1\in(s,t],Z_1=z|\,T_0=s-u,\,Z_0=y).\end{aligned} \tag{2.25.5}$$

Further, the left-hand side of (2. 25. 1) is given by

$$Q_{yz}(s,t,u) \overset{\text{a.s.}}{=} P(T(s) \le t, X_{T(s)} = z|\, X_s = y,\, U_s = u). \tag{2.25.6}$$

To verify (2. 25. 1), it must be shown that the right-hand side of (2. 25. 5) forms a version of the right-hand side of (2. 25. 6). This can be done in almost the same manner as in the proof of theorem 2. 23. Therefore, we restrict ourselves to the verification of the corresponding Radon-Nikodym equation, which is for $\{\eta\} \times [0,w] \in \mathfrak{D}_0$ given by

$$\begin{aligned} &P(T(s) \le t, X_{T(s)} = z, X_s = \eta, U_s \le w) \\ &= \int_{\{\eta\}\times[0,w]} P(T(s) \le t, X_{T(s)} = z|\, X_s = y, U_s = u)\, \mathcal{L}(X_s, U_s|\, P)(dy, du). \end{aligned} \tag{2.25.7}$$

Starting from the left-hand side and employing (2. 4. 5), (2. 5. 5), (2. 5. 9), and (2. 4. 6) gives

$$\begin{aligned} &P(T(s) \le t, X_{T(s)} = z, X_s = \eta, U_s \le w) \\ &= P(T_{N_s+1} \le t, Z_{N_s+1} = z, Z_{N_s} = \eta, s - T_{N_s} \le w) \\ &= \sum_{k\in\mathbb{N}_0} P(\{N_s = k\} \cap \{T_{k+1} \le t, Z_{k+1} = z, Z_k = \eta, s - T_k \le w\}) \\ &= \sum_{k\in\mathbb{N}_0} P(\{T_k \le s < T_{k+1}\} \cap \{T_{k+1} \le t, Z_{k+1} = z, Z_k = \eta, s - T_k \le w\}) \\ &= \sum_{k\in\mathbb{N}_0} P(s < T_{k+1} \le t, Z_{k+1} = z, Z_k = \eta, s - T_k \in [0,w]) \\ &= \sum_{k\in\mathbb{N}_0} \int_{\{\eta\}\times[0,w]} P(T_{k+1} \in (s,t], Z_{k+1} = z|\, Z_k = y, s - T_k = u)\, \mathcal{L}(Z_k, s - T_k|\, P)(dy, du) \\ &= \sum_{k\in\mathbb{N}_0} \int_{\{\eta\}\times[0,w]} P(T_{k+1} \in (s,t], Z_{k+1} = z|\, T_k = r, Z_k = \xi)|_{r=s-u,\xi=y}\, \mathcal{L}(Z_k, s - T_k|\, P)(dy, du) \\ &= \sum_{k\in\mathbb{N}_0} \int_{\{\eta\}\times[0,w]} P(T_1 \in (s,t], Z_1 = z|\, T_0 = r, Z_0 = \xi)|_{r=s-u,\xi=y}\, \mathcal{L}(Z_k, s - T_k|\, P)(dy, du), \\ &= \sum_{k\in\mathbb{N}_0} \int_{\{Z_k=\eta, s-T_k\in[0,w]\}} P(T_1 \in (s,t], Z_1 = z|\, T_0 = r, Z_0 = \xi)|_{r=s-s+T_k,\xi=Z_k}\, dP \\ &= \sum_{k\in\mathbb{N}_0} \int_{\{Z_{N_s}=\eta, s-T_{N_s}\in[0,w]\}\cap\{N_s=k\}} P(T_1 \in (s,t], Z_1 = z|\, T_0 = r, Z_0 = \xi)|_{r=s-s+T_{N_s},\xi=Z_{N_s}}\, dP \\ &= \sum_{k\in\mathbb{N}_0} \int_{\{X_s=\eta, U_s\in[0,w]\}\cap\{N_s=k\}} P(T_1 \in (s,t], Z_1 = z|\, T_0 = r, Z_0 = \xi)|_{r=s-U_s,\xi=X_s}\, dP \\ &= \int_{\{X_s=\eta, U_s\in[0,w]\}} P(T_1 \in (s,t], Z_1 = z|\, T_0 = r, Z_0 = \xi)|_{r=s-U_s,\xi=X_s}\, dP \\ &= \int_{\{\eta\}\times[0,w]} P(T_1 \in (s,t], Z_1 = z|\, T_0 = r, Z_0 = \xi)|_{r=s-u,\xi=y}\, \mathcal{L}(X_s, U_s|\, P)(dy, du) \\ &= \int_{\{\eta\}\times[0,w]} P(T_1 \in (s,t], Z_1 = z|\, T_0 = s - u, Z_0 = y)\, \mathcal{L}(X_s, U_s|\, P)(dy, du). \end{aligned} \tag{2.25.8}$$

Hence, we have for $\mathcal{L}(X_s, U_s|\, P)$-a.e. (y,u)

$$P(T(s) \le t, X_{T(s)} = z|\, X_s = y, U_s = u) = P(T_1 \in (s,t], Z_1 = z|\, T_0 = s - u, Z_0 = y),$$

which verifies (2. 25. 1). Analogously, one obtains for $\mathcal{L}(X_s, U_s|\, P)$-a.e. (y,u)

$$\begin{aligned} 1 - Q_{yz}(s,t,u) &= P(T(s) > t, X_{T(s)} = z|\, X_s = y,\, U_s = u) \\ &= P(T_1 > t, Z_1 = z|\, T_0 = s - u, Z_0 = y) \\ &= 1 - \hat{Q}_{yz}(s - u, t), \end{aligned} \tag{2.25.9}$$

and with that, due to (2. 8. 2) as well as (2. 24. 2),

$$1 - Q_y(s,t,u) = 1 - \hat{Q}_y(s-u,t). \tag{2. 25. 10}$$

According to their definition (2. 24. 4), the cumulative transition intensities q_{yz} are determined by the conditional probabilities Q_{yz}. Hence, (2. 25. 2) is a consequence of (2. 25. 1): By employing (2. 25. 1) and (2. 25. 10), we obtain for $\mathcal{L}(X_s, U_s|\, P)$-a.e. (y,u)

$$q_{yz}(s,t,u) \;=\; \int_{(s,t]} \frac{Q_{yz}(s,d\tau,u)}{1-Q_y(s,\tau-0,u)} = \int_{(s,t]} \frac{\hat{Q}_{yz}(s-u,d\tau)}{1-\hat{Q}_y(s-u,\tau-0)} = \int_{(s,t]} \hat{q}_{yz}(s-u,d\tau),$$

where the last equation is due to the definition of $\hat{q}_{yz}$. This yields (2. 25. 2).
Regarding (2. 25. 3), we will verify that

$$Q_{yz}(s,t,0) = P(T(s) \le t, X_{T(s)} = z|\, X_s = y,\, U_s = 0) \tag{2. 25. 11}$$

satisfies the Radon-Nikodym equation for $\hat{Q}_{yz}(s,t)$, which is for $m \in \mathbb{N}_0$ given by

$$\begin{aligned} &P(T_{m+1} \le t, Z_{m+1} = z, T_m \le w, Z_m = \eta) \\ &= \int_{[0,w]\times\{\eta\}} P(T_{m+1} \le t, Z_{m+1} = z|\, T_m = s, Z_m = y)\, \mathcal{L}(T_m, Z_m|P)(ds,dy). \end{aligned} \tag{2. 25. 12}$$

In order to verify that (2. 25. 11) solves the above equation, it is inserted into the right-hand side of (2. 25. 12). Applying (2. 4. 5), (2. 5. 5), as well as (2. 5. 9) gives

$$\begin{aligned} &\int_{[0,w]\times\{\eta\}} P(T(s) \le t, X_{T(s)} = z|\, X_s = y,\, U_s = 0)\; \mathcal{L}(T_m, Z_m|P)(ds,dy) \\ &= \int_{[0,w]\times\{\eta\}} P(T_{N_s+1} \le t, Z_{N_s+1} = z|\, Z_{N_s} = y,\, s - T_{N_s} = 0)\; \mathcal{L}(T_m, Z_m|P)(ds,dy) \\ &= \int_{\{T_m\in[0,w],Z_m=\eta\}} P(T_{N_s+1} \le t, Z_{N_s+1} = z|\, T_{N_s} = s, Z_{N_s} = y)|_{s=T_m,y=Z_m}\; dP, \end{aligned}$$

which is due to the homogeneity of the marked point process (T,Z) equal to

$$\begin{aligned} &= \int_{\{T_m\in[0,w],Z_m=\eta\}} P(T_{m+1} \le t, Z_{m+1} = z|\, T_m = s, Z_m = y)|_{s=T_m,y=Z_m}\; dP \\ &= \int_{[0,w]\times\{\eta\}} P(T_{m+1} \le t, Z_{m+1} = z|\, T_m = s, Z_m = y)\; \mathcal{L}(T_m, Z_m|P)(ds,dy). \end{aligned}$$

The last equation yields the right-hand side of (2. 25. 12). Hence, (2. 25. 11) satisfies the Radon-Nikodym equation for $\hat{Q}_{yz}(s,t)$.

The assertion concerning the corresponding cumulative intensities likewise follows from their definitions (2. 24. 4) and (2. 8. 3):

$$\hat{q}_{yz}(s,t) = \int_{(s,t]} \frac{\hat{Q}_{yz}(s,d\tau)}{1-\hat{Q}_y(s,\tau-0)} = \int_{(s,t]} \frac{Q_{yz}(s,d\tau,0)}{1-Q_y(s,\tau-0,0)} = q_{yz}(s,t,0).$$

□

2. 26 Remark. Note that in addition to (2. 25. 1), $Q_{yz}(s,t,u)$, $0 \le u \le s \le t$, $(y,z) \in \mathcal{J}$ also satisfies for $\mathcal{L}(X_s, U_s|\, P)$-a.e. (y,u) the following:

$$Q_{yz}(s,t,u) = P(T_1 \le t, Z_1 = z|\, T_1 > s, T_0 = s-u,\, Z_0 = y). \tag{2. 26. 1}$$

By using (2. 25. 1), this can be confirmed in the same manner as (2. 23. 10). Further, similarly to (2. 25. 3), it can be proved that for $m \in \mathbb{N}_0$ and $\mathcal{L}(T_m, Z_m | P)$-a.e. (s, y)

$$\begin{aligned}
&P(\exists l \in \mathbb{N}_0 : T_{m+l+1} > t, T_{m+l} \in [t-v, t], Z_{m+l} = z | T_m = s, Z_m = y) \\
&= P(\exists l \in \mathbb{N}_0 : T_{l+1} > t, T_l \in [t-v, t], Z_l = z | T_1 > s, T_0 = s, Z_0 = y) \\
&= p_{yz}(s, t, 0, v). \qquad (2.26.2)
\end{aligned}$$

We now formulate the properties of the cumulative transition intensities.

2. 27 Lemma. *For* $0 \le u \le s$ *and* $(y, z) \in \mathcal{J}$ *let* $q_{yz}(s, \cdot, u)$ *and* $q_{yy}(s, \cdot, u)$ *be cumulative intensities according to definition* 2. 24. *Then, for* $\mathcal{L}(X_s, U_s | P)$*-a.e.* (y, u)

$$q_{yz}(s, r, u) + q_{yz}(r, t, r - s + u) = q_{yz}(s, t, u), \quad s \le r \le t, \qquad (2.27.1)$$

$$q_{yz}(s, s, u) = 0, \qquad (2.27.2)$$

$$\lim_{h \searrow 0} q_{yz}(s, t + h, u) = q_{yz}(s, t, u), \quad s \le t \qquad (2.27.3)$$

$$q_{yz}(s, t, u) \ge 0, \quad \text{and} \quad q_{yy}(s, t, u) \le 0, \quad s \le t, \qquad (2.27.4)$$

$$-q_{yy}(s, t, u) = \sum_{\substack{z \in \mathcal{S} \\ z \ne y}} q_{yz}(s, t, u), \quad s \le t, \qquad (2.27.5)$$

$$-q_{yy}(s, \{t\}, u) \le 1, \text{ and } -q_{yy}(s, \{t\}, u) = 1 \Longrightarrow -q_{yy}(s, \tau, u) = -q_{yy}(s, t, u), \tau \ge t \ge s. \qquad (2.27.6)$$

The properties (2. 27. 2) - (2. 27. 6) are similar to the corresponding properties of cumulative transition intensities for a Markovian pure jump process as stated by Milbrodt and Helbig ([1999], Hilfssatz 4.29: (4.29.2) - (4.29.7)). In principle, the dependence on the time spent in the current state must simply be added. However, the property (4.29.1) there,

$$q_{yz}(s, r) + q_{yz}(r, t) = q_{yz}(s, t), \quad s \le r \le t,$$

(see also (2. 11. 2)) has to be adapted in a different way, leading to (2. 27. 1). The distinction between these two properties is due to the fact that in a semi-Markov approach the time elapsed since the current state was entered must always be taken into account. In contrast, in a Markov set-up, it is only the state at a certain time that matters. The property (2. 27. 1) can be proved by employing the relation (2. 25. 2).

Proof. (of lemma 2. 27). Let $0 \le u \le s \le r \le t$ and $(y, z) \in \mathcal{J}$. Regarding (2. 27. 1), we get by employing (2. 25. 2)

$$\begin{aligned}
q_{yz}(s, t, u) &= \hat{q}_{yz}(s - u, (s, t]) \\
&= \int_{(s,t]} \hat{q}_{yz}(s - u, d\tau) \\
&= \int_{(s,r]} \hat{q}_{yz}(s - u, d\tau) + \int_{(r,t]} \hat{q}_{yz}(s - u, d\tau) \\
&= \int_{(s,r]} \hat{q}_{yz}(s - u, d\tau) + \int_{(r,t]} \hat{q}_{yz}(r - r + s - u, d\tau). \\
&= \hat{q}_{yz}(s - u, (s, r]) + \hat{q}_{yz}(r - r + s - u, (r, t]). \qquad (2.27.7)
\end{aligned}$$

Another application of (2. 25. 2) yields $\hat{q}_{yz}(r - r + s - u, (r, t]) = q_{yz}(r, t, r - s + u)$. Inserting this into (2. 27. 7) gives

$$q_{yz}(s, t, u) = \hat{q}_{yz}(s - u, (s, r]) + q_{yz}(r, t, r - s + u) = q_{yz}(s, r, u) + q_{yz}(r, t, r - s + u),$$

which is (2. 27. 1).
The properties (2. 27. 2) - (2. 27. 6) follow immediately from definition 2. 24, with (2. 27. 3) being arrived at by means of lemma A.4 (cf. appendix). □

Regarding the conditional probabilities of remaining in a certain state $\bar{p}_y, y \in \mathcal{S}$, the property (2. 27. 1) yields along with the exponential formula (2. 24. 7) the following: For $\mathcal{L}(X_s, U_s|\, P)$- a.e. $(y, u) \in \mathcal{S} \times [0, \infty)$

$$\bar{p}_y(s,t,u) = \bar{p}_y(s,r,u)\,\bar{p}_y(r,t,r-s+u), \;\; s \le r \le t. \tag{2.27.8}$$

Likewise in order to avoid difficulties with null sets, the concept of regular cumulative transition intensities is also adapted to the bivariate Markov process $((X_t, U_t))_{t\ge 0}$. Similarly to definition 2. 11, we define a cumulative transition intensity matrix as follows.

2. 28 Definition. A matrix-valued map

$$q = (q_{yz})_{(y,z)\in\mathcal{S}^2} : \{(s,t,u)|\, 0 \le u \le s \le t < \infty\} \to \mathbb{R}^{\mathcal{S}^2} \tag{2.28.1}$$

possessing the properties (2. 27. 1) - (2. 27. 6) without exceptional sets is called a *regular cumulative transition intensity matrix for a bivariate process* $((X_t, U_t))_{t\ge 0}$.

Given such a regular cumulative transition intensity matrix q, one obtains by means of

$$q_{yz}(s,(r,t],u) := q_{yz}(s,t,u) - q_{yz}(s,r,u), \;\; 0 \le u \le s \le r \le t < \infty, (y,z) \in \mathcal{S}^2$$

Borel measures on $\mathfrak{B}([0,\infty))$, for $s \ge 0$ concentrated on (s,∞). These measures are related according to (2. 27. 1). This means, for example, for $0 \le u \le s \le r \le t$ and $(y,z) \in \mathcal{S}^2$ the following:

$$\begin{aligned} q_{yz}(s,(r,t],u) &= q_{yz}(s,t,u) - q_{yz}(s,r,u) \\ &= q_{yz}(s,r,u) + q_{yz}(r,t,r-s+u) - q_{yz}(s,r,u) \\ &= q_{yz}(r,(r,t],r-s+u). \end{aligned} \tag{2.28.2}$$

Hence, the measures $q_{yz}(s, d\tau, u)$ and $q_{yz}(r, d\tau, r-s+u)$ coincide on (r,∞).

The question of whether or not there exists a regular cumulative transition intensity matrix q for the process $((X_t, U_t))_{t\ge 0}$ can be answered by referring lemma 2. 25. Due to (2. 25. 2), a regular cumulative transition intensity matrix $\hat{q}$ for the appertaining marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$ determines a cumulative transition intensity matrix q for the process $((X_t, U_t))_{t\ge 0}$. This intensity matrix q satisfies the requirements for a regular cumulative transition intensity matrix for the bivariate Markov process. More precisely, a regular cumulative transition intensity matrix q can be specified by means of

$$q(s,t,u) := \hat{q}(s-u,(s,t]), \;\; 0 \le u \le s \le t < \infty, \tag{2.28.3}$$

with $\hat{q}$ being a regular cumulative transition intensity matrix for the appertaining marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$. Then, according to lemma 2. 10 and (2. 27. 7), the properties (2. 27. 1) - (2. 27. 6) are identically satisfied.

For actuarial modelling, the (cumulative) transition intensities usually form the starting point of consideration. Thus, most of the calculations in the sequel are based on a given regular cumulative transition intensity matrix q for the bivariate Markov process $((X_t, U_t))_{t\ge 0}$. The appertaining versions of the conditional distributions Q_* can then for $0 \le u \le s \le t < \infty$ and $(y,z) \in \mathcal{J}$ be obtained as

$$Q_{yz}(s,t,u) = \int_{(s,t]} \bar{p}_y(s,\tau-0,u)\, q_{yz}(s,d\tau,u), \tag{2.28.4}$$

and

$$Q_y(s,t,u) = -\int_{(s,t]} \bar{p}_y(s,\tau-0,u)\, q_{yy}(s,d\tau,u), \tag{2.28.5}$$

where the integrand $\bar{p}_y(s,\cdot,u) = 1 - Q_y(s,\cdot,u)$ is specified by the exponential formula (2. 24. 7) and the cumulative intensity of decrement $-q_{yy}$. As we point out next, a given regular cumulative transition intensity matrix q even solves the modelling problem, meaning that there is a semi-Markov process $(X_t)_{t\geq 0}$ with the appertaining bivariate process $((X_t,U_t))_{t\geq 0}$ possessing the cumulative transition intensity matrix q. Further, the distribution of a semi-Markov process and of the appertaining bivariate Markov process $((X_t,U_t))_{t\geq 0}$ are uniquely specified by a regular cumulative transition intensity matrix and an initial distribution.

2. 29 Theorem. *Let π be a probability measure on $2^{\mathcal{S}}$ and $q = (q_{yz})_{(y,z)\in\mathcal{S}^2}$ a regular cumulative transition intensity matrix according to definition* 2. 28. *Then:*

- *There is a semi-Markov process $(\Omega, \mathfrak{F}, P, (X_t)_{t\geq 0})$ such that the appertaining bivariate Markov process $((X_t,U_t))_{t\geq 0}$ possesses the intitial distribution $\mathcal{L}(X_0,U_0) = \pi\otimes\varepsilon_0$ and the cumulative transition intensity matrix q.*
- *The distribution of a semi-Markov process $(X_t)_{t\geq 0}$ as well as the distribution of the appertaining bivariate Markov process $((X_t,U_t))_{t\geq 0}$ are uniquely determined by π and q.*

Proof. In view of (2. 25. 4), we specify by

$$\hat{q}_{yz}(s,t) := q_{yz}(s,t,0), \;\; 0\leq s\leq t<\infty, (y,z)\in\mathcal{J}, \tag{2.29.1}$$

cumulative transition intensities for a marked point process (T,Z). Since the cumulative transition intensities $q_{yz}, (y,z)\in\mathcal{J}$, are assumed to be regular, the cumulative transition intensities $\hat{q}_{yz}$, $(y,z)\in\mathcal{S}^2$, form a regular cumulative transition intensity matrix for a marked point process (T,Z) (cf. lemma 2. 10). According to theorem 2. 12, there exists a homogeneous Markovian marked point process $((T_m,Z_m))_{m\in\mathbb{N}_0}$ with cumulative transition intensity matrix $\hat{q} = (\hat{q}_{yz})_{(y,z)\in\mathcal{S}^2}$ and initial distribution $\mathcal{L}(T_0,Z_0|P) := \varepsilon_0\otimes\pi$. Further, the distribution of this process is uniquely determined by π and $\hat{q}$. Hence, according to $X = G^{-1}(T,Z)$, the appertaining pure jump process exists and its distribution is likewise uniquely determined. As the appertaining pure jump process for a given homogeneous Markovian marked point process, $(X_t)_{t\geq 0}$ is by definition a semi-Markov process.
Now consider the appertaining bivariate process $((X_t,U_t))_{t\geq 0}$, which is according to lemma 2. 21 a Markov process. (As we will see later (corollary 2. 33), this process is even strong Markovian.) The distribution of the process $((X_t,U_t))_{t\geq 0}$ is also specified by the distribution of the above marked point process (for the corresponding isomorphism see lemma 2. 4). Further, by means of (2. 28. 3), regular cumulative transition intensities $\bar{q}_*$ for the process $((X_t,U_t))_{t\geq 0}$ are given by

$$\bar{q}_{yz}(s,t,u) := \hat{q}_{yz}(s-u,(s,t]), \;\; 0\leq u\leq s\leq t<\infty, (y,z)\in\mathcal{J}. \tag{2.29.2}$$

Since the right-hand side of (2. 29. 2) is specified by means of (2. 29. 1), we have

$$\hat{q}_{yz}(s-u,(s,t]) = q_{yz}(s-u,t,0) - q_{yz}(s-u,s,0), \;\; 0\leq u\leq s\leq t<\infty, (y,z)\in\mathcal{J}.$$

The cumulative transition intensities q_* are assumed to be regular. Hence, the property (2. 27. 1) holds without exceptional sets. Using this, we finally get

$$\begin{aligned}\bar{q}_{yz}(s,t,u) &= q_{yz}(s-u,t,0) - q_{yz}(s-u,s,0)\\ &= q_{yz}(s-u,s,0) + q_{yz}(s,t,u) - q_{yz}(s-u,s,0)\\ &= q_{yz}(s,t,u), \;\; 0\leq u\leq s\leq t<\infty.\end{aligned}$$

Hence, the process $((X_t, U_t))_{t\geq 0}$ possesses the cumulative transition intensity matrix q. □

We finish this section by verifying that the strong Markov property of the process $((X_t, U_t))_{t\geq 0}$ that appertains to a pure jump process $(X_t)_{t\geq 0}$ implies the appertaining marked point process to be a homogeneous Markov chain (cf. figure 11). Yet before, a sufficient condition for the process $((X_t, U_t))_{t\geq 0}$ to be strong Markovian is given. For some general information concerning the strong Markov property, we refer to section 2B.

2. 30 Lemma. *Let $(X_t, \mathfrak{F}_t)_{t\geq 0}$ be a pure jump process and $((X_t, U_t))_{t\geq 0}$ the appertaining bivariate process being Markovian with regular transition probability matrix p, the components of which satisfy for $0 \leq u \leq s \leq t < \infty, v \geq 0, (y, z) \in \mathcal{S}^2$*

$$\lim_{h\searrow 0, k\searrow 0} p_{yz}(s+h, t+k, u+h, v) = p_{yz}(s, t, u, v). \tag{2.30.1}$$

Then, the process $((X_t, U_t), \mathfrak{F}_t)_{t\geq 0}$ is strong Markovian, meaning that for any $(\mathfrak{F}_t)_{t\geq 0}$-stopping time T $p_{yz}(s, \tau+s, u, v)|_{s=T,(y,u)=(X_T,U_T)}$ is $\mathfrak{F}_T$-measurable and satisfies

$$P(X_{\tau+T} = z, U_{\tau+T} \leq v | \mathfrak{F}_T) \;=\; p_{yz}(s, \tau+s, u, v)|_{s=T,(y,u)=(X_T,U_T)} \;\; P-a.s. \tag{2.30.2}$$

Proof. Let $0 \leq u \leq s \leq t < \infty, v \geq 0, (y, z) \in \mathcal{S}^2$. As a regular conditional probability, $p_{yz}(s, t, u, v)$ is for fixed (s, t, v) a $2^{\mathcal{S}} \otimes \mathfrak{B}([0, \infty))$-measurable function of $(y, u) \in \mathcal{S} \times [0, \infty)$. Further, it follows from the right continuity (2. 30. 1) that for each v, $p_{yz}(s, t, u, v)$ is even $2^{\mathcal{S}} \otimes \mathfrak{B}([0, \infty)) \otimes \mathfrak{B}([0, \infty)) \otimes \mathfrak{B}([0, \infty))$-measurable on $\{(y, u, s, t) | \, y \in \mathcal{S}, 0 \leq u \leq s < t < \infty\}$. Since the paths of $((X_t, U_t))_{t\geq 0}$ are likewise right continuous, this process is progressively measurable. Hence, for any $(\mathfrak{F}_t)_{t\geq 0}$-stopping time T, the pair (X_T, U_T) is $\mathfrak{F}_T$-measurable. Along with the measurability of $p_{yz}(s, t, u, v)$, the $\mathfrak{F}_T$-measurability of $p_{yz}(s, \tau+s, u, v)|_{s=T,(y,u)=(X_T,U_T)}$ is granted.

Next we verify (2. 30. 2) in a manner similar to the proof of Satz 4.27 in Milbrodt and Helbig [1999], namely by approximating a $(\mathfrak{F}_t)_{t\geq 0}$-stopping time T by a decreasing sequence of $(\mathfrak{F}_t)_{t\geq 0}$-stopping times T_n taking at most countable many values. For $n \in \mathbb{N}$, T_n can be specified by

$$T_n := \sum_{k=1}^{\infty} \frac{k}{2^n} \cdot \mathbf{1}_{\{\frac{k-1}{2^n} < T \leq \frac{k}{2^n}\}} + \infty \cdot \mathbf{1}_{\{T=\infty\}}.$$

By employing the fact that the strong Markov property always holds for stopping times with countable range (cf. (2. 7. 7)), we obtain

$$\begin{aligned}
P(X_{\tau+T} = z, U_{\tau+T} \leq v | \mathfrak{F}_T) &= \lim_{n\to\infty} P(X_{\tau+T_n} = z, U_{\tau+T_n} \leq v | \mathfrak{F}_T) \\
&= \lim_{n\to\infty} \mathbf{E}\left[\mathbf{E}\big[\mathbf{1}_{\{X_{\tau+T_n}=z, U_{\tau+T_n}\leq v\}} | \mathfrak{F}_{T_n}\big] | \mathfrak{F}_T\right] \\
&= \lim_{n\to\infty} \mathbf{E}\left[P\left(X_{\tau+T_n} = z, U_{\tau+T_n} \leq v | \mathfrak{F}_{T_n}\right) | \mathfrak{F}_T\right] \\
&= \lim_{n\to\infty} \mathbf{E}\left[p_{yz}(s, \tau+s, u, v)|_{s=T_n,(y,u)=(X_{T_n},U_{T_n})} | \mathfrak{F}_T\right] \\
&= \lim_{n\to\infty} \mathbf{E}\left[p_{X_{T_n}z}(T_n, \tau+T_n, U_{T_n}, v) | \mathfrak{F}_T\right] \\
&= \mathbf{E}\left[\lim_{n\to\infty} p_{X_{T_n}z}(T_n, \tau+T_n, U_{T_n}, v) | \mathfrak{F}_T\right],
\end{aligned}$$

where the last equation is due to the dominated convergence theorem. According to the properties of the paths of $((X_t, U_t))_{t\geq 0}$, we get for each $\omega \in \Omega$ and sufficiently large n

$$X_{T_n(\omega)}(\omega) = X_{T(\omega)}(\omega) \quad \text{and} \quad U_{T_n(\omega)}(\omega) = U_{T(\omega)}(\omega) + T_n(\omega) - T(\omega).$$

By further employing (2.30.1), the above chain of equations can be continued as

$$
\begin{aligned}
&= \mathbf{E}\left[\lim_{n\to\infty} p_{X_T z}(T_n, \tau + T_n, U_T + (T_n - T), v) | \mathfrak{F}_T\right] \\
&= \mathbf{E}\left[p_{yz}(s, \tau + s, u, v)|_{s=T,(y,u)=(X_T,U_T)} | \mathfrak{F}_T\right].
\end{aligned}
$$

Using the $\mathfrak{F}_T$-measurability of $p_{yz}(s, \tau + s, u, v)|_{s=T,(y,u)=(X_T,U_T)}$, we finally get

$$
P(X_{\tau+T} = z, U_{\tau+T} \le v | \mathfrak{F}_T) = p_{yz}(s, \tau + s, u, v)|_{s=T,(y,u)=(X_T,U_T)} \quad P - a.s.,
$$

which is (2.30.2). □

We turn to the announced implication of the strong Markov property of the bivariate process $((X_t, U_t))_{t\ge 0}$.

2.31 Lemma. *Let $(X_t, \mathfrak{F}_t)_{t\ge0}$ be a pure jump process and $((X_t, U_t))_{t\ge0}$ the appertaining bivariate process being strong Markovian with regular transition probability matrix p. Then, the appertaining marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$ is a homogenous Markov chain with one-step transition probabilities*

$$
P(T_{m+1} \le t, Z_{m+1} = z | T_m = s,\ Z_m = y) = Q_{yz}(s, t, 0) \tag{2.31.1}
$$

$(0 \le s \le t < \infty,\ (y, z) \in \mathcal{J},\ m \in \mathbb{N}_0)$ for $\mathcal{L}(T_m, Z_m | P)$-a.e. (s, y), where the right-hand side is given by (2.24.1).

Proof. The proof is similar to the proof of Hilfssatz 4.33 in Milbrodt and Helbig [1999] stating that the Markov property of a pure jump process implies the Markov property of the appertaining marked point process.
Employing the strong Markov property of $((X_t, U_t), \mathfrak{F}_t)_{t\ge0}$, we verify that the one-step transition probabilities of the appertaining marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$ satisfy

$$
P((T_{m+1}, Z_{m+1}) \in C | (T_i, Z_i)_{i\in\mathbb{N}_0, i\le m}) = P((T_{m+1}, Z_{m+1}) \in C | T_m, Z_m)
$$

$(C \in \mathfrak{B}([0, \infty]) \otimes 2^{\mathcal{S}}, m \in \mathbb{N}_0)$, where the right-hand side is also independent of m. As in the proof of lemma 2.21, it is sufficient to show this for sets $\{\eta\} \times [0, w] \in \mathfrak{D}_0$.
For $s \ge 0$, let $P_{(y,u)s}$ denote the regular version of the conditional distribution of $((X_t, U_t))_{t\ge s}$ given $(X_s, U_s) = (y, u)$, which appertains to the regular transition probability matrix p. Then, according to (2.7.5) as well as (2.7.6), the strong Markov property of $((X_t, U_t), \mathfrak{F}_t)_{t\ge0}$ yields almost surely for any $(\mathfrak{F}_t)_{t\ge0}$-stopping time T

$$
\mathbf{E}\left[f(T, (X_\tau, U_\tau)_{\tau\ge T}) | \mathfrak{F}_T\right] = \int_{\mathcal{U}_{\mathcal{X}}} f(s, \cdot)\, dP_{(y,u)s}|_{s=T,(y,u)=(X_T,U_T)} \tag{2.31.2}
$$

for any measurable mapping $f : ([0, \infty)\times\mathcal{U}_{\mathcal{X}}, \mathfrak{B}([0, \infty))\otimes\mathfrak{U}_{\mathcal{X}}) \to ([0, \infty), \mathfrak{B}([0, \infty)))$. By applying (2.31.2) to

$$
f : [0, \infty) \times \mathcal{U}_{\mathcal{X}} \ni (s, (x, u(x))) \mapsto \mathbf{1}_{\{s+t_1(x)=[0,w], pr(x)_{t_1(x)}=\eta\}} \in \{0, 1\},
$$

we get P-a.s. on $\{T_m < \infty\}$

$$
\begin{aligned}
&P(T_{m+1} \in [0, w], Z_{m+1} = \eta | (T_i, Z_i)_{i\in\mathbb{N}_0, i\le m}) \\
&= P(T_{m+1} \in [0, w], Z_{m+1} = \eta | \mathfrak{F}_{T_m}) \\
&= P(T_m + t_1((X_{\tau+T_m})_{\tau\ge0}) \in [0, w], pr((X_{\tau+T_m})_{\tau\ge0})_{t_1} = \eta | \mathfrak{F}_{T_m}) \\
&= \mathbf{E}\left[f(T_m, (X_\tau, U_\tau)_{\tau\ge T_m}) | \mathfrak{F}_{T_m}\right] \\
&= P_{(y,u)s}(s + t_1((X_{\tau+s})_{\tau\ge0}) \in [0, w], pr((X_{\tau+s})_{\tau\ge0})_{t_1} = \eta)|_{s=T_m,(y,u)=(X_{T_m},U_{T_m})}.
\end{aligned}
$$

According to the definition of $P_{(y,u)s}$ and (2. 24. 1), we have for $\mathcal{L}(X_s, U_s| P)$-a.e. (y,u)

$$
\begin{aligned}
&P_{(y,u)s}(s + t_1((X_{\tau+s})_{\tau\geq 0}) \in [0,w], pr((X_{\tau+s})_{\tau\geq 0})_{t_1} = \eta) \\
&= \mathbf{1}_{\{y\neq\eta\}}\, P(T(s) \leq w, X_{T(s)} = \eta | X_s = y, U_s = u) \\
&= \mathbf{1}_{\{y\neq\eta\}}\, Q_{y\eta}(s,w,u).
\end{aligned}
$$

Hence, on $\{T_m < \infty\}$

$$
\begin{aligned}
P(T_{m+1} \in [0,w], Z_{m+1} = \eta |\, (T_i, Z_i)_{i\in\mathbb{N}_0, i\leq m}) &= \mathbf{1}_{\{y\neq\eta\}}\, Q_{yz}(s,w,u)|_{s=T_m,(y,u)=(X_{T_m},U_{T_m})} \\
&= \mathbf{1}_{\{Z_m\neq\eta\}}\, Q_{Z_m\eta}(T_m, w, 0) \;\; P-a.s.,
\end{aligned}
$$

where the last equation takes into account that $X_{T_m} = Z_m$ and $U_{T_m} = T_m - T_m = 0, m \in \mathbb{N}$. Thus, the one-step transition probabilities satisfy the requirements for the marked point process to be homogeneous and Markovian and it holds (2. 31. 1). □

The assertion of this lemma seems to coincide with formula (2. 25. 3) in lemma 2. 25. However, lemma 2. 25 starts from a given homogeneous marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$ and shows that the one-step transition probabilities of this process can be expressed by means of the conditional probabilities Q_* for the appertaining bivariate Markov process. Lemma 2. 31, however, states that by starting from a bivariate Markov process $((X_t, U_t), \mathfrak{F}_t)_{t\geq 0}$ the strong Markov property of this process implies the appertaining marked point process to be homogeneous and Markovian, with one-step transition probabilities determined by Q_*.

D.1 Backward and forward equations

The aim of this section is to establish so-called Kolmogorov backward and forward equations for the bivariate Markov process $((X_t, U_t), \mathfrak{F}_t)_{t\geq 0}$ that appertains to a semi-Markov process $(X_t, \mathfrak{F}_t)_{t\geq 0}$. In a non-smooth framework, these equations must be given as integral equations. Further, according to the state space $\mathcal{S}$, these equations correspond to systems of integral equations. These systems allow one to obtain transition probabilities from given cumulative transition intensities.

Analogously to the derivation of two different types of backward and forward integral equations for a non-smooth Markov process $(X_t)_{t\geq 0}$ (cf. Milbrodt and Helbig [1999], section 4C), two types of backward integral equations and two types of forward integral equations are derived here. Likewise similar to the Markov set-up, it turns out that the backward equations are special cases of the systems of integral equations for the prospective reserve. For this reason, we abstain here from either investigating the solvability of the backward integral equations or giving numerical examples. Both will be done in the context of Thiele's integral equations for the prospective reserve. By inserting certain parameters into these equations, the corresponding results can be obtained for the backward integral equations. Regarding the forward equations, a concept of the retrospective reserve in a semi-Markov framework is sketched in chapter 5, such that the appertaining systems of integral equations are related to the forward integral equations for the transition probabilities of the bivariate Markov process (cf. Norberg [1991]).

If the cumulative transition intensities are absolutely continuous with respect to the Lebesgue measure, the forward and backward equations for the Markov process $((X_t, U_t), \mathfrak{F}_t)_{t\geq 0}$ can be established as systems of differential equations. Rewritten as integral equations, they can already be found in Hoem [1972]. Backward differential equations and so-called integro-differential equations were established by Möller and Zwiesler [1996]. Backward differential equations can also be obtained by the work of Møller [1993], namely as special cases of the corresponding differential equations for the prospective reserve in a smooth semi-Markov framework.

As in the preceding sections, let $(\Omega, \mathfrak{F}, P, ((T_m, Z_m))_{m\in\mathbb{N}_0})$ be a homogeneous Markovian marked point process with paths in $\mathcal{K}$ and $(X_t, \mathfrak{F}_t)_{t\geq 0}$ the appertaining semi-Markov process with paths in $\mathcal{X}$. Further, $((X_t, U_t), \mathfrak{F}_t)_{t\geq 0}$ is the associated bivariate Markov process with transition probability matrix

$$p(s,t,u,v) = \big(p_{yz}(s,t,u,v)\big)_{(y,z)\,\in\,\mathcal{S}^2}, \quad 0 \leq u \leq s \leq t < \infty, v \geq 0,$$

and cumulative transition intensity matrix

$$q(s,t,u) = \big(q_{yz}(s,t,u)\big)_{(y,z)\,\in\,\mathcal{S}^2}, \quad 0 \leq u \leq s \leq t < \infty.$$

The following lemma establishes a system of backward integral equations for the transition probabilities of the process $((X_t, U_t), \mathfrak{F}_t)_{t\geq 0}$.

2.32 Lemma. [Backward integral equations] *Let $(X_t, \mathfrak{F}_t)_{t\geq 0}$ be a semi-Markov process and $((X_t, U_t))_{t\geq 0}$ the bivariate Markov process with conditional distributions Q_* according to definition 2.24 and transition probability matrix p. Further, let $0 \leq u \leq s \leq t < \infty, v \geq 0$, and $(y,z) \in \mathcal{S}^2$. Then, the components of p satisfy for $\mathcal{L}(X_s, U_s|\, P)$-a.e. (y,u) the backward integral equations*

$$p_{yz}(s,t,u,v) \;=\; \delta_{yz}\, \mathbf{1}_{(v\geq t-s+u)}\,(1 - Q_y(s,t,u)) + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \int_{(s,t]} p_{\xi z}(\tau,t,0,v)\, Q_{y\xi}(s,d\tau,u). \quad (2.32.1)$$

If regular cumulative transition intensities q_* exist, one obtains

$$\begin{aligned} p_{yz}(s,t,u,v) \;=\;& \delta_{yz}\, \mathbf{1}_{(v\geq t-s+u)}\,(1 - Q_y(s,t,u)) \\ &+ \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \int_{(s,t]} p_{\xi z}(\tau,t,0,v)\, \big(1 - Q_y(s,\tau-0,u)\big)\, q_{y\xi}(s,d\tau,u). \end{aligned} \quad (2.32.2)$$

Proof. With the aid of (2.28.4), (2.32.2) follows immediately from (2.32.1). Hence, it is sufficient to derive (2.32.1). Let $0 \leq u \leq s \leq t < \infty, v \geq 0$, and $(y,z) \in \mathcal{S}^2$. According to (2.23.4) and (2.23.5), we have for $\mathcal{L}(X_s, U_s|\, P)$-a.e. (y,u)

$$p_{yz}(s,t,u,v) \;=\; \sum_{l=0}^{\infty} p_{yz}^{(l)}(s,t,u,v) = p_{yz}^{(0)}(s,t,u,v) + \sum_{l=1}^{\infty} p_{yz}^{(l)}(s,t,u,v), \quad (2.32.3)$$

with

$$p_{yz}^{(l)}(s,t,u,v) \;=\; P(T_{l+1} > t, T_l \in [t-v,t], Z_l = z|\, T_1 > s, T_0 = s-u, Z_0 = y). \quad (2.32.4)$$

For the first addend of the right-hand side of (2.32.3), we get according to (2.25.9) and (2.26.1) for $\mathcal{L}(X_s, U_s|\, P)$-a.e. (y,u)

$$\begin{aligned} p_{yz}^{(0)}(s,t,u,v) \;&=\; P(T_1 > t, T_0 \in [t-v,t], Z_0 = z|\, T_1 > s, T_0 = s-u, Z_0 = y) \\ &=\; \delta_{yz}\, \mathbf{1}_{(v\geq t-s+u)}\, P(T_1 > t|\, T_1 > s, T_0 = s-u, Z_0 = y) \\ &=\; \delta_{yz}\, \mathbf{1}_{(v\geq t-s+u)}\, P(T(s) > t|\, X_s = y,\, U_s = u) \\ &=\; \delta_{yz}\, \mathbf{1}_{(v\geq t-s+u)}\,(1 - Q_y(s,t,u)). \end{aligned} \quad (2.32.5)$$

In the case of $l > 0$, we obtain - according to remark 2.16 - by conditioning on (T_1, Z_1) and applying the homogeneity of (T, Z), for $\mathcal{L}(X_s, U_s|\, P)$-a.e. (y,u)

$$\begin{aligned} p_{yz}^{(l)}(s,t,u,v) \;&=\; P(T_{l+1} > t, T_l \in [t-v,t], Z_l = z|\, T_1 > s, T_0 = s-u, Z_0 = y) \\ &=\; \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \int_{(s,t]} P(T_{l+1} > t, T_l \in [t-v,t], Z_l = z|\, T_1 = \tau, Z_1 = \xi) \\ &\qquad\qquad P(T_1 \in d\tau, Z_1 = \xi|\, T_1 > s, T_0 = s-u, Z_0 = y) \\ &=\; \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \int_{(s,t]} p_{\xi z}^{(l-1)}(\tau,t,0,v)\, Q_{y\xi}(s,d\tau,u), \end{aligned} \quad (2.32.6)$$

where the last equation is due to (2.32.4), (2.26.2), and (2.26.1). According to (2.32.3), we get by adding (2.32.5) and (2.32.6)

$$
\begin{aligned}
p_{yz}(s,t,u,v) &= \delta_{yz}\,\mathbf{1}_{(v\geq t-s+u)}\,(1-Q_y(s,t,u)) + \sum_{l=1}^{\infty}\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,t]} p_{\xi z}^{(l-1)}(\tau,t,0,v)\,Q_{y\xi}(s,d\tau,u)\\
&= \delta_{yz}\,\mathbf{1}_{(v\geq t-s+u)}\,(1-Q_y(s,t,u)) + \sum_{l=0}^{\infty}\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,t]} p_{\xi z}^{(l)}(\tau,t,0,v)\,Q_{y\xi}(s,d\tau,u).
\end{aligned}
$$

Finally, by interchanging the infinite sum and the integral using the monotone convergence theorem, and afterwards applying (2.32.3) again, we obtain (2.32.1). □

By specifying the state space as $\mathcal{S}\times[0,\infty)$, the system of integral equations (2.32.1) corresponds to equation (4) in Cinlar [1969]. The proof of this system of integral equations is based on the fact that the marked point process $((T_m,Z_m))_{m\in\mathbb{N}_0}$ that appertains to a semi-Markov process $(X_t)_{t\geq 0}$ provides by means of (2.23.1) and (2.23.2) a set of transition probabilities for the bivariate Markov process $((X_t,U_t))_{t\geq 0}$. Consequently, this set of transition probabilities satisfies (2.32.1) without exceptional sets, provided that Q_* is selected according to (2.25.1). Further, this set of transition probabilities forms a regular set of transition probabilities. Taking these facts into account, it can be shown that the bivariate process $((X_t,U_t))_{t\geq 0}$ appertaining to a homogeneous Markovian marked point process is even strong Markovian. For this, it is according to lemma 2.30 sufficient to verify (2.30.1).

2.33 Corollary. *Let $(X_t,\mathfrak{F}_t)_{t\geq 0}$ be a semi-Markov process and $((T_m,Z_m))_{m\in\mathbb{N}_0}$ the appertaining homogeneous Markovian marked point process with transition probabilities $\hat{Q}_*$ according to definition* 2.8*. Then, the appertaining bivariate process $((X_t,U_t))_{t\geq 0}$ is a strong Markov process.*

Proof. Let $0\leq u\leq s\leq t<\infty, v\geq 0, (y,z)\in\mathcal{S}^2$. According to lemma 2.21, $((X_t,U_t))_{t\geq 0}$ is a Markov process. A regular set of transition probabilities of this process is given by (2.23.1) and (2.23.2). The conditional distributions Q_* according to definition 2.24 are selected by means of (2.25.1), i.e. $Q_{yz}(s,t,u)=\hat{Q}_{yz}(s-u,(s,t])$. Then, the system of integral equations (2.32.1) is satisfied without exceptional sets.
Employing lemma 2.30, the strong Markov property of the process $((X_t,U_t))_{t\geq 0}$ follows from condition (2.30.1). We will verify that this condition is satisfied in the present situation. In doing so, we restrict ourselves to the case $y\neq z$, such that (2.32.1) can be reduced to

$$
p_{yz}(s,t,u,v) \;=\; \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,\hat{Q}_{y\xi}(s-u,d\tau). \tag{2.33.1}
$$

The case $y=z$ can be investigated analogously. By employing (2.33.1), one obtains for $h\geq 0$

$$
\begin{aligned}
p_{yz}(s,s+h,u,v) &= \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,s+h]} p_{\xi z}(\tau,s+h,0,v)\,\hat{Q}_{y\xi}(s-u,d\tau)\\
&\leq \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\hat{Q}_{y\xi}(s-u,(s,s+h]). \qquad (2.33.2)
\end{aligned}
$$

Thus, it follows $\lim_{h\searrow 0} p_{yz}(s,s+h,u,v)=0=p_{yz}(s,s,u,v)$, the last due to (2.22.2) and $y\neq z$. For $y=z$, one gets $\lim_{h\searrow 0} p_{yy}(s,s+h,u,v)=p_{yy}(s,s,u,v)=\mathbf{1}_{(v\geq u)}$. A regular set of transition probabilities $p_{yz}(s,t,u,v)$ also satisfies the Chapman-Kolmogorov equations

(2.22.1) without exceptional sets. Using these equations, we further conclude by employing the dominated convergence theorem

$$\begin{aligned}\lim_{h\searrow 0} p_{yz}(s,t+h,u,v) &= \lim_{h\searrow 0}\sum_{\xi\in\mathcal{S}}\int_{[0,\infty)} p_{\xi z}(t,t+h,l,v)\,p_{y\xi}(s,t,u,dl)\\ &= \sum_{\xi\in\mathcal{S}}\int_{[0,\infty)} \lim_{h\searrow 0} p_{\xi z}(t,t+h,l,v)\,p_{y\xi}(s,t,u,dl)\\ &= p_{yz}(s,t,u,v). \qquad (2.33.3)\end{aligned}$$

Hence, $p_{yz}(s,t,u,v)$ is right continuous as function of t. By again using (2.33.1), we obtain (2.30.1) as follows:

$$\begin{aligned}&|p_{yz}(s,t,u,v) - p_{yz}(s+h,t+k,u+h,v)|\\ &= \Big|\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,\hat{Q}_{y\xi}(s-u,d\tau) - \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s+h,t+k]} p_{\xi z}(\tau,t+k,0,v)\,\hat{Q}_{y\xi}(s+h-u-h,d\tau)\Big|\\ &= \Big|\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,\hat{Q}_{y\xi}(s-u,d\tau) - \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s+h,t+k]} p_{\xi z}(\tau,t+k,0,v)\,\hat{Q}_{y\xi}(s-u,d\tau)\Big|\\ &= \Big|\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,s+h]} p_{\xi z}(\tau,t,0,v)\,\hat{Q}_{y\xi}(s-u,d\tau) + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s+h,t]} p_{\xi z}(\tau,t,0,v)\,\hat{Q}_{y\xi}(s-u,d\tau)\\ &\quad - \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s+h,t]} p_{\xi z}(\tau,t+k,0,v)\,\hat{Q}_{y\xi}(s-u,d\tau) - \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(t,t+k]} p_{\xi z}(\tau,t+k,0,v)\,\hat{Q}_{y\xi}(s-u,d\tau)\Big|\\ &\leq \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,s+h]} p_{\xi z}(\tau,t,0,v)\,\hat{Q}_{y\xi}(s-u,d\tau) + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(t,t+k]} p_{\xi z}(\tau,t+k,0,v)\,\hat{Q}_{y\xi}(s-u,d\tau)\\ &\quad + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,t]} \Big|p_{\xi z}(\tau,t,0,v) - p_{\xi z}(\tau,t+k,0,v)\Big|\,\hat{Q}_{y\xi}(s-u,d\tau), \qquad (2.33.4)\end{aligned}$$

where, by arguing analogously as in (2.33.2) and by employing the right continuity of the transition probabilities according to (2.33.3), each of the addends converge against zero for either $h\searrow 0$ or $k\searrow 0$. □

2.34 Remark. Similarly as in the proof of the above corollary, the following can be worked out: Let $(X_t,\mathfrak{F}_t)_{t\geq 0}$ be a pure jump process and $((X_t,U_t))_{t\geq 0}$ the appertaining bivariate process being Markovian with regular cumulative transition intensity matrix q. Further, let p be a regular set of transition probabilities of $((X_t,U_t))_{t\geq 0}$ satisfying (2.32.2) without exceptional sets. Then the process $((X_t,U_t))_{t\geq 0}$ is strong Markovian.

The system of integral equations (2.32.1) generalizes the backward integral equations for the transition probabilities of a non-smooth Markovian pure jump process (cf. Milbrodt and Helbig, Hilfssatz 4.46). In almost the same manner, a generalization of the forward integral equations in a Markov set-up (cf. Milbrodt and Helbig, Hilfssatz 4.48) can be established. For this, however, it must be additionally assumed that the cumulative transition intensities are dominated, such that the dominating measure does not depend on the time elapsed since the current state was entered. In proving the forward integral equations, this assumption is essential for changing the order of integrations according to Fubini's theorem.

2.35 Assumption. The regular cumulative transition intensities q_{yz}, $(y,z)\in\mathcal{J}$ for the bivariate Markov process $((X_t,U_t))_{t\geq 0}$ are assumed to be dominated by Borel measures $\boldsymbol{\Lambda}_{yz}$, $(y,z)\in$

$\mathcal{J}$, such that for $0 \le u \le s$ and $(y,z) \in \mathcal{J}$

$$q_{yz}(s, d\tau, u) \ll \mathbf{\Lambda}_{yz}(d\tau) \quad \text{with} \quad q_{yz}(s, d\tau, u) = \lambda_{yz}(\tau, \tau - s + u)\, \mathbf{\Lambda}_{yz}(d\tau). \qquad (2.35.1)$$

Thus, the dominating measure $\mathbf{\Lambda}_{yz}(d\tau)$ for the cumulative transition intensity from y to z does not depend on the time elapsed since state y was entered.

Recall that when considering $q_{yz}(s, d\tau, u)$, $0 \le u \le s \le \tau$, the state y was entered at time $s-u$. Hence, at time τ, the time elapsed since this state was entered is given by $\tau - s + u$. Under assumption 2.35, the dependence on the time elapsed since entering the current state is only taken into account by the corresponding densities.

The following lemma establishes the forward integral equations for the transition probabilities of the bivariate process $((X_t, U_t))_{t\ge 0}$.

2.36 Lemma. [Forward integral equations] *Let $(X_t, \mathfrak{F}_t)_{t\ge0}$ be a semi-Markov process and $((X_t, U_t))_{t\ge0}$ the bivariate Markov process with conditional distributions Q_* according to definition 2.24 and transition probability matrix p. q_* are regular cumulative transition intensities for which 2.35 is assumed. Further, let $0 \le u \le s \le t < \infty, v \ge 0$, and $(y,z) \in \mathcal{S}^2$. Then, the components of p satisfy for $\mathcal{L}(X_s, U_s|\, P)$-a.e. (y,u) the forward integral equations*

$$\begin{aligned} p_{yz}(s,t,u,v) &= \delta_{yz}\, \mathbf{1}_{(v\ge t-s+u)}\,(1 - Q_y(s,t,u)) \\ &\quad + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{[t-v,t]} \int_{(0,\infty)} (1 - Q_z(\tau,t,0))\, \lambda_{\xi z}(\tau, l)\, p_{y\xi}(s, \tau-0, u, dl)\, \mathbf{\Lambda}_{\xi z}(d\tau). \end{aligned} \qquad (2.36.1)$$

Proof. Analogously to the proof of lemma 2.32, we obtain for $\mathcal{L}(X_s, U_s|\, P)$-a.e. (y,u)

$$p_{yz}(s,t,u,v) = \delta_{yz}\, \mathbf{1}_{(v\ge t-s+u)}\,(1 - Q_y(s,t,u)) + \sum_{l=1}^{\infty} p_{yz}^{(l)}(s,t,u,v), \qquad (2.36.2)$$

with

$$p_{yz}^{(l)}(s,t,u,v) = P(T_{l+1} > t, T_l \in [t-v,t], Z_l = z|\, T_1 > s, T_0 = s-u, Z_0 = y). \quad (2.36.3)$$

For $l > 0$, we likewise proceed according to remark 2.16. By conditioning on (T_{l-1}, Z_{l-1}), using the properties of (T, Z), inserting (2.8.1), (2.8.2) and afterwards employing (2.25.3), we get for $\mathcal{L}(X_s, U_s|\, P)$-a.e. (y,u)

$$\begin{aligned} & p_{yz}^{(l)}(s,t,u,v) \\ &= \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{(s,t]} \int_{[t-v,t]} P\big(T_{l+1} > t|\, T_l = x, Z_l = z\big)\, P(T_l \in dx, Z_l = z|\, T_{l-1} = \tau, Z_{l-1} = \xi) \\ &\qquad P(T_{l-1} \in d\tau, Z_{l-1} = \xi|\, T_1 > s, T_0 = s-u, Z_0 = y) \\ &= \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{(s,t]} \int_{[t-v,t]} (1 - \hat{Q}_z(x,t))\, \hat{Q}_{\xi z}(\tau, dx) \\ &\qquad P(T_{l-1} \in d\tau, Z_{l-1} = \xi|\, T_1 > s, T_0 = s-u, Z_0 = y) \\ &= \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{(s,t]} \int_{[t-v,t]} (1 - Q_z(x,t,0))\, Q_{\xi z}(\tau, dx, 0) \\ &\qquad P(T_{l-1} \in d\tau, Z_{l-1} = \xi|\, T_1 > s, T_0 = s-u, Z_0 = y) \\ &= \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{(s,t]} \int_{[t-v,t]} \big(1 - Q_z(x,t,0)\big)\, \big(1 - Q_\xi(\tau, x-0, 0)\big)\, q_{\xi z}(\tau, dx, 0) \\ &\qquad P(T_{l-1} \in d\tau, Z_{l-1} = \xi|\, T_1 > s, T_0 = s-u, Z_0 = y), \end{aligned} \qquad (2.36.4)$$

where the last equation is by an application of (2. 28. 4). According to assumption 2. 35, we have

$$q_{\xi z}(\tau, dx, 0) = \lambda_{\xi z}(x, x-\tau)\, \mathbf{\Lambda}_{\xi z}(dx).$$

Inserting this into the last line of the above chain of equations allows us to change the order of integrations according to Fubini's theorem. Doing so, we obtain for $\mathcal{L}(X_s, U_s|\, P)$-a.e. (y, u)

$$\begin{aligned} p^{(l)}_{yz}(s,t,u,v) &= \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{(s,t]} \int_{[t-v,t]} \big(1-Q_z(x,t,0)\big)\big(1-Q_\xi(\tau,x-0,0)\big)\,\lambda_{\xi z}(x,x-\tau)\,\mathbf{\Lambda}_{\xi z}(dx) \\ &\qquad P(T_{l-1}\in d\tau, Z_{l-1}=\xi|\, T_1>s, T_0=s-u, Z_0=y) \\ &= \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{[t-v,t]} \big(1-Q_z(x,t,0)\big) \int_{(s,t]} \big(1-Q_\xi(\tau,x-0,0)\big)\,\lambda_{\xi z}(x,x-\tau) \\ &\qquad P(T_{l-1}\in d\tau, Z_{l-1}=\xi|\, T_1>s, T_0=s-u, Z_0=y)\,\mathbf{\Lambda}_{\xi z}(dx). \end{aligned} \tag{2. 36. 5}$$

Consider the interior integral. For $\tau \geq 0$, the measure $q_{\xi z}(\tau, dx, 0) = \lambda_{\xi z}(x, x-\tau)\,\mathbf{\Lambda}_{\xi z}(dx)$ is concentrated on (τ, ∞). Hence, the appertaining density satisfies for $\mathbf{\Lambda}_{\xi z}$-a.e. $x \in [0, \infty)$

$$\lambda_{\xi z}(x, x-\tau) = 0, \;\; x \leq \tau.$$

Therefore, the integration interval $(s, t]$ can be restricted to (s, x). By employing (2. 25. 3) as well as (2. 8. 2) again, and afterwards inserting (2. 36. 3), we get

$$\begin{aligned} &\int_{(s,t]} \big(1-Q_\xi(\tau,x-0,0)\big)\,\lambda_{\xi z}(x,x-\tau)\,P(T_{l-1}\in d\tau, Z_{l-1}=\xi|\, T_1>s, T_0=s-u, Z_0=y) \\ &= \int_{(s,x)} \lambda_{\xi z}(x,x-\tau)\,P(T_l\geq x|\, T_{l-1}=\tau, Z_{l-1}=\xi) \\ &\qquad P(T_{l-1}\in d\tau, Z_{l-1}=\xi|\, T_1>s, T_0=s-u, Z_0=y) \\ &= \int_{(s,x)} \lambda_{\xi z}(x,x-\tau)\,P(T_l\geq x, T_{l-1}\in d\tau, Z_{l-1}=\xi|\, T_1>s, T_0=s-u, Z_0=y) \\ &= \int_{(s,x)} \lambda_{\xi z}(x,x-\tau)\,p^{(l-1)}_{y\xi}(s,x-0,u,x-d\tau). \end{aligned}$$

Now, apply the theorem on integration with respect to an image measure to the measurable function $I : \tau \mapsto x-\tau$ with $\mathbf{1}_{(s,x)}(\tau) = \mathbf{1}_{(0,x-s)}(I(\tau))$. This results in

$$\begin{aligned} &\int_{(s,t]} \big(1-Q_\xi(\tau,x-0,0)\big)\,\lambda_{\xi z}(x,x-\tau)\,P(T_{l-1}\in d\tau, Z_{l-1}=\xi|\, T_1>s, T_0=s-u, Z_0=y) \\ &= \int_{(0,x-s)} \lambda(x,\tau)\,p^{(l-1)}_{y\xi}(s,x-0,u,d\tau) \\ &= \int_{(0,\infty)} \lambda(x,\tau)\,p^{(l-1)}_{y\xi}(s,x-0,u,d\tau). \end{aligned}$$

Inserting the above into (2. 36. 5) leads for $\mathcal{L}(X_s, U_s|\, P)$-a.e. (y, u) to

$$p^{(l)}_{yz}(s,t,u,v) = \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{[t-v,t]} \int_{(0,\infty)} \big(1-Q_z(x,t,0)\big)\,\lambda_{\xi z}(x,\tau)\,p^{(l-1)}_{y\xi}(s,x-0,u,d\tau)\,\mathbf{\Lambda}_{\xi z}(dx).$$

This yields, along with (2. 36. 2) and the same argumentation as at the end of the proof of lemma 2. 32, the forward integral equations (2. 36. 1). □

Assumption 2. 35 is not only required to establish forward integral equations for the transition probabilities of the process $((X_t, U_t))_{t\geq 0}$, but also for some results concerning the prospective reserve in a semi-Markov set-up. Further, it is essential for the investigation of the retrospective reserve. A situation where assumption 2. 35 is satisfied is given by the framework of the investigations of Hoem [1972], Møller [1993], and Möller & Zwiesler [1996]. They assumed the existence of transition intensities

$$q_{yz}(s, d\tau, u) = \mu_{yz}(\tau, \tau - s + u)\,\lambda^1(d\tau), \quad (y, z) \in \mathcal{J}, 0 \leq u \leq s, \tag{2.36.6}$$

and

$$\mu_{yy}(\tau, \tau - s + u) := -\sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \mu_{yz}(\tau, \tau - s + u). \tag{2.36.7}$$

All results concerning the transition probabilities of the process $((X_t, U_t))_{t\geq 0}$, which were established by the above authors, can be obtained by either simplification or differentiation of (2. 32. 2) and (2. 36. 1), respectively. Note, however, that Möller and Zwiesler used the more actuarial *Hamza notation.* For this reason, the corresponding results must be converted according to the replacing rules (1. 0. 1) - (1. 0. 4).

For the sake of completeness, we now state differential equations for the transition probabilities of the bivariate Markov process. The backward differential equations correspond to (2. 32. 2) and the forward differential equations to (2. 36. 1). The proofs can be found in A.7 in the appendix.

2. 37 Corollary. [Backward and forward differential equations] *Let $(X_t, \mathfrak{F}_t)_{t\geq 0}$ be a semi-Markov process and $((X_t, U_t))_{t\geq 0}$ the appertaining bivariate Markov process with transition probability matrix p and regular cumulative transition intensity matrix q. The components of the latter are assumed to satisfy assumption* 2. 35 *by means of* (2. 36. 6) *and* (2. 36. 7) *with the corresponding densities meet the following requirements: for* $0 \leq u \leq s \leq t < \infty$, $(y, z) \in \mathcal{S}^2$

- $u \mapsto \mu_{yz}(s, u)$ *is differentiable*
- $(s, u) \mapsto \mu_{yz}(s, u)$ *as well as* $(s, u) \mapsto \frac{\partial}{\partial u}\mu_{yz}(s, u)$ *are continuous.*

Then, the transition probabilities $p_{yz}(s, t, u, v), v \geq 0$, are partial differentiable with respect to s, t, u, v and they satisfy for $\mathcal{L}(X_s, U_s|\, P)$-a.e. (y, u) the backward differential equations

$$\frac{\partial p_{yz}(s,t,u,v)}{\partial s} = -\sum_{\xi\in\mathcal{S}} \mu_{y\xi}(s,u)\, p_{\xi z}(s,t,0,v) - \frac{\partial p_{yz}(s,t,u,v)}{\partial u}, \tag{2.37.1}$$

as well as the forward differential equations

$$\begin{aligned}\frac{\partial p_{yz}(s,t,u,v)}{\partial t} &= \int_0^v \mu_{yy}(t,l)\, p_{yz}(s,t,u,dl) \\ &\quad + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_0^\infty \mu_{\xi z}(t,l)\, p_{y\xi}(s,t,u,dl) - \frac{\partial p_{yz}(s,t,u,v)}{\partial v}.\end{aligned} \tag{2.37.2}$$

Obviously, assumption 2. 35 is also generally satisfied in the framework of a non-smooth Markov model. Moreover, it remains satisfied when manufacturing a semi-Markov model by starting from a Markov process with cumulative transition intensities $\bar{q}_{yz}, (y, z) \in \mathcal{S}^2$, and multiplying densities describing the duration-dependence of the cumulative transition intensities. More precisely, consider a Markov process $(Y_t)_{t\geq 0}$ with regular cumulative transition intensity matrix $\bar{q}(s,t) = (\bar{q}_{yz}(s,t))_{(y,z)\in\mathcal{S}^2}$ and measurable functions $\lambda_{yz} : [0,\infty)\times[0,\infty) \to [0,\infty)$. Then, duration-depending cumulative transition intensities q can be specified by means of

$$q_{yz}(s,t,u) \;:=\; \int_{(s,t]} \lambda_{yz}(\tau, \tau - s - u)\, \bar{q}_{yz}(d\tau), \quad (y,z) \in \mathcal{S}^2. \tag{2.37.3}$$

These cumulative transition intensities q_* are regular in terms of definition 2. 28. Hence, according to theorem 2. 29, there exists a semi-Markov process $(X_t)_{t\geq 0}$ such that the appertaining bivariate Markov process $((X_t, U_t))_{t\geq 0}$ possesses the cumulative transition intensities q_*, which satisfy by definition assumption 2. 35.

Beside the above models, there are more situations of practical interest with assumption 2. 35 being automatically satisfied. For example, when transitions are only annually recorded, $(X_t)_{t\geq 0}$ is a pure jump process for which jumps are only allowed at integer times. Let $((X_t, U_t))_{t\geq 0}$ be the appertaining bivariate Markov process with regular cumulative transition intensity matrix q. Then, one obtains for $t \geq 0$, $(X_t, U_t) = (X_{\lfloor t\rfloor}, U_{\lfloor t\rfloor})$ P-a.s. In that situation, a dominating measure is given by $\mathbf{\Lambda}_{yz}(d\tau) = \sum_{l=0}^{\infty} \varepsilon_l(d\tau)$: Let $(y,z) \in \mathcal{J}$, $n, u \in \mathbb{N}$ and $P(X_n = y, U_n = u) > 0$. Then for $t \in (n, n+1]$

$$\begin{aligned} Q_{yz}(n,t,u) &= P(T(n) \leq t, X_{T(n)} = z \,|\, X_n = y, U_n = u) \\ &= \begin{cases} 0, & t < n+1 \\ p_{yz}(n, n+1, u), & t = n+1 \end{cases}. \end{aligned} \tag{2.37.4}$$

According to (2. 24. 4), one obtains

$$q_{yz}(n,t,u) \;=\; \int_{(n,t]} \frac{Q_{yz}(n, d\tau, u)}{1 - Q_y(n, \tau - 0, u)} = p_{yz}(n, n+1, u)\, \varepsilon_{n+1}((n,t]), \;\; t \in (n, n+1]. \tag{2.37.5}$$

Hence, for $t \geq n$, it follows with (2. 27. 1)

$$\begin{aligned} q_{yz}(n,t,u) &= \int_{(n,t]} q_{yz}(n, d\tau, u) \\ &= \sum_{l=0}^{\infty} \int_{(n+l\wedge t, n+l+1\wedge t]} q_{yz}(n, d\tau, u) \\ &= \sum_{l=0}^{\infty} \int_{(n+l\wedge t, n+l+1\wedge t]} q_{yz}(n+l, d\tau, u+l) \\ &= \sum_{l=0}^{\infty} p_{yz}(n+l, n+l+1, u+l)\, \varepsilon_{n+l+1}((n+l\wedge t, n+l+1\wedge t]) \\ &= \sum_{l=0}^{\infty} \int_{(n,t]} p_{yz}(\tau-1, \tau, \tau-n+u-1)\, \varepsilon_{n+l+1}(d\tau) \\ &= \int_{(n,t]} p_{yz}(\tau-1, \tau, \tau-n+u-1) \sum_{l=0}^{\infty} \varepsilon_l(d\tau). \end{aligned}$$

The density according to assumption 2. 35 is here given as

$$\lambda_{yz}(\tau, \tau - n + u) = p_{yz}(\tau-1, \tau, \tau-n+u-1).$$

A discrete model for which transitions are only allowed at integer times is, for example, used for describing the decrement of a portfolio for German private health insurance (cf. Milbrodt [2005], chapter 4). The following example introduces a corresponding model that additionally allows the transition probabilities to depend on the previous contract duration. For this, we restrict ourselves to the consideration of a PKV model for the so-called "old world". In contrast, the "new world" refers to PKV modelling that takes into account the new regulations due to the GKV-WSG (see [2007]). The differences between the models for both the old and the new world are discussed at the end of example 4. 13. Especially regarding the decrement cause *withdrawal*, some things will be different. This, however, does not devaluate the basic ideas behind the following approach.

2. 38 Example. Modelling the decrement of a PKV portfolio is currently based on a multiple state model with states a (*active*), w (*withdrawn*) and d (*dead*) (i.e. $\mathcal{S} = \{a, w, d\}$). Only two possible transitions are allowed: decrement due to withdrawal and decrement due to death (i.e. $\mathcal{J} = \{(a, w), (a, d)\}$). Figure 12 illustrates the state space and the possible transitions.

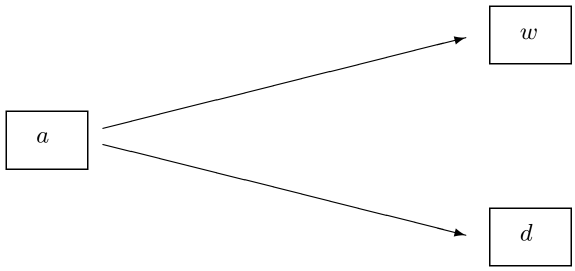

Figure 12: A multiple state model for the decrement of a PKV portfolio

Note that *active* is the only state that matters. The other states simply represent the possible causes of decrement. For the sake of convenience, such a simple structured multiple state model is usually investigated by means of a multiple decrement model. This means here that two random variables are considered: the future sojourn time of an insured $T \in [0, \infty)$, and the corresponding cause of decrement $J \in \{w, d\}$. Analogously to Milbrodt ([2005], section 4.1), both quantities are understood as random variables on a probability space $(\Omega, \mathfrak{F}, P)$. Regarding the assumptions 4.1(1) - 4.1(4) in Milbrodt ([2005], section 4.1), similar assumptions concerning the joint distribution of (T, J) are stated here, yet adapted to the additional duration-dependence. Due to the fact that only staying in state *active* is considered, the time elapsed since entering this state corresponds to the time since issue of the contract. This time will also be referred to as previous sojourn time in the portfolio or previous contract duration.
The assumptions concerning the joint distribution of (T, J) are

- stochastic independence of (T, J) for different insured,
- identical distributions of (T, J) for insured of the same age and sex if the previous sojourn times in the portfolio coincide,
- the joint distribution $\mathcal{L}((T, J))$ only depends on age and sex of an insured as well as on the previous sojourn time in the portfolio,
- condition of stationarity: Let $AB := \{x_{MIN}, ..., x_{MAX}\}$ be the age range. Further, let T_x be the future sojourn time and J_x the corresponding cause of decrement of an insured with attained age x and previous sojourn time U_x, $x \in AB$. For $j \in \{d, w\}, u \in \{0, ..., x - x_{MIN}\}, t \geq 0$, and $h \in \{0, ..., x_{MAX} - x\}$, we assume

$$P\Big(T_x > t + h, J_x = j \,|\, T_x > h, U_x = u\Big) = P\Big(T_{x+h} > t, J_{x+h} = j \,|\, U_{x+h} = u + h\Big). \quad (2.38.1)$$

This means that the observation of survival at age $x+h$ with previous sojourn time $u+h$ yields the same conditional distribution of future sojourn time and cause of decrement, as the hypothesis that an x-year insured with previous sojourn time u attains the age $x+h$.

According to these assumptions, the decrement of a PKV portfolio can be described by means of a pair of random variables (T_x, J_x), $x \in AB$ on $(\Omega, \mathfrak{F}, P)$ with values in $[0,\infty) \times \{w, d\}$. The conditional probabilities of decrement, given $U_x = u \in \{0, ..., x - x_{MIN}\}$, are for $t \geq 0$ determined by

$$ {}_tq_{x,u} := P(T_x \leq t, J_x = d |\, U_x = u) \quad \text{and} \quad {}_tw_{x,u} := P(T_x \leq t, J_x = w |\, U_x = u). $$

In comparison with definition 2. 24, the conditional probability ${}_tq_{x,u}$ corresponds for an insured with age at issue $x-u$ to $Q_{ad}(u, u+t, u)$. ${}_tw_{x,u}$ corresponds to $Q_{aw}(u, u+t, u)$. The total conditional probability of decrement, corresponding to $Q_a(u, u+t, u)$, is specified by

$$ F_{x,u} : [0,\infty) \ni t \mapsto {}_tq_{x,u} + {}_tw_{x,u}. $$

Hence, one obtains for the conditional probability of remaining in the portfolio

$$ {}_tp_{x,u} := 1 - F_{x,u}(t) = 1 - {}_tq_{x,u} - {}_tw_{x,u}, \;\; t \geq 0. $$

This probability corresponds to $\bar{p}_a(u, u+t, u)$.

In order to base the model on annual mortality and withdrawal rates, only integer times $k \in \{0, ..., x_{MAX} - x\}$ for insured of age $x \in AB$ are considered. By employing the exponential formula (2. 24. 7), the relation (2. 37. 5), and the condition (2. 38. 1), one obtains

$$ {}_kp_{x,u} = \prod_{j=0}^{k-1} p_{x+j,u+j} = \prod_{j=0}^{k-1} (1 - q_{x+j,u+j} - w_{x+j,u+j}), \tag{2. 38. 2} $$

where, as usual, $q_{x+j,u+j} := {}_1q_{x+j,u+j}$ and $w_{x+j,u+j} := {}_1w_{x+j,u+j}$. Further, let

$$ \mathcal{K}_x := [T_x - 0] = \sum_{k=0}^{\infty} k \cdot \mathbf{1}_{\{k < T_x \leq k+1\}} $$

be the curtate-future-lifetime of an insured of age x. It turns out that for $x \in AB$ and $u \in \{0, ..., x - x_{MIN}\}$

$$ P(\mathcal{K}_x > k |\, U_x = u) = P(T_x > k |\, U_x = u) = {}_kp_{x,u}, \;\; k \in \{0, ..., x_{MAX} - x\}, $$

and

$$ v_{x,u}(k) := P(\mathcal{K}_x = k |\, U_x = u) = P(k < T_x \leq k+1 |\, U_x = u) = {}_kp_{x,u} - {}_{k+1}p_{x,u}. $$

Hence, the conditional distribution of $\mathcal{K}_x$ is, by means of (2. 38. 2), fully specified by annual mortality and annual withdrawal rates. For the joint distribution of $\mathcal{K}_x$ and the corresponding cause of decrement, (2. 38. 1) yields

$$ P(\mathcal{K}_x = k, J_x = d |\, U_x = u) = {}_kp_{x,u}\; q_{x+k,u+k} \tag{2. 38. 3} $$

as well as

$$ P(\mathcal{K}_x = k, J_x = w |\, U_x = u) = {}_kp_{x,u}\; w_{x+k,u+k}. \tag{2. 38. 4} $$

In order to specify the distribution of T_x, additional assumptions about the distribution between integer times must be added. Usually, however, it is assumed that benefits as well as premiums are payable at integer times. To derive present values of future benefits and premiums, it is therefore sufficient to provide the probabilities of remaining in the portfolio ${}_kp_{x,u}$, $k \in \{0, ..., x_{MAX} - x\}$. Under the above assumptions, this can according to (2.38.2) be granted by providing select-and-ultimate tables containing annual mortality rates

$$(q_{x,u})_{x \in AB, u \in \{0,...,x-x_{MIN}\}} \tag{2.38.5}$$

and annual withdrawal rates

$$(w_{x,u})_{x \in AB, u \in \{0,...,x-x_{MIN}\}}, \tag{2.38.6}$$

respectively. These annual rates depend on both the attained age of an insured and the previous sojourn time in the portfolio.

To provide numerical examples and corresponding illustrations (recall figures 1 and 2 from the introduction), age- and duration-depending annual withdrawal rates are used. These withdrawal rates originate from a real existing PKV portfolio of adults that was structured with respect to age $x \in AB := \{x_{MIN} = 21, ..., x_{MAX} = \omega = 100\}$ and previous sojourn time in the portfolio $U_x = u \in \{0, ..., 15\}$. For each of these groups of insured having same age and previous sojourn time in the portfolio, the relative frequency of withdrawal was derived. Afterwards, a two-dimensional kernel smoothing estimator (Nadaraya-Watson) was applied. Safety margins were omitted. According to this procedure, we obtained a select-and-ultimate table $({}^{\mathbf{u}}\hat{w}_{x,u})_{x \in AB, u \in \{0,...,15\}}$ of so-called *independent* withdrawal probabilities. The number 15 corresponds here to the select period r. For previous sojourn times greater than $r = 15$, we set ${}^{\mathbf{u}}\hat{w}_{x,u} := {}^{\mathbf{u}}\hat{w}_{x,15}$, $u \geq 15$. Incidentally, the word *independent* means in the present context that the decrement cause *withdrawal* does not compete with the other cause *dead* in determining ${}^{\mathbf{u}}\hat{w}_{x,u}$. For more details concerning this issue, Bowers et al. ([1997], section 10.5) or Milbrodt and Helbig ([1999], section 3C) can be consulted.

Independent annual mortality rates $({}^{\mathbf{u}}q_x)_{x \in AB}$ are taken from the PKV mortality table 2007. These mortality rates do not depend on the previous sojourn time in the portfolio. Hence, a durational effect is here only implemented for the decrement cause *withdrawal.* However, by means of the conversion of independent annual rates into dependent annual rates, according to the formulas

$$w_{x,u} = {}^{\mathbf{u}}\hat{w}_{x,u} \left(1 - \frac{1}{2}\,{}^{\mathbf{u}}q_x\right) \tag{2.38.7}$$

and

$$q_{x,u} = {}^{\mathbf{u}}q_x \left(1 - \frac{1}{2}\,{}^{\mathbf{u}}\hat{w}_{x,u}\right), \tag{2.38.8}$$

the withdrawal rates as well as the mortality rates depend on both the age of an insured and the previous sojourn time in the portfolio. The duration-dependence of the annual mortality rates will be referred to as secondary dependence, since it is only caused by the duration-dependence of the annual withdrawal rates. The formulas (2.38.7) and (2.38.8) can also be found in Bowers et al. ([1997], section 10.5) or Milbrodt and Helbig ([1999], section 3C, see also exercise 3.17(c)). Note that these formulas are usually approximations, yet are exact in case of two decrement causes. Further, we should mention that the dependent and the independent probabilities do not differ significantly in size. For this reason, the conversion from independent rates into dependent rates is often omitted. In summary, we obtain in the above described manner tables according to (2.38.5) and (2.38.6). These tables can be used to derive the probabilities ${}_kp_{x,u}$ with the aid of (2.38.2).

In order to quantify the durational effects due to age- and duration-depending withdrawal rates, a corresponding model based on mere age-depending decrement rates is additionally computed. For this model, the independent mortality rates are also taken from the above mortality table. Concerning mere age-depending withdrawal rates, it would be usual to disregard the structure of the PKV portfolio with respect to the previous contract duration. This could be achieved by adding the number of all insured within the portfolio groups that pertain to the same age. (To be precise, for the estimation of withdrawal rates in practise, only insured with a previous contract duration of more that three years are taken into account.) After aggregating these portfolio groups, relative frequencies could be derived and a smoothing estimator would yield independent withdrawal rates which are merely age-depending. Yet, in order to ensure that all possible differences between the results of both models are purely caused by regarding or disregarding the durational effects - and not by changing parameters of the kernel smoothing estimator or by the use of a completely different smoothing approach (one-dimensional) - the age-depending withdrawal rates are derived as follows: The portion of withdrawing insured is for each group containing insured of the same age calculated with the aid of the (independent) age- and duration-depending withdrawal rates. Let $\tilde{L}_{x,u}$ be the number of insured with attained age $x \in AB$ and previous sojourn time $u \in \{0, ..., r-1 = 14\}$. Further, let $\tilde{L}_{x,\geq r}$ be the number of insured with attained age $x \in AB$ and previous sojourn time equal to or greater than $r = 15$. Hence, $\tilde{L}_x = \sum_{u=0}^{r-1} \tilde{L}_{x,u} + \tilde{L}_{x,\geq r}$ is the number of all insured at the age x. Then, the independent withdrawal rates are specified by

$$^{\mathbf{u}}w_x \quad := \quad \frac{\sum_{u=0}^{r-1} \tilde{L}_{x,u}\ {}^{\mathbf{u}}\hat{w}_{x,u} + \tilde{L}_{x,\geq r}\ {}^{\mathbf{u}}\hat{w}_{x,r}}{\tilde{L}_x}, \quad x \in AB. \tag{2.38.9}$$

The corresponding dependent annual decrement rates q_x and w_x can afterwards be obtained by applying the formulas (2. 38. 7) and (2. 38. 8).

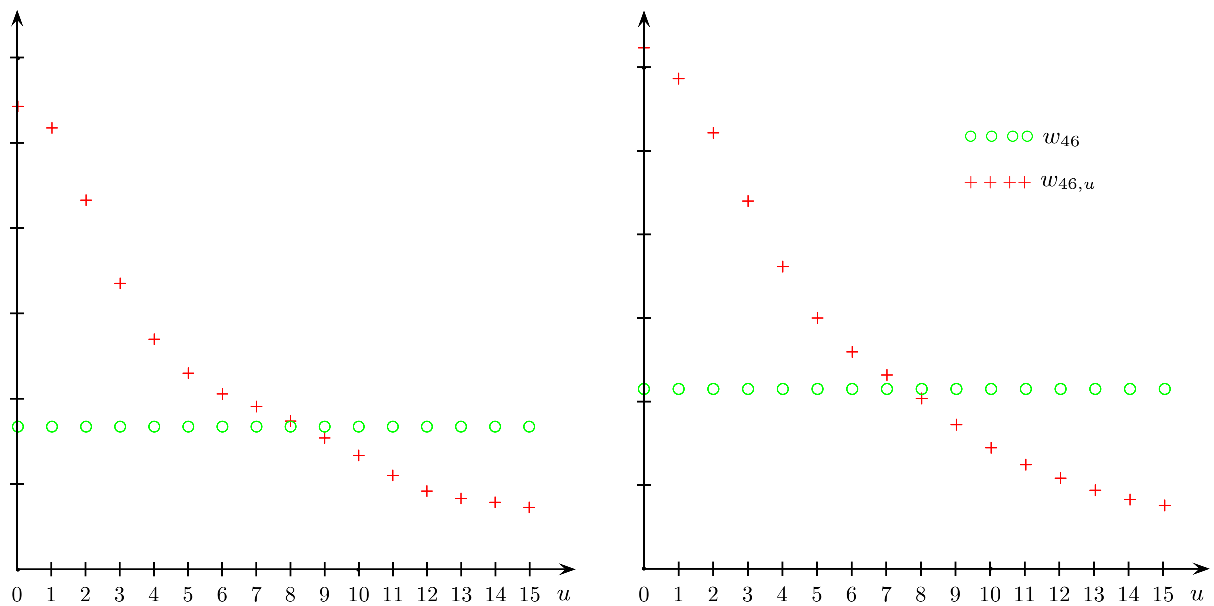

Figure 13: Age- and duration-depending annual withdrawal rates compared with mere age-depending withdrawal rates with respect to the previous contract duration u, for men (left) and women (right) at the age of $x = y = 46$

Figure 1 has outlined, in case of insured at the age of 36, both mere age-depending and age- and duration-depending withdrawal rates with respect to the previous contract duration u. In addition, figure 13 illustrates the same for insured at the age of 46. Compared to figure 1, it can be observed that here the mere age-depending withdrawal rates underestimate the actual withdrawal rates for previous contract durations of almost ten years, instead of three or four years. Hence, the actual withdrawal rates are overestimated only for insured with previous sojourn time in the portfolio of more than a decade. As exemplified later, this results in the fact that premiums in a model relying on age- and duration-depending decrement rates, when compared to premiums in a corresponding model that relies on mere age-depending decrement rates, become higher for new entries of younger ages. Premiums for older new entries, however, become lower.

As mentioned above, the estimation of mere age-depending withdrawal rates usually disregards insured with previous contract durations of less than three years. Thus, the relatively high withdrawal rates for these insured are not taken into account. As we will see, proceeding this way results in lower mere age-depending withdrawal rates. But the problems due to the actual age- and duration-dependence of withdrawal rates largely remain. To disregard insured with previous contract durations of less than three years, (2. 38. 9) must be replaced by

$$ {}^{\mathbf{u}}\tilde{w}_x \;:=\; \frac{\sum_{u=3}^{r-1} \tilde{L}_{x,u}\, {}^{\mathbf{u}}\hat{w}_{x,u} + \tilde{L}_{x,\geq r}\, {}^{\mathbf{u}}\hat{w}_{x,r}}{\sum_{u=3}^{r-1} \tilde{L}_{x,u} + \tilde{L}_{x,\geq r}}, \quad x \in AB. \tag{2. 38. 10} $$

The dependent withdrawal rates based on (2. 38. 10) are for female insured of the ages $y = 36$ and $y = 46$ illustrated in figure 14. This figure allows one to compare these withdrawal rates with the mere age-dependent withdrawal rates based on (2. 38. 9) as well as the age- and duration-depending withdrawal rates, both of which have been illustrated in figure 1 and figure 13. It can be observed that the withdrawal rates based on (2. 38. 10) are actually lower than the withdrawal rates relying on (2. 38. 9). Yet, the age-and duration-depending withdrawal rates remain overestimated for insured with long previous contract durations. Hence, by using (2. 38. 10) instead of (2. 38. 9), only quantitatively different results would be obtained.

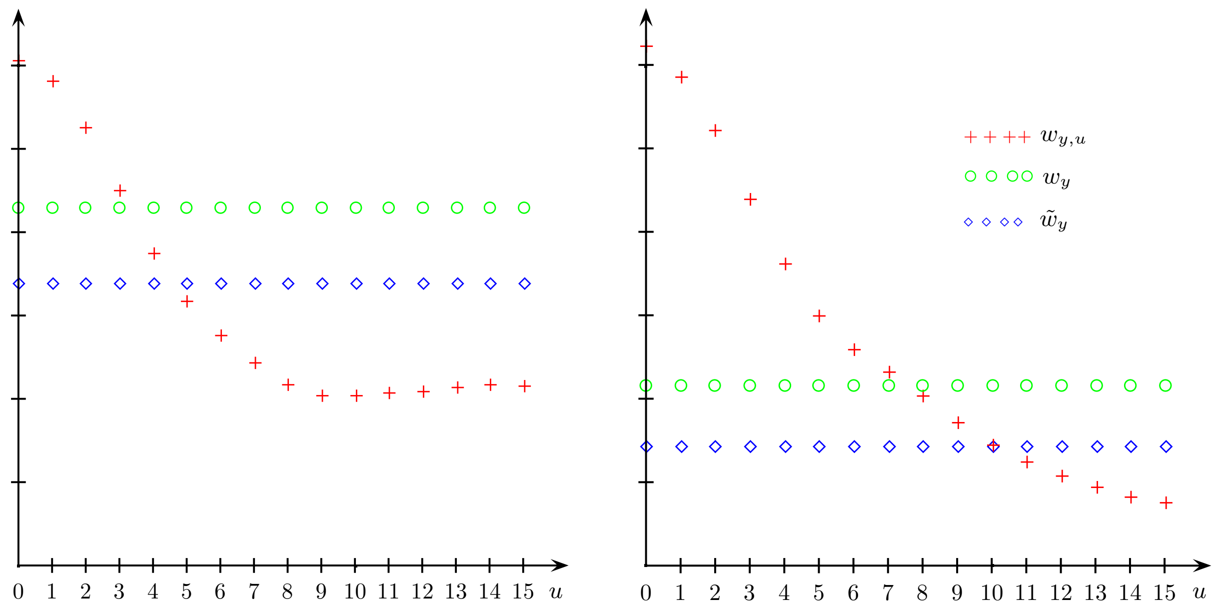

Figure 14: Age- and duration-depending annual withdrawal rates, mere age-depending withdrawal rates based on (2. 38. 9), and mere age-depending withdrawal rates based on (2. 38. 10) with respect to the previous contract duration u, for women at the age of $y = 36$ (left) and women at the age of $y = 46$ (right)

Our results (see also example 4. 13) are derived by using (2. 38. 9). The reason for this is that we want to work out the durational effects as they appear for the portfolio under consideration. Further, a more effective approach of taking into account the duration-dependence of certain model parameters is presented. For this, we refer to the continuation of this example, example 4. 13.

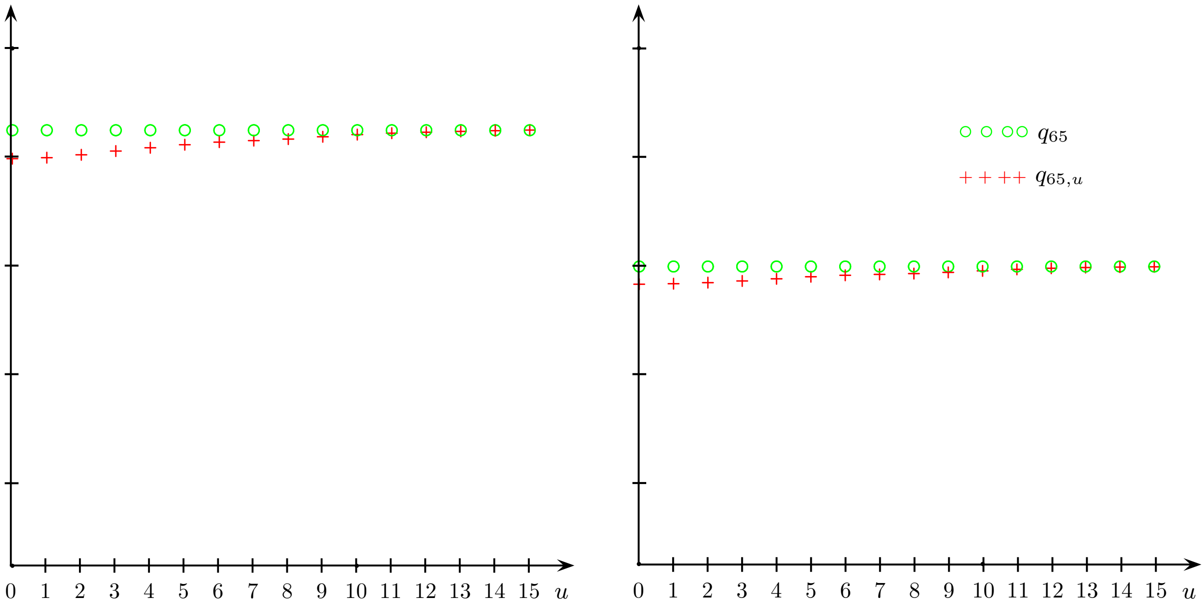

Figure 15: Age- and (secondarily) duration-depending annual mortality rates compared with mere age-depending mortality rates with respect to the previous duration u, for men (left) and women (right) at the age of $x = y = 65$

Figure 15 illustrates the secondary dependence of the annual mortality rates on the previous contract duration. With increasing previous contract duration, the smaller age- and duration-depending annual mortality rates converge against the mere age-depending annual mortality rates. Hence, this effect corresponds to the growth of mortality rates for new entries with time since selection due to the health examination during the application procedure of PKV companies.

The probabilities shown in figure 2 are probabilities of remaining k-years in the portfolio, $k \in \{0, ..., \omega - x\}$. According to (2. 38. 2), these probabilities can be derived by using both the annual mortality rates and the annual withdrawal rates. Regarding figure 2, we have assumed for both models an attained age of 31. △

We now return to backward and forward integral equations for the transition probabilities of the bivariate Markov process $((X_t, U_t))_{t\geq 0}$ appertaining to a semi-Markov process $(X_t)_{t\geq 0}$. Milbrodt and Helbig ([1999], Folgerung 4.49) established another two systems of integral equations for the transition probabilities of a Markov process. These equations hold for transition probabilities which identically satisfy the backward integral equations and the forward integral equations, respectively. Referred to as backward and forward integral equations of type 2, these equations turned out to be formal integrations of the appertaining differential equations. In almost the same manner, backward integral equations as well as forward integral equations of type 2 can be established for the transition probabilities of the bivariate Markov process $((X_t, U_t))_{t\geq 0}$. In order to verify the backward integral equations of type 2, the following relationship must be provided.

2. 39 Lemma. *Let $(X_t, \mathfrak{F}_t)_{t\geq 0}$ be a semi-Markov process and $((X_t, U_t))_{t\geq 0}$ the appertaining bivariate Markov process with conditional distributions Q_* according to definition* 2. 24 *and regular cumulative transition intensities q_*. Further, let $0 \leq u \leq s \leq t < \infty$, and $y \in \mathcal{S}$. Then*

$$Q_y(s, t, u) \quad = \quad -\int_{(s,t]} \big(1 - Q_y(\tau, t, \tau - s + u)\big)\, q_{yy}(s, d\tau, u). \qquad (2.\,39.\,1)$$

PROOF. Let $0 \le u \le s \le t < \infty$, and $y \in \mathcal{S}$. In the case of $Q_y(s,t,u) < 1$, (2. 27. 8) and (2. 24. 5) give for the right-hand side of (2. 39. 1)

$$\begin{aligned}
-\int_{(s,t]} \bar{p}_y(\tau,t,\tau-s+u)\, q_{yy}(s,d\tau,u) &= -\int_{(s,t]} \frac{\bar{p}_y(s,t,u)}{\bar{p}(s,\tau,u)}\, q_{yy}(s,d\tau,u) \\
&= \bar{p}_y(s,t,u) \int_{(s,t]} \frac{Q_y(s,d\tau,u)}{\bar{p}_y(s,\tau-0,u)\,\bar{p}_y(s,\tau,u)} \\
&= \bar{p}_y(s,t,u) \int_{(s,t]} -\frac{\bar{p}_y(s,d\tau,u)}{\bar{p}(s,\tau-0,u)\,\bar{p}_y(s,\tau,u)} \\
&= \bar{p}_y(s,t,u) \int_{(s,t]} \Big(\frac{1}{\bar{p}_y(s,\cdot,u)}\Big)(d\tau). \qquad (2.39.2)
\end{aligned}$$

The last equation is an application of corollary A.2, with the right continuity of $\bar{p}_y(s,\cdot,u)$ following from the right continuity of $-q_{yy}(s,\cdot,u)$ and (2. 24. 7). Finally, (2. 39. 2) can be continued as

$$= \bar{p}_y(s,t,u)\Big(\frac{1}{\bar{p}_y(s,t,u)} - \frac{1}{\bar{p}_y(s,s,u)}\Big) = 1 - \bar{p}_y(s,t,u) = Q_y(s,t,u).$$

By employing the regularity of q_{yy}, it can simply be checked that (2. 39. 1) remains valid in the case of $Q_y(s,t,u) = 1$. □

The following backward integral equations of type 2 are generalizations of the backward integral equations provided by Folgerung 4.49 in Milbrodt and Helbig [1999]. They allow the transition probabilities and the cumulative transition intensities to additionally depend on the time spent in the current state. The proof of the subsequent lemma can be found in A.8 in the appendix.

2. 40 Lemma. [Backward integral equations of type 2] *Let $(X_t, \mathfrak{F}_t)_{t\ge 0}$ be a semi-Markov process and $((X_t, U_t))_{t\ge 0}$ the appertaining bivariate Markov process with transition probability matrix p and regular cumulative transition intensity matrix q. The components of p are assumed to satisfy the backward integral equations* (2. 32. 2) *identically. Further, let $0 \le u \le s \le t < \infty, v \ge 0$, and $(y,z) \in \mathcal{S}^2$. Then, the components of p likewise satisfy*

$$\begin{aligned}
p_{yz}(s,t,u,v) &= \delta_{yz}\mathbf{1}_{(v \ge t-s+u)} + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\ne y}} \int_{(s,t]} p_{\xi z}(\tau,t,0,v)\, q_{y\xi}(s,d\tau,u) \\
&\quad + \int_{(s,t]} p_{yz}(\tau,t,\tau-s+u,v)\, q_{yy}(s,d\tau,u). \qquad (2.40.1)
\end{aligned}$$

To generalize the forward integral equations of type 2 provided by Folgerung 4.49 in Milbrodt and Helbig [1999], the next lemma establishes forward integral equations of type 2 for versions of the transition probabilities of the bivariate Markov process $((X_t, U_t))_{t\ge 0}$. However, analogously to lemma 2. 36, assumption 2. 35 must be fulfilled. The proof can be found in A.9 in the appendix.

2. 41 Lemma. [Forward integral equations of type 2] *Let $(X_t, \mathfrak{F}_t)_{t\ge 0}$ be a semi-Markov process and $((X_t, U_t))_{t\ge 0}$ the appertaining bivariate Markov process with transition matrix p and regular cumulative intensity matrix q. For the transition intensities q_* we assume* 2. 35*. Further, let $0 \le u \le s \le t < \infty, v \ge 0$, and $(y,z) \in \mathcal{S}^2$. In cases where the components of p satisfy the* *forward integral equations* (2. 36. 1) *identically, they likewise satisfy*

$$p_{yz}(s,t,u,v) \;=\; \delta_{yz}\boldsymbol{1}_{(v\geq t-s+u)} + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{[t-v,t]}\int_{(0,\infty)} \lambda_{\xi z}(\tau,l)\,p_{y\xi}(s,\tau-0,u,dl)\,\boldsymbol{\Lambda}_{\xi z}(d\tau)$$
$$+\int_{[t-v,t]}\int_{(0,\tau-t+v]} \lambda_{zz}(\tau,l)\,p_{yz}(s,\tau-0,u,dl)\,\boldsymbol{\Lambda}_{zz}(d\tau). \qquad (2.41.1)$$

D.2 A special case - Markovian pure jump processes

Recall that a non-homogeneous semi-Markov process X with state space $\mathcal{S}$ was defined as the appertaining pure jump process for a given homogeneous Markovian marked point process (T, Z) with space of marks S. Further, this process is not necessarily Markovian although, conversely, the Markov property of a pure jump process X implies the Markov property of the appertaining marked point process (T, Z) (see Milbrodt and Helbig [1999], Hilfssatz 4.33). Hence, pure jump processes that are Markovian form a subclass of all semi-Markovian pure jump processes.

A Markovian pure jump process $(X_t)_{t\geq 0}$ is characterized by possessing the elementary Markov property (cf. (2. 5. 10)). According to the finite state space $\mathcal{S}$, the transition probabilities can be specified for each pair of states $(y, z) \in \mathcal{S}^2$ by

$$p_{yz}(s,t) = P(X_t = z|\, X_s = y) \;\; \mathcal{L}(X_s|P)-a.s., \;\; 0\leq s\leq t. \qquad (2.41.2)$$

Hence, in contrast to (2. 21. 8), the conditional probability of X being in state $y \in \mathcal{S}$ at time t depends only on the state occupied at time s, and not on the time elapsed since entering this state. In this situation, the conditional distribution of the time of the first jump after s, $T(s)$, and the corresponding destination state, $X_{T(s)}$, also does not depend on the time since entering the current state. The same holds for the appertaining cumulative transition intensities. In the same manner as Milbrodt and Helbig ([1999], chapter 4), we define for $(y, z) \in \mathcal{J}$ and $s \geq 0$

$$E_{yz}(s,\cdot):\; [s,\infty)\ni t\mapsto P(T(s)\leq t, X_{T(s)} = z|\, X_s = y), \qquad (2.41.3)$$
$$E_{y}(s,\cdot):\; [s,\infty)\ni t\mapsto \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} E_{yz}(s,t) = P(T(s)\leq t|\, X_s = y), \qquad (2.41.4)$$
$$\bar{p}_y(s,t):\; = 1 - E_y(s,t) = P(X_\tau = y, \tau\in[s,t]|\, X_s = y),\;\; t\geq s, \qquad (2.41.5)$$
$$q_{yz}(s,\cdot):\; [s,\infty)\ni t\mapsto \int_{(s,t]} \frac{E_{yz}(s,d\tau)}{1-E_y(s,\tau-0)} \in [0,\infty]\,, \qquad (2.41.6)$$
$$q_{yy}(s,\cdot):\; [s,\infty)\ni t\mapsto -\int_{(s,t]} \frac{E_{y}(s,d\tau)}{1-E_y(s,\tau-0)} \in [-\infty,0]. \qquad (2.41.7)$$

According to Milbrodt and Helbig ([1999], Hilfssatz 4.29), regular cumulative transition intensities q_* for a Markov process $(X_t)_{t\geq 0}$ identically satisfy, among other things, the additivity property

$$q_{yz}(s,r) + q_{yz}(r,t) = q_{yz}(s,t), \;\; s\leq r\leq t, \qquad (2.41.8)$$

instead of (2. 27. 1). Hence, every regular cumulative transition intensity q_{yz}, $(y, z) \in \mathcal{J}$ can - by means of $q_{yz}((s, t]) := q_{yz}(s, t)$ - be viewed as a Borel measure on $\mathfrak{B}((0, \infty))$. Further, instead of (2. 27. 8), the following holds:

$$\bar{p}_y(s,t) = \bar{p}_y(s,r)\,\bar{p}_y(r,t), \;\; s\leq r\leq t. \qquad (2.41.9)$$

In section 2.C, a condition (cf. (2. 11. 2)) was mentioned which was supposed to ensure, for a given marked point process with cumulative transition intensity matrix $\hat{q}$, that the appertaining pure jump process is Markovian. This condition is given by

$$\hat{q}_{yz}(s,t) = \hat{q}_{yz}(s,r) + \hat{q}_{yz}(r,t), \;\; 0 \le s \le r \le t, (y,z) \in \mathcal{S}^2. \tag{2. 41. 10}$$

The following theorem establishes that (2. 41. 10) is indeed sufficient for the process $X = G^{-1}(T,Z)$ to be Markovian. It can be proved (see appendix A.10) by using some of the results provided by Milbrodt and Helbig ([1999], section 4B), which themselves were inspired by Jacobsen [1972]. Further, it turns out that the cumulative transition intensities for the marked point process and the corresponding cumulative transition intensities for the pure jump process coincide.

2. 42 Theorem. *Let $((T_m, Z_m))_{m\in\mathbb{N}_0}$ be a homogeneous Markovian marked point process with transition probabilities $\hat{Q}_*$ according to definition* 2. 8*, and regular cumulative transition intensities $\hat{q}_*$ satisfying* (2. 41. 10)*. Then, the appertaining pure jump process $X = G^{-1}(T,Z)$ is a Markov process with transition probabilities*

$$p_{yz}(s,t) \;\; = \;\; \sum_{l=0}^{\infty} p_{yz}^{(l)}(s,t), \;\;\; 0 \le s \le t, (y,z) \in \mathcal{S}^2, \tag{2. 42. 1}$$

where $p_{yz}^{(l)}(s,t) := P(T_l \le t < T_{l+1}, Z_l = z |\, T_0 = s, Z_0 = y)$. Further, for $s \ge 0$ and $\mathcal{L}(X_s|P)$-a.e. $y \in \mathcal{S}$

$$E_{yz}(s,t) = \hat{Q}_{yz}(s,t), \;\; t \ge s, z \in \mathcal{S}, \tag{2. 42. 2}$$

and

$$q_{yz}(s,t) = \hat{q}_{yz}(s,t), \;\; t \ge s, z \in \mathcal{S}. \tag{2. 42. 3}$$

Due to (2. 41. 8) and (2. 42. 3), the property (2. 41. 8) is likewise satisfied by the cumulative transition intensities $\hat{q}_*$ for a marked point process (T,Z) appertaining to a given Markovian pure jump process X with regular cumulative transition intensity matrix q. Hence, the pure jump process $X = G^{-1}(T,Z)$ is a Markov process if and only if the cumulative transition intensities of the marked point process (T,Z) satisfy (2. 41. 10). Further, due to (2. 42. 2) and (2. 42. 3), the first three assertions of lemma 2. 19 also hold for E_* and q_*. Finally, by employing (2. 42. 2) and (2. 42. 3) as well as the last assertion of lemma 2. 19, the integral equations in section 2.D.1 can be reduced to the known integral equations for the transition probabilities within a non-smooth Markov framework (cf. Milbrodt and Helbig [1999], section 4C). For example, the transition probabilities of $X = G^{-1}(T,Z)$ according to (2. 41. 2) satisfy, for $0 \le s \le t$ and $(y,z) \in \mathcal{S}^2$ with $P(X_s = y) > 0$, the backward integral equations

$$p_{yz}(s,t) = \delta_{yz}\,(1 - E_y(s,t)) + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \int_{(s,t]} p_{\xi z}(\tau,t)\,(1 - E_y(s,\tau - 0))\,q_{y\xi}(d\tau), \tag{2. 42. 4}$$

as well as the forward integral equations

$$p_{yz}(s,t) = \delta_{yz}(1 - E_y(s,t)) + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{(s,t]} (1 - E_z(\tau,t))\,p_{y\xi}(s,\tau - 0)\;q_{\xi z}(dr). \tag{2. 42. 5}$$

Chapter 3

Interest, payments and reserves

Firstly, we are concerned with defining capital accumulation functions, the appertaining discounting functions, and general payments streams, thus allowing us to describe and evaluate cash flows which represent payments of premiums and benefits. Further, some relations between these quantities are established which will be used in the sequel. For proofs we refer to Milbrodt and Helbig ([1999], sections 2A and 2C).

3. 1 Definition. A *capital accumulation function* $K : [0,\infty) \to [1,\infty)$ is a non-decreasing, right continuous function of locally bounded variation. For $t \geq 0$, $K(t)$ is interpreted as the value at time t of an initial capital $K(0) = 1$. The reciprocal $v := \frac{1}{K}$ is referred to as the *discounting function.* The *cumulative interest intensity* Φ is given by

$$\Phi : [0,\infty) \ni t \mapsto \int_{(0,t]} \frac{K(d\tau)}{K(\tau - 0)}. \tag{3. 1. 1}$$

3. 2 Corollary. *It follows from the above definition that, on* $\mathfrak{B}([0,\infty))$*,*

$$K(d\tau) = K(\tau - 0)\,\Phi(d\tau), \tag{3. 2. 1}$$

and

$$v(d\tau) = -v(\tau)\,\Phi(d\tau). \tag{3. 2. 2}$$

Further, one obtains for $0 \leq s \leq t < \infty$

$$v(t) \cdot K(s) = 1 - v(t) \int_{(s,t]} K(\tau - 0)\,\Phi(d\tau). \tag{3. 2. 3}$$

3. 3 Definition. A *directed payment stream* is a distribution function $Z : [0,\infty) \to [0,\infty)$. Each function $Z = Z_1 - Z_2$, where the components Z_* are directed payment streams satisfying $Z_1(\infty) \wedge Z_2(\infty) < \infty$, is referred to as a *payment stream.* For each $t \geq 0$, $Z(t)$ is regarded as the accumulated value of all payments in $[0, t]$. The directed payment streams Z_1 and Z_2 are interpreted as net outgoings and net incomes, respectively.

A Reserves in a non-random set-up

The aim of this section is to introduce prospective and retrospective reserves in a deterministic set-up. For this, we follow Milbrodt and Helbig ([1999], section 2C). It turns out that for equivalent payment streams both reserves coincide. This, however, is not the case within a random framework. Before defining prospective and retrospective reserves, present values and accumulated values of payment streams must be provided.

3. 4 Definition. Let K be a capital accumulation function and v the appertaining discounting function. For a given payment stream Z, the *accumulated value* of all payments up to and including time $t \geq 0$ is given as

$$s(Z)(t) := K(t) \int_{[0,t]} v(\tau)\, Z(d\tau). \tag{3.4.1}$$

Accordingly, the *present value* is given as

$$a(Z)(t) := \int_{[0,t]} v(\tau)\, Z(d\tau). \tag{3.4.2}$$

The present value of the overall payment stream is defined by

$$a(Z) := \int_{[0,\infty)} v(\tau)\, Z(d\tau). \tag{3.4.3}$$

3. 5 Definition. Two deterministic payment streams $Z_i, i = 1, 2$ are called *equivalent* with respect to a capital accumulation function K if their present values with respect to K are finite and coincide, i.e. $a(Z_1) = a(Z_2) < \infty$.

3. 6 Definition. [Prospective and retrospective reserve] Let K be a capital accumulation function. Further, let Z_b and Z_p be directed payment streams satisfying $a(Z_b) \wedge a(Z_p) < \infty$, which are assumed to describe the accumulation of benefits and premiums, respectively. Then, the *prospective reserve* at time $t \geq 0$ is defined as

$$V^+(t) := K(t) \left(\int_{[t,\infty)} v(\tau)\, Z_b(d\tau) - \int_{[t,\infty)} v(\tau)\, Z_p(d\tau) \right). \tag{3.6.1}$$

Thus, the prospective reserve is the present value at time t of the difference between all future benefits and premiums. Accordingly, the *retrospective reserve* at time $t \geq 0$ is given as

$$V^-(t) := K(t) \left(\int_{[0,t)} v(\tau)\, Z_p(d\tau) - \int_{[0,t)} v(\tau)\, Z_b(d\tau) \right). \tag{3.6.2}$$

This is the accumulated value of the difference between all previously payed premiums and benefits.

In life insurance one often has to deal with long-term contracts. Though the actual insured risk changes over time, the net premiums are usually uniformly distributed over the policy term. Due to this compensation of risk over the term of a contract, reserves are generated. For example, in cases where the risk increases over time, the net premium is at first higher than the natural premium which covers the actual risk for a certain period. Thus, the net premium contains a portion referred to as savings premium which generates the prospective reserve. In later years of the contract term, the prospective reserve is used to fill the gap between the natural premium and the net premium. At that time, the former will prevail. In contrast, short-term insurance contracts for which the net premiums are simply the expected benefits for the contract term do not generate reserves. This is often the case in non-life insurance.

By interpreting two directed payment streams as accumulated premiums and accumulated benefits respectively, the prospective reserve is the present value at time $t \geq 0$ of the difference between future benefits and future premiums. For equivalent payment streams, the balance between the present values of all premiums and all benefits at policy issue shifts into a balance at time $t > 0$. Regarding the prospective reserve, this balance is between future benefits on the one side, and future premiums along with the prospective reserve on the other side. Regarding the retrospective reserve, this balance is between previously payed premiums, and previously payed benefits along with the retrospective reserve.

The following corollary states the equality of both reserves in a deterministic set-up, provided that the corresponding payment streams Z_b and Z_p are equivalent.

3. 7 Corollary. *Let Z_b and Z_p be two deterministic directed payments streams which are equivalent with respect to the capital accumulation function K. Then, for each $t \geq 0$, the prospective and the retrospective reserves coincide, i.e. $V^-(t) = V^+(t)$.*

In a random set-up for life insurance modelling, however, the situation is different. For payment streams that are assumed to have equal expected present values at policy issue, the equality of both reserves must be replaced by the equality of the expected values of the prospective reserve and the retrospective reserve. Further, the prospective reserve is defined as conditional expectation of future benefits less future premiums. This concept is not a matter of dispute. Regarding the retrospective reserve, however, different definitions can be given that satisfy the condition for the expectation of the retrospective reserve to coincide at any time with the expectation of the prospective reserve. Different concepts for the retrospective reserve will be discussed in section 5A.

B Relating payment streams and probabilities

Here, a technical lemma is provided which introduces some relationships between payments streams, capital accumulation functions, and the probabilities of remaining in a certain state $\bar{p}_y,\ y \in \mathcal{S}$. The proof can be found in A.11 in the appendix.

3. 8 Lemma. *Let Z be a payment stream, q a regular cumulative transition intensity matrix according to definition 2. 28, K a capital accumulation function with cumulative interest intensity Φ, and v the appertaining discounting function. Then, for all $0 \leq u \leq s \leq t$, and $y \in \mathcal{S}$*

$$\begin{aligned}\int_{[s,t]} Z(d\tau) &= K(s) \int_{[s,t]} v(\tau)\, \bar{p}_y(s,\tau,u)\, Z(d\tau) \\ &\quad + K(s) \int_{(s,t]} v(\tau)\, \bar{p}_y(s,\tau-0,u)\, Z([\tau,t])\, \Phi(d\tau) \\ &\quad - K(s) \int_{(s,t]} v(\tau)\, \bar{p}_y(s,\tau-0,u)\, Z([\tau,t])\, q_{yy}(s,d\tau,u), \qquad (3.8.1)\end{aligned}$$

and

$$\begin{aligned}\int_{(s,t]} Z(d\tau) &= K(s) \int_{(s,t]} v(\tau)\, \bar{p}_y(s,\tau-0,u)\, Z(d\tau) \\ &\quad + K(s) \int_{(s,t]} v(\tau)\, \bar{p}_y(s,\tau-0,u)\, Z([\tau,t])\, \Phi(d\tau) \\ &\quad - K(s) \int_{(s,t]} v(\tau)\, \bar{p}_y(s,\tau-0,u)\, Z((\tau,t])\, q_{yy}(s,d\tau,u). \qquad (3.8.2)\end{aligned}$$

Equation (3. 8. 1) is a generalization of equation (10.17.3) in Milbrodt and Helbig [1999]. Analogously, equation (3. 8. 2) generalizes (10.17.4). Both equations will be used in order to verify that Thiele's integral equations of type 2 imply Thiele's integral equations of type 1.

C Duration-depending actuarial payment functions

Recall that $(\mathcal{V},\mathfrak{V})$ denotes the measurable space of all possible patterns of states. An actuarial payment function for a policy (p) is here defined in a manner similar to Milbrodt and Helbig ([1999], section 6A). Yet, analogously to Stracke [1997], we omit for the sake of convenience the employment of another measurable space $(\mathcal{W},\mathfrak{W})$ and an isomorphism $I:(\mathcal{V},\mathfrak{V})\to(\mathcal{W},\mathfrak{W})$. Both were used by Milbrodt and Helbig [1999] to simplify, if possible, the description of a single risk by means of random patterns of states. For example, in a one-life-one-risk set-up, it is sufficient to record the time of death instead of the entire pattern of states.

Since actuarial payments are usually considered separately as sojourn payments SA and payments due to transitions DA, an actuarial payment function is given as $A=SA+DA$. This decomposition is not necessarily unique. Yet, in practise one usually starts by specifying the addends of the right-hand side. These addends, referred to as annuity payment function and assurance payment function, respectively, will be defined next. Note that, in contrast to Stracke [1997] and Milbrodt and Helbig [1999], both the sojourn payments and the payments due to transitions are additionally allowed to depend on the time since the current state was entered or on the time of the last previous jump. Recall that for $t\in(t_m(x),t_{m+1}(x)]$, the time of the last previous jump of $x\in\mathcal{V}$ is given by $t_m(x)$.

3. 9 Definition. Let $D_{yz}:(0,\infty)\times(0,\infty)\to[0,\infty),(y,z)\in\mathcal{J}$, be a measurable lump sum and $DT:(0,\infty)\to(0,\infty)$ a non-decreasing mapping with $DT\geq Id$. Then,

$$DA:=\sum_{(y,z)\in\mathcal{J}}DA_{yz}:\mathcal{V}\times[0,\infty)\to[0,\infty),\tag{3.9.1}$$

with

$$DA_{yz,x}(t):=\sum_{m\in\mathbb{N}_0}\int_{(0,t]}\mathbf{1}_{\{t_m(x)<\tau\leq t_{m+1}(x)\}}\,\mathbf{1}_{\{DT(\tau)\leq t\}}\,D_{yz}(\tau,\tau-t_m(x))\,H_{yz,d\tau}(x)\tag{3.9.2}$$

($t\geq0,x\in\mathcal{V}$, $(y,z)\in\mathcal{J}$, $H_{yz,\cdot}(x)$ according to (2. 3. 1)) is an *assurance payment function.* $DA_x(t)$ corresponds to the total amount of lump sums due to the pattern of states x, payed up to and including time t.
Further, let $\hat{F}_z:=\hat{F}_z^+-\hat{F}_z^-$, $z\in\mathcal{S}$, with the components $\hat{F}_z^+,\hat{F}_z^-:[0,\infty)\times\mathfrak{B}([0,\infty))\to[0,\infty)$ being kernels such that for $s\geq0$:

- $\hat{F}_z^+(s,\infty)\wedge\hat{F}_z^-(s,\infty)<\infty$,
- $\hat{F}_z^+(s,B)=\hat{F}_z^-(s,B)=0\ \forall B\in\mathfrak{B}([0,s))$, and
- $\hat{F}_z^+(s,B)=\hat{F}_z^-(s,B)=0,\ \forall B\in\mathfrak{B}([0,\infty))$ if $z\in\mathcal{S}$ is absorbing.

Then

$$SA:=\sum_{z\in\mathcal{S}}SA_z:\mathcal{V}\times[0,\infty)\to[0,\infty),\tag{3.9.3}$$

with

$$SA_{z,x}(t):=\sum_{m\in\mathbb{N}_0}\int_{[0,t]}\mathbf{1}_{\{t_m(x)\leq\tau<t_{m+1}(x)\}}\,\mathbf{1}_{\{x_\tau=z\}}\,\hat{F}_z(t_m(x),d\tau).\tag{3.9.4}$$

($t\geq0,x\in\mathcal{V}$, $z\in\mathcal{S}$) is an *annuity payment function.* $SA_x(t)$ is the total amount of sojourn payments due to the pattern of states x, payed up to and including time t. Positive payments are regarded as benefits and negative payments as premiums. The sum $A:=SA+DA$ is called an *actuarial payment function.*

Analogously to the usual understanding of an actuarial payment function, A according to the above definition defines a kernel

$$A : \mathcal{V} \times [0,\infty) \ni (x,t) \mapsto A_x(t) \in \mathbb{R} \tag{3.9.5}$$

such that

- $t \mapsto A_x(t)$ is the distribution function of a signed measure on $\mathfrak{B}([0,\infty))$ for every $x \in \mathcal{V}$
- $x \mapsto A_x(t)$ is $\mathfrak{V} - \mathfrak{B}(\mathbb{R})-$measurable for every $t \geq 0$.

An actuarial payment function A is called natural if for each $t \geq 0$, the annuity payment function as well as the assurance payment function are $\mathfrak{V}_t - \mathfrak{B}(\mathbb{R})$-measurable (cf. Milbrodt and Helbig [1999], section 5D). Thus, the word *natural* states that the payments up to a certain time do not depend on future values of the pattern of states. For an actuarial payment function A according to definition 3.9, this is granted since $t_m(x)$, $m \in \mathbb{N}$, are stopping times with respect to $(\mathfrak{V}_t)_{t\geq 0}$.

For a given actuarial payment function A, we now derive present values according to definition 3.4. The risk due to the policy (p) is described by the pure jump process $(X_t)_{t\geq 0}$ having paths in $\mathcal{V}$. With K being a capital accumulation function, the random present value of all payments due to the actuarial payment function A for the policy (p) is given by

$$B_{(p)}(\omega) := a(A_{X(\omega)}) = \int_{[0,\infty)} v(\tau)\, A_{X(\omega)}(d\tau). \tag{3.9.6}$$

For the present value at time $s \geq 0$ of all payments after and including time s, one obtains

$$B_{(p),s}(\omega) := K(s) \int_{[s,\infty)} v(\tau)\, A_{X(\omega)}(d\tau). \tag{3.9.7}$$

Beside the process $(X_t)_{t\geq 0}$, we also consider within a duration-depending framework the processes $((X_t, U_t))_{t\geq 0}$ and $((T_m, Z_m))_{m\in\mathbb{N}_0}$. In view of these processes, we introduce two different representations of present values of future actuarial payments. The first is based on pattern of states and the corresponding pattern of previous durations $(x, u(x)) \in (\mathcal{U}_{\mathcal{V}}, \mathfrak{U}_{\mathcal{V}})$. The second representation relies on the corresponding chains of jumps $((t_m, z_m))_{m\in\mathbb{N}_0} \in (\mathcal{T}, \mathfrak{T})$. Both representations are related by means of the mapping $\bar{G}$ according to (2.3.6).

For the subsequent representations of present values for an actuarial payment function A being well defined, the following assumption is stipulated:

3.10 Assumption.

$$\mathbf{A2}: \quad \mathbf{E}\Big[\sum_{z\in\mathcal{S}}\sum_{m\in\mathbb{N}_0}\int_{[0,\infty)} \mathbf{1}_{\{T_m\leq\tau<T_{m+1}\}}\, v(\tau)\, |\hat{F}_z|(T_m, d\tau)\Big] < \infty. \tag{3.10.1}$$

3.11 Lemma. *Let (p) be a policy described by a semi-Markovian pure jump process $(X_t)_{t\geq 0}$. $((X_t, U_t))_{t\geq 0}$ as well as $((T_m, Z_m))_{m\in\mathbb{N}_0}$ denote the appertaining bivariate Markov process as well as the homogeneous Markovian marked point process. Further, let K be a capital accumulation function, v the appertaining discounting function, and A an actuarial payment function according to definition 3.9 that satisfies (3.10.1). Then, for $s \geq 0$,*

$$\begin{aligned} C(s,(x,u(x))) \;:=\; & K(s)\sum_{z\in\mathcal{S}}\sum_{m\in\mathbb{N}_0}\int_{[s,\infty)} \mathbf{1}_{\{t_m(x)\leq\tau<t_{m+1}(x)\}}\mathbf{1}_{\{x_\tau=z\}}\, v(\tau)\, \hat{F}_z(t_m(x), d\tau) \\ & + K(s)\sum_{(y,z)\in\mathcal{J}}\int_{(s,\infty)} v(DT(\tau))\, D_{yz}(\tau, u_{\tau-0}(x))\, H_{yz,d\tau}(x) \end{aligned} \tag{3.11.1}$$

is $\mathfrak{B}([0,\infty))\otimes\mathfrak{U}_{\mathcal{V}}-\mathfrak{B}((-\infty,\infty])$*-measurable and satisfies* $B_{(p),s}=C(s,((X_t,U_t))_{t\geq s})$.
With $((X_t,U_t))_{t\geq s}$ *being considered as a process starting in* s *and* $\bar{G}(((X_t,U_t))_{t\geq s})$ *yielding the appertaining marked point process* $((T_m,Z_m))_{m\in\mathbb{N}_0}$ *with* $(T_0,Z_0):=(s-U_s,X_s)$*, one obtains*

$$B_{(p),s}=C(s,((X_t,U_t))_{t\geq s})=\hat{C}(s,\bar{G}(((X_t,U_t))_{t\geq s}))=\hat{C}(s,((T_m,Z_m))_{m\in\mathbb{N}_0}) \qquad (3.11.2)$$

with the $\mathfrak{B}([0,\infty))\otimes\mathfrak{T}-\mathfrak{B}((-\infty,\infty])$*-measurable mapping*

$$\begin{aligned}\hat{C}(s,((t_m,z_m))_{m\in\mathbb{N}_0}) &:= K(s)\sum_{m\in\mathbb{N}_0}\int_{[t_m\vee s,t_{m+1}\vee s)} v(\tau)\,\hat{F}_{z_m}(t_m,d\tau)\\ &\quad +K(s)\sum_{m\in\mathbb{N}_0}\frac{D_{z_m z_{m+1}}(t_{m+1},t_{m+1}-t_m)}{K(DT(t_{m+1}))}\mathbf{1}_{\{s<t_{m+1}<\infty\}}. \qquad (3.11.3)\end{aligned}$$

Proof. The first assertion follows immediately from definition 3.9, definition 3.4, and (2.3.5). (3.11.2) follows from (2.4.3), (2.5.9), and $\bar{G}((X_t,U_t)_{t\geq s})=((T_m,Z_m))_{m\in\mathbb{N}_0}$ by means of

$$\begin{aligned}v(s)\,C(s,((X_t,U_t))_{t\geq s}) &= \sum_{z\in\mathcal{S}}\sum_{m\in\mathbb{N}_0}\int_{[s,\infty)}\mathbf{1}_{\{T_m\leq\tau<T_{m+1}\}}\mathbf{1}_{\{X_\tau=z\}}\,v(\tau)\,\hat{F}_z(T_m,d\tau)\\ &\quad+\sum_{(y,z)\in\mathcal{J}}\int_{(s,\infty)} v(DT(\tau))\,D_{yz}(\tau,U_{\tau-0})\,N_{yz,d\tau}\\ &= \sum_{z\in\mathcal{S}}\sum_{m\in\mathbb{N}_0}\int_{[T_m\vee s,T_{m+1}\vee s)}\mathbf{1}_{\{X_\tau=z\}}\,v(\tau)\,\hat{F}_z(T_m,d\tau)\\ &\quad+\sum_{(y,z)\in\mathcal{J}}\ \sum_{\{\tau>s|\,X_{\tau-0}=y\wedge X_\tau=z\}} v(DT(\tau))\,D_{yz}(\tau,U_{\tau-0})\\ &= \sum_{m\in\mathbb{N}_0}\int_{[T_m\vee s,T_{m+1}\vee s)} v(\tau)\,\hat{F}_{Z_m}(T_m,d\tau)\\ &\quad+\sum_{m\in\mathbb{N}_0}\frac{D_{Z_m Z_{m+1}}(T_{m+1},T_{m+1}-T_m)}{K(DT(T_{m+1}))}\mathbf{1}_{\{s<T_{m+1}<\infty\}}\\ &= v(s)\,\hat{C}(s,\bar{G}(((X_t,U_t))_{t\geq s})).\end{aligned}$$

□

For $t\geq 0$, the overall payment stream due to a policy (p) can be split up into payments after time t and payments up to time t. Hence, the present value of all payments also splits up into two addends. In view of the definition of both prospective reserves and retrospective reserves, we define similarly to Norberg [1991] the following decomposition of the present value of all payments.

3.12 Definition. Let K be a capital accumulation function and v the appertaining discounting function. Further, let A be an actuarial payment function according to definition 3.9 such that (3.10.1) is granted. For $t\geq 0$, we define

$$\begin{aligned}V^-(t,A) &:= -K(t)\sum_{m\in\mathbb{N}_0}\int_{[T_m\wedge t,T_{m+1}\wedge t)} v(\tau)\,\hat{F}_{Z_m}(T_m,d\tau)\\ &\quad -K(t)\sum_{m\in\mathbb{N}_0}\frac{D_{Z_m Z_{m+1}}(T_{m+1},T_{m+1}-T_m)}{K(DT(T_{m+1}))}\mathbf{1}_{\{T_{m+1}\leq t\}}, \qquad (3.12.1)\end{aligned}$$

and accordingly

$$\begin{aligned} V^+(t,A) \;\; := \;\; & K(t) \sum_{m\in\mathbb{N}_0} \int_{[T_m\vee t, T_{m+1}\vee t)} v(\tau)\, \hat{F}_{Z_m}(T_m, d\tau) \\ & + K(t) \sum_{m\in\mathbb{N}_0} \frac{D_{Z_m Z_{m+1}}(T_{m+1}, T_{m+1}-T_m)}{K(DT(T_{m+1}))}\, \mathbf{1}_{\{t<T_{m+1}<\infty\}}. \end{aligned} \qquad (3.12.2)$$

Thus, the present value of all payments $V_0 := B_{(p)} = B_{(p),0}$ can be represented as

$$V_0 = -v(t)\, V^-(t,A) + v(t)\, V^+(t,A). \qquad (3.12.3)$$

$V^-(t,A)$ can be regarded as accumulated value of past net incomes (premiums less benefits) for the insurer. Analogously, $V^+(t,A)$ is the present value of future net outgoings (benefits less premiums).

In contrast to the above definition, Norberg [1991] defined the quantity $V^-(t,A)$ by using the accumulated value at time t of all annuity payments up to and including time t, and all assurance payments up to and including time t. Accordingly, $V^+(t,A)$ takes the payments strictly after time t into account. Hence, his definition differs from our approach only by the assignment of the sojourn payments at time t.

Chapter 4

Prospective reserves and prospective losses

Consider a single policy (p) modelled by a pure jump process $(\Omega, \mathfrak{F}, P, (X_t)_{t\geq 0})$. $((X_t, U_t))_{t\geq 0}$ is the appertaining bivariate process. In the following sections, let K be a capital accumulation function, v the appertaining discounting function, and $A = SA + DA$ an actuarial payment function according to definition 3.9 that satisfies (3.10.1). $B_{(p),s} = C(s, ((X_t, U_t))_{t\geq s})$ is the present value at time $s \geq 0$ of future actuarial payments for $((X_t, U_t))_{t\geq 0}$. Further, let $(\mathfrak{F}_t)_{t\geq 0}$ be a filtration such that the bivariate process $((X_t, U_t))_{t\geq 0}$ is adapted to $(\mathfrak{F}_t)_{t\geq 0}$. Then, the prospective reserve is specified as follows.

4.1 Definition. The *prospective reserve of the actuarial payment function A for* $((X_t, U_t))_{t\geq 0}$ *at time* $s \geq 0$ *with respect to the filtration* $(\mathfrak{F}_t)_{t\geq 0}$ is defined as

$$V^+(s) := \mathbf{E}\left[B_{(p),s} \middle| \mathfrak{F}_s\right]. \tag{4.1.1}$$

Due to (3.10.1), the right-hand side of (4.1.1) is well defined. We generally assume in what follows that for $t \geq 0$, $\mathfrak{F}_t$ is the σ-field generated by $(X_s)_{s\leq t}$. Further, the pure jump process $(X_t)_{t\geq 0}$ is assumed to be semi-Markovian and hence, the bivariate process $((X_t, U_t))_{t\geq 0}$ is Markovian. Then, the following holds for the prospective reserve at time $s \geq 0$:

$$V^+(s) = \mathbf{E}\left[B_{(p),s} \middle| \mathfrak{F}_s\right] = \mathbf{E}\left[C(s, ((X_t, U_t))_{t\geq s}) \middle| X_s, U_s\right] \quad P-a.s. \tag{4.1.2}$$

We now consider the appertaining factorized conditional expectations.

4.2 Definition. The *prospective reserve of the actuarial payment function A for* $((X_t, U_t))_{t\geq 0}$ *at time* $s \geq 0$ *in state* $(y, u) \in \mathcal{S} \times [0, \infty)$ is the conditional expectation of the present value of future actuarial payments given $(X_s, U_s) = (y, u)$:

$$V^+_{(y,u)}(s) := \mathbf{E}\left[C(s, ((X_t, U_t))_{t\geq s}) \middle| X_s = y, U_s = u\right]. \tag{4.2.1}$$

For $s \geq 0$ and $(y, u) \in \mathcal{S} \times [0, \infty)$, $V^+_{(y,u)}(s)$ is regarded as prospective reserve at time s belonging to a policy which is at time s in state y with this state being entered at time $s - u$. In general, the prospective reserve is understood as amount of money covering, along with future premiums and interest gains, the future benefits. For almost all results concerning the prospective reserve, we will employ the following representation.

4. 3 Lemma. *Let $s \geq 0$, and assume (3. 10. 1). Then, the prospective reserve according to definition 4. 2 is for $\mathcal{L}(X_s, U_s)$-a.e. $(y, u) \in \mathcal{S} \times [0, \infty)$ given by*

$$V^+_{(y,u)}(s) = \int_{\mathcal{K}} \hat{C}(s, ((T_m, Z_m))_{m \geq 0})\, d\mathcal{L}(((T_m, Z_m))_{m \geq 0} |\, T_1 > s, T_0 = s - u, Z_0 = y). \qquad (4.3.1)$$

Proof. Consider the bivariate Markov process $((X_t, U_t))_{t \geq 0}$ with transition probabilities according to (2. 21. 6). Theorem 2. 23 provided by means of

$$p_{yz}(s, t, u, v) = P(\exists l \in \mathbb{N}_0 : T_{l+1} > t, T_l \in [t - v, t], Z_l = z |\, T_1 > s, T_0 = s - u, Z_0 = y)$$

for $0 \leq u \leq s < t, v \geq 0$, and $(y, z) \in \mathcal{S}^2$ a regular set of transition probabilities. Due to the regularity of these transition probabilities, they can be used to manufacture regular versions of the conditional distribution of $((X_t, U_t))_{t \geq s}$ given (X_s, U_s). Hence, the conditional distribution $\mathcal{L}(((X_t, U_t))_{t \geq s} |\, T_1 > s, T_0 = s - u, Z_0 = y)$ forms a regular version of the conditional distribution of $((X_t, U_t))_{t \geq s}$ given $(X_s, U_s) = (y, u)$. For $s \geq 0$, this version is uniquely determined up to $\mathcal{L}(X_s, U_s | P)$-exceptional sets. Using this regular conditional distribution, one obtains according to (4. 2. 1) and (2. 6. 1)

$$\begin{aligned} V^+_{(y,u)}(s) &= \mathbf{E}\left[C(s, ((X_t, U_t))_{t \geq s}) |\, X_s = y, U_s = u\right] \\ &= \int_{\mathcal{U}_{\mathcal{X}}} C(s, ((X_t, U_t))_{t \geq s})\, d\mathcal{L}(((X_t, U_t))_{t \geq s} |\, T_1 > s, T_0 = s - u, Z_0 = y) \end{aligned}$$

for $\mathcal{L}(X_s, U_s |\, P)$-a.e. (y, u). By using (3. 11. 2), and afterwards applying lemma 2. 7 to the measurable mapping $\bar{G} : \mathcal{U}_{\mathcal{X}} \to \mathcal{K}$ according to (2. 3. 6), we finally get

$$\begin{aligned} V^+_{(y,u)}(s) &= \int_{\mathcal{U}_{\mathcal{X}}} C(s, ((X_t, U_t))_{t \geq s})\, d\mathcal{L}(((X_t, U_t))_{t \geq s} |\, T_1 > s, T_0 = s - u, Z_0 = y) \\ &= \int_{\mathcal{U}_{\mathcal{X}}} \hat{C}(s, \bar{G}(((X_t, U_t))_{t \geq s}))\, d\mathcal{L}(((X_t, U_t))_{t \geq s} |\, T_1 > s, T_0 = s - u, Z_0 = y) \\ &= \int_{\mathcal{K}} \hat{C}(s, ((T_m, Z_m))_{m \geq 0})\, d\mathcal{L}(((T_m, Z_m))_{m \geq 0} |\, T_1 > s, T_0 = s - u, Z_0 = y). \end{aligned}$$

□

In the sequel, versions of the prospective reserve $V^+_{(y,u)}(s)$ will be derived by using (4. 3. 1). But before investigating the prospective reserve in more detail, another convention is introduced. Recall that according to definition 3. 9, the annuity payments accumulated in state $z \in \mathcal{S}$ are described by a kernel $\hat{F}_z = \hat{F}^+_z - \hat{F}^-_z$. Using these kernels, we define cumulative annuity payment rates as follows. Their definition is inspired by the cumulative transition intensities for the bivariate Markov process $((X_t, U_t))_{t \geq s}$ (cf. definition 2. 24). It is intended for standardizing the notation to simplify future calculations.

4. 4 Definition. For $z \in \mathcal{S}$ and $0 \leq u \leq s$ we define

$$F_z(s, \cdot, u) : [s, \infty) \ni t \mapsto \int_{(s,t]} \hat{F}_z(s - u, d\tau). \qquad (4.4.1)$$

F_z is referred to as *cumulative annuity payment rate for state y*. It corresponds to the accumulated amount of sojourn payments in state z payable in the interval $(s, t]$, with the state z being entered at time $s - u$.

With this convention, one obtains similar properties for both the cumulative annuity payment rates and the cumulative transition intensities. For example, we obtain as counterpart to (2. 25. 2)

$$F_z(s, t, u) = \hat{F}_z(s - u, (s, t]), \;\; 0 \leq u \leq s \leq t, z \in \mathcal{S}, \qquad (4.4.2)$$

and, according to (2. 27. 1),

$$F_z(s,r,u) + F_z(r,t,r-s+u) = F_z(s,t,u), \;\; s \le r \le t. \qquad (4.4.3)$$

Recall the proof of the forward integral equations for the transition probabilities of the process $((X_t,U_t))_{t\ge 0}$ (lemma 2. 36). There, an essential assumption for the interchange of the order of integrations according to Fubini's theorem was given by the cumulative transition intensities being dominated (assumption 2. 35). Similar to assumption 2. 35, we sometimes stipulate for the cumulative annuity payment rates the following.

4. 5 Assumption. The cumulative annuity payment rates $F_z, z\in\mathcal{S}$ according to (4. 4. 1) are assumed to be dominated by Borel measures $\mathbf{F}_z, z\in\mathcal{S}$, such that for $0\le u\le s$ and $z\in\mathcal{S}$

$$F_z(s,d\tau,u) \ll \mathbf{F}_z(d\tau) \quad \text{with} \quad F_z(s,d\tau,u) = f_z(\tau,\tau-s+u)\,\mathbf{F}_z(d\tau). \qquad (4.5.1)$$

Thus, the dominating measure $\mathbf{F}_z$ for the cumulative annuity payment rate for state z does not depend on the time elapsed since this state was entered.

Under assumption 4. 5, the present value at policy issue of all actuarial payments $V_0 = B_{(p),0}$ according to (3. 11. 1) can be represented as

$$V_0 = \sum_{z\in\mathcal{S}} \int_{[0,\infty)} \mathbf{1}_{\{X_\tau = z\}}\, v(\tau)\, f_z(\tau,U_\tau)\,\mathbf{F}_z(d\tau) + \sum_{(y,z)\in\mathcal{J}} \int_{(0,\infty)} v(DT(\tau))\, D_{yz}(\tau,U_{\tau-0})\, N_{yz,d\tau}.$$

In cases where $\mathbf{F}_y$ corresponds for each $y\in\mathcal{S}$ to λ^1, V_0 is equal to

$$\sum_{z\in\mathcal{S}} \int_{[0,\infty)} \mathbf{1}_{\{X_\tau = z\}}\, v(\tau)\, f_z(\tau,U_\tau)\,d\tau + \sum_{(y,z)\in\mathcal{J}} \int_{(0,\infty)} v(DT(\tau))\, D_{yz}(\tau,U_{\tau-0})\, N_{yz,d\tau}.$$

This is, for example, by setting $DT = Id$ and $v : [0,\infty) \ni \tau \mapsto r^{-\tau}$ with $r\ge 1$ the starting point of Møller's [1993] investigation.

4. 6 Theorem. [Definition formula of the prospective reserve] *Assume* (3. 10. 1)*, and let* q *be a regular cumulative transition intensity matrix for the bivariate Markov process* $((X_t,U_t))_{t\ge 0}$*. Then, the prospective reserve* $V^+_{(y,u)}(s)$*,* $s\ge 0$*, according to definition* 4. 2 *is for* $\mathcal{L}(X_s,U_s|\,P)$*-a.e.* $(y,u)\in\mathcal{S}\times[0,\infty)$ *given as*

$$\begin{aligned} V^+_{(y,u)}(s) \;=\; & K(s)\sum_{l=0}^{\infty}\sum_{\xi\in\mathcal{S}} \int_{[s-u,\infty)}\int_{[s,\infty)} v(\tau)\,(1-Q_\xi(r,\tau,0))\,F_\xi(r,d\tau,0) \\ & \cdot P(T_l\in dr, Z_l=\xi|\,T_1>s, T_0=s-u, Z_0=y) \\ & + K(s)\sum_{l=0}^{\infty}\sum_{(\xi,z)\in\mathcal{J}} \int_{[s-u,\infty)}\int_{(s,\infty)} \frac{D_{\xi z}(\tau,\tau-r)}{K(DT(\tau))}\,(1-Q_\xi(r,\tau-0,0))\,q_{\xi z}(r,d\tau,0) \\ & \cdot P(T_l\in dr, Z_l=\xi|\,T_1>s, T_0=s-u, Z_0=y). \end{aligned} \qquad (4.6.1)$$

Further, if the assumptions 2. 35 *and* 4. 5 *are fulfilled, one gets*

$$\begin{aligned} V^+_{(y,u)}(s) \;=\; & K(s)\sum_{\xi\in\mathcal{S}} \int_{[s,\infty)}\int_{[0,\infty)} v(\tau)\, f_\xi(\tau,l)\, p_{y\xi}(s,\tau,u,dl)\,\mathbf{F}_\xi(d\tau) \\ & + K(s)\sum_{(\xi,z)\in\mathcal{J}} \int_{(s,\infty)}\int_{(0,\infty)} \frac{D_{\xi z}(\tau,l)}{K(DT(\tau))}\,\lambda_{\xi z}(\tau,l)\, p_{y\xi}(s,\tau-0,u,dl)\,\mathbf{\Lambda}_{\xi z}(d\tau). \end{aligned} \qquad (4.6.2)$$

PROOF. The above assertions can be verified by deriving $V^+_{(y,u)}(s)$, $s \geq 0$, with the aid of (4. 3. 1). A detailed computation of (4. 6. 1) and (4. 6. 2) can be found in the appendix (see A.12). □

Formula (4. 6. 2) is a generalization of the corresponding results by Hoem [1972] in a smooth framework. There, the assumptions 2. 35 and 4. 5 are obviously satisfied. Further, these assumptions are also satisfied in a series of scenarios of practical interest. Recall the comments subsequent to assumption 2. 35 and lemma 2. 36. Hence, the main disadvantage of formula (4. 6. 2) is not rooted in assuming 2. 35 and 4. 5. The main disadvantage rather rests in the structure of the above formula which is caused by the duration-dependence of transition intensities and actuarial payments. Since integrations also have to be performed over the times elapsed since the current state was entered, the computation of prospective reserves by means of (4. 6. 2) is cumbersome. This is the same problem that was mentioned by Norberg [1992] regarding the computation of the variance of the prospective loss under non-Markov assumptions (recall our introduction, p.11). Further, the transition probabilities are usually not available and must be provided by solving either (2. 32. 2) or (2. 36. 1). In all, (4. 6. 2) does not provide an appropriate way to derive versions of the prospective reserve. Analogously to the Markov model by Milbrodt and Helbig [1999], the best way to derive versions of prospective reserves is given by solving Thiele's integral equations of type 1. We shall see in the next section that by proceeding in this way the structural problems due to the duration-dependence do not appear. Before this, another assumption will be introduced.

4. 7 Assumption.

$$\mathbf{A3}: \ \mathbf{E}\Big[\sum_{(y,z)\in\mathcal{J}}\sum_{m\in N_0}\int_{(0,\infty)} \mathbf{1}_{\{T_m<\tau\leq T_{m+1}\}}\, v(DT(\tau))\, D_{yz}(\tau,\tau-T_m)\, \hat{q}_{yz}(T_m,d\tau)\Big] < \infty \quad (4.7.1)$$

In cases where this assumption is additionally fulfilled, the following holds: $\mathbf{E}[|V_0|] < \infty$. Then, the prospective reserve according to definition 4. 2 is finite for $\mathcal{L}(X_s, U_s)$-a.e. $(y, u) \in \mathcal{S} \times [0, \infty)$.

A Thiele's integral equations of type 1

4. 8 Theorem. [Thiele's integral equations of type 1] *Assume* (3. 10. 1). *Then, the prospective reserve* $V^+_{(y,u)}(s)$, $s \geq 0$, *according to definition* 4. 2 *satisfies for* $\mathcal{L}(X_s, U_s|\, P)$*-a.e.* $(y, u) \in \mathcal{S} \times [0, \infty)$ *the following system of integral equations*

$$\begin{aligned} V^+_{(y,u)}(s) &= K(s)\int_{[s,\infty)} v(\tau)\,\bar{p}_y(s,\tau,u)\,F_y(s,d\tau,u) \\ &\quad + K(s)\sum_{\substack{z\in\mathcal{S}\\ z\neq y}}\int_{(s,\infty)}\Big(\frac{D_{yz}(\tau,\tau-s+u)}{K(DT(\tau))} + v(\tau)\,V^+_{(z,0)}(\tau)\Big)\,Q_{yz}(s,d\tau,u), \end{aligned} \quad (4.8.1)$$

which is referred to as Thiele's integral equations of type 1. If regular cumulative transition intensities q_* *exist, one obtains*

$$\begin{aligned} V^+_{(y,u)}(s) &= K(s)\int_{[s,\infty)} v(\tau)\,\bar{p}_y(s,\tau,u)\,F_y(s,d\tau,u) \\ &\quad + K(s)\sum_{\substack{z\in\mathcal{S}\\ z\neq y}}\int_{(s,\infty)}\Big(\frac{D_{yz}(\tau,\tau-s+u)}{K(DT(\tau))} + v(\tau)\,V^+_{(z,0)}(\tau)\Big)\,\bar{p}_y(s,\tau-0,u)\,q_{yz}(s,d\tau,u). \end{aligned} \quad (4.8.2)$$

Proof. It is sufficient to verify (4. 8. 1), since (4. 8. 2) follows immediately from (4. 8. 1) by employing (2. 28. 4). In order to prove (4. 8. 1), we likewise start for $s \geq 0$ with (4. 3. 1). In contrast to the proof of (4. 6. 1), however, we split up the integrand according to

$$\begin{aligned}
&\hat{C}(s, ((T_m, Z_m))_{m\in\mathbb{N}_0}) \\
&= K(s) \int_{[T_0\vee s, T_1\vee s)} v(\tau)\, \hat{F}_{Z_0}(T_0, d\tau) + K(s)\, \frac{D_{Z_0 Z_1}(T_1, T_1 - T_0)}{K(DT(T_1))}\, \mathbf{1}_{\{s<T_1<\infty\}} \\
&\quad + K(s) \sum_{m=1}^{\infty} \int_{[T_m\vee s, T_{m+1}\vee s)} v(\tau)\, \hat{F}_{Z_m}(T_m, d\tau) \\
&\quad + K(s) \sum_{m=1}^{\infty} \frac{D_{Z_m Z_{m+1}}(T_{m+1}, T_{m+1} - T_m)}{K(DT(T_{m+1}))}\, \mathbf{1}_{\{s<T_{m+1}<\infty\}}. \qquad (4.8.3)
\end{aligned}$$

Computing the corresponding integrals and applying (4. 3. 1) again, yields the assertion. For a detailed calculation, we refer to A.13 in the appendix. □

Compared with formula (4. 6. 2), Thiele's integral equations of type 1 (according to either (4. 8. 1) or (4. 8. 2)) have some advantages. These advantages are: (1) the transition probabilities of the process $((X_t, U_t))_{t\geq 0}$ do not have to be provided, (2) there are no additional integrals due to the duration-dependence of transition intensities and actuarial payments, (3) the assumptions 2. 35 and 4. 5 can be avoided, and (4) the second integral of the right-hand side always contains only prospective reserves for $u = 0$. The latter result in the following procedure to solve Thiele's integral equation of type 1. It consists of two steps. For the first step, the duration-dependence of the prospective reserve can be disregarded by setting $u = 0$. Starting from (4. 8. 2), this yields the following system of linear integral equations for $s \geq 0$ and $y \in \mathcal{S}$:

$$\begin{aligned}
V^+_{(y,0)}(s) \;=\;& K(s) \int_{[s,\infty)} v(\tau)\, \bar{p}_y(s, \tau, 0)\, F_y(s, d\tau, 0) \qquad (4.8.4) \\
&+ K(s) \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} \Big(\frac{D_{yz}(\tau, \tau - s)}{K(DT(\tau))} + v(\tau)\, V^+_{(z,0)}(\tau) \Big)\, \bar{p}_y(s, \tau - 0, 0)\, q_{yz}(s, d\tau, 0).
\end{aligned}$$

A solution of this system is given by a vector of prospective reserves $(V^+_{(y,0)}(\cdot))_{y\in\mathcal{S}}$. Due to the duration-dependence being disregarded, (4. 8. 4) can be solved in almost the same manner as the corresponding system in a non-smooth Markov set-up (cf. Milbrodt and Helbig [1999], Satz 10.13). In section E of this chapter, it will be established that the system of integral equations (4. 8. 4) is likewise uniquely solvable. For this, the corresponding results by Milbrodt and Helbig ([1999], Satz 10.24) can be adapted to the system (4. 8. 4).

In a second step, for each state $y \in \mathcal{S}$ and each previous duration $u \leq s$, the prospective reserve $V^+_{(y,u)}(s)$ can be derived by computing an integral according to right-hand side of (4. 8. 2). This integral contains the solutions $V^+_{(z,0)}(\cdot)$, $z \in \mathcal{S}, z \neq y$, which are provided by the first step.

The following section is concerned with computing prospective reserves with the aid of Thiele's integral equations of type 1. We will consider two previously mentioned models, namely the PKV model and the PHI model. But before we turn to this, the relationships between Thiele's integral equations of type 1 and the backward integral equations (cf. (2. 32. 1) and (2. 32. 2)) for the transition probabilities of $((X_t, U_t))_{t\geq 0}$ are investigated. Afterwards, a system of differential equations is provided which appertains to (4. 8. 2).

4. 9 Lemma. *Thiele's integral equations of type 1 according to* (4. 8. 1) *are generalizations of the backward integral equations* (2. 32. 1). *(If regular cumulative transition intensities exist, this* *relation also holds for* (4. 8. 2) *and* (2. 32. 2)*.) Conversely, if regular cumulative transition intensities exist and if further the assumptions* 2. 35 *and* 4. 5 *are satisfied, the backward integral equations* (2. 32. 1) *imply Thiele's integral equations* (4. 8. 1)*.*

Proof. Consider the following specified parameters

$$D_{\xi z} \equiv 0,\ (\xi, z) \in \mathcal{J}, \quad v \equiv 1,$$

and

$$F_\xi(s, d\tau, u) = f_\xi(\tau, \tau - s + u)\, \mathbf{F}_\xi(d\tau) := \mathbf{1}_{(v \geq \tau - s + u)}\, \delta_{\xi z}\ \varepsilon_t(d\tau),\ \ \xi \in \mathcal{S},\ 0 \leq u \leq s, v \geq 0.$$

Inserting these parameters into (4. 6. 1) yields for $\mathcal{L}(X_s, U_s |\, P)$-a.e. $(y, u) \in \mathcal{S} \times [0, \infty)$

$$V^+_{(y,u)}(s) \ = \ \sum_{l=0}^{\infty} P(T_{l+1} > t, T_l \in [t - v, t], Z_l = z |\, T_1 > s, T_0 = s - u, Z_0 = y)\, \mathbf{1}_{(t \geq s)}.$$

Thus, we obtain by means of (2. 23. 6) for $\mathcal{L}(X_s, U_s |\, P)$-a.e. $(y, u) \in \mathcal{S} \times [0, \infty)$

$$V^+_{(y,u)}(s) \ = \ p_{yz}(s, t, u, v)\, \mathbf{1}_{(t \geq s)}. \tag{4. 9. 1}$$

Further, we get by inserting the above parameters as well as (4. 9. 1) into the right-hand side of (4. 8. 1)

$$\begin{aligned} & K(s) \int_{(s,\infty)} v(\tau)\, \bar{p}_y(s, \tau, u)\, F_y(s, d\tau, u) \\ & \quad + K(s) \sum_{\substack{\xi \in \mathcal{S} \\ \xi \neq y}} \int_{(s,\infty)} \Big(\frac{D_{y\xi}(\tau, \tau - s + u)}{K(DT(\tau))} + v(\tau)\, V^+_{(\xi,0)}(\tau) \Big)\, Q_{y\xi}(s, d\tau, u), \\ & = \int_{(s,\infty)} \bar{p}_y(s, \tau, u)\, \mathbf{1}_{(v \geq \tau - s + u)}\, \delta_{yz}\, \varepsilon_t(d\tau) + \sum_{\substack{\xi \in \mathcal{S} \\ \xi \neq y}} \int_{(s,\infty)} p_{\xi z}(\tau, t, 0, v)\, \mathbf{1}_{(t \geq \tau)}\, Q_{y\xi}(s, d\tau, u), \\ & = \mathbf{1}_{(t \geq s)}\, (1 - Q_y(s, t, u))\, \mathbf{1}_{(v \geq t - s + u)}\, \delta_{yz} + \sum_{\substack{\xi \in \mathcal{S} \\ \xi \neq y}} \int_{(s,t]} p_{\xi z}(\tau, t, 0, v)\, Q_{y\xi}(s, d\tau, u). \end{aligned} \tag{4. 9. 2}$$

Finally, by inserting (4. 9. 1) into the left-hand side of (4. 8. 1), we have verified that Thiele's integral equations of type 1 can for $s \leq t$ be reduced to

$$p_{yz}(s, t, u, v) = \delta_{yz}\, \mathbf{1}_{(v \geq t - s + u)}\, (1 - Q_y(s, t, u)) + \sum_{\substack{\xi \in \mathcal{S} \\ \xi \neq y}} \int_{(s,t]} p_{\xi z}(\tau, t, 0, v)\, Q_{y\xi}(s, d\tau, u),$$

which corresponds to (2. 32. 1).
Conversely, that the backward integral equations (2. 32. 1) imply Thiele's integral equations (4. 8. 1) can be verified by inserting (2. 32. 1) into (4. 6. 2). This yields (4. 8. 1). Yet in order to employ (4. 6. 2), regular cumulative transition intensities must exist and the assumptions 2. 35 and 4. 5 must be fulfilled. For a detailed proof of this relationship, we refer to the appendix (see A.14). □

4. 10 Definition. For $(y, z) \in \mathcal{J}$ and $0 \leq u \leq t$ we define

$$R_{yz}(t, u) := K(t)\, v(DT(t))\, D_{yz}(t, u) + V^+_{(z,0)}(t) - V^+_{(y,u)}(t). \tag{4. 10. 1}$$

R_{yz} is referred to as the *amount of risk associated with transition from state y to z.*

The amount of risk is a common quantity in life insurance (cf. Ramlau-Hansen [1988], Norberg [1992], Milbrodt and Helbig [1999] (10.19.1), etc.). Usually, however, it only depends on the corresponding states and the time of transition. Analogously to Møller [1993], the amount of risk R_{yz} is here also allowed to depend on the time elapsed since the state y was entered.

4. 11 Theorem. [Thiele's differential equations] *Let $(q_{yz})_{(y,z)\in\mathcal{S}^2}$ be regular cumulative transition intensities with densities $(\mu_{yz})_{(y,z)\in\mathcal{S}^2}$ according to* (2. 36. 6) *and* (2. 36. 7)*, $(F_y)_{y\in\mathcal{S}}$ cumulative annuity payment rates according to* (4. 5. 1) *with Lebesgue densities $(f_y)_{y\in\mathcal{S}}$, and K a capital accumulation function with interest intensity ϕ. Further, assume that all densities are continuous with $(\mu_{yz})_{(y,z)\in\mathcal{S}^2}$ and $(f_y)_{y\in\mathcal{S}}$ being additionally continuously differentiable with respect to the second variable. Then, under assumptions* (3. 10. 1) *and* (4. 7. 1)*, the prospective reserve at time $s \geq 0$ satisfies for $\mathcal{L}(X_s, U_s|\, P)$-a.e. $(y,u) \in \mathcal{S}\times[0,\infty)$ the following system of differential equations*

$$\frac{\partial V^+_{(y,u)}(s)}{\partial s} = -\frac{\partial V^+_{(y,u)}(s)}{\partial u} - f_y(s,u) + \phi(s)\, V^+_{(y,u)}(s) - \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} R_{yz}(t,u)\,\mu_{yz}(s,u). \qquad (4.\,11.\,1)$$

Proof. Starting from Thiele's integral equations (4. 8. 2), the assertion can be verified in a manner similar to the proof of (2. 37. 1) (see A.7). □

4. 12 Corollary. *Thiele's differential equations according to* (4. 11. 1) *are equivalent to the backward differential equations* (2. 37. 1)*.*

Proof. In the situation of theorem 4. 11, the assumptions 2. 35 and 4. 5 are fulfilled. Hence, lemma 4. 9 states that Thiele's integral equations of type 1 are equivalent to the backward integral equations of type 1. Differentiating with respect to s yields the assertion. □

B Premiums and reserves for PKV and PHI

The aim of this section is to exemplify the application of Thiele's integral equations (4. 8. 1) regarding a model for German private health insurance (PKV) and a model for permanent health insurance (PHI). Both models have been presented in the introduction. First, the PKV model is considered. In a discrete set-up, we derive premiums and reserves based on a decrement model for which the withdrawal rates are additionally allowed to depend on the previous contract duration. The following issues regarding this model form a further development of example 2. 38. For more details on precisely this topic, Helwich and Milbrodt [2007] can be consulted.

Afterwards, a non-discrete approach for PHI modelling is investigated. For this, duration-depending recovery and mortality rates for disabled insured are implemented. These age- and duration-depending recovery and mortality rates are originated from the select-and-ultimate tables DAV-SRT 1997 RI M and DAV-SST TI 1997 M.

4. 13 Example. [example 2. 38 continued] Consider the situation of example 2. 38. For the age range $AB = \{x_{MIN}, ..., x_{MAX}\}$, we assume that select-and-ultimate tables

$$(q_{x,u})_{x\in AB, u\in\{0,...,x-x_{MIN}\}} \quad \text{and} \quad (w_{x,u})_{x\in AB, u\in\{0,...,x-x_{MIN}\}}$$

are given for both annual mortality rates and annual withdrawal rates. According to (2. 38. 2), these rates can be used to derive conditional probabilities of remaining k-years in the portfolio:

$${}_kp_{x,u}, \quad x\in AB, u\in\{0,...,x-x_{MIN}\}, k\in\{0,...,x_{MAX}-x\}.$$

As previously mentioned, the PKV model forms a special case of a multiple state model, which is reduced to the recording of future sojourn time and cause of decrement. Benefits incurred by decrement are not intended. Hence, only sojourn payments for the state *active* are implemented. The annual benefits are specified by the claims amounts per risk (Kopfschäden) which correspond to the expected annual costs due to health services (e.g. medical services, drugs, remedies, medical devices, etc.). These benefits are assumed to depend only on the attained age of an insured and they are denoted by $(K_x)_{x\in AB}$. It should be mentioned here that another interesting issue can be covered by the present approach. Namely, beside allowing annual decrement rates to additionally depend on the previous contract duration, the claims amounts per risk can also be modelled in a similar manner, since durational effects actually also appear on this level. More precisely, for certain PKV products, the claims amounts per risk for new entrants are often smaller than the claims amounts per risk for insured of the same age, yet with a longer previous sojourn time in the portfolio.

We go on by assuming that the benefits as well as the annual premiums are payable in advance. Finally, let $r := 1 + i = 1.035$ be the accumulation factor which is currently used in Germany, and $v := 1/r$ the corresponding discounting factor. Then, the actuarial present value of all future benefits, for an insured with attained age x and previous sojourn time in the portfolio u, is given by

$$A_{x,u} := \sum_{\nu=0}^{x_{MAX}-x} v^{\nu}\, K_{x+\nu}\ {}_{\nu}p_{x,u}. \tag{4.13.1}$$

According to this, the actuarial present value of an annual payment of a currency unit, as long as the insured remains in the portfolio, is given by

$$\ddot{a}_{x,u} := \sum_{\nu=0}^{x_{MAX}-x} v^{\nu}\ {}_{\nu}p_{x,u}. \tag{4.13.2}$$

The previous sojourn time in the portfolio u always refers to the attained age x. It corresponds to the difference $x-x_0$. Hence, if x coincides with the age at policy issue x_0, the previous sojourn time in the portfolio is given by $u = 0$.

Due to the duration-dependence of the above actuarial present values, the annual net premiums also depend on both age and previous sojourn time in the portfolio. For if the principle of equivalence is invoked, the annual net premiums are specified by

$$P_{x,u} := \frac{A_{x,u}}{\ddot{a}_{x,u}},\ \ x \in AB,\ u \in \{0, ..., x - x_{MIN}\}. \tag{4.13.3}$$

The same effect appears for the prospective reserves (aging provisions). They also depend on both age and previous sojourn time in the portfolio. Consider an insured with age x and previous sojourn time u. Then, the prospective reserve at time $m \in \mathbb{N}_0$, satisfying $x + m \leq x_{MAX}$, is given as the difference between the actuarial present value of all benefits and of all premiums at and after time m:

$$\begin{aligned} {}_{m}V_{x,u} &:= \sum_{\nu=0}^{x_{MAX}-x-m} v^{\nu}\, K_{x+m+\nu}\ {}_{\nu}p_{x+m,u+m} - \sum_{\nu=0}^{x_{MAX}-x-m} v^{\nu}\, P_{x,u}\ {}_{\nu}p_{x+m,u+m} \\ &= \sum_{\nu=0}^{x_{MAX}-x-m} v^{\nu} \Big(K_{x+m+\nu} - P_{x,u}\Big)\ {}_{\nu}p_{x+m,u+m}. \end{aligned} \tag{4.13.4}$$

The previous sojourn time in the portfolio u is likewise related to the age x. The attained age at time m is accordingly given as $x + m$. In the sequel, we consider the prospective reserve

${}_{m}V_{x_0,0}$, $m \in \mathbb{N}$. This reserve corresponds to the aging provision at time m which appertains to an insured with age at issue x_0.

For the prospective reserve ${}_{m}V_{x_0,0}$, $m \in \mathbb{N}$, formula (4.13.4) is a special case of Thiele's integral equations (4.8.1) (or (4.8.2)), namely by means of

$$ {}_{m}V_{x_0,0} = V^{+}_{(a,m)}(m). \tag{4.13.5} $$

The right-hand side of (4.13.5) corresponds to the prospective reserve according to definition 4.2 for a discrete process $(X_t, U_t) = (X_{\lfloor t \rfloor}, U_{\lfloor t \rfloor}), t \geq 0$, modelling a policy of an insured with age at issue x_0 allowing the policy states $\mathcal{S} = \{a, w, d\}$. To obtain (4.13.4), the following settings must be inserted into (4.8.1):

$$ D_{\xi z} \equiv 0,\ (\xi, z) \in \mathcal{J}, \quad V^{+}_{(z,0)} \equiv 0,\ z \neq a, \quad v(k) := r^{-k},\ k \in \mathbb{N}_0, \quad \bar{p}_a(m, m+\nu, m) := {}_{\nu}p_{x_0+m,m}, $$

and

$$ F^{+}_{a}(m, d\tau, m) := \sum_{k=m}^{x_{MAX}-x} K_{x_0+k}\, \varepsilon_k(d\tau) \quad \text{as well as} \quad F^{-}_{a}(m, d\tau, m) = \sum_{k=m}^{x_{MAX}-x} P_{x_0,0}\, \varepsilon_k(d\tau). $$

The annual increment of the prospective reserve ${}_{m}V_{x_0,0}$ is given by

$$ \begin{aligned} &{}_{m+1}V_{x_0,0} - {}_{m}V_{x_0,0} \\ &= (P_{x_0,0} - K_{x_0+m})(1+i) + i\ {}_{m}V_{x_0,0} + (q_{x_0+m,m} + w_{x_0+m,m})\ {}_{m+1}V_{x_0,0} \end{aligned} \tag{4.13.6} $$

(cf. Milbrodt [2005], formula (7.6.3)). This means that the increment of the reserve can be divided into (1) the savings part of the net premiums, (2) the interest gain of the preexisting reserve, and (3) the part of the increment coming from reserves that are left by insured who drop out of the portfolio due to death or withdrawal. According to the arguments in the introduction, the last addend in (4.13.6) - or more precisely, the expression $w_{x_0+m,m}\ {}_{m+1}V_{x_0,0}$ which refers to the portion of the reserve that is left by withdrawing insured - is exactly the quantity that we are interested in. We previously argued that the mere age-depending withdrawal rates often overestimate the actual withdrawal rates which depend on both age and previous contract duration. Thus, the portion of the overall portfolio reserve that is left when insured withdraw will also be overestimated and the resulting net premiums might be too low.

In the following we will compare net premiums and reserves for two different approaches, with the actuarial basis for both being originated from the same portfolio. The first is based on an age- and duration-depending decrement model and the second on a mere age-depending decrement model. Regarding the notation, we use the same symbols for both models. In cases where quantities do not depend on the previous sojourn time in the portfolio, the indication of this time is simply omitted. Concerning the annual mortality rates and the annual withdrawal rates, we refer to example 2.38. For the age range $AB := \{21, ..., \omega = 100\}$, example 2.38 provides tables $(q_{x,u})_{x \in AB, u \in \{0,...,15\}}$ and $(w_{x,u})_{x \in AB, u \in \{0,...,15\}}$ as well as $(q_x)_{x \in AB}$ and $(w_x)_{x \in AB}$. The underlying (independent) mortality rates were taken from the PKV mortality table 2007 and the (independent) withdrawal rates originated from a real existing PKV portfolio.

We consider a PKV tariff for adults with the above age range AB. The benefits by means of the claims amounts per risk $(K_x)_{x \in AB}$ for a so-called Kompakttarif - including: out-patient treatment services (almost 100% absorption of costs), in-patient treatment services (two-bed room), dental treatment (100% absorption of costs), and dental prostheses (50-65% absorbtion of cost) - are taken, along with the profiles and the standardized claims amounts per risk (male 1761 Euros and female 3325 Euros), from the VerBaFin 12/2006 (cf. http://www.bafin.de). The corresponding net premiums are obtained by employing the principle of equivalence and (4.13.3). The prospective reserves are derived by means of (4.13.4).

Figure 16 compares for both approaches annual net premiums for new entrants in the above described Kompakttarif, in case of male insured (left) and female insured (right) with ages at issue $x_0 = y_0 = 25, ..., 50$. It can be observed that approximately up to age 35 the net premiums in the model with mere age-depending withdrawal rates are actually too low. The differences can sometimes add up to several hundred Euros. Note that ages up to the mid-thirties are the most likely ages of new entrants for German private health insurance, at least for the portfolio under consideration. For ages beyond the mid-thirties, however, the annual net premiums in the model with mere age-depending withdrawal rates are higher than the premiums relying on an age- and duration-depending decrement model. The reason for is given by the effect that can be observed by comparing figure 13 and figure 1. Namely, the higher the age of the insured, the longer the previous contract duration for which the mere age-depending withdrawal rates underestimate the actual withdrawal rates. However, that the annual net premiums in the model with mere age-depending withdrawal rates prevail over the premiums relying on an age- and duration-depending decrement model mainly appears when considering the annual net premiums for new entrants.

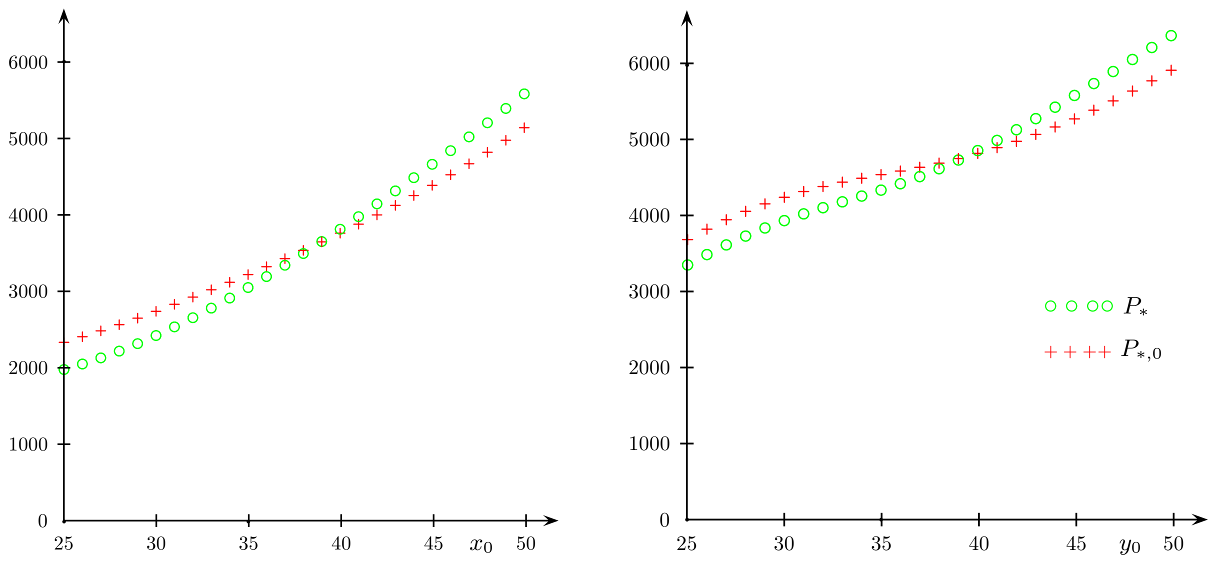

Figure 16: Comparison of annual net premiums relying on an age- and duration-depending decrement model with the corresponding premiums relying on a mere age-depending decrement model with respect to the age at issue $x_0 = y_0 = 25, ..., 50$, for male (left) and female (right) new entrants in the previously described Kompakttarif

Figure 17 illustrates annual net premiums not only for new entrants. It outlines for certain attained ages the duration-dependence of premiums relying on an age- and duration-depending decrement model. Obviously, even for ages for which the age-and duration-depending net premiums for new entrants fall below the mere age-depending net premiums, they finally prevail over the mere age-depending premiums, since they increase with the previous duration. Thus, even for higher ages, losses due to premiums that do not cover the actual risk are possible.

Recall the dilemma mentioned in the introduction, namely that, on the one hand, it is actually a matter of actuarial necessity to model the decrement of a PKV portfolio age- and duration-depending, but that, on the other hand, the use of such a model is forbidden by several regulations. An appropriate solution of this dilemma consists in adapting a mere age-depending model in a manner such that the durational effects are taken into account in the best possible way. For example, Rudolph [2005] suggested the recalculation of mere age-depending withdrawal rates by using the portion of the overall portfolio reserve which has been left by withdrawing insured, instead of using relative frequencies of withdrawal. This method has some disadvantages, with the major drawback being that the reserves must be derived by using the incorrect mere age-depending withdrawal rates. In contrast, we suggest to adapt a mere age-depending model by using the results from a calculation based on age- and duration-depending decrement rates. Further, we focus on premiums and choose a mere age-depending model such that the appertaining premiums are the best approximation of the actual age- and duration-depending premiums (cf. figure 17). But before this is discussed in more detail, we will illustrate the prospective reserves for both of the above approaches. Afterwards the expressions $(q_{x_0+m} + w_{x_0+m})\ {}_{m+1}V_{x_0}$ and $(q_{x_0+m,m} + w_{x_0+m,m})\ {}_{m+1}V_{x_0,0}$ are compared. These quantities correspond to the portion of the annual increments of the prospective reserves due to reserves which are left by insured dropping out of the portfolio. The quantity $(q_{x_0+m} + w_{x_0+m})\ {}_{m+1}V_{x_0}$ appertains to a decrement model with mere age-depending annual rates and $(q_{x_0+m,m} + w_{x_0+m,m})\ {}_{m+1}V_{x_0,0}$ to a model with age- and duration-depending annual rates.

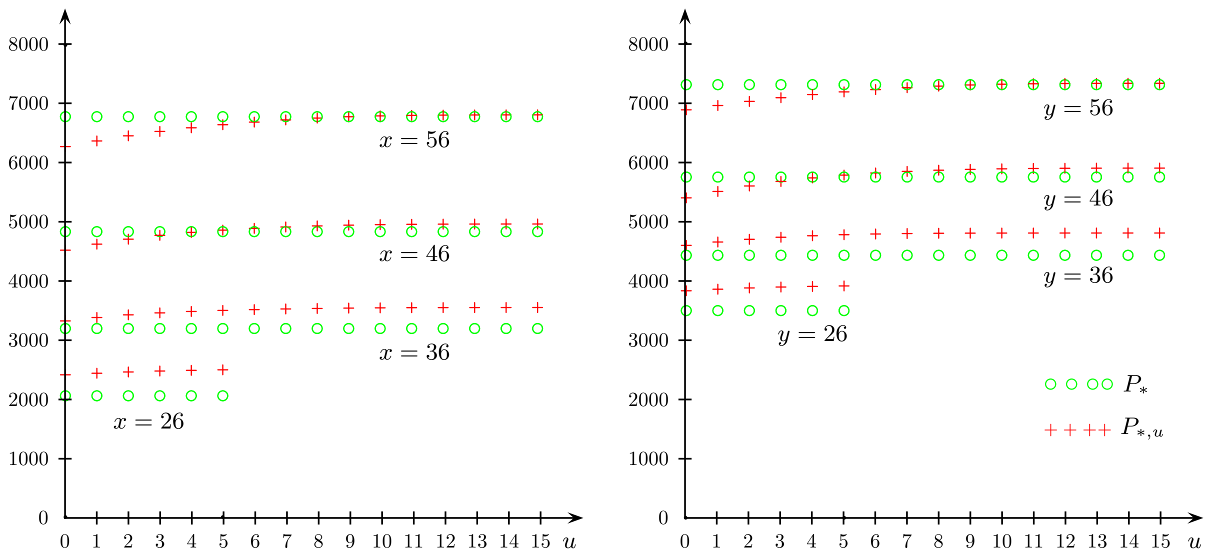

Figure 17: Comparison of mere age-depending annual net premiums with age- and duration-depending annual net premiums with respect to the previous contract duration u, for male (left) and female (right) insured at the attained ages $x = y = 26, 36, 46, 56$

Figure 18 illustrates the prospective reserves ${}_mV_{x_0,0}$ and ${}_mV_{x_0}$ with respect to the attained age x - or, which is equivalent, with respect to $m = x - x_0$ - for insured with age at issue $x_0 = 30$ (for female insured, y and y_0 are used). Realize that, for small m, the age- and duration-depending reserves prevail, since the corresponding premiums are higher. In cases of m being beyond a certain time, however, the mere age-depending prospective reserves prevail, even though the mere age-depending premiums are lower. According to (4.13.6), the annual increments of the prospective reserve split up into three parts. In view of lower premiums and equal interest, the increments of the mere age-depending reserve only prevail due to the third addend in (4.13.6). This addend is for both reserves illustrated in figure 19. Regarding higher ages, it is obvious that the cause of decrement *death* is less important for women than for men. Further, there is no significant difference between the values relying on an age- and duration-depending decrement model and those relying on a mere age-depending model. In contrast, for younger ages, the considered portion of the annual increments of the prospective reserves differ significantly. Due to the overestimation of the actual withdrawal rates by the mere age-depending withdrawal rates, the portion of the overall reserve that is left by withdrawing insured is also overestimated. This is the main reason for the difference between the annual net premiums for both approaches applied to the same portfolio.

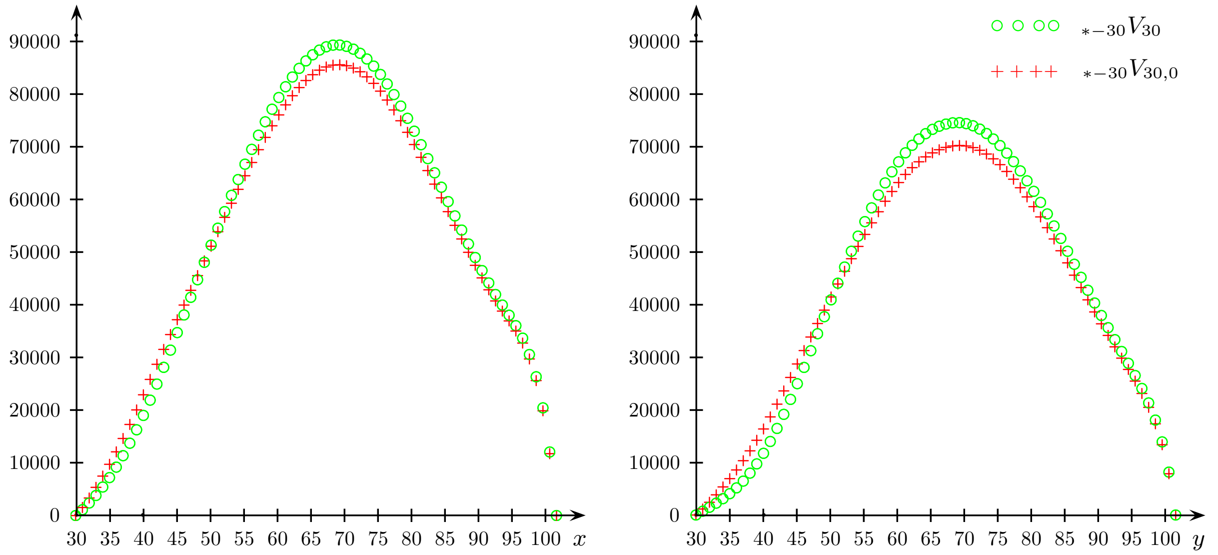

Figure 18: Comparison of mere age-depending prospective reserves with age- and duration-depending prospective reserves with respect to the attained age x (y), for male (left) and female (right) insured with age at issue $x_0 = y_0 = 30$

We now turn to the adaption of a model with mere age-depending annual decrement rates. Our starting point is a portfolio of insured structured by age and previous sojourn time in the portfolio. Then, according to example 2.38 (see also Helwich and Milbrodt [2007]), age- and duration-depending annual decrement rates can be specified. By using these decrement rates, age-and duration-depending premiums $P_{x,u}$ can be derived for each age $x \in AB$ and each previous duration in the portfolio $u \in \{0, ..., x - x_{MIN}\}$. Now let P_x^{opt}, $x \in AB$, denote mere age-depending premiums - the notation *opt* refers to the word *optimal* - such that for each $x \in AB$ the squared error (SE) is minimized:

$$SE_x : \mathbb{R} \ni P_x^{opt} \mapsto \sum_{u \in \{0,\dots,x-x_{MIN}\}} (P_x^{opt} - P_{x,u})^2. \tag{4.13.7}$$

It is a matter of common sense that the method of least squares results in the average value. Hence, for given premiums $(P_{x,u})_{u\in\{0,\dots,x-x_{MIN}\}}$, the minimum of (4.13.7) is for each $x \in AB$ attained by

$$P_x^{opt} = \frac{\sum_{u\in\{0,\dots,x-x_{MIN}\}} P_{x,u}}{(x - x_{MIN} + 1)}. \tag{4.13.8}$$

Thus, the optimal premium P_x^{opt} can simply be obtained by averaging the age- and duration-depending premiums $P_{x,u}$ with respect to all possible previous sojourn times u. Given these optimal premiums $(P_x^{opt})_{x\in AB}$, the corresponding annual withdrawal rates $w_x^{opt}, x \in AB$ can be derived recursively, provided that the mortality rates are known. Doing so, we set $q_x^{opt} + w_x^{opt} = 1$ for $x = x_{MAX}$. For $x \in AB \setminus \{x_{MAX}\}$, we get by employing the mere age-depending versions of (4.13.1), (4.13.2), and (4.13.3), along with (2.38.2),

$$\begin{aligned}
P_x^{opt} &= \frac{\sum_{\nu=0}^{x_{MAX}-x} v^\nu \, K_{x+\nu} \prod_{j=0}^{\nu-1}(1 - q_{x+j}^{opt} - w_{x+j}^{opt})}{\sum_{\nu=0}^{x_{MAX}-x} v^\nu \prod_{j=0}^{\nu-1}(1 - q_{x+j}^{opt} - w_{x+j}^{opt})} \\
&= \frac{K_x + \sum_{\nu=1}^{x_{MAX}-x} v^\nu \, K_{x+\nu} \prod_{j=0}^{\nu-1}(1 - q_{x+j}^{opt} - w_{x+j}^{opt})}{1 + \sum_{\nu=1}^{x_{MAX}-x} v^\nu \prod_{j=0}^{\nu-1}(1 - q_{x+j}^{opt} - w_{x+j}^{opt})} \\
&= \frac{K_x + (1 - q_x^{opt} - w_x^{opt}) \sum_{\nu=1}^{x_{MAX}-x} v^\nu \, K_{x+\nu} \prod_{j=1}^{\nu-1}(1 - q_{x+j}^{opt} - w_{x+j}^{opt})}{1 + (1 - q_x^{opt} - w_x^{opt}) \sum_{\nu=1}^{x_{MAX}-x} v^\nu \prod_{j=1}^{\nu-1}(1 - q_{x+j}^{opt} - w_{x+j}^{opt})}.
\end{aligned}$$

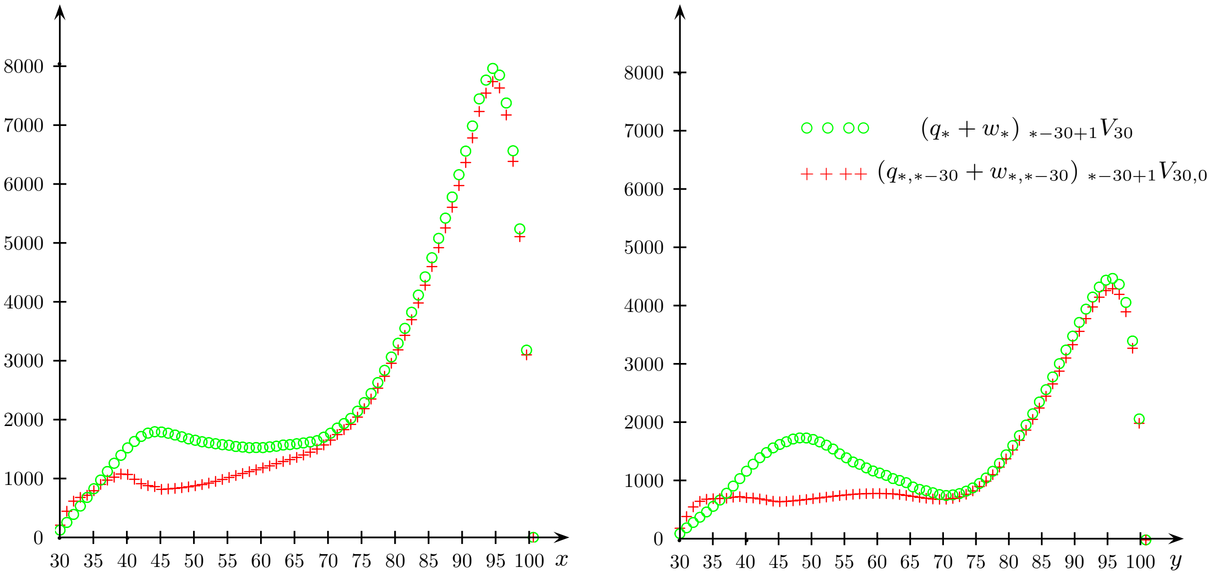

Figure 19: Comparison of the corresponding third addends of (4.13.6) for both prospective reserves illustrated in figure 18

This yields

$$\begin{aligned} &P_x^{opt} + (1 - q_x^{opt} - w_x^{opt}) \sum_{\nu=1}^{x_{MAX}-x} v^\nu\, P_x^{opt} \prod_{j=1}^{\nu-1}(1 - q_{x+j}^{opt} - w_{x+j}^{opt}) \\ &= K_x + (1 - q_x^{opt} - w_x^{opt}) \sum_{\nu=1}^{x_{MAX}-x} v^\nu\, K_{x+\nu} \prod_{j=1}^{\nu-1}(1 - q_{x+j}^{opt} - w_{x+j}^{opt}), \end{aligned}$$

from which the following can be obtained:

$$(1 - q_x^{opt} - w_x^{opt}) \quad = \quad \frac{K_x - P_x^{opt}}{\sum_{\nu=1}^{x_{MAX}-x} v^\nu\, (P_x^{opt} - K_{x+\nu}) \prod_{j=1}^{\nu-1}(1 - q_{x+j}^{opt} - w_{x+j}^{opt})}.$$

Finally, the sum of both annual mortality rates and annual withdrawal rates is for $x \in AB \setminus \{x_{MAX}\}$ given by

$$q_x^{opt} + w_x^{opt} \quad = \quad 1 - \frac{K_x - P_x^{opt}}{\sum_{\nu=1}^{x_{MAX}-x} v^\nu\, (P_x^{opt} - K_{x+\nu}) \prod_{j=1}^{\nu-1}(1 - q_{x+j}^{opt} - w_{x+j}^{opt})}. \qquad (4.13.9)$$

Employing (4.13.9), the sums $q_x^{opt} + w_x^{opt}$ can be derived recursively, starting with $x = x_{MAX} - 1$.

For given annual mortality rates, both independent and dependent annual withdrawal rates can be determined according to the following procedure. Recall that the differences between both types of withdrawal rates are fairly small, and hence, the corresponding conversion is often omitted. In the latter case, the annual withdrawal rates can be determined by means of the given annual mortality rates and the sums $q_x^{opt} + w_x^{opt}, x \in AB \setminus \{x_{MAX}\}$ according to (4.13.9). Recall that for $x = x_{MAX}$, the corresponding sum equals 1. When the withdrawal rates shall be determined exactly, the formulas (2.38.7) and (2.38.8) must be employed in the following way. By means of (2.38.7) and (2.38.8), it holds for the left-hand side of (4.13.9)

$$q_x^{opt} + w_x^{opt} \quad = \quad {}^{\mathbf{u}}q_x \left(1 - \frac{1}{2}\, {}^{\mathbf{u}}w_x^{opt}\right) + {}^{\mathbf{u}}w_x^{opt} \left(1 - \frac{1}{2}\, {}^{\mathbf{u}}q_x\right) = {}^{\mathbf{u}}q_x + {}^{\mathbf{u}}w_x^{opt}\,(1 - {}^{\mathbf{u}}q_x), \;\; x \in AB,$$

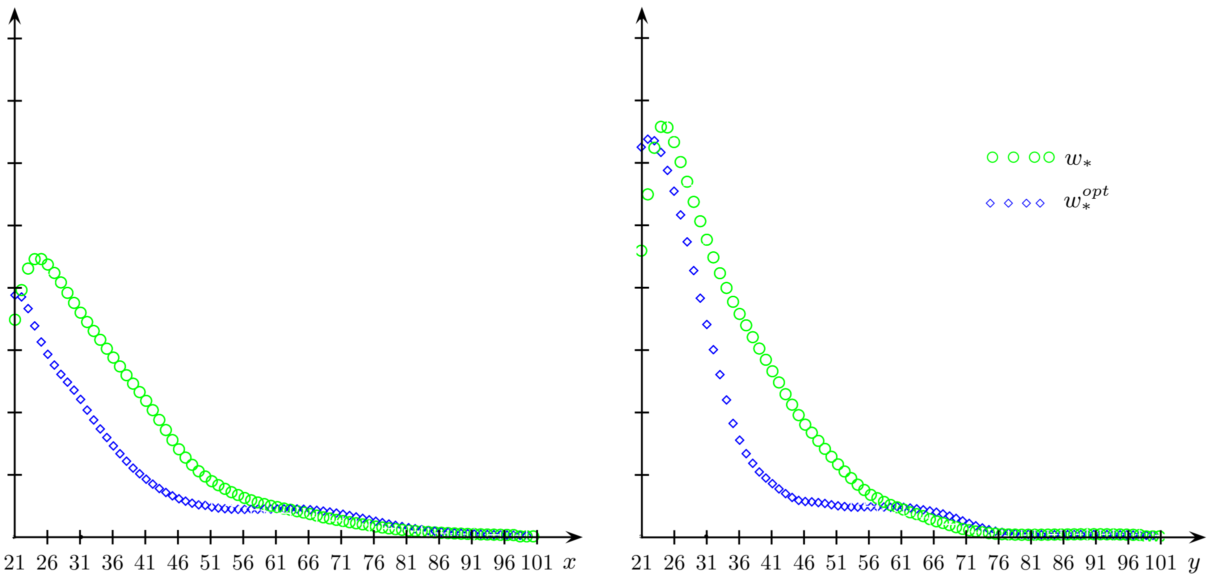

Figure 20: Comparison of original mere age-depending annual withdrawal rates w_* with optimal mere age-depending annual withdrawal rates w^{opt} with respect to the attained age, for male insured (left) and female insured (right)

where the values ${}^{\mathbf{u}}q_x$ are assumed to be given. Hence, we get for the independent annual withdrawal rates

$$ {}^{\mathbf{u}}w_x^{opt} = \frac{(q_x^{opt} + w_x^{opt}) - {}^{\mathbf{u}}q_x}{(1 - {}^{\mathbf{u}}q_x)}, \quad x \in AB. \tag{4.13.10} $$

For the values $(q_x^{opt} + w_x^{opt}), x \in AB \setminus \{x_{MAX}\}$, the results of (4.13.9) must be inserted. Another application of (2.38.7) and (2.38.8) finally gives q_x^{opt} and w_x^{opt} for $x \in AB$.

Figure 20 outlines the original mere age-depending annual withdrawal rates w_* and the optimal annual withdrawal rates w_*^{opt} which appertain to the optimal premiums $(P_x^{opt})_{x \in AB}$. The withdrawal rates w_*^{opt} are derived according to the procedure introduced above. For most ages, these optimal withdrawal rates are smaller than the original withdrawal rates. However, especially for the ages sixty to seventy, the optimal withdrawal rates prevail over the original withdrawal rates. The reason for this is the same effect as described above: For higher ages, the mere age-depending withdrawal rates underestimate the actual age- and duration-depending withdrawal rates not only for short previous contract durations, but also for previous contract durations of up to ten years or more. For those ages, therefore, the age- and duration-depending premiums are sometimes smaller than the mere age-depending premiums. Hence, by determining the optimal premiums as average of the age- and duration-depending premiums, these premiums can also be smaller than the original mere age-depending premiums. For given mortality rates, this results for the corresponding ages in optimal withdrawal rates being higher that the withdrawal rates which appertain to the original mere age-depending premiums.

Figure 21 compares the original net premiums P_* with the optimal net premiums P_*^{opt} for the ages at issue $x_0 = y_0 = 25, ..., 50$. Recall figure 16. Up to the mid-thirties, the optimal net premiums are closely related to the age- and duration-depending premiums $P_{*,0}$ for new entrants. For higher ages, however, the optimal net premiums almost coincide with the original net premiums.

Incidentally, in Helwich and Milbrodt [2007], both the original net premiums and the optimal net premiums are also compared with the premiums of the approach by Rudolph [2005]. Most of the premiums for the latter approach turn out to be between the original net premiums and

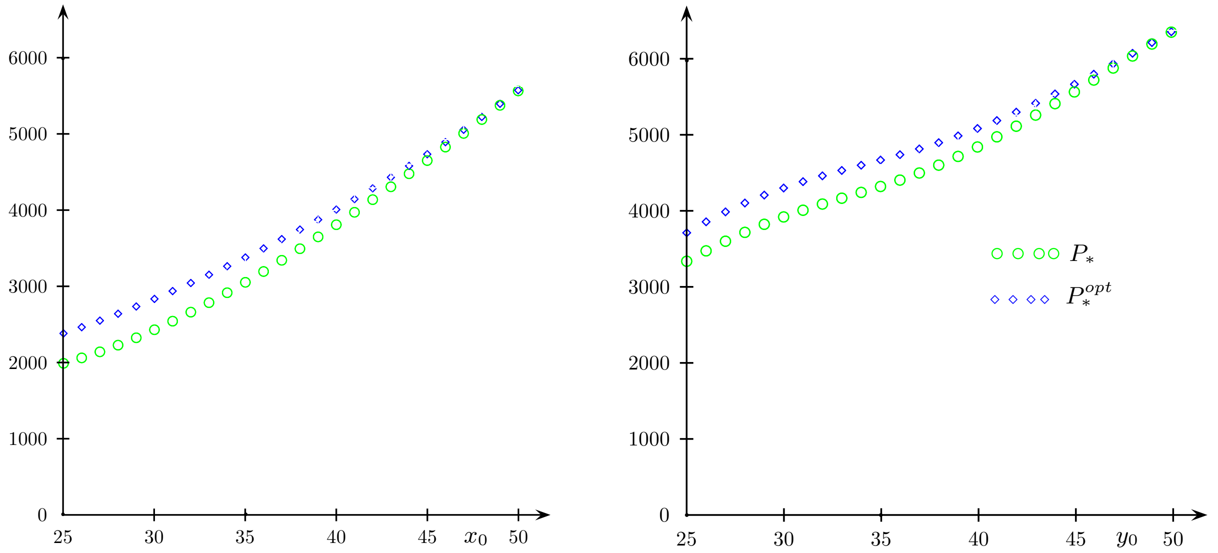

Figure 21: Comparison of original mere age-depending annual net premiums with optimal mere age-depending premiums, with respect to the age at issue $x_0 = y_0 = 25, ..., 50$, for male insured (left) and female insured (right)

the optimal net premiums. Hence, Rudolph's approach takes the durational effects partially into account, yet not in the best possible way. Note that the *best possible way* depends on the criterion for the adaption of one model to another. Instead of the least square method (cf. (4. 13. 7)), one could prefer another way, for example by weighting the differences between the age- and duration-depending premiums and the optimal premiums (to be specified) by the size of the corresponding groups of the portfolio. The latter and other criteria, however, are not further discussed here.

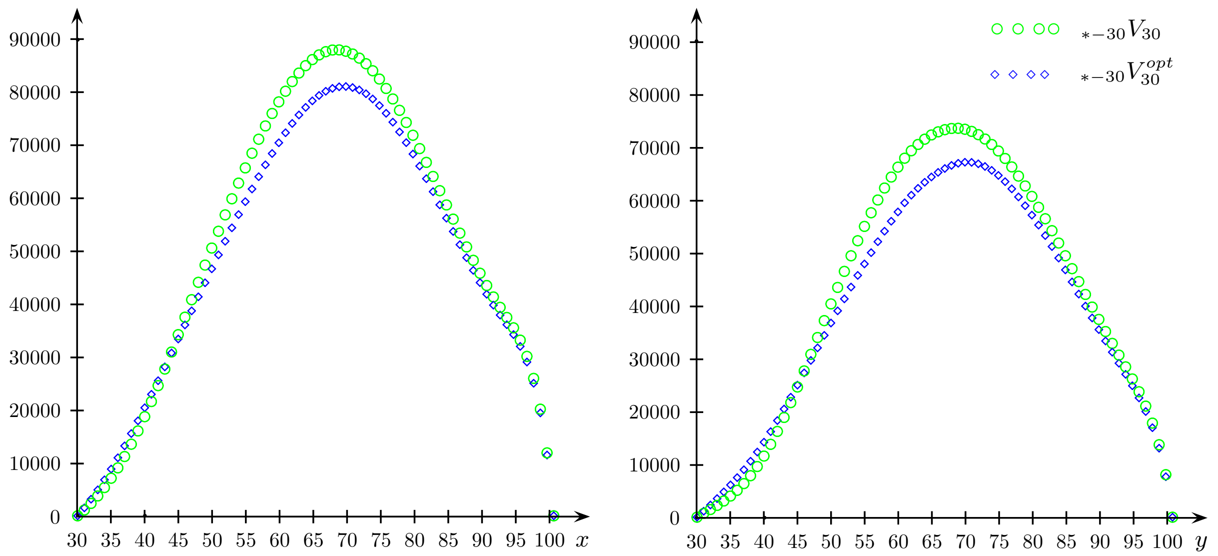

Figure 22: Comparison of prospective reserves appertaining to original net premiums with prospective reserves that appertain to optimal net premiums with respect to the attained age x (y), for male (left) and female (right) insured with age at issue $x_0 = y_0 = 30$

Before discussing PKV modelling for the "new world", we want to have a look at the prospective reserves and the portions of the annual increments of the prospective reserves caused by withdrawing insured. The prospective reserves that appertain to the original net premiums and the optimal net premiums, respectively are illustrated in figure 22 for insured with age at issue

30. For this age, the development of the prospective reserve that appertains to the optimal net premium is similar to the development of the aging provision that appertains to $P_{30,0}$ for the age- and duration-depending approach (cf. figure 18). The reason for this is that the optimal net premium for insured at the age of 30 almost coincides with the age- and duration-depending net premium $P_{30,0}$. Nevertheless, the values of both of these prospective reserves are different. Note that due to the use of different withdrawal rates, the values of these prospective reserves would also be different if the premiums were exactly equal. According to figure 23, the overestimation of the portions of the annual increments of the prospective reserve due to withdrawing insured is here almost completely avoided (cf. figure 19).

In summary we have illustrated that the implementation of durational effects leads to more realistic modelling, even in situations for which the use of an age- and duration-depending model is not permitted. After introducing a corresponding model allowing age- and duration-depending annual decrement rates, we have suggested how to take durational effects in a mere age-depending model into account in the best possible way. By this procedure, a method is provided that does not confront with binding regulations, but nevertheless makes the PKV calculation more realistic. Thus, the dilemma mentioned in the introduction can be solved. Yet, one has to realize that even the best adaption of a model relying on a mere age-depending actuarial basis to a model allowing age-and duration-depending annual rates forms a falsification of the actual structure of the underlying data.

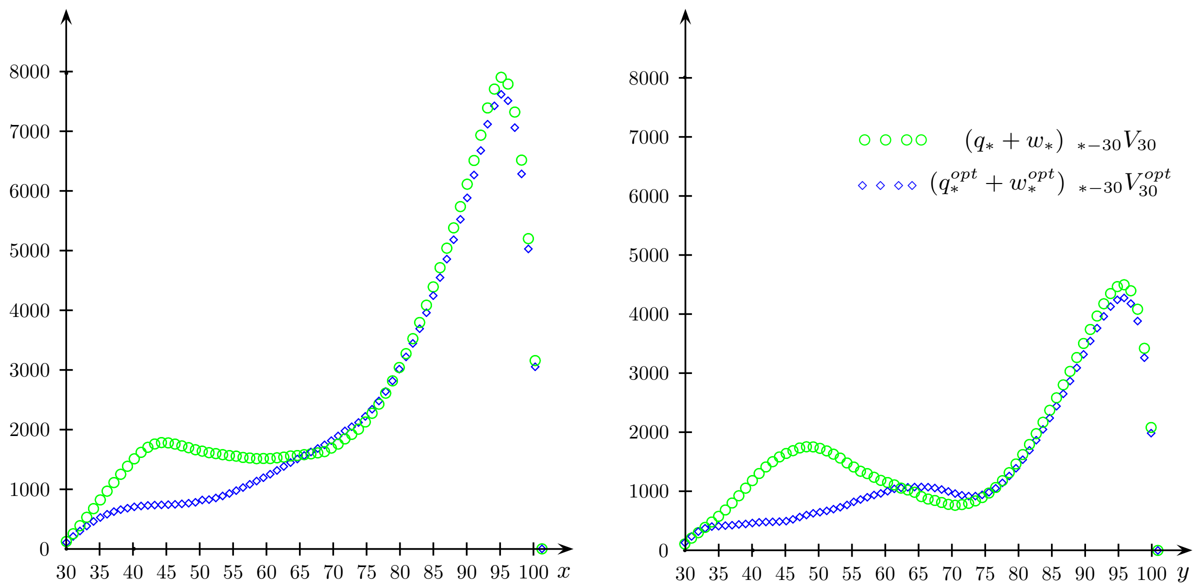

Figure 23: Comparison of the corresponding third addends of (4.13.6) for both prospective reserves illustrated in figure 22

We now turn to the "new world", meaning PKV modelling that takes into account the regulations due to the GKV-WSG. Though the GKV-WSG is concerned with improving the competition for the statutory health insurance system (GKV) in Germany, it also demands substantial reforms for the capital funded part of the German health insurance system, private health insurance. Important issues are, for example, concerned with changing the insurance system (GKV $\leftrightarrow$ PKV), defining full coverage for private health insurance, and establishing a standardized tariff, the so-called Basistarif. Regarding PKV modelling, however, one of the main issues of the GKV-WSG consists in granting a transfer of the aging provision (or at least a substantial part of it) in case of changing the insurer. Hence, in the situation of the "new world", the role of the cause of decrement *withdrawal* changes. For the "new world", this cause of decrement splits up into two possible scenarios with different consequences. The first scenario is given by the change of an insured into another PKV company, and the second corresponds to a change into the statutory health insurance system. Since the GKV-WSG allows the cancellation of an insurance contract with full coverage only in cases where a new contract can be attested, there are no further possibilities. The change into the statutory health insurance system can be considered analogously to the cause of decrement *withdrawal* within the "old world", meaning that insured lose their full aging provision. Hence, the calculation regarding the standardized tariff - according to the revised KalV (see [1996]), the calculation for this tariff only allows the causes of decrement *change into GKV* and *death* - corresponds to the PKV modelling for the "old world". For a change into another PKV company, however, the GKV-WSG intends a transfer value, the so-called Übertragungswert. This transfer value depends on the prospective reserve for the changing insured, but it does not necessarily correspond to the full aging provision. Since according to the revised KalV (§ 13a Abs.1), the maximum amount of this value is closely related to the ageing provision that pertains to an insured with equal previous contract duration in the standardized tariff. It is a matter of current actuarial discussions how to define the transfer value exactly and how to derive it best possible. Such discussions are not enforced here.

Our aim is simply to consider the above introduced approach in the light of the "new world". Regarding this approach to handle durational effects, the situation for the "new world" raises two questions. Firstly, there is the question of whether or not the problem, consisting in the risk of possible losses (or in the shifting of risks) due to disregarding the actual duration-dependence of withdrawal rates, will continue to exist. According to the theorem of Cantelli (cf. Milbrodt and Helbig [1999], section 10D), the cause of decrement *withdrawal* can completely be disregarded in cases where the overall prospective reserve serves as withdrawal benefit. This, however, is not fulfilled for the situation of the "new world". By splitting up the former cause of decrement *withdrawal* into two causes of decrement, *transition into GKV* and *transition into PKV*, it is obvious that *transition into GKV* implicates the loss of the full ageing provision. Further, for *transition into PKV*, insured may also lose a part of there aging provision. Hence, none of these two causes of decrement can be disregarded. Concerning the duration-dependence of the rates for these transitions, future investigations must be enforced to analyze the impacts of the dependence on age and previous contract duration.

The second question raised by the new situation is concerned with the adaption of our approach. Due to the intended transfer values - which will here for the sake of simplicity be regarded as age- and duration-depending transition benefits - the PKV model introduced in example 2. 38 must be generalized. In what follows, we will generally consider the implementation of age- and duration-depending transition benefits, which can, for example, be specified as transfer values. For PKV modelling, there are two different ways of implementing the financing of a transition benefit for decrement due to *transition into PKV*. The first is given by considering a model with three causes of decrement, with the decrement due to *transition into PKV* incurring a transition benefit. The second is given by the usual model with two causes of decrement, yet with the withdrawal rates being reduced such that they reflect the reserve that is actually left by withdrawing insured.

We start by considering an age- and duration-depending model with three causes of decrement. For $x \in AB$ and $u \in \{0, ..., x - x_{MIN}\}$ let $q_{x,u}$ be annual mortality rates, $w_{x,u}^{GKV}$ annual rates for *transition into GKV*, and $w_{x,u}^{PKV}$ annual rates for *transition into PKV*. For an insured with attained age x and previous contract duration u, let ${}_mT_{x,u}$ denote the transition benefit at time $m \in \mathbb{N}_0$, with $x + m \leq x_{MAX}$, which is assumed to be payable at the end of the year of transition. By assuming (2. 38. 1), and adapting (2. 38. 2) as well as (2. 38. 4) to the present situation, one obtains according to (4. 8. 2) for the present value of benefits (cf. (4. 13. 1))

$$A_{x,u} := \sum_{\nu=0}^{x_{MAX}-x} v^{\nu} K_{x+\nu} \; {}_{\nu}p_{x,u} + \sum_{\nu=0}^{x_{MAX}-x} v^{\nu+1} \; {}_{\nu+1}T_{x,u} \; {}_{\nu}p_{x,u} \; w_{x+\nu,u+\nu}^{PKV}, \qquad (4.13.11)$$

where

$$ {}_{\nu}p_{x,u} = \prod_{j=0}^{\nu-1} p_{x+j,u+j} = \prod_{j=0}^{\nu-1}(1 - q_{x+j,u+j} - w_{x+j,u+j}^{GKV} - w_{x+j,u+j}^{PKV}). \tag{4.13.12} $$

By taking (4. 13. 12) into account, the annual net premiums can be obtained by means of (4. 13. 3), with $\ddot{a}_{x,u}$ being specified by (4. 13. 2). For a given mortality table, given annual age- and duration-depending decrement rates (which can be specified as explained in example 2. 38), and given transition benefits, the above formulas yield age- and duration-depending premiums $P_{x,u}, x \in AB, u \in \{0, ..., x - x_{MIN}\}$. Using these premiums, the optimal mere age-depending premiums $P_x^{opt}, x \in AB$, can be obtained as average with respect to all possible previous durations.

The optimal premiums $P_x^{opt}, x \in AB$, can be used to determine the corresponding decrement rates. For this, we consider the model with two causes of decrement and reduced withdrawal rates, with *withdrawal* including both *transition into PKV* and *transition into GKV*. For this model, the mortality rates $q_{x,u}$ are the same as for the above model. Regarding the withdrawal rates, we follow the approach of Richter [2007], yet generalized by allowing age- and duration-depending parameters. According to this approach, the withdrawal rates are given by $w_{x,u} = w_{x,u}^{GKV} + \tilde{w}_{x,u}^{PKV}$, where $\tilde{w}_{x,u}^{PKV} := w_{x,u}^{PKV}(1 - \kappa_{x,u})$ denote the reduced annual rates for the decrement due to *transition into PKV*. The quantity $\kappa_{x,u}$ is implemented to reduce the corresponding decrement rates in order to finance the transition benefit. The present value of benefits is for this approach given by (4. 13. 1), with ${}_{\nu}p_{x,u}$ being specified by

$$ {}_{\nu}p_{x,u} = \prod_{j=0}^{\nu-1}(1 - q_{x+j,u+j} - w_{x+j,u+j}^{GKV} - w_{x+j,u+j}^{PKV}(1 - \kappa_{x+j,u+j})). \tag{4.13.13} $$

Obviously, for $\kappa_{x,u} \equiv 0$ and ${}_{m}T_{x,u} \equiv 0$, both models coincide.

Considering the second model, one obtains, by disregarding the duration-dependence of the parameters, for the present value of benefits (4. 13. 1), with

$$ {}_{\nu}p_{x} = \prod_{j=0}^{\nu-1}(1 - q_{x+j} - w_{x+j}^{GKV} - w_{x+j}^{PKV}(1 - \kappa_{x+j})). \tag{4.13.14} $$

Using the optimal premiums $P_x^{opt}, x \in AB$, and the above introduced recursion to obtain corresponding optimal decrement rates, one gets analogously to (4. 13. 9) the following sums of optimal parameters

$$ q_x^{opt} + w_x^{GKV\,opt} + w_x^{PKV\,opt}(1 - \kappa_x^{opt}), \;\; x \in AB. \tag{4.13.15} $$

From these sums, the single components must be derived. As in the situation of the "old world", the parameters which do not depend on the previous contract duration must be used to specify the other optimal parameters. Thus, by omitting the conversion from independent rates into dependent rates, the above sums can be reduced by the given mortality rates. Further, we assume that mere age-depending rates for decrement due to *transition into GKV* are also given. This can be granted by assuming that these decrement rates do not depend on the previous contract duration. As mentioned above, the impact of the previous contract duration on the rates for decrement due to *transition into GKV* should be investigated. Since such transitions seem to be rather motivated by administrational regulations than by the own choice of an insured, we could imagine that the duration-dependence of these rates is not significant. Another way to get optimal mere age-depending rates for *transition into GKV* - even in case that they are actually age- and duration-depending - is given by assuming that these rates are the same as the rates for *transition into GKV* used for the calculation of the standardized tariff. Since for the latter, optimal mere age-depending rates for *transition into GKV* can be simply derived according to our introduced method for the "old world".

In summarizing the above we assume that it remains to specify $w_x^{PKV\,opt}$ and κ_x^{opt} for $x \in AB$. This can be done by firstly setting ${}_mT_{x,u} \equiv 0$ and $\kappa_{x,u} \equiv 0$, and afterwards deriving $w_x^{PKV\,opt}$ with the corresponding optimal premiums based on the above first model without transition benefit. Thus, $w_x^{PKV\,opt}$ can be obtained from the sums

$$q_x^{opt} + w_x^{GKV\,opt} + w_x^{PKV\,opt}, \;\; x \in AB. \tag{4.13.16}$$

Finally, κ_x^{opt} can be specified by using (4. 13. 15) as well as q_x^{opt}, $w_x^{GKV\,opt}$, and $w_x^{PKV\,opt}$.

Regarding the "new world", we altogether maintain, that firstly, it must be investigated which parameters do significantly depend on the previous duration of an insured in the portfolio. Beside the decrement rates, also the claims amounts per risk can be taken into consideration. After that, an age- and duration-depending model must be used to obtain the actual age- and duration-depending premiums, from which the optimal premiums can be derived. The crucial step is then to get the optimal values for all parameters that actually depend on the previous contract duration. As we have seen, this can be realized with the approach introduced above, provided that not too many parameters must be specified. △

As announced, the second example is concerned with a non-discrete model for permanent health insurance.

4. 14 Example. Consider a single policy (p) modelled by a pure jump process $(X_t)_{t\geq 0}$ with states $a \sim$*active*, $i \sim$*invalid* and $d \sim$*dead*. Recovery is implemented in the model. Hence, the set of transitions is given by $\mathcal{J} := \{(a,i),(a,d),(i,a),(i,d)\}$. The set of states and the possible transitions for this model are sketched in figure 3. Further, the following tables are used: The mortality rates for active insured are taken from the mortality table DAV-ST 1994 T; for the disability rates, we use the table DAV-IT 1997. The mortality rates as well as the recovery rates for disabled insured are originated from the select-and-ultimate tables DAV-SST 1997 TI and DAV-SRT 1997 RI, respectively. Note that we only use the corresponding tables for male insured. The rates of transitions from the state *active* are assumed to depend only on the attained age of an insured. The rates of transitions from the state *invalid*, however, are allowed to depend additionally on the time elapsed since disablement.

According to the previous example, tables of annual rates determine the distribution of a discrete model $(X_t, U_t) = (X_{\lfloor t\rfloor}, U_{\lfloor t\rfloor}), t \geq 0$, provided that a condition of stationarity (cf. (2. 38. 1)) holds. For such a discrete model, transitions are only allowed at integer times or at integer ages. In order to specify the joint distribution of a jump time and the corresponding destination state between integers, further assumptions must be added. Doing so, it is often assumed in actuarial science that the distribution of transitions within each year of age is a uniform distribution. To achieve this, linear interpolation can be employed according to what follows. Before this, however, an appropriate condition of stationarity will be formulated. For this, we closely follow Milbrodt and Helbig ([1999], section 6B), who introduced a condition of stationarity for a generalized Markov approach.

Let $x \geq 0$ be the age at issue for the holder of the policy (p), and $(\Omega, \mathfrak{F}, P, ((X_t^{(x)}, U_t^{(x)}))_{t\geq 0})$ a bivariate Markov process recording the current state of the policy and the corresponding time elapsed since entering this state. Further, for $s \geq 0$, the time of the first jump after s and the corresponding destination state are denoted by $T^{(x)}(s)$ and $X^{(x)}_{T^{(x)}(s)}$, respectively. According to definition 2. 24, we define the conditional distribution of $T^{(x)}(s)$ and $X^{(x)}_{T^{(x)}(s)}$, given the the current state $X_s^{(x)}$ and the time elapsed since entering this state $U_s^{(x)}$, as

$$Q_{yz}^{(x)}(s,t,u) \quad := \quad P\left(T^{(x)}(s) \le t, X_{T(s)}^{(x)} = z |\, X_s^{(x)} = y,\, U_s^{(x)} = u\right),\ \ 0 \le u \le s \le t, (y,z) \in \mathcal{J}.$$

Further,

$$Q_y^{(x)}(s,t,u) \quad := \quad \sum_{\substack{z \in \mathcal{S} \\ z \neq y}} Q_{yz}^{(x)}(s,t,u),\ \ 0 \le u \le s \le t, y \in \mathcal{S}. \tag{4.14.1}$$

$q_{yz}^{(x)}$ is assumed to be the appertaining cumulative transition intensity for the transition from y to z, and $-q_{yy}^{(x)}$ the cumulative intensity of decrement for state y. In view of (2.38.1), we state the following assumption:

Assumption.
For $x \ge 0$ and $h \ge 0$ we stipulate

$$Q_{yz}^{(x)}(s+h, t+h, u) = Q_{yz}^{(x+h)}(s,t,u),\ \ 0 \le u \le s \le t, (y,z) \in \mathcal{J}, \tag{4.14.2}$$

which corresponds to

$$\begin{aligned} &P\left(T^{(x)}(s+h) \le t+h, X_{T(s+h)}^{(x)} = z |\, X_{s+h}^{(x)} = y,\, U_{s+h}^{(x)} = u\right) \\ &= P\left(T^{(x+h)}(s) \le t, X_{T(s)}^{(x+h)} = z |\, X_s^{(x+h)} = y,\, U_s^{(x+h)} = u\right). \end{aligned} \tag{4.14.3}$$

To avoid confusion when comparing (4.14.3) and (2.38.1), realize that the age x denotes here the age at issue of the contract. In the situation of (2.38.1), however, x denotes the attained age of an insured. Further, the time elapsed since the current state was entered is for the PKV model always related to the state *active* and corresponds to the difference of attained age and age at issue. From (4.14.2) and (4.14.1), it follows that

$$Q_y^{(x)}(s+h, t+h, u) = Q_y^{(x+h)}(s,t,u),\ \ 0 \le u \le s \le t, y \in \mathcal{S},$$

as well as

$$q_{yz}^{(x)}(s+h, t+h, u) = q_{yz}^{(x+h)}(s,t,u),\ \ 0 \le u \le s \le t, (y,z) \in \mathcal{S}^2. \tag{4.14.4}$$

The latter results in the fact that a single cumulative transition intensity matrix $q^{(0)}$ is sufficient to obtain the corresponding transition intensities for each age x, namely by means of

$$q_{yz}^{(x)}(s,t,u) = q_{yz}^{(0)}(x+s, x+t, u),\ \ 0 \le u \le s \le t, (y,z) \in \mathcal{S}^2.$$

In contrast to Milbrodt and Helbig ([1999], section 6B), this transition intensity matrix does not correspond to a Borel measure on the positive real line, but it relies on a kernel in accordance to the additional duration-dependence.

By employing the exponential formula (2.24.7) and the property (2.27.1), one obtains with (4.14.4) a counterpart to the usual product formula for the probabilities of remaining in a certain state (cf. (2.38.2)). Thus, for $y \in \mathcal{S}$, $x, u \ge 0$, and $k, l \in \mathbb{N}$ satisfying $u \le k$, the following holds:

$$\bar{p}_y^{(x)}(k, k+l, u) = \prod_{j=0}^{l-1} \bar{p}_y^{(x)}(k+j, k+j+1, u+j) = \prod_{j=0}^{l-1} \bar{p}_y^{(x+k+j)}(0, 1, u+j). \tag{4.14.5}$$

Regarding select-and-ultimate tables, the right-hand side of the above equation reflects for $u \in \mathbb{N}$ the use of annual rates along the arrows in figure 4, since the time spent in a certain state increases with increasing age, provided that no transition occurs. Note that annual rates, contained in a certain table, correspond to probabilities of certain transitions between two successive integer times. This means, for example, for annual mortality rates of disabled insured - as gathered in the table DAV-SST 1997 TI - the following: Let $x \in \mathbb{N}$ be the technical age of an insured and $n, u \in \mathbb{N}$ satisfying $u \leq n$. Then

$$Q_{id}^{(x)}(n, n+1, u) := q_{[x+n-u]+u}^{ii}.$$

According to condition (4.14.2), the age at issue is in the sequel not explicitly indicated. Doing so, we write Q_* instead of $Q_*^{(x)}$. Further, realize that the entries of select-and-ultimate tables actually correspond to independent probabilities ${}^u q_*^{ii}$, which should be converted into dependent probabilities q_*^{ii} (cf. Bowers et al. [1997], section 10.5.1 or Milbrodt and Helbig [1999], section 3 C as well as exercise 3.17 (c)). But as mentioned above, the numerical differences are fairly small, so that the conversion is often omitted.

After all these preliminaries, we will now explain how to derive the corresponding distributions of transitions between integer ages. Doing so, we consider for $(y, z) \in \mathcal{J}$ and $n, u \in \mathbb{N}$ satisfying $u \leq n$

$$Q_{yz}(n, n+1, u) = P(T(n) \leq n+1, X_{T(n)} = z | \, X_n = y, U_n = u),$$

and define

$$Q_{yz}(n, t, u) := (t-n)\, Q_{yz}(n, n+1, u), \;\; t \in (n, n+1]. \tag{4.14.6}$$

This definition yields the following consequences (cf. Milbrodt and Helbig [1999], Satz 6.24).

Theorem. *Let* $(y, z) \in \mathcal{J}$, $n, u \in \mathbb{N}$ *satisfying* $u \leq n$, *and* $n < s \leq t \leq n+1$. *Further, assume* (4.14.6) *and* $P(X_n = y, U_n = u) > 0$. *Then*

$$Q_{yz}(s, t, s-n+u) = \frac{(t-s)\, Q_{yz}(n, n+1, u)}{1-(s-n)\, Q_y(n, n+1, u)}, \tag{4.14.7}$$

and

$$Q_y(s, t, s-n+u) = \frac{(t-s)\, Q_y(n, n+1, u)}{1-(s-n)\, Q_y(n, n+1, u)}. \tag{4.14.8}$$

Further, the corresponding cumulative transition intensities satisfy (2.36.6) *and* (2.36.7) *by means of the densities*

$$\mu_{yz}(t, t-n+u) = \frac{Q_{yz}(n, n+1, u)}{1-(t-n)\, Q_y(n, n+1, u)}, \;\; t \in (n, n+1], \tag{4.14.9}$$

as well as

$$\mu_{yy}(t, t-n+u) = -\frac{Q_y(n, n+1, u)}{1-(t-n)\, Q_y(n, n+1, u)}, \;\; t \in (n, n+1]. \tag{4.14.10}$$

Thus, $q_{yy}|_{\mathfrak{B}((n,n+1])}$ *is finite as long as* $Q_y(n, n+1, u) < 1$.

Proof. We start by verifying (4.14.7). Let $(y, z) \in \mathcal{J}$, $n, u \in \mathbb{N}$ satisfying $u \leq n$, and $n < s \leq t \leq n+1$. Consider $Q_{yz}(n, (s, t], u)$. On the one hand, we obtain according to (4.14.6)

$$\begin{aligned} Q_{yz}(n, (s, t], u) &= Q_{yz}(n, t, u) - Q_{yz}(n, s, u) \\ &= (t-n)\, Q_{yz}(n, n+1, u) - (s-n)\, Q_{yz}(n, n+1, u) \\ &= (t-s)\, Q_{yz}(n, n+1, u). \end{aligned} \tag{4.14.11}$$

On the other hand, one gets by employing (2.28.4), (2.27.8), (2.27.1), and likewise (4.14.6),

$$\begin{aligned} Q_{yz}(n,(s,t],u) &= \int_{(s,t]} \bar{p}_y(n,r-0,u)\,q_{yz}(n,dr,u) \\ &= \int_{(s,t]} \bar{p}_y(n,s,u)\,\bar{p}_y(s,r-0,s-n+u)\,q_{yz}(s,dr,s-n+u) \\ &= (1-Q_y(n,s,u)) \int_{(s,t]} \bar{p}_y(s,r-0,s-n+u)\,q_{yz}(s,dr,s-n+u) \\ &= (1-(s-n)\,Q_y(n,n+1,u))\,Q_{yz}(s,t,s-n+u). \end{aligned} \tag{4.14.12}$$

Together with (4.14.11), (4.14.12) yields (4.14.7). (4.14.8) follows from (4.14.7) by deriving $\sum_{\substack{z\in S\\ z\neq y}} Q_{yz}(s,t,s-n+u)$.

In order to prove (4.14.9), we differentiate (4.14.7) with respect to t. This yields

$$\frac{\partial Q_{yz}(s,t,s-n+u)}{\partial t} = \frac{Q_{yz}(n,n+1,u)}{1-(s-n)\,Q_y(n,n+1,u)}. \tag{4.14.13}$$

According to (2.28.4) and (2.36.6), the left-hand side of (4.14.13) also corresponds to

$$\bar{p}_y(s,t,s-n+u)\,\mu_{yz}(t,t-s+s-n+u) = (1-Q_y(s,t,s-n+u))\,\mu_{yz}(t,t-n+u).$$

Hence,

$$\mu_{yz}(t,t-n+u) = \frac{Q_{yz}(n,n+1,u)}{(1-(s-n)\,Q_y(n,n+1,u))\,(1-Q_y(s,t,s-n+u))},$$

which yields the assertion (4.14.9) by applying (4.14.8). Further, (4.14.10) follows according to (2.36.7). □

As mentioned above, select-and-ultimate tables only provide transition probabilities for integer times. In the same manner, the duration-dependence is likewise only recorded for integer values. In order to obtain a model for continuous time, we will therefore also employ linear interpolation on this level. In doing so, we define in addition to (4.14.6) for $u \in (m-1,m]$ with $m \in \mathbb{N}$ satisfying $m \le n$

$$\begin{aligned} Q_{yz}(n,n+1,u) &:= (m-u)\,Q_{yz}(n,n+1,m-1) + (u-m+1)\,Q_{yz}(n,n+1,m) \\ &= (m-u)\,Q_{yz}(n,n+1,m-1) - (m-u-1)\,Q_{yz}(n,n+1,m). \end{aligned} \tag{4.14.14}$$

By inserting (4.14.14) into the right-hand sides of (4.14.7) and (4.14.9), the corresponding formulas also become applicable for $u \notin \mathbb{N}$. By afterwards substituting $l := s-n+u$, we obtain for $n \le s \le t \le n+1$ and $l \ge s-n$ satisfying $l-s+n \in (m-1,m]$

$$\begin{aligned} &Q_{yz}(s,t,l) \\ &= \frac{(t-s)\Big((m-l+s-n)\,Q_{yz}(n,n+1,m-1) - (m-l+s-n-1)\,Q_{yz}(n,n+1,m)\Big)}{1-(s-n)\Big((m-l+s-n)\,Q_y(n,n+1,m-1) - (m-l+s-n-1)\,Q_y(n,n+1,m)\Big)}. \end{aligned} \tag{4.14.15}$$

For $l < s-n$, we set

$$Q_{yz}(s,t,l) = \frac{(t-s)\Big(m\,Q_{yz}(n,n+1,m-1) - (m-1)\,Q_{yz}(n,n+1,m)\Big)}{1-(s-n)\Big(m\,Q_y(n,n+1,m-1) - (m-1)\,Q_y(n,n+1,m)\Big)}.$$

This corresponds to the case of $l = s - n$. Regarding the densities, we get, by substituting $l := t - n + u$, for $t \in (n, n+1]$ and $l \geq t - n$ satisfying $l - t + m \in (m-1, m]$

$$\mu_{yz}(t,l) \tag{4.14.16}$$
$$= \frac{\Big((m-l+t-n)\, Q_{yz}(n,n+1,m-1) - (m-l+t-n-1)\, Q_{yz}(n,n+1,m)\Big)}{1-(t-n)\Big((m-l+t-n)\, Q_{y}(n,n+1,m-1) - (m-l+t-n-1)\, Q_{y}(n,n+1,m)\Big)}.$$

The case of $l < t - n$ is likewise put on a level with $l = t - n$, resulting in

$$\mu_{yz}(t,l) = \frac{\Big(m\, Q_{yz}(n,n+1,m-1) - (m-1)\, Q_{yz}(n,n+1,m)\Big)}{1-(t-n)\Big(m\, Q_{y}(n,n+1,m-1) - (m-1)\, Q_{y}(n,n+1,m)\Big)}. \tag{4.14.17}$$

In summary, the assumptions (4.14.2), (4.14.6), and (4.14.14) allow us to fully base the calculations concerning a single policy (p) described by a bivariate Markov process $((X_t, U_t))_{t\geq 0}$ on corresponding select-and-ultimate tables. Under these assumptions, the cumulative transition intensities are continuous and piecewise continuously differentiable. The appertaining densities can be obtained by (4.14.16), (4.14.17), and

$$\mu_{yy}(t,l) = -\sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \mu_{yz}(t,l).$$

We turn back to the above PHI policy (p) and assume a finite policy term which is in the sequel denoted by n. The annual sojourn benefits, payable in advance, are given by

$$F_y(s,B,u) \;=\; \int_B f_y(\tau,\tau-s+u) \sum_{k\geq s}^{n-1} \varepsilon_k(d\tau), \quad 0\leq u\leq s,\ y\in\mathcal{S},\ k\in\mathbb{N},\ B\in\mathfrak{B}([s,n)).$$

Then, by inserting

$$\bar{p}_y(s,\tau,u) \;=\; \exp\{q^{(c)}_{yy}(s,\tau,u)\} = \exp\Big\{\int_{(s,\tau]} \mu_{yy}(r,r-s+u)\, dr\Big\}, \quad \tau\geq s$$

(cf. (2.24.7)), Thiele's integral equation of type 1 is for $s \geq 0$ and $y \in \mathcal{S}$ of the form

$$v(s)\, V^{+}_{(y,u)}(s) \tag{4.14.18}$$
$$= \int_{[s,n)} v(\tau)\, e^{\int_{(s,\tau]} \mu_{yy}(r,r-s+u)\,dr}\, f_y(\tau,\tau-s+u) \sum_{k\geq s}^{n-1} \varepsilon_k(d\tau)$$
$$+\sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,n]} \Big(\frac{D_{yz}(\tau,\tau-s+u)}{K(DT(\tau))} + v(\tau)\, V^{+}_{(z,0)}(\tau)\Big)\, e^{\int_{(s,\tau]} \mu_{yy}(r,r-s+u)\,dr}\, \mu_{yz}(\tau,\tau-s+u)\, d\tau.$$

According to (4.8.4), the above system of integral equations must only be solved for $u = 0$. Upon setting $u = 0$, we obtain

$$v(s)\, V^{+}_{(y,0)}(s) \tag{4.14.19}$$
$$= \int_{[s,n)} v(\tau)\, e^{\int_{(s,\tau]} \mu_{yy}(r,r-s)\,dr}\, f_y(\tau,\tau-s) \sum_{k\geq s}^{n-1} \varepsilon_k(d\tau)$$
$$+\sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,n]} \Big(\frac{D_{yz}(\tau,\tau-s)}{K(DT(\tau))} + v(\tau)\, V^{+}_{(z,0)}(\tau)\Big)\, e^{\int_{(s,\tau]} \mu_{yy}(r,r-s)\,dr}\, \mu_{yz}(\tau,\tau-s)\, d\tau.$$

We recommend solving this system of integral equations numerically, by applying a simple quadrature rule. Doing so, we choose the right-hand side rectangle formula. In order to solve an integral of the form $\int_{(a,b]} I(\tau)\,d\tau$, let $l \in \mathbb{N}$, $h := b - a/l$ and $x_k = a + k \cdot h, k = 0, ..., l$. Then, one obtains

$$\int_{(a,b]} I(\tau)\,d\tau \approx h\Big(\sum_{j=1}^{l} I(x_j)\Big).$$

This quadrature rule allows us, by setting $a = s$ and $b = n$, to solve the system of integral equations (4. 14. 19) recursively, by means of

$$\begin{aligned}
& v(x_{k-1})\, V^{+}_{(y,0)}(x_{k-1}) \qquad (4.\,14.\,20)\\
&= \int_{[x_{k-1},n)} v(\tau)\, e^{\int_{(x_{k-1},\tau]} \mu_{yy}(r,r-x_{k-1})\,dr} f_y(\tau,\tau-x_{k-1}) \sum_{k \geq x_{k-1}}^{n-1} \varepsilon_k(d\tau)\\
&\quad + h \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \sum_{j=k}^{l} \Big(\frac{D_{yz}(x_j, x_j - x_{k-1})}{K(DT(x_j))} + v(x_j)\, V^{+}_{(z,0)}(x_j)\Big)\, e^{\int_{(x_{k-1},x_j]} \mu_{yy}(r,r-x_{k-1})\,dr} \mu_{yz}(x_j, x_j - x_{k-1})\\
&= \sum_{k \geq x_{k-1}}^{n-1} v(k)\, e^{\int_{(x_{k-1},k]} \mu_{yy}(r,r-x_{k-1})\,dr} f_y(k, k-x_{k-1})\\
&\quad + h \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \sum_{j=k}^{l} \Big(\frac{D_{yz}(x_j, x_j - x_{k-1})}{K(DT(x_j))} + v(x_j)\, V^{+}_{(z,0)}(x_j)\Big)\, e^{\int_{(x_{k-1},x_j]} \mu_{yy}(r,r-x_{k-1})\,dr} \mu_{yz}(x_j, x_j - x_{k-1}),
\end{aligned}$$

for $k = 1, ..., l$. The starting point for this recursion is the point of termination of the contract n. The initial condition is given by $v(x_l)\, V^{+}_{(y,0)}(x_l) = v(n)\, V^{+}_{(y,0)}(n) = 0, \;\; \forall y \in \mathcal{S}$. With the recursion (4. 14. 20), we obtain for all states $y \in \mathcal{S}$ the values

$$v(x_k)\, V^{+}_{(y,0)}(x_k), \;\; k = 0, ..., l. \qquad (4.\,14.\,21)$$

Afterwards, in order to derive a specific prospective reserve $V^{+}_{(y,u)}(s)$, $0 \leq u \leq s, y \in \mathcal{S}$ according to (4. 14. 18), the integral of the right-hand side of (4. 14. 18) can be likewise numerically computed by inserting the values (4. 14. 21).

We want to continue by specifying the policies under consideration and the surrounding framework for this example. The capital accumulation function K is determined by assuming compound interest with force of interest (interest intensity) $\phi = \ln(1.04)$. The assurance payments according to definition 3. 9 are chosen as independent of the time spent in the current state, since this example is presented to quantify the impact of durational effects due to recovery and mortality rates of disabled insured. In order to achieve this, the results of a calculation based on a semi-Markov model shall be compared with the results relying on the corresponding Markov model. (For the latter, the ultimate tables of the select-and-ultimate tables DAV-SST 1997 TI and DAV-SRT 1997 RI are used.) Consequently, in order to ensure that the differences are solely caused by regarding or disregarding the duration-dependence of the corresponding rates, we assume equal actuarial payments for both models. Doing so, we define the transition benefits according to

$$D_{ai} \equiv D_{ia} \equiv 0, \quad D_{ad} \equiv D_{id} \equiv 1000, \quad \text{and} \quad DT = Id.$$

In the case of disability, the insured is provided an annual annuity of 100 currency units, payable in advance. This corresponds to

$$F_i(s, d\tau, u) = 100 \cdot \sum_{k \geq s}^{n-1} \varepsilon_k(d\tau), \;\; 0 \leq u \leq s \leq n.$$

The net premium P is only payable in state *active*, likewise annually in advance:

$$F_a(s, d\tau, u) = P \cdot \sum_{k \geq s}^{n-1} \varepsilon_k(d\tau), \quad 0 \leq u \leq s \leq n.$$

P is determined by the principle of equivalence. It can be obtained upon dividing the actuarial present value of benefits PV_B by the actuarial present value of premiums PV_P, both of which can be obtained by computing $V^+_{(a,0)}(0)$ according to (4. 14. 19). The former follows by assuming zero premiums, and the latter by assuming zero benefits.

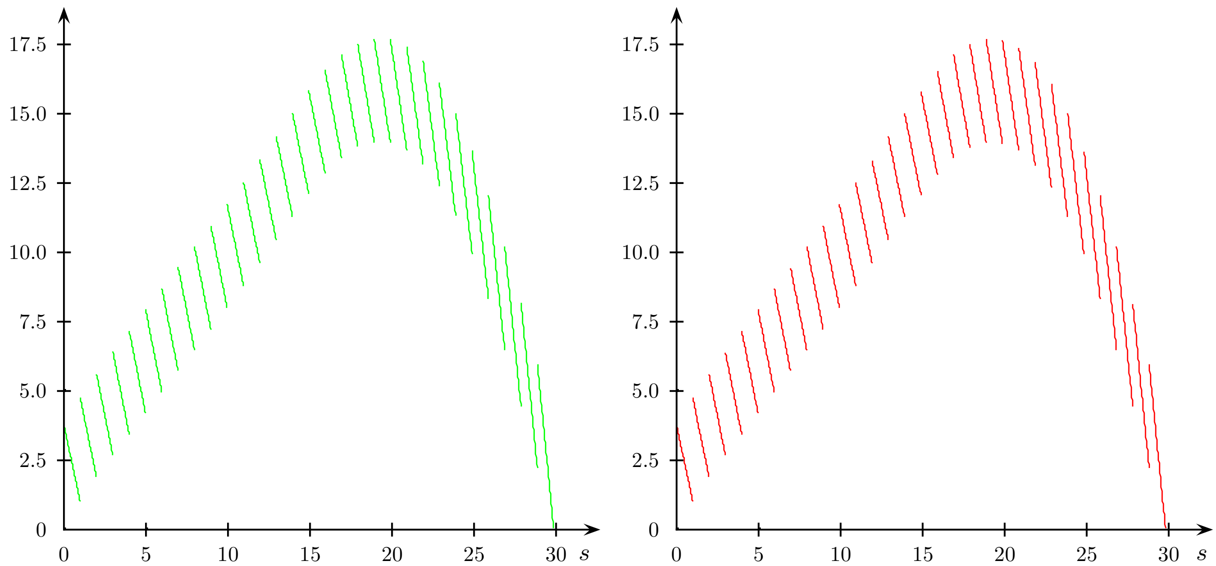

Figure 24: Comparison of prospective reserves in state *active* for male insured with age at issue $x = 20$; $V^+_{(a)}(s)$ (left) is the prospective reserve within a Markov model, and $V^+_{(a,s)}(s)$ (right) is the prospective reserve relying on the corresponding semi-Markov model

We consider policies with term $n = 30$ for insured of three different ages at issue, $x = 20, 30, 40$. Figure 24 illustrates, for both the Markov model and the semi-Markov model, the development of the prospective reserve for insured with age at issue $x = 20$ who remain in state *active*. In case of a Markov model, this prospective reserve only depends on the time since policy issue. Yet, in case of a semi-Markov model, the reserve in state *active* also depends on the time elapsed since the current state was entered. As long as the insured remains in state *active*, however, this time corresponds to the time since policy issue. According to figure 24, no difference can be observed between the prospective reserve based on a Markov model (left) and the prospective reserve based on a semi-Markov model (right). The reason for this is that the rates of transitions from the state *active* are for both models assumed to be independent of the time elapsed since entering this state. Yet, though no difference can be observed, the numerical values differ by a magnitude of up to five percent. The duration-dependence of both recovery and mortality rates causes - particularly for short durations elapsed since onset of disability - higher probabilities of leaving the state *invalid* (cf. figure 5). Further, most of the benefits are usually intended for this state. Hence, the actuarial present values of benefits decrease when a semi-Markov model is used. The table below presents the actuarial present values of benefits PV_B and the net premiums P for all three ages at policy issue being considered. The last row contains the percentage of the premiums based on the semi-Markov approach when compared with the premiums of the Markov approach. Note that the use of a semi-Markov model yielding smaller premiums is not generally valid. Rather, it depends on the construction of the policy specifications.

age at issue	$x = 20$		$x = 30$		$x = 40$	
	PV_B	P	PV_B	P	PV_B	P
Markov model	67.05	3.87	132.96	7.94	304.62	20.37
semi-Markov model	64.99	3.75	128.04	7.62	298.77	19.92
percentage		97%		96%		97.8%

We now turn to the prospective reserves in state *invalid*. The figures 25a - 25c outline, likewise for both models, the prospective reserves in this state with respect to the time since policy issue and accordingly with respect to both the time since policy issue and the time elapsed since disablement. Since a policy is usually commencing in state *active*, we consider the development of the reserves beginning five years after policy issue.

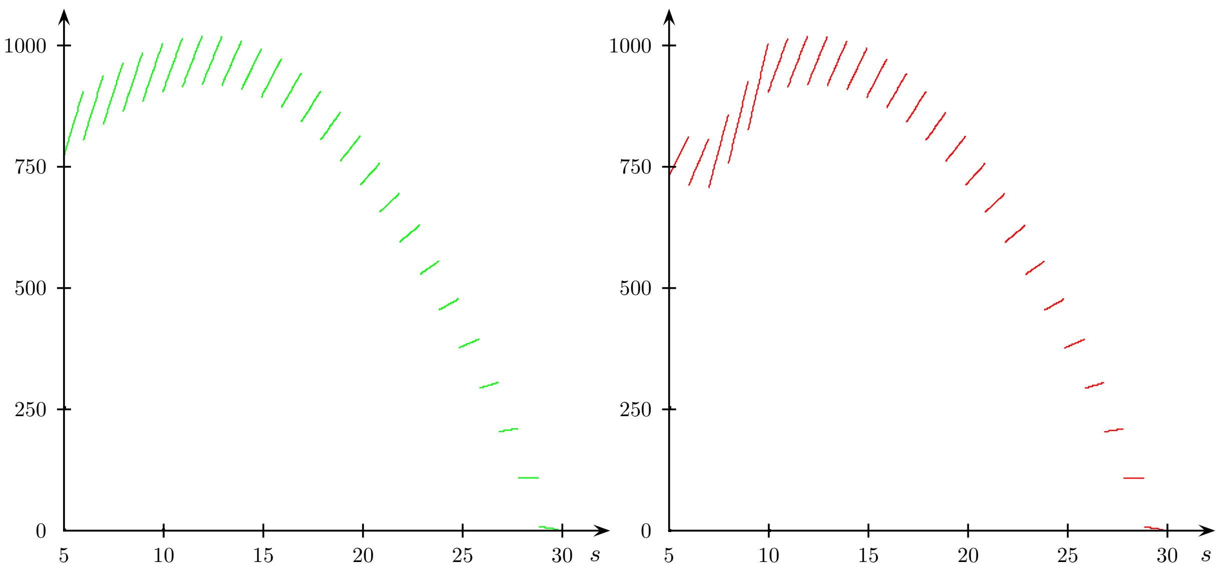

Figure 25a: Comparison of prospective reserves in state *invalid* for male insured with age at issue $x = 20$; $V^+_{(i)}(s)$ (left) is the prospective reserve within a Markov model, and $V^+_{(i,s-5)}(s)$ (right) is the prospective reserve relying on the corresponding semi-Markov model

Figure 25a compares the prospective reserve for the Markov approach with the prospective reserve relying on the semi-Markov model with previous duration $u = 0$ at time $s = 5$. This corresponds to the situation in which the policyholder, who is at the age of 25 at time $s = 5$, became disabled at time $s = 5$. Then, for the first years of disablement, the prospective reserve is smaller within the semi-Markov framework than within the Markov framework. This is caused by higher rates for both recovery and mortality for disabled insured with a short previous duration of disablement, when compared with the corresponding ultimate tables. The latter quantify the probabilities of recovery or death of disabled insured for whom the onset of disability dates back more than five years.

Figure 25b compares the prospective reserve for the Markov approach with the prospective reserve relying on the semi-Markov model with previous duration $u = 2$ at time $s = 5$. This corresponds to the situation in which the policyholder became disabled at the age 23, such that at time $s = 5$, the onset of disability dates back two years. Finally, figure 25c compares the prospective reserve for the Markov approach with the prospective reserve relying on the semi-Markov model with previous duration $u = 5$ at time $s = 5$. In this situation, the insured became disabled right after the issue of the policy, such that at time $s = 5$, the insured has been invalid for five years. Due the select period of $r = 5$, the duration-dependence has no more impact on the results in cases where the corresponding event dates back five years or more. Hence, the prospective reserves in figure 25c coincide at each time.

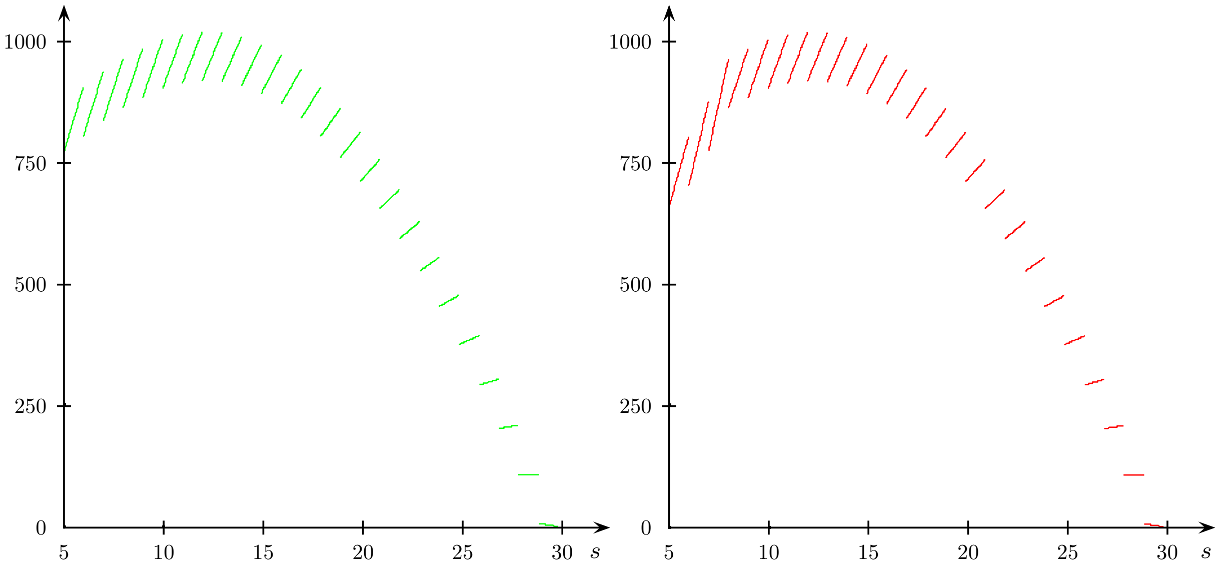

Figure 25b: Comparison of prospective reserves in state *invalid* for male insured with age at issue $x = 20$; $V^+_{(i)}(s)$ (left) is the prospective reserve within a Markov model, and $V^+_{(i,s-3)}(s)$ (right) is the prospective reserve relying on the corresponding semi-Markov model

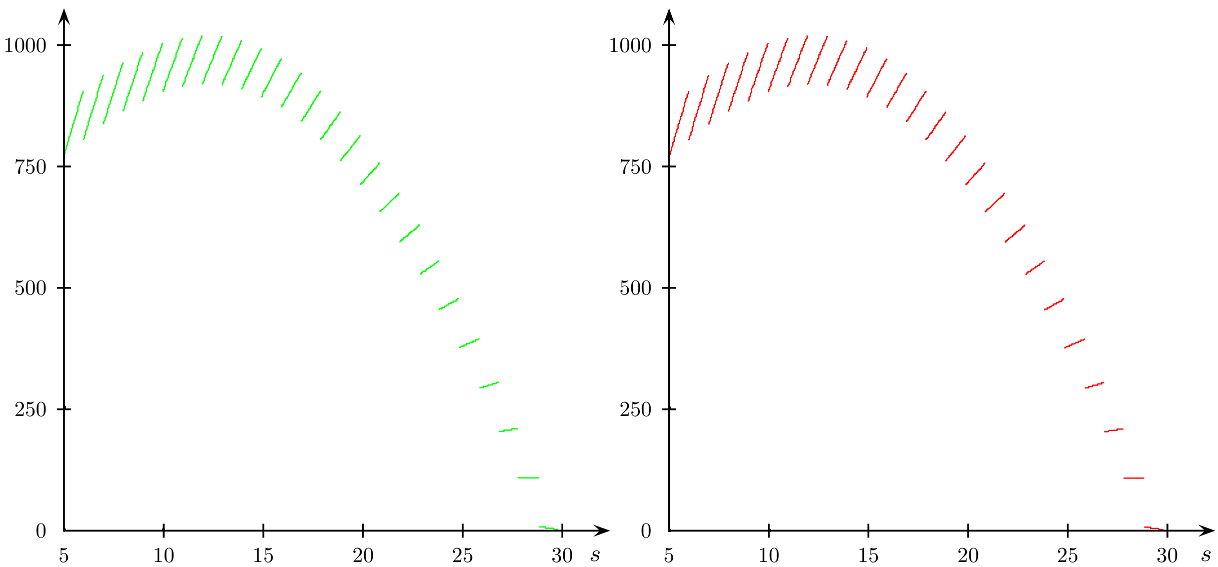

Figure 25c: Comparison of prospective reserves in state *invalid* for male insured with age at issue $x = 20$; $V^+_{(i)}(s)$ (left) is the prospective reserve within a Markov model, and $V^+_{(i,s)}(s)$ (right) is the prospective reserve relying on the corresponding semi-Markov model

In summary one can say that for the impact of durational effects concerning a single policy three things are important. Firstly, it must be clarified as to what kind of effect is caused by allowing the annual rates to depend additionally on the time elapsed since the current state was entered. Do the probabilities of remaining in this state increase, or do they decrease by taking into account the time elapsed since entering the current state? Secondly, it is important what payments are intended for the state considered. Do the actuarial present values of both premiums and benefits increase or decrease with respect to the changes of the probabilities of remaining in the state under consideration. Thirdly, it is obvious that the length of the select period r, compared with the policy term n, does significantly intensify the impact of durational effects. △

C Martingale representation of the prospective loss

We start again by considering a single policy (p) modelled by a semi-Markovian pure jump process $(X_t)_{t\geq 0}$ with finite state space $\mathcal{S}$ and transition space $\mathcal{J}$. $((T_m, Z_m))_{m\in\mathbb{N}_0}$ is the appertaining homogeneous Markovian marked point process with regular cumulative transition intensity matrix $\hat{q}$. $(\mathbf{N}_t)_{t\geq 0}$ is the appertaining multivariate counting process, and $((X_t, U_t))_{t\geq 0}$ the appertaining bivariate Markov process with regular cumulative transition intensity matrix q. From now on, the cumulative transition intensity matrix q will - according to (2. 28. 3) - always be specified by means of $\hat{q}$. The corresponding conditional distributions Q_* as well as the conditional probabilities of remaining $\bar{p}_*$ can then be determined with the aid of (2. 28. 4), (2. 28. 5), and (2. 24. 7). Regarding the cumulative annuity payment rates F_*, we turn back to the notation by means of $\hat{F}_*$ (cf. (4. 4. 1)), and assume

$$\mathbf{A4}: \quad \sum_{z\in\mathcal{S}} \int_{[s,\infty)} v(\tau)\,|\hat{F}_z|(s, d\tau) < \infty, \quad s \geq 0. \tag{4. 14. 22}$$

Thus, Thiele's integral equations of type 1 can for $0 \leq u \leq s$ and $y \in \mathcal{S}$ be represented as

$$\begin{aligned} V^+_{(y,u)}(s) \;=\; & K(s) \int_{[s,\infty)} v(\tau)\,\bar{p}_y(s,\tau,u)\,\hat{F}_y(s-u, d\tau) \\ & + K(s) \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} \Big(\frac{D_{yz}(\tau, \tau-s+u)}{K(DT(\tau))} + v(\tau)\,V^+_{(z,0)}(\tau) \Big)\, \bar{p}_y(s, \tau-0, u)\, \hat{q}_{yz}(s-u, d\tau). \end{aligned} \tag{4. 14. 23}$$

As customary, K denotes a capital accumulation function, and $v = \frac{1}{K}$ the appertaining discounting function.

We start by defining the prospective loss. For this, the reader may recall the expressions $V^-(t, A)$, $V^+(t, A)$, and V_0 according to definition 3. 12, with A being an actuarial payment function according to definition 3. 9. Then, for a filtration $(\mathfrak{F}_t)_{t\geq 0}$ such that the process $((X_t, U_t))_{t\geq 0}$ is adapted to $(\mathfrak{F}_t)_{t\geq 0}$, the prospective reserve according to definition 4. 1 is generally given as

$$V^+(t) = \mathbf{E}[V^+(t, A)|\,\mathfrak{F}_t],\ t \geq 0.$$

4. 15 Definition. The *prospective loss of an insurer up to time $t \geq 0$ due to an actuarial payment function A for a policy* (p), is defined as

$$\mathcal{L}(t) := -v(t)\,V^-(t, A) + v(t)\,V^+(t) - \mathbf{E}[V_0].$$

Note that in cases where the principle of equivalence is assumed, i.e. $\mathbf{E}[V_0] = 0$, the prospective loss is equal to

$$\mathcal{L}(t) = -v(t)\,V^-(t, A) + v(t)\,V^+(t), \quad t \geq 0,$$

which corresponds to the present value of all payments up to time t plus the present value of the prospective reserve at time t.

4.16 Lemma. *Assume* (3.10.1) *and* (4.7.1) *yielding* $\mathbf{E}[|V_0|] < \infty$. *Then, the prospective loss* $(\mathcal{L}(t))_{t\geq 0}$ *is a centered uniformly integrable martingale with respect to the filtration* $(\mathfrak{F}_t)_{t\geq 0}$.

Proof. For each $t \geq 0$, both $V^-(t, A)$ and $V^+(t)$ are $\mathfrak{F}_t$-measurable. Hence,

$$\mathcal{L}(t) \;=\; \mathbf{E}[-v(t)\,V^-(t,A)|\,\mathfrak{F}_t] + \mathbf{E}[v(t)\,V^+(t)|\,\mathfrak{F}_t] - \mathbf{E}[V_0] = \mathbf{E}[V_0|\,\mathfrak{F}_t] - \mathbf{E}[V_0]. \quad (4.16.1)$$

Thus, the prospective loss $(\mathcal{L}(t))_{t\geq 0}$ is adapted to the filtration $(\mathfrak{F}_t)_{t\geq 0}$. Further, the expectation of $\mathcal{L}(t)$ is given by $\mathbf{E}[\mathcal{L}(t)] = \mathbf{E}[V_0] - \mathbf{E}[V_0]$, $t \geq 0$, with $\mathbf{E}[V_0] < \infty$. Hence, $(\mathcal{L}(t))_{t\geq 0}$ is integrable with zero mean. It remains to verify the martingale property $\mathbf{E}[\mathcal{L}(t)|\,\mathfrak{F}_s] = \mathcal{L}(s),\ s \leq t$. This obviously holds due to the smoothing property of conditional expectations, i.e. $\mathbf{E}[\mathbf{E}[V_0|\,\mathfrak{F}_t]|\,\mathfrak{F}_s] = \mathbf{E}[V_0|\,\mathfrak{F}_s]$ for $\mathfrak{F}_s \subseteq \mathfrak{F}_t$. That $(\mathcal{L}(t))_{t\geq 0}$ is even a uniformly integrable martingale also follows from the representation (4.16.1). According to this, $(\mathcal{L}(t))_{t\geq 0}$ is a so-called Doob martingale (cf. Klebaner [1998], theorem 7.8) for which the uniform integrability holds in general. □

Note that the above lemma does not employ specific properties of the actuarial payment function or the underlying process modelling a single risk. It is only required that both are adapted to the corresponding filtration.

Theorem 4.18 states that the prospective loss possesses a certain integral representation. The corresponding proof is closely related to the establishment of the integral representation of point process martingales in Brémaud ([1981], theorem T9), and likewise to the proof of theorem 3.1 in Møller [1993]. Møller adapted the ideas of Brémaud to the prospective loss due to a single policy, which can also be considered as a point process martingale. However, both authors assumed the existence of intensities. Regarding Møller [1993], this means that the cumulative transition intensities, the cumulative annuity payment rates, and the discounting function are assumed to be absolutely continuous with respect to the Lebesgue measure λ^1. In order to prove the corresponding integral representation for the prospective loss $(\mathcal{L}(t))_{t\geq 0}$ within the present more general set-up, a bit more is involved. Particularly, Thiele's integral equations of type 1 must be employed. In order to do this without problems due to null sets, it must be granted that there are versions of prospective reserves which satisfy Thiele's integral equations of type 1 without exceptional sets.

4.17 Lemma. *Let* $((T_m, Z_m))_{m\in\mathbb{N}_0}$ *be a homogeneous Markovian marked point process with regular cumulative transition intensity matrix* $\hat{q}$, *and* $((X_t, U_t))_{t\geq 0}$ *the appertaining bivariate Markov process with regular cumulative transition intensity matrix* q *specified by means of* $\hat{q}$. *Then, there are versions of the prospective reserves* $V^+_{(y,u)}(s), 0 \leq u \leq s, y \in \mathcal{S}$, *that satisfy Thiele's integral equations of type 1 without exceptional sets. Such versions can be specified by the right-hand side of* (4.3.1).

Proof. The proof of Thiele's integral equation of type 1 (cf. theorem 4.8) basically relies on the fact that the right-hand side of (4.3.1) provides versions of the prospective reserves $V^+_{(y,u)}(s)$, $0 \leq u \leq s, y \in \mathcal{S}$. Along with the specification of q, Q_*, and $\bar{p}_*$ by means of $\hat{q}$, it follows by reproducing the proof of theorem 4.8 (see A.13) that the prospective reserves specified by the right-hand side of (4.3.1) satisfy Thiele's integral equations of type 1 without exceptional sets. □

In the following, we will always specify versions of the prospective reserves $V^+_{(y,u)}(s)$, $0 \leq u \leq s, y \in \mathcal{S}$, by the right-hand side of (4.3.1). In a certain sense, this corresponds to the specification of a set of transition probabilities $p_{yz}(s,t,u,v)$, $0 \leq u \leq s \leq t, v \geq 0, (y,z) \in \mathcal{S}^2$, for the bivariate Markov process $((X_t, U_t))_{t\geq 0}$ by means of (2.23.1) and (2.23.2). Recall that this set of transition probabilities satisfies the backward integral equations (2.32.2) identically. Conversely, inserting these transition probabilities into (4.6.2) yields, under assumptions 2.35

and 4. 5, the above versions of the prospective reserves which satisfy Thiele's integral equations of type 1 identically.

We now turn to the previously mentioned integral representation of the prospective loss.

4. 18 Theorem. *Let* $(X_t, \mathfrak{F}_t)_{t\geq 0}$ *be a semi-Markov process and* $((T_m, Z_m))_{m\in\mathbb{N}_0}$ *the appertaining homogeneous Markovian marked point process with regular cumulative transition intensity matrix* $\hat{q}$. $(\mathbf{N}_t)_{t\geq 0}$ *denotes the associated multivariate counting process, and* $((X_t, U_t))_{t\geq 0}$ *the appertaining bivariate Markov process with regular cumulative transition intensity matrix* q *specified by* $\hat{q}$. *Further, for* $(y,z) \in \mathcal{J}$, *let* R_{yz} *be the amount of risk according to definition* 4. 10, *and* M_{yz} *the innovation process according to* (2. 17. 2). *Then, under the assumptions* (2. 14. 1), (3. 10. 1), (4. 7. 1), *and for* $t \geq 0$

$$\mathbf{A5}: \ \mathbf{E}\Big[\sum_{(y,z)\in\mathcal{J}}\sum_{m\in\mathbb{N}_0}\int_{(T_m\wedge t, T_{m+1}\wedge t]} v(\tau)\,|V^+_{(z,0)}(\tau) - V^+_{(y,\tau-T_m)}(\tau)|\ \hat{q}_{yz}(T_m, d\tau)\Big] < \infty, \quad (4.\,18.\,1)$$

the prospective loss $(\mathcal{L}(t))_{t\geq 0}$ *satisfies* P*-a.s. the following representation*

$$\mathcal{L}(t) \ = \ \sum_{(y,z)\in\mathcal{J}}\sum_{m\in\mathbb{N}_0}\int_{(T_m\wedge t, T_{m+1}\wedge t]} v(\tau)\,R_{yz}(\tau, \tau - T_m)\,M_{yz,d\tau}, \ \ t \geq 0. \quad (4.\,18.\,2)$$

Proof. For the proof of this theorem we refer to A.15 in the appendix. □

Assumption (4. 18. 1) ensures, along with (4. 7. 1), that the right-hand side of (4. 18. 2) is P-a.s. well defined. With the aid of (2. 5. 9), the right-hand side of (4. 18. 2) can be rewritten as

$$\sum_{(y,z)\in\mathcal{J}}\int_{(0,t]} v(\tau)\,R_{yz}(\tau, U_{\tau-0})\,M_{yz,d\tau}, \quad t \geq 0. \quad (4.\,18.\,3)$$

Generally, both (4. 18. 2) and (4. 18. 3) form under the assumptions stipulated above stochastic integrals of predictable processes with respect to square integrable martingales. These representations of the prospective loss turn out to be appropriate tools for further investigations. In order to do this, we start by discussing some consequences of the integral representation of the prospective loss. Firstly, it follows from theorem 4. 18 that the prospective loss $(\mathcal{L}(t))_{t\geq 0}$, as well as the loss in a certain state $y \in \mathcal{S}$, $(\mathcal{L}_y(t))_{t\geq 0}$, are square integrable martingales, provided that the following is also satisfied:

4. 19 Assumption.

$$\mathbf{A6}: \quad \mathbf{E}\Big[\sum_{(y,z)\in\mathcal{J}}\sum_{m\in\mathbb{N}_0}\int_{(T_m\wedge t, T_{m+1}\wedge t]} \big(v(\tau)\,R_{yz}(\tau, \tau - T_m)\big)^2\ \hat{q}_{yz}(T_m, d\tau)\Big] < \infty, \ \ t \geq 0. \quad (4.\,19.\,1)$$

For $y \in \mathcal{S}$, the loss in state y is here defined as

$$\mathcal{L}_y(t) \ := \ \sum_{\substack{z\in\mathcal{S}\\ z\neq y}}\sum_{m\in\mathbb{N}_0}\int_{(T_m\wedge t, T_{m+1}\wedge t]} v(\tau)\,R_{yz}(\tau, \tau - T_m)\,M_{yz,d\tau}, \ \ t \geq 0. \quad (4.\,19.\,2)$$

The losses in state obviously satisfy P-almost surely

$$\mathcal{L}(t) = \sum_{y\in\mathcal{S}}\mathcal{L}_y(t), \ \ t \geq 0. \quad (4.\,19.\,3)$$

4.20 Corollary. *Assume the situation of theorem* 4.18. *Further, assume* (4.19.1). *Then,* $(\mathcal{L}(t))_{t\geq 0}$ *and* $(\mathcal{L}_y(t))_{t\geq 0}$, $y \in \mathcal{S}$, *are square integrable martingales.*

Proof. In view of the representations (4.18.2) and (4.19.2), the assertions follow immediately from theorem A.5. Hence, it remains to verify the assumptions of this theorem. Let $(y,z) \in \mathcal{J}$. Under assumption (2.14.1), the innovation process M_{yz} according to (2.17.2) is a square integrable martingale (cf. Milbrodt and Helbig [1999], theorem 12.27). The $(\mathfrak{F}_t)_{t\geq 0}$-predictability of the corresponding integrands, which are in both cases given by the process

$$\Omega \times [0,\infty) \ni (\omega, t) \mapsto \sum_{m\in\mathbb{N}_0} v(t)\, R_{yz}(t, t - T_m(\omega))\, \mathbf{1}_{\{T_m(\omega)<t\leq T_{m+1}(\omega)\}},$$

is granted by the last assertion of theorem 2.5. Finally, we obtain with the aid of lemma 2.18 and (4.19.1) the following

$$\begin{aligned}
&\mathbf{E}\Big[\sum_{m\in\mathbb{N}_0}\int_{(T_m\wedge t, T_{m+1}\wedge t]} \big(v(\tau)\, R_{yz}(\tau, \tau - T_m)\big)^2 \langle M_{yz}\rangle_{d\tau}\Big] \\
&= \mathbf{E}\Big[\sum_{m\in\mathbb{N}_0}\int_{(T_m\wedge t, T_{m+1}\wedge t]} \big(v(\tau)\, R_{yz}(\tau, \tau - T_m)\big)^2\, \mathbf{1}_{\{Z_m=y\}}\, \hat{q}_{yz}(T_m, d\tau)\Big] \\
&\quad -\mathbf{E}\Big[\sum_{m\in\mathbb{N}_0}\int_{(T_m\wedge t, T_{m+1}\wedge t]} \big(v(\tau)\, R_{yz}(\tau, \tau - T_m)\big)^2\, \mathbf{1}_{\{Z_m=y\}}\, \hat{q}_{yz}(T_m, \{\tau\})\;\; \hat{q}_{yz}(T_m, d\tau)\Big] \\
&\leq \mathbf{E}\Big[\sum_{m\in\mathbb{N}_0}\int_{(T_m\wedge t, T_{m+1}\wedge t]} \big(v(\tau)\, R_{yz}(\tau, \tau - T_m)\big)^2\, \mathbf{1}_{\{Z_m=y\}}\, \hat{q}_{yz}(T_m, d\tau)\Big] \\
&\leq \mathbf{E}\Big[\sum_{m\in\mathbb{N}_0}\int_{(T_m\wedge t, T_{m+1}\wedge t]} \big(v(\tau)\, R_{yz}(\tau, \tau - T_m)\big)^2\, \hat{q}_{yz}(T_m, d\tau)\Big] \\
&< \infty.
\end{aligned}$$

Thus, the assumptions of theorem A.5 are satisfied and it follows that

$$\sum_{m\in\mathbb{N}_0}\int_{(T_m\wedge t, T_{m+1}\wedge t]} v(\tau)\, R_{yz}(\tau, \tau - T_m)\, M_{yz,d\tau}, \;\; t \geq 0, \tag{4.20.1}$$

is a square integrable martingale. The sum of square integrable martingales is also a square integrable martingale. Hence, it follows that both $(\mathcal{L}(t))_{t\geq 0}$ and $(\mathcal{L}_y(t))_{t\geq 0}$, $y \in \mathcal{S}$, are likewise square integrable martingales. □

Corollary 4.20 will later be used to prove Hattendorff's theorem for the present framework. Another consequence of theorem 4.18 is given by Thiele's integral equations of type 2, to which we turn next.

D Thiele's integral equations of type 2

Consider prospective reserves $V^+_{(y,u)}(s), 0 \leq u \leq s, y \in \mathcal{S}$, specified by the right-hand side of (4.3.1), and assume (4.14.22). Then, these prospective reserves satisfy the system of integral equations (4.14.23) without exceptional sets. Under stronger integrability conditions, another system of integral equations is satisfied. This system of integral equations is referred to as Thiele's integral equations of type 2. For Thiele's integral equations of type 2 in a Markov set-up, we refer to Milbrodt and Helbig ([1999], section 10C). Our systems of integral equations

(4. 22. 1) and (4. 22. 2) form generalizations of (10.21.3) and (10.18.5) in Milbrodt and Helbig [1999].

The following integrability conditions ensuring Thiele's integral equations of type 2 to be well defined rely on Thiele's integral equations of type 1. Thus, the duration-dependence of the prospective reserves according to either (4. 6. 1) or (4. 6. 2) can be neglected. Doing so, we assume that for $y \in \mathcal{S}$, the prospective reserves $V^{+}_{(y,0)}(s)$, $s \geq 0$, are bounded, with $\bar{V}_y(s)$, $s \geq 0$, being an upper bound satisfying $|V^{+}_{(y,0)}(s)| \leq \bar{V}_y(s)$, $s \geq 0$. By using the right-hand side of (4. 6. 1) - which identically corresponds to a prospective reserve specified by the right-hand side of of (4. 3. 1) - such an upper bound for $V^{+}_{(y,0)}(s)$, $s \geq 0$, can be obtained by taking into account the specific given context. Thus, for example, by employing the restrictions concerning the policy term (usually finite), the eventual satisfaction of the assumptions 2. 35 and 4. 5 (cf. (2. 15. 3)), and the specific actuarial payment function, an upper bound according to $\bar{V}_y(s)$, $s \geq 0$, can be specified. By using this, we state the following integrability conditions.

4. 21 Assumption. Assume for all $s \geq 0$

$$\begin{aligned} \mathbf{A7}: \quad & \sum_{z \in \mathcal{S}} \int_{[s,\infty)} |\hat{F}_z|(s, d\tau) < \infty, \\ & \sum_{(y,z) \in \mathcal{J}} \int_{(s,\infty)} \Big(K(\tau)\, v(DT(\tau))\, D_{yz}(\tau, \tau - s) + \bar{V}_z(\tau) \Big)\, \hat{q}_{yz}(s, d\tau) < \infty, \end{aligned} \tag{4. 21. 1}$$

and

$$\begin{aligned} \mathbf{A8}: \quad & \sum_{(y,z) \in \mathcal{J}} \int_{[s,\infty)} \hat{q}_{yz}(s, (s,\tau])\, v(\tau)\, |\hat{F}_y|(s, d\tau) < \infty \\ & \sum_{y \in \mathcal{S}} \sum_{z,\xi \in \mathcal{S}\setminus\{y\}} \int_{(s,\infty)} \hat{q}_{yz}(s, (s,\tau]) \Big(v(DT(\tau))\, D_{y\xi}(\tau, \tau - s) + v(\tau)\, \bar{V}_\xi(\tau) \Big)\, \hat{q}_{y\xi}(s, d\tau) < \infty. \end{aligned} \tag{4. 21. 2}$$

A stronger condition than **A8** is given by the following: for $s \geq 0$ assume

$$\begin{aligned} \mathbf{A9}: \quad & \sum_{(y,z) \in \mathcal{J}} \int_{[s,\infty)} \hat{q}_{yz}(s, (s,\tau])\, |\hat{F}_y|(s, d\tau) < \infty \\ & \sum_{y \in \mathcal{S}} \sum_{z,\xi \in \mathcal{S}\setminus\{y\}} \int_{(s,\infty)} \hat{q}_{yz}(s, (s,\tau]) \Big(K(\tau)\, v(DT(\tau))\, D_{y\xi}(\tau, \tau - s) + \bar{V}_\xi(\tau) \Big)\, \hat{q}_{y\xi}(s, d\tau) < \infty. \end{aligned} \tag{4. 21. 3}$$

Under assumption (4. 21. 1), the expressions $v(s)\, V^{+}_{(y,u)}(s), 0 \leq u \leq s, y \in \mathcal{S}$, are bounded, since according to (4. 14. 23)

$$\begin{aligned} & v(s)\, V^{+}_{(y,u)}(s) \\ & \leq \int_{[s,\infty)} |\hat{F}_y|(s-u, d\tau) + \sum_{\substack{z \in \mathcal{S} \\ z \neq y}} \int_{(s,\infty)} \Big(K(\tau) \frac{D_{yz}(\tau, \tau - s + u)}{K(DT(\tau))} + |V^{+}_{(z,0)}(\tau)| \Big)\, \hat{q}_{yz}(s-u, d\tau) \\ & \leq \int_{[s-u,\infty)} |\hat{F}_y|(s-u, d\tau) + \sum_{\substack{z \in \mathcal{S} \\ z \neq y}} \int_{(s-u,\infty)} \Big(K(\tau) \frac{D_{yz}(\tau, \tau - s + u)}{K(DT(\tau))} + \bar{V}_z(\tau) \Big)\, \hat{q}_{yz}(s-u, d\tau). \end{aligned}$$

Further, one obtains

$$\lim_{t\to\infty} v(s+t)\, V^+_{(y,u+t)}(s+t) = 0. \tag{4.21.4}$$

The following consequence of theorem 4.18 presents Thiele's integral equations of type 2.

4.22 Corollary. *Assume the situation of theorem* 4.18. *Further, assume* (4.21.1) *and* (4.21.2). *Then, the versions of the prospective reserves* $V^+_{(y,u)}(s), 0 \le u \le s, y \in \mathcal{S}$, *specified by the right-hand side of* (4.3.1) *satisfy*

$$v(s)\, V^+_{(y,u)}(s) \;=\; \int_{[s,\infty)} v(\tau)\, \hat{F}_y(s-u, d\tau) + \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} v(\tau)\, R_{yz}(\tau, \tau-s+u)\; \hat{q}_{yz}(s-u, d\tau). \tag{4.22.1}$$

Further, under assumption (4.21.3), *the above prospective reserves likewise satisfy*

$$\begin{aligned} V^+_{(y,u)}(s) \;&=\; \hat{F}_y(s-u, [s,\infty)) + \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} R_{yz}(\tau, \tau-s+u)\; \hat{q}_{yz}(s-u, d\tau) \\ &\quad - \int_{(s,\infty)} V^+_{(y,\tau-s+u)}(\tau-0)\, \Phi(d\tau), \end{aligned} \tag{4.22.2}$$

which is referred to as Thiele's integral equations of type 2.

Proof. (4.22.1) is a consequence of the integral representation for the prospective loss (4.18.2). (4.22.2) follows from (4.22.1) by employing (3.2.3). For a detailed proof of (4.22.1) and (4.22.2), we refer to A.16 in the appendix. □

Assumption (4.21.1) ensures that the last addend of (4.22.2) is finite. This can be verified by employing (4.14.23), (3.2.1), and Fubini's theorem: For $s \ge 0$ and $y \in \mathcal{S}$ one obtains

$$\begin{aligned} &\int_{(s,\infty)} V^+_{(y,r-s+u)}(r-0)\, \Phi(dr) \\ &\le \int_{(s,\infty)} \int_{[r,\infty)} v(\tau)\, |\hat{F}_y|(s-u, d\tau)\, K(r-0)\, \Phi(dr) \\ &\quad + \int_{(s,\infty)} \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{[r,\infty)} \Big(\frac{D_{yz}(\tau, \tau-s+u)}{K(DT(\tau))} + v(\tau)\, |V^+_{(z,0)}(\tau)| \Big)\, \hat{q}_{yz}(s-u, d\tau)\, K(r-0)\, \Phi(dr) \\ &= \int_{(s,\infty)} \int_{(s,\tau]} K(r-0)\, \Phi(dr)\, v(\tau)\, |\hat{F}_y|(s-u, d\tau) \\ &\quad + \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} \int_{(s,\tau]} K(r-0)\, \Phi(dr)\, v(\tau) \Big(K(\tau)\, \frac{D_{yz}(\tau, \tau-s+u)}{K(DT(\tau))} + |V^+_{(z,0)}(\tau)| \Big)\, \hat{q}_{yz}(s-u, d\tau) \\ &\le \int_{[s-u,\infty)} |\hat{F}_y|(s-u, d\tau) \\ &\quad + \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s-u,\infty)} \Big(K(\tau)\, \frac{D_{yz}(\tau, \tau-s+u)}{K(DT(\tau))} + \bar{V}_z(\tau) \Big)\, \hat{q}_{yz}(s-u, d\tau). \end{aligned}$$

Assumption (4.21.2) ensures for $0 \le s \le u$ and $(y,z) \in \mathcal{J}$ that $v(\cdot)\, V^+_{(y,\cdot-s+u)}(\cdot)$ is integrable with respect to $\hat{q}_{yz}(s-u, \cdot)$. This implies, along with (4.21.1), that the right-hand side of

(4.22.1) is finite. In almost the same manner, (4.21.3) ensures that $V^+_{(y,\cdot-s+u)}(\cdot)$ is integrable with respect to $\hat{q}_{yz}(s-u,\cdot)$. By using Fubini's theorem, this can be established as follows

$$\begin{aligned}
&\int_{(s,\infty)} V^+_{(y,r-s+u)}(r)\,\hat{q}_{yz}(s-u,dr)\\
&\le \int_{(s,\infty)} K(r)\int_{[r,\infty)} v(\tau)\,|\hat{F}_y|(s-u,d\tau)\,\hat{q}_{yz}(s-u,dr)\\
&\quad+\int_{(s,\infty)} K(r)\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(r,\infty)}\Big(\frac{D_{y\xi}(\tau,\tau-s+u)}{K(DT(\tau))}+v(\tau)\,|V^+_{(\xi,0)}(\tau)|\Big)\,\hat{q}_{y\xi}(s-u,d\tau)\,\hat{q}_{yz}(s-u,dr)\\
&=\int_{(s,\infty)}\int_{(s,\tau]} K(r)\,\hat{q}_{yz}(s-u,dr)\,v(\tau)\,|\hat{F}_y|(s-u,d\tau)\\
&\quad+\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,\infty)}\int_{(s,\tau)} K(r)\,\hat{q}_{yz}(s-u,dr)\Big(\frac{D_{y\xi}(\tau,\tau-s+u)}{K(DT(\tau))}+v(\tau)\,|V^+_{(\xi,0)}(\tau)|\Big)\,\hat{q}_{y\xi}(s-u,d\tau)\\
&\le\int_{(s,\infty)} \hat{q}_{yz}(s-u,(s,\tau])\,|\hat{F}_y|(s-u,d\tau)\\
&\quad+\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,\infty)} \hat{q}_{yz}(s-u,(s,\tau])\Big(K(\tau)\,\frac{D_{y\xi}(\tau,\tau-s+u)}{K(DT(\tau))}+|V^+_{(\xi,0)}(\tau)|\Big)\,\hat{q}_{y\xi}(s-u,d\tau)\\
&\le\int_{[s-u,\infty)} \hat{q}_{yz}(s-u,(s-u,\tau])\,|\hat{F}_y|(s-u,d\tau)\\
&\quad+\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s-u,\infty)} \hat{q}_{yz}(s-u,(s-u,\tau])\Big(K(\tau)\,\frac{D_{y\xi}(\tau,\tau-s+u)}{K(DT(\tau))}+\bar{V}_z(\tau)\Big)\,\hat{q}_{y\xi}(s-u,d\tau).
\end{aligned}$$

According to this, it follows with (4.21.1) that all integrals on the right-hand side of (4.22.2) are finite.

As counterpart to lemma 4.9, the following lemma is concerned with the relation between Thiele's integral equations of type 2 and the backward integral equations of type 2.

4.23 Lemma. *Thiele's integral equations of type 2 according to* (4.22.2) *are generalizations of the backward integral equations of type 2* (2.40.1)*. Conversely, under the assumptions* 2.35, 4.5, (4.21.1) *and* (4.21.3)*, the backward integral equations* (2.40.1) *imply Thiele's integral equations* (4.22.2)*.*

Proof. In order to prove the first assertion, we use the same parameters as in the proof of lemma 4.9. By inserting these parameters into (4.6.1), we obtain

$$V^+_{(y,u)}(s) \quad = \quad p_{yz}(s,t,u,v)\,\mathbf{1}_{(t\ge s)},$$

which holds without exceptional sets for versions of prospective reserves $V^+_{(y,u)}(s), 0\le u\le s, y\in\mathcal{S}$, specified by the right-hand side of (4.3.1). By inserting this into the right-hand side of (4.22.2), and afterwards applying (2.28.3) as well as (2.27.5), we get

$$\begin{aligned}
&p_{yz}(s,t,u,v)\,\mathbf{1}_{(t\ge s)}\\
&=\mathbf{1}_{(v\ge t-s+u)}\,\delta_{yz}\,\mathbf{1}_{(t\ge s)}\\
&\quad+\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,\infty)}\Big(p_{\xi z}(\tau,t,0,v)\,\mathbf{1}_{(t\ge\tau)}-p_{yz}(\tau,t,\tau-s+u,v)\,\mathbf{1}_{(t\ge\tau)}\Big)\,\hat{q}_{y\xi}(s-u,d\tau)\\
&=
\end{aligned}$$

$$= \mathbf{1}_{(v \geq t-s+u)}\, \delta_{yz}\, \mathbf{1}_{(t \geq s)} + \sum_{\substack{\xi \in \mathcal{S} \\ \xi \neq y}} \int_{(s,t]} p_{\xi z}(\tau, t, 0, v)\; q_{y\xi}(s, d\tau, u)$$

$$+ \int_{(s,t]} p_{yz}(\tau, t, \tau - s + u, v)\, q_{yy}(s, d\tau, u),$$

which corresponds for $s \leq t$ to the backward integral equations (2. 40. 1).
That the backward integral equations (2. 40. 1) imply Thiele's integral equations (4. 22. 2) can under the assumptions stipulated above be verified by inserting the backward integral equations (2. 40. 1) into (4. 6. 2). This yields (4. 22. 1), provided that the cumulative transition intensity matrix q is specified by $\hat{q}$. Thiele's integral equations (4. 22. 2) can then be obtained by following the same argumentation as for proving (4. 22. 2). For more details, confer A.17. □

So far, we have seen that the versions of the prospective reserves which are specified by the right-hand side of (4. 3. 1), and hence identically satisfy Thiele's integral equations of type 1, also satisfy Thiele's integral equations of type 2, provided that the corresponding integrability conditions are fulfilled. The next lemma states that for versions of the prospective reserves which satisfy Thiele's integral equations of type 2 identically, Thiele's integral equations of type 1 also hold. Hence, under the assumptions (4. 21. 1) and (4. 21. 3), both systems of integral equations are equivalent.

As we will see in the next section, both systems of integral equations uniquely determine the prospective reserves. Thus, there is only one set of prospective reserves which fulfills both Thiele's integral equations of type 1 and Thiele's integral equations of type 2 identically. Since both backward integral equations can be obtained by either inserting specific parameters into Thiele's integral equations of type 1 or by inserting the same parameters into Thiele's integral equations of type 2, the above relationships also hold for them. In principle, all these relationships were established by Milbrodt and Helbig [1999] in a non-smooth Markov set-up. So it turns out that they remain valid in a non-smooth semi-Markov set-up.

Further, we have seen by establishing Thiele's integral equations of type 2 that the integral representation of the prospective loss (4. 18. 2) implies these integral equations. Conversely, it can also be shown that Thiele's integral equations of type 2 imply the integral representation of the prospective loss. In order to prove this, the corresponding result within a Markov set-up established by Milbrodt and Helbig ([1999], Hilfssatz 10.36) can be reproduced almost literally. The starting point for this is given by (4. 22. 1), which must then be understood as consequence of Thiele's integral equation of type 2.

We now state that Thiele's integral equations of type 2 imply Thiele's integral equations of type 1. For this, lemma 3. 8 will be used.

4. 24 Lemma. *Assume* (4. 21. 1) *as well as* (4. 21. 3)*, and let* $V^{+}_{(y,u)}(s), 0 \leq s \leq u, y \in \mathcal{S}$*, be versions of the prospective reserves that satisfy Thiele's integral equations of type 2 identically. Then, they also satisfy Thiele's integral equations of type 1 identically.*

Proof. For prospective reserves $V^{+}_{(y,u)}(s), 0 \leq s \leq u, y \in \mathcal{S}$, that satisfy (4. 22. 2) identically, Thiele's integral equations of type 1 can then be obtained as follows. Apply (3. 8. 1) to $Z(d\tau) = \hat{F}_y(s-u, d\tau)$, and (3. 8. 2) to $Z(d\tau) = V^{+}_{(y,\tau-s+u)}(\tau - 0)\, \Phi(d\tau)$ as well as to $Z(d\tau) = R_{yz}(\tau, \tau - s + u)\, \hat{q}_{yz}(s-u, d\tau)$. Afterwards, (4. 22. 2) and definition 4. 10 must be applied twice. This yields Thiele's integral equations of type 1 which are here given by (4. 14. 23). For a detailed proof, we refer to A.18 in the appendix. □

E Solvability of integral equations for the prospective reserve

This section is concerned with establishing the uniqueness of solutions of Thiele's integral equations of type 1 and Thiele's integral equations of type 2. Recall that the regular cumulative transition intensity matrix q for the bivariate Markov process $((X_t, U_t))_{t\geq 0}$ is specified by the regular cumulative transition intensity matrix $\hat{q}$ for the marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$. Further, regarding Thiele's integral equations of type 1, the assumption (4. 14. 22) must be satisfied. Regarding to Thiele's integral equations of type 2, (4. 21. 1) and (4. 21. 3) must hold. According to the relationship between both types of integral equations - prospective reserves which identically satisfy the former also satisfy the latter identically (theorem 4. 18 and corollary 4. 22), and conversely, prospective reserves which satisfy the latter identically, also satisfy the former identically (lemma 4. 24) - the unique solvability of Thiele's integral equations of type 2 follows from the unique solvability of Thiele's integral equations of type 1. Hence, it is sufficient to investigate (4. 14. 23). In the following, we will verify that this system of integral equations uniquely determines prospective reserves $V^+_{(y,u)}(s)$, $0 \leq s \leq u$, $y \in \mathcal{S}$. In order to achieve this, we firstly disregard the duration-dependence by setting $u = 0$. Doing so, we consider the system of integral equations (cf. (4. 8. 4))

$$V^+_{(y,0)}(s) = K(s) \int_{[s,\infty)} v(\tau)\, \bar{p}_y(s,\tau,0)\, \hat{F}_y(s,d\tau) + K(s) \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} \Big(\frac{D_{yz}(\tau,\tau-s)}{K(DT(\tau))} + v(\tau)\, V^+_{(z,0)}(\tau) \Big)\, \bar{p}_y(s,\tau-0,0)\, \hat{q}_{yz}(s,d\tau). \tag{4. 24. 1}$$

For any $u \in [0,s]$, the prospective reserve $V^+_{(y,u)}(s)$ can then be derived by computing an integral according to the right-hand side of (4. 14. 23), the integrand of which only contains prospective reserves $V^+_{(z,0)}(\cdot)$, $z \neq y$. Hence, the prospective reserves $V^+_{(y,u)}(s)$, $0 \leq s \leq u$, $y \in \mathcal{S}$, are uniquely determined by $(V^+_{(y,0)}(\cdot))_{y\in\mathcal{S}}$. As we will see, the latter can be uniquely determined by solving the system of integral equations (4. 24. 1). Investigating the solvability of this system of linear integral equations is closely related to the proceedings in Stracke [1997] or Milbrodt and Helbig [1999], who establish the unique solvability of Thiele's integral equations in a non-smooth Markov set-up. Incidentally, Thiele's integral equations of type 2, which are for $u = 0$ of the form

$$V^+_{(y,0)}(s) = \hat{F}_y(s,[s,\infty)) - \int_{(s,\infty)} V^+_{(y,\tau-s)}(\tau-0)\, \Phi(d\tau) + \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} \Big(K(t)\, v(DT(t))\, D_{yz}(\tau,\tau-s) + V^+_{(z,0)}(\tau) - V^+_{(y,\tau-s)}(\tau) \Big)\, \hat{q}_{yz}(s,d\tau), \tag{4. 24. 2}$$

are not appropriate for determining $(V^+_{(y,0)}(\cdot))_{y\in\mathcal{S}}$, because the right-hand side also depends on prospective reserves with strictly positive previous durations.

4. 25 Theorem. *Let q be a regular cumulative transition intensity matrix according to* (2. 28. 3). *Further, let A be an actuarial payment function according to definition* 3. 9*, K a capital accumulation function, and v the appertaining discounting function. Then, under assumption* (4. 21. 1)*, the system of integral equations* (4. 14. 23) *uniquely determines prospective reserves $V^+_{(y,u)}(s), 0 \leq u \leq s, y \in \mathcal{S}$, which are bounded. Further, if* (4. 21. 3) *is additionally satisfied, these prospective reserves also form the one and only solution of Thiele's integral equations of type 2* (4. 22. 2).

Proof. Since the regular cumulative transition intensity matrix q is specified by $\hat{q}$, the prospective reserves determined by the right-hand side of (4. 3. 1) satisfy Thiele's integral equations

identically (lemma 4. 17). Hence, there is a solution of (4. 14. 23), which is bounded in case that the corresponding integrability conditions are fulfilled. Under the assumptions (4. 21. 1) and (4. 21. 3), this solution also satisfies Thiele's integral equations of type 2 identically (corollary 4. 22). Hence, there is also a solution of (4. 22. 2).
Regarding the uniqueness of these solutions, recall the above arguments. According to them, we only need to investigate the unique solvability of (4. 24. 1). The lemmata 4. 26 and 4. 27 provide this for actuarial payment functions that appertain to policies with finite terms. Using the result from lemma 4. 27, the unique solvability of the homogeneous system generally follows (lemma 4. 28). Hence, by using the same arguments as in the proof of lemma 4. 27, the unique solvability of (4. 24. 1) can be verified. □

4. 26 Lemma. *Let $\hat{q}$ be a regular cumulative transition intensity matrix according to definition 2. 8, and $n > 0$. Further, let K be a capital accumulation function, and v the appertaining discounting function. Then, the system of integral equations*

$$h_y(s) = K(s) \sum_{\substack{z \in \mathcal{S} \\ z \neq y}} \int_{(s,n]} v(\tau)\, h_z(\tau)\, \bar{p}_y(s, \tau - 0, 0)\, \hat{q}_{yz}(s, d\tau), \quad s \in [0, n],\, y \in \mathcal{S} \qquad (4.26.1)$$

has a unique Borel-measurable and bounded solution, namely $h_y \equiv 0, y \in \mathcal{S}$.

Proof. This proof reproduces the corresponding result in Stracke [1997] or Milbrodt and Helbig ([1999], Hilfsfssatz 10.25), but it takes into account the additional dependence of the cumulative transition intensities. In view of (4. 26. 1), $h_y \equiv 0$, $y \in \mathcal{S}$, is obviously a solution. We now demonstrate that this is the one and only solution. Obviously, for each $y \in \mathcal{S}$, h_y is right continuous and satisfies $h_y(n) = 0$. Assume $h_y \neq 0$ for any $y \in \mathcal{S}$. Further, let

$$t_0 := \min\{s \in [0, n] \,\big|\, h_y\big|_{[s,n]} \equiv 0, y \in \mathcal{S}\}.$$

Then, by assumption $t_0 > 0$. Now select an $s_0 \in (0, t_0)$ satisfying

$$c := \sup\Big\{ \sum_{(y,z) \in \mathcal{S}} \hat{q}_{yz}(s, (s, t_0)) \,\big|\, s \in [s_0, t_0) \Big\} < 1.$$

By using (4. 26. 1), this yields the following contradiction:

$$\begin{aligned}
0 &< \sup\big\{ |v(s)\, h_y(s)| \,\big|\, s \in [s_0, n], y \in \mathcal{S} \big\} \\
&\leq \sup\Big\{ \sum_{\substack{z \in \mathcal{S} \\ z \neq y}} \int_{(s_0, t_0)} |v(t)\, h_z(t)|\, \hat{q}_{yz}(s, dt) \,\big|\, s \in [s_0, t_0], y \in \mathcal{S} \Big\} \\
&= \sup\Big\{ \sum_{\substack{z \in \mathcal{S} \\ z \neq y}} \int_{(s, t_0)} |v(t)\, h_z(t)|\, \hat{q}_{yz}(s, dt) \,\big|\, s \in [s_0, t_0], y \in \mathcal{S} \Big\} \\
&\leq c \cdot \sup\big\{ |v(s)\, h_y(s)| \,\big|\, s \in [s_0, n], y \in \mathcal{S} \big\}.
\end{aligned}$$

□

4. 27 Lemma. *Let $\hat{q}$ be a regular cumulative transition intensity matrix according to definition 2. 8, A an actuarial payment function with finite term $n > 0$, and K a capital accumulation function with appertaining discounting function v. Further, assume (4. 21. 1). Then, for $0 \leq s \leq n$, and $y \in \mathcal{S}$, the following system of integral equations*

$$\begin{aligned}
V^+_{(y,0)}(s) &= K(s) \int_{[s,n]} v(\tau)\, \bar{p}_y(s, \tau, 0)\, \hat{F}_y(s, d\tau) \qquad (4.27.1) \\
&\quad + K(s) \sum_{\substack{z \in \mathcal{S} \\ z \neq y}} \int_{(s,n]} \Big(\frac{D_{yz}(\tau, \tau - s)}{K(DT(\tau))} + v(\tau)\, V^+_{(z,0)}(\tau) \Big)\, \bar{p}_y(s, \tau - 0, 0)\, \hat{q}_{yz}(s, d\tau)
\end{aligned}$$

has a unique Borel-measurable and bounded solution.

Proof. According to the argumentation in the proof of theorem 4.25, there is at least one solution of (4.27.1). In order to verify the uniqueness of this solution, assume that there are two different solutions $V^+_{(y,0)}, y \in \mathcal{S}$, and $\tilde{V}^+_{(y,0)}, y \in \mathcal{S}$. Then, the difference $h_y := V^+_{(y,0)} - \tilde{V}^+_{(y,0)}, y \in \mathcal{S}$, must be a solution of (4.26.1), since (4.26.1) is the homogeneous counterpart of (4.27.1). According to lemma 4.26, however, $h_y \equiv 0, y \in \mathcal{S}$, is the one and only solution of (4.26.1). Hence, $V^+_{(y,0)} \equiv \tilde{V}^+_{(y,0)}$. Thus, (4.27.1) has only one solution. □

4.28 Lemma. *Let $\hat{q}$ be a regular cumulative transition intensity matrix according to definition 2.8. Further, let K be a capital accumulation function, and v the appertaining discounting function. Then, the system of integral equations*

$$h_y(s) = K(s) \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} v(\tau)\, h_z(\tau)\, \bar{p}_y(s, \tau - 0, 0)\, \hat{q}_{yz}(s, d\tau), \quad s \geq 0, y \in \mathcal{S}, \tag{4.28.1}$$

has a unique Borel-measurable and bounded solution satisfying

$$h_y(s) = \sum_{z\in\mathcal{S}} \int_{(s,\infty)} v(t)\, |h_z(t)|\, \hat{q}_{yz}(s, dt) < \infty \ \ \forall s \geq 0, \tag{4.28.2}$$

namely $h_y \equiv 0, y \in \mathcal{S}$.

Proof. For the proof we refer to A.19 in the appendix. □

With lemma 4.28, theorem 4.25 is completely verified. Recall that the backward integral equations (2.32.1), (2.32.2), and (2.40.1) are special cases of Thiele's integral equations of type 1 and Thiele's integral equations of type 2, respectively. Hence, it also follows from theorem 4.25 that these systems of integral equations are uniquely solvable. Further, the derivation of solutions to (2.32.2) can be performed in a manner similar to solving Thiele's integral equations of type 1, namely by firstly disregarding the duration-dependence and solving the corresponding system.

F Loss variances and Hattendorff's theorem

Now we want to derive variances of the prospective loss by using a version of Hattendorff's theorem which is adapted to our semi-Markov set-up. We consider the prospective loss according to (4.18.2), and assume (4.19.1). For the corresponding results in a generalized Markov set-up, we likewise refer to Milbrodt and Helbig ([1999], section 10E).

4.29 Theorem. [Hattendorff's theorem] *Assume the situation of theorem 4.18. Further, assume (4.19.1). Then, the prospective losses in state $(\mathcal{L}_y(t))_{t\geq 0}$, $y \in \mathcal{S}$, are centered square integrable martingales with predictable quadratic variation (up to P-indistinguishability)*

$$\begin{aligned} \langle \mathcal{L}_y \rangle_t \;=\; & \sum_{z\neq y} \sum_{m\in\mathbb{N}_0} \int_{(T_m\wedge t, T_{m+1}\wedge t]} \Big(v(\tau)\, R_{yz}(\tau, \tau - T_m) \Big)^2 \, \mathbf{1}_{\{Z_m = y\}}\, \hat{q}_{yz}(T_m, d\tau) \\ & - \sum_{z,\xi\in\mathcal{S}\setminus\{y\}} \sum_{m\in\mathbb{N}_0} \int_{(T_m\wedge t, T_{m+1}\wedge t]} (v(\tau))^2\, R_{yz}(\tau, \tau - T_m)\, R_{y\xi}(\tau, \tau - T_m) \\ & \cdot \mathbf{1}_{\{Z_m = y\}}\, \hat{q}_{yz}(T_m, \{\tau\})\, \hat{q}_{y\xi}(T_m, d\tau), \quad t \geq 0, \end{aligned} \tag{4.29.1}$$

and zero covariation. Further, we obtain for the quadratic variation of $(\mathcal{L}(t))_{t\geq 0}$ the following:

$$\langle \mathcal{L} \rangle_t = \sum_{y\in\mathcal{S}} \langle \mathcal{L}_y \rangle_t. \tag{4.29.2}$$

PROOF. According to the proof of corollary 4. 20, the assumptions of theorem A.5 are satisfied in cases where (4. 19. 1) is fulfilled. It follows from theorem A.5 that the predictable covariation of $\langle \mathcal{L}_y, \mathcal{L}_\eta \rangle_t$ is for $t \geq 0$ and $(y, \eta) \in \mathcal{S}^2$ given by

$$\langle \mathcal{L}_y, \mathcal{L}_\eta \rangle_t = \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq\eta}} \sum_{m\in\mathbb{N}_0} \int_{(T_m\wedge t, T_{m+1}\wedge t]} (v(\tau))^2 R_{yz}(\tau, \tau - T_m)\, R_{\eta\xi}(\tau, \tau - T_m)\, \langle M_{yz}, M_{\eta\xi} \rangle_{d\tau}.$$

By inserting (2. 18. 1), the above can represented as

$$\begin{aligned}\langle \mathcal{L}_y, \mathcal{L}_\eta \rangle_t &= \delta_{y\eta}\, \delta_{z\xi} \sum_{z\neq y} \sum_{m\in\mathbb{N}_0} \int_{(T_m\wedge t, T_{m+1}\wedge t]} \Big(v(\tau)\, R_{yz}(\tau, \tau - T_m)\Big)^2\, \mathbf{1}_{\{Z_m = y\}}\, \hat{q}_{yz}(T_m, d\tau) \\ &\quad -\delta_{y\eta} \sum_{z\neq y} \sum_{\xi\neq y} \sum_{m\in\mathbb{N}_0} \int_{(T_m\wedge t, T_{m+1}\wedge t]} (v(\tau))^2 R_{yz}(\tau, \tau - T_m)\, R_{y\xi}(\tau, \tau - T_m) \\ &\qquad \cdot \mathbf{1}_{\{Z_m = y\}}\, \hat{q}_{yz}(T_m, \{\tau\})\; \hat{q}_{y\xi}(T_m, d\tau).\end{aligned}$$

Thus, one obtains (4. 29. 1) for $y = \eta$, and otherwise $\langle \mathcal{L}_y, \mathcal{L}_\eta \rangle_t = 0$, $t \geq 0$. According to (4. 19. 3),

$$\langle \mathcal{L} \rangle_t = \sum_{(y,\eta)\in\mathcal{S}^2} \langle \mathcal{L}_y, \mathcal{L}_\eta \rangle_t = \sum_{y\in\mathcal{S}} \langle \mathcal{L}_y \rangle_t, \quad t \geq 0 \tag{4. 29. 3}$$

follows immediately from the above. □

Another consequence of theorem A.5 is that the variance of the prospective loss $(\mathcal{L}_y(t))_{t\geq 0}$ in state $y \in \mathcal{S}$ can simply be obtained by means of $\mathbf{V}[\mathcal{L}_y(t)] = \mathbf{E}[\langle \mathcal{L}_y \rangle_t], t \geq 0, y \in \mathcal{S}$. Since the prospective losses are centered martingales, this follows from (A. 5. 3). Hence, according to (4. 29. 3), one obtains

$$\mathbf{V}[\mathcal{L}(t)] = \mathbf{E}[\langle \mathcal{L} \rangle_t] = \sum_{y\in\mathcal{S}} \mathbf{E}[\langle \mathcal{L}_y \rangle_t] = \sum_{y\in\mathcal{S}} \mathbf{V}[\mathcal{L}_y(t)], \quad t \geq 0. \tag{4. 29. 4}$$

For the variance of the prospective loss $\mathcal{L}(t)$ up to time $t \geq 0$, it is therefore sufficient to derive the variance of the corresponding losses in state $(\mathcal{L}_y(t))_{t\geq 0}, y \in \mathcal{S}$. This can be obtained by deriving the expectation of (4. 29. 1). Note that deriving the expectation of (4. 29. 1) corresponds in a sense to the derivation of the prospective reserve according to theorem 4. 6.

4. 30 Corollary. *Assume the situation of theorem* 4. 18*. Further assume* (4. 19. 1) *and* 2. 35*. Then, the variance of* $\mathcal{L}_y(t), t \geq 0, y \in \mathcal{S}$*, is given by*

$$\mathbf{V}[\mathcal{L}_y(t)] = \sum_{\eta\in\mathcal{S}} P(Z_0 = \eta) \cdot \mathbf{E}[\langle \mathcal{L}_y \rangle_t | T_0 = 0, Z_0 = \eta], \tag{4. 30. 1}$$

where

$$\begin{aligned}&\mathbf{E}[\langle \mathcal{L}_y \rangle_t | T_0 = 0, Z_0 = \eta] \\ &= \sum_{z\neq y} \int_{(0,t]} \int_{(0,\infty)} \Big(v(\tau)\, R_{yz}(\tau, l)\Big)^2\, \lambda_{yz}(\tau, l)\, p_{\eta y}(0, \tau - 0, 0, dl)\, \mathbf{\Lambda}_{yz}(d\tau) \\ &\quad - \sum_{z,\xi\in\mathcal{S}\setminus\{y\}} \int_{(0,t]} \int_{(0,\infty)} (v(\tau))^2 R_{yz}(\tau, l)\, R_{y\xi}(\tau, l)\, \lambda_{yz}(\tau, l)\; \lambda_{y\xi}(\tau, l)\, p_{\eta y}(0, \tau - 0, 0, dl) \\ &\qquad \cdot \mathbf{\Lambda}_{yz}(\{\tau\})\, \mathbf{\Lambda}_{y\xi}(d\tau)\end{aligned} \tag{4. 30. 2}$$

for $\eta \in \mathcal{S}$ *satisfying* $P(Z_0 = \eta) > 0$.

Proof. For the proof, we refer to A.20 in the appendix. □

Formula (4. 30. 2) generalizes the corresponding result in Milbrodt and Helbig ([1999], Folgerung 10.39). Yet, analogously to formula (4. 6. 2) for the prospective reserve, (4. 30. 2) is due to the additional duration-dependence of the cumulative transition intensities not suitable to derive the variance of prospective losses. This drawback corresponds to the problem formulated by Norberg [1992] (see our introduction). Regarding the derivation of prospective reserves, however, one observes that though this problem appears when using formula (4. 6. 2), it does not appear for Thiele's integral equations of type 1. Driven by this observation, we will establish a system of integral equations in order to derive $\mathbf{E}[\langle \mathcal{L}_y \rangle_t | T_0 = 0, Z_0 = \eta]$ for $t \geq 0$ and $(\eta, y) \in \mathcal{S}^2$. This is what the next section is concerned with.

G Integral equations for the loss variance

In order to establish a system of integral equations allowing us to derive the variance of prospective losses more comfortably, we consider

$$\mathbf{V}_{(\eta)}^{\mathcal{L}_y(t)}(s) := \mathbf{E}[\langle \mathcal{L}_y \rangle_t - \langle \mathcal{L}_y \rangle_s | T_0 = s, Z_0 = \eta], \;\; 0 \leq s \leq t, (\eta, y) \in \mathcal{S}^2. \tag{4. 30. 3}$$

According to (4. 29. 1), it turns out that for $\mathcal{L}(T_0, Z_0 |\, P)$-a.e. $(s, \eta) \in [0, \infty) \times \mathcal{S}$

$$\mathbf{V}_{(\eta)}^{\mathcal{L}_y(s)}(s) = 0.$$

For $s = 0$, one obtains $\mathcal{L}(Z_0 |\, P)$-a.s.

$$\mathbf{V}_{(\eta)}^{\mathcal{L}_y(t)}(s) = \mathbf{E}[\langle \mathcal{L}_y \rangle_t | T_0 = 0, Z_0 = \eta]. \tag{4. 30. 4}$$

This yields

$$\mathbf{V}\left[\mathcal{L}_y(t)\right] = \mathbf{E}\left[\langle \mathcal{L}_y \rangle_t\right] \;\; = \;\; \sum_{\eta \in \mathcal{S}} \mathbf{V}_{(\eta)}^{\mathcal{L}_y(t)}(0)\, P(Z_0 = \eta). \tag{4. 30. 5}$$

The following theorem establishes a system of integral equations for the expressions $\mathbf{V}_*^{\mathcal{L}_*}$ according to (4. 30. 3). In contrast to (4. 30. 2), this system of integral equations allows us to derive (4. 30. 4) without an additional integration due to the duration-dependence of the cumulative transition intensities.

4. 31 Theorem. *Assume the situation of theorem* 4. 18. *Further assume* (4. 19. 1)*, and let* $0 \leq s \leq t < \infty$*, and* $(\eta, y) \in S^2$*. Then, the conditional expectations* $\mathbf{V}_*^{\mathcal{L}_*}$ *satisfy for* $\mathcal{L}(T_0, Z_0|P)$*-a.e.* (s, η) *the system of integral equations*

$$\begin{aligned}
&\mathbf{V}_{(\eta)}^{\mathcal{L}_y(t)}(s) \qquad\qquad (4.31.1)\\
&= \sum_{\substack{z \in \mathcal{S} \\ z \neq \eta}} \int_{(s,t]} \Big(\big(\delta_{\eta y}\, v(\tau)\, R_{\eta z}(\tau, \tau - s)\big)^2 + \mathbf{V}_{(z)}^{\mathcal{L}_y(t)}(\tau) \Big)\, \bar{p}_\eta(s, \tau - 0, 0)\, \hat{q}_{\eta z}(s, d\tau) \\
&\quad - \sum_{z, \xi \in \mathcal{S} \setminus \{\eta\}} \int_{(s,t]} \delta_{\eta y}\, (v(\tau))^2\, R_{\eta z}(\tau, \tau - s)\, R_{\eta \xi}(\tau, \tau - s)\, \hat{q}_{\eta z}(s, \{\tau\})\, \bar{p}_\eta(s, \tau - 0, 0)\, \hat{q}_{\eta \xi}(s, d\tau).
\end{aligned}$$

Proof. For the proof we refer to A.21 in the appendix. □

As mentioned above, by using the system of integral equations (4. 31. 1), the additional integration due to the duration-dependence can be avoided. Further, the transition probabilities of the bivariate Markov process $((X_t, U_t))_{t\geq 0}$ do not have to be provided, and assumption 2.35 can be omitted. To summarize, (4.31.1) has almost all advantages of Thiele's integral equations of type 1 when compared to (4.6.2).

A solution to the system (4.31.1) for $\mathbf{V}_{(*)}^{\mathcal{L}_y(t)}(\cdot)$ yields by means of (4.30.5) the variance of the prospective loss in state y up to time t. Hence, in order to compute, for example, the variances of all annual losses, the loss variances must be derived for all integer times and all possible states, each of which refers to a system of the form (4.31.1). This procedure can be compared with that of deriving transition probabilities of a non-smooth Markovian pure jump process by using a system of integral equations according to (4.46.2) or (4.48.2) in Milbrodt and Helbig [1999]. There, each solution likewise refers to a fixed point in time and a fixed state.

By continuing example 4.14, we will derive the variances of annual losses for the PHI policies under consideration. Further, we compare these variances for both models, the semi-Markov model and the Markov model. Analogously to Thiele's integral equations in example 4.14, the system of integral equations (4.31.1) is likewise solved numerically by using the same approach. Incidentally, in a Markov set-up, it would be more convenient to derive the variances of prospective losses by using the counterpart to (4.30.2) (cf. Milbrodt and Helbig [1999], Folgerung 10.39), especially in cases where the transition probabilities are available.

Before we turn to the example, we will just mention that each system of the form (4.31.1) is uniquely solvable. To verify this, one can argue in a manner similar to the proof of lemma 4.27, since the difference of two presumed solutions of (4.31.1),

$$h_\eta := \tilde{\mathbf{V}}_{(\eta)}^{\mathcal{L}_y(t)} - \mathbf{V}_{(\eta)}^{\mathcal{L}_y(t)}, \ \ \eta \in \mathcal{S},$$

must be a solution of (4.26.1), with $K \equiv v \equiv 1$. The system (4.26.1), however, was proved to be uniquely solved by $h_\eta \equiv 0, \eta \in \mathcal{S}$. Hence, both presumed solutions of (4.31.1) must coincide.

4.32 Example. [example 4.14 continued] Recall the situation of example 4.14. Different PHI policies have been considered for which the decrement for state *invalid* was modelled duration-depending. By assuming (4.14.2) and employing linear interpolation, the select-and-ultimate tables DAV-SST 1997 TI and DAV-SRT 1997 RI have provided transition intensities $\mu_{iz}, \ z \in \{a, d\}$, according to (4.14.16) and (4.14.17), respectively. These transition intensities additionally depend on the time elapsed since onset of disability. Regarding the corresponding Markov model, the ultimate tables of both above tables were used. Thus, the corresponding transition intensities are not duration-depending. The decrement for state *active* is in both models also described by mere time-depending intensities. These intensities were derived from the tables DAV-ST 1994 T and DAV-IT 1997. We have considered policies with term $n = 30$ for insured of three different ages at issue, $x = 20, 30, 40$. For these policies, we will now compare the variances of the annual losses in the states *active* and *invalid*, and the variance of the overall loss. The latter can be obtained by adding the variances of the annual losses with respect to all possible states, i.e. *active*, *invalid*, *dead*. Due to the fact that the state *dead* is absorbing, the intensities for transitions from this state vanish. Hence, the corresponding loss variances likewise vanish (cf. (4.30.2)).

In order to derive the variances of annual losses, we employ the system of integral equations (4.31.1). In the present situation, this is given by

$$\mathbf{V}_{(\eta)}^{\mathcal{L}_y(t)}(s) \ = \ \sum_{z\neq\eta} \int_{(s,t]} \Big(\big(\delta_{\eta y}\, v(\tau)\, R_{\eta z}(\tau, \tau - s)\big)^2 + \mathbf{V}_{(z)}^{\mathcal{L}_y(t)}(\tau) \Big)\, \bar{p}_\eta(s, \tau - 0, 0)\, \hat{q}_{\eta z}(s, d\tau) \quad (4.32.1)$$

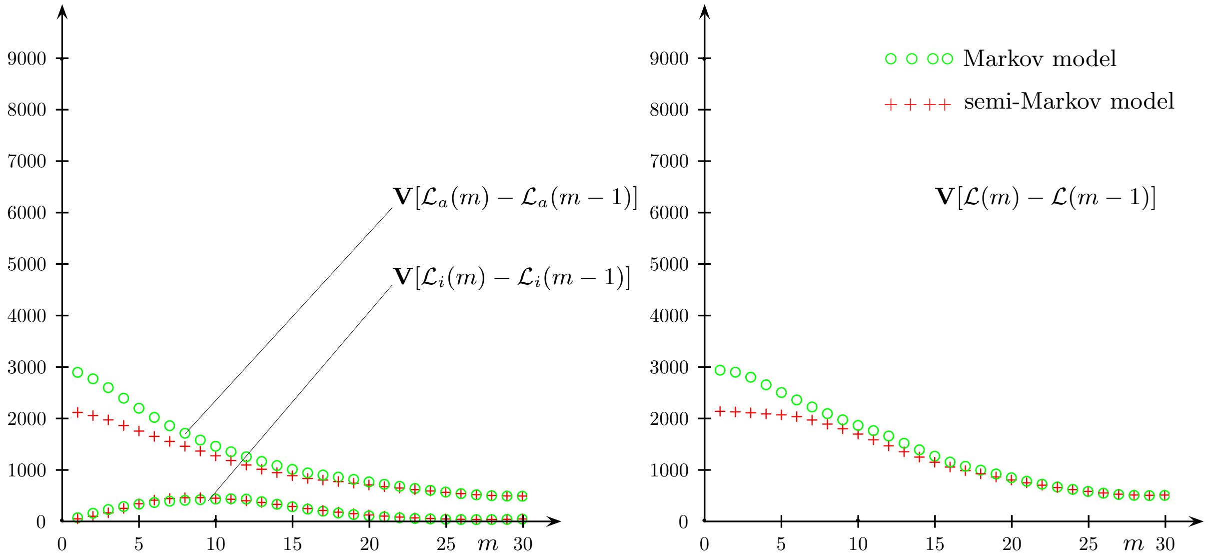

Figure 26: Comparison of the variances of annual losses in a semi-Markov model with the corresponding variances for the appertaining Markov model, for male insured with age at issue $x = 20$

($0 \leq s \leq t$, $(\eta, y) \in \mathcal{S}^2$), since the cumulative transition intensities are continuous. This system of integral equations shall likewise be solved numerically by using the right-hand side rectangle formula (cf. example 4.14). By setting $a = s$ and $b = t$, we solve the system of integral equations (4.32.1) recursively by means of ($l \in \mathbb{N}$, $h := b - a/l$ and $x_k = a + k \cdot h, k = 0, ..., l$)

$$\mathbf{V}_{(\eta)}^{\mathcal{L}_y(t)}(x_{k-1}) = h \sum_{\substack{z\in\mathcal{S}\\ z\neq\eta}} \sum_{j=k}^{l} \Big(\big(\delta_{\eta y}\, v(x_j)\, R_{\eta z}(x_j, x_j - x_{k-1})\big)^2 + \mathbf{V}_{(z)}^{\mathcal{L}_y(t)}(x_j)\Big) \cdot e^{\int_{(x_{k-1},x_j]} \mu_{\eta\eta}(r, r-x_{k-1})\,dr} \mu_{yz}(x_j, x_j - x_{k-1}),$$

for $k = 1, ..., l$. The starting point is given by t, and the initial condition is

$$\mathbf{V}_{(\eta)}^{\mathcal{L}_y(t)}(t) = 0, \ \ \forall (\eta, y) \in \mathcal{S}^2.$$

The variances of annual losses in state $y \in \mathcal{S}$ can according to $\mathbf{V}\left[\mathcal{L}_y(t)\right] = \mathbf{E}\left[\langle\mathcal{L}_y\rangle_t\right], t \geq 0$, be derived as

$$\mathbf{V}[\mathcal{L}_y(m) - \mathcal{L}_y(m-1)] = \mathbf{V}[\mathcal{L}_y(m)] - \mathbf{V}[\mathcal{L}_y(m-1)], \ \ m \in \mathbb{N}, m \leq n.$$

Since the policies are commencing in state *active*, (4.30.5) reduces to

$$\mathbf{V}[\mathcal{L}_a(m)] = \mathbf{V}_{(a)}^{\mathcal{L}_a(m)}(0) \quad \text{and} \quad \mathbf{V}[\mathcal{L}_i(m)] = \mathbf{V}_{(a)}^{\mathcal{L}_i(m)}(0).$$

In order to derive $\mathbf{V}_{(a)}^{\mathcal{L}_a(m)}(0)$, we must determine $\mathbf{V}_{(a)}^{\mathcal{L}_a(m)}(\cdot), \mathbf{V}_{(i)}^{\mathcal{L}_a(m)}(\cdot)$, and $\mathbf{V}_{(d)}^{\mathcal{L}_a(m)}(\cdot)$. In order to derive $\mathbf{V}_{(a)}^{\mathcal{L}_i(m)}(0)$, the quantities $\mathbf{V}_{(a)}^{\mathcal{L}_i(m)}(\cdot), \mathbf{V}_{(i)}^{\mathcal{L}_i(m)}(\cdot)$, and $\mathbf{V}_{(d)}^{\mathcal{L}_i(m)}(\cdot)$ must be used. Note that $\mathbf{V}_{(d)}^{\mathcal{L}_a(m)}(\cdot) \equiv \mathbf{V}_{(d)}^{\mathcal{L}_i(m)}(\cdot) \equiv 0$.

Figure 26 illustrates the variances of the annual losses for both the semi-Markov model and the appertaining Markov model in case of an insured with age at issue $x = 20$. One can observe that the variances of the annual losses in state *invalid* almost coincide for both models. In state *active*, however, the variances of the annual losses for the semi-Markov model are smaller, particularly for the first half of the policy term. This results in the variances of the annual overall losses being lower for the semi-Markov model than for the Markov model. For the policies pertaining to insured of age at issue $x = 30$ and $x = 40$ (cf. figure 27 and 28), this remains the same. Yet for these policies, the variances of the annual losses in state *invalid* are likewise higher for the first half of the policy term than the corresponding values for the Markov model.

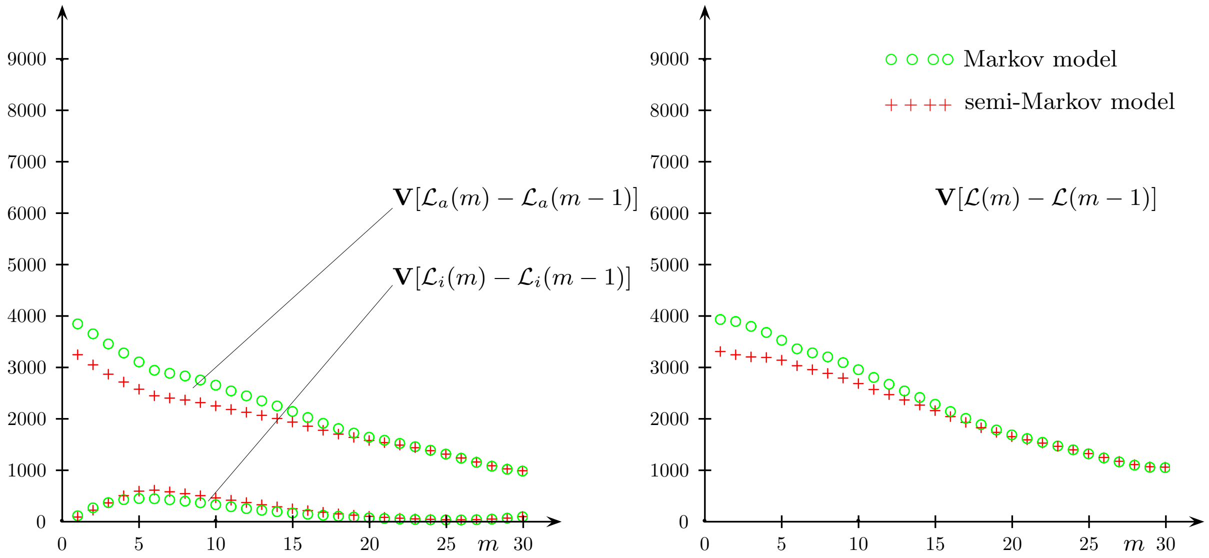

Figure 27: COMPARISON OF THE VARIANCES OF ANNUAL LOSSES IN A SEMI-MARKOV MODEL WITH THE CORRESPONDING VARIANCES FOR THE APPERTAINING MARKOV MODEL, FOR MALE INSURED WITH AGE AT ISSUE $x = 30$

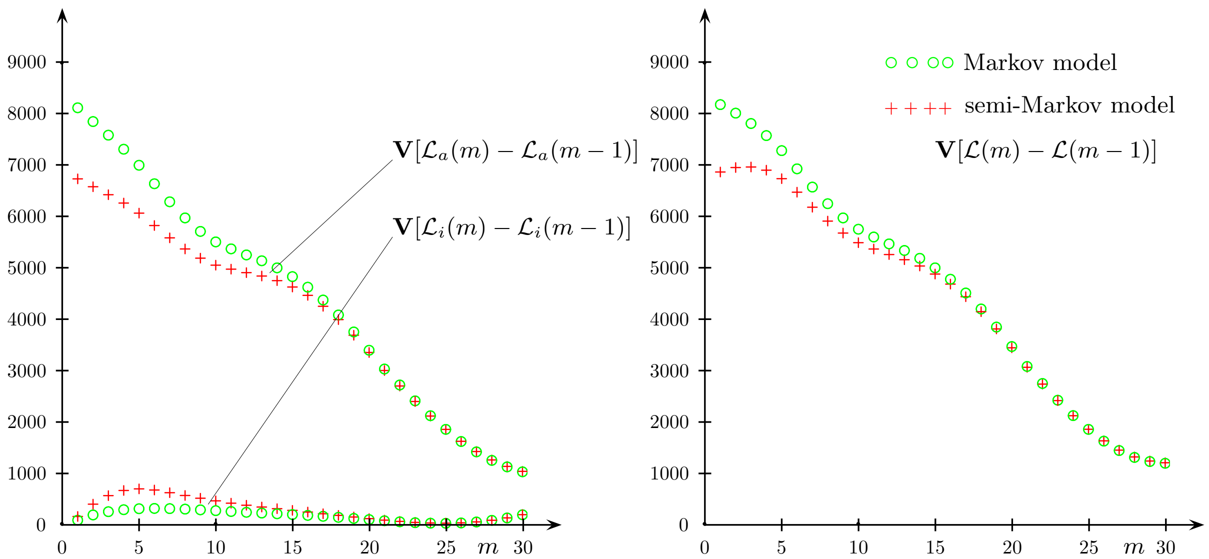

Figure 28: COMPARISON OF THE VARIANCES OF ANNUAL LOSSES IN A SEMI-MARKOV MODEL WITH THE CORRESPONDING VARIANCES FOR THE APPERTAINING MARKOV MODEL, FOR MALE INSURED WITH AGE AT ISSUE $x = 40$

Note that the sum of all variances of the annual overall losses corresponds to the variance of the present value of all premiums and benefits for the policy considered. Due to lower values of the variances of the annual overall losses for the semi-Markov model, the variances of the present values of all premiums and benefits likewise turn out to be smaller. Hence, for this example, the semi-Markov approach seems to result in a less risky calculation when compared with the corresponding Markov approach.

age at issue	$x = 20$		$x = 30$		$x = 40$	
	$\sqrt{\mathbf{V}[\mathcal{L}_i]}$	$\sqrt{\mathbf{V}[\mathcal{L}]}$	$\sqrt{\mathbf{V}[\mathcal{L}_i]}$	$\sqrt{\mathbf{V}[\mathcal{L}]}$	$\sqrt{\mathbf{V}[\mathcal{L}_i]}$	$\sqrt{\mathbf{V}[\mathcal{L}]}$
Markov model	79.58	205.89	73.65	275.34	68.74	387.95
semi-Markov model	78.84	200.42	85.80	264.83	93.15	377.22
percentage	99.0%	97.3%	116.5%	96.0%	135.5%	97.2%
percentage (premium)		97%		96%		97.8%

The above table contains the standard deviation of the sums of all annual losses in state *invalid* and the standard deviation of the present values of all benefits and premiums. Further, the differences between these values for both models are quantified by deriving the percentage of the values within the semi-Markov framework when compared with the values obtained from the appertaining Markov model. It turns out that these percentages are almost of the same magnitude as the corresponding percentages for the premiums of both models. Yet, as for the premiums, these results are not valid in general. Rather, they also depend on the policy specifications. $\triangle$

Chapter 5

Retrospective reserves

This chapter is concerned with the concept of retrospective reserves which forms a counterpart to the theory of prospective reserves. Our aim of investigating retrospective reserves is, analogously to Norberg [1991], to find generalizations of the forward integral equations for the transition probabilities of the underlying Markov process, which have a meaningful interpretation in view of a single policy. We start by introducing different approaches for retrospective reserves in a Markov set-up. By adapting Norberg's [1991] definition, we afterwards derive retrospective reserves in a non-smooth Markov set-up as well as in a non-smooth semi-Markov set-up. In a semi-Markov set-up, however, we must additionally assume 2. 35 and 4. 5. Recall that these assumptions were also essential for establishing formula (4. 6. 2) for the prospective reserve. Further, we state, at least for the Markov set-up, systems of integral equations which correspond to either Thiele's integral equations of type 1 or Thiele's integral equations of type 2 for the prospective reserve. These systems of integral equations generalize the corresponding forward integral equations. In a semi-Markov framework, one can also establish systems of integral equations that correspond to either Thiele's integral equations of type 1 or of type 2. Yet, as it already can be observed when comparing backward and forward equations in a semi-Markov set-up (compare (2. 32. 2) and (2. 36. 1), as well as (2. 40. 1) and (2. 41. 1)), the integration over the times elapsed since entering the current state does also appear in the integral equations for the retrospective reserve. Regarding the prospective reserve, it was the main advantage of the use of integral equations that this problem, caused by the duration-dependence of the cumulative transition intensities, could be avoided.

We now turn to the introduction of different definitions of prospective reserves. For this, we firstly assume a Markov set-up according to section 2.D.2. This means that the pure jump process $(\Omega, \mathfrak{F}, P, (X_t)_{t\geq 0})$ is Markovian with transition probability matrix

$$p(s,t) = (p_{yz}(s,t))_{(y,z)\in\mathcal{S}^2},\ 0 \leq s \leq t < \infty,$$

and cumulative transition intensity matrix

$$q(s,t) = (q_{yz}(s,t))_{(y,z)\in\mathcal{S}^2}.$$

As customary, $((T_m, Z_m))_{m\in\mathbb{N}_0}$ is the appertaining marked point process, homogeneous and likewise Markovian. Further, we consider actuarial payments which are not duration-depending. Doing so, we get according to definition 3. 12 for a capital accumulation function K and the appertaining discounting function v

$$\begin{aligned} V^-(t,A) &= -K(t) \sum_{m\in\mathbb{N}_0} \int_{[T_m\wedge t, T_{m+1}\wedge t)} v(\tau)\, F_{Z_m}(d\tau) \\ &\quad -K(t) \sum_{m\in\mathbb{N}_0} \frac{D_{Z_m Z_{m+1}}(T_{m+1})}{K(DT(T_{m+1}))}\, \mathbf{1}_{\{T_{m+1}\leq t\}}, \qquad (5.0.1) \end{aligned}$$

and

$$V^+(t,A) = K(t)\sum_{m\in\mathbb{N}_0}\int_{[T_m\vee t,T_{m+1}\vee t)} v(\tau)\,F_{Z_m}(d\tau) + K(t)\sum_{m\in\mathbb{N}_0}\frac{D_{Z_mZ_{m+1}}(T_{m+1})}{K(DT(T_{m+1}))}\,\mathbf{1}_{\{t<T_{m+1}<\infty\}}. \tag{5.0.2}$$

Further, we assume $\mathbf{E}[|V_0|]<\infty$ as well as $\mathbf{E}[V_0]=0$. The prospective reserve is given as

$$V^+(t) = \mathbf{E}[V^+(t,A)|\,\mathfrak{F}_t] = \mathbf{E}[V^+(t,A)|\,X_t], \quad t\geq 0,$$

where $(\mathfrak{F}_t)_{t\geq 0}$ is a filtration such that the process $(X_t)_{t\geq 0}$ is adapted to $(\mathfrak{F}_t)_{t\geq 0}$. The prospective reserve in state $y\in\mathcal{S}$ is then defined as $V_y^+(t) := \mathbf{E}[V^+(t,A)|\,X_t=y]$, and it holds

$$V^+(t) = \sum_{y\in\mathcal{S}}\mathbf{1}_{\{X_t=y\}}V_y^+(t) \quad P-a.s. \tag{5.0.3}$$

A Different definitions of the retrospective reserve

According to definition 3.6, the retrospective reserve is in a non-random set-up defined as accumulated value of past premiums minus benefits. Recall that in that situation, for deterministic payment streams which are assumed to be equivalent, the retrospective reserve coincides at any time with the prospective reserve (cf. corollary 3.7). Considering a random set-up, meaning here a single policy described by a pure jump process, the retrospective reserve must satisfy the basic requirement that at any time $t\geq 0$ the expectation of it coincides with the expected accumulated value of all premiums minus benefits up to time t (cf. Wolthuis [1994], chapter 9). For the latter, one obtains, by assuming the principle of equivalence, $\mathbf{E}[V^-(t,A)]=\mathbf{E}[V^+(t,A)]$, $t\geq 0$, where the right-hand side of this equation is according to (5.0.3) given by

$$\mathbf{E}[V^+(t,A)] = \mathbf{E}[V^+(t)] = \sum_{y\in\mathcal{S}} V_y^+(t)\,P(X_t=y).$$

For the basic requirement for the retrospective reserve, denoted as $V^-(t)$, $t\geq 0$, we then get

$$\mathbf{E}[V^-(t)] = \mathbf{E}[V^-(t,A)] = \mathbf{E}[V^+(t,A)] = \mathbf{E}[V^+(t)] = \sum_{y\in\mathcal{S}} V_y^+(t)\,P(X_t=y),\ t\geq 0.$$

By defining the retrospective reserve for each state $y\in\mathcal{S}$, the basic requirement is finally given by

$$\sum_{y\in\mathcal{S}} V_y^-(t)\,P(X_t=y) = \sum_{y\in\mathcal{S}} V_y^+(t)\,P(X_t=y), \quad t\geq 0. \tag{5.0.4}$$

Thus, the expectation of the retrospective reserve must at any time coincide with the expectation of the prospective reserve. The relation (5.0.4), however, does not uniquely determine the retrospective reserves $V_y^-(\cdot)$, $y\in\mathcal{S}$. For example, it can already be satisfied by setting

$$V_y^-(t) := V_y^+(t),\ \forall y\in\mathcal{S}, t\geq 0 \ \text{ with } \ P(X_t=y)>0. \tag{5.0.5}$$

This is basically the definition of retrospective reserves introduced by Hoem [1988]. With exception of the initial state $X_0=a$ satisfying $P(X_0=a)=1$, Hoem defined $V_y^-(t) := V_y^+(t)$, $t\geq 0$, for each state $y\in\mathcal{S}\setminus\{a\}$. For the initial state, the following was defined:

$$V_a^-(t) := V_a^+(t) - \frac{\mathbf{E}[V_0]}{v(t)\,p_{aa}(0,t)}, \quad t\geq 0, \tag{5.0.6}$$

where $p_{aa}(0,t) = P(X_t = a|\, X_0 = a)$ must be strictly positive. In summary, the $|\mathcal{S}|$-dimensional vector of retrospective reserves defined by Hoem [1988] is given by

$$\mathbf{V}^-(t) = \mathbf{V}^+(t) - (v(t)\, p_{aa}(0,t))^{-1} \begin{pmatrix} V_a^+(0) \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}. \tag{5.0.7}$$

According to this definition, it is granted that the expectation of the retrospective reserve $\mathbf{E}[V^-(t)]$ coincides for each $t \geq 0$ with $\mathbf{E}[V^-(t,A)]$. This equality also holds in cases where the principle of equivalence is not stipulated. If the principle of equivalence is fulfilled, the retrospective reserves by Hoem [1988] even coincide for each state with the corresponding prospective reserves.

To explain the concept of Wolthuis and Hoem [1990], we likewise start from the basic requirement for retrospective reserves, $\sum_{y\in\mathcal{S}} V_y^-(t)\, P(X_t = y) = \mathbf{E}[V^-(t)] = \mathbf{E}[V^-(t,A)], t \geq 0$. Wolthuis and Hoem [1990] allow several initial states $v \in \mathcal{S}$ and convert - by conditioning - this requirement into

$$\mathbf{E}[V^-(t,A)|\, X_0 = v] = \sum_{y\in\mathcal{S}} V_y^-(t)\, P(X_t = y|\, X_0 = v), \;\; t \geq 0, v \in \mathcal{S}. \tag{5.0.8}$$

Discounting the above yields

$$\mathbf{E}[v(t)\, V^-(t,A)|\, X_0 = v] = \sum_{y\in\mathcal{S}} v(t)\, V_y^-(t)\, P(X_t = y|\, X_0 = v),$$

where the left-hand side is also equal to $\sum_{y\in\mathcal{S}} v(t)\, p_{vy}(0,t)\, V_y^+(t) - V_v^+(0)$. This leads to retrospective reserves $V_y^-(\cdot)$, $y \in \mathcal{S}$, determined by

$$\sum_{y\in\mathcal{S}} v(t)\, p_{vy}(0,t)\, V_y^-(t) = -V_v^+(0) + \sum_{y\in\mathcal{S}} v(t)\, p_{vy}(0,t)\, V_y^+(t).$$

If the transition probability matrix p is nonsingular, the above results in

$$\mathbf{V}^-(t) = \mathbf{V}^+(t) - K(t)\, p^{-1}(0,t)\, \mathbf{V}^+(0). \tag{5.0.9}$$

The retrospective reserves according to Wolthuis and Hoem [1990] differ from the retrospective reserves according to Hoem [1988] (cf. (5.0.7)). For a more detailed discussion of these approaches, we again refer to Wolthuis ([1994], chapter 9 & 10) and additionally to Norberg [1991]. The latter introduced a third definition of the retrospective reserves $V_y^-(t), t \geq 0, y \in \mathcal{S}$, simply by means of

$$V_y^-(t) := \mathbf{E}[V^-(t,A)|\, X_t = y]. \tag{5.0.10}$$

Obviously, prospective reserves defined this way also satisfy (5.0.4). According to Norberg [1991], who only investigated the smooth case, this approach yields retrospective reserves for which the corresponding differential equations generalize the forward differential equations for the transition probabilities of the Markov process $(X_t)_{t\geq 0}$. In what follows, Norberg's approach shall be adapted to a non-smooth Markov set-up, and afterwards to a semi-Markov set-up.

B Retrospective reserves in a non-smooth Markov set-up

Consider a policy (p) modelled by a Markovian pure jump process $(X_t)_{t\geq 0}$ commencing in state a. $((T_m, Z_m))_{m\in\mathbb{N}_0}$ is the appertaining marked point process satisfying $P(T_0 = 0, Z_0 = a) = 1$. Actuarial payments are specified by an actuarial payment function $A = DA + SA$ with $V^-(\cdot, A)$ according to (5. 0. 1). Then, the following representation of the retrospective reserves according to Norberg's approach can be established. For the proof we refer to A.22 in the appendix.

5. 1 Theorem. *In the present framework, the retrospective reserves according to* (5. 0. 10) *are for* $t \geq 0$ *and* $y \in \mathcal{S}$ *satisfying* $P(X_t = y) > 0$ *given as*

$$\begin{aligned} V_y^-(t) &= -K(t) \sum_{\xi\in\mathcal{S}} \int_{[0,t)} v(\tau)\, \frac{p_{a\xi}(0,\tau)\, p_{\xi y}(\tau,t)}{p_{ay}(0,t)}\, F_\xi(d\tau) \\ &\quad -K(t) \sum_{(\xi,z)\in\mathcal{J}} \int_{(0,t]} v(DT(\tau))\, D_{\xi z}(\tau)\, \frac{p_{a\xi}(0,\tau-0)\, p_{zy}(\tau,t)}{p_{ay}(0,t)}\, q_{\xi z}(d\tau). \qquad (5.1.1) \end{aligned}$$

The appertaining system of (Thiele's) integral equations of type 1 is given by

$$\begin{aligned} V_y^-(t) &= -K(t) \int_{[0,t)} v(\tau)\, \frac{p_{ay}(0,\tau)\, \bar{p}_y(\tau,t)}{p_{ay}(0,t)}\, F_y(d\tau) \qquad (5.1.2) \\ &\quad -K(t) \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(0,t]} \Big(\frac{D_{zy}(\tau)}{K(DT(\tau))} - v(\tau-0)\, V_z^-(\tau-0)\Big) \frac{p_{az}(0,\tau-0)\, \bar{p}_y(\tau,t)}{p_{ay}(0,t)}\, q_{zy}(d\tau), \end{aligned}$$

for $t \geq 0$ and $y \in \mathcal{S}$ satisfying $P(X_t = y) > 0$. This system of integral equations generalizes the system of forward integral equations (2. 42. 5). In order to verify this, the parameters

$$D_{\xi z} \equiv 0,\ (\xi, z) \in \mathcal{J}, \quad v \equiv 1, \quad \text{and} \quad F_\xi(d\tau) := \delta_{\xi z}\, \varepsilon_s(d\tau),\ \xi \in \mathcal{S},\ s \geq 0,$$

must be used. With these parameters, one obtains from (5. 1. 1)

$$V_y^-(t) = -\mathbf{1}_{(s<t)} \frac{p_{az}(0,s)\, p_{zy}(s,t)}{p_{ay}(0,t)},$$

and hence,

$$-p_{ay}(0,t)\, V_y^-(t) = \mathbf{1}_{(s<t)}\, p_{az}(0,s)\, p_{zy}(s,t). \qquad (5.1.3)$$

From (5. 1. 2), it follows

$$-p_{ay}(0,t)\, V_y^-(t) = p_{az}(0,s) \sum_{\substack{v\in\mathcal{S}\\ v\neq y}} \int_{(s,t]} p_{zv}(s,\tau-0)\, \bar{p}_y(\tau,t)\, q_{vy}(d\tau) + p_{ay}(0,s)\, \bar{p}_y(s,t)\, \delta_{yz}\, \mathbf{1}_{(s<t)}.$$

Hence, along with (5. 1. 3), this is

$$\mathbf{1}_{(s<t)}\, p_{az}(0,s)\, p_{zy}(s,t) = p_{az}(0,s) \Big(\mathbf{1}_{(s<t)}\, \delta_{zy}\, \bar{p}_z(s,t) + \sum_{\substack{v\in\mathcal{S}\\ v\neq y}} \int_{(s,t]} p_{zv}(s,\tau-0)\, \bar{p}_y(\tau,t)\, q_{vy}(d\tau)\Big),$$

from which for $s < t$ (2. 42. 5) follows.

Similar to Thiele's integral equations for the prospective reserve, (5. 1. 2) can be proved in two different ways. One the one hand, (5. 1. 2) can simply be obtained by inserting (2. 42. 5) into (5. 1. 1) (substitution of $p_{\xi y}(\tau, t)$ and $p_{zy}(\tau, t)$, respectively), and afterwards applying (5. 1. 1) again. On the other hand, (5. 1. 2) can also be proved without employing the forward integral equations (2. 42. 5). The latter can then be obtained by inserting the above parameters. Thus, it can be verified that both systems of integral equations are equivalent.

A system of integral equations for the retrospective reserves according to (5. 0. 10), which generalizes the forward integral equations of type 2 (cf. Milbrodt and Helbig [1999], (4.49.2)), is given by

$$\begin{aligned} &V_y^-(t) \\ &= -\int_{[0,t)} F_y(dr) + \int_{(0,t]} V_y^-(\tau-0)\,\Phi(d\tau) \qquad (5.1.4)\\ &-\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \int_{(0,t]} \Big(K(r-0)\,v(DT(r))\,D_{\xi y}(r) + V_y^-(r-0) - V_\xi^-(r-0)\Big)\frac{p_{a\xi}(0,r-0)}{p_{ay}(0,r)}\,q_{\xi y}(dr). \end{aligned}$$

In order to prove this counterpart of Thiele's integral equations of type 2 (cf. Milbrodt and Helbig [1999], Satz 10.18) in a non-smooth Markov set-up, one can likewise start from (5. 1. 1) and insert the forward integral equations of type 2. Doing so, one obtains with corollary A.3

$$\begin{aligned} &v(t)\,V_y^-(t) \\ &= -\int_{[0,t)} v(r)\,F_y(dr) \\ &-\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \int_{(0,t]} v(r-0)\Big(K(r-0)\,v(DT(r))\,D_{\xi y}(r) + V_y^-(r-0) - V_\xi^-(r-0)\Big)\frac{p_{a\xi}(0,r-0)}{p_{ay}(0,r)}\,q_{\xi y}(dr). \end{aligned}$$

By afterwards applying

$$K(t)\cdot v(s) = 1 + v(s)\int_{(s,t]} K(\tau-0)\,\Phi(d\tau),\;\; s\le t, \qquad (5.1.5)$$

(5. 1. 4) follows. Similar to (3. 2. 3), (5. 1. 5) is also a consequence of (3. 2. 1).

C Retrospective reserves in a non-smooth semi-Markov set-up

Now assume $(X_t)_{t\ge 0}$ to be a semi-Markovian pure jump process modelling a policy (p). This process is likewise assumed to start from state a. Then, the appertaining marked point process $((T_m, Z_m))_{m\in\mathbb{N}_0}$ also satisfies $P(T_0 = 0, Z_0 = a) = 1$. Let this process be a homogeneous Markov chain with regular cumulative transition intensity matrix $\hat{q}$. $((X_t, U_t))_{t\ge 0}$ is the appertaining bivariate Markov process with regular cumulative transition intensity matrix q. Actuarial payments are given by an actuarial payment function $A = DA + SA$ with $V^-(\cdot, A)$ according to (3. 12. 1). Instead of (5. 0. 10), we define the retrospective reserves in the present semi-Markov framework by

$$V_{(y,v)}^-(t) := \mathbf{E}[V^-(t,A)\,|\,X_t = y, U_t \le v],\;\; 0\le v\le t, y\in\mathcal{S}. \qquad (5.1.6)$$

They can be represented as follows.

5. 2 Theorem. *Assume* 2. 35 *and* 4. 5*. Then, the retrospective reserve according to* (5. 1. 6) *is in the present framework for* $0\le v\le t$ *and* $y\in\mathcal{S}$ *satisfying* $P(X_t = y, U_t \le v) > 0$ *given as*

$$\begin{aligned} &V_{(y,v)}^-(t) \\ &= -K(t)\sum_{\xi\in\mathcal{S}}\int_{[0,t)}\int_{[0,\infty)} v(\tau)\,f_\xi(\tau,l)\,\frac{p_{a\xi}(0,\tau,0,dl)\,p_{\xi y}(\tau,t,l,v)}{p_{ay}(0,t,0,v)}\,\mathbf{F}_\xi(d\tau) \qquad (5.2.1)\\ &\quad -K(t)\sum_{(\xi,z)\in\mathcal{J}}\int_{(0,t]}\int_{(0,\infty)} \frac{D_{\xi z}(\tau,l)}{K(DT(\tau))}\,\lambda_{\xi z}(\tau,l)\,\frac{p_{a\xi}(0,\tau-0,0,dl)\,p_{zy}(\tau,t,0,v)}{p_{ay}(0,t,0,v)}\,\mathbf{\Lambda}_{\xi z}(d\tau). \end{aligned}$$

Proof. The proof can be found in A.23 in the appendix. □

Starting from (5. 2. 1) and inserting the forward integral equations of type 1 (2. 36. 1), a system of integral equations corresponding to Thiele's integral equations of type 1 for the prospective reserve in a semi-Markov framework (cf. theorem 4. 8) can be obtained. Similarly, by using the forward integral equations of type 2 (2. 41. 1), a system of integral equations can be obtained, which corresponds to Thiele's integral equations of type 2 for the prospective reserve (cf. corollary 4. 22). As mentioned at the beginning of this chapter, however, the additional integration due to the duration-dependence of cumulative transition intensities and actuarial payments cannot be avoided. Hence, computing the retrospective reserves by using the corresponding systems of integral equations seems not to have any advantage when compared with (5. 2. 1). Regarding the prospective reserve in a semi-Markov framework, however, the use of Thiele's integral equations of type 1 allows one to derive prospective reserves more comfortable.

Appendix A

Tools and Proofs

1 Tools

The following tools are taken from Last and Brandt ([1995], appendix A4). Proofs are omitted.

A. 1 Theorem. *Let $F, G : \mathbb{R} \to \mathbb{R}$ be two right continuous functions of locally bounded variation. Then for $s \le t$*

$$\begin{aligned} F(t)\,G(t) - F(s)\,G(s) &= \int_{(s,t]} F(\tau)\,G(d\tau) + \int_{(s,t]} G(\tau-0)\,F(d\tau) \\ &= \int_{(s,t]} F(\tau-0)\,G(d\tau) + \int_{(s,t]} G(\tau-0)\,F(d\tau) \\ &\quad + \sum_{s<\tau\le t} \Delta F(\tau)\,\Delta G(\tau), \end{aligned} \tag{A. 1. 1}$$

where $F(\tau-0) = \lim_{s \nearrow \tau} F(s)$ and $\Delta F(s) = F(s) - F(s-0)$.

A. 2 Corollary. *Let $F : \mathbb{R} \to \mathbb{R}$ be a right continuous function of locally bounded variation with*

$$\inf_{0\le s\le t} |F(s)| > 0$$

for all t. Then

$$d\left(\frac{1}{F(t)}\right) = -\frac{F(dt)}{F(t)\,F(t-0)}. \tag{A. 2. 1}$$

A. 3 Corollary. *Let $F, G : \mathbb{R} \to \mathbb{R}$ be two right continuous functions of locally bounded variation. Further, assume that $\frac{1}{F}$ is also of locally bounded variation. Then for $s \le t$*

$$\frac{G(t)}{F(t)} - \frac{G(s)}{F(s)} = \int_{(s,t]} \frac{G}{F}(d\tau) = \int_{(s,t]} \frac{G(d\tau)}{F(\tau-0)} - \int_{(s,t]} \frac{G(\tau)\,F(d\tau)}{F(\tau)\,F(\tau-0)}.$$

A. 4 Lemma. *Let $f : \mathbb{R} \to \mathbb{R}$ be a right continuous function of locally bounded variation, g a measurable function, and*

$$F(t) := \int_{(s,t]} |g(\tau)|\,|f|(d\tau) < \infty \tag{A. 4. 1}$$

for $s < t \le t^*$. Then $F(t)$ is right continuous on $[s, t^*)$ and $F(s) = 0$.

For the following theorem, we refer to theorem 12.26 in Milbrodt and Helbig [1999] or Jacod and Shiryaev ([1987], theorem I4.40).

A. 5 Theorem. *Let X and Y be square integrable càdlàg martingales. Further, let U and V be predictable càdlàg processes satisfying*

$$\mathbf{E}\Big[\int_{[0,t]} U_s^2\, d\langle X\rangle_s\Big] < \infty, \quad \textit{and} \quad \mathbf{E}\Big[\int_{[0,t]} V_s^2\, d\langle Y\rangle_s\Big] < \infty, \;\; t \geq 0. \tag{A. 5. 1}$$

Then, the stochastic integrals $\int U\, dX$ and $\int V\, dY$ are square integrable martingales with covariation

$$\Big\langle \int U\, dX, \int V\, dY \Big\rangle_t = \int_{[0,t]} U_s\, V_s\, d\langle X, Y\rangle_s, \;\; t \geq 0, \tag{A. 5. 2}$$

(up to P- indistinguishability). Further,

$$\mathbf{E}\Big[\Big(\int_{[0,t]} U_s\, dX_s\Big)\Big(\int_{[0,t]} V_s\, dY_s\Big)\Big] = \mathbf{E}\Big[\int_{[0,t]} U_s V_s\, d\langle X, Y\rangle_s\Big], \;\; t \geq 0. \tag{A. 5. 3}$$

2 Proofs

A. 6 Proof of theorem 2. 23:
It remains to verify that the provided versions of $p_{yz}(s,t,u,v)$ satisfy (2. 22. 1) without exceptional sets. Due to the right-hand side of (2. 23. 2) being specified with a regular version of the conditional distribution of $((T_l, Z_l))_{l\in\mathbb{N}_0}$ given (T_0, Z_0), one obtains for each $s \geq 0$ and $(y,u) \in \mathcal{S} \times [0,\infty)$ a measure on $\mathcal{K}$. According to this, (2. 22. 1) can be derived as follows, with all equations holding without exceptional sets. Let $s \leq r \leq t$. Then

$$\begin{aligned}
p_{yz}(s,t,u,v) &= P(\exists l \in \mathbb{N}_0 : T_{l+1} > t, T_l \in [t-v,t], Z_l = z |\, T_1 > s, T_0 = s-u, Z_0 = y) \\
&= \sum_{l\in\mathbb{N}_0} P(T_{l+1} > t, T_l \in [t-v,t], Z_l = z |\, T_1 > s, T_0 = s-u, Z_0 = y) \\
&= \sum_{l\in\mathbb{N}_0} \sum_{k\leq l} \sum_{\xi\in\mathcal{S}} P(T_{l+1} > t, T_l \in [t-v,t], Z_l = z, \\
&\qquad\qquad T_{k+1} > r, s-u \leq T_k \leq r, Z_k = \xi |\, T_1 > s, T_0 = s-u, Z_0 = y) \\
&= \sum_{l\in\mathbb{N}_0} \sum_{k\leq l} \sum_{\xi\in\mathcal{S}} \int_{[t-v,t]} \int_{[s-u,r]} P(T_{l+1} > t, T_l \in dl, Z_l = z, \\
&\qquad\qquad T_{k+1} > r, T_k \in dx, Z_k = \xi |\, T_1 > s, T_0 = s-u, Z_0 = y).
\end{aligned}$$

Upon successive conditioning and applying the Markov property of the marked point process (T, Z), the above chain of equations can be continued as

$$\begin{aligned}
&= \sum_{l\in\mathbb{N}_0} \sum_{k\leq l} \sum_{\xi\in\mathcal{S}} \int_{[t-v,t]} \int_{[s-u,r]} P(T_{l+1} > t, T_l \in dl, Z_l = z |\, T_{k+1} > r, T_k = x, Z_k = \xi) \\
&\qquad\qquad \cdot P(T_{k+1} > r, T_k \in dx, Z_k = \xi |\, T_1 > s, T_0 = s-u, Z_0 = y) \\
&= \sum_{l\in\mathbb{N}_0} \sum_{k\leq l} \sum_{\xi\in\mathcal{S}} \int_{[s-u,r]} P(T_{l+1} > t, T_l \in [t-v,t], Z_l = z |\, T_{k+1} > r, T_k = x, Z_k = \xi) \\
&\qquad\qquad \cdot P(T_{k+1} > r, T_k \in dx, Z_k = \xi |\, T_1 > s, T_0 = s-u, Z_0 = y).
\end{aligned}$$

Now change the order of the double series according to Cauchy's theorem on double series. Afterwards, by interchanging the infinite sums and the integral using the monotone convergence theorem, we get

$$\begin{aligned}
&p_{yz}(s,t,u,v)\\
&=\sum_{\xi\in\mathcal{S}}\int_{[s-u,r]}\sum_{k\in\mathbb{N}_0}\sum_{l\geq k}P(T_{l+1}>t,T_l\in[t-v,t],Z_l=z|\,T_{k+1}>r,T_k=x,Z_k=\xi)\\
&\qquad\qquad\cdot P(T_{k+1}>r,T_k\in dx,Z_k=\xi|\,T_1>s,T_0=s-u,Z_0=y)\\
&=\sum_{\xi\in\mathcal{S}}\int_{[s-u,r]}\sum_{k\in\mathbb{N}_0}P(\exists l\geq k:T_{l+1}>t,T_l\in[t-v,t],Z_l=z|\,T_{k+1}>r,T_k=r-r+x,Z_k=\xi)\\
&\qquad\qquad\cdot P(T_{k+1}>r,T_k\in dx,Z_k=\xi|\,T_1>s,T_0=s-u,Z_0=y),
\end{aligned}$$

which is due to the homogeneity of (T,Z) and $(2.23.2)$ equal to

$$\sum_{\xi\in\mathcal{S}}\int_{[s-u,r]}p_{\xi z}(r,t,r-x,v)\,p_{y\xi}(s,r,u,r-dx).$$

Using the measurable function $I:x\mapsto r-x$ with $\mathbf{1}_{[s-u,r]}(x)=\mathbf{1}_{[0,r-s+u]}(I(x))$, we obtain by integration with respect to the corresponding image measure

$$\begin{aligned}
p_{yz}(s,t,u,v)&=\sum_{\xi\in\mathcal{S}}\int_{[s-u,r]}p_{\xi z}(r,t,r-x,v)\,p_{y\xi}(s,r,u,r-dx)\\
&=\sum_{\xi\in\mathcal{S}}\int_{[0,r-s+u]}p_{\xi z}(r,t,x,v)\,p_{y\xi}(s,r,u,dx)\\
&=\sum_{\xi\in\mathcal{S}}\int_{[0,\infty)}p_{\xi z}(r,t,x,v)\,p_{y\xi}(s,r,u,dx),
\end{aligned}$$

which is $(2.22.1)$. □

A. 7 Proof of corollary 2. 37:
Let $0\leq u\leq s\leq t<\infty, v\geq 0$, and $(y,z)\in\mathcal{S}^2$. In the situation of $(2.36.6)$ and $(2.36.7)$, the exponential formula $(2.24.7)$ gives

$$\bar{p}_y(s,t,u)=e^{q_{yy}(s,t,u)}=\exp\Big\{\int_s^t\mu_{yy}(r,r-s+u)\,dr\Big\}. \qquad (A.7.1)$$

Differentiating this with respect to t yields according to the fundamental theorem of calculus

$$\frac{\partial\bar{p}_y(s,t,u)}{\partial t}=\bar{p}_y(s,t,u)\,\mu_{yy}(t,t-s+u). \qquad (A.7.2)$$

In order to differentiate (A. 7. 1) w.r.t. u, the so called differentiation lemma (cf. e.g. Bauer, [2001] § 16) must be applied. According to this, one obtains

$$\frac{\partial\bar{p}_y(s,t,u)}{\partial u}=\bar{p}_y(s,t,u)\Big(\int_s^t\frac{\partial}{\partial l}\mu_{yy}(r,l)|_{l=r-s+u}\,dr\Big), \qquad (A.7.3)$$

with the differentiation under the integral sign being allowed by the existence of an λ^1-integrable function h with $\sup_{l\in[0,r]}|\frac{\partial}{\partial l}\mu_{yy}(r,l)|\leq h(r)$, which is a consequence of the stipulated continuity. In order to differentiate (A. 7. 1) w.r.t. s, we define for fixed (t,u)

$$F(x,y):=\exp\Big\{\int_x^t\mu_{yy}(r,r-y+u)\,dr\Big\},\quad x,y\leq t, \qquad (A.7.4)$$

with

$$\frac{\partial F}{\partial x}(x,y) = -F(x,y)\,\mu_{yy}(x, x-y+u), \tag{A.7.5}$$

and

$$\frac{\partial F}{\partial y}(x,y) = -F(x,y)\Big(\int_x^t \frac{\partial}{\partial l}\mu_{yy}(r,l)|_{l=r-y+u}\,dr\Big). \tag{A.7.6}$$

Due to the continuity of both $\mu_{yy}(\cdot,\cdot)$ and $\frac{\partial}{\partial l}\mu_{yy}(\cdot,\cdot)$, it follows from the continuity lemma (see also e.g. Bauer, [2001] § 16) that both partial derivatives are continuous. Hence, $F(x,y)$ is differentiable and we get by applying the chain rule

$$\begin{aligned}\frac{\partial \bar{p}_y(s,t,u)}{\partial s} &= \frac{dF(s,s)}{ds} = \frac{\partial F}{\partial x}(s,s) + \frac{\partial F}{\partial y}(s,s) \qquad (A.7.7)\\ &= -\bar{p}_y(s,t,u)\,\mu_{yy}(s,u) - \bar{p}_y(s,t,u)\Big(\int_s^t \frac{\partial}{\partial l}\mu_{yy}(r,l)|_{l=r-s+u}\,dr\Big),\end{aligned}$$

which is by inserting (A.7.3) equal to

$$\frac{\partial \bar{p}_y(s,t,u)}{\partial s} = -\bar{p}_y(s,t,u)\,\mu_{yy}(s,u) - \frac{\partial \bar{p}_y(s,t,u)}{\partial u}. \tag{A.7.8}$$

Now consider the backward integral equations (2.32.2). In the present framework they are for $\mathcal{L}(X_s, U_s|\,P)$-a.e. (y,u) given as

$$\begin{aligned}p_{yz}(s,t,u,v) &= \delta_{yz}\,\mathbf{1}_{(v\ge t-s+u)}\,\exp\Big\{\int_s^t \mu_{yy}(r, r-s+u)\,dr\Big\}\\ &\quad+\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,\exp\Big\{\int_s^\tau \mu_{yy}(r,r-s+u)\,dr\Big\}\,\mu_{y\xi}(\tau,\tau-s+u)\,d\tau\\ &= \delta_{yz}\,\mathbf{1}_{(v\ge t-s+u)}\,\bar{p}_y(s,t,u) + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,\bar{p}_y(s,\tau,u)\,\mu_{y\xi}(\tau,\tau-s+u)\,d\tau.\end{aligned}$$

In case of either $y \neq z$ or $v < t-s+u$, this reduces to

$$p_{yz}(s,t,u,v) = \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,\bar{p}_y(s,\tau,u)\,\mu_{y\xi}(\tau,\tau-s+u)\,d\tau. \tag{A.7.9}$$

We start by differentiating (A.7.9) w.r.t. u. According to the stated assumptions and the above arguments, it can be differentiated under the integral sign yielding

$$\begin{aligned}\frac{p_{yz}(s,t,u,v)}{\partial u} &= \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}}\int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,\frac{\partial \bar{p}_y(s,\tau,u)}{\partial u}\,\mu_{y\xi}(\tau,\tau-s+u)\,d\tau\\ &\quad+\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}}\int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,\bar{p}_y(s,\tau,u)\,\frac{\partial}{\partial l}\,\mu_{y\xi}(\tau,l)|_{l=\tau-s+u}\,d\tau. \qquad (A.7.10)\end{aligned}$$

Differentiating (A.7.9) w.r.t. s in the same manner as in (A.7.7), we get

$$\begin{aligned}\frac{\partial p_{yz}(s,t,u,v)}{\partial s} &= -\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} p_{\xi z}(s,t,0,v)\,\bar{p}_y(s,s,u)\,\mu_{y\xi}(s,s-s+u)\\ &\quad+\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}}\int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,\frac{\partial \bar{p}_y(s,\tau,u)}{\partial s}\,\mu_{y\xi}(\tau,\tau-s+u)\,d\tau\\ &\quad-\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}}\int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,\bar{p}_y(s,\tau,u)\,\frac{\partial}{\partial l}\,\mu_{y\xi}(\tau,l)|_{l=\tau-s+u}\,d\tau. \qquad (A.7.11)\end{aligned}$$

By employing $\bar{p}_y(s,s,u)=1$, (A. 7. 8), (A. 7. 10), and (A. 7. 9), one obtains

$$\begin{aligned}
\frac{\partial p_{yz}(s,t,u,v)}{\partial s} &= -\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} p_{\xi z}(s,t,0,v)\,\mu_{y\xi}(s,u)\\
&\quad -\mu_{yy}(s,u)\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}}\int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,\bar{p}_y(s,\tau,u)\,\mu_{y\xi}(\tau,\tau-s+u)\,d\tau\\
&\quad -\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}}\int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,\frac{\partial \bar{p}_y(s,\tau,u)}{\partial u}\,\mu_{y\xi}(\tau,\tau-s+u)\,d\tau\\
&\quad -\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}}\int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,\bar{p}_y(s,\tau,u)\,\frac{\partial}{\partial l}\,\mu_{y\xi}(\tau,l)|_{l=\tau-s+u}\,d\tau\\
&= -\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} p_{\xi z}(s,t,0,v)\,\mu_{y\xi}(s,u)\\
&\quad -\mu_{yy}(s,u)\,p_{yz}(s,t,u,v)-\frac{p_{yz}(s,t,u,v)}{\partial u}\\
&= -\sum_{\xi\in\mathcal{S}}\mu_{y\xi}(s,u)\,p_{\xi z}(s,t,0,v)\;-\frac{p_{yz}(s,t,u,v)}{\partial u}, \qquad (A.\,7.\,12)
\end{aligned}$$

which is the assertion (2. 37. 1). In the case $y=z$ and $v\geq t-s+u$, the backward integral equation (2. 32. 2) is given by

$$p_{yz}(s,t,u,v) = \bar{p}_y(s,t,u)+\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,\bar{p}_y(s,\tau,u)\,\mu_{y\xi}(\tau,\tau-s+u)\,d\tau.$$

Differentiating this w.r.t. s also yields, by employing the above results, the assertion (2. 37. 1).

In order to establish the forward differential equations (2. 37. 2), we also start by differentiating the corresponding integral equations in the case of either $y\neq z$ or $v<t-s+u$. The forward integral equations (2. 36. 1) are then for $\mathcal{L}(X_s,U_s|\,P)$-a.e. (y,u) given by

$$p_{yz}(s,t,u,v) = \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}}\int_{[t-v,t]}\int_{(0,\infty)} \bar{p}_z(\tau,t,0)\,\mu_{\xi z}(\tau,l)\,p_{y\xi}(s,\tau,u,dl)\,\lambda^1(d\tau). \qquad (A.\,7.\,13)$$

We firstly differentiate (A. 7. 13) with respect to v. For this, it is represented in a different way. According to the theorem on integration with respect to an image measure applied to the measurable function $I:\tau\mapsto t-\tau$, we get

$$\begin{aligned}
p_{yz}(s,t,u,v) &= \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}}\int_{[0,v]}\int_{(0,\infty)} \bar{p}_z(t-\tau,t,0)\,\mu_{\xi z}(t-\tau,l)\,p_{y\xi}(s,t-\tau,u,dl)\,\lambda^1(t-d\tau)\\
&= \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}}\int_{[0,v]}\int_{(0,\infty)} \bar{p}_z(t-\tau,t,0)\,\mu_{\xi z}(t-\tau,l)\,p_{y\xi}(s,t-\tau,u,dl)\,\lambda^1(d\tau), \quad (A.\,7.\,14)
\end{aligned}$$

where the last equation is due to the translation-invariance of the Lebesgue measure λ^1. Differentiating this with respect to v gives

$$\frac{\partial p_{yz}(s,t,u,v)}{\partial v} = \bar{p}_z(t-v,t,0)\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}}\int_{(0,\infty)} \mu_{\xi z}(t-v,l)\,p_{y\xi}(s,t-v,u,dl). \quad (A.\,7.\,15)$$

Now, (A. 7. 13) is differentiated with respect to t, with the differentiation under the integral sign being allowed according to the assumptions stipulated. Doing so, we get

$$\begin{aligned}
&\frac{\partial p_{yz}(s,t,u,v)}{\partial t} \\
&= \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{[t-v,t]} \int_{(0,\infty)} \frac{\partial}{\partial t}\Big(\exp\Big\{ \int_{\tau}^{t} \mu_{zz}(r,r-\tau)\,dr\Big\}\,\mu_{\xi z}(\tau,l)\,p_{y\xi}(s,\tau,u,dl)\Big)\,\lambda^1(d\tau) \\
&\quad + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{(0,\infty)} \mu_{\xi z}(t,l)\,p_{y\xi}(s,t,u,dl) \\
&\quad - \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{(0,\infty)} \exp\Big\{ \int_{t-v}^{t} \mu_{zz}(r,r-t-v)\,dr\Big\}\,\mu_{\xi z}(t-v,l)\,p_{y\xi}(s,t-v,u,dl) \\
&= \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{[t-v,t]} \int_{(0,\infty)} \exp\Big\{ \int_{\tau}^{t} \mu_{zz}(r,r-\tau)\,dr\Big\}\,\mu_{z}(t,t-\tau)\,\mu_{\xi z}(\tau,l)\,p_{y\xi}(s,\tau,u,dl)\,\lambda^1(d\tau) \\
&\quad + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{(0,\infty)} \mu_{\xi z}(t,l)\,p_{y\xi}(s,t,u,dl) \\
&\quad - \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{(0,\infty)} \exp\Big\{ \int_{t-v}^{t} \mu_{zz}(r,r-t-v)\,dr\Big\}\,\mu_{\xi z}(t-v,l)\,p_{y\xi}(s,t-v,u,dl) \\
&= \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{[0,v]} \int_{(0,\infty)} \bar{p}_z(t-\tau,t,0)\,\mu_{zz}(t,\tau)\,\mu_{\xi z}(t-\tau,l)\,p_{y\xi}(s,t-\tau,u,dl)\,\lambda^1(d\tau) \\
&\quad + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{(0,\infty)} \mu_{\xi z}(t,l)\,p_{y\xi}(s,t,u,dl) \\
&\quad - \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{(0,\infty)} \bar{p}_z(t-v,t,0)\,\mu_{\xi z}(t-v,l)\,p_{y\xi}(s,t-v,u,dl). \qquad \text{(A. 7. 16)}
\end{aligned}$$

For the last equations, the first addend was represented in a manner similar to (A. 7. 14). Altogether, we get from (A. 7. 16)

$$\begin{aligned}
\frac{\partial p_{yz}(s,t,u,v)}{\partial t} &= \int_{[0,v]} \mu_{zz}(t,\tau)\,\bar{p}_z(t-\tau,t,0) \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{(0,\infty)} \mu_{\xi z}(t-\tau,l)\,p_{y\xi}(s,t-\tau,u,dl)\,d\tau \\
&\quad + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{(0,\infty)} \mu_{\xi z}(t,l)\,p_{y\xi}(s,t,u,dl) \\
&\quad - \bar{p}_z(t-v,t,0) \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{(0,\infty)} \mu_{\xi z}(t-v,l)\,p_{y\xi}(s,t-v,u,dl). \qquad \text{(A. 7. 17)}
\end{aligned}$$

By inserting (A. 7. 15), we obtain the forward differential equations (2. 37. 2) for $\mathcal{L}(X_s, U_s|\, P)$-a.e. (y,u) in the case of $y \neq z$ or $v < t-s+u$. Regarding the remaining, one can argue in almost the same manner as for the backward differential equations. □

A. 8 Proof of lemma 2. 40:

In order to prove the backward integral equations (2. 40. 1), we start with the equations (2. 32. 2) which are assumed to be identically satisfied. Thus, we have for $0 \leq u \leq s \leq t, v \geq 0$, and

$(y,z) \in \mathcal{S}^2$

$$\begin{aligned}
&p_{yz}(s,t,u,v) \\
&= \delta_{yz}\mathbf{1}_{(v\geq t-s+u)}\big(1-Q_y(s,t,u)\big) \\
&\quad+\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,\big(1-Q_y(s,\tau-0,u)\big)\,q_{y\xi}(s,d\tau,u) \\
&= \delta_{yz}\mathbf{1}_{(v\geq t-s+u)} - \delta_{yz}\mathbf{1}_{(v\geq t-s+u)}\,Q_y(s,t,u) \\
&\quad+\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,q_{y\xi}(s,d\tau,u) - \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,Q_y(s,\tau-0,u)\,q_{y\xi}(s,d\tau,u).
\end{aligned}$$

Inserting (2. 39. 1) gives

$$\begin{aligned}
&p_{yz}(s,t,u,v) \\
&= \delta_{yz}\,\mathbf{1}_{(v\geq t-s+u)} + \delta_{yz}\,\mathbf{1}_{(v\geq t-s+u)}\int_{(s,t]}\big(1-Q_y(r,t,r-s+u)\big)\,q_{yy}(s,dr,u) \\
&\quad+\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,q_{y\xi}(s,d\tau,u) \\
&\quad+\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,t]}\int_{(s,\tau)} p_{\xi z}(\tau,t,0,v)\,\big(1-Q_y(r,\tau-0,r-s+u)\big)\,q_{yy}(s,dr,u)\,q_{y\xi}(s,d\tau,u) \\
&= \delta_{yz}\mathbf{1}_{(v\geq t-s+u)} + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,q_{y\xi}(s,d\tau,u) \\
&\quad+\int_{(s,t]}\delta_{yz}\mathbf{1}_{(v\geq t-s+u)}\,\big(1-Q_y(r,t,r-s+u)\big)\,q_{yy}(s,dr,u) \\
&\quad+\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,t]}\int_{(s,t]}\mathbf{1}_{(r<\tau)}\,p_{\xi z}(\tau,t,0,v)\,\big(1-Q_y(r,\tau-0,r-s+u)\big)\,q_{yy}(s,dr,u)\,q_{y\xi}(s,d\tau,u).
\end{aligned}$$

We now employ Fubini's theorem to change the order of integrations in the last addend. This yields

$$\begin{aligned}
&p_{yz}(s,t,u,v) \\
&= \delta_{yz}\,\mathbf{1}_{(v\geq t-s+u)} + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,q_{y\xi}(s,d\tau,u) \\
&\quad+\int_{(s,t]}\delta_{yz}\,\mathbf{1}_{(v\geq t-s+u)}\,\big(1-Q_y(r,t,r-s+u)\big)\,q_{yy}(s,dr,u) \\
&\quad+\int_{(s,t]}\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(r,t]} p_{\xi z}(\tau,t,0,v)\,\big(1-Q_y(r,\tau-0,r-s+u)\big)\,q_{y\xi}(s,d\tau,u)\,q_{yy}(s,dr,u).
\end{aligned}$$

Due to the property (2. 27. 1), the measures $q_{y\xi}(s,d\tau,u) = q_{y\xi}(r,d\tau,r-s+u)$ coincide on $\mathfrak{B}((r,\infty))$ (cf. (2. 28. 2)). Hence, by inserting the backward integral equations (2. 32. 2), we get the assertion by means of

$$\begin{aligned}
&p_{yz}(s,t,u,v)\\
&= \delta_{yz}\,\mathbf{1}_{(v\geq t-s+u)} + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,q_{y\xi}(s,d\tau,u)\\
&\quad + \int_{(s,t]} \Bigg(\delta_{yz}\mathbf{1}_{(v\geq t-r+r-s+u)} \big(1-Q_y(r,t,r-s+u)\big)\\
&\quad + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \int_{(r,t]} p_{\xi z}(\tau,t,0,v)\,\big(1-Q_y(r,\tau-0,r-s+u)\big)\,q_{y\xi}(r,d\tau,r-s+u)\Bigg)\,q_{yy}(s,dr,u)\\
&= \delta_{yz}\mathbf{1}_{(v\geq t-s+u)} + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,q_{y\xi}(s,d\tau,u) + \int_{(s,t]} p_{yz}(r,t,r-s+u,v)\,q_{yy}(s,dr,u).
\end{aligned}$$

□

A. 9 Proof of lemma 2. 41:
In order to prove the forward integral equations of type 2, we start with the forward integral equations (2. 36. 1). Since these equations are assumed to be identically satisfied, we have for $0 \leq u \leq s \leq t, v \geq 0$, and $(y,z) \in \mathcal{S}^2$

$$\begin{aligned}
p_{yz}(s,t,u,v) &= \delta_{yz}\,\mathbf{1}_{(v\geq t-s+u)}\big(1-Q_y(s,t,u)\big) \qquad \text{(A. 9. 1)}\\
&\quad + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{[t-v,t]} \int_{(0,\infty)} \big(1-Q_z(\tau,t,0)\big)\,\lambda_{\xi z}(\tau,l)\,p_{y\xi}(s,\tau-0,u,dl)\,\mathbf{\Lambda}_{\xi z}(d\tau)\\
&= \delta_{yz}\,\mathbf{1}_{(v\geq t-s+u)} - \delta_{yz}\,\mathbf{1}_{(v\geq t-s+u)}\,Q_y(s,t,u)\\
&\quad + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{[t-v,t]} \int_{(0,\infty)} \lambda_{\xi z}(\tau,l)\,p_{y\xi}(s,\tau-0,u,dl)\,\mathbf{\Lambda}_{\xi z}(d\tau)\\
&\quad - \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{[t-v,t]} \int_{(0,\infty)} Q_z(\tau,t,0)\,\lambda_{\xi z}(\tau,l)\,p_{y\xi}(s,\tau-0,u,dl)\,\mathbf{\Lambda}_{\xi z}(d\tau).
\end{aligned}$$

Applying (2. 28. 5) to both $Q_y(s,t,u)$ and $Q_z(\tau,t,0)$ results in

$$\begin{aligned}
p_{yz}(s,t,u,v) &= \delta_{yz}\,\mathbf{1}_{(v\geq t-s+u)} + \delta_{yz}\,\mathbf{1}_{(v\geq t-s+u)} \int_{(s,t]} \big(1-Q_y(s,r-0,u)\big)\,q_{yy}(s,dr,u)\\
&\quad + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{[t-v,t]} \int_{(0,\infty)} \lambda_{\xi z}(\tau,l)\,p_{y\xi}(s,\tau-0,u,dl)\,\mathbf{\Lambda}_{\xi z}(d\tau)\\
&\quad + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{[t-v,t]} \int_{(0,\infty)} \int_{(\tau,t]} \big(1-Q_z(\tau,r-0,0)\big)\,q_{zz}(\tau,dr,0)\\
&\qquad \cdot\lambda_{\xi z}(\tau,l)\,p_{y\xi}(s,\tau-0,u,dl)\,\mathbf{\Lambda}_{\xi z}(d\tau).
\end{aligned}$$

According to assumption 2. 35, we have $q_{zz}(\tau,dr,0) = \lambda_{zz}(r,r-\tau)\,\mathbf{\Lambda}_{zz}(dr)$. This allows us to change the order of integrations in the last added according to Fubini's theorem. Doing so, we get

$$
\begin{aligned}
&p_{yz}(s,t,u,v)\\
&= \delta_{yz}\,\mathbf{1}_{(v\geq t-s+u)} + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{[t-v,t]}\int_{(0,\infty)} \lambda_{\xi z}(\tau,l)\,p_{y\xi}(s,\tau-0,u,dl)\,\mathbf{\Lambda}_{\xi z}(d\tau)\\
&\quad + \int_{(s,t]} \delta_{yz}\,\mathbf{1}_{(v\geq t-s+u)}\big(1-Q_y(s,r-0,u)\big)\,\lambda_{yy}(r,r-s+u)\,\mathbf{\Lambda}_{yy}(dr)\\
&\quad + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{[t-v,t]}\int_{(0,\infty)}\int_{[t-v,t]} \mathbf{1}_{(\tau<r\leq t)}\big(1-Q_z(\tau,r-0,0)\big)\,\lambda_{zz}(r,r-\tau)\,\mathbf{\Lambda}_{zz}(dr)\\
&\qquad\qquad\qquad \cdot\lambda_{\xi z}(\tau,l)\,p_{y\xi}(s,\tau-0,u,dl)\,\mathbf{\Lambda}_{\xi z}(d\tau)\\
&= \delta_{yz}\,\mathbf{1}_{(v\geq t-s+u)} + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{[t-v,t]}\int_{(0,\infty)} \lambda_{\xi z}(\tau,l)\,p_{y\xi}(s,\tau-0,u,dl)\,\mathbf{\Lambda}_{\xi z}(d\tau)\\
&\quad + \int_{(s,t]} \delta_{yz}\,\mathbf{1}_{(v\geq t-s+u)}\big(1-Q_z(s,r-0,u)\big)\,\lambda_{zz}(r,r-s+u)\,\mathbf{\Lambda}_{zz}(dr)\\
&\quad + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{[t-v,t]}\int_{[t-v,r)}\int_{(0,\infty)} \big(1-Q_z(\tau,r-0,0)\big)\,\lambda_{zz}(r,r-\tau)\\
&\qquad\qquad\qquad \cdot\lambda_{\xi z}(\tau,l)\,p_{y\xi}(s,\tau-0,u,dl)\,\mathbf{\Lambda}_{\xi z}(d\tau)\,\mathbf{\Lambda}_{zz}(dr)\\
&= \delta_{yz}\,\mathbf{1}_{(v\geq t-s+u)} + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{[t-v,t]}\int_{(0,\infty)} \lambda_{\xi z}(\tau,l)\,p_{y\xi}(s,\tau-0,u,dl)\,\mathbf{\Lambda}_{\xi z}(d\tau)\\
&\quad + \int_{(s,t]} \lambda_{zz}(r,r-s+u)\,\delta_{yz}\,\mathbf{1}_{(v\geq t-s+u)}\big(1-Q_z(s,r-0,u)\big)\,\mathbf{\Lambda}_{zz}(dr)\\
&\quad + \int_{[t-v,t]}\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}}\int_{[t-v,r)} \lambda_{zz}(r,r-\tau)\,\big(1-Q_z(\tau,r-0,0)\big)\\
&\qquad\qquad \cdot\int_{(0,\infty)} \lambda_{\xi z}(\tau,l)\,p_{y\xi}(s,\tau-0,u,dl)\,\mathbf{\Lambda}_{\xi z}(d\tau)\,\mathbf{\Lambda}_{zz}(dr). \qquad (\mathrm{A.\,9.\,2})
\end{aligned}
$$

The following equality will be verified below:

$$
\begin{aligned}
&\int_{[t-v,t]}\int_{(0,r-t+v]} \lambda_{zz}(r,l)\,p_{yz}(s,r-0,u,dl)\,\mathbf{\Lambda}_{zz}(dr)\\
&= \int_{(s,t]} \lambda_{zz}(r,r-s+u)\,\delta_{yz}\,\mathbf{1}_{(v\geq t-s+u)}\big(1-Q_z(s,r-0,u)\big)\,\mathbf{\Lambda}_{zz}(dr)\\
&\quad + \int_{[t-v,t]}\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}}\int_{[t-v,r)} \lambda_{zz}(r,r-\tau)\,\big(1-Q_z(\tau,r-0,0)\big)\\
&\qquad\qquad \cdot\int_{(0,\infty)} \lambda_{\xi z}(\tau,l)\,p_{y\xi}(s,\tau-0,u,dl)\,\mathbf{\Lambda}_{\xi z}(d\tau)\,\mathbf{\Lambda}_{zz}(dr). \qquad (\mathrm{A.\,9.\,3})
\end{aligned}
$$

Inserting this into (A. 9. 2) yields the forward integral equations of type 2 (2. 41. 1). Hence, it remains to confirm (A. 9. 3). For this, consider $p_{yz}(s,r-0,u,l)$, $0\leq u\leq s\leq r, l\geq 0$, and rewrite the corresponding forward integral equations (2. 36. 1) (or (A. 9. 1)) as follows:

$$
\begin{aligned}
p_{yz}(s,r-0,u,l) &= \delta_{yz}\,\mathbf{1}_{[r-s+u,\infty)}(l)\big(1-Q_y(s,r-0,u)\big)\\
&\quad + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq z}} \int_{[r-l,r)}\int_{(0,\infty)} \big(1-Q_z(\tau,r-0,0)\big)\,\lambda_{\xi z}(\tau,\sigma)\,p_{y\xi}(s,\tau-0,u,d\sigma)\,\mathbf{\Lambda}_{\xi z}(d\tau).
\end{aligned}
$$

By application of the theorem on integration with respect to an image measure to the measurable function $I : \tau \mapsto r - \tau$ with $\mathbf{1}_{[r-l,r)}(\tau) = \mathbf{1}_{(0,l]}(I(\tau))$, we get

$$\begin{aligned} & p_{yz}(s, r-0, u, l) \\ & = \delta_{yz}\, \mathbf{1}_{[r-s+u,\infty)}(l)\big(1 - Q_y(s, r-0, u)\big) \\ & \quad + \sum_{\substack{\xi \in \mathcal{S} \\ \xi \neq z}} \int_{(0,l]} \int_{(0,\infty)} \big(1 - Q_z(r-\tau, r-0, 0)\big)\, \lambda_{\xi z}(r-\tau, \sigma)\, p_{y\xi}(s, r-\tau-0, u, d\sigma)\, \mathbf{\Lambda}_{\xi z}(r - d\tau). \end{aligned}$$

According to $\mathbf{1}_{[r-s+u,\infty)}(dl) = \varepsilon_{r-s+u}(dl)$, with ε_a being the Dirac measure at a, it follows from the above

$$\begin{aligned} & p_{yz}(s, r-0, u, dl) \\ & = \delta_{yz}\, \varepsilon_{r-s+u}(dl)\big(1 - Q_y(s, r-0, u)\big) \\ & \quad + \sum_{\substack{\xi \in \mathcal{S} \\ \xi \neq z}} \int_{(0,\infty)} \big(1 - Q_z(r-l, r-0, 0)\big)\, \lambda_{\xi z}(r-l, \sigma)\, p_{y\xi}(s, r-l-0, u, d\sigma)\, \mathbf{\Lambda}_{\xi z}(r - dl). \end{aligned}$$

Inserting this into the left-hand side of (A. 9. 3), and afterwards applying the theorem on integration with respect to an image measure to the above function, result in

$$\begin{aligned} & \int_{[t-v,t]} \int_{(0,r-t+v]} \lambda_{zz}(r, l)\, p_{yz}(s, r-0, u, dl)\, \mathbf{\Lambda}_{zz}(dr) \\ & = \int_{[t-v,t]} \int_{(0,r-t+v]} \lambda_{zz}(r, l)\, \delta_{yz}\, \varepsilon_{r-s+u}(dl)\big(1 - Q_y(s, r-0, u)\big)\, \mathbf{\Lambda}_{zz}(dr) \\ & \quad + \int_{[t-v,t]} \sum_{\substack{\xi \in \mathcal{S} \\ \xi \neq z}} \int_{(0,r-t+v]} \lambda_{zz}(r, l)\, \big(1 - Q_z(r-l, r-0, 0)\big) \\ & \qquad \cdot \int_{(0,\infty)} \lambda_{\xi z}(r-l, \sigma)\, p_{y\xi}(s, r-l-0, u, d\sigma)\, \mathbf{\Lambda}_{\xi z}(r - dl)\, \mathbf{\Lambda}_{zz}(dr) \\ & = \int_{[t-v,t]} \lambda_{zz}(r, r-s+u)\, \delta_{yz}\, \mathbf{1}_{(r-s+u \geq r-t+v)}\, \big(1 - Q_z(s, r-0, u)\big)\, \mathbf{\Lambda}_{zz}(dr) \\ & \quad + \int_{[t-v,t]} \sum_{\substack{\xi \in \mathcal{S} \\ \xi \neq z}} \int_{[t-v,r)} \lambda_{zz}(r, r-l)\, \big(1 - Q_z(l, r-0, 0)\big) \\ & \qquad \cdot \int_{(0,\infty)} \lambda_{\xi z}(l, \sigma)\, p_{y\xi}(s, l-0, u, d\sigma)\, \mathbf{\Lambda}_{\xi z}(dl)\, \mathbf{\Lambda}_{zz}(dr) \\ & = \int_{(s,t]} \lambda_{zz}(r, r-s+u)\, \delta_{yz}\, \mathbf{1}_{(v \geq t-s+u)}\, \big(1 - Q_z(s, r-0, u)\big)\, \mathbf{\Lambda}_{zz}(dr) \\ & \quad + \int_{[t-v,t]} \sum_{\substack{\xi \in \mathcal{S} \\ \xi \neq z}} \int_{[t-v,r)} \lambda_{zz}(r, r-l)\, \big(1 - Q_z(l, r-0, 0)\big) \\ & \qquad \cdot \int_{(0,\infty)} \lambda_{\xi z}(l, \sigma)\, p_{y\xi}(s, l-0, u, d\sigma)\, \mathbf{\Lambda}_{\xi z}(dl)\, \mathbf{\Lambda}_{zz}(dr). \end{aligned}$$

The last equation holds since $v \geq t-s+u$ implies $t-v \leq s$, and since the measure $q_{zz}(s, dr, u) = \lambda_{zz}(r, r-s+u)\, \mathbf{\Lambda}_{zz}(dr)$ is concentrated on $\mathfrak{B}((s,\infty))$. Thus, the integration interval $[t-v, t]$ can be restricted to $(s, t]$. Hence, (A. 9. 3) is verified and the proof of the forward integral equations of type 2 is complete. □

A. 10 Proof of theorem 2. 42:
As a consequence of (2. 41. 10), the one-step transition probabilities of the homogeneous Markovian marked point process (T, Z) are according to lemma 2. 19 for $s \geq 0, B \in \mathfrak{B}([0, \infty))$ and $(y, z) \in \mathcal{J}$ given by

$$\hat{Q}_{yz}(s, B) \quad = \quad \int_B -\frac{d\hat{q}_{yz}}{d\hat{q}_{yy}}(\tau)\, \hat{Q}_y(s, d\tau), \tag{A. 10. 1}$$

where

$$\hat{Q}_y(s, t) := 1 - \exp\{\hat{q}_{yy}^{(c)}(s, t)\} \prod_{s<\tau\leq t} (1 + \Delta\hat{q}_{yy}(s, \tau)), \;\; 0 \leq s \leq t.$$

Since G^{-1} (see (2. 3. 4)) is a measurable mapping $G^{-1} : \mathcal{K} \to \mathcal{X}$, the distribution of the appertaining pure jump process $X = G^{-1}(T, Z)$ is uniquely determined by the distribution of (T, Z). Recall that according to theorem 2. 12, the distribution of the latter is uniquely determined by the cumulative transition intensity matrix $\hat{q}$, along with an initial distribution $\mathcal{L}(T_0, Z_0) = \varepsilon_0 \otimes \pi$. Since the paths of (T, Z) are elements of $\mathcal{K}$, the paths of X are contained in $\mathcal{X}$. Under the assumptions stipulated above, it then follows with (2. 19. 5) (which corresponds to Hilfssatz 4.39 in Milbrodt and Helbig [1999]), Hilfssatz 4.40, and Hilfssatz 4.41 (both in Milbrodt and Helbig [1999]) that the process $X = G^{-1}((T, Z))$ is a Markovian pure jump process process with transition probabilities according to (2. 42. 1).
In order to verify the assertion (2. 42. 3), we argue in almost the same manner as Milbrodt and Helbig ([1999], Hilfssatz 4.42): For $s < t$ and $(y, z) \in \mathcal{J}$ with $P(X_s = y) > 0$, one obtains by employing the definition of E_{yz}, (2. 8. 1), and (A. 10. 1)

$$\begin{aligned}
& E_{yz}(s, [t, \infty]) \\
& = P(T(s) \geq t, X_{T(s)} = z |\, X_s = y) \\
& = \frac{1}{P(X_s = y)} \sum_{m=0}^{\infty} P(T_m \leq s < t \leq T_{m+1}, Z_{m+1} = z, Z_m = y) \\
& = \frac{1}{P(X_s = y)} \sum_{m=0}^{\infty} \int_{[0,s]} \int_{[t,\infty)} P(T_{m+1} \in d\tau, Z_{m+1} = z |\, T_m = \sigma, Z_m = y)\, P(T_m \in d\sigma, Z_m = y) \\
& = \frac{1}{P(X_s = y)} \sum_{m=0}^{\infty} \int_{[0,s]} \int_{[t,\infty)} -\frac{d\hat{q}_{yz}}{d\hat{q}_{yy}}(\tau)\, \hat{Q}_y(\sigma, d\tau)\, P(T_m \in d\sigma, Z_m = y).
\end{aligned} \tag{A. 10. 2}$$

Further, employing (2. 9. 2) gives

$$\begin{aligned}
E_{yz}(s, [t, \infty]) &= \frac{1}{P(X_s = y)} \sum_{m=0}^{\infty} \int_{[0,s]} \int_{[t,\infty)} (1 - \hat{Q}_y(\sigma, \tau - 0))\, \hat{q}_{yz}(d\tau)\, P(T_m = \sigma, Z_m = y) \\
&= \int_{[t,\infty)} \frac{1}{P(X_s = y)} \sum_{m=0}^{\infty} \int_{[0,s]} (1 - \hat{Q}_y(\sigma, \tau - 0))\, P(T_m = \sigma, Z_m = y)\, \hat{q}_{yz}(d\tau).
\end{aligned}$$

Hence, it holds $E_{yz}(s, \cdot) \ll \hat{q}_{yz}(\cdot)$ with density

$$\frac{E_{yz}(s, \cdot)}{\hat{q}_{yz}(\cdot)} : \tau \mapsto \frac{1}{P(X_s = y)} \sum_{m=0}^{\infty} \int_{[0,s]} (1 - \hat{Q}_y(\sigma, \tau - 0))\, P(T_m = \sigma, Z_m = y).$$

On the other side, one obtains from (A. 10. 2) by using (2. 10. 5) the following:

$$\begin{aligned} 1 - E_y(s, t-0) &= E_y(s, [t, \infty]) = \sum_{\substack{\xi \in \mathcal{S} \\ \xi \neq y}} E_{y\xi}(s, [t, \infty]) \\ &= \sum_{\substack{\xi \in \mathcal{S} \\ \xi \neq y}} \frac{1}{P(X_s = y)} \sum_{m=0}^{\infty} \int_{[0,s]} \int_{[t,\infty)} -\frac{d\hat{q}_{y\xi}}{d\hat{q}_{yy}}(\tau)\, \hat{Q}_y(\sigma, d\tau)\, P(T_m = \sigma, Z_m = y) \\ &= \frac{1}{P(X_s = y)} \sum_{m=0}^{\infty} \int_{[0,s]} (1 - \hat{Q}_y(\sigma, \tau - 0))\, P(T_m = \sigma, Z_m = y). \end{aligned}$$

Thus, we get

$$\frac{E_{yz}(s, \cdot)}{\hat{q}_{yz}(\cdot)}(\tau) = 1 - E_y(s, \tau - 0) \tag{A. 10. 3}$$

and, due to (2. 10. 5),

$$-\frac{E_y(s, \cdot)}{\hat{q}_{yy}(\cdot)}(\tau) = 1 - E_y(s, \tau - 0). \tag{A. 10. 4}$$

Hence, $\hat{q}_*$ are the cumulative transition intensities of the Markov process $X = G^{-1}((T, Z))$. Due the right continuity of $\hat{q}_*$, this is also true for $s = t$.
It remains to verify (2. 42. 2). By applying the exponential formula (2. 10. 7) to E_y, it follows

$$\begin{aligned} E_y(s, t) &= 1 - \exp\{q_{yy}^{(c)}(s, t)\} \prod_{s < \tau \leq t} (1 + \Delta q_{yy}(s, \tau)) \\ &= 1 - \exp\{\hat{q}_{yy}^{(c)}(s, t)\} \prod_{s < \tau \leq t} (1 + \Delta \hat{q}_{yy}(s, \tau)) \\ &= \hat{Q}_y(s, t) \end{aligned}$$

for $0 \leq s \leq t$ and $y \in \mathcal{S}$ satisfying $P(X_s = y) > 0$. Thus, one obtains with (2. 9. 1) and (A. 10. 3) the assertion (2. 42. 2) by means of

$$\hat{Q}_{yz}(s, d\tau) = (1 - \hat{Q}_y(s, \tau - 0))\, \hat{q}_{yz}(d\tau) = (1 - E_y(s, \tau - 0))\, q_{yz}(d\tau) = E_{yz}(s, d\tau).$$

□

A. 11 Proof of lemma 3. 8:
For $0 \leq u \leq s \leq t$ and $y \in \mathcal{S}$, we get by using the product formula (cf. theorem A.1)

$$\begin{aligned} & v(t)\, \bar{p}_y(s, t, u) - v(s) \\ & = v(t)\, \bar{p}_y(s, t, u) - v(s)\, \bar{p}_y(s, s, u) \\ & = \int_{(s,t]} \bar{p}_y(s, \tau - 0, u)\, v(d\tau) + \int_{(s,t]} v(\tau)\, \bar{p}_y(s, d\tau, u) \\ & = -\int_{(s,t]} \bar{p}_y(s, \tau - 0, u)\, v(\tau)\, \Phi(d\tau) + \int_{(s,t]} v(\tau)\, \bar{p}_y(s, \tau - 0, u)\, q_{yy}(s, d\tau, u), \end{aligned} \tag{A. 11. 1}$$

where the last equation is due to (3. 2. 2) and (2. 28. 5). Now consider

$$K(s) \int_{[s,t]} v(r)\, \bar{p}_y(s, r, u)\, Z(dr) - \int_{[s,t]} Z(dr) = K(s) \int_{[s,t]} \Big(v(r)\, \bar{p}_y(s, r, u) - v(s)\Big)\, Z(dr),$$

insert (A. 11. 1) into the right-hand side, and employ Fubini's theorem. This gives

$$\begin{aligned}
&K(s) \int_{[s,t]} v(r)\, \bar{p}_y(s,r,u)\, Z(dr) - \int_{[s,t]} Z(dr) \\
&= -K(s) \int_{[s,t]} \int_{(s,r]} \bar{p}_y(s,\tau-0,u)\, v(\tau)\, \Phi(d\tau)\, Z(dr) \\
&\quad +K(s) \int_{[s,t]} \int_{(s,r]} v(\tau)\, \bar{p}_y(s,\tau-0,u)\, q_{yy}(s,d\tau,u)\, Z(dr) \\
&= -K(s) \int_{(s,t]} \bar{p}_y(s,\tau-0,u)\, Z([\tau,t])\, v(\tau)\, \Phi(d\tau) \\
&\quad +K(s) \int_{(s,t]} v(\tau)\, \bar{p}_y(s,\tau-0,u)\, Z([\tau,t])\, q_{yy}(s,d\tau,u).
\end{aligned}$$

In almost the same manner, assertion (3. 8. 2) follows by using

$$\begin{aligned}
&v(t)\, \bar{p}_y(s,t-0,u) - v(s) \\
&= v(t-0)\, \bar{p}_y(s,t-0,u) - v(s) + \bar{p}_y(s,t-0,u)\, \Delta v(t) \\
&= -\int_{(s,t)} \bar{p}_y(s,\tau-0,u)\, v(\tau)\, \Phi(d\tau) + \int_{(s,t)} v(\tau)\, \bar{p}_y(s,\tau-0,u)\, q_{yy}(s,d\tau,u) \\
&\quad - \bar{p}_y(s,t-0,u)\, v(t)\, \Phi(\{t\}) \\
&= -\int_{(s,t]} \bar{p}_y(s,\tau-0,u)\, v(\tau)\, \Phi(d\tau) + \int_{(s,t)} v(\tau)\, \bar{p}_y(s,\tau-0,u)\, q_{yy}(s,d\tau,u).
\end{aligned}$$

□

A. 12 Proof of theorem 4. 6:
Let $s \geq 0$. In order to verify (4. 6. 1), we employ (4. 3. 1) and insert $\hat{C}(s, ((T_l, Z_l))_{l\geq 0})$ according to (3. 11. 3). This yields for $\mathcal{L}(X_s, U_s|\, P)$-a.e. $(y,u) \in \mathcal{S} \times [0,\infty)$

$$\begin{aligned}
V^+_{(y,u)}(s) &= \int_{\mathcal{K}} \hat{C}(s, ((T_l, Z_l))_{l\geq 0})\, d\mathcal{L}(((T_l, Z_l))_{l\geq 0}|\, T_1 > s, T_0 = s-u, Z_0 = y) \\
&= \int_{\mathcal{K}} \Big[K(s) \sum_{l\in\mathbb{N}_0} \int_{[T_l\vee s, T_{l+1}\vee s)} v(\tau)\, \hat{F}_{Z_l}(T_l, d\tau) \\
&\qquad + K(s) \sum_{l\in\mathbb{N}_0} \frac{D_{Z_l Z_{l+1}}(T_{l+1}, T_{l+1}-T_l)}{K(DT(T_{l+1}))}\, \mathbf{1}_{\{s<T_{l+1}<\infty\}} \Big] \\
&\qquad\qquad \cdot \mathcal{L}(((T_l, Z_l))_{l\geq 0}|\, T_1 > s, T_0 = s-u, Z_0 = y). \qquad \text{(A. 12. 1)}
\end{aligned}$$

Upon multiplying the discounting function, and interchanging the integrals and the infinite sums using the monotone convergence theorem (note that the annuity payments can be split up by means of $\hat{F}_z = \hat{F}^+_z - \hat{F}^-_z,\ z \in \mathcal{S}$), we obtain for $\mathcal{L}(X_s, U_s|\, P)$-a.e. $(y,u) \in \mathcal{S} \times [0,\infty)$

$$\begin{aligned}
&v(s)\, V^+_{(y,u)}(s) \\
&= \sum_{l\in\mathbb{N}_0} \int_{\mathcal{K}} \Big[\int_{[T_l\vee s, T_{l+1}\vee s)} v(\tau)\, \hat{F}_{Z_l}(T_l, d\tau) \Big]\, \mathcal{L}(((T_l, Z_l))_{l\geq 0}|\, T_1 > s, T_0 = s-u, Z_0 = y) \\
&\quad + \sum_{l\in\mathbb{N}_0} \int_{\mathcal{K}} \Big[\frac{D_{Z_l Z_{l+1}}(T_{l+1}, T_{l+1}-T_l)}{K(DT(T_{l+1}))}\, \mathbf{1}_{\{s<T_{l+1}<\infty\}} \Big]\, \mathcal{L}(((T_l, Z_l))_{l\geq 0}|\, T_1 > s, T_0 = s-u, Z_0 = y).
\end{aligned}$$

Deriving the integrals with respect to the corresponding conditional distributions yields for

$\mathcal{L}(X_s, U_s| P)$-a.e. $(y,u) \in \mathcal{S} \times [0,\infty)$

$$\begin{aligned} v(s)\, V^+_{(y,u)}(s) \;=\; & \sum_{l=0}^{\infty} \sum_{\xi \in \mathcal{S}} \int_{[0,\infty)\times[0,\infty)} \int_{[s\vee r, s\vee x)} v(\tau)\, \hat{F}_\xi(r, d\tau) \\ & \quad \cdot P(T_{l+1} \in dx, T_l \in dr, Z_l = \xi | \, T_1 > s, T_0 = s-u, Z_0 = y) \\ & + \sum_{l=0}^{\infty} \sum_{(\xi,z)\in\mathcal{J}} \int_{[0,\infty)\times[0,\infty)} \frac{D_{\xi z}(x, x-r)}{K(DT(x))} \mathbf{1}_{(s<x)} \\ & \quad \cdot P(T_{l+1} \in dx, Z_{l+1} = z, T_l \in dr, Z_l = \xi | \, T_1 > s, T_0 = s-u, Z_0 = y). \end{aligned}$$

By successive conditioning, employing the Markov property of the marked point process (T, Z), and inserting the transition probabilities (2. 8. 1) as well as (2. 8. 2), the above equation can be continued as

$$\begin{aligned} = \; & \sum_{l=0}^{\infty} \sum_{\xi \in \mathcal{S}} \int_{[0,\infty)} \int_{[0,\infty)} \int_{[0,\infty)} \mathbf{1}_{(s\vee r \le \tau < x)}\, v(\tau)\, \hat{F}_\xi(r, d\tau) \\ & \quad \cdot \hat{Q}_\xi(r, dx)\, P(T_l \in dr, Z_l = \xi | \, T_1 > s, T_0 = s-u, Z_0 = y) \\ & + \sum_{l=0}^{\infty} \sum_{(\xi,z)\in\mathcal{J}} \int_{[0,\infty)} \int_{[0,\infty)} \frac{D_{\xi z}(x, x-r)}{K(DT(x))} \mathbf{1}_{(s<x)} \\ & \quad \cdot \hat{Q}_{\xi z}(r, dx)\, P(T_l \in dr, Z_l = \xi | \, T_1 > s, T_0 = s-u, Z_0 = y). \end{aligned} \tag{A. 12. 2}$$

We now consider the first addend of (A. 12. 2). For this, Fubini's theorem gives

$$\begin{aligned} & \sum_{l=0}^{\infty} \sum_{\xi \in \mathcal{S}} \int_{[0,\infty)} \int_{[0,\infty)} \int_{[0,\infty)} \mathbf{1}_{(s\vee r \le \tau < x)}\, v(\tau)\, \hat{F}_\xi(r, d\tau) \\ & \quad \cdot \hat{Q}_\xi(r, dx)\, P(T_l \in dr, Z_l = \xi | \, T_1 > s, T_0 = s-u, Z_0 = y) \\ = & \sum_{l=0}^{\infty} \sum_{\xi \in \mathcal{S}} \int_{[0,\infty)} \int_{[s\vee r,\infty)} v(\tau)\, \hat{Q}_\xi(r, (\tau,\infty))\, \hat{F}_\xi(r, d\tau) \\ & \quad \cdot P(T_l \in dr, Z_l = \xi | \, T_1 > s, T_0 = s-u, Z_0 = y) \\ = & \sum_{l=0}^{\infty} \sum_{\xi \in \mathcal{S}} \int_{[0,\infty)} \int_{[s,\infty)} v(\tau)\, (1 - \hat{Q}_\xi(r, \tau))\, \hat{F}_\xi(r, d\tau) \\ & \quad \cdot P(T_l \in dr, Z_l = \xi | \, T_1 > s, T_0 = s-u, Z_0 = y), \end{aligned}$$

where the last equation is due to $\hat{F}_\xi(r, dt)$ being concentrated on $[r, \infty)$ and (2. 8. 2). Inserting this into (A. 12. 2) yields for $\mathcal{L}(X_s, U_s| P)$-a.e. $(y,u) \in \mathcal{S} \times [0,\infty)$

$$\begin{aligned} v(s)\, V^+_{(y,u)}(s) \;=\; & \sum_{l=0}^{\infty} \sum_{\xi \in \mathcal{S}} \int_{[0,\infty)} \int_{[s,\infty)} v(\tau)\, (1 - \hat{Q}_\xi(r, \tau))\, \hat{F}_\xi(r, d\tau) \\ & \quad \cdot P(T_l \in dr, Z_l = \xi | \, T_1 > s, T_0 = s-u, Z_0 = y) \\ & + \sum_{l=0}^{\infty} \sum_{(\xi,z)\in\mathcal{J}} \int_{[0,\infty)} \int_{[0,\infty)} \frac{D_{\xi z}(x, x-r)}{K(DT(x))} \mathbf{1}_{(s<x)}\, \hat{Q}_{\xi z}(r, dx) \\ & \quad \cdot P(T_l \in dr, Z_l = \xi | \, T_1 > s, T_0 = s-u, Z_0 = y) \\ = \; & \end{aligned}$$

$$= \sum_{l=0}^{\infty} \sum_{\xi \in \mathcal{S}} \int_{[s-u,\infty)} \int_{[s,\infty)} v(\tau)\,(1-\hat{Q}_\xi(r,\tau))\,\hat{F}_\xi(r,d\tau) \cdot P(T_l \in dr, Z_l = \xi |\, T_1 > s, T_0 = s-u, Z_0 = y) + \sum_{l=0}^{\infty} \sum_{(\xi,z)\in\mathcal{J}} \int_{[s-u,\infty)} \int_{(s,\infty)} \frac{D_{\xi z}(\tau,\tau-r)}{K(DT(\tau))}\,\hat{Q}_{\xi z}(r,d\tau) \cdot P(T_l \in dr, Z_l = \xi |\, T_1 > s, T_0 = s-u, Z_0 = y). \qquad (A.12.3)$$

Upon applying (2.25.3), (4.4.1), and afterwards (2.28.4), we obtain for $\mathcal{L}(X_s, U_s|\, P)$-a.e. $(y,u) \in \mathcal{S} \times [0,\infty)$

$$v(s)\,V^+_{(y,u)}(s) = \sum_{l=0}^{\infty} \sum_{\xi \in \mathcal{S}} \int_{[s-u,\infty)} \int_{[s,\infty)} v(\tau)\,(1-Q_\xi(r,\tau,0))\,F_\xi(r,d\tau,0) \cdot P(T_l \in dr, Z_l = \xi |\, T_1 > s, T_0 = s-u, Z_0 = y) + \sum_{l=0}^{\infty} \sum_{(\xi,z)\in\mathcal{J}} \int_{[s-u,\infty)} \int_{(s,\infty)} \frac{D_{\xi z}(\tau,\tau-r)}{K(DT(\tau))}\,(1-Q_\xi(r,\tau-0,0))\,q_{\xi z}(r,d\tau,0) \cdot P(T_l \in dr, Z_l = \xi |\, T_1 > s, T_0 = s-u, Z_0 = y), \qquad (A.12.4)$$

from which (4.6.1) follows by multiplying $K(s)$.
Further, if the assumptions 2.35 and 4.5 are fulfilled, the right-hand side of (A.12.4) can be written as

$$= \sum_{l=0}^{\infty} \sum_{\xi \in \mathcal{S}} \int_{[s-u,\infty)} \int_{[s,\infty)} v(\tau)\,(1-Q_\xi(r,\tau,0))\,f_\xi(\tau,\tau-r)\,\mathbf{F}_\xi(d\tau) \cdot P(T_l \in dr, Z_l = \xi |\, T_1 > s, T_0 = s-u, Z_0 = y) + \sum_{l=0}^{\infty} \sum_{(\xi,z)\in\mathcal{J}} \int_{[s-u,\infty)} \int_{(s,\infty)} \frac{D_{\xi z}(\tau,\tau-r)}{K(DT(\tau))}\,(1-Q_\xi(r,\tau-0,0))\,\lambda_{\xi z}(\tau,\tau-r)\,\mathbf{\Lambda}_{\xi z}(d\tau) \cdot P(T_l \in dr, Z_l = \xi |\, T_1 > s, T_0 = s-u, Z_0 = y). \qquad (A.12.5)$$

According to these assumptions, the dominating measures $\mathbf{\Lambda}_{\xi z}$ and $\mathbf{F}_\xi$ do not depend on the time elapsed since state ξ was entered. Hence, the order of integrations can be interchanged according to Fubini's theorem. Applying this theorem to both addends in (A.12.5) gives

$$v(s)\,V^+_{(y,u)}(s) = \sum_{l=0}^{\infty} \sum_{\xi \in \mathcal{S}} \int_{[s,\infty)} v(\tau) \int_{[s-u,\infty)} (1-Q_\xi(r,\tau,0))\,f_\xi(\tau,\tau-r) \cdot P(T_l \in dr, Z_l = \xi |\, T_1 > s, T_0 = s-u, Z_0 = y)\,\mathbf{F}_\xi(d\tau) + \sum_{l=0}^{\infty} \sum_{(\xi,z)\in\mathcal{J}} \int_{(s,\infty)} \int_{[s-u,\infty)} \frac{D_{\xi z}(\tau,\tau-r)}{K(DT(\tau))}\,(1-Q_\xi(r,\tau-0,0))\,\lambda_{\xi z}(\tau,\tau-r) \cdot P(T_l \in dr, Z_l = \xi |\, T_1 > s, T_0 = s-u, Z_0 = y)\,\mathbf{\Lambda}_{\xi z}(d\tau) \qquad (A.12.6)$$

for $\mathcal{L}(X_s, U_s|\, P)$-a.e. $(y,u) \in \mathcal{S} \times [0,\infty)$. Further, we argue in almost the same manner as in the proof of lemma 2.36. Doing so, we consider both interior integrals and restrict the integration intervals according to the fact that for $r \geq 0$, the measures $F_\xi(r,d\tau,0) = f_\xi(\tau,\tau-r)\,\mathbf{F}_\xi(d\tau)$ and $q_{\xi z}(r,d\tau,0) = \lambda_{\xi z}(\tau,\tau-r)\,\mathbf{\Lambda}_{\xi z}(d\tau)$ are concentrated on $[r,\infty)$ and (r,∞), respectively. Thus, we get for the first interior integral

$$\begin{aligned}
&\int_{[s-u,\infty)} (1-Q_\xi(r,\tau,0))\, f_\xi(\tau,\tau-r)\, P(T_l\in dr, Z_l=\xi|\, T_1>s, T_0=s-u, Z_0=y)\\
&=\int_{[s-u,\tau]} (1-Q_\xi(r,\tau,0))\, f_\xi(\tau,\tau-r)\, P(T_l\in dr, Z_l=\xi|\, T_1>s, T_0=s-u, Z_0=y)\\
&=\int_{[s-u,\tau]} f_\xi(\tau,\tau-r)\, P(T_{l+1}>\tau, T_l\in dr, Z_l=\xi|\, T_1>s, T_0=s-u, Z_0=y)\\
&=\int_{[s-u,\tau]} f_\xi(\tau,\tau-r)\, p^{(l)}_{y\xi}(s,\tau,u,\tau-dr). \qquad \text{(A. 12. 7)}
\end{aligned}$$

Using the theorem on integration with respect to an image measure applied to the function $I: r\mapsto \tau-r$ with $\mathbf{1}_{[s-u,\tau]}(r)=\mathbf{1}_{[0,\tau-s+u]}(I(r))$, the above chain of equations can be continued as

$$\begin{aligned}
&=\int_{[0,\tau-s+u]} f_\xi(\tau,r)\, p^{(l)}_{y\xi}(s,\tau,u,dr)\\
&=\int_{[0,\infty)} f_\xi(\tau,r)\, p^{(l)}_{y\xi}(s,\tau,u,dr). \qquad \text{(A. 12. 8)}
\end{aligned}$$

For the interior integral of the second addend of (A. 12. 6) we get

$$\begin{aligned}
&\int_{[s-u,\infty)} \frac{D_{\xi z}(\tau,\tau-r)}{K(DT(\tau))}\,(1-Q_\xi(r,\tau-0,0))\,\lambda_{\xi z}(\tau,\tau-r)\\
&\qquad \cdot P(T_l\in dr, Z_l=\xi|\, T_1>s, T_0=s-u, Z_0=y)\\
&=\int_{(0,\tau-s+u]} \frac{D_{\xi z}(\tau,r)}{K(DT(\tau))}\,\lambda_{\xi z}(\tau,r)\, p^{(l)}_{y\xi}(s,\tau-0,u,dr)\\
&=\int_{(0,\infty)} \frac{D_{\xi z}(\tau,r)}{K(DT(\tau))}\,\lambda_{\xi z}(\tau,r)\, p^{(l)}_{y\xi}(s,\tau-0,u,dr). \qquad \text{(A. 12. 9)}
\end{aligned}$$

By inserting (A. 12. 8) and (A. 12. 9) into (A. 12. 6), we obtain for $\mathcal{L}(X_s,U_s|\,P)$-a.e. $(y,u)\in\mathcal{S}\times[0,\infty)$

$$\begin{aligned}
v(s)\,V^+_{(y,u)}(s) \;=\; &\sum_{l=0}^{\infty}\sum_{\xi\in\mathcal{S}}\int_{[s,\infty)}\int_{[0,\infty)} v(\tau)\, f_\xi(\tau,r)\, p^{(l)}_{y\xi}(s,\tau,u,dr)\,\mathbf{F}_\xi(d\tau)\\
&+\sum_{l=0}^{\infty}\sum_{(\xi,z)\in\mathcal{J}}\int_{(s,\infty)}\int_{(0,\infty)} \frac{D_{\xi z}(\tau,r)}{K(DT(\tau))}\,\lambda_{\xi z}(\tau,r)\, p^{(l)}_{y\xi}(s,\tau-0,u,dr)\boldsymbol{\Lambda}_{\xi z}(d\tau).
\end{aligned}$$

Interchanging the infinite sums and the integrals using the monotone convergence theorem, applying (2. 23. 4), and afterwards multiplying $K(s)$, result in the assertion (4. 6. 2). □

A. 13 Proof of theorem 4. 8:
In order to verify (4. 8. 1), we likewise start from (4. 3. 1) and insert (3. 11. 3). Upon splitting up the integrand according to (4. 8. 3) and multiplying the discounting function, we get for $s\geq 0$

and $\mathcal{L}(X_s, U_s|\, P)$-a.e. $(y,u) \in \mathcal{S} \times [0,\infty)$

$$\begin{aligned}
& v(s)\, V^{+}_{(y,u)}(s) \\
&= \int_{\mathcal{K}} \Big[\int_{[T_0 \vee s, T_1 \vee s)} v(\tau)\, \hat{F}_{Z_0}(T_0, d\tau) \Big]\, \mathcal{L}(((T_m, Z_m))_{m\geq 0}|\, T_1 > s, T_0 = s-u, Z_0 = y) \\
&\quad + \int_{\mathcal{K}} \Big[\frac{D_{Z_0 Z_1}(T_1, T_1 - T_0)}{K(DT(T_1))}\, \mathbf{1}_{\{s<T_1<\infty\}} \Big]\, \mathcal{L}(((T_m, Z_m))_{m\geq 0}|\, T_1 > s, T_0 = s-u, Z_0 = y) \\
&\quad + \int_{\mathcal{K}} \Big[\sum_{m=1}^{\infty} \int_{[T_m \vee s, T_{m+1} \vee s)} v(\tau)\, \hat{F}_{Z_m}(T_m, d\tau) \Big]\, \mathcal{L}(((T_m, Z_m))_{m\geq 0}|\, T_1 > s, T_0 = s-u, Z_0 = y) \\
&\quad + \int_{\mathcal{K}} \Big[\sum_{m=1}^{\infty} \frac{D_{Z_m Z_{m+1}}(T_{m+1}, T_{m+1} - T_m)}{K(DT(T_{m+1}))}\, \mathbf{1}_{\{s<T_{m+1}<\infty\}} \Big] \\
&\qquad\qquad \cdot \mathcal{L}(((T_m, Z_m))_{m\geq 0}|\, T_1 > s, T_0 = s-u, Z_0 = y). \qquad \text{(A. 13. 1)}
\end{aligned}$$

Now consider the first two addends of the right-hand side of the above equation. They can directly be derived upon integrating with respect to the corresponding conditional distributions. For the first addend, we get, by inserting (2. 24. 2) (according to the argumentation in (2. 32. 5)) and (4. 4. 1), as well as afterwards employing Fubini's theorem for $\mathcal{L}(X_s, U_s|\, P)$-a.e. $(y,u) \in \mathcal{S} \times [0,\infty)$

$$\begin{aligned}
& \int_{\mathcal{K}} \Big[\int_{[T_0 \vee s, T_1 \vee s)} v(\tau)\, \hat{F}_{Z_0}(T_0, d\tau) \Big]\, \mathcal{L}(((T_m, Z_m))_{m\geq 0}|\, T_1 > s, T_0 = s-u, Z_0 = y) \\
&= \int_{(s,\infty)} \int_{[s,r)} v(\tau)\, \hat{F}_y(s-u, d\tau)\, P(T_1 \in dr \big|\, T_1 > s, T_0 = s-u, Z_0 = y) \\
&= \int_{(s,\infty)} \int_{[s,r)} v(\tau)\, F_y(s, d\tau, u)\, Q_y(s, dr, u) \\
&= \int_{[s,\infty)} v(\tau)\, Q_y(s, (\tau,\infty), u)\, F_y(s, d\tau, u) \\
&= \int_{[s,\infty)} v(\tau)\, (1 - Q_y(s, \tau, u))\, F_y(s, d\tau, u) \\
&= \int_{[s,\infty)} v(\tau)\, \bar{p}_y(s, \tau, u)\, F_y(s, d\tau, u). \qquad \text{(A. 13. 2)}
\end{aligned}$$

By inserting (2. 24. 1), the second addend gives for $\mathcal{L}(X_s, U_s|\, P)$-a.e. $(y,u) \in \mathcal{S} \times [0,\infty)$

$$\begin{aligned}
& \int_{\mathcal{K}} \Big[\frac{D_{Z_0 Z_1}(T_1, T_1 - T_0)}{K(DT(T_1))}\, \mathbf{1}_{\{s<T_1<\infty\}} \Big]\, \mathcal{L}(((T_m, Z_m))_{m\geq 0}|\, T_1 > s, T_0 = s-u, Z_0 = y) \\
&= \sum_{\substack{z\in\mathcal{S} \\ z\neq y}} \int_{(s,\infty)} \frac{D_{yz}(r, r-s+u)}{K(DT(r))}\, P(T_1 \in dr, Z_1 = z \big|\, T_1 > s, T_0 = s-u, Z_0 = y) \\
&= \sum_{\substack{z\in\mathcal{S} \\ z\neq y}} \int_{(s,\infty)} \frac{D_{yz}(r, r-s+u)}{K(DT(r))}\, Q_{yz}(s, dr, u). \qquad \text{(A. 13. 3)}
\end{aligned}$$

We now consider the remaining addends of (A. 13. 1). By using the same arguments yielding (A. 12. 3) in the proof of theorem 4. 6 (see A.12), and afterwards conditioning on (T_1, Z_1) as well as employing the Markov property of (T, Z), the following equations can be verified:

$$
\begin{aligned}
&\int_{\mathcal{K}} \Big[\sum_{m=1}^{\infty} \int_{[T_m \vee s, T_{m+1} \vee s)} v(\tau)\, \hat{F}_{Z_m}(T_m, d\tau) \Big]\, \mathcal{L}(((T_m, Z_m))_{m\geq 0} |\, T_1 > s, T_0 = s-u, Z_0 = y) \\
&+ \int_{\mathcal{K}} \Big[\sum_{m=1}^{\infty} \frac{D_{Z_m Z_{m+1}}(T_{m+1}, T_{m+1} - T_m)}{K(DT(T_{m+1}))}\, \mathbf{1}_{\{s < T_{m+1} < \infty\}} \Big] \\
&\quad \cdot \mathcal{L}(((T_m, Z_m))_{m\geq 0} |\, T_1 > s, T_0 = s-u, Z_0 = y) \\
&= \sum_{m=1}^{\infty} \sum_{\xi \in \mathcal{S}} \int_{[s,\infty)} \int_{[s,\infty)} v(\tau)\,(1 - \hat{Q}_{\xi}(x,\tau))\, \hat{F}_{\xi}(x, d\tau) \\
&\qquad \cdot P(T_m \in dx, Z_m = \xi |\, T_1 > s, T_0 = s-u, Z_0 = y) \\
&\quad + \sum_{m=1}^{\infty} \sum_{(\xi,z)\in\mathcal{J}} \int_{[s,\infty)} \int_{(s,\infty)} \frac{D_{\xi z}(\tau, \tau - x)}{K(DT(\tau))}\, \hat{Q}_{\xi z}(x, d\tau) \\
&\qquad \cdot P(T_m \in dx, Z_m = \xi |\, T_1 > s, T_0 = s-u, Z_0 = y) \\
&= \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} \sum_{m=1}^{\infty} \sum_{\xi\in\mathcal{S}} \int_{[s,\infty)} \int_{[s,\infty)} v(\tau)\,(1 - \hat{Q}_{\xi}(x,\tau))\, \hat{F}_{\xi}(x, d\tau) \\
&\qquad \cdot P(T_m \in dx, Z_m = \xi |\, T_1 = r, Z_1 = z)\, P(T_1 \in dr, Z_1 = z |\, T_1 > s, T_0 = s-u, Z_0 = y) \\
&\quad + \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} \sum_{m=1}^{\infty} \sum_{(\xi,z)\in\mathcal{J}} \int_{[s,\infty)} \int_{(s,\infty)} \frac{D_{\xi z}(\tau, \tau - x)}{K(DT(\tau))}\, \hat{Q}_{\xi z}(x, d\tau) \\
&\qquad \cdot P(T_m \in dx, Z_m = \xi |\, T_1 = r, Z_1 = z)\, P(T_1 \in dr, Z_1 = z |\, T_1 > s, T_0 = s-u, Z_0 = y) \\
&= \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} \Big(\sum_{m=1}^{\infty} \sum_{\xi\in\mathcal{S}} \int_{[s,\infty)} \int_{[s,\infty)} v(\tau)\,(1 - \hat{Q}_{\xi}(x,\tau))\, \hat{F}_{\xi}(x, d\tau) \\
&\qquad \cdot P(T_m \in dx, Z_m = \xi |\, T_1 = r, Z_1 = z) \\
&\quad + \sum_{m=1}^{\infty} \sum_{(\xi,z)\in\mathcal{J}} \int_{[s,\infty)} \int_{(s,\infty)} \frac{D_{\xi z}(\tau, \tau - x)}{K(DT(\tau))}\, \hat{Q}_{\xi z}(x, d\tau)\, P(T_m \in dx, Z_m = \xi |\, T_1 = r, Z_1 = z) \Big) \\
&\qquad \cdot P(T_1 \in dr, Z_1 = z |\, T_1 > s, T_0 = s-u, Z_0 = y). \qquad \text{(A. 13. 4)}
\end{aligned}
$$

Due to the homogeneity of the marked point process (T, Z), the summation indices within the bracket expression can be reduced by 1. Doing so, the bracket expression can, according to (A. 12. 3) and an argumentation similar to (2. 26. 2), be replaced by $v(r)\, V^{+}_{(z,0)}(r)$. Upon afterwards inserting, (2. 24. 1) it turns out that (A. 13. 4) equals for $\mathcal{L}(X_s, U_s |\, P)$-a.e. $(y, u) \in \mathcal{S} \times [0, \infty)$

$$
\sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} v(r)\, V^{+}_{(z,0)}(r)\, Q_{yz}(s, dr, u). \qquad \text{(A. 13. 5)}
$$

Adding (A. 13. 2), (A. 13. 3), and (A. 13. 5) yields for $\mathcal{L}(X_s, U_s |\, P)$-a.e. $(y, u) \in \mathcal{S} \times [0, \infty)$

$$
\begin{aligned}
v(s) V^{+}_{(y,u)}(s) &= \int_{[s,\infty)} v(\tau)\, \bar{p}_y(s, \tau, u)\, F_y(s, d\tau, u) \\
&\quad + \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} \Big(\frac{D_{yz}(\tau, \tau - s + u)}{K(DT(\tau))} + v(\tau)\, V^{+}_{(z,0)}(\tau) \Big)\, Q_{yz}(s, d\tau, u),
\end{aligned}
$$

from which the assertion (4. 8. 1) follows upon multiplying $K(s)$. □

A. 14 Proof of lemma 4. 9:
In order to complete the proof of lemma 4. 9, it remains to verify that (2. 32. 1) inserted into (4. 6. 2) yields the integral equations (4. 8. 1). Under the assumptions 2. 35 and 4. 5, we obtain with (4. 6. 2)

$$v(s)\,V^{+}_{(y,u)}(s) = \sum_{z\in\mathcal{S}}\int_{[s,\infty)}\int_{[0,\infty)} v(\tau)\,f_z(\tau,l)\,p_{yz}(s,\tau,u,dl)\,\mathbf{F}_z(d\tau) + \sum_{(z,\eta)\in\mathcal{J}}\int_{(s,\infty)}\int_{(0,\infty)}\frac{D_{z\eta}(\tau,l)}{K(DT(\tau))}\,\lambda_{z\eta}(\tau,l)\,p_{yz}(s,\tau-0,u,dl)\,\mathbf{\Lambda}_{z\eta}(d\tau) \tag{A. 14. 1}$$

for $\mathcal{L}(X_s,U_s|\,P)$-a.e. $(y,u)\in\mathcal{S}\times[0,\infty)$. Further, the backward integral equations (2. 32. 1) can be written as

$$p_{yz}(s,t,u,v) = \delta_{yz}\,\mathbf{1}_{[t-s+u,\infty)}(v)\,\bar{p}_y(s,t,u) + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,Q_{y\xi}(s,d\tau,u),$$

and accordingly,

$$p_{yz}(s,t-0,u,v) = \delta_{yz}\,\mathbf{1}_{[t-s+u,\infty)}(v)\,\bar{p}_y(s,t-0,u) + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,t)} p_{\xi z}(\tau,t-0,0,v)\,Q_{y\xi}(s,d\tau,u).$$

According to this, we get for $\mathcal{L}(X_s,U_s|\,P)$-a.e. $(y,u)\in\mathcal{S}\times[0,\infty)$

$$p_{yz}(s,\tau,u,dl) = \delta_{yz}\,\bar{p}_y(s,\tau,u)\,\varepsilon_{\tau-s+u}(dl) + \Big(\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,\tau]} p_{\xi z}(r,\tau,0,\cdot)\,Q_{y\xi}(s,dr,u)\Big)(dl),$$

and

$$p_{yz}(s,\tau-0,u,dl) = \delta_{yz}\,\bar{p}_y(s,\tau-0,u)\,\varepsilon_{\tau-s+u}(dl) + \Big(\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,\tau)} p_{\xi z}(r,\tau-0,0,\cdot)\,Q_{y\xi}(s,dr,u)\Big)(dl),$$

which leads by inserting into (A. 14. 1) to the following:

$$\begin{aligned} v(s)\,V^{+}_{(y,u)}(s) &\\ = \int_{[s,\infty)} v(\tau)\,\bar{p}_y(s,\tau,u)\,f_y(\tau,\tau-s+u)\,\mathbf{F}_y(d\tau) &\\ +\sum_{\substack{\eta\in\mathcal{S}\\ \eta\neq y}}\int_{(s,\infty)}\frac{D_{y\eta}(\tau,\tau-s+u)}{K(DT(\tau))}\,\bar{p}_y(s,\tau-0,u)\,\lambda_{y\eta}(\tau,\tau-s+u)\,\mathbf{\Lambda}_{y\eta}(d\tau) &\\ +\sum_{z\in\mathcal{S}}\int_{[s,\infty)}\int_{[0,\infty)} v(\tau)\,f_z(\tau,l)\Big(\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,\tau]} p_{\xi z}(r,\tau,0,\cdot)\,Q_{y\xi}(s,dr,u)\Big)(dl)\,\mathbf{F}_z(d\tau) &\\ +\sum_{(z,\eta)\in\mathcal{J}}\int_{(s,\infty)}\int_{(0,\infty)}\frac{D_{z\eta}(\tau,l)}{K(DT(\tau))}\,\lambda_{z\eta}(\tau,l) &\\ \cdot\Big(\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,\tau)} p_{\xi z}(r,\tau-0,0,\cdot)\,Q_{y\xi}(s,dr,u)\Big)(dl)\,\mathbf{\Lambda}_{z\eta}(d\tau) & \end{aligned} \tag{A. 14. 2}$$

for $\mathcal{L}(X_s, U_s| P)$-a.e. $(y,u) \in \mathcal{S} \times [0,\infty)$. Now consider the latter two addends of the above equation. By repeatedly using Fubini's theorem, and employing (A. 14. 1) again, we get

$$\begin{aligned}
&\sum_{z\in\mathcal{S}} \int_{[s,\infty)} \int_{[0,\infty)} v(\tau)\, f_z(\tau,l) \Big(\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \int_{(s,\tau]} p_{\xi z}(r,\tau,0,\cdot)\, Q_{y\xi}(s,dr,u)\Big)(dl)\, \mathbf{F}_z(d\tau) \\
&+ \sum_{(z,\eta)\in\mathcal{J}} \int_{(s,\infty)} \int_{(0,\infty)} \frac{D_{z\eta}(\tau,l)}{K(DT(\tau))}\, \lambda_{z\eta}(\tau,l) \Big(\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \int_{(s,\tau)} p_{\xi z}(r,\tau-0,0,\cdot)\, Q_{y\xi}(s,dr,u)\Big)(dl)\, \mathbf{\Lambda}_{z\eta}(d\tau) \\
&= \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \sum_{z\in\mathcal{S}} \int_{[s,\infty)} \int_{(s,\tau]} \int_{[0,\infty)} v(\tau)\, f_z(\tau,l)\, p_{\xi z}(r,\tau,0,dl)\, Q_{y\xi}(s,dr,u)\, \mathbf{F}_z(d\tau) \\
&\quad + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \sum_{(z,\eta)\in\mathcal{J}} \int_{(s,\infty)} \int_{(s,\tau)} \int_{(0,\infty)} \frac{D_{z\eta}(\tau,l)}{K(DT(\tau))}\, \lambda_{z\eta}(\tau,l)\, p_{\xi z}(r,\tau-0,0,dl)\, Q_{y\xi}(s,dr,u)\, \mathbf{\Lambda}_{z\eta}(d\tau) \\
&= \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \sum_{z\in\mathcal{S}} \int_{(s,\infty)} \int_{[r,\infty)} \int_{[0,\infty)} v(\tau)\, f_z(\tau,l)\, p_{\xi z}(r,\tau,0,dl)\, \mathbf{F}_z(d\tau)\, Q_{y\xi}(s,dr,u) \\
&\quad + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \sum_{(z,\eta)\in\mathcal{J}} \int_{(s,\infty)} \int_{(r,\infty)} \int_{(0,\infty)} \frac{D_{z\eta}(\tau,l)}{K(DT(\tau))}\, \lambda_{z\eta}(\tau,l)\, p_{\xi z}(r,\tau-0,0,dl)\, \mathbf{\Lambda}_{z\eta}(d\tau)\, Q_{y\xi}(s,dr,u) \\
&= \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \int_{(s,\infty)} \Big(\sum_{z\in\mathcal{S}} \int_{[r,\infty)} \int_{[0,\infty)} v(\tau)\, f_z(\tau,l)\, p_{\xi z}(r,\tau,0,dl)\, \mathbf{F}_z(d\tau) \\
&\qquad + \sum_{(z,\eta)\in\mathcal{J}} \int_{(r,\infty)} \int_{(0,\infty)} \frac{D_{z\eta}(\tau,l)}{K(DT(\tau))}\, \lambda_{z\eta}(\tau,l)\, p_{\xi z}(r,\tau-0,0,dl)\, \mathbf{\Lambda}_{z\eta}(d\tau)\Big)\, Q_{y\xi}(s,dr,u) \\
&= \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \int_{(s,\infty)} v(r)\, V^{+}_{(\xi,0)}(r)\, Q_{y\xi}(s,dr,u). \qquad \text{(A. 14. 3)}
\end{aligned}$$

Inserting this into the right-hand side of (A. 14. 2), recalling the assumptions 2. 35 and 4. 5, and finally using (2. 28. 4) yield for $\mathcal{L}(X_s, U_s| P)$-a.e. $(y,u) \in \mathcal{S} \times [0,\infty)$

$$\begin{aligned}
v(s)\, V^{+}_{(y,u)}(s) &= \int_{[s,\infty)} v(\tau)\, \bar{p}_y(s,\tau,u)\, f_y(\tau,\tau-s+u)\, \mathbf{F}_y(d\tau) \\
&\quad + \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} \frac{D_{yz}(\tau,\tau-s+u)}{K(DT(\tau))}\, \bar{p}_y(s,\tau-0,u)\, \lambda_{yz}(\tau,\tau-s+u)\, \mathbf{\Lambda}_{yz}(d\tau) \\
&\quad + \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} v(r)\, V^{+}_{(z,0)}(r)\, Q_{yz}(s,dr,u) \\
&= \int_{[s,\infty)} v(\tau)\, \bar{p}_y(s,\tau,u)\, F_y(s,d\tau,u) \\
&\quad + \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} \Big(\frac{D_{yz}(\tau,\tau-s+u)}{K(DT(\tau))} + v(\tau)\, V^{+}_{(z,0)}(\tau)\Big)\, Q_{yz}(s,d\tau,u),
\end{aligned}$$

which corresponds to the system of integral equations (4. 8. 1). □

A. 15 Proof of theorem 4. 18:

For the most part, this proof is closely related to the proof of theorem T9 in Brémaud [1981], and likewise to the proof of theorem 3.1 in Møller [1993]. For $t \geq 0$, consider the prospective loss

$$\mathcal{L}(t) = -v(t)\, V^{-}(t,A) + v(t)\, V^{+}(t) - \mathbf{E}[V_0].$$

According to lemma 4.16, $(\mathcal{L}(t))_{t\geq 0}$ is under the assumptions (3.10.1) and (4.7.1) yielding $\mathbf{E}[|V_0|] < \infty$ a centered uniformly integrable martingale. Further, by (3.12.1), (4.2.1) and (2.5.9), the prospective loss up to time $t \geq 0$, $\mathcal{L}(t)$, can be represented as

$$\begin{aligned}\mathcal{L}(t) \;&=\; \sum_{m\in\mathbb{N}_0} \int_{[T_m\wedge t, T_{m+1}\wedge t)} v(\tau)\, \hat{F}_{Z_m}(T_m, d\tau) \;+\; \sum_{m\in\mathbb{N}_0} \frac{D_{Z_m Z_{m+1}}(T_{m+1}, T_{m+1}-T_m)}{K(DT(T_{m+1}))}\, \mathbf{1}_{\{T_{m+1}\leq t\}}\\ &\quad + v(t) \sum_{m\in\mathbb{N}_0} V^{+}_{(Z_m,\, t-T_m)}(t)\, \mathbf{1}_{\{T_m\leq t<T_{m+1}\}} - \mathbf{E}[V_0]. \qquad \text{(A. 15. 1)}\end{aligned}$$

By using (2.4.4), we now decompose $\mathcal{L}(t)$ as follows:

$$\mathcal{L}(t) \;=\; \sum_{m\in\mathbb{N}_0} \mathcal{L}^{(m)}(t)\, \mathbf{1}_{\{N_t=m\}} = \sum_{m\in\mathbb{N}_0} \mathcal{L}^{(m)}(t)\, \mathbf{1}_{\{T_m\leq t<T_{m+1}\}}\,, \qquad \text{(A. 15. 2)}$$

where

$$\begin{aligned}\mathcal{L}^{(m)}(t) \;&:=\; \sum_{n=0}^{m-1} \int_{[T_n, T_{n+1})} v(\tau)\, \hat{F}_{Z_n}(T_n, d\tau) + \int_{[T_m, t)} v(\tau)\, \hat{F}_{Z_m}(T_m, d\tau)\\ &\quad + \sum_{n=0}^{m-1} \frac{D_{Z_n Z_{n+1}}(T_{n+1}, T_{n+1}-T_n)}{K(DT(T_{n+1}))}\\ &\quad + v(t)\, V^{+}_{(Z_m,\, t-T_m)}(t) - \mathbf{E}[V_0], \quad T_m \leq t < T_{m+1}, T_m<\infty. \qquad \text{(A. 15. 3)}\end{aligned}$$

Thus, we obtain on $\{T_m<\infty\}$

$$\mathcal{L}(t)\, \mathbf{1}_{\{T_m\leq t<T_{m+1}\}} = \mathcal{L}^{(m)}(t)\, \mathbf{1}_{\{T_m\leq t<T_{m+1}\}}.$$

Further, the mapping $\Omega\times[0,\infty) \ni (\omega,t) \mapsto \mathcal{L}^{(m)}(t,\omega)$ is $\mathfrak{F}_{T_m}\otimes\mathfrak{B}([0,\infty))$-measurable.
For each $z\in\mathcal{S}$, we define another $\mathfrak{F}_{T_m}\otimes\mathfrak{B}([0,\infty))$-measurable mapping

$$\Omega\times[0,\infty) \ni (\omega,s) \mapsto f_z^{(m)}(s,\omega)$$

by means of

$$\begin{aligned}f_z^{(m)}(s) \;&:=\; \sum_{n=0}^{m-1} \int_{[T_n, T_{n+1})} v(\tau)\, \hat{F}_{Z_n}(T_n, d\tau) + \int_{[T_m, T_m+s)} v(\tau)\, \hat{F}_{Z_m}(T_m, d\tau)\\ &\quad + \sum_{n=0}^{m-1} \frac{D_{Z_n Z_{n+1}}(T_{n+1}, T_{n+1}-T_n)}{K(DT(T_{n+1}))} + \frac{D_{Z_m z}(T_m+s, s)}{K(DT(T_m+s))}\\ &\quad + v(T_m+s)\, V^{+}_{(z,0)}(T_m+s) - \mathbf{E}[V_0], \quad s\geq 0, T_m<\infty. \qquad \text{(A. 15. 4)}\end{aligned}$$

On $\{T_m<\infty\}$, this mapping satisfies for $V_m = T_{m+1}-T_m$ the following:

$$f^{(m)}_{Z_{m+1}}(V_m) = \mathcal{L}(T_{m+1}). \qquad \text{(A. 15. 5)}$$

For the stopping times $T_{m+1}\wedge t$ and T_{m+1} satisfying $T_{m+1}\wedge t \leq T_{m+1}$, it follows from the optional sampling theorem that

$$\mathbf{E}[\mathcal{L}(T_{m+1}\wedge t)\, \mathbf{1}_A] = \mathbf{E}[\mathcal{L}(T_{m+1})\, \mathbf{1}_A], \;\; A\in\mathfrak{F}_{T_{m+1}\wedge t},$$

and therefore,

$$\mathbf{E}[\mathcal{L}(t)\,\mathbf{1}_A\,\mathbf{1}_{\{t<T_{m+1}\}}] = \mathbf{E}[\mathcal{L}(T_{m+1})\,\mathbf{1}_A\,\mathbf{1}_{\{t<T_{m+1}\}}].$$

Now, for each $B \in \mathfrak{F}_{T_{m+1}\wedge t}$, the set $B \cap \{T_m \leq t\}$ is also contained in $\mathfrak{F}_{T_{m+1}\wedge t}$, since $\{T_m \leq t\} = \{T_m \leq t \wedge T_{m+1}\}$ is contained in $\mathfrak{F}_{T_{m+1}\wedge t}$. Hence, for each $B \in \mathfrak{F}_{T_{m+1}\wedge t}$,

$$\mathbf{E}[\mathcal{L}(t)\,\mathbf{1}_B\,\mathbf{1}_{\{T_m\leq t<T_{m+1}\}}] = \mathbf{E}[\mathcal{L}(T_{m+1})\,\mathbf{1}_B\,\mathbf{1}_{\{T_m\leq t<T_{m+1}\}}]. \tag{A. 15. 6}$$

According to the third assertion of theorem 2. 5, the event $C \cap \{T_m \leq t < T_{m+1}\}$ can for each $C \in \mathfrak{F}_{T_m}$ be written as $B \cap \{T_m \leq t < T_{m+1}\}$ for some $B \in \mathfrak{F}_{T_{m+1}\wedge t}$, since $\{T_m \leq t < T_{m+1}\} = \{T_m \leq T_{m+1} \wedge t < T_{m+1}\}$. Hence, equation (A. 15. 6) remains valid for each $C \in \mathfrak{F}_{T_m}$ with $C \subset \{T_m < \infty\}$:

$$\mathbf{E}[\mathcal{L}(t)\,\mathbf{1}_C\,\mathbf{1}_{\{T_m\leq t<T_{m+1}\}}] = \mathbf{E}[\mathcal{L}(T_{m+1})\,\mathbf{1}_C\,\mathbf{1}_{\{T_m\leq t<T_{m+1}\}}]. \tag{A. 15. 7}$$

Now recall that on $\{T_m \leq t < T_{m+1}\}$, the prospective loss $\mathcal{L}(t)$ equals $\mathcal{L}^{(m)}(t)$. Further, since $\mathcal{L}^{(m)}(t)$, $\mathbf{1}_C$, and $\mathbf{1}_{\{T_m\leq t\}}$ are $\mathfrak{F}_{T_m}$-measurable, we get for the left-hand side of (A. 15. 7)

$$\begin{aligned}
\mathbf{E}[\mathcal{L}(t)\,\mathbf{1}_C\,\mathbf{1}_{\{T_m\leq t<T_{m+1}\}}] &= \mathbf{E}[\mathcal{L}^{(m)}(t)\,\mathbf{1}_C\,\mathbf{1}_{\{T_m\leq t<T_{m+1}\}}] \\
&= \mathbf{E}[\mathcal{L}^{(m)}(t)\,\mathbf{1}_C\,\mathbf{1}_{\{T_m\leq t\}}\,\mathbf{1}_{\{T_{m+1}>t\}}] \\
&= \mathbf{E}[\mathbf{E}[\mathcal{L}^{(m)}(t)\,\mathbf{1}_C\,\mathbf{1}_{\{T_m\leq t\}}\,\mathbf{1}_{\{T_{m+1}>t\}}|\mathfrak{F}_{T_m}]] \\
&= \mathbf{E}[\mathcal{L}^{(m)}(t)\,\mathbf{1}_C\,\mathbf{1}_{\{T_m\leq t\}}\,P(T_{m+1}>t|\mathfrak{F}_{T_m})].
\end{aligned}$$

By employing the second assertion of theorem 2. 5, the Markov property of (T, Z), and (2. 8. 2), we finally obtain for the left-hand side of (A. 15. 7)

$$\mathbf{E}[\mathcal{L}(t)\,\mathbf{1}_C\,\mathbf{1}_{\{T_m\leq t<T_{m+1}\}}] = \mathbf{E}[\mathcal{L}^{(m)}(t)\,\mathbf{1}_C\,\mathbf{1}_{\{T_m\leq t\}}\,(1-\hat{Q}_{Z_m}(T_m,t))].$$

For the right-hand side of (A. 15. 7), the following holds by inserting (A. 15. 5) and arguing in a manner similar as above:

$$\begin{aligned}
&\mathbf{E}[\mathcal{L}(T_{m+1})\,\mathbf{1}_C\,\mathbf{1}_{\{T_m\leq t<T_{m+1}\}}] \\
&= \mathbf{E}\left[f^{(m)}_{Z_{m+1}}(V_{m+1})\,\mathbf{1}_C\,\mathbf{1}_{\{T_m\leq t\}}\,\mathbf{1}_{\{T_{m+1}>t\}}\right] \\
&= \mathbf{E}\left[\mathbf{1}_C\,\mathbf{1}_{\{T_m\leq t\}}\,\mathbf{E}\left[f^{(m)}_{Z_{m+1}}(T_{m+1}-T_m)\,\mathbf{1}_{\{T_{m+1}>t\}}\middle|\,\mathfrak{F}_{T_m}\right]\right] \\
&= \mathbf{E}\left[\mathbf{1}_C\,\mathbf{1}_{\{T_m\leq t\}}\,\mathbf{E}\left[f^{(m)}_{Z_{m+1}}(T_{m+1}-T_m)\,\mathbf{1}_{\{T_{m+1}>t\}}\middle|\,T_m, Z_m\right]\right] \\
&= \mathbf{E}\left[\mathbf{1}_C\,\mathbf{1}_{\{T_m<t\}}\int_{(t,\infty]} f^{(m)}_z(s-T_m)\,\mathcal{L}(T_{m+1},Z_{m+1}|\,T_m,Z_m)(ds,dz)\right] \\
&= \mathbf{E}\left[\mathbf{1}_C\,\mathbf{1}_{\{T_m\leq t\}}\sum_{\substack{z\in\mathcal{S}\\ z\neq Z_m}}\int_{(t,\infty]} f^{(m)}_z(s-T_m)\,\hat{Q}_{Z_m z}(T_m,ds)\right].
\end{aligned}$$

Altogether, we obtain from (A. 15. 7)

$$\begin{aligned}
&\mathbf{E}[\mathcal{L}^{(m)}(t)\,\mathbf{1}_C\,\mathbf{1}_{\{T_m\leq t\}}\,(1-\hat{Q}_{Z_m}(T_m,t))] \\
&= \mathbf{E}\left[\mathbf{1}_C\,\mathbf{1}_{\{T_m\leq t\}}\sum_{\substack{z\in\mathcal{S}\\ z\neq Z_m}}\int_{(t,\infty]} f^{(m)}_z(s-T_m)\,\hat{Q}_{Z_m z}(T_m,ds)\right],
\end{aligned} \tag{A. 15. 8}$$

which means that, with the convention $0/0 := 0$,

$$\begin{aligned}
\mathbf{E}\left[\mathcal{L}^{(m)}(t)\,\mathbf{1}_{\{T_m\leq t\}}\,|\mathfrak{F}_{T_m}\right] &= \mathcal{L}^{(m)}(t)\,\mathbf{1}_{\{T_m\leq t\}} \\
&= \frac{\sum_{\substack{z\in\mathcal{S}\\ z\neq Z_m}}\int_{(t,\infty]} f^{(m)}_z(s-T_m)\,\hat{Q}_{Z_m z}(T_m,ds)}{1-\hat{Q}_{Z_m}(T_m,t)}\,\mathbf{1}_{\{T_m\leq t\}} \quad P-a.s.
\end{aligned} \tag{A. 15. 9}$$

Hence, we get on $\{T_m \le t < T_{m+1}\} \cap \{T_m < \infty\}$

$$\begin{aligned}\mathcal{L}(t) - \mathcal{L}(T_m) &= \mathcal{L}^{(m)}(t) - \mathcal{L}^{(m)}(T_m) \\ &= \frac{\sum_{\substack{z\in\mathcal{S}\\ z\neq Z_m}} \int_{(t,\infty]} f_z^{(m)}(s-T_m)\,\hat{Q}_{Z_m z}(T_m, ds)}{1-\hat{Q}_{Z_m}(T_m,t)} \\ &\quad - \frac{\sum_{\substack{z\in\mathcal{S}\\ z\neq Z_m}} \int_{(T_m,\infty]} f_z^{(m)}(s-T_m)\,\hat{Q}_{Z_m z}(T_m, ds)}{1-\hat{Q}_{Z_m}(T_m,T_m)} \quad P-a.s.\end{aligned} \tag{A.15.10}$$

By applying corollary A.3 to the functions

$$F(t) = 1 - \hat{Q}_{Z_m}(T_m, t) \quad \text{and } G(t) = \sum_{\substack{z\in\mathcal{S}\\ z\neq Z_m}} \int_{(t,\infty]} f_z^{(m)}(s-T_m)\,\hat{Q}_{Z_m z}(T_m, ds),$$

(A. 15. 10) yields on $\{T_m \le t < T_{m+1}\} \cap \{T_m < \infty\}$

$$\begin{aligned}&\mathcal{L}^{(m)}(t) - \mathcal{L}^{(m)}(T_m) \\ &= -\sum_{\substack{z\in\mathcal{S}\\ z\neq Z_m}} \int_{(T_m,t]} \frac{f_z^{(m)}(s-T_m)\,\hat{Q}_{Z_m z}(T_m, ds)}{1-\hat{Q}_{Z_m}(T_m, s-0)} \\ &\quad + \int_{(T_m,t]} \frac{\sum_{\substack{z\in\mathcal{S}\\ z\neq Z_m}} \int_{(s,\infty]} f_z^{(m)}(\tau-T_m)\,\hat{Q}_{Z_m z}(T_m, d\tau)}{(1-\hat{Q}(T_m,s))\,(1-\hat{Q}_{Z_m}(T_m,s-0))}\,\hat{Q}_{Z_m}(T_m, ds) \quad P-a.s.\end{aligned} \tag{A.15.11}$$

By inserting (A. 15. 9), (2. 8. 3), and (2. 8. 4), we obtain from (A. 15. 10) and (A. 15. 11) P-a.s. on $\{T_m \le t < T_{m+1}\} \cap \{T_m \le t\}$

$$\begin{aligned}\mathcal{L}(t) - \mathcal{L}(T_m) &= -\sum_{\substack{z\in\mathcal{S}\\ z\neq Z_m}} \int_{(T_m,t]} f_z^{(m)}(s-T_m)\,\hat{q}_{Z_m z}(T_m, ds) \\ &\quad - \sum_{\substack{z\in\mathcal{S}\\ z\neq Z_m}} \int_{(T_m,t]} \mathcal{L}^{(m)}(s)\,\hat{q}_{Z_m z}(T_m, ds) \\ &= -\sum_{\substack{z\in\mathcal{S}\\ z\neq Z_m}} \int_{(T_m,t]} \left(f_z^{(m)}(s-T_m) - \mathcal{L}^{(m)}(s)\right)\hat{q}_{Z_m z}(T_m, ds).\end{aligned} \tag{A.15.12}$$

According to (A. 15. 3) and (A. 15. 4), we also obtain on $\{T_m \le t < T_{m+1}\} \cap \{T_m \le t\}$

$$\begin{aligned}f_z^{(m)}(s-T_m) - \mathcal{L}^{(m)}(s) &= \frac{D_{Z_m z}(s, s-T_m)}{K(DT(s))} + v(s)\,V^+_{(z,0)}(s) - v(s)\,V^+_{(Z_m, s-T_m)}(s), \\ &= v(s)\,R_{Z_m z}(s, s-T_m),\end{aligned}$$

where $R_{Z_m z}$ is the amount of risk according to definition 4. 10. Hence, we finally obtain for the difference $\mathcal{L}(t) - \mathcal{L}(T_m)$ according to (A. 15. 12) on $\{T_m \le t < T_{m+1}\} \cap \{T_m \le t\}$,

$$\mathcal{L}(t) - \mathcal{L}(T_m) = -\sum_{\substack{z\in\mathcal{S}\\ z\neq Z_m}} \int_{(T_m,t]} v(\tau)\,R_{Z_m z}(\tau, \tau-T_m)\,\hat{q}_{Z_m z}(T_m, d\tau) \quad P-a.s. \tag{A.15.13}$$

Now consider $\mathcal{L}(T_{m+1}-0) - \mathcal{L}(T_m) = \lim_{t\nearrow T_{m+1}} \mathcal{L}^{(m)}(t) - \mathcal{L}(T_m)$ which is according to the above P-a.s. given by

$$\mathcal{L}(T_{m+1}-0) - \mathcal{L}(T_m) = -\sum_{\substack{z\in\mathcal{S}\\ z\neq Z_m}} \int_{(T_m,T_{m+1})} v(\tau)\,R_{Z_m z}(\tau, \tau-T_m)\,\hat{q}_{Z_m z}(T_m, d\tau). \tag{A.15.14}$$

Further, consider the difference

$$\mathcal{L}(T_{m+1}) - \mathcal{L}(T_{m+1}-0) = \mathcal{L}(T_{m+1}) - \lim_{t \nearrow T_{m+1}} \mathcal{L}^{(m)}(t). \tag{A. 15. 15}$$

Due to the underlying smoothness conditions in the framework Møller [1993], the corresponding difference causes no further problems. Here, however, Thiele's integral equations of type 1 must be employed to terminate the proof. In doing so, we will later verify that

$$\Delta\left(v(\cdot)\, V^{+}_{(Z_m,\,\cdot - T_m)}(\cdot)\right)(T_{m+1}) \;=\; \sum_{\substack{z\in\mathcal{S}\\ z\neq Z_m}} v(T_{m+1})\, R_{Z_m z}(T_{m+1}, T_{m+1}-T_m)\, \hat{q}_{Z_m z}(T_m, \{T_{m+1}\}), \tag{A. 15. 16}$$

where the left-hand side means

$$v(T_{m+1})\, V^{+}_{(Z_m,\, T_{m+1}-T_m)}(T_{m+1}) - \lim_{t\nearrow T_{m+1}} v(t)\, V^{+}_{(Z_m,\, t-T_m)}(t). \tag{A. 15. 17}$$

Then, by using (A. 15. 1) and (A. 15. 3), we get for the difference (A. 15. 15)

$$\begin{aligned}
&\mathcal{L}(T_{m+1}) - \mathcal{L}(T_{m+1}-0)\\
&= \frac{D_{Z_m Z_{m+1}}(T_{m+1}, T_{m+1}-T_m)}{K(DT(T_{m+1}))} + v(T_{m+1})\, V^{+}_{(Z_{m+1},\,0)}(T_{m+1})\\
&\quad - \lim_{t\nearrow T_{m+1}} v(t)\, V^{+}_{(Z_m,\, t-T_m)}(t)\\
&= \frac{D_{Z_m Z_{m+1}}(T_{m+1}, T_{m+1}-T_m)}{K(DT(T_{m+1}))} + v(T_{m+1})\, V^{+}_{(Z_{m+1},\,0)}(T_{m+1})\\
&\quad - v(T_{m+1})\, V^{+}_{(Z_m,\, T_{m+1}-T_m)}(T_{m+1}) + \Delta\left(v(\cdot)\, V^{+}_{(Z_m,\,\cdot - T_m)}(\cdot)\right)(T_{m+1})\\
&= v(T_{m+1})\, R_{Z_m Z_{m+1}}(T_{m+1}, T_{m+1}-T_m) + \Delta\left(v(\cdot)\, V^{+}_{(Z_m,\,\cdot - T_m)}(\cdot)\right)(T_{m+1}).
\end{aligned} \tag{A. 15. 18}$$

By adding (A. 15. 18) and (A. 15. 14) , and afterwards substituting (A. 15. 16), we get the assertion (4. 18. 2) as follows:

$$\begin{aligned}
&\mathcal{L}(T_{m+1}) - \mathcal{L}(T_m)\\
&= \mathcal{L}(T_{m+1}) - \mathcal{L}(T_{m+1}-0) + \mathcal{L}(T_{m+1}-0) - \mathcal{L}(T_m)\\
&= v(T_{m+1})\, R_{Z_m Z_{m+1}}(T_{m+1}, T_{m+1}-T_m) + \Delta\left(v(\cdot)\, V^{+}_{(Z_m,\,\cdot - T_m)}(\cdot)\right)(T_{m+1})\\
&\quad - \sum_{\substack{z\in\mathcal{S}\\ z\neq Z_m}} \int_{(T_m, T_{m+1})} v(\tau)\, R_{Z_m z}(\tau, \tau - T_m)\, \hat{q}_{Z_m z}(T_m, d\tau)\\
&= \int_{(T_m, T_{m+1}]} v(\tau)\, R_{Z_m Z_{m+1}}(\tau, \tau - T_m)\, N_{Z_m Z_{m+1}, d\tau}\\
&\quad - \sum_{\substack{z\in\mathcal{S}\\ z\neq Z_m}} \int_{(T_m, T_{m+1}]} v(\tau)\, R_{Z_m z}(\tau, \tau - T_m)\, \hat{q}_{Z_m z}(T_m, d\tau)\\
&= \sum_{(y,z)\in\mathcal{J}} \int_{(T_m, T_{m+1}]} v(\tau)\, R_{yz}(\tau, \tau - T_m)\, N_{yz, d\tau}\\
&\quad - \sum_{(y,z)\in\mathcal{J}} \int_{(T_m, T_{m+1}]} v(\tau)\, R_{yz}(\tau, \tau - T_m)\, \mathbf{1}_{\{Z_m = y\}}\, \hat{q}_{yz}(T_m, d\tau)\\
&= \sum_{(y,z)\in\mathcal{J}} \int_{(T_m, T_{m+1}]} v(\tau)\, R_{yz}(\tau, t - T_m)\, M_{yz, d\tau} \quad P-a.s.,
\end{aligned} \tag{A. 15. 19}$$

where the last equation holds due to (2. 13. 1) and (2. 17. 2). Finally, with $\mathcal{L}(0) = \mathcal{L}(T_0) = 0$, (A. 15. 19) and (A. 15. 13), the assertion can be obtained by

$$\begin{aligned}
\mathcal{L}(t) &= \mathcal{L}(t) - \mathcal{L}(0) \\
&= \sum_{m\in\mathbb{N}_0} \Big(\mathcal{L}(T_m) - \mathcal{L}(0) + \mathcal{L}(t) - \mathcal{L}(T_m)\Big) \mathbf{1}_{\{T_m \le t < T_{m+1}\}} \\
&= \sum_{m\in\mathbb{N}_0} \left(\sum_{n=0}^{m-1} \big(\mathcal{L}(T_{n+1}) - \mathcal{L}(T_n)\big) + \mathcal{L}(t) - \mathcal{L}(T_m) \right) \mathbf{1}_{\{T_m \le t < T_{m+1}\}} \\
&= \sum_{(y,z)\in\mathcal{J}} \sum_{m\in\mathbb{N}_0} \int_{(T_m\wedge t, T_{m+1}\wedge t]} v(\tau)\, R_{yz}(\tau, t - T_m)\, M_{yz,d\tau} \quad P-a.s.
\end{aligned}$$

It remains to verify (A. 15. 16). In order to achieve this, consider $\bar{p}_{Z_m}(T_m, t, 0)\, v(t)\, V^+_{(Z_m, t-T_m)}(t)$ for $t > T_m$ and apply Thiele's integral equations of type 1 (4. 14. 23) which hold without exceptional sets. This yields

$$\begin{aligned}
&\bar{p}_{Z_m}(T_m, t, 0)\, v(t)\, V^+_{(Z_m, t-T_m)}(t) \\
&= \bar{p}_{Z_m}(T_m, t, 0) \int_{[t,\infty)} v(\tau)\, \bar{p}_{Z_m}(t, \tau, t - T_m)\, \hat{F}_{Z_m}(T_m, d\tau) \qquad \text{(A. 15. 20)} \\
&\quad + \bar{p}_{Z_m}(T_m, t, 0) \sum_{\substack{z\in\mathcal{S} \\ z\neq Z_m}} \int_{(t,\infty)} \Big(\frac{D_{Z_m z}(\tau, \tau - T_m)}{K(DT(\tau))} + v(\tau)\, V^+_{(z,0)}(\tau)\Big) \\
&\qquad \cdot \bar{p}_{Z_m}(t, \tau - 0, t - T_m)\, \hat{q}_{Z_m z}(T_m, d\tau),
\end{aligned}$$

and, by further employing (2. 27. 8),

$$\begin{aligned}
&\bar{p}_{Z_m}(T_m, t, 0)\, v(t)\, V^+_{(Z_m, t-T_m)}(t) \qquad \text{(A. 15. 21)} \\
&= \int_{[t,\infty)} v(\tau)\, \bar{p}_{Z_m}(T_m, \tau, 0)\, \hat{F}_{Z_m}(T_m, d\tau) \\
&\quad + \sum_{\substack{z\in\mathcal{S} \\ z\neq Z_m}} \int_{(t,\infty)} \Big(\frac{D_{Z_m z}(\tau, \tau - T_m)}{K(DT(\tau))} + v(\tau)\, V^+_{(z,0)}(\tau)\Big)\, \bar{p}_{Z_m}(T_m, \tau - 0, 0)\, \hat{q}_{Z_m z}(T_m, d\tau).
\end{aligned}$$

Thus, we have according to (A. 15. 21) on the one hand

$$\begin{aligned}
&\Delta\Big(\bar{p}_{Z_m}(T_m, \cdot, 0)\, v(\cdot)\, V^+_{(Z_m, \cdot - T_m)}(\cdot)\Big)(t) \\
&= \sum_{\substack{z\in\mathcal{S} \\ z\neq Z_m}} \Big(\frac{D_{Z_m z}(t, t - T_m)}{K(DT(t))} + v(t)\, V^+_{(z,0)}(t)\Big)\, \bar{p}_{Z_m}(T_m, t - 0, 0)\, \hat{q}_{Z_m z}(T_m, \{t\}).
\end{aligned}$$

On the other hand, it holds

$$\begin{aligned}
&\Delta\Big(\bar{p}_{Z_m}(T_m, \cdot, 0)\, v(\cdot)\, V^+_{(Z_m, \cdot - T_m)}(\cdot)\Big)(t) \\
&= \bar{p}_{Z_m}(T_m, \{t\}, 0)\, v(t)\, V^+_{(Z_m, t-T_m)}(t) + \bar{p}_{Z_m}(T_m, t - 0, 0)\, \Delta\Big(v(\cdot)\; V^+_{(Z_m, \cdot - T_m)}(\cdot)\Big)(t). \qquad \text{(A. 15. 22)}
\end{aligned}$$

Hence, we get in view of (A. 15. 16)

$$\begin{aligned}\Delta\Big(v(\cdot)\,V^+_{(Z_m,\,\cdot-T_m)}(\cdot)\Big)(t) &= \sum_{\substack{z\in\mathcal{S}\\ z\neq Z_m}}\Big(\frac{D_{Z_m z}(t,t-T_m)}{K(DT(t))}+v(t)\,V^+_{(z,0)}(t)\Big)\,\hat{q}_{Z_m z}(T_m,\{t\})\\ &\quad -v(t)\,V^+_{(Z_m,\,t-T_m)}(t)\,\frac{\bar{p}_{Z_m}(T_m,\{t\},0)}{\bar{p}_{Z_m}(T_m,t-0,0)}\\ &= \sum_{\substack{z\in\mathcal{S}\\ z\neq Z_m}}\Big(\frac{D_{Z_m z}(t,t-T_m)}{K(DT(t))}+v(T_{m+1})\,V^+_{(z,0)}(t)\Big)\,\hat{q}_{Z_m z}(T_m,\{t\})\\ &\quad +v(t)\,V^+_{(Z_m,\,t-T_m)}(t)\,\hat{q}_{Z_m Z_m}(T_m,\{t\}),\end{aligned}$$

where the last equation is due to (2. 24. 3) and (2. 24. 5). By applying (2. 27. 5) and inserting $t=T_{m+1}$, one obtains

$$\begin{aligned}&\Delta\Big(v(\cdot)\,V^+_{(Z_m,\,\cdot-T_m)}(\cdot)\Big)(T_{m+1})\\ &= \sum_{\substack{z\in\mathcal{S}\\ z\neq Z_m}}\Big(\frac{D_{Z_m z}(T_{m+1},T_{m+1}-T_m)}{K(DT(T_{m+1}))}+v(T_{m+1})\,V^+_{(z,0)}(T_{m+1})\\ &\quad -v(T_{m+1})\,V^+_{(Z_m,\,T_{m+1}-T_m)}(T_{m+1})\Big)\,\hat{q}_{Z_m z}(T_m,\{T_{m+1}\}).\end{aligned}$$

Thus, (A. 15. 16) is verified and the proof of (4. 18. 2) is complete. □

A. 16 Proof of corollary 4. 22:

Consider the situation of theorem 4. 18. On the one hand, the prospective loss $\mathcal{L}(t)$, $t\geq 0$, can according to (A. 15. 1) be represented as

$$\begin{aligned}\mathcal{L}(t) &= \sum_{m\in\mathbb{N}_0}\int_{[T_m\wedge t,T_{m+1}\wedge t)} v(\tau)\,\hat{F}_{Z_m}(T_m,d\tau)+\sum_{m\in\mathbb{N}_0}\frac{D_{Z_m Z_{m+1}}(T_{m+1},T_{m+1}-T_m)}{K(DT(T_{m+1}))}\,\mathbf{1}_{\{T_{m+1}\leq t\}}\\ &\quad +v(t)\sum_{m\in\mathbb{N}_0}V^+_{(Z_m,\,t-T_m)}(t)\,\mathbf{1}_{\{T_m\leq t<T_{m+1}\}}-\mathbf{E}[V_0].\end{aligned}\qquad \text{(A. 16. 1)}$$

On the other hand, it admits the integral representation (4. 18. 2),

$$\mathcal{L}(t) = \sum_{(y,z)\in\mathcal{J}}\sum_{m\in\mathbb{N}_0}\int_{(T_m\wedge t,T_{m+1}\wedge t]} v(\tau)\,R_{yz}(\tau,\tau-T_m)\,M_{yz,d\tau},$$

which can by employing (2. 17. 2), (2. 13. 1) and definition 4. 10 be written as

$$\begin{aligned}\mathcal{L}(t) &= \sum_{(y,z)\in\mathcal{J}}\sum_{m\in\mathbb{N}_0}\int_{(T_m\wedge t,T_{m+1}\wedge t]} v(\tau)\,R_{yz}(\tau,\tau-T_m)\,N_{yz,d\tau}\\ &\quad -\sum_{(y,z)\in\mathcal{J}}\sum_{m\in\mathbb{N}_0}\int_{(T_m\wedge t,T_{m+1}\wedge t]} v(\tau)\,R_{yz}(\tau,\tau-T_m)\,\mathbf{1}_{\{Z_m=y\}}\,\hat{q}_{yz}(T_m,d\tau)\\ &= \sum_{(y,z)\in\mathcal{J}}\sum_{m\in\mathbb{N}_0}\int_{(T_m\wedge t,T_{m+1}\wedge t]}\Big(v(\tau)\,V^+_{(z,\,0)}(\tau)-v(\tau)\,V^+_{(y,\,\tau-T_m)}(\tau)\Big)\,N_{yz,d\tau}\\ &\quad +\sum_{m\in\mathbb{N}_0}\frac{D_{Z_m Z_{m+1}}(T_{m+1},T_{m+1}-T_m)}{K(DT(T_{m+1}))}\,\mathbf{1}_{\{T_{m+1}\leq t\}}\\ &\quad -\sum_{(y,z)\in\mathcal{J}}\sum_{m\in\mathbb{N}_0}\int_{(T_m\wedge t,T_{m+1}\wedge t]} v(\tau)\,R_{yz}(\tau,\tau-T_m)\,\mathbf{1}_{\{Z_m=y\}}\,\hat{q}_{yz}(T_m,d\tau).\end{aligned}\qquad \text{(A. 16. 2)}$$

Bringing (A. 16. 1) and (A. 16. 2) together, we get

$$\begin{aligned}
& v(t) \sum_{m\in\mathbb{N}_0} V^+_{(Z_m,\, t-T_m)}(t)\, \mathbf{1}_{\{T_m\le t<T_{m+1}\}} - \mathbf{E}[V_0] \\
&= - \sum_{m\in\mathbb{N}_0} \int_{[T_m\wedge t, T_{m+1}\wedge t)} v(\tau)\, \hat{F}_{Z_m}(T_m, d\tau) \\
&\quad + \sum_{(y,z)\in\mathcal{J}} \sum_{m\in\mathbb{N}_0} \int_{(T_m\wedge t, T_{m+1}\wedge t]} \Big(v(\tau)\, V^+_{(z,\,0)}(\tau) \;-\; v(\tau)\, V^+_{(y,\,\tau-T_m)}(\tau) \Big)\, N_{yz,d\tau} \\
&\quad - \sum_{(y,z)\in\mathcal{J}} \sum_{m\in\mathbb{N}_0} \int_{(T_m\wedge t, T_{m+1}\wedge t]} v(\tau)\, R_{yz}(\tau, \tau-T_m)\, \mathbf{1}_{\{Z_m=y\}}\, \hat{q}_{yz}(T_m, d\tau).
\end{aligned}$$

This yields, for $s,t\ge 0$ satisfying $T_m \le s \le s+t < T_{m+1}$, the following:

$$\begin{aligned}
& v(s+t)\, V^+_{(Z_m,\, s-T_m+t)}(s+t) - v(s)\, V^+_{(Z_m,\, s-T_m)}(s) \\
&= - \int_{[s,s+t)} v(\tau)\, \hat{F}_{Z_m}(T_m, d\tau) \\
&\quad + \sum_{(y,z)\in\mathcal{J}} \int_{(s,s+t]} \Big(v(\tau)\, V^+_{(z,\,0)}(\tau) \;-\; v(\tau)\, V^+_{(y,\,\tau-T_m)}(\tau) \Big)\, N_{yz,d\tau} \\
&\quad - \sum_{(y,z)\in\mathcal{J}} \int_{(s,s+t]} v(\tau)\, R_{yz}(\tau, \tau-T_m)\, \mathbf{1}_{\{Z_m=y\}}\, \hat{q}_{yz}(T_m, d\tau) \\
&= - \int_{[s,t)} v(\tau)\, \hat{F}_{Z_m}(T_m, d\tau) \\
&\quad - \sum_{(y,z)\in\mathcal{J}} \int_{(s,s+t]} v(\tau)\, R_{yz}(\tau, \tau-T_m)\, \mathbf{1}_{\{Z_m=y\}}\, \hat{q}_{yz}(T_m, d\tau).
\end{aligned}$$

Given $(T_m, Z_m) = (s-u, y)$, we obtain

$$\begin{aligned}
v(s+t)\, V^+_{(y,u+t)}(s+t) - v(s)\, V^+_{(y,u)}(s) \quad &= \quad - \int_{[s,s+t)} v(\tau)\, \hat{F}_y(s-u, d\tau) \\
&\quad - \sum_{\substack{z\in\mathcal{S}\\ z\ne y}} \int_{(s,s+t]} v(\tau)\, R_{yz}(\tau, \tau-s+u)\; \hat{q}_{yz}(s-u, d\tau).
\end{aligned}$$

Now let $t\to\infty$. Then, under the assumptions stipulated, it follows according to (4. 21. 4)

$$v(s)\, V^+_{(y,u)}(s) \quad = \quad \int_{[s,\infty)} v(\tau)\, \hat{F}_y(s-u, d\tau) + \sum_{\substack{z\in\mathcal{S}\\ z\ne y}} \int_{(s,\infty)} v(\tau)\, R_{yz}(\tau, \tau-s+u)\; \hat{q}_{yz}(s-u, d\tau). \tag{A. 16. 3}$$

This is assertion (4. 22. 1).

For (4. 22. 2), assumption (4. 21. 3) must additionally be satisfied. (4. 22. 2) can then be obtained, by inserting (3. 2. 3) into (A. 16. 3) and afterwards applying Fubini's theorem, as follows:

$$\begin{aligned}
&V^+_{(y,u)}(s)\\
&= \hat{F}_y(s-u,[s,\infty)) - \int_{[s,\infty)}\int_{(s,\tau]} v(\tau)\,K(r-0)\,\Phi(dr)\,\hat{F}_y(s-u,d\tau)\\
&\quad+\sum_{\substack{z\in\mathcal{S}\\ z\neq y}}\int_{(s,\infty)} R_{yz}(\tau,\tau-s+u)\ \hat{q}_{yz}(s-u,d\tau)\\
&\quad-\sum_{\substack{z\in\mathcal{S}\\ z\neq y}}\int_{(s,\infty)}\int_{(s,\tau]} v(\tau)\,K(r-0)\,\Phi(dr)\,R_{yz}(\tau,\tau-s+u)\ \hat{q}_{yz}(s-u,d\tau)\\
&= \hat{F}_y(s-u,[s,\infty)) + \sum_{\substack{z\in\mathcal{S}\\ z\neq y}}\int_{(s,\infty)} R_{yz}(\tau,\tau-s+u)\ \hat{q}_{yz}(s-u,d\tau)\\
&\quad-\int_{(s,\infty)}\int_{[r,\infty)} v(\tau)\,K(r-0)\,\hat{F}_y(s-u,d\tau)\,\Phi(dr)\\
&\quad-\sum_{\substack{z\in\mathcal{S}\\ z\neq y}}\int_{(s,\infty)}\int_{[r,\infty)} v(\tau)\,K(r-0)\,R_{yz}(\tau,\tau-s+u)\ \hat{q}_{yz}(s-u,d\tau)\,\Phi(dr)\\
&= \hat{F}_y(s-u,[s,\infty)) + \sum_{\substack{z\in\mathcal{S}\\ z\neq y}}\int_{(s,\infty)} R_{yz}(\tau,\tau-s+u)\ \hat{q}_{yz}(s-u,d\tau)\\
&\quad-\int_{(s,\infty)}\left(\int_{[r,\infty)} v(\tau)\,K(r-0)\,\hat{F}_y(r-r+s-u,d\tau)\right.\\
&\quad\left.-\sum_{\substack{z\in\mathcal{S}\\ z\neq y}}\int_{[r,\infty)} v(\tau)\,K(r-0)\,R_{yz}(\tau,\tau-r-r+s+u)\ \hat{q}_{yz}(r-r+s-u,d\tau)\right)\Phi(dr). \qquad \text{(A. 16. 4)}
\end{aligned}$$

Inserting (A. 16. 3), (A. 16. 4) turns out to be

$$\begin{aligned}
V^+_{(y,u)}(s) &= \hat{F}_y(s-u,[s,\infty)) + \sum_{\substack{z\in\mathcal{S}\\ z\neq y}}\int_{(s,\infty)} R_{yz}(\tau,\tau-s+u)\ \hat{q}_{yz}(s-u,d\tau)\\
&\quad-\int_{(s,\infty)} V^+_{(y,r-s+u)}(r-0)\,\Phi(dr),
\end{aligned}$$

which is (4. 22. 2). □

A. 17 Proof of lemma 4. 23:

Here, the remaining direction of the proof of lemma 4. 23 is verified, namely that the backward integral equations of type 2 imply Thiele's integral equations of type 2. Let $0\le u\le s\le t, v\ge 0$, and $(y,z)\in\mathcal{S}^2$. The prospective reserve $V^+_{(y,u)}(s)$ can under assumptions 2. 35 and 4. 5 be represented as

$$\begin{aligned}
V^+_{(y,u)}(s) &= K(s)\sum_{z\in\mathcal{S}}\int_{[s,\infty)}\int_{[0,\infty)} v(\tau)\,f_z(\tau,l)\,p_{yz}(s,\tau,u,dl)\,\mathbf{F}_z(d\tau) \qquad \text{(A. 17. 1)}\\
&\quad+K(s)\sum_{(z,\eta)\in\mathcal{J}}\int_{(s,\infty)}\int_{(0,\infty)} \frac{D_{z\eta}(\tau,l)}{K(DT(\tau))}\,\lambda_{z\eta}(\tau,l)\,p_{yz}(s,\tau-0,u,dl)\,\mathbf{\Lambda}_{z\eta}(d\tau).
\end{aligned}$$

Further, the backward integral equations of type 2 are given as

$$p_{yz}(s,t,u,v) = \delta_{yz}\,\mathbf{1}_{(v\geq t-s+u)} + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,q_{y\xi}(s,d\tau,u) + \int_{(s,t]} p_{yz}(\tau,t,\tau-s+u,v)\,q_{yy}(s,d\tau,u).$$

They are assumed to hold without exceptional sets. Rewriting the above, we get

$$p_{yz}(s,t,u,v) = \delta_{yz}\,\mathbf{1}_{[t-s+u,\infty)}(v) + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \int_{(s,t]} p_{\xi z}(\tau,t,0,v)\,q_{y\xi}(s,d\tau,u) + \int_{(s,t]} p_{yz}(\tau,t,\tau-s+u,v)\,q_{yy}(s,d\tau,u) \qquad \text{(A. 17. 2)}$$

and

$$p_{yz}(s,t-0,u,v) = \delta_{yz}\,\mathbf{1}_{[t-s+u,\infty)}(v) + \sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \int_{(s,t)} p_{\xi z}(\tau,t-0,0,v)\,q_{y\xi}(s,d\tau,u) + \int_{(s,t)} p_{yz}(\tau,t-0,\tau-s+u,v)\,q_{yy}(s,d\tau,u), \qquad \text{(A. 17. 3)}$$

respectively. Thus, one obtains the following measures

$$p_{yz}(s,\tau,u,dl) = \delta_{yz}\,\varepsilon_{\tau-s+u}(dl) + \Big(\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \int_{(s,\tau]} p_{\xi z}(r,\tau,0,\cdot)\,q_{y\xi}(s,dr,u)\Big)(dl) + \Big(\int_{(s,\tau]} p_{yz}(r,\tau,r-s+u,\cdot)\,q_{yy}(s,dr,u)\Big)(dl) \qquad \text{(A. 17. 4)}$$

as well as

$$p_{yz}(s,\tau-0,u,dl) = \delta_{yz}\,\varepsilon_{\tau-s+u}(dl) + \Big(\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}} \int_{(s,\tau)} p_{\xi z}(r,\tau-0,0,\cdot)\,q_{y\xi}(s,dr,u)\Big)(dl) + \Big(\int_{(s,\tau)} p_{yz}(r,\tau-0,r-s+u,\cdot)\,q_{yy}(s,dr,u)\Big)(dl). \qquad \text{(A. 17. 5)}$$

Inserting these equations into (A. 17. 1), and afterwards applying Fubini's theorem twice, yield

$$\begin{aligned}
&v(s)\,V^+_{(y,u)}(s)\\
&= \int_{[s,\infty)} v(\tau)\,f_y(\tau,\tau-s+u)\,\mathbf{F}_y(d\tau)\\
&\quad+\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\sum_{z\in\mathcal{S}}\int_{[s,\infty)}\int_{(s,\tau]}\int_{[0,\infty)} v(\tau)\,f_z(\tau,l)\,p_{\xi z}(r,\tau,0,dl)\,q_{y\xi}(s,dr,u)\,\mathbf{F}_z(d\tau)\\
&\quad+\sum_{z\in\mathcal{S}}\int_{[s,\infty)}\int_{(s,\tau]}\int_{[0,\infty)} v(\tau)\,f_z(\tau,l)\,p_{yz}(r,\tau,r-s+u,dl)\,q_{yy}(s,dr,u)\,\mathbf{F}_z(d\tau)\\
&\quad+\sum_{\substack{\eta\in\mathcal{S}\\ \eta\neq y}}\int_{(s,\infty)}\frac{D_{y\eta}(\tau,\tau-s+u)}{K(DT(\tau))}\,\lambda_{y\eta}(\tau,\tau-s+u)\,\mathbf{\Lambda}_{y\eta}(d\tau)\\
&\quad+\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\sum_{(z,\eta)\in\mathcal{J}}\int_{(s,\infty)}\int_{(s,\tau)}\int_{(0,\infty)}\frac{D_{z\eta}(\tau,l)}{K(DT(\tau))}\,\lambda_{z\eta}(\tau,l)\,p_{\xi z}(r,\tau-0,0,dl)\,q_{y\xi}(s,dr,u)\,\mathbf{\Lambda}_{z\eta}(d\tau)\\
&\quad+\sum_{(z,\eta)\in\mathcal{J}}\int_{(s,\infty)}\int_{(s,\tau)}\int_{(0,\infty)}\frac{D_{z\eta}(\tau,l)}{K(DT(\tau))}\,\lambda_{z\eta}(\tau,l)\,p_{yz}(r,\tau-0,r-s+u,dl)\,q_{yy}(s,dr,u)\,\mathbf{\Lambda}_{z\eta}(d\tau)\\
&= \int_{[s,\infty)} v(\tau)\,f_y(\tau,\tau-s+u)\,\mathbf{F}_y(d\tau)+\sum_{\substack{\eta\in\mathcal{S}\\ \eta\neq y}}\int_{(s,\infty)}\frac{D_{y\eta}(\tau,\tau-s+u)}{K(DT(\tau))}\,\lambda_{y\eta}(\tau,\tau-s+u)\,\mathbf{\Lambda}_{y\eta}(d\tau)\\
&\quad+\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\sum_{z\in\mathcal{S}}\int_{(s,\infty)}\int_{[r,\infty)}\int_{[0,\infty)} v(\tau)\,f_z(\tau,l)\,p_{\xi z}(r,\tau,0,dl)\,\mathbf{F}_z(d\tau)\,q_{y\xi}(s,dr,u)\\
&\quad+\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\sum_{(z,\eta)\in\mathcal{J}}\int_{(s,\infty)}\int_{(r,\infty)}\int_{(0,\infty)}\frac{D_{z\eta}(\tau,l)}{K(DT(\tau))}\,\lambda_{z\eta}(\tau,l)\,p_{\xi z}(r,\tau-0,0,dl)\,\mathbf{\Lambda}_{z\eta}(d\tau)\,q_{y\xi}(s,dr,u)\\
&\quad+\sum_{z\in\mathcal{S}}\int_{(s,\infty)}\int_{[r,\infty)}\int_{[0,\infty)} v(\tau)\,f_z(\tau,l)\,p_{yz}(r,\tau,r-s+u,dl)\,q_{yy}(s,dr,u)\,\mathbf{F}_z(d\tau)\\
&\quad+\sum_{(z,\eta)\in\mathcal{J}}\int_{(s,\infty)}\int_{(r,\infty)}\int_{(0,\infty)}\frac{D_{z\eta}(\tau,l)}{K(DT(\tau))}\,\lambda_{z\eta}(\tau,l)\,p_{yz}(r,\tau-0,r-s+u,dl)\,\mathbf{\Lambda}_{z\eta}(d\tau)\,q_{yy}(s,dr,u).
\end{aligned}$$

A twofold application of (A. 17. 1) result in

$$\begin{aligned}
v(s)\,V^+_{(y,u)}(s) &= \int_{[s,\infty)} v(\tau)\,\mathbf{F}_y(s,d\tau,u)+\sum_{\substack{\eta\in\mathcal{S}\\ \eta\neq y}}\int_{(s,\infty)}\frac{D_{y\eta}(\tau,\tau-s+u)}{K(DT(\tau))}\,q_{y\eta}(s,d\tau,u)\\
&\quad+\sum_{\substack{\xi\in\mathcal{S}\\ \xi\neq y}}\int_{(s,\infty)} v(r)\,V^+_{(\xi,0)}(r)\,q_{y\xi}(s,dr,u)+\int_{(s,\infty)} v(r)\,V^+_{(y,r-s+u)}(r)\,q_{yy}(s,dr,u),
\end{aligned}$$

which is according to (2. 27. 5), the selection of q by means of (2. 28. 3), and definition 4. 10 equal to

$$v(s)\,V^+_{(y,u)}(s) = \int_{[s,\infty)} v(\tau)\,\hat{F}_y(s-u,d\tau)+\sum_{\substack{z\in\mathcal{S}\\ z\neq y}}\int_{(s,\infty)} v(\tau)\,R_{yz}(\tau,\tau-s+u)\;\hat{q}_{yz}(s-u,d\tau).$$

This corresponds to (4. 22. 1). Thiele's integral equation (4. 22. 2) can then be obtained by following the same argumentation as in the proof of (4. 22. 2) (cf. A.16). □

A. 18 Proof of lemma 4. 24:
In order to accomplish the proof of lemma 4. 24, we start with Thiele's integral equations of type 2, which are for $0 \leq u \leq s$ and $y \in \mathcal{S}$ given by

$$\begin{aligned} V^+_{(y,u)}(s) &= \hat{F}_y(s-u,[s,\infty)) - \int_{(s,\infty)} V^+_{(y,r-s+u)}(r-0)\,\Phi(dr) \\ &\quad + \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} R_{yz}(\tau,\tau-s+u)\ \hat{q}_{yz}(s-u,d\tau), \end{aligned} \tag{A. 18. 1}$$

and apply (3. 8. 1) to $Z(d\tau) = \hat{F}_y(s-u,d\tau)$, and (3. 8. 2) to $Z(d\tau) = V^+_{(y,\tau-s+u)}(\tau-0)\,\Phi(d\tau)$ as well as to $Z(d\tau) = R_{yz}(\tau,\tau-s+u)\,\hat{q}_{yz}(s-u,d\tau)$. Thus, we obtain

$$\begin{aligned} & V^+_{(y,u)}(s) \qquad\qquad \text{(A. 18. 2)} \\ &= K(s) \int_{[s,\infty)} v(\tau)\,\bar{p}_y(s,\tau,u)\,\hat{F}_y(s-u,d\tau) \\ &\quad + K(s) \int_{(s,\infty)} v(\tau)\,\bar{p}_y(s,\tau-0,u)\,\hat{F}_y(s-u,[\tau,\infty))\,\Phi(d\tau) \\ &\quad - K(s) \int_{(s,\infty)} v(\tau)\,\bar{p}_y(s,\tau-0,u)\,\hat{F}_y(s-u,[\tau,\infty))\,q_{yy}(s,d\tau,u) \\ &\quad - K(s) \int_{(s,\infty)} v(\tau)\,\bar{p}_y(s,\tau-0,u)\,V^+_{(y,\tau-s+u)}(\tau-0)\,\Phi(d\tau) \\ &\quad - K(s) \int_{(s,\infty)} v(\tau)\,\bar{p}_y(s,\tau-0,u) \int_{[\tau,\infty)} V^+_{(y,r-s+u)}(r-0)\,\Phi(dr)\,\Phi(d\tau) \\ &\quad + K(s) \int_{(s,\infty)} v(\tau)\,\bar{p}_y(s,\tau-0,u) \int_{(\tau,\infty)} V^+_{(y,r-s+u)}(r-0)\,\Phi(dr)\ q_{yy}(s,d\tau,u) \\ &\quad + K(s) \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} v(\tau)\,\bar{p}_y(s,\tau-0,u)\,R_{yz}(\tau,\tau-s+u)\,\hat{q}_{yz}(s-u,d\tau) \\ &\quad + K(s) \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} v(\tau)\,\bar{p}_y(s,\tau-0,u) \int_{[\tau,\infty)} R_{yz}(r,r-s+u)\,\hat{q}_{yz}(s-u,dr)\,\Phi(d\tau) \\ &\quad - K(s) \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} v(\tau)\,\bar{p}_y(s,\tau-0,u) \int_{(\tau,\infty)} R_{yz}(r,r-s+u)\,\hat{q}_{yz}(s-u,dr)\ q_{yy}(s,d\tau,u). \end{aligned}$$

Another application of Thiele's integral equations of type 2 (A. 18. 1) yields

$$\begin{aligned} & K(s) \int_{(s,\infty)} v(\tau)\,\bar{p}_y(s,\tau-0,u)\,\hat{F}_y(s-u,[\tau,\infty))\,\Phi(d\tau) \\ & - K(s) \int_{(s,\infty)} v(\tau)\,\bar{p}_y(s,\tau-0,u) \int_{[\tau,\infty)} V^+_{(y,r-s+u)}(r-0)\,\Phi(dr)\,\Phi(d\tau) \\ & + K(s) \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} v(\tau)\,\bar{p}_y(s,\tau-0,u) \int_{[\tau,\infty)} R_{yz}(r,r-s+u)\,\hat{q}_{yz}(s-u,dr)\,\Phi(d\tau) \\ & = \end{aligned}$$

$$
\begin{aligned}
&= K(s) \int_{(s,\infty)} v(\tau)\, \bar{p}_y(s,\tau-0,u)\, \hat{F}_y(\tau-\tau+s-u,[\tau,\infty))\, \Phi(d\tau) \\
&\quad -K(s) \int_{(s,\infty)} v(\tau)\, \bar{p}_y(s,\tau-0,u) \int_{[\tau,\infty)} V^+_{(y,r-\tau+\tau-s+u)}(r-0)\, \Phi(dr)\, \Phi(d\tau) \\
&\quad +K(s) \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} v(\tau)\, \bar{p}_y(s,\tau-0,u) \\
&\qquad\qquad \cdot \int_{[\tau,\infty)} R_{yz}(r,r-\tau+\tau-s+u)\, \hat{q}_{yz}(\tau-\tau+s-u,dr)\, \Phi(d\tau) \\
&= K(s) \int_{(s,\infty)} v(\tau)\, \bar{p}_y(s,\tau-0,u)\, V^+_{(y,\tau-s+u)}(\tau-0)\, \Phi(d\tau).
\end{aligned}
$$

In almost the same manner, the following can be confirmed:

$$
\begin{aligned}
&-K(s) \int_{(s,\infty)} v(\tau)\, \bar{p}_y(s,\tau-0,u)\, \hat{F}_y(s-u,[\tau,\infty))\, q_{yy}(s,d\tau,u) \\
&\quad +K(s) \int_{(s,\infty)} v(\tau)\, \bar{p}_y(s,\tau-0,u) \int_{(\tau,\infty)} V^+_{(y,r-s+u)}(r-0)\, \Phi(dr)\ q_{yy}(s,d\tau,u) \\
&\quad -K(s) \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} v(\tau)\, \bar{p}_y(s,\tau-0,u) \int_{(\tau,\infty)} R_{yz}(r,r-s+u)\, \hat{q}_{yz}(s-u,dr)\ q_{yy}(s,d\tau,u) \\
&= -K(s) \int_{(s,\infty)} v(\tau)\, \bar{p}_y(s,\tau-0,u)\, V^+_{(y,\tau-s+u)}(\tau)\ q_{yy}(s,d\tau,u).
\end{aligned}
$$

Inserting both of the above relations, (A. 18. 2) can be reduced to

$$
\begin{aligned}
V^+_{(y,u)}(s) \;=\; & K(s) \int_{[s,\infty)} v(\tau)\, \bar{p}_y(s,\tau,u)\, \hat{F}_y(s-u,d\tau) \\
& +K(s) \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} v(\tau)\, \bar{p}_y(s,\tau-0,u)\, R_{yz}(\tau,\tau-s+u)\, \hat{q}_{yz}(s-u,d\tau) \\
& -K(s) \int_{(s,\infty)} v(\tau)\, \bar{p}_y(s,\tau-0,u)\, V^+_{(y,\tau-s+u)}(\tau)\ q_{yy}(s,d\tau,u). \qquad \text{(A. 18. 3)}
\end{aligned}
$$

According to definition 4. 10, (2. 27. 5), and the selection of q by means of (2. 28. 3), this is equal to Thiele's integral equation of type 1, i.e.

$$
\begin{aligned}
V^+_{(y,u)}(s) \;=\; & K(s) \int_{[s,\infty)} v(\tau)\, \bar{p}_y(s,\tau,u)\, \hat{F}_y(s-u,d\tau) \\
& +K(s) \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,\infty)} \Big(\frac{D_{yz}(\tau,\tau-s+u)}{K(DT(\tau))} + v(\tau)\, V^+_{(z,0)}(\tau)\Big)\, \bar{p}_y(s,\tau-0,u)\, \hat{q}_{yz}(s-u,d\tau).
\end{aligned}
$$

□

A. 19 Proof of lemma 4. 28:
In order to prove the assertion, we likewise follow the proceedings in Stracke [1997] or Milbrodt and Helbig ([1999], Hilfsfssatz 10.27), and select a solution $(h_y)_{y\in\mathcal{S}}$ of the system (4. 28. 1) that also fulfills (4. 28. 2). Then, we define

$$h_y^{(n)} := h_y \cdot \mathbf{1}_{[0,n]}, \;\; y\in\mathcal{S}, n>0.$$

Since $(h_y)_{y\in\mathcal{S}}$ is a solution of (4. 28. 1), we obtain for $s \in [0,n], y \in \mathcal{S}$

$$\begin{aligned} h_y^{(n)}(s) = h_y(s) &= K(s) \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,n]} v(\tau)\, h_z^{(n)}(\tau)\, \bar{p}_y(s,\tau-0,0)\, \hat{q}_{yz}(s,d\tau) \\ &\quad + K(s) \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(n,\infty)} v(\tau)\, h_z(\tau)\, \bar{p}_y(s,\tau-0,0)\, \hat{q}_{yz}(s,d\tau). \end{aligned}$$

By employing (2. 27. 8), one obtains from the above

$$\begin{aligned} h_y^{(n)}(s) &= K(s) \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(s,n]} v(\tau)\, h_z^{(n)}(\tau)\, \bar{p}_y(s,\tau-0,0)\, \hat{q}_{yz}(s,d\tau) \\ &\quad + K(s)\, v(n)\, \bar{p}_y(s,n,0)\, H_y^{(n)}(s), \end{aligned} \tag{A. 19. 1}$$

where

$$H_y^{(n)} : [0,n] \ni s \mapsto K(n) \sum_{\substack{z\in\mathcal{S}\\ z\neq y}} \int_{(n,\infty)} v(\tau)\, h_z(\tau)\, \bar{p}_y(n,\tau-0,n-s)\, \hat{q}_{yz}(s,d\tau). \tag{A. 19. 2}$$

Thus, (A. 19. 1) corresponds to (4. 27. 1) with $D_{yz} \equiv 0$ and $\hat{F}_y^{(n)}(s,d\tau) = H_y^{(n)}(s)\, \varepsilon_n(d\tau)$. Hence, according to lemma 4. 27, $(h_y^{(n)}(\cdot))_{y\in\mathcal{S}}$ is the one and only solution of (A. 19. 1). By inserting the above parameters in formula (4. 6. 1), we learn that this solution must for $s \in [0,n]$ and $y \in \mathcal{S}$ be of the form

$$\begin{aligned} & h_y^{(n)}(s) \\ &= K(s) \sum_{l=0}^{\infty} \sum_{\xi\in\mathcal{S}} \int_{[s,\infty)} \int_{[s,\infty)} v(\tau)\, (1 - Q_\xi(r,\tau,0))\, H_\xi^{(n)}(r)\, \varepsilon_n(d\tau)\, P(T_l \in dr, Z_l = \xi |\, T_0 = s, Z_0 = y) \\ &= K(s) \sum_{l=0}^{\infty} \sum_{\xi\in\mathcal{S}} \int_{[s,\infty)} v(n)\, p_\xi(r,n,0)\, H_\xi^{(n)}(r)\, P(T_l \in dr, Z_l = \xi |\, T_0 = s, Z_0 = y). \end{aligned} \tag{A. 19. 3}$$

Since $(h_y)_{y\in\mathcal{S}}$ was selected as a solution of (4. 28. 1) which fulfills (4. 28. 2), we get according to (A. 19. 2)

$$\lim_{n\to\infty} |v(n)\, p_\xi(s,n,0)\, H_\xi^{(n)}(s)| \leq \lim_{n\to\infty} \sum_{z\in\mathcal{S}} \int_{(n,\infty)} v(t)\, |h_z(t)|\, \hat{q}_{yz}(s,dt) = 0 \ \ \forall s \geq 0.$$

Hence, it follows by applying the theorem on dominated convergence to the last line of (A. 19. 3) that

$$h_y(s) = \lim_{n\to\infty} h_y^{(n)}(s) = 0, \quad s \geq 0, y \in \mathcal{S}.$$

□

A. 20 Proof of corollary 4. 30:
As mentioned subsequent to theorem 4. 29, the variance $\mathbf{V}[(\mathcal{L}_y(t))]$, $t \geq 0$, can be derived as expectation of the predictable variation. Thus we obtain for $t \geq 0$

$$\mathbf{V}[\mathcal{L}(t)] = \mathbf{E}[\langle\mathcal{L}\rangle_t] = \sum_{\eta\in\mathcal{S}} P(Z_0 = \eta)\, \mathbf{E}[\langle\mathcal{L}_y\rangle_t | T_0 = 0, Z_0 = \eta].$$

In order to derive $\mathbf{E}[\langle\mathcal{L}_y\rangle_t | T_0 = 0, Z_0 = \eta]$, we start from (4. 29. 1) and proceed in almost the same manner as in the proof of theorem 4. 6 (cf. A.12), namely by deriving a version of the conditional expectation by integration with respect to the corresponding conditional expectations according to (2. 17. 1). Thus, we obtain for $\mathcal{L}(Z_0|P)$-a.e. $\eta \in \mathcal{S}$

$$\begin{aligned}
&\mathbf{E}[\langle \mathcal{L}_y \rangle_t | T_0 = 0, Z_0 = \eta] \\
&= \mathbf{E}\Big[\sum_{z \neq y} \sum_{l \in \mathbb{N}_0} \int_{(T_l \wedge t, T_{l+1} \wedge t]} \Big(v(\tau)\, R_{yz}(\tau, \tau - T_l) \Big)^2 \, \mathbf{1}_{\{Z_l = y\}}\, \hat{q}_{yz}(T_l, d\tau) \\
&\quad - \sum_{z,\xi \in \mathcal{S}\setminus\{y\}} \sum_{l \in \mathbb{N}_0} \int_{(T_l \wedge t, T_{l+1} \wedge t]} (v(\tau))^2 \, R_{yz}(\tau, \tau - T_l)\, R_{y\xi}(\tau, \tau - T_l) \\
&\qquad \cdot \mathbf{1}_{\{Z_l = y\}}\, \hat{q}_{yz}(T_l, \{\tau\})\; \hat{q}_{y\xi}(T_l, d\tau) \big| T_0 = 0, Z_0 = \eta \Big] \\
&= \sum_{l=0}^{\infty} \sum_{z \neq y} \int_{(0,t] \times \mathcal{S} \times (0,t]} \int_{(r,x]} \Big(v(\tau)\, R_{yz}(\tau, \tau - r) \Big)^2 \, \mathbf{1}_{\{Z_l = y\}}\, \hat{q}_{yz}(r, d\tau) \\
&\qquad \cdot P(T_{l+1} \in dx, T_l \in dr, Z_l \in d\nu | T_0 = 0, Z_0 = \eta) \\
&\quad - \sum_{l=0}^{\infty} \sum_{z,\xi \in \mathcal{S}\setminus\{y\}} \int_{(0,t] \times \mathcal{S} \times (0,t]} \int_{(r,x]} (v(\tau))^2 \, R_{yz}(\tau, \tau - r)\, R_{y\xi}(\tau, \tau - r)\, \mathbf{1}_{\{Z_l = y\}}\, \hat{q}_{yz}(r, \{\tau\})\; \hat{q}_{y\xi}(r, d\tau) \\
&\qquad \cdot P(T_{l+1} \in dx, T_l \in dr, Z_l \in d\nu | T_0 = 0, Z_0 = \eta). \qquad (A.\,20.\,1)
\end{aligned}$$

By successive conditioning and applying the Markov property of the marked point process (T, Z), we get for $\mathcal{L}(Z_0|P)$-a.e. $\eta \in \mathcal{S}$

$$\begin{aligned}
&\mathbf{E}[\langle \mathcal{L}_y \rangle_t | T_0 = 0, Z_0 = \xi] \\
&= \sum_{l=0}^{\infty} \sum_{z \neq y} \int_{(0,t]} \int_{(0,t]} \int_{(r,x]} \Big(v(\tau)\, R_{yz}(\tau, \tau - r) \Big)^2 \, \hat{q}_{yz}(r, d\tau) \\
&\qquad \cdot P(T_{l+1} \in dx |\, T_l = r, Z_l = y)\, P(T_l \in dr, Z_l = y | T_0 = 0, Z_0 = \eta) \\
&\quad - \sum_{l=0}^{\infty} \sum_{z,\xi \in \mathcal{S}\setminus\{y\}} \int_{(0,t]} \int_{(0,t]} \int_{(r,x]} (v(\tau))^2 \, R_{yz}(\tau, \tau - r)\, R_{y\xi}(\tau, \tau - r)\, \hat{q}_{yz}(r, \{\tau\})\; \hat{q}_{y\xi}(r, d\tau) \\
&\qquad \cdot P(T_{l+1} \in dx |\, T_l = r, Z_l = y)\, P(T_l \in dr, Z_l = y | T_0 = 0, Z_0 = \eta).
\end{aligned}$$

Applying Fubini's theorem to both addends, and afterwards inserting (2. 8. 2), the above equation can be continued as

$$\begin{aligned}
&= \sum_{l=0}^{\infty} \sum_{z \neq y} \int_{(0,t]} \int_{(r,t]} (1 - \hat{Q}_y(r, \tau - 0)) \Big(v(\tau)\, R_{yz}(\tau, \tau - r) \Big)^2 \, \hat{q}_{yz}(r, d\tau) \\
&\qquad \cdot P(T_l \in dr, Z_l = y | T_0 = 0, Z_0 = \eta) \\
&\quad - \sum_{l=0}^{\infty} \sum_{z,\xi \in \mathcal{S}\setminus\{y\}} \int_{(0,t]} \int_{(r,t]} (1 - \hat{Q}_y(r, \tau - 0))\, (v(\tau))^2 \, R_{yz}(\tau, \tau - r)\, R_{y\xi}(\tau, \tau - r)\, \hat{q}_{yz}(r, \{\tau\})\; \hat{q}_{y\xi}(r, d\tau) \\
&\qquad \cdot P(T_l \in dr, Z_l = y | T_0 = 0, Z_0 = \eta).
\end{aligned}$$

According to assumption 2. 35, Fubini's theorem can also be applied to the remaining integrals.

Thus, we achieve for $\mathcal{L}(Z_0|P)$-a.e. $\eta \in \mathcal{S}$

$$\begin{aligned}
&\mathbf{E}[\langle \mathcal{L}_y \rangle_t | T_0 = 0, Z_0 = \eta] \\
&= \sum_{l=0}^{\infty} \sum_{z \neq y} \int_{(0,t]} \int_{(0,\tau)} (1 - \hat{Q}_y(r, \tau - 0)) \Big(v(\tau)\, R_{yz}(\tau, \tau - r) \Big)^2 \lambda_{yz}(\tau, \tau - r) \\
&\qquad \cdot P(T_l \in dr, Z_l = y | T_0 = 0, Z_0 = \eta)\, \mathbf{\Lambda}_{yz}(d\tau) \\
&\quad - \sum_{l=0}^{\infty} \sum_{z,\xi \in \mathcal{S} \setminus \{y\}} \int_{(0,t]} \int_{(0,\tau)} (1 - \hat{Q}_y(r, \tau - 0))\, (v(\tau))^2\, R_{yz}(\tau, \tau - r)\, R_{y\xi}(\tau, \tau - r)\, \lambda_{yz}(\tau, \tau - r) \\
&\qquad \cdot \lambda_{y\xi}(\tau, \tau - r)\, P(T_l \in dr, Z_l = y | T_0 = 0, Z_0 = \eta)\, \mathbf{\Lambda}_{yz}(\{\tau\})\, \mathbf{\Lambda}_{y\xi}(d\tau),
\end{aligned}$$

where the right-hand side can be represented, by proceeding in a similar manner as in (A. 12. 7), (A. 12. 8) and (A. 12. 9), as

$$\begin{aligned}
&= \sum_{l=0}^{\infty} \sum_{z \neq y} \int_{(0,t]} \int_{(0,\tau)} \Big(v(\tau)\, R_{yz}(\tau, r) \Big)^2 \lambda_{yz}(\tau, r)\, p^{(l)}_{\eta y}(0, \tau - 0, 0, dr)\, \mathbf{\Lambda}_{yz}(d\tau) \\
&\quad - \sum_{l=0}^{\infty} \sum_{z,\xi \in \mathcal{S} \setminus \{y\}} \int_{(0,t]} \int_{(0,\tau)} (v(\tau))^2\, R_{yz}(\tau, r)\, R_{y\xi}(\tau, r)\, \lambda_{yz}(\tau, r)\, \lambda_{y\xi}(\tau, r)\, p^{(l)}_{\eta y}(0, \tau - 0, 0, dr) \\
&\qquad \cdot \mathbf{\Lambda}_{y\xi}(\{\tau\})\, \mathbf{\Lambda}_{yz}(d\tau).
\end{aligned}$$

By interchanging the infinite sums and the integrals according to the monotone convergence theorem as well as applying (2. 23. 4) the assertion follows. □

A. 21 Proof of theorem 4. 31:
The following proof is closely related to the proof of Thiele's integral equations of type 1 (cf. theorem 4. 8). Let $0 \leq s \leq t$, and $(\eta, y) \in S^2$. We start by inserting (4. 29. 1) into

$$\mathbf{E}[\langle \mathcal{L}_y \rangle_t - \langle \mathcal{L}_y \rangle_s | T_0 = s, Z_0 = \eta].$$

Thus, we obtain for $\mathcal{L}(T_0, Z_0|P)$-a.e. (s, η)

$$\begin{aligned}
&\mathbf{E}[\langle \mathcal{L}_y \rangle_t - \langle \mathcal{L}_y \rangle_s | T_0 = s, Z_0 = \eta] \\
&= \mathbf{E}\Big[\sum_{z \neq y} \sum_{l=0}^{\infty} \int_{(T_l \vee s, T_{l+1} \wedge t]} \Big(v(\tau)\, R_{yz}(\tau, \tau - T_l) \Big)^2 \mathbf{1}_{\{Z_l = y\}}\, \hat{q}_{yz}(T_l, d\tau) \\
&\qquad - \sum_{z,\xi \in \mathcal{S} \setminus \{y\}} \sum_{l=0}^{\infty} \int_{(T_l \vee s, T_{l+1} \wedge t]} (v(\tau))^2\, R_{yz}(\tau, \tau - T_l)\, R_{y\xi}(\tau, \tau - T_l) \\
&\qquad\qquad \cdot \mathbf{1}_{\{Z_l = y\}}\, \hat{q}_{yz}(T_l, \{\tau\})\ \hat{q}_{y\xi}(T_l, d\tau) \big| T_0 = s, Z_0 = \eta \Big],
\end{aligned} \tag{A. 21. 1}$$

for which we split up both addends in the same way as in (4. 8. 3). This yields for $\mathcal{L}(T_0, Z_0|P)$-a.e.

(s,η)

$$\begin{aligned}
&\mathbf{E}[\langle\mathcal{L}_y\rangle_t - \langle\mathcal{L}_y\rangle_s | T_0 = s, Z_0 = \eta] \\
&= \sum_{z\neq y} \int_{(s,t]} \int_{(s,x]} \Big(v(\tau)\, R_{yz}(\tau, \tau - s) \Big)^2 \, \delta_{\eta y}\, \hat{q}_{yz}(s, d\tau)\, P(T_1 \in dx | T_0 = s, Z_0 = \eta) \\
&\quad - \sum_{z,\xi\in\mathcal{S}\setminus\{y\}} \int_{(s,t]} \int_{(s,x]} (v(\tau))^2\, R_{yz}(\tau, \tau - s)\, R_{y\xi}(\tau, \tau - s) \\
&\qquad\qquad \cdot\, \delta_{\eta y}\, \hat{q}_{yz}(s, \{\tau\})\ \hat{q}_{y\xi}(s, d\tau)\, P(T_1 \in dx | T_0 = s, Z_0 = \eta) \\
&\quad + \sum_{l=1}^{\infty} \sum_{z\neq y} \int_{(s,t]\times\mathcal{S}\times(s,t]} \int_{(r,x]} \Big(v(\tau)\, R_{yz}(\tau, \tau - r) \Big)^2\, \mathbf{1}_{\{Z_l = y\}}\, \hat{q}_{yz}(r, d\tau) \\
&\qquad\qquad \cdot P(T_{l+1} \in dx, T_l \in dr, Z_l \in d\nu | T_0 = s, Z_0 = \eta) \\
&\quad - \sum_{l=1}^{\infty} \sum_{z,\xi\in\mathcal{S}\setminus\{y\}} \int_{(s,t]\times\mathcal{S}\times(s,t]} \int_{(r,x]} (v(\tau))^2\, R_{yz}(\tau, \tau - r)\, R_{y\xi}(\tau, \tau - r)\, \mathbf{1}_{\{Z_l = y\}}\, \hat{q}_{yz}(r, \{\tau\})\ \hat{q}_{y\xi}(r, d\tau) \\
&\qquad\qquad \cdot P(T_{l+1} \in dx, T_l \in dr, Z_l \in d\nu | T_0 = s, Z_0 = \eta). \qquad \text{(A. 21. 2)}
\end{aligned}$$

Applying Fubini's theorem and afterwards inserting (2. 8. 2) yield for the first two addends of (A. 21. 2) the following:

$$\begin{aligned}
&\sum_{z\neq y} \int_{(s,t]} \int_{(s,x]} \Big(v(\tau)\, R_{yz}(\tau, \tau - s) \Big)^2 \, \delta_{\eta y}\, \hat{q}_{yz}(s, d\tau)\, P(T_1 \in dx | T_0 = s, Z_0 = \eta) \qquad \text{(A. 21. 3)} \\
&\quad - \sum_{z,\xi\in\mathcal{S}\setminus\{y\}} \int_{(s,t]} \int_{(s,x]} (v(\tau))^2\, R_{yz}(\tau, \tau - s)\, R_{y\xi}(\tau, \tau - s) \\
&\qquad\qquad \cdot\, \delta_{\eta y}\, \hat{q}_{yz}(s, \{\tau\})\ \hat{q}_{y\xi}(s, d\tau)\, P(T_1 \in dx | T_0 = s, Z_0 = \eta) \\
&= \sum_{z\neq y} \int_{(s,t]} \Big(v(\tau)\, R_{yz}(\tau, \tau - s) \Big)^2 \, \delta_{\eta y}\, (1 - \hat{Q}_y(s, \tau - 0))\, \hat{q}_{yz}(s, d\tau) \\
&\quad - \sum_{z,\xi\in\mathcal{S}\setminus\{y\}} \int_{(s,t]} (v(\tau))^2\, R_{yz}(\tau, \tau - s)\, R_{y\xi}(\tau, \tau - s)\, \delta_{\eta y}\, \hat{q}_{yz}(s, \{\tau\})\, (1 - \hat{Q}_y(s, \tau - 0))\, \hat{q}_{y\xi}(s, d\tau).
\end{aligned}$$

By conditioning on (T_1, Z_1) and employing the Markov property of (T, Z), we get for the remaining addends of (A. 21. 2)

$$\begin{aligned}
&\sum_{\substack{\zeta\in\mathcal{S}\\ \zeta\neq\eta}} \int_{(s,t]} \sum_{l=1}^{\infty} \sum_{z\neq y} \int_{(s,t]\times\mathcal{S}\times(s,t]} \int_{(r,x]} \Big(v(\tau)\, R_{yz}(\tau, \tau - r) \Big)^2\, \mathbf{1}_{\{Z_l = y\}}\, \hat{q}_{yz}(r, d\tau) \\
&\qquad \cdot P(T_{l+1} \in dx, T_l \in dr, Z_l = \nu |\, T_1 = \sigma, Z_1 = \zeta)\, P(T_1 \in d\sigma, Z_1 = \zeta | T_0 = s, Z_0 = \eta) \\
&\quad - \sum_{\substack{\zeta\in\mathcal{S}\\ \zeta\neq\eta}} \int_{(s,t]} \sum_{l=1}^{\infty} \sum_{z,\xi\in\mathcal{S}\setminus\{y\}} \int_{(s,t]\times\mathcal{S}\times(s,t]} \int_{(r,x]} (v(\tau))^2\, R_{yz}(\tau, \tau - r)\, R_{y\xi}(\tau, \tau - r)\, \mathbf{1}_{\{Z_l = y\}}\, \hat{q}_{yz}(r, \{\tau\}) \\
&\qquad \cdot \hat{q}_{y\xi}(r, d\tau)\, P(T_{l+1} \in dx, T_l \in dr, Z_l = \nu |\, T_1 = \sigma, Z_1 = \zeta)\, P(T_1 \in d\sigma, Z_1 = \zeta | T_0 = s, Z_0 = \eta),
\end{aligned}$$

which is, by inserting (A. 20. 1) according to the fact that (T, Z) is homogeneous, equal to

$$\begin{aligned}
&\sum_{\substack{\zeta\in\mathcal{S}\\ \zeta\neq\eta}} \int_{(s,t]} \mathbf{V}^{\mathcal{L}_y(t)}_{(\zeta)}(\sigma)\, P(T_1 \in d\sigma, Z_1 = \zeta | T_0 = s, Z_0 = \eta) \\
&= \sum_{\substack{z\in\mathcal{S}\\ z\neq\eta}} \int_{(s,t]} \mathbf{V}^{\mathcal{L}_y(t)}_{(z)}(\tau)\, (1 - \hat{Q}_\eta(s, \tau - 0))\, \hat{q}_{\eta z}(s, d\tau), \qquad \text{(A. 21. 4)}
\end{aligned}$$

where the last equation is due to (2. 8. 1), and afterwards applying (2. 9. 1). Altogether, we get by adding (A. 21. 3) and (A. 21. 4) for $\mathcal{L}(T_0, Z_0|P)$-a.e. (s,η)

$$\begin{aligned}
&\mathbf{V}_{(\eta)}^{\mathcal{L}_y(t)}(s) \qquad\qquad (A.\,21.\,5)\\
&= \sum_{\substack{z\in\mathcal{S}\\ z\neq\eta}} \int_{(s,t]} \mathbf{V}_{(z)}^{\mathcal{L}_y(t)}(\tau)\,(1-\hat{Q}_\eta(s,\tau-0))\,\hat{q}_{\eta z}(s,d\tau)\\
&\quad + \sum_{\substack{z\in\mathcal{S}\\ z\neq\eta}} \int_{(s,t]} \Big(v(\tau)\,R_{\eta z}(\tau,\tau-s)\Big)^2\,\delta_{\eta y}\,(1-\hat{Q}_\eta(s,\tau-0))\,\hat{q}_{\eta z}(s,d\tau)\\
&\quad - \sum_{z,\xi\in\mathcal{S}\setminus\{\eta\}} \int_{(s,t]} (v(\tau))^2\,R_{\eta z}(\tau,\tau-s)\,R_{\eta\xi}(\tau,\tau-s)\,\delta_{\eta y}\,\hat{q}_{\eta z}(s,\{\tau\})\,(1-\hat{Q}_\eta(s,\tau-0))\,\hat{q}_{\eta\xi}(s,d\tau).
\end{aligned}$$

By substituting $\bar{p}_\eta(s,\tau-0,0)$ for $(1-\hat{Q}_\eta(s,\tau-0))$ according to (2. 28. 3), the system of integral equations (4. 31. 1) is verified. □

A. 22 Proof of theorem 5. 1:

According to (5. 0. 10), the retrospective reserve is defined as $V_y^-(t) = \mathbf{E}[V^-(t,A)|X_t=y]$, $t \geq 0, y \in \mathcal{S}$. In case that $P(X_t = y) > 0$, this is equal to

$$V_y^-(t) = \frac{\mathbf{E}\left[V^-(t,A)\,\mathbf{1}_{\{X_t=y\}}\right]}{P(X_t=y)}. \qquad (A.\,22.\,1)$$

Hence, we consider $\mathbf{E}\left[V^-(t,A)\,\mathbf{1}_{\{X_t=y\}}\right]$. According to (2. 4. 2), we get

$$\mathbf{E}\left[V^-(t,A)\,\mathbf{1}_{\{\exists m\in\mathbb{N}_0: T_m\leq t<T_{m+1}, Z_m=y\}}\right] = \sum_{m\in\mathbb{N}_0} \mathbf{E}\left[V^-(t,A)\,\mathbf{1}_{\{T_m\leq t<T_{m+1}, Z_m=y\}}\right]. \qquad (A.\,22.\,2)$$

Inserting (5. 0. 1) yields

$$\begin{aligned}
\mathbf{E}\left[V^-(t,A)\,\mathbf{1}_{\{X_t=y\}}\right] &= \sum_{m\in\mathbb{N}_0} \mathbf{E}\Big[-K(t)\sum_{l\in\mathbb{N}_0}\int_{[T_l\wedge t, T_{l+1}\wedge t)} v(\tau)\,F_{Z_l}(d\tau)\,\mathbf{1}_{\{T_m\leq t, T_{m+1}>t, Z_m=y\}}\Big]\\
&\quad + \sum_{m\in\mathbb{N}_0} \mathbf{E}\Big[-K(t)\sum_{l\in\mathbb{N}_0}\frac{D_{Z_l Z_{l+1}}(T_{l+1})}{K(DT(T_{l+1}))}\,\mathbf{1}_{\{T_{l+1}\leq t\}}\,\mathbf{1}_{\{T_m\leq t, T_{m+1}>t, Z_m=y\}}\Big],
\end{aligned}$$

which can also be derived in a manner similar to the proof of theorem 4. 6 (cf. A.13). In doing so, we obtain

$$\begin{aligned}
&-v(t)\,\mathbf{E}\left[V^-(t,A)\,\mathbf{1}_{\{X_t=y\}}\right]\\
&= \sum_{m\in\mathbb{N}_0}\sum_{l<m}\int_{[0,\infty)\times\mathcal{S}\times[0,\infty)\times\mathcal{S}}\int_{[r\wedge t, x\wedge t)} v(\tau)\,F_\xi(d\tau)\\
&\qquad\qquad \cdot P(T_{l+1}\in dx, Z_{l+1}\in dz, T_l\in dr, Z_l\in d\xi, T_m\leq t, T_{m+1}>t, Z_m=y)\\
&\quad + \sum_{m\in\mathbb{N}_0}\sum_{l<m}\int_{[0,\infty)\times\mathcal{S}\times[0,\infty)\times\mathcal{S}} \frac{D_{\xi z}(x)}{K(DT(x))}\,\mathbf{1}_{\{x\leq t\}}\\
&\qquad\qquad \cdot P(T_{l+1}\in dx, Z_{l+1}\in dz, T_l\in dr, Z_l\in d\xi, T_m\leq t, T_{m+1}>t, Z_m=y)\\
&= \sum_{(\xi,z)\in\mathcal{J}}\sum_{m\in\mathbb{N}_0}\sum_{l<m}\int_{[0,t)}\int_{(0,t]}\int_{[r,x)} v(\tau)\,F_\xi(d\tau)\\
&\qquad\qquad \cdot P(T_{l+1}\in dx, Z_{l+1}\in dz, T_l\in dr, Z_l=\xi, T_m\leq t, T_{m+1}>t, Z_m=y)\\
&\quad + \sum_{(\xi,z)\in\mathcal{J}}\sum_{m\in\mathbb{N}_0}\sum_{l<m}\int_{[0,t)}\int_{(0,t]} \frac{D_{\xi z}(x)}{K(DT(x))}\\
&\qquad\qquad \cdot P(T_{l+1}\in dx, Z_{l+1}\in dz, T_l\in dr, Z_l\in d\xi, T_m\leq t, T_{m+1}>t, Z_m=y). \qquad (A.\,22.\,3)
\end{aligned}$$

Upon successive conditioning, applying the Markov property of (T, Z) along with $P(T_0 = 0, Z_0 = a) = 1$, and changing the order of the double series, the above equation can be continued as

$$
\begin{aligned}
&= \sum_{(\xi,z)\in\mathcal{J}} \sum_{m\in\mathbb{N}_0} \sum_{l<m} \int_{[0,t)} \int_{(0,t]} \int_{[r,x)} v(\tau)\, F_\xi(d\tau)\, P(T_m \le t, T_{m+1} > t, Z_m = y |\, T_{l+1} = x, Z_{l+1} = z) \\
&\qquad \cdot P(T_{l+1} \in dx, Z_{l+1} \in dz |\, T_l = r, Z_l = \xi)\, P(T_l \in dr, Z_l = \xi | T_0 = 0, Z_0 = a) \\
&\quad + \sum_{(\xi,z)\in\mathcal{J}} \sum_{m\in\mathbb{N}_0} \sum_{l<m} \int_{[0,t)} \int_{(0,t]} \frac{D_{\xi z}(x)}{K(DT(x))}\, P(T_m \le t, T_{m+1} > t, Z_m = y |\, T_{l+1} = x, Z_{l+1} = z) \\
&\qquad \cdot P(T_{l+1} \in dx, Z_{l+1} \in dz |\, T_l = r, Z_l = \xi)\, P(T_l \in dr, Z_l = \xi | T_0 = 0, Z_0 = a) \\
&= \sum_{(\xi,z)\in\mathcal{J}} \sum_{l\in\mathbb{N}_0} \sum_{m>l} \int_{[0,t)} \int_{(0,t]} \int_{[r,x)} v(\tau)\, F_\xi(d\tau)\, P(T_m \le t, T_{m+1} > t, Z_m = y |\, T_{l+1} = x, Z_{l+1} = z) \\
&\qquad \cdot P(T_{l+1} \in dx, Z_{l+1} \in dz |\, T_l = r, Z_l = \xi)\, P(T_l \in dr, Z_l = \xi | T_0 = 0, Z_0 = a) \\
&\quad + \sum_{(\xi,z)\in\mathcal{J}} \sum_{l\in\mathbb{N}_0} \sum_{m>l} \int_{[0,t)} \int_{(0,t]} \frac{D_{\xi z}(x)}{K(DT(x))}\, P(T_m \le t, T_{m+1} > t, Z_m = y |\, T_{l+1} = x, Z_{l+1} = z) \\
&\qquad \cdot P(T_{l+1} \in dx, Z_{l+1} \in dz |\, T_l = r, Z_l = \xi)\, P(T_l \in dr, Z_l = \xi | T_0 = 0, Z_0 = a).
\end{aligned}
$$

By the monotone convergence theorem, the order of the sums and the integrals can also be changed. This, along with the homogeneity of (T, Z), yield

$$
\begin{aligned}
&-v(t)\, \mathbf{E}\left[V^-(t, A)\, \mathbf{1}_{\{X_t = y\}}\right] \\
&= \sum_{(\xi,z)\in\mathcal{J}} \sum_{l\in\mathbb{N}_0} \int_{[0,t)} \int_{(0,t]} \int_{[r,x)} v(\tau)\, F_\xi(d\tau) \sum_{m\in\mathbb{N}_0} P(T_m \le t, T_{m+1} > t, Z_m = y |\, T_0 = x, Z_0 = z) \\
&\qquad \cdot P(T_{l+1} \in dx, Z_{l+1} \in dz |\, T_l = r, Z_l = \xi)\, P(T_l \in dr, Z_l = \xi | T_0 = 0, Z_0 = a) \\
&\quad + \sum_{(\xi,z)\in\mathcal{J}} \sum_{l\in\mathbb{N}_0} \int_{[0,t)} \int_{(0,t]} \frac{D_{\xi z}(x)}{K(DT(x))} \sum_{m\in\mathbb{N}_0} P(T_m \le t, T_{m+1} > t, Z_m = y |\, T_0 = x, Z_0 = z) \\
&\qquad \cdot P(T_{l+1} \in dx, Z_{l+1} \in dz |\, T_l = r, Z_l = \xi)\, P(T_l \in dr, Z_l = \xi | T_0 = 0, Z_0 = a).
\end{aligned}
$$

Employing (2. 42. 1) and (2. 42. 2), and afterwards applying (2. 41. 6) as well as (2. 41. 5), the above can be continued as

$$
\begin{aligned}
&-v(t)\, \mathbf{E}\left[V^-(t, A)\, \mathbf{1}_{\{X_t = y\}}\right] \\
&= \sum_{(\xi,z)\in\mathcal{J}} \sum_{l\in\mathbb{N}_0} \int_{[0,t)} \int_{(0,t]} \int_{[r,x)} v(\tau)\, F_\xi(d\tau)\, p_{zy}(x,t)\, \bar{p}_\xi(r, x-0)\, q_{\xi z}(dx) \\
&\qquad \cdot P(T_l \in dr, Z_l = \xi | T_0 = 0, Z_0 = a) \\
&\quad + \sum_{(\xi,z)\in\mathcal{J}} \sum_{l\in\mathbb{N}_0} \int_{[0,t)} \int_{(0,t]} \frac{D_{\xi z}(x)}{K(DT(x))}\, p_{zy}(x,t)\, \bar{p}_\xi(r, x-0)\, q_{\xi z}(dx) \\
&\qquad \cdot P(T_l \in dr, Z_l = \xi | T_0 = 0, Z_0 = a). \qquad \text{(A. 22. 4)}
\end{aligned}
$$

A repeatedly application of Fubini's theorem, and, for the first addend, applying (2. 41. 9) and inserting (2. 42. 4), result in the following:

$$
\begin{aligned}
&-v(t)\,\mathbf{E}\left[V^-(t,A)\,\mathbf{1}_{\{X_t=y\}}\right]\\
&=\sum_{\xi\in\mathcal{S}}\sum_{l\in\mathbb{N}_0}\int_{[0,t)}\int_{[r,t)} v(\tau)\,\bar{p}_\xi(r,\tau)\sum_{\substack{z\in\mathcal{S}\\ z\neq\xi}}\int_{(\tau,t]} p_{zy}(x,t)\,\bar{p}_\xi(\tau,x-0)\,q_{\xi z}(dx)\,F_\xi(d\tau)\\
&\qquad\cdot P(T_l\in dr, Z_l=\xi|T_0=0,Z_0=a)\\
&\quad+\sum_{(\xi,z)\in\mathcal{J}}\sum_{l\in\mathbb{N}_0}\int_{[0,t)}\int_{(0,t]}\frac{D_{\xi z}(x)}{K(DT(x))}\,p_{zy}(x,t)\,\bar{p}_\xi(r,x-0)\,q_{\xi z}(dx)\\
&\qquad\cdot P(T_l\in dr, Z_l=\xi|T_0=0,Z_0=a)\\
&=\sum_{\xi\in\mathcal{S}}\sum_{l\in\mathbb{N}_0}\int_{[0,t)}\int_{[r,t)} v(\tau)\,\bar{p}_\xi(r,\tau)\,p_{\xi y}(\tau,t)\,F_\xi(d\tau)\\
&\qquad\cdot P(T_l\in dr, Z_l=\xi|T_0=0,Z_0=a)\\
&\quad+\sum_{(\xi,z)\in\mathcal{J}}\sum_{l\in\mathbb{N}_0}\int_{[0,t)}\int_{(0,t]}\frac{D_{\xi z}(x)}{K(DT(x))}\,p_{zy}(x,t)\,\bar{p}_\xi(r,x-0)\,q_{\xi z}(dx)\\
&\qquad\cdot P(T_l\in dr, Z_l=\xi|T_0=0,Z_0=a)\\
&=\sum_{\xi\in\mathcal{S}}\sum_{l\in\mathbb{N}_0}\int_{[0,t)} v(\tau)\int_{[0,\tau]}\bar{p}_\xi(r,\tau)\,P(T_l\in dr, Z_l=\xi|T_0=0,Z_0=a)\,p_{\xi y}(\tau,t)\,F_\xi(d\tau)\\
&\quad+\sum_{(\xi,z)\in\mathcal{J}}\sum_{l\in\mathbb{N}_0}\int_{(0,t]}\frac{D_{\xi z}(x)}{K(DT(x))}\,p_{zy}(x,t)\\
&\qquad\cdot\int_{[0,t)}\bar{p}_\xi(r,x-0)\,P(T_l\in dr, Z_l=\xi|T_0=0,Z_0=a)\,q_{\xi z}(dx)\\
&=\sum_{\xi\in\mathcal{S}}\sum_{l\in\mathbb{N}_0}\int_{[0,t)} v(\tau)\,p^{(l)}_{a\xi}(0,\tau)\,p_{\xi y}(\tau,t)\,F_\xi(d\tau)\\
&\quad+\sum_{(\xi,z)\in\mathcal{J}}\sum_{l\in\mathbb{N}_0}\int_{(0,t]}\frac{D_{\xi z}(x)}{K(DT(x))}\,p_{zy}(x,t)\,p^{(l)}_{a\xi}(0,x-0)\,q_{\xi z}(dx).
\end{aligned}
$$

Interchanging the infinite sum and the integral again yields along with (2. 42. 1)

$$
\begin{aligned}
-v(t)\,\mathbf{E}\left[V^-(t,A)\,\mathbf{1}_{\{X_t=y\}}\right] &= \sum_{\xi\in\mathcal{S}}\int_{[0,t)} v(\tau)\,p_{a\xi}(0,\tau)\,p_{\xi y}(\tau,t)\,F_\xi(d\tau)\\
&\quad+\sum_{(\xi,z)\in\mathcal{J}}\int_{(0,t]}\frac{D_{\xi z}(\tau)}{K(DT(\tau))}\,p_{a\xi}(0,\tau-0)\,p_{zy}(\tau,t)\,q_{\xi z}(d\tau),
\end{aligned}
$$

from which, the assertion (5. 1. 1) immediately follows by using (A. 22. 1) together with $P(T_0 = 0, Z_0 = a) = 1$. □

A. 23 Proof of theorem 5. 2:
According to definition (5. 1. 6), one gets for $0 \le v \le t$ and $y \in \mathcal{S}$ satisfying $P(X_t = y, U_t \le v) > 0$

$$
V^-_{(y,v)}(t) = \frac{\mathbf{E}\left[V^-(t,A)\,\mathbf{1}_{\{X_t=y,U_t\le v\}}\right]}{P(X_t=y,U_t\le v)}. \qquad \text{(A. 23. 1)}
$$

Hence, we likewise consider the numerator and obtain by using (2.4.2) and inserting (3.12.1)

$$\begin{aligned}
&\mathbf{E}\left[V^-(t,A)\,\mathbf{1}_{\{X_t=y,U_t\le v\}}\right]\\
&=\sum_{m\in\mathbb{N}_0}\mathbf{E}\left[-K(t)\sum_{l\in\mathbb{N}_0}\int_{[T_l\wedge t,T_{l+1}\wedge t)} v(\tau)\,\hat{F}_{Z_l}(T_l,d\tau)\,\mathbf{1}_{\{T_{m+1}>t,T_m\in[t-v,t],Z_m=y\}}\right]\\
&\quad+\sum_{m\in\mathbb{N}_0}\mathbf{E}\left[-K(t)\sum_{l\in\mathbb{N}_0}\frac{D_{Z_lZ_{l+1}}(T_{l+1},T_{l+1}-T_l)}{K(DT(T_{l+1}))}\,\mathbf{1}_{\{T_{l+1}\le t\}}\,\mathbf{1}_{\{T_{m+1}>t,T_m\in[t-v,t],Z_m=y\}}\right].
\end{aligned}$$

Further, we proceed in almost the same manner as in the proof of theorem 5.1. Thus, we obtain

$$\begin{aligned}
&-v(t)\,\mathbf{E}\left[V^-(t,A)\,\mathbf{1}_{\{X_t=y,U_t\le v\}}\right]\\
&=\sum_{m\in\mathbb{N}_0}\sum_{l<m}\int_{[0,\infty)\times\mathcal{S}\times[0,\infty)\times\mathcal{S}}\int_{[r\wedge t,x\wedge t)} v(\tau)\,F_\xi(r,d\tau)\\
&\qquad\cdot P(T_{l+1}\in dx,Z_{l+1}\in dz,T_l\in dr,Z_l\in d\xi,T_{m+1}>t,T_m\in[t-v,t],Z_m=y)\\
&\quad+\sum_{m\in\mathbb{N}_0}\sum_{l<m}\int_{[0,\infty)\times\mathcal{S}\times[0,\infty)\times\mathcal{S}}\frac{D_{\xi z}(x,x-r)}{K(DT(x))}\,\mathbf{1}_{\{x\le t\}}\\
&\qquad\cdot P(T_{l+1}\in dx,Z_{l+1}\in dz,T_l\in dr,Z_l\in d\xi,T_{m+1}>t,T_m\in[t-v,t],Z_m=y)\\
&=\sum_{(\xi,z)\in\mathcal{J}}\sum_{m\in\mathbb{N}_0}\sum_{l<m}\int_{[0,t)}\int_{(0,t]}\int_{[r,x)} v(\tau)\,F_\xi(r,d\tau)\\
&\qquad\cdot P(T_{l+1}\in dx,Z_{l+1}\in dz,T_l\in dr,Z_l=\xi,T_{m+1}>t,T_m\in[t-v,t],Z_m=y)\\
&\quad+\sum_{(\xi,z)\in\mathcal{J}}\sum_{m\in\mathbb{N}_0}\sum_{l<m}\int_{[0,t)}\int_{(0,t]}\frac{D_{\xi z}(x,x-r)}{K(DT(x))}\\
&\qquad\cdot P(T_{l+1}\in dx,Z_{l+1}\in dz,T_l\in dr,Z_l\in d\xi,T_{m+1}>t,T_m\in[t-v,t],Z_m=y).
\end{aligned} \tag{A.23.2}$$

Upon successive conditioning, applying the Markov property of (T,Z) together with $P(T_0=0,Z_0=a)=1$, and changing the order of the double series, the above equation can be continued as

$$\begin{aligned}
&=\sum_{(\xi,z)\in\mathcal{J}}\sum_{m\in\mathbb{N}_0}\sum_{l<m}\int_{[0,t)}\int_{(0,t]}\int_{[r,x)} v(\tau)\,F_\xi(r,d\tau)\,P(T_{m+1}>t,T_m\in[t-v,t],Z_m=y|\,T_{l+1}=x,Z_{l+1}=z)\\
&\qquad\cdot P(T_{l+1}\in dx,Z_{l+1}\in dz|\,T_l=r,Z_l=\xi)\,P(T_l\in dr,Z_l=\xi|T_0=0,Z_0=a)\\
&\quad+\sum_{(\xi,z)\in\mathcal{J}}\sum_{m\in\mathbb{N}_0}\sum_{l<m}\int_{[0,t)}\int_{(0,t]}\frac{D_{\xi z}(x,x-r)}{K(DT(x))}\,P(T_{m+1}>t,T_m\in[t-v,t],Z_m=y|\,T_{l+1}=x,Z_{l+1}=z)\\
&\qquad\cdot P(T_{l+1}\in dx,Z_{l+1}\in dz|\,T_l=r,Z_l=\xi)\,P(T_l\in dr,Z_l=\xi|T_0=0,Z_0=a)\\
&=\sum_{(\xi,z)\in\mathcal{J}}\sum_{l\in\mathbb{N}_0}\sum_{m>l}\int_{[0,t)}\int_{(0,t]}\int_{[r,x)} v(\tau)\,F_\xi(r,d\tau)\,P(T_{m+1}>t,T_m\in[t-v,t],Z_m=y|\,T_{l+1}=x,Z_{l+1}=z)\\
&\qquad\cdot P(T_{l+1}\in dx,Z_{l+1}\in dz|\,T_l=r,Z_l=\xi)\,P(T_l\in dr,Z_l=\xi|T_0=0,Z_0=a)\\
&\quad+\sum_{(\xi,z)\in\mathcal{J}}\sum_{l\in\mathbb{N}_0}\sum_{m>l}\int_{[0,t)}\int_{(0,t]}\frac{D_{\xi z}(x,x-r)}{K(DT(x))}\,P(T_{m+1}>t,T_m\in[t-v,t],Z_m=y|\,T_{l+1}=x,Z_{l+1}=z)\\
&\qquad\cdot P(T_{l+1}\in dx,Z_{l+1}\in dz|\,T_l=r,Z_l=\xi)\,P(T_l\in dr,Z_l=\xi|T_0=0,Z_0=a).
\end{aligned}$$

According to the monotone convergence theorem, the order of summation and integration can also be changed. By additionally applying the homogeneity of (T, Z), we obtain

$$-v(t)\,\mathbf{E}\left[V^{-}(t,A)\,\mathbf{1}_{\{X_t=y,U_t\le v\}}\right]$$
$$= \sum_{(\xi,z)\in\mathcal{J}} \sum_{l\in\mathbb{N}_0} \int_{[0,t)} \int_{(0,t]} \int_{[r,x)} v(\tau)\, F_\xi(d\tau) \sum_{m\in\mathbb{N}_0} P(T_{m+1} > t, T_m \in [t-v,t], Z_m = y|\, T_0 = x, Z_0 = z)$$
$$\cdot P(T_{l+1}\in dx, Z_{l+1}\in dz|\, T_l = r, Z_l = \xi)\, P(T_l \in dr, Z_l = \xi | T_0 = 0, Z_0 = a)$$
$$+ \sum_{(\xi,z)\in\mathcal{J}} \sum_{l\in\mathbb{N}_0} \int_{[0,t)} \int_{(0,t]} \frac{D_{\xi z}(x)}{K(DT(x))} \sum_{m\in\mathbb{N}_0} P(T_{m+1} > t, T_m \in [t-v,t], Z_m = y|\, T_0 = x, Z_0 = z)$$
$$\cdot P(T_{l+1}\in dx, Z_{l+1}\in dz|\, T_l = r, Z_l = \xi)\, P(T_l \in dr, Z_l = \xi | T_0 = 0, Z_0 = a).$$

Now we insert (2. 26. 2) and (2. 8. 1), and apply (2. 9. 1). Doing so, the above equation can be continued as

$$= \sum_{(\xi,z)\in\mathcal{J}} \sum_{l\in\mathbb{N}_0} \int_{[0,t)} \int_{(0,t]} \int_{[r,x)} v(\tau)\, F_\xi(r,d\tau)\, p_{zy}(x,t,0,v)\, (1-\hat{Q}_\xi(r,x-0))\, \hat{q}_{\xi z}(r,dx)$$
$$\cdot P(T_l \in dr, Z_l = \xi | T_0 = 0, Z_0 = a)$$
$$+ \sum_{(\xi,z)\in\mathcal{J}} \sum_{l\in\mathbb{N}_0} \int_{[0,t)} \int_{(0,t]} \frac{D_{\xi z}(x,x-r)}{K(DT(x))}\, p_{zy}(x,t,0,v)\, (1-\hat{Q}_\xi(r,x-0))\, \hat{q}_{\xi z}(r,dx)$$
$$\cdot P(T_l \in dr, Z_l = \xi | T_0 = 0, Z_0 = a),$$

which is according to the specification of q by means of $\hat{q}$ equal to

$$-v(t)\,\mathbf{E}\left[V^{-}(t,A)\,\mathbf{1}_{\{X_t=y,U_t\le v\}}\right]$$
$$= \sum_{(\xi,z)\in\mathcal{J}} \sum_{l\in\mathbb{N}_0} \int_{[0,t)} \int_{(0,t]} \int_{[r,x)} v(\tau)\, F_\xi(r,d\tau)\, p_{zy}(x,t,0,v)\, \bar{p}_\xi(r,x-0,0)\, q_{\xi z}(r,dx,0)$$
$$\cdot P(T_l \in dr, Z_l = \xi | T_0 = 0, Z_0 = a)$$
$$+ \sum_{(\xi,z)\in\mathcal{J}} \sum_{l\in\mathbb{N}_0} \int_{[0,t)} \int_{(0,t]} \frac{D_{\xi z}(x,x-r)}{K(DT(x))}\, p_{zy}(x,t,0,v)\, \bar{p}_\xi(r,x-0,0)\, q_{\xi z}(r,dx,0)$$
$$\cdot P(T_l \in dr, Z_l = \xi | T_0 = 0, Z_0 = a). \qquad \text{(A. 23. 3)}$$

By employing assumption 4. 5, Fubini's theorem, (2. 27. 1), (2. 27. 8), and the backward integral equations (2. 32. 2), the first addend of the above equation turns out to be

$$\sum_{\xi\in\mathcal{S}} \sum_{l\in\mathbb{N}_0} \int_{[0,t)} \int_{[r,t)} v(\tau)\, f_\xi(\tau,\tau-r)\, \bar{p}_\xi(r,\tau,0) \sum_{\substack{z\in\mathcal{S}\\ z\neq\xi}} \int_{(\tau,t]} p_{zy}(x,t,0,v)\, (1-Q_\xi(\tau,x-0,\tau-r))$$
$$\cdot\, q_{\xi z}(\tau,dx,\tau-r)\, \mathbf{F}_\xi(d\tau)\, P(T_l \in dr, Z_l = \xi | T_0 = 0, Z_0 = a)$$
$$= \sum_{\xi\in\mathcal{S}} \sum_{l\in\mathbb{N}_0} \int_{[0,t)} \int_{(0,t]} v(\tau)\, f_\xi(\tau,\tau-r)\, \bar{p}_\xi(r,\tau,0)\, p_{\xi y}(\tau,t,\tau-r,v)\, \mathbf{F}_\xi(d\tau)$$
$$\cdot\, P(T_l \in dr, Z_l = \xi | T_0 = 0, Z_0 = a).$$

Along with assumption 2. 35, this yields for (A. 23. 3)

$$-v(t)\,\mathbf{E}\left[V^{-}(t,A)\,\mathbf{1}_{\{X_t=y,U_t\le v\}}\right]$$
$$= \sum_{\xi\in\mathcal{S}} \sum_{l\in\mathbb{N}_0} \int_{[0,t)} \int_{[r,t)} v(\tau)\, f_\xi(\tau,\tau-r)\, (1-Q_\xi(r,\tau,0))\, p_{\xi y}(\tau,t,\tau-r,v)\, \mathbf{F}_\xi(d\tau)$$
$$\cdot\, P(T_l \in dr, Z_l = \xi | T_0 = 0, Z_0 = a)$$
$$+ \sum_{(\xi,z)\in\mathcal{J}} \sum_{l\in\mathbb{N}_0} \int_{[0,t)} \int_{(0,t]} \frac{D_{\xi z}(x,x-r)}{K(DT(x))}\, p_{zy}(x,t,0,v)\, (1-Q_\xi(r,x-0,0))\, \lambda_{\xi z}(x,x-r)\, \mathbf{\Lambda}_{\xi z}(dx)$$
$$\cdot P(T_l \in dr, Z_l = \xi | T_0 = 0, Z_0 = a). \qquad \text{(A. 23. 4)}$$

After employing Fubini's theorem again, the same argumentation as in (A. 12. 7), (A. 12. 8), and (A. 12. 9) results in

$$
\begin{aligned}
&-v(t)\,\mathbf{E}\left[V^{-}(t,A)\,\mathbf{1}_{\{X_t=y,U_t\le v\}}\right]\\
&=\sum_{\xi\in\mathcal{S}}\sum_{l\in\mathbb{N}_0}\int_{[0,t)}\int_{[0,\tau]} v(\tau)\,f_\xi(\tau,\tau-r)\;p_{\xi y}(\tau,t,\tau-r,v)\,p_{a\xi}^{(l)}(0,\tau,0,\tau-dr)\,\mathbf{F}_\xi(d\tau)\\
&\quad+\sum_{(\xi,z)\in\mathcal{J}}\sum_{l\in\mathbb{N}_0}\int_{(0,t]}\int_{[0,t)}\frac{D_{\xi z}(\tau,\tau-r)}{K(DT(\tau))}\,p_{zy}(\tau,t,0,v)\;\lambda_{\xi z}(\tau,\tau-r)\,p_{a\xi}^{(l)}(0,\tau-0,0,\tau-dr)\,\mathbf{\Lambda}_{\xi z}(d\tau)\\
&=\sum_{\xi\in\mathcal{S}}\sum_{l\in\mathbb{N}_0}\int_{[0,t)}\int_{[0,\tau]} v(\tau)\,f_\xi(\tau,\tau-r)\;p_{\xi y}(\tau,t,\tau-r,v)\,p_{a\xi}^{(l)}(0,\tau,0,\tau-dr)\,\mathbf{F}_\xi(d\tau)\\
&\quad+\sum_{(\xi,z)\in\mathcal{J}}\sum_{l\in\mathbb{N}_0}\int_{(0,t]}\int_{[0,\tau)}\frac{D_{\xi z}(\tau,\tau-r)}{K(DT(\tau))}\,p_{zy}(\tau,t,0,v)\;\lambda_{\xi z}(\tau,\tau-r)\,p_{a\xi}^{(l)}(0,\tau-0,0,\tau-dr)\,\mathbf{\Lambda}_{\xi z}(d\tau)\\
&=\sum_{\xi\in\mathcal{S}}\int_{[0,t)}\int_{[0,\infty)} v(\tau)\,f_\xi(\tau,r)\;p_{\xi y}(\tau,t,r,v)\,p_{a\xi}(0,\tau,0,dr)\,\mathbf{F}_\xi(d\tau)\\
&\quad+\sum_{(\xi,z)\in\mathcal{J}}\int_{(0,t]}\int_{(0,\infty)}\frac{D_{\xi z}(\tau,r)}{K(DT(\tau))}\,p_{zy}(\tau,t,0,v)\;\lambda_{\xi z}(\tau,r)\,p_{a\xi}(0,\tau-0,0,dr)\,\mathbf{\Lambda}_{\xi z}(d\tau).
\end{aligned}
\tag{A. 23. 5}
$$

Consider the second of the above equations. The restriction of the integration interval for the interior integral of the second addend is due to the fact that the measure

$$q_{\xi z}(r,d\tau,0)=\lambda_{\xi z}(\tau,\tau-r)\,\mathbf{\Lambda}_{\xi z}(d\tau)$$

is concentrated on (r,∞). The third equation is by the interchange of infinite sums and integrals according to the monotone convergence theorem, along with (2. 23. 4). Finally, inserting (A. 23. 5) into (A. 23. 1) verifies together with $P(T_0=0,Z_0=a)=1$ the assertion. □

Bibliography

[1993] **P.K. Andersen, O. Borgan, R.D. Gill and N. Keiding** *Statistical Models Based on Counting Processes,* Springer, Berlin 1993

[2001] **H. Bauer** *Measure and Integration Theorie,* de Gruyter, Berlin 2001

[1997] **N.L. Bowers, H.U. Gerber, J.C. Hickman, D.A. Jones & C.J. Nesbitt** *Actuarial Mathematics,* The Society of Actuaries, Schaumburg (Illinois) 1997

[1968] **L. Breiman** *Probability,* Addison-Wesley, Reading 1968

[1981] **P. Brémaud** *Point Processes and Queues - Martingale Dynamics,* Springer, Berlin 1981

[2001] **K.L. Chung** *A Course in Probability Theory,* Academic Press, San Diego 2001, third edition

[1969] **E. Cinlar** *On semi-Markov processes on arbitrary spaces,* in Proc. Camb. Phil. Soc. 66, 1969, 381-392

[1955] **D.R. Cox** *The analysis of non-Markovian stochastic processes by the inclusion of supplementary variables,* in Proc. Camb. Philos. Soc. 51, 1955, 433-441

[2004] *Gesetz über die Beaufsichtigung der Versicherungsunternehmen (Versicherungsaufsichtsgesetz - VAG),* siehe z.B. http://www.versicherungsgesetze.de/versicherungsaufsichtsgesetz, Stand 01.12.2007

[2007] *Gesetz zur Stärkung des Wettbewerbs in der gesetzlichen Krankenversicherung (GKV-Wettbewerbsstärkungsgesetz - GKV-WSG),* siehe z.B. http://www.buzer.de/gesetz/7655, Stand 01.04.2007

[1975] **I.I. Gihman & A.V. Skorohod** *The Theory of Stochastic Processes II,* Springer, Heidelberg 1975

[1999] **S. Haberman & E. Pitacco** *Actuarial Models for Disability Insurance,* Chapman & Hall, Boca Raton 1999

[2007] **M. Helwich & H. Milbrodt** *Verweildauereffekte in der deutschen Privaten Krankenversicherung,* Preprint an der Universität Rostock, Juli 2007

[1969] **J.M. Hoem** *Markov chain models in life insurance,* in Blätter der DGVM IX, 1969, 91-107

[1972] **J.M. Hoem** *Inhomogeneous semi-Markov processes, select actuarial tables, and duration-depedence in demography,* in Population Dynamics, Academic Press New York, 1972, 251-296

[1988] **J.M. Hoem** *The versatility of the Markov chain as a tool in the mathematics of life insurance,* in Transactions of the 23rd International Congress of Actuaries, Keynote lecture for subject 3, 141-202

[1972] **M. Jacobsen** *A characterization of minimal Markov jump processes,* in Zur Wahrscheinlichkeitstheorie verw. Geb. 23, 1972, 32-46

[1975] **J. Jacod** *Multivariate point processes: Predictable projection, Radon-Nikodym derivatives, representation of martingales,* in Zur Wahrscheinlichkeitstheorie verw. Geb. 31, 1975, 235 -253

[1987] **J. Jacod & A.N. Shiryaev** *Limit Theorems for Stochastic Processes,* Springer, Berlin 1987

[1984] **J. Janssen & R. De Dominicis** *Finite non-homogeneous semi-Markov processes: theoretical and computational aspects,* in Insurance: Mathematics and Economics 3, 1984, 157-165

[1998] **F.C. Klebaner** *Introduction to Stochastic Calculus with Applications,* Imperial College Press, London 1998

[1974] **V.S. Korolyuk, S.M. Brodi, A.F. Turbin** *Semi-Markov processes and their applications,* in Journal of Soviet Mathematics (4), 1974, 244-280

[1995] **G. Last & A. Brandt** *Marked Point Processes on the Real Line - The Dynamic Approach,* Springer, Berlin 1995

[1997] **H. Milbrodt & A. Stracke** *Markov models and Thiele's integral equations for the prospective reserve,* in Insurance: Mathematics and Economics, Nr. 19, 1997, 187-235

[1999] **H. Milbrodt & M. Helbig** *Mathematische Methoden der Personenversicherung,* Walter de Gruyter, Berlin 1999

[2005] **H. Milbrodt** *Aktuarielle Methoden der Deutschen Privaten Krankenversicherung,* Verlag Versicherungswirtschaft GmbH, Karlsruhe 2005

[1993] **C.M. Møller** *A stochastic version of Thiele's differential equation,* in Scandinavian Actuarial Journal, 1993, 1-16

[1996] **H.G. Möller & H.J. Zwiesler** *Mehrzustandsmodelle in der Berufsunfähigkeitsversicherung,* in Blätter der DGVM XXII, 1996, 479-499

[1980] **V. Nollau** *Semi-Markovsche Prozesse,* Akademie-Verlag, Berlin 1980

[1991] **R. Norberg** *Reserves in life and pension insurance,* in Scandinavian Actuarial Journal, 1991, 3-24

[1992] **R. Norberg** *Hattendorf's theorem and Thiele's differential equation generalized,* in Scandinavian Actuarial Journal, 1992, 2-14

[1995] **E. Pitacco** *Actuarial models for pricing disability benefits: Towards a unifying approach,* in Insurance: Mathematics and Economics, Nr. 16, 1995, 39-62

[1964] **R. Pyke & R.A. Schaufele** *Limit theorems for Markov renewal processes,* in Ann. Math. Stat. 35, No. 4, 1964

[1988] **H. Ramlau-Hansen** *Hattendorff's theorem: A Markov chain and counting process approach,* in Scandinavian Actuarial Journal, 1988, 143-156

[2007] **H.-W. Richter** *GKV-WSG: Anmerkungen zur Portabilität der Alterungsrückstellungen,* Vortrag bei der Jahrestagung der DAV, Kranken-Gruppe, Berlin 2007

[2000] **F. Rudolph** *Anwendungen der Überlebenszeitanalyse in der Pflegeversicherung,* Diplomarbeit im Zentrum Mathematik der Technischen Universität München, München 2000

[2005] **J. Rudolph** *Stornowahrscheinlichkeiten in der Privaten Krankenversicherung,* in Der Aktuar 11, Heft 3, 2005

[2007] **W. Seger, N.-A. Sittaro, R. Lohse, J. Rabba & J. Post** *Hannover Morbiditäts- und Mortalitäts-Pflegestudie (HMMPS): Langzeitverläufe, Pflegestufenübergänge und Reaktivierungen in der gesetzlichen Pflegeversicherung,* Preprint der E+S Rückversicherungs AG Hannover, Oktober 2007

[1993] **G. Segerer** *The actuarial treatment of the disability risk in Germany, Austria and Switzerland,* in Insurance: Mathematics and Economics, Nr. 13, 1993, 131-140

[1997] **A. Stracke** *Markov-Modelle und Thielesche Integralgleichungen für das prospektive Deckungskapital,* Dissertation an der Universität zu Köln, Köln 1997

[1996] *Verordnung über die versicherungsmathematischen Methoden zur Prämienkalkulation und zur Berechnung der Alterungsrückstellung in der privaten Krankenversicherung (Kalkulationsverordnung - KalV),* siehe z.B. http://www.buzer.de/gesetz/5735, Stand 27.11.2007

[1984] **H.R. Waters** *An approach to the study of multiple state models,* in Journal of the Institute of Actuaries, Nr. 111, 1984, 363-375

[1989] **H.R. Waters** *Some aspects of the modelling of permanent health insurance,* in Journal of the Institute of Actuaries, Nr. 116, 1989, 611-624

[2002] **C. Wetzel** *Modellierung eines speziellen Lebensversicherungsproduktes mit Hilfe von Semi-Markov-Prozessen,* ifa-Schriftenreihe, Ulm 2002

[2003] **C. Wetzel & H.J. Zwiesler** *Einsatzmöglichkeiten von Markovschen Prozessen bei der Kalkulation von Lebensversicherungsprodukten,* in Blätter der DGVFM, Band XXVI, 2003, 239-253

[1990] **H. Wolthuis & J.M. Hoem** *The retrospective premium reserve,* in Insurance: Mathematics and Economics, Nr. 9, 1990, 229-234

[1994] **H. Wolthuis** *Life Insurance Mathematics,* CAIRE, Brussels 1994

Zusammenfassung

Bei der Betrachtung von Einzelrisiken in der Personenversicherung können Verweildauereffekte auf zwei Ebenen auftreten. Einerseits gibt es Verweildauerabhängigkeiten von Übergangswahrscheinlichkeiten wie zum Beispiel die Abhängigkeit von Reaktivierungswahrscheinlichkeiten von der bisherigen Dauer einer Invalidisierung. Andererseits gibt es aber auch den Bedarf, verweildauerabhängige Versicherungsleistungen anzubieten. Derartige Abhängigkeiten lassen sich auf Basis der üblicherweise verwendeten Markovschen Sprungrozesse nicht explizit beschreiben. Mit Hilfe des vorgestellten Modells, basierend auf semi-Markov Prozessen, können Verweildauereffekte jedoch direkt modelliert werden. Der allgemein gehaltene Ansatz erlaubt eine gleichermaßene Behandlung von diskreten und kontinuierlichen Versicherungsmodellen. Beispiele aus der Berufsunfähigkeits- und der Privaten Krankenversicherung verdeutlichen anhand von realen Daten den Einfluss der Einbeziehung von verweildauerabhängigen Übergangswahrscheinlichkeiten.

Summary

In considering life insurance contracts, durational effects may appear at two levels. The first is concerned with the underlying biometrical risk, meaning that dependencies of transition probabilities on the previous duration in a certain state can be observed. An example is given by recovery rates of disabled insured which additionally depend on the time elapsed since disablement. Secondly, there is a need for duration-depending actuarial payments. Both of these durational effects cannot be explicitly implemented in the common model relying on Markovian pure jump processes. The model presented here, based on semi-Markov processes, allows one to directly model dependencies on the previous duration. Due to the generality of the approach, discrete and continuous life insurance models can be similarly discussed. Relying on real data, numerical examples dealing with disability insurance as well as German private health insurance outline the impact of using duration-depending transition rates.

Selbstständigkeitserklärung

Ich versichere hiermit an Eides statt, dass ich die vorliegende Arbeit selbstständig 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, im Dezember 2007 Marko Helwich

Lebenslauf

Marko Helwich
Hermannstr. 15
18055 Rostock
fon 0381/ 127 34 83
mobile 0163/ helwich
mail marko.helwich@uni-rostock.de

Weitere Persönliche Angaben:

Geboren	09. Dezember 1976
Staatsangehörigkeit	deutsch

Schulausbildung:

1983 - 1991	Polytechnische Oberschule Heinrich-Heine in Gadebusch
1991 - 1995	Gymnasium Gadebusch, Schulabschluss Abitur mit den Leistungskursen Mathematik und Physik

Hochschulausbildung:

1996 - 2003	Universität Rostock, Studiengang Diplom-Mathematik mit Nebenfach Philosophie Spezialisierung: Wahrscheinlichkeitstheorie und Mathematische Statistik, Finanz- und Versicherungsmathematik Diplomarbeit: *Über den Vergleich des Zinsrisikos mit dem biometrischen Risiko bei Lebensversicherungen*, ausgezeichnet mit dem GAUSS-Preis 2003 (Nachwuchspreis) der DAV/ DGVFM

Tätigkeiten:

seit 2004	Wissenschaftlicher Mitarbeiter bei Prof. Dr. Hartmut Milbrodt, Lehrstuhl für Finanz- und Versicherungsmathematik an der Universität Rostock, Zusätzlich eigenverantwortliche Lehrtätigkeiten im Bereich Wirtschaftsstatistik an der School of Management and Innovation (Steinbeis-Hochschule Berlin) und der University of Latvia, Riga

Rostock, im Dezember 2007