Markov chain backward stochastic differential equations in modeling insurance policy

Henri Hänninen

Master's thesis

University of Jyväskylä
Department of mathematics and statistics
July 28, 2022

Tiivistelmä
Henri Hänninen, *Markov chain backward stochastic differential equations in modeling insurance policy*, Jyväskylän yliopisto, Matematiikan ja tilastotieteen laitos, matematiikan pro gradu -tutkielma, 48 s., heinäkuu 2022.

Tässä tutkielmassa tarkastelemme henkivakuutuksen varantoa. Mallinnamme henkivakuutusta Markovin prosessin avulla, ja varannon määrittelyyn ja mallintamiseen käytämme Markovin ketju BSDE:itä (Markovin ketju takaperoinen stokastinen differentiaaliyhtälö). Seuraamme ensisijaisena lähteenä Boualem Djehichen ja Björn Löfdahlin artikkelia *Nonlinear reserving in life insurance: Aggregation and mean-field approximation.* Muotoilemme ja todistamme ensimmäisten lukujen väitteet, osittain eri oletuksin.

Markovin ketju BSDE:iden määrittelyä varten tarvitsemme sopivan yleistä stokastisen integroinnin ja Markovin prosessien teoriaa. Annamme tarvittavat esitiedot todennäköisyysteoriasta ja integroinnin teoriasta. Esittelemme martingaalien teoriaa, jotta voimme määritellä stokastisen integraalin semimartingaalien suhteen.

Todistamme olemassaolon ja yksikäsitteisyyden Markovin ketju BSDE:iden ratkaisulle. Todistus mukailee vastaavaa Brownin liikkeen tapausta. Tutkimme myös erityistapausta, jossa Markovin ketju BSDE:iden ensimmäisen asteen termin kerroinfunktio on deterministinen Markovin ketjun ja varannon funktio. Osoitamme, että tällöin varanto on deterministinen Markovin ketjun funktio. Todistamme, että tässä tapauksessa varanto toteuttaa epälineaarisen Thielen yhtälön.

Abstract
In this thesis we introduce Markov chain backward stochastic differential equations (BSDE), in aim to let us model insurance policies with payments dependent on the policy reserve. We prove the existence and uniqueness of a solution to the BSDEs. In the case of a deterministic driver for the BSDE, we prove that the modeled reserve is a solution to a nonlinear Thiele equation. For our main results we follow the article *Nonlinear reserving in life insurance: Aggregation and mean-field approximation* by Boualem Djehiche and Björn Löfdahl.

To define Markov chain BSDEs and prove our main results, we need suitably general theory of stochastic integration and Markov processes. After preliminary results, we define the stochastic integral with respect to semimartingales. Then we introduce Markov processes to study the model of the insurance policy.

Contents

1 Introduction

A life insurance contract specifies states (for example healthy, disabled and dead), and the payments that depend on the state and transitions between the states. We model these states and transitions using a stochastic process to account for randomness. We make the assumption that the process is Markov. A Markov process has versatility to represent the complexity of reality, but also simplicity and structure to allow computations. See [13, 7.4.A] for more discussion on the suitability of this assumption and the model.

An insurance company must provide a reserve to match liabilities. This is modeled with the prospective reserve - the expected (current) value of future payments. In this thesis we are interested in the case where we allow the payment process to be dependent on the reserve. This is a natural extension: an example of nonlinear dependence in insurance is a policy that pays guaranteed benefits based on the maximum between the accumulated reserve and a guaranteed amount.

However, when the payment process and the prospective reserve are dependent on each other, a problem of recursivity arises in the definition. This problem can be reformulated as a Markov chain backward stochastic differential equation (BSDE). Since the nature of stochastic differentials differs greatly from ordinary differentials we need theory on martingales, local martingales and quadratic variation. For this and the definition of the stochastic integral we follow [10].

Our primary source for this thesis is the article [3]. We formulate and prove most claims of the first chapters of the article. After introducing Markov processes, we construct the Markov chain model of the insurance policy in line with the article. Then we prove our main theorems - in more detail and in some cases with different assumptions.

We have two main results in this thesis. First, we prove the existence and uniqueness of the solution to the BSDEs. The proofs are adaptations of those of the Brownian motion case. Then we narrow our scope to the case where the driver of the BSDEs is a deterministic function of the prospective reserve and the Markov chain, as is natural for the application to insurance. We will show that under this restriction the prospective reserve can be represented as a deterministic function of the Markov chain. Our other goal is to prove that in this case the prospective reserve fulfills a nonlinear Thiele equation.

2 Preliminaries

We introduce the definitions and theorems that are used in this thesis. We begin from basics of stochastic processes, but assume familiarity with measure theory.

In this thesis we consider the time interval $I = [0, T]$, for some $T > 0$. We will modify some definitions accordingly and note the difference.

2.1 Notation

We list some notation and terminology used in the thesis.

- For a set A we use both $\#A$ and $|A|$ to denote its cardinality.
- A function f on a interval I is increasing, if $f(x) \geq f(y)$ for $x \geq y$, $x, y \in I$.
- A function is càdlàg if it is right-continuous and has left limits.
- We define the indicator function

$$\mathbb{1}_A(x) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A. \end{cases}$$

- We denote

$$\delta_{ij}(x) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

- The power set of a non-empty set X

$$2^X = \{A : \ A \subset X\}.$$

- We denote $\mathbb{N} = \{1, 2, 3, \dots\}$.

2.2 Stochastic Processes

Stochastic processes are used to model random development. The whole of the information in the events of the development is the filtration.

Definition 2.1 ([7, Definition 2.1.8])**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $(\mathcal{F}_t)_{t \in I}$ be a family of σ-algebras of $\mathcal{F}$. If $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ for all $s, t \in I$ such that $s \leq t$, then the family $(\mathcal{F}_t)_{t \in I}$ is a filtration.*

We call the quadruple $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \in I})$ a stochastic basis.

Definition 2.2. *A probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *is complete, if for all* $B \in \mathcal{F}$ *such that* $\mathbb{P}(B) = 0$ *and for all* $A \subset B$ *it holds that* $A \in \mathcal{F}$*.*

The following condition will be required of the stochastic basis for many results and it will be a standard assumption.

Definition 2.3 ([7, Definition 2.4.11])**.** *A stochastic basis* $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in I})$ *satisfies the usual conditions provided that the following holds:*

(i) $(\Omega, \mathcal{F}, \mathbb{P})$ *is a complete probability space,*

(ii) for all $A \in \mathcal{F}$ *such that* $\mathbb{P}(A) = 0$*, it holds that* $A \in \mathcal{F}_t$ *for all* $t \in I$*,*

(iii) the filtration is right-continuous:

$$\mathcal{F}_t = \cap_{s>t, s\in I} \mathcal{F}_s$$

for all $t \in [0, T)$*.*

Definition 2.4. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space. A map* $X : \Omega \to \mathbb{R}$ *is a random variable, if* X *is measurable from* $(\Omega, \mathcal{F})$ *to* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$*, where* $\mathcal{B}(\mathbb{R})$ *is the Borel* σ*-algebra.*

Definition 2.5 ([7, Definition 2.1.1])**.** *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space. The family* $X = (X_t)_{t\in I}$ *is a stochastic process, if* $X_t : \Omega \to \mathbb{R}^d$ *is a random variable for all* $t \in I$*.*

We defined a stochastic process as a family of random variables, however the stochastic process $X = (X_t)_{t\in I}$ also describes the functions $f_\omega = \{t \mapsto X_t(\omega) : I \to \mathbb{R}\}$ for all $\omega \in \Omega$. The function f_ω is called a path of X.

We will say that a process X is càdlàg, if all its paths are càdlàg.

There are different measurability concepts for stochastic processes. We introduce here the concepts we need: adapted, measurable, progressively measurable and predictable processes.

Definition 2.6 ([7, Definition 2.1.9])**.** *Let* $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in I})$ *be a stochastic basis and* $X = (X_t)_{t\in I}$ *be a stochastic process.*

(i) The stochastic process X *is adapted with respect to the filtration* $(\mathcal{F}_t)_{t\in I}$*, if* X_t *is* $\mathcal{F}_t$*-measurable for all* $t \in I$*.*

(ii) The stochastic process X *is measurable if it is* $\mathcal{F} \otimes \mathcal{B}(I)$*-measurable.*

(iii) The stochastic process X is progressively measurable with respect to the filtration $(\mathcal{F}_t)_{t\in I}$, if for all $S \in I$ the restriction $(\omega, t) \mapsto X_t(\omega) : \Omega \times [0, S] \to \mathbb{R}$ is $\mathcal{F}_S \otimes \mathcal{B}([0, S])$-measurable. Alternatively, we define the σ-algebra

$$\mathcal{PM}_T = \sigma(A_t \times B_t : \; A_t \in \mathcal{B}([0, t]), \; B_t \in \mathcal{F}_t, \; t \in [0, T])$$

on $[0, T] \times \Omega$. Then X is progressively measurable, if X is $\mathcal{PM}_T$-measurable.

We have the following relations between the concepts.

Theorem 2.7 ([7, Definition 2.1.10])**.** *If a stochastic process $X = (X_t)_{t\in I}$ is progressively measurable, then it is measurable and adapted.*

Theorem 2.8 ([7, Definition 2.1.11])**.** *Let $X = (X_t)_{t\in I}$ be a stochastic process. If X is adapted and has right-continuous paths, then it is progressively measurable.*

Definition 2.9 ([10, Definition 3.15])**.** *The σ-algebra $\mathcal{P}$ on $[0, T] \times \Omega$, generated by adapted processes with left-continuous paths, is called the predictable σ-algebra. We say that a process $(X_t)_{t\in[0,T]}$ is predictable, if it is $\mathcal{P}$-measurable.*

We remark that predictable processes are progressively measurable, [10, Theorem 3.11].

We have two concepts of equality for stochastic processes.

Definition 2.10 ([15, Definition p.3])**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $X = (X_t)_{t\in I}$ and $Y = (Y_t)_{t\in I}$ be stochastic processes. We say that X and Y are indistinguishable if*

$$\{\omega \in \Omega : X_t = Y_t, t \in I\}$$

is measurable and

$$\mathbb{P}(X_t = Y_t, t \in I) = 1.$$

Also, we say that X and Y are modifications of each other if

$$\mathbb{P}(X_t = Y_t) = 1.$$

for all $t \in I$.

For modifications the set of difference, N_t, is dependent on t, whereas for indistinguishable it is not. Therefore, indistinguishable processes are modifications. The converse is not true in general however, we have the next theorem.

Theorem 2.11 ([15, Theorem I.2])**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $X = (X_t)_{t\in I}$ and $Y = (Y_t)_{t\in I}$ be modifications of each other. If X and Y have right continuous paths, then X and Y are indistinguishable.*

2.3 Integration theorems

Let us now introduce definitions and theorems regarding integrals for later use. All results in this chapter are in [4].

First we have the monotone convergence theorem.

Theorem 2.12 ([4, Theorem 4.3.2])**.** *Let* $(\Omega, \mathcal{F}, \mu)$ *be a measure space. Let* $(f_n)_{n\in\mathbb{N}}$ *be a sequence of measurable functions* $f_n : \Omega \to [-\infty, \infty]$ *for all* $n \in \mathbb{N}$ *such that for* $n \in \mathbb{N}$, $\omega \in \Omega$

$$f_n(\omega) \leq f_{n+1}(\omega).$$

Assume

$$\int_\Omega f_1 \, d\mu > -\infty.$$

Then $f = \lim_{n\to\infty} f_n$ *is measurable and*

$$\lim_{n\to\infty} \int_\Omega f_n \, d\mu = \int_\Omega f \, d\mu$$

We introduce the spaces of functions we need, and then give some of the essential results.

Definition 2.13 ([4, Definition p.153])**.** *Let* $(\Omega, \mathcal{F}, \mu)$ *be a measure space and* $1 \leq p < \infty$. *We denote* $\mathcal{L}^p(\Omega, \mathcal{F}, \mu)$ *the set of all measurable* $f : \Omega \to \mathbb{R}$ *such that the* L^p*-norm is finite:*

$$||f||_p = \Big(\int_\Omega |f|^p \, d\mu \Big)^{\frac{1}{p}} < \infty.$$

A consequence of the fundamental Hölder's inequality, the Cauchy-Schwartz inequality.

Theorem 2.14 ([4, Theorem 5.1.4])**.** *Let* $(\Omega, \mathcal{F}, \mu)$ *be a measure space, and* $f, g \in \mathcal{L}^2(\Omega, \mathcal{F}, \mu)$. *Then* $fg \in \mathcal{L}^1(\Omega, \mathcal{F}, \mu)$ *and*

$$\Big| \int fg \, d\mu \Big| \leq ||f||_2 ||g||_2.$$

Next Minkowski inequality, the triangle inequality for the L^p-norm.

Theorem 2.15 ([4, Theorem 5.1.5])**.** *Let* $(\Omega, \mathcal{F}, \mu)$ *be a measure space,* $1 \leq p < \infty$ *and* $f, g \in \mathcal{L}^p(\Omega, \mathcal{F}, \mu)$. *Then* $f + g \in \mathcal{L}^p(\Omega, \mathcal{F}, \mu)$ *and*

$$||f + g||_p \leq ||f||_p + ||g||_p.$$

Now we can define the normed linear space L^p.

Definition 2.16 ([4, Definition p.158]). *Let $f, g \in \mathcal{L}^p(\Omega, \mathcal{F}, \mu)$. Define the equivalence relation $f \sim g$ if and only if $f = g$ μ-a.e. Now denote*

$$L^p(\Omega, \mathcal{F}, \mu) = \{[f] : f \in \mathcal{L}^p(\Omega, \mathcal{F}, \mu)\},$$

where

$$[f] = \{g : f \sim g\}$$

is the equivalence class of f.

Definition 2.17 ([4, Definition p.158]). *Let E be a vector space and $||\cdot||$ a norm on it. We say that the pair $(E, ||\cdot||)$ is a normed linear space. The normed linear space $(E, ||\cdot||)$ is called a Banach space, if it is complete with the metric given by the norm.*

The completeness of L^p will be used multiple times in finding solutions to differential equations.

Theorem 2.18 ([4, Theorem 5.2.1]). *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $1 \leq p < \infty$. Then $(L^p(\Omega, \mathcal{F}, \mu), ||\cdot||_p)$ is a Banach space.*

We give two version of Jensen's inequality in the form we need. As the (Lebesgue) measure space $([0, T], \mathcal{B}([0, T]), \lambda)$ can be normalized to be a probability space we first give Jensen's inequality as:

Theorem 2.19 ([4, Theorem 10.2.6]). *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space such that $\mu(\Omega) < \infty$. Let $f \in \mathcal{L}^2(\Omega, \mathcal{F}, \mu)$. Then*

$$\Big|\int_\Omega f \, d\mu\Big|^2 \leq \mu(\Omega) \int_\Omega |f|^2 \, d\mu.$$

Another version is the conditional Jensen's inequality.

Theorem 2.20 ([4, Theorem 10.2.7]). *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $\mathcal{G}$ be a sub-σ-algebra, and $X \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$. Then*

$$\mathbb{E}[X|\mathcal{G}]^2 \leq \mathbb{E}[|X|^2|\mathcal{G}] \;\; a.s.$$

We can interchange integral and conditional expectation.

Theorem 2.21 ([9]). *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and $\mathcal{H} \subset \mathcal{F}$ a sub-σ-algebra. Let $X \in \mathcal{L}^1([0, T] \times \Omega, \mathcal{B}([0, T]) \otimes \mathcal{F}, \lambda \otimes \mathbb{P})$. Then for $A \in \mathcal{B}([0, T])$*

$$\mathbb{E}[\int_A X(s) \, ds|\mathcal{H}] = \int_A \mathbb{E}[X(s)|\mathcal{H}] \, ds \;\; a.s.$$

2.4 Lebesgue-Stieltjes Integral

We give a primer on Lebesgue-Stieltjes integration. We will later use Lebesgue-Stieltjes integration in defining the stochastic integral. Also for many of our discussed processes later the stochastic integrals and Lebesgue-Stieltjes integrals agree.

First we define the total variation of a function.

Definition 2.22 ([12, Definition 15.8])**.** *Let $f : [0, T] \to \mathbb{R}$ be càdlàg. Then*

$$\int_{[0,t]} |\, df| = \sup_{0=t_0 \leq \ldots \leq t_n = t} \sum_{k=1}^{n} |f(t_k) - f(t_{k-1})|$$

is the total variation of f on $[0, t]$.

We will often handle functions of bounded variation.

Definition 2.23. *Let $f : [0, T] \to \mathbb{R}$ be càdlàg. Then f is of bounded variation (on $[0, T]$), if $\int_{[0,T]} |\, df| < \infty$.*

Functions of bounded variation can be decomposed.

Theorem 2.24 ([12, Theorem 15.11])**.** *Let $f : [0, T] \to \mathbb{R}$ be of bounded variation. Then there is a pair of functions $f_1, f_2 : [0, T] \to \mathbb{R}$ such that*

(i) f_1, f_2 are non-decreasing,

(ii) $f = f_1 - f_2$.

For formulating the definition of Lebesgue-Stieltjes integration we need the following.

Theorem 2.25 ([5, Proposition 1.2.17])**.** *(Carathéodory) Let Ω be non-empty and $\mathcal{G}$ be an algebra such that $\mathcal{F} = \sigma(\mathcal{G})$. Assume $\mu_0 : \mathcal{G} \to [0, \infty)$ such that*

(i) $\mu_0(\Omega) < \infty$.

(ii) For every sequence of pair-wise disjoint $A_n \in \mathcal{G}$, such that $\cup_{n \in \mathbb{N}} A_n \in \mathcal{G}$, it holds that

$$\mu_0(\cup_{n \in \mathbb{N}} A_n) = \sum_{n \in \mathbb{N}} \mu_0(A_n)$$

Now there exists a unique measure $\mu : \mathcal{F} \to [0, \infty)$ such that

$$\mu(A) = \mu_0(A)$$

for $A \in \mathcal{G}$.

For defining the Lebesgue-Stieltjes integral, first assume that the càdlàg $f : [0,T] \to \mathbb{R}$ is non-decreasing. Define $\mu_0((s,t]) = f(t) - f(s)$ and $\mu_0(\{t\}) = f(t) - f(t-)$. As the half open intervals generate $\mathcal{B}([0,T])$, we have by Theorem 2.25 that there exists a unique measure $\mu_f : \mathcal{B}([0,T]) \to [0,\infty)$ such that

$$\mu_f((s,t]) = \mu_0((s,t])$$

for $(s,t] \subset [0,T]$.

Definition 2.26. *Let $f : [0,T] \to \mathbb{R}$ be of bounded variation and $H : [0,T] \to \mathbb{R}$ be Borel-measurable. If*

$$\int_{[0,t]} |H(s)||\,df(s)| = \int_{[0,t]} |H(x)|\,d\mu_{f_1}(x) + \int_{[0,t]} |H(x)|\,d\mu_{f_2}(x) < \infty,$$

then we define the Lebesgue-Stieltjes integral with respect to f as

$$\int_{[0,t]} H(x)\,df(x) = \int_{[0,t]} H(x)\,d\mu_{f_1}(x) - \int_{[0,t]} H(x)\,d\mu_{f_2}(x),$$

where f_1, f_2 are given by Theorem 2.24, and the integrals are Lebesgue integrals.

Now we can define the Stieltjes-Lebesgue integral w.r.t. processes with bounded variation.

Definition 2.27. *Let the stochastic process $H : [0,T] \times \Omega \to \mathbb{R}$ be jointly measurable, and let the stochastic process X have paths of bounded variation. We define ω-by-ω, $X(s,\omega) = f_\omega(s)$. If*

$$\int_{[0,t]} |H(s)||\,dX(s)| = \int_{[0,t]} |H(s,\omega)||\,df_\omega(s)| < \infty,$$

then we define the Lebesgue-Stieltjes integral as the random variable

$$\int_{[0,t]} H(s)\,dX(s)(\omega) = \int_{[0,t]} H(s,\omega)\,df_\omega(s).$$

We also very naturally define the following.

Definition 2.28 ([10, Definition 15.8])**.** *Let X be a process with paths of bounded variation. If*

$$\int_{[0,T]} |\,dX(s)|$$

is an integrable random variable, we say that X is a process with integrable variation. We denote $\mathcal{A}$ the set of adapted processes of integrable variation, and $\mathcal{A}^+$ the set of adapted integrable processes with increasing paths.

3 Stochastic Integration

Stochastic processes may have unbounded variation and be nowhere differentiable. Due to this nature stochastic differential equations are defined via integral equations. We follow [10] in defining the stochastic integral. Before the main definitions we need theory on martingales, local martingales and quadratic variation. Since we consider the finite time interval $I = [0, T]$, our definitions differ accordingly from those in the literature. We assume a stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in I})$ satisfying the usual conditions for this chapter.

3.1 Martingales

Martingales have a fundamental role in the theory of stochastic processes and stochastic integration. We will go over the needed definitions and theorems.

Definition 3.1 ([11, Definition 7.1])**.** *Let $X = (X_t)_{t\in I}$ be a $(\mathcal{F}_t)_{t\in I}$-adapted stochastic process such that $\mathbb{E}|X_t| < \infty$ for all $t \in I$. Then X is a martingale if for $s \leq t \in I$*

$$\mathbb{E}[X_t|\mathcal{F}_s] = X_s \text{ a.s.}$$

Martingales are special in that they have a constant expectation.

It is important to have martingales with càdlàg paths. We have the following existence result.

Theorem 3.2 ([15, Theorem I.9])**.** *Let $X = (X)_{t\in I}$ be a martingale. Then there exists a unique modification Y of X which is càdlàg. Specifically, a martingale with right continuous paths is càdlàg.*

We give shorthands for our most used classes of martingales.

Definition 3.3. *A martingale X such that $\mathbb{E}[X_t^2] < \infty$, for $t \in [0, T]$, is called a square integrable martingale.*

We denote by $\mathcal{M}$ the space of càdlàg martingales and by $\mathcal{M}^2$ the space of square integrable càdlàg martingales.

Our definition differs from [10] since we consider martingales on $[0, T]$. In finite time martingales are uniformly integrable ([10, Theorem 1.8]), and $\sup_{t\in[0,T]} \mathbb{E}[M_t^2] = \mathbb{E}[M_T^2] < \infty$.

For square integrable martingales we have Doob's maximal quadratic inequality.

Theorem 3.4 ([15, Theorem I.20]). *Let X be a càdlàg martingale. Then we have*

$$\mathbb{E}[\sup_{0\le s\le t} X_s^2] \le 4\mathbb{E}[X_t^2]$$

for $t \in [0, T]$.

To begin the study of stochastic integration, we need key results for martingales with integrable variation. But before that we define, for a process X,

$$X_{t-} = X(t-) = \lim_{s\uparrow t} X_s, \ \ X_{0-} = X_0,$$

where the limit is the left limit. Also, define the jump process

$$\Delta X_t = X_t - X_{t-}.$$

Theorem 3.5 ([10, Theorem 6.4]). *Let $M \in \mathcal{M}$ be bounded with integrable variation. Then*

$$M_t^2 - \sum_{s\le t} (\Delta M)^2 \in \mathcal{M}.$$

The class of predictable processes has a central role as the integrands. To this end, the following result says that the Lebesgue-Stieltjes integral of predictable processes preserves the property of being a càdlàg martingale with integrable variation.

Theorem 3.6 ([10, Theorem 6.5]). *Let $M \in \mathcal{M}$ have integrable variation. Let H be a predictable process such that*

$$\mathbb{E}[\int_{[0,T]} |H_s| |\, dM_s|] < \infty.$$

Then for the Lebesgue-Stieltjes integral $H \cdot M$, we have $H \cdot M \in \mathcal{M}$ with integrable variation.

Theorem 3.7 ([10, Theorem 6.8]). *The space $\mathcal{M}^2$, equipped with the inner product $(M, N) = \mathbb{E}[M_T N_T]$, is a Hilbert space.*

We also have isomorphism to $L(\Omega, \mathcal{F}_T, \mathbb{P})$ via the map $M \mapsto M_T$.

The fact that $\mathcal{M}^2$ is a Hilbert space allows the usage of the Riesz representation theorem in the proof of Theorem 3.23 to determine the existence of the stochastic integral under certain assumptions. It also, together with [10, Corollary 6.17], lets us define the following.

Definition 3.8 ([10, Definition 6.18])**.** *Let $\mathcal{M}^{2,c}$ be the set of all continuous $M \in \mathcal{M}^2$. Denote $\mathcal{M}^{2,d} = (\mathcal{M}^{2,c})^{\perp}$. Let $M \in \mathcal{M}^2$. Then M has a unique decomposition:*

$$M = M(0) + M^c + M^d,$$

where $M^c \in \mathcal{M}_0^{2,c}$ is the continuous martingale part and $M^d \in \mathcal{M}^{2,d}$ is the purely discontinuous martingale part.

To determine purely discontinuous martingales we have the following.

Theorem 3.9 ([10, Theorem 6.22.1])**.** *Let $M \in \mathcal{M}_0^2$. Then*

$$\mathbb{E}[\sum_{0 \le s \le T} (\Delta M_s)^2] = \mathbb{E}[M_T^2]$$

if and only if $M \in \mathcal{M}^{2,d}$.

3.2 Local martingales

The general theory of stochastic integration needs local martingales. First, we have the concepts of a stopping time and localization.

Definition 3.10 ([15, Definition I.2])**.** *Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \in I})$ be a stochastic basis. A random variable $\tau : \Omega \to [0, \infty)$ is stopping time if*

$$\{\tau \le t\} \in \mathcal{F}_t$$

for all $t \in I$.

Let τ be a stopping time and X a process. We denote X^τ the stopped process defined by $X_t^\tau(\omega) = X_{min\{t,\tau(\omega)\}}(\omega)$.

Definition 3.11 ([10, Definition 7.1])**.** *Let $\mathcal{D}$ be a class of processes. A process X is an element of the localized class $\mathcal{D}_{loc}$ if and only if $X_0 \in \mathcal{F}_0$ and there exists a sequence (T_n) of stopping times such that $T_n \uparrow \infty$ and for each n the stopped process $X^{T_n} - X_0 \in \mathcal{D}$.*

Definition 3.12 ([10, Definition 7.11])**.** *We say that a process $M \in \mathcal{M}_{loc}$ is a local martingale. Other localized classes of processes we consider are:*

- *$\mathcal{A}_{loc}$, adapted processes with locally integrable variation, Definition 2.28,*
- *$\mathcal{A}_{loc}^+$, adapted locally integrable increasing processes, Definition 2.28,*
- *$\mathcal{M}_{loc}^2$, locally square integrable martingales,*

- $\mathcal{M}^{2,c}_{loc}$, *continuous locally square integrable martingales,*

- $\mathcal{M}^{2,d}_{loc}$, *purely discontinuous locally square integrable martingales, Definition 3.8.*

We remark that a local martingale is an adapted càdlàg process and a càdlàg martingale is a local martingale.

We need results showing that specific local martingales are martingales. Indeed, local martingales bounded by integrable random variables are martingales:

Theorem 3.13 ([11, Theorem 7.21]). *Let X be a local martingale. If there exists a random variable Y such that $\mathbb{E}[Y] < \infty$ and $|X_t| \leq Y$ a.s. for all $t \in I$, then X is a martingale.*

Definition 3.14 ([10, Definition 7.21]). *Let M be a local martingale. If $M_0 = 0$ and M has a decomposition as follows:*

$$M = U + V$$

where $U \in \mathcal{M}^{2,d}_{loc}$ and $V \in \mathcal{A}_{loc}$ is a local martingale, we say that M purely discontinuous, and denote $\mathcal{M}^{d}_{loc}$.

We also have the localized class of continuous martingales $\mathcal{M}^{c}_{loc}$.

We omit much of the theory developed in Chapters 6 and 7 in [10] that lead to the previous definition and the following result. The following decomposition relies on that of square integrable martingales, and the results on the nature of the localization. For example, we note that $\mathcal{M}^{c}_{loc} = \mathcal{M}^{2,c}_{loc}$, since we can localize a $M \in \mathcal{M}^{c}_{loc}$ by stopping at bounds,

$$T_n = \inf\{t \geq 0 : \ |M_t| \geq n\}, \ n \geq 1,$$

to get square integrability.

Theorem 3.15 ([10, Theorem 7.25]). *A local martingale M has the unique decomposition:*

$$M = M_0 + M^c + M^d$$

where $M^c \in \mathcal{M}^{c}_{loc,0}$ is the continuous martingale part and $M^d \in \mathcal{M}^{d}_{loc}$ the purely discontinuous martingale part of M respectively.

3.3 Quadratic variation

The quadratic variation plays a key role in the difference between stochastic calculus and standard calculus. It shows up in integration by parts, Itô's formula and the Burkholder-Davis-Gundy inequality. To define it for local martingales we need the following two results.

Theorem 3.16 ([10, Lemma 7.27]). *For a local martingale M,*

$$\sum_{s \leq T} (\Delta M_s)^2 < \infty \text{ a.s.}$$

The following result uses the Doob-Meyer decomposition theorem, [10, Theorem 5.44], to show the existence:

Theorem 3.17 ([10, Lemma 7.28]). *If $M \in \mathcal{M}^2_{loc}$, then there exists a unique predictable locally integrable increasing process, $\langle M \rangle \in \mathcal{A}^+_{loc}$, such that $M^2 - \langle M \rangle \in \mathcal{M}^2_{loc,0}$.*

Define $\langle M, N \rangle = \frac{1}{2}(\langle M + N \rangle - \langle M \rangle - \langle N \rangle)$. Now we can define the quadratic (co)variation.

Definition 3.18 ([10, Definition 7.29]). *Let M and N be local martingales. We define*

$$[M, N] = M_0 N_0 + \langle M^c, N^c \rangle + \sum_{s \leq \cdot} \Delta M_s \Delta N_s.$$

The quadratic variation has the following properties, which we will use without explicit mention. For local martingales M and N:

- the quadratic variation $[M, N]$ is symmetric and bilinear,
- we have the polarization identity $[M, N] = \frac{1}{2}([M+N, M+N] - [M, M] - [N, N])$.

Quadratic variation has the fundamental property:

Theorem 3.19 ([10, Theorem 7.31]). *If M and N are local martingales, then $[M, N]$ is the unique adapted process of bounded variation such that $MN - [M, N] \in \mathcal{M}_{loc,0}$ and $\Delta[M, N] = \Delta M \Delta N$.*

We can use the finiteness of the quadratic variation to determine whether a local martingale is a square integrable martingale.

Theorem 3.20 ([10, Theorem 7.32]). *For a local martingale M, $M \in \mathcal{M}^2$ if and only if $\mathbb{E}[M]_T < \infty$.*

Last, we give the following definition and result:

$$\mathcal{F}_\tau = \{A \in \mathcal{F}_T : \text{ for every } t \in [0,T],\ A \cap \{\tau \leq t\} \in \mathcal{F}_t\}.$$

Theorem 3.21 ([10, Theorem 7.38]). *Let M be a local martingale, τ a stopping time and ξ be a $\mathcal{F}_\tau$-measurable random variable. Now $N = \xi(M - M^\tau)$ is a local martingale. We also have*

$$[N, L] = \xi([M, L] - [M, L]^\tau),$$

for every local martingale L.

3.4 The Stochastic Integral

Now we define the stochastic integral and state its basic properties. The approach is to define the stochastic integral for elementary processes and observe that it has a characterizing property. We then define the stochastic integral as the process satisfying this property.

Let S, τ be stopping times such that $S \leq \tau \leq T$ and ξ a $\mathcal{F}_S$-measurable random variable. Then $H = \xi I_{(S,\tau]}$ is a predictable process, [10, Theorem 3.16.2]. We call a process of this form elementary predictable.

Let M be a local martingale. Define the stochastic integral $H \cdot M$ of elementary predictable H w.r.t. M as the process:

$$(H \cdot M)_t = \xi(M_{min\{t,\tau\}} - M_{min\{t,S\}}),\ t \in [0,T].$$

By Theorem 3.21 and the above definition w.r.t $[M, N]$, $H \cdot M$ is a local martingale with the following property:

$$[H \cdot M, N] = \xi([M, N]^\tau - [M, N]^S) = H \cdot [M, N],$$

for every local martingale N, where $H \cdot [M, N]$ is a Lebesgue-Stieltjes integral. The critical observation is that by Theorem 3.19 the integral $H \cdot M$ is the unique local martingale with this, now characteristic, property.

We now define the stochastic integral via this property.

Definition 3.22 ([10, Definition 9.1]). *Let M be a local martingale and H a predictable process. If H is Lebesgue-Stieltjes integrable w.r.t. $[M, N]$, for all $N \in \mathcal{M}_{loc}$, and there exists a local martingale L such that*

$$[L, N] = H \cdot [M, N], \tag{1}$$

for all $N \in \mathcal{M}_{loc}$, then H is integrable w.r.t. M. We define the unique (by Theorem 3.19) local martingale $L = H \cdot M$ the stochastic integral of H w.r.t. M. We denote $L(M)$ the set of predictable processes integrable w.r.t. M.

To see the uniqueness in the definition, let $L, L' \in \mathcal{M}_{loc}$ such that for any $N \in \mathcal{M}_{loc}$, $[L, N] = H \cdot [M, N]$ and $[L', N] = H \cdot [M, N]$. Then we have $[(L - L'), N] = 0$, and by Theorem 3.19

$$(L - L')N = (L - L')N - [(L - L'), N] \in \mathcal{M}_{loc,0}.$$

With $N = (L - L') \in \mathcal{M}_{loc}$, we have $(L - L')^2 \in \mathcal{M}_{loc,0}$, which is satisfied only for $L = L'$.

The definition of the stochastic integral does not tell us when the integral exists or how to begin computing one. For the former we have the following characterization of $L(M)$.

Theorem 3.23 ([10, Theorem 9.2])**.** *A predictable process H is integrable w.r.t. a local martingale M, $H \in L(M)$, if and only if $(H^2 \cdot [M])^{\frac{1}{2}} \in \mathcal{A}^+_{loc}$.*

We note for future usage that when $\mathbb{E}[(H^2 \cdot [M]_T)^{\frac{1}{2}}] < \infty$, the above condition is satisfied.

To help us compute stochastic integrals we state the following fundamental properties. We may later use them liberally without referencing.

Theorem 3.24 ([10, Theorem 9.3])**.** *Let M be a local martingale and $H, K \in L(M)$.*

(i) $(H \cdot M)^d = H \cdot M^d$, $(H \cdot M)^c = H \cdot M^c$ *and* $(H \cdot M)_0 = H_0 M_0$.

(ii) $\Delta(H \cdot M) = H \Delta M$.

(iii) $(H + K) \in L(M)$ *and* $(H + K) \cdot M = H \cdot M + K \cdot M$.

(iv) Let G be a predictable process. Then $G \in L(H \cdot M)$ if and only if $(HG) \in L(M)$. In this case, we have

$$G \cdot (H \cdot M) = (GH) \cdot M.$$

We denote

$$(H \cdot M)_t = \int_{[0,t]} H(s)\, dM(s).$$

And when $M_0 = 0$, we have

$$\int_{[0,t]} H(s)\, dM(s) = \int_{(0,t]} H(s)\, dM(s).$$

Under certain restrictions on the variation the stochastic integral agrees with the Lebesgue-Stieltjes integral.

Theorem 3.25 ([10, Theorem 9.5])**.** *Let $M \in \mathcal{A}_{loc}$ be a local martingale and H a predictable process. If $\sum |H\Delta M| \in \mathcal{A}^+_{loc}$, then $H \in L(M)$ and $H \cdot M$ agrees with the Lebesgue-Stieltjes integral.*

Let us extend the stochastic integral to semimartingales.

Definition 3.26 ([10, Definition 8.1])**.** *We say that the process X is a semimartingale if it has a decomposition:*

$$X = M + A,$$

where M is a local martingale and A is an adapted process of bounded variation.

Let X be a semimartingale and $X = M + A$ a decomposition as above. Let us further decompose $X = M^c + M^d + A$ by Theorem 3.15. By [10, Theorem 7.19.] a local martingale of bounded variation is purely discontinuous. Therefore, M^c is uniquely determined by X, and we define $X^c = M^c$ the continuous part of X.

We will define the stochastic integral w.r.t. a semimartingale as the sum of integrals w.r.t. the parts of the decomposition. This requires that the sum of the integrals is independent of the decomposition of the semimartingale.

Theorem 3.27 ([10, Lemma 9.12])**.** *Let X be a semimartingale and H a predictable process. Let $X = M + A$ and $X = N + B$ be decompositions of X, where $M, N \in \mathcal{M}_{loc}$ and A, B adapted of bounded variation. If $H \in L(M) \cap L(N)$ and the Lebesgue-Stieltjes integrals $H \cdot A, H \cdot B$ exist, then*

$$H \cdot M + H \cdot A = H \cdot N + H \cdot B.$$

Definition 3.28 ([10, Definition 9.13])**.** *Let X be a semimartingale and H a predictable process. Suppose $X = M + A$ is a decomposition of X, where $M \in \mathcal{M}_{loc}$ and A adapted and of bounded variation, such that $H \in L(M)$ and the Lebesgue-Stieltjes integral $H \cdot A$ exists. Then H is integrable w.r.t. X and*

$$H \cdot X = H \cdot M + H \cdot A.$$

3.5 Itô's formula and BDG

We present two important results in this section. First we have Itô's formula, which is a change of variables formula for stochastic integrals. We will use it later to compute integrals. We present it in the multidimensional case for semimartingales.

Theorem 3.29 ([10, Theorem 9.35]). *Let $X = (X^1, \ldots, X^n)$ be an n-tuple of semimartingales and $f : \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable. Then $f(X)$ is a semimartingale and*

$$\begin{aligned} f(X_T) - f(X_t) = &\sum_{i=1}^{n} \int_{(t,T]} \frac{\partial}{\partial x_i} f(X_{s-})\, dX_s^i \\ &+ \sum_{1 \le i,j \le n} \frac{1}{2} \int_{(t,T]} \frac{\partial^2}{\partial x_i \partial x_j} f(X_{s-})\, d\langle (X^i)^c, (X^j)^c \rangle_s \\ &+ \sum_{t < s \le T} (\Delta f(X_s) - \frac{\partial}{\partial x_i} f(X_{s-}) \Delta X_s^i). \end{aligned}$$

The next result is the Burkholder-Davis-Gundy inequality. It lets us bound the maximum by the quadratic variation, linking them. Indeed, we will later use this to show that a specific integral is a martingale.

Theorem 3.30 (Burkholder-Davis-Gundy inequality, [10, Theorem 10.36]). *Let X be a local martingale. Let $p \ge 1$. There exists $c_p, C_p > 0$ such that for any X*

$$c_p \mathbb{E}([X]_t^{\frac{p}{2}}) \le \mathbb{E}(\sup_{0 \le s \le t} |X_s|^p) \le C_p \mathbb{E}([X]_t^{\frac{p}{2}})$$

for all $t \in [0, T]$.

4 The Markov chain model

4.1 Markov processes

Since the processs modeling the insurance policy is driven by a Markov chain, we state the needed definitions and results in this section. We begin by following [8]. Then, as the Markov chains we use are allowed to be inhomogenous, we will state the fundamental Dynkin's formula in the required generality. For this we follow [2], as well as for the definitions of the Feller evolution and infinitesimal generator.

For this section let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space and $(E, \mathcal{E})$ be a measure space, where E is a complete separable metric space.

Definition 4.1 ([8, Definition 2.1])**.** *Let X be a stochastic process. If X is adapted to the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ and if for all $t \in [0, T]$*

$$\mathbb{P}(A \cap B|X_t) = \mathbb{P}(A|X_t)\mathbb{P}(B|X_t)$$

almost surely for $A \in \mathcal{F}_t$ and $B \in \sigma(X_s; s \geq t)$, then X is a Markov process. The conditional probability is given by $\mathbb{P}(C|X_t) := \mathbb{P}(C|\sigma(X_t)) = \mathbb{E}[\mathbb{1}_C|\sigma(Xt)]$.

A Markov process has the following characteristic properties.

Theorem 4.2 ([8, Theorem 2.3])**.** *Let X be a stochastic process adapted to the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$. Then the following assertions are equivalent:*

(i) X is a Markov process.

(ii) For $s \in I$ and bounded $\sigma(X_t; T \geq t \geq s)$-measurable Y it holds

$$\mathbb{E}[Y|\mathcal{F}_s] = \mathbb{E}[Y|\sigma(X_s)] \text{ a.s.}$$

(iii) For $0 \leq s \leq t \leq T$ and all bounded $f : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$

$$\mathbb{E}[f(X_t)|\mathcal{F}_s] = \mathbb{E}[f(X_t)|\sigma(X_s)] \text{ a.s.}$$

To define our model we need the transition function.

Definition 4.3 ([8, Definition 3.1])**.** *We say that the map $(s, t, x, A) \mapsto P(s, t, x, A)$, where $0 \leq s \leq t \leq T, x \in E, A \in \mathcal{E}$ is a Markov transition function if:*

(i) $A \mapsto P(s, t, x, A)$ is a probability measure on $(E, \mathcal{E})$ for $0 \leq s \leq t \leq T, x \in E$,

(ii) $x \mapsto P(s,t,x,A)$ *is* $\mathcal{E}$*-measurable for* $0 \le s \le t \le T, A \in \mathcal{E}$,

(iii) $P(t,t,x,A) = \delta_x(A)$ *for* $t \in [0,T], x \in E, A \in \mathcal{E}$,

(iv) *(Chapman-Kolmogorov)* $P(s,u,x,A) = \int_E P(t,u,y,A)P(s,t,x,dy)$ *for* $0 \le s \le t \le u \le T, x \in E, A \in \mathcal{E}$.

Now, let $(x,A) \mapsto P(s,t,x,A)$*, where* $0 \le s \le t \le T, x \in E, A \in \mathcal{E}$*, be a Markov transition function. A stochastic process* X *adapted to* $\mathbb{F}$ *is a Markov process (with respect to* $\mathbb{F}$*) having* $P(t,s,x,A)$ *as a transition function if*

$$\mathbb{E}[f(X_t)|\mathcal{F}_s] = \int_E f(y)P(s,t,X_s,dy) \text{ a.s.}$$

for $0 \le s \le t \le T$ *and bounded* $f : (E,\mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

A probability measure μ *on* $(E,\mathcal{E})$ *is called the initial distribution of* X *if* $\mu(A) = \mathbb{P}(X(0) \in A)$.

To continue, we briefly note the following.

Definition 4.4. *We denote* $C(E)$ *the space of continuous functions* $f : E \to \mathbb{R}$ *with the uniform norm* $||f|| = \sup_{x\in E} |f(x)|$.

Compared with [2, Definition 2.4] our definition of Feller evolution has stricter assumptions and simpler form. This is done according to [2, Remark 2.5] and [2, Proposition 2.6].

Definition 4.5 ([2, Definition 2.4])**.** *A family* $\{P(s,t) : 0 \le s \le t \le T\}$ *of bounded linear operators* $P(s,t) : C(E) \to C(E)$ *is called a Feller evolution on* $C(E)$ *if the following conditions hold:*

(i) $P(t,t) = Id$ *for* $t \in [0,T]$,

(ii) $P(\tau,t) = P(\tau,s) \circ P(s,t)$ *for all* $0 \le \tau \le s \le t \le T$,

(iii) $0 \le P(s,t)f \le 1$ *for* $f \in C(E)$ *with* $f(E) \subset [0,1]$,

(iv) *For every* $f \in C(E)$ *and* $t \in [0,T]$ *the function* $(s,x) \mapsto P(s,t)f(x)$ *is continuous,*

(v) $\lim_{t\downarrow s, y\to x} P(s,t)f(y) = f(x)$.

Definition 4.6 ([2, Definition 2.7])**.** *Let us define*

$$G(s)f = \lim_{t\downarrow s} \frac{P(s,t)f - f}{t-s}$$

for every $s \in [0, T]$ *and* $f \in C(E)$ *such that the limit, say* g*, exists in* $C(E)$*:*

$$||\frac{P(s,t)f - f}{t - s} - g|| \to 0,$$

as $t \downarrow s$*.*

The limit may not always exist, so we define the domain $D(G(s)) = \{f \in C(E) : G(s)f \text{ exists}\}$*,* $s \in [0, T]$*, giving us the operator* $G(s) : D(G(s)) \to C(E)$*.*

Then the family of operators $G(s)$*,* $s \in [0, T]$*, is said to be the (infinitesimal) generator of the Feller evolution* $\{P(s, t) : 0 \le s \le t \le T\}$*. We also write* $G(s)f(s, x)$ *for* $(s, x) \mapsto G(s)f(s, \cdot)(x)$*.*

A Feller evolution gives rise to a Markov process with desirable properties.

Theorem 4.7 ([2, Theorem 2.9])**.** *Let* E *be a complete metric space,* $\{P(s, t) : 0 \le s \le t \le T\}$ *be a Feller evolution on* $C(E)$*, and* μ *an initial distribution on* $(E, \mathcal{E})$*. Now there exists a complete probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and a Markov process* $(X_t)_{t\in[0,T]}$ *w.r.t* $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ *such that*

- $P(s, t)f(X(s)) = \mathbb{E}[f(X(t))|X(s)]$ *for* $0 \le s \le t \le T$ *and bounded* $f \in C(E)$*,*
- $\mathbb{F}$ *is the completion of the natural filtration of* X*,* $\mathcal{F}_t = \sigma(\{X(s) : 0 \le s \le t\} \cup \{A \in \mathcal{F} : \mathbb{P}(A) = 0\})$
- $\mathbb{F}$ *is right-continuous,*
- X *is càdlàg.*

Next we have Dynkin's formula, which plays a key role in identifying martingales for the rest of the thesis.

Theorem 4.8 ([2, Theorem 2.11])**.** *Let* $\{P(s, t) : 0 \le s \le t \le T\}$ *be a Feller evolution on* $C(E)$*, and* μ *an initial distribution on* $(E, \mathcal{E})$*. Let* $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ *be the stochastic basis and* X *the Markov process given by Theorem 4.7.*

Let the family of operators $(G(s))_{s\in[0,T]}$ *be the generator of the Feller evolution* $\{P(s, t) : 0 \le s \le t \le T\}$*. Let* $f \in C([0, T] \times E)$ *be such that* $f(s, \cdot) \in D(G(s))$ *for* $s \in [0, T]$*,* $(s, x) \mapsto G(s)f(s, x)$ *is continuous, and* $s \mapsto f(s, k)$ *is continuously differentiable. Now the process*

$$M_t = f(t, X_t) - f(0, X_0) - \int_0^t G(s)f(s, X_s) + \frac{\partial}{\partial s} f(s, X_s)\, ds$$

is an $\mathbb{F}$*-martingale.*

4.2 The model

In this section we introduce the Markov chain model and study its properties. Then we define the payment process and the prospective reserve. We follow [3], but give proofs.

A probability transition function specifies the Markov process and the probability space. So, for the model we assume the probability transition function and transition intensities as given. We make smoothness assumptions that might not exactly reflect reality, but let us use the transition intensities. The transition intensities help us in computations and in interpreting the Markov process.

For the model we consider finite state space $E = \mathcal{S} = \{1, \ldots, S\}$, so that $|\mathcal{S}| = S$. Let $p_{ij}(s,t)$, $0 \leq s \leq t \leq T$, $i, j \in \mathcal{S}$, be probability transition functions such that the following conditions hold:

- $s \mapsto p_{ij}(s,t)$ for $s \in [0,t]$, is continuously differentiable for all $i, j \in \mathcal{S}$,
- the transition intensities μ_{ij} defined by

  $$\mu_{ij}(s) = \lim_{t \downarrow s} \frac{p_{ij}(s,t) - \delta_{ij}(s)}{t - s}$$

  for $s \in [0,T]$, $i, j \in \mathcal{S}$ are continuous.

Since $\mathcal{S}$ is finite we can define a family of linear operators, $\{P(s,t) : 0 \leq s \leq t \leq T\}$, by

$$P(s,t)f(i) = \int_{\mathcal{S}} f(j)p(s,t,i,dj) = \sum_{j \in \mathcal{S}} f(j)p_{ij}(s,t).$$

In matrix form

$$P(s,t)f = [p_{ij}(s,t)]_{i,j \in \mathcal{S}}(f(j))_{j \in \mathcal{S}}.$$

We will show that $P(s,t)$ defines a Feller evolution. The condition (i) follows from the definition of transition probability functions. In matrix form the Chapman-Kolmogorov equation gives the evolution condition, (ii),

$$p_{ij}(t,\tau) = \sum_k p_{ik}(t,s)p_{kj}(s,\tau),$$

entry-by-entry. The calculation

$$P(s,t)f(i) = \sum_{j \in \mathcal{S}} f(j)p_{ij}(s,t) \leq \sum_{j \in \mathcal{S}} p_{ii}(s,t) = 1.$$

gives us (iii). (iv) follows from the continuity of $s \mapsto p_{ij}(s,t)$. Last, from the existence of μ_{ij} we get that $\lim_{t\downarrow s} p_{ij}(s,t) = \delta_{ij}$ and (v) follows.

We have that $\{P(s,t) : 0 \le s \le t \le T\}$ is a Feller evolution. Then by Theorem 4.7 we have a stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ satisfying the usual conditions, and a càdlàg Markov process X having $p_{ij}(s,t)$ as transition functions such that $\mathbb{F}$ is the completion of the natural filtration of X.

We show that the probability transition function satisfies the Kolmogorov backward equation. Since

$$\sum_j \mu_{ij}(t) = \sum_j \lim_{t\downarrow s} \frac{p_{ij}(s,t) - \delta_{ij}(s)}{t-s} = 0$$

we get the identity

$$G(t)f(t,i) = \sum_j \mu_{ij}(t) f(t,j) = \sum_{j,j\neq i} \mu_{ij}(t)(f(t,j) - f(t,i)). \tag{2}$$

Using the evolution property (ii) we have

$$\frac{P(t,\tau) - P(s,\tau)}{t-s} = \frac{P(t,t) - P(s,t)}{t-s} P(t,\tau),$$

where letting $t \downarrow s$, we have

$$\frac{\partial P}{\partial s}(s,t) = -G(s)P(s,t),$$

which, entry-by-entry, is the Kolmogorov backward equation

$$\frac{\partial p_{ij}}{\partial s}(s,t) + \sum_k \mu_{ik}(s) p_{kj}(s,t) = 0. \tag{3}$$

We need processes expressing the stay at states and those counting the jumps between states. For the former denote $\mathbb{1}_{\{X(t)=i\}} = I_i(t)$. And for the latter define the counting processes,

$$N_{ij}(t) = \#\{s \in (0,t] : X(s-) = i, X(s) = j\},\ N_{ij}(0) = 0, \text{ for } i \neq j.$$

We note that both of these processes are càdlàg.

Theorem 4.9. *For $i \neq j$ the compensated process,*

$$M_{ij}(t) = N_{ij}(t) - \int_{(0,t]} I_i(s-)\mu_{ij}(s)\, ds,\ M_{ij}(0) = 0,$$

is a square integrable càdlàg martingale. We also have for the quadratic variation

$$[M_{ij}](t) = N_{ij}(t),$$

and for the predictable quadratic variation

$$\langle M_{ij}\rangle(t) = \int_{(0,t]} I_i(s-)\mu_{ij}(s)\, ds,$$

Proof. By Theorem 4.8 for $f(t,k) = \mathbb{1}_{\{k=j\}}$, we get that the process,

$$M^j(t) = I_j(t) - I_j(0) - \int_0^t G(s)I_j(s)\, ds = I_j(t) - I_j(0) - \int_0^t \sum_{k\in\mathcal{S}} \mu_{kj}(s)I_k(s)\, ds,$$

is a martingale. As $I_j(t)$ is càdlàg, so is M^j. Because the intensities $\mu_{ij}(t)$ are uniformly bounded, $M^j(t)$ is also square integrable. Now by Theorem 3.5 and Theorem 3.19

$$[M^j](t) = \sum_{s\le t}(\Delta M^j)^2 = \sum_{s\le t}\Delta I_j(s).$$

Since $[M^j]$ is finite $I_i(s-) \in L(M^j)$ by Theorem 3.23 and the integral $I_i(s-)\cdot M^j \in \mathcal{M}^2$ by Theorem 3.20. Because $I_k(s) = I_k(s-)$ a.s. we have for $i \neq j$

$$\begin{aligned} M_{ij}(t) &= \int_{(0,t]} I_i(s-)\, dM^j(s) \\ &= \sum_{0<s\le t} I_i(s-)I_j(s) - \int_{(0,t]} \sum_k \mu_{kj}(s)I_i(s-)I_k(s)\, ds \\ &= \sum_{0<s\le t} I_i(s-)I_j(s) - \int_{(0,t]} \sum_k \mu_{kj}(s)I_i(s-)I_k(s-)\, ds \\ &= N_{ij}(t) - \int_{(0,t]} \mu_{ij}(s)I_i(s-)\, ds. \end{aligned}$$

By the characteristic property of the integral, (1),

$$\begin{aligned} [M_{ij}](t) &= \int_{(0,t]} I_i(s-)\, d[M^j] \\ &= \sum_{s\in(0,t]} I_i(s-)I_j(s) \\ &= N_{ij}(t). \end{aligned}$$

The predictable quadratic variation follows from theorems Theorem 3.17 and Theorem 3.19. □

Let us denote $M = \{M_{ij} : \; i \neq j\}$ for the martingales accompanying the Markov process X.

Theorem 4.10. *Let $Z = (Z_{ij})_{i \neq j}$ be a family of predictable processes $Z_{ij} \in \mathcal{P}$. Define (the random variable)*

$$||Z(s)||_{\mu}^{2} = \sum_{i,j:i \neq j} Z_{ij}^{2}(s) I_i(s-) \mu_{ij}(s). \tag{4}$$

Now, if

$$\mathbb{E}[\int_{(0,T]} ||Z(s)||_{\mu}^{2} \, ds] < \infty,$$

then

$$U(t) = \int_{(0,t]} Z(s) \, dM(s) = \sum_{i,j:i \neq j} \int_{(0,t]} Z_{ij}(s) \, dM_{ij}(s),$$

is a square integrable càdlàg martingale . We also have

$$[U](t) = \sum_{0<s \leq t} \sum_{i,j:i \neq j} |Z_{ij}(s) \Delta M_{ij}(s)|^2,$$

and

$$\mathbb{E}[[U](t)] = \mathbb{E}[\sum_{0<s \leq t} \sum_{i,j;i \neq j} |Z_{ij}(s) \Delta M_{ij}(s)|^2] = \mathbb{E}[\int_{(0,t]} ||Z(s)||_{\mu}^{2} \, ds]. \tag{5}$$

Proof. First let $H_{ij}^{n}(s) = \min\{Z_{ij}^{2}(s), n\}$, $n \in \mathbb{N}$, so that H_{ij}^{n} is a bounded predictable process. By Theorem 4.9 $M_s = ([M_{ij}]_s - \langle M_{ij} \rangle_s)$ is a martingale with integrable variation, therefore $H_{ij}^{n} \cdot M$ is a martingale by Theorem 3.6. Now

$$\mathbb{E} \int_{(0,T]} H_n(s) \, d([M_{ij}]_s - \langle M_{ij} \rangle_s) = 0.$$

Since

$$\langle M_{ij} \rangle(t) = \int_{(0,t]} I_i(s-) \mu_{ij}(s) \, ds,$$

we have

$$\begin{aligned}
\sum_{i,j;i \neq j} \mathbb{E} \int_{(0,T]} H_{ij}^{n}(s) \, d[M_{ij}]_s &= \sum_{i,j;i \neq j} \mathbb{E} \int_{(0,T]} H_{ij}^{n}(s) \, d\langle M_{ij} \rangle_s \\
&\leq \sum_{i,j;i \neq j} \mathbb{E} \int_{(0,T]} Z_{ij}^{2}(s) \, d\langle M_{ij} \rangle_s \\
&= \sum_{i,j;i \neq j} \mathbb{E} \int_{(0,T]} Z_{ij}^{2}(s) I_i(s-) \mu_{ij}(s) \, ds \\
&= \mathbb{E}[\int_{(0,T]} ||Z(t)||_{\mu}^{2} \, ds] < \infty.
\end{aligned}$$

We can use the monotone convergence Theorem 2.12 on the Lebesgue-Stieltjes integrals to get

$$\begin{aligned}\sum_{i,j;i\neq j}\mathbb{E}\int_{(0,T]} Z_{ij}^2(s)\,d[M_{ij}]_s &= \lim_{n\to\infty}\sum_{i,j;i\neq j}\mathbb{E}\int_{(0,T]} H_{ij}^n(s)\,d[M_{ij}]_s\\ &= \lim_{n\to\infty}\sum_{i,j;i\neq j}\mathbb{E}\int_{(0,T]} H_{ij}^n(s)\,d\langle M_{ij}\rangle_s\\ &= \sum_{i,j;i\neq j}\mathbb{E}\int_{(0,T]} Z_{ij}^2(s)\,d\langle M_{ij}\rangle_s\\ &= \mathbb{E}[\int_{(0,T]} ||Z(t)||_\mu^2\,ds] < \infty.\end{aligned}$$

Now by Theorem 3.23 $Z_{ij}\in L(M_{ij})$, and by (1)

$$[U](t) = \sum_{i,j:i\neq j}\int_{(0,t]} Z_{ij}^2(s)\,d[M_{ij}](s).$$

It follows from Theorem 3.20 that

$$U(t) = \int_{(0,t]} Z(s)\,dM(s) = \sum_{i,j:i\neq j}\int_{(0,t]} Z_{ij}(s)\,dM_{ij}(s)$$

is a square integrable càdlàg martingale .

By Theorem 3.9 and Theorem 4.9 M_{ij}, $i\neq j$, are purely discontinuous. Using Theorem 3.24, $Z\cdot M$ is also purely discontinuous. It follows from the Definition 3.18 with Theorem 3.24, we get

$$[U](t) = \sum_{i,j;i\neq j}\sum_{0<s\leq t}|Z_{ij}(s)\Delta M_{ij}(s)|^2.$$

Last, we have

$$\mathbb{E}[\sum_{0<s\leq t}\sum_{i,j;i\neq j}|Z_{ij}(s)\Delta M_{ij}(s)|^2] = \mathbb{E}[\int_{(0,t]}||Z(s)||_\mu^2\,ds].$$

□

4.2.1 Modeling the prospective reserve with a BSDE

To define the prospective reserve, we first define the payment process to express income and outgoes:

$$A(t) = \sum_{i\in\mathcal{S}}\int_{(0,t]} I_i(s)\,dA_i(s) + \sum_{i,j:i\neq j}\int_{(0,t]} a_{ij}(s-)\,dN_{ij}(s),$$

where a_{ij} is an adapted càdlàg process expressing payments upon transitions between states and A_i is an adapted process of bounded variation describing payments accumulated during a stay at state i. In addition we assume that A_i has the following decomposition:

$$A_i(t) = \int_{(0,t]} a_i(s)\, ds + \sum_{0<s\leq t} \Delta A_i(t),$$

where a_i is a progressively measurable process.

Now, we define the prospective reserve as

$$Y(t) = \mathbb{E}[\int_{(t,T]} e^{-\int_t^s \delta(u)\, du}\, dA(s)|\mathcal{F}_t],$$

for the payment process A, a (deterministic) discount rate δ and the Markov process X.

To begin studying the reserve we use the martingales, M, associated with the Markov process X. But first as in [3], for the sake of simplicity we make the additional assumptions for the rest of the thesis:

- $\mathbb{E}[\int_{(0,t]} |a_i(s)|^2\, ds] < \infty$ and $\mathbb{E}[\int_{(0,t]} ||(a_{ij})_{i\neq j}(s)||_\mu^2\, ds] < \infty$,
- $t \mapsto a_i(t,\omega)$ and $t \mapsto a_{ij}(t,\omega)$, for $\omega \in \Omega$, are continuous,
- The process A_i is continuous, that is $\Delta A_i = 0$.

Because X is càdlàg, it follows that $X(s-) = X(s)$ ds-a.e. and so $I_i(s-) = I_i(s)$ ds-a.e. Now setting

$$b(s,\omega,i) = \sum_{j:j\neq i} a_{ij}(s,\omega)\mu_{ij}(s),\ \ a(s,i) = a_i(s),$$

and using the continuity of a_{ij} and the fact that $I_i(s-) = I_i(s)$ ds-a.e. it follows that the payment process has the form

$$\begin{aligned}
A(t) &= \sum_{i\in\mathcal{S}} \int_{(0,t]} I_i(s)a(s,i)\, ds + \sum_{i\in\mathcal{S}} \int_{(0,t]} I_i(s)b(s,i))\, ds \\
&+ \sum_{i,j:i\neq j} \int_{(0,t]} a_{ij}(s)\, dM_{ij}(s) \\
&= \int_{(0,t]} a(s,X(s))\, ds + \int_{(0,t]} b(s,X(s))\, ds \\
&+ \sum_{i,j:i\neq j} \int_{(0,t]} a_{ij}(s)\, dM_{ij}(s).
\end{aligned}$$

Since, by Theorem 4.10,

$$\sum_{i,j:i\neq j} \int_{(0,t]} a_{ij}(s))\, dM_{ij}(s)$$

is a martingale, we get

$$\mathbb{E}[\sum_{i,j:i\neq j} \int_{(t,T]} a_{ij}(s)\, dM_{ij}(s)|\mathcal{F}_t] = 0.$$

With the above, and setting $g(s,\omega) = e^{-\int_t^s \delta(u)\, du}(a(s,\omega,X(s))+b(s,\omega,X(s))$, we can formulate the prospective reserve as

$$Y(t) = \mathbb{E}[\int_{(t,T]} g(s)\, ds|\mathcal{F}_t],$$

for the process g specified by the payment process A and the discount rate δ.

We are interested in the case where g is dependent on the reserve. However, due to the recursivity of the definition, it is not a given that there exists a process, Y, such that

$$Y(t) = \mathbb{E}[\int_{(t,T]} g(s, Y(s))\, ds|\mathcal{F}_t].$$

This problem can be framed as a backward stochastic differential equation as follows. Suppose we have adapted processes Y and $Z = (Z_{ij})_{i\neq j}$, satisfying appropriate measurability and integrability conditions later specified, such that

$$Y(t) = \int_{(t,T]} g(s, Y(s))\, ds - \sum_{i,j:i\neq j} \int_{(t,T]} Z_{ij}(s)\, dM_{ij}(s).$$

Let us take the conditional expectation. As Y is $\mathcal{F}_t$-adapted, we have

$$Y(t) = \mathbb{E}[\int_{(t,T]} g(s, Y(s))\, ds|\mathcal{F}_t] - \mathbb{E}[\sum_{i,j:i\neq j} \int_{(t,T]} Z_{ij}(s)\, dM_{ij}(s)|\mathcal{F}_t].$$

By the martingale property we have

$$\mathbb{E}[\sum_{i,j:i\neq j} \int_{(t,T]} Z_{ij}(s)\, dM_{ij}(s)|\mathcal{F}_t] = 0.$$

Now

$$Y(t) = \mathbb{E}[\int_{(t,T]} g(s, Y(s))\, ds|\mathcal{F}_t]. \tag{6}$$

This formulation is yet informal, but in the next chapter we introduce Markov chain BSDEs and lay out conditions under which there exists a solution. Then by the above argument, under these conditions (6) is defined.

5 Markov chain BSDEs

For this chapter we consider a stochastic basis $(\Omega, \mathcal{F}_T, \mathbb{P}, (\mathcal{F}_t)_{t\in I})$ satisfying the usual conditions and a Markov process X given by the model assumptions and Theorem 4.7.

Before we define the BSDEs, we give some notation and state the martingale representation theorem. Denote $J = \{(i,j) : \ i \neq j, \ i,j \in \mathcal{S}\}$. We use the following notation for the needed spaces of stochastic processes:

- $L^2_{\mathcal{F}_T}(\mathbb{R}) = \{\xi : \Omega \to \mathbb{R}, (\mathcal{F}_T, \mathcal{B}(\mathbb{R})) -$ random variable with $\mathbb{E}|\xi|^2 < \infty\}$,
- $L^2_{\mathcal{PM}_T}([0,T],\mathbb{R}) = \{Y : \Omega \times [0,T] \to \mathbb{R},$ progressively measurable, $\mathbb{E}[\int_{(0,T]} |Y(s)|^2\, ds] < \infty\}$,
- $\mathcal{G}^2 = \{Y : \Omega \times [0,T] \to \mathbb{R}, \mathbb{F} -$ adapted and càdlàg, $||Y||^2_{\mathcal{G}^2} = \mathbb{E}[\sup_{t\in[0,T]} |Y(t)|^2] < \infty\}$,
- $\mathcal{H}^2_\mu = \{(Z = (Z_{ij})_{i\neq j}; Z_{ij} : \Omega \times [0,T] \to \mathbb{R},$ predictable, $\mathbb{E}[\int_{(0,T]} ||Z(s)||^2_\mu\, ds] < \infty\}$, where $||\cdot||_\mu$ is defined in (4).

Denote $\hat{\Omega} = \Omega \times [0,T] \times J$. Define the measure $\nu : \mathcal{P} \otimes 2^J \to \mathbb{R}$ by

$$d\nu(\omega, t, (k,l)) = \sum_{(i,j)\in J} I_i(t-)\mu_{ij}(t)\, dt\, d\mathbb{P}(\omega)\delta_{(i,j)}(k,l).$$

Then, we have

$$||Z||^2_{\mathcal{H}^2_\mu} = \mathbb{E}\int_{(0,T]} \sum_{(i,j)\in J} Z^2_{ij}(t) I_i(t-)\mu_{ij}(t)\, dt = \int_{\hat{\Omega}} |Z(\hat{\omega})|^2\, d\nu(\hat{\omega}).$$

By Theorem 2.18, $\mathcal{H}^2_\mu = L^2(\hat{\Omega}, \mathcal{P} \otimes 2^J, \nu)$ is a Banach space. Here we made the convention that an element is unique in $\mathcal{H}^2_\mu$ if it is $I_i(t-)\mu_{ij}(t)\, dt\, d\mathbb{P}$-a.e. unique.

Also $L^2_{\mathcal{PM}_T}([0,T],\mathbb{R}) = L^2([0,T] \times \Omega, \mathcal{PM}_T, \lambda \otimes \mathbb{P}))$, where λ is the Lebesgue measure on $[0,T]$, is complete.

We need the following martingale representation theorem to find processes in the proof for the existence of a solution to the BSDE. Since $\{N_{ij} : \ i \neq \ j\}$ and X generate the same filtration we can state it in the following form.

Theorem 5.1 ([1, T11]). *Let L be a càdlàg square integrable $\mathbb{F}$-martingale. Then there exists a unique family of predictable processes $(Z_{ij})_{i\neq j} \in \mathcal{H}^2_\mu$, i.e. with*

$$\mathbb{E}[\int_{(0,T]} ||Z(t)||^2_\mu\, ds] < \infty,$$

such that

$$L(t) = L(0) + \sum_{i,j:i\neq j} \int_{(0,t]} Z_{ij}(s)\, dM_{ij}(s), dM(s), \text{ for } t \in [0,T], \text{ a.s.}$$

While this theorem is not constructive, the uniqueness makes it so that we can later use the following corollary to make Z explicit.

Corollary 5.2. *Let $L \in \mathcal{M}_0^2$. If $Z\Delta M = \Delta L$, then*

$$L(t) = \int_{(0,t]} Z(s)\, dM(s), \text{ for } t \in [0,T], \text{ a.s.}$$

Proof. Let $Z' \in \mathcal{H}_\mu^2$ such that $\int_{(0,t]} Z'(s)\, dM(s) = L(t)$, for $t \in [0,T]$, a.s. Then $Z'\Delta M = \Delta L$, and

$$(Z - Z')\Delta M = 0.$$

Now, by (5)

$$\mathbb{E}[\int_{(0,T]} ||(Z - Z')(s)||_\mu^2\, ds] = \mathbb{E}[\sum_{0<s\leq T} |(Z - Z')(s)\Delta M(s)|^2] = 0.$$

Therefore, $Z' = Z$ in $\mathcal{H}_\mu^2$, and the claim follows. □

We make formal the definition of the Markov chain backward stochastic differential equations.

Definition 5.3. *Let $\xi \in L^2_{\mathcal{F}_T}(\mathbb{R})$. Assume a map $g : [0,T] \times \Omega \times \mathbb{R} \times \mathbb{R}^J \to \mathbb{R}$ that is $\mathcal{PM}_T \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}^J)$-measurable.*

A pair $(Y, Z) \in \mathcal{G}^2 \times \mathcal{H}_\mu^2$ is a solution of the backward stochastic differential equation,

$$-\, dY(t) = g(t, Y(t), Z(t))\, dt - Z(t)\, dM(t),\ Y(T) = \xi,$$

if the following conditions hold:

(i) $Y_T = \xi$

(ii)

$$Y_t = \xi + \int_{(t,T]} g(s, Y(s), Z(s))\, ds - \int_{(t,T]} Z_s\, dM_s, \text{ for all } t \in [0,T), \text{ a.s.}$$

We call g the driver and the pair (ξ, g) the data of the BSDE.

5.1 Existence and uniqueness theorem for BSDEs

Let us define the following norms, in preparation for the proof on the existence of a solution to the BSDE. For $\beta \geq 0$ and $Z \in \mathcal{H}^2_\mu$

$$||Z||^2_{\mathcal{H}^2_\mu,\beta} = \mathbb{E}[\int_{(0,T]} e^{\beta s} ||Z(s)||^2_\mu \, ds],$$

and for $\beta \geq 0$ and $Y \in L^2_{\mathbb{F}}([0,T],\mathbb{R})$

$$||Y||^2_{2,\beta} = \mathbb{E}[\int_{(0,T]} e^{\beta s} Y_s^2 \, ds].$$

Also, we make the following assumptions, usual for differential equations.

(H1) There exists $C \geq 0$ such that for all $s \in [0,T]$, $y_1, y_2 \in \mathbb{R}$, $z_1 = (z^1_{ij})$, $z_2 = (z^2_{ij})$, $z^2_{ij}, z^1_{ij} \in \mathbb{R}$

$$|g(s,\omega,y_1,z_1) - g(s,\omega,y_2,z_2)| \leq C(|y_1 - y_2| + ||z_1 - z_2||_\mu(s,\omega))$$

(H2)

$$\mathbb{E}[\int_{(0,T]} |g(t,\omega,0,0)|^2 \, dt] < \infty$$

First we prove the following estimate. The proof is a close adaptation of [7, Proposition 5.3.1].

Theorem 5.4. *Let $(Y,Z), (\hat{Y},\hat{Z}) \in \mathcal{G}^2 \times \mathcal{H}^2_\mu$ be solutions to*

$$Y_t = \xi + \int_{(t,T]} g(s, Y_{s-}, Z_s) \, ds - \int_{(t,T]} Z_s \, dM_s,$$

for $t \in [0,T]$, and

$$\hat{Y}_t = \hat{\xi} + \int_{(t,T]} \hat{g}(s, \hat{Y}_{s-}, \hat{Z}_s) \, ds - \int_{(t,T]} \hat{Z}_s \, dM_s,$$

for $t \in [0,T]$, with data (ξ, g), $(\hat{\xi}, \hat{g})$ respectively. Assume that (H1-2) hold for both data. Let $A > 0$. Then for $\beta \geq A + 2C + 2C^2 - \frac{1}{2}$ we have

$$||Y - \hat{Y}||^2_{2,\beta} + ||Z - \hat{Z}||^2_{\mathcal{H}^2_\mu,\beta} \leq 2e^{\beta T}\mathbb{E}|Y_T - \hat{Y}_T|^2 + \frac{2}{A}||g(\cdot, Y(\cdot), Z(\cdot)) - \hat{g}(\cdot, Y(\cdot), Z(\cdot))||^2_{2,\beta}.$$

Proof. Assume $\beta > 0$. Consider $X_t = Y_t - \hat{Y}_t$ and $f(t, X_t) = e^{\beta t} X_t^2$ for Itô's formula, Theorem 3.29. As shown in Theorem 4.10 $Z \cdot M$ is purely discontinuous, therefore $X^c = 0$ and the second order term of the Itô's formula vanishes. Because $Y_{s-} = Y_s$ ds-almost everywhere, we have

$$\begin{aligned}
e^{\beta T}(Y_T - \hat{Y}_T)^2 = e^{\beta t}(Y_t - \hat{Y}_t)^2 \\
&+ \int_{(t,T]} \beta e^{\beta s}(Y_s - \hat{Y}_s)^2 \, ds \\
&- \int_{(t,T]} 2e^{\beta s}(Y_s - \hat{Y}_s)(g(s, Y_s, Z_s) - \hat{g}(s, \hat{Y}_s, \hat{Z}_s)) \, ds \\
&+ \int_{(t,T]} 2e^{\beta s}(Y_{s-} - \hat{Y}_{s-}) \, dU_s \\
&+ \sum_{t<s\leq T} (\Delta(e^{\beta s}(Y_s - \hat{Y}_s)^2) - 2e^{\beta s}(Y_{s-} - \hat{Y}_{s-})\Delta(Y_s - \hat{Y}_s)),
\end{aligned}$$

where

$$U_t = \int_{(0,t]} (Z_s - \hat{Z}_s) \, dM_s.$$

First, we have the computation

$$\begin{aligned}
&\sum_{t<s\leq T} \Delta(e^{\beta s}(X_s)^2) - 2e^{\beta s} X_{s-} \Delta X_s \\
&= \sum_{t<s\leq T} e^{\beta s}(X_s^2 - X_{s-}^2 - 2X_{s-}(X_s - X_{s-})) \\
&= \sum_{t<s\leq T} e^{\beta s}(X_s^2 - 2X_{s-}X_s + X_{s-}^2) \\
&= \sum_{t<s\leq T} e^{\beta s}(X_s - X_{s-})^2 \\
&= \sum_{t<s\leq T} e^{\beta s}(\Delta X_s)^2 \\
&= \sum_{t<s\leq T} e^{\beta s}|(Z_s - \hat{Z}_s)\Delta M_s|^2,
\end{aligned}$$

where the last line is by Theorem 3.24. Now, since

$$\mathbb{E} \int_{(t,T]} ||e^{\frac{1}{2}\beta s}(Z_s - \hat{Z}_s)||_\mu^2 \, ds \leq e^{\beta T} \mathbb{E} \int_{(0,T]} ||Z_s - \hat{Z}_s||_\mu^2 \, ds < \infty,$$

it follows from Theorem 4.10 that

$$\mathbb{E} \sum_{t<s\leq T} e^{\beta s}|(Z_s - \hat{Z}_s)\Delta M_s|^2 = \mathbb{E} \int_{(t,T]} e^{\beta s}||Z_s - \hat{Z}_s||_\mu^2 \, ds.$$

Next, estimating $e^{2\beta s}(Y_s - \hat{Y}_s)^2$, using the Cauchy-Schwarz inequality Theorem 2.14 and then Theorem 4.10 we get

$$\begin{aligned}
&\mathbb{E}[\Big(\int_{(0,T]} e^{2\beta s}(Y_{s-} - \hat{Y}_{s-})^2\, d[U_s]\Big)^{\frac{1}{2}}] \\
&\leq e^{2\beta T}\mathbb{E}[(\sup_{s\in[0,T]}(Y_{s-} - \hat{Y}_{s-})^2)^{\frac{1}{2}}(\int_{(0,T]} d[U_s])^{\frac{1}{2}}] \\
&= e^{2\beta T}\mathbb{E}[(\sup_{s\in[0,T]}(Y_{s-} - \hat{Y}_{s-})^2)^{\frac{1}{2}}([U](T))^{\frac{1}{2}}] \\
&\leq e^{2\beta T}\Big(\mathbb{E}[(\sup_{s\in[0,T]}(Y_{s-} - \hat{Y}_{s-})^2]\mathbb{E}[[U](T)]\Big)^{\frac{1}{2}} \\
&\leq e^{2\beta T}||Y_s - \hat{Y}_s||_{\mathcal{G}^2}||Z - \hat{Z}||_{\mathcal{H}^2_\mu} < \infty.
\end{aligned}$$

By Theorem 3.23 the following integral is defined, and therefore is a local martingale:

$$\int_{(0,t]} e^{\beta s}(Y_{s-} - \hat{Y}_{s-})\, dU_s. \tag{7}$$

Now, using the Burkholder-Davis-Gundy inequality Theorem 3.30 and the characteristic property of the stochastic integral, (1), we have

$$\begin{aligned}
&\mathbb{E}[\sup_{t\in[0,T]}|\int_{(0,t]} e^{\beta s}(Y_{s-} - \hat{Y}_{s-})\, dU_s|] \\
&\leq C\mathbb{E}[[\int_{(0,T]} e^{\beta s}(Y_{s-} - \hat{Y}_{s-})\, dU_s]^{\frac{1}{2}}] \\
&= C\mathbb{E}[(\int_{(0,T]} e^{2\beta s}(Y_{s-} - \hat{Y}_{s-})^2\, d[U_s])^{\frac{1}{2}}] < \infty,
\end{aligned}$$

for some constant $C > 0$. By Theorem 3.13 the integral (7) is a martingale, and so we get

$$\mathbb{E}\int_{(t,T]} 2e^{\beta s}(Y_{s-} - \hat{Y}_{s-})\, dU_s = 0.$$

Now we have

$$\begin{aligned}
\mathbb{E}e^{\beta T}(Y_T - \hat{Y}_T)^2 = {}& \mathbb{E}e^{\beta t}(Y_t - \hat{Y}_t)^2 \\
&+ \mathbb{E}\int_{(t,T]} \beta e^{\beta s}(Y_s - \hat{Y}_s)^2\, ds \\
&- \mathbb{E}\int_{(t,T]} 2e^{\beta s}(Y_s - \hat{Y}_s)(g(s, Y_s, Z_s) - \hat{g}(s, \hat{Y}_s, \hat{Z}_s))\, ds \\
&+ \mathbb{E}\int_{(t,T]} e^{\beta s}||Z_s - \hat{Z}_s||^2_\mu\, ds.
\end{aligned}$$

We estimate,

$$
\begin{aligned}
&2e^{\beta s}|Y_s - \hat{Y}_s||g(s, Y_s, Z_s) - \hat{g}(s, \hat{Y}_s, \hat{Z}_s)| \\
&\leq 2e^{\beta s}|Y_s - \hat{Y}_s||g(s, Y_s, Z_s) - \hat{g}(s, Y_s, Z_s)| \\
&+ 2e^{\beta s}|Y_s - \hat{Y}_s||\hat{g}(s, Y_s, Z_s) - \hat{g}(s, \hat{Y}_s, \hat{Z}_s)|.
\end{aligned}
$$

Using Young's inequality, for $A > 0$,

$$
2ab \leq Aa^2 + \frac{1}{A}b^2,
$$

with $a = |Y_s - \hat{Y}_s|$ and $b = |g(s, Y_s, Z_s) - \hat{g}(s, Y_s, Z_s)|$, we get

$$
\begin{aligned}
&2e^{\beta s}|Y_s - \hat{Y}_s||g(s, Y_s, Z_s) - \hat{g}(s, Y_s, Z_s)| \\
&\leq Ae^{\beta s}|Y_s - \hat{Y}_s|^2 + \frac{1}{A}e^{\beta s}|g(s, Y_s, Z_s) - \hat{g}(s, Y_s, Z_s)|^2.
\end{aligned}
$$

With the Lipschitz condition (H1),

$$
|\hat{g}(s, \omega, Y_s, Z_s) - \hat{g}(s, \omega, \hat{Y}_s, \hat{Z}_s)| \leq C(|Y_s - \hat{Y}_s| + ||Z_s - \hat{Z}_s||_\mu(\omega)),
$$

we have

$$
\begin{aligned}
&2e^{\beta s}|Y_s - \hat{Y}_s||\hat{g}(s, Y_s, Z_s) - \hat{g}(s, \hat{Y}_s, \hat{Z}_s)| \\
&\leq 2Ce^{\beta s}|Y_s - \hat{Y}_s|^2 + 2Ce^{\beta s}|Y_s - \hat{Y}_s|||Z_s - \hat{Z}_s||_\mu).
\end{aligned}
$$

Using Young's equality with $a = 2^{\frac{1}{2}}C|Y_s - \hat{Y}_s|$, $b = (\frac{1}{2})^{\frac{1}{2}}||Z_s - \hat{Z}_s||$, and $A = 1$, we get

$$
2Ce^{\beta s}|Y_s - \hat{Y}_s|||Z_s - \hat{Z}_s||_\mu \leq 2C^2e^{\beta s}|Y_s - \hat{Y}_s|^2 + \frac{1}{2}e^{\beta s}||Z_s - \hat{Z}_s||_\mu^2.
$$

Combining the above estimates, we have the inequality

$$
\begin{aligned}
&2e^{\beta s}|Y_s - \hat{Y}_s||g(s, Y_s, Z_s) - \hat{g}(s, \hat{Y}_s, \hat{Z}_s)| \\
&\leq Ae^{\beta s}|Y_s - \hat{Y}_s|^2 + \frac{1}{A}e^{\beta s}|g(s, Y_s, Z_s) - \hat{g}(s, Y_s, Z_s)|^2 \\
&+ 2Ce^{\beta s}|Y_s - \hat{Y}_s|^2 + 2C^2e^{\beta s}|Y_s - \hat{Y}_s|^2 + \frac{1}{2}e^{\beta s}||Z_s - \hat{Z}_s||_\mu^2.
\end{aligned}
$$

Using the above inequality, we have

$$\begin{aligned}
&\mathbb{E}e^{\beta t}|Y_t - \hat{Y}_t|^2 \\
&+ \mathbb{E}\int_{(t,T]} e^{\beta s}||Z_s - \hat{Z}_s||_\mu^2 \, ds \\
&\leq \mathbb{E}e^{\beta T}|Y_T - \hat{Y}_T|^2 \\
&+ (A + 2C + 2C^2 - \beta)\mathbb{E}\int_{(t,T]} e^{\beta s}(Y_s - \hat{Y}_s)^2 \, ds \\
&+ \frac{1}{A}\mathbb{E}\int_{(t,T]} e^{\beta s}|g(s, Y_s, Z_s) - \hat{g}(s, Y_s, Z_s)|^2 \, ds \\
&+ \frac{1}{2}\mathbb{E}\int_{(t,T]} e^{\beta s}||Z_s - \hat{Z}_s||_\mu^2 \, ds.
\end{aligned}$$

Now for $\beta \geq A + 2C + 2C^2 + \frac{1}{2}$ we have

$$\begin{aligned}
&\mathbb{E}e^{\beta t}|Y_t - \hat{Y}_t|^2 \\
&+ \frac{1}{2}\mathbb{E}\int_{(t,T]} e^{\beta s}||Z_s - \hat{Z}_s||_\mu^2 \, ds \\
&\leq \mathbb{E}e^{\beta T}|Y_T - \hat{Y}_T|^2 \\
&- \frac{1}{2}\mathbb{E}\int_{(t,T]} e^{\beta s}(Y_s - \hat{Y}_s)^2 \, ds \\
&+ \frac{1}{A}\mathbb{E}\int_{(t,T]} e^{\beta s}|g(s, Y_s, Z_s) - \hat{g}(s, Y_s, Z_s)|^2 \, ds.
\end{aligned}$$

Where letting $t = 0$ we get

$$\begin{aligned}
||Y - \hat{Y}||_{2,\beta}^2 + ||Z - \hat{Z}||_{\mathcal{H}_\mu^2,\beta}^2 \leq{}& 2e^{\beta T}\mathbb{E}|Y_T - \hat{Y}_T|^2 \\
&+ \frac{2}{A}||g(\cdot, Y(\cdot), Z(\cdot)) - \hat{g}(\cdot, Y(\cdot), Z(\cdot))||_{2,\beta}^2
\end{aligned}$$

□

We establish the existence and uniqueness of the solution to the BSDE, again closely adapting [7, Theorem 5.3.2].

Theorem 5.5. *Let the driver g be such that the conditions (H1-2) are satisfied. Then for any $\xi \in L^2_{\mathcal{F}_T}(\mathbb{R})$, for $t \in [0, T]$,*

$$Y_t = \xi + \int_{(t,T]} g(s, Y_s, Z_s) \, ds - \int_{(t,T]} Z_s \, dM_s,$$

has a solution $(Y, Z) \in \mathcal{G}^2 \times \mathcal{H}_\mu^2$. The solution (Y, Z) is unique in $\mathcal{G}^2 \times \mathcal{H}_\mu^2$.

Proof. Let $(\mathcal{Y},\mathcal{Z}) \in L^2_{\mathcal{P}\mathcal{M}_T}([0,T],\mathbb{R}) \times \mathcal{H}^2_\mu$. We construct a map $F : L^2_{\mathcal{P}\mathcal{M}_T}([0,T],\mathbb{R}) \times \mathcal{H}^2_\mu \to L^2_{\mathcal{P}\mathcal{M}_T}([0,T],\mathbb{R}) \times \mathcal{H}^2_\mu$, $F(\mathcal{Y},\mathcal{Z}) = (Y,Z)$ such that

$$Y_t = \xi + \int_{(t,T]} g(s,\mathcal{Y}_s,\mathcal{Z}_s)\,ds - \int_{(t,T]} Z_s\,dM_s. \tag{8}$$

To begin, set

$$N_t = \mathbb{E}[\xi + \int_{(0,T]} g(s,\mathcal{Y}_s,\mathcal{Z}_s)\,ds|\mathcal{F}_t].$$

Using Theorem 2.20 and Theorem 2.19 we have

$$\begin{aligned}
&\mathbb{E}[\mathbb{E}[\xi + \int_{(0,T]} g(s,\mathcal{Y}_s,\mathcal{Z}_s)\,ds|\mathcal{F}_t]^2]\\
&\leq \mathbb{E}[\mathbb{E}[|\xi + \int_{(0,T]} g(s,\mathcal{Y}_s,\mathcal{Z}_s)\,ds|^2|\mathcal{F}_t]]\\
&= \mathbb{E}[|\xi + \int_{(0,T]} g(s,\mathcal{Y}_s,\mathcal{Z}_s)\,ds|^2]\\
&\leq 2\mathbb{E}[|\xi|^2] + 2\mathbb{E}[|\int_{(0,T]} g(s,\mathcal{Y}_s,\mathcal{Z}_s)\,ds|^2]\\
&\leq 2\mathbb{E}[|\xi|^2] + 2T\mathbb{E}[\int_{(0,T]} |g(s,\mathcal{Y}_s,\mathcal{Z}_s) - g(s,0,0) + g(s,0,0)|^2\,ds]\\
&\leq 2\mathbb{E}[|\xi|^2] + 2T\mathbb{E}[\int_{(0,T]} 2C^2(|\mathcal{Y}_s|^2 + ||\mathcal{Z}_s||^2_\mu) + 2(g(s,0,0))^2\,ds]\\
&< \infty,
\end{aligned}$$

where we used the conditions (H1-2) and $\mathbb{E}\xi^2 < \infty$.

Since N_t is adapted it is a square integrable martingale, and by Theorem 3.2 we can choose a càdlàg version.

By Theorem 5.1 there exists a unique $Z \in \mathcal{H}^2_\mu$ such that

$$N_t = N_0 + \int_{(0,t]} Z_s\,dM_s$$

for $t \in [0,T]$. Now set

$$Y_t = \mathbb{E}[\xi + \int_{(t,T]} g(s,\mathcal{Y}_s,\mathcal{Z}_s)\,ds|\mathcal{F}_t].$$

By Fubini's theorem [5, Proposition 3.5.5] progressive measurability gives $\mathcal{F}_t$-measurability of

$$\omega \mapsto \int_{(0,t]} g(s,\omega,Y_s(\omega),Z_s(\omega))\,ds.$$

Therefore, we can calculate

$$\begin{aligned}
Y_t &= N_t - \int_{(0,t]} g(s, \mathcal{Y}_s, \mathcal{Z}_s)\, ds \\
&= N_0 + \int_{(0,t]} Z_s\, dM_s - \int_{(0,t]} g(s, \mathcal{Y}_s, \mathcal{Z}_s)\, ds \\
&= N_0 + \int_{(0,T]} Z_s\, dM_s - \int_{(0,T]} g(s, \mathcal{Y}_s, \mathcal{Z}_s)\, ds \\
&+ \int_{(t,T]} g(s, \mathcal{Y}_s, \mathcal{Z}_s)\, ds - \int_{(t,T]} Z_s\, dM_s \\
&= \xi + \int_{(t,T]} g(s, \mathcal{Y}_s, \mathcal{Z}_s)\, ds - \int_{(t,T]} Z_s\, dM_s,
\end{aligned}$$

which is (8).

Now let $\mathcal{Y} \in L^2_{\mathcal{P}\mathcal{M}_T}([0,T], \mathbb{R})$. From the construction of F we have that

$$Y_t = N_t - \int_{(0,t]} g(s, \mathcal{Y}_s, \mathcal{Z}_s)\, ds.$$

Since N_t is càdlàg, Y_t is also.

We have to show that the map is well defined. Let $\mathcal{Z}' \in \mathcal{H}^2_\mu$ such that $\mathcal{Z} = \mathcal{Z}'$ $I_i(t-)\mu_{ij}(t)\, dt\, d\mathbb{P}$-a.e. and $\mathcal{Y}' \in L^2_{\mathcal{P}\mathcal{M}_T}([0,T], \mathbb{R})$ such that $\mathcal{Y} = \mathcal{Y}'$ $ds\, d\mathbb{P}$-a.e. Then

$$\begin{aligned}
\mathbb{E}[|N_t - N'_t|] &= \mathbb{E}[|\mathbb{E}[\int_{(0,T]} g(s, \mathcal{Y}_s, \mathcal{Z}_s)\, ds - \int_{(0,T]} g(s, \mathcal{Y}'_s, \mathcal{Z}'_s)\, ds|\mathcal{F}_t]|] \\
&\leq \mathbb{E}[\int_{(0,T]} C(|\mathcal{Y}_s - \mathcal{Y}'_s| + ||\mathcal{Z}_s - \mathcal{Z}'_s||_\mu)\, ds] = 0
\end{aligned}$$

for all $t \in [0,T]$. We have $N_t = N'_t$ a.s. for all $t \in [0,T]$. Since N and N'_t are càdlàg modifications of each other, they are equal. Then by construction $Z = Z'$. Similarly

$$\begin{aligned}
\mathbb{E}[|Y_t - Y'_t|] &= \mathbb{E}[|\mathbb{E}[\int_{(0,T]} g(s, \mathcal{Y}_s, \mathcal{Z}_s)\, ds - \int_{(0,T]} g(s, \mathcal{Y}'_s, \mathcal{Z}'_s)\, ds|\mathcal{F}_t]|] \\
&\leq \mathbb{E}[\int_{(t,T]} C(|\mathcal{Y}_s - \mathcal{Y}'_s| + ||\mathcal{Z}_s - \mathcal{Z}'_s||_\mu)\, ds] = 0,
\end{aligned}$$

for all $t \in [0,T]$, gives $Y = Y'$. We have shown that (Y, Z) exists and is unique.

As Y is càdlàg, we have by Doob's maximal inequality Theorem 3.4

$$\begin{aligned}
||Y||_{\mathcal{G}^2}^2 &= \mathbb{E}\sup_t |Y_t|^2 \\
&\leq 2\mathbb{E}\sup_t |N_t|^2 + 2\mathbb{E}|\int_{(0,T]} |g(s,\mathcal{Y}_s,\mathcal{Z}_s)|\,ds|^2 \\
&\leq 8\mathbb{E}|N_T|^2 + 2\mathbb{E}|\int_{(0,T]} |g(s,\mathcal{Y}_s,\mathcal{Z}_s)|\,ds|^2 \\
&\leq 16\mathbb{E}|Y_T|^2 + 18\mathbb{E}|\int_{(0,T]} |g(s,\mathcal{Y}_s,\mathcal{Z}_s)|\,ds|^2 \\
&< \infty.
\end{aligned}$$

Now we have that $Y \in \mathcal{G}^2$. We have shown $F(L^2_{\mathcal{P}\mathcal{M}_T}([0,T],\mathbb{R}) \times \mathcal{H}^2_\mu) \subset \mathcal{G}^2 \times \mathcal{H}^2_\mu$.

Next we will show that the map F is a contraction. Let $(\mathcal{Y},\mathcal{Z}),(\hat{\mathcal{Y}},\hat{\mathcal{Z}}) \in L^2_{\mathcal{P}\mathcal{M}_T}([0,T],\mathbb{R}) \times \mathcal{H}^2_\mu$. Define

$$f(s,\omega) = g(s,\omega,\mathcal{Y}_s(\omega),\mathcal{Z}_s(\omega))$$

for $(Y,Z) = F(\mathcal{Y},\mathcal{Z})$ and

$$\hat{f}(s,\omega) = g(s,\omega,\hat{\mathcal{Y}}_s(\omega),\hat{\mathcal{Z}}_s(\omega))$$

for $(\hat{Y},\hat{Z}) = F(\hat{\mathcal{Y}},\hat{\mathcal{Z}})$. Set $\xi = \hat{\xi}$, then by the construction of F, (Y,Z) and $(\hat{Y},\hat{Z})$ solve

$$Y_t = \xi + \int_{(t,T]} f(s)\,ds - \int_{(t,T]} Z_s\,dM_s,$$

for $t \in [0,T]$, and

$$\hat{Y}_t = \xi + \int_{(t,T]} \hat{f}(s)\,ds - \int_{(t,T]} \hat{Z}_s\,dM_s,$$

for $t \in [0,T]$, respectively. For these equations the assumptions of Theorem 5.4 are satisfied and therefore

$$\begin{aligned}
||Y-\hat{Y}||_{2,\beta}^2 + ||Z-\hat{Z}||_{\mathcal{H}^2_\mu,\beta}^2 &\leq \frac{2}{A}||f(\cdot,\cdot)-\hat{f}(\cdot,\cdot)||_{2,\beta}^2 \\
&\leq \frac{2}{A}||g(\cdot,\cdot,\mathcal{Y}(\cdot),\mathcal{Z}(\cdot)) - g(\cdot,\cdot,\hat{\mathcal{Y}}(\cdot),\hat{\mathcal{Z}}(\cdot))||_{2,\beta}^2 \\
&\leq \frac{4C^2}{A}(||\mathcal{Y}-\hat{\mathcal{Y}}||_{2,\beta}^2 + ||\mathcal{Z}-\hat{\mathcal{Z}}||_{\mathcal{H}^2_\mu,\beta}^2).
\end{aligned}$$

By taking A such that $A > 4C^2$, and β according to Theorem 5.4 we have that F is a contraction.

Since $1 \leq e^{\beta t} \leq e^{\beta T}$, we have that the norms $||\cdot||_{\mathcal{H}^2_\mu}$, $||\cdot||_{\mathcal{H}^2_\mu,\beta}$ induce equivalent metrics. It follows that the space $L^2_{\mathcal{PM}_T}([0,T],\mathbb{R}) \times \mathcal{H}^2_\mu$ is a complete space with the β-metric.

Because F is a contraction, there exists a fixed point, $F(Y,Z) = (Y,Z)$, by the Banach fixed point theorem, [16, Theorem 12.8].
Since $F(L^2_{\mathcal{PM}_T}([0,T],\mathbb{R}) \times \mathcal{H}^2_\mu) \subset \mathcal{G}^2 \times \mathcal{H}^2_\mu$, we have $(Y,Z) \in \mathcal{G}^2 \times \mathcal{H}^2_\mu$.

Last, for $\xi = \hat{\xi}$ and $g = \hat{g}$ by Theorem 5.4

$$
\begin{aligned}
||Y - \hat{Y}||^2_{2,\beta} + ||Z - \hat{Z}||^2_{\mathcal{H}^2_\mu,\beta} &\leq 2e^{\beta T}\mathbb{E}|\xi - \hat{\xi}|^2 \\
&\quad + \frac{2}{A}||g(\cdot, Y(\cdot), Z(\cdot)) - \hat{g}(\cdot, Y(\cdot), Z(\cdot))||^2_{2,\beta} \\
&= 0,
\end{aligned}
$$

giving us the uniqueness. □

5.2 Markovian BSDEs

In this section we introduce Markovian BSDEs and prove a nonlinear Thiele equation for the prospective reserve, when the driver is a deterministic function of the Markov process. We primarily follow [3] for this section, but we have altered the assumptions for the proofs. Also, the proof for the existence of the solution to the Thiele equation adapts that of [14, Theorem 5.1].

We say that the BSDE

$$
\begin{cases}
-dY(t) = g(t, \cdot, Y(t), Z(t))\,dt - Z(t)\,dM(t), \\
Y(T) = \xi,
\end{cases}
$$

where $\xi = \phi(X(T))$, for a given (deterministic) $\phi : \mathcal{S} \to \mathbb{R}$, is Markovian if $Y(t) = V(t, X(t))$ for a deterministic function $V : [0,T] \times \mathcal{S} \to \mathbb{R}$.

As mentioned, we assume that the driver g specifying the payment process is a deterministic function of $(t, X(t), Y(t), Z(t))$, now defined as a function $g : [0,T] \times \mathcal{S} \times \mathbb{R} \times \mathbb{R}^J \to \mathbb{R}$ that is $\mathcal{PM}_T \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}^J)$-measurable.

Instead of (H1-2) we have:

(G1) There exists $C \geq 0$ such that for all $s \in [0,T]$, $i \in \mathcal{S}$, $y_1, y_2 \in \mathbb{R}$, $z_1 = (z^1_{ij})$, $z_2 = (z^2_{ij})$, $z^2_{ij}, z^1_{ij} \in \mathbb{R}$

$$
|g(s, i, y_1, z_1) - g(s, i, y_2, z_2)| \leq C(|y_1 - y_2| + ||z_1 - z_2||_\mu(s, \omega)).
$$

(G2) For all $j \in \mathcal{S}$

$$\int_0^T |g(s, j, 0, 0)| \, ds < \infty.$$

Theorem 5.6. *Let $g : [0, T] \times \mathcal{S} \times \mathbb{R} \times \mathbb{R}^J \to \mathbb{R}$ be such that G1-2 hold and $t \mapsto g(t, x, y, z)$ is continuous for $x \in \mathbb{R}, y \in \mathbb{R}, z \in \mathbb{R}^J$, and $\phi : \mathcal{S} \to \mathbb{R}$. Then there exists a solution $V : [0, T] \times \mathcal{S} \to \mathbb{R}$, $t \mapsto V(t, i)$ differentiable, to the nonlinear Thiele equation:*

$$\begin{cases} \frac{\partial V}{\partial t}(t, i) + g(t, i, V(t, i), (V(t, k) - V(t, j))_{jk}) \\ + \sum_{j : j \neq i} \mu_{ij}(t)(V(t, j) - V(t, i)) = 0, \\ V(T, i) = \phi(i), \ i \in \mathcal{S}. \end{cases}$$

Proof. We endow $\mathcal{S}$ with the discrete topology. First we assume that there exists $V \in C([0, T] \times \mathcal{S}, \mathbb{R})$ such that

$$\begin{aligned} V(t, i) = &\sum_{j \in \mathcal{S}} p_{ij}(t, T)\phi(j) \\ &+ \int_t^T \sum_{j \in \mathcal{S}} p_{ij}(t, s) g(s, j, V(s, j), (V(s, k) - V(s, j))_{jk}) \, ds. \end{aligned} \tag{9}$$

Since (G1) holds and V and g are continuous in their first arguments we have the continuity of $s \mapsto g(s, j, V(s, j), (V(s, k) - V(s, j))_{jk})$. Because $p_{ij}(t, s)$ is continuously differentiable in t, we can use the Leibniz integral rule to get

$$\begin{aligned} \frac{\partial V}{\partial t}(t, i) = &\sum_{j \in \mathcal{S}} \frac{\partial p_{ij}}{\partial t}(t, T)\phi(j) \\ &- \sum_{j \in \mathcal{S}} p_{ij}(t, t) g(t, j, V(t, j), (V(t, k) - V(t, j))_{jk}) \\ &+ \int_t^T \sum_{j \in \mathcal{S}} \frac{\partial p_{ij}}{\partial t}(t, s) g(s, j, V(s, j), (V(s, k) - V(s, j))_{jk}) \, ds. \end{aligned} \tag{10}$$

By the Kolmogorov backward equation, (3), we get

$$\begin{aligned} \frac{\partial V}{\partial t}(t, i) = &- \sum_{j \in \mathcal{S}} \sum_k \mu_{ik}(t) p_{kj}(t, T)\phi(j) \\ &- g(t, i, V(t, i)), (V(t, k) - V(t, j))_{jk}) \\ &- \int_t^T \sum_{j \in \mathcal{S}} \sum_k \mu_{ik}(t) p_{kj}(t, s) g(s, j, V(s, j)), (V(s, k) - V(s, j))_{jk}) \, ds. \end{aligned}$$

Using (9), we have

$$\frac{\partial V}{\partial t}(t,i) = -g(t,i,V(t,i)),(V(t,k)-V(t,j))_{jk})$$
$$-\sum_k \mu_{ik}(t)V(t,k).$$

We reformulate this using (2),

$$\begin{aligned}
&\frac{\partial V}{\partial t}(t,i) + g(t,i,V(t,i)),(V(t,k)-V(t,j))_{jk}) + G(t)V(t,i) \\
&= \frac{\partial V}{\partial t}(t,i) + g(t,i,V(t,i)),(V(t,k)-V(t,j))_{jk}) \\
&+ \sum_{j:j\neq i} \mu_{ij}(t)(V(t,j)-V(t,i)) \\
&= 0.
\end{aligned}$$

Now we prove that such V exists. Let $\mathcal{V} \in C([0,T]\times\mathcal{S},\mathbb{R})$. Define $H(\mathcal{V}) = V$ by

$$\begin{aligned}
V(t,i) &= \sum_{j\in\mathcal{S}} p_{ij}(t,T)\phi(j) \\
&\quad + \int_t^T \sum_{j\in\mathcal{S}} p_{ij}(t,s)g(s,j,\mathcal{V}(s,j)),(\mathcal{V}(s,k)-\mathcal{V}(s,j))_{jk})\,ds.
\end{aligned}$$

Similarly as in (10), we get the differentiability of $t \mapsto V(t,i)$. Therefore $V \in C([0,T]\times\mathcal{S},\mathbb{R})$ and $H : C([0,T]\times\mathcal{S},\mathbb{R}) \to C([0,T]\times\mathcal{S},\mathbb{R})$ is well defined.

Let $\mathcal{V}_1,\mathcal{V}_2 \in C([0,T]\times\mathcal{S},\mathbb{R})$. Define $\mathcal{Z}_r = (\mathcal{Z}^r_{jk})_{(j,k)\in J}$, for $r=1,2$, by

$$\mathcal{Z}^r_{jk}(s) = \mathcal{V}_r(s,k) - \mathcal{V}_r(s,j), \text{ for } t\in[0,T],\ \ j,k\in\mathcal{S}.$$

Since μ_{jk} are uniformly bounded on $[0,T]$, we can define

$$\mu^* = \sup_{j,k:j\neq k}\ \sup_{s\in[0,T]} \mu_{jk}(s).$$

Now we can estimate

$$\begin{aligned}
||\mathcal{Z}_1(s)-\mathcal{Z}_2(s)||^2_\mu &= \sum_{j\neq k} |\mathcal{V}_1(s,k)-\mathcal{V}_2(s,k)-(\mathcal{V}_1(s,j)-\mathcal{V}_2(s,j))|^2 I_j(s-)\mu_{jk}(s) \\
&\leq \mu^* \sum_{j\neq k} 2|\mathcal{V}_1(s,k)-\mathcal{V}_2(s,k)|^2 + 2|\mathcal{V}_1(s,j)-\mathcal{V}_2(s,j)|^2 \\
&\leq 4\mu^*|J| \sup_{k\in\mathcal{S}}\ \sup_{r\in[0,s]} |\mathcal{V}_1(r,k)-\mathcal{V}_2(r,k)|^2,
\end{aligned}$$

and, for some $L > 0$,

$$||\mathcal{Z}_1(s) - \mathcal{Z}_2(s)||_\mu \le L \sup_{k\in\mathcal{S}} \sup_{s\in[0,s]} |\mathcal{V}_1(s,k) - \mathcal{V}_2(s,k)|.$$

Now we can compute, for $H(\mathcal{V}_1) = V_1$ and $H(\mathcal{V}_2) = V_2$,

$$\begin{aligned}
&|V_1(t,i) - V_2(t,i)| \\
&\le \int_t^T \sum_{j\in\mathcal{S}} p_{ij}(t,s)|g(s,j,\mathcal{V}_1(s,j),\mathcal{Z}_1(s)) - g(s,j,\mathcal{V}_2(s,j),\mathcal{Z}_2(s))|\, ds \\
&\le \int_t^T \sum_{j\in\mathcal{S}} |g(s,j,\mathcal{V}_1(s,j),\mathcal{Z}_1(s)) - g(s,j,\mathcal{V}_2(s,j),\mathcal{Z}_2(s))|\, ds \\
&\le \int_t^T \sum_{j\in\mathcal{S}} C(|\mathcal{V}_1(s,j) - \mathcal{V}_2(s,j)| + ||\mathcal{Z}_1(s) - \mathcal{Z}_2(s)||_\mu)\, ds \\
&\le \int_t^T \sum_{j\in\mathcal{S}} C(\sup_{k\in\mathcal{S}} \sup_{s\in[0,s]} |\mathcal{V}_1(s,k) - \mathcal{V}_2(s,k)| + L \sup_{k\in\mathcal{S}} \sup_{s\in[0,s]} |\mathcal{V}_1(s,k) - \mathcal{V}_2(s,k)|)\, ds \\
&\le CS(1+L) \int_t^T \sup_{k\in\mathcal{S}} \sup_{s\in[0,s]} |\mathcal{V}_1(s,k) - \mathcal{V}_2(s,k)|\, ds \\
&= M_1 \int_t^T \sup_{k\in\mathcal{S}} \sup_{s\in[0,s]} |\mathcal{V}_1(s,k) - \mathcal{V}_2(s,k)|\, ds,
\end{aligned}$$

where $M_1 = CS(1+L)$. This implies

$$\sup_{k\in\mathcal{S}} \sup_{s\in[0,T]} |V_1(t,k) - V_2(t,k)| \le M_1 \int_0^T \sup_{k\in\mathcal{S}} \sup_{s\in[0,s]} |\mathcal{V}_1(s,k) - \mathcal{V}_2(s,k)|\, ds. \quad (11)$$

First we get the continuity of H:

$$\begin{aligned}
\sup_{k\in\mathcal{S}} \sup_{s\in[0,T]} |V_1(t,k) - V_2(t,k)| &\le M_1 \int_0^T \sup_{k\in\mathcal{S}} \sup_{s\in[0,s]} |\mathcal{V}_1(s,k) - \mathcal{V}_2(s,k)|\, ds \\
&\le M_1 T \sup_{k\in\mathcal{S}} \sup_{s\in[0,T]} |\mathcal{V}_1(s,k) - \mathcal{V}_2(s,k)|.
\end{aligned}$$

Then define a sequence V^n by $V^0(t,i) = \sum_{j\in\mathcal{S}} p_{ij}(t,T)\phi(j)$ and $V^{n+1} = H(V^n)$. By the above inequality (11),

$$\begin{aligned}
&\sup_{i\in\mathcal{S}} \sup_{t\in[0,T]} |V^{n+1}(t,i) - V^n(t,i)| \\
&\le M_1 \int_0^T (\sup_{k\in\mathcal{S}} \sup_{u\in[0,s]} |V^n(u,k) - V^{n-1}(u,k)|)\, ds.
\end{aligned}$$

Similarly as before, now also using (G2) and the definition of V^0

$$
\begin{aligned}
&|V^1(t,i) - V^0(t,i)| \\
&\leq \int_t^T \sum_{j\in\mathcal{S}} p_{ij}(t,s)|g(s,j,V^0(s,j),\mathcal{Z}^0(s))|\,ds \\
&\leq \int_t^T \sum_{j\in\mathcal{S}} |g(s,j,V^0(s,j),Z^0(s)) - g(s,j,0,0) + g(s,j,0,0)|\,ds \\
&\leq \int_t^T \Big(\sum_{j\in\mathcal{S}} C(|V^0(s,j)| + ||Z^0(s)||_\mu) + \sum_{j\in\mathcal{S}} |g(s,j,0,0)|\Big)\,ds \\
&\leq CS(1+L)\int_0^T \sup_{k\in\mathcal{S}} \sup_{u\in[0,T]} |V^0(u,k)|\,ds + \sum_{j\in\mathcal{S}} \int_0^T |g(s,j,0,0)|\,ds \\
&\leq CS(1+L)T\sum_{j\in\mathcal{S}} \phi(j) + \sum_{j\in\mathcal{S}} \int_0^T |g(s,j,0,0)|\,ds \\
&\leq M_2 T,
\end{aligned}
$$

for some $M_2 > 0$. Giving us

$$
\sup_{i\in\mathcal{S}} \sup_{t\in[0,T]} |V^1(t,i) - V^0(t,i)| \leq M_2 T.
$$

Now we get recursively

$$
\begin{aligned}
&\sup_{k\in\mathcal{S}} \sup_{s\in[0,T]} |V^{n+1}(t,i) - V^n(t,i)| \\
&\leq M_1 \int_0^T (\sup_{k\in\mathcal{S}} \sup_{u\in[0,s_{n-1}]} |V^n(u,k) - V^{n-1}(u,k)|)\,ds_{n-1} \\
&\leq M_1^2 \int_0^T \int_0^{s_{n-1}} (\sup_{k\in\mathcal{S}} \sup_{u\in[0,s_{n-2}]} |V^{n-1}(u,k) - V^{n-2}(u,k)|)\,ds_{n-2}\,ds_{n-1} \\
&\dots \\
&\leq M_1^n \int_0^T \cdots \int_0^{s_1} M_2 T\,ds_0 \dots\,ds_{n-1} \\
&= \frac{M_1^n M_2 T^{n+1}}{n!}.
\end{aligned}
$$

And finally,

$$\begin{aligned}
&\sup_{k\in\mathcal{S}}\sup_{s\in[0,T]}|V^m(t,i)-V^n(t,i)| \\
&\leq \sum_{i=n}^{m-1}\sup_{k\in\mathcal{S}}\sup_{s\in[0,T]}|V^{i+1}(t,i)-V^i(t,i)| \\
&\leq \sum_{i=n}^{m-1}\frac{M_1^iM_2T^{i+1}}{i!},
\end{aligned}$$

which goes to 0 as $n\to\infty$. We have shown that (V^n) is Cauchy in $C([0,T]\times\mathcal{S},\mathbb{R})$, and therefore we have $V\in C([0,T]\times\mathcal{S},\mathbb{R})$ $V^n\to V$. Because H is continuous we also have $H(V^n)\to V$, and so $H(V)=V$. $\square$

Theorem 5.7. *Let the driver g be such that (G1-2) hold and $t\mapsto g(t,x,y,z)$ is continuous for $x\in\mathbb{R}, y\in\mathbb{R}, z\in\mathbb{R}^J$, and $\phi:\mathcal{S}\to\mathbb{R}$. Let $V:[0,T]\times\mathcal{S}\to\mathbb{R}$, $t\mapsto V(t,i)$ be differentiable, and solve the nonlinear Thiele equation*

$$\begin{cases}\dfrac{\partial V}{\partial t}(t,i)+g(t,i,V(t,i),(V(t,k)-V(t,j))_{jk}) \\ +\displaystyle\sum_{j:j\neq i}\mu_{ij}(t)(V(t,j)-V(t,i))=0, \\ V(T,i)=\phi(i),\;\; i\in\mathcal{S}.\end{cases}$$

Now the pair (Y,Z) where $Y(t)=V(t,X(t))$ and $Z_{jk}(t)=V(t,k)-V(t,j)$ for $t\in[0,T]$, $j,k\in\mathcal{S}$, solves the BSDE

$$\begin{cases}-\,dY(t)=g(t,X(t),Y(t),Z(t))\,dt-Z(t)\,dM(t)\\ Y(T)=\phi(X(T)).\end{cases}$$

Then, we have the following representation for the prospective reserve as a deterministic function of t and $X(t)$:

$$V(t,X(t))=\mathbb{E}[\phi(X(T))+\int_{(t,T]}g(s,X(s),Y(s),Z(s))\,ds|X(t)],$$

for $t\in[0,T]$.

Proof. The proof follows [3, Theorem 3.4.2].

By Dynkin's formula Theorem 4.8

$$\begin{aligned}M^V(t)&=V(t,X(t))-V(0,X(0))\\ &\quad-\int_0^t\Big(G(s)V(s,X(s))+\frac{\partial V}{\partial s}(s,X(s))\Big)\,ds\end{aligned}\tag{12}$$

is a martingale, for which we have the jump process

$$\Delta M^V(t) = \Delta V(t, X(t)) = \sum_{j:i\neq j} (V(t,j) - V(t,i)) I_i(t-) I_j(t)$$
$$= \sum_{j:i\neq j} (V(t,j) - V(t,i)) \Delta M_{ij}(t).$$

Define $Z = (Z_{ij})$, $Z_{ij}(t) = V(t,j) - V(t,i)$. Using Corollary 5.2 we get

$$M^V(t) = \int_{(0,t]} Z(s)\, dM(s) = \sum_{j:i\neq j} \int_{(0,t]} (V(t,j) - V(t,i))\, dM(s).$$

Now by (12) we get

$$\int_{(t,T]} Z(s)\, dM(s) = V(T, X(T)) - V(t, X(t))$$
$$- \int_{(t,T]} G(s) V(s, X(s)) + \frac{\partial V}{\partial s}(s, X(s))\, ds.$$

Using the Thiele equation in the form

$$\frac{\partial V}{\partial t}(t,i) + g(t, i, V(t,i), (V(t,k) - V(t,j))_{jk}) + G(t)V(t,i) = 0,$$

and rearranging we get

$$V(t, X(t)) = V(T, X(T)) + \int_{(t,T]} g(t, X(t), V(t, X(t)), (V(t,k) - V(t,j))_{jk})\, ds$$
$$- \int_{(t,T]} Z(s)\, dM(s),$$

which is the BSDE for (Y, Z), where $Y(t) = V(t, X(t))$.

Taking conditional expectation with respect to $\mathcal{F}_t$,

$$Y(t) = \mathbb{E}[\phi(X(T)) + \int_{(t,T]} g(s, X(s), Y(s), Z(s))\, ds | \mathcal{F}_t]. \tag{13}$$

Because X, Y and Z are uniformly bounded on $\Omega \times [0,T]$ we can use (G1) and the continuity of $t \mapsto g(t,i,y,z)$ to get that $(\omega, t) \mapsto g(t, X(\omega,t), Y(\omega,t), Z(\omega,t))$ is uniformly bounded on $\Omega \times [0,T]$. We also have $\sigma(X(t))$-measurability of $\omega \mapsto g(t, X(t,\omega), Y(t,\omega), Z(t,\omega))$. Now we can interchange the integral and conditional expectation by Theorem 2.21 and use the Markov property Theorem 4.2 to get

$$
\begin{aligned}
&\mathbb{E}[\int_{(t,T]} g(s, X(s), Y(s), Z(s))\, ds|\mathcal{F}_t] \\
&= \int_{(t,T]} \mathbb{E}[g(s, X(s), Y(s), Z(s))|\mathcal{F}_t]\, ds \\
&= \int_{(t,T]} \mathbb{E}[g(s, X(s), Y(s), Z(s))|X(t)]\, ds \\
&= \mathbb{E}[\int_{(t,T]} g(s, X(s), Y(s), Z(s))\, ds|X(t)].
\end{aligned}
$$

Finally, with the above and (13) we have

$$
Y(t) = \mathbb{E}[\phi(X(T)) + \int_{(t,T]} g(s, X(s), Y(s), Z(s))\, ds)|X(t)].
$$

□

References

[1] P. Brémaud, *Point Processes and Queues: Martingale Dynamics*, Springer, 1981.

[2] J. Casteren, *Markov Processes, Feller Semigroups and Evolution Equations*, World Scientific Publishing, 2011.

[3] B. Djehiche, B. Löfdahl, *Nonlinear reserving in life insurance: Aggregation and mean-field approximation*,
Insurance: Mathematics and Economics Volume 69, 2016. https://doi.org/10.1016/j.insmatheco.2016.04.002.

[4] R. Dudley, *Real Analysis and Probability*, Cambridge University Press, 2004.

[5] C. Geiss and S. Geiss, *An introduction to probability theory*, Department of Mathematics and Statistics, University of Jyväskylä, 2014.

[6] S. Geiss, *An introduction to probability theory II*, Department of Mathematics and Statistics, University of Jyväskylä, 2014.

[7] S. Geiss, *Stochastic Differential Equation*, Department of Mathematics and Statistics, University of Jyväskylä, 2020.

[8] C. Geiss and S. Geiss, *Markov Processes*, Department of Mathematics and Statistics, University of Jyväskylä, 2021.

[9] D. Giraudo,
https://math.stackexchange.com/users/9849/davide-giraudo,
Exchange integral and conditional expectation,
URL (version: 2017-04-13): https://math.stackexchange.com/q/310037,
downloaded 17.7.2022.

[10] S. He, J. Wang, J. Yan, *Semimartingale Theory and Stochastic Calculus*, Science Press, 1992.

[11] F. Klebaner, *Introduction To Stochastic Calculus With Applications*, Imperial College Press, 2005.

[12] J. Lehrbäck, *Mitta- Ja Integraaliteoria*, Department of Mathematics and Statistics, University of Jyväskylä, 2018.

[13] R. Norberg, *Basic Life Insurance Mathematics*, 2002, http://web.math.ku.dk/ mogens/lifebook.pdf.

[14] J. Nykänen, *On the uniqueness of a solution and stability of McKean-Vlasov stochastic differential equations*, Department of Mathematics and Statistics, University of Jyväskylä, 2020.

[15] P. Protter, *Stochastic Integration and Differential Equations*, Springer, 2004.

[16] J. Väisälä, *Topologia I*, Limes ry, 2012.