g-Expectations with application to risk measures

Sonja C. Offwood

Programme in Advanced Mathematics of Finance,
University of the Witwatersrand,
Johannesburg.

5 August 2012

Abstract

Peng introduced a typical filtration consistent nonlinear expectation, called a g-expectation in [40]. It satisfies all properties of the classical mathematical expectation besides the linearity. Peng's conditional g-expectation is a solution to a backward stochastic differential equation (BSDE) within the classical framework of Itô's calculus, with terminal condition given at some fixed time T. In addition, this g-expectation is uniquely specified by a real function g satisfying certain properties. Many properties of the g-expectation, which will be presented, follow from the specification of this function. Martingales, super- and submartingales have been defined in the nonlinear setting of g-expectations. Consequently, a nonlinear Doob-Meyer decomposition theorem was proved.

Applications of g-expectations in the mathematical financial world have also been of great interest. g-Expectations have been applied to the pricing of contingent claims in the financial market, as well as to risk measures. Risk measures were introduced to quantify the riskiness of any financial position. They also give an indication as to which positions carry an acceptable amount of risk and which positions do not. Coherent risk measures and convex risk measures will be examined. These risk measures were extended into a nonlinear setting using the g-expectation. In many cases due to intermediate cashflows, we want to work with a multi-period, dynamic risk measure. Conditional g-expectations were then used to extend dynamic risk measures into the nonlinear setting.

The Choquet expectation, introduced by Gustave Choquet, is another nonlinear expectation. An interesting question which is examined, is whether there are incidences when the g-expectation and the Choquet expectation coincide.

Acknowledgements

I would like to express my appreciation to Coenraad Labuschagne for his assistance and supervision. I would also like to thank my mother for her continued support and encouragement throughout the year, as well as my sister for her time in assisting and supporting me. My appreciation also goes out to everyone at Wits who provided support and encouragement.

Contents

Chapter 1

Introduction

In 1933 Andrei Kolmogorov set out the axiomatic basis for modern probability theory in his book 'Foundation of Probability Theory' (Grundbegriffe der Wahrscheinlichkeitsrechnung). It was built on the measure theory introduced by Émile Borel and Henry Lebesgue, and extended by Radon and Fréchet. Nowadays the probability measure space $(\Omega, \mathcal{F}, \mathbb{P})$ is a standard concept and appears in most papers on probability and mathematical finance. Another important notion which was introduced, was the concept of expectation. The linear expectation of a $\mathcal{F}$-measurable random variable X is defined as the integral $\int_\Omega X d\mathbb{P}$. In his book 'Foundation of Probability Theory', Kolmogorov used the Radon-Nikodym theorem to introduce the conditional probability and the conditional expectation under a certain σ-algebra $\mathcal{G} \subset \mathcal{F}$.

An interesting problem which has come about, is whether a nonlinear expectation could be developed under which we still have such a conditional expectation. Peng introduced a typical filtration consistent nonlinear expectation, called g-expectation in [40]. It satisfies all properties of the classical mathematical expectation besides the linearity. Peng's conditional g-expectation is a solution to a backward stochastic differential equation (BSDE) within the classical framework of Itô's calculus, with terminal condition given at some fixed time T. In addition, this g-expectation is uniquely specified by a real function g satisfying certain properties. Many properties of the g-expectation follow from the specification of this function. For example, when g is a convex (resp. concave) function, it can be shown that the g-expectation is also convex (resp. concave). It can also be shown that the classical mathematical expectation corresponds to the case when $g = 0$. The Girsanov transformations are also contained in the concept of g-expectations.

g-Expectations are a fairly new research topic in mathematics and finance. The properties and behaviour of g-expectations have been studied extensively by Peng [40, 41, 42, 43, 44]. In a mathematical sense, g-expectations are of particular interest.

The concept of martingales do not require the linearity assumption of the classical mathematical expectation. Hence, martingales, super- and submartingales have been defined in the nonlinear setting of g-expectations. Consequently, a nonlinear Doob-Meyer decomposition theorem was proved. Other results on g-expectations were obtained by Briand et al. [3], Jiang [32, 33], Chen et al. [5, 6, 7] and Pardoux and Peng [38].

Applications of g-expectations in the mathematical financial world have also been of great interest. g-Expectations have been applied to the pricing of contingent claims in the financial market. This was studied by El Karoui et al. [35] amongst others. Risk measures are another application of g-expectations. Risk measures were introduced to quantify the riskiness of any financial position. They also give an indication as to which positions carry an acceptable amount of risk and which positions do not. A well known and popular risk measure is Value at Risk (VaR). This risk measure has been of great interest in financial and mathematical research. However, due to the drawbacks of VaR, Artzner et al. [1] introduced some desirable properties which lead to the concept of coherent risk measures. Föllmer and Schied then generalised this concept and introduced convex risk measures in general probability spaces. These risk measures were extended into a nonlinear setting using the g-expectation. The risk measures mentioned thus far have been one-period risk measures. However, in many cases due to intermediate cashflows, we want to work with a multi-period, dynamic risk measure. Risk measures were first introduced in a dynamic setting by Cvitanic and Karatzas [14] and Wang [49]. Frittelli and Rosazza Gianin [26] and Riedel [46] amongst others have done more recent studies on dynamic risk measures. Conditional g-expectations were then used to extend dynamic risk measures into the nonlinear setting. In fact the assumptions on the function g turn out to give an ideal characterisation of the dynamical behaviour of risk measures.

The Choquet expectation, introduced by Gustave Choquet, is another nonlinear expectation. An interesting question that arises is whether there are incidences when the g-expectation and the Choquet expectation coincide. What assumptions do we require for a g-expectation to be a Choquet expectation, and alternatively for a Choquet expectation to be a g-expectation?

The concept of g-expectations and many of the results in this dissertation have been generalised to a general framework by Samuel Cohen and Robert Elliot [9, 10, 11].

This masters dissertation has been set up in the following way: in Chapter 2 we recall some important mathematical preliminaries required in the remainder of the dissertation. It covers mathematical expectation as well as stochastic differential equations and Brownian motions. Chapter 3 introduces the BSDE and defines g-expectations and conditional g-expectations. Properties applying to g-expectations are given. We also consider special cases for the function g and examine its effect on the g-expectation. Section 3.5 gives the representation lemma for g-expectations. The resemblance between the classical mathematical expectation and the g-expectation should become clear in this chapter. This chapter also shows how the Black-Scholes option pricing formula can be retained from g-expectations. An introduction to coherent and convex, static and dynamic risk measures follows in Chapter 4. Chapter 5 combines the results obtained in Chapters 3 and 4. This chapter begins by analysing the properties positive homogeneity, subadditivity, convexity, translation invariance and monotonicity with regards to g-expectations. Next we define risk measures in terms of g-expectations. Chapter 6 defines the Choquet integral and gives general properties of the Choquet expectation. We also show that the classical mathematical expectation is a special case of the Choquet expectation and link this nonlinear Choquet expectation to Peng's g-expectation. The masters dissertation ends off with Chapter 7 on the nonlinear Doob-Meyer decomposition theorem. We recall the classical Doob-Meyer decomposition theorem and any related definitions. Thereafter, we define and outline some properties of g-solutions, g-super- and subsolutions as well as of g-martingales, g-super- and g-submartingales. Section 7.4 presents the nonlinear Doob-Meyer decomposition theorem.

Appendix A contains some mathematical results required for various proofs in the dissertation. For ease of reference Appendix B give an overview of the different spaces we are working in.

Chapter 2

Mathematical preliminaries

This chapter covers some mathematical preliminaries required for the remainder of the masters dissertation. Section 2.1 focuses on the concept of mathematical expectation. Section 2.2 defines stochastic differential equations, Brownian motions and states some of the important properties of stochastic calculus.

2.1 Mathematical expectation

The majority of the information in this section has been taken from Shreve [47]. We begin by recalling the definition of the mathematical expectation in a general probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We also recall the definition of the conditional mathematical expectation. Some important and useful properties of the mathematical expectation will be stated.

Definition 2.1. Let X be a random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The *expectation* (or *expected value*) of X is defined to be

$$\mathbb{E}[X] = \int_{\Omega} X(\omega)\, d\mathbb{P}(\omega). \tag{2.1}$$

Definition 2.2. The random variable X is *integrable* if and only if

$$\mathbb{E}[|X|] < \infty.$$

The basic properties of the expected value will be stated next.

Proposition 2.3. *Let X and Y be random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.*

(i) If $X \leq Y$ almost surely and X and Y are integrable (or almost surely nonnegative), then

$$\mathbb{E}[X] \leq \mathbb{E}[Y].$$

In particular, if $X = Y$ almost surely and one of the random variables is integrable (or almost surely nonnegative), then they are both integrable (or almost surely nonnegative), and

$$\mathbb{E}[X] = \mathbb{E}[Y].$$

(ii) If α and β are real constants and X and Y are integrable (or if α and β are nonnegative constants and X and Y are nonnegative), then

$$\mathbb{E}[\alpha X + \beta Y] = \alpha \mathbb{E}[X] + \beta \mathbb{E}[Y].$$

(iii) If ϑ is a convex, real-valued function and X is an integrable random variable, then

$$\mathbb{E}[\vartheta(X)] \geq \vartheta(\mathbb{E}[X]).$$

Property (ii) from the previous proposition tells us that the mathematical expectation is a linear operator. Property (iii) is the well known *Jensen's inequality.*

We next define the conditional mathematical expectation.

Definition 2.4. Let X be an integrable or almost surely nonnegative random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. The *conditional expectation* of X given $\mathcal{G}$, denoted $\mathbb{E}[X|\mathcal{G}]$, is defined to be the random variable that satisfies the following properties:

(i) $\mathbb{E}[X|\mathcal{G}]$ is $\mathcal{G}$-measurable, and

(ii)
$$\int_A \mathbb{E}[X|\mathcal{G}](\omega)\, d\mathbb{P}(\omega) = \int_A X(\omega)\, d\mathbb{P}(\omega), \text{ for all } A \in \mathcal{G}. \tag{2.2}$$

Equation (2.2) is equivalent to having for all $A \in \mathcal{G}$

$$\mathbb{E}[\mathbf{1}_A \mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[\mathbf{1}_A X]. \tag{2.3}$$

The basic properties of the conditional mathematical expectation will be stated next.

Proposition 2.5. *Let X and Y be random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{G}$ be a sub-σ-algebra of F.*

(i) If α and β are real constants and X and Y are integrable (or if α and β are nonnegative constants and X and Y are almost surely nonnegative), then

$$\mathbb{E}[\alpha X + \beta Y|\mathcal{G}] = \alpha \mathbb{E}[X|\mathcal{G}] + \beta \mathbb{E}[Y|\mathcal{G}].$$

(ii) *If X and Y are integrable (or X is positive and Y is $\mathcal{G}$-measurable), Y and XY are integrable, and X is $\mathcal{G}$-measurable, then*

$$\mathbb{E}[XY|\mathcal{G}] = X\mathbb{E}[Y|\mathcal{G}].$$

(iii) *If $\mathcal{H}$ is a sub-σ-algebra of $\mathcal{G}$ and X is integrable (or almost surely nonnegative), then*

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}].$$

(iv) *If X is integrable and independent of $\mathcal{G}$, then*

$$\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X].$$

(v) *If ϑ is a convex, real-valued function and X is an integrable random variable, then*

$$\mathbb{E}[\vartheta(X)|\mathcal{G}] \geq \vartheta(\mathbb{E}[X|\mathcal{G}]).$$

Property (i) from Proposition 2.5 tells us that the mathematical conditional expectation is a linear operator. Property (iii) is also known as the ‘tower property’ or as ‘iterated conditioning’. Property (v) is the well known *conditional Jensen's inequality.*

This classical, linear mathematical expectation is a powerful tool for dealing with stochastic processes. However, there are many uncertain mathematical phenomena that are not easily modelled using the classical mathematical expectation. Thus nonlinear operators were introduced. Choquet [8] introduced the concept of a capacity and subsequently defined the nonlinear Choquet expectation based on this concept. Peng [40] introduced the so-called g-expectation, a nonlinear expectation based on a backward stochastic differential equation.

We are interested in comparing the g-expectation, defined in Chapter 3, to the classical mathematical expectation. The resemblance between the two will become clear once we have defined and stated the properties of the g-expectation. In fact, we will see that most of the basic properties of the classical mathematical expectation, besides the linearity, are preserved with the g-expectation.

2.2 Stochastic differential equations

For completeness, this section will cover stochastic differential equations, the definition of a Brownian motion and that of a martingale, super- and submartingale. Results regarding Brownian motions and martingales, such as the martingale representation theorem and Girsanov's theorem, will also be stated. The work in this section has been taken from Shreve [47].

Consider a stochastic differential equation

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dB_t, \tag{2.4}$$

where B_t is a Brownian motion. The terms $\mu(t, X_t)$ and $\sigma(t, X_t)$ represent the drift term and the volatility term of X_t respectively. The definition of a Brownian motion follows. We define it in one dimension, it can however easily be extended to d dimensions.

Definition 2.6. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. For each $\omega \in \Omega$, suppose there is a continuous function $B_t : \mathbb{R}^+ \times \Omega \to \mathbb{R}$ that satisfies $B(0) = 0$. Then $(B_t)_{t\geq 0}$ is a *Brownian motion* if for all $0 = t_0 < t_1 < \ldots < t_m$, the increments

$$B_{t_1} = B_{t_1} - B_{t_0},\, B_{t_2} - B_{t_1},\, \ldots,\, B_{t_m} - B_{t_{m-1}}$$

are independent and each of these increments is normally distributed with

$$\mathbb{E}[B_{t_{i+1}} - B_{t_i}] = 0,$$
$$\mathrm{Var}[B_{t_{i+1}} - B_{t_i}] = t_{i+1} - t_i.$$

Paths of Brownian motions are continuous and nowhere differentiable. Consequently normal integration and differentiation rules do not apply to Brownian motions. The most common stochastic differential equation involving Brownian motions is *geometric Brownian motion.* If the process X_t follows geometric Brownian motion, it satisfies the stochastic differential equation

$$dX_t = \mu X_t dt + \sigma X_t dB_t.$$

Here, X_t is lognormally distributed with parameters μ and σ. Subsequently $\log X_t$ is normally distributed with mean μ and standard deviation σ.

Definition 2.7. Let Ω be equipped with the filtration $(\mathcal{F}_t)_{t\in[0,T]}$. The stochastic process $(X_t)_{t\in[0,T]}$ is said to be *adapted* if, for each t, the random variable X_t is $\mathcal{F}_t$-measurable.

Theorem 2.8 (Itô's formula)**.** *Let $(X_t)_{t\geq 0}$ be a stochastic process with dynamics given by (2.4). Consider a function $f(t, x)$ which is continuous, once differentiable with respect to time and twice differentiable with respect to x. Then, for any $t \geq 0$, $f(t, X_t)$ satisfies*

$$\begin{aligned} df(t, X_t) = &\left[\frac{\partial}{\partial t} f(t, x) + \mu(t, X_t)\frac{\partial}{\partial x} f(t, x) + \frac{1}{2}\sigma^2(t, X_t)\frac{\partial^2}{\partial^2 x} f(t, x)\right] dt \\ &+ \sigma(t, X_t)\frac{\partial}{\partial x} f(t, x) dB_t. \end{aligned}$$

Definition 2.9. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, T a fixed positive number and $(\mathcal{F}_t)_{t\in[0,T]}$ a filtration of sub-σ-algebras of $\mathcal{F}$. Let $(M_t)_{t\in[0,T]}$ be an adapted stochastic process.

- The process, M_t, is a *martingale* if for all $0 \leq s \leq t \leq T$
$$\mathbb{E}[M_t|\mathcal{F}_s] = M_s \text{ a.s.}$$
- The process, M_t, is a *submartingale* if for all $0 \leq s \leq t \leq T$
$$\mathbb{E}[M_t|\mathcal{F}_s] \geq M_s \text{ a.s.}$$
- The process, M_t, is a *supermartingale* if for all $0 \leq s \leq t \leq T$
$$\mathbb{E}[M_t|\mathcal{F}_s] \leq M_s \text{ a.s.}$$

Note that a process that is both a sub- and a supermartingale, is a martingale. Conversely, a martingale is simultaneously a sub- and a supermartingale.

The martingale representation theorem connects Brownian motions and martingales. We state the martingale representation theorem in one dimension. It can also easily be extended to d dimensions. The one-dimensional martingale representation theorem makes use of a one-dimensional Brownian motion, whereas the d-dimensional equivalent uses the d-dimensional Brownian motion.

Theorem 2.10 (Martingale representation theorem, one dimension)**.** *Let $(B_t)_{t\in[0,T]}$ be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $(\mathcal{F}_t)_{t\in[0,T]}$ be the filtration generated by this Brownian motion. Let $(M_t)_{t\in[0,T]}$ be a martingale with respect to this filtration. Then there is an adapted process $(\Gamma_t)_{t\in[0,T]}$ such that*
$$M_t = M_0 + \int_0^t \Gamma_u dB_u, \quad 0 \leq t \leq T.$$

Suppose we are working with the filtration generated by the Brownian motion. The martingale representation theorem states that any martingale with respect to this filtration can be represented as an initial condition plus an Itô integral with respect to the Brownian motion. We can see here that a stochastic differential equation with a zero drift term is a martingale.

Lastly we state a fundamental result in stochastic calculus, called Girsanov's theorem. Girsanov's theorem describes how the dynamics of a stochastic process change under a change in measure. Girsanov's theorem is particularly important in financial mathematics as it allows us to change from the real-world probability measure into the risk-neutral measure. Under the risk-neutral probability measure, contingent claims can be priced fairly.

Theorem 2.11 (Girsanov's theorem). *Let $(B_t)_{t\in[0,T]}$ be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $(\mathcal{F}_t)_{t\in[0,T]}$ be a filtration for this Brownian motion. Let θ_t be an adapted process. Define*

$$Z_t = exp\left[\int_0^t \theta_u dB_u - \frac{1}{2}\int_0^t \theta_u^2 du\right],$$
$$\tilde{B}_t = B_t + \int_0^t \theta_u du,$$

and assume that

$$\mathbb{E}\int_0^T \theta_u^2 Z_u^2 du < \infty.$$

Under the probability measure defined by

$$\mathbb{Q}(A) = \int_A Z(\omega) d\mathbb{P}(\omega) \quad \textit{for all } A \in \mathcal{F}$$

the process $(\tilde{B}_t)_{t\in[0,T]}$ is a Brownian motion.

Chapter 3

g-Expectations

In this chapter we define the concept of g-expectation and conditional g-expectation. We also state and prove properties applying to these g-expectations. Section 3.1 introduces the notation required for the remainder of the dissertation. In Section 3.2 we define the general concept of a filtration consistent nonlinear expectation, before continuing on to the definition of g-expectations in Section 3.4. Properties of the filtration consistent nonlinear expectation, as well as of the g-expectation will be proved. We also consider special cases for the function g and examine its effect on the g-expectation. Section 3.5 gives the representation lemma for g-expectations. The resemblance between the classical mathematical expectation and the g-expectation will become clear in this chapter. The chapter ends off with a financial application of the g-expectation.

3.1 Notation

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $(B_t)_{t\geq 0}$ be a standard d-dimensional Brownian motion on this space. Let $(\mathcal{F}_t^B)_{t\geq 0}$ be the filtration generated by the Brownian motion, i.e. $\mathcal{F}_t^B = \sigma\{B_s; 0 \leq s \leq t\}$ for any $t \geq 0$. Denote by $(\mathcal{F}_t)_{t\geq 0}$ the augmented filtration such that $\mathcal{F}_t = \sigma\{B_s; 0 \leq s \leq t\} \cup \mathcal{N}$ for any $t \geq 0$, where $\mathcal{N}$ is the collection of all $\mathbb{P}$-null sets.

Let $T > 0$ be a fixed time horizon and for simplicity of notation let $\mathcal{F} = \mathcal{F}_T$. Without loss of generality consider processes indexed by $t \in [0, T]$.

Let $L^2(\mathcal{F}_t) = L^2(\Omega, \mathcal{F}_t, \mathbb{P})$, with $t \in [0, T]$, denote the space of all real-valued, $\mathcal{F}_t$-measurable and square integrable random variables applying with the L^2-norm $\|\cdot\|_2$. Let X be a real-valued random variable, then the L^2-norm $\|\cdot\|_2$ is defined by

$$\|X\|_2 = \left(\int_0^T |X_t|^2 dt\right)^{\frac{1}{2}}.$$

Let $L^2_{\mathcal{F}}(T, \mathbb{R}^n)$ denote the space of all $\mathbb{R}^n$-valued, $\mathcal{F}_T$-adapted processes $(V_t)_{t\in[0,T]}$ with

$$\mathbb{E}\int_0^T |V_t|^2 dt < \infty,$$

where $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^n$. The Euclidean norm on $\mathbb{R}^n$ is defined by

$$|(v_1, \dots v_n)| = \left(\sum_{i=1}^n v_i^2\right)^{\frac{1}{2}}.$$

Also let $M_{\mathcal{F}}(\mathbb{R}^n)$ denote the space of all $\mathbb{R}^n$-valued, $\mathcal{F}_t$-progressively measurable processes $(\psi_t)_{t\in[0,T]}$. Let $H^q_{\mathcal{F}}(T, \mathbb{R}^n)$ denote the space of all $(\psi_t)_{t\in[0,T]} \in M_{\mathcal{F}}(\mathbb{R}^n)$ with

$$\mathbb{E}\int_0^T |\psi_t|^q dt < \infty.$$

Before we introduce the backward stochastic differential equation (BSDE) which will be used in defining the g-expectation and conditional g-expectation, we define nonlinear expectations and filtration consistent nonlinear expectations.

3.2 Filtration consistent nonlinear expectation

The majority of this section is based on the work by Coquet, Hu, Mémin and Peng [12].

Definition 3.1. A *(nonlinear) expectation* is a function

$$\mathcal{E}[\,\cdot\,] : L^2(\mathcal{F}_T) \to \mathbb{R}$$

which satisfies

(i) *Strict monotonicity*: For all $X_1, X_2 \in L^2(\mathcal{F}_T)$, if $X_1 \geq X_2$ a.s. then $\mathcal{E}[X_1] \geq \mathcal{E}[X_2]$, and if $X_1 \geq X_2$ a.s. then $\mathcal{E}[X_1] = \mathcal{E}[X_2]$ if and only if $X_1 = X_2$ a.s..

(ii) *Constancy*: For all $c \in \mathbb{R}$, we have $\mathcal{E}[\mathbf{1}c] = c$, where $\mathbf{1} : \Omega \to \mathbb{R}$ is defined by $\mathbf{1} = 1$ a.e. As is customary in literature, we write $\mathcal{E}[\mathbf{1}c] = \mathcal{E}[c]$.

Lemma 3.2. *Let $t \in [0, T]$ and let $\eta_1, \eta_2 \in L^2(\mathcal{F}_t)$. If for all $A \in \mathcal{F}_t$*

$$\mathcal{E}[\eta_1 \mathbf{1}_A] = \mathcal{E}[\eta_2 \mathbf{1}_A],$$

then

$$\eta_1 = \eta_2 \ \textit{a.s.}$$

Proof. Choose $A = \{\eta_1 \geq \eta_2\} \in \mathcal{F}_t$. Since $\eta_1 \mathbf{1}_A \geq \eta_2 \mathbf{1}_A$ and $\mathcal{E}[\eta_1 \mathbf{1}_A] = \mathcal{E}[\eta_2 \mathbf{1}_A]$, it follows from the definition of (nonlinear) expectation, that $\eta_1 \mathbf{1}_A = \eta_2 \mathbf{1}_A$ a.s.. Therefore we have that $\eta_1 \leq \eta_2$ a.s.. Similarly, we can show that $\eta_2 \leq \eta_1$ a.s.. The result follows and the proof is complete. □

Definition 3.3. For the given filtration $(\mathcal{F}_t)_{t\in[0,T]}$, a (nonlinear) expectation is called an *$\mathcal{F}_t$-consistent expectation* (*$\mathcal{F}$-expectation*) if for each $X \in L^2(\mathcal{F}_T)$ and for each $t \in [0,T]$ there exists a random variable $\eta \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$, such that for all $A \in \mathcal{F}_t$

$$\mathcal{E}[\mathbf{1}_A X] = \mathcal{E}[\mathbf{1}_A \eta]. \tag{3.1}$$

By Lemma 3.2 this η, which is denoted by $\mathcal{E}[X|\mathcal{F}_t]$, is uniquely defined. $\mathcal{E}[X|\mathcal{F}_t]$ is called the *conditional $\mathcal{F}$-expectation* of X under $\mathcal{F}_t$ and by Definition 3.3 it is characterised such that for all $A \in \mathcal{F}_t$, we have

$$\mathcal{E}[\mathbf{1}_A X] = \mathcal{E}[\mathbf{1}_A \mathcal{E}[X|\mathcal{F}_t]]. \tag{3.2}$$

The properties of $\mathcal{F}_t$-consistent expectations are listed below. We will soon see that the g-expectation is an $\mathcal{F}_t$-consistent expectation. Thus these properties also apply to the g-expectation.

Proposition 3.4 (Properties of $\mathcal{F}$-expectation)**.**

(i) For each $0 \leq r \leq t \leq T$, we have $\mathcal{E}[\mathcal{E}[X|\mathcal{F}_t]|\mathcal{F}_r] = \mathcal{E}[X|\mathcal{F}_r]$. In particular, we have $\mathcal{E}[\mathcal{E}[X|\mathcal{F}_t]] = \mathcal{E}[X]$.

(ii) For all $t \in [0,T]$ and $X_1, X_2 \in L^2(\mathcal{F}_T)$ with $X_1 \geq X_2$, we have $\mathcal{E}[X_1|\mathcal{F}_t] \geq \mathcal{E}[X_2|\mathcal{F}_t]$. If moreover $\mathcal{E}[X_1|\mathcal{F}_t] = \mathcal{E}[X_2|\mathcal{F}_t]$ a.s. for some $t \in [0,T]$, then $X_1 = X_2$ a.s..

(iii) For all $B \in \mathcal{F}_t$, we have $\mathcal{E}[\mathbf{1}_B X|\mathcal{F}_t] = \mathbf{1}_B \mathcal{E}[X|\mathcal{F}_t]$.

Proof. (i) Let $0 \leq r \leq t \leq T$. For each $A \in \mathcal{F}_r$, we also have $A \in \mathcal{F}_t$. Thus applying Equation (3.2) twice, we get

$$\mathcal{E}[\mathbf{1}_A X] = \mathcal{E}[\mathbf{1}_A \mathcal{E}[\mathcal{E}[X|\mathcal{F}_t]|\mathcal{F}_r]].$$

Now $\mathcal{E}[X|\mathcal{F}_r]$ is the unique random variable η in $L^2(\mathcal{F}_r)$ such that $\mathcal{E}[\mathbf{1}_A X] = \mathcal{E}[\mathbf{1}_A \eta]$, hence it follows that

$$\mathcal{E}[\mathcal{E}[X|\mathcal{F}_t]|\mathcal{F}_r] = \mathcal{E}[X|\mathcal{F}_r].$$

The particular case follows by setting $r = 0$ and from the fact that $\mathcal{F}_0$ is the trivial σ-algebra.

(ii) Set $\eta_i := \mathcal{E}[X_i|\mathcal{F}_t]$, $i = 1, 2$. Since $X_1 \geq X_2$, we have by Equation (3.2) and by the monotonicity of the (nonlinear) expectation that for each $A \in \mathcal{F}_t$,

$$\mathcal{E}[\mathbf{1}_A\eta_1] = \mathcal{E}[\mathbf{1}_A X_1] \geq \mathcal{E}[\mathbf{1}_A X_2] = \mathcal{E}[\mathbf{1}_A\eta_2].$$

In particular, we know that $\{\eta_1 \leq \eta_2\}$ is $\mathcal{F}_t$-measurable and thus setting

$$A = \{\eta_1 \leq \eta_2\},$$

we get $\mathcal{E}[\mathbf{1}_{\{\eta_1\leq\eta_2\}}\eta_1] \geq \mathcal{E}[\mathbf{1}_{\{\eta_1\leq\eta_2\}}\eta_2]$. It is clear that $\mathbf{1}_{\{\eta_1\leq\eta_2\}}\eta_1 \leq \mathbf{1}_{\{\eta_1\leq\eta_2\}}\eta_2$. By the definition of the (nonlinear) expectation, the above two relations imply that $\mathbf{1}_{\{\eta_1\leq\eta_2\}}\eta_1 = \mathbf{1}_{\{\eta_1\leq\eta_2\}}\eta_2$ which consequently implies that $\eta_1 \geq \eta_2$ i.e. $\mathcal{E}[X_1|\mathcal{F}_t] \geq \mathcal{E}[X_2|\mathcal{F}_t]$.

Now assume that $\mathcal{E}[X_1|\mathcal{F}_t] = \mathcal{E}[X_2|\mathcal{F}_t]$ a.s. for some $t \in [0, T]$. Applying $\mathcal{E}[\,\cdot\,]$ on both sides, we get by (i) that $\mathcal{E}[X_1] = \mathcal{E}[X_2]$ and by the monotonicity of $\mathcal{E}[\,\cdot\,]$, it follows that $X_1 = X_2$ a.s..

(iii) Applying Equation (3.2) in two consecutive steps, we have for each $A \in \mathcal{F}_t$

$$\begin{aligned}\mathcal{E}[\mathbf{1}_A\mathcal{E}[\mathbf{1}_B X|\mathcal{F}_t]] &= \mathcal{E}[\mathbf{1}_A\mathbf{1}_B X]\\ &= \mathcal{E}[\mathbf{1}_A\{\mathbf{1}_B\mathcal{E}[X|\mathcal{F}_t]\}].\end{aligned}$$

It follows easily that $\mathcal{E}[\mathbf{1}_B X|\mathcal{F}_t] = \mathbf{1}_B\mathcal{E}[X|\mathcal{F}_t]$ and consequently we have proved (iii). □

3.3 Backward stochastic differential equations

Consider a function

$$\begin{aligned}g : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d &\to \mathbb{R}\\ (\omega, t, y, z) &\mapsto g(\omega, t, y, z).\end{aligned}$$

In any operation involving $g(\omega, t, y, z)$ we assume the operation holds $\mathbb{P}$-almost surely, for any fixed $t \in [0, T]$. For simplicity we omit the '$\mathbb{P}$-almost surely' and write $g(t, y, z)$ instead of $g(\omega, t, y, z)$.

Pardoux and Peng [38] introduced the BSDE

$$-dY_t = g(t, Y_t, Z_t)dt - Z_t dB_t, \qquad 0 \leq t \leq T \tag{3.3}$$

with terminal condition

$$Y_T = X, \tag{3.4}$$

where $X \in L^2(\mathcal{F}_T)$. We can write this BSDE in integral form as

$$Y_t = X + \int_t^T g(s, Y_s, Z_s)\, ds - \int_t^T Z_s\, dB_s, \qquad 0 \leq t \leq T. \tag{3.5}$$

Remark 3.5. BSDE (3.3) can equivalently be written as

$$dY_t = -g(t, Y_t, Z_t)dt + Z_t dB_t, \qquad 0 \le t \le T$$

with terminal condition

$$Y_T = X.$$

However the literature on g-expectations prefers expressing the BSDE in the form of Equation (3.3)-(3.4). We are working with a terminal condition at time T instead of an initial condition. Hence $-\int_t^T dY_s$ integrates to $Y_t - Y_T$ which makes it easy to solve the equation for Y_t yielding the integral form seen in Equation (3.5).

In the above we have $(Y_t, Z_t)_{t\in[0,T]} \in L^2_{\mathcal{F}}(T, \mathbb{R}) \times L^2_{\mathcal{F}}(T, \mathbb{R}^d)$, which denotes the solution to BSDE (3.3) with terminal condition (3.4) i.e. Y_t is a $\mathbb{R}$-valued, $\mathcal{F}_T$-adapted process and Z_t is a $\mathbb{R}^d$-valued, $\mathcal{F}_T$-predictable process.

Assume that the function g satisfies the following conditions.

(A1) g is *uniformly Lipschitz continuous* in (y, z), i.e. there exists a constant $C > 0$ such that for all $y_1, y_2 \in \mathbb{R}$ and $z_1, z_2 \in \mathbb{R}^d$,

$$|g(t, y_1, z_1) - g(t, y_2, z_2)| \le C(|y_1 - y_2| + |z_1 - z_2|); \tag{3.6}$$

(A2) $g(\cdot, y, z) \in L^2_{\mathcal{F}}(T, \mathbb{R})$ for each $y \in \mathbb{R}$, $z \in \mathbb{R}^d$;

(A3) $g(\cdot, y, 0) \equiv 0$ for each $y \in \mathbb{R}$.

Note that the Lipschitz continuity in assumption (A1) is equivalent to the following: there exists a constant $C' > 0$ such that for all $y_1, y_2 \in \mathbb{R}$ and $z_1, z_2 \in \mathbb{R}^d$,

$$|g(t, y_1, z_1) - g(t, y_2, z_2)|^2 \le C'(|y_1 - y_2|^2 + |z_1 - z_2|^2).$$

To see this, we square Equation (3.6) and note that we have $(a + b)^2 \le 2(a^2 + b^2)$ for $a, b \ge 0$. Therefore

$$\begin{aligned} |g(t, y_1, z_1) - g(t, y_2, z_2)|^2 &\le C^2(|y_1 - y_2| + |z_1 - z_2|)^2 \\ &\le C^2(2|y_1 - y_2|^2 + 2|z_1 - z_2|^2) \\ &= C'(|y_1 - y_2|^2 + |z_1 - z_2|^2). \end{aligned}$$

Also note that assumption (A3) is equivalent to having

$$g(\cdot, \mathbf{1}_A y, \mathbf{1}_A z) \equiv g(\cdot, y, \mathbf{1}_A z) \equiv \mathbf{1}_A g(\cdot, y, z) \tag{3.7}$$

for all $(y, z) \in \mathbb{R} \times \mathbb{R}^d$ and $A \in \mathcal{F}_t$. We see this by examining the individual functions in (3.7). We have that

$$g(\cdot, \mathbf{1}_A y, \mathbf{1}_A z) = \begin{cases} g(\cdot, y, z) & \text{for } \omega \in A \\ g(\cdot, 0, 0) & \text{for } \omega \notin A. \end{cases}$$

Also

$$g(\cdot, y, \mathbf{1}_A z) = \begin{cases} g(\cdot, y, z) & \text{for } \omega \in A \\ g(\cdot, y, 0) & \text{for } \omega \notin A, \end{cases}$$

and

$$\mathbf{1}_A g(\cdot, y, z) = \begin{cases} g(\cdot, y, z) & \text{for } \omega \in A \\ 0 & \text{for } \omega \notin A. \end{cases}$$

These three are equivalent if and only if $g(\cdot, y, 0) \equiv 0$ for each $y \in \mathbb{R}$. Assumptions (A1), (A2) and (A3) are the usual ones when working with g-expectations and will be assumed throughout this chapter, unless otherwise stated. A typical example of such a g, satisfying assumptions (A1), (A2) and (A3), is $g(t, y, z) = \mu|z|$ where $\mu \in \mathbb{R}$. This function will be of particular importance later on.

Additional assumptions which are often imposed are listed below. When any of these assumptions is required for a specific result, they will be explicitly stated.

(B1) $g(t, y, z)$ is continuous in t for each $y \in \mathbb{R}$ and $z \in \mathbb{R}^d$;

(B2) g does not depend on y;

(B3) g is positive homogeneous in (y, z), i.e. for all $t \in [0, T]$, $\alpha \geq 0$ and $(y, z) \in \mathbb{R} \times \mathbb{R}^d$, we have
$$g(t, \alpha y, \alpha z) = \alpha g(t, y, z); \tag{3.8}$$

(B4) g is subadditive in (y, z), i.e. for all $t \in [0, T]$ and $(y_0, z_0), (y_1, z_1) \in \mathbb{R} \times \mathbb{R}^d$, we have
$$g(t, y_0 + y_1, z_0 + z_1) \leq g(t, y_0, z_0) + g(t, y_1, z_1); \tag{3.9}$$

(B5) g is convex in (y, z), i.e. for all $t \in [0, T]$, $(y_0, z_0), (y_1, z_1) \in \mathbb{R} \times \mathbb{R}^d$ and $\alpha \in [0, 1]$, we have
$$g(t, \alpha y_0 + (1-\alpha) y_1, \alpha z_0 + (1-\alpha) z_1) \leq \alpha g(t, y_0, z_0) + (1-\alpha) g(t, y_1, z_1). \tag{3.10}$$

The following two propositions will be stated without proof. The proof of the first proposition can be found in Pardoux and Peng [38].

Proposition 3.6. *Let assumptions (A1) and (A2) hold true for g. Then there exists a unique pair of processes $(Y_t, Z_t)_{t\in[0,T]} \in L^2_{\mathcal{F}}(T,\mathbb{R}) \times L^2_{\mathcal{F}}(T,\mathbb{R}^d)$ that solves the BSDE* (3.3) *with terminal condition* (3.4).

The solution to (3.3) with final condition (3.4) will also be denoted by $(Y_t(g,T,X), Z_t(g,T,X))_{t\in[0,T]}$.

Proposition 3.7. *Let assumptions (A1) and (A2) hold true for the function g and let $X_1, X_2 \in L^2(\mathcal{F}_T)$ and $(\phi^1_t), (\phi^2_t) \in L^2_{\mathcal{F}}(T,\mathbb{R})$ be given. Furthermore, let $(Y^{(i)}_t, Z^{(i)}_t)_{t\in[0,T]} \in L^2_{\mathcal{F}}(T,\mathbb{R}) \times L^2_{\mathcal{F}}(T,\mathbb{R}^d)$ for $i = 1, 2$ be solutions of*

$$-dY^{(i)}_t = [g(t, Y^{(i)}_t, Z^{(i)}_t) + \phi^i_t]\,dt - Z^{(i)}_t\,dB_t, \qquad 0 \le t \le T$$

$$Y^{(i)}_T = X_i.$$

Then we have the following 'continuous dependence property':

$$\sup_{0\le t\le T} \mathbb{E}[|Y^{(1)}_t - Y^{(2)}_t|^2] + \mathbb{E}\int_0^T |Z^{(1)}_s - Z^{(2)}_s|^2 ds \le C\mathbb{E}[|X_1 - X_2|^2] + C\mathbb{E}\int_0^T |\phi^1_s - \phi^2_s|^2 ds.$$

The previous proposition as well as the following theorem have been taken from Peng [40].

Theorem 3.8 (Comparison theorem)**.** *We suppose the same assumptions as in Proposition 3.7. If we furthermore assume that*

$$Y^{(1)}_t \ge Y^{(2)}_t \text{ a.s. and } \phi^1_t \ge \phi^2_t \text{ a.s., a.e.} \tag{3.11}$$

then for each $t \in [0,T]$, we have

$$Y^{(1)}_t \ge Y^{(2)}_t \text{ a.s.} \tag{3.12}$$

Also, we have

$$Y^{(1)}_t = Y^{(2)}_t \text{ a.s. if and only if } Y^{(1)}_T = Y^{(2)}_T \text{ a.s. and } \phi^1_t = \phi^2_t \text{ a.s., a.e.} \tag{3.13}$$

Proof. For simplicity we will only prove the case where $t = 0$, $d = 1$ and $\phi^1_t \equiv \phi^2_t$. Since

$$dY^{(1)}_t = -[g(t, Y^{(1)}_t, Z^{(1)}_t) + \phi^1_t]\,dt + Z^{(1)}_t\,dB_t$$

and

$$-dY^{(2)}_t = [g(t, Y^{(2)}_t, Z^{(2)}_t) + \phi^2_t]\,dt - Z^{(2)}_t\,dB_t$$

we can clearly see that

$$d(Y^{(1)}_t - Y^{(2)}_t) = -[g(t, Y^{(1)}_t, Z^{(1)}_t) - g(t, Y^{(2)}_t, Z^{(2)}_t)]\,dt + (Z^{(1)}_t - Z^{(2)}_t)\,dB_t.$$

Consequently we have

$$d(Y_t^{(1)} - Y_t^{(2)}) = -[a_t(Y_t^{(1)} - Y_t^{(2)}) + b_t(Z_t^{(1)} - Z_t^{(2)})]\, dt + (Z_t^{(1)} - Z_t^{(2)})\, dB_t,$$

where

$$a_t := \frac{g(t, Y_t^{(1)}, Z_t^{(1)}) - g(t, Y_t^{(2)}, Z_t^{(1)})}{Y_t^{(1)} - Y_t^{(2)}} \mathbf{1}_{\{Y_t^{(1)} \neq Y_t^{(2)}\}}$$

and

$$b_t := \frac{g(t, Y_t^{(2)}, Z_t^{(1)}) - g(t, Y_t^{(2)}, Z_t^{(2)})}{Z_t^{(1)} - Z_t^{(2)}} \mathbf{1}_{\{Z_t^{(1)} \neq Z_t^{(2)}\}}.$$

Since g is uniformly Lipschitz, we have

$$|g(t, Y_t^{(1)}, Z_t^{(1)}) - g(t, Y_t^{(2)}, Z_t^{(1)})| \leq C|Y_t^{(1)} - Y_t^{(2)}|$$

and

$$|g(t, Y_t^{(2)}, Z_t^{(1)}) - g(t, Y_t^{(2)}, Z_t^{(2)})| \leq C|Z_t^{(1)} - Z_t^{(2)}|,$$

giving

$$|a_t| \leq C \text{ and } |b_t| \leq C.$$

We now set

$$Q_t := e^{\int_0^t b_s\, dB_s - \frac{1}{2}\int_0^t |b_s|^2\, ds} \tag{3.14}$$

and apply Itô's formula to

$$(Y_t^{(1)} - Y_t^{(2)})Q_t e^{\int_0^t a_s\, ds}.$$

We know Q_t, $e^{\int_0^t a_s\, ds}$ and $Y_t^{(1)} - Y_t^{(2)}$ respectively satisfy

$$\begin{aligned}
&dQ_t = Q_t b_t dB_t, \\
&d(e^{\int_0^t a_s\, ds}) = a_t e^{\int_0^t a_s\, ds} dt, \\
&d(Y_t^{(1)} - Y_t^{(2)}) = -[a_t(Y_t^{(1)} - Y_t^{(2)}) + b_t(Z_t^{(1)} - Z_t^{(2)})]\, dt + (Z_t^{(1)} - Z_t^{(2)})\, dB_t.
\end{aligned}$$

Applying Itô's product rule, first to $Q_t e^{\int_0^t a_s\, ds}$ and then to $(Y_t^{(1)} - Y_t^{(2)})Q_t e^{\int_0^t a_s\, ds}$, we get

$$d(Q_t e^{\int_0^t a_s\, ds}) = Q_t e^{\int_0^t a_s\, ds}[a_t dt + b_t dB_t]$$

and

$$\begin{aligned}
d[(Y_t^{(1)} - Y_t^{(2)})Q_t e^{\int_0^t a_s\, ds}] &= (Y_t^{(1)} - Y_t^{(2)}) d(Q_t e^{\int_0^t a_s\, ds}) + Q_t e^{\int_0^t a_s\, ds} d(Y_t^{(1)} - Y_t^{(2)}) \\
&\quad + d(Y_t^{(1)} - Y_t^{(2)}) d(Q_t e^{\int_0^t a_s\, ds}) \\
&= Q_t e^{\int_0^t a_s\, ds}[(Y_t^{(1)} - Y_t^{(2)})b_t + (Z_t^{(1)} - Z_t^{(2)})]dB_t.
\end{aligned}$$

From this we know that $(Y_t^{(1)} - Y_t^{(2)})Q_t e^{\int_0^t a_s\,ds}$ is a local martingale giving

$$Y_0^1 - Y_0^2 = \mathbb{E}[(Y_t^{(1)} - Y_t^{(2)})Q_T e^{\int_0^T a_s\,ds}].$$

We note here that

$$\begin{aligned} &Q_0 = 1, \text{ and} \\ &e^{\int_0^0 a_s\,ds} = 1. \end{aligned}$$

Since $|a_t| \leq C$ for all $t \in [0,T]$, we have $e^{\int_0^T a_s\,ds} \geq e^{-Ct}$, giving

$$Y_0^1 - Y_0^2 \geq e^{-Ct}\mathbb{E}[(Y_T^{(1)} - Y_T^{(2)})Q_T] = e^{-Ct}\mathbb{E}^{\mathbb{Q}}[(Y_T^{(1)} - Y_T^{(2)})].$$

Here $\mathbb{E}[\,\cdot\,]$ denotes the classical mathematical expectation under the probability measure $\mathbb{P}$ and $\mathbb{E}^{\mathbb{Q}}[\,\cdot\,]$ denotes the classical mathematical expectation under the probability measure given by

$$\mathbb{Q}(A) = \int_A Q_T(\omega)d\mathbb{P}(\omega) \quad \text{for all } A \in \mathcal{F}.$$

Thus $Y_T^{(1)} \geq Y_T^{(2)}$ implies $Y_0^1 \geq Y_0^2$. Furthermore, since the probability measures $\mathbb{P}[\,\cdot\,]$ and $\mathbb{Q}[\,\cdot\,]$ are mutually absolutely continuous, it follows that

$$Y_0^{(1)} = Y_0^{(2)} \quad \text{if and only if} \quad Y_T^{(1)} = Y_T^{(2)} \text{ a.s.}$$

□

3.4 g-Expectations

We can now introduce g-expectations and conditional g-expectations as defined by Peng [40] and outline some properties of the g-expectation. We assume that the function g satisfies assumptions (A1), (A2) and (A3).

Definition 3.9. For any $X \in L^2(\mathcal{F}_T)$, let $(Y_t, Z_t)_{t\in[0,T]} \in L^2_{\mathcal{F}}(T,\mathbb{R}) \times L^2_{\mathcal{F}}(T,\mathbb{R}^d)$ be the solution of the BSDE (3.3) with terminal condition X. The *g-expectation* $\mathcal{E}_g$ of X is defined by

$$\mathcal{E}_g[X] = Y_0.$$

Note that Y_0 is a deterministic number that depends on the final condition X. When $g(t,y,z) = \mu|z|$ where $\mu > 0$, we will denote $\mathcal{E}_g[\,\cdot\,]$ by $\mathcal{E}^\mu[\,\cdot\,]$ and $\mathcal{E}_g[\,\cdot\,|\mathcal{F}_t]$ by $\mathcal{E}^\mu[\,\cdot\,|\mathcal{F}_t]$.

Even though in general g-expectations are not linear, the same basic properties as for the classical, linear mathematical expectation are preserved.

Proposition 3.10 (Properties of g-expectation)**.**

(i) *$\mathcal{E}_g[c] = c$ for all $c \in \mathbb{R}$. In particular, we have $\mathcal{E}_g[0] = 0$ and $\mathcal{E}_g[1] = 1$.*

(ii) *If $X_1 \geq X_2$ a.s., then $\mathcal{E}_g[X_1] \geq \mathcal{E}_g[X_2]$. In this case $\mathcal{E}_g[X_1] = \mathcal{E}_g[X_2]$ if and only if $X_1 = X_2$ a.s..*

(iii) *There exists a constant $C > 0$ such that for all $X_1, X_2 \in L^2(\mathcal{F}_t)$, we have*

$$|\mathcal{E}_g[X_1] - \mathcal{E}_g[X_2]|^2 \leq C\mathbb{E}[|X_1 - X_2|^2].$$

Proof. (i) Due to assumption (A3), when $X = c$, the solution of Equation (3.3) $(Y_t, Z_t)_{t\in[0,T]}$ is identically equal to $(c, 0)$. Thus we have that $\mathcal{E}_g[c] = c$. The particular cases follow by setting $c = 0$ and $c = 1$.

(ii) This is a direct consequence of the comparison theorem, theorem 3.8.

(iii) This is a direct consequence of the 'continuous dependence property' in Proposition 3.7. □

From the definition of g-expectation, Peng [40] introduces the concept of conditional g-expectation of a random variable $X \in L^2(\mathcal{F}_T)$ under $\mathcal{F}_t$. Analoguously to the definition of classical conditional expectation, we are looking for a random variable η satisfying the following properties:

1. η is $\mathcal{F}_t$-measurable, $\eta \in L^2(\mathcal{F}_t)$;
2. $\mathcal{E}_g[\mathbf{1}_A X] = \mathcal{E}_g[\mathbf{1}_A \eta]$ for all $A \in \mathcal{F}_t$.

Definition 3.11. For any $X \in L^2(\mathcal{F}_T)$, let $(Y_t, Z_t)_{t\in[0,T]} \in L^2_{\mathcal{F}}(T, \mathbb{R}) \times L^2_{\mathcal{F}}(T, \mathbb{R}^d)$ be the solution of the BSDE (3.3) with terminal condition X. The *conditional g-expectation* $\mathcal{E}_g$ of X for any $t \in [0, T]$ is defined by

$$\mathcal{E}_g[X|\mathcal{F}_t] = Y_t.$$

Proposition 3.12. *$\mathcal{E}_g[X|\mathcal{F}_t]$ is the unique random variable η in $L^2(\mathcal{F}_t)$ such that for all $A \in \mathcal{F}_t$*

$$\mathcal{E}_g[\mathbf{1}_A X] = \mathcal{E}_g[\mathbf{1}_A \eta]. \tag{3.15}$$

Proof. Uniqueness: Suppose η_1, $\eta_2 \in \mathcal{F}_t$ and both satisfy Equation (3.15), then we have that for all $A \in \mathcal{F}_t$

$$\mathcal{E}_g[\mathbf{1}_A \eta_1] = \mathcal{E}_g[\mathbf{1}_A \eta_2].$$

In particular, we have that $\{\eta_1 \geq \eta_2\}$ and $\{\eta_1 \leq \eta_2\}$ are $\mathcal{F}_t$-measurable and thus setting $A = \{\eta_1 \geq \eta_2\}$ or $A = \{\eta_1 \leq \eta_2\}$ we respectively get

$$\mathcal{E}_g[1_{\{\eta_1\geq\eta_2\}}\eta_1] = \mathcal{E}_g[1_{\{\eta_1\geq\eta_2\}}\eta_2],$$
$$\mathcal{E}_g[1_{\{\eta_1\leq\eta_2\}}\eta_1] = \mathcal{E}_g[1_{\{\eta_1\leq\eta_2\}}\eta_2].$$

But

$$\mathbf{1}_{\{\eta_1 \geq \eta_2\}}\eta_1 \geq \mathbf{1}_{\{\eta_1 \geq \eta_2\}}\eta_2,$$
$$\mathbf{1}_{\{\eta_1 \leq \eta_2\}}\eta_1 \leq \mathbf{1}_{\{\eta_1 \leq \eta_2\}}\eta_2.$$

From these above relations and from the monotonicity property of $\mathcal{E}_g[\,\cdot\,]$, Proposition 3.10 (ii), we have that

$$\mathbf{1}_{\{\eta_1 \geq \eta_2\}}\eta_1 = \mathbf{1}_{\{\eta_1 \geq \eta_2\}}\eta_2, \text{ a.s.}$$
$$\mathbf{1}_{\{\eta_1 \leq \eta_2\}}\eta_1 = \mathbf{1}_{\{\eta_1 \leq \eta_2\}}\eta_2, \text{ a.s.}$$

Thus it follows that $\eta_1 = \eta_2$ a.s..

Existence: For each $t \geq 0$, let $T > t$ and let $X \in L^2(\mathcal{F}_T)$. Let $(Y_t, Z_t)_{t\in[0,T]} \in L^2_{\mathcal{F}}(T, \mathbb{R}) \times L^2_{\mathcal{F}}(T, \mathbb{R}^d)$ be the solution of the BSDE (3.3) with terminal condition X, then

$$Y_u = X + \int_u^T g(s, Y_s, Z_s)\, ds - \int_u^T Z_s\, dB_s, \qquad 0 \leq u \leq T.$$

By (3.7), we can write for all $A \in \mathcal{F}_t$

$$\mathbf{1}_A Y_u = \mathbf{1}_A X + \int_u^T g(s, \mathbf{1}_A Y_s, \mathbf{1}_A Z_s)\, ds - \int_u^T \mathbf{1}_A Z_s\, dB_s, \qquad t \leq u \leq T.$$

From the definition of $\mathcal{E}_g[\,\cdot\,]$ and from the above equation, noting that in this case $\mathbf{1}_A Y_t$ is deterministic, it follows that

$$\mathcal{E}_g[\mathbf{1}_A X] = \mathbf{1}_A Y_t = \mathcal{E}_g[\mathbf{1}_A Y_t].$$

Thus $\eta = Y_t$ satisfies (3.15) and is $\mathcal{F}_t$-measurable, which completes the proof. □

Similarly to Lemma 3.2 with (nonlinear) expectations, this proposition tells us that for all $A \in \mathcal{F}_t$, we have

$$\mathcal{E}_g[\mathbf{1}_A X] = \mathcal{E}_g[\mathbf{1}_A \mathcal{E}_g[X|\mathcal{F}_t]]. \tag{3.16}$$

The conditional g-expectation also preserves the essential properties of the classical conditional expectation, except for the linearity.

Proposition 3.13 (Properties of conditional g-expectation)**.**

(i) Let X be an $\mathcal{F}_t$-measurable random variable. Then $\mathcal{E}_g[X|\mathcal{F}_t] = X$. In particular, we have $\mathcal{E}_g[0|\mathcal{F}_t] = 0$ and $\mathcal{E}_g[1|\mathcal{F}_t] = 1$.

(ii) For all $X \in L^2(\mathcal{F}_T)$ and $r, t \in [0, T]$, we have $\mathcal{E}_g[\mathcal{E}_g[X|\mathcal{F}_t]|\mathcal{F}_r] = \mathcal{E}_g[X|\mathcal{F}_{t\wedge r}]$, where $t \wedge r$ denotes the minimum of t and r.

(iii) If $X_1 \geq X_2$, then $\mathcal{E}_g[X_1|\mathcal{F}_t] \geq \mathcal{E}_g[X_2|\mathcal{F}_t]$.

(iv) For all $B \in \mathcal{F}_t$, we have $\mathcal{E}_g[\mathbf{1}_B X|\mathcal{F}_t] = \mathbf{1}_B \mathcal{E}_g[X|\mathcal{F}_t]$.

Proof. (i) This follows directly from the definition of the conditional g-expectation.

(ii) We only consider the case where $t > r$ since if $t \leq r$, we can apply (i) and the result follows easily. For $t > r$ we have that for each $A \in \mathcal{F}_r$, we also have $A \in \mathcal{F}_t$. Thus applying Equation (3.16) twice, we get

$$\mathcal{E}_g[\mathbf{1}_A \mathcal{E}_g[\mathcal{E}_g[X|\mathcal{F}_t]|\mathcal{F}_r]] = \mathcal{E}_g[\mathbf{1}_A X].$$

Now $\mathcal{E}_g[X|\mathcal{F}_t]$ is the unique random variable η in $L^2(\mathcal{F}_t)$ such that $\mathcal{E}_g[\mathbf{1}_A X] = \mathcal{E}_g[\mathbf{1}_A \eta]$, thus by Proposition 3.12, it follows that

$$\begin{aligned} \mathcal{E}_g[\mathcal{E}_g[X|\mathcal{F}_t]|\mathcal{F}_r] &= \mathcal{E}_g[X|\mathcal{F}_r] \\ &= \mathcal{E}_g[X|\mathcal{F}_{t\wedge r}]. \end{aligned}$$

(iii) Set $\eta_i := \mathcal{E}_g[X_i|\mathcal{F}_t]$, $i = 1, 2$. Since $X_1 \geq X_2$, we know by Equation (3.16) and from Proposition 3.10 (ii) that for each $A \in \mathcal{F}_t$,

$$\mathcal{E}_g[\mathbf{1}_A \eta_1] = \mathcal{E}_g[\mathbf{1}_A X_1] \geq \mathcal{E}_g[\mathbf{1}_A X_2] = \mathcal{E}_g[\mathbf{1}_A \eta_2].$$

In particular, we know that $\{\eta_1 \leq \eta_2\}$ is $\mathcal{F}_t$-measurable and thus setting $A = \{\eta_1 \leq \eta_2\}$, we get $\mathcal{E}_g[\mathbf{1}_{\{\eta_1 \leq \eta_2\}} \eta_1] \geq \mathcal{E}_g[\mathbf{1}_{\{\eta_1 \leq \eta_2\}} \eta_2]$. It is clear that $\mathbf{1}_{\{\eta_1 \leq \eta_2\}} \eta_1 \leq \mathbf{1}_{\{\eta_1 \leq \eta_2\}} \eta_2$. Again by Proposition 3.10 (ii), the above two relations imply that $\mathbf{1}_{\{\eta_1 \leq \eta_2\}} \eta_1 = \mathbf{1}_{\{\eta_1 \leq \eta_2\}} \eta_2$ which consequently implies that $\eta_1 \geq \eta_2$ i.e. $\mathcal{E}_g[X_1|\mathcal{F}_t] \geq \mathcal{E}_g[X_2|\mathcal{F}_t]$.

(iv) For each $A \in \mathcal{F}_t$

$$\begin{aligned} \mathcal{E}_g[\mathbf{1}_A \mathcal{E}_g[\mathbf{1}_B X|\mathcal{F}_t]] &= \mathcal{E}_g[\mathbf{1}_A \mathbf{1}_B X] \\ &= \mathcal{E}_g[\mathbf{1}_A \{\mathbf{1}_B \mathcal{E}_g[X|\mathcal{F}_t]\}]. \end{aligned}$$

Now from Proposition 3.12 applying to conditional g-expectations, we get that $\mathcal{E}_g[\mathbf{1}_B X|\mathcal{F}_t] = \mathbf{1}_B \mathcal{E}_g[X|\mathcal{F}_t]$ and consequently we have proved (iv). □

Having seen the similarities between the classical mathematical expectation and the g-expectation, a natural question which Briand et al. [3] pose is the following: if $X \in L^2(\mathcal{F}_T)$ is independent of $\mathcal{F}_t$, do we have that $\mathcal{E}_g[X|\mathcal{F}_t] = \mathcal{E}_g[X]$? Briand et al. [3] proved the following proposition.

Proposition 3.14. *Let g satisfy assumptions (A1), (A2) and (A3) and let g be a deterministic function. Also let $X \in L^2(\mathcal{F}_T)$. If X is independent of $\mathcal{F}_t$, then $\mathcal{E}_g[X|\mathcal{F}_t] = \mathcal{E}_g[X]$.*

Proof. By Proposition 3.13 (ii) we know that $\mathcal{E}_g[X] = \mathcal{E}_g[\mathcal{E}_g[X|\mathcal{F}_t]]$ and by the constancy of the g-expectation, Proposition 3.10 (i), we also know that for all $c \in \mathbb{R}$, $\mathcal{E}_g[c] = c$. It is hence sufficient to show that $\mathcal{E}_g[X|\mathcal{F}_t]$ is deterministic.

To prove that $\mathcal{E}_g[X|\mathcal{F}_t]$ is deterministic, we use the shift method. Let $0 \leq s \leq T - t$, and let $B'_s = B_{t+s} - B_t$. Then $\{B'_s, 0 \leq s \leq T-t\}$ is a Brownian motion with respect to its filtration $\mathcal{F}'_s$, which is the σ-algebra generated by the increments of the Brownian motion after time t. Now X is $\mathcal{F}_T$-measurable and independent of $\mathcal{F}_t$, and hence it is measurable with respect to $\mathcal{F}'_{T-t}$. We can thus construct the solution $(Y'_s, Z'_s)_{s\in[0,T-t]}$ of the BSDE

$$Y'_s = X + \int_s^{T-t} g(t+u, Y'_u, Z'_u)\, du - \int_s^{T-t} Z'_u\, dB'_u, \qquad 0 \leq s \leq T-t.$$

Setting $s = v - t$, we get

$$Y'_{v-t} = X + \int_{v-t}^{T-t} g(t+u, Y'_u, Z'_u)\, du - \int_{v-t}^{T-t} Z'_u\, dB'_u, \qquad t \leq v \leq T.$$

Making the change of variable $r = t + u$ in the integrals, we get

$$\begin{aligned} Y'_{v-t} &= X + \int_v^T g(r, Y'_{r-t}, Z'_{r-t})\, dr - \int_v^T Z'_{r-t}\, dB'_{r-t}, \qquad t \leq v \leq T \\ &= X + \int_v^T g(r, Y'_{r-t}, Z'_{r-t})\, dr - \int_v^T Z'_{r-t}\, dB_r, \qquad t \leq v \leq T. \end{aligned} \tag{3.17}$$

We know that $(Y_v, Z_v)_{v\in[t,T]}$ is the solution to the BSDE

$$Y_v = X + \int_v^T g(s, Y_s, Z_s)\, ds - \int_v^T Z_s\, dB_s, \qquad t \leq v \leq T. \tag{3.18}$$

From Equation (3.17), we see that $(Y'_{v-t}, Z'_{v-t})_{v\in[t,T]}$ is also a solution of the BSDE (3.18). Hence by uniqueness, we have that $(Y'_{v-t}, Z'_{v-t})_{v\in[t,T]} = (Y_v, Z_v)_{v\in[t,T]}$. In particular, setting $v = t$, gives us $Y'_0 = Y_t = \mathcal{E}_g[X|\mathcal{F}_t]$. Since Y'_0 is deterministic, $\mathcal{E}_g[X|\mathcal{F}_t]$ is deterministic, which completes the proof.

□

The definition of the conditional g-expectation and the previous propositions lead to the following obvious lemma.

Lemma 3.15. *Consider the function g. The related g-expectation $\mathcal{E}_g[\,\cdot\,]$ is an $\mathcal{F}_t$-consistent expectation.*

We now consider two specific cases of the function g. Firstly we will show that the classical mathematical expectation $\mathbb{E}[\,\cdot\,]$ corresponds to the case of $g = 0$. Subsequently we state and prove the lemma known as the risk aversion property of the g-expectation, $\mathcal{E}_g[\,\cdot\,]$, shown by Peng [40]. Secondly, we consider the case of $g = b_s z$, where $(b_s)_{s\in[0,\infty)}$ is a uniformly bounded and $\mathcal{F}_t$-adapted process. This case shows that the concept of g-expectations contains the Girsanov transformations.

Let X be $\mathcal{F}_T$-measurable. Let $g = 0$ and consider the following BSDE

$$dY_t = 0dt + Z_t dB_t, \qquad 0 \le t \le T$$

with boundary condition

$$Y_T = X.$$

Integrating over the interval $[t, T]$, we get

$$Y_t = X - \int_t^T Z_s\, dB_s.$$

Now Y_t is an $\mathcal{F}_t$-measurable random variable, thus taking conditional expectation and noting that the expectation of an Itô integral is 0, we get

$$\mathcal{E}_g[X|\mathcal{F}_t] = Y_t = \mathbb{E}[X|\mathcal{F}_t].$$

In particular, setting $t = 0$, we get that

$$\mathcal{E}_g[X] = Y_0 = \mathbb{E}[X].$$

This provides an alternative explanation for the classical mathematical conditional expectation. In the framework of a Brownian filtration, conditional expectations with respect to $\mathcal{F}_t$ are solutions of simple BSDEs. The classical mathematical expectation is the value of this solution when $t = 0$.

Lemma 3.16 (Risk aversion)**.** *Assume that for all (t, y, z),*

$$g(t, y, z) \le 0,$$

then

$$\mathcal{E}_g[X|\mathcal{F}_t] \le \mathbb{E}[X|\mathcal{F}_t].$$

In particular, when $t = 0$, we have,

$$\mathcal{E}_g[X] \le \mathbb{E}[X].$$

Proof. Let X be $\mathcal{F}_T$-measurable and consider the following BSDE

$$d\hat{Y}_t = \hat{Z}_t dB_t, \qquad 0 \le t \le T \tag{3.19}$$

$$\hat{Y}_T = X. \tag{3.20}$$

The solution of this BSDE is $\hat{Y}_t := \mathbb{E}[X|\mathcal{F}_t]$. From the comparison theorem, Theorem 3.8, we have that $Y_t \le \hat{Y}_t$ which gives us that $\mathcal{E}_g[X|\mathcal{F}_t] \le \mathbb{E}[X|\mathcal{F}_t]$ and completes the proof. The particular case follows by setting $t = 0$. $\square$

This risk aversion property gives a connection between g-expectations and risk measures. Clearly the choice of g is important in the construction of the risk measure. Risk measures, as well as the choice and interpretation of the function g in connection with risk measures, will be discussed in Chapter 5.

We now consider the case $g(s,z) = b_s z$, where $(b_s)_{s\in[0,\infty]}$ is a uniformly bounded and $\mathcal{F}_t$-adapted process. We let X be $\mathcal{F}_T$-measurable and let

$$Q_t := \exp\left[\int_0^t b_s dB_s - \frac{1}{2}\int_0^t |b_s|^2 ds\right].$$

It is well known that Q_t is the solution to the Itô equation

$$\begin{aligned} dQ_s &= b_s Q_s dB_s, \text{ with} \\ Q_0 &= 1. \end{aligned}$$

We also know that $Y_s = \mathcal{E}_g[X|\mathcal{F}_s]$ is the solution to the backward equation

$$\begin{aligned} dY_s &= -b_s Z_s ds + Z_s dB_s, \\ Y_T &= X. \end{aligned}$$

Applying Itô's formula to $Y_s Q_s$, we get

$$\begin{aligned} d(Y_s Q_s) &= Y_s b_s Q_s dB_s - Q_s b_s Z_s ds + Q_s Z_s dB_s + b_s Z_s Q_s ds \\ &= Q_s(Y_s b_s + Z_s) dB_s. \end{aligned} \tag{3.21}$$

Integrating over the interval $[0,T]$, yields

$$Y_T Q_T = Y_0 + \int_0^T Q_s(Y_s b_s + Z_s) dB_s.$$

By taking the classical mathematical expectation on both sides and noting that $Y_T = X$ and that the expectation of an Itô integral is 0, we get

$$\mathbb{E}[XQ_T] = Y_0 = \mathcal{E}_g[X].$$

We therefore have that

$$\mathcal{E}_g[X] = \mathbb{E}^{\mathbb{Q}}[X]$$

where $\mathbb{E}^{\mathbb{Q}}[\,\cdot\,]$ is the expectation under the probability measure given by

$$\mathbb{Q}(A) = \int_A Q_T(\omega) d\mathbb{P}(\omega) \quad \text{for all } A \in \mathcal{F}$$

and

$$\mathbb{E}^{\mathbb{Q}}[X] = \mathbb{E}[XQ_T].$$

This clearly shows that the concept of g-expectations contains the Girsanov transformations.

Similarly, we can find an expression for the conditional g-expectation, $\mathcal{E}_g[\,\cdot\,|\mathcal{F}_t]$ in terms of the conditional expectation, $\mathbb{E}^{\mathbb{Q}}[\,\cdot\,|\mathcal{F}_t]$, under $\mathbb{Q}$. Integrating over the interval $[t,T]$ on both sides of Equation (3.21) yields

$$Y_T Q_T = Y_t Q_t + \int_t^T Q_s(Y_s b_s + Z_s)dB_s.$$

Taking the conditional expectation with respect to the filtration $\mathcal{F}_t$ and noting that $Y_t Q_t$ is $\mathcal{F}_t$-measurable, we have that

$$\mathbb{E}[XQ_T|\mathcal{F}_t] = Y_t Q_t.$$

This gives us

$$\mathcal{E}_g[X|\mathcal{F}_t] = Y_t = \frac{1}{Q_t}\mathbb{E}[XQ_T|\mathcal{F}_t].$$

Thus we have that

$$\mathcal{E}_g[X|\mathcal{F}_t] = \mathbb{E}^{\mathbb{Q}}[X|\mathcal{F}_t]$$

where $\mathbb{E}^{\mathbb{Q}}[\,\cdot\,|\mathcal{F}_t]$ is the expectation under the probability measure $\mathbb{Q}$ and

$$\mathbb{E}^{\mathbb{Q}}[X|\mathcal{F}_t] = \frac{1}{Q_t}\mathbb{E}[XQ_T|\mathcal{F}_t].$$

This is again consistent with the Girsanov transformations.

The next theorem is due to the author. It combines the classical mathematical expectation and the g-expectation under a change of measure.

Theorem 3.17. *Let $(B_t)_{t\in[0,T]}$ be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $(\mathcal{F}_t)_{t\in[0,T]}$ be a filtration for this Brownian motion. Consider the following BSDE*

$$dY_t = -g(t, Y_t, Z_t)dt + Z_t dB_t, \qquad 0 \leq t \leq T$$

with terminal condition

$$Y_T = X,$$

where $X \in L^2(\mathcal{F}_T)$. Assume that $Z_t > 0$ or $Z_t < 0$ a.e. Define

$$\theta_t = \frac{g(t, Y_t, Z_t)}{Z_t}.$$

Also define

$$Z_t = exp\left[\int_0^t \theta_u dB_u - \frac{1}{2}\int_0^t \theta_u^2 du\right],$$

$$\tilde{B}_t = B_t + \int_0^t \theta_u du,$$

and assume that

$$\mathbb{E}\int_0^T \theta_u^2 Z_u^2 du < \infty.$$

Consider the probability measure defined by

$$\mathbb{Q}(A) = \int_A Z(\omega) d\mathbb{P}(\omega) \qquad \text{for all } A \in \mathcal{F}.$$

The g-expectation of X, $\mathcal{E}_g[X|\mathcal{F}_t]$, corresponds to the classical conditional expectation under $\mathbb{Q}$. In particular, $\mathcal{E}_g[X]$ corresponds to the classical expectation under $\mathbb{Q}$.

Proof. Under $\mathbb{Q}$, we have that

$$dY_t = Z_t d\tilde{B}_t, \qquad 0 \le t \le T$$

with terminal condition

$$Y_T = X.$$

By Girsanov's theorem, we know that $\tilde{B}_t$ is a Brownian motion under $\mathbb{Q}$. Hence, under $\mathbb{Q}$, we see that $g(t, y, z) = 0$. We know that the classical mathematical expectation $\mathbb{E}^{\mathbb{Q}}[\,\cdot\,|\mathcal{F}_t]$ corresponds to the case of $g = 0$. The particular case follows by setting $t = 0$. This completes the proof. □

We can now give a counterexample to Proposition 3.14, when g is not deterministic. This example has been taken from Briand et al. [3].

Fix $T > 0$ and let $t \in [0, T]$. Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous and bounded function. Define $g(s, z) := f(B_{s\wedge t})z$. Furthermore, let $X = B_T - B_t$. X is $\mathcal{F}_T$-measurable and independent of $\mathcal{F}_t$. We now have

$$Q_t := \exp\left[\int_0^t f(B_{s\wedge t}) dB_s - \frac{1}{2}\int_0^t f^2(B_{s\wedge t}) ds\right].$$

By the above result, we know that

$$\begin{aligned}\mathcal{E}_g[X|\mathcal{F}_t] &= \mathbb{E}^{\mathbb{Q}}[X|\mathcal{F}_t] \\ &= \mathbb{E}^{\mathbb{Q}}[B_T - B_t|\mathcal{F}_t]. \end{aligned} \tag{3.22}$$

Girsanov's theorem tells us that under the probability measure $\mathbb{Q}$, the process $\bar{B}_r$, $0 \le r \le T$ given by

$$\bar{B}_r = B_r - \int_0^r f(B_{s\wedge t}) ds$$

is a Brownian motion. Moreover,

$$\begin{aligned} B_T - B_t &= \bar{B}_T - \bar{B}_t + \int_t^T f(B_{s\wedge t}) ds \\ &= \bar{B}_T - \bar{B}_t + \int_t^T f(B_t) ds \\ &= \bar{B}_T - \bar{B}_t + (T - t) f(B_t). \end{aligned} \tag{3.23}$$

Plugging (3.23) into (3.22), it follows that

$$\begin{aligned}\mathcal{E}_g[X|\mathcal{F}_t] &= \mathbb{E}^{\mathbb{Q}}[\bar{B}_T - \bar{B}_t + (T-t)f(B_t)|\mathcal{F}_t] \\ &= (T-t)f(B_t).\end{aligned}$$

Hence, if f is not constant, $\mathcal{E}_g[B_T - B_t|\mathcal{F}_t]$ is not deterministic, giving the desired result.

Proposition 3.18. *Let g be convex in (y,z). Then for all $X, \eta \in L^2(\mathcal{F}_T)$, $t \in [0,T]$ we have*

$$\mathcal{E}_g[X+\eta|\mathcal{F}_t] \leq \mathcal{E}_g[X|\mathcal{F}_t] + \mathcal{E}_g[\eta|\mathcal{F}_t].$$

Proof. Consider $Y_t^* = \mathcal{E}_g[X|\mathcal{F}_t]$ which is the solution of

$$Y_t^* = X + \int_t^T g(s, Y_s^*, Z_s^*)ds - \int_t^T Z_s^* dB_s, \qquad 0 \leq t \leq T.$$

Also $Y_t^{**} = \mathcal{E}_g[\eta|\mathcal{F}_t]$ is the solution of

$$Y_t^{**} = \eta + \int_t^T g(s, Y_s^{**}, Z_s^{**})ds - \int_t^T Z_s^{**} dB_s, \qquad 0 \leq t \leq T.$$

Now

$$\mathcal{E}_g[X+\eta|\mathcal{F}_t] = X + \eta + \int_t^T g(s, Y_s^* + Y_s^{**}, Z_s^* + Z_s^{**})ds - \int_t^T (Z_s^* + Z_s^{**})dB_s.$$

Since g is convex in (y,z) we have that

$$\begin{aligned}\mathcal{E}_g[X+\eta|\mathcal{F}_t] \leq& X + \eta + \int_t^T g(s, Y_s^*, Z_s^*) + g(s, Y_s^{**}, Z_s^{**})ds \\ &- \int_t^T (Z_s^* + Z_s^{**})dB_s.\end{aligned}$$

Hence, it follows that

$$\mathcal{E}_g[X+\eta|\mathcal{F}_t] \leq Y_t^* + Y_t^{**},$$

giving the desired result. □

Remark 3.19. Similarly, if g be concave in (y,z), then for all $X, \eta \in L^2(\mathcal{F}_T)$, $t \in [0,T]$ we have

$$\mathcal{E}_g[X+\eta|\mathcal{F}_t] \geq \mathcal{E}_g[X|\mathcal{F}_t] + \mathcal{E}_g[\eta|\mathcal{F}_t].$$

The proof of the concave case follows identically to that in the convex case with the $\leq$ inequality replaced by a $\geq$ inequality.

Before going on to results in the y-independent case, another useful result giving the representation for generators of BSDEs, will be proved.

3.5 Representation lemma for generators of BSDEs

Recall that the solution to the BSDE

$$Y_t = X + \int_t^T g(s, Y_s, Z_s)\, ds - \int_t^T Z_s\, dB_s, \qquad 0 \leq t \leq T,$$

will also be denoted by $(Y_t(g, T, X), Z_t(g, T, X))_{t \in [0,T]}$. Consequently, assuming (A1), (A2) and (A3) hold true, we have that the g-expectation of X is given by

$$\mathcal{E}_g[X|\mathcal{F}_t] = Y_t(g, T, X).$$

In studying a converse comparison theorem for BSDEs, Briand et al. [3] showed that under (A1), (A2), assuming that $g(t, y, z)$ is continuous in t for each $y \in \mathbb{R}$ and $z \in \mathbb{R}^d$ and assuming that

$$\mathbb{E}\left[\sup_{t \in [0,T]} |g(t, 0, 0)|^2\right] < \infty,$$

we have that for any $(y, z) \in \mathbb{R} \times \mathbb{R}^d$ and $t \in [0, T)$

$$g(t, y, z) = L^2 - \lim_{\epsilon \to 0+} \frac{1}{\epsilon}[Y_t(g, t + \epsilon, y + z \cdot (B_{t+\epsilon} - B_t)) - y].$$

The notation 'L^2-' implies that the limit is taken in L^2. While studying Jensen's inequality for g-expectations, Jiang [33] got the following proposition, which gives a more general representation.

Proposition 3.20. *Let (A1) and (A2) hold true for the function g and let $1 \leq p \leq 2$. Then for any $(t, y, z) \in [0, T) \times \mathbb{R} \times \mathbb{R}^d$, the following are equivalent:*

(i) $g(t, y, z) = L^p - \lim_{\epsilon \to 0+} \frac{1}{\epsilon} [Y_t(g, t + \epsilon, y + z \cdot (B_{t+\epsilon} - B_t)) - y]$*;*

(ii) $g(t, y, z) = L^p - \lim_{\epsilon \to 0+} \mathbb{E}\left[\frac{1}{\epsilon} \int_t^{t+\epsilon} g(s, y, z) ds | \mathcal{F}_t\right]$.

Many problems on BSDEs are related to this kind of representation problem. We next state a general representation lemma for generators of BSDEs under assumptions (A1) and (A2). This generalises the result of Briand et al. [3]. The representation lemma is taken from Jiang [32].

Lemma 3.21 (Representation lemma)**.** *Let (A1) and (A2) hold true for the function g and let $1 \leq p \leq 2$. Then for any $(y, z) \in \mathbb{R} \times \mathbb{R}^d$, we have that the equality*

$$g(t, y, z) = L^p - \lim_{\epsilon \to 0+} \frac{1}{\epsilon} [Y_t(g, t + \epsilon, y + z \cdot (B_{t+\epsilon} - B_t)) - y]$$

holds for almost every $t \in [0, T)$.

If we set

$$S_y^z(g) := \left\{ t \in [0,T) \,:\, g(t,y,z) = L^1 - \lim_{\epsilon \to 0+} \frac{1}{\epsilon} \left[Y_t(g, t+\epsilon, y + z \cdot (B_{t+\epsilon} - B_t)) - y \right] \right\},$$

the representation lemma tells us that

$$\lambda([0,T] \backslash S_y^z(g)) = 0, \tag{3.24}$$

where λ denotes the Lebesgue measure.

Before being able to prove Lemma 3.21, we require some additional results. We first state the Lebesgue lemma taken from Fan and Hu [22].

Theorem 3.22 (Lebesgue lemma)**.** *Let f be a Lebesgue integrable function on the interval $[0,T]$. Then, in $[0,T]$, we have*

$$\lim_{n \to \infty} n \int_t^{t+\frac{1}{n}} |f(u) - f(t)| du = 0 \quad dt \ a.s. \tag{3.25}$$

Equivalently we can write Equation (3.25) as follows.

$$\begin{aligned} &\lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_t^{t+\epsilon} |f(u) - f(t)| du = 0 \\ \Longrightarrow \quad &\lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_t^{t+\epsilon} f(u) du - \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_t^{t+\epsilon} f(t) du = 0, \end{aligned}$$

which implies that

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_t^{t+\epsilon} f(u) du = \lim_{\epsilon \to 0} \frac{1}{\epsilon} f(t)(t + \epsilon - t),$$

giving

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_t^{t+\epsilon} f(u) du = f(t) \quad \text{dt a.s.}$$

We next give a proposition required in the proof of the representation lemma.

Proposition 3.23. *Let $q > 1$ and let $1 \leq p < q$. For any $(\psi_t)_{t \in [0,T]} \in H^q_{\mathcal{F}}(T, \mathbb{R})$ we have*

$$\psi_t = L^p - \lim_{\epsilon \to 0+} \frac{1}{\epsilon} \int_t^{t+\epsilon} \psi_s ds \qquad a.e.$$

Proof. Since $(\psi_t)_{t \in [0,T]} \in H^q_{\mathcal{F}}(T, \mathbb{R})$, Fubini's Theorem A.1 (see Theorem A.1 in the appendix) yields

$$\int_0^T \mathbb{E}\left[|\psi_t|^q\right] dt = \mathbb{E} \int_0^T |\psi_t|^q dt < \infty.$$

Thus $\int_0^T \mathbb{E}\left[|\psi_t|^q\right] dt < \infty$ a.e. for $t \in [0,T]$. By the Lebesgue lemma, we have for almost every $t \in [0,T)$

$$\lim_{\epsilon \to 0+} \frac{1}{\epsilon} \int_t^{t+\epsilon} \mathbb{E}\left[|\psi_s|^q\right] ds = \mathbb{E}\left[|\psi_t|^q\right].$$

Also since $(\psi)_{t\in[0,T]} \in H^q_{\mathcal{F}}(T,\mathbb{R})$ we have that $\left|\int_0^T \psi_t dt\right| < \infty$, a.s. By the Lebesgue lemma and by Fubini's theorem, we have

$$\lim_{\epsilon\to 0+} \frac{1}{\epsilon}\int_t^{t+\epsilon} \psi_s ds = \psi_t, \quad \text{a.s., a.e.}$$

Thus there exists a subset $S \subseteq [0,T)$ such that the Lebesgue measure of $[0,T]\backslash S$ equals 0, and for each $t \in S$, we have that

$$\lim_{\epsilon\to 0+} \frac{1}{\epsilon}\int_t^{t+\epsilon} \psi_s ds = \psi_t \quad \text{a.s.} \tag{3.26}$$

We also have that

$$\mathbb{E}\left[|\psi_t|^q\right] < \infty \text{ and } \lim_{\epsilon\to 0+} \frac{1}{\epsilon}\int_t^{t+\epsilon} \mathbb{E}\left[|\psi_s|^q\right] ds = \mathbb{E}\left[|\psi_t|^q\right].$$

For any $t \in S$, we know there exists a constant $\delta_t > 0$ such that for all $\epsilon \in (0, \delta_t]$

$$\frac{1}{\epsilon}\int_t^{t+\epsilon} \mathbb{E}\left[|\psi_s|^q\right] ds \le \mathbb{E}\left[|\psi_t|^q\right] + 1. \tag{3.27}$$

Now for any $t \in S$ and $\epsilon \in (0, \delta_t]$, set $X_t^\epsilon := \left|\frac{1}{\epsilon}\int_t^{t+\epsilon} \psi_s ds\right|$. For any a natural number $N \ge 1$, we want to integrate over $X_t^\epsilon > N$. Then for $X_t^\epsilon > N \ge 1$ and for $1 \le p < q$, we have $\frac{1}{N}X_t^\epsilon > 1$. Therefore

$$\left|\frac{1}{N}X_t^\epsilon\right|^p < \left|\frac{1}{N}X_t^\epsilon\right|^q,$$

giving us

$$|X_t^\epsilon|^p < \frac{1}{N^{q-p}}\,|X_t^\epsilon|^q. \tag{3.28}$$

By Equation (3.28), Fubini's theorem and Equation (3.27), we have

$$\begin{aligned}
\int_{\{X_t^\epsilon > N\}} \left|\frac{1}{\epsilon}\int_t^{t+\epsilon} \psi_s ds\right|^p d\mathbb{P} &\le \int_{\{X_t^\epsilon > N\}} \frac{1}{N^{q-p}} \left|\frac{1}{\epsilon}\int_t^{t+\epsilon} \psi_s ds\right|^q d\mathbb{P} \\
&\le \int_{\{X_t^\epsilon > N\}} \frac{1}{N^{q-p}} \left[\frac{1}{\epsilon}\int_t^{t+\epsilon} |\psi_s|^q ds\right] d\mathbb{P} \\
&\le \frac{1}{N^{q-p}}\mathbb{E}\left[\frac{1}{\epsilon}\int_t^{t+\epsilon} |\psi_s|^q ds\right] \\
&\le \frac{1}{N^{q-p}}\frac{1}{\epsilon}\int_t^{t+\epsilon} \mathbb{E}\left[|\psi_s|^q\right] ds \\
&\le \frac{1}{N^{q-p}}\left[\mathbb{E}\left[|\psi_t|^q\right] + 1\right] \\
&< \infty.
\end{aligned}$$

Thus $\left\{\left|\frac{1}{\epsilon}\int_t^{t+\epsilon}\psi_s ds\right|^p; \epsilon \in (0,\delta_t]\right\}$ are uniformly integrable. Combining this with Equation (3.26) we conclude that for each $t \in S$, we have that

$$\psi_t = L^p - \lim_{\epsilon\to 0+} \frac{1}{\epsilon}\int_t^{t+\epsilon} \psi_s ds.$$

This completes the proof. □

We can now prove Lemma 3.21 which is restated below.

Lemma 3.24 (Representation lemma). *Let (A1) and (A2) hold true for the function g and let $1 \leq p \leq 2$. Then for any $(y,z) \in \mathbb{R}\times\mathbb{R}^d$, we have that the equality*

$$g(t,y,z) = L^p - \lim_{\epsilon\to 0+} \frac{1}{\epsilon}\left[Y_t(g, t+\epsilon, y + z\cdot(B_{t+\epsilon} - B_t)) - y\right]$$

holds for almost every $t \in [0,T)$.

Proof. Since $(g(t,0,0))_{t\in[0,T]} \in H^2_{\mathcal{F}}(T,\mathbb{R})$ and g satisfies (A1), we know that for each $(y,z) \in \mathbb{R}\times\mathbb{R}^d$, $(g(t,y,z))_{t\in[0,T]} \in H^2_{\mathcal{F}}(T,\mathbb{R})$. Then for any $1 \leq p \leq 2$ and any $(y,z) \in \mathbb{R}\times\mathbb{R}^d$, Proposition 3.23 and the monotone convergence theorem yield for $t \in [0,T)$

$$g(t,y,z) = L^p - \lim_{\epsilon\to 0+} \mathbb{E}\left[\frac{1}{\epsilon}\int_t^{t+\epsilon} g(s,y,z)ds|\mathcal{F}_t\right] \quad \text{a.e.}$$

Thus, using the equivalence of Proposition 3.20, the result of the representation lemma follows, which completes the proof. □

Under the additional assumption of (A3), the representation lemma has an alternate formulation. To see this, we first note that

$$Y_t(g, t+\epsilon, y + z\cdot(B_{t+\epsilon} - B_t) = \mathcal{E}_g[y + z\cdot(B_{t+\epsilon} - B_t)|\mathcal{F}_t],$$

and that

$$\begin{aligned}\mathbb{E}[Y_t + Z_t\cdot(B_s - B_t)|\mathcal{F}_t] &= \mathbb{E}[Y_t|\mathcal{F}_t] + \mathbb{E}[Z_t\cdot(B_s - B_t)|\mathcal{F}_t] \\ &= Y_t + Z_t\mathbb{E}[(B_s - B_t)|\mathcal{F}_t] \\ &= Y_t.\end{aligned}$$

Now Lemma 3.21 states that

$$g(t,Y_t,Z_t) = \lim_{s\to t} \frac{Y_t(g,s,Y_t + Z_t\cdot(B_s - B_t)) - Y_t}{s-t}. \tag{3.29}$$

We can rewrite Equation (3.29) as

$$g(t,Y_t,Z_t) = \lim_{s\to t} \frac{\mathcal{E}_g[Y_t + Z_t\cdot(B_s - B_t)|\mathcal{F}_t] - \mathbb{E}[Y_t + Z_t\cdot(B_s - B_t)|\mathcal{F}_t]}{s-t},$$

which gives us the alternate formulation of the representation lemma.

3.6 The y-independent case

We now examine and state some results in the case where the generator g does not depend on the variable y.

Lemma 3.25. *Let g be a function independent of y i.e. $g = g(t, z)$. Then for all $X \in L^2(\mathcal{F}_T)$, $\eta \in L^2(\mathcal{F}_t)$, we have*

$$\mathcal{E}_g[X + \eta|\mathcal{F}_t] = \mathcal{E}_g[X|\mathcal{F}_t] + \eta. \tag{3.30}$$

Moreover, if g is continuous in t, then (3.30) *holds for any $t \in [0, T]$ if and only if g does not depend on y.*

Proof. Consider $Y_t = \mathcal{E}_g[X + \eta|\mathcal{F}_t]$ which is the solution of

$$Y_t = X + \eta + \int_t^T g(s, Z_s)ds - \int_t^T Z_s dB_s, \qquad 0 \leq t \leq T. \tag{3.31}$$

Also $Y_t^* = \mathcal{E}_g[X|\mathcal{F}_t]$ is the solution of

$$Y_t^* = X + \int_t^T g(s, Z_s^*)ds - \int_t^T Z_s^* dB_s, \qquad 0 \leq t \leq T.$$

Consequently we have that $Y_t^* + \eta$ satisfies the equation

$$Y_t^* + \eta = X + \eta + \int_t^T g(s, Z_s^*)ds - \int_t^T Z_s^* dB_s.$$

By Peng [40], we know, however, that the solution of Equation (3.31) is unique giving us

$$\mathcal{E}_g[X + \eta|\mathcal{F}_t] = Y_t = Y_t^* + \eta = \mathcal{E}_g[X|\mathcal{F}_t] + \eta.$$

In addition we assume that g is continuous in t. We now want to show that if (3.30) holds, then g does not depend on y. This is a direct application of the representation lemma for generators of BSDEs. If we pick a triple $(t, y, z) \in [0, T) \times \mathbb{R} \times \mathbb{R}^d$, then by Lemma 3.21 we have

$$\begin{aligned} g(t, y, z) &= \lim_{\epsilon \to 0+} \frac{1}{\epsilon} \left[Y_t(g, t + \epsilon, y + z \cdot (B_{t+\epsilon} - B_t)) - y\right] \\ &= \lim_{\epsilon \to 0+} \frac{1}{\epsilon} \{\mathcal{E}_g[y + z \cdot (B_{t+\epsilon} - B_t)|\mathcal{F}_t] - y\}, \end{aligned}$$

where the limit is taken in L^2. On the other hand, by hypothesis we have

$$\mathcal{E}_g[y + z \cdot (B_{t+\epsilon} - B_t)|\mathcal{F}_t] = y + \mathcal{E}_g[z \cdot (B_{t+\epsilon} - B_t)|\mathcal{F}_t]$$

which yields

$$g(t, y, z) = \lim_{\epsilon \to 0+} \frac{1}{\epsilon} \mathcal{E}_g[z \cdot (B_{t+\epsilon} - B_t)|\mathcal{F}_t].$$

Hence g does not depend on y, which completes the proof. ☐

A fundamental property of convex functions also leads to an interesting y-independence result. We will first state and prove this elementary property, adapted from Briand et al. [3].

Lemma 3.26. *Let $k : \mathbb{R}^n \to \mathbb{R}$ be a convex function and let k be bounded from above. Then k is constant.*

Proof. Since k is convex, we know that for each $\alpha \in [0,1]$ and for each y, $k(\alpha y) \leq \alpha k(y) + (1-\alpha)k(0)$. Choosing $y = \frac{x}{\alpha}$ in the previous equation, we get $k(x) \leq \alpha k(\frac{x}{\alpha}) + (1-\alpha)k(0)$. Since k is bounded from above, say by M, we have that $k(x) \leq \alpha M + (1-\alpha)k(0)$. Letting α tend towards 0 yields

$$k(x) \leq k(0). \tag{3.32}$$

Also, again using the convexity of k and setting $\alpha = \frac{1}{2}$, we get $k(\frac{1}{2}x - \frac{1}{2}x) \leq \frac{1}{2}k(x) + \frac{1}{2}k(-x)$, yielding $2k(0) \leq k(x) + k(-x)$. By Equation (3.32), we have that $2k(0) \leq k(x) + k(0)$, which gives us

$$k(0) \leq k(x). \tag{3.33}$$

Combining Equations (3.32) and (3.33), we find that

$$k(0) = k(x). \tag{3.34}$$

Thus k is constant. □

From this lemma, we can deduce that, if a function $g(t,y,z)$ satisfies assumptions (A1) and (A3) and is convex in (y,z), then g does not depend on y. Clearly if z is fixed, the function $g(t,y,z)$ is convex in y. Moreover, using the Lipschitz continuity of the function g, we have that

$$|g(t,y,z) - g(t,y,0)| \leq C(|y-y| + |z-0|).$$

From assumption (A3), we know that $g(t,y,0) = 0$, giving

$$|g(t,y,z)| \leq C(|z|).$$

Hence g is convex in y and bounded from above. Applying the previous lemma, we have that g is constant in the variable y. Thus g is independent of y, giving us the desired result.

3.7 Converse comparison theorem

A natural extension to the comparison theorem is the reverse thereof. If we have that $\mathcal{E}_{g_1}[X] \leq \mathcal{E}_{g_2}[X]$ or $\mathcal{E}_{g_1}[X|\mathcal{F}_t] \leq \mathcal{E}_{g_2}[X|\mathcal{F}_t]$, then do we necessarily have that $g_1 \leq g_2$? Also, if $\mathcal{E}_{g_1}[X] = \mathcal{E}_{g_2}[X]$, do we have $g_1 = g_2$? Chen [7] and Briand et al. [3] worked on this converse comparison theorem and found the following results.

The first theorem is taken from Chen [7] and the following two from Briand et al. [3].

Theorem 3.27. *Let assumptions (A1), (A2) and (A3) hold true for g_1 and g_2 and assume that g_1 and g_2 are continuous in t. Also assume that for all $X \in L(\mathcal{F}_T)$, $\mathcal{E}_{g_1}[X] = \mathcal{E}_{g_2}[X]$. Then $\mathbb{P}$-a.s. for all $(t, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^d$, we have $g_1(t, y, z) = g_2(t, y, z)$.*

Proof. For any $X \in L(\mathcal{F}_T)$ and $A \in L(\mathcal{F}_t)$, we have

$$\mathcal{E}_{g_2}[\mathbf{1}_A\mathcal{E}_{g_2}[X|\mathcal{F}_t]] = \mathcal{E}_{g_2}[\mathbf{1}_A X] = \mathcal{E}_{g_1}[\mathbf{1}_A X] = \mathcal{E}_{g_1}[\mathbf{1}_A\mathcal{E}_{g_1}[X|\mathcal{F}_t]] = \mathcal{E}_{g_2}[\mathbf{1}_A\mathcal{E}_{g_1}[X|\mathcal{F}_t]].$$

This follows by the tower property of the g-expectation and by hypothesis. Then by the uniqueness of the conditional g-expectation we have that

$$\mathcal{E}_{g_1}[X|\mathcal{F}_t] = \mathcal{E}_{g_2}[X|\mathcal{F}_t].$$

Set $Y_t^X = \mathcal{E}_{g_1}[X|\mathcal{F}_t] = \mathcal{E}_{g_2}[X|\mathcal{F}_t]$. By the definition of $\mathcal{E}_{g_i}[X|\mathcal{F}_t]$, $i = 1, 2$, there exists $Z_t^{(1)}$ and $Z_t^{(2)}$ such that $(Y_t^X, Z_t^{(i)})$, $i = 1, 2$, are the respective solutions to the BSDEs

$$Y_t^X = X + \int_t^T g_i(s, Y_s^X, Z_s^{(i)})\, ds - \int_t^T Z_s^{(i)}\, dB_s, \qquad i = 1, 2. \tag{3.35}$$

We get that

$$Z_t^X := Z_t^{(1)} = Z_t^{(2)} = \frac{d\left\langle Y_t^X, B_t\right\rangle}{dt}, \tag{3.36}$$

where $\left\langle Y_t^X, B_t\right\rangle$ is the quadratic variation of Y_t^X and B_t. Moreover, for any $X \in L^2(\mathcal{F}_T)$ we know that the solution of BSDE (3.35) satisfies for all $t \in [0, T]$

$$g_1(t, Y_t^X, Z_t^X) = g_2(t, Y_t^X, Z_t^X).$$

We still need to show that this equality holds for arbitrary $(t_0, a, b) \in [0, T] \times \mathbb{R} \times \mathbb{R}^d$. Let $(a, b) \in \mathbb{R} \times \mathbb{R}^d$. We need to show that there exists a random variable $X \in L^2(\mathcal{F}_T)$ such that the solution of the BSDE (3.35) satisfies

$$(Y_t^X, Z_t^X) = (a, b). \tag{3.37}$$

For any $(t_0, a, b) \in [0,T] \times \mathbb{R} \times \mathbb{R}^d$, we consider the stochastic differential equation (SDE)

$$\begin{aligned}
\bar{Y}_t &= a - \int_{t_0}^t g(s, \bar{Y}_s, b)\, ds + \int_{t_0}^t b\, dB_s, \quad t_0 \le t \le T, \\
\bar{Y}_t &= a, \quad 0 \le t \le t_0.
\end{aligned}$$

This SDE has a unique solution $(\bar{Y}_t)_{t\in[0,T]}$ for g. Choose $X := \bar{Y}_T$. Obviously $X \in L^2(\mathcal{F}_T)$. Also we have that $(\bar{Y}_t, b)_{t\in[t_0,T]}$ solves the BSDE (3.35) when $X = \bar{Y}_T$. We see this by the following. The left hand side of Equation (3.35) is

$$a - \int_{t_0}^t g(s, \bar{Y}_s, b)\, ds + \int_{t_0}^t b\, dB_s.$$

The right hand side of the equation yields

$$\begin{aligned}
X + \int_t^T g(s, Y_s^X, Z_s^{(i)})\, ds - \int_t^T Z_s^{(i)}\, dB_s &= \bar{Y}_T + \int_t^T g(s, Y_s^X, b)\, ds - \int_t^T b\, dB_s \\
&= a - \int_{t_0}^T g(s, \bar{Y}_s, b)\, ds + \int_{t_0}^T b\, dB_s \\
&\quad + \int_t^T g(s, Y_s^X, b)\, ds - \int_t^T b\, dB_s \\
&= a - \int_{t_0}^t g(s, \bar{Y}_s, b)\, ds + \int_{t_0}^t b\, dB_s.
\end{aligned}$$

We then have that the solution (Y_t^X, Z_t^X) of BSDE (3.35) satisfies

$$(Y_t^X, Z_t^X) = (\bar{Y}_t, b), \quad t_0 \le t \le T.$$

It follows from this equation and from (3.37) that $\bar{Y}_{t_0} = a$, which gives us the desired result. $\square$

Theorem 3.28. *Let assumptions (A1), (A2) and (A3) hold true for g_1 and g_2 and assume that g_1 and g_2 are continuous in t. Also assume that for all $X \in L(\mathcal{F}_T)$, $\mathcal{E}_{g_1}[X] \le \mathcal{E}_{g_2}[X]$. Then we have for all $t \in [0,T]$ and for all $(y,z) \in \mathbb{R} \times \mathbb{R}^d$ $g_1(t,y,z) \le g_2(t,y,z)$.*

Proof. We fix $(t,y,z) \in [0,T] \times \mathbb{R} \times \mathbb{R}^d$, and for large enough $n \in \mathbb{N}$, we consider $X_n = y + z \cdot (B_{t+\frac{1}{n}} - B_t)$. Let $(Y_t^{n(i)}, Z_t^{n(i)})_{t\in[0,T]}$ be the solution of the BSDE

$$Y_t^{n(i)} = X_n + \int_t^T g_i(s, Y_s^{n(i)}, Z_s^{n(i)})\, ds - \int_t^T Z_s^{n(i)}\, dB_s, \qquad 0 \le t \le T, \tag{3.38}$$

for $i = 1, 2$. We have by the representation lemma, Lemma 3.21 that

$$g(t,y,z) = \lim_{\epsilon \to 0+} \frac{1}{\epsilon} \left[Y_t(g, t+\epsilon, y + z \cdot (B_{t+\epsilon} - B_t)) - y \right],$$

or equivalently, letting $\epsilon = \frac{1}{n}$, we have

$$g(t, y, z) = \lim_{n\to\infty} n \left[Y_t(g, t + \frac{1}{n}, y + z \cdot (B_{t+\frac{1}{n}} - B_t)) - y \right],$$

where the limit is taken in L^2. Therefore we have that in L^2

$$n\{\mathcal{E}_{g_i}[X_n|\mathcal{F}_t] - y\} \to g_i(t, y, z).$$

On the other hand, since X_n is independent of $\mathcal{F}_t$, we know by Proposition 3.14 that $\mathcal{E}_{g_i}[X_n|\mathcal{F}_t] = \mathcal{E}_{g_i}[X_n]$. By hypothesis, we know that $\mathcal{E}_{g_1}[X] \leq \mathcal{E}_{g_2}[X]$ and hence

$$n\{\mathcal{E}_{g_1}[X_n|\mathcal{F}_t] - y\} \leq n\{\mathcal{E}_{g_2}[X_n|\mathcal{F}_t] - y\}.$$

Letting $n \to \infty$ on both sides, and noting that g_1 and g_2 are deterministic, we obtain $g_1(t, y, z) \leq g_2(t, y, z)$. This concludes the proof since (t, y, z) is arbitrary and both $g_1(\cdot, y, z)$ and $g_2(\cdot, y, z)$ are continuous. □

Theorem 3.29. *Let assumptions (A1), (A2) and (A3) hold true for g_1 and g_2 and assume that g_1 and g_2 are continuous in t. Also assume that for all $X \in L(\mathcal{F}_T)$, and for all $t \in [0, T]$, $\mathcal{E}_{g_1}[X|\mathcal{F}_t] \leq \mathcal{E}_{g_2}[X|\mathcal{F}_t]$. Then we have $\mathbb{P}$-a.s. for all $t \in [0, T]$ and for all $(y, z) \in \mathbb{R} \times \mathbb{R}^d$ $g_1(t, y, z) \leq g_2(t, y, z)$.*

Proof. This proof follows similarly to the previous one. We fix $(t, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^d$, and for large enough $n \in \mathbb{N}$, we consider $X_n = y + z \cdot (B_{t+\frac{1}{n}} - B_t)$. Let $(Y_t^{n(i)}, Z_t^{n(i)})_{t\in[0,T]}$ be the solution of the BSDE

$$Y_t^{n(i)} = X_n + \int_t^T g_i(s, Y_s^{n(i)}, Z_s^{n(i)})\, ds - \int_t^T Z_s^{n(i)}\, dB_s, \qquad 0 \leq t \leq T. \tag{3.39}$$

We have by the representation lemma, Lemma 3.21 that

$$g(t, y, z) = \lim_{\epsilon\to 0+} \frac{1}{\epsilon} \left[Y_t(g, t + \epsilon, y + z \cdot (B_{t+\epsilon} - B_t)) - y \right],$$

or equivalently, letting $\epsilon = \frac{1}{n}$, we have

$$g(t, y, z) = \lim_{n\to\infty} n \left[Y_t(g, t + \frac{1}{n}, y + z \cdot (B_{t+\frac{1}{n}} - B_t)) - y \right],$$

where the limit is taken in L^2. Therefore we have that in L^2

$$n\{\mathcal{E}_{g_i}[X_n|\mathcal{F}_t] - y\} \to g_i(t, y, z).$$

By hypothesis, we know that $\mathcal{E}_{g_1}[X|\mathcal{F}_t] \leq \mathcal{E}_{g_2}[X|\mathcal{F}_t]$ and hence

$$n\{\mathcal{E}_{g_1}[X_n|\mathcal{F}_t] - y\} \leq n\{\mathcal{E}_{g_2}[X_n|\mathcal{F}_t] - y\}.$$

Extracting a subsequence to get the convergence $\mathbb{P}$-a.s. and letting $n \to \infty$ on both sides we obtain that $\mathbb{P}$-a.s. $g_1(t, y, z) \leq g_2(t, y, z)$. By the continuity, we obtain that $\mathbb{P}$-a.s. for all $(t, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^d$, we have $g_1(t, y, z) \leq g_2(t, y, z)$. This completes the proof. □

If we do not assume that g_1 and g_2 are continuous in t, the above results still hold, however we note that they now hold for all $(y, z) \in \mathbb{R} \times \mathbb{R}^d$ for almost every $t \in [0, T]$. In fact, combining these results with the representation lemma, we have that for any $(y, z) \in \mathbb{R} \times \mathbb{R}^d$ and for any $t \in S_y^z(g_1) \cap S_y^z(g_2)$ the results hold $\mathbb{P}$-almost surely.

3.8 On Jensen's inequality

In this section, taken from Briand et al. [3], we examine Jensen's inequality for g-expectations and show that in general Jensen's inequality does not hold for g-expectations. This will be shown by a counterexample.

Consider the function $g : \mathbb{R} \to \mathbb{R}$ defined as follows

$$g(z) = \begin{cases} z^2 & \text{if } z \in [-1, 1]; \\ 2|z| - 1 & \text{if } |z| > 1. \end{cases}$$

Introduce $\xi := -\sigma^2 T + \sigma B_T$ for a fixed $\sigma \in (0, 1]$. We then consider the BSDE

$$Y_t = \xi + \int_t^T g(Z_s)\, ds - \int_t^T Z_s\, dB_s, \qquad 0 \leq t \leq T. \tag{3.40}$$

The solution to this BSDE is

$$(-\sigma^2 t + \sigma B_t, \sigma)_{t \in [0,T]};$$

that is $Y_t = -\sigma^2 t + \sigma B_t$ and $Z_t = \sigma$. We see this by the following. On the left hand side of Equation (3.40) we have $-\sigma^2 t + \sigma B_t$. The right hand side of the equality in Equation (3.40) yields

$$\begin{aligned} \xi + \int_t^T g(Z_s)\, ds - \int_t^T Z_s\, dB_s &= -\sigma^2 T + \sigma B_T + \int_t^T \sigma^2\, ds - \int_t^T \sigma\, dB_s \\ &= -\sigma^2 T + \sigma B_T + \sigma^2 (T - t) - \sigma (B_T - B_t) \\ &= -\sigma^2 t + \sigma B_t. \end{aligned}$$

We note here that since $\sigma \in (0, 1]$, we have $g(z) = z^2$. Hence

$$\frac{1}{2} \mathcal{E}_g[\xi | \mathcal{F}_t] = -\frac{\sigma^2}{2} t + \frac{\sigma}{2} B_t.$$

We now consider the BSDE

$$Y_t = \frac{1}{2}\xi + \int_t^T g(Z_s)\,ds - \int_t^T Z_s\,dB_s, \qquad 0 \le t \le T. \tag{3.41}$$

The solution to this BSDE is

$$\left(-\frac{\sigma^2}{2}T + \frac{\sigma}{2}B_t + \left(\frac{\sigma}{2}\right)^2 (T-t), \frac{\sigma}{2}\right)_{t\in[0,T]};$$

that is $Y_t = -\frac{\sigma^2}{2}T + \frac{\sigma}{2}B_t + \left(\frac{\sigma}{2}\right)^2 (T-t)$ and $Z_t = \frac{\sigma}{2}$. We see this by the following. On the left hand side of Equation (3.41) we have $-\frac{\sigma^2}{2}T + \frac{\sigma}{2}B_t + \left(\frac{\sigma}{2}\right)^2 (T-t)$. The right hand side of the equality in Equation (3.41) yields

$$\begin{aligned}\frac{1}{2}\xi + \int_t^T g(Z_s)\,ds - \int_t^T Z_s\,dB_s &= -\frac{\sigma^2}{2}T + \frac{\sigma}{2}B_T + \int_t^T \left(\frac{\sigma}{2}\right)^2 ds - \int_t^T \frac{\sigma}{2}\,dB_s \\ &= -\frac{\sigma^2}{2}T + \frac{\sigma}{2}B_T + \left(\frac{\sigma}{2}\right)^2 (T-t) - \frac{\sigma}{2}(B_T - B_t) \\ &= -\frac{\sigma^2}{2}T + \frac{\sigma}{2}B_t + \left(\frac{\sigma}{2}\right)^2 (T-t).\end{aligned}$$

Hence, we have

$$\mathcal{E}_g[\frac{1}{2}\xi|\mathcal{F}_t] = -\frac{\sigma^2}{2}T + \frac{\sigma}{2}B_t + \left(\frac{\sigma}{2}\right)^2 (T-t).$$

Consequently it follows that

$$\frac{1}{2}\mathcal{E}_g[\xi|\mathcal{F}_t] - \mathcal{E}_g[\frac{1}{2}\xi|\mathcal{F}_t] = (T-t)\frac{\sigma^2}{4},$$

which is positive if $t < T$. This contradicts Jensen's inequality in the simplest case. We considered the linear function $\vartheta : x \to \frac{x}{2}$ and a convex generator g and found that for $t \le T$ we have

$$\vartheta\left(\mathcal{E}_g[\xi|\mathcal{F}_t]\right) \ge \mathcal{E}_g[\vartheta\left(\xi\right)|\mathcal{F}_t].$$

Briand et al. [3] has established that Jensen's inequality does hold under certain additional assumptions on the generator g and the g-expectation $\mathcal{E}_g[\,\cdot\,]$. For more information regarding this, we refer the reader to Briand et al. [3].

3.9 Financial application of g-expectations

BSDEs are useful in finance and have been applied in the evaluation of contingent claims, especially in constrained cases, as well as in the theory of recursive utilities, researched extensively by Duffie and Epstein [20, 21]. Most notably, BSDEs appeared in the valuation of contingent claims in complete markets, studied by Black and Scholes [2] and Merton [36] amongst others. The problem posed was to determine the price of a contingent claim with payoff X and maturity T. In a complete market it was possible to construct a portfolio which replicates the payoff of the contingent claim. The dynamics of the value of the replicating portfolio are given by such a BSDE with a linear generator. The BSDE theory showed that there exists a unique price and a unique hedging portfolio which replicates the payoff of the contingent claim.

In this section, we specifically look at the financial application, interpretation and implication of Peng's g-expectation. g-Expectations give a dynamic evaluation of pricing contigent claims with payoff X and maturity T in complete and incomplete markets. As seen in Chapter 5, g-expectations can also be used to create risk measures. A significant feature of the g-expectation $\mathcal{E}_g[\,\cdot\,|\mathcal{F}_t]$ is that its value and behaviour is uniquely and entirely determined by the generator function g. When pricing contingent claims, we are concerned about finding this generator g. With risk measures, this function can be chosen to suit the risk preference of the investor and would hence indicate how conservative the risk measure is.

The work in this section has been based on El Karoui et al. [35], Rosazza Gianin [30], Peng [44], Schroder [48] and Finch [23].

3.9.1 Black-Scholes option pricing formula

We first consider the Black-Scholes framework of pricing contingent claims in complete or incomplete markets. Let $(S_t)_{t\in[0,T]}$ be the value process of a risky asset with dynamics:

$$dS_t = \nu_t S_t dt + \sigma_t S_t dB_t,$$

or equivalently

$$\frac{dS_t}{S_t} = \nu_t dt + \sigma_t dB_t.$$

In general $(S_t)_{t\in[0,T]}$ can be a vector $S_t = (S_t^1, \ldots, S_t^d)$, representing the vector value process of d risky assets and where S_t^i represents the value process of risky asset i. Let π_t denote the portion of wealth invested in the risky asset at time t. Let $(V_t)_{t\in[0,T]}$ be the wealth process of the portfolio. We now consider a contingent claim with payoff X at time T where X is a function of S_T i.e. $X = f(S_T)$. As mentioned previously, in complete markets we know that the price at time 0 of such a contingent claim is the initial value of the self-financing replicating portfolio, i.e. V_0. We consequently have

$$dV_t = \pi_t dS_t = \pi_t \nu_t S_t dt + \pi_t \sigma_t S_t dB_t, \tag{3.42}$$

with terminal condition

$$V_T = X. \tag{3.43}$$

Setting $Y_t := V_t$ and $Z_t := \pi_t \sigma_t S_t$ we get

$$dY_t = -g(t, Y_t, Z_t)dt + Z_t dB_t$$

with terminal condition

$$Y_T = X$$

and where g is given by

$$g(t, y, z) = -\frac{\nu_t}{\sigma_t} z.$$

The initial value of the replicating portfolio is then given by $V_0 = Y_0 = \mathcal{E}_g[X]$. Note that the function g is independent of y and linear in z. Also note that the function depends on the drift, ν_t, and the volatility, σ_t, of the risky asset S_t. In this case the static risk measure $\rho_g(X) = \mathcal{E}_g[-X]$ represents the initial value of the replicating portfolio of a contingent claim with payoff $-X$. This risk measure can hence be seen as the 'natural' risk measure. We also note that the expression $-\frac{\nu_t}{\sigma_t}$ resembles the market price of risk.

We know that the solution to the BSDE (3.42) with terminal condition (3.43) is given by the well known Black-Scholes option pricing formula. Hence we see here that Peng's g-expectation contains the Black-Scholes option pricing formula. In this case g is a linear function in z.

We now look at two other processes used for modelling equity prices. For each of these processes, we set up the option pricing formula using the g-expectation. The work in the following two subsections is due to the author.

3.9.2 Constant elasticity of variance option pricing formula

An alternative process used to model the dynamics of the stock price is the constant elasticity of variance process. Empirical evidence showed there exists a relationship between the stock price and the variance of the stock price returns. More precisely, studies showed that an increase in stock price may reduce the variance of the stock price return. Hence the constant elasticity of variance (CEV) model was proposed as an alternative to the Black-Scholes model. The CEV process suggests the following deterministic relationship between the stock price, S_t, and the volatility of the stock price $\sigma(S, t)$:

$$\sigma(S, t) = \delta S_t^{\frac{(\beta-2)}{2}}.$$

If $\beta < 2$ in the CEV model, then the stock price and the volatility are inversely related. Under the constant elasticity of variance model, the stock price is assumed to follow the dynamics

$$dS_t = \nu_t S_t dt + \delta S_t^{\frac{\beta}{2}} dB_t.$$

When $\beta = 2$, the stock price is lognormally distributed and the variance of the return is constant. In this case the elasticity of the variance with respect to the price, which is given by $\beta - 2$, is zero. The stock then follows geometric Brownian motion and the Black-Scholes framework applies. Again, in general $(S_t)_{t\in[0,T]}$ can be a vector $S_t = (S_t^1, \ldots, S_t^d)$. We let π_t denote the portion of wealth invested in the risky asset at time t. Let $(V_t)_{t\in[0,T]}$ be the wealth process of the portfolio. We again consider a contingent claim with payoff X at time T where X is a function of S_T i.e. $X = f(S_T)$. Consequently we have that the wealth process satisfies

$$dV_t = \pi_t dS_t = \pi_t \nu_t S_t dt + \pi_t \delta S_t^{\frac{\beta}{2}} dB_t,$$

with terminal condition

$$V_T = X.$$

In this case we set $Y_t := V_t$ and $Z_t := \pi_t \delta S_t^{\frac{\beta}{2}}$ giving us

$$dY_t = -g(t, Y_t, Z_t)dt + Z_t dB_t$$

with terminal condition

$$Y_T = X$$

and where g is given by

$$g(t, y, z) = -\nu_t z S_t^{\frac{(\beta-2)}{2}}.$$

The initial value of the replicating portfolio and hence the value of the contingent claim at time 0, is then given by $V_0 = Y_0 = \mathcal{E}_g[X]$.

3.9.3 Ornstein-Uhlenbeck option pricing formula

Mean reversion processes are widely seen in finance. They are most commonly used for modelling interest rates, however are also used for modelling currency exchanges, convenience yields, volatilities of asset prices and even commodity prices. The most popular mean reversion model is the Ornstein-Uhlenbeck process with following dynamics:

$$dX_t = \lambda(\mu - X_t)dt + \sigma dB_t.$$

The Vasicek model is a well-known example of an Ornstein-Uhlenbeck process, used in interest rate modelling. If we assume that the stock price is modelled using the Ornstein-Uhlenbeck process, we have

$$dV_t = \pi_t dX_t = \pi_t \lambda(\mu - X_t)dt + \pi_t \sigma dB_t.$$

with terminal condition

$$V_T = X.$$

Here we set $Y_t := V_t$ and $Z_t := \pi_t \sigma$ giving us

$$dY_t = -g(t, Y_t, Z_t)dt + Z_t dB_t$$

with terminal condition

$$Y_T = X$$

and where g is given by

$$g(t, y, z) = -\frac{z}{\sigma}\lambda(\mu - X_t).$$

The initial value of the replicating portfolio and hence the value of the contingent claim at time 0, is again given by $V_0 = Y_0 = \mathcal{E}_g[X]$.

Chapter 4

Risk measures

A considerable amount of research has been done on different methods of measuring the riskiness of financial positions, both in a theoretical and a practical setting. Risk measures were introduced to quantify the riskiness of any financial position. They also give an indication as to which positions carry an acceptable amount of risk and which positions do not. A well known and popular risk measure is Value at Risk (VaR). This risk measure has been of great interest in financial and mathematical research. Value at Risk, for a given time horizon and probability α, also denoted by VaR_α, is the maximum loss in market value of a financial position over the time horizon that is exceeded with a confidence of $1-\alpha$. For an extensive overview of VaR, we refer the reader to Duffie and Pan [19]. VaR, however, has quite a few drawbacks and has been subject to a lot of criticism in the literature, see for example Artzner et al. [1]. Consequently, Artzner et al. [1] introduced some desirable axioms for risk measures which led to the concept of coherent risk measures. Delbaen [15] generalised the concept of coherent risk measures to general probability spaces. Furthermore Artzner et al. [1] and Delbaen [15] proved a representation theorem for coherent risk measures.

A more general and more desirable property was then established which brought about the concept of convex risk measures, first introduced by Heath [31] in finite probability spaces and later by Föllmer and Schied [24] in general probability spaces. Fritelli and Rosazza Gianin [27] independently defined convex risk measures in general probability spaces. Fritelli and Rosazza Gianin [27] also independently to Föllmer and Schied [24], proved an analogous representation theorem for convex risk measures that generalises the representation theorem of coherent risk measures.

All risk measures mentioned thus far have been one-period risk measures. However, in many cases due to intermediate cashflows, we want to work with a multi-period risk measure. Therefore, as an extension to quantifying the risk today of a financial position at some fixed point in the future, we would like to be able to quantify the riskiness of a position at intermediate timepoints. Thus risk measures were first introduced in a dynamic setting by Cvitanic and Karatzas [14] and Wang [49]. More recent and more extensive studies on dynamic risk measures have been done by Frittelli and Rosazza Gianin [26] and Riedel [46] amongst others.

In this chapter, we recall the basic concepts of static and dynamic risk measures, as well as what is meant by coherent and convex risk measures and state the corresponding representation theorems.

4.1 Static risk measures

A large portion of the work in the next two sections has been adapted from Föllmer and Schied [25], Rosazza Gianin [30] and Frittelli and Rosazza Gianin [26].

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We let T be a fixed future date. Denote by $\mathcal{X}$ a vector subspace of the space of all measurable functions $f : \Omega \to \mathbb{R}$ which contains the constant functions. The space $\mathcal{X}$ is interpreted as the space of all financial positions which we are interested in. A financial position is a mapping $X : \Omega \to \mathbb{R}$, where $X(\omega)$ is the discounted net worth of any position at time T, the end of the trading period.

A static measure of risk is a mapping

$$\rho : \mathcal{X} \to \mathbb{R}.$$

Many properties have been imposed on risk measures which consequently led to the concepts of coherent and convex risk measures. The financial interpretation of these properties makes them desirable. We list the axioms for static risk measures below.

Axiom 4.1 (Axioms for static risk measures)**.**

(a) Positivity *: For all $X \in \mathcal{X}$ with $X \geq 0$, we require $\rho(X) \leq \rho(0)$.*

(b) Monotonicity *: For all $X, Y \in \mathcal{X}$ with $X \leq Y$, we require $\rho(Y) \leq \rho(X)$.*

(c) Subadditivity *: For all $X, Y \in \mathcal{X}$, we require $\rho(X + Y) \leq \rho(X) + \rho(Y)$.*

(d) Positive Homogeneity *: For all $\lambda \geq 0$ and $X \in \mathcal{X}$, we require $\rho(\lambda X) = \lambda\rho(X)$.*

(e) Translation Invariance *: For all $X \in \mathcal{X}$ and $\alpha \in \mathbb{R}$, we require $\rho(X + \alpha) = \rho(X) - \alpha$.*

(f) Constancy *: For all $c \in \mathbb{R}$, we require $\rho(c) = -c$.*

(g) Convexity *: For all* $\lambda \in [0,1]$ *and* $X, Y \in \mathcal{X}$*, we require* $\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y)$.

(h) Lower semi-continuity *:* $\{X \in \mathcal{X} : \rho(X) \leq \gamma\}$ *is closed in* $\mathcal{X}$ *for any* $\gamma \in \mathbb{R}$.

Definition 4.2. A *monetary measure of risk* is a mapping $\rho : \mathcal{X} \to \mathbb{R}$ which satisfies Axioms 4.1 (b) monotonicity and (e) translation invariance.

The number $\rho(X)$ represents the riskiness of the position X and can be interpreted as the minimum extra capital that needs to be added to the original risky position to make the position acceptable from the viewpoint of a supervising agency. If the risk measure is negative, the capital amount $-\rho(X)$ can be withdrawn from the position and the resulting position will remain acceptable.

The financial interpretation of both monotonicity and of translation invariance, also known as cash invariance, is clear. Monotonocity ensures that if a position always leads to a worse outcome than another position, then its riskiness is greater than the riskiness of the other position. Translation invariance guarantees that adding a sure amount to the initial position decreases the riskiness of the position by that amount. In addition this property implies that $\rho(X + \rho(X)) = 0$. This confirms the financial interpretation of a risk measure given earlier: $\rho(X)$ is the amount of money needed to add to the risky position X to make the position neutrally acceptable.

Note that in Axiom 4.1 (a) positivity, (b) monotonicity, (e) translation invariance and in (f) constancy, the risk measure inverts signs. This is due to the interpretation of ρ. Axiom 4.1 (c) subadditivity and (d) positive homogeneity are together also known as sublinearity. Positive homogeneity in particular tells us that if we increase the size of a position by λ, the riskiness of the position increases by a factor of λ. In other words, the size of the position directly influences the riskiness of the position. In many situations, however, the riskiness of a financial position may increase in a nonlinear fashion with the size of the position. If, for example, the position is multiplied by a large factor, an additional liquidity risk may arise. This suggests that the condition of positive homogeneity should be relaxed. Subadditivity is seen as a natural requirement for coherency of risk measures as it ensures that diversification of a portfolio holds. This means that combining risks into one portfolio may lead to a lower resulting risk: the loss of one position may offset the gains on other positions and the total risk may be reduced. Subadditivity also ensures that there is no motivation to break into separate affiliates if the combined risk of the positions is higher than the individual risk. As 'natural' as this condition may seem, subadditivity has also been subject to criticsm in literature. This suggests that it should also be relaxed. We elaborate on this later on in this chapter. For an extensive discussion of the financial motivation of the above axioms, we refer the reader to Artzner et al. [1] and Frittelli and Rosazza Gianin [26, 27].

When dealing with risk measures, we are concerned about having acceptable financial positions. In a mathematical sense, we mean the following when referring to an acceptable position.

Definition 4.3. The *acceptance set* associated to a risk measure, ρ, is the set, $\mathcal{A}_\rho$, defined by

$$\mathcal{A}_\rho = \{X \in \mathcal{X} \mid \rho(X) \leq 0\}.$$

In other words, a financial position X is acceptable if $\rho(X) < 0$, and unacceptable otherwise. The relationship between risk measures and their acceptance sets will be given shortly.

We require an additional result in the proof of the following proposition. We first state and prove this result; Proposition 4.5 follows.

Lemma 4.4. *Any monetary risk measure, ρ, is Lipschitz continuous with respect to the supremum norm $\|\cdot\|$, that is*

$$|\rho(X) - \rho(Y)| \leq ||X - Y||.$$

Proof. Clearly $X \leq Y + \|X - Y\|$. By monotonicity and translation invariance, noting that $\|X - Y\|$ is a constant, we have $\rho(Y) - \|X - Y\| \leq \rho(X)$. This gives us $\rho(Y) - \rho(X) \leq \|X - Y\|$. Reversing the roles of X and Y, we get $\rho(X) - \rho(Y) \leq \|Y - X\|$. Since $\|X - Y\| = \|Y - X\|$, we get $|\rho(X) - \rho(Y)| \leq \|X - Y\|$ which completes the proof. $\square$

The relationship between a monetary measure of risk and the acceptance set of the risk measure is given by the following.

Proposition 4.5. *Let ρ be a monetary measure of risk with acceptance set $\mathcal{A} = \mathcal{A}_\rho$. Then we have*

(i) $\mathcal{A}$ is non-empty and satisfies the following conditions:

$$\inf\{m \in \mathbb{R} \,|\, m \in \mathcal{A}\} > -\infty. \tag{4.1}$$

$$X \in \mathcal{A}, Y \in \mathcal{X}, Y \geq X \Rightarrow Y \in \mathcal{A}. \tag{4.2}$$

Moreover, $\mathcal{A}$ has the following closure property. Let $X \in \mathcal{A}, Y \in \mathcal{X}$,

$$\{\lambda \in [0,1] \,|\, \lambda X + (1-\lambda)Y \in \mathcal{A}\} \textit{ is closed in } [0,1].$$

(ii) The risk measure ρ can be recovered from the acceptance set $\mathcal{A}$ by

$$\rho(X) = \inf\{m \in \mathbb{R} \,|\, m + X \in \mathcal{A}\}$$

Proof. (i) The first two properties are straightforward. By Lemma 4.4, the function $\lambda \mapsto \rho(\lambda X + (1-\lambda)Y)$ is continuous. Hence the set of $\lambda \in [0,1]$ such that $\rho(\lambda X + (1-\lambda)Y) \leq 0$ is closed.

(ii) Translation invariance implies that for $X \in \mathcal{X}$,

$$\begin{aligned} \inf\{m \in \mathbb{R} \,|\, m + X \in \mathcal{A}\} &= \inf\{m \in \mathbb{R} \,|\, \rho(m + X) \leq 0\} \\ &= \inf\{m \in \mathbb{R} \,|\, \rho(X) \leq m\} \\ &= \rho(X). \end{aligned}$$

□

On the other hand, we can also consider a given class $\mathcal{A} \subset \mathcal{X}$ of acceptable positions. For any financial position X, we can define the risk measure $\rho_{\mathcal{A}}(X)$ by

$$\rho_{\mathcal{A}}(X) := \inf\{m \in \mathbb{R} \,|\, m + X \in \mathcal{A}\}.$$

Here m is again the minimal capital amount that needs to be added to the position X so that $m + X$ is acceptable. We then have that $\rho = \rho_{\mathcal{A}_\rho}$.

Proposition 4.6. *Assume that $\mathcal{A}$ is a non-empty subset of $\mathcal{X}$ which satisfies conditions* (4.1) *and* (4.2). *Then $\rho_{\mathcal{A}}$ is a monetary risk measure.*

For more information and an extensive study on the relationship between monetary risk measures and their acceptance sets, we refer the reader to Föllmer and Schied [25], and to Artzner et al. [1].

Due to the financial convenience, we concentrate on monetary risk measures with additonal properties. Risk measures which satisfy (d) positive homogeneity and (c) subadditivity, and risk measures which satisfies (g) convexity are of particular interest. These two classes of risk measures will be dealt with in the following two sections.

4.1.1 Coherent risk measures

Having stated the important properties imposed on risk measures, we can now define the concept of a coherent measure of risk in general probability spaces.

Definition 4.7. A function $\rho : \mathcal{X} \to \mathbb{R}$ is a *coherent measure of risk* if it satisfies Axiom 4.1 (b) monotonocity, (c) subadditivity, (d) positive homogeneity and (e) translation invariance.

Delbaen [15] defined and proved the general characterisation of a coherent measure of risk. Typically, any coherent measure of risk ρ arises from some family $\mathcal{P}$ of probability measures on Ω.

Theorem 4.8. *Consider the function $\rho : \mathcal{X} \to \mathbb{R}$. Then ρ is a coherent measure of risk if and only if there exists a closed convex set $\mathcal{P}$ of $\mathbb{P}$-continuous probability measures such that*

$$\rho(X) = \sup_{Q \in \mathcal{P}} \mathbb{E}_Q[-X]. \tag{4.3}$$

For a proof of this theorem, see Delbaen [15].

By this representation any coherent risk measure ρ can be represented as the maximum expected loss over a set $\mathcal{P}$ of generalised scenarios i.e. any coherent risk measure can be seen as the 'worst case method' in a framework of generalised scenarios. The more scenarios one considers, the more conservative the resulting risk measure.

The relationship between the risk measure and the acceptance set in the class of coherent risk measures will next be given, however, we first recall the definition of convexity and define the concept of a cone. The following definition has been taken from Offwood [37].

Definition 4.9.

(i) A set A is *convex* if for all $\lambda \in [0,1]$ and $x, y \in A$, we have that $\lambda x + (1-\lambda)y \in A$.

(ii) A set C in a vector space is said to be a *cone* with vertex at the origin if $x \in C$ implies that $\alpha x \in C$ for all $\alpha \geq 0$. A cone with vertex p is defined as a translation $p + C$ of a cone C with vertex at the origin. If the vertex of a cone is not explicitly mentioned then it is assumed to be the origin.

Proposition 4.10. *Suppose ρ is a monetary risk measure with acceptance set $\mathcal{A}$. Then ρ is positively homogeneous if and only if $\mathcal{A}$ is a cone. In particular, ρ is a coherent risk measure if and only if $\mathcal{A}$ is a convex cone.*

For a proof of this theorem, see Föllmer and Schied [25].

As explained previously, positive homogeneity and subadditivity are not always applicable or suitable properties for risk measures. Thus it is suggested that the sublinearity property of risk measures be relaxed. Consequently convex risk measures were introduced as a generalisation of coherent risk measures.

4.1.2 Convex risk measures

In the class of convex measures of risk the positive homogeneity and subadditivity properties are replaced with the weaker property of convexity.

Definition 4.11. A monetary risk measure $\rho : \mathcal{X} \to \mathbb{R}$ is a *convex measure of risk* if it satisfies Axiom 4.1 (g) convexity, (h) lower semi-continuity and $\rho(0) = 0$.

Remark 4.12. In the literature, the definition of convex risk measures differs depending on the author. The above definition, which will be used throughout the dissertation, is a combination of definitions by Föllmer and Schied [25] and by Rosazza Gianin [30].

We still interpret $\rho(X)$ as the minimum capital requirement, which, if added to a position, makes it acceptable. An analogous representation theorem as for coherent risk measures exists for convex risk measures.

Theorem 4.13. *The function $\rho : \mathcal{X} \to \mathbb{R}$ is a convex measure of risk if and only if there exists a convex set $\mathcal{P}$ of $\mathbb{P}$-continuous probability measures such that*

$$\rho(X) = \sup_{Q \in \mathcal{P}} \{\mathbb{E}_Q[-X] - F(\mathcal{Q})\}. \tag{4.4}$$

For a proof of this theorem, see Föllmer and Schied [25].

This representation also has a financial interpretation. By this characterisation, any convex risk measure ρ can be represented as the maximum expected loss over a set $\mathcal{P}$ of generalised scenarios corrected with a penalty term F, which depends on the scenarios. Whilst the set of possible scenarios may be exogenously determined, the penalty function F may be determined by the investor, depending on his or her own preference.

In the class of convex risk measures, the relationship between the risk measure and the acceptance set is given by the following.

Proposition 4.14. *Suppose ρ is a monetary risk measure with acceptance set $\mathcal{A}$. The risk measure ρ is convex if and only if $\mathcal{A}$ is convex.*

For a proof of this theorem, see Föllmer and Schied [25].

The risk measures discussed this far deal with quantifying the risk today of a financial position with fixed maturity date T. The natural question arises as to how we can quantify the riskiness of a financial position at different timepoints between 0 and T i.e. in a dynamic setting.

4.2 Dynamic risk measures

The risk measures proposed thus far have all been one-period risk measures and consequently do not measure the riskiness of a position at intermediate time points. In many cases, however, particularly when intermediary cashflows take place, we are interested in a multi-period framework. We are therefore now concerned about monitoring the riskiness of a financial position X at any intermediate time t between the initial time 0 and the maturity T. To define a dynamic risk measure, it seems reasonable to define a map, ρ_t, indexed by time, where $\rho_t(X)$ denotes the riskiness of the financial position X at time t conditional to the information available at time t. To define dynamic risk measures, we furthermore require boundary conditions for $\rho(X)$ at the initial time 0 and at the final time T.

We now consider a general filtration $\{\mathcal{F}_t\}_{t\geq 0}$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $L_p(\mathcal{F}_t) = L_p(\Omega, \mathcal{F}_t, \mathbb{P})$ for $p \geq 1$ denote the space of all real-valued, $\mathcal{F}_t$-measurable and p-integrable random variables. Let $L_0(\Omega, \mathcal{F}_t, \mathbb{P})$ denote the space of all $\mathcal{F}_t$-measurable random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$. We still have $\mathcal{X}$ denoting the space of all financial positions which we are interested in and T a fixed future date representing the period of uncertainty. For simplicity, we will assume that all elements of $\mathcal{X}$ are $\mathcal{F}_T$-measurable, i.e. we assume that $\mathcal{X} = L_0(\Omega, \mathcal{F}_T, \mathbb{P})$.

We now define dynamic risk measures.

Definition 4.15. A *dynamic risk measure* is any map satisfying the following conditions:

(i) $\rho_t : \mathcal{X} \to L_0(\Omega, \mathcal{F}_t, \mathbb{P})$ for all $t \in [0, T]$;

(ii) ρ_0 is a static risk measure;

(iii) $\rho_T(X) = -X$ for all $X \in \mathcal{X}$.

Analogously to static risk measures, we wish to define coherent and convex dynamic risk measures. To do this, some properties of the dynamic risk measure $(\rho_t)_{t\in[0,T]}$ will be listed.

Axiom 4.16 (Axioms for dynamic risk measures)**.**

(a) Positivity *: For all $t \in [0, T]$ and $X \in \mathcal{X}$ with $X \geq 0$, we require $\rho_t(X) \leq \rho_t(0)$.*

(b) Monotonicity *: For all $t \in [0, T]$ and $X, Y \in \mathcal{X}$ with $X \leq Y$, we require $\rho_t(Y) \leq \rho_t(X)$.*

(c) Subadditivity *: For all $X, Y \in \mathcal{X}$ and $t \in [0, T]$, we require $\rho_t(X + Y) \leq \rho_t(X) + \rho_t(Y)$.*

(d) Positive Homogeneity *: For all $t \in [0, T]$, $\lambda \geq 0$ and $X \in \mathcal{X}$, we require $\rho_t(\lambda X) = \lambda \rho_t(X)$.*

(e) Translation Invariance *: For all $t \in [0, T]$, $X \in \mathcal{X}$ and an $\mathcal{F}_t$-measurable random variable ξ in $\mathcal{X}$, we require $\rho_t(X + \xi) = \rho_t(X) - \xi$.*

(f) Constancy *: For all $c \in \mathbb{R}$ and $t \in [0,T]$, we require $\rho_t(c) = -c$.*

(g) Convexity *: For all $t \in [0,T]$, ρ_t is convex i.e. for all $\lambda \in [0,1]$ and $X, Y \in \mathcal{X}$, we require $\rho_t(\lambda X + (1-\lambda)Y) \leq \lambda\rho_t(X) + (1-\lambda)\rho_t(Y)$.*

Note that the interpretation of most of the axioms is analogous to those explained in the static case. In fact most of these axioms are identical to the axioms defined for static risk measures, however we are now working in a dynamic setting. The dynamic translation invariance axiom is however stronger than the static translation invariance axiom. In the dynamic setting, translation invariance does not only apply with respect to constants, but also with respect to any $\mathcal{F}_t$-measurable random variable. In other words, translation invariance applies with respect to a risky position that is completely determined by the information available to the market at time t.

We can now define coherent and convex risk measures in the dynamic setting.

Definition 4.17. A dynamic risk measure $(\rho_t)_{t\in[0,T]}$ is called

(i) *coherent* if it satisfies Axiom 4.16 (b) monotonicity, (c) subadditivity, (d) positive homogeneity and (e) translation invariance;

(ii) *convex* if for each t, ρ_t is a monetary risk measure and it satisfies Axiom 4.16 (g) convexity and $\rho_t(0) = 0$;

(iii) *time-consistent* if for all $t \in [0,T]$, $X \in \mathcal{X}$ and $A \in \mathcal{F}_t$,

$$\rho_0(X\mathbf{1}_A) = \rho_0[-\rho_t(X)\mathbf{1}_A]. \tag{4.5}$$

The coherent and convex risk measures are the dynamic equivalent to coherent and convex risk measures in the static sense. The new concept of time-consistency is however now introduced. The time-consistency condition gives us two approaches to quantify the riskiness of a financial position at the initial time 0:

- computing the static risk measure $\rho_0(X)$ directly;
- evaluating $\rho_0(X)$ in two steps, i.e. first evaluating the riskiness of the financial position X at an intermediate time t and then quantifying the risk of $-\rho_t(X)$ at time 0.

The negative sign in the time-consistency property is as a consequence of the financial interpretation of the risk measure $\rho_t(X)$.

As a natural extension to the representation of coherent and convex risk measures in the static sense, we present an example of a dynamic coherent risk measure and an example of a dynamic convex risk measure taken from Rosazza Gianin [30].

Example 4.18 (Dynamic coherent risk measure)**.** Let $\mathcal{P}$ be a convex set of $\mathbb{P}$-absolutely continuous probability measures defined on $(\Omega, \mathcal{F}_T)$. Then for all $X \in \mathcal{X}$ and $t \in [0, T]$,

$$\rho_t(X) = \operatorname*{ess.sup}_{Q \in \mathcal{P}} \mathbb{E}_Q[-X \mid \mathcal{F}_t] \tag{4.6}$$

is a dynamic coherent risk measure.

It is easy to see that this is a dynamic risk measure as $\rho_t : \mathcal{X} \to L_0(\Omega, \mathcal{F}_t, \mathbb{P})$ for all $t \in [0, T]$, $\rho_0 = \text{ess.sup}_{Q \in \mathcal{P}} \mathbb{E}_Q[-X]$ is a static risk measure and $\rho_T(X) = \text{ess.sup}_{Q \in \mathcal{P}} \mathbb{E}_Q[-X \mid \mathcal{F}_T] = -X$ for all $X \in \mathcal{X}$. Also it is easy to see that it satisfies coherency due to the properties of the essential supremum and those of the expectation.

Example 4.19 (Dynamic convex risk measure)**.** Let $\mathcal{P}$ be a convex set of $\mathbb{P}$-absolutely continuous probability measures defined on $(\Omega, \mathcal{F}_T)$. For any $t \in [0, T]$ let $F_t : \mathcal{P} \to \mathbb{R}$ be a convex function such that $\inf_{Q \in \mathcal{P}} F_t(Q) = 0$. Then for all $X \in \mathcal{X}$ and $t \in [0, T]$

$$\rho_t(X) = \operatorname*{ess.sup}_{Q \in \mathcal{P}} \{\mathbb{E}_Q[-X \mid \mathcal{F}_t] - F_t(Q)\} \tag{4.7}$$

is a dynamic convex risk measure satisfying Axiom 4.16 (a) positivity, (e) translation invariance and (f) constancy.

Again, going through the properties as in the previous example and keeping Theorem 4.13 in mind, it is easy to see that this is a dynamic convex risk measure satisfying the mentioned properties of Axiom 4.16.

Similarly to the static convex risk measure, this dynamic convex risk measure ρ_t is represented as the essential supremum over a set $\mathcal{P}$ of generalised scenarios corrected with a penalty term F, which depends on the scenarios. Also we remark here that not all dynamic risk measures as defined in (4.6) satisfy the time-consistency property. Since the risk measure as defined in (4.6) is a special case of the risk measure defined in (4.7), we know that (4.7) is in general not a time-consistent risk measure.

Another class of dynamic risk measure arises from the conditional g-expectation introduced by Peng [40] and which is elaborated further on in the dissertation.

For an extensive discussion on dynamic risk measures, we refer the reader to Frittelli and Gianin [26].

All static risk measures discussed thus far have been based on this classical mathematical expectation and all dynamic risk measures discussed thus far have been based on this conditional mathematical expectation. However, none of the risk measures mentioned have been ideal in measuring the riskiness of a financial position. Consequently, in an attempt to find a better risk measure, Rosazza Gianin

[30] defines a static risk measure in terms of the g-expectation and a dynamic risk measure in terms of the conditional g-expectation.

Chapter 5

Risk measures via g-expectations

In Chapter 4 we introduced the concept of static and dynamic risk measures, as well as coherent and convex risk measures. Having also introduced the concept of g-expectations, we can now define a risk measure in terms of the g-expectation. We begin by analysing the properties positive homogeneity, subadditivity, convexity, translation invariance and monotonicity with regards to g-expectations in Section 5.1. Following this, we define risk measures in terms of g-expectations.

5.1 Properties for g-expectations

The necessary and sufficient conditions for the properties positive homogeneity, subadditivity, convexity, translation invariance and monotonicity will be determined so that the properties can be applied to g-expectations. The link between risk measures and g-expectation may become clear, however will be explicitly stated in the next section.

Recall that for any pair $(y, z) \in \mathbb{R} \times \mathbb{R}^d$ we set

$$S_y^z(g) := \left\{ t \in [0, T) \,:\, g(t, y, z) = \lim_{\epsilon \to 0+} \frac{1}{\epsilon} \left[Y_t(g, t+\epsilon, y + z \cdot (B_{t+\epsilon} - B_t)) - y \right] \right\}.$$

If g is independent of y, then for any $z \in \mathbb{R}^d$, we set

$$S^z(g) := \left\{ t \in [0, T) \,:\, g(t, y, z) = \lim_{\epsilon \to 0+} \frac{1}{\epsilon} Y_t(g, t+\epsilon, z \cdot (B_{t+\epsilon} - B_t)) \right\}.$$

The results in this section are obtained under assumptions (A1) and (A3). The results also hold true under the additional assumption of continuity, even though this assumption is not necessary.

5.1.1 Positive homogeneity for g-expectations

We begin with a theorem regarding the positive homogeneity for g-expectations.

Theorem 5.1. *Let assumptions (A1), (A2) and (A3) hold for g, then the following conditions are equivalent:*

(i) $\mathcal{E}_g[\,\cdot\,]$ *is positively homogeneous;*

(ii) $\mathcal{E}_g[\,\cdot\,|\mathcal{F}_t]$ *is positively homogeneous for any $t \in [0,T]$; i.e. for all $X \in L^2(\mathcal{F}_T)$ and $\alpha \geq 0$,*

$$\mathcal{E}_g[\alpha X|\mathcal{F}_t] = \alpha \mathcal{E}_g[X|\mathcal{F}_t];$$

(iii) g *is positively homogeneous with respect to (y,z); i.e. for any $y \in \mathbb{R}, z \in \mathbb{R}^d$ and $\alpha \geq 0$,*

$$g(t, \alpha y, \alpha z) = \alpha g(t, y, z).$$

Proof. (iii)⇒(ii) For $\alpha = 0$ the proof is trivial, we thus assume that $\alpha > 0$. By definition $\mathcal{E}_g[X|\mathcal{F}_t] = Y_t$ and $\mathcal{E}_g[\alpha X|\mathcal{F}_t] = Y_t^*$ where we have that (Y_t, Z_t) and (Y_t^*, Z_t^*) are the respective solutions of

$$Y_t = X + \int_t^T g(s, Y_s, Z_s)ds - \int_t^T Z_s dB_s, \tag{5.1}$$

$$Y_t^* = \alpha X + \int_t^T g(s, Y_s^*, Z_s^*)ds - \int_t^T Z_s^* dB_s. \tag{5.2}$$

Now g is positively homogeneous in (y,z), thus we have that $g(t, \alpha y, \alpha z) = \alpha g(t,y,z)$ or equivalently, as $\alpha > 0$, we can write $\alpha g(t, \frac{y}{\alpha}, \frac{z}{\alpha}) = g(t,y,z)$. Thus (5.2) can be rewritten as

$$Y_t^* = \alpha X + \alpha \int_t^T g(s, \frac{Y_s^*}{\alpha}, \frac{Z_s^*}{\alpha})ds - \int_t^T Z_s^* dB_s. \tag{5.3}$$

Dividing by $\alpha > 0$ yields

$$\frac{Y_t^*}{\alpha} = X + \int_t^T g(s, \frac{Y_s^*}{\alpha}, \frac{Z_s^*}{\alpha})ds - \int_t^T \frac{Z_s^*}{\alpha} dB_s. \tag{5.4}$$

From (5.4) we notice that $(\frac{Y_t^*}{\alpha}, \frac{Z_t^*}{\alpha})_{t\in[0,T]}$ solves (5.1), however by Peng [40] we know that the solution to (5.1) is unique. Thus it follows that for any $t \in [0,T]$

$$\mathcal{E}_g[\alpha X|\mathcal{F}_t] = Y_t^* = \alpha Y_t = \alpha \mathcal{E}_g[X|\mathcal{F}_t]. \tag{5.5}$$

(ii)⇒(i) is trivial.

(i)⇒(iii) Suppose that (i) holds for $\mathcal{E}_g[\,\cdot\,]$. Let $\alpha > 0$ and define a new function, $\bar{g}^\alpha$, such that for all $(t, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^d$,

$$\bar{g}^\alpha(t, y, z) := \alpha g(t, \frac{y}{\alpha}, \frac{z}{\alpha}). \tag{5.6}$$

It is clear that $\bar{g}^\alpha$ also satisfies assumptions (A1), (A2) and (A3). Now

$$(Y_t(\bar{g}^\alpha, T, \alpha X), Z_t(\bar{g}^\alpha, T, \alpha X))_{t\in[0,T]}$$

is the solution of

$$Y_t = \alpha X + \int_t^T \bar{g}^\alpha(s, Y_s, Z_s)ds - \int_t^T Z_s dB_s \tag{5.7}$$

$$\Longrightarrow \qquad Y_t = \alpha X + \int_t^T \alpha g(s, \frac{Y_s}{\alpha}, \frac{Z_s}{\alpha})ds - \int_t^T Z_s dB_s. \tag{5.8}$$

Also, we have that $(Y_t(g, T, X), Z_t(g, T, X))_{t\in[0,T]}$ is the solution of

$$Y_t = X + \int_t^T g(s, Y_s, Z_s)ds - \int_t^T Z_s dB_s \tag{5.9}$$

$$\Longrightarrow \qquad \alpha Y_t = \alpha X + \int_t^T \alpha g(s, \frac{\alpha Y_s}{\alpha}, \frac{\alpha Z_s}{\alpha})ds - \int_t^T \alpha Z_s dB_s. \tag{5.10}$$

However, by Peng [40], we know that for any $X \in L^2(\mathcal{F}_T)$, the solution of the BSDE (5.8) is unique, and thus

$$(Y_t(\bar{g}^\alpha, T, \alpha X), Z_t(\bar{g}^\alpha, T, \alpha X))_{t\in[0,T]} = \alpha(Y_t(g, T, X), Z_t(g, T, X))_{t\in[0,T]}.$$

We therefore have that for any given $\alpha > 0$ and for all $X \in L^2(\mathcal{F}_T)$

$$\mathcal{E}_{\bar{g}^\alpha}[\alpha X] = \alpha \mathcal{E}_g[X]. \tag{5.11}$$

Combining this equality with (i), we get

$$\alpha \mathcal{E}_{\bar{g}^\alpha}[X] = \alpha \mathcal{E}_g[X]. \tag{5.12}$$

Dividing by $\alpha > 0$ gives

$$\mathcal{E}_{\bar{g}^\alpha}[X] = \mathcal{E}_g[X]. \tag{5.13}$$

Then by Theorem 3.27 and by the representation lemma, we have that for any $(y, z) \in \mathbb{R} \times \mathbb{R}^d$ and for any $t \in S_y^z(\bar{g}^\alpha) \cap S_y^z(g)$ $\mathbb{P}$-almost surely

$$\bar{g}^\alpha(t, y, z) = g(t, y, z), \tag{5.14}$$

where we know that

$$\lambda([0, T]\backslash(S_y^z(\bar{g}^\alpha) \cap S_y^z(g))) = 0, \tag{5.15}$$

and λ denotes the Lebesgue measure. It follows from (5.14) and (5.15) that

$$d\mathbb{P} \times dt \ \text{ a.s.}, \quad \bar{g}^{\alpha}(t, y, z) = g(t, y, z).$$

Since $\bar{g}^{\alpha}$ and g are both Lipschitz continuous with respect to (y, z), it follows that for all $y \in \mathbb{R}$ and $z \in \mathbb{R}^d$

$$d\mathbb{P} \times dt \ \text{ a.s.}, \quad \bar{g}^{\alpha}(t, y, z) = g(t, y, z).$$

Hence for any $\alpha > 0$, we conclude that for all $(y, z) \in \mathbb{R} \times \mathbb{R}^d$

$$\bar{g}^{\alpha} = g, \quad \text{i.e. } g(t, y, z) = \alpha g(t, \frac{y}{\alpha}, \frac{z}{\alpha}), \tag{5.16}$$

and (iii) follows. ☐

5.1.2 Translation invariance for g-expectations

Theorem 5.2. *Let assumptions (A1), (A2) and (A3) hold for g, then the following conditions are equivalent:*

(i) for all $X \in L^2(\mathcal{F}_T)$ and $c \in \mathbb{R}$, we have $\mathcal{E}_g[X + c] = \mathcal{E}_g[X] + c$;

(ii) for all $X \in L^2(\mathcal{F}_T), c \in \mathbb{R}$ and $t \in [0, T]$, we have $\mathcal{E}_g[X + c|\mathcal{F}_t] = \mathcal{E}_g[X|\mathcal{F}_t] + c$;

(iii) g is independent of y.

Proof. (iii)$\Rightarrow$(ii) This follows from Lemma 3.25.

(ii)$\Rightarrow$(i) is trivial.

(i)$\Rightarrow$(iii) Suppose that (i) holds. For any $c \in \mathbb{R}$, we define a new generator such that for all $(t, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^d$

$$g^c(t, y, z) := g(t, y - c, z).$$

It is clear that g^c satisfies the assumptions (A1), (A2) and (A3). Now $(Y_t(g^c, T, X + c), Z_t(g^c, T, X + c))_{t \in [0,T]}$ is the solution of

$$Y_t = X + c + \int_t^T g^c(s, Y_s, Z_s) ds - \int_t^T Z_s dB_s \tag{5.17}$$

$$\Longrightarrow \qquad Y_t - c = X + \int_t^T g(s, Y_s - c, Z_s) ds - \int_t^T Z_s dB_s. \tag{5.18}$$

Also, we have that $(Y_t(g, T, X) + c, Z_t(g, T, X))_{t \in [0,T]}$ is the solution of

$$Y_t + c = X + \int_t^T g(s, Y_s + c, Z_s) ds - \int_t^T Z_s dB_s. \tag{5.19}$$

However, by Peng [40], we know that for any $X \in L^2(\mathcal{F}_T)$, the solution of the BSDE (5.18) is unique, and thus

$$(Y_t(g^c, T, X + c), Z_t(g^c, T, X + c))_{t\in[0,T]} = (Y_t(g, T, X) + c, Z_t(g, T, X))_{t\in[0,T]}.$$

We consequently have for all $X \in L^2(\mathcal{F}_T)$

$$\mathcal{E}_{g^c}[X + c] = Y_0(g^c, T, X + c) = Y_0(g, T, X) + c = \mathcal{E}_g[X] + c.$$

Combining this equality with (i), we get

$$\mathcal{E}_{g^c}[X + c] = \mathcal{E}_g[X + c].$$

Therefore, we have for all $\eta \in L^2(\mathcal{F}_T)$

$$\mathcal{E}_{g^c}[\eta] = \mathcal{E}_g[\eta].$$

Then by Theorem 3.27 and the representation lemma, we have for any $(y, z) \in \mathbb{R}\times\mathbb{R}^d$ and for any $t \in S^z_y(g^c) \cap S^z_y(g)$ that $\mathbb{P}$-almost surely

$$g^c(t, y, z) = g(t, y, z). \tag{5.20}$$

Here we know that

$$\lambda([0, T]\backslash(S^z_y(g^c) \cap S^z_y(g))) = 0, \tag{5.21}$$

where λ denotes the Lebesgue measure. It follows from (5.20) and (5.21) that

$$d\mathbb{P} \times dt \ \text{ a.s.,} \quad g^c(t, y, z) = g(t, y, z).$$

Since g^c and g are both Lipschitz continuous with respect to (y, z), it follows that for all $y \in \mathbb{R}$ and $z \in \mathbb{R}^d$

$$d\mathbb{P} \times dt \ \text{ a.s.,} \quad g^c(t, y, z) = g(t, y, z).$$

That is for any given $c \in \mathbb{R}$, we have $g^c = g$, i.e. $g^c(t, y, z) = g(t, y - c, z) = g(t, y, z)$. In particular this equation holds for $c = y$ for any $y \in \mathbb{R}$. Thus for any $y \in \mathbb{R}$, and for all $z \in \mathbb{R}^d$ we have

$$d\mathbb{P} \times dt \ \text{ a.s.,} \quad g^c(t, y, z) = g(t, y - y, z) = g(t, 0, z).$$

This shows that g is independent of y. and part (iii) follows. □

5.1.3 Convexity for g-expectations

Theorem 5.3. *Let assumptions (A1), (A2) and (A3) hold for g, then the following conditions are equivalent:*

(i) $\mathcal{E}_g[\,\cdot\,]$ *is convex.*

(ii) $\mathcal{E}_g[\,\cdot\,|\mathcal{F}_t]$ *is convex for all* $t \in [0,T]$, *i.e. for all* $X, \eta \in L^2(\mathcal{F}_T)$ *and* $\alpha \in [0,1]$

$$\mathcal{E}_g[\alpha X + (1-\alpha)\eta|\mathcal{F}_t] \leq \alpha\mathcal{E}_g[X|\mathcal{F}_t] + (1-\alpha)\mathcal{E}_g[\eta|\mathcal{F}_t].$$

(iii) g *is independent of* y *and* g *is convex with respect to* z, *i.e. for any* $z_1, z_2 \in \mathbb{R}^d$ *and* $\alpha \in [0,1]$

$$g(t, \alpha z_1 + (1-\alpha)z_2) \leq \alpha g(t, z_1) + (1-\alpha)g(t, z_2).$$

Proof. (iii)⇒(ii) This follows from the comparison theorem, Theorem 3.8.

(ii)⇒(i) is trivial.

(i)⇒(iii) Suppose that (i) holds. We first prove that (i) implies the translation invariance property, i.e. given that for all $X, \eta \in L^2(\mathcal{F}_T)$ and $\alpha \in [0,1]$,

$$\mathcal{E}_g[\alpha X + (1-\alpha)\eta] \leq \alpha\mathcal{E}_g[X] + (1-\alpha)\mathcal{E}_g[\eta], \tag{5.22}$$

we want to prove that for any $c \in \mathbb{R}$,

$$\mathcal{E}_g[X + c] = \mathcal{E}_g[X] + c. \tag{5.23}$$

By (i) and by Proposition 3.13 (i), we have for all $X \in L^2(\mathcal{F}_T), c \in \mathbb{R}$ and $\alpha \in [0,1]$

$$\begin{aligned}\mathcal{E}_g[\alpha X + (1-\alpha)c] &\leq \alpha\mathcal{E}_g[X] + (1-\alpha)\mathcal{E}_g[c] \\ &= \alpha\mathcal{E}_g[X] + (1-\alpha)c.\end{aligned}$$

Now we set $\alpha = (1 - \frac{1}{n}) \in [0,1]$, where n is any positive integer. Noting that if $c \in \mathbb{R}$, then $nc \in \mathbb{R}$, we get that for all $X \in L^2(\mathcal{F}_T)$ and $c \in \mathbb{R}$

$$\begin{aligned}\mathcal{E}_g[(1-\frac{1}{n})X + c] &= \mathcal{E}_g[(1-\frac{1}{n})X + \frac{1}{n}(nc)] \\ &\leq (1-\frac{1}{n})\mathcal{E}_g[X] + c.\end{aligned}$$

The operator $\mathcal{E}_g[\,\cdot\,]$ is continuous in an L^2 sense, and thus, taking the limit in L^2 we have

$$\begin{aligned}\mathcal{E}_g[X + c] &= \lim_{n\to\infty} \mathcal{E}_g[(1-\frac{1}{n})X + c] \\ &\leq \lim_{n\to\infty}(1-\frac{1}{n})\mathcal{E}_g[X] + c \\ &= \mathcal{E}_g[X] + c.\end{aligned}$$

We thus get for all $X \in L^2(\mathcal{F}_T)$ and $c \in \mathbb{R}$

$$\mathcal{E}_g[X + c] \leq \mathcal{E}_g[X] + c, \tag{5.24}$$

giving

$$\mathcal{E}_g[X] = \mathcal{E}_g[X + c - c] \leq \mathcal{E}_g[X + c] - c. \tag{5.25}$$

Equation (5.25) can be rewritten as

$$\mathcal{E}_g[X] + c \leq \mathcal{E}_g[X + c]. \tag{5.26}$$

From (5.24) and (5.26) we infer that for all $X \in L^2(\mathcal{F}_T), c \in \mathbb{R}$

$$\mathcal{E}_g[X] + c = \mathcal{E}_g[X + c].$$

We have thus shown that the g-expectation $\mathcal{E}_g[\,\cdot\,]$ satisfies the translation invariance property. By Theorem 5.2 we can conclude that g is independent of y.

We now need to prove that for all $t \in [0, T], X, \eta \in L^2(\mathcal{F}_T)$ and $\alpha \in [0, 1]$

$$\mathcal{E}_g[\alpha X + (1 - \alpha)\eta|\mathcal{F}_t] \leq \alpha\mathcal{E}_g[X|\mathcal{F}_t] + (1 - \alpha)\mathcal{E}_g[\eta|\mathcal{F}_t]. \tag{5.27}$$

Consider the event

$$A := \{\mathcal{E}_g[\alpha X + (1 - \alpha)\eta|\mathcal{F}_t] > \alpha\mathcal{E}_g[X|\mathcal{F}_t] + (1 - \alpha)\mathcal{E}_g[\eta|\mathcal{F}_t]\}.$$

Clearly $A \in \mathcal{F}_t$. Now suppose that $\mathbb{P}(A) > 0$. Then

$$\mathbf{1}_A\mathcal{E}_g[\alpha X + (1 - \alpha)\eta|\mathcal{F}_t] - \mathbf{1}_A(\alpha\mathcal{E}_g[X|\mathcal{F}_t] + (1 - \alpha)\mathcal{E}_g[\eta|\mathcal{F}_t]) \geq 0$$

and

$$\mathbb{P}(\mathbf{1}_A\mathcal{E}_g[\alpha X + (1 - \alpha)\eta|\mathcal{F}_t] - \mathbf{1}_A(\alpha\mathcal{E}_g[X|\mathcal{F}_t] + (1 - \alpha)\mathcal{E}_g[\eta|\mathcal{F}_t]) > 0) > 0.$$

Since $A \in \mathcal{F}_t$, it is obvious that for all $X \in L^2(\mathcal{F}_T)$

$$\mathcal{E}_g[\mathbf{1}_A X|\mathcal{F}_t] = \mathbf{1}_A\mathcal{E}_g[X|\mathcal{F}_t]. \tag{5.28}$$

Now since g is independent of y and $A \in \mathcal{F}_t$, we have by Proposition 3.13 (ii), Lemma 3.25 and by Equality (5.28) respectively that

$$\begin{aligned}
&\mathcal{E}_g[\mathbf{1}_A(\alpha X + (1 - \alpha)\eta) - \mathbf{1}_A(\alpha\mathcal{E}_g[X|\mathcal{F}_t] + (1 - \alpha)\mathcal{E}_g[\eta|\mathcal{F}_t])] \\
&\quad = \mathcal{E}_g\{\mathcal{E}_g[\mathbf{1}_A(\alpha X + (1 - \alpha)\eta) - \mathbf{1}_A(\alpha\mathcal{E}_g[X|\mathcal{F}_t] + (1 - \alpha)\mathcal{E}_g[\eta|\mathcal{F}_t])|\mathcal{F}_t]\} \\
&\quad = \mathcal{E}_g\{\mathcal{E}_g[\mathbf{1}_A(\alpha X + (1 - \alpha)\eta)|\mathcal{F}_t] - \mathbf{1}_A(\alpha\mathcal{E}_g[X|\mathcal{F}_t] + (1 - \alpha)\mathcal{E}_g[\eta|\mathcal{F}_t])\} \\
&\quad = \mathcal{E}_g\{\mathbf{1}_A\mathcal{E}_g[(\alpha X + (1 - \alpha)\eta)|\mathcal{F}_t] - \mathbf{1}_A(\alpha\mathcal{E}_g[X|\mathcal{F}_t] + (1 - \alpha)\mathcal{E}_g[\eta|\mathcal{F}_t])\} \\
&\quad > 0.
\end{aligned}$$

On the other hand we have, using the fact that $\mathcal{E}_g[\,\cdot\,]$ is convex, Proposition 3.13 (ii), Lemma 3.25 (noting that g is independent of y) and Equality (5.28) respectively, that

$$\begin{aligned}
&\mathcal{E}_g[\mathbf{1}_A(\alpha X + (1-\alpha)\eta) - \mathbf{1}_A(\alpha\mathcal{E}_g[X|\mathcal{F}_t] + (1-\alpha)\mathcal{E}_g[\eta|\mathcal{F}_t])] \\
&\qquad = \mathcal{E}_g[\alpha(\mathbf{1}_A X - \mathbf{1}_A\mathcal{E}_g[X|\mathcal{F}_t]) + (1-\alpha)(\mathbf{1}_A\eta - \mathbf{1}_A\mathcal{E}_g[\eta|\mathcal{F}_t])] \\
&\qquad \le \alpha\mathcal{E}_g[(\mathbf{1}_A X - \mathbf{1}_A\mathcal{E}_g[X|\mathcal{F}_t])] + (1-\alpha)\mathcal{E}_g[(\mathbf{1}_A\eta - \mathbf{1}_A\mathcal{E}_g[\eta|\mathcal{F}_t])] \\
&\qquad = \alpha\mathcal{E}_g\{\mathcal{E}_g[(\mathbf{1}_A X - \mathbf{1}_A\mathcal{E}_g[X|\mathcal{F}_t])|\mathcal{F}_t]\} + (1-\alpha)\mathcal{E}_g\{\mathcal{E}_g[(\mathbf{1}_A\eta - \mathbf{1}_A\mathcal{E}_g[\eta|\mathcal{F}_t])\mathcal{F}_t]\} \\
&\qquad = \alpha\mathcal{E}_g\{\mathbf{1}_A\mathcal{E}_g[X|\mathcal{F}_t] - \mathbf{1}_A\mathcal{E}_g[X|\mathcal{F}_t]\} + (1-\alpha)\mathcal{E}_g\{\mathbf{1}_A\mathcal{E}_g[\eta|\mathcal{F}_t] - \mathbf{1}_A\mathcal{E}_g[\eta|\mathcal{F}_t]\} \\
&\qquad = 0 + 0 \\
&\qquad = 0.
\end{aligned}$$

This leads to a contradiction and thus we cannot have $\mathbb{P}(A) > 0$, giving us $\mathbb{P}(A) = 0$. Consequently we have that Equation (5.27) holds.

For any $z_1, z_2 \in \mathbb{R}^d$, $\alpha \in [0,1]$, if $t \in S^{\alpha z_1 + (1-\alpha)z_2}(g) \cap S^{z_1}(g) \cap S^{z_2}(g)$, by Theorem 3.29 we deduce that $\mathbb{P}$-almost surely we have

$$g(t, \alpha z_1 + (1-\alpha)z_2) \le \alpha g(t, z_1) + (1-\alpha)g(t, z_2).$$

For any $z_1, z_2 \in \mathbb{R}^d$, $\alpha \in [0,1]$, by the representation lemma we have that

$$\lambda([0,T]\backslash S^{\alpha z_1 + (1-\alpha)z_2}(g) \cap S^{z_1}(g) \cap S^{z_2}(g)) = 0.$$

Thus for any $z_1, z_2 \in \mathbb{R}^d$, $\alpha \in [0,1]$, we have

$$\mathbb{P} \times dt \;\; \text{a.s.}, \quad g(t, \alpha z_1 + (1-\alpha)z_2) \le \alpha g(t, z_1) + (1-\alpha)g(t, z_2).$$

Thus (iii) holds, which completes the proof. □

5.1.4 Subadditivity for g-expectations

Theorem 5.4. *Let assumptions (A1), (A2) and (A3) hold for g, then the following conditions are equivalent:*

(i) $\mathcal{E}_g[\,\cdot\,]$ *is subadditive.*

(ii) $\mathcal{E}_g[\,\cdot\,|\mathcal{F}_t]$ *is subadditive for all $t \in [0,T]$, i.e. for all $X, \eta \in L^2(\mathcal{F}_T)$*

$$\mathcal{E}_g[X + \eta|\mathcal{F}_t] \le \mathcal{E}_g[X|\mathcal{F}_t] + \mathcal{E}_g[\eta|\mathcal{F}_t].$$

(iii) g is independent of y and g is subadditive with respect to z, i.e. for any $z_1, z_2 \in \mathbb{R}^d$

$$g(t, z_1 + z_2) \le g(t, z_1) + g(t, z_2).$$

Proof. As subadditivity follows directly from convexity and positive homogeneity the proof of this theorem follows from Theorem 5.3 and Theorem 5.1. □

5.2 Risk measures using g-expectations

Let g satisfy assumptions (A1), (A2) and (A3) and let $\rho^g : L^2(\mathcal{F}_T) \to \mathbb{R}$ and $\rho_t^g : L^2(\mathcal{F}_T) \to \mathbb{R}$. Let $X \in L^2(\mathcal{F}_T)$. We define a static measure of risk in terms of the g-expectation as

$$\rho^g(X) := \mathcal{E}_g[-X], \tag{5.29}$$

and, for all $t \in [0,T]$, a dynamic measure of risk in terms of the conditional g-expectation as

$$\rho_t^g(X) := \mathcal{E}_g[-X|\mathcal{F}_t]. \tag{5.30}$$

By Proposition 3.10 (i) and (ii) respectively, the risk measure as described in (5.29) satisfies Axiom 4.1 (a) positivity, (b) monotonicity and (f) constancy.

The following two propositions have been based on Rosazza Gianin [30].

Proposition 5.5. *The risk measure ρ^g defined in* (5.29) *satisfies the following properties.*

(i) Coherency*: If g is positively homogeneous and subadditive in (y, z), then ρ^g is a coherent risk measure satisfying Axiom 4.1 (h) lower semi-continuity.*

(ii) Convexity*: If g is convex in (y, z), then ρ^g is a convex risk measure satisfying Axiom 4.1 (f) constancy.*

Proof. As (i) is a particular case of (ii), we first prove case (ii) and case (i) follows.

(ii) By Theorem 5.2 we have that ρ^g satisfies translation invariance. Positivity, monotonicity and constancy follow directly from Proposition 3.10 (i) and (ii) respectively. Hence ρ^g is a monetary risk measure. Now suppose g is convex in (y, z). By the remark following Lemma 3.26, we know that if g satisfies convexity, then g does not depend on y. Hence by Theorem 5.3, we know that $\mathcal{E}_g[\,\cdot\,]$ is convex and thus ρ^g satisfies convexity. We still need the lower semi-continuity of ρ^g and $\rho^g(0) = 0$. The latter follows from the constancy of ρ^g. Now by Delbaen [15] and Rosazza Gianin [30] we have that, assuming translation invariance, ρ^g is lower semi-continuous if and only if the acceptance set $\mathcal{A} = \{X \in L^2(\mathcal{F}_T) \,|\, \rho(X) \leq 0\}$ is closed in L^2. The closure of the set $\mathcal{A}$ follows directly from Proposition 3.10 (iii) and from the definition of $\rho(X)$. We know that $\rho(X) = \mathcal{E}_g[-X]$ and Proposition 3.10 (iii) tells us that

$$|\rho(X_1) - \rho(X_2)|^2 \leq C\mathbb{E}[|X_1 - X_2|^2].$$

Hence $\mathcal{A}$ is a closed set. Consequently we have that ρ^g is lower semi-continuous which completes the proof.

(i) Suppose that g is positively homogeneous and subadditive in (y, z). Thus g is also convex in (y, z) and the reasoning of part (ii) applies. Hence we only need to verify that ρ^g satisfies subadditivity and positive homogeneity. These two properties follow from Theorems 5.4 and 5.1 respectively and hence the result is proved. $\square$

In the dynamic setting, we have the following.

Proposition 5.6. *The dynamic risk measure* $(\rho_t^g)_{t\in[0,T]}$ *defined in* (5.30) *satisfies the following properties.*

(i) Continuous-time recursivity*: For* $0 \leq s \leq t \leq T$ *and for all* $X \in L^2(\mathcal{F}_T)$*,*

$$\rho_s^g(X) = \rho_s^g[-\rho_t^g(X)]. \tag{5.31}$$

(ii) Let $X, Y \in L^2(\mathcal{F}_T)$*. If for some* $t \in (0,T]$*,* $\rho_t^g(X) \leq \rho_t^g(Y)$*, then for any* $s \in [0,t]$ *we have* $\rho_s^g(X) \leq \rho_s^g(Y)$*.*

(iii) Coherency*: If* g *is positively homogeneous and subadditive in* (y,z)*, then we have that* $(\rho_t^g)_{t\in[0,T]}$ *is a dynamic coherent and time-consistent risk measure.*

(iv) Convexity*: If* g *is convex in* (y,z)*, then we have that* $(\rho_t^g)_{t\in[0,T]}$ *is a dynamic convex and time-consistent risk measure satisfying Axiom 4.16 (f) dynamic constancy.*

In (i) and (ii) above, we can replace deterministic times s and t by stopping times σ and τ with $0 \leq \sigma \leq \tau \leq T$.

Proof. (i) Continuous-time recursivity follows from the definition of ρ_t^g and from Proposition 3.13 (ii). Let $0 \leq s \leq t \leq T$, then

$$\begin{aligned}
\rho_s^g(X) &= \mathcal{E}_g[-X|\mathcal{F}_s] \\
&= \mathcal{E}_g[\mathcal{E}_g[-X|\mathcal{F}_t]|\mathcal{F}_s] \\
&= \mathcal{E}_g[\rho_t^g(X)|\mathcal{F}_s] \\
&= \rho_s^g(-\rho_t^g(X)).
\end{aligned}$$

(ii) Suppose $\rho_t^g(X) \leq \rho_t^g(Y)$ for some $t \in (0,T]$. Thus $\mathcal{E}_g[-X|\mathcal{F}_t] \leq \mathcal{E}_g[-Y|\mathcal{F}_t]$. Using Proposition 3.13 (ii), we get that for any $s \in [0,t]$,

$$\rho_s^g(X) = \mathcal{E}_g[\mathcal{E}_g[-X|\mathcal{F}_t]|\mathcal{F}_s] \leq \mathcal{E}_g[\mathcal{E}_g[-Y|\mathcal{F}_t]|\mathcal{F}_s] = \rho_s^g(Y). \tag{5.32}$$

In the proofs of (i) and (ii), the same reasoning applies to stopping times σ and τ.

As (iii) is a particular case of (iv), we first prove case (iv) and case (iii) follows.

(iv) The proof follows similarly to the proof of Proposition 5.5. Dynamic constancy, dynamic positivity and dynamic monotonicity follow from Proposition 3.13 (i) and (iii) respectively. Since g is independent of y, dynamic translability follows from Lemma 3.25. Hence for each t, ρ_g is a monetary risk measure. Now suppose g is convex in (y, z). By the remark following Lemma 3.26, we know that if g satisfies convexity, then g does not depend on y. Hence by Theorem 5.3, we know that $\mathcal{E}_g[\,\cdot\,|\mathcal{F}_t]$ is convex and thus ρ_t^g satisfies convexity. Let $t \in [0, T]$, $X \in L^2(\mathcal{F}_T)$ and $A \in \mathcal{F}_t$. Note that since $\mathcal{F}_0$ is the trivial filtration, we have that $\rho_0^g(X) = \mathcal{E}_g[X|\mathcal{F}_0] = \mathcal{E}_g[X]$. The time-consistency follows from the definition of ρ_t^g, from Proposition 3.13 (ii) and using the fact that A is $\mathcal{F}_t$-measurable:

$$\begin{aligned}
\rho_0^g(X\mathbf{1}_A) &= \mathcal{E}_g[-X\mathbf{1}_A] \\
&= \mathcal{E}_g[\mathcal{E}_g[-X\mathbf{1}_A|\mathcal{F}_t]] \\
&= \rho_0^g(-\mathcal{E}_g[-X\mathbf{1}_A|\mathcal{F}_t]) \\
&= \rho_0^g(-\mathcal{E}_g[-X|\mathcal{F}_t]\mathbf{1}_A) \\
&= \rho_0^g[-\rho_t^g(X)\mathbf{1}_A].
\end{aligned}$$

(iii) The proof follows identically to that of Proposition 5.5 (i). Suppose that g is positively homogeneous and subadditive in (y, z). Thus g is also convex in (y, z) and the reasoning of part (iv) can be applied. We hence only need to verify that ρ_t^g satisfies subadditivity and positive homogeneity. These two properties follow from Theorems 5.4 and 5.1 respectively and hence the result is proved. □

Note the connection between the recursivity of the dynamic risk measure and the time-consistency condition of a dynamic risk measure.

We have seen by the previous proposition, that a conditional g-expectation can induce a dynamic, time-consistent risk measure. The aim of this next section is to find the conditions under which a dynamic time-consistent risk measure is induced by a conditional g-expectation. The following section has been taken from Rosazza Gianin [30] and Coquet et al. [12].

5.3 Risk measures induced by g-expectations

Before stating and proving the main proposition of this section, the concept of $\mathcal{E}^\mu$-domination needs to be defined. Recall that when $g(t, y, z) = \mu|z|$ where $\mu > 0$, we will denote $\mathcal{E}_g[\,\cdot\,]$ by $\mathcal{E}^\mu[\,\cdot\,]$ and $\mathcal{E}_g[\,\cdot\,|\mathcal{F}_t]$ by $\mathcal{E}^\mu[\,\cdot\,|\mathcal{F}_t]$. Let $X \in L^2(\mathcal{F}_T)$. We denote the corresponding static risk measure by ρ^μ and the dynamic risk measure by $(\rho_t^\mu)_{t\in[0,T]}$ and define them by

$$\rho^\mu(X) := \mathcal{E}^\mu[-X] \tag{5.33}$$

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

$$\rho_t^\mu(X) := \mathcal{E}^\mu[-X|\mathcal{F}_t]. \tag{5.34}$$

Also when $g(t,y,z) = -\mu|z|$ where $\mu > 0$, we will denote $\mathcal{E}_g[\,\cdot\,]$ by $\mathcal{E}^{-\mu}[\,\cdot\,]$ and $\mathcal{E}_g[\,\cdot\,|\mathcal{F}_t]$ by $\mathcal{E}^{-\mu}[\,\cdot\,|\mathcal{F}_t]$.

Definition 5.7. An $(\mathcal{F}_t)_{t\in[0,T]}$-consistent expectation, $\mathcal{E}[\,\cdot\,]$, is said

(i) to be $\mathcal{E}^{\mu}$-*dominated*, where $\mu > 0$, if for all $X, Y \in L^2(\mathcal{F}_T)$, we have

$$\mathcal{E}[X+Y] - \mathcal{E}[X] \leq \mathcal{E}^{\mu}[Y];$$

(ii) to satisfy the *translability condition* if for any $X \in L^2(\mathcal{F}_T)$ and $\eta \in L^2(\mathcal{F}_t)$, we have

$$\mathcal{E}[X+\eta \mid \mathcal{F}_t] = \mathcal{E}[X \mid \mathcal{F}_t] + \eta.$$

We say that the risk measure, ρ, is $\mathcal{E}^{\mu}$-dominated if for any $X, Y \in L^2(\mathcal{F}_T)$ we have

$$\rho(X+Y) - \rho(X) \leq \rho^{\mu}(Y) = \mathcal{E}^{\mu}[-Y].$$

The property of $\mathcal{E}^{\mu}$-domination is the generalisation of the 'domination' true for any static convex risk measure. Using the representation of a convex risk measure (4.4), properties of the supremum and the linearity of the classical expectation, $\mathbb{E}[\,\cdot\,]$, we know that

$$\begin{aligned}
\rho(X+Y) - \rho(X) &= \sup_{Q\in\mathcal{P}} \{\mathbb{E}_Q[-(X+Y)] - F(\mathcal{Q})\} - \sup_{Q\in\mathcal{P}} \{\mathbb{E}_Q[-X] - F(\mathcal{Q})\} \\
&\leq \sup_{Q\in\mathcal{P}} \{\mathbb{E}_Q[-X-Y] - \mathbb{E}_Q[-X]\} \\
&= \sup_{Q\in\mathcal{P}} \{\mathbb{E}_Q[-Y]\} \\
&= \hat{\rho}(Y)
\end{aligned}$$

which is a coherent risk measure by Theorem 4.8. Hence we have that any convex risk measure, ρ, is dominated by a suitable coherent risk measure, $\hat{\rho}$.

The following lemma from Rosazza Gianin [30] provides necessary and sufficient conditions for $\mathcal{E}^{\mu}$-domination, as a consequence of the representation of static coherent and convex risk measures. Note that $g(z) = \mu|z|$ is convex when $\mu > 0$ and concave when $\mu < 0$. Hence $\mathcal{E}^{\mu}[X]$ is convex when $\mu > 0$ and concave when $\mu < 0$.

Lemma 5.8. *Let ρ_0 be a static coherent risk measure satisfying lower semi-continuity and let $\rho^{\mu}(X) = \mathcal{E}^{\mu}[-X]$ for some $\mu > 0$. Denote by $\mathcal{P}_0$ and $\mathcal{P}_{\mu}$ the convex sets in the representation of ρ_0 and ρ^{μ} respectively. The risk measure ρ_0 is $\mathcal{E}^{\mu}$-dominated for some $\mu > 0$ if and only if $\mathcal{P}_0 \subseteq \mathcal{P}_{\mu}$.*

Proof. Suppose that ρ_0 is $\mathcal{E}^\mu$-dominated. By the representation of a static coherent risk measure (4.3), we have that

$$\rho_0(X) = \sup_{Q\in\mathcal{P}_0} \{\mathbb{E}_Q[-X]\}.$$

Also, since $\rho^\mu(X) = \mathcal{E}^\mu[-X]$ is convex and consequently a convex risk measure, we have by the representation of a static convex risk measure (4.4) that

$$\rho^\mu(X) = \sup_{Q\in\mathcal{P}_\mu} \{\mathbb{E}_Q[-X] - F(Q)\}. \tag{5.35}$$

Let $Q \in \mathcal{P}_0$ be arbitrary. Then we know by the definition of supremum that $\mathbb{E}_Q[X] \leq \rho_0(-X)$. By the dominance of ρ_0, we also know that for any $X \in L^2(\mathcal{F}_T)$

$$\rho(0 - X) - \rho(0) \leq \rho^\mu(-X) = \mathcal{E}^\mu[X],$$

where $\rho(0) = 0$. Combining these inequalities, we have

$$\mathbb{E}_Q[X] \leq \rho_0(-X) \leq \rho^\mu(-X) = \mathcal{E}^\mu(X).$$

Hence in the representation of ρ^μ, (5.35), we have that $F(Q) = 0$ and consequently $Q \in \mathcal{P}_\mu$, giving us $\mathcal{P}_0 \subseteq \mathcal{P}_\mu$.

Conversely, let us suppose that $\mathcal{P}_0 \subseteq \mathcal{P}_\mu$. Also let $X, Y \in L^2(\mathcal{F}_T)$. Then, using the fact that ρ_0 is a static coherent risk measure and hence subadditive, from the representation of ρ_0 and using $\mathcal{P}_0 \subseteq \mathcal{P}_\mu$, we have

$$\begin{aligned}
\rho_0(X+Y) - \rho_0(X) &\leq \rho_0(Y) \\
&= \sup_{Q\in\mathcal{P}_0} \{\mathbb{E}_Q[-Y]\} \\
&\leq \sup_{Q\in\mathcal{P}_\mu} \{\mathbb{E}_Q[-Y]\} \\
&= \mathcal{E}^\mu[-Y]
\end{aligned}$$

and hence ρ_0 is $\mathcal{E}^\mu$-dominated for some $\mu > 0$. □

The following theorem from Coquet et al. [12] gives the conditions under which a $(\mathcal{F}_t)_{t\in[0,T]}$-consistent expectation is induced by a g-expectation. The result is needed for the main proposition of this section and will be stated without proof. The interested reader can refer to Coquet et al. [12] for the proof.

Theorem 5.9. *Let $\mathcal{E} : L^2(\mathcal{F}_T) \to \mathbb{R}$ be an $(\mathcal{F}_t)_{t\in[0,T]}$-consistent expectation. If $\mathcal{E}[\,\cdot\,]$ is $\mathcal{E}^\mu$-dominated for some $\mu > 0$ and if it satisfies the translability condition, then* *there exists a unique function g, independent of y, i.e. $g = g(t,z)$ satisfying the usual assumptions (A1), (A2) and (A3), such that for all $X \in L^2(\mathcal{F}_T)$*

$$\mathcal{E}[X] = \mathcal{E}_g[X]$$
$$\text{and} \quad \mathcal{E}[X|\mathcal{F}_t] = \mathcal{E}_g[X|\mathcal{F}_t].$$

Moreover we have that $|g(t,z)| \leq \mu|z|$ for all $t \in [0,T]$.

Under the further assumption of continuity on the function g, we have that the converse implication of Theorem 5.9 holds too. This is stated and proved next.

Lemma 5.10. *Let g satisfy the usual assumptions (A1), (A2) and (A3), be continuous in t, independent of y and let $|g(t,z)| \leq \mu|z|$ for some $\mu > 0$. Then $\mathcal{E}_g[\,\cdot\,]$ is an $(\mathcal{F}_t)_{t\in[0,T]}$-consistent expectation satisfying $\mathcal{E}^\mu$-domination and the translability condition.*

Proof. *$(\mathcal{F}_t)_{t\in[0,T]}$-consistent expectation:* By Lemma 3.15, $\mathcal{E}_g[\,\cdot\,]$ is a $(\mathcal{F}_t)_{t\in[0,T]}$-consistent expectation.

$\mathcal{E}^\mu$-domination: By hypothesis, g is continuous in t and independent of y and hence it is progressively measurable. Moreover we know that $|g(t,z)| \leq \mu|z|$ or equivalently that $-\mu|z| \leq g(t,z) \leq \mu|z|$. By the definition of a g-expectation, we have that

$$\mathcal{E}_g[X+Y] = X + Y + \int_0^T g(s,Z_s)ds - \int_0^T Z_s dB_s \tag{5.36}$$
$$\mathcal{E}_g[X] = X + \int_0^T g(s,\hat{Z}_s)ds - \int_0^T \hat{Z}_s dB_s. \tag{5.37}$$

Now set $Z_t^* := Z_t - \hat{Z}_t$ and $g_{\hat{Z}}(t,Z_t^*) := g(t,Z_t) - g(t,\hat{Z}_t)$ for all $t \in [0,T]$. Let $X \in L^2(\mathcal{F}_T)$. Subtracting (5.37) from (5.36), gives

$$\begin{aligned}\mathcal{E}_g[X+Y] - \mathcal{E}_g[X] &= Y + \int_0^T [g(s,Z_s) - g(s,\hat{Z}_s)]ds - \int_0^T [Z_s - \hat{Z}_s]dB_s \\ &= Y + \int_0^T g_{\hat{Z}}(s,Z_s^*)ds - \int_0^T Z_s^* dB_s. \end{aligned}\tag{5.38}$$

By hypothesis, we have that $g(t,Z_t) \leq \mu|Z_t|$ and that $-\mu|\hat{Z}_t| \leq g(t,\hat{Z}_t)$, giving

$$g_{\hat{Z}}(t,Z_t^*) = g(t,Z_t) - g(t,\hat{Z}_t) \leq \mu(|Z_t| - |\hat{Z}_t|) \leq \mu|Z_t - \hat{Z}_t| = \mu|Z_t^*|.$$

Clearly $g_{\hat{Z}}$ satisfies the usual assumptions (A1), (A2) and (A3). Now using the comparison theorem, Theorem 3.8 and Equation (5.38), we have that for all $Y \in L^2(\mathcal{F}_T)$

$$\mathcal{E}_g[X+Y] - \mathcal{E}_g[X] \leq \mathcal{E}^\mu(Y). \tag{5.39}$$

Since X is arbitrary and the bound in (5.39) is independent of X, we have the $\mathcal{E}^\mu$-domination of $\mathcal{E}_g[\,\cdot\,]$.

Translability: Since g is independent of y, we know by Theorem 5.2 that $\mathcal{E}[\,\cdot\,|\mathcal{F}_t]$ satisifes the translability condition. $\square$

This leads us to the main proposition of this section, giving the conditions under which a dynamic time-consistent risk measure is induced by a conditional g-expectation.

Proposition 5.11. *Let $(\rho_t)_{t\in[0,T]}$ be a dynamic coherent time-consistent risk measure (resp. convex). Now if $\mathcal{E}[X] = \rho_0(-X)$ is strictly monotone and $\mathcal{E}^\mu$-dominated for some $\mu > 0$, then there exists a unique function g, independent of y, satisfying the usual assumptions (A1), (A2) and (A3) and with $|g(t,z)| \le \mu|z|$, such that for all $X \in L^2(\mathcal{F}_T)$*

$$\rho_0(X) = \mathcal{E}_g[-X] \tag{5.40}$$

$$\text{and} \quad \rho_t(X) = \mathcal{E}_g[-X|\mathcal{F}_t]. \tag{5.41}$$

If in addition g is continuous in t for all $z \in \mathbb{R}^d$, then g is also positively homogeneous and subadditive (resp. convex) in z.

The proof of this proposition relies heavily on Theorem 5.9. Having stated all the results needed for the proof of Proposition 5.11, we can end this section with the proof.

Proof. Set $\mathcal{E}[X] = \rho_0(-X)$. We want to apply Theorem 5.9 to $\mathcal{E}[\,\cdot\,]$. To do so, we need to check that $\mathcal{E}[\,\cdot\,]$ satisfies all the hypothesis of Theorem 5.9.

Nonlinear expectation: For $\mathcal{E}[\,\cdot\,]$ to be a nonlinear expectation, it needs to satisfy constancy and strict monotonicity. The latter is assumed by hypothesis. It is easy to verify that for both the coherent and the convex case $\mathcal{E}[\,\cdot\,]$ satisfies constancy. Thus it follows that $\mathcal{E}[\,\cdot\,]$ is a nonlinear expectation.

$(\mathcal{F}_t)_{t\in[0,T]}$-consistent expectation: By the time-consistency of the risk measure, we have that for all $t \in [0,T]$, $X \in L^2(\mathcal{F}_T)$ and $A \in \mathcal{F}_t$,

$$\begin{aligned} \rho_0(-X\mathbf{1}_A) &= \rho_0[-\rho_t(-X)\mathbf{1}_A] \\ \Longrightarrow \qquad \mathcal{E}[X\mathbf{1}_A] &= \mathcal{E}[\rho_t(-X)\mathbf{1}_A]. \end{aligned}$$

Thus for each $X \in L^2(\mathcal{F}_T)$ and for each $t \in [0,T]$ there exists a random variable $\eta \in L^2(\mathcal{F}_t)$, such that for all $A \in \mathcal{F}_t$

$$\mathcal{E}[\mathbf{1}_A X] = \mathcal{E}[\mathbf{1}_A \eta].$$

More precisely $\eta = \mathcal{E}[X|\mathcal{F}_t] = \rho_t(-X)$ and consequently we have that $\mathcal{E}[\,\cdot\,]$ is a $(\mathcal{F}_t)_{t\in[0,T]}$-consistent expectation.

$\mathcal{E}^\mu$-domination and translability: The $\mathcal{E}^\mu$-domination is assumed by hypothesis and the translability condition of $\mathcal{E}[\,\cdot\,]$ is a consequence of the dynamic translability of the risk measure $(\rho_t)_{t\in[0,T]}$.

Thus $\mathcal{E}[\,\cdot\,]$ satisfies all the hypothesis of Theorem 5.9 and the first part follows i.e.

$$\begin{aligned} &\rho_0(X) = \mathcal{E}_g[-X] \\ \text{and} \quad &\rho_t(X) = \mathcal{E}_g[-X|\mathcal{F}_t]. \end{aligned}$$

Now we furthermore assume that g is continuous in t for all $z \in \mathbb{R}^d$. From Equations (5.40) and (5.41), and by the dynamic coherency (resp. convexity) of $(\rho_t)_{t\in[0,T]}$, we have that $\mathcal{E}_g[\,\cdot\,]$ and $\mathcal{E}_g[\,\cdot\,|\mathcal{F}_t]$ are positively homogeneous and subadditive (resp. convex). By Theorem 5.1 and Theorem 5.4 (resp. Theorem 5.3), we have that g is also positively homogeneous and subadditive (resp. convex) in z. This completes the proof. $\square$

We lastly look at the financial interpretation of the function g in risk measures. The information in this next section has been taken from Rosazza Gianin [30].

5.4 Financial interpretation of g in risk measures

The choice of the function g is important when defining risk measures in terms of g-expectations. The construction of the risk measure depends on the choice of g and thus the conservativeness of the risk measures is largely dependent on this choice. Clearly the preference of the investor also plays an important role in determining what function to use for g.

From Property 3.10 (ii) and Property 3.13 (iii), we recall the monotonicity of g-expectations and of conditional g-expectations. Thus g-expectation is 'increasing with respect to g'. This intuitively tells us that the bigger the function g is, the more conservative the corresponding static risk measure ρ^g and the corresponding dynamic risk measure ρ_t^g are. Thus if we consider two functions, g and $\hat{g}$ with $g \leq \hat{g}$, and the respective risk measures constructed from the g-expectation, ρ^g and $\rho^{\hat{g}}$, then, if a financial position X is $\rho^{\hat{g}}$-acceptable, it is also ρ^g-acceptable. In contrast, if a risky position X is not ρ^g-acceptable, it is also not $\rho^{\hat{g}}$-acceptable. A minimum amount of additional cash $\rho^{\hat{g}}(X)$ needs to be added to the position to make it $\rho^{\hat{g}}$-acceptable.

Chapter 6

Choquet expectation

Many uncertain mathematical phenomena cannot be explained by the linear, classical mathematical expectation. In an attempt to deal with these phenomena, Choquet [8] extended the idea of a probability measure to a nonlinear probability measure, also known as a capacity. He consequently defined the nonlinear Choquet expectation, also known as the Choquet integral. The Choquet integral is a nonlinear generalisation of the Lebesgue integral. It has several properties that allow it to be suitable for pricing insurance contracts or financial assets. Choquet pricing has recently been introduced as an alternative to traditional pricing both in insurance and in finance. For more information on this see Wang [50] and Chateauneuf [4] respectively. Choquet expectations have been very useful in economics, mathematics, finance and physics, however defining conditional Choquet expectations has been an issue in research. In this section we define the concept of a capacity and of Choquet expectation and state some properties of the Choquet integral.

The work in this chapter is based on Chen et al. [5]. Definitions, propositions and proofs have been taken from Chen et al. [5] and from Offwood [37]. Section 6.1 defines the Choquet integral and gives general properties of the Choquet expectation. We also show that the classical mathematical expectation is a special case of the Choquet expectation. Section 6.2 links this nonlinear Choquet expectation to Peng's g-expectation.

6.1 Choquet integral

We again consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 6.1. A set function $V : \mathcal{F} \to [0, 1]$ is called a *capacity* if

(i) $V(\emptyset) = 0$, $V(\Omega) = 1$, and

(ii) if $A, B \in \mathcal{F}$ and $A \subseteq B$, then $V(A) \leq V(B)$.

Noticeably all probability measures are capacities. Also, if we consider a non-decreasing function $h : [0,1] \to [0,1]$ with $h(0) = 0$ and $h(1) = 1$, and a probability measure $\mathbb{P}$, then we can easily verify that $V = h \circ \mathbb{P}$ is a capacity. Clearly $V : \mathcal{F} \to [0,1]$ and

(i) $V(\emptyset) = h \circ \mathbb{P}(\emptyset) = h(0) = 0$, $V(\Omega) = h \circ \mathbb{P}(\Omega) = h(1) = 1$;

(ii) if $A, B \in \mathcal{F}$ and $A \subseteq B$, then $V(A) = h \circ \mathbb{P}(A) \leq h \circ \mathbb{P}(B) = V(B)$. This follows from the definition of a probability measure and since h is a non-decreasing function.

The function h is called a *distortion function* or a *distortion operator*. Let $F_\xi(x) = \mathbb{P}(\xi \leq x)$ be the cumulative distribution function of the random variable ξ on the probability space $\mathbb{P}$. The distortion function h transforms the original probability distribution function F_ξ into a new distribution function $h(F_\xi)$.

Recall that for any probability measure $\mathbb{P}$ and $A, B \in \mathcal{F}$, we know that

$$\mathbb{P}(A \cup B) + \mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B).$$

Following on from this relationship, we can define a concave set function.

Definition 6.2. A set function $V : \mathcal{F} \to [0,1]$ is said to be *concave* if for all $A, B \in \mathcal{F}$ we have

$$V(A \cup B) + V(A \cap B) \leq V(A) + V(B).$$

Definition 6.3. Let V be a capacity and let $\xi \in L^2(\mathcal{F}_T)$. We define the *Choquet expectation* of ξ with respect to the capacity V, denoted $C_V(\xi)$, by

$$C_V(\xi) := \int_{-\infty}^{0} (V(\xi \geq t) - 1)\, dt + \int_{0}^{\infty} V(\xi \geq t)\, dt. \tag{6.1}$$

Definition 6.4. Random variables ξ and $\eta \in L^2(\mathcal{F}_T)$ are said to be *comonotonic* if

$$[\xi(\omega) - \xi(\omega')][\eta(\omega) - \eta(\omega')] \geq 0 \quad \text{for all } \omega, \omega' \in \Omega. \tag{6.2}$$

The following lemma has been taken from Parker [39]. The proofs are a direct application of the definition of comonotonicity, and will not be included.

Lemma 6.5.

(a) Any function f and any constant function are comonotonic.

(b) If f and g are comonotonic, then so are αf and g for all $\alpha > 0$.

(c) *If f and h are comonotonic and g and h are comonotonic, then $f+g$ and h are comonotonic.*

(d) *If f,g and h are pairwise comonotonic, then $\max\{f,g\}$ and h are comonotonic, as are $\min\{f,g\}$ and h.*

(e) *Let $A, B \in \Omega$. Then $\mathbf{1}_A$ and $\mathbf{1}_B$ are comonotonic if and only if $A \subseteq B$ or $B \subseteq A$.*

(f) *Comonotonicity is not transitive as can be seen by considering the indicator functions of subsets of Ω.*

Definition 6.6. A real function F is said to be *comonotonic additive* if for any comonotonic random variables ξ and $\eta \in L^2(\mathcal{F}_T)$ we have

$$F(\xi + \eta) = F(\xi) + F(\eta). \tag{6.3}$$

For a function to be represented by a Choquet expectation, Dellacherie [16] proved that comonotonic additivity is a necessary condition.

The Choquet expectation also satisfies certain properties, which will be listed below.

Proposition 6.7. *The Choquet expectation with respect to the capacity V has the following properties:*

(a) Monotonicity *: For all $\xi, \eta \in L^2(\mathcal{F}_T)$ with $\xi \geq \eta$, we have $C_V(\xi) \geq C_V(\eta)$.*

(b) Positive Homogeneity *: For all $\lambda \geq 0$ and $\xi \in L^2(\mathcal{F}_T)$, we have $C_V(\lambda\xi) = \lambda C_V(\xi)$.*

(c) Translation Invariance *: For all $\xi \in L^2(\mathcal{F}_T)$ and $\alpha \in \mathbb{R}$, we have $C_V(\xi+\alpha) = C_V(\xi) + \alpha$.*

(d) Subadditivity *: For all $\xi, \eta \in L^2(\mathcal{F}_T)$, we have $C_V(\xi + \eta) \leq C_V(\xi) + C_V(\eta)$ if and only if V is concave.*

Remark 6.8. Note that for $A \in \mathcal{F}_T$ we have $C_V(\mathbf{1}_A) = V(A)$. For more information on this, and on any of the other properties of the Choquet integral, see Denneberg [17].

From the properties in Proposition 6.7, it is clear that $\rho(X) = C_V(-X)$ defines a coherent risk measure if and only if the capacity V is concave.

As mentioned before, the Choquet expectation is an extension of the classical, linear expectation, $\mathbb{E}[\,\cdot\,]$. This can be shown by the following. Again let F_ξ be the cumulative distribution function of a random variable ξ with respect to the probability measure $\mathbb{P}$, i.e. $F_\xi(x) = \mathbb{P}(\xi \leq x)$.

Proposition 6.9. *The expected value of ξ can be represented as*

$$\mathbb{E}[\xi] = \int_{-\infty}^{0} (\mathbb{P}(\xi \geq t) - 1)dt + \int_{0}^{\infty} \mathbb{P}(\xi \geq t)dt. \tag{6.4}$$

Proof. Firstly, let $\xi \geq 0$. Then ξ can be represented by $\xi = \int_0^\infty \mathbf{1}_{\{\xi > t\}} dt$. Using Fubini's theorem to change the order of integration, we have

$$\begin{aligned}\mathbb{E}[\xi] &= \int_0^\infty \mathbb{E}[\mathbf{1}_{\{\xi > t\}}]dt \\ &= \int_0^\infty \mathbb{P}(\xi \geq t)dt.\end{aligned}$$

Next, let $\xi \leq 0$. In this case, ξ can be written as $\xi = \int_{-\infty}^{0} -\mathbf{1}_{\{\xi < t\}} dt$. Then

$$\begin{aligned}\mathbb{E}[\xi] &= \int_{-\infty}^{0} -\mathbb{E}[\mathbf{1}_{\{\xi < t\}}]dt \\ &= -\int_{-\infty}^{0} \mathbb{P}(\xi < t)dt \\ &= -\int_{-\infty}^{0} (1 - \mathbb{P}(\xi \geq t))dt \\ &= \int_{-\infty}^{0} (\mathbb{P}(\xi \geq t) - 1)dt.\end{aligned}$$

As $\xi = \xi^+ + \xi^-$, the result follows. □

This representation implies the relationship between the linear mathematical expectation $\mathbb{E}[\cdot]$ and the probability measure $\mathbb{P}$. When Choquet integrals are concerned, we are no longer working in a linear framework. However, we can clearly see the resemblence between the classical mathematical expectation, represented by (6.4) and the Choquet expectation, as seen in Equation (6.1). This representation shows that any linear mathematical expectation can be written as a Choquet expectation. The Choquet integral is thus the natural extension of the mathematical expectation in the nonlinear framework.

Alternatively, Proposition 6.9 can be restated in terms of the cumulative distribution function of ξ.

Proposition 6.10. *The expected value of the random variable ξ can be represented as*

$$\mathbb{E}[\xi] = -\int_{-\infty}^{0} F_\xi(t)dt + \int_{0}^{\infty} (1 - F_\xi(t))dt. \tag{6.5}$$

6.2 Linking Choquet expectation and g-expectation

Having defined Peng's g-expectation in Chapter 3 and the Choquet expectation, a natural question that arises, is what the relationship between these two expectations is. Do there exist conditions under which these two definitions coincide? Chen et al. [5] proved a necessary and sufficient condition to answer this question in the one-dimensional Brownian motion case. It turns out that the classical, linear expectation is the only expectation that falls within both definitions.

We note that Peng's g-expectation is defined only in a BSDE framework. Choquet expectations are defined in a more general setting. Consequently, to explore the relationship between the two, we make suitable restrictions to the Choquet expectations. More precisely, we restrict the Choquet expectations to the domain $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$.

This section deals explicitly with the one-dimensional Brownian motion case. We thus assume, using our previous notation, that $d = 1$ and let y and $z \in \mathbb{R}$. Before we continue, the concept of g-probability needs to be defined.

Definition 6.11. Let $A \in \mathcal{F}_T$. Define the *g-probability* of A by

$$\mathbb{P}_g(A) = \mathcal{E}_g[\mathbf{1}_A]. \tag{6.6}$$

Clearly $\mathbb{P}_g(\,\cdot\,)$ is a capacity.

The main result of this section which is taken from Chen et al. [5] can now be stated. The proof of this theorem, however, relies on several lemmas. We first state and prove these lemmas before proving the following main result.

Theorem 6.12. *Suppose g satisfies assumptions (A1), (A2) and (A3). Then there exists a Choquet expectation, whose restriction to the domain $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$ is equal to a g-expectation if and only if g does not depend on y and is linear in z, i.e. there exists a continuous function υ_t such that*

$$g(t, y, z) = \upsilon_t z. \tag{6.7}$$

We show that if $\mathcal{E}_g[\,\cdot\,]$ is a Choquet expectation on the set of random variables of the form $y + zB_T$, then g is of the form $g(t, z) = \mu(t)|z| + \upsilon_t z$, where $\mu(t)$ and υ_t are continuous functions. The first point in the following lemma shows the uniqueness of a capacity. We also note that in the proofs of Lemma 6.13 (ii) and Theorem 6.12 we only use random variables of the form $y + zB_T$ and $\mathbf{1}_{B_T \in (a,b)}$. Hence Lemma 6.13 (ii) and Theorem 6.12 actually state that if and only if g is linear in z, then the g-expectation is a Choquet expectation on the set of all random variables of the form $f(B_T) \in L(\mathcal{F}_T)$.

Lemma 6.13. *If there exists a capacity, V, such that the associated Choquet expectation on $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$ is equal to the g-expectation, then*

(i) $V(A) = \mathbb{P}_g(A)$ for all $A \in \mathcal{F}_T$.

(ii) There exist two continuous functions $\mu(t)$ and υ_t on $[0, T]$ such that g is of the form

$$g(t, z) = \mu(t)|z| + \upsilon_t z.$$

Proof. (i) Let $\xi \in L^2(\mathcal{F}_T)$. Also let $C_V(\xi)$ be the Choquet expectation of ξ with respect to the capacity V. By hypothesis, we have that for all $\xi \in L^2(\mathcal{F}_T)$

$$\mathcal{E}_g[\xi] = C_V(\xi). \tag{6.8}$$

If we choose $\xi = \mathbf{1}_A$ where $A \in \mathcal{F}_T$, then we have $\mathcal{E}_g[\mathbf{1}_A] = C_V(\mathbf{1}_A)$. By the definition of Choquet expectation we get $C_V(\mathbf{1}_A) = V(A)$. Combining these two equations, we have that $V(A) = C_V(\mathbf{1}_A) = \mathcal{E}_g[\mathbf{1}_A] = \mathbb{P}_g(A)$, where the third equality follows from the definition of the g-probability.

(ii) By hypothesis, $\mathcal{E}_g[\,\cdot\,]$ is a Choquet expectation. Hence, by Dellacherie [16], $\mathcal{E}_g[\,\cdot\,]$ is comonotonic additive, i.e. for comonotonic random variables ξ, $\eta \in L^2(\mathcal{F}_T)$ we have

$$\mathcal{E}_g[\xi + \eta] = \mathcal{E}_g[\xi] + \mathcal{E}_g[\eta]. \tag{6.9}$$

Choose constants (t, y_1, z_1), $(t, y_2, z_2) \in [0, T] \times \mathbb{R}^2$ such that $z_1 z_2 > 0$. Denote $\xi = y_1 + z_1(B_\tau - B_t)$ and $\eta = y_2 + z_2(B_\tau - B_t)$ for $\tau \in [t, T]$. Now clearly ξ and η are independent of $\mathcal{F}_t$ and we can show that ξ and η are comonotonic random variables. Let $\omega, \omega' \in \Omega$. Then

$$\begin{aligned}[\xi(\omega) - \xi(\omega')][\eta(\omega) - \eta(\omega')] &= z_1 z_2 [B_\tau(\omega) - B_t(\omega) + B_\tau(\omega') - B_t(\omega')]^2 \\ &\geq 0.\end{aligned}$$

Note that g satisfies assumptions (A1) and (A3) and is deterministic. Also $y_i, z_i, i = 1, 2$ are constants. Applying Proposition 3.14 we have

$$\mathcal{E}_g[\xi|\mathcal{F}_t] = \mathcal{E}_g[\xi], \quad \mathcal{E}_g[\eta|\mathcal{F}_t] = \mathcal{E}_g[\eta], \quad \mathcal{E}_g[\xi + \eta|\mathcal{F}_t] = \mathcal{E}_g[\xi + \eta]. \tag{6.10}$$

Combining this with Equation (6.9) gives

$$\mathcal{E}_g[\xi + \eta|\mathcal{F}_t] = \mathcal{E}_g[\xi|\mathcal{F}_t] + \mathcal{E}_g[\eta|\mathcal{F}_t],$$

and consequently

$$\frac{\mathcal{E}_g[\xi + \eta|\mathcal{F}_t] - \mathbb{E}[\xi + \eta|\mathcal{F}_t]}{\tau - t} = \frac{\mathcal{E}_g[\xi|\mathcal{F}_t] - \mathbb{E}[\xi|\mathcal{F}_t]}{\tau - t} - \frac{\mathcal{E}_g[\eta|\mathcal{F}_t] - \mathbb{E}[\eta|\mathcal{F}_t]}{\tau - t}.$$

Let $\tau \to t$ on both sides and apply Lemma 3.21, noting the alternate formulation, to get that for all $z_1 z_2 \geq 0$, $y_1, y_2 \in \mathbb{R}$

$$g(t, y_1 + y_2, z_1 + z_2) = g(t, y_1, z_1) + g(t, y_2, z_2). \tag{6.11}$$

This implies that g is linear with respect to y in $\mathbb{R}$ and z in $\mathbb{R}^+$ (or $\mathbb{R}^-$). Applying Equation (6.11) repeatedly gives

$$\begin{aligned} g(t, y, z) &= g(t, y + 0, z\mathbf{1}_{\{z \geq 0\}} + z\mathbf{1}_{\{z \leq 0\}}) \\ &= g(t, y, z\mathbf{1}_{\{z \geq 0\}}) + g(t, 0, z\mathbf{1}_{\{z \leq 0\}}) \\ &= g(t, y + 0, 0 + z\mathbf{1}_{\{z \geq 0\}}) + g(t, 0, -(-z)\mathbf{1}_{\{z \leq 0\}}) \\ &= g(t, y, 0) + g(t, 0, z\mathbf{1}_{\{z \geq 0\}}) + g(t, 0, -(-z)\mathbf{1}_{\{z \leq 0\}}) \\ &= g(t, 0, 1)z\mathbf{1}_{\{z \geq 0\}} - g(t, 0, -1)z\mathbf{1}_{\{z \leq 0\}} \\ &= g(t, 0, 1)z^+ + g(t, 0, -1)(-z)^+ \\ &= g(t, 0, 1)\frac{|z| + z}{2} + g(t, 0, -1)\frac{|z| - z}{2} \\ &= \frac{g(t, 0, 1) + g(t, 0, -1)}{2}|z| + \frac{g(t, 0, 1) + g(t, 0, -1)}{2}z. \end{aligned}$$

Note that $g(t, y, 0) = 0$. Also the second equality follows since $z\mathbf{1}_{\{z \geq 0\}}.z\mathbf{1}_{\{z \leq 0\}} = 0$. Setting

$$\mu(t) = \frac{g(t, 0, 1) + g(t, 0, -1)}{2} \text{ and } \upsilon_t = \frac{g(t, 0, 1) + g(t, 0, -1)}{2}$$

completes the proof. □

We next need to show that for $t \in [0, T]$, $\mu(t) = 0$. To prove this result we require additional lemmas. The following lemma, which is a special case of the comonotonic theorem in Chen, Kulperger and Wei [6], will be stated without proof. The interested reader can refer to Chen et al. [5] for a sketch of the proof.

Lemma 6.14. *Suppose Φ is a function such that $\Phi(B_T) \in L^2(\mathcal{F}_T)$. Let (Y_t, Z_t) be the solution to*

$$Y_t = \Phi(B_T) + \int_t^T \mu(s)|Z_s|ds - \int_t^T Z_s dB_s,$$

where $\mu(t)$ is a continuous function on $[0, T]$. Then

(i) if Φ is increasing, then $Z_t \geq 0$ a.e. for all $t \in [0, T]$, and

(ii) if Φ is decreasing, then $Z_t \leq 0$ a.e. for all $t \in [0, T]$.

Lemma 6.15. *Let $\mu(t)$ be a continuous function on $[0, T]$, $\xi \in L^2(\mathcal{F}_T)$ and let (Y_t, Z_t) be the solution to*

$$Y_t = \xi + \int_t^T \mu(s)|Z_s|ds - \int_t^T Z_s dB_s.$$

Then the following holds.

(i) If $\xi = \mathbf{1}_{\{B_T \geq 1\}}$*, then* $Z_t > 0$ *for all* $t \in [0, T)$*.*

(ii) If $\xi = \Phi(B_T)$*, where* Φ *is a bounded function with strictly positive derivative* Φ'*, then* $Z_t > 0$ *for all* $t \in [0, T)$*.*

(iii) If $\xi = \mathbf{1}_{\{2 \geq B_T \geq 1\}}$*, then* $\mathbb{P} \times \lambda(\{(\omega, t) : Z_t(\omega) < 0\}) > 0$*, where* λ *denotes the Lebesgue measure on* $[0, T)$*, and* $\mathbb{P} \times \lambda$ *denotes the product of the probability space* $\mathbb{P}$ *and the Lebesgue measure* λ*.*

Proof. (i) Since the indicator function $\mathbf{1}_{\{x \geq 1\}}$ is increasing, we can apply Lemma 6.14 (i) and have that $Z_t \geq 0$ a.e. $t \in [0, T]$. This gives us that $|Z_t| = Z_t$. We still need to show that the strict inequality holds.

The BSDE

$$Y_t = \mathbf{1}_{\{B_T \geq 1\}} + \int_t^T \mu(s)|Z_s|ds - \int_t^T Z_s dB_s,$$

is actually a linear BSDE in this case since we have that

$$Y_t = \mathbf{1}_{\{B_T \geq 1\}} + \int_t^T \mu(s)Z_s ds - \int_t^T Z_s dB_s.$$

Let

$$\bar{B}_t = B_t - \int_0^t \mu(s)ds, \tag{6.12}$$

giving us

$$Y_t = \mathbf{1}_{\{B_T \geq 1\}} - \int_t^T Z_s d\bar{B}_s. \tag{6.13}$$

Consider

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left[-\frac{1}{2}\int_0^T \mu^2(s)ds + \int_0^T \mu(s)dB_s\right] \tag{6.14}$$

and define a probability measure $\mathbb{Q}$ by

$$\mathbb{Q}(A) = \int_A \frac{d\mathbb{Q}}{d\mathbb{P}}(\omega)d\mathbb{P}(\omega) \quad \text{for all } A \in \mathcal{F}. \tag{6.15}$$

Using Girsanov's theorem, we know that $(\bar{B}_t)_{t\in[0,T]}$ is a Brownian motion under the probability measure $\mathbb{Q}$.

Taking the conditional expectation $\mathbb{E}^{\mathbb{Q}}[\,\cdot\,|\mathcal{F}_t]$ on both sides of Equation (6.13), gives us

$$\begin{aligned}
Y_t &= \mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\{B_T \geq 1\}}|\mathcal{F}_t] \\
&= \mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\{\bar{B}_T \geq 1 - \int_0^T \mu(s)ds\}}|\mathcal{F}_t] \\
&= \mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\{\bar{B}_T - \bar{B}_t \geq 1 - \int_0^T \mu(s)ds - \bar{B}_t\}}|\mathcal{F}_t] \\
&= \mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\{\bar{B}_T - \bar{B}_t \geq 1 - \int_0^T \mu(s)ds - \bar{B}_t\}}|\sigma(B_t)] \\
&= \mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\{\bar{B}_T - \bar{B}_t \geq 1 - \int_0^T \mu(s)ds - \bar{B}_t\}}|\sigma(\bar{B}_t)].
\end{aligned}$$

For the second equality, we use Equation (6.12). Also we know that $\sigma\{B_s; s \leq t\} = \sigma\{\bar{B}_s; s \leq t\}$ since $\mu(t)$ is a real function in t.

Since $\bar{B}_T - \bar{B}_t$ and $\bar{B}_t$ are independent, we know that $\mathbf{1}_{\{\bar{B}_T - \bar{B}_t \geq 1 - \int_0^T \mu(s)ds - \bar{B}_t\}}$ is independent of $\sigma(\bar{B}_t)$ and hence

$$Y_t = \mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\{\bar{B}_T - \bar{B}_t \geq 1 - \int_0^T \mu(s)ds - k\}}] \mid_{k=\bar{B}_t}.$$

But we know that $\bar{B}_T - \bar{B}_t \sim \mathcal{N}(0, T-t)$, and hence

$$Y_t = \int_{1-\int_0^T \mu(s)ds - k}^{\infty} \phi(x)dx \Big|_{k=\bar{B}_t},$$

where $\phi(x)$ is the density of the normal distribution $\mathcal{N}(0, T-t)$, i.e.

$$\phi(x) = \frac{1}{\sqrt{2\pi(T-t)}} \exp\left[-\frac{x^2}{2(T-t)}\right].$$

Using Corollary 4.1 in El Karoui, Peng and Quenez [35], we get

$$Z_t = \frac{\partial Y_t}{\partial k} \Big|_{k=\bar{B}_t} = \phi\left(1 - \int_0^T \mu(s)ds - \bar{B}_t\right) > 0.$$

Hence $Z_t > 0$ a.e. for all $t \in [0, T]$, which completes the proof.

(ii) This case follows similarly to case (i). We first note that since $\Phi' > 0$, we can again apply Lemma 6.14 (i) and have that $Z_t \geq 0$ a.e. for all $t \in [0, T]$. We again need to show that the inequality is strictly greater than 0.

In this case we have the linear BSDE

$$Y_t = \Phi(B_T) + \int_t^T \mu(s)|Z_s|ds - \int_t^T Z_s dB_s, \qquad 0 \leq t \leq T.$$

Again we let $\bar{B}_t$ be defined by Equation (6.12) giving us

$$Y_t = \Phi(B_T) - \int_t^T Z_s d\bar{B}_s. \tag{6.16}$$

Consider $\frac{d\mathbb{Q}}{d\mathbb{P}}$ and the probability measure $\mathbb{Q}$ as defined by Equations (6.14) and (6.15) respectively. Similarly to part (i), taking the conditional expectation $\mathbb{E}^{\mathbb{Q}}[\,\cdot\,|\mathcal{F}_t]$ on both sides of Equation (6.16), gives us

$$\begin{aligned}
Y_t &= \mathbb{E}^{\mathbb{Q}}[\Phi(B_T)|\mathcal{F}_t] \\
&= \mathbb{E}^{\mathbb{Q}}\left[\Phi\left(\bar{B}_T + \int_0^T \mu(s)ds\right) \Big| \mathcal{F}_t\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[\Phi\left(\bar{B}_T - \bar{B}_t + \int_0^T \mu(s)ds + \bar{B}_t\right) \Big| \mathcal{F}_t\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[\Phi\left(\bar{B}_T - \bar{B}_t + \int_0^T \mu(s)ds + \bar{B}_t\right) \Big| \sigma(B_t)\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[\Phi\left(\bar{B}_T - \bar{B}_t + \int_0^T \mu(s)ds + \bar{B}_t\right) \Big| \sigma(\bar{B}_t)\right].
\end{aligned}$$

Since $\bar{B}_T - \bar{B}_t$ and $\bar{B}_t$ are independent, we know that $\Phi(\bar{B}_T - \bar{B}_t + \int_0^T \mu(s)ds + \bar{B}_t)$ is independent of $\sigma(\bar{B}_t)$ and hence

$$Y_t = \mathbb{E}^{\mathbb{Q}}\left[\Phi\left(\bar{B}_T - \bar{B}_t + \int_0^T \mu(s)ds + k\right)\right] \Big|_{k=\bar{B}_t}.$$

But we still have that $\bar{B}_T - \bar{B}_t \sim \mathcal{N}(0, T-t)$, and hence

$$Y_t = \int_{-\infty}^{\infty} \Phi\left(x + \int_0^T \mu(s)ds + k\right)\phi(x)dx \Big|_{k=\bar{B}_t},$$

where $\phi(x)$ is the density of the normal distribution $\mathcal{N}(0, T-t)$. Using Corollary 4.1 in El Karoui, Peng and Quenez [35], we get that for all $t \in [0, T)$

$$Z_t = \frac{\partial Y_t}{\partial k}\Big|_{k=\bar{B}_t} = \int_{-\infty}^{\infty} \Phi'\left(x + \int_0^T \mu(s)ds + \bar{B}_t\right)\phi(x)dx > 0.$$

This completes the proof.

(iii) Let $\xi = \mathbf{1}_{\{2 \geq B_T \geq 1\}}$. We assume that the conclusion of (ii) is false i.e. we assume that $Z_t \leq 0$ a.e. However by Lemma 6.14 (i) we have that $Z_t \geq 0$ a.e. for all $t \in [0, T]$. Hence we assume that $Z_t = 0$ for all $t \in [0, T)$. This again implies that the BSDE

$$Y_t = \mathbf{1}_{\{2 \geq B_T \geq 1\}} + \int_t^T \mu(s)|Z_s|ds - \int_t^T Z_s dB_s, \qquad 0 \leq t \leq T$$

is a linear BSDE.

Again we let $\bar{B}_t$ be defined by Equation (6.12) giving us

$$Y_t = \mathbf{1}_{\{2 \geq B_T \geq 1\}} - \int_t^T Z_s d\bar{B}_s. \tag{6.17}$$

Consider $\frac{d\mathbb{Q}}{d\mathbb{P}}$ and the probability measure $\mathbb{Q}$ as defined by Equations (6.14) and (6.15) respectively. Using Girsanov's theorem, we know that $(\bar{B}_t)_{t\in[0,T]}$ is a Brownian motion under the probability measure $\mathbb{Q}$.

Set conditional expectation $\mathbb{E}^{\mathbb{Q}}[\,\cdot\,|\mathcal{F}_t]$ on both sides of Equation (6.17). Note that, by the same reasoning as in part (i) and (ii), we have that $\sigma\{B_s; s \leq t\} = \sigma\{\bar{B}_s; s \leq t\}$.

$$\begin{aligned}
Y_t &= \mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\{2 \geq B_T \geq 1\}}|\mathcal{F}_t] \\
&= \mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\{2-\int_0^T \mu(s)ds - \bar{B}_t \geq \bar{B}_T - \bar{B}_t \geq 1 - \int_0^T \mu(s)ds - \bar{B}_t\}}|\mathcal{F}_t] \\
&= \mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\{2-\int_0^T \mu(s)ds - \bar{B}_t \geq \bar{B}_T - \bar{B}_t \geq 1 - \int_0^T \mu(s)ds - \bar{B}_t\}}|\sigma(\bar{B}_t)] \\
&= \mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\{2-\int_0^T \mu(s)ds - k \geq \bar{B}_T - \bar{B}_t \geq 1 - \int_0^T \mu(s)ds - k\}}] \Big|_{k=\bar{B}_t}.
\end{aligned}$$

Now $\bar{B}_T - \bar{B}_t \sim \mathcal{N}(0, T-t)$, and hence

$$Y_t = \int_{1-\int_0^T \mu(s)ds-k}^{2-\int_0^T \mu(s)ds-k} \phi(x)dx \Big|_{k=\bar{B}_t},$$

where $\phi(x)$ is the density of the normal distribution $\mathcal{N}(0, T-t)$. Using the relation between Y_t and Z_t by El Karoui, Peng and Quenez [35], we get that for all $t \in [0, T)$

$$\begin{aligned}
Z_t &= \frac{\partial Y_t}{\partial k} \Big|_{k=\bar{B}_t} \\
&= \phi\left(1 - \int_0^T \mu(s)ds - \bar{B}_t\right) - \phi\left(2 - \int_0^T \mu(s)ds - \bar{B}_t\right) \\
&= \frac{1}{\sqrt{2\pi(T-t)}}\exp\left[-\frac{(1-\int_0^T \mu(s)ds - \bar{B}_t)^2}{2(T-t)}\right] \\
&\quad - \frac{1}{\sqrt{2\pi(T-t)}}\exp\left[-\frac{(2-\int_0^T \mu(s)ds - \bar{B}_t)^2}{2(T-t)}\right].
\end{aligned}$$

From this it is easy to check that for all $t \in [0, T)$

$$\begin{aligned}
Z_t > 0 \quad &\text{when } \bar{B}_t < \frac{3}{2} - \int_0^T \mu(s)ds; \\
Z_t < 0 \quad &\text{when } \bar{B}_t > \frac{3}{2} - \int_0^T \mu(s)ds.
\end{aligned}$$

This implies that for all $t \in [0, T)$,

$$\mathbb{P}(Z_t > 0) > 0, \quad \mathbb{P}(Z_t < 0) > 0 \text{ a.e.}$$

Hence $\mathbb{P} \times \lambda((\omega, t) : Z_t(\omega) < 0) > 0$, which contradicts our original assumption. This completes the proof. ☐

Lemma 6.16. *Suppose g is a convex (or concave) function. If $\mathcal{E}_g[\,\cdot\,]$ is comonotonic additive on $L^2_+(\Omega, \mathcal{F}_T, \mathbb{P})$ (resp. $L^2_-(\Omega, \mathcal{F}_T, \mathbb{P})$), then $\mathcal{E}_g[\,\cdot\,|\mathcal{F}_t]$ is also comonotonic additive on $L^2_+(\Omega, \mathcal{F}_T, \mathbb{P})$ (resp. $L^2_-(\Omega, \mathcal{F}_T, \mathbb{P})$) for all $t \in [0, T]$.*

Proof. We prove the result on $L^2_+(\Omega, \mathcal{F}_T, \mathbb{P})$; the proof on $L^2_-(\Omega, \mathcal{F}_T, \mathbb{P})$ follows similarly. Denote $L^2_+(\Omega, \mathcal{F}_T, \mathbb{P})$ by $L^2_+(\mathcal{F}_T)$ and $L^2_-(\Omega, \mathcal{F}_T, \mathbb{P})$ by $L^2_-(\mathcal{F}_T)$.

Since $\mathcal{E}_g[\,\cdot\,]$ is comonotonic additive on $L^2_+(\Omega, \mathcal{F}_T, \mathbb{P})$, we have that for all comonotonic random variables ξ, $\eta \in L^2_+(\mathcal{F}_T)$

$$\mathcal{E}_g[\xi + \eta] = \mathcal{E}_g[\xi] + \mathcal{E}_g[\eta].$$

We want to show that for all $t \in [0, T]$, and for all comonotonic random variables ξ, $\eta \in L^2_+(\mathcal{F}_T)$

$$\mathcal{E}_g[\xi + \eta|\mathcal{F}_t] = \mathcal{E}_g[\xi|\mathcal{F}_t] + \mathcal{E}_g[\eta|\mathcal{F}_t]. \tag{6.18}$$

We first consider the case where g is a convex function. By Proposition 3.18, we know that for all $t \in [0,T]$

$$\mathcal{E}_g[\xi + \eta|\mathcal{F}_t] \leq \mathcal{E}_g[\xi|\mathcal{F}_t] + \mathcal{E}_g[\eta|\mathcal{F}_t].$$

Assume that (6.18) is false. Then there exists a $t \in [0,T]$ such that

$$\mathbb{P}(\omega; \mathcal{E}_g[\xi + \eta|\mathcal{F}_t] < \mathcal{E}_g[\xi|\mathcal{F}_t] + \mathcal{E}_g[\eta|\mathcal{F}_t]) > 0.$$

Let

$$A = \{\omega; \mathcal{E}_g[\xi + \eta|\mathcal{F}_t] < \mathcal{E}_g[\xi|\mathcal{F}_t] + \mathcal{E}_g[\eta|\mathcal{F}_t]\}.$$

Clearly $A \in \mathcal{F}_T$ and

$$\mathbf{1}_A\mathcal{E}_g[\xi + \eta|\mathcal{F}_t] < \mathbf{1}_A\mathcal{E}_g[\xi|\mathcal{F}_t] + \mathbf{1}_A\mathcal{E}_g[\eta|\mathcal{F}_t].$$

By the comparison theorem, Theorem 3.8, we can take the g-expectation $\mathcal{E}_g[\,\cdot\,]$ on either side of the above inequality, giving

$$\mathcal{E}_g[\mathbf{1}_A\mathcal{E}_g[\xi + \eta|\mathcal{F}_t]] < \mathcal{E}_g\{\mathbf{1}_A\mathcal{E}_g[\xi|\mathcal{F}_t] + \mathbf{1}_A\mathcal{E}_g[\eta|\mathcal{F}_t]\}. \tag{6.19}$$

Applying Proposition 3.18 and the tower property of the g-expectation to the right-hand side of Equation (6.19), we get

$$\begin{aligned}\mathcal{E}_g\{\mathbf{1}_A\mathcal{E}_g[\xi|\mathcal{F}_t] + \mathbf{1}_A\mathcal{E}_g[\eta|\mathcal{F}_t]\} &\leq \mathcal{E}_g\{\mathbf{1}_A\mathcal{E}_g[\xi|\mathcal{F}_t]\} + \mathcal{E}_g\{\mathbf{1}_A\mathcal{E}_g[\eta|\mathcal{F}_t]\} \\ &= \mathcal{E}_g[\mathbf{1}_A\xi] + \mathcal{E}_g[\mathbf{1}_A\eta].\end{aligned}$$

Applying the Tower Property to the left-hand side of (6.19), we have that

$$\mathcal{E}_g[\mathbf{1}_A\mathcal{E}_g[\xi + \eta|\mathcal{F}_t]] = \mathcal{E}_g[\mathbf{1}_A\xi + \mathbf{1}_A\eta].$$

Hence, combining these we get

$$\mathcal{E}_g[\mathbf{1}_A\xi + \mathbf{1}_A\eta] < \mathcal{E}_g[\mathbf{1}_A\xi] + \mathcal{E}_g[\mathbf{1}_A\eta]. \tag{6.20}$$

Since ξ and η are positive and comonotonic, we have that $\mathbf{1}_A\xi$ and $\mathbf{1}_A\eta$ are also positive and comonotonic. However by the assumption that $\mathcal{E}_g[\,\cdot\,]$ is comonotonic additive, we have that

$$\mathcal{E}_g[\mathbf{1}_A\xi + \mathbf{1}_A\eta] = \mathcal{E}_g[\mathbf{1}_A\xi] + \mathcal{E}_g[\mathbf{1}_A\eta]. \tag{6.21}$$

Inequality (6.20) contradicts Equation (6.21) and hence our original assumption that (6.18) is false, cannot hold. Thus for all $t \in [0,T]$,

$$\mathcal{E}_g[\xi + \eta|\mathcal{F}_t] = \mathcal{E}_g[\xi|\mathcal{F}_t] + \mathcal{E}_g[\eta|\mathcal{F}_t]. \tag{6.22}$$

which proves the case when g is convex.

Now let g be a concave function. By Proposition 3.18, we know that for all $t \in [0,T]$

$$\mathcal{E}_g[\xi + \eta|\mathcal{F}_t] \geq \mathcal{E}_g[\xi|\mathcal{F}_t] + \mathcal{E}_g[\eta|\mathcal{F}_t].$$

The rest of the proof follows similarly to the convex case. □

Combining this lemma with Dellacherie's Theorem [16], stating that comonotonic additivity is a necessary condition for a function to be represented by a Choquet expectation, results in the following corollary.

Corollary 6.17. *Suppose g is a convex (or concave) function. If $\mathcal{E}_g[\,\cdot\,]$ is a Choquet expectation on $L^2_+(\Omega, \mathcal{F}_T, \mathbb{P})$ (resp. $L^2_-(\Omega, \mathcal{F}_T, \mathbb{P})$), then $\mathcal{E}_g[\,\cdot\,|\mathcal{F}_t]$ is also a Choquet expectation on $L^2_+(\Omega, \mathcal{F}_T, \mathbb{P})$ (resp. $L^2_-(\Omega, \mathcal{F}_T, \mathbb{P})$) for all $t \in [0,T]$.*

The next couple of lemmas deal with the case when g is of the form $g(t,z) = \mu(t)|z|$, where $\mu(t)$ is a continuous function in t. Clearly, if $\mu(t) \geq 0$ for all $t \in [0,T]$, then g is a convex function and if $\mu(t) \leq 0$ for all $t \in [0,T]$, then g is a concave function.

Lemma 6.18. *Let $\mu(t) \neq 0$ be a continuous function on $[0,T]$ and $g(t,z) = \mu(t)|z|$. Then there exists no Choquet expectation agreeing with $\mathcal{E}^\mu[\,\cdot\,]$ on $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$.*

Proof. Assume that the result is false. Then there exists a Choquet expectation agreeing with $\mathcal{E}^\mu$ on $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$. By Dellacherie's Theorem [16], $\mathcal{E}^\mu$ is comonotonic additive on $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$.

We choose two particular random variables $\xi_1 = \mathbf{1}_{\{B_T \geq 1\}}$ and $\xi_2 = \mathbf{1}_{\{2 \geq B_T \geq 1\}}$. Let $(Y_t^{(i)}, Z_t^{(i)})_{t\in[0,T]}$, $i = 1,2$ be the respective solutions to the following BSDEs

$$Y_t^{(i)} = \xi_i + \int_t^T \mu(s)|Z_s^i|ds - \int_t^T Z_s^i dB_s.$$

Let $(\bar{Y}, \bar{Z})_{t\in[0,T]}$ be the solution to the BSDE

$$\bar{Y}_t = \xi_1 + \xi_2 + \int_t^T \mu(s)|\bar{Z}_s|ds - \int_t^T \bar{Z}_s dB_s.$$

Then

$$Y_t^{(1)} = \mathcal{E}^\mu[\xi_1|\mathcal{F}_t], \quad Y_t^{(2)} = \mathcal{E}^\mu[\xi_2|\mathcal{F}_t], \quad \bar{Y}_t = \mathcal{E}^\mu[\xi_1 + \xi_2|\mathcal{F}_t]$$

Clearly ξ_1 and ξ_2 are positive and comonotonic. Hence by our assumption, $\mathcal{E}^\mu[\,\cdot\,]$ is comonotonic additive with respect to ξ_1 and ξ_2. Lemma 6.16 tells us that $\mathcal{E}^\mu[\,\cdot\,|\mathcal{F}_t]$ is also comonotonic additive with respect to ξ_1 and ξ_2, i.e. for all $t \in [0,T]$

$$\mathcal{E}^\mu[\xi_1 + \xi_2|\mathcal{F}_t] = \mathcal{E}^\mu[\xi_1|\mathcal{F}_t] + \mathcal{E}^\mu[\xi_2|\mathcal{F}_t]$$

which can also be written as

$$\bar{Y}_t = Y_t^{(1)} + Y_t^{(2)}. \tag{6.23}$$

Following from (6.23), we have that for all $t \in [0,T]$

$$\langle \bar{Y}_t, B_t \rangle = \langle Y_t^1, B_t \rangle + \langle Y_t^2, B_t \rangle,$$

where $\langle Y_t, B_t \rangle$ is the finite variation process between the process $Y = (Y_t)_{t\in[0,T]}$ and the Brownian motion $B = (B_t)_{t\in[0,T]}$. But

$$\bar{Z}_t = \frac{d\langle \bar{Y}_t, B_t \rangle}{dt}, \quad Z_t^{(1)} = \frac{d\left\langle Y_t^{(1)}, B_t \right\rangle}{dt}, \quad Z_t^{(2)} = \frac{d\left\langle Y_t^{(2)}, B_t \right\rangle}{dt}.$$

Thus for all $t \in [0,T]$ we have that a.e.

$$\bar{Z}_t = Z_t^{(1)} + Z_t^{(2)}. \tag{6.24}$$

Note that (6.23) can be written as

$$\xi_1 + \xi_2 + \int_t^T \mu(s)|\bar{Z}_s|ds - \int_t^T \bar{Z}_s dB_s = \sum_{i=1}^{2} \left(\xi_i + \int_t^T \mu(s)|Z_s^i|ds - \int_t^T Z_s^i dB_s \right).$$

This gives us

$$\begin{aligned} \int_t^T \mu(s)|\bar{Z}_s|ds &= \int_t^T \mu(s)|Z_s^1|ds + \int_t^T \mu(s)|Z_s^2|ds \\ &= \int_t^T \mu(s)(|Z_s^1| + |Z_s^2|)ds. \end{aligned}$$

We consequently obtain that a.e. for all $t \in [0,T]$

$$\mu(t)|Z_t^{(1)} + Z_t^{(2)}| = \mu(t)|Z_t^{(1)}| + \mu(t)|Z_t^{(2)}|.$$

Since $\mu(t) \neq 0$, we have that a.e.

$$|Z_t^{(1)} + Z_t^{(2)}| = |Z_t^{(1)}| + |Z_t^{(2)}|. \tag{6.25}$$

Clearly (6.25) is only true if $Z_t^{(1)} Z_t^{(2)} \geq 0$. However from Lemma 6.15 we know that $Z_t^{(1)} > 0$ a.e. for all $t \in [0,T]$ and $\mathbb{P} \times \lambda((\omega,t) : Z_t^{(2)}(\omega) < 0) > 0$. Thus $\mathbb{P} \times \lambda((\omega,t) : Z_t^{(1)}(\omega) Z_t^{(2)}(\omega) < 0) > 0$ which implies that

$$\mathbb{P} \times \lambda((\omega,t) : |Z_t^{(1)} + Z_t^{(2)}| < |Z_t^{(1)}| + |Z_t^{(2)}|) > 0,$$

which contradicts (6.25). This completes the proof. □

Lemma 6.19. *Let $\mu(t) \neq 0$ be a continuous function on $[0,T]$. Let $\xi_1 = \mathbf{1}_{\{B_T \geq 1\}}$ and $\xi_2 = \mathbf{1}_{\{2 \geq B_T \geq 1\}}$. Then ξ_1 and ξ_2 are comonotonic, but $\mathcal{E}^\mu[\xi_1 + \xi_2] < \mathcal{E}^\mu[\xi_1] + \mathcal{E}^\mu[\xi_2]$.*

We can finally prove the main theorem of this section, which will first be stated again.

Theorem 6.20. *Suppose g satisfies assumptions (A1), (A2) and (A3). Then there exists a Choquet expectation, whose restriction to the domain $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$ is equal to a g-expectation if and only if g does not depend on y and is linear in z, i.e. there exists a continuous function υ_t such that*

$$g(t, y, z) = \upsilon_t z. \tag{6.26}$$

Proof. Sufficiency: Let $g(t, y, z) = \upsilon_t z$ and let $\xi \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$. We need to show that there exists a Choquet expectation which is equal to the g-expectation. Consider the BSDE

$$Y_t = \xi + \int_t^T \upsilon_s Z_s ds - \int_t^T Z_s dB_s, \qquad 0 \leq t \leq T. \tag{6.27}$$

Let

$$\bar{B}_t = B_t - \int_0^t \upsilon_s ds, \tag{6.28}$$

giving

$$Y_t = \xi - \int_t^T Z_s d\bar{B}_s. \tag{6.29}$$

Define $\frac{d\mathbb{Q}}{d\mathbb{P}}$ and a probability measure $\mathbb{Q}$ by Equations (6.14) and (6.15) respectively. Using Girsanov's theorem, we know that $(\bar{B}_t)_{t\in[0,T]}$ is a Brownian motion under the probability measure $\mathbb{Q}$. Thus

$$\mathcal{E}_g[\xi|\mathcal{F}_t] = \mathbb{E}^{\mathbb{Q}}[\xi|\mathcal{F}_t].$$

In particular, setting $t = 0$, we have that

$$\mathcal{E}_g[\xi] = \mathcal{E}_g[\xi|\mathcal{F}_0] = \mathbb{E}^{\mathbb{Q}}[\xi|\mathcal{F}_0] = \mathbb{E}^{\mathbb{Q}}[\xi].$$

This implies that the g-expectation is a classical mathematical expectation. Also, clearly the classical mathematical expectation can be represented by the Choquet expectation. Consequently there exists a Choquet expectation which coincides with the g-expectation. This completes the sufficiency part of the proof.

Necessity: Let $\xi \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$ and suppose there exists a Choquet expectation which coincides with a g-expectation. By Dellacherie's Theorem [16], $\mathcal{E}_g[\cdot]$ is comonotonic additive on $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$. We need to show that g does not depend on y and has the form $g(t, z) = \upsilon_t z$. By Lemma 6.13, there exist two continuous functions on $[0, T]$ such that

$$g(t, y, z) = \mu(t)|z| + \upsilon_t z.$$

Without loss of generality, we assume that $\upsilon_t = 0$, for $t \in [0,T]$, otherwise by Girsanov's theorem, we can rewrite

$$Y_t = \xi + \int_t^T (\mu(s)|Z_s| + \upsilon_s Z_s)ds - \int_t^T Z_s dB_s,$$

as

$$Y_t = \xi + \int_t^T \mu(s)|Z_s|ds - \int_t^T Z_s d\bar{B}_s,$$

where $\bar{B}_t = B_t - \int_0^t \upsilon_s ds$ and $(\bar{B}_t)_{t\in[0,T]}$ is a Brownian motion under the probability measure $\mathbb{Q}$ as defined previously.

Assume $\mu(t) \neq 0$ for some $t \in [0,T]$. There exists t_0 such that $\mu(t_0) \neq 0$. Without loss of generality, assume $\mu(t_0) > 0$. Since $\mu(t)$ is continuous, there exists a region of t_0, say $[t_1, t_2] \subset [0,T]$ such that $\mu(t) > 0$ for all $t \in [t_1, t_2]$.

Let $\xi_1 = \mathbf{1}_{\{B_{t_2} - B_{t_1} \geq 1\}}$ and $\xi_2 = \mathbf{1}_{\{2 \geq B_{t_2} - B_{t_1} \geq 1\}}$. Then ξ_1 and ξ_2 are comonotonic random variables.

Next we show that

$$\mathcal{E}^\mu[\xi_1 + \xi_2] < \mathcal{E}^\mu[\xi_1] + \mathcal{E}^\mu[\xi_2],$$

which shows that $\mathcal{E}^\mu$ is not comonotonic additive for comonotonic random variable ξ_1 and ξ_2.

Let $\bar{B}_s = \bar{B}_{t_1+s} - \bar{B}_{t_1}$, then $\{\bar{B}_s : 0 \leq s \leq t_2 - t_1\}$ is a Brownian motion under the filtration $(\mathcal{F}'_s)$, where

$$\mathcal{F}'_s = \sigma\{\bar{B}_r : 0 \leq r \leq s\} = \sigma\{\bar{B}_{t_1+r} - \bar{B}_{t_1} : 0 \leq r \leq s\}.$$

Using the above notation, we can write $\xi_1 = \mathbf{1}_{\{\bar{B}_{t_2-t_1} \geq 1\}}$ and $\xi_2 = \mathbf{1}_{\{2 \geq \bar{B}_{t_2-t_1} \geq 1\}}$. Let $a(t) = \mu(t + t_1)$ and let $(y_t^{(i)}, z_t^{(i)}), i = 1, 2$ be the solutions of the following BSDEs with terminal values ξ_1 and ξ_2 respectively:

$$y_t^{(i)} = \xi_i + \int_t^{t_2-t_1} a(s)|z_s^{(i)}|ds - \int_t^{t_2-t_1} z_s^{(i)} d\bar{B}_s, \qquad 0 \leq t \leq t_2 - t_1. \tag{6.30}$$

Let $(\bar{y}_t, \bar{z}_t)$ be the solution of the BSDE with terminal value $\xi_1 + \xi_2$:

$$\bar{y}_t = \xi_1 + \xi_2 + \int_t^{t_2-t_1} a(s)|\bar{z}_s|ds - \int_t^{t_2-t_1} \bar{z}_s d\bar{B}_s, \qquad 0 \leq t \leq t_2 - t_1. \tag{6.31}$$

Since $a(t) = \mu(t + t_1) \neq 0$ for all $t \in [0, t_2 - t_1]$, by Corollary 6.19 we have

$$\bar{y}_t < \bar{y}_t^1 + \bar{y}_t^2. \tag{6.32}$$

On the other hand, consider the BSDEs

$$Y_t^{(i)} = \xi_i + \int_t^T \mu(s)|Z_s^{(i)}|ds - \int_t^T Z_s^{(i)} dB_s, \qquad 0 \leq t \leq T \tag{6.33}$$

and

$$\bar{Y}_t = \xi_1 + \xi_2 + \int_t^T \mu(s)|\bar{Z}_s|ds - \int_t^T \bar{Z}_s dB_s, \qquad 0 \le t \le T. \tag{6.34}$$

Comparing Equation (6.30) with (6.33) and Equation (6.31) with (6.34), we find that for all $t \in [0, t_2 - t_1]$

$$y_t^{(1)} = Y_t^{(1)}, \quad y_t^{(2)} = Y_t^{(1)}, \quad \bar{y}_t = \bar{Y}_t.$$

But

$$Y_t^{(1)} = \mathcal{E}^\mu[\xi_1|\mathcal{F}_t], \quad Y_t^{(2)} = \mathcal{E}^\mu[\xi_2|\mathcal{F}_t], \quad \bar{Y}_t = \mathcal{E}^\mu[\xi_1 + \xi_2|\mathcal{F}_t].$$

Thus, we have that

$$\bar{y}_0^1 = \mathcal{E}^\mu[\xi_1], \quad \bar{y}_0^2 = \mathcal{E}^\mu[\xi_2], \quad \bar{y}_0 = \mathcal{E}^\mu[\xi_i + \xi_2].$$

Applying Equation (6.32),

$$\mathcal{E}^\mu[\xi_i + \xi_2] < \mathcal{E}^\mu[\xi_1] + \mathcal{E}^\mu[\xi_2].$$

This contradicts the comonotonic additivity of $\mathcal{E}_g[\,\cdot\,] = \mathcal{E}^\mu[\,\cdot\,]$. Hence $\mu(t) = 0$ for all $t \in [0, T]$, which completes the proof. $\square$

Theorem 6.12 tells us that if g is nonlinear in z, then the g-expectation is not a Choquet expectation on $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$.

Since the classical mathematical expectation is linear, we know that for all $\xi, \eta \in L^2(\mathcal{F}_T)$

$$\mathbb{E}[\xi + \eta] = \mathbb{E}[\xi] + \mathbb{E}[\eta].$$

For the Choquet expectation, this equality is still true when ξ and η are comonotonic random variables. However, if we consider Peng's g-expectation, with g being a nonlinear function, then the above equality no longer holds, even when ξ and η are comonotonic. We can thus informally say that Peng's g-expectation is 'more nonlinear' than the Choquet expectation on $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$.

Chapter 7

Doob-Meyer decomposition

In Chapter 3 we defined the concept of Peng's g-expectation and showed that many powerful properties still hold true in the nonlinear setting. In fact the notion of martingales, sub- and supermartingales does not need the linearity assumption. Consequently we can define martingales, sub- and supermartingales in the nonlinear setting. This leads to the question, if the well known Doob-Meyer decomposition theorem also holds true in the nonlinear setting. The Doob-Meyer decomposition theorem in the classical theory of martingales shows that certain submartingales can be written as the sum of a martingale and an increasing process. However, since the classical demonstration of the Doob-Meyer decomposition theorem is based on the fact that the expectation $\mathbb{E}[\,\cdot\,]$ is a linear operator, the related nonlinear Doob-Meyer decomposition is not as straightforward.

In this chapter we begin by recalling the classical Doob-Meyer decomposition theorem and any related definitions. Thereafter, in Section 7.2, we define and outline some properties of g-solutions, g-super- and g-subsolutions as well as of g-martingales, g-super- and g-submartingales. Consequently we attempt to find an equivalent Doob-Meyer decomposition for nonlinear g-expectations. This is presented in Section 7.4. To prove the nonlinear version of the Doob-Meyer decomposition however, we require a convergence and a limit theorem. These can be found in Section 7.3. The work in this chapter has been taken from Karatzas and Shreve [34], Cohen [9], Peng [42] and Peng and Xu [41].

7.1 Doob-Meyer decomposition theorem

We begin by defining some concepts required for the Doob-Meyer decomposition theorem and consequently state the original Doob-Meyer decomposition theorem.

Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and assume a continuous setting. The following definitions are taken from Shreve [47] and Karatzas and Shreve [34].

Definition 7.1. An adapted process $(A_t)_{t\in[0,\infty)}$ is called *increasing* $\mathbb{P}$-a.s. if we have that $\mathbb{E}[A_t] < \infty$ for all $t \in [0,\infty)$ and

(a) $A_0 = 0$;

(b) $t \mapsto A_t$ is a nondecreasing, right-continuous function.

An increasing process is called *integrable* if $\mathbb{E}\left[\lim_{t\to\infty} A_t\right] < \infty$.

Definition 7.2. An increasing process $(A_t)_{t\in[0,\infty)}$ is called *natural* if for every bounded right-continuous martingale $(M_t)_{t\in[0,\infty)}$ we have for every $t \in [0,\infty)$

$$\mathbb{E}\int_0^t M_s dA_s = \mathbb{E}\int_0^t M_{s^-} dA_s,$$

where s^- denotes the left limit of s, i.e.

$$M_{s^-} = \lim_{t\to s^-} M_t.$$

Definition 7.3. Consider the class $\mathcal{S}$ and $\mathcal{S}_a$ of all stopping times τ of the filtration $(\mathcal{F}_t)$ which respectively satisfy $\mathbb{P}(\tau < \infty) = 1$ and $\mathbb{P}(\tau < a) = 1$ for a given finite number $a > 0$. The right-continuous process $(X_t)_{t\in[0,\infty)}$ is said to be of class (D) if the family $(X_T)_{T\in\mathcal{S}}$ is uniformly integrable and of class (DL) if the family $(X_\tau)_{\tau\in\mathcal{S}_a}$ is uniformly integrable for every $0 < a < \infty$.

Theorem 7.4 (Doob-Meyer decomposition)**.** *Let $X = (X_t)_{t\in[0,\infty)}$ be a continuous submartingale of class (DL). Then there exists a continuous martingale $M = (M_t)_{t\in[0,\infty)}$ and a continuous increasing process $A = (A_t)_{t\in[0,\infty)}$ such that for all $t > 0$ we have almost surely*

$$X_t = M_t + A_t. \tag{7.1}$$

If, in addition, the increasing process A is taken to be natural, then the processes M and A are uniquely determined up to indistinguishability. Furthermore, if X is of class (D), then M is a uniformly integrable martingale and A is integrable.

The proof of this theorem is given in Karatzas and Shreve [34]. For an extensive study on the Doob-Meyer decomposition, we refer the reader to Karatzas and Shreve [34]. For the purpose of this dissertation, however, the statement of the classical Doob-Meyer decomposition theorem suffices. We want to find a related nonlinear Doob-Meyer decomposition which can be applied to the nonlinear case of Peng's g-expectation.

7.2 g-Solutions and g-martingales

We are working in the framework outlined in Chapter 3. Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ endowed with the filtration $(\mathcal{F}_t^B)_{t\geq 0}$. Let $(B_t)_{t\geq 0}$ be a standard d-dimensional Brownian motion on this space. The filtration is the one generated by this Brownian motion, i.e. $\mathcal{F}_t^B = \sigma\{B_s; 0 \leq s \leq t\}$ for any $t \geq 0$. We also consider a function g satisfying assumptions (A1) and (A2) and a BSDE (3.3) with terminal condition (3.4). Let $D_{\mathcal{F}}^2(T, \mathbb{R}^n)$ denote the space of all $\mathbb{R}^n$-valued, RCLL $\mathcal{F}_t$-progressively measurable processes $(V_t)_{t\in[0,T]}$ with

$$\mathbb{E}\left[\sup_{0\leq s\leq T} |V_s|^2\right] < \infty,$$

where $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^n$. A process $(V_t)_{t\in[0,T]}$ is said to be RCLL if it a.s. has right continuous sample paths which have a left limit. An RCLL process is more commonly known as a *càdlàg process*. Also let $A_{\mathcal{F}}^2(T, \mathbb{R})$ denote the space of all increasing processes $(A_t)_{t\in[0,T]}$ in $D_{\mathcal{F}}^2(T, \mathbb{R})$ with $A_0 = 0$.

We can now define the g-supersolution taken from Peng and Xu [41].

Definition 7.5. A process $(Y_t)_{t\in[0,T]} \in D_{\mathcal{F}}^2(T, \mathbb{R})$ is called a *g-supersolution* if there exists a predictable process $(Z_t)_{t\in[0,T]} \in L_{\mathcal{F}}^2(T, \mathbb{R}^d)$ and an increasing RCLL process $(A_t)_{t\in[0,T]} \in A^2(T, \mathbb{R})$ such that for $t \in [0, T]$

$$Y_t = Y_T + \int_t^T g(s, Y_s, Z_s)ds + A_T - A_t - \int_t^T Z_s dB_s. \tag{7.2}$$

In the above $(Z_t)_{t\in[0,T]}$ is called the martingale part and $(A_t)_{t\in[0,T]}$ is called the increasing part. If $A_t = 0$ for all $t \in [0, T]$, then $(Y_t)_{t\in[0,T]}$ is called a *g-solution*. In this case, we retrieve Equation (3.5). Hence, when $A_t \equiv 0$, the g-solution is equivalent to the g-expectation $\mathcal{E}_g[\,\cdot\,|\mathcal{F}_t]$.

We can also define a g-supersolution and a g-solution in terms of stopping times. If we replace the deterministic terminal time T with a stopping time $\tau \leq T$, we obtain the following definition.

Definition 7.6. For a given stopping time τ, we consider the BSDE

$$Y_t = X + \int_{t\wedge\tau}^{\tau} g(s, Y_s, Z_s)ds + A_\tau - A_{t\wedge\tau} - \int_{t\wedge\tau}^{\tau} Z_s dB_s, \tag{7.3}$$

where $(A_t)_{t\in[0,\tau]} \in A^2(\tau, \mathbb{R})$ is an increasing RCLL process and $Y_\tau = X \in L^2(\mathcal{F}_\tau)$ is the terminal condition of the BSDE. A process $(Y_t)_{t\in[0,\tau]} \in D_{\mathcal{F}}^2(\tau, \mathbb{R})$ is called a *g-supersolution* on $[0, \tau]$ if it solves Equation (7.3). If $A_t = 0$ for all $t \in [0, \tau]$, then $(Y_t)_{t\in[0,\tau]}$ is called a *g-solution*.

A g-solution is uniquely determined if its terminal condition Y_T (or equivalently Y_τ) is given. A g-supersolution is uniquely determined if its terminal condition Y_T and the process $(A_t)_{t\in[0,T]}$ (or equivalently Y_τ and $(A_t)_{t\in[0,\tau]}$) are given. This gives us the following proposition.

Proposition 7.7. *Given $(Y_t)_{t\in[0,\tau]}$ a g-supersolution on $[0,\tau]$, there is a unique $(Z_t)_{t\in[0,\tau]} \in L^2_{\mathcal{F}}(\tau,\mathbb{R}^d)$ and a unique increasing RCLL process $(A_t)_{t\in[0,\tau]} \in A^2(\tau,\mathbb{R})$ such that the triple $(Y_t,Z_t,A_t)_{t\in[0,\tau]}$ satisfies Equation (7.3).*

Proof. Suppose both $(Y_t,Z_t,A_t)_{t\in[0,\tau]}$ and $(Y_t,Z'_t,A'_t)_{t\in[0,\tau]}$ satisfy Equation (7.3). We note that

$$\begin{aligned} Y_t - Y_t = \int_{t\wedge\tau}^{\tau} [g(s,Y_s,Z_s) - g(s,Y_s,Z'_s)]ds + A_\tau - A_{t\wedge\tau} \\ - (A'_\tau - A'_{t\wedge\tau}) - \int_{t\wedge\tau}^{\tau} [Z_s - Z'_s]dB_s, \end{aligned}$$

from which we get $d(Y_t - Y_t)$. Applying Itô's formula to $(Y_t - Y_t)^2$, we get

$$\begin{aligned} d(Y_t - Y_t)^2 &= d(Y_t - Y_t)d(Y_t - Y_t) \\ &= (\Delta(A_t - A'_t))^2 + |Z_t - Z'_t|^2 dt. \end{aligned}$$

Integrating over $[0,\tau]$, subsequently taking expectations and noting that $d(Y_t - Y_t)^2 = 0$, we have that

$$\mathbb{E}\int_0^\tau |Z_s - Z'_s|^2 ds + \mathbb{E}\sum_{t\in(0,\tau]} (\Delta(A_t - A'_t))^2 = 0.$$

Thus we get $Z_t = Z'_t$ for all $t \in [0,\tau]$ and $A_t = A'_t$ for all $t \in [0,\tau]$. $\square$

Definition 7.8. Let $(Y_t)_{t\in[0,\tau]}$ be a g-supersolution on $[0,\tau]$, and let $(Y_t,Z_t,A_t)_{t\in[0,\tau]}$ be the related unique triple from BSDE (7.3). Then we call $(Z_t,A_t)_{t\in[0,\tau]}$ the *decomposition* of $(Y_t)_{t\in[0,\tau]}$.

Given a g-supersolution $(Y_t)_{t\in[0,\tau]}$, the previous proposition showed that this decomposition is unique.

Similarly to the case of g-expectations, we also have the following proposition.

Proposition 7.9. *Let (A1) and (A2) hold true for g and let $(A_t)_{t\in[0,T]} \in A^2(T,\mathbb{R})$. Then there exists a unique pair of processes $(Y_t,Z_t)_{t\in[0,T]} \in L^2_{\mathcal{F}}(T,\mathbb{R}) \times L^2_{\mathcal{F}}(T,\mathbb{R}^d)$ that solves the BSDE (7.2) with terminal condition $Y_T \in L^2(\mathcal{F}_T)$ such that $(Y_t + A_t)_{t\in[0,T]}$ is continuous and such that*

$$\mathbb{E}\left[\sup_{0\le s\le T} |Y_s|^2\right] < \infty.$$

Proof. When $A_t \equiv 0$, we refer to Proposition 3.6. Otherwise, we make the change of variable $\bar{Y}_t := Y_t + A_t$ and get the BSDE

$$\bar{Y}_t = Y_T + A_T + \int_t^T g(s, \bar{Y}_s - A_s, Z_s)ds - \int_t^T Z_s dB_s. \tag{7.4}$$

We again refer back to Proposition 3.6. Also, we have

$$\mathbb{E}\left[\sup_{0\le s\le T} |Y_s|^2\right] < \infty,$$

since

$$\mathbb{E}\left[\sup_{0\le s\le T} |A_s|^2\right] < \infty,$$

$$\mathbb{E}\left[\sup_{0\le t\le T} \left|\int_0^t Z_s dB_s\right|^2\right] < \infty,$$

and

$$\mathbb{E}\int_0^T |g(s, Y_s, Z_s)|^2 ds < \infty.$$

This completes the proof. □

The comparison theorem for g-supersolutions is also of particular interest. It was first introduced by Peng [43]. Two improved version were then given by El Karoui et al. [35]. One of these improved versions is stated below. The proof and additional information can be found in El Karoui et al. [35]. The result of strict comparison was given by Peng [40].

We again consider $(Y_t)_{t\in[0,T]}$ which is the solution of Equation (7.2) recalled here.

$$Y_t = Y_T + \int_t^T g(s, Y_s, Z_s)ds + A_T - A_t - \int_t^T Z_s dB_s.$$

Theorem 7.10 (Comparison theorem)**.** *Let assumptions (A1) and (A2) hold true for g. Let $(\bar{Y}_t, \bar{Z}_t)_{t\in[0,T]}$ be the solution to*

$$\bar{Y}_t = \bar{Y}_T + \int_t^T \bar{g}_t(s, \bar{Y}_s, \bar{Z}_s)ds + \bar{A}_T - \bar{A}_t - \int_t^T \bar{Z}_s dB_s,$$

where $(\bar{g}_t)_{t\in[0,T]} \in L^2_{\mathcal{F}}(T,\mathbb{R})$, $(\bar{A}_t)_{t\in[0,T]} \in A^2(T,\mathbb{R})$ and $\bar{Y}_T \in L^2(\mathcal{F}_T)$ are given such that

(i) $Y_T - \bar{Y}_T \ge 0$,

(ii) $g(\bar{Y}_t, \bar{Z}_t, t) - \bar{g}_t(\bar{Y}_t, \bar{Z}_t, t) \ge 0$ *a.s., a.e.,*

(iii) $A_t - \bar{A}_t$ *is an increasing RCLL process.*

Then we have for all $t \in [0,T]$

$$Y_t \geq \bar{Y}_t \quad a.s., \ a.e.$$

If in addition we assume that $\mathbb{P}(Y_T - \bar{Y}_T > 0) > 0$*, then* $\mathbb{P}(Y_t > \bar{Y}_t) > 0$*. In particular,* $Y_0 > \bar{Y}_0$*.*

We can now define the concept of a g-martingale, g-super- and g-submartingale in a strong sense, as well as in a weak sense.

Definition 7.11. Let $(M_t)_{t\in[0,T]}$ be an $\mathcal{F}_t$-progressively measurable real-valued process such that $\mathbb{E}[|M_t|^2] < \infty$.

(i) The process, M_t, is a *g-martingale* on $[0,T]$, if it is a g-solution on $[0,T]$.

(ii) The process, M_t, is a *g-supermartingale* on $[0,T]$ *in a strong sense* if, for each stopping time $\tau \leq T$, we have that the g-solution $(Y_t)_{t\in[0,\tau]}$ with terminal condition $Y_\tau = M_\tau$ satisfies
$$Y_\sigma \leq M_\sigma$$
for all stopping times $\sigma \leq \tau$.

(iii) The process, M_t, is a *g-submartingale* on $[0,T]$ *in a strong sense* if, for each stopping time $\tau \leq T$, we have that the g-solution $(Y_t)_{t\in[0,\tau]}$ with terminal condition $Y_\tau = M_\tau$ satisfies
$$Y_\sigma \geq M_\sigma$$
for all stopping times $\sigma \leq \tau$.

If we replace the stopping times σ and τ in the previous definition with deterministic times s and t, then we have defined a g-supermartingale and a g-submartingale *in a weak sense.*

We know that a g-solution is a g-martingale. If we assume (A3) in addition to (A1) and (A2), and use the notion of g-expectations, we have that the process $(M_t)_{t\in[0,T]}$ defined by
$$M_t = \mathcal{E}_g[X|\mathcal{F}_t] \tag{7.5}$$
for $t \in [0,T]$ is a g-martingale, where X is the terminal condition of the BSDE. Using this, we have an alternate, more familiar definition for g-martingales, g-supermartingales and g-submartingales.

Definition 7.12. Consider a function g satisfying assumptions (A1), (A2) and (A3). Let the process $(M_t)_{t\in[0,T]}$ satisfy $\mathbb{E}[|M_t|^2] < \infty$.

(i) The process, M_t, is a *g-martingale* in a weak sense if and only if for all $0 \leq s \leq t \leq T$ we have

$$\mathcal{E}_g[M_t|\mathcal{F}_s] = M_s.$$

(ii) The process, M_t, is a *g-submartingale* in a weak sense if and only if for all $0 \leq s \leq t \leq T$ we have

$$\mathcal{E}_g[M_t|\mathcal{F}_s] \geq M_s.$$

(iii) The process, M_t, is a *g-supermartingale* in a weak sense if and only if for all $0 \leq s \leq t \leq T$ we have

$$\mathcal{E}_g[M_t|\mathcal{F}_s] \leq M_s.$$

For the equivalent definition of a g-martingale, g-super- and g-submartingale in a strong sense, we replace the deterministic times s and t with stopping times σ and τ. Clearly a g-supermartingale in a strong sense is also a g-supermartingale in a weak sense. Under certain assumptions a stopped g-martingale, g-supermartingale or g-submartingale remains such. This result corresponds to the optional stopping theorem in the classical theory of martingales. It has been taken from Peng [44].

Proposition 7.13. *Let g satisfy assumptions (A1), (A2) and (A3) and let $(Y_t)_{t\in[0,T]} \in D^2_{\mathcal{F}}(T, \mathbb{R})$ be a g-martingale (resp. g-super-, g-submartingale). Then for all stopping times $0 \leq \sigma \leq \tau \leq T$, we have*

$$\mathcal{E}_g[Y_\tau|\mathcal{F}_\sigma] = Y_\sigma \ \ (resp. \ \leq Y_\sigma, \ \geq Y_\sigma).$$

Note that if $(Y_t, Z_t)_{t\in[0,T]}$ solves the equation

$$Y_t = Y_T + \int_t^T g(s, Y_s, Z_s)ds + A_T - A_t - \int_t^T Z_s dB_s,$$

it is clear that $(-Y_t, -Z_t)_{t\in[0,T]}$ solves the equation

$$-Y_t = -Y_T + \int_t^T -g(s, -(-Y_s), -(-Z_s))ds + (-A_T) - (-A_t) - \int_t^T (-Z_s)dB_s.$$

Hence if $(Y_t)_{t\in[0,T]}$ is a g-martingale (resp. g-super- or g-submartingale), then we have that $(-Y_t)_{t\in[0,T]}$ is a g*-martingale (resp. g*-sub- or g*-supermartingale) where

$$g^*(t, y, z) = -g(t, -y, -z).$$

Consequently, any results applying to g-supermartingales can also be applied to g-submartingales and similarly, any results applying to g-submartingales can be applied to g-supermartingales.

7.3 Limit theorem of g-supersolutions

This section is based on the work done by Peng [41]. We begin by stating the 'convergence theorem' taken from Peng [41]. We omit the proof of this theorem in this dissertation. Using this convergence theorem, we can prove the limit theorem of g-supersolutions. This result is needed for the main theorem of this chapter, the nonlinear Doob-Meyer decomposition theorem.

We first consider the family of semi-martingales

$$Y_t^i = Y_0^i + \int_0^t g_s^i(s, Y_s^i, Z_s^i)ds - A_t^i + \int_0^t Z_s^i dB_s. \tag{7.6}$$

A semi-martingale is defined as follows. The definition is taken from Offwood [37].

Definition 7.14. A regular càdlàg adapted process S_t is a *semi-martingale* if it can be represented as a sum of two processes: a local martingale M_t with $M_0 = 0$ and a process of finite variation A_t with $A_0 = 0$, and

$$S_t = S_0 + M_t + A_t.$$

In the semi-martingale given by Equation (7.6) the process $(g_t^i)_{t\in[0,T]} \in L^2_{\mathcal{F}}(T, \mathbb{R})$ is given and adapted for each i. We also assume that for each i, $(A_t^i)_{t\in[0,T]}$ is a continuous increasing process with $\mathbb{E}[(A_T^i)^2] < \infty$. Furthermore, we assume that

(i) $(g_t^i)_{t\in[0,T]}$ and $(Z_t^i)_{t\in[0,T]}$ are bounded in $L^2_{\mathcal{F}}(T)$ i.e. $\mathbb{E}\int_0^T [|g_s^i|^2 + |Z_s^i|^2]ds \leq C$;

(ii) $\{(Y_t^i)_{t\in[0,T]}\}$ increasingly converges to $(Y_t)_{t\in[0,T]}$ with $\mathbb{E}\left[\sup_{0\leq t\leq T} |Y_t|^2\right] < \infty$.

The limit of $\{(Y_t^i)_{t\in[0,T]}\}$ has the following form

$$Y_t = Y_0 + \int_0^t g_s^0(s, Y_s, Z_s)ds - A_t + \int_0^t Z_s dB_s, \tag{7.7}$$

where $(g_t^0)_{t\in[0,T]}$, $(Z_t)_{t\in[0,T]}$ and $(A_t)_{t\in[0,T]}$ are respectively the L^2-weak limits of $\{(g_t^i)_{t\in[0,T]}\}$, $\{(Z_t^i)_{t\in[0,T]}\}$ and $\{(A_t^i)_{t\in[0,T]}\}$.

Theorem 7.15 (Convergence theorem)**.** *Assume $(g_t^i)_{t\in[0,T]}$ and $(Z_t^i)_{t\in[0,T]}$ are bounded in $L^2_{\mathcal{F}}(T)$ i.e. $\mathbb{E}\int_0^T [|g_s^i|^2 + |Z_s^i|^2]ds \leq C$. Also assume that $\{(Y_t^i)_{t\in[0,T]}\}$ increasingly converges to $(Y_t)_{t\in[0,T]}$ with $\mathbb{E}\left[\sup_{0\leq t\leq T} |Y_t|^2\right] < \infty$ and that $(A_t^i)_{t\in[0,T]}$ is a continuous increasing process with $\mathbb{E}[(A_T^i)^2] < \infty$. The limit $(Y_t)_{t\in[0,T]}$ of $\{(Y_t^i)_{t\in[0,T]}\}$ has the form (7.7), where $(g_t^0)_{t\in[0,T]} \in L^2_{\mathcal{F}}(T, \mathbb{R})$, $(Z_t)_{t\in[0,T]}$ is the weak limit of $\{(Z_t^i)_{t\in[0,T]}\}$ and $(A_t)_{t\in[0,T]}$ is an RCLL square-integrable increasing process. Furthermore, for any $p \in [0, 2)$, $\{(Z_t^i)_{t\in[0,T]}\}$ converges strongly to $(Z_t)_{t\in[0,T]}$ in $L^p_{\mathcal{F}}(T, \mathbb{R}^d)$, i.e. for all $p \in [0, 2)$*

$$\lim_{i\to\infty} \mathbb{E}\int_0^T |Z_s^i - Z_s|^p ds = 0.$$

Consider the sequence of g-supersolutions $(Y_t^i)_{t\in[0,T]}$ solving the BSDEs for $i = 1, 2, \dots$

$$Y_t^i = Y_T^i + \int_t^T g(s, Y_s^i, Z_s^i)ds + A_T^i - A_t^i - \int_t^T Z_s^i dB_s. \tag{7.8}$$

Here the function g satisfies assumptions (A1) and (A2) and $(A_t^i)_{t\in[0,T]}$ is a continuous increasing process for each $i = 1, 2, \dots$ with $\mathbb{E}[(A_T^i)^2] < \infty$. From Proposition 7.7 there exists a unique pair $(Y_t^i, Z_t^i)_{t\in[0,T]} \in L^2_{\mathcal{F}}(T, \mathbb{R}) \times L^2_{\mathcal{F}}(T, \mathbb{R}^d)$ for each $i = 1, 2, \dots$ satisfying BSDE (7.8).

The next theorem tells us that the limit of $\{(Y_t^i)_{t\in[0,T]}\}$ is still a g-supersolution.

Theorem 7.16. *Consider a function g satisfying assumptions (A1) and (A2) and a continuous increasing process $(A_t^i)_{t\in[0,T]}$ with $\mathbb{E}[(A_T^i)^2] < \infty$. For $i = 1, 2, \dots$ let $(Y_t^i, Z_t^i)_{t\in[0,T]}$ be the solution of BSDE (7.8) with $\mathbb{E}\left[\sup_{0\le t\le T} |Y_t^i|^2\right] < \infty$. If $\{(Y_t^i)_{t\in[0,T]}\}$ increasingly converges to $(Y_t)_{t\in[0,T]}$ with $\mathbb{E}\left[\sup_{0\le t\le T} |Y_t|^2\right] < \infty$, then $(Y_t)_{t\in[0,T]}$ is a g-supersolution, i.e. there exist a predictable process $(Z_t)_{t\in[0,T]} \in L^2_{\mathcal{F}}(T, \mathbb{R}^d)$ and an increasing square-integrable RCLL process $(A_t)_{t\in[0,T]}$ such that the pair $(Y_t, Z_t)_{t\in[0,T]}$ is the solution of the BSDE*

$$Y_t = Y_T + \int_t^T g(s, Y_s, Z_s)ds + A_T - A_t - \int_t^T Z_s dB_s. \tag{7.9}$$

Here $(Z_t)_{t\in[0,T]}$ is the weak (resp. strong) limit of $\{(Z_t^i)_{t\in[0,T]}\}$ in $L^2_{\mathcal{F}}(T, \mathbb{R}^d)$ (resp. in $L^p_{\mathcal{F}}(T, \mathbb{R}^d)$ for $p < 2$) and for each $t \in [0, T]$, $(A_t)_{t\in[0,T]}$ is the weak limit of $\{(A_t^i)_{t\in[0,T]}\}$ in $A^2(T, \mathbb{R})$.

Note that Equation (7.8) can be rewritten as

$$Y_t^i = Y_0^i - \int_0^t g(s, Y_s^i, Z_s^i)ds - A_t^i + \int_0^t Z_s^i dB_s$$

and Equation (7.9) can be rewritten as

$$Y_t = Y_0 - \int_0^t g(s, Y_s, Z_s)ds - A_t + \int_0^t Z_s dB_s.$$

To prove this theorem, we need the following additional lemma, telling us that both $\{(Z_t^i)_{t\in[0,T]}\}$ and $\{(A_T^i)^2_{t\in[0,T]}\}$ are uniformly bounded in L^2. The proof of this lemma is based on the proof by Peng [42]. Parts of the proof are based on Cohen's proof of the same theorem in the general framework, given in [9].

Lemma 7.17. *Under the assumptions of Theorem 7.16, there exists a constant C that is independent of i such that*

$$(i)\ \ \mathbb{E}\int_0^T |Z_s^i|^2 ds \le C,$$
$$(ii)\ \ \mathbb{E}[(A_T^i)^2] \le C.$$

Proof. From BSDE (7.8), we have

$$\begin{aligned}
A^i_T &= Y^i_0 - Y^i_T - \int_0^T g(s, Y^i_s, Z^i_s)ds + \int_0^T Z^i_s dB_s \\
&\le |Y^i_0| + |Y^i_T| + \left|\int_0^T g(s, Y^i_s, Z^i_s)ds\right| + \left|\int_0^T Z^i_s dB_s\right| \\
&\le |Y^i_0| + |Y^i_T| + \int_0^T |g(s, Y^i_s, Z^i_s)|ds + \left|\int_0^T Z^i_s dB_s\right|.
\end{aligned}$$

Thus

$$\begin{aligned}
|A^i_T|^2 =& (|Y^i_0| + |Y^i_T|)^2 + \left(\int_0^T |g(s, Y^i_s, Z^i_s)|ds + \left|\int_0^T Z^i_s dB_s\right|\right)^2 \\
&+ 2(|Y^i_0| + |Y^i_T|)\left(\int_0^T |g(s, Y^i_s, Z^i_s)|ds + \left|\int_0^T Z^i_s dB_s\right|\right).
\end{aligned}$$

We know that for $a, b \in \mathbb{R}$, we have that

$$2ab \le \epsilon a^2 + \frac{1}{\epsilon} b^2, \quad \epsilon > 0. \tag{7.10}$$

Thus setting

$$\begin{aligned}
a &= (|Y^i_0| + |Y^i_T|), \\
b &= \left(\int_0^T |g(s, Y^i_s, Z^i_s)|ds + \left|\int_0^T Z^i_s dB_s\right|\right),
\end{aligned}$$

we get that

$$\begin{aligned}
|A^i_T|^2 &\le 2(|Y^i_0| + |Y^i_T|)^2 + 2\left(\int_0^T |g(s, Y^i_s, Z^i_s)|ds + \left|\int_0^T Z^i_s dB_s\right|\right)^2 \\
&= 2(|Y^i_0| + |Y^i_T|)^2 + 2\left(\int_0^T |g(s, Y^i_s, Z^i_s)|ds\right)^2 + 2\left|\int_0^T Z^i_s dB_s\right|^2 \\
&\quad + 4\left(\int_0^T |g(s, Y^i_s, Z^i_s)|ds\right)\left|\int_0^T Z^i_s dB_s\right| \\
&\le 2(|Y^i_0| + |Y^i_T|)^2 + 4\left(\int_0^T |g(s, Y^i_s, Z^i_s)|ds\right)^2 + 4\left|\int_0^T Z^i_s dB_s\right|^2.
\end{aligned}$$

By Jensen's inequality for integrals, we know that for a convex function φ

$$\varphi\left(\int_a^b f(x)dx\right) \le \int_a^b \varphi((b-a)f(x))\frac{1}{b-a}dx.$$

Thus we have that

$$\begin{aligned}
\left(\int_0^T |g(s, Y^i_s, Z^i_s)|ds\right)^2 &\le \int_0^T T^2 |g(s, Y^i_s, Z^i_s)|^2 \frac{1}{T} ds \\
&= T\int_0^T |g(s, Y^i_s, Z^i_s)|^2 ds.
\end{aligned}$$

Consequently

$$|A_T^i|^2 \leq 2(|Y_0^i| + |Y_T^i|)^2 + 4T \int_0^T |g(s, Y_s^i, Z_s^i)|^2 ds + 4 \left| \int_0^T Z_s^i dB_s \right|^2. \tag{7.11}$$

By the triangle inequality and by the Lipschitz continuity of the function g, we have that there exists a $\mu > 0$ such that

$$\begin{aligned} |g(s, Y_s^i, Z_s^i)| - |g(s, 0, 0)| &\leq |g(s, Y_s^i, Z_s^i) - g(s, 0, 0)| \\ &\leq \mu(|Y_s^i| + |Z_s^i|). \end{aligned}$$

Hence

$$|g(s, Y_s^i, Z_s^i)| \leq \mu(|Y_s^i| + |Z_s^i|) + |g(s, 0, 0)|.$$

Squaring the above expression, gives

$$|g(s, Y_s^i, Z_s^i)|^2 \leq \mu^2(|Y_s^i| + |Z_s^i|)^2 + |g(s, 0, 0)|^2 + 2\mu(|Y_s^i| + |Z_s^i|)|g(s, 0, 0)|.$$

We get using Equation (7.10) with $a = \mu(|Y_s^i| + |Z_s^i|)$ and $b = |g(s, 0, 0)|$, that

$$\begin{aligned} |g(s, Y_s^i, Z_s^i)|^2 &\leq 2\mu^2(|Y_s^i| + |Z_s^i|)^2 + 2|g(s, 0, 0)|^2 \\ &\leq 2\mu^2|Y_s^i|^2 + 2\mu^2|Z_s^i|^2 + 4\mu^2|Y_s^i||Z_s^i| + 2|g(s, 0, 0)|^2 \\ &\leq 4\mu^2|Y_s^i|^2 + 4\mu^2|Z_s^i|^2 + 2|g(s, 0, 0)|^2. \end{aligned}$$

Therefore there exists a constant μ' such that

$$|g(s, Y_s^i, Z_s^i)|^2 \leq \mu'|Y_s^i|^2 + \mu'|Z_s^i|^2 + 2|g(s, 0, 0)|^2. \tag{7.12}$$

Plugging Equation (7.12) into Equation (7.11), we get that

$$|A_T^i|^2 \leq 2(|Y_0^i| + |Y_T^i|)^2 + 4T \int_0^T (\mu'|Y_s^i|^2 + \mu'|Z_s^i|^2 + 2|g(s, 0, 0)|^2) ds + 4 \left| \int_0^T Z_s^i dB_s \right|^2.$$

Taking expectation on both sides, leads to

$$\begin{aligned} \mathbb{E}\left[|A_T^i|^2\right] \leq & 2\mathbb{E}\left[(|Y_0^i| + |Y_T^i|)^2\right] + 4T\mu'\mathbb{E}\int_0^T |Y_s^i|^2 ds + 4T\mu'\mathbb{E}\int_0^T |Z_s^i|^2 ds \\ & + 8T\mathbb{E}\int_0^T |g(s, 0, 0)|^2 ds + 4\mathbb{E}\left| \int_0^T Z_s^i dB_s \right|^2. \end{aligned}$$

By Itô isometry, we know that

$$\mathbb{E}\left| \int_0^T Z_s^i dB_s \right|^2 = \mathbb{E}\int_0^T |Z_s^i|^2 ds.$$

This gives us

$$\mathbb{E}\left[|A_T^i|^2\right] \leq 2\mathbb{E}\left[(|Y_0^i| + |Y_T^i|)^2\right] + 4T\mu'\mathbb{E}\int_0^T |Y_s^i|^2 ds + 8T\mathbb{E}\int_0^T |g(s,0,0)|^2 ds$$
$$+ (4T\mu' + 4)\mathbb{E}\int_0^T |Z_s^i|^2 ds.$$

Since $Y_t^1 \leq Y_t^i \leq Y_t$ we observe that $|Y_t^i|$ is dominated by $|Y_t^1| + |Y_t|$. Thus there exists a constant independent of i such that

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |Y_t^i|^2\right] \leq C. \tag{7.13}$$

It follows that there exists a constant C_1, independent of i, such that

$$\mathbb{E}[|A_T^i|^2] \leq C_1 + (4T\mu' + 4)\mathbb{E}\int_0^T |Z_s^i|^2 ds. \tag{7.14}$$

On the other hand, we use Itô's formula applied to $|Y_t^i|^2$. First, we know that

$$dY_t^i = -g(t, Y_t^i, Z_t^i)dt - dA_t^i + Z_t^i dB_t.$$

Applying Itô's formula yields

$$d|Y_t^i|^2 = -2Y_t^i g(t, Y_t^i, Z_t^i)dt - 2Y_t^i dA_t^i + 2Y_t^i Z_t^i dB_t + |Z_t^i|^2 dt.$$

Hence

$$|Y_T^i|^2 - |Y_0^i|^2 = -2\int_0^T Y_s^i g(s, Y_s^i, Z_s^i)ds - 2\int_0^T Y_s^i dA_s^i + 2\int_0^T Y_s^i Z_s^i dB_s + \int_0^T |Z_s^i|^2 ds,$$

giving us that

$$|Y_0^i|^2 + \int_0^T |Z_s^i|^2 ds = |Y_T^i|^2 + 2\int_0^T Y_s^i g(s, Y_s^i, Z_s^i)ds + 2\int_0^T Y_s^i dA_s^i - 2\int_0^T Y_s^i Z_s^i dB_s.$$

Taking expectations and noting that Y_0 is deterministic and that the expected value of an Itô integral is 0, we get using the Lipschitz continuity of g

$$\begin{aligned}
&|Y_0^i|^2 + \mathbb{E}\int_0^T |Z_s^i|^2 ds \\
&\quad = \mathbb{E}[|Y_T^i|^2] + 2\mathbb{E}\int_0^T Y_s^i g(s, Y_s^i, Z_s^i)ds + 2\mathbb{E}\int_0^T Y_s^i dA_s^i \\
&\quad \leq \mathbb{E}[|Y_T^i|^2] + 2\mathbb{E}\int_0^T [|Y_s^i|(\mu|Y_s^i| + \mu|Z_s^i| + |g(s,0,0)|)]ds + 2\mathbb{E}\int_0^T |Y_s^i| dA_s^i \\
&\quad = \mathbb{E}[|Y_T^i|^2] + 2\mathbb{E}\int_0^T [\mu|Y_s^i|^2 + \mu|Y_s^i||Z_s^i| + |Y_s^i||g(s,0,0)|]ds + 2\mathbb{E}\int_0^T |Y_s^i| dA_s^i \\
&\quad \leq \mathbb{E}[|Y_T^i|^2] + 2\mathbb{E}\int_0^T [\mu|Y_s^i|^2 + 4\mu^2|Y_s^i|^2 + \frac{1}{4}|Z_s^i|^2 + |Y_s^i||g(s,0,0)|]ds \\
&\qquad + 2\mathbb{E}\int_0^T |Y_s^i| dA_s^i.
\end{aligned}$$

Therefore, setting

$$a = 4\mathbb{E}\left[\sup_{0\le t\le T}|Y_t^i|^2\right]^{\frac{1}{2}},$$
$$b = \frac{1}{2}\mathbb{E}[|A_T^i|^2]^{\frac{1}{2}},$$

we get using Equation (7.10) that

$$\begin{aligned}
&|Y_0^i|^2 + \mathbb{E}\int_0^T |Z_s^i|^2 ds \\
&\quad\le \mathbb{E}[|Y_T^i|^2] + 2\mathbb{E}\int_0^T [(\mu+4\mu^2)|Y_s^i|^2 + \frac{1}{4}|Z_s^i|^2 + |Y_s^i||g(s,0,0)|]ds \\
&\qquad + 2\mathbb{E}\left[A_T^i \sup_{0\le s\le T}|Y_s^i|\right] \\
&\quad\le C_2 + \frac{1}{2}\mathbb{E}\int_0^T |Z_s^i|^2 ds + 2\mathbb{E}\left[\sup_{0\le t\le T}|Y_t^i|^2\right]^{\frac{1}{2}}\mathbb{E}[|A_T^i|^2]^{\frac{1}{2}}.
\end{aligned}$$

From this we get that

$$\begin{aligned}
\mathbb{E}\int_0^T |Z_s^i|^2 ds &\le 2C_2 - 2|Y_0^i|^2 + 4\mathbb{E}\left[\sup_{0\le t\le T}|Y_t^i|^2\right]^{\frac{1}{2}}\mathbb{E}[|A_T^i|^2]^{\frac{1}{2}} \\
&\le 2C_2 - 2|Y_0^i|^2 + 16\mathbb{E}\left[\sup_{0\le t\le T}|Y_t^i|^2\right] + \frac{1}{4}\mathbb{E}[|A_T^i|^2] \\
&= C_3 + \frac{1}{4}\mathbb{E}[|A_T^i|^2],
\end{aligned}$$

where from (7.13), the constants C_2 and C_3 are independent of i. Combining the previous inequality with (7.14) and setting $C = \max(C_1, C_2)$ gives us that (i) holds true and subsequently (ii) holds true, which completes the proof. $\square$

We can finally prove Theorem 7.16.

Proof. We want to apply the Convergence Theorem 7.15. In BSDE (7.8), we set $g_t^i := -g(t, Y_t^i, Z_t^i)$ which satisfies assumptions (A1) and (A2). By hypothesis we have that $\{(Y_t^i)_{t\in[0,T]}\}$ increasingly converges to $(Y_t)_{t\in[0,T]}$ with $\mathbb{E}[\sup_{0\le t\le T}|Y_t|^2] < \infty$. By Lemma 7.17 we have that $(Z_t^i)_{t\in[0,T]}$ is bounded in $L^2_{\mathcal{F}}(T,\mathbb{R}^d)$. Hence by Theorem 7.15, we have that there exists a $(Z_t^i)_{t\in[0,T]}$ in $L^2_{\mathcal{F}}(T,\mathbb{R}^d)$ such that $\{(Z_t^i)_{t\in[0,T]}\}$ strongly converges to $(Z_t)_{t\in[0,T]}$ in $L^p_{\mathcal{F}}(T,\mathbb{R}^d)$ for all $p\in[0,2)$. As a result $\{(g_t^i)_{t\in[0,T]}\}$ strongly converges in $L^p_{\mathcal{F}}(T,\mathbb{R}^d)$ to $(g_t^0)_{t\in[0,T]}$ and

$$g_s^0 = -g(s, Y_s, Z_s), \quad \text{a.s., a.e.}$$

It follows that $(Y_t, Z_t)_{t\in[0,T]}$ is the solution of the BSDE (7.9). This completes the proof. $\square$

7.4 Nonlinear Doob-Meyer decomposition theorem

By the comparison theorem, Theorem 7.10 we can easily see that a g-supersolution on $[0,T]$ is also a g-supermartingale in both a strong and a weak sense. We are now concerned with the inverse problem: is a right-continuous g-supermartingale also a g-supersolution? This question leads to the nonlinear version of the Doob-Meyer decomposition theorem.

Before stating and proving the main theorem of this chapter, we require an additional lemma. Consider the family of BSDEs parameterised by $i = 1, 2 \dots$

$$y_t^i = Y_T + \int_t^T g(s, y_s^i, Z_s^i)ds + i\int_t^T (Y_s - y_s^i)ds - \int_t^T Z_s^i dB_s. \tag{7.15}$$

We observe in the next lemma that for each $i > 0$, $(y_t^i)_{t\in[0,T]}$ is bounded from above by $(Y_t)_{t\in[0,T]}$. Hence $(y_t^i)_{t\in[0,T]}$ is a g-supersolution.

Lemma 7.18. *Consider the BSDE* (7.15) *for $i = 1, 2 \dots$. Then we have for each i, for $t \in [0,T]$*

$$Y_t \geq y_t^i.$$

Proof. Suppose this is not the case. Then there exists a $\delta > 0$ and a positive integer i such that the measure of $\{(\omega, t) : y_t^i - Y_t - \delta \geq 0\} \subset \Omega \times [0,T]$ is nonzero. We can then define the stopping times

$$\sigma := \inf_t \{y_t^i \geq Y_t + \delta\},$$
$$\tau := \min[T, \inf_{t\geq\sigma}\{y_t^i = Y_t\}].$$

We can see that $\sigma \leq \tau \leq T$ and $\mathbb{P}(\tau > 0) > 0$. Since $Y_t - y_t^i$ is right-continuous, we have

$$(i)\ \ y_\sigma^i \geq Y_\sigma + \delta,$$
$$(ii)\ \ y_\tau^i = Y_\tau.$$

Let $(y_t^i)_{t\in[0,\tau]}$ denote the g-solution with terminal condition y_τ^i and $(Y_t)_{t\in[0,\tau]}$ denote the g-solution with terminal condition Y_τ. By the comparison theorem, Theorem 7.10, we have that $y_\sigma^i = Y_\sigma$. However, on the other hand, since $(Y_t)_{t\in[0,\tau]}$ is a g-submartingale, we have

$$Y_\sigma \geq y_\sigma^i.$$

This contradicts equation (i) above, which completes the proof. □

We can now prove the following nonlinear version of the Doob-Meyer decomposition theorem.

Theorem 7.19. *Consider a function g satisfying assumptions (A1) and (A2). Let $(Y_t)_{t\in[0,T]} \in D^2_{\mathcal{F}}(T,\mathbb{R})$ be a g-supermartingale in a strong sense. Then $(Y_t)_{t\in[0,T]}$ is a g-supersolution on $[0,T]$, i.e. there exists a unique process $(A_t)_{t\in[0,T]} \in A^2(T,\mathbb{R})$ with $\mathbb{E}[(A_T)^2] < \infty$ and a process $(Z_t)_{t\in[0,T]} \in L^2_{\mathcal{F}}(T,\mathbb{R}^d)$ such that*

$$Y_t = Y_T + \int_t^T g(s, y_s, Z_s)ds + A_T - A_t - \int_t^T Z_s dB_s.$$

Proof. *Existence:* Let $(Y_t)_{t\in[0,T]}$ be a g-supermartingale. We consider a sequence of g-supermartingales given by BSDE (7.15). Note that BSDE (7.15) can be rewritten as

$$y_t^i = Y_T + \int_t^T g(s, y_s^i, Z_s^i)ds + A_T^i - A_t^i - \int_t^T Z_s^i dB_s,$$

where

$$A_t^i := i \int_0^t (Y_s - y_s^i)ds.$$

From Lemma 7.17 we have that there exists a constant C independent of i such that

$$\mathbb{E}[|A_T^i|^2] = i^2 \mathbb{E}\left[\left(\int_0^T |Y_s - y_s^i|ds\right)^2\right] \le C.$$

Therefore we have that

$$\mathbb{E}\left[\left(\int_0^T |Y_t - y_t^i|dt\right)^2\right] = 0.$$

It follows that y_t^i converges to Y_t for all $t \in [0,T]$.

Now from Lemma 7.18 we have that $Y_t - y_t^i = |Y_t - y_t^i|$. From the comparison theorem, Theorem 7.10, it follows that $y_t^i \le y_t^{i+1}$. Thus $\{(y_t^i)_{t\in[0,T]}\}$ is a sequence of continuous g-supermartingales, that is monotonically converging to the process $(Y_t)_{t\in[0,T]}$. It is easy to check that all conditions from Thereom 7.16 are satisfied. Hence $(Y_t)_{t\in[0,T]}$ is a g-supersolution on $[0,T]$ of the following form:

$$Y_t = Y_T + \int_t^T g(s, Y_s, Z_s)ds + A_T - A_t - \int_t^T Z_s dB_s.$$

where $(A_t)_{t\in[0,T]}$ is a RCLL increasing process.

Uniqueness: The uniqueness is due to the uniqueness of g-supersolutions, i.e. Proposition 7.7. This completes the proof. □

Corollary 7.20. *Consider a function g independent of y and satisfying assumption (A1), (A2) and (A3). Let $(X_t)_{t\in[0,T]}$ be a g-submartingale on $[0,T]$ in a strong sense satisfying $\mathbb{E}\left[\sup_{0\le s\le T} |X_s|^2\right] < \infty$. Then $(X_t)_{t\in[0,T]}$ has the following decomposition*

$$X_t = M_t - A_t.$$

In this representation $(M_t)_{t\in[0,T]}$ *is a g-martingale of the form* (7.5) *and* $(A_t)_{t\in[0,T]}$ *is an RCLL increasing process with* $A_0 = 0$ *and* $\mathbb{E}[(A_T)^2] < \infty$*. Furthermore, such a decomposition is unique.*

Proof. By Theorem 7.19, we know that the g-submartingale $(X_t)_{t\in[0,T]}$ can be decomposed as follows: there exists a unique RCLL increasing process $(A_t)_{t\in[0,T]} \in A^2(T,\mathbb{R})$ with $\mathbb{E}[(A_T)^2] < \infty$ such that for $t \in [0,T]$ we have

$$X_t = X_T + \int_t^T g(s, Z_s)ds + A_T - A_t - \int_t^T Z_s dB_s. \tag{7.16}$$

Set $M_t = X_t + A_t$. Then for $t \in [0,T]$

$$M_t = X_T + A_T + \int_t^T g(s, Z_s)ds - \int_t^T Z_s dB_s. \tag{7.17}$$

Plugging Equation (7.17) into (7.16), we get

$$X_t = M_t - A_t.$$

Also, noting that g is independent of y, we have from Equation (7.17) that

$$M_t = \mathcal{E}_g[X_T + A_T|\mathcal{F}_t].$$

Clearly $(M_t)_{t\in[0,T]}$ has the form (7.5). Also since g-expectations are g-martingales, we have that $(M_t)_{t\in[0,T]}$ is a g-martingale. This completes the proof. $\square$

Appendix A

Appendix

Theorem A.1 (Fubini's theorem). *Let $(X, \mathcal{F}_X, \mu_X)$ and $(Y, \mathcal{F}_Y, \mu_Y)$ be measure spaces with σ-finite complete measures μ_X and μ_Y defined on the σ-algebras $\mathcal{F}_X$ and $\mathcal{F}_Y$ respectively. If the function $f(x,y)$ is integrable on the product $X \times Y$ of X and Y with respect to the product measure $\mu = \mu_X \times \mu_Y$ of μ_X and μ_Y, then for almost all $y \in Y$ the function $f(x,y)$ of the varibale x is integrable on X with respect to μ_X, the function $g(y) = \int_X f(x,y)d\mu_X$ is integrable on Y with respect to μ_Y and one has the equality*

$$\int_{X\times Y} f(x,y)d\mu = \int_Y d\mu_Y \int_X f(x,y)d\mu_X.$$

Fubini's theorem is valid, in particular for the case when μ_X, μ_Y and μ are the Lebesgue measures in the Euclidean spaces $\mathbb{R}^m$, $\mathbb{R}^n$ and $\mathbb{R}^{m+n}$ respectively.

Fubini's theorem was established by Guido Fubini [29]. An important consequence of Fubini's theorem allows the order of integration to be reversed in iterated integrals.

Theorem A.2 (Lebesgue lemma). *Let f be a Lebesgue integrable function on the interval $[0,T]$. Then, in $[0,T]$, we have*

$$\lim_{n\to\infty} n \int_t^{t+\frac{1}{n}} |f(u) - f(t)|du = 0 \quad dt \text{ a.s.}$$

Appendix B

List of Spaces

For ease of reference, we give an overview of the different spaces we are working in, in this masters dissertation.

$\mathbf{L^2}(\mathcal{F}_\mathbf{t}) = \mathbf{L^2}(\boldsymbol{\Omega}, \mathcal{F}_\mathbf{t}, \mathbb{P})$ denotes the space of all real-valued, $\mathcal{F}_t$-measurable and square integrable random variables applying with the L^2-norm.

$\mathbf{L^2_{\mathcal{F}}}(\mathbf{T}, \mathbb{R}^\mathbf{n})$ denotes the space of all $\mathbb{R}^n$-valued, $\mathcal{F}_T$-adapted processes $(V_t)_{t\in[0,T]}$ with

$$\mathbb{E}\int_0^T |V_t|^2 dt < \infty.$$

$\mathbf{M}_{\mathcal{F}}(\mathbb{R}^\mathbf{n})$ denotes the space of all $\mathbb{R}^n$-valued, $\mathcal{F}_t$-progressively measurable processes $(\psi_t)_{t\in[0,T]}$.

$\mathbf{H^q_{\mathcal{F}}}(\mathbf{T}, \mathbb{R}^\mathbf{n})$ denotes the space of all $(\psi_t)_{t\in[0,T]} \in M_{\mathcal{F}}(\mathbb{R}^n)$ with

$$\mathbb{E}\int_0^T |\psi_t|^q dt < \infty.$$

$\mathbf{D^2_{\mathcal{F}}}(\mathbf{T}, \mathbb{R}^\mathbf{n})$ denotes the space of all $\mathbb{R}^n$-valued, RCLL $\mathcal{F}_t$-progressively measurable processes $(V_t)_{t\in[0,T]}$ with

$$\mathbb{E}\left[\sup_{0\le s\le T} |V_s|^2\right] < \infty.$$

$\mathbf{A^2_{\mathcal{F}}}(\mathbf{T}, \mathbb{R})$ denotes the space of all increasing processes $(A_t)_{t\in[0,T]}$ in $D^2_{\mathcal{F}}(T, \mathbb{R})$ with $A_0 = 0$.

Bibliography

[1] Artzner, P., Delbaen F., Eber, J.M., Heath, D. (1999). *Coherent risk measures.* Mathematical Finance 4, 203-228.

[2] Black, M., Scholes, M. (1973). *The pricing of options and corporate liabilities.* Journal of Political Economy 3, 637-654.

[3] Briand, P., Coquet, F., Hu, Y., Mémin, J., Peng, S. (2000). *A converse comparison theorem for BSDEs and related properties of g-expectation.* Electronic Communications in Probability 5, 101-117.

[4] Chateauneuf, A., Kast, R., Lapied, A. (1996). *Choquet pricing for financial markets with frictions.* Mathematical Finance 6, 323-330.

[5] Chen, Z., Chen, T., Davison, M. (2005). *Choquet expectation and Peng's g-expectation.* The Annals of Probability 3, 1179-1199.

[6] Chen, Z., Kulperger, R., Wei, G. (2005). *A comonotonic theorem of BSDEs.* Stochastic Processes and their Applications 115, 41-54.

[7] Chen, Z. (1998). *A property of backward stochastic differential equations.* Comptes Rendus de l'Academie des Sciences Paris, Série I 326, 483-488.

[8] Choquet, G. (1953). *Theory of capacities.* Annales de l'institut Fourier (Grenoble), 131-195.

[9] Cohen, S.N. (2011). *Representing filtration consistent nonlinear expectations as g-expectations in general probability spaces.* Preprint submitted to Stochastic Processes and their Applications.

[10] Cohen, S.N., Elliot, R.J., Pearce, C.E.M. (2010). *A general comparison theorem for backward stochastic differential equations.* Advances in Applied Probability 42, 878-898.

[11] Cohen, S.N., Elliott, R.J. (2008). *Solutions of backward stochastic differential equations on Markov chains.* Communications on Stochastic Analysis 2, 251-262.

[12] Coquet, F., Hu, Y., Mémin, J., Peng, S. (2002). *Filtration-consistent nonlinear expectations and related g-expectations.* Probability Theory and Related Fields 123, 381-411.

[13] Coquet, F., Hu, Y., Mémin, J., Peng, S. (2001). *A general converse comparison theorem for backward stochastic differential equations.* Comptes Rendus de l'Academie des Sciences Paris, Série I 333, 577-581.

[14] Cvitanic, J., Karatzas, I. (1999). *On dynamic measures of risk.* Finance and Stochastics 3, 451-482.

[15] Delbaen, F. (1998). *Coherent risk measures on general probability spaces.* In: Sandmann, K., Schönbucher, P.J. (Eds.), Advances in Finance and Stochastics, Springer-Verlag, 1-37.

[16] Dellacherie, C. (1970). *Quelques commentaires sur les prolongements de capacites.* Séminaire de Probabilités V. Strasbourg. Lecture Notes in Math. 191, 77-81.

[17] Denneberg, D. (1994). *Non-additive measure and integral.* Kluwer Academic Publishers, Boston.

[18] Detlefsen, K., Scandolo, G. (2005). *Conditional and dynamic convex risk measures.* Finance and Stochastics 9, 539-561.

[19] Duffie, D., Pan, J. (1997). *An overview of value at risk.* The Journal of Derivatives 4, 7-49.

[20] Duffie, D., Epstein, L. (1992). *Stochastic differential utility.* Econometrica 60, 353-394.

[21] Duffie, D., Epstein, L. (1992). *Asset pricing with stochastic differential utility.* Review of Financial Studies 5, 411-436.

[22] Fan, S., Hu, J. (2007). *A limit theorem for solutions to BSDEs in the space of processes.* Statistics and Probability Letters 78, 1024-1033.

[23] Finch, S. (2004). *Ornstein-Uhlenbeck process.* Availabe at paulllac.inria.fr/algo/csolve/ou.pdf.

[24] Föllmer, H., Schied, A. (2002). *Convex measures of risk and trading constraints.* Finance and Stochastic 6, 429-447.

[25] Föllmer, H., Schied, A. (2004). *Stochastic finance, an introduction in discrete time.* Berlin: de Gruyter Studies in Mathematics.

[26] Frittelli, M., Rosazza Gianin, E. (2004). *Dynamic convex risk measures.* In: Szegö, G. (Ed.), Risk measures for the 21st century. J. Wiley, 227-248.

[27] Frittelli, M., Rosazza Gianin, E. (2002). *Putting order in risk measures.* Journal of Banking and Finance 26, 1473-1486.

[28] Frittelli, M. (2000). *Representing sublinear risk measures and pricing rules.* Working paper n. 10 April 2000, Universitá di Milano Bicocca, Italy.

[29] Fubini, G. (1958). *Sugli integrali multipli.* Opere scelte 2, 243-249.

[30] Rosazza Gianin, E. (2006). *Risk measures via g-expectations.* Insurance: Mathematics and Economics 39, 19-34.

[31] Heath, D. (2000). *Back to the future.* In: Plenary lecture at the First World Congress of the Bachelier Society, June 2000, Paris.

[32] Jiang, L. (2008). *Convexity, translation invariance and subadditivity for g-expectations and related risk measures.* The Annals of Applied Probability 18, 245-258.

[33] Jiang, L. (2005). *Representation theorems for generators of backward stochastic differential equations and their applications.* Stochastic Processes and their Applications 115, 1883-1903.

[34] Karatzas, I., Shreve, E.S. (1991). *Brownian motion and stochastic calculus.* 2nd Edition, Springer Verlag, 21-30.

[35] El Karoui, N., Peng, S., Quenez, M. C. (1997). *Backward stochastic differential equations in finance.* Mathematical Finance 7, 1-71.

[36] Merton, R. (1973). *Theory of rational option pricing.* Bell J. Economic and Management Science 4, 141-183.

[37] Offwood, T. (2011). *No free lunch and risk measures on Orlicz spaces.* PhD thesis. University of the Witwatersrand.

[38] Pardoux, E., Peng, S. (1990). *Adapted solution of a backward stochastic differential equation.* Systems and Control Letters 14, 55-61.

[39] Parker, J.M (1996). *A note on comonotonic additivity.* Glasgow Mathematical Journal 38, 199 - 205.

[40] Peng, S. (1997). *Backward SDE and related g-expectations.* In: El Karoui, N., Mazliak, L. (Eds.), Backward Stochastic Differential Equations. In: Pitman Research Notes Mathematics Series 364. Longman, Harlow, 141-159.

[41] Peng, S., Xu, M. (2007). *g_Γ-expectations and related nonlinear Doob-Meyer decomposition theorem.* In: Tang, S.,Yong, J., Control theory and related topics. World Scientific, 122-138.

[42] Peng, S. (1998). *Monotonic limit theorem of BSDE and nonlinear decomposition theorem of Doob-Meyer's type.* Probability Theory and Related Fields 113, 473-499.

[43] Peng, S. (1992). *A generalized dynamic programming principle and Hamilton-Jacobi-Bellman equation.* Stochastics 38, 119-134.

[44] Peng, S. (2005). *Dynamically consistent nonlinear evaluations and expectations.* Preprint available at arXiv:math.PR/0501415.

[45] Peng, S. (2007). *G-expectation, G-Brownian motion and related stochastic calculus of Itô type.* In Stochastic Analysis and Applications, The Able Symposium 2005, Able Symposia 2, Edit Benth et al., 541-567.

[46] Riedel, F. (2004). *Dynamic coherent risk measures.* Stochastic Processes and their Applications 112, 185-200.

[47] Shreve, E.S. (2004). *Stochastic calculus for finance II, continuous-time models.* New York: Springer.

[48] Schroder, M. (1989). *Computing the constant elasticity of variance option pricing formula.* The Journal of Finance 44, 211-215.

[49] Wang, T. (1999). *A class of dynamic risk measures.* Working Paper. University of British Columbia.

[50] Wang, S.S. (1996). *Premium calculation by transforming the layer premium density.* ASTIN Bulletin 26, 71-92.