RECURSIVE UTILITY AND STOCHASTIC DIFFERENTIAL UTILITY: FROM DISCRETE TO CONTINUOUS TIME

THOMAS SEIFERLING

TECHNISCHE UNIVERSITÄT KAISERSLAUTERN

Vom Fachbereich Mathematik der Technischen Universität Kaiserslautern zur Verleihung des akademischen Grades Doktor der Naturwissenschaften (Doctor rerum naturalium, Dr. rer. nat.) genehmigte Dissertation.

1. Gutachter: Prof. Dr. Frank Thomas Seifried
2. Gutachter: Prof. Dr. Mark Schroder

Datum der Disputation: 28. April 2016

D386

Thomas Seiferling
Fachbereich Mathematik
Technische Universität Kaiserslautern
Erwin-Schrödinger-Straße
67663 Kaiserslautern
thomas.seiferling@web.de

Thomas Seiferling: *Recursive Utility and Stochastic Differential Utility: From Discrete to Continuous Time,* vom Fachbereich Mathematik der Technischen Universität Kaiserslautern zur Verleihung des akademischen Grades Doktor der Naturwissenschaften (Doctor rerum naturalium, Dr. rer. nat.) genehmigte Dissertation. 28. April 2016

REFEREES:
Prof. Dr. Frank Thomas Seifried
Prof. Dr. Mark Schroder

ABSTRACT

In this thesis, mathematical research questions related to recursive utility and stochastic differential utility (SDU) are explored.

First, a class of backward equations under nonlinear expectations is investigated: Existence and uniqueness of solutions are established, and the issues of stability and discrete-time approximation are addressed. It is then shown that backward equations of this class naturally appear as a continuous-time limit in the context of recursive utility with nonlinear expectations.

Then, the Epstein-Zin parametrization of SDU is studied. The focus is on specifications with both relative risk aversion and elasitcity of intertemporal substitution greater that one. A concave utility functional is constructed and a utility gradient inequality is established.

Finally, consumption-portfolio problems with recursive preferences and unspanned risk are investigated. The investor's optimal strategies are characterized by a specific semilinear partial differential equation. The solution of this equation is constructed by a fixed point argument, and a corresponding efficient and accurate method to calculate optimal strategies numerically is given.

ZUSAMMENFASSUNG

Die vorliegende Arbeit beschäftigt sich mit mathematischen Fragestellungen aus dem Themenfeld des rekursiven und stochastischen differenziellen Nutzens (SDU).

Zunächst werden Rückwärtsgleichungen unter nichtlinearen Erwartungswerten untersucht. Die Fragen nach Existenz und Eindeutigkeit von Lösungen sowie nach deren Stabilität und Diskretisierbarkeit werden beantwortet. Diese Rückwärtsgleichungen tauchen – wie gezeigt wird – in natürlicher Weise als zeitstetiger Grenzwert des rekursiven Nutzens mit nichtlinearen Erwartungswerten auf.

Der zweite Teil der Arbeit beschäftigt sich mit SDU in der Epstein-Zin-Parametrisierung im Fall von relativer Risikoaversion und intertemporaler Substitutionselastizität größer als Eins. Das zugehörige Nutzenfunktional wird konstruiert; Konkavität und eine Gradientenungleichung werden nachgewiesen.

Die Arbeit schließt mit der Optimierung von Konsum- und Handelsstrategien in unvollständigen Märkten bezüglich rekursiver Präferenzen. Die Lösung ist durch eine semilineare partielle Differentialgleichung gegeben, die mit Fixpunktmethoden behandelt wird. Eine zugehörige schnelle und präzise numerische Methode zur Berechnung der optimalen Strategien wird bereitgestellt.

PUBLICATIONS

The centerpiece of this thesis is composed of three chapters on

- backward nonlinear expectation equations,
- stochastic differential utility, and
- consumption-portfolio optimization with recursive preferences.

The structure and style of the present thesis reflect its origins: Each of the three main chapters consists of a revised and partly extended version of one of the following original research articles, each of which has been submitted for publication in a scientific journal.

Belak, C., T. Seiferling, and F. T. Seifried (2015): "Backward nonlinear expectation equations," available at SSRN: http://ssrn.com/abstract=2547940.

Seiferling, T. and F. T. Seifried (2015): "Stochastic differential utility with preference for information: Existence, uniqueness, concavity, and utility gradients," available at SSRN: http://ssrn.com/abstract=2625800.

Kraft, H., T. Seiferling, and F. T. Seifried (2015): "Asset pricing and consumption-portfolio choice with recursive utility and unspanned risk," available at at SSRN: http://ssrn.com/abstract=2424706.

The idea that the future is unpredictable is undermined every day by the ease with which the past is explained.

— Kahneman (2011)

ACKNOWLEDGMENTS

For his continuous support and guidance and his sustained interest in my work, I am very grateful to my supervisor Prof. Frank Seifried. I owe special thanks to Prof. Mark Schroder for agreeing to act as a referee of my thesis.

I sincerely thank Christoph Belak, Holger Kraft and Frank Seifried for very fruitful collaborations, which have led to three research papers. I would also like to thank my colleagues at the Department of Mathematics and, in particular, the Financial Mathematics Group for the excellent working atmosphere and many productive discussions.

Finally, I am very grateful and honored to have received a PhD grant from the German Academic Scholarship Foundation (Studienstiftung des deutschen Volkes).

CONTENTS

1

INTRODUCTION

This thesis deals with mathematical research questions within the context of *recursive utility* and its continuous-time counterpart *stochastic differential utility*.

Recursive utility is a powerful concept to mathematically describe dynamic risk preferences in *discrete time*. It is an extension of the concept of intertemporal expected utility in the sense of von Neumann and Morgenstern (1944) and has been developed by Kreps and Porteus (1978), Epstein and Zin (1989), Weil (1990), and others. By abandoning the independence axiom, recursive utility allows for a greater flexibility in the modeling of preferences that have both a risk and a time dimension. More precisely, in contrast to the standard discounted expected utility paradigm, recursive utility does not impose a strict relationship between preferences for smoothing across time and smoothing across states.

Stochastic differential utility has been proposed as a *continuous-time* version of recursive utility by Duffie and Epstein (1992b).[1] Duffie and Epstein (1992b) provide heuristic arguments that make a very convincing case for stochastic differential utility as the continuous-time analog of recursive utility; see also Svensson (1989), Skiadas (2008), and the references therein. Kraft and Seifried (2014) were the first to prove a convergence result which yields a mathematically rigorous link between the two concepts: They show that stochastic differential utility emerges in the continuous-time limit of recursive utility in combination with static certainty equivalents of Kreps-Porteus type. One major result of this thesis is the *extension* of their *convergence theorem* to recursive utility in the context of *robust certainty equivalents.*

Recursive preferences play an important role in the asset pricing literature. In the theory of equilibrium asset pricing, the shortcomings of time-additive (non-recursive) preferences become particularly manifest; indeed, the classical consumption-based asset pricing model is well-known to produce inconsistencies when confronted with empirical data. One famous instance is the *equity premium puzzle* of Mehra and Prescott (1985): The excess return of stocks implied by expected utility is significantly too high. In response to these challenges, recursive preferences have become an important tool and are by now ubiquitous in the asset pricing literature; see, among many others, Weil (1989), Duffie and Epstein (1992a), Obstfeld (1994), Duffie et al. (1997), Tallarini (2000), Bansal and Yaron (2004), Hansen et al. (2008),

1 And before that by Epstein (1987), in a deterministic setting.

Guvenen (2009), Kaltenbrunner and Lochstoer (2010), Borovička et al. (2011), Gabaix (2012), and Wachter (2013).

In nearly all of the above publications, the *Epstein-Zin-Weil* parametrization of recursive or stochastic differential utility plays an important role. Despite its prevalent usage, fundamental questions about existence, uniqueness and concavity of stochastic differential utility in the Epstein-Zin parametrization have so far not been entirely solved: The general result of Duffie and Epstein (1992b) relies on a global Lipschitz condition which is not satisfied by the Epstein-Zin specification. Schroder and Skiadas (1999) prove existence and uniqueness of continuous-time Epstein-Zin utility under some parameter restrictions[2] in a Brownian framework. The present thesis complements their findings by establishing existence and uniqueness as well as monotonicity and concavity of Epstein-Zin stochastic differential utility with relative risk aversion (RRA) and elasticity of intertemporal substitution (EIS) both exceeding one in a fully general semimartingale setting.

Following the pioneering work of Duffie and Skiadas (1994), *utility gradients* have proven to be an indispensable tool in the analysis of optimal portfolio allocations and in equilibrium asset pricing. Advancing ideas of Harrison and Kreps (1979), the far-reaching insight of Duffie and Skiadas (1994) is that the first-order optimality condition in the maximization of a utility functional can be formulated as a martingale property of prices, after normalization by the relevant utility gradient. This implies on the one hand that portfolio optimization problems can be addressed directly in terms of the implied first-order conditions; see, e.g., Bank and Riedel (2001a), Schroder and Skiadas (1999, 2003, 2008), Kallsen and Muhle-Karbe (2010), Skiadas (2013), Levental et al. (2013), and the references therein. On the other hand, applying the same line of reasoning to the representative agent's portfolio in general equilibrium, it follows that the state-price deflator in the underlying economy can be represented as a utility gradient. This makes it possible to investigate asset prices, allocations, and efficiency issues in equilibrium. For implementations of this approach, see, among many others, Duffie and Epstein (1992a), Duffie et al. (1994, 1997), Bank and Riedel (2001b), Chen and Epstein (2002), Epstein and Ji (2013) as well as Campbell (2003) and the references therein. In the literature, utility gradients for continuous-time recursive utility have been derived by Duffie and Skiadas (1994) and proven by Schroder and Skiadas (1999) in a Brownian setting (under the same parameter restrictions as for their existence and uniqueness result). The present thesis supplements their analysis by providing

2 These restrictions exclude parameterizations with RRA and EIS both greater than one. These specifications are frequently used (i.e., in the literature on asset pricing with long-run risk started by Bansal and Yaron (2004)); however, Schroder and Skiadas (1999) remark that those parameter restrictions can be finessed by introducing a bequest utility.

the corresponding results for RRA and EIS greater than one within a general semimartingale framework.

In this thesis, utility gradients are used in the context of optimal consumption and portfolio selection, to characterize the value function of an agent with continuous-time Epstein-Zin preferences by a semilinear partial differential equation (PDE). The incomplete-market consumption-portfolio problem under consideration comprises suitably truncated[3] versions of classical frameworks such as Kim and Omberg (1996), Campbell and Viceira (1999), Barberis (2000), Wachter (2002), Chacko and Viceira (2005) and Liu (2007), among others.[4] The nonlinear PDEs that arise in the context of these consumption-portfolio problems are quite challenging, and it has not been clear whether they admit unique smooth solutions at all. So far, researchers have been forced to resort to linearization techniques of unclear precision and to restrict attention to models with affine dynamics; however, even in the presence of affine dynamics, solutions in closed form are available only in the case of unit EIS or similar parameter restrictions, see, e.g., Kraft et al. (2013).

In the present thesis, existence and uniqueness of a classical solution of the relevant semilinear PDE are established by a fixed point argument. Moreover, a numerical method that guarantees a fast and accurate calculation of both indirect utility and – even more importantly – optimal strategies is provided. Hereby, this thesis presents a tractable approach to incomplete-market consumption-portfolio problems with recursive preferences when closed-form solutions are not available.

The remainder of this introduction is structured as follows: First, we briefly review the concepts of recursive utility and stochastic differential utility as well as the convergence result which links the two. Then, we explain the great extent to which this convergence result remains valid in the context of nonlinear expectations. We briefly elaborate on nonlinear expectations in general and, in particular, on the associated backward equations that are the main topic of Chapter 2. After that, the topics of Chapter 3 and Chapter 4 are sketched out. Finally, an outline of the thesis is presented.

---

3 Our analysis in Chapter 4 imposes no structural conditions on the underlying model coefficients, but requires them to be bounded. This formally rules out some popular asset price dynamics, including affine models. However, our results do apply to all such models once they are suitably truncated.

4 With the exception of Chacko and Viceira (2005), these authors assume time-additive CRRA preferences.

1.1 RECURSIVE UTILITY IN DISCRETE AND CONTINUOUS TIME

The general situation is as follows: Given some set $\mathcal{C}$ of $\mathfrak{C}$-valued consumption plans (c_t), one is interested in valuation mappings

$$\nu : \mathcal{C} \to \mathbb{R}, \quad c \mapsto \nu(c)$$

that induce preferences on $\mathcal{C}$ by saying that c' is (weakly) preferred to c if and only if $\nu(c') \geqslant \nu(c)$. A classical example of such a map is the expected utility functional

$$\begin{aligned} & \nu(c) = u^{-1}\big(E\big[\textstyle\sum \beta_{t_k} u(c_{t_k})(t_{k+1} - t_k)\big]\big) && \text{(discrete time)} \quad (1.1) \\ \text{or} \quad & \nu(c) = u^{-1}\big(\textstyle\int \beta_t u(c_t)\, dt\big), && \text{(continuous time)} \end{aligned}$$

defined in terms of subjective discount factors (β_t) and a utility function u like $u(c) = c^{1-\gamma}/(1-\gamma)$.

Recursive utility in discrete time

Recursive utility is a paradigm to construct such valuation mappings/ utility functionals in discrete time. Following Kreps and Porteus (1978) (in the presentation of Kraft and Seifried (2014)), the two main objects in the construction are

(i) an *intertemporal aggregator* W, that is a mapping

$$W : [0, T] \times \mathfrak{C} \times \mathfrak{C} \to \mathfrak{C}, \quad (\Delta, c, \nu) \mapsto W(\Delta, c, \nu)$$

with $W(0, c, \nu) = \nu$ for all $c, \nu \in \mathfrak{C}$, and

(ii) a family $(\mathfrak{m}_t)$ of *certainty equivalents*, i.e., a family of mappings

$$\mathfrak{m}_t : \mathcal{X} \to \mathcal{X}_t, \quad X \mapsto \mathcal{X}_t \quad \text{with} \quad \mathfrak{m}_t(X) = X \quad \text{for all } X \in \mathcal{X}_t,$$

where $\mathcal{X}$ is a suitable set of $\mathfrak{C}$-valued risky (and hence unknown future) payments and $\mathcal{X}_t \subset \mathcal{X}$ is the subset of payments whose values are known at time t.

The current value V_t of a consumption of $c_t\Delta$ over a time-interval $[t, t+\Delta]$ and a risky payment of $V_{t+\Delta}$ at the end of that time-period is calculated as

$$(\Delta, c_t, V_{t+\Delta}) \mapsto V_t = W(\Delta, c_t, \mathfrak{m}_t(V_{t+\Delta})).$$

This aggregation consists of two components: The certainty equivalent (statically) evaluates the risk at time t, while the intertemporal aggregator W assesses the timing of events. This separation allows to disentangle risk and time-preferences; for a discussion hereof, we refer to Skiadas (1998) and the references therein.

Now, the *recursive utility process* V^π on the time-grid $\pi : 0 = t_0 < \cdots < t_n = T$ is given by the recursion

$$V^\pi_{t_k} = W\big(t_{k+1} - t_k, c_{t_k}, m_{t_k}(V^\pi_{t_{k+1}})\big),$$

and then the corresponding *recursive utility functional* is defined via

$$\nu^\pi(c) = V^\pi_0. \tag{1.2}$$

The most widely used certainty equivalents are of the form characterized by Kreps and Porteus (1978) and are given by

$$m_t(X) = u^{-1}\big(E_t[u(X)]\big), \tag{1.3}$$

where $u : \mathfrak{C} \to \mathbb{R}$ is a strictly increasing continuous function and E_t denotes the conditional expectation given information available at time t. Usually, u will be a von Neumann-Morgenstern utility function. A corresponding general class of discrete-time aggregators is given in terms of a strictly increasing function $g : \mathfrak{C} \subset [0, \infty) \to \mathbb{R}$, via

$$W(\Delta, c, \nu) = g^{-1}\big(\delta\Delta g(c) + (1 - \delta\Delta)g(\nu)\big). \tag{1.4}$$

Note that with the choice $u = g$, the recursive utility functional (1.2) becomes the expected utility functional in (1.1). The popular *Epstein-Zin-Weil parametrization* of recursive utility is obtained by choosing $\mathfrak{C} = (0, \infty)$ and the isoelastic utility functions

$$\begin{aligned} u(c) = \tfrac{1}{1-\gamma}c^{1-\gamma}, \qquad g(c) = \tfrac{1}{1-\phi}c^{1-\phi}, \qquad & \gamma, \phi \in (0, \infty),\ \gamma, \phi \neq 1, \\ u(c) = g(c) = \log c, \qquad & \gamma = \phi = 1, \end{aligned}$$

in (1.3) and (1.4). An agent with the corresponding recursive utility functional has *relative risk aversion* (RRA) γ and *elasticity of intertemporal substitution* (EIS) $\psi = 1/\phi$; see Epstein and Zin (1989) and Weil (1990).

Stochastic differential utility

In the context of stochastic differential utility (SDU), as introduced by Duffie and Epstein (1992b), the utility functional is defined as

$$\nu(c) = V_0,$$

where $V = (V_t)$ is given in terms of a *continuous-time aggregator* function $f : \mathfrak{C} \times \mathbb{R} \to \mathbb{R}$, as the solution of the backward stochastic differential equation (BSDE)

$$\begin{aligned} dV_t = -f(c_t, V_t)dt + dM_t \quad & \text{for a martingale } M, & V_T = \xi, \\ \text{or, equivalently,} \quad & V_t = E_t\Big[\int_t^T f(c_s, V_s)ds + \xi\Big], & t \in [0, T]. \end{aligned} \tag{1.5}$$

In fact, Duffie and Epstein (1992b) were the first to consider BSDEs in the context of general semimartingales. In a Brownian framework, BSDEs had previously been introduced by Pardoux and Peng (1990). In the Brownian setting, BSDEs take the form

$$dV_t = -f(t, V_t, Z_t)dt + Z_t dW_t, \qquad V_T = \xi,$$

where the aggregator f may also depend on the martingale integrand Z of the solution.[5] For an overview of the theory of BSDEs and their applications in finance, we refer to El Karoui et al. (1997).

The *Epstein-Zin* parameterization of SDU (for non-unit EIS and RRA) corresponds to the continuous-time aggregator

$$f(c, v) \triangleq \delta v \frac{1-\gamma}{1-\frac{1}{\psi}}\left[\left(\frac{c}{((1-\gamma)v)^{\frac{1}{1-\gamma}}}\right)^{1-\frac{1}{\psi}} - 1\right], \tag{1.6}$$

see, e.g., Duffie and Epstein (1992b).

Convergence of recursive utility

Under appropriate continuity, differentiability and growth conditions on the intertemporal aggregator W and the utility function u inducing the certainty equivalent via (1.3), Kraft and Seifried (2014) prove that the recursive utility process V^π converges to the stochastic differential utility process V in the limit of vanishing grid size, i.e.,

$$V^\pi \to V \quad \text{as} \quad |\pi| = \max_{k=1,\dots,n} |t_k - t_{k-1}| \to 0.$$

More precisely, they show that $\overline{V}^\pi = u(V^\pi) \to u(V) = \overline{V}$, where

- $\overline{V}^\pi$ is the recursive utility process corresponding to the *normalized* intertemporal aggregator

  $$\overline{W}(\Delta, c, \bar{v}) = u \circ W(\Delta, c, u^{-1}(\bar{v}))$$

  and the family of *normalized* certainty equivalents

  $$\overline{m}_t(\bar{X}) = u\big(m_t(u^{-1}(\bar{X}))\big) = E_t\left[\bar{X}\right] \qquad \text{and}$$

- $\overline{V}$ is given by (1.5) with continuous-time aggregator

  $$f(c, \bar{v}) = \tfrac{\partial}{\partial \Delta}\overline{W}(\Delta, c, \bar{v})|_{\Delta=0},$$

  confirming a formula derived by Epstein (1987).

As long as only bounded consumption plans are considered, Kraft and Seifried (2014) cover all specifications of recursive utility with certainty equivalents and aggregators as in (1.3) and (1.4) and, in particular, the Epstein-Zin-Weil parametrization.

5 This has been exploited (e.g., in Lazrak and Quenez (2003)) to formulate SDU with source-dependent risk aversion.

1.2 NONLINEAR EXPECTATIONS

In this thesis, we generalize the convergence result of Kraft and Seifried (2014) to normalized certainty equivalents $\overline{m}_t(X) = \mathcal{E}_t[X]$, where $(\mathcal{E}_t)_{t\in[0,T]}$ is a (dynamic) *nonlinear expectation*. Prominent examples of (static) nonlinear expectations are given by *lower expectations*,[6]

$$\mathcal{E}_t[X] = \inf_{q\in Q_t} E_t^q[X],$$

where Q_t is a family of probability measures. Lower expectations $\mathcal{E}_t$ emerge as normalized versions of *(max)min expected utility* certainty equivalents under multiple priors,

$$m_t(X) = \inf_{q\in Q_t} u^{-1}\big(E_t^q[u(X)]\big),$$

as proposed by Gilboa and Schmeidler (1989). This robust approach is a conservative alternative to classical expected utility certainty equivalents in situations where precise subjective probabilities cannot be assigned. Epstein and Wang (1994), Epstein and Schneider (2003) and Chen and Epstein (2002) combine (max)min expected utility certainty equivalents with recursive utility in discrete time. To consider corresponding recursive utilites in continuous time, the family $(\mathcal{E}_t)_{t\in[0,T]}$ must be *dynamically consistent*, i.e., it must satisfy the tower property

$$\mathcal{E}_s\big[\mathcal{E}_t[X]\big] = \mathcal{E}_s[X] \quad \text{for all } s \leqslant t. \tag{TC}$$

This is a serious restriction on the family $(Q_t)_{t\in[0,T]}$ of probability measures. In their continuous-time intertemporal version of multiple-priors utility, Chen and Epstein (2002) guarantee this *time-consistency* property by a so-called *rectangularity condition*. In the version of stochastic differential utility for ambiguous volatility proposed recently by Epstein and Ji (2014),[7] time-consistency is guaranteed by optimal control techniques; see in particular Nutz (2012, 2013) for the precise constructions and the dynamic programming principle.

In this thesis, we focus on (dynamic) nonlinear expectations which satisfy the time-consistency condition (TC). Under assumptions on the intertemporal aggregator paralleling those of Kraft and Seifried (2014), we prove that the corresponding recursive utility processes $\overline{V}^\pi$ converge to a limiting object $\overline{V}$ (Theorem 2.80), where $\overline{V}$ is the solution of a *backward nonlinear expectation equation* (BNEE) of the form

$$\overline{V}_t = \mathcal{E}_t\Big[\int_t^T f(c_s, \overline{V}_s)\, ds + \xi\Big], \quad t\in[0,T],$$

given in terms of the continuous-time aggregator

$$f(t,\bar{v}) = \tfrac{\partial}{\partial\Delta}\overline{W}(\Delta, c, \bar{v})|_{\Delta=0}.$$

In particular, we substantiate the results of Chen and Epstein (2002) and Epstein and Ji (2014) by showing that their models emerge in the continuous-time limit of recursive utilities under multiple priors.

6 See, e.g., Huber (1981).
7 See Epstein and Ji (2013) for applications of this model.

Backward nonlinear expectation equations

To prove the convergence result advertised above, in Chapter 2,[8] we study general backward nonlinear expectation equations (BNEEs),

$$X_t = \mathcal{E}_t\Big[\int_t^T g(s,X)\,\mu(ds) + \xi\Big], \quad t \in [0,T]. \tag{1.7}$$

In our analysis of (1.7), we build on an abstract framework for dynamic nonlinear expectations. This setting comprises g-expectations, see, e.g., Peng (1997, 2004b), G-expectations as put forward by Peng (2007, 2008), and random G-expectations as introduced and analyzed by Nutz (2012, 2013). In particular, the model of Chen and Epstein (2002) and a version of then one of Epstein and Ji (2014) are included in our abstract framework.

Equations of the form (1.7) should be looked upon as backward stochastic differential equations (BSDEs) in the context of nonlinear expectations. In the literature, such equations have previously been investigated in particular situations: For instance, Peng (2004b, 2005a) considers BSDEs under g-expectations, and Peng (2010) and Hu et al. (2014a,b) study BSDEs under G-expectations. Moreover, with so-called second-order BSDEs, Cheridito et al. (2007) and Soner et al. (2012, 2013) have introduced a related class of equations. In the economics literature, BNEEs have appeared in the context of dynamic robust risk preferences; see, e.g., Chen and Epstein (2002), Hayashi (2005) and Epstein and Ji (2014).

Results

In Chapter 2, we prove existence and uniqueness for solutions of (1.7) under a Lipschitz condition on the aggregator g (Theorem 2.26), and we show that solutions are stable under perturbations of the aggregator function (Theorem 2.29). Moreover, we demonstrate that (1.7) emerges in the continuous-time limit of the discrete aggregations

$$X_k^\Delta \triangleq \mathcal{E}_{t_k}\Big[m\big((t_k,t_{k+1}]\big)\, g\big(t_k, \mathcal{E}_{t_k}[X_{k+1}^\Delta]\big) + X_{k+1}^\Delta\Big], \quad k = N-1,\dots,0,$$

where $\Delta : 0 = t_0 < t_1 < \cdots < t_N = T$ (Theorem 2.32). Building on these results, convergence of recursive utilities (Theorem 2.80) is proven.

1.3 EPSTEIN-ZIN SDU AND UTILITY GRADIENTS

In Chapter 3,[9] we return to the classical probabilistic framework and work on a filtered probability space $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t\in[0,T]}, P)$ satisfying the usual conditions. We consider the Epstein-Zin-Weil parameterization of stochastic differential utility as introduced in (1.6), focusing on the case where RRA and EIS are greater than one.

8 Chapter 2 is largely based on Belak, Seiferling, and Seifried (2015).
9 Chapter 3 is largely based on Seiferling and Seifried (2015).

For every consumption plan c with

$$\mathrm{E}\big[\int_0^T c_t^r dt + c_T^r\big] < \infty \qquad \text{for all } r \in \mathbb{R},$$

we show that there exists a unique semimartingale V^c satsifying (1.5) (Theorem 3.2). Moreover, we prove that the mapping $c \mapsto V^c$ is increasing and concave (Theorem 3.4). Finally, we establish a utility gradient inequality (Theorem 3.5) of the form

$$\nu(c) \leqslant \nu(\hat{c}) + \langle m(\hat{c}), c - \hat{c} \rangle.$$

1.4 CONSUMPTION-PORTFOLIO OPTIMIZATION WITH SDU

In Chapter 4,[10] we consider the consumption and portfolio selection problem of an investor with continuous-time Epstein-Zin preferences in an incomplete financial market consisting of a locally risk-free asset M and a risky asset S with dynamics[11]

$$dM_t = r(Y_t)M_t dt, \qquad dS_t = S_t\left[(r + \lambda(Y_t))dt + \sigma(Y_t)dW_t\right].$$

The investor consumes at rate c and invests the fraction π of his capital into S, and so his wealth follows the dynamics

$$dX_t^{\pi,c} = X_t^{\pi,c}\left[(r(Y_t) + \pi_t\lambda(Y_t))dt + \pi_t\sigma(Y_t)dW_t\right] - c_t dt.$$

Utility of terminal wealth and consumption is given by

$$\nu(c) \triangleq V_0^c, \quad \text{where } V_t^c \triangleq \mathrm{E}_t\left[\int_t^T f(c_s, V_s^c)ds + U(X_T^{\pi,c})\right] \text{ for } t \in [0,T].$$

Here, f is the continuous-time Epstein-Zin aggregator as in (1.6). We study the associated optimal consumption and portfolio optimization problem, i.e., we aim to

$$\text{find} \quad (\pi^\star, c^\star) \quad \text{such that} \quad \nu(c^\star) = \sup_{(\pi,c)} \nu(c),$$

where the supremum is taken over all integrable strategies (π, c) the investor can implement without going bankrupt. We prove a verification theorem (Theorem 4.9) which shows that the optimal strategies $(\pi^\star, c^\star)$ are given in terms of a semilinear partial differential equation (PDE) of the form

$$0 = h_t - \tilde{r}h + \tilde{\alpha}h_y + \tfrac{1}{2}\beta^2 h_{yy} + \tfrac{\delta^\psi}{1-q}h^q, \quad h(T,\cdot) = \hat{\varepsilon}. \tag{1.8}$$

Note that equations of this form frequently arise in asset pricing, too. The main result of Chapter 4 is a general existence and uniqueness theorem (Theorem 4.10) for semilinear PDEs of the form as in (1.8).

10 Chapter 4 is largely based on Kraft, Seiferling, and Seifried (2015).
11 We impose standard regularity and boundedness conditions on the coefficients; see assumptions (A1) and (A2) on p. 110 and assumption (A1′) on p. 129.

The proof of this result is based on fixed point arguments which entail the following: Set $h^0 = \hat{\varepsilon}$ and iteratively define h^n as the unique solution g of the linear PDE

$$0 = g_t - \tilde{r}g + \tilde{\alpha}g_y + \frac{1}{2}\beta^2 g_{yy} + \frac{\delta\psi}{1-q}(h^{n-1})^q, \quad g(T,\cdot) = \hat{\varepsilon}.$$

Then the sequence $(h_n)_{n\in\mathbb{N}}$ converges to the solution h of (1.8), with respect to the uniform norm $\|h\|_{C^{0,1}} = \|h\|_\infty + \|h_y\|_\infty$ (Theorem 4.29). This convergence result immediately yields a fast numerical method which enables us to efficiently determine optimal strategies via iteratively solving linear PDEs; see the informal users' guide in Section 4.6, where we also provide numerical examples to illustrate our approach.

1.5 OUTLINE OF THIS THESIS

This thesis is structured as follows: Chapter 2 deals with backward nonlinear expectation equations (BNEEs) and convergence of corresponding recursive utilities, as explained in Section 1.2 above. In Chapter 3, we investigate the Epstein-Zin parametrization of stochastic differential utility with RRA and EIS both greater than one. The consumption-portfolio problem of an investor with stochastic differential utility in an incomplete market is studied in Chapter 4. Each chapter in the main part of this thesis is complemented by a corresponding chapter in the appendix. The bibliography is provided at the end of the appendix.

Chapter 2 Sections 2.1 and 2.2 set the scene for Chapter 2 by introducing a general framework for the study of backward nonlinear expectation equations. These equations are then studied in Section 2.3, where the issues of existence, uniqueness, stability and discretization are addressed. Section 2.4 provides general conditions which can be used to verify whether a given nonlinear expectation fits into our framework. Concrete examples of such nonlinear expectations are presented in Section 2.5. For the case of linear expectations, Section 2.6 relates our findings on BNEEs to the theory of BSDEs. The convergence result for recursive utilities is established in Section 2.7. Where appropriate, the end of each section provides bibliographical notes that relate its contents to the literature.

Chapter 3 Section 3.1 formally introduces continuous-time Epstein-Zin preferences. Our main results are stated in Section 3.2. The following two sections prepare for the proofs: In Section 3.3, utility functionals are constructed for bounded consumption plans, and Section 3.4 provides an excursion on stochastic Gronwall inequalities. Then the proofs of our main results are given in Section 3.5.

Chapter 4 The consumption and portfolio selection problem is formulated in Section 4.1, and the structure of the optimal strategy is announced in terms of a semilinear PDE. In Section 4.2, existence and uniqueness of appropriate solutions for this semilinear PDE are established by fixed point arguments. A verification result which ensures that the associated strategies are indeed optimal is proven in Section 4.3. Section 4.4 briefly comments on the relevance of our PDE results for asset pricing. Section 4.5 provides the basis for our numerical method. Section 4.6 offers an informal user's guide on how to apply our results numerically and illustrates the power of the solution method by a number of numerical examples.

2 BACKWARD NONLINEAR EXPECTATION EQUATIONS

In the first chapter of this thesis, we study backward nonlinear expectation equations (BNEEs) of the form

$$X_t = \mathcal{E}_t\Big[\int_t^T g(s, X)\,\mu(ds) + \xi\Big], \quad t \in [0, T]. \tag{2.1}$$

Here, $\mathcal{E}_t$ is a nonlinear expectation operator, g is a generator function, μ is a suitable integrator, and ξ is the terminal value.

Our analysis is built on a general framework for dynamic nonlinear expectations that comprises several well-known nonlinear expectations from the literature. In this framework, we establish existence, uniqueness and stability for solutions of (2.1). Moreover, we show that (2.1) emerges in the continuous-time limit of the discrete aggregations

$$X_k^\Delta \triangleq \mathcal{E}_{t_k}\Big[m\big((t_k, t_{k+1}]\big)\, g\big(t_k, \mathcal{E}_{t_k}[X_{k+1}^\Delta]\big) + X_{k+1}^\Delta\Big], \quad k = N-1, \dots, 0,$$

where $\Delta : 0 = t_0 < t_1 < \dots < t_N = T$, and we provide an application to recursive utility under ambiguity. These parts of the exposition are largely based on Belak, Seiferling, and Seifried (2015). In addition, we provide a sufficient criterion for nonlinear expectations to fit in our general framework, and we apply that criterion to several examples from the literature. Moreover, we relate BNEEs to classical backward stochastic differential equations (BSDEs) in the case when $\mathcal{E}_t$ is a *linear* expectation.

This chapter is structured as follows: Section 2.1 recalls the definition of a (time-consistent) nonlinear expectation. Section 2.2 introduces the notion of appropriate domains for sublinear expectations and provides our general framework for dynamic nonlinear expectations. Backward nonlinear expectation equations within that framework are studied in Section 2.3, where we prove existence, uniqueness, and stability results and establish convergence of the associated discrete-time nonlinear aggregations. Section 2.4 provides a general result on the existence of appropriate domains for sublinear expectations satisfying a Fatou property. In Section 2.5, this result is used to embed several examples of nonlinear expectations into the setting of Section 2.3. The connection between BNEEs and BSDEs is examined in Section 2.6. Section 2.7 concludes the chapter with an application of our results to recursive utility in discrete and continuous time.

2.1 NONLINEAR EXPECTATIONS

We give the definition of a *nonlinear expectation*. A nonlinear expectation will always be defined on a *domain (for expectations)*.

Domain for expectations

DEFINITION 2.1. Let $T > 0$ and let $(\mathcal{H}_t)_{t\in[0,T]}$ be an increasing family of linear spaces. Suppose that $(\mathcal{H}_T, \leqslant)$ is an ordered vector space and that $(\mathcal{H}_0, \leqslant)$ is order-isomorphic to $\mathbb{R}$.

Then $((\mathcal{H}_t)_{t\in[0,T]}, \leqslant)$ is a *domain for expectations*. ◇

DEFINITION 2.2. Let $((\mathcal{H}_t)_{t\in[0,T]}, \leqslant)$ be a domain for expectations and let $(\mathcal{E}_t)_{t\in[0,T]}$ be a family of operators

$$\mathcal{E}_t : \mathcal{H}_T \to \mathcal{H}_t, \qquad t \in [0, T].$$

Then $(\mathcal{E}_t)_{t\in[0,T]}$ is called a (time-consistent) *nonlinear expectation* on $(\mathcal{H}_t)_{t\in[0,T]}$, if the following four conditions are satisfied:

For all $X, Y \in \mathcal{H}_T$ and $t \in [0, T]$, the map $\mathcal{E}_t$ is

Nonlinear expectation

(M) *monotone*, i.e., $X \leqslant Y$ implies $\mathcal{E}_t[X] \leqslant \mathcal{E}_t[Y]$,

(SI) *shift-invariant*, i.e., $\mathcal{E}_t[X + Y] = X + \mathcal{E}_t[Y]$ if $X \in \mathcal{H}_t$,

(TC) *time-consistent*, i.e., $\mathcal{E}_t[\mathcal{E}_s[X]] = \mathcal{E}_t[X]$ for every $s \in [t, T]$,

(N) and *normalized*, i.e., $\mathcal{E}_t[0] = 0$.

A nonlinear expectation $(\mathcal{E}^{\mathrm{sub}}_t)_{t\in[0,T]}$ is called *sublinear* if each $\mathcal{E}^{\mathrm{sub}}_t$ is

Sublinear expectation

(PH) *positively homogeneous*, i.e., $\mathcal{E}^{\mathrm{sub}}_t[\lambda X] = \lambda\mathcal{E}^{\mathrm{sub}}_t[X]$ for all $\lambda \geqslant 0$,

(SUB) and *subadditive*, i.e., $\mathcal{E}^{\mathrm{sub}}_t[X + Y] \leqslant \mathcal{E}^{\mathrm{sub}}_t[X] + \mathcal{E}^{\mathrm{sub}}_t[Y]$.

Finally, a nonlinear expectation $(\mathcal{E}^{\mathrm{sup}}_t)_{t\in[0,T]}$ is called *superlinear* if $(-\mathcal{E}^{\mathrm{sup}}_t[-\cdot])_{t\in[0,T]}$ is a sublinear expectation. If $(\mathcal{E}_t)_{t\in[0,T]}$ is a nonlinear expectation on $(\mathcal{H}_t)_{t\in[0,T]}$, we shall refer to $(\mathcal{H}_t)_{t\in[0,T]}$ simply as the *domain* of $(\mathcal{E}_t)_{t\in[0,T]}$. ◇

From (SI) and (N) it is obvious that a nonlinear expectation also

(PC) *preserves constants*, i.e. $\mathcal{E}_t[X] = X$ for all $X \in \mathcal{H}_t$ and $t \in [0, T]$.

For subadditive operators the converse is also true, i.e. (PC) implies (SI) and (N). This is a consequence of the *self-domination* property of sublinear expectations:

Self-domination of sublinear expectations

LEMMA 2.3. *Let* $(\mathcal{E}^{\mathrm{sub}}_t)_{t\in[0,T]}$ *be a sublinear expectation on* $((\mathcal{H}_t)_{t\in[0,T]}, \leqslant)$. *For all* $t \in [0, T]$, *we have*

$$\mathcal{E}^{\mathrm{sub}}_t[X] - \mathcal{E}^{\mathrm{sub}}_t[Y] \leqslant \mathcal{E}^{\mathrm{sub}}_t[X - Y] \qquad \text{for all } X, Y \in \mathcal{H}_T. \tag{2.2}$$

Proof. For all $X, Y \in \mathcal{H}_T$ and $t \in [0, T]$, (SUB) implies that

$$\mathcal{E}^{\mathrm{sub}}_t[X] = \mathcal{E}^{\mathrm{sub}}_t[X - Y + Y] \leqslant \mathcal{E}^{\mathrm{sub}}_t[X - Y] + \mathcal{E}^{\mathrm{sub}}_t[Y],$$

which rearranges to (2.2). □

COROLLARY 2.4. *A family of operators* $\mathcal{E}_t : \mathcal{H}_T \to \mathcal{H}_t$, $t \in [0, T]$, *on a domain for expectations is a sublinear expectation if and only if it satisfies* (M), (TC), (PC), (SUB), *and* (PH).

Proof. Since (PC) implies (N), it remains to check (SI). Thus let $X \in \mathcal{H}_t$, $Y \in \mathcal{H}_T$, and note that Lemma 2.3 and (PC) imply

$$X + \mathcal{E}_t[Y] - \mathcal{E}_t[X+Y] \leqslant X + \mathcal{E}_t[Y - (X+Y)] = X + \mathcal{E}_t[-X] = 0,$$

i.e., $X + \mathcal{E}_t[Y] \leqslant \mathcal{E}_t[X+Y]$. On the other hand, (PC) and (SUB) yield $\mathcal{E}_t[X+Y] \leqslant \mathcal{E}_t[X] + \mathcal{E}_t[Y] = X + \mathcal{E}_t[Y]$, and (SI) is established. □

A simple example of a nonlinear expectation is given by the classical conditional expectation, of course. For truly nonlinear examples, we refer to Section 2.5 below.

Classical conditional expectation as nonlinear expectation

EXAMPLE 2.5. Let $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t\in[0,T]}, P)$ be a complete filtered probability space and let $\mathcal{F}_0$ be P-trivial. Then the family of *linear* operators $(E_t)_{t\in[0,T]}$, given by

$$E_t : L^1(\Omega, \mathcal{F}_T, P) \to L^1(\Omega, \mathcal{F}_t, P), \quad X \mapsto E^P[X \mid \mathcal{F}_t],$$

is both a sub- and superlinear expectation on $(L^1(\Omega, \mathcal{F}_t, P))_{t\in[0,T]}$, in the sense of Definition 2.2. Here, $L^1(\Omega, \mathcal{F}_t, P)$ is the Lebesgue space of all $\mathcal{F}_t$-measurable, P-integrable random variables and $E^P[X \mid \mathcal{F}_t]$ is the conditional expectation of X w.r.t. P, given $\mathcal{F}_t$.

2.2 SUBLINEAR EXPECTATIONS AND APPROPRIATE DOMAINS

Domains for expectations (as in Definition 2.1) have – except for the presence of additive inverses – just the right amount of structure to write down the general Definition 2.2 of sublinear expectations. Ideally, one would want to work with Banach spaces of "random variables" as in the case of linear expectations (recall Example 2.5). There, the family $\{L^p(\Omega, \mathcal{F}_t, P) \,:\, t \in [0,T],\ 1 \leqslant p \leqslant \infty\}$ of classical Lebesgue spaces enjoys a great deal more of structure: The spaces consist of equivalence classes of random variables, where two random variables are considered to be equivalent if their values differ only outside a set in $\mathcal{N}_P = \{N \subset \Omega : P(N) = 0\}$, and have a topology which is consistent with the pointwise order; all the spaces $L^p(\Omega, \mathcal{F}_0, P)$, $p \geqslant 1$, coincide and are isometrically order-isomorphic to $\mathbb{R}$; the spaces are related by Hölder's inequality, increasing in t and decreasing in p. Finally, the (sub)linear expectation generates the norms on the spaces on which it is defined by the expression

$$\|\cdot\|_{L,p} : L^p_T \to [0,\infty), \quad X \mapsto E^P_0\left[|X|^p\right]^{\frac{1}{p}}. \tag{2.3}$$

Essentially, we want the same thing for sublinear expecations: First, their domain should have the structural Lebesgue space-type features as outlined above. This leads to the notion of a *Lebesgue family*, which is defined in the following Subsection 2.2.1. Second, we want sublinear expectations to generate their own norms via the L^p-type expression (2.3). In this case, we have an *appropriate domain* for the sublinear expectation, as defined in Subsection 2.2.2 below.

2.2.1 *Lebesgue families*

Preliminaries

Let Ω be a non-empty set. To *axiomatically define* the concept of a *Lebesgue family*, we begin with the equivalence relation: Let $\mathcal{N} \subset 2^\Omega$ be a non-empty collection of subsets of Ω which is closed under countable unions and does not contain Ω; we refer to such an $\mathcal{N}$ as a *collection of negligible sets*.[1] A collection of negligible sets $\mathcal{N}$ induces an equivalence relation $\sim_\mathcal{N}$ on $\Omega^\mathbb{R}$ by

$$f \sim_\mathcal{N} g \iff f(\omega) = g(\omega) \quad \text{for all } \omega \in \Omega \setminus N \text{ and some } N \in \mathcal{N}.$$

The set $\{f \in \Omega^\mathbb{R} : f \sim_\mathcal{N} 0\}$ forms a subspace of $\Omega^\mathbb{R}$; the corresponding quotient space is denoted by $\Omega^\mathbb{R}/\mathcal{N}$. By definition, pointwise operations on $\Omega^\mathbb{R}/\mathcal{N}$ are well-behaved: If $\varphi : \mathbb{R} \to \mathbb{R}$ and $f \sim_\mathcal{N} g$, then $\varphi \circ f \sim_\mathcal{N} \varphi \circ g$. Moreover, $\Omega^\mathbb{R}/\mathcal{N}$ inherits the pointwise partial order from $\Omega^\mathbb{R}$, i.e.,

$$X \leqslant Y \text{ in } \Omega^\mathbb{R}/\mathcal{N} \iff f(\omega) \leqslant g(\omega) \text{ for all } \omega \in \Omega \text{ and some } f \in X,\ g \in Y,$$

and it is immediate that the canonical injection

$$i : \mathbb{R} \to \Omega^\mathbb{R}/\mathcal{N}, \quad x \mapsto [\Omega \to \mathbb{R},\ \omega \mapsto x]_{\sim_\mathcal{N}}$$

is order-preserving. With these preliminaries, we now give the definition of a Lebesgue family:

Lebesgue family

DEFINITION 2.6. Let Ω be a non-empty set and let $\mathcal{N} \subset 2^\Omega$ be a collection of negligible sets. Suppose that

$$\{(L^p_t, \|\cdot\|_{L,p}) \ :\ t \in [0,T],\ p \geqslant 1\}$$

is a family of Banach spaces such that the following statements hold true for all $t \in [0,T]$ and $p \geqslant 1$:

(L1) $L^p_t \leqslant \Omega^\mathbb{R}/\mathcal{N}$, i.e., L^p_t is a linear subspace of $\Omega^\mathbb{R}/\mathcal{N}$,

(L2) $L^p_s \subset L^q_t$ for all $1 \leqslant q \leqslant p$ and $0 \leqslant s \leqslant t$,

(L3) $L^p_t = \{X \in L^1_t : |X|^p \in L^1_t\}$,

(L4) $XY \in L^1_t$, if $X \in L^p_t$, $Y \in L^q_t$ and $p, q \in (1,\infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$,

(L5) $L^p_0 = i(\mathbb{R})$ for all $p \geqslant 1$ and $i : \mathbb{R} \to L^1_0$ is an isometry,

1 In the above example of classical Lebesgue spaces, this collection consists precisely of all P-null sets.

(L6) $L^1_+ \triangleq \{X \in L^1 \,:\, X \geqslant 0\} \subset L^1$ is closed.

Then $\{(L^p_t, \|\cdot\|_{L,p}) \,:\, t \in [0,T],\ p \geqslant 1\}$ is said to be a *Lebesgue family*. For brevity, we write $L^p \triangleq L^p_T$, $p \geqslant 1$. $\diamond$

Assumption (L1) simply states that we are working with equivalence classes of functions with respect to one fixed, pointwise defined equivalence relation. Condition (L2) guarantees that the spaces become larger as time passes and "more becomes known" and smaller as p increases, i.e., "as integrability gets harder." Assumption (L3) is certainly expected to hold for anything that claims to be similar to the classical L^p-spaces and implies in particular that all spaces are closed under taking the absolute value. The desired relationship between Hölder conjugate spaces is ensured by (L4). Condition (L5) entails that L^p_0 is isometrically order-isomorphic to $\mathbb{R}$: "Nothing is known at time 0." This allows us to identify L^p_0 and $\mathbb{R}$, which we will do in the following. The last assumption, (L6), guarantees that the ordering is preserved under taking limits. Intuitively, one should think of L^p_t as the space of time-t measurable, p-integrable random variables.

If $\{(L^p_t, \|\cdot\|_{L,p}) \,:\, t \in [0,T],\ p \geqslant 1\}$ is a Lebesgue family, then, for every $p \geqslant 1$, $((L^p_t)_{t\in[0,T]}, \leqslant)$ is a domain for expectations and may thus carry a nonlinear expectation (c.f. Definition 2.2). Sublinear expectations on Lebesgue families are investigated in the next subsection.

Note: For our results on backward nonlinear expectation equations and the convergence of recursive utilities in Section 2.3 and Section 2.7 below, the full Lebesgue family is not needed: One can fix $p = 1$ and work only on $(L^1_t)_{t\in[0,t]}$. In this case, the conditions in Definition 2.6 reduce to the requirements

$$L^1_t \leqslant L^1_s \leqslant \Omega^{\mathbb{R}}/\mathcal{N} \quad \text{and} \quad X \in L^1_t \implies |X| \in L^1_t \qquad \text{for all} \quad 0 \leqslant t \leqslant s \leqslant T$$

and the conditions that

$$\iota : \mathbb{R} \to L^1 \text{ is a surjective isometry} \quad \text{and that} \quad L^1_+ = \{X \in L^1 \,:\, X \geqslant 0\} \subset L^1 \text{ is closed.}$$

2.2.2 *Appropriate domains for sublinear expectations*

Suppose we are given a Lebesgue family

$$\{(L^p_t, \|\cdot\|_{L,p}) \,:\, t \in [0,T],\ p \geqslant 1\},$$

and let $(\mathcal{E}^{\text{sub}}_t)_{t\in[0,T]}$ be a sublinear expectation on $(L^1_t)_{t\in[0,T]}$. Then, for each $p \geqslant 1$, the function

$$\rho_p : L^p \to [0,\infty), \quad X \mapsto \mathcal{E}^{\text{sub}}_0\big[|X|^p\big]^{\frac{1}{p}} \tag{2.4}$$

defines a seminorm on L^p: Indeed, ρ_p is well-defined by (L3) and (M), and homogeneity is obvious from (PH). The triangle inequality is proven in exactly the same manner as the classical Minkowski inequality, making use of Hölder's inequality for ρ_p. For the sake of completeness, the proofs are provided below.

Hölder inequality

LEMMA 2.7. *Let* $t \in [0,T]$ *and* $p, q > 1$ *with* $\frac{1}{p} + \frac{1}{q} = 1$. *If* $X \in L_t^p$ *and* $Y \in L_t^q$, *then* $XY \in L_t^1$ *and*

$$\mathcal{E}_0^{sub}[|XY|] = \rho_1(XY) \leqslant \rho_p(X)\rho_q(Y) = \mathcal{E}_0^{sub}\left[|X|^p\right]^{\frac{1}{p}} \mathcal{E}_0^{sub}\left[|Y|^q\right]^{\frac{1}{q}}.$$

Proof. We have $XY \in L_t^1$ by (L4). For $\varepsilon > 0$ we put

$$\bar{X} \triangleq \frac{X}{\left(\varepsilon + \mathcal{E}_0^{sub}[|X|^p]\right)^{\frac{1}{p}}} \in L^p \quad \text{and} \quad \bar{Y} \triangleq \frac{Y}{\left(\varepsilon + \mathcal{E}_0^{sub}[|Y|^q]\right)^{\frac{1}{q}}} \in L^q.$$

Choose some $g \in \bar{X}$ and some $h \in \bar{Y}$. Then Young's inequality implies

$$|g(\omega)h(\omega)| \leqslant \tfrac{1}{p}|g(\omega)|^p + \tfrac{1}{q}|h(\omega)|^q \quad \text{for all } \omega \in \Omega.$$

Hence $|\bar{X}\bar{Y}| \leqslant \frac{1}{p}|\bar{X}|^p + \frac{1}{q}|\bar{Y}|^q$, where both $\frac{1}{p}|\bar{X}|^p$ and $\frac{1}{q}|\bar{Y}|^q$ are in L^1 by (L3), and thus we obtain

$$\mathcal{E}_0^{sub}\left[|\bar{X}\bar{Y}|\right] \leqslant \mathcal{E}_0^{sub}\left[\tfrac{1}{p}|\bar{X}|^p + \tfrac{1}{q}|\bar{Y}|^q\right] \leqslant \tfrac{1}{p}\mathcal{E}_0^{sub}\left[|\bar{X}|^p\right] + \tfrac{1}{q}\mathcal{E}_0^{sub}\left[|\bar{Y}|^q\right]$$

by (M), (SUB) and (PH). Moreover, (PH) yields

$$\mathcal{E}_0^{sub}\left[|\bar{X}|^p\right] = \mathcal{E}_0^{sub}[|X|^p] / \left(\varepsilon + \mathcal{E}_0^{sub}[|X|^p]\right) \leqslant 1 \quad \text{and} \quad \mathcal{E}_0^{sub}\left[|\bar{Y}|^q\right] \leqslant 1,$$

and we get $\mathcal{E}_0^{sub}[|\bar{X}\bar{Y}|] \leqslant \frac{1}{p} + \frac{1}{q} = 1$. Therefore, we have

$$\mathcal{E}_0^{sub}[|XY|] \leqslant \left(\varepsilon + \mathcal{E}_0^{sub}[|X|^p]\right)^{\frac{1}{p}} \left(\varepsilon + \mathcal{E}_0^{sub}[|Y|^q]\right)^{\frac{1}{q}},$$

and letting $\varepsilon \to 0$ yields the claim. □

From Hölder's inequality, one can prove Minkowski's inequality using the classical argument.

Minkowski inequality

LEMMA 2.8. *Let* $p \geqslant 1$ *and* $X, Y \in L^p$. *Then* $\rho_p(X+Y) \leqslant \rho_p(X) + \rho_p(Y)$.

Proof. We first prove the claim for $p = 1$. The triangle inequality in $\mathbb{R}$ implies that $|X+Y| \leqslant |X| + |Y|$, where $|X+Y|, |X|, |Y| \in L^1$ by (L3). Now monotonicity (M) and subadditivity (SUB) yield

$$\mathcal{E}_0^{sub}\left[|X+Y|\right] \leqslant \mathcal{E}_0^{sub}\left[|X|+|Y|\right] \leqslant \mathcal{E}_0^{sub}\left[|X|\right] + \mathcal{E}_0^{sub}\left[|Y|\right],$$

i.e., $\rho_1(X+Y) \leqslant \rho_1(X) + \rho_1(Y)$.

Now let $p > 1$. Making use of (L3), we set

$$\bar{X} \triangleq |X| \in L^p, \quad \bar{Y} \triangleq |Y| \in L^p, \quad \text{and } \bar{Z} \triangleq |X+Y| \in L^p.$$

The triangle inequality in $\mathbb{R}$ implies

$$|\bar{Z}|^p = |X+Y| \cdot |\bar{Z}|^{p-1} \leqslant \bar{X} \cdot |\bar{Z}|^{p-1} + \bar{Y} \cdot |Z|^{p-1}.$$

Note that $|\bar{Z}|^p \in L^1$ and $|\bar{Z}|^{p-1} \in L^q$ by (L3), where $q = \frac{p}{p-1} > 1$. Since $\frac{1}{p} + \frac{1}{q} = 1$, Lemma 2.7 yields $\bar{X} \cdot |\bar{Z}|^{p-1}$, $\bar{Y} \cdot |\bar{Z}|^{p-1} \in L^1$ with

$$\rho_1(\bar{X} \cdot |\bar{Z}|^{p-1}) \leqslant \rho_p(X)\rho_q\left(|\bar{Z}|^{p-1}\right), \quad \rho_1(\bar{Y} \cdot |\bar{Z}|^{p-1}) \leqslant \rho_p(Y)\rho_q\left(|\bar{Z}|^{p-1}\right).$$

Thus with (M) and the already proven inequality for $p = 1$, we get

$$\rho_1\left(|\bar{Z}|^p\right) \leqslant \rho_1\left(\bar{X}\cdot|\bar{Z}|^{p-1}+\bar{Y}\cdot|Z|^{p-1}\right) \leqslant \rho_1\left(\bar{X}\cdot|\bar{Z}|^{p-1}\right)+\rho_1\left(\bar{Y}\cdot|Z|^{p-1}\right).$$

Combining the above, we arrive at

$$\rho_p(X+Y)^p = \rho_1\left(|\bar{Z}|^p\right) \leqslant \left(\rho_p(X)+\rho_p(Y)\right)\rho_q\left(|\bar{Z}|^{p-1}\right).$$

For $\rho_p(X+Y) = 0$, there was nothing to show, anyway. Otherwise, we divide by $\rho_p(X+Y)^{p-1} = \rho_q(|\bar{Z}|^{p-1})$ to complete the proof. □

As hinted at above, a Lebesgue family is an appropriate domain for a sublinear expectation $(\mathcal{E}_t^{\text{sub}})_{t\in[0,T]}$ if $(\mathcal{E}_t^{\text{sub}})_{t\in[0,T]}$ generates its norms via (2.4).

Appropriate domain

DEFINITION 2.9. Let

$$\mathcal{L} = \left\{(L_t^p, \|\cdot\|_{L,p}) : t\in[0,T],\ p\geqslant 1\right\}$$

be a Lebesgue family and suppose that $(\mathcal{E}_t^{\text{sub}})_{t\in[0,T]}$ is a sublinear expectation on $(L_t^1)_{t\in[0,T]}$. If

$$\mathcal{E}_t^{\text{sub}} : L^p \to L_t^p \quad \text{for all } t\in[0,T] \text{ and } p\geqslant 1,$$

and if the seminorms in (2.4) satisfy

$$\rho_p = \|\cdot\|_{L,p} \quad \text{for all } p\geqslant 1,$$

then we say that $\mathcal{L}$ is an *appropriate domain* for $(\mathcal{E}_t^{\text{sub}})_{t\in[0,T]}$. ◇

Note: If one restricts attention to $p = 1$ (see the note on p. 17), it suffices to require that $\rho_1 = \|\cdot\|_{L,1}$.

Note: In our article Belak, Seiferling, and Seifried (2015), on which the current section is partially based, the concept of an appropriate domain was referred to as an appropriate family of L^p-spaces.

Remark. In this thesis, we focus on nonlinear expectations that are defined on appropriate domains. This restriction is necessary if one insists that the domain of the nonlinear expectation carry a linear structure. For instance, Nutz and van Handel (2013) show that it is not possible to construct G-expectations on a linear space containing all Borel functions on Wiener space. ◇

General results on the existence of appropriate domains for sublinear expectations satisfying a *Fatou property* will be presented in Subsection 2.4. Appropriate domains for the concrete examples of g-, G- and random G-expectations can be found below, in Subsections 2.5.1, 2.5.2 and 2.5.3. For linear expectations, appropriate domains come without surprises:

EXAMPLE 2.10. Let $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t\in[0,T]}, P)$ be a complete filtered probability space and let $\mathcal{F}_0$ be P-trivial. Then the family of classical Lebesgue spaces $\{L^p(\Omega,\mathcal{F}_t,P) : t\in[0,T],\ p\geqslant 1\}$ is an appropriate domain for the linear expectation $(E_t)_{t\in[0,T]}$ of Example 2.5.

The definition of an appropriate domain entails in particular that $(\mathcal{E}_t^{\mathrm{sub}})_{t\in[0,T]}$ is a sublinear expectation on $(L_t^p)_{t\in[0,T]}$. Moreover, as a straightforward consequence of the next lemma (Jensen's inequality), the operators $\mathcal{E}_t^{\mathrm{sub}} : L^p \to L_t^p$ are continuous. We record this fact explicitly as Lemma 2.15 in the next subsection.

Jensen's inequality

LEMMA 2.11. *Let* $\{(L_t^p, \|\cdot\|_{L,p}) : t \in [0,T],\ p \geqslant 1\}$ *be a Lebesgue family, and let* $(\mathcal{E}_t^{\mathrm{sub}})_{t\in[0,T]}$ *be a sublinear expectation on* $(L_t^1)_{t\in[0,T]}$. *Then Jensen's inequality holds:*

For all $t \in [0,T]$ *and any convex function* $\varphi : \mathbb{R} \to \mathbb{R}$, *we have*

$$\varphi\big(\mathcal{E}_t^{\mathrm{sub}}[X]\big) \leqslant \mathcal{E}_t^{\mathrm{sub}}\big[\varphi(X)\big] \quad \textit{whenever } X \in L^1 \textit{ with } \varphi(X) \in L^1.$$

In particular, for all $p \geqslant 1$ *and* $t \in [0,T]$, *it follows that*

$$\big|\mathcal{E}_t^{\mathrm{sub}}[X]\big|^p \leqslant \mathcal{E}_t^{\mathrm{sub}}\big[|X|^p\big] \quad \textit{for all } X \in L^p.$$

Proof. We first prove that

$$a\mathcal{E}_t^{\mathrm{sub}}[X] + b \leqslant \mathcal{E}_t^{\mathrm{sub}}[aX + b] \quad \text{for all } a, b \in \mathbb{R}. \tag{2.5}$$

Normalization (N) and self-domination (2.2) give

$$-\mathcal{E}_t^{\mathrm{sub}}[X] = \mathcal{E}_t^{\mathrm{sub}}[0] - \mathcal{E}_t^{\mathrm{sub}}[X] \leqslant \mathcal{E}_t^{\mathrm{sub}}[0 - X] = \mathcal{E}_t^{\mathrm{sub}}[-X].$$

Together with positive homogeneity (PH), this implies

$$a\mathcal{E}_t^{\mathrm{sub}}[X] = \mathrm{sign}(a)\mathcal{E}_t^{\mathrm{sub}}\big[|a|X\big] \leqslant \mathcal{E}_t^{\mathrm{sub}}\big[aX\big] \quad \text{for all } a \in \mathbb{R}.$$

Now shift-invariance (SI) yields (2.5).

We denote by L_φ the collection of all $(a,b) \in \mathbb{Q}^2$ with $ay + b \leqslant \varphi(y)$ for all $y \in \mathbb{R}$. Then, for all $(a,b) \in L_\varphi$, monotonicity (M) and inequality (2.5) imply

$$a\mathcal{E}_t^{\mathrm{sub}}[X] + b \leqslant \mathcal{E}_t^{\mathrm{sub}}[aX + b] \leqslant \mathcal{E}_t^{\mathrm{sub}}[\varphi(X)].$$

Thus, since $\mathcal{N}$ is closed under countable unions, we find a negligible set $N \in \mathcal{N}$ and representatives $f, g \in \Omega^{\mathbb{R}}$ of $\mathcal{E}_t^{\mathrm{sub}}[X]$ and $\mathcal{E}_t^{\mathrm{sub}}[\varphi(X)]$ in $L^1 \leqslant \Omega^{\mathbb{R}}/\mathcal{N}$ such that

$$af(\omega) + b \leqslant g(\omega) \quad \text{for all } \omega \in \Omega \setminus N \text{ and all } (a,b) \in L_\varphi.$$

Taking the pointwise supremum, we obtain

$$\varphi(f(\omega)) = \sup_{(a,b)\in L_\varphi} af(\omega) + b \leqslant g(\omega) \quad \text{for all } \omega \in \Omega \setminus N$$

by convexity of φ, and the first claim follows. The second claim is an immediate consequence of the first since $x \mapsto |x|^p$ is convex and $|X|^p \in L^1$ for all $X \in L^p$ by (L3). ☐

We stress that the above version of Jensen's inequality does not claim that $\varphi(\mathcal{E}_t^{\mathrm{sub}}[X]) \in L_t^1$. It only asserts that the desired inequality holds in $\Omega^{\mathbb{R}}/\mathcal{N}$. In the following, we will only use the second part, which may also pass as a consequence of Hölder's inequality. In that specific case, the membership $\mathcal{E}_t^{\mathrm{sub}}[X] \in L_t^p$ is guaranteed if the Lebesgue family $\{(L_t^p, \|\cdot\|_{L,p}) : t \in [0,T],\ p \geqslant 1\}$ is an appropriate domain for $(\mathcal{E}_t^{\mathrm{sub}})_{t\in[0,T]}$.

2.2.3 *Dominated nonlinear expectations on appropriate domains*

Nonlinear expectations (which are not sublinear) do not generate their own norm via the L^p-type expression (2.4) from Subsection 2.2.2 above. To accommodate them, we assume that they come together with a sublinear expectation that dominates them in the sense of the following definition:

Dominated nonlinear expectations

DEFINITION 2.12. Let $(\mathcal{E}_t)_{t\in[0,T]}$, $(\bar{\mathcal{E}}_t)_{t\in[0,T]}$ be nonlinear expectations defined on the same domain $((\mathcal{H}_t)_{t\in[0,T]}, \leqslant)$. We say that $(\bar{\mathcal{E}}_t)_{t\in[0,T]}$ *dominates* $(\mathcal{E}_t)_{t\in[0,T]}$ if

$$\mathcal{E}_t[X] - \mathcal{E}_t[Y] \leqslant \bar{\mathcal{E}}_t[X-Y] \quad \text{for all } X, Y \in \mathcal{H}_T \text{ and } t \in [0,T]. \qquad \diamond$$

Remark. Lemma 2.3 (on p. 14 above) implies that every sublinear expectation dominates itself. ◇

The domination property has useful implications for nonlinear expectations that are dominated by a sublinear expectation:

Dominated nonlinear expectations on appropriate domain

DEFINITION 2.13. Let $(\mathcal{E}^{\text{sub}}_t)_{t\in[0,T]}$ be a sublinear expectation on an appropriate domain $\mathcal{L} = \{(L^p_t, \|\cdot\|_{L,p}) \ : \ t\in[0,T],\ p\geqslant 1\}$. Suppose that $(\mathcal{E}_t)_{t\in[0,T]}$ is a nonlinear expectation on $(L^1_t)_{t\in[0,T]}$ such that

$$(\mathcal{E}^{\text{sub}}_t)_{t\in[0,T]} \text{ dominates } (\mathcal{E}_t)_{t\in[0,T]} \quad \text{and} \quad \mathcal{E}_t : L^p \to L^p_t \qquad (2.6)$$

for all $t\in[0,T]$ and $p\geqslant 1$. Then we call $(\mathcal{E}_t)_{t\in[0,T]}$ a *dominated nonlinear expectation carried by* $(\mathcal{E}^{\text{sub}}_t)_{t\in[0,T]}$ *on (the appropriate domain)* $\mathcal{L}$.

Note: If one restricts to $p=1$, it suffices to require that $(\mathcal{E}^{\text{sub}}_t)_{t\in[0,T]}$ dominates $(\mathcal{E}_t)_{t\in[0,T]}$ on $(L^1_t)_{t\in[0,T]}$.

Dominated nonlinear expectations satisfy a triangle inequality:

Triangle inequality

LEMMA 2.14. *Let $(\mathcal{E}_t)_{t\in[0,t]}$ be a nonlinear expectation that is carried by $(\mathcal{E}^{\text{sub}}_t)_{t\in[0,T]}$ on the appropriate domain $\{(L^p_t, \|\cdot\|_{L,p}) \ : \ t\in[0,T],\ p\geqslant 1\}$. Then, for all $X, Y \in L^1$, the triangle inequality*

$$|\mathcal{E}_t[X] - \mathcal{E}_t[Y]| \leqslant \mathcal{E}^{\text{sub}}_t[|X-Y|], \qquad t\in[0,T],$$

is satisfied. In particular, we have $|\mathcal{E}_t[X]| \leqslant \mathcal{E}^{\text{sub}}_t[|X|]$ and

$$|\mathcal{E}^{\text{sub}}_t[X] - \mathcal{E}^{\text{sub}}_t[Y]| \leqslant \mathcal{E}^{\text{sub}}_t[|X-Y|] \quad \textit{as well as} \quad |\mathcal{E}^{\text{sub}}_t[X]| \leqslant \mathcal{E}^{\text{sub}}_t[|X|].$$

Proof. First note that $|X-Y| \in L^1$ by property (L3) of a Lebesgue family. Now, $(\mathcal{E}^{\text{sub}}_t)_{t\in[0,T]}$ dominates $(\mathcal{E}_t)_{t\in[0,T]}$, and hence monotonicity (M) of $\mathcal{E}^{\text{sub}}_t$ implies

$$\mathcal{E}_t[X] - \mathcal{E}_t[Y] \leqslant \mathcal{E}^{\text{sub}}_t[X-Y] \leqslant \mathcal{E}^{\text{sub}}_t\big[|X-Y|\big].$$

In the same fashion, it follows that $\mathcal{E}_t[Y] - \mathcal{E}_t[X] \leqslant \mathcal{E}^{\text{sub}}_t[|X-Y|]$, which proves the first claim. The second is implied by the normalization property (N) upon setting $Y = 0$. The remaining statements follow from self-domination of $(\mathcal{E}^{\text{sub}}_t)_{t\in[0,T]}$, see Lemma 2.3 above. □

Together with Jensen's inequality, the triangle inequality shows that $\mathcal{E}_t : L^p \to L^p_t$ is a continuous mapping.

Continuity of dominated nonlinear expectations

LEMMA 2.15. *Let* $(\mathcal{E}_t)_{t\in[0,T]}$ *be a nonlinear expectation that is carried by* $(\mathcal{E}^{\mathrm{sub}}_t)_{t\in[0,T]}$ *on the appropriate domain* $\{(L^p_t, \|\cdot\|_{L,p}) : t \in [0,T],\ p \geqslant 1\}$. *Then, for all* $p \geqslant 1$ *and* $t \in [0,T]$, *we have the* contraction property

$$\left\|\mathcal{E}_t[X] - \mathcal{E}_t[Y]\right\|_{L,p} \leqslant \|X - Y\|_{L,p}, \qquad X, Y \in L^p,$$

and, in particular,

$$\left\|\mathcal{E}^{\mathrm{sub}}_t[X] - \mathcal{E}^{\mathrm{sub}}_t[Y]\right\|_{L,p} \leqslant \|X - Y\|_{L,p}, \qquad X, Y \in L^p.$$

Proof. The triangle inequality (Lemma 2.14) implies

$$|\mathcal{E}_t[X] - \mathcal{E}_t[Y]| \leqslant \mathcal{E}^{\mathrm{sub}}_t\left[|X - Y|\right],$$

where $|\mathcal{E}_t[X] - \mathcal{E}_t[Y]| \in L^p$ and $|X - Y| \in L^p$ by (2.6) and (L3). Hence, Jensen's inequality (Lemma 2.11) yields

$$\left|\mathcal{E}_t[X] - \mathcal{E}_t[Y]\right|^p \leqslant \left|\mathcal{E}^{\mathrm{sub}}_t[|X - Y|]\right|^p \leqslant \mathcal{E}^{\mathrm{sub}}_t\left[|X - Y|^p\right],$$

where $|\mathcal{E}_t[X] - \mathcal{E}_t[Y]|^p \in L^1$ and $\mathcal{E}^{\mathrm{sub}}_t[|X - Y|^p] \in L^1$. Now, monotonicity (M) ensures that

$$\|\mathcal{E}_t[X] - \mathcal{E}_t[Y]\|^p_{L,p} = \mathcal{E}^{\mathrm{sub}}_0\left[\left|\mathcal{E}_t[X] - \mathcal{E}_t[Y]\right|^p\right] \leqslant \mathcal{E}^{\mathrm{sub}}_0\left[\mathcal{E}^{\mathrm{sub}}_t\left[|X - Y|^p\right]\right],$$

where the last term is precisely $\mathcal{E}^{\mathrm{sub}}_0[|X - Y|^p] = \|X - Y\|^p_{L,p}$ by time-consistency (TC). This proves the first claim. Self-dominance of sublinear expectations (Lemma 2.3) immediately yields the second. □

2.2.4 *Processes and integrals*

By definition, a stochastic process is a family $X = (X_t)_{t\in[0,T]}$ of random variables. In line with our notion of L^p_t as the space of time-t measurable, p-integrable random variables, the natural definition of processes in the context of nonlinear expectations is as follows.

Measurable and adapted processes

DEFINITION 2.16. An L^p*-process* is a function $X : [0,T] \to L^p$, where $\mathcal{L} \triangleq \{(L^p_t, \|\cdot\|_{L,p}) : t \in [0,T],\ p \geqslant 1\}$ is a Lebesgue family. It is *measurable* if the function X is $\mathcal{B}([0,T])$-$\mathcal{B}(L^p)$-measurable, and it is *adapted* if $X_t \in L^p_t$ for all $t \in [0,T]$. The space of all measurable and adapted L^p-processes is denoted by $\mathcal{X}^p(\mathcal{L})$. We will simply write $\mathcal{X}^p$ if the Lebesgue family $\mathcal{L}$ is clear from the context. ⋄

Remark. In the above definition (and the remainder of this text), the symbol $\mathcal{B}(S)$ denotes the Borel σ-algebra on a topological space S. ⋄

We will consider (Lebesgue-Bochner) integrals of measurable L^p-processes with respect to a finite positive measure μ on $\mathcal{B}([0,T])$, as outlined below. For additional details, we refer to Appendix A.3 (p. 163ff.) and Dunford and Schwartz (1958).

DEFINITION 2.17. Let μ be a positive finite measure on $\mathcal{B}([0,T])$. For each $p \geqslant 1$, we define the *space of μ-integrable adapted L^p-processes* as

Space of integrable processes $\mathcal{P}^p$

$$\mathcal{P}^p(\mathcal{L},\mu) \triangleq \{X \in \mathcal{X}^p(\mathcal{L}) : X1_{[0,t]} \in \mathcal{L}^1(\mu; L^p_t) \text{ for all } t \in [0,T]\},$$

where $\mathcal{L}^1(\mu; L^p_t)$ denotes the space of L^p_t-valued μ-integrable functions. A *seminorm* on $\mathcal{P}^p(\mathcal{L},\mu)$ is defined via

Norm on $\mathcal{P}^p$

$$\|\cdot\|_{P,p} : \mathcal{X}^p(\mathcal{L}) \to [0,\infty], \quad \|X\|_{P,p} \triangleq \int_{[0,T]} \|X_t\|_{L,p}\,\mu(dt).$$

Identifying $X, Y \in \mathcal{P}^p(\mathcal{L},\mu)$ if $\|X - Y\|_{P,p} = 0$, we obtain the associated quotient space $P^p(\mathcal{L},\mu)$. $\diamond$

Note: The condition $X1_{[0,t]} \in \mathcal{L}^1(\mu; L^p_t)$ is imposed in the definition of $\mathcal{P}^p$ in order to ensure that $\int_A X\,d\mu \in L^p_t$ whenever $A \subset \mathcal{B}([0,t])$. In the presence of the integrability condition $\|X\|_{P,p} = \int_{[0,T]} \|X_t\|_{L,p}\,\mu(dt) < \infty$, this is equivalent to $X1_{[0,t]}$ being μ-essentially separably valued, see Lemma A.39 (p. 163). If L^p is a separable Banach space, this condition is automatically satisfied, and then $\mathcal{P}^p$ consists precisely of all norm-integrable $X \in \mathcal{X}^p(\mathcal{L})$, i.e., $\mathcal{P}^p(\mathcal{L},\mu) = \{X \in \mathcal{X}^p(\mathcal{L}) : \|X\|_{P,p} < \infty\}$.

PROPOSITION 2.18. *Let μ be a positive finite measure on $\mathcal{B}([0,T])$, and let $\mathcal{L}$ be a Lebesgue family. Then $(P^p, \|\cdot\|_{P,p})$ is a Banach space.*

P^p is complete

Proof. This is a straightforward consequence of Theorem III.3.6 in Dunford and Schwartz (1958), p. 146. See Lemma A.40 (p. 164) in Appendix A.3 for some particularities of the present situation. □

For every process $X \in P^p(\mathcal{L},\mu)$, the Bochner integral

Bochner integral

$$\int_A X\,d\mu \in L^p, \quad A \in \mathcal{B}([0,T]),$$

of X over A with respect to μ is defined. As is the usual convention, for $(a,b] \subset [0,T]$, we write $\int_a^b X\,d\mu \triangleq \int_{(a,b]} X\,d\mu$. To emphasize the integration variable, we shall also write $\int_A X_t\,\mu(dt) \triangleq \int_A X\,d\mu$.

LEMMA 2.19. *The integral*

$$\int_A : P^p \to L^p, \quad X \mapsto \int_A X\,d\mu$$

is a continuous linear operator satisfying

$$\left\|\int_A X\,d\mu\right\|_{L,p} \leqslant \int_A \|X_t\|_{L,p}\,\mu(dt) \leqslant \|X\|_{P,p}, \qquad X \in P^p, \qquad (2.7)$$

for all $A \in \mathcal{B}([0,T])$.

Proof. See Lemma A.41, p. 165. □

Since $X \in \mathcal{P}^p(\mathcal{L},\mu)$ is adapted, $X1_{[0,t]}$ is an L^p_t-valued Bochner integrable function, and we have

$$\int_A X\,d\mu = \int_A X1_{[0,t]}\,d\mu \in L^p_t \quad \text{for all } A \in \mathcal{B}([0,t]). \qquad (2.8)$$

Thus, the integral process $[0,T] \to L^p$, $t \mapsto \int_0^t X\,d\mu \in L^p$ is adapted. With Lemma 2.21, we show that it is also right-continuous and, in particular, measurable.

We conclude this section by introducing subspaces of P^p that represent natural domains for backward nonlinear expectation equations.

The spaces S^p and D^p

DEFINITION 2.20. Let $\mathcal{L}$ be a Lebesgue family. For $1 \leqslant p < \infty$ we define

$$S^p(\mathcal{L}, \mu) \triangleq \{X \in \mathcal{P}^p(\mathcal{L}, \mu) : \|X\|_{S,p} < \infty\}, \qquad \|X\|_{S,p} \triangleq \sup_{t \in [0,T]} \|X_t\|_{L,p}.$$

Then $(S^p, \|\cdot\|_{S,p})$ is a Banach space that contains the closed subspace

$$D^p(\mathcal{L}) \triangleq \{X \in \mathcal{X}^p(\mathcal{L}) : [0,T] \to L^p,\ t \mapsto X_t \text{ is càdlàg}\}.$$

If $\mathcal{L}$ and μ are clear from the context, we will only write S^p and D^p. ⋄

We note that every càdlàg adapted L^p-process automatically is a member of $S^p(\mathcal{L}, \mu)$, regardless of μ, that is, $D^p(\mathcal{L}) \subset S^p(\mathcal{L}, \mu)$ for all finite Borel measures μ on $[0,T]$.

Note: The necessary simple function approximation is easily produced: Let X be a càdlàg adapted L^p-process; then $t \mapsto \|X_t\|_{L,p}$ is also càdlàg and, in particular, $\|X\|_{S,p} < \infty$. We set $X^n \triangleq X_0 1_{\{0\}} + \sum_{k=1}^n X_{t_k^n} 1_{(t_{k-1}^n, t_k^n]}$, $t_k^n \triangleq kt/n$, $n \in \mathbb{N}$. Then X^n is an L_t^p-valued measurable simple function with $\|X_s^n - X_s\|_{L,p} \to 0$ and $\|X_s^n - X_s\|_{L^p} \leqslant 2\|X\|_{S,p}$ for all $s \in [0,t]$. Dominated convergence implies $X 1_{[0,t]} \in \mathcal{L}^1(\mu; L_t^p)$ for all $t \in [0,T]$; thus, $X \in \mathcal{P}^p(\mathcal{L}, \mu)$, and hence $X \in S(\mathcal{L}, \mu)$.

For every finite positive measure μ, we have

$$\|X\|_{P,p} \leqslant \mu([0,T])\|X\|_{S,p}, \qquad X \in S^p,$$

and thus, we can regard $S^p(\mathcal{L}, \mu) \subseteq P^p(\mathcal{L}, \mu)$ as a subspace. It is noteworthy that the integral induces a continuous linear operator $P^p \to D^p$.

Integral process

LEMMA 2.21. *For each $1 \leqslant p < \infty$, the integral*

$$I : P^p \to D^p, \quad X \mapsto IX, \quad \textit{where } (IX)_t \triangleq \textstyle\int_0^t X_s\, \mu(ds),\ t \in [0,T],$$

is a continuous linear operator.

Proof. By (2.8), the integral IX is an adapted L^p-process. We have $\sup_{t \in [0,T]} \|IX_t\|_{L,p} \leqslant \|X\|_{P,p}$ by (2.7), and, for $t \leqslant s$, the inequality

$$\|(IX)_s - (IX)_t\|_{L,p} \leqslant \textstyle\int_0^T 1_{(t,s]}(u) \|X_u\|_{L,p}\, \mu(du)$$

implies that IX is càdlàg; thus, $IX \in D^p$. □

2.2.5 *Bibliographical notes*

Upper expectations, given by $\mathcal{E}_0^{\text{sub}}[\cdot] = \sup_{P \in \mathbb{P}} E^P[\cdot]$ (in terms of a set $\mathbb{P}$ of probability measures), are the most natural examples of (static) *sublinear expectations*, and they have a long history in robust statistics, see, e.g., Huber and Strassen (1973), the book by Huber (1981) and the references therein. *Upper* and *lower probabilities* yield a popular extension of the classical probabilistic framework and allow the quantification of uncertainty in the presence of inconsistent information, when the Kolmogorov (1933) axioms reach their limitations. Instead of associating a single number – a precise probability – with an event, a whole interval of plausible probabilities between the upper and the lower probability is assigned. Hence, upper and lower probabilities are highly related to *upper* and *lower previsions*, as proposed by Walley (1991) in his book on statistical reasoning with *imprecise probabilities*. For a recent survey of the topic, we refer to Miranda (2008). An extensive account of the history of nonlinear probabilities in statistics can be found in Hampel (2009). For additional background on imprecise probabilities, we refer to Coolen et al. (2011) and the references therein.

Upper and lower previsions can be interpreted as maximum buying prices and minimum selling prices for gambles; decisions on whether a gamble is accepted or rejected are based on these prices and obey two principles: They are *coherent*, meaning that a positive linear combination of acceptable positions is still acceptable, and they *avoid sure loss*. Thus, it is apparent that the concept of imprecise probabilities is strongly connected to that of *coherent risk measures*, as proposed by Artzner et al. (1999); see also Delbaen (2002) and Föllmer and Schied (2004). On the other hand, there is a one-to-one correspondence between coherent risk measures and (static) sublinear expectations: A mapping ρ is a coherent risk measure if and only if $\nu \triangleq \rho[-\cdot]$ is a (static) sublinear expectation. In that context, ν is often referred to as a *valuation*.

A dynamically consistent extension of these concepts has been proposed by Peng (1999, 2004a) and Coquet et al. (2002), under the name *filtration-consistent nonlinear expectation*, in the context of g-expectations. A filtration-consistent valuation or nonlinear expectation $\mathcal{E}_t$, $t \in [0, T]$, maps $\mathcal{F}_T$-measurable random variables to $\mathcal{F}_t$-measurable random variables, where $(\mathcal{F}_t)_{t\in[0,T]}$ is a filtration, and it is supposed to satisfy the tower property

$$\mathcal{E}_t[\mathcal{E}_s[X]] = \mathcal{E}_t[X], \quad t \leqslant s.$$

This strong notion of *time-consistency* is the same one we use in our Definition 2.2 of a nonlinear expectation; however, by directly modeling the flow of information by an increasing family of linear spaces $(\mathcal{H}_t)_{t\in[0,T]}$ we abstract from the filtration. In the context of dynamic risk measures, there are various weaker concepts of time-consistency in the literature. For these and further developments of the theory, we refer to Cheridito et al. (2004, 2006), Artzner et al. (2007), Jobert and Rogers (2008), Stadje (2010), Acciaio et al. (2012), Pelsser and Stadje (2014) and the references therein.

Our general framework of Section 2.2 comprises g-expectations, see, e.g., Peng (1997, 2004b), G-expectations as put forward by Peng (2007, 2008), and random G-expectations as introduced and analyzed by Nutz (2012, 2013). By construction, the G- and random G-expectation are defined on appropriate domains in the sense of Definition 2.9. *Appropriate domains* are an abstraction of these concrete situations which has been proposed in Belak, Seiferling, and Seifried (2015) and which generalizes the constructions of Peng (2005b, 2008, 2010).

All of the observations made in Section 2.2 have been made before, in similar frameworks for nonlinear expectations; see, for example, Peng (2008) for Hölder's inequality (Lemma 2.7) and Minkowski's inequality (Lemma 2.8), and Cohen et al. (2011) for Jensen's inequality (Lemma 2.11). The notion of domination from Definition 2.12 is also standard in the theory of nonlinear expectations and so are its implications, as the triangle inequality (Lemma 2.14) and continuity (Lemma 2.15); see, e.g., Coquet et al. (2002).

2.3 BACKWARD NONLINEAR EXPECTATION EQUATIONS

In Section 2.1 and Section 2.2, we have set up the mathematical framework for our study of backward nonlinear expectation equations:

STANDING ASSUMPTIONS Throughout this section,

$$(\mathcal{E}_t)_{t\in[0,T]} \quad \text{is a nonlinear expectation, carried by a}$$

sublinear expectation $(\mathcal{E}_t^{\mathrm{sub}})_{t\in[0,T]}$ on an appropriate domain

$$\mathcal{L} \triangleq \{(L_t^p, \|\cdot\|_{L,p}) \ : \ t\in[0,T],\ p\geqslant 1\},$$

in the sense of Definition 2.13. Moreover, μ is a finite Borel measure on $[0,T]$; we consider the corresponding space of (μ-equivalence classes of) μ-integrable L^p-processes

$$P^p = \left\{[X] \ : \ X \text{ adapted } L^p\text{-process with } X1_{[0,t]} \in \mathcal{L}^1(\mu; L_t^p)\ \forall t\in[0,T]\right\},$$

as in Definition 2.17. Recall that P^p is a Banach space with norm

$$\|X\|_{P,p} = \int_{[0,T]} \|X_t\|_{L,p}\, \mu(dt), \qquad X\in P^p.$$

Moreover, as in Definition 2.20, we consider the process spaces

$$S^p = \left\{X\in \mathcal{P}^p(\mathcal{L},\mu) \ : \ \|X\|_{S,p} < \infty\right\} \qquad \text{and}$$
$$D^p = \{X \text{ adapted } L^p\text{-process} \ : \ [0,T]\to L^p,\ t\mapsto X_t \text{ is càdlàg}\} \subset S^p$$

equipped with the norm

$$\|X\|_{S,p} \triangleq \sup_{t\in[0,T]} \|X_t\|_{L,p}, \qquad X\in S^p.$$

In the following, we are interested in solutions of *backward nonlinear expectation equations* (BNEEs) of the form

$$X_t = \mathcal{E}_t\left[\int_t^T g(s,X)\,\mu(ds) + \xi\right], \quad t\in[0,T], \tag{2.9}$$

where X is an L^p-process and (g,ξ) is a suitable parameter (see Definition 2.24 below).

First, we note that it is natural to require a priori that $X \in S^p$. Indeed, for the nonlinear expectation and the integral on the right-hand side of (2.9) to be well-defined, $g(\cdot, X)$ must be an integrable L^p-process. On the other hand, a solution X of the equation is automatically adapted, so $X \in \mathcal{X}^p$. Moreover, by the contraction property (Lemma 2.15) and continuity of the integral (Lemma 2.19), we have

$$\sup_{t\in[0,T]} \|X_t\|_{L,p} \leqslant \int_0^T \|g(t,X)\|_{L,p}\, \mu(dt) + \|\xi\|_{L,p} < \infty,$$

see also Lemma 2.27 below.

In summary, the BNEE (2.9) only makes sense for $X \in S^p$. However, the measurability condition implied by $X \in S^p$ is not automatically satisfied: For a general nonlinear expectation and $\xi \in L^p$, the mapping $t \mapsto \mathcal{E}_t[\xi]$ need not be well-behaved, and so even the simplest BNEE

$$X_t = \mathcal{E}_t[\xi], \quad t \in [0,T],$$

may not admit a solution in S^p. It is thus necessary to restrict the class of nonlinear expectations under consideration:

Measurable and regular nonlinear expectations

DEFINITION 2.22. A nonlinear expectation $(\mathcal{E}_t)_{t\in[0,T]}$ is said to be *measurable* if the L^p-process

$$X : [0,T] \to L^p, \quad t \mapsto \mathcal{E}_t[\xi]$$

is a member of S^p for each $\xi \in L^p$. If X even is a member of D^p, then $(\mathcal{E}_t)_{t\in[0,T]}$ is called *regular*. ◇

Note: The naming in Definition 2.22 probably warrants some explanations. Note that $X_t = \mathcal{E}_t[\xi]$ is always an adapted L^p-process with $\|X\|_{S,p} = \sup_{t\in[0,T]} \|X_t\|_{L,p} = \|\xi\|_{L,p} < \infty$. Thus insisting that X be a member of D^p is equivalent to requiring that the mapping $[0,T] \to L^p$, $t \mapsto \mathcal{E}_t[\xi]$ be càdlàg. Hence the use of the word *regular* seems to be justified in this context. The use of the term *measurable* is somewhat more arbitrary. However, if the appropriate domain under consideration consists solely of separable Banach spaces (as is often the case), then asking that X be in S^p is the same as requiring that the map $[0,T] \to L^p$, $t \mapsto \mathcal{E}_t[\xi]$ be $\mathcal{B}([0,T])$-$\mathcal{B}(L^p)$-measurable; so at least in that case, the wording is meaningful.

By continuity of nonlinear expectations, it suffices to check measurability and regularity on a dense subset of L^p:

LEMMA 2.23. *For each $\xi \in L^p$, let $X^\xi : [0,T] \to L^p_t$, $t \mapsto \mathcal{E}_t[\xi]$. Then $(\mathcal{E}_t)_{t\in[0,T]}$ is regular (measurable) if and only if there exists a dense subset $M \subseteq L^p$ such that X^ξ is càdlàg ($X^\xi \in S^p$) for each $\xi \in M$.*

Proof. By the contraction property from Lemma 2.15, the operator $X : L^p \to \times_{t\in[0,T]} L^p_t$, $\xi \mapsto X^\xi$ satisfies

$$\sup_{t\in[0,T]} \|X^\xi(t) - X^\eta(t)\|_{L,p} \leqslant \|\xi - \eta\|_{L,p} \quad \text{for all } \xi, \eta \in L^p.$$

Hence, the claim follows by a density argument. □

In Section 2.5, we will show that the most widely used specifications of nonlinear expectations are regular in the sense of Definition 2.22.

2.3.1 Existence and uniqueness

For $X \in \mathfrak{X}^p$ and $t \in [0,T]$, we briefly write $X^t \triangleq 1_{[t,T]}X \in \mathfrak{X}^p$.

BNEE^p-parameter

DEFINITION 2.24. Let $\xi \in L^p$, and let

$$g : [0,T] \times S^p \to L^p \quad \text{be such that} \quad g(\cdot, X) \in P^p \text{ for every } X \in S^p. \tag{2.10}$$

Suppose there exists a constant $L > 0$ such that g satisfies the Lipschitz condition

$$\|g(t,X) - g(t,Y)\|_{L,p} \leqslant L\|X^t - Y^t\|_{S,p} \quad \text{for all } X, Y \in S^p, t \in [0,T]. \tag{2.11}$$

Then, the pair (g, ξ) is called a *BNEE^p-parameter*. ⋄

For future reference, we note that (2.11) implies

$$\|g(\cdot, X)\|_{P,p} \leqslant \mu([0,T])L\|X\|_{S,p} + \|g(\cdot, 0)\|_{P,p}, \quad X \in S^p. \tag{2.12}$$

Note: In view of (2.12), the requirement $g(\cdot, X) \in P^p$ for every $X \in S^p$ in Definition 2.24 can be relaxed to $g(\cdot, 0) \in P^p$ if L^p is separable.

The BNEE associated to a BNEE^p-parameter (g, ξ) is given by

$$X_t = \mathcal{E}_t\Big[\int_t^T g(s,X)\,\mu(ds) + \xi\Big], \quad t \in [0,T]. \tag{2.13}$$

Frequently, the aggregator function g depends only on the current value of the solution process:

BNEE^p-standard parameter

DEFINITION 2.25. Let $\xi \in L^p$ and $f : [0,T] \times L^p \to L^p$. Suppose that

$$\|f(t,\zeta) - f(t,\eta)\|_{L,p} \leqslant L\|\zeta - \eta\|_{L,p} \quad \text{for all } \zeta, \eta \in L^p, \ t \in [0,T], \tag{2.14}$$

and that $f(t, L^p_t) \subset L^p_t$ for all $t \in [0,T]$.

If (f, ξ) induces a BNEE^p-standard parameter (g, ξ) via

$$g : [0,T] \times S^p \to L^p, \quad (t, X) \mapsto f(t, X_t), \tag{2.15}$$

then (f, ξ) is called a *BNEE^p-standard parameter*. ⋄

In line with (2.15) and (2.13), the BNEE associated to a BNEE^p-standard parameter (f, ξ) takes the form

$$X_t = \mathcal{E}_t\big[\int_t^T f(s,X_s)\,\mu(ds) + \xi\big], \quad t \in [0,T]. \tag{2.16}$$

The remainder of this subsection addresses the proof of the following first main result:

Existence and uniqueness for BNEEs

THEOREM 2.26. *Suppose that $(\mathcal{E}_t)_{t\in[0,T]}$ is measurable and let (g, ξ) be a BNEE^p-parameter. Then the BNEE (2.9) has a unique solution $X \in S^p$. If $(\mathcal{E}_t)_{t\in[0,T]}$ is regular, then $X \in D^p$.*

The proof is based on a classical fixed point approach. Before we show that the corresponding iteration operator is well-defined, we prove a norm estimate that will be frequently used in the following.

LEMMA 2.27. *Let* $Y, \bar{Y} \in P^p$ *and* $\xi, \bar{\xi} \in L^p$, *and set* *Useful estimate*

$$\Delta_t \triangleq \mathcal{E}_t \left[\int_t^T Y_s \, \mu(ds) + \xi \right] - \mathcal{E}_t \left[\int_t^T \bar{Y}_s \, \mu(ds) + \bar{\xi} \right].$$

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

$$\sup_{s \in [t,T]} \|\Delta_s\|_{L,p} \leqslant \int_t^T \|Y_s - \bar{Y}_s\|_{L,p} \, \mu(ds) + \|\xi - \bar{\xi}\|_{L,p}. \tag{2.17}$$

Proof. Let $0 \leqslant t \leqslant s \leqslant T$. Clearly $U_s \triangleq \int_s^T Y_r \, \mu(dr) + \xi \in L^p$ and $\bar{U}_s \triangleq \int_s^T \bar{Y}_r \, \mu(dr) + \bar{\xi} \in L^p$. Thus $\Delta_s \in L^p_s$, and the triangle inequality (Lemma 2.14) implies that

$$|\Delta_s| = |\mathcal{E}_s[U_s] - \mathcal{E}_s[\bar{U}_s]| \leqslant \mathcal{E}_s^{\text{sub}}\big[|U_s - \bar{U}_s|\big].$$

Hence, applying the contraction property (Lemma 2.15), we obtain

$$\|\Delta_s\|_{L,p} \leqslant \|U_s - \bar{U}_s\|_{L,p} = \left\| \int_s^T (Y_r - \bar{Y}_r) \, \mu(dr) + (\xi - \bar{\xi}) \right\|_{L,p}.$$

By continuity of the integral (2.7), this entails

$$\|\Delta_s\|_{L,p} - \|\xi - \bar{\xi}\|_{L,p} \leqslant \int_s^T \|Y_r - \bar{Y}_r\|_{L,p} \, \mu(dr) \leqslant \int_t^T \|Y_r - \bar{Y}_r\|_{L,p} \, \mu(dr)$$

for all $t \in [t, T]$. Thus (2.17) is established. □

LEMMA 2.28. *Under the assumptions of Theorem 2.26, the formula* *Fixed point operator*

$$(\Phi X)_t \triangleq \mathcal{E}_t \Big[\int_t^T g(s, X) \, \mu(ds) + \xi \Big], \quad t \in [0, T], \tag{2.18}$$

defines an operator $\Phi : S^p \to S^p$. *If* $(\mathcal{E}_t)_{t \in [0,T]}$ *is regular, then* $\Phi(S^p) \subseteq D^p$.

Proof. Let $X \in S^p$. Then $Y \triangleq g(\cdot, X) \in P^p$ by (2.10), and, in particular, $\int_0^T Y_t \, \mu(dt) \in L^p$. Since $(\mathcal{E}_t)_{t \in [0,T]}$ is measurable, the L^p-process M given by

$$M_t = \mathcal{E}_t\big[\int_0^T Y_s \, \mu(ds) + \xi\big], \quad t \in [0, T],$$

is a member of S^p (and of D^p if $(\mathcal{E}_t)_{t \in [0,T]}$ is regular). By Lemma 2.21, the process I given by $I_t \triangleq \int_0^t Y_s \, \mu(ds)$, $t \in [0, T]$, is in D^p. Thus $\Phi X \triangleq M - I \in S^p$ (and $\Phi X \in D^p$ if $(\mathcal{E}_t)_{t \in [0,T]}$ is regular). By shift invariance (SI), we have

$$(\Phi X)_t = \mathcal{E}_t\big[\int_0^T Y_s \, \mu(ds) + \xi\big] - \int_0^t Y_s \, \mu(ds) = \mathcal{E}_t\big[\int_t^T Y_s \, \mu(ds) + \xi\big],$$

which establishes (2.18) and completes the proof. □

Proof of Theorem 2.26. Let $U, V \in S^p$. Lemma 2.27, applied to $Y \triangleq g(\cdot, U)$ and $\bar{Y} \triangleq g(\cdot, V)$, shows that

$$\sup_{s \in [t,T]} \|(\Phi U)_s - (\Phi V)_s\|_{L,p} \leqslant \int_t^T \|Y_s - \bar{Y}_s\|_{L,p} \, \mu(ds)$$

for all $t \in [0, T]$. By Lipschitz continuity (2.11) of g, we have

$$\|Y_s - \bar{Y}_s\|_{L,p} = \|g(s, X) - g(s, Y)\|_{L,p} \leqslant L \|U^s - V^s\|_{S,p}$$

for all $s \in [0,T]$ (where $U^s = 1_{[s,T]}U$). We thus get

$$\|(\Phi U)^t - (\Phi V)^t\|_{S,p} \leqslant L\int_t^T \|U^s - V^s\|_{S,p}\,\mu(ds) \quad \text{for all } t \in [0,T].$$

Iterating this estimate, we see that $\Delta^n \triangleq \|\Phi^n U - \Phi^n V\|_{S,p}$ can be bounded above as follows:

$$\begin{aligned}\Delta^n &\leqslant L^n \int_0^T\int_{t_1}^T \cdots \int_{t_{n-1}}^T \|U^{t_n} - V^{t_n}\|_{S,p}\,\mu(dt_n)\cdots\mu(dt_2)\mu(dt_1)\\ &\leqslant \mu([0,T])^n \|U - V\|_{S,p}\frac{L^n}{n!}.\end{aligned}$$

Thus Φ^n is a contraction for sufficiently large $n \in \mathbb{N}$, and hence Φ has a unique fixed point $X \in S^p$ that satisfies (2.9). If $(\mathcal{E}_t)_{t\in[0,T]}$ is regular, then $X = \Phi X \in D^p$ by Lemma 2.28. □

2.3.2 *Stability of BNEEs*

With Theorem 2.26 above, we have established existence and uniqueness for solutions of BNEEs. It is now natural to investigate their stability under perturbations of the aggregator and the terminal value. We have the following result:

Stability of BNEEs

THEOREM 2.29. *Let $(\mathcal{E}_t)_{t\in[0,T]}$ be measurable, and let (g^n, ξ^n), $n \in \mathbb{N}$, and (g, ξ) be BNEE^p-parameters. Suppose there is a constant $L > 0$ such that*

$$\|g^n(t,X) - g^n(t,Y)\|_{L,p} \leqslant L\|X^t - Y^t\|_{S,p} \quad \textit{for all } X, Y \in S^p,\ t \in [0,T],$$

and all $n \in \mathbb{N}$. Let X^n, $n \in \mathbb{N}$, and X denote the solutions of the associated BNEEs and suppose that

$$\int_0^T \|g^n(t,X) - g(t,X)\|_{L,p}\,\mu(dt) \to 0 \quad \textit{and} \quad \xi^n \to \xi \ \textit{in } L^p.$$

Then $X^n \to X$ in S^p.

Proof. Lemma 2.27, applied to $Y = g(\cdot, X)$ and $\bar{Y} = g^n(\cdot, X^n)$ as well as ξ and $\bar{\xi} = \xi^n$, shows that

$$\sup_{s\in[t,T]} \|X_s - X_s^n\|_{L,p} \leqslant \int_t^T \|g(s,X) - g^n(s,X_s^n)\|_{L,p}\,\mu(ds) + \|\xi - \xi^n\|_{L,p}.$$

By assumption, we have

$$\|g^n(s,X) - g^n(s,X^n)\|_{L,p} \leqslant L\|X^s - (X^n)^s\|_{S,p} \quad \text{for all } s \in [0,T],$$

and therefore

$$\|g(s,X) - g^n(s,X^n)\|_{L,p} \leqslant \|g(s,X) - g^n(s,X)\|_{L,p} + L\|X^s - (X^n)^s\|_{S,p}.$$

Inserting this into the first inequality, we obtain

$$\|X^t - (X^n)^t\|_{S^p} = \sup_{s\in[t,T]} \|X_s - X_s^n\|_{L,p} \leqslant L\int_t^T \|X^s - (X^n)^s\|_{S,p}\,\mu(ds) + \delta_n$$

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

$$\delta_n \triangleq \int_0^T \|g(s,X) - g^n(s,X)\|_{L,p}\,\mu(ds) + \|\xi^n - \xi\|_{L,p} \to 0.$$

We conclude by Gronwall's inequality. □

2.3.3 *Discretization of BNEEs*

Building on the stability result of Theorem 2.29, we now address the discrete-time approximation of BNEEs.

In the following, we fix a BNEE^p-standard parameter (f, ξ) and let

$$X_t = \mathcal{E}_t\Big[\int_t^T f(s, X_s)\,\mu(ds) + \xi\Big], \quad t \in [0, T], \tag{2.19}$$

denote the unique solution of the corresponding BNEE. To ease notation, we assume without loss that $m(\{0\}) = m(\{T\}) = 0$. We are interested in a suitable discrete-time approximation of X that converges in the continuous-time limit of vanishing grid size. More specifically, given a partition $\Delta\colon\ 0 = t_0 < t_1 < \cdots < t_{N(\Delta)} = T$ of $[0, T]$, we set

$$|\Delta| \triangleq \max_{k=1,\dots,N(\Delta)} \big(\mu\big((t_{k-1}, t_k]\big) + |t_k - t_{k-1}|\big).$$

We consider the Δ-discretization scheme $X^\Delta \in \times_{k=0}^{N(\Delta)} L^p_{t_k}$ that is defined by a suitable approximation of the terminal value $X^\Delta_{N(\Delta)} \triangleq \xi^\Delta$ and then iteratively, for $k = N(\Delta) - 1, \dots, 0$, via

$$X^\Delta_k \triangleq \mathcal{E}_{t_k}\Big[\mu\big((t_k, t_{k+1}]\big) f^\Delta\big(t_k, \mathcal{E}_{t_k}[X^\Delta_{k+1}]\big) + X^\Delta_{k+1}\Big]. \tag{2.20}$$

We are interested in the convergence $X^\Delta \to X$ for vanishing grid size $|\Delta|$. In (2.20), the mapping f^Δ is a BNEE^p-standard parameter that may depend on the grid Δ and approximates f as $|\Delta| \to 0$ in the sense made precise in the following definition.

DEFINITION 2.30. Let $(\Delta^n)_{n\in\mathbb{N}}$ be a sequence of partitions *Exhausting sequences*

$$\Delta^n : 0 = t^n_0 < t^n_1 < \cdots < t^n_{N_n} = T, \quad n \in \mathbb{N},$$

and let $(f^n, \xi^n)_{n\in\mathbb{N}}$ be a sequence of BNEE^p-standard parameters. Suppose there is a constant $L > 0$ such that

$$\|f^n(t,\zeta) - f^n(t,\eta)\|_{L,p} \leqslant L\|\zeta - \eta\|_{L,p} \quad \text{for all } \zeta, \eta \in L^p,\ n \in \mathbb{N}, \tag{2.21}$$

and let X denote the unique solution of (2.19).

If $|\Delta^n| \to 0$, $\xi^n \to \xi$ in L^p and

$$\sum_{k=0}^{N_n-1} \int_{[t^n_k, t^n_{k+1})} \|f(s, X_s) - f^n(t^n_k, X_s)\|_{L,p}\,\mu(ds) \to 0, \tag{2.22}$$

then $(\Delta^n, f^n, \xi^n)_{n\in\mathbb{N}}$ is said to be (f,ξ)-*exhausting*. ◇

A natural special case of exhausting sequences is the following:

LEMMA 2.31. *Let $(\Delta^n)_{n\in\mathbb{N}}$ be a sequence of partitions with $|\Delta^n| \to 0$ and suppose that the mapping $[0, T] \to L^p$, $t \mapsto f(t, \eta)$ is left-continuous for every $\eta \in L^p$. Then $(\Delta^n, f, \xi)_{n\in\mathbb{N}}$ is (f,ξ)-exhausting.*

Proof. We have to establish (2.22). Setting $g^n(s, X_s) \triangleq f(t^n_k, X_s)$ for $s \in [t^n_k, t^n_{k+1})$, we can rewrite (2.22) as

$$\int_0^T \|f(s, X_s) - g^n(s, X_s)\|_{L,p}\,\mu(ds) \to 0.$$

Lipschitz continuity (2.14) of f yields

$$\|f(s,X_s) - g^n(s,X_s)\|_{L,p} \leqslant 2L\|X_s\|_{L^p} \leqslant 2L\|X\|_{S,p},$$

and, by dominated convergence, it suffices to show that $g^n(s,X_s) \to f(s,X_s)$ for all $s \in [0,T]$: Choose k_n such that $s \in [t^n_{k_n}, t^n_{k_n+1})$. Since $|\Delta^n| \to 0$, we see that $t^n_{k_n}$ converges to s from the left, and hence

$$g^n(s,X_s) = f(t^n_{k_n}, X_s) \to f(s,X_s)$$

because f is left-continuous. □

For simplicity of notation, we write $X^n \triangleq X^{\Delta^n}$ in the sequel. The main result of this subsection is the following convergence theorem:

THEOREM 2.32. *Suppose that* $(\mathcal{E}_t)_{t\in[0,T]}$ *is regular. Let* (f,ξ) *be a BNEE*^{p}*-standard parameter and let* $(\Delta^n, f^n, \xi^n)_{n\in\mathbb{N}}$ *be* (f,ξ)*-exhausting. Let* X *denote the solution of* (2.19) *and let* $X^n \triangleq X^{\Delta^n}$ *be given by* (2.20)*. Then*

$$\max_{k=0,\dots,N_n} \|X^n_k - X_{t^n_k}\|_{L,p} \to 0.$$

The remainder of this subsection is concerned with the proof of Theorem 2.32. It will be useful to introduce the following continuous-time interpolation of X^n (also denoted by X^n, with a slight abuse of notation): For $n \in \mathbb{N}$, we put $X^n_T \triangleq \xi^n$ and, for $k = N_n - 1, \dots, 0$, we set

$$X^n_t \triangleq \mathcal{E}_t\Big[\mu\big((t, t^n_{k+1}]\big) f^n\big(t^n_k, \mathcal{E}_{t^n_k}[X^n_{t^n_{k+1}}]\big) + X^n_{t^n_{k+1}}\Big], \quad t \in [t^n_k, t^n_{k+1}). \quad (2.23)$$

To prove Theorem 2.32, we first identify X^n as the unique D^p-solution of the BNEE driven by the aggregator g^n, where

$$g^n(\cdot, Y) \triangleq \textstyle\sum_{k=0}^{N_n-1} 1_{[t^n_k, t^n_{k+1})} f^n(t^n_k, \mathcal{E}_{t^n_k}[Y_{t^n_{k+1}}]), \quad Y \in S^p. \quad (2.24)$$

Then we show $\|X^n - X\|_{S^p} \to 0$ via the stability result of Theorem 2.29.

LEMMA 2.33. *For each* $n \in \mathbb{N}$*, the pair* (g^n, ξ^n) *is a BNEE*^{p}*-parameter. Moreover, with the constant* $L > 0$ *from* (2.21)*, we have*

$$\|g^n(t,U) - g^n(t,V)\|_{L,p} \leqslant L\|U^t - V^t\|_{S,p} \quad \textit{for all } U, V \in S^p,$$

all $t \in [0,T]$*, and each* $n \in \mathbb{N}$*.*

Proof. Let $U, V \in S^p$. Then $f^n(t^n_k, \mathcal{E}_{t^n_k}[U_{t^n_{k+1}}]) \in L^p_{t^n_k}$ since f is a BNEE^p standard parameter. It thus becomes apparent from (2.24) that $g^n(\cdot, U)$ is a right-continuous and adapted L^p-step process. In particular, $g^n(\cdot, U) \in P^p$ for all $U \in S^p$, as required for a BNEE^p-parameter. It remains to verify the Lipschitz condition.

For an arbitrary $t \in [t^n_k, t^n_{k+1})$, the definition of g^n and the uniform Lipschitz continuity (2.21) of $(f^n)_{n\in\mathbb{N}}$ imply

$$\|g^n(t,U) - g^n(t,V)\|_{L,p} \leqslant L\big\|\mathcal{E}_{t^n_k}[U_{t^n_{k+1}}] - \mathcal{E}_{t^n_k}[V_{t^n_{k+1}}]\big\|_{L,p}.$$

Hence the contraction property (Lemma 2.15) yields

$$\|g^n(t,U) - g^n(t,V)\|_{L,p} \leqslant L\|U_{t^n_{k+1}} - V_{t^n_{k+1}}\|_{L,p} \leqslant L\|U^t - V^t\|_{S,p}. \quad □$$

LEMMA 2.34. *For each* $n \in \mathbb{N}$, *the* L^p*-process* X^n *given by* (2.23) *is a member of* D^p *and satisfies the BNEE*

$$X_t^n = \mathcal{E}_t\left[\int_t^T g^n(s, X^n)\,\mu(ds) + \xi^n\right], \quad t \in [0, T].$$

Proof. Let $n \in \mathbb{N}$ and $k \in \{0, \ldots, N_n - 1\}$. By (2.24) we have

$$f^n(t_k^n, \mathcal{E}_{t_k^n}[X^n_{t^n_{k+1}}]) = g^n(t, X^n) \quad \text{for all } t \in [t_k^n, t_{k+1}^n),$$

and hence (2.23) yields

$$X_t^n = \mathcal{E}_t\left[\int_t^{t^n_{k+1}} g^n(s, X^n)\,\mu(ds) + X^n_{t^n_{k+1}}\right] \quad \text{for all } t \in [t_k^n, t_{k+1}^n)$$

and all $k \in \{0, \ldots, N_n - 1\}$. Plugging the above representation of $X^n_{t^n_{k+1}}$ into the one for X_t^n, we get

$$X_t^n = \mathcal{E}_t\left[\int_t^{t^n_{k+1}} g^n(s, X^n)\,\mu(ds) + \mathcal{E}_{t^n_{k+1}}\left[\int_{t^n_{k+1}}^{t^n_{k+2}} g^n(s, X^n)\,\mu(ds) + X^n_{t^n_{k+2}}\right]\right].$$

Now, shift-invariance (SI) implies

$$X_t^n = \mathcal{E}_t\left[\mathcal{E}_{t^n_{k+1}}\left[\int_t^{t^n_{k+2}} g^n(s, X^n)\,\mu(ds) + X^n_{t^n_{k+2}}\right]\right],$$

and time-consistency (TC) yields

$$X_t^n = \mathcal{E}_t\left[\int_t^{t^n_{k+2}} g^n(s, X^n)\,\mu(ds) + X^n_{t^n_{k+2}}\right].$$

Iterating this procedure, we obtain

$$X_t^n = \mathcal{E}_t\left[\int_t^T g^n(s, X^n)\,\mu(ds) + \xi^n\right], \quad t \in [0, T].$$

Finally, Lemma 2.28 shows that $X^n \in D^p$ since $(\mathcal{E}_t)_{t \in [0,T]}$ is regular and (g^n, ξ) is a BNEE^p-parameter by Lemma 2.33. □

We now provide the proof of the convergence result. By Lemma 2.34, the discrete-time approximations X^n solve BNEEs associated with the BNEE^p-parameters (g^n, ξ^n). Lemma 2.33 shows that the aggregators g^n have a common Lipschitz constant, as required by the stability result of Theorem 2.29. We know that $\xi^n \to \xi$ in L^p because (Δ^n, f^n, ξ^n) is (f, ξ)-exhausting, and hence the stability result will imply that $X_n \to X$ in S^p, provided that $g^n(\cdot, X) \to g(\cdot, X)$ in P^p.

Proof of Theorem 2.32. In view of Theorem 2.29, Lemma 2.33, and Lemma 2.34 (and the above discussion), it remains to prove that

$$\int_0^T \|g^n(t, X) - f(t, X_t)\|_{L,p}\,\mu(dt) \to 0.$$

Let $t \in [0, T]$ and $k_n \in \mathbb{N}$ such that $t \in [t^n_{k_n}, t^n_{k_n+1})$, $n \in \mathbb{N}$. We note that

$$\begin{aligned} &\|g^n(t, X) - f(t, X_t)\|_{L,p} \\ &\leqslant \|g^n(t, X) - f^n(t^n_{k_n}, X_t)\|_{L,p} + \|f^n(t^n_{k_n}, X_t) - f(t, X_t)\|_{L,p}. \end{aligned} \tag{2.25}$$

By definition, $g^n(t,X) = f^n(t^n_{k_n}, \mathcal{E}_{t^n_{k_n}}[X_{t^n_{k_n+1}}])$, and hence by (2.21)

$$D^n_t \triangleq \|g^n(t,X) - f^n(t^n_{k_n}, X_t)\|_{L,p} \leqslant L\|\mathcal{E}_{t^n_{k_n}}[X_{t^n_{k_n+1}}] - X_t\|_{L,p}.$$

We further estimate this by

$$D^n_t \leqslant L\big(\|\mathcal{E}_{t^n_{k_n}}[X_{t^n_{k_n+1}}] - \mathcal{E}_{t^n_{k_n}}[X_t]\|_{L,p} + \|\mathcal{E}_{t^n_{k_n}}[X_t] - X_t\|_{L,p}\big),$$

and so the contraction property (Lemma 2.15) yields

$$D^n_t \leqslant L\|X_{t^n_{k_n+1}} - X_t\|_{L,p} + L\|\mathcal{E}_{t^n_{k_n}}[X_t] - X_t\|_{L,p}. \tag{2.26}$$

In particular, we have the uniform bound $D^n_t \leqslant 4L\|X\|_{S,p}$. Coming back to (2.25), (2.22) implies

$$\begin{aligned}\int_0^T \|g^n(t,X) - f(t,X_t)\|_{L,p}\,\mu(dt) - \int_0^T D^n_t\,\mu(dt)\\ \leqslant \sum_{k=0}^{N_n-1}\int_{[t^n_k,t^n_{k+1})}\|f^n(t^n_k,X_t) - f(t,X_t)\|_{L,p}\,\mu(dt) \to 0\end{aligned}$$

since (Δ^n, f^n, ξ^n) is (f,ξ)-exhausting. To complete the proof, it thus remains to argue that $\int_0^T D^n_t\,\mu(dt) \to 0$. By dominated convergence, it suffices to show that $D^n_t \to 0$ for μ-a.e. $t \in [0,T]$, as we shall do in the following.

Since $(\mathcal{E}_t)_{t\in[0,T]}$ is regular, for each $t \in [0,T]$, the mappings

$$[0,T] \to L^p,\ s \mapsto X_s \quad\text{and}\quad [0,T] \to L^p,\ s \mapsto \mathcal{E}_s[X_t]$$

are càdlàg. Moreover, as $(\Delta^n, f^n, \xi^n)_{n\in\mathbb{N}}$ is (f,ξ)-exhausting, we have

$$|\Delta^n| = \max_{k=1,\dots,N_n}\big(\mu\big((t^n_{k-1}, t^n_k)\big) + |t^n_k - t^n_{k-1}|\big) \to 0, \tag{2.27}$$

and thus $t^n_{k_n}, t^n_{k_n+1} \to t$ as $n \to \infty$ for all $t \in [0,T]$. Hence, we obtain $\|X_{t^n_{k_n+1}} - X_t\|_{L,p} \to 0$ for all $t \in [0,T]$ and $\|\mathcal{E}_{t^n_{k_n}}[X_t] - X_t\|_{L,p} \to 0$ for all $t \in [0,T]$ outside some countable set $N \subset [0,T]$. In view of (2.26), it remains to prove that N is a μ-null set.

If $\mu(\{t\}) > 0$, we must have $t^n_{k_n} = t$ for all but finitely many $n \in \mathbb{N}$ by (2.27). Otherwise, we would have $t \in (t_{k_{n_\ell}}, t_{k_{n_\ell}+1})$ for some increasing sequence $(n_\ell)_{\ell\in\mathbb{N}} \subset \mathbb{N}$, and thus $|\Delta^{n_\ell}| \geqslant \mu(\{t\}) > 0$, which is a contradiction. Hence we have

$$\|\mathcal{E}_{t^n_{k_n}}[X_t] - X_t\|_{L,p} = |\mathcal{E}_t[X_t] - X_t\|_{L,p} = 0$$

for all but finitely many $n \in \mathbb{N}$ whenever $\mu(\{t\}) > 0$. This shows that N contains no atoms of μ. Since N is countable, it thus is a μ-null set, and the proof is complete. □

2.3.4 *Bibliographical notes*

Backward equations of the form (2.9) and, in particular, of the form (2.16) may be regarded as generalizations of backward stochastic differential equations (BSDEs) under nonlinear expectations. In the literature, equations of that form have previously been studied in specific settings: Peng (2004b, 2005a) considers BSDEs under g-expectations, and G-BSDEs have been investigated in, among others, Peng (2010), and Hu et al. (2014a,b), the latter building on the G-martingale representation results of Soner et al. (2011a).

A related class of equations, known as second-order BSDEs, has been introduced by Cheridito et al. (2007) and Soner et al. (2012, 2013); see also Soner et al. (2011b) for related results. In the economics literature, BNEEs have appeared in the context of dynamic robust risk preferences; see, e.g., Chen and Epstein (2002), Hayashi (2005) and Epstein and Ji (2014).

When $\mathcal{E}_t$ is a *linear* expectation, backward equations of the form (2.16) have been studied extensively in the literature; see, e.g., Pardoux and Peng (1990), Duffie and Epstein (1992b), Antonelli (1993) and El Karoui et al. (1997). To the best of our knowledge, Chen and Epstein (2002) and Peng (2004b) are the first to formulate equations of the form (2.9) under nonlinear (g-)expectations.

In the case of linear expectations, the questions of stability and discrete-time approximation of BSDEs are very well studied. Stability results related to the one from Subsection 2.3.2 can be found in Antonelli (1996), El Karoui et al. (1997), Barles et al. (1997), as well as Peng (2004b) and the references therein. Discrete-time approximation results related to the one from Subsection 2.3.3 can be found in, e.g., Zhang (2004), Bouchard and Touzi (2004) and Cheridito and Stadje (2013); see also Bouchard and Elie (2008) and the references therein.

2.4 EXISTENCE OF APPROPRIATE DOMAINS

This section is concerned with the existence and construction of appropriate domains for sublinear expectations. The idea is the following: Suppose that, by some procedure, a sublinear expectation has been defined on a "small" space of bounded random variables $\mathcal{H}$. This is the case, for instance, for the G-expectation and the random G-expectation. Then a classical representation result guarantees that $\mathcal{E}_0^{\text{sub}}$ can be represented via

$$\mathcal{E}_0^{\text{sub}}[X] = \sup_{q \in Q} \int_\Omega X(\omega) q(d\omega) \quad \text{for all } X \in \mathcal{H},$$

where Q is a family of finitely additive probabilities. This representation allows us to extend $\mathcal{E}_0^{\text{sub}}$ to all positive random variables. If this extension satisfies the *Fatou property*

$$\mathcal{E}_0^{\text{sub}}\left[\liminf_{n\to\infty} X_n\right] \leqslant \liminf_{n\to\infty} \mathcal{E}_0^{\text{sub}}[X_n],$$

then completeness of L^p-type spaces can be shown along the lines of a classical argument. Subsequently, an appropriate domain can be obtained by taking closures of the original domain $\mathcal{H}$ in those L^p-type spaces. In concrete examples, the family Q will consist of countably additive probabilities, and hence the Fatou property will always be satisfied (see Example 2.39, p. 38).

STANDING ASSUMPTIONS Throughout this entire section, $(\Omega, \mathcal{A})$ is a measurable space and $(\mathcal{F}_t)_{t\in[0,T]}$ is a filtration with $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_T = \mathcal{A}$. Unless explicitly stated otherwise, all notions requiring a measurable space are to be understood with respect to $(\Omega, \mathcal{A})$. If $\mathcal{G} \subset 2^\Omega$ is a σ-algebra and $(S, \mathcal{S})$ is a measurable space, we denote the space of all measurable functions $\Omega \to S$ by $\mathcal{L}^0(\mathcal{G}; S)$. If $S = \mathbb{R}$, we simply write $\mathcal{L}^0(\mathcal{G}) \triangleq \mathcal{L}^0(\mathcal{G}; \mathbb{R})$; the subspace of all bounded measurable functions is denoted by $\mathcal{L}^\infty(\mathcal{G}) \subset \mathcal{L}^0(\mathcal{G})$.

From here on out, $\mathcal{H} \subset \mathcal{L}^\infty(\mathcal{A})$ is a $|\cdot|^p$-stable algebra of functions containing all constants, i.e.,

(H1) $\mathcal{H} \leqslant \mathcal{L}^\infty(\mathcal{A})$,

(H2) $\alpha g + \beta h \in \mathcal{H}$, if $\alpha, \beta \in \mathbb{R}$, $g, h \in \mathcal{H}$,

(H3) $fg \in \mathcal{H}$, if $f, g \in \mathcal{H}$,

(H4) $|h|^p \in \mathcal{H}$ for all $p \geqslant 1$, if $h \in \mathcal{H}$,

(H5) $1 \in \mathcal{H}$.

Note: If one seeks to construct an appropriate domain $(L^1_t)_{t\in[0,T]}$ with $p = 1$ fixed, condition (H3) can be dropped and condition (H4) simplifies to the requirement $h \in \mathcal{H} \implies |h| \in \mathcal{H}$. In other words, it suffices that $\mathcal{H} \leqslant \mathcal{L}^\infty(\mathcal{A})$ be a Riesz space.

Remark. A trivial example of such a set $\mathcal{H}$ is of course $\mathcal{L}^\infty(\mathcal{A})$ itself. If Ω is a topological space (e.g., the Wiener space), a frequent choice for $\mathcal{H}$ is the space of bounded (uniformly) continuous functions on Ω. ⋄

We associate a domain for expectations $((\mathcal{H}_t)_{t\in[0,T]}, \leqslant)$ with $\mathcal{H}$ as follows: We take $\mathcal{H}_t \triangleq \mathcal{H} \cap \mathcal{L}^0(\mathcal{F}_t)$, $t \in [0, T]$, as the $|\cdot|^p$-stable subalgebra of all $\mathcal{F}_t$-measurable functions in $\mathcal{H}$ and $\leqslant$ as the pointwise order on $\mathcal{L}^0(\mathcal{A})$. Since $\mathcal{F}_0$ is trivial, we have $\mathcal{H}_0 \cong \mathbb{R}$, and $((\mathcal{H}_t)_{t\in[0,T]}, \leqslant)$ is indeed a domain for expectations. In the following, we consider a

$$\text{sublinear expectation} \quad (\mathcal{E}^{\text{sub}}_t)_{t\in[0,T]} \quad \text{on} \quad ((\mathcal{H}_t)_{t\in[0,T]}, \leqslant).$$

We will prove the existence of an appropriate domain for $(\mathcal{E}^{\text{sub}}_t)_{t\in[0,T]}$, under the assumption that $(\mathcal{E}^{\text{sub}}_t)_{t\in[0,T]}$ has the Fatou property:

THEOREM 2.35. *Suppose that $\mathcal{E}^{\text{sub}}_0$ has the Fatou property as set forth in Definition 2.38. Then $(\mathcal{E}^{\text{sub}}_t)_{t\in[0,T]}$ can be extended to an appropriate domain $\{(L^p_t, \|\cdot\|_{L,p}) : t \in [0, T],\ p \geqslant 1\}$.*

Proof. This will be a consequence of Theorem 2.45 below. □

The following subsections address the proof of Theorem 2.35.

2.4.1 *Static sublinear expectation operators and their representation*

As the norms on an appropriate domain are generated by $\mathcal{E}_0^{\text{sub}}$, it is hardly surprising that the central object in the construction of an appropriate domain for $(\mathcal{E}_t^{\text{sub}})_{t\in[0,T]}$ is the static sublinear expectation operator $\mathcal{E}_0^{\text{sub}}$.

DEFINITION 2.36. Let $K \subset \mathcal{L}^0(\mathcal{A}; [-\infty, \infty])$ be a convex cone containing all constants. A *(static) sublinear expectation operator* is a function $\mathcal{E} : K \to [-\infty, \infty]$ that is

- *sublinear*, i.e., $\mathcal{E}[X + Y] \leqslant \mathcal{E}[X] + \mathcal{E}[Y]$ and $\mathcal{E}[aX] = a\mathcal{E}[X]$ for all $X, Y \in \mathcal{H}$ and $a \geqslant 0$,
- *monotone*, i.e., $\mathcal{E}[X] \leqslant \mathcal{E}[Y]$, for all $X, Y \in \mathcal{H}$ with $X \leqslant Y$, and
- *constant-preserving*, i.e., $\mathcal{E}[a] = a$ for all $a \in \mathbb{R}$. ◇

Here, by a *convex cone* $K \subset \mathcal{L}^0(\mathcal{A}; [-\infty, \infty])$ we mean nothing more than a subset K of $\mathcal{L}^0(\mathcal{A}; [-\infty, \infty])$ which satisfies

$$\alpha X + \beta Y \in K, \qquad \text{whenever} \quad X, Y \in K \quad \text{and} \quad 0 \leqslant \alpha, \beta < \infty.$$

Note: If $\mathcal{E} : \mathcal{H} \subset \mathcal{L}^\infty(\mathcal{A})$ is a sublinear expectation operator, then $\rho : \mathcal{H} \subset \mathcal{L}^\infty(\mathcal{A}) \to \mathbb{R}$, $X \mapsto \mathcal{E}[-X]$ is a *coherent risk measure* (and vice versa); see Artzner et al. (1999) as well as Föllmer and Schied (2004), and the references therein.

It is immediate from the definition of sublinear expectations that $\mathcal{E}_0^{\text{sub}}$ is a static sublinear expectation operator on $\mathcal{H}_T$. The next step is to represent $\mathcal{E}_0^{\text{sub}}$ by a family of finitely additive probabilities.

Sublinear expectation operators and additive probabilities

Every (non-empty) family M of finitely additive probability measures on $(\Omega, \mathcal{A})$ induces a (static) sublinear expectation operator

$$M : \mathcal{K}(M) \to (-\infty, \infty], \quad g \mapsto M[g] \triangleq \sup_{m\in M} \left(\int g^+ dm - \int g^- dm\right)$$

on the convex cone

$$\mathcal{K}(M) \triangleq \left\{ g \in \mathcal{L}^0(\mathcal{A}; (-\infty, \infty]) : \sup_{m\in M} \int g^- dm < \infty \right\} \subset \mathcal{L}^0(\mathcal{A}; [-\infty, \infty]).$$

We refer to Lemma A.14 (on p. 147) for a detailed proof of this fact. For spaces of bounded random variables, the converse is also true by a well-known representation theorem: Every sublinear expectation operator is given by a family of additive probabilities.

THEOREM 2.37. *Let $\mathcal{V} \subset \mathcal{L}^\infty(\mathcal{A})$ be a linear space of bounded measurable functions containing all constants, and let $\mathcal{E}$ be a real-valued function on $\mathcal{V}$. Then $\mathcal{E}$ is a sublinear expectation operator if and only if there exists a (non-empty) family M of finitely additive probability measures such that $\mathcal{E} = M[\cdot]|_{\mathcal{V}}$, i.e.,*

$$\mathcal{E}[h] = \sup_{m\in M} \int h \, dm \qquad \textit{for all } h \in \mathcal{V}.$$

Proof. See Theorem A.36 in Appendix A.2.6 (p. 160ff.). □

Note: The definition of the operator M relies on an integral $\int g\,dm$ of measurable functions $g : \Omega \to [0,\infty]$ with respect to a finitely additive probability measure m. The construction of such integrals is detailed in Appendix A.1, p. 141ff. We shall briefly outline the construction here: The integral of a simple function $g = \sum_{i=1}^{n} a_i 1_{A_i}$ is defined in the usual way, of course, as $\int g\,dm = \sum_{i=1}^{n} a_i m(A_i)$. Clearly, this definition gives rise to a positive linear operator on the space of all simple functions. As in the construction of the Lebesgue integral, this operator is extended to all positive measurable functions by setting $\int g\,dm \triangleq \sup\{\int h\,dm : h \text{ is a simple function with } h \leqslant g\}$. Of course, monotone convergence fails for this integral (if m is not countably additive), and thus some extra work is required to prove additivity of the extension.

2.4.2 *The Fatou property and its ramifications*

The domain of the sublinear expectation $(\mathcal{E}_t^{\text{sub}})_{t\in[0,T]}$ under consideration in this section is contained in $\mathcal{L}^\infty(\mathcal{A})$, and hence Theorem 2.37 shows that there exists a family Q of finitely additive probability measures such that

$$\mathcal{E}_0^{\text{sub}}[X] = Q[X] = \sup_{q\in Q} \int X(\omega)\,q(d\omega) \quad \text{for all } X \in \mathcal{H}. \tag{2.28}$$

We refer to such a set Q as a *representation* of $\mathcal{E}_0^{\text{sub}}$. In general, (2.28) does not uniquely determine the representation Q, of course. Any representation provides a way to extend $\mathcal{E}_0^{\text{sub}}$ to all measurable functions. For our purposes, it suffices if one of these extensions is well-behaved.

Fatou property

DEFINITION 2.38. A family of finitely additive probability measures Q has the *Fatou property* if

$$Q\big[\liminf_{n\to\infty} X_n\big] \leqslant \liminf_{n\to\infty} Q\big[X_n\big] \quad \text{for all } (X_n)_{n\in\mathbb{N}} \subset \mathcal{L}^0(\mathcal{A}; [0,\infty]).$$

Correspondingly, we say that $(\mathcal{E}_t^{\text{sub}})_{t\in[0,T]}$ has the *Fatou property* if $\mathcal{E}_0^{\text{sub}}$ admits a representation Q which has the Fatou property. ◇

Fatou property by countable additivity

EXAMPLE 2.39. If Q is a family of countably additive probabilities, then Q has the Fatou property: Indeed, for every countably additive probability $q \in Q$, Fatou's lemma implies $q[\liminf_{n\to\infty} X_n] \leqslant \liminf_{n\to\infty} q[X_n]$ whenever $(X_n)_{n\in\mathbb{N}} \subset \mathcal{L}_+^0(\mathcal{A}; [0,\infty])$, and hence

$$Q[X] = \sup_{q\in Q} q[X] \leqslant \sup_{q\in Q} \liminf_{n\to\infty} q[X_n] \leqslant \liminf_{n\to\infty} \sup_{q\in Q} q[X_n],$$

which establishes the Fatou property. ◇

Banach space of Q-integrable random variables

Let Q be a family of finitely additive probabilities. As a domain for the associated sublinear expectation operator $Q[\cdot]$, it is natural to consider

$$\mathcal{L}^p(Q) \triangleq \left\{X \in \mathcal{L}^0(\mathcal{A}) : \|X\|_{L,p} < \infty\right\}, \text{ where } \|X\|_{L,p} \triangleq \big(Q\big[|X|^p\big]\big)^{\frac{1}{p}}.$$

Properties of $\mathcal{L}^p(Q)$

PROPOSITION 2.40. *For each* $p \geqslant 1$, $(\mathcal{L}^p(Q), \|\cdot\|_{L,p})$ *is a seminormed space. The spaces are related by Hölder's inequality: If* $\frac{1}{p}+\frac{1}{q}=1$, *then*

$$\|XY\|_{L,1} \leqslant \|X\|_{L^p}\|Y\|_{L,q} \quad \textit{for all } X,Y \in \mathcal{L}^0(\mathcal{A}),$$

and thus $\mathcal{L}^p(Q)\cdot\mathcal{L}^q(Q) \subset \mathcal{L}^1(Q)$.

If Q *has the Fatou property, then* $\mathcal{L}^p(Q)$ *is complete and*

$$\|X\|_{L,p} = 0 \iff Q\big(X \neq 0\big) = \sup_{q\in Q} q\big(\{\omega \in \Omega : X(\omega) \neq 0\}\big) = 0.$$

Proof. The proofs of these results can be found in Appendix A.2 (p. 147ff.). See Lemma A.15 for Hölder's inequality, Corollary A.18 for the seminorms and Corollary A.27 for the completeness result. The key steps in the proof of the completeness result are based on the Fatou property of Q and are thus very close to the classical Riesz-Fischer theorem; for details see Appendix A.2.3 (p. 152ff.). □

Let us refer to a measurable set $N \in \mathcal{A}$ as Q-negligible if

$$Q(N) = \sup_{q\in Q} q(N) = 0,$$

and let us write $\mathcal{N}_Q$ for the collection of all Q-negligible sets. Then $\mathcal{N}_Q$ is a collection of negligible sets, as in the context of Lebesgue families (see Subsection 2.2.1 above, p.16ff.), and thus induces an equivalence relation $\sim_Q$ on $\Omega^{\mathbb{R}}$ via

$$f \sim_Q g \iff f(\omega) = g(\omega) \quad \text{for all } \omega \in \Omega \setminus N \text{ and some } N \in \mathcal{N}_Q.$$

Given $f : \Omega \to \mathbb{R}$, we write $[f]_Q$ for its $\sim_Q$-equivalence class and

$$L^p(Q) \triangleq \big\{[f]_Q : f \in \mathcal{L}^p(Q)\big\} \leqslant \Omega^{\mathbb{R}}/\mathcal{N}_Q \triangleq \Omega^{\mathbb{R}}/\sim_Q$$

for the set of all equivalence classes containing a member of $\mathcal{L}^p(Q)$.

Note: The above procedure is slightly different to simply identifying two functions $f, g \in \mathcal{L}^p(Q)$ if $\|f-g\|_{L,p} = 0$ because an equivalence class $[f]_Q$ may contain a function g which is not measurable. This is due to fact that the notion of a Lebesgue family is, by design, not a measure theoretic one. Instead, a Lebesgue family specifically asks its spaces to consist of equivalence classes in $\Omega^{\mathbb{R}}$ (rather than, for instance, in $\mathcal{L}^0(\mathcal{A})$). This is a rather technical point that is somewhat similar to the completion of a measure space and which is – as Lemma 2.41 below shows – of no importance in the following.

The sublinear expectation operator $Q[\cdot]$ can be considered as a mapping defined on $L^1(Q)$ without any difficulties:

LEMMA 2.41. *For all* $p \geqslant 1$, *the operator*

$$Q : L^1(Q) \to \mathbb{R}, \quad X \mapsto Q[f], \quad \textit{where } f \in X \cap \mathcal{L}^0(\mathcal{A}),$$

is well-defined, and $(L^p, \|\cdot\|_{L,p})$ *is a normed space.*

Proof. This is essentially trivial; see Lemma A.28, p. 155. □

Combining this result with Proposition 2.40, we get

COROLLARY 2.42. *If* Q *has the Fatou property, then* $L^p(Q) \leqslant \Omega^{\mathbb{R}}/\mathcal{N}_Q$ *is a Banach space for all* $p \geqslant 1$.

We have thus completed the first part of the program. It is now straightforward to extend $\mathcal{E}_t^{\text{sub}}$ to the closure of $\mathcal{H}_t$ in $\mathcal{L}^p(Q)$.

2.4.3 *Appropriate domains*

We now consider the domain $(\mathcal{H}_t)_{t\in[0,T]}$ of $(\mathcal{E}_t^{\text{sub}})_{t\in[0,T]}$ as a family of subspaces of $\Omega^{\mathbb{R}}/\mathcal{N}_Q$, using the same procedure as for $\mathcal{L}^p(Q)$ above: For each $t \in [0,T]$, we put

$$H_t(Q) \triangleq \left\{[h]_Q \,:\, h \in \mathcal{H}_t\right\}.$$

LEMMA 2.43. *For all* $t \in [0,T]$ *and each* $p \geqslant 1$,

$$H_t(Q) \subset L^p(Q) \subset \Omega^{\mathbb{R}}/\mathcal{N}_Q$$

is a $|\cdot|^p$*-stable algebra containing all constants, that is, it satisfies* (H2)-(H5). *Moreover,* $(\mathcal{H}_t)_{t\in[0,T]}$ *is increasing.*

Proof. By choice of $(\mathcal{H}_t)_{t\in[0,T]}$, we have $\mathcal{H}_t \subset \mathcal{L}^\infty(\mathcal{F}_t)$, and thus $\mathcal{H}_t \subset \mathcal{L}^p(Q)$. This immediately yields $H_t(Q) \subset L^p(Q)$. Since both $\mathcal{H}$ and $\mathcal{L}^0(\mathcal{F}_t)$ satisfy (H2)-(H5), so does their intersection $\mathcal{H}_t$. Since $\sim_Q$-equivalence is compatible with pointwise operations, it is immediate that $H_t(Q)$ also satisfies (H2)-(H5). □

Every nonlinear expectation $(\mathcal{E}_t)_{t\in[0,T]}$ dominated by $(\mathcal{E}_t^{\text{sub}})_{t\in[0,T]}$ is continuous for the seminorm $\mathcal{E}_0^{\text{sub}}[|\cdot|^p]^{\frac{1}{p}}$, and this seminorm coincides with $\|\cdot\|_{L,p}$ by the probabilistic representation (2.28); thus, $(\mathcal{E}_t)_{t\in[0,T]}$ will extend continuously (and hence uniquely) to a family of operators defined on the L^p-closures of $H_t(Q)$, $t \in [0,T]$. We now study these closures.

THEOREM 2.44. *The family* $\left\{\overline{H_t(Q)}^p \,:\, t \in [0,T],\ p \geqslant 1\right\}$ *of* $L^p(Q)$*-closures*

$$\overline{H_t(Q)}^p \triangleq \operatorname{clos}(H_t(Q); L^p(Q)) \subset L^p(Q)$$

is a Lebesgue family if Q *has the Fatou property.*

Proof. By Lemma 2.43, $H_t(Q) \subset L^p(Q)$ is a linear space; thus,

$$\overline{H_t(Q)}^p = \operatorname{clos}(H_t(Q); L^p(Q)) \subset \Omega^{\mathbb{R}}/\mathcal{N}_Q$$

is a Banach space by completeness of $L^p(Q)$ (Corollary 2.42). Now, we verify properties (L1)-(L6) from Definition 2.6. We sketch the key steps here and refer to Theorem A.35 (p. 159) for a detailed proof

(L1) By construction.

(L2) Since $\mathcal{H}_s(Q) \subset \mathcal{H}_t(Q)$ if $s \leqslant t$, and $L^p(Q)$-convergence implies $L^q(Q)$-convergence if $p \geqslant q$, we have $\overline{H_s(Q)}^p \subset \overline{H_t(Q)}^q$.

(L3) We obtain $L_t^p = \{X \in L_t^1 : |X|^p \in L_t^1\}$ because $\mathcal{H}$ is $|\cdot|^p$-stable.

(L4) Since $\mathcal{H}$ is stable under multiplication, the inclusion $\overline{H_t(Q)}^p \cdot \overline{H_t(Q)}^q \subset \overline{H_t(Q)}^1$ follows from Hölder's inequality.

(L5) Since $\mathcal{F}_0$ is trivial, $H_t(Q) = \mathcal{H} \cap \mathcal{L}^0(\mathcal{F}_0)$ consists solely of constants and (L5) is obvious.

(L6) The positive cone is norm-closed, because $L^p(Q)$-convergence implies pointwise convergence for a subsequence (outside of a Q-negligible set).

□

It is straightforward to extend $(\mathcal{E}_t^{\text{sub}})_{t\in[0,T]}$ to $\overline{H_T(Q)}^1$ in order to obtain a sublinear expectation on an appropriate domain:

THEOREM 2.45. *Let Q be a representation of $\mathcal{E}_0^{\text{sub}}$ which has the Fatou property. Then every nonlinear expectation $(\mathcal{E}_t)_{t\in[0,T]}$ on $(\mathcal{H}_t)_{t\in[0,T]}$ which is dominated by $(\mathcal{E}_t^{\text{sub}})_{t\in[0,T]}$ admits a unique extension*

$$\overline{\mathcal{E}_t} : \overline{H_T(Q)}^1 \to \overline{H_t(Q)}^1$$

that is characterized by the following properties:

(i) $\left\|\overline{\mathcal{E}_t}[X] - \overline{\mathcal{E}_t}[Y]\right\|_{L,1} \leqslant \|X - Y\|_{L,1}$ *for all* $X, Y \in \overline{H_T(Q)}^1$,

(ii) $\overline{\mathcal{E}_t}\big[\,[h]_Q\,\big] = \mathcal{E}_t[h]$ *for all* $h \in \mathcal{H}$.

In view of (ii), this extension is denoted by $(\mathcal{E}_t)_{t\in[0,T]}$, as well. It is a nonlinear expectation on $(\overline{H_t(Q)}^1)_{t\in[0,T]}$. If $(\mathcal{E}_t)_{t\in[0,T]}$ is sublinear or superlinear on $(\mathcal{H}_t)_{t\in[0,T]}$, then so is its extension.

(a) The Lebesgue family $\{\overline{H_t(Q)}^p : t \in [0,T],\ p \geqslant 1\}$ is an appropriate domain for the extension of $(\mathcal{E}_t^{\text{sub}})_{t\in[0,T]}$.

(b) The extension of $(\mathcal{E}_t)_{t\in[0,T]}$ is a dominated nonlinear expectation carried by $(\mathcal{E}_t^{\text{sub}})_{t\in[0,T]}$ on $\{\overline{H_t(Q)}^p : t \in [0,T],\ p \geqslant 1\}$.

Proof. Since $(\mathcal{E}_t^{\text{sub}})_{t\in[0,T]}$ dominates $(\mathcal{E}_t)_{t\in[0,T]}$ on $(\mathcal{H}_t)_{t\in[0,T]}$, we get $\mathcal{E}_t[g] - \mathcal{E}_t[h] \leqslant \mathcal{E}_t^{\text{sub}}[g-h]$ for all $g, h \in \mathcal{H}$. Moreover, $|g-h| \in \mathcal{H}$ by (H4) and (H2), and hence monotonicity (M) of $\mathcal{E}_t^{\text{sub}}$ implies $\mathcal{E}_t[g] - \mathcal{E}_t[h] \leqslant \mathcal{E}_t^{\text{sub}}[|g-h|]$. Reversing the roles of g and h, we obtain

$$\big|\mathcal{E}_t[g] - \mathcal{E}_t[h]\big| \leqslant \mathcal{E}_t^{\text{sub}}\big[|g-h|\big] \quad \text{for all } g, h \in \mathcal{H}. \tag{2.29}$$

Thus monotonicity (M) and time-consistency (TC) yield

$$\mathcal{E}_0^{\text{sub}}\big[\big|\mathcal{E}_t[g] - \mathcal{E}_t[h]\big|\big] \leqslant \mathcal{E}_0^{\text{sub}}\big[|g-h|\big].$$

Now, $\mathcal{E}_0^{\text{sub}} = Q[\cdot]|_{\mathcal{H}}$, so we have shown that

$$\left\|\mathcal{E}_t[g] - \mathcal{E}_t[h]\right\|_{L,1} \leqslant \|g-h\|_{L,1}.$$

This implies that the operator

$$\mathcal{E}_t : H_T(Q) \to H_t(Q), \quad X \mapsto \left[\mathcal{E}_t[h]\right]_Q, \quad \text{where } h \in \mathcal{H} \cap X,$$

is well-defined and $\|\cdot\|_{L,1}$-continuous; hence, it admits an extension $\overline{\mathcal{E}_t} : \overline{H_T(Q)}^1 \to \overline{H_t(Q)}^1$ which satisfies (i) and (ii) and is thus uniquely determined.

It is straightforward to extend the properties of nonlinear expectations from $(\mathcal{E}_t)_{t\in[0,T]}$ to $(\overline{\mathcal{E}_t})_{t\in[0,T]}$: For brevity, we write $L_t^1 \triangleq \overline{H_t(Q)}^1$. Let $t \in [0,T]$ and $X, Y \in L^1$, and take $X_n, Y_n \in \mathcal{H}$ such that $(X_n, Y_n) \to (X,Y)$ in $L^1(Q)$.

(M) Suppose that $X \leqslant Y$ and set $\xi^n \triangleq X_n \wedge Y_n$ and $\eta^n \triangleq X_n \vee Y_n$. Then $\xi^n, \eta^n \in \mathcal{H}$ since $\mathcal{H}$ is $|\cdot|$-stable, $\xi^n \leqslant \eta^n$ and $(\xi^n, \eta^n) \to (X,Y)$ in $L^1(Q)$; (M) on $(\mathcal{H}_t)_{t\in[0,T]}$ implies $\mathcal{E}_t^{\text{sub}}[\xi^n] \leqslant \mathcal{E}_t^{\text{sub}}[\eta^n]$, and we conclude $\mathcal{E}_t^{\text{sub}}[X] \leqslant \mathcal{E}_t^{\text{sub}}[Y]$ since the positive cone in L^1 is closed (property (L6) of a Lebesgue family).

(SI) If $X \in L_t^1$, we can assume $X_n \in \mathcal{H}_t$. Then (SI) on $(\mathcal{H}_t)_{t\in[0,T]}$ yields $\mathcal{E}_t^{\text{sub}}[X_n + Y_n] = \left(X_n + \mathcal{E}_t^{\text{sub}}[Y_n]\right)$, and $\mathcal{E}_t^{\text{sub}}[X+Y] = X + \mathcal{E}_t^{\text{sub}}[Y]$ follows upon sending $n \to \infty$.

(TC) For every $s \in [0,T]$, we have $\mathcal{E}_s^{\text{sub}}[X_n] \in \mathcal{H}_s$ and $\mathcal{E}_s^{\text{sub}}[X_n] \to \mathcal{E}_s^{\text{sub}}[X]$; therefore, (TC) on $(\mathcal{H}_t)_{t\in[0,T]}$ and continuity imply

$$\mathcal{E}_t^{\text{sub}}[\mathcal{E}_s^{\text{sub}}[X]] = \lim_{n\to\infty} \mathcal{E}_t^{\text{sub}}[\mathcal{E}_s^{\text{sub}}[X_n]] = \lim_{n\to\infty} \mathcal{E}_t^{\text{sub}}[X_n] = \mathcal{E}_t^{\text{sub}}[X].$$

(N) Obvious since $0 \in \mathcal{H}$.

We have thus shown that $(\overline{\mathcal{E}_t})_{t\in[0,T]}$ is a nonlinear expectation on $(L_t^1)_{t\in[0,T]}$. If $(\mathcal{E}_t)_{t\in[0,T]}$ is sublinear, then so is $(\overline{\mathcal{E}_t})_{t\in[0,T]}$:

(PH) For $\lambda \geqslant 0$, we have $\|\lambda X_n - \lambda X\|_{L,1} = \lambda\|X_n - X\|_{L,1} \to 0$ and $\lambda X_n \in \mathcal{H}$; thus, $\mathcal{E}_t[\lambda X] = \lambda \mathcal{E}_t[X]$ follows by (PH) on $(\mathcal{H}_t)_{t\in[0,T]}$.

(SUB) By (SUB) on $(\mathcal{H}_t)_{t\in[0,T]}$, we have $\mathcal{E}_t[X_n + Y_n] \leqslant \mathcal{E}_t[X_n] + \mathcal{E}_t[Y_n]$ and this inequality is preserved in the limit by (L6).

If $(\mathcal{E}_t)_{t\in[0,T]}$ is superlinear, we apply the previous reasoning to the sublinear expectation $(-\mathcal{E}_t[-\cdot])_{t\in[0,T]}$, to conclude that $(\overline{\mathcal{E}_t})_{t\in[0,T]}$ is superlinear.

We have completed the first part of the proof. In particular, we have seen that $(\mathcal{E}_t^{\text{sub}})_{t\in[0,T]}$ extends to a sublinear expectation on $(L_t^1)_{t\in[0,T]}$ (again denoted by $(\mathcal{E}_t^{\text{sub}})_{t\in[0,T]}$). By Theorem 2.44, the Banach spaces $L_t^p \triangleq \overline{H_t(Q)}^p$ form a Lebesgue family. Next, we show that the extensions $\overline{\mathcal{E}_t}$ map L^p into L_t^p.

Since the positive cone in L^1 is closed (L6) and $(\overline{\mathcal{E}_t})_{t\in[0,T]}$ is continuous, the triangle inequality from (2.29) extends to all of L^1, i.e.,

$$\left|\overline{\mathcal{E}_t}[X]-\overline{\mathcal{E}_t}[Y]\right| \leqslant \mathcal{E}^{\text{sub}}_t\left[|X-Y|\right] \quad \text{for all } X,Y\in L^1.$$

Now, Jensen's inequality 2.11 implies

$$\mathcal{E}^{\text{sub}}_t\left[|X-Y|\right]^p \leqslant \mathcal{E}^{\text{sub}}_t\left[|X-Y|^p\right] \quad \text{for all } X,Y\in L^p,$$

and thus we have

$$\left\|\overline{\mathcal{E}_t}[X]-\overline{\mathcal{E}_t}[Y]\right\|_{L,p} \leqslant \|X-Y\|_{L,p}. \tag{2.30}$$

Indeed, $Q[\cdot]: L^1(Q)\to\mathbb{R}$ is a continuous operator with $Q[h]=\mathcal{E}^{\text{sub}}_0[h]$ for all $h\in\mathcal{H}$ and thus coincides with the extension $\mathcal{E}^{\text{sub}}_0$ on the closure L^1 of $\mathcal{H}$. Hence, applying $Q[\cdot]$ to the inequality

$$\left|\overline{\mathcal{E}_t}[X]-\overline{\mathcal{E}_t}[Y]\right|^p \leqslant \mathcal{E}^{\text{sub}}_t\left[|X-Y|^p\right] \quad \text{in } \mathcal{L}^0(\mathcal{A}),$$

we obtain

$$Q\left[\left|\overline{\mathcal{E}_t}[X]-\overline{\mathcal{E}_t}[Y]\right|^p\right] \leqslant Q\left[\mathcal{E}^{\text{sub}}_t\left[|X-Y|^p\right]\right] = \mathcal{E}^{\text{sub}}_0\left[\mathcal{E}^{\text{sub}}_t\left[|X-Y|^p\right]\right],$$

because $\mathcal{E}^{\text{sub}}_t : L^1 \to L^1_t$ and $|X-Y|^p \in L^1$ by property (L3) of a Lebesgue family. Now, $(\mathcal{E}^{\text{sub}}_t)_{t\in[0,T]}$ is a sublinear expectation on $(L^1_t)_{t\in[0,T]}$, and thus (2.30) follows by time-consistency (TC); we have shown that $\overline{\mathcal{E}_t}$ is $\|\cdot\|_{L,p}$-continuous. Now, it is easy to prove that $\overline{\mathcal{E}_t}: L^p\to L^p_t$:

Let $X\in L^p$ and take $X_n\in\mathcal{H}$ with $\|X_n-X\|_{L,p}\to 0$. Then $\overline{\mathcal{E}_t}[X_n]\in H_t(Q)\subset L^p_t$, and (2.30) shows that $\overline{\mathcal{E}_t}[X_n]$ converges in $L^p(Q)$. Thus $\overline{\mathcal{E}_t}[X]=\lim_{n\to\infty}\overline{\mathcal{E}_t}[X_n]\in L^p_t$ and

$$\overline{\mathcal{E}_t}: L^p\to L^p_t \quad \text{and} \quad \mathcal{E}^{\text{sub}}_t: L^p\to L^p_t.$$

Moreover, $\|X\|_{L,p}=\mathcal{E}^{\text{sub}}_0[|X|^p]^{\frac{1}{p}}$ for $X\in L^p$ since $Q[\cdot]=\mathcal{E}^{\text{sub}}_0$ on L^1. Hence $\{L^p_t : t\in[0,T],\ p\geqslant 1\}$ is an appropriate domain for $(\mathcal{E}^{\text{sub}}_t)_{t\in[0,T]}$, and (a) is proven. To establish (b), it remains to show that $(\mathcal{E}^{\text{sub}}_t)_{t\in[0,T]}$ dominates $(\overline{\mathcal{E}_t})_{t\in[0,T]}$. This is the case on $\mathcal{H}$, i.e.,

$$\mathcal{E}_t[g]-\mathcal{E}_t[h] \leqslant \mathcal{E}^{\text{sub}}_t[g-h] \quad \text{for all } g,h\in\mathcal{H},$$

and extends readily to L^1 by (L6) and continuity. □

If Q and M are representations of $\mathcal{E}^{\text{sub}}_0$ which have the Fatou property, then the spaces $\overline{H_t(Q)}^p$ and $\overline{H_t(M)}^p$ are isometrically order-isomorphic, and both of them are concrete representations of the abstract completion of $\mathcal{H}_t$ with respect to the seminorm $\mathcal{E}^{\text{sub}}_0[|\cdot|^p]^{\frac{1}{p}}$. Moreover, the values of the extension of $\mathcal{E}_t$ are uniquely determined by their values on $\mathcal{H}$. Hence it is completely justified to refer to the extension of $(\mathcal{E}_t)_{t\in[0,T]}$ provided by Theorem 2.45 as the *unique continuous extension*.

2.4.4 *Bibliographical notes*

The representation result (Theorem 2.37) can be found in Huber (1981), or, in the context of coherent risk measures, in Artzner et al. (1999). These authors assume that $|\Omega| < \infty$ in order to obtain a representation in terms of genuine, "countably additive" probability measures. A formulation for general measurable spaces can be found in Föllmer and Schied (2004); see their Proposition 4.14.

The completeness result for the spaces $L^p(Q)$ (Proposition 2.40) is probably very well-known; however, we have not been able to locate a reference in the literature. By the Fatou property, the proof of completeness is virtually the same as for the classical Lebesgue spaces; see, e.g., Rudin (1974), Theorem 3.11, p. 69ff. Under some conditions on $(\Omega, \mathcal{A})$ and Q, the set function $A \mapsto Q(A)$ can be viewed as a Choquet capacity (see, e.g., Choquet (1954, 1959) and Dellacherie (1972)). This is exploited in Denis and Martini (2006) and Denis et al. (2011).

The general extension result from Theorem 2.45 is based on simple arguments that appear repeatedly in the context of nonlinear expectations; see, e.g., Peng (2004b, 2005b, 2008, 2010), and Nutz (2012, 2013).

2.5 EXAMPLES OF NONLINEAR EXPECTATIONS

In this section, we illustrate that the general framework of Section 2.2 and Section 2.3 is very well suited for the study of backward nonlinear expectation equations: We investigate several concrete examples of nonlinear expectations from the literature and show that these are – in the sense of this thesis – *regular* nonlinear expectations defined on *appropriate domains*.

In Subsection 2.5.1, a class of random G-expectations is considered. The classical G-expectation is the topic of Subsection 2.5.2. Subsection 2.5.3 deals with the g-expectation, and an application to robust expectations with ambiguity about the drift and the intensity of jumps is provided in Subsection 2.5.4.

In each subsection, we briefly outline the construction of the nonlinear expectation under consideration and record some of its important properties. Then we make use of the general results from the previous Section 2.4 to construct an *appropriate domain* for the nonlinear expectation. Finally, we prove that the nonlinear expectation is *regular*. As pointed out in Section 2.3 above, this regularity property is crucial in the theory of backward nonlinear expectation equations.

2.5.1 *A class of random G-expectations*

This subsection is devoted to a class of random G-expectations which are defined in terms of non-Markovian control problems. These sublinear expectations have been introduced by Nutz (2012) and can be seen as variants of the Nutz (2013) random G-expectations; see also Nutz and Soner (2012) for related results. The classical, non-random G-expectations developed in Peng (2007, 2008), and Denis et al. (2011), which will be considered separately, in Subsection 2.5.2 below, also belong to this class of random G-expectations. Moreover, random G-expectations are also of interest in economics; see, e.g., Epstein and Ji (2014) and Section 2.7 below.

Following the program outlined above, we give a brief review of the construction of the relevant random G-expectations and show that they are sublinear expectations on an appropriate domain. Then, as the main result of this subsection, we prove that these random G-expectations are *regular*, provided that the coefficients of the dynamics of the state process are *bounded*.

Preliminaries

Following Nutz (2012), we briefly review the construction of the relevant random G-expectation $(\mathcal{E}_t^{\text{sub}})_{t\in[0,T]}$. For additional details we refer to Nutz (2012).

For $0 \leqslant t \leqslant s \leqslant T$ consider the canonical Wiener space

$$\Omega_s^t \triangleq \{\omega : [t,s] \to \mathbb{R}^d \ : \ \omega \text{ is continuous with } \omega(t) = 0\}$$

equipped with the supremum norm $\|\cdot\|_\infty$ and write $\Omega^t \triangleq \Omega_T^t$ and $\Omega_s \triangleq \Omega_s^0$. Let $W_s^t(\omega) = \omega_s$ be the canonical process on Ω^t, P^t the Wiener measure on Ω^t and let $\mathcal{F}^t$ be the raw filtration generated by W^t. Moreover, we denote by $\Omega \triangleq C([0,T];\mathbb{R}^d)$ the space of all continuous paths.

Next, let U be a non-empty Borel subset of $\mathbb{R}^m$ and fix functions

$$\mu : [0,T] \times \Omega \times U \to \mathbb{R}^d \text{ and } \sigma : [0,T] \times \Omega \times U \to \mathbb{R}^{d\times d}$$

such that $(r,\omega) \mapsto \mu(r, X(\omega), \nu_r(\omega))$ and $(r,\omega) \mapsto \sigma(r, X(\omega), \nu_r(\omega))$ are progressively measurable whenever X is continuous and adapted and ν is progressive. The functions $\mu(r,\cdot,u)$ and $\sigma(r,\cdot,u)$ are assumed to be Lipschitz continuous, uniformly in (r,u). For $\eta \in \Omega$ and $t \in [0,T]$, the conditioned coefficients are defined as

$$\begin{aligned} \mu^{t,\eta} &: [0,T] \times \Omega^t \times U \to \mathbb{R}^d, & \mu^{t,\eta}(r,\omega,u) &\triangleq \mu(r, \eta \otimes_t \omega, u), \\ \sigma^{t,\eta} &: [0,T] \times \Omega^t \times U \to \mathbb{R}^{d\times d}, & \sigma^{t,\eta}(r,\omega,u) &\triangleq \sigma(r, \eta \otimes_t \omega, u), \end{aligned}$$

where $(\eta \otimes_t \omega)_r \triangleq \eta_r 1_{[0,t)}(r) + (\eta_t + \omega_r) 1_{[t,T]}(r)$ for $r \in [0,T]$.

Let $\mathcal{U}^t$ denote the set of all $\mathcal{F}^t$-progressively measurable, U-valued processes ν such that

$$\int_t^T |\mu(r, X, \nu_r)| + |\sigma(r, X, \nu_r)|^2 \, dr < \infty$$

for all $\mathcal{F}^t$-adapted continuous processes $(X_t)_{t\in[0,T]}$. Then, for every $\nu \in \mathcal{U}^t$ and under each P^t, the SDE

$$X_s = \eta_t + \int_t^s \mu^{t,\eta}(r, X, \nu_r)dr + \int_t^s \sigma^{t,\eta}(r, X, \nu_r)dW_r^t, \quad s \in [t, T], \quad (2.31)$$

admits a P^t-a.s. unique solution $X = X(t, \eta, \nu)$. Below, the sublinear expectation will be constructed in terms of the measures

$$P(t, \eta, \nu) \triangleq P^t \circ (X(t, \eta, \nu) - \eta_t)^{-1}. \quad (2.32)$$

Now, an initial condition $x \in \mathbb{R}^d$ is fixed, and it is assumed that $\mathcal{F} \subseteq \mathcal{F}^X$, where $\mathcal{F}^X$ is the P-augmentation of the filtration generated by $\{X(0, x, \nu) : \nu \in \mathcal{U}^0\}$.

Note: The property $\mathcal{F} \subseteq \mathcal{F}^X$ is not automatically satisfied. In fact, one can even give examples where the filtration $\mathcal{F}^{\tilde{W}}$ generated by drift-changed Brownian motion $\tilde{W}$ – although Brownian – is strictly contained in the filtration $\mathcal{F}^W$ generated by the original Brownian motion W; see, for instance, Feldman and Smorodinsky (1997). Intuitively, having $\mathcal{F} \subseteq \mathcal{F}^X$ means that one is able to distinguish between drift and volatility by observing paths. For a sufficient condition, see Remark 2.2 in Nutz (2012).

Construction of the random G-expectation for regular random variables

Let $\Omega_t^x = x + \Omega_t$ be the space of all continuous paths $\omega : [0, t] \to \mathbb{R}^d$ with $\omega_0 = x$ and abbreviate $\Omega^x = \Omega_T^x$. We write $UC_b(\Omega_t^x)$ for the space of bounded, uniformly continuous functions $\Omega_t^x \to \mathbb{R}$.

One may view $\Omega_t^x \subset \Omega^x$ as a closed subspace via the inclusion $\omega \mapsto \omega_{\cdot\wedge t}$; correspondingly, one may identify $UC_b(\Omega_t^x)$ with the closed subspace $\mathcal{H}_t \triangleq UC_b(\Omega^x) \cap \mathcal{L}^0(\mathcal{F}_t)$ of $UC_b(\Omega^x)$.

We note that $UC_b(\Omega^x) \subset \mathcal{L}^\infty(\mathcal{F}_T)$ is a $|\cdot|^p$-stable algebra of functions which contains all constants, as required for the general existence theory for appropriate domains of Section 2.4 above. In particular, $(\mathcal{H}_t)_{t\in[0,T]}$ forms a domain for expectations as considered in Section 2.4.

Following Nutz (2012), for $\xi \in UC_b(\Omega^x)$, the random G-expectation with fixed initial condition x is defined ω by ω as the value function

$$V_t^x(\xi, \omega) \triangleq \sup_{\nu \in \mathcal{U}^t} E^{P(t,\omega,\nu)}[\xi^{t,\omega}], \quad (t, \omega) \in [0, T] \times \Omega^x, \quad (2.33)$$

where $\xi^{t,\omega}(\bar{\omega}) \triangleq \xi(\omega \otimes_t \bar{\omega})$ for $\bar{\omega} \in \Omega^t$ and $t \in [0, T]$. The measures $P(t, \omega, \nu)$ are given in (2.32).

The following result is implicitly contained in Nutz (2012).

PROPOSITION 2.46. *The family of operators $(V_t^x)_{t\in[0,T]}$, given by (2.33), forms a sublinear expectation on $(\mathcal{H}_t)_{t\in[0,T]}$.*

Proof. Lemma 4.3 in Nutz (2012) guarantees that V_t^x maps $\mathcal{H}_T = UC_b(\Omega^x)$ to $\mathcal{H}_t = UC_b(\Omega_t^x)$. It is immediate from the definition of V_t^x (as a supremum of normalized positive linear operators) that V_t^x is monotone (M), normalized (N), positively homogeneous (PH) and subadditive (SUB) for each $t \in [0, T]$. If $\xi \in \mathcal{H}_t = UC_b(\Omega_t^x)$, then $\xi^{t,\omega} = \xi(\omega)$;

hence $V_t^x(\xi) = \xi$ and V_t preserves constants (PC). Thus V_t^x is shift-invariant (SI) by Corollary 2.4. Time-consistency (TC) is guaranteed by a deep result of Nutz (2012); see his Theorem 3.2. □

Extension to an appropriate domain

For $t = 0$, the definition of the random G-expectation (2.33) becomes

$$V_0^x(\xi) = \sup_{\nu \in \mathcal{U}^0} E^{P(0,x,\nu)}[\xi^{0,x}] = \sup_{P \in \mathbb{P}} E^P[\xi], \qquad \xi \in \mathcal{H}_T,$$

where $\mathbb{P} \triangleq \{P(0,x,\nu) \circ (W^0 + x)^{-1} : \nu \in \mathcal{U}^0\}$. Since it consists solely of countably additive probabilities, the sublinear expectation operator $\mathbb{P}[\cdot]$ has the Fatou property. We are thus precisely in the situation of Section 2.4 above:

- The collection $L^p(\mathbb{P})$ of (equivalence classes of) $\mathcal{F}_T$-measurable random variables is a Banach space for the norm $\|\xi\|_{L,p} = \sup_{P \in \mathbb{P}} E^P[|\xi|^p]^{\frac{1}{p}}$ by Corollary 2.42.
- The closures $L_t^p \triangleq \mathrm{clos}(\mathcal{H}_t) \subset L^p(\mathbb{P})$ form a Lebesgue family by Theorem 2.44.
- The sublinear expectation $(V_t^x)_{t \in [0,T]}$ on $(\mathcal{H}_t)_{t \in [0,T]}$ extends continuously to a sublinear expectation $(\mathcal{E}_t^{\text{sub}})_{t \in [0,T]}$ on the *appropriate domain* $\{L_t^p : t \in [0,T],\ p \geqslant 1\}$; see Theorem 2.45.

An equivalent extension is also carried out in Nutz (2012); see his Lemma 4.3. With the above, the random G-expectation $(\mathcal{E}_t^{\text{sub}})_{t \in [0,T]}$ is constructed on an appropriate domain.

Regularity

We now prove that the random G-expectation of Nutz (2012) is a regular nonlinear expectation in our sense, provided that the coefficients μ and σ of the dynamics of the state process (2.31) are bounded. This boundedness ensures that the set $\mathbb{P}$ is relatively compact, which is crucial for the proof of the regularity result (see Lemma 2.49 below).

LEMMA 2.47. *If μ and σ are bounded, the family $\mathbb{P}$ is uniformly tight.* *Uniform tightness*

Proof. Let μ and σ be bounded by $K > 0$, and let $\nu \in \mathcal{U}^0$. For $X = X(0,x,\nu)$ (given by (2.31)) a simple calculation shows that

$$E^{P^0}\left[|X_s - X_t|^2\right] \leqslant K^2\left(2d|s-t|^2 + d^2 2^{d-1}|s-t|\right), \quad s,t \in [0,T].$$

Thus, the family $\{X(0,x,\nu) : \nu \in \mathcal{U}^t\}$ is uniformly tight on Ω^0 by the moment criterion for tightness on Wiener space; see, e.g., Corollary 16.9 in Kallenberg (2002), p. 313. Hence the family $\mathbb{P}$ is uniformly tight on Ω^x. □

To establish regularity, by Lemma 2.23, it suffices to show that $[0,T] \to L^p, t \to \mathcal{E}_t^{\mathrm{sub}}[\xi] = V_t^x(\xi)$ is càdlàg for all $\xi \in \mathcal{H}_T = UC_b(\Omega^x)$. Lemma 4.7 in Nutz (2012) implies that $V^x(\xi)$ is a $(P, \mathcal{F}^0)$-supermartingale for each $P \in \mathbb{P}$; thus, under each P, the event on which limits along monotone rational sequences exist has full measure. A careful examination also reveals this event to be a closed subset of Ω^x.

LEMMA 2.48. *Let $\xi \in UC_b(\Omega^x)$ and let A be the set of all $\omega \in \Omega^x$ such that*

$$\lim_{q \uparrow t, q \in \mathbb{Q}} V_q(\xi, \omega) \quad \textit{and} \quad \lim_{q \downarrow t, q \in \mathbb{Q}} V_q(\xi, \omega) \quad \textit{exist for all } t \in [0,T].$$

Then A is closed and its complement is $\mathbb{P}$-negligible.

Proof. Let $\omega \in A^c$ and $s_1 < s_2$ be an upcrossing of $V(\xi, \omega)$ through $[a,b]$. By Proposition 2.5 in Nutz (2012), there exists a modulus of continuity ρ_ξ with

$$|V_t(\xi, \omega^1) - V_t(\xi, \omega^2)| \leqslant \rho_\xi(\|\omega^1 - \omega^2\|_\infty) \quad \text{for all } \omega^1, \omega^2 \in \Omega^x.$$

Choosing $r > 0$ such that $\rho_\xi(r) < (b-a)/4$, we see that $s_1 < s_2$ is an upcrossing of $V(\xi, \bar{\omega})$ through $[a + \rho(r), b - \rho(r)]$, for every $\bar{\omega}$ with $\|\bar{\omega} - \omega\|_\infty < r$.

Since $V(\xi, \omega)$ is bounded and $\omega \in A^c$, it follows that $V(\xi, \omega)$ has infinitely many upcrossings through some non-empty interval and hence so does $V(\xi, \bar{\omega})$ for every $\bar{\omega}$ in the r-neighborhood of ω. This implies that A^c is open.

Finally, by Lemma 4.7 in Nutz (2012), the process $V(\xi)$ is a P-supermartingale for each $P \in \mathbb{P}$, and so $P(A^c) = 0$ for every $P \in \mathbb{P}$ by supermartingale regularity. □

LEMMA 2.49. *Suppose that μ and σ are bounded. For every monotone sequence $(t_n)_{n \in \mathbb{N}} \subset [0,T]$ and every $\xi \in UC_b(\Omega^x)$, the sequence $(V_{t_n}(\xi))_{n \in \mathbb{N}}$ is Cauchy in L^p, $1 \leqslant p < \infty$.*

Proof. It suffices to show that $(V_{t_n}(\xi))_{n \in \mathbb{N}}$ is Cauchy in L^p for every strictly monotone sequence $(t_n)_{n \in \mathbb{N}} \subset \mathbb{Q} \cap [0,T]$. Since $V(\xi)$ is bounded, for this, it is enough to prove that

$$\sup_{P \in \mathbb{P}} P(|V_{t_n}(\xi) - V_{t_m}(\xi)| \geqslant \eta) \to 0 \quad \text{for all } \eta > 0 \quad \text{as } m, n \to \infty,$$

see, e.g., Lemma A.31, p. 156.

Thus let $\varepsilon > 0$ and $\omega \in A$ (where A is given in Lemma 2.48). Then we find $N(\omega)$ such that $|V_{t_n}(\xi, \omega) - V_{t_m}(\xi, \omega)| < \eta/2$ for all $m, n \geqslant N(\omega)$. We choose $r > 0$ such that $\rho_\xi(r) < \eta/4$; for all $\bar{\omega} \in \Omega$ with $\|\bar{\omega} - \omega\|_\infty < r$ and $m, n \geqslant N(\omega)$, it follows that

$$|V_{t_n}(\xi, \bar{\omega}) - V_{t_m}(\xi, \bar{\omega})| < \eta.$$

By Lemma 2.47, there is some compact set K with $\sup_{P \in \mathbb{P}} P(K^c) < \varepsilon$. Clearly, the family of r-balls $\{B_r(\omega) : \omega \in A\}$ is an open covering of the compact set $A \cap K$, and hence there exist $\omega^1, \ldots, \omega^M \in A$ such that $A \cap K \subset \bigcup_{i=1}^M B_r(\omega^i)$. Setting $N \triangleq \max_{i=1,\ldots,M} N(\omega^i)$, we have

$$|V_{t_n}(\xi, \omega) - V_{t_m}(\xi, \omega)| < \eta \quad \text{for all } \omega \in A \cap K \text{ and } m, n \geqslant N.$$

Writing $A_{m,n} \triangleq \{|V_{t_n}(\xi) - V_{t_m}(\xi)| \geqslant \eta\}$, we thus have $K \cap A \cap A_{m,n} = \emptyset$ for all $m, n \geqslant N$, and it follows that

$$\sup_{P \in \mathbb{P}} P(A_{m,n}) \leqslant \sup_{P \in \mathbb{P}} P(K^c) + \sup_{P \in \mathbb{P}} P(A^c) < \varepsilon \quad \text{for all } m, n \geqslant N$$

since $\sup_{P \in \mathbb{P}} P(A^c) = 0$ by Lemma 2.48. □

THEOREM 2.50. *If μ and σ are bounded, then the random G-expectation $(\mathcal{E}_t^{\text{sub}})_{t \in [0,T]}$ constructed above is regular.*

Regularity of random G-expectations

Proof. In view of Lemma 2.23, it suffices to prove that the function $[0,T] \to L^p$, $t \to \mathcal{E}_t^{\text{sub}}[\xi] = V_t(\xi)$ is càdlàg for all $\xi \in UC_b(\Omega^x)$. Lemma 2.49 shows that it is làdlàg. But now, Theorem 5.1 in Nutz (2012) implies that $\lim_{q \downarrow t, q \in Q} V_q(\xi)$ and $V_t(\xi)$ coincide outside a $\mathbb{P}$-negligible set and hence in L_t^p. □

We have seen that the random G-expectation $(\mathcal{E}_t^{\text{sub}})_{t \in [0,T]}$ of Nutz (2012) is a regular sublinear expectation on an appropriate domain if the coefficients of the dynamics of the state process are bounded. Thus the full extent of our theory of backward nonlinear expectation equations (BNEEs) from Section 2.3 applies to this class of nonlinear expectations.

As an additional consequence of the fact that $\mathbb{P}$ is relatively compact, we now prove that the appropriate domain of $(\mathcal{E}_t^{\text{sub}})_{t \in [0,T]}$ consists solely of separable Banach spaces. This is very convenient, for then every measurable and adapted L^p-process X is μ-integrable (and thus in P^p) if and only if

$$\|X\|_{P,p} = \int \|X\|_{L,p} \, d\mu < \infty.$$

First, we state a refined version of Tietze's extension theorem due to Mandelkern (1990), which will be used to prove separability of L_t^p in Lemma 2.52 below.

LEMMA 2.51. *Let (X, d) be a metric space, $A \subset X$ a closed subspace and $f : A \to \mathbb{R}$ bounded and uniformly continuous. Then $g : X \to \mathbb{R}$, given by*

Tietze extension

$$g(x) \triangleq f(x) \quad \text{if } x \in A \quad \text{and} \quad g(x) \triangleq \inf_{a \in A} f(a) \frac{d(x,a)}{d(x,A)} \quad \text{if } x \notin A,$$

is uniformly continuous with $\sup_{x \in X} |g(x)| \leqslant \sup_{a \in A} |f(a)|$.

Proof. We refer to Mandelkern (1990). □

LEMMA 2.52. *All of the spaces L_t^p are separable.*

Proof. By Lemma 2.47, the family $\mathbb{P}$ is uniformly tight; hence for each $n \in \mathbb{N}$ there is some compact set $K_n \subset \Omega^x$ such that

$$\mathbb{P}(\Omega^x \setminus K_n) = \sup_{P \in \mathbb{P}} P(\Omega^x \setminus K_n) \leqslant 1/n.$$

For $t \in [0, T]$, $\Omega^x_t \subset \Omega^x$ is closed so that $K^t_n \triangleq \Omega^x_t \cap K_n \subset \Omega$ is compact. Thus the space $C(K^t_n)$ of continuous real functions on K^t_n is separable; see, e.g., Lemma 3.99 in Aliprantis and Border (2006), p. 125.

For each $n \in \mathbb{N}$, we pick a countable dense subset $F_n \subset C(K^t_n)$. By Lemma 2.51, every $f \in C(K^n_t)$ can be extended to some $f \in UC_b(\Omega^x_t)$ with the same norm, and we may view $F_n \subset UC_b(\Omega^x_t)$ as a subset. We set $F \triangleq \bigcup_{n \in \mathbb{N}} F_n \subset UC_b(\Omega^x_t) = \mathcal{H}_t$. It is straightforward to check that $F \subset L^p_t$ is dense. □

In the (non-random) G-expectation setting, BSDEs are intimately related to so-called second-order BSDEs (2BSDEs); see Cheridito et al. (2007); Soner et al. (2012, 2013) for background on 2BSDEs and, e.g., Hu et al. (2014a) for BSDEs under G-expectations and their relation to 2BSDEs. Thus we conclude our consideration of the Nutz (2012) random G-expectation with a brief remark on the relationship between BNEEs and second-order BSDEs.

Remark. To link 2BSDEs to the notion of BNEEs with respect to random G-expectations, let (f, ξ) be a BNEE^p-standard parameter, where f is induced path-by-path by a measurable Lipschitz function $f_0 : [0, T] \times \Omega \times \mathbb{R} \to \mathbb{R}$ via

$$[0, T] \times L^p \ni (t, \eta) \mapsto f_0(t, \cdot, \eta(\cdot)) \in L^p,$$

and let $X \in D^p$ be the unique solution of the BNEE

$$X_t = \mathcal{E}^{\text{sub}}_t\left[\int_t^T f(s, X_s) ds + \xi\right], \quad t \in [0, T], \tag{2.34}$$

where $(\mathcal{E}^{\text{sub}}_t)_{t \in [0,T]}$ denotes the random G-expectation constructed above. An application of Theorem 5.1 in Nutz (2012) shows that the process

$$M_t \triangleq \mathcal{E}^{\text{sub}}_t\left[\int_0^T f(s, X_s) ds + \xi\right], \quad t \in [0, T], \tag{2.35}$$

admits a $\mathbb{P}$-modification Y which is càdlàg and $\bar{\mathcal{F}}$-adapted; here, $\bar{\mathcal{F}}$ denotes the minimal right-continuous filtration containing $\mathcal{F}$, augmented by the collection of all $\mathbb{P}$-negligible sets. Moreover, Theorem 6.4 in Nutz (2012) shows that there exists an $\bar{\mathcal{F}}$-predictable process Z such that (Y, Z) is the solution of the 2BSDE

$$Y_t = \left[\xi + \int_0^T f(s, X_s) ds\right] - \int_t^T Z_s dM^{W,P}_s + K^P_T - K^P_t, \quad t \in [0, T], \text{ P-a.s.}$$

for each $P \in \mathbb{P}$.

Note: Here, $M^{W,P}$ is the local martingale part in the canonical semimartingale decomposition of $W^0(\cdot - x)$ under $P \in \mathbb{P}$ and $(K^P)_{P \in \mathbb{P}}$ is a family of increasing processes satisfying a minimality condition. For further details, we refer to Section 6 in Nutz (2012).

One can show that X admits a càdlàg and $\bar{\mathcal{F}}$-adapted modification $\bar{X}$ which satisfies

$$\bar{X}_t = Y_t - \oint_0^t f_0(s, \bar{X}_s)\, ds \quad P\text{-a.s. for all } P \in \mathbb{P} \text{ and } t \in [0,T],$$

and so $(\bar{X}, Z)$ solves the 2BSDE

$$\bar{X}_t = \xi + \oint_t^T f_0(s, \bar{X}_s)\, ds - \int_t^T Z_s dM_s^{W,P} + K_T^P - K_t^P, \qquad t \in [0,T],\ P\text{-a.s.} \tag{2.36}$$

Here, we write $\oint$ to emphasize that the pathwise Lebesgue integral for measurable stochastic processes is used, and not the Bochner intergral on P^p.

Note: One way to obtain such a modification $\bar{X}$ is as follows: For $n \in \mathbb{N}$ and $k \leqslant n$, put $t_k^n = \lceil T \rceil \frac{k}{n} \wedge T$ and set $X^n \triangleq X_0 1_{\{0\}} + \sum_{k=1}^n X_{t_k^n} 1_{(t_{k-1}^n, t_k^n]}$. Recalling that $[0,T] \to L^p$, $t \mapsto X_t$ is càdlàg since $X \in D^p$, it is obvious that $X^n \to X$ in P^p and $X_t^n \to X_t$ in L^p for all $t \in [0,T]$. Moreover, Fubini's theorem yields

$$\mathbb{P}\left[\oint_0^T |X_s^n - X_s^m|\, ds\right] \leqslant \int_0^T \mathbb{P}[|X_s^n - X_s^m|]\, ds \leqslant \|X^n - X^m\|_{\mathbb{P},p} \to 0 \quad \text{as } (n,m) \to \infty.$$

Hence $(X^n)_{n\in\mathbb{N}}$ is a Cauchy sequence in the complete space $\mathcal{P}^{1,1}(\mathbb{P}, dt; \bar{\mathcal{F}}_T \otimes \mathcal{B}([0,T]))$; see (A.4) on p. 150 for the definition of this space and Theorem A.26 for the proof of completeness. Thus, there exists an $\bar{\mathcal{F}}_T \otimes \mathcal{B}([0,T])$-measurable process $\tilde{X}$ such that $\mathbb{P}[\oint_0^T |X_s^n - \tilde{X}_s|\, ds] \to 0$. This entails in particular that

$$\oint_0^t f_0(s, X_s^n)\, ds \to \oint_0^t f_0(s, \tilde{X}_s)\, ds \quad \text{in } L^1 \quad \text{for all } t \in [0,T].$$

At the same time, we have

$$\oint_0^t f_0(s, X_s^n)\, ds = \int_0^t f(s, X_s^n)\, ds \to \int_0^t f(s, X_s)\, ds \quad \text{in } L^p \quad \text{for all } t \in [0,T]$$

since $X^n \to X$ in P^p. Setting $\bar{X} \triangleq Y - \oint_0^t f_0(s, \tilde{X}_s)\, ds$, where Y is the càdlàg, $\bar{\mathcal{F}}$-adapted $\mathbb{P}$-modification of the process M from (2.35), it thus follows that

$$X = M_t - \int_0^t f(s, X_s) ds = Y_t - \oint_0^t f_0(s, \tilde{X}_s)\, ds = \bar{X}_t \quad \text{in } L^p \quad \text{for all } t \in [0,T],$$

and we have constructed a càdlàg, $\bar{\mathcal{F}}$-adapted $\mathbb{P}$-modification $\bar{X}$ of X.

The above discussion shows that the unique solution X of the BNEE (2.34) induces a solution of the 2BSDE (2.36), which is unique in the sense of Theorem 6.4 in Nutz (2012). Finally, note that the 2BSDE (2.36) is not included in the class of 2BSDEs studied by Soner et al. (2012, 2013), as the domain of the conjugate of the nonlinear generator is possibly path-dependent. ◇

2.5.2 *Classical G-expectations*

As pointed out by Nutz (2012), the classical G-expectation of Peng (2008, 2010) can be considered as the special case of the random G-expectation presented above where $\mu = 0$, $\sigma(r, X, \nu_r) = \nu_r$ and the SDE for the state process (2.31) is just a stochastic integral; the stochastic control representation of the G-expectation which implies this is studied in great detail by Denis et al. (2011). Thus, the results from the previous Subsection 2.5.1 directly show that Peng's G-expectation is a regular sublinear expectation on an appropriate domain. In this subsection, we present an alternative, more standard construction of G-expectations and a more elementary proof of their regularity.

To introduce the G-expectation, we follow Peng (2008, 2010) and Denis et al. (2011) and work on the canonical Wiener space

$$\Omega_t \triangleq \{\omega : [0,t] \to \mathbb{R}^d \, : \, \omega \text{ is continuous with } \omega(0) = 0\}, \quad t \in [0,T].$$

As usual, Ω_t is equipped with the topology of uniform convergence. The canonical process is denoted by $B_s(\omega) = \omega_s$; the d-dimensional Wiener measure is denoted by P. We write $C_{lLip}(\mathbb{R}^{d\times n})$ for the space of all functions $\varphi : \mathbb{R}^{d\times n} \to \mathbb{R}$ such that

$$\exists\, k, C > 0 \, : \, |\varphi(x) - \varphi(y)| \leqslant C(1 + |x|^k + |y|^k)|x-y| \quad \forall x, y \in \mathbb{R}^{d\times n}.$$

The G-expectation is defined first for members of the increasing family $(\mathcal{H}_t)_{t\in[0,T]}$ of cylindrical random variables

$$\mathcal{H}_t \triangleq \left\{\varphi(B_{t_1}, \dots, B_{t_n}) \, : \, n \in \mathbb{N}, \; t_1, \dots, t_n \in [0,t], \; \varphi \in C_{lLip}(\mathbb{R}^{d\times n})\right\}.$$

For this, a function $G : \mathbb{S}^d \to \mathbb{R}$ of the form

$$G(A) = \frac{1}{2} \sup_{\gamma\in\Gamma} \operatorname{tr}[\gamma\gamma^\top A], \qquad A \in \mathbb{S}^d,$$

is chosen, where $\Gamma \subseteq \mathbb{R}^{d\times d}$ is bounded, non-empty and closed, and $\mathbb{S}^d$ denotes the set of symmetric $d \times d$ matrices.

The sublinear expectation $(\mathcal{E}^G_t)_{t\in[0,T]}$ – the G-expectation – is then defined on $(\mathcal{H}_t)_{t\in[0,T]}$ by the condition that, for each $\varphi \in C_{lLip}(\mathbb{R}^{d\times n})$ and all $0 \leqslant t_1 \leqslant \cdots \leqslant t_n \leqslant T$, we have

$$\mathcal{E}^G_{t_{n-1}}\left[\varphi(B_{t_1}, B_{t_2}, \dots, B_{t_n} - B_{t_{n-1}})\right] = \psi(B_{t_1}, \dots, B_{t_{n-1}}), \tag{2.37}$$

where $\quad \psi(x_1, \dots, x_{n-1}) \triangleq \mathcal{E}^G[\varphi(x_1, x_2, \dots, \sqrt{t_n - t_{n-1}} B_1)].$

Here, B_1 is G-normal, that is, for each $\varphi \in C_{lLip}(\mathbb{R}^d)$ one defines

$$\mathcal{E}^G[\varphi(x + \sqrt{t} B_1)] \triangleq u(t,x),$$

where u is the unique continuous viscosity solution of the so-called G-heat equation (see Peng (2008))

$$u_t - G(u'') = 0, \quad u(0,\cdot) = \varphi. \tag{2.38}$$

It is straightforward to check that the definition of the G-expectation on $\mathcal{H}$ is independent of the choice of representatives, and thus it is well-defined as an operator $\mathcal{E}^G_t : \mathcal{H} \to \mathcal{H}_t$. From there, time-consistency (TC) is a consequence of the recursive definition (2.37).

Theorem 4.3.1 and Theorem 4.4.5 in Pham (2009) characterize the unique continuous viscosity solution u of (2.38) as the value function

$$u(T-t,x) = v(t,x) \triangleq \sup_{\alpha\in\mathcal{A}} E^{t,x}\left[\varphi(B_T^\alpha)\right], \quad t \in [0,T],\ x \in \mathbb{R}^d, \qquad (2.39)$$

where $\mathcal{A}$ is the set of all Γ-valued progressive processes and $dB_t^\alpha = \alpha_t dB_t$; see also Theorem 48 in Denis et al. (2011). Here, the connection to the random G-expectation from Section 2.5.1 and, in particular, to (2.31) and (2.33) becomes apparent.

Note: The optimal control characterization (2.39) is obtained as follows: Standard SDE estimates show that the value function $\sup_{\alpha\in\mathcal{A}} E^{t,x}[\varphi(B_T^\alpha)]$ is continuous and locally bounded; thus, Theorem 4.3.1 in Pham (2009) implies that v is a continuous viscosity solution of $-v_t - G(v'') = 0$, $v(T,\cdot) = \varphi$. A comparison principle (see, e.g., Theorem 4.4.5 in Pham (2009)) shows that there is at most one such solution. We note that the definition of a viscosity solution used in Pham (2009) is different from the one used in Peng (2008) and Denis et al. (2011). In Denis et al. (2011), the function v is a viscosity solution of $v_t + G(v'') = 0$; see their Theorem 48. This is equivalent to v being a viscosity solution of $-v_t - G(v'') = 0$ in the sense of Pham (2009).

The representation (2.39) directly implies that the G-expectation is monotone, sublinear, and preserves constants; therefore, $(\mathcal{E}_t^G)_{t\in[0,T]}$ is a sublinear expectation on $(\mathcal{H}_t)_{t\in[0,T]}$. To apply the results of Section 2.4, consider $(\mathcal{E}_t^G)_{t\in[0,T]}$ as a sublinear expectation on $(\mathcal{H}_t^b)_{t\in[0,T]}$, where

$$\mathcal{H}_t^b \triangleq \left\{\varphi(B_{t_1},\dots,B_{t_n}) : n \in \mathbb{N},\ t_1,\dots t_n \in [0,t],\ \varphi \in C_{bLip}(\mathbb{R}^{d\times n})\right\},$$

and where $C_{bLip}(\mathbb{R}^{d\times n})$ is the space of all bounded and uniformly Lipschitz continuous functions $\varphi : \mathbb{R}^{d\times n} \to \mathbb{R}$. This is possible by (2.39), which even yields the probabilistic representation

$$\mathcal{E}_0^G[\xi] = \sup_{P\in\mathbb{P}} E^P[\xi], \quad \xi \in \mathcal{H}_T, \quad \text{where } \mathbb{P} \triangleq \{P \circ (B^\alpha)^{-1} : \alpha \in \mathcal{A}\}. \qquad (2.40)$$

Once again, we are precisely in the situation of Section 2.4. As in Subsection 2.5.1 above, we obtain the following:

- The collection $L^p(\mathbb{P})$ of (equivalence classes of) $\mathcal{F}_T$-measurable random variables is a Banach space for the norm $\|\xi\|_{L,p} = \sup_{P\in\mathbb{P}} E^P[|\xi|^p]^{\frac{1}{p}}$ by Corollary 2.42.
- The closures $L_t^p \triangleq \text{clos}(\mathcal{H}_t^b) \subset L^p(\mathbb{P}^p)$ form a Lebesgue family by Theorem 2.44.
- The sublinear expectation $(\mathcal{E}_t^G)_{t\in[0,T]}$ on $(\mathcal{H}_t^b)_{t\in[0,T]}$ extends continuously to a sublinear expectation $(\bar{\mathcal{E}}_t^G)_{t\in[0,T]}$ on the *appropriate domain* $\{L_t^p : t \in [0,T],\ p \geqslant 1\}$; see Theorem 2.45.

Moreover, Lemma 2.47 implies that the family $\mathbb{P}$ is uniformly tight. As a consequence, we get the following lemma; see also Hu and Peng (2009) and Denis et al. (2011).

LEMMA 2.53. *For all* $t \in [0,T]$ *and* $p \geqslant 1$, *we have* $L^p_t = \mathrm{clos}(\mathcal{H}^b_t; L^p(\mathbb{P})) = \mathrm{clos}(\mathcal{H}_t; L^p(\mathbb{P}))$ *and* $\bar{\mathcal{E}}^G_t|_{\mathcal{H}_T} = \mathcal{E}^G_t$; *moreover* L^p_t *is separable.*

Proof. We note that (2.39) and standard SDE techniques show that $v^n(t,x) \triangleq \sup_{\alpha\in\mathcal{A}} E^{t,x}[\varphi^n(B^\alpha_T)] \to v(t,x) = \sup_{\alpha\in\mathcal{A}} E^{t,x}[\varphi(B^\alpha_T)]$ uniformly on compacts whenever $\varphi^n \to \varphi$ uniformly on compacts. By the definition (2.37) of the G-expectation and uniform tightness of the representing set $\mathbb{P}$ (2.40), we obtain the first two assertions. For the last, note that for every compact set $K \subset \Omega$, the set of restricted functions $\{\xi|_K : \xi \in \mathcal{H}^b_t\} \subset C(K)$ is dense by the Stone-Weierstrass Theorem (see, e.g., Aliprantis and Burkinshaw (1998), Theorem 11.5, p. 89), and hence the claim follows as in Lemma 2.52 above. □

In view of Lemma 2.53, we shall denote the continuous extension $(\bar{\mathcal{E}}^G_t)_{t\in[0,T]}$ of $(\mathcal{E}^G_t)_{t\in[0,T]}$ by $(\mathcal{E}^G_t)_{t\in[0,T]}$, as well. We now give an elementary argument which shows that the G-expectation $(\mathcal{E}^G_t)_{t\in[0,T]}$ is a regular sublinear expectation in the sense of Definition 2.22. In fact, $t \mapsto \mathcal{E}^G_t[\xi]$ is even continuous.

Regularity

LEMMA 2.54. *Let* $\xi = \varphi(B_{t_1}, \ldots B_{t_n}) \in \mathcal{H}^b_T$. *Then there exists a constant* $K > 0$ *such that*

$$\left\|\mathcal{E}^G_s[\xi] - \mathcal{E}^G_t[\xi]\right\|_{L,p} \leqslant K|s-t|^{\frac{1}{2}} \quad \textit{for all } s,t \in [0,T].$$

Proof. Let $L > 0$ denote the Lipschitz constant of φ and let $0 \leqslant t < s \leqslant T$. Without loss of generality,[2] we may assume that $t_i = s < t = t_j$ for $1 \leqslant i < j \leqslant n$. Iterating (2.37), we see that $\mathcal{E}^G_{t_k}[\xi] = \psi^k(B_{t_1}, \ldots, B_{t_k})$, where $\psi^n \triangleq \varphi$ and

$$\psi^k(x_1, \ldots, x_k) \triangleq \mathcal{E}^G\left[\psi^{k+1}(x_1, \ldots, x_k, \sqrt{t_{k+1}-t_k}B_1 + x_k)\right]$$

for $k = n-1, \ldots, 1$, and where each ψ^k has the same Lipschitz constant L as φ. Therefore, we obtain

$$\left|\mathcal{E}^G_{t_k}[\xi] - \mathcal{E}^G_{t_{k-1}}[\xi]\right| \leqslant L|B_{t_k} - B_{t_{k-1}}| + L\sqrt{t_k - t_{k-1}}\,\mathcal{E}^G[|B_1|].$$

By a telescoping sum argument and Jensen's inequality (Lemma 2.11), we thus arrive at

$$\left|\mathcal{E}^G_s[\xi] - \mathcal{E}^G_t[\xi]\right|^p \leqslant 2^p L^p k^p \left(\sum_{\ell=1}^k |B_{t_{i+\ell}} - B_{t_{i+\ell-1}}|^p + |s-t|^{\frac{p}{2}}\mathcal{E}^G[|B_1|^p]\right).$$

Now, $|B_{u+\Delta} - B_u|^p \in \mathcal{H}_T$, and hence we have $\mathcal{E}^G[|B_{u+\Delta} - B_u|^p] = \Delta^{\frac{p}{2}}\mathcal{E}^G[|B_1|^p]$ by (2.37). Since $|B_1|^p \in \mathcal{H}_T$, it has finite sublinear expectation $c \triangleq \mathcal{E}^G[|B_1|^p]$, and the proof is complete. □

COROLLARY 2.55. *The* G*-expectation is a regular nonlinear expectation. In particular, so is its superlinear counterpart* $-\mathcal{E}^G[-\cdot]$.

Proof. By Lemma 2.54, the map $t \mapsto \mathcal{E}^G_t[\xi]$ is uniformly continuous for every $\xi \in \mathcal{H}^b_T$. We conclude by Lemma 2.23. □

2 Every $\xi \in \mathcal{H}$ admits a representation $\varphi^\xi_{\min}$ with a minimal number of time points. An arbitrary representation φ and $\varphi^\xi_{\min}$ share the same Lipschitz constant.

2.5.3 *The g-expectation*

In this subsection, we demonstrate how g-expectations can be subsumed into our general nonlinear expectations framework; g-expectations are defined in terms of backward stochastic differential equations and were first put forward by Peng (1997). Since then, g-expectations have been studied extensively in the literature; see, among others, Briand et al. (2000), Chen and Peng (2000), Coquet et al. (2002), Chen and Epstein (2002) as well as Peng (2004b) and the references therein. Royer (2006) studies g-expectations on probability spaces which, in addition to a Brownian motion, also carry a Poisson random measure; see also Delong (2013) for an overview.

Preliminaries

Let $(\Omega, \mathcal{A}, P)$ be a complete probability space that supports both a d-dimensional Brownian motion W and a Poisson random measure

$$\Gamma : \Omega \times \mathcal{B}([0,T] \times E) \to \mathbb{N}_0 \cup \{\infty\}.$$

Here $E \triangleq \mathbb{R}^\ell \setminus \{0\}$, and $\mathcal{B}([0,T] \times E)$ denotes the Borel-σ-algebra on $[0,T] \times E$. We assume that the compensator Λ of Γ takes the form $\Lambda(dt, de) = \gamma(de)dt$, where γ is a σ-finite measure on $(E, \mathcal{B}(E))$ such that $\int_E (1 \wedge |e|^2)\gamma(de) < \infty$. Finally, we denote by $\tilde{\Gamma} \triangleq \Gamma - \Lambda$ the compensated jump measure and by $(\mathcal{F}_t)_{t \in [0,T]}$ the augmented filtration generated by W and Γ; we assume that $\mathcal{F}_T = \mathcal{A}$. Since we are now working in an honestly probabilistic framework, we adopt a usual convention and identify two random variables if they coincide almost surely. Moreover, we will not distinguish between stochastic processes which coincide $P \otimes dt$-almost everywhere.

BSDE

We consider BSDEs of the form

$$X_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Y_s^\top dW_s - \int_t^T \int_E Z_s(e)\tilde{\Gamma}(ds, de), \quad (2.41)$$

where we assume that

$$f : \Omega \times [0,T] \times \mathbb{R}^d \times L^2(\gamma) \to \mathbb{R}$$

satisfies the following conditions:

(D1) The process $f(\cdot, Y, Z)$ is predictable for every $\mathbb{R}^d \times L^2(\gamma)$-valued predictable process (Y, Z).

(D2) For all $y_1, y_2 \in \mathbb{R}^d$, $z_1, z_2 \in L^2(\gamma)$ and some $L > 0$, we have

$$|f(\,\cdot\,, y_1, z_1) - f(\,\cdot\,, y_2, z_2)| \leqslant L\big(|y_1 - y_2| + \|z_1 - z_2\|_{L^2(\gamma)}\big).$$

(D3) We have $E\big[\int_0^T |f(t,0,0)|^2 dt\big] < \infty$.

Let us refer to such a function f as a *driver*. The next two lemmas provide well-known existence, uniqueness and stability results for the corresponding BSDEs.

Existence, uniqueness

LEMMA 2.56. *Let f be a driver and $\xi \in L^2(P)$. Then the BSDE (2.41) has a unique solution (X, Y, Z) satisfying*

$$E\left[\sup_{t\in[0,T]}|X_t|^2 + \int_0^T|Y_t|^2dt + \int_0^T\|Z_t(e)\|^2_{L^2(\gamma)}dt\right] < \infty. \tag{2.42}$$

Proof. See Lemma 2.4 in Tang and Li (1994). An inspection of their proof reveals that f need not be measurable: It suffices if $f(\cdot, Y, Z)$ admits a predictable a.e.-modification for all $\mathbb{R}^d \times L^2(\gamma)$-valued predictable processes (Y, Z). □

Stability

LEMMA 2.57. *Let f be a driver, $\xi^1, \xi^2 \in L^2(P)$ and let (X^i, Y^i, Z^i) denote the unique solutions of the BSDE (2.41) with terminal condition $\xi = \xi^i$, $i = 1, 2$. Then there exists a constant $c > 0$ (depending only on f and T) such that*

$$E\left[\sup_{t\in[0,T]}|X_t^1 - X_t^2|^2 + \int_0^T|Y_t^1 - Y_t^2|^2dt + \int_0^T\|Z^1 - Z^2\|^2_{L^2(\gamma)}dt\right] \leqslant c\,E\left[|\xi^1 - \xi^2|^2\right].$$

Proof. See Proposition 2.2. in Barles et al. (1997). □

Construction of g-expectation

To construct a g-expectation, let g and h be two drivers, where h is sublinear, and suppose that h dominates g, i.e., for all $y_1, y_2 \in \mathbb{R}^d$ and $z_1, z_2 \in L^2(\gamma)$, we have

$$g(\cdot, y_1, z_1) - g(\cdot, y_2, z_2) \leqslant h(\cdot, y_1 - y_2, z_1 - z_2). \tag{2.43}$$

Moreover, suppose that the following two conditions are met:

(g1) For both $f = g$ and $f = h$, the BSDE (2.41) satisfies the following comparison principle: Let $\xi \in L^2(P)$, let (X, Y, Z) denote the unique solution of (2.41), and suppose that $(\overline{X}, \overline{Y}, \overline{Z})$ satisfies the integrability condition in (2.42). If $\xi \geqslant \eta = \overline{X}_T \in L^2(P)$ and

$$d\overline{X}_t = -\left[f(t, \overline{Y}_t, \overline{Z}_t) + \beta_t\right]dt + \overline{Y}_t^\top dW_t + \int_E \overline{Z}_t(e)\tilde{\Gamma}(dt, de)$$

with $\beta \leqslant 0$, then $\overline{X} \leqslant X$. Similarly, if $\xi \leqslant \eta$ and $\beta \geqslant 0$, then $\overline{X} \geqslant X$.

(g2) The drivers g and h are normalized, i.e., $g(\cdot, 0, 0) = 0 = h(\cdot, 0, 0)$.

In the Brownian setting, assumption (D2) implies a comparison theorem for BSDEs which guarantees (g1); see, e.g., El Karoui et al. (1997). In jump-diffusion settings, comparison principles for BSDEs are a more complicated matter. A comparison theorem was first obtained by Barles et al. (1997) and later improved by Royer (2006) and Quenez and Sulem (2013). For sufficient conditions for (g1), we refer to these papers.

Note: In a purely Brownian setting without jumps, one may take any suitably measurable, normalized function g which satisfies the Lipschitz condition $|g(\cdot, y) - g(\cdot, \bar{y})| \leqslant L|y - \bar{y}|$ and put $h(\cdot, y) \triangleq L|y|$. Then the conditions (g1)-(g2) are satisfied; see, e.g, Peng (2004b).

Under the assumptions (g1)-(g2), the operators

$$\mathcal{E}_t^g : L^2(P) \to L^2(\Omega, \mathcal{F}_t, P), \quad \mathcal{E}_t^g[\xi] \triangleq X_t, \qquad t \in [0, T],$$

where X is given as the unique solution of (2.41) with $f = g$, satisfy the axioms (SI), (TC) and (N) as direct consequences of uniqueness (Lemma 2.56) and normalization (g2). Monotonicity (M) follows from the comparison principle (g1). Additionally, if g is positively homogeneous or subadditive, then, by (g1), so is $\mathcal{E}_t^g$; for details, we refer to Royer (2006). The nonlinear expectation $(\mathcal{E}_t^g)_{t\in[0,T]}$ is called a g-expectation. The same procedure (taking $f = h$ in (2.41)) yields a sublinear expectation $(\mathcal{E}_t^h)_{t\in[0,T]}$. Since h dominates g (2.43), the comparison principle (g1) implies that

$$\mathcal{E}_t^g[\xi] - \mathcal{E}_t^g[\eta] \leqslant \mathcal{E}_t^h[\xi - \eta], \quad t \in [0, T], \ \xi, \eta \in L^2(P),$$

that is, $(\mathcal{E}_t^h)_{t\in[0,T]}$ dominates $(\mathcal{E}_t^g)_{t\in[0,T]}$.

In the following, we construct an appropriate domain for the g-expectation. Since $\mathcal{E}_0^h : L^2(P) \to \mathbb{R}$ is a sublinear expectation operator, Theorem 2.37 yields a family $\mathbb{P}$ of finitely additive probability measures on $\mathcal{A}$ such that

$$\mathcal{E}_0^h[\xi] = \mathbb{P}[\xi] = \sup_{p\in\mathbb{P}} E^p[\xi] \qquad \text{for all } \xi \in \mathcal{L}^\infty(\mathcal{A}). \qquad (2.44)$$

The stability result in Lemma 2.57 shows that $\mathcal{E}_0^h[\xi^n] \to \mathcal{E}_0^h[\xi]$ if $\xi^n \to \xi$ in $L^2(P)$. In particular,

$$\mathbb{P}(A_n) = \sup_{p\in\mathbb{P}} E^p[1_{A_n}] = \mathcal{E}_0^h[1_{A_n}] \to 0 \quad \text{whenever } A_n \downarrow \emptyset.$$

Hence, each $p \in \mathbb{P}$ is σ-continuous from above and thus countably additive. As it consists solely of genuine probabilities, the family $\mathbb{P}$ has the Fatou property. Moreover, $\mathbb{P}(A) = \mathcal{E}_0^h[1_A] = 0$ if $P(A) = 0$, i.e., P is absolutely continuous with respect to each $p \in \mathbb{P}$.

As in Subsection 2.4, we introduce the Banach space

$$L^p(\mathbb{P}) \triangleq \left\{ [\xi]_{\mathbb{P}} \, : \, \xi \in \mathcal{L}^0(\mathcal{A}) \text{ with } \|\xi\|_{L,p} = \mathbb{P}\left[|\xi|^p\right]^{\frac{1}{p}} < \infty \right\} \leqslant \Omega^{\mathbb{R}} / \mathcal{N}_{\mathbb{P}}.$$

If $(\xi^n)_{n\in\mathbb{N}} \subset \mathcal{L}^\infty(\mathcal{A})$ with $\xi^n \to \xi$ in $L^{2p}(P)$, then the stability result in Lemma 2.57 and the probabilistic representation (2.44) imply

$$\mathbb{P}\left[|\xi^n - \xi^m|^p\right] = \mathcal{E}_0^h\left[|\xi^n - \xi^m|^p\right] \to 0 \quad \text{as } (n, m) \to \infty.$$

It follows that $\xi^n \to \xi$ in $L^p(\mathbb{P})$. In particular,

$$L^{2p}(P) \hookrightarrow L^p(\mathbb{P}) \quad \text{for all } p \geqslant 1 \text{ and } \quad \mathbb{P}[\xi] = \mathcal{E}_0^h[\xi] \quad \text{for all } \xi \in L^2(P).$$

Hence, defining

$$L_t^p \quad \text{as the closure of} \quad L^{2p}(\Omega, \mathcal{F}_t, P) \quad \text{in } L^p(\mathbb{P}), \quad t \in [0, T], \ p \geqslant 1,$$

we obtain a Lebesgue family $\mathcal{L} = \{(L_t^p, \|\cdot\|_{L,p}) \, : \, t \in [0, T], \ p \geqslant 1\}$.

Since $\mathcal{E}_0^h = \mathbb{P}[\cdot]|_{L^2(P)}$, time consistency (TC) implies

$$\mathbb{P}\left[|\mathcal{E}_t^g[\xi^1] - |\mathcal{E}_t^g[\xi^2]|\right] \leqslant \mathbb{P}\left[\mathcal{E}_t^h\left[|\xi^1 - \xi^2|\right]\right] = \mathbb{P}\left[|\xi^1 - \xi^2|\right].$$

Thus, both $\mathcal{E}_t^h$ and $\mathcal{E}_t^g$ extend continuously (and hence uniquely) to operators $\mathcal{E}_t^h, \mathcal{E}_t^g : L^1 \to L_t^1$, $t \in [0, T]$. Now, arguing as in the proof of Theorem 2.45, it is straightforward to verify that $(\mathcal{E}_t^g)_{t\in[0,T]}$ is a nonlinear expectation, carried by the sublinear expectation $(\mathcal{E}_t^h)_{t\in[0,T]}$ on the appropriate domain $\mathcal{L}$.

g-expectations

THEOREM 2.58. *The* g*-expectation* $(\mathcal{E}_t^g)_{t\in[0,T]}$ *is a regular nonlinear expectation carried by* $(\mathcal{E}_t^h)_{t\in[0,T]}$ *on* $\mathcal{L}$*.*

Proof. It remains to establish regularity. By Lemma 2.23, it suffices to check that $[0, T] \to L^p$, $t \mapsto \mathcal{E}_t[\xi]$ is càdlàg for every $\xi \in L^{2p}(P)$ and each $p \geqslant 1$. Thus let $p \geqslant 1$ and $\xi \in L^{2p}(P)$, and let (X, Y, Z) be the unique solution of (2.41) with $f = g$. Then $\mathcal{E}_t^g[\xi] = X_t$, and $E[\sup_{t\in[0,T]} |X_t|^{2p}] < \infty$.

If $t_n \uparrow t$ or $t_n \downarrow t$, then $X_{t_n} \to \eta$ a.s., where $\eta = X_t$ or $\eta = X_{t-}$, resp., since X has a càdlàg paths. But then $\eta^n \triangleq |X_{t_n} - \eta|^p \to 0$ in $L^2(P)$ by dominated convergence. It remains to show that $\mathbb{P}[\eta^n] \to 0$.

Let (X^n, Y^n, Z^n) be the unique solution of (2.41) with $f = h$ and terminal condition η^n. Then $X^n \to 0$ by Lemma 2.57, and hence $\mathbb{P}[\eta^n] = \mathcal{E}^h[\eta^n] = X_0^n \to 0$. □

Remark. Theorem 2.58 shows that it is possible to consider BNEEs under g-expectations; however, this does not result in new objects: Indeed, let (X, Y, Z) be the unique solution of the BSDE

$$X_t = -\left[f(t, X_t) + g(t, Y_t, Z_t)\right]dt + Y_t^\top dW_t + \int_E Z_t(e)\tilde{\Gamma}(dt, de) \qquad (2.45)$$

with terminal condition $X_T = \xi$. Setting $\tilde{f}(t, x) \triangleq f(t, x) + g(t, Y_t, Z_t)$, the semimartingale X is the unique solution of the BSDE associated with $(\tilde{f}, \xi)$; hence, X (considered as a mapping $t \to L^p$) is in D^p by Lemma 2.70 below.

Now, let (M, N, O) be the unique solution of the BSDE

$$M_t = -g(t, N_t, O_t)dt + N_t^\top dW_t + \int_E O_t(e)\tilde{\Gamma}(dt, de)$$

with terminal condition $M_T = \xi + \int_0^T f(s, X_s)ds$. By uniqueness, it follows that $M_t - \int_0^t f(s, X_s)ds = X_t$, and hence

$$X_t = \mathcal{E}_t^g\left[\int_t^T f(s, X_s)ds + \xi\right], \qquad t \in [0, T]. \qquad (2.46)$$

Thus the unique solution of the BNEE (2.46) (as in Theorem 2.26) is given by the solution X of the classical BSDE (2.45): BNEEs under g-expectations are BSDEs. Nevertheless, the discretization result of Theorem 2.32 or the convergence result for recursive utilities (Theorem 2.80 below) are of interest in the context of g-expectations. ⋄

2.5.4 *Drift and intensity uncertainty*

Robust expectations under drift and intensity uncertainty are an interesting special case of g-expectations, and hence they fit into the abstract nonlinear expectation framework of this thesis.

The construction of dynamic robust expectations under uncertainty about the drift is due to Chen and Epstein (2002). In the following, we briefly sketch out a natural generalization of their construction which also includes uncertainty about the intensity of jumps.

Let $(\Omega, \mathcal{A}, P)$ be a complete probability space which carries both an n-dimensional Brownian motion W and an m-dimensional standard Poisson process N. The associated compensated Poisson process is denoted by $\bar{N}$. Let $(\mathcal{F}_t)_{t\in[0,T]}$ be the augmented filtration generated by W and N and assume that $\mathcal{A} = \mathcal{F}_T$.

We fix a set $\mathcal{D}$ of predictable $E \triangleq \mathbb{R}^n \times (-1, \infty)^m$-valued processes $\theta = (\alpha, \beta)$ for which the stochastic exponential

$$dZ_t^\theta = Z_{t-}^\theta(\alpha_t^\top dW_t + \beta_t^\top d\bar{N}_t), \quad Z_0^\theta = 1,$$

is a martingale. Then each $\theta = (\alpha, \beta) \in \mathcal{D}$ gives rise to an equivalent probability P^θ with P-density Z^θ. Moreover, under P^θ the process

$$W^\theta \triangleq W - \textstyle\int_0^\cdot \alpha_s ds \quad \text{is a standard } n\text{-dimensional Brownian motion}$$

and N is a counting process with intensity $\mathbf{1}_m + \beta$, i.e., the process

$$\bar{N}^\theta \triangleq N - \textstyle\int_0^\cdot (\mathbf{1}_m + \beta_s) ds \quad \text{is a local martingale.}$$

Here $\mathbf{1}_m$ denotes $(1, 1, \ldots, 1)^\top \in \mathbb{R}^m$.

We write $E^\theta[\cdot] \triangleq E[\cdot Z_T^\theta]$ for the expectation with respect to P^θ and denote by $\mathbb{P}$ the collection of all P^θ, $\theta \in \mathcal{D}$. Following Chen and Epstein (2002), the set $\mathbb{P}$ is said to be *rectangular* if there exists a 2^E-valued process Θ such that

$$\theta \in \mathcal{D} \quad \Longleftrightarrow \quad \begin{array}{l} \theta \text{ is a predictable process such that} \\ \theta_t(\omega) \in \Theta_t(\omega) \quad \text{for } dt \otimes P\text{-a.e. } (t, \omega). \end{array} \tag{R}$$

We note that this property implies that $\mathcal{D}$ is stable under pasting, i.e., if $\theta^1, \theta^2 \in \mathcal{D}$ and $t \in [0, T]$, then $\theta \triangleq 1_{[0,t)}\theta^1 + 1_{[t,T]}\theta^2 \in \mathcal{D}$. The process Θ is further supposed to satisfy the following conditions:

(a) *Weak measurability:* For every open set $G \subseteq \mathbb{R}^n \times \mathbb{R}^m$ the lower inverse $\Theta^\ell(G) = \{(t, \omega) : \Theta_t(\omega) \cap G \neq \emptyset\}$ is a predictable set.

(b) *Boundedness:* There exists a compact set $K \subseteq \mathbb{R}^{n+m}$ such that $\Theta_t(\omega) \subseteq K$ for all $(t, \omega) \in [0, T] \times \Omega$.

(c) *Closedness:* For all $(t, \omega) \in [0, T] \times \Omega$, the set $\Theta_t(\omega) \subseteq E$ is closed.

(d) *Normalization:* For all $(t, \omega) \in [0, T] \times \Omega$, we have $0 \in \Theta_t(\omega)$.

We note that the boundedness condition (b) ensures that

$$\{Z^\theta_T : \theta \in \mathcal{D}\} \quad \text{is uniformly bounded in } L^p(P) \text{ for all } p \geqslant 1. \tag{2.47}$$

More importantly, the above conditions yield the following measurable selection result:

Predictable maximizer

LEMMA 2.59. *Let ρ be an $\mathbb{R}^{n\times m}$-valued predictable process. Then there exists $\bar\theta \in \mathcal{D}$ with*

$$\rho^\top \bar\theta = \max_{\theta\in\mathcal{D}} \rho^\top \theta \qquad P\otimes dt\text{-}a.e. \tag{2.48}$$

Proof. By the above assumptions, Θ is a weakly measurable correspondence with non-empty compact values, from the measurable space $\Omega \times [0,T]$ (endowed with the predictable σ-algebra) into the (separable and metrizable) space E. Since $((\omega,t),x) \mapsto \rho_t(\omega)^\top x$ is predictable in (ω,t) and continuous in x, the argmax-correspondence $(\omega,t) \mapsto \{y \in \Theta_t(\omega) : \rho_t(\omega)^\top y = \max_{x\in\Theta_t(\omega)} \rho_t(\omega)^\top x\}$ admits a predictable selector $\bar\theta$ by Theorem 18.19 in Aliprantis and Border (2006). Now, the rectangularity condition (R) implies $\theta \in \mathcal{D}$ and (2.48). □

Being able to select the maximizer is a convenient feature of the present situation that simplifies the construction of a *time-consistent* nonlinear expectation in the present context. The reason for this is that all involved measures are mutually absolutely continuous; thus a process that is well-defined for a single measure is automatically well-defined for all relevant measures: There is no need to aggregate a whole *family of objects* (being defined only under one measure) into one meaningful (i.e., measurable) *single object* (being defined for all measures) (as, e.g., in Soner et al. (2011b)). The following simple martingale representation result is another manifestation of the comforts of absolute continuity; it is the second important ingredient in the construction of the robust expectation.

Martingale representation

LEMMA 2.60. *Let $\xi \in L^2(P)$. For each $\theta = (\alpha,\beta) \in \mathcal{D}$, there exists a predictable $\mathbb{R}^n \times \mathbb{R}^m$-valued process (H,K) such that*

$$\begin{aligned} E^\theta_t[\xi] &= E^\theta[\xi] - \textstyle\int_0^t (H_s^\top \alpha_s + K_s^\top \beta_s)ds + \int_0^t H_s^\top dW_s + \int_0^t K_s^\top d\bar N_s \\ &= E^\theta[\xi] + \textstyle\int_0^t H_s^\top dW^\theta_s + \int_0^t K_s^\top d\bar N^\theta_s \quad \textit{for all } t\in[0,T] \; a.s. \end{aligned} \tag{2.49}$$

Proof. Note that $Z^\theta_T \xi \in L^r(P)$ and $\xi \in L^r(P^\theta)$ for $r \in (1,2)$ by (2.47). Now, we set $M_t \triangleq E^\theta_t[\xi]$. Then $X_t \triangleq Z^\theta_t M_t = E^P_t[Z^\theta_T \xi]$ is an $L^r(P)$-martingale. By martingale representation, we obtain

$$X_t = X_0 + \textstyle\int_0^t (\bar H_s)^\top dW_s + \int_0^t (\bar K_s)^\top d\bar N_s$$

for some predictable $\mathbb{R}^n \times \mathbb{R}^m$-valued process $(\bar H, \bar K)$. Now, the assertion follows upon computing the dynamics of $M_t = X_t / Z^\theta_t$. □

Note: A suitable martingale representation result for L^2-martingales can be found in Tang and Li (1994); see their Lemma 2.3. Making use of the Burkholder-Davis-Gundy inequality and Doob's L^p-inequality, the result is easily extended to all $p > 1$. For a further discussion of the predictable representation property of Lévy processes, we refer to Nualart and Schoutens (2000) and Di Tella and Engelbert (2015).

Equation (2.49) shows that the conditional expectation process $t \mapsto E_t^\theta[\xi]$ solves a linear BSDE under P. Using the predictable maximizer from Lemma 2.59, one can solve the "maximal" BSDE.

PROPOSITION 2.61. *For every $\xi \in L^2(P)$, there exist a unique adapted càdlàg process X and a unique $\mathbb{R}^n \times \mathbb{R}^m$-valued predictable process (H, K) such that*

$$dX_t = -\max_{(\alpha,\beta)\in\mathcal{D}} (H_t^\top \alpha_t + K_t^\top \beta_t)dt + H_t^\top dW_t + K_t^\top d\bar{N}_t, \; X_T = \xi. \tag{2.50}$$

Proof. Setting

$$g(t, \omega, h, k) \triangleq \max_{(\alpha,\beta)\in\mathcal{D}} [h^\top \alpha_t(\omega) + k^\top(\omega)\beta_t(\omega)], \tag{2.51}$$

we have that g is Lipschitz continuous in (h, k), uniformly in a.e. (t, ω). Moreover, Lemma 2.59 yields predictable a.e.-modifications of $g(\cdot, H, K)$ for all predictable processes H, K; hence, g satisfies the conditions (D1)-(D3) (on p. 55) and Lemma 2.56 shows that (2.50) has a unique solution. □

In view of Proposition 2.61 and the results of Subsection 2.5.3, one can construct a g-expectation corresponding to the BSDE (2.50). Indeed, choosing $h = g$ as in (2.51), the conditions (D1)-(D3) and the normalization property (g2) (on p. 56) are satisfied; the comparison principle (g1) is implied by Lemma 4.1 in Quenez and Sulem (2013).

Following Chen and Epstein (2002), we take a more direct route here and define the nonlinear expectation under drift and intensity uncertainty via its explicit robust representation (2.52).

PROPOSITION 2.62. *For each $\xi \in L^2(P)$, there is some $\bar{\theta} \in \mathcal{D}$ such that*

$$\mathcal{E}_t[\xi] = E_t^{\bar{\theta}}[\xi] = \max_{\theta\in\mathcal{D}} E_t^\theta[\xi] \qquad \textit{for all } t \in [0, T]. \tag{2.52}$$

Proof. Let $\xi \in L^2(P)$ and let (X, H, K) be given as the unique solution of (2.50). Lemma 2.59 applied to $\rho = (H, K)$ yields $\bar{\theta} = (\bar{\alpha}, \bar{\beta}) \in \mathcal{D}$ with

$$H^\top \bar{\alpha} + K^\top \bar{\beta} = \max_{(\alpha,\beta)\in\mathcal{D}} \left(H^\top \alpha + K^\top \beta\right) \quad P \otimes dt\text{-a.e.},$$

and hence X satisfies

$$dX_t = -\left[H_t^\top \bar{\alpha}_t + K_t^\top \bar{\beta}_t\right]dt + H_t^\top dW_t + \bar{K}_t^\top d\bar{N}_t = H_t^\top dW_t^{\bar{\theta}} + \bar{K}_t^\top d\bar{N}_t^{\bar{\theta}}$$

with terminal value $X_T = \xi$. It follows that $X_t = E_t^{\bar{\theta}}[\xi]$.

To show that $\bar{\theta}$ is in fact a maximizer, let $\theta = (\alpha, \beta) \in \mathcal{D}$ and put $Y_t \triangleq E_t^\theta[\xi]$. Lemma 2.60 shows that

$$dY_t = -(I_t^\top \alpha_t + L_t^\top \beta_t)dt + I_t^\top dW_t + L_t^\top d\bar{N}_t, \quad Y_T = \xi.$$

Now, $H^\top \bar{\alpha} + K^\top \bar{\beta} \geqslant H^\top \alpha + K^\top \beta$ holds $P \otimes dt$-a.e., and hence the comparison result from Lemma 4.1 in Quenez and Sulem (2013) implies that $X_t \geqslant Y_t$ for all $t \in [0, T]$. □

Proposition 2.62 allows us to define the operator

$$\mathcal{E}_t : \mathcal{H}_T \to \mathcal{H}_t, \quad \xi \mapsto \max_{\theta \in \mathcal{D}} E_t^\theta[\xi], \quad \text{where } \mathcal{H}_t \triangleq L^\infty(\Omega, \mathcal{F}_t, P),\ t \in [0, T].$$

Note: As in Chen and Epstein (2002), we could also consider $\mathcal{E}_t : L^2(P) \to L^2(\Omega, \mathcal{F}_t, P)$ because the solution $X_t = \max_{\theta \in \mathcal{D}} E_t^\theta[\xi]$ of (2.50) is in $L^2(P)$. Here, we start from bounded random variables in order to use the extension machinery from Section 2.4. Classical stability results for BSDEs (see e.g., Proposition 2.2 in Barles et al. (1997)) show that this makes no difference.

We have, by now, constructed a dynamic robust expectation.

THEOREM 2.63. *The family of operators $(\mathcal{E}_t)_{t \in [0,T]}$ is a sublinear expectation on $(\mathcal{H}_t)_{t \in [0,T]}$.*

Proof. The properties (M), (PC), (SUB) and (PH) are satisfied since $\mathcal{E}_t$ is given as a maximum of linear expectations. Time consistency (TC) follows from (2.52) and stability of $\mathcal{D}$ under pasting: Indeed, let $0 \leqslant s \leqslant t \leqslant T$ and $\xi \in \mathcal{H}_T$. By Proposition 2.62, there are $\theta, \bar{\theta} \in \mathcal{D}$ such that $\mathcal{E}_s[\xi] = E_s^\theta[\xi]$ and $\mathcal{E}_s[\mathcal{E}_t[\xi]] = E_s^{\bar{\theta}}[E_t^\theta[\xi]]$. Then

$$\mathcal{E}_s\left[\mathcal{E}_t[\xi]\right] = E_s^{\bar{\theta}}\left[E_t^\theta[\xi]\right] = E_s\left[\frac{Z_T^{\bar{\theta}}}{Z_t^{\bar{\theta}}}\frac{Z_t^\theta}{Z_s^\theta}\xi\right].$$

We note that $\frac{Z_T^{\bar{\theta}}}{Z_t^{\bar{\theta}}}\frac{Z_t^\theta}{Z_s^\theta} = \frac{Z_T^\rho}{Z_s^\rho}$, where $\rho = 1_{[0,t)}\theta + 1_{[t,T]}\bar{\theta} \in \mathcal{D}$ by stability under pasting. Therefore, we have $\mathcal{E}_s[\mathcal{E}_t[\xi]] = E_s^\rho[\xi] \leqslant \mathcal{E}_s[\xi]$. The converse inequality is obvious. □

Once more, Theorem 2.45 shows that $(\mathcal{E}_t)_{t \in [0,T]}$ extends to an appropriate domain $\{L_t^p : p \geqslant 1,\ t \in [0,T]\}$, where $L_t^p \subset L^1(\Omega, \mathcal{F}_t, P)$ is defined as the closure of $\mathcal{H}_t = L^\infty(\Omega, \mathcal{F}_t, P)$ with respect to the norm $\|\cdot\|_{L,p} \triangleq \sup_{\theta \in \mathcal{D}} \|\cdot\|_{L^p(P^\theta)}$.

Drift and jump intensity uncertainty

THEOREM 2.64. *The nonlinear expectation $(\mathcal{E}_t)_{t \in [0,T]}$ is a regular sublinear expectation on an appropriate domain.*

Proof. Let $\xi \in L^\infty(P)$ and note that $X_t = \mathcal{E}_t[\xi] = E_t^{\bar{\theta}}[\xi]$ for some $\bar{\theta} \in \mathcal{D}$ by Proposition 2.62. Hence X can be chosen to have càdlàg paths and, for all $\theta \in \mathcal{D}$, we have

$$E^\theta\left[|X_{t+h} - X_t|^p\right] \leqslant E\left[|X_{t+h} - X_t|^{2p}\right]^{\frac{1}{2}} E\left[|Z_T^\theta|^2\right]^{\frac{1}{2}}$$

by (2.47). Now, bounded convergence shows that the mapping $t \mapsto L^p$, $t \mapsto X_t$ is càdlàg, and regularity follows with Lemma 2.23. □

2.6 THE LINEAR CASE

The purpose of this section is to demonstrate that the BNEE theory from Section 2.3 is an honest extension of the classical theory of backward stochastic differential equations (BSDEs). To make that point, we show that the solution of a BSDE solves the corresponding BNEE, and that a suitable modification of the solution of a BNEE solves the corresponding BSDE.

Thus let $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t\in[0,T]}, P)$ be a filtered probability space that is both right-continuous and complete. Suppose that $\mathcal{F}_0$ is P-trivial and that $\mathcal{A} = \mathcal{F}_T$. Whenever a martingale appears, we work with a càdlàg version. Moreover, we identify random variables that coincide P-a.s. and stochastic processes that are indistinguishable from one another. We write $L^p_t \triangleq L^p(\Omega, \mathcal{F}_t, P)$, $p \geqslant 1$, $t \in [0, T]$, for the Banach space of p-integrable, $\mathcal{F}_t$-measurable random variables. Clearly,

$$\{(L^p_t, \|\cdot\|_{L,p}) : t \in [0,T],\ p \geqslant 1\} \quad \text{is an appropriate domain}$$

(in the sense of Definition 2.9) for the linear (conditional) expectations

$$E_t : L^p \to L^p_t, \quad X \mapsto E^P_t[X \,|\, \mathcal{F}_t], \quad t \in [0,T].$$

LEMMA 2.65. *The linear expectation $(E_t)_{t\in[0,T]}$ is regular.*

Proof. Let $\xi \in L^p$. We have to prove that $[0,T] \to L^p$, $t \mapsto E_t[\xi]$ is càdlàg. Define a (càdlàg) martingale M by $M_t \triangleq E_t[\xi]$, $t \in [0,T]$. Then $\{|M_t|^p : t \in [0,T]\}$ is uniformly integrable, and M_{t_n} converges pointwisely along all monotone sequences $(t_n)_{n\in\mathbb{N}}$; by Vitali's theorem this convergence is in L^p. □

In this section, we consider the classical conditional expectations $(E_t)_{t\in[0,T]}$ as a regular nonlinear expectation on its appropriate domain $\{(L^p_t, \|\cdot\|_{L,p}) : t \in [0,T],\ p \geqslant 1\}$ of classical Lebesgue spaces. We focus on the case where $\mu(dt) = dt$ is the Lebesgue measure on $[0,T]$ and consider the corresponding Banach space of (equivalence classes of) dt-integrable L^p-processes, as in Definition 2.17,

$$P^p = \big\{[X] : \text{the } L^p\text{-process } [0,T] \to L^p,\ t \mapsto X_t \text{ is measurable}$$
$$\text{and adapted with norm } \|X\|_{P,p} = \textstyle\int_{[0,T]} \|X_t\|_{L,p}\, dt < \infty\big\}.$$

Moreover, as in Definition 2.20, we consider the process space

$$D^p = \{X \text{ adapted } L^p\text{-process} : [0,T] \to L^p,\ t \mapsto X_t \text{ is càdlàg}\} \hookrightarrow P^p,$$

which is a Banach space with norm $\|X\|_{S,p} \triangleq \sup_{t\in[0,T]} \|X_t\|_{L,p}$.

Our nonlinear expectation framework relies on the above spaces of L^p-processes and Bochner integrals. In the classical probabilistic setting of this section, one would, of course, prefer to work with stochastic processes and pathwise integrals. Fortunately, the transition between the two frameworks causes now difficulties:

We identify stochastic processes that coincide $P \otimes dt$-a.e. and write $\mathbb{P}^p$ for the Banach space of all real-valued, progressively measurable stochastic processes X with finite norm $\|X\|_{\mathbb{P},p} = \int_0^T \|X_t\|_{L,p}\, dt$.

PROPOSITION 2.66. *The mapping*

One-to-one correspondence of $\mathbb{P}^p$ and P^p

$$\alpha : \mathbb{P}^p \to P^p, \quad X \mapsto X 1_{[0,\infty)}\big(E[|X|^p]\big)$$

is an isometric isomorphism.

Proof. Clearly, $Y_t \triangleq X_t 1_{[0,\infty)}(E[|X_t|^p])$, $t \in [0,T]$, defines an adapted L^p-process. Using Tonelli's theorem, it is straightforward to prove that the distance function

$$d_\xi : [0,T] \to \mathbb{R},\ t \mapsto E\big[|Y_t - \xi|^p\big]^{\frac{1}{p}} \quad \text{is } \mathcal{B}([0,T])\text{-}\mathcal{B}(\mathbb{R})\text{-measurable}$$

for every $\xi \in L^p$. Since L^p is a separable Banach space, it follows that Y is $\mathcal{B}([0,T])$-$\mathcal{B}(L^p)$-measurable. Moreover, $Y_t = X_t$ in L^p_t for dt-a.e. $t \in [0,T]$ because $E[|X_t|^p] < \infty$ for dt-a.e. $t \in [0,T]$. In particular, $\|Y\|_{P,p} = \|X\|_{P,p} < \infty$ and $Y \in P^p$. By the above, $Y = \alpha(X) \in P^p$ is uniquely characterized by the property

$$Y_t = X_t \quad \text{in } L^p \text{ for } dt\text{-a.e. } t \in [0,T].$$

Hence, α is well-defined, linear and isometric. An elementary approximation argument shows that α is onto; see Lemma A.44 (p. 167) □

The isometric isomorphism α from Proposition 2.66 is consistent with integration:

Consistency of integration

PROPOSITION 2.67. *Let $X \in \mathbb{P}^p$ and $Y \in P^p$ such that $X_t = Y_t$ in L^p for dt-a.e. $t \in [0,T]$. Then, the* pathwise *Lebesgue integral*

$$\oint_A X_t\, dt \triangleq \big[\omega \mapsto \int_A X_s(\omega)\, dt\big], \quad A \in \mathcal{B}([0,T]),$$

defines a member of L^p which coincides with the Bochner integral $\int_A Y_t\, dt$. In particular, we have

$$\oint_A X_t\, dt = \int_A [\alpha(X)]_t\, dt \quad \textit{in } L^p \textit{ for all } X \in \mathbb{P}^p.$$

Proof. Both integrals coincide for step processes. Approximating X by predictable step processes $(X^n)_{n\in\mathbb{N}}$ with $\|X^n - X\|_{P,p} \to 0$, the claim follows by $\|\cdot\|_{P,p}$-continuity of the integral operators. For details, we refer to Proposition A.46 on p. 168 of the appendix. □

With the above one-to-one correspondences between P^p and $\mathbb{P}^p$, we obtain several results of the classical linear theory as consequences of our theory of backward nonlinear expectation equations.

Note: Proposition 2.66 can also be found in Subsection A.3.2 in the appendix, p. 165ff. There, a more detailed proof is provided.

The use of the space $\mathbb{P}^p$ is non-standard. Usually, one would work with the space of progressively measurable processes that are p-integrable with respect to the product measure $P \otimes dt$; however, this space is contained in $\mathbb{P}^p$ so this causes no problems.

2.6.1 *Existence and uniqueness for BSDEs via the theory of BNEEs*

Now, we consider classical backward stochastic differential equations (BSDEs) of the form

$$X_t = E_t\left[\int_t^T f(s, X_s) ds + \xi\right], \quad t \in [0,T], \tag{2.53}$$

where the aggregator f is a BSDE^p-standard parameter in the sense of the following definition.

BSDE^p-standard parameter

DEFINITION 2.68. Let $\xi \in L^p = L^p(P)$, and let

$$f : \Omega \times [0,T] \times \mathbb{R} \to \mathbb{R}$$

be $\mathcal{G} \otimes \mathcal{B}(\mathbb{R})$-measurable, where $\mathcal{G}$ denotes the σ-algebra of progressively measurable sets in $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t\in[0,T]})$. If

(B1) $|f(t,x) - f(t,y)| \leqslant L|x-y|$ for all $x, y \in \mathbb{R}$, $t \in [0,T]$, and

(B2) $\int_0^T E\big[|f(t,0)|^p\big]^{\frac{1}{p}} dt < \infty$,

then (f, ξ) is called a *BSDE^p-standard parameter*. ⋄

If (f,ξ) is a BSDE^p-standard parameter, then $f(\cdot, X)$ is a progressively measurable process for every progressive X. Moreover, (B1) implies that

$$\|f(\cdot,X)\|_{\mathbb{P},p} \leqslant L\|X\|_{\mathbb{P},p} + \int_0^T E\big[|f(t,0)|^p\big]^{\frac{1}{p}} dt, \qquad (2.54)$$

and thus $f(\cdot, X) \in \mathbb{P}^p$ by (B2) whenever $X \in \mathbb{P}^p$.

DEFINITION 2.69. Let (f,ξ) be a BSDE^p-standard parameter. A semimartingale X is called a *solution of the BSDE associated with* (f,ξ) if $\sup_{t\in[0,T]} \|X_t\|_{L,p} < \infty$ and X satisfies (2.53). ⋄

LEMMA 2.70. *Let (f,ξ) be a BSDE^p-standard parameter and X a semimartingale with $\sup_{t\in[0,T]} \|X_t\|_{L,p} < \infty$. Then X is a solution of the BSDE associated with (f,ξ) if and only if there exists a martingale M such that*

$$dX_t = -f(t, X_t)dt + dM_t, \qquad X_T = \xi. \qquad (2.55)$$

In that case, we have

$$X_\sigma = E_\sigma\left[\int_\sigma^\tau f(s, X_s)ds + X_\tau\right] \quad \textit{for all stopping times } \sigma \leqslant \tau, \qquad (2.56)$$

and $\sup\{E[|X_\tau|] \ :\ \tau$ *is a* $[0,T]$*-valued stopping time*$\} < \infty$*; moreover, if* $p > 1$*, then* M *is an* L^p*-martingale and* $E\big[\sup_{t\in[0,T]}|X_t|^p\big] < \infty$.
Finally, X *(considered as a mapping* $[0,T] \to L^p$*) is a member of* D^p.

Proof. Suppose that X is a solution of the BSDE, and define a closed martingale M by

$$M_t \triangleq E_t\left[\int_0^T f(s, X_s)\, ds + \xi\right], \quad t \in [0,T].$$

This is possible by (2.54) and Proposition 2.67. We note that M is an L^p-martingale if $p > 1$. Since X satisfies (2.53), we get

$$M_t - \int_0^t f(s, X_s)\, ds = E_t\left[\int_t^T f(s, X_s)\, ds + \xi\right] = X_t, \quad t \in [0,T], \qquad (2.57)$$

which establishes (2.55).

Now, suppose that (2.55) holds for some martingale M. Let $\sigma \leqslant \tau$ be two stopping times, and integrate (2.55) from time σ up to time τ to obtain

$$X_\tau + \int_\sigma^\tau f(s, X_s)\, ds = X_\sigma + M_\tau - M_\sigma$$

Taking $\mathcal{F}_\sigma$-conditional expectations, we get (2.56). In particular, choosing $\sigma = t \in [0,T]$ and $\tau = T$, we see that X is a solution of the BSDE associated with (f,ξ). Moreover, (2.56) implies

$$|X_\tau| \leqslant E_\tau\left[\int_0^T |f(s,X_s)|\,ds + |\xi|\right] =: N_\tau \quad \text{for all stopping times } \tau.$$

Once again, Proposition 2.67 and (2.54) show that N is a uniformly integrable martingale and, provided that $p > 1$, even an L^p-martingale. Now, $E[|X_\tau|] = E[N_\tau] = N_0 < \infty$, proving that $\sup E[|X_\tau|] < \infty$, where the supremum is taken over all stopping times. Moreover, if $p > 1$, then Doob's L^p-inequality yields

$$E\big[\sup_{t\in[0,T]}|X_t|^p\big] = E\big[\sup_{t\in[0,T]}|N_t|^p\big] \leqslant \big(\tfrac{p}{p-1}\big)^p\, E\big[|N_T|^p\big] < \infty.$$

For the last assertion, recall the decomposition (2.57). The map $[0,T] \to L^p, t \mapsto M_t$ is càdlàg by Lemma 2.65. Moreover, the function $[0,T] \to L^p, t \mapsto \int_0^t f(s,X_s)\,ds$ is continuous by dominated convergence; thus, the mapping $t \mapsto X_t = M_t - \int_0^t f(s,X_s)\,ds$ is càdlàg, and hence $X \in D^p$. □

By Lemma 2.70, we can equivalently say that X is a solution of the BSDE associated with (f,ξ) if X satisfies (2.55) for some martingale M. This definition is of course much more in line with the name backward *stochastic differential* equation. Indeed, it will often (but not always) be convenient to represent the solution of a BSDE by (2.55) rather than by (2.53). We use both terminologies interchangeably.

We now clarify the relationship between the well-established concept of BSDEs and our notion of BNEEs under $(E_t)_{t\in[0,T]}$. For every BSDE^p-standard parameter (f,ξ), we will prove the following :

1. "(f,ξ) can be translated into the 'nonlinear' setting:" There exists a BNEE^p-standard parameter (g,ξ) with
$$\alpha\big(f(\cdot,X)\big) = g\big(\cdot,\alpha(X)\big) \quad \text{in } P^p \text{ for every } X \in \mathbb{P}^p.$$

2. Every solution of the BSDE associated with (f,ξ) is a solution of the BNEE associated with (g,ξ).

3. A modification of the solution of the BNEE associated with (g,ξ) solves the BSDE associated with (f,ξ).

The above three points are hardly surprising. For the sake of mathematical completeness, we nevertheless provide complete, rigorous proofs of these three statements. Together, they show in particular that the BSDE (2.53) has a unique solution (if one only aims at this result, it is, of course, easier to prove it directly). We start with the first point in the above list. To obtain the corresponding BNEE-parameter, we need to modify f in order to ensure that it always maps to L^p.

LEMMA 2.71. *Let* (f, ξ) *be a BSDE*^{p}*-standard parameter. Then, there exists a BNEE*^{p}*-standard parameter* $g : [0,T] \times L^p \to L^p$ *such that*

BNEE parameter from BSDE parameter

$$\alpha\big(f(\cdot, X)\big) = g\big(\cdot, \alpha(X)\big) \qquad \text{for all } X \in \mathbb{P}^p. \tag{2.58}$$

Proof. Property (B2) of a BSDE^p-standard parameter guarantees that $\|f(t,0)\|_{L,p} < \infty$ for all $t \in N^c$ and some dt-null set $N \in \mathcal{B}([0,T])$. Now, it is straightforward to verify that

$$g(t,\eta) \triangleq f(\cdot,t,\eta)1_{N^c}(t) \in L^p$$

defines a BNEE^p-standard parameter with the desired properties; for the details, we refer to Lemma A.48, p.169. □

Next, we show that solutions of BSDEs are solutions of BNEEs.

PROPOSITION 2.72. *Let* (f, ξ) *be a BSDE*^{p}*-standard parameter, and let* (g, ξ) *be a BNEE*^{p}*-standard parameter that satisfies* (2.58)*. If* X *is a solution of the BSDE associated with* (f, ξ)*, then* X *(considered as a mapping* $t \mapsto L^p$*) is the unique solution of the BNEE associated with* (g, ξ)*.*

Solution of BSDE solves BNEE

Proof. Let X be a solution of the BSDE associated with (f, ξ). Then the semimartingale X satisfies

$$X_t = E_t\big[\oint_t^T f(s, X_s)\,ds + \xi\big], \qquad t \in [0,T],$$

where we write $\oint$ to emphasize that the pathwise Lebesgue integral is used. Lemma 2.70 shows that X (considered as a mapping $t \mapsto L^p$) is a member of $D^p \subset P^p$. In particular, $X = \alpha(X) \in P^p$, and hence $\alpha(f(\cdot, X)) = g(\cdot, X)$. Thus, we have

$$X_t = E_t\big[\oint_t^T f(s, X_s)\,ds + \xi\big] = E_t\big[\int_t^T g(s, X_s)\,ds + \xi\big] \quad \text{in } L^p, \quad t \in [0,T],$$

by Proposition 2.67. Consequently, X is the unique solution of the BNEE associated with (g, ξ). □

Finally, we show that the solution of a BNEE admits a modification which solves the corresponding BSDE.

PROPOSITION 2.73. *Let* (f, ξ) *be a BSDE*^{p}*-standard parameter. Suppose that* (g, ξ) *is a BNEE*^{p}*-standard parameter that satisfies* (2.58)*, and let* $Y \in D^p$ *denote the corresponding unique solution given by Theorem* 2.26*.*

Modification of solution of BNEE solves BSDE

There exists a semimartingale X *which solves the BSDE associated with* (f, ξ) *and satisfies* $X_t = Y_t$ *in* L^p *for all* $t \in [0,T]$*.*

Proof. Proposition 2.66 shows that $Y = \alpha(\overline{X})$ in P^p for some $\overline{X}$ in $\mathbb{P}^p$, and hence $\alpha(f(\cdot, \overline{X})) = g(\cdot, Y)$ in P^p by (2.58). In particular, $f(\cdot, \overline{X}) \in \mathbb{P}^p$ and Proposition 2.67 implies that

$$I_t \triangleq \oint_0^t f(s, \bar{X}_s)\,ds = \int_0^t g(s, Y_s)\,ds \quad \text{in } L^p \text{ for all } t \in [0,T].$$

Let M be the martingale with terminal value $M_T = I_T + \xi \in L^p$, and define a semimartingale X by $X_t = M_t - I_t$, $t \in [0,T]$. Then we have

$$X_t = M_t - I_t = E_t\left[I_T - I_t + \xi\right] = E_t\left[\int_t^T g(s, Y_s)\, ds + \xi\right] = Y_t \quad \text{in } L^p$$

for all $t \in [0,T]$. In particular, $X = Y \in D^p$; hence $\alpha(X) = Y = \alpha(\overline{X})$ in $\mathbb{P}^p$, and therefore $X = \overline{X}$ in $\mathbb{P}^p$. Thus, $I_t = \oint_0^t f(s, X_s)\, ds$ in L^p for all $t \in [0,T]$, and we have

$$X_t = E_t\left[I_T - I_t + \xi\right] = E_t\left[\oint_t^T f(s, X_s)\, ds + \xi\right], \qquad t \in [0,T]. \qquad \square$$

Combining the above, we obtain the classical existence and uniqueness result for BSDEs. In the present semimartingale setting, it was first proven by Duffie and Epstein (1992b).

Existence and uniqueness for BSDEs

THEOREM 2.74. *Let (f, ξ) be a BSDE^p-standard parameter. Then the BSDE*

$$X_t = E_t\left[\int_t^T f(s, X_s) ds + \xi\right], \quad t \in [0,T], \tag{2.53}$$

has a unique solution $X \in D^p$.

Proof. Lemma 2.71 yields a BNEE^p-standard parameter (g, ξ) with

$$\alpha(f(\cdot, X)) = g(\cdot, \alpha(X)), \qquad \text{for all } X \in \mathbb{P}^p,$$

and hence Proposition 2.73 yields a solution X of the BSDE. Proposition 2.72 shows that every solution is a modification of the unique solution of the BNEE associated with (g, ξ); thus, any two solutions are indistinguishable. $\square$

Similarly as in Theorem 2.74, combining Lemma 2.71, Proposition 2.72 and Proposition 2.73 allows us to translate all of the results from Section 2.3 into the linear setting. For instance, we have

Stability of BSDEs

COROLLARY 2.75. *Let (f^n, ξ^n), $n \in \mathbb{N}$, and (f, ξ) be BSDE^p-parameters. Suppose there is a constant $L > 0$ such that*

$$|f^n(t,x) - f^n(t,y)| \leqslant L|x-y| \quad \text{for all } x, y \in \mathbb{R},\ t \in [0,T],$$

and all $n \in \mathbb{N}$. Let X^n, $n \in \mathbb{N}$, and X denote the solutions of the associated BSDEs and suppose that

$$\int_0^T E\left[|f^n(t, X_t) - f(t, X_t)|^p\right]^{\frac{1}{p}} dt \to 0 \quad \text{and} \quad \xi^n \to \xi \ \text{in } L^p.$$

Then, $X^n \to X$ in S^p, and $E[\sup_{t\in[0,T]} |X^n_t - X_t|^p] \to 0$ if $p > 1$.

The discretization result of Theorem 2.32 also has its analog, of course; however, it will not be needed in the following, and so we do not record it.

2.6.2 *A comparison theorem for BSDEs*

A nice feature of BSDEs is that their solutions can be expected to satisfy a comparison principle. This means that solutions of BSDEs can be compared simply by comparing their aggregators and terminal values. We conclude our brief discussion of BSDEs by proving such a comparison result.

Comparison theorem

THEOREM 2.76. *Let (f, ξ) be a $BSDE^p$-standard parameter and let X denote the solution of the correpsonding BSDE*

$$dX_t = -f(t, X_t)dt + dM_t, \qquad X_T = \xi.$$

Let $g : \Omega \times [0, T] \times \mathbb{R} \to \mathbb{R}$ be a $\mathcal{G} \otimes \mathcal{B}(\mathbb{R})$-measurable function, and suppose that Y is a semimartingale that satisfies

$$dY_t = -g(t, Y_t)dt + dN_t, \qquad Y_T = \overline{\xi} \in L^p,$$

for some martingale N. Then $X \leqslant Y$, provided that

$$\xi \leqslant \overline{\xi}, \qquad f(t, Y_t) \leqslant g(t, Y_t) \quad \text{for } dt\text{-a.e. } t \in [0, T]$$

and $E\left[\int_0^T (\|g(t, Y_t)\|_{L,p} + \|Y_t\|_{L,p})dt\right] < \infty$.

Proofs of comparison theorems usually rely on explicit representations of solutions of linear BSDEs or closely related stochastic Gronwall inequalities (see Section 3.4 below). If the underlying expectation is truly nonlinear, backward nonlinear expectation equations that look linear are, in fact, nonlinear, and thus one cannot expect a comparison principle as powerful as the one in Theorem 2.76 to hold. Linear BSDEs under linear expectations can, however, be solved explicitly:

Linear BSDEs

PROPOSITION 2.77. *Let α be a bounded progressively measurable process, and let $\beta \in \mathbb{P}^p$, $\xi \in L^p$. Then, the unique solution of the linear BSDE*

$$dX_t = -\left[\alpha_t X_t + \beta_t\right] + dM_t, \qquad X_T = \xi, \tag{2.59}$$

is given by

$$X_t = E_t\left[\int_t^T e^{\int_t^s \alpha_u du}\beta_s ds + e^{\int_t^T \alpha_u du}\xi\right], \quad t \in [0, T]. \tag{2.60}$$

Proof. Setting $f(t, x) = \alpha_t x + \beta_t$, it is immediate that (f, ξ) is a $BSDE^p$-standard parameter; thus, (2.59) has a unique solution (X, M).

Now, suppose that $p > 1$. Then M is an L^p-martingale, and integrating by parts, we get

$$de^{\int_0^t \alpha_u du}X_t = -e^{\int_0^t \alpha_u du}\beta_t dt + e^{\int_0^t \alpha_u du}dM_t. \tag{2.61}$$

Note that $e^{\int_0^t \alpha_u du}dM_t$ is also an L^p-martingale since α is bounded and hence (2.61) yields

$$e^{\int_0^t \alpha_u du}X_t = E_t\left[\int_t^T e^{\int_0^s \alpha_u du}\beta_s ds + e^{\int_0^T \alpha_u du}X_T\right], \quad t \in [0, T],$$

which is equivalent to (2.60).

For $p = 1$, we apply the above to the truncated parameters $\beta^n \triangleq (-n \vee \beta) \wedge n$, $\xi^n \triangleq (-n \vee \xi) \wedge n$, to see that the semimartingales

$$X_t^n = E_t\left[\int_t^T e^{\int_t^s \alpha_u du} \beta_s^n ds + e^{\int_t^T \alpha_u du} \xi^n\right], \quad n \in \mathbb{N},$$

are the unique solutions of $dX_t^n = -[\alpha_t X_t^n + \beta_t^n]dt + dM_t^n$, $X_T^n = \xi^n$. The assumptions of Corollary 2.75 (stability result) are fulfilled, and hence $X^n \to X$ in S^p, where X denotes the unique solution of the linear BSDE (2.59). On the other hand, dominated convergence implies $X_t^n \to E_t[\int_t^T e^{\int_t^s \alpha_u du} \beta_s ds + e^{\int_t^T \alpha_u du} \xi]$ in L^1; thus (2.61) is established for $p = 1$, as well. □

With Proposition 2.77, we can give the proof of Theorem 2.76.

Proof of Theorem 2.76. We set $\Delta \triangleq X - Y$ and note that

$$\alpha_t \triangleq \frac{f(t,X_t) - f(t,Y_t)}{X_t - Y_t} 1_{\{X_t \neq Y_t\}}, \qquad t \in [0,T],$$

is a bounded process by the Lipschitz condition (B1). Now, we have

$$d\Delta_t = -\left[\alpha_t \Delta_t + \beta_t\right]dt + d(M_t - N_t), \quad \Delta_T = \xi - \bar{\xi} \leqslant 0, \tag{2.62}$$

where $\beta_t = f(t, Y_t) - g(t, Y_t) \leqslant 0$ and $\beta \in \mathbb{P}^p$. Since Δ satisfies the linear BSDE (2.62), Proposition 2.77 implies that

$$\Delta_t = E_t\left[\int e^{\int_t^s \alpha_u du} \beta_s ds + e^{\int_t^T \alpha_u du}(\xi - \bar{\xi})\right] \leqslant 0. \qquad \square$$

2.6.3 *Bibliographical notes*

Backward stochastic differential equations (BSDEs) have first appeared in Bismut (1973). The generalized modern formulation of BSDEs is due to Pardoux and Peng (1990), who have also settled the question of existence and uniqueness under a global Lipschitz condition in a Brownian setting. In the special case of BSDEs driven by Brownian motion, their result implies our Theorem 2.74. In a general semimartingale framework, existence and uniqueness of BSDEs under a global Lipschitz condition (Theorem 2.74) were first proven by Duffie and Epstein (1992b), in the context of their seminal contribution of stochastic differential utility. Their results were generalized by Antonelli (1993), who replaces the integrator dt in the BSDE by an increasing process dA_t and provides an L^1-theory. Stability results for such BSDEs in a semimartingale setting can be found in Antonelli (1996) and Antonelli and Kohatsu-Higa (2000).

The literature on BSDEs mostly focuses on Brownian filtrations, where there is a strong connection to quasilinear partial differential equations, see, e.g., Pardoux and Peng (1992), Ma et al. (1994, 2012) and the references therein. The theory has undergone several generalizations and is now able to deal with quadratic and convex drivers;

moreover, the Lipschitz assumption has been replaced by monotonicity and polynomial growth conditions; see, e.g., Kobylanski (2000), Briand and Carmona (2000), Delarue (2002), Briand and Hu (2008) and Delbaen et al. (2011). BSDEs have also been investigated in jump-diffusion settings, see e.g., Tang and Li (1994), Barles et al. (1997), Nualart and Schoutens (2001), Becherer (2006), Royer (2006), Quenez and Sulem (2013), Delong (2013), Kharroubi and Pham (2015) and the references therein.

2.7 RECURSIVE UTILITY WITH NONLINEAR EXPECTATIONS

In the following, as advertised in the introduction of this thesis, we apply our theory of backward nonlinear expectation equations to utility theory; see, in particular, Section 1.1 (p. 4ff.) and Section 1.2 (p. 7ff.).

Throughout this section, let $(\mathcal{E}_t)_{t\in[0,T]}$ denote a regular nonlinear expectation, carried by a sublinear expectation $(\mathcal{E}_t^{\mathrm{sub}})_{t\in[0,T]}$ on an appropriate domain $\{(L_t^p, \|\cdot\|_{L,p}) : t \in [0,T],\ p \geqslant 1\}$. A nonlinear expectation can be used as a certainty equivalent in the construction of recursive utilities: Given a suitable *discrete-time aggregator* $W : [0,T] \times L^p \to L^p$, we can define a *recursive utility process* $(U_{t_k})_{k=0,\dots,N}$ via

$$U_{t_k} \triangleq W\big(t_{k+1} - t_k, c_{t_k}, \mathcal{E}_{t_k}[U_{t_{k+1}}]\big), \quad U_{t_N} = \xi,$$

where $0 = t_0 < \cdots < t_N = T$ are time-periods between which a consumption of $(t_{k+1} - t_k)c_{t_k} \in L^p_{t_k}$ takes place.

In Section 2.5, we have given several examples of regular nonlinear expectations, and any one of these may be used here. For instance, we may choose $(\mathcal{E}_t)_{t\in[0,T]}$ as the superlinear counterpart of the expectation under drift and jump intensity uncertainty from Subsection 2.5.4,

$$\mathcal{E}_t[\xi] = \min_{\theta\in\mathcal{D}} E_t^{\theta}[\xi],$$

which corresponds to recursive utility under multiple priors, as investigated by Epstein and Wang (1994), Chen and Epstein (2002), Epstein and Schneider (2003) and Hayashi (2005). Alternatively, we may take $\mathcal{E}_t = -\mathcal{E}_t^{\mathrm{sub}}[-\cdot]$ as the superlinear counterpart of the random G-expectation from Subsection 2.5.1, which corresponds to a discrete-time analog of the continuous-time utility with drift and volatility uncertainty as proposed in Epstein and Ji (2013, 2014). Of course, we may also simply take $(\mathcal{E}_t)_{t\in[0,T]}$ as the classical conditional expectation; this corresponds to recursive utility with expected utility certainty equivalents, see Kraft and Seifried (2014).

In continuous time, Chen and Epstein (2002) construct stochastic differential utility under multiple priors. Epstein and Ji (2014) propose a continuous-time version of stochastic differential utility in the setting of Nutz (2012); note that Theorems 2.26 and 2.50 guarantee existence and uniqueness of the associated utility process

$$U_t = \mathcal{E}_t\left[\int_t^T f(c_s, U_s)ds + \xi\right], \quad t \in [0,T]. \tag{2.63}$$

The main purpose of this section is to prove that – under suitable conditions on W and in the limit of vanishing grid size – the recursive utility process converges to a solution of the BNEE (2.63), where the *continuous-time aggregator* f is given by the derivative $f(c,u) = \frac{\partial}{\partial\Delta}W(\Delta,c,u)|_{\Delta=0}$. Thus, we substantiate the axiomatic definition of stochastic differential utility under nonlinear expectations in continuous time, generalizing the results of Kraft and Seifried (2014).

A convergence result

Denote by $P^p = P^p(dt)$ the space of all dt-integrable adapted L^p-processes as defined in Subsection 2.2.4.

Let $(\Delta^n)_{n\in\mathbb{N}}$ be a sequence of partitions,

$$\Delta^n : 0 = t_0^n < t_1^n < \cdots < t_{N_n}^n = T, \qquad \text{set} \quad \Delta_k^n \triangleq t_k^n - t_{k-1}^n,$$

and suppose that $|\Delta^n| = \max_{k=1,\dots,N_n} \Delta_k^n \to 0$; see also Section 2.3.3. We consider *consumption plans* (c,ξ), where $\xi \in L^p$ is a terminal payoff and $c \in P^p$ is a consumption rate process. We suppose that there exists an approximating sequence $(c^n)_{n\in\mathbb{N}} \subseteq P^p$ as follows:

The consumption plan c^n is piecewise constant on Δ^n, i.e., $c_t^n = c_{t_k^n}^n$ for $t \in [t_k n, t_{k+1}^n)$, and we have

$$c^n \to c \quad \text{in } P^p \quad \text{and} \quad c_t^n \to c_t \quad \text{for } dt\text{-a.e. } t \in [0,T].$$

Recursive utility is constructed via a discrete-time aggregator

$$W : [0,T] \times L^p \times L^p \to L^p, \qquad W(\Delta,c,u) \triangleq u + \Delta f^{\Delta}(c,u)$$

that satisfies the following conditions:

(A1) There exists a modulus of continuity[3] $h : [0,T] \to \mathbb{R}$ such that

$$\|f^{\Delta}(c,u) - f^0(c,u)\|_{L,p} \leqslant h(\Delta)(1 + \|c\|_{L,p} + \|u\|_{L,p})$$

for all $c,u \in L^p$.

(A2) There exists $L > 0$ such that f^0 satisfies the Lipschitz property

$$\|f^0(c,u_1) - f^0(c,u_2)\|_{L,p} \leqslant L\|u_1 - u_2\|_{L,p} \quad \text{for all } c,u_1,u_2 \in L^p.$$

(A3) $f^0(\cdot,u)$ is continuous for every $u \in L^p$.

(A4) There exists a constant $K > 0$ such that

$$\|f^0(c,0)\|_{L,p} \leqslant K(1 + \|c\|_{L,p}) \quad \text{for every } c \in L^p.$$

(A5) $f^0(c,u) \in L_t^p$ whenever $c,u \in L_t^p$, $t \in [0,T]$.

3 That is, h is a continuous increasing function with $h(0) = 0$.

For brevity, we write $f \triangleq f^0$ in the sequel. The mapping $f : L^p \times L^p \to L^p$ is called the *continuous-time aggregator*.

Remark. Writing $W(\Delta, c, u) \triangleq u + \Delta f^\Delta(c, u)$ is without loss of generality, of course. With that notation, $(W(\Delta, c, u) - W(0, c, u))/\Delta = f^\Delta(c, u)$, and hence (A1) guarantees that $\frac{\partial}{\partial \Delta} W(\Delta, c, u)|_{\Delta=0} = f(c, u)$ exists as a sufficiently uniform limit. ◇

LEMMA 2.78. *Under the preceding assumptions, for each consumption plan* (c, ξ)*, the function*

$$f_c : [0, T] \times L^p \mapsto L^p, \quad (t, \eta) \mapsto f(c_t, \eta),$$

defines a BNEE^p*-standard parameter* (f_c, ξ).

Proof. The requirement $f_c(t, L^p_t) \subset L^p_t$ for all $t \in [0, T]$ and the uniform Lipschitz condition (2.14) are fulfilled by (A2) and (A5). Thus, it remains to show that $f_c(\cdot, X) \in P^p$ for every $X \in S^p$. For this, by (A2) and a closure argument, it suffices to check that $f_c(\cdot, X)$ is a $\mathcal{B}([0, t])$-$\mathcal{B}(L^p_t)$-measurable simple function for every measurable simple function $X : [0, t] \to L^p_t$. This, in turn, is obvious from (A5). □

DEFINITION 2.79. Let (c, ξ) be a consumption plan with an associated approximating sequence $(c^n)_{n \in \mathbb{N}}$. The discrete-time *recursive utility* process U^n is defined on the time grid Δ^n via

$$U^n_k \triangleq W\big(\Delta^n_k, c^n_{t^n_k}, \mathcal{E}_{t^n_k}[U^n_{k+1}]\big), \quad \text{where } U^n_{N_n} = \xi,$$

and $\Delta^n : 0 = t^n_0 < \ldots < t^n_{N_n} = T$, $\Delta^n_k \triangleq t^n_{k+1} - t^n_k$, $k = 0, \ldots, N_n - 1$.

The continuous-time *stochastic differential utility* process U is given as the unique solution of the BNEE

$$U_t = \mathcal{E}_t\big[\textstyle\int_t^T f(c_s, U_s)ds + \xi\big] = \mathcal{E}_t\big[\textstyle\int_t^T f_c(s, U_s)ds + \xi\big]. \qquad ◇$$

Note that existence and uniqueness of a solution $U \in D^p$ are guaranteed by Theorem 2.26. The main result of this section demonstrates that the recursive utility processes U^n converge to the stochastic differential utility process U:

THEOREM 2.80. *Let* U^n, $n \in \mathbb{N}$, *be the discrete-time recursive utility processes, and let* U *be the stochastic differential utility process from Definition 2.79. Then,*

Continuous-time limit of recursive utility

$$\max_{k=0,\ldots,N_n} \|U^n_k - U_{t^n_k}\|_{L^p} \to 0.$$

Remark. Theorem 2.80 implies in particular that the recursive utility values $U^n_0 \in \mathbb{R}$ converge to the stochastic differential utility value $U_0 \in \mathbb{R}$. ◇

The proof proceeds in two steps. Introducing the BNEE^p-standard parameters $f^n \triangleq f_{c^n} = \sum_{k=0}^{N_n-1} 1_{[t^n_k, t^n_{k+1})} f(c^n_{t^n_k}, \cdot)$, $n \in \mathbb{N}$, we have

LEMMA 2.81. *With* f^n *as above,* $(\Delta^n, f^n, \xi)_{n \in \mathbb{N}}$ *is* (f_c, ξ)*-exhausting.*

Proof. Lemma 2.78 shows that (f^n, ξ) is a BNEE^p-standard parameter for every $n \in \mathbb{N}$. The definition of f^n and (A2) imply

$$\|f^n(t,\zeta) - f^n(t,\eta)\|_{L,p} = \|f(c^n_t, \zeta) - f(c^n_t, \eta)\|_{L,p} \leqslant L\|\zeta - \eta\|_{L,p}$$

for all $\zeta, \eta \in L^p$ and all $n \in \mathbb{N}$. Thus the Lipschitz condition (2.21), as required for an exhausting sequence, is satisfied. Since $|\Delta^n| \to 0$, it remains to establish (2.22), i.e., we must prove that

$$\textstyle\sum_{k=0}^{N_n-1} \int_{[t^n_k, t^n_{k+1})} \|f_c(t, U_t) - f^n(t^n_k, U_t)\|_{L,p}\, dt \to 0. \tag{2.64}$$

The approximating consumption plan c^n is piecewise constant on Δ^n, and hence we have

$$f^n(t, U_t) = f(c^n_t, U_t) = f(c^n_{t^n_k}, U_t) = f^n(t^n_k, U_t) \quad \text{for } t \in [t^n_k, t^n_{k+1}).$$

Therefore, the sum in (2.64) can be written as

$$\textstyle\int_0^T \|f_c(t, U_t) - f^n(t, U_t)\|_{L,p}\, dt. \tag{2.65}$$

Now, we have $c^n_t \to c_t$ for dt-a.e. $t \in [0, T]$, and thus

$$f^n(t, U_t) = f(c^n_t, U_t) \to f(c_t, U_t) = f_c(t, U_t) \quad \text{for } dt\text{-a.e. } t \in [0, T],$$

by (A3) – the aggregator function f is continuous in the first component. Hence, we have pointwise convergence inside the integral in (2.65). To establish (2.64), by Vitali's theorem, it remains to show that the sequence $(\|f_c(t, U_t) - f^n(t, U_t)\|_{L,p})_{n \in \mathbb{N}}$ is uniformly integrable.

To do so, we first estimate $\|f^\Delta(c, u)\|_{L,p}$ for $\Delta \geqslant 0$ and $c, u \in L^p$. By (A1), we have

$$\|f^\Delta(c, u) - f^0(c, u)\|_{L,p} \leqslant h(\Delta)(1 + \|c\|_{L,p} + \|u\|_{L,p}),$$

and (A2) implies $\|f^0(c, u) - f^0(c, 0)\|_{L,p} \leqslant L\|u\|_{L,p}$. Finally, (A4) yields $\|f^0(c, 0)\|_{L,p} \leqslant K(1 + \|c\|_{L,p})$; hence, we get

$$\|f^\Delta(c, u)\|_{L,p} \leqslant C_0(1 + \|c\|_{L,p} + \|u\|_{L,p}) \quad \text{for all } c, u \in L^p, \tag{2.66}$$

where $C_0 \triangleq h(T) + L + K$. In particular, we have

$$\|f^n(t, U_t) - f_c(t, U_t)\| \leqslant C_0(2 + \|c_t\|_{L,p} + \|c^n_t\|_{L,p} + 2\|U\|_{S,p}).$$

Now, $c^n \to c$ in P^p, and thus the right-hand side of the above inequality converges in $L^1([0, T], dt)$; hence, it is uniformly integrable by Vitali's theorem, and then, so is the left-hand side. □

With Lemma 2.81, we have shown that $(\Delta^n, f^n, \xi)_{n \in \mathbb{N}}$ is (f_c, ξ)-exhausting. Therefore our convergence result for discrete-time approximations of Theorem 2.32 applies, and shows that

$$\max_{k=0,\ldots,N_n} \|X^n_k - U_{t^n_k}\|_{L,p} \to 0,$$

where $X^n_{N_n} = \xi$ and

$$X^n_k \triangleq \mathcal{E}_{t^n_k}\Big[\Delta^n_k f^n\big(t_k, \mathcal{E}_{t^n_k}[X^n_{k+1}]\big) + X^n_{k+1}\Big], \quad k = N_n - 1, \ldots, 0. \quad (2.67)$$

To conclude the proof of Theorem 2.80, it thus remains to show that

$$\max_{k=0,\ldots,N_n} \|U^n_k - X^n_k\|_{L,p} \to 0.$$

This will be accomplished with Lemma 2.83 below, the proof of which relies on the following a priori estimate.

LEMMA 2.82. *There is a constant* $C_1 > 0$ *such that for all but finitely many* $n \in \mathbb{N}$ *we have* *A priori estimate*

$$\max_{k=0,\ldots,N_n} \|U^n_k\|_{L,p} \leqslant C_1\big(1 + \|\xi\|_{L,p} + \|c\|_{P,p}\big).$$

Proof. We recall that $U^n_k = W(\Delta^n_k, c^n_{t^n_k}, \mathcal{E}_{t^n_k}[U^n_{k+1}])$, where $W(\Delta, c, v) = v + \Delta f^\Delta(c, v)$. Hence, for every $k = 0, \ldots, N_n - 1$, we have

$$U^n_k = \mathcal{E}_{t^n_k}[U^n_{k+1}] + \Delta^n_k f^{\Delta^n_k}(c^n_{t^n_k}, \mathcal{E}_{t^n_k}[U^n_{k+1}]). \quad (2.68)$$

Thus estimate (2.66) from the proof of Lemma 2.81 yields

$$\|U^n_k\|_{L,p} \leqslant \big\|\mathcal{E}_{t^n_k}[U^n_{k+1}]\big\|_{L,p} + \Delta^n_k C_0\big(1 + \|c^n_{t^n_k}\|_{L,p} + \big\|\mathcal{E}_{t^n_k}[U^n_{k+1}]\big\|_{L,p}\big).$$

Using the contraction property (Lemma 2.15), we get

$$\|U^n_k\|_{L,p} \leqslant (1 + C_0\Delta^n_k)\|U^n_{k+1}\|_{L,p} + C_0\Delta^n_k\big(1 + \|c^n_{t^n_k}\|_{L,p}\big) \quad (2.69)$$

for all $k = 0, \ldots, N_n - 1$. Iterating (2.69), we arrive at

$$\|U^n_k\|_{L,p} \leqslant P^{N_n}_k \|U^n_{N_n}\|_{L,p} + \textstyle\sum_{\ell=k}^{N_n-1} P^\ell_k C_0 \Delta^n_\ell (1 + \|c^n_{t^n_\ell}\|_{L,p}),$$

where $P^m_k \triangleq \prod_{k \leqslant \ell < m}(1 + C_0\Delta^n_\ell)$. Since $1 + x \leqslant e^x$, we have

$$P^m_k \leqslant \exp\Big(C_0 \textstyle\sum_{\ell=k}^{m-1} \Delta^n_\ell\Big) \leqslant e^{C_0 T} \quad \text{for all } k, m = 0, \ldots, N_n - 1.$$

For all $k = 0, \ldots, N_n$, we hence obtain

$$\|U^n_k\|_{L,p} \leqslant e^{C_0 T}\|\xi\|_{L,p} + C_0 e^{C_0 T} \textstyle\sum_{\ell=0}^{N_n-1} \Delta^n_\ell (1 + \|c^n_{t^n_\ell}\|_{L,p}).$$

Since $\sum_{\ell=0}^{N_n-1} \Delta^n_\ell \|c^n_{t^n_\ell}\|_{L,p} \to \|c\|_{P,p}$, the proof is complete. □

LEMMA 2.83. *The processes* X^n *from* (2.67) *satisfy*

$$\max_{k=0,\ldots,N_n} \|U^n_k - X^n_k\|_{L,p} \to 0.$$

Proof. By shift-invariance (SI), equation (2.68) can be rewritten as

$$U^n_k = \mathcal{E}_{t^n_k}\big[\Delta^n_k f^{\Delta^n_k}(c^n_{t^n_k}, \mathcal{E}_{t^n_k}[U^n_{k+1}]) + U^n_{k+1}\big].$$

The definition of X_k^n from (2.67) reads

$$X_k^n \triangleq \mathcal{E}_{t_k^n}\left[\Delta_k^n f^n\left(t_k, \mathcal{E}_{t_k^n}[X_{k+1}^n]\right) + X_{k+1}^n\right].$$

We set $D_k^n \triangleq \|U_k^n - X_k^n\|_{L,p}$. The contraction property (Lemma 2.15) implies

$$D_k^n \leqslant \Delta_k^n \|f^{\Delta_k^n}(c_{t_k^n}^n, \mathcal{E}_{t_k^n}[U_{k+1}^n]) - f(c_{t_k^n}^n, \mathcal{E}_{t_k^n}[X_{k+1}^n])\|_{L,p} + D_{k+1}^n.$$

By (A1), (A2) and the contraction property, we get

$$\begin{aligned}&\|f^{\Delta_k^n}(c_{t_k^n}^n, \mathcal{E}_{t_k^n}[U_{k+1}^n]) - f(c_{t_k^n}^n, \mathcal{E}_{t_k^n}[X_{k+1}^n])\|_{L,p}\\ &\leqslant h(\Delta_k^n)(1 + \|c_{t_k^n}^n\|_{L,p} + \|U_{k+1}^n\|_{L,p}) + L\|U_{k+1}^n - X_{k+1}^n\|_{L,p}.\end{aligned}$$

Hence, we have

$$D_k^n \leqslant h(|\Delta^n|)\Delta_k^n(1 + \|c_{t_k^n}^n\|_{L,p} + \|U_{k+1}^n\|_{L,p}) + (1 + \Delta_k^n L)D_{k+1}^n \quad (2.70)$$

for all $k = 0, \dots, N_n - 1$. Iterating inequality (2.70), we find

$$D_k^n \leqslant h(|\Delta^n|)\sum_{\ell=k}^{N_n-1} P_k^\ell \Delta_\ell\left(1 + \|c_{t_\ell^n}^n\|_{L,p} + \|U_{\ell+1}^n\|_{L,p}\right) + P_k^{N_n} D_n^n,$$

where $P_k^\ell = \prod_{k\leqslant j\leqslant \ell-1}(1 + \Delta_j^n L)$ and $D_n^n = \|U_n^n - X_n^n\|_{L,p} = 0$. Using $1 + x \leqslant e^x$ again, we get

$$P_k^\ell \leqslant \exp\left(L\sum_{j=k}^{\ell-1}\Delta_j^n\right) \leqslant e^{LT} \quad \text{for all } k, \ell = 0, \dots, N_n - 1,$$

and therefore

$$D_k^n \leqslant h(|\Delta^n|)e^{LT}\sum_{\ell=0}^{N_n-1}\Delta_\ell^n(1 + \|c_{t_\ell^n}^n\|_{L,p} + \|U_{\ell+1}^n\|_{L,p}).$$

For all but finitely many $n \in \mathbb{N}$, we have $\max_{k=0,\dots,N_n}\|U_k^n\|_{L,p} \leqslant C_1(1 + \|\xi\|_{L,p} + \|c\|_{\mathbb{P},p})$ by Lemma 2.82 as well as $\sum_{\ell=0}^{N_n-1}\Delta_\ell^n\|c_{t_\ell^n}^n\|_{L,p} = \int_0^T \|c_t^n\|_{L,p}dt \leqslant 1 + \|c\|_{\mathbb{P},p}$. Setting $K \triangleq C_1(1 + \|\xi\|_{L,p} + \|c\|_{\mathbb{P},p}) + 1 + \|c\|_{\mathbb{P},p}$, we thus obtain

$$\|U_k^n - X_k^n\|_{L,p} = D_k^n \leqslant e^{LT}KTh(|\Delta^n|), \qquad k = 0, \dots, N_n,$$

for all but finitely many $n \in \mathbb{N}$, and the claim follows. □

In view of the discussion preceding Lemma 2.82, this also completes the proof of Theorem 2.80.

3 STOCHASTIC DIFFERENTIAL UTILITY

In this chapter, we investigate the Epstein-Zin parameterization of continuous-time recursive utility with relative risk aversion (RRA) $\gamma > 1$ and elasticity of intertemporal substitution (EIS) $\psi > 1$. The material of this chapter is largely based on Seiferling and Seifried (2015).

Specifications of Epstein-Zin utility with RRA $\gamma > 1$ and EIS $\psi > 1$ are important in a number of both theoretical and empirical applications. In spite of their widespread usage, the fundamental questions of existence, uniqueness and concavity of the corresponding utility functionals, which will be addressed below, have so far remained unresolved for these parameters.

In a fully general semimartingale setting, we establish existence and uniqueness as well as monotonicity and concavity of continuous-time Epstein-Zin recursive utility with RRA $\gamma > 1$ and EIS $\psi > 1$. Moreover, we will provide a corresponding utility gradient inequality.

This chapter is organized as follows: Section 3.1 introduces continuous-time Epstein-Zin utility. Our main results are stated in Section 3.2, which also provides links to the literature. We then take a slight detour: In Section 3.3, Epstein-Zin utility for bounded consumption plans is investigated, and Section 3.4 is concerned with stochastic Gronwall inequalities. Then we focus on the specification RRA $\gamma > 1$ and EIS $\psi > 1$ again: The proofs of our results are provided in Section 3.5.

3.1 DYNAMIC RISK PREFERENCES AND EPSTEIN-ZIN UTILITY

Let $(\Omega, \mathcal{F}_T, P)$ be a probability space, endowed with a right-continuous and complete filtration $\mathcal{F} = (\mathcal{F}_t)_{t\in[0,T]}$. We denote by $\mathcal{S}$ the space of $(\mathcal{F}, P)$-semimartingales and, for $\beta \geqslant 1$, by $\mathcal{S}^\beta$ the space of all semimartingales X with $E[\sup_{t\in[0,T]} |X_t|^\beta] < \infty$. Adopting a standard convention, we identify random variables that coincide almost surely and stochastic processes that are indistinguishable.

DEFINITION 3.1. A predictable process c that takes values in $\mathfrak{C} \triangleq (0, \infty)$ is called a continuous-time *consumption plan* if

Consumption plans

$$E\left[\int_0^T c_t^\beta dt + c_T^\beta\right] < \infty \qquad \text{for all } \beta \in \mathbb{R}.$$

The family of all consumption plans is denoted by $\mathcal{C}$. ⋄

Here, c_t represents the time-t consumption rate and c_T models lump-sum consumption at time T. A partial order on the set $\mathcal{C}$ of consumption plans is defined via

$$c \preccurlyeq \bar{c} \quad \Longleftrightarrow \quad c_t \leqslant \bar{c}_t \quad \text{for } dt\text{-a.e. } t \in [0,T] \quad \text{and} \quad c_T \leqslant \bar{c}_T.$$

Accordingly, two consumption plans c and $\bar{c}$ are identified if $c \preccurlyeq \bar{c}$ and $\bar{c} \preccurlyeq c$. We investigate intertemporal risk preferences on the lattice $(\mathcal{C}, \preccurlyeq)$ of continuous-time consumption plans.

Preferences

In general, an agent's subjective preferences can be described in terms of a *utility index* functional $\nu : \mathcal{C} \to \mathbb{R}$ such that c is weakly preferred to $\bar{c}$ if and only if $\nu(c) \geqslant \nu(\bar{c})$; here, ν will typically be monotone in the sense that $c \preccurlyeq \bar{c}$ implies $\nu(c) \leqslant \nu(\bar{c})$. Note that ν need not be monotone a priori. In a continuous-time recursive utility context, the utility index functional ν is defined via

$$\nu : \mathcal{C} \to \mathbb{R}, \quad \nu(c) \triangleq V_0(c),$$

Stochastic differential utility

where the *(stochastic differential) utility process* $V = V(c)$ satisfies a backward stochastic differential equation (BSDE) of the form

$$V_t = E_t\big[\int_t^T f(c_s, V_s)ds + U(c_T)\big], \qquad t \in [0,T], \tag{3.1}$$

see Duffie and Epstein (1992b). We investigate the *Epstein-Zin parameterization* of recursive utility with

$$\begin{aligned} \text{relative risk aversion (RRA)} \quad & \gamma > 0, \quad \gamma \neq 1, \\ \text{elasticity of intertemporal substitution (EIS)} \quad & \psi > 0, \quad \psi \neq 1, \end{aligned}$$

Epstein-Zin parametrization

see Epstein and Zin (1989) and Weil (1990). Then utility takes values in $\mathfrak{U} \triangleq (1-\gamma)\mathfrak{C}$, the *continuous-time Epstein-Zin aggregator* is given by

$$f : \mathfrak{C} \times \mathfrak{U} \to \mathbb{R}, \quad (c, \nu) \mapsto \delta \tfrac{1-\gamma}{1-\phi} \nu \left[\left(\frac{c}{((1-\gamma)\nu)^{\frac{1}{1-\gamma}}} \right)^{1-\frac{1}{\psi}} - 1 \right],$$

and terminal utility satisfies $U(x) \triangleq \frac{1}{1-\gamma}(\varepsilon x)^{1-\gamma}$, $x \in \mathfrak{C}$. The coefficients $\delta > 0$ and $\varepsilon > 0$ capture the agent's rate of time preference and weight on terminal consumption, respectively. Setting

$$\phi \triangleq 1/\psi, \quad \theta \triangleq \tfrac{1-\gamma}{1-\phi}, \quad \text{and} \quad q \triangleq \tfrac{\theta - 1}{\theta} = \tfrac{\phi - \gamma}{1-\gamma},$$

the aggregator f can be written as

$$f(c, \nu) \triangleq \tfrac{\delta}{1-\phi} c^{1-\phi} [(1-\gamma)\nu]^q - \delta\theta\nu. \tag{3.2}$$

In this chapter, we mostly focus on

$$\gamma, \psi > 1, \quad \text{so that} \quad \phi < 1, \quad \theta < 0, \quad \text{and} \quad q > 1. \tag{3.3}$$

Then (3.2) shows that f is well-defined on $\mathfrak{C} \times (\mathfrak{U} \cup \{0\}) = (0, \infty) \times (-\infty, 0]$, and the relevant class of utility processes is defined via

$$\mathcal{V} \triangleq \{V \in \mathcal{S}^\cap : (1-\gamma)V \geqslant 0\}, \qquad \text{where } \mathcal{S}^\cap \triangleq \textstyle\bigcap_{\beta > 1} \mathcal{S}^\beta. \tag{3.4}$$

Note: To establish existence and uniqueness for the stochastic differential utility process, it will be convenient to work with $\mathcal{V}$ as defined above, instead of the intuitively more appealing $\{V \in \mathcal{S}^{\cap} : (1-\gamma)V > 0\}$; however, one can verify *ex post* that $(1-\gamma)V > 0$, see Proposition 3.1.

Remark. For the parameterization as in (3.3), one may set $\varepsilon \triangleq +\infty$ to produce a zero weight on bequest utility (in this case $\varepsilon^{1-\gamma} = 0$). This leads to trivial preferences, because then the utility process $V = V(c) \in \mathcal{V}$ of every consumption plan is given by $V_t = 0$, $t \in [0, T]$. ◇

3.2 THE MAIN RESULTS

In this section, we present our main results on the Epstein-Zin parameterization of stochastic differential utility with

$$\text{RRA} \quad \gamma > 1 \quad \text{and} \quad \text{EIS} \quad \psi > 1.$$

Our first main result establishes existence and uniqueness of utility processes and, in particular, of utility index functionals for continuous-time Epstein-Zin recursive utility.

THEOREM 3.2. *For every consumption plan $c \in \mathcal{C}$, there exists a unique semimartingale $V = V(c) \in \mathcal{V}$ that satisfies (3.1).* *Existence and uniqueness*

Since $f(\lambda c, \lambda^{1-\gamma} v) = \lambda^{1-\gamma} f(c, v)$ for all $\lambda > 0$, the uniqueness part of Theorem 3.2 yields:

COROLLARY 3.3. *The mapping $\mathcal{C} \to \mathcal{V}$, $c \mapsto V(c)$ is homothetic, i.e., we have $V(\lambda c) = \lambda^{1-\gamma} V(c)$ for every consumption plan $c \in \mathcal{C}$ and every $\lambda > 0$.* *Homotheticity*

As utility processes are defined implicitly, in terms of the BSDE (3.1), and the aggregator f fails to be either concave or monotone with respect to v, it is not clear a priori whether the implied utility index functional v is concave or monotone.

Our second main result guarantees these crucial properties.

THEOREM 3.4. *The mapping* *Monotonicity and concavity*

$$\mathcal{C} \to \mathcal{V}, \quad c \mapsto V(c)$$

is concave and increasing, i.e., for all $c, \bar{c} \in \mathcal{C}$ and $\lambda \in [0, 1]$, we have

$$V(\lambda c + (1-\lambda)\bar{c}) \geqslant \lambda V(c) + (1-\lambda) V(\bar{c}) \quad \textit{and} \quad c \preccurlyeq \bar{c} \implies V(c) \leqslant V(\bar{c})$$

In particular, the utility index functional $v : \mathcal{C} \to \mathbb{R}$, $c \mapsto v(c) = V_0(c)$ is increasing and concave.

Finally, our third main result provides a general utility gradient inequality for Epstein-Zin recursive utility in continuous time.

THEOREM 3.5. *Suppose that $\hat{c} \in \mathcal{C}$ satisfies* *Utility gradient inequality*

$$E\left[\int_0^T f_c(\hat{c}_t, V_t(\hat{c}))\hat{c}_t dt + U'(\hat{c}_T)\hat{c}_T\right] < \infty.$$

Then, the following utility gradient inequality holds:

For every $c \in \mathcal{C}$ *and all* $t \in [0, T]$*, we have*

$$V_t(c) \leqslant V_t(\hat{c}) + \langle m^t(\hat{c}), c - \hat{c} \rangle_t,$$

where $\langle m, y \rangle_t \triangleq E_t[\int_t^T m_s y_s ds + m_T y_T]$ *and the time-*t *utility gradient* $m^t(\hat{c}) \triangleq \hat{m}^t$ *is given by*

$$\hat{m}_s^t \triangleq e^{\int_t^s f_v(\hat{c}_\tau, V_\tau(\hat{c}))d\tau}\big(f_c(\hat{c}_s, V_s(\hat{c}))1_{[0,T)}(s) + U'(\hat{c}_T)1_{\{T\}}(s)\big), \quad s \in [t, T].$$

In particular, with $\hat{m} \triangleq \hat{m}^0$*, the utility index* $\nu : \mathcal{C} \to \mathbb{R}$*,* $c \mapsto \nu(c)$ *satisfies*

$$\nu(c) \leqslant \nu(\hat{c}) + \langle \hat{m}, c - \hat{c} \rangle \quad \textit{for all } c \in \mathcal{C}.$$

The above utility gradient will be used in Chapter 4 to prove a verification theorem for a consumption-portfolio optimization problem of an agent with Epstein-Zin utility. For a further discussion of utility gradients and their applications, we refer to the bibliographical notes below and to Duffie and Skiadas (1994).

Bibliographical notes

Recursive utility has been developed in discrete time in the seminal contributions of Kreps and Porteus (1978, 1979), Epstein and Zin (1989) and Weil (1990), as a model of dynamic risk preferences that permits risk attitudes to be disentangled from the elasticity of intertemporal substitution; its continuous-time version, stochastic differential utility, was introduced by Duffie and Epstein (1992b). Kraft and Seifried (2014) show that discrete-time recursive utility converges to stochastic differential utility in the continuous-time limit of vanishing grid size and thus provide a solid mathematical link between the two concepts. The *Epstein-Zin* parametrization of recursive utility can be regarded as a non-additive generalization of the classical discounted expected utility paradigm with a power utility function, where the standard setting is subsumed as the special case $\gamma\psi = 1$.

As shown by Skiadas (1998), agents with $\gamma, \psi > 1$ exhibit a preference for information, i.e., they prefer early resolution of uncertainty to late resolution. This specification is important in a number of both theoretical and empirical applications, including the literature on asset pricing with long-run risk that was initiated by Bansal and Yaron (2004). Despite their widespread usage, fundamental questions like existence, uniqueness and concavity of stochastic differential utility have so far remained unresolved for these parameters.

General existence and uniqueness results for recursive utility are provided by Marinacci and Montrucchio (2010) in discrete time, and by Duffie and Epstein (1992b) and Ma (2000) in continuous time; however, these continuous-time results rely on global Lipschitz conditions that are violated for the Epstein-Zin specification. Schroder and Skiadas (1999) prove existence and uniqueness of Epstein-Zin utility in a Brownian framework for $\gamma > 1$ and $\psi \leqslant 1$ and parameterizations with $\gamma < 1$. Xing (2015) also addresses existence of Epstein-Zin utility in a Brownian framework.

Utility gradients (also known as superdifferentials or supergradient densities) have become an indispensable tool in the analysis of both optimal portfolio allocations (see, e.g., Schroder and Skiadas (1999, 2003, 2008), Bank and Riedel (2001a), Kallsen and Muhle-Karbe (2010), Skiadas (2008, 2013), and the references therein) and equilibrium asset pricing (see, e.g., Duffie and Epstein (1992a), Duffie et al. (1994), Bank and Riedel (2001b), Chen and Epstein (2002), Epstein and Ji (2013) as well as Campbell (2003) and the references therein). The reason for this is the far-reaching insight that the first-order optimality condition in the maximization of a utility functional can be formulated as a martingale property of prices, after they have been deflated by the associated utility gradient; see Duffie and Skiadas (1994) and Harrison and Kreps (1979). From a mathematical perspective, utility gradients are intimately related to the stochastic maximum principle in intertemporal optimization problems, where the agent's utility gradient appears naturally as the minimizer in the associated dual problem; see, e.g., Cox and Huang (1989), Karatzas et al. (1991), Kramkov and Schachermayer (1999), El Karoui et al. (2001), and Levental et al. (2013). In the literature, utility gradients for continuous-time recursive utility have been derived by Duffie and Skiadas (1994), and proven in a Brownian framework by Schroder and Skiadas (1999).

3.3 EPSTEIN-ZIN UTILITY FOR BOUNDED CONSUMPTION PLANS

In this section, we focus on bounded consumption plans. Accordingly, we do *not* impose the parameter restrictions (3.3) and allow *for all* non-unit values of relative risk aversion and elasticity of intertemporal substitution

$$\gamma > 0,\ \gamma \neq 1, \quad \text{and} \quad \psi > 0,\ \psi \neq 1.$$

We say that a consumption plan $c \in \mathcal{C}$ is *bounded* if

$$k_0 \preccurlyeq c \preccurlyeq k_1 \quad \text{for some constans } k_0, k_1 > 0,$$

and we denote the class of all bounded consumption plans by $\mathcal{C}_b$. Correspondingly, we refer to a semimartingale V as $\mathfrak{U}$-bounded if it takes values in a compact subset of $\mathfrak{U}$, and we write $\mathcal{V}_b$ for the collection of all $\mathfrak{U}$-bounded semimartingales. By a *bounded utility process associated with a consumption plan* $c \in \mathcal{C}$, we mean a $\mathfrak{U}$-bounded semimartingale $V \in \mathcal{V}_b$ which satisfies

$$V_t = E_t\left[\int_t^T f(c_s, V_s)ds + U(c_T)\right], \qquad t \in [0, T]. \tag{3.5}$$

Our first goal is the construction of a well-behaved utility process functional $\mathbf{V} : \mathcal{C}_b \to \mathcal{V}_b$ mapping bounded consumption plans to their associated bounded utility process.

3.3.1 *Existence of utility for bounded consumption plans*

We show that every bounded consumption plan admits one and only one bounded utility process. To this end, for each $n \in \mathbb{N}$, we consider the truncated Epstein-Zin aggregator

$$f^n(c, v) \triangleq f\big(c, (u_\gamma(1/n) \vee v) \wedge u_\gamma(n)\big), \tag{3.6}$$

where $u_\gamma(x) \triangleq \frac{1}{1-\gamma}x^{1-\gamma}$, $x \in \mathfrak{C}$; note that $u_\gamma(\mathfrak{C}) = \mathfrak{U}$.

LEMMA 3.6. *The pair* $(f^n(c, \cdot), U(c_T))$ *is a BSDE*^{2}*-standard parameter for every bounded consumption plan* $c \in \mathcal{C}_b$.

Proof. The Epstein-Zin aggregator is differentiable; hence, it is Lipschitz continuous and bounded on any compact subset of its domain, and thus all conditions imposed by Definition 2.68 are satisfied. □

COROLLARY 3.7. *For every bounded consumption plan* $c \in \mathcal{C}_b$ *and each* $n \in \mathbb{N}$, *there exists a unique* $V^n \triangleq V^n(c) \in \mathcal{S}^2$ *satisfying*

$$V_t^n = E_t\left[\int_t^T f^n(c_s, V_s^n)ds + U(c_T)\right], \qquad t \in [0, T]. \tag{3.7}$$

Proof. In view of Lemma 3.6, this is guaranteed by the existence and uniqueness result of Theorem 2.74. □

Uniqueness for bounded plans

LEMMA 3.8. *If* V *is a bounded utility process associated with* $c \in \mathcal{C}$, *then* V *satisfies* (3.7) *for all but finitely many* $n \in \mathbb{N}$. *In particular, every* $c \in \mathcal{C}_b$ *admits at most one bounded utility process.*

Proof. Let V be a bounded utility process associated with $c \in \mathcal{C}$. Then $u_\gamma(1/n) \leqslant V \leqslant u_\gamma(n)$ for all but finitely many $n \in \mathbb{N}$, and hence

$$f(c, V_t) = f^n(c_t, V_t) \quad \text{for } dt\text{-a.e. } t \in [0, T],$$

so that (3.5) implies that V satisfies (3.7). If additionally $c \in \mathcal{C}_b$, then V coincides with the unique solution V^n of (3.7) given by Corollary 3.7; thus if $\bar{V}$ is another bounded utility process associated with c, then we have $V = V^n = \bar{V}^n$ for all but finitely many $n \in \mathbb{N}$. □

As a next step, we establish the converse for bounded consumption plans: If c is a bounded consumption plan, then the unique solution V^n of (3.7) is, in fact, a solution of (3.5) and thus a utility process associated with c.

Existence for bounded plans

LEMMA 3.9. *Let* $c \in \mathcal{C}_b$ *be a bounded consumption plan. For each* $n \in \mathbb{N}$, *denote by* V^n *the unique solution of* (3.7), *with* f^n *given by* (3.6). *Then there exist constants* $\ell, m \in \mathfrak{U}$ *such that* $\ell \leqslant V^n \leqslant m$ *for all but finitely many* $n \in \mathbb{N}$. *In particular,* $V^{n+1} = V^n$ *for all but finitely* $n \in \mathbb{N}$, *and* $V^c \triangleq \lim_{n\to\infty} V^n$ *is a bounded utility process associated with* c.

Proof. A simple calculation shows that $f(k, u_\gamma(k)) = 0$ for all $k > 0$, and hence $f^n(k, u_\gamma(k)) = 0$ for all $k \in [1/n, n]$ by (3.6). We choose $k_1 > k_0 > 0$ such that

$$k_0 \preccurlyeq c \preccurlyeq k_1 \quad \text{and} \quad k_0 \preccurlyeq \varepsilon c \preccurlyeq k_1$$

and take $n_0 \in \mathbb{N}$ with $1/n_0 < k_0$ and $n_0 > k_1$.

Then, for all $n \geqslant n_0$, the constant processes $\ell \triangleq u_\gamma(k_0)$ and $m \triangleq u_\gamma(k_1)$ are the unique solutions of the BSDEs associated to the parameters $(f^n(k_0, \cdot), u_\gamma(k_0))$ and $(f^n(k_1, \cdot), u_\gamma(k_1))$, respectively.

By Lemma B.1 (p. 171), the Epstein-Zin aggregator is increasing in c, and hence we have

$$f(k_0, Y_t) \leqslant f(c_t, Y_t) \leqslant f(k_1, Y_t) \quad \text{for } dt\text{-a.e. } t \in [0, T]$$

and every semimartingale Y. Using the above inequality with $Y = (u_\gamma(1/n) \vee V^n) \wedge u_\gamma(n)$, we get

$$f^n(k_0, V_t^n) \leqslant f^n(c_t, V_t^n) \leqslant f^n(k_1, V_t^n) \quad \text{for } dt\text{-a.e. } t \in [0, T]$$

and all $n \geqslant n_0$. Since $u_\gamma(k_0) \leqslant u_\gamma(\varepsilon c_T) = V_T^n \leqslant u_\gamma(k_1)$, the comparison theorem for BSDEs (Theorem 2.76) implies $\ell \leqslant V^n \leqslant m$ for all $n \geqslant n_0$. In particular, we have

$$f^n(c_t, V_t^n) = f(c, V_t^n) = f^{n_0}(c_t, V_t^n) \quad \text{for } dt\text{-a.e. } t \in [0, T],\ n \geqslant n_0,$$

and hence $V^n = V^{n_0}$ for all $n \geqslant n_0$ by the uniqueness part of Corollary 3.7; the semimartingale $V^c \triangleq V^{n_0} = \lim_{n \to \infty} V^n$ satisfies the BSDE (3.5) as well as $\ell \leqslant V \leqslant m$. □

COROLLARY 3.10. *Let $c \in \mathcal{C}_b$ be a bounded consumption plan. There exists precisely one bounded utility process V associated with c.*

At this point, we have successfully constructed a mapping

$$\mathbf{V} : \mathcal{C}_b \to \mathcal{V}_b, \quad c \mapsto V^c, \quad \text{where } V_t^c = E_t\left[\int_t^T f(c_s, V_s^c) ds + U(c_T)\right]$$

is the unique bounded utility process associated with c.

3.3.2 *Some properties of utility for bounded consumption plans*

First, we show that the utility process functional is increasing, which is a direct consequence of the comparison theorem for BSDEs.

PROPOSITION 3.11. *The mapping $\mathbf{V} : \mathcal{C}_b \to \mathcal{V}_b$ is monotone, i.e.,* *Monotonicity*

$$\mathbf{V}(c^1) \leqslant \mathbf{V}(c^2) \quad \textit{whenever} \quad c^1 \preccurlyeq c^2.$$

Proof. Let $c^1, c^2 \in \mathcal{C}_b$ and let $V^1 \triangleq \mathbf{V}(c^1)$ and $V^2 \triangleq \mathbf{V}(c^2)$ denote the associated bounded utility processes. Lemma 3.9 yields $n \in \mathbb{N}$ such that V^1 and V^2 solve the BSDEs

$$dV_t^i = -f^n(c_s^i, V_s^i) dt + dM_t^i, \quad V_T^i = U(c_T^i), \qquad i = 1, 2,$$

with the truncated Epstein-Zin aggregator f^n given by (3.6).

Lemma B.1 shows that the Epstein-Zin aggregator is increasing in c and hence $c^1 \preccurlyeq c^2$ implies

$$f^n(c_t^1, V_t^2) \leqslant f^n(c_t^2, V_t^2) \quad \text{for } dt\text{-a.e. } t \in [0,T] \quad \text{and} \quad U(c_T^1) \leqslant U(c_T^2).$$

Recalling that $(f^n, U(c_T^i))$, $i = 1,2$, are BSDE^2-standard parameters by Lemma 3.6, the comparison theorem for BSDEs (Theorem 2.76) implies that $V^1 \leqslant V^2$. □

Next, we show how additive CRRA utilities yield natural bounds for stochastic differential utilities. For $1 \neq \rho > 0$, let $u_\rho(c) = \frac{1}{1-\rho}c^{1-\rho}$ denote the corresponding CRRA utility function. For every consumption plan $c \in \mathcal{C}_b$ and $\rho \in \{\gamma, \phi\}$, we consider the additive utility

$$\mathbf{U}_\rho : \mathcal{C}_b \to \mathcal{S}_b, \quad c \mapsto u_\gamma \circ u_\rho^{-1}(Y^\rho(c)),$$

where $Y^\rho(c) = Y^\rho$ is given by

$$Y_t^\rho = e^{\delta t} E_t\big[\int_t^T \delta e^{-\delta s} u_\rho(c_s) ds + e^{-\delta T} u_\rho(\varepsilon c_T)\big]. \tag{3.8}$$

Thus $\mathbf{U}_\rho(c)$ represents additive power utility with parameter ρ, transformed onto a γ-power utility scale. Intuitively, we would expect that

$$\mathbf{U}_{\gamma\vee\phi}(c) \leqslant \mathbf{V}(c) \leqslant \mathbf{U}_{\gamma\wedge\phi}(c). \tag{3.9}$$

In the following, we confirm that this is indeed the case:

Let $c \in \mathcal{C}_b$ be a bounded consumption plan and put

$$Y^\phi \triangleq u_\gamma^{-1} \circ u_\phi(\mathbf{U}_\phi(c)) \quad \text{as well as} \quad Y^\gamma \triangleq \mathbf{U}_\gamma(c),$$

i.e., Y^ϕ and Y^γ satsify (3.8) with $\rho = \phi$ and $\rho = \gamma$, respectively. The explicit representation result for linear BSDEs (Proposition 2.59) shows that the processes Y^ϕ and Y^γ satisfy the BSDEs

$$\begin{aligned} dY_t^\phi &= -\delta\big[u_\phi(c_t) - Y_t^\phi\big]dt + dM_t^\phi, \quad && Y_T^\phi = u_\phi(\varepsilon c_T), \quad \text{and} \\ dY_t^\gamma &= -\delta\big[u_\gamma(c_t) - Y_t^\gamma\big]dt + dM_t^\gamma, \quad && Y_T^\gamma = u_\gamma(\varepsilon c_T). \end{aligned} \tag{3.10}$$

Moreover, let $V \triangleq \mathbf{V}(c) \in \mathcal{V}_b$ and recall that V solves the BSDE

$$dV_t = -f(c_t, V_t)dt + dM_t, \qquad V_T = U(c_T) = u_\gamma(\varepsilon c_T),$$

where $M_t = E_t[\int_0^T f(c_s, V_s)ds + U(c_T)]$ is a bounded martingale.

One of the inequalities in (3.9) is a direct consequence of elementary properties of the Epstein-Zin aggregator and the comparison theorem for BSDEs.

LEMMA 3.12. *If $\gamma \geqslant \phi$, then $\mathbf{U}_\gamma(c) \leqslant V$, and $V \leqslant \mathbf{U}_\gamma(c)$ if $\gamma \leqslant \phi$.*

Proof. With $g(t,\nu) \triangleq \delta[u_\gamma(c_t) - \nu]$, the pair $(g, u_\gamma(\varepsilon c_T))$ is a BSDE^2-standard parameter. Moreover, by Lemma B.3, the Epstein-Zin aggregator satisfies

$$\begin{aligned} f(c_t, V_t) &\geqslant \delta[u_\gamma(c_t) - V_t] \quad \text{if} \quad \gamma \geqslant \phi, \quad \text{and} \\ f(c_t, V_t) &\leqslant \delta[u_\gamma(c_t) - V_t] \quad \text{if} \quad \gamma \leqslant \phi \end{aligned}$$

The comparison theorem for BSDEs (Theorem 2.76) yields

$$Y^\gamma \leqslant V \quad \text{if} \quad \gamma \geqslant \phi \qquad \text{and} \qquad V \leqslant Y^\gamma \quad \text{if} \quad \gamma \leqslant \phi,$$

since $V_T = U(c_T) = u_\gamma(\varepsilon c_T) = Y^\gamma_T$. □

The other inequality requires more work: A calculation using Itō's formula shows that the process $Y \triangleq u_\phi \circ u_\gamma^{-1}(V)$ has dynamics

$$\begin{aligned} dY_t = -\left(\delta\left[u_\phi(c_t) - Y_t\right]dt + dA_t\right) + \left(u_\phi \circ u_\gamma^{-1}\right)'(V_{t-})dM_t, \\ \text{where } A \text{ is increasing if } \phi \geqslant \gamma \text{ and decreasing if } \phi \leqslant \gamma. \end{aligned} \tag{3.11}$$

A detailed proof is provided in the appendix (Lemma B.4, p. 173ff.).

LEMMA 3.13. *If $\gamma \geqslant \phi$, then $V \leqslant \mathbf{U}_\gamma(c)$, and $V \geqslant \mathbf{U}_\gamma(c)$, if $\gamma \leqslant \phi$.*

Proof. The claim is equivalent to $Y \geqslant Y^\phi$ if $\phi \geqslant \gamma$ and $Y \leqslant Y^\phi$ if $\phi \leqslant \gamma$. The dynamics (3.11) show that Y is (in the terminology of BSDE-theory, see, e.g., Peng (1999)) a supersolution (subsolution) of the BSDE (3.10) if $\phi \geqslant \gamma$ (if $\phi \leqslant \gamma$); indeed, $(u_\phi \circ u_\gamma^{-1})'(V_{t-})$ is bounded since $V \in \mathcal{V}_b$, and hence the stochastic integral $N \triangleq \int_0^\cdot (u_\phi \circ u_\gamma^{-1})'(V_{t-})dM_t$ is still a martingale. Now, Corollary 3.23 below shows that supersolutions (subsolutions) Y of (3.10) are bigger (smaller) than the solution Y^ϕ, and the proof is complete. □

Combining the above results we obtain

COROLLARY 3.14. *For every bounded consumption plan $c \in \mathcal{C}_b$, we have*

Power utility bounds

$$\mathbf{U}_{\gamma \vee \phi}(c) \leqslant \mathbf{V}(c) \leqslant \mathbf{U}_{\gamma \wedge \phi}(c). \tag{3.9}$$

The crucial step in the proof of Lemma 3.13 makes use of a refined comparison result for BSDEs. Its proof, which will be given at the end of the next section, is based on a stochastic Gronwall inequality. Such stochastic Gronwall inequalites will play an important role in the following: They yield a comparison principle which allows us to extend the utility process functional $\mathbf{V} : \mathcal{C}_b \to \mathcal{S}_b$ to a mapping $\mathcal{C} \to \mathcal{V}$ defined for all consumption plans. Thus, we interrupt our investigation of stochastic differential utility to provide an excursion on stochastic Gronwall inequalities.

3.4 STOCHASTIC GRONWALL INEQUALITIES

A stochastic Gronwall inequality is a result which guarantees that a process X, which satisfies a conditional linear integral inequality of the form

$$X_t \leqslant E_t\left[\int_t^T (\alpha_s X_s + H_s)ds + Z\right] \qquad \text{for all } t \in [0,T], \tag{G}$$

is bounded above by the solution of the corresponding linear integral equation, that is, X satisfies

$$X_t \leqslant E_t\left[\int_t^T e^{\int_t^s \alpha_u du} H_s ds + e^{\int_t^T \alpha_u du} Z\right] \qquad \text{for all } t \in [0,T].$$

Stochastic Gronwall inequalities are of great importance in the theory of backward stochastic differential equations (see Antonelli (1996)) and stochastic differential utility (see Duffie and Epstein (1992b) and Schroder and Skiadas (1999)), and they will be used frequently in the remainder of this thesis.

In this section, we provide a general stochastic Gronwall inequality under weak integrability conditions in a semimartingale framework. Related results can be found in Duffie and Epstein (1992b), Antonelli (1996), Schroder and Skiadas (1999) and Kraft and Seifried (2010). While Duffie and Epstein (1992b) and Schroder and Skiadas (1999) assume a continuous filtration, Kraft and Seifried (2010) do allow for general filtrations, but they only consider homogeneous linear integral inequalities with a constant coefficient. The most general result can be found in Antonelli (1996), where integral inequalities with bounded increasing processes are considered. We restrict ourselves to processes which are absolutely continuous with respect to the Lebesgue measure and impose weak integrability conditions on the density process. Under these integrability conditions, the result is, to the best of our knowledge, new and not contained in the literature.

3.4.1 *A general theorem and some ramifications*

As above, $(\Omega, \mathcal{F}_T, P)$ is a probability space endowed with a right-continuous and complete filtration $(\mathcal{F}_t)_{t\geqslant 0}$; whenever a martingale appears, we work with a càdlàg version.

For the linear integral inequality

$$X_t \leqslant E_t\left[\int_t^T (\alpha_s X_s + H_s)ds + Z\right] \qquad \text{for all } t \in [0,T], \tag{G}$$

to yield a meaningful upper bound, we should at least require that the quantities on the right-hand side are integrable. Moreover, α should be integrable as the following simple deterministic example shows: With $\alpha_t = (T-t)^{-1}$, $X_t = T-t$, $H = 0$ and $Z = 0$, we have $X_t = T - t = \int_t^T \alpha_s X_s ds$, but $X_t \nleqslant 0$. In the following, we thus restrict ourselves to tuples (α, X, H, Z) such that

- $\alpha \geqslant 0$ is a progressively measurable process with $\int_0^T \alpha_s ds < \infty$,
- X and H are progressive processes such that $\alpha X + H$ is integrable,
- Z is an integrable random variable.

For brevity, a tuple (α, X, H, Z) satisfying the above will be referred to as a Gronwall parameter. For a fixed Gronwall parameter (α, X, H, Z) and $0 \leqslant t \leqslant s \leqslant T$ and $n \in \mathbb{N}$, we set

$$I_s^{t,n} \triangleq \tfrac{1}{n!}\left(\int_t^s \alpha_u du\right)^n \quad \text{and} \quad \mathcal{E}_s^{t,n} \triangleq \sum_{k=0}^n I_s^{t,k},$$

and we formulate the following statements: For all $t \in [0, T]$,

$$E\left[\int_t^T (\mathcal{E}_s^{t,n}|H_s| + I_s^{t,n}\alpha_s|X_s|)ds + \mathcal{E}_T^{t,n}|Z|\right] < \infty \quad \text{for all } n \in \mathbb{N}, \tag{I}$$

$$\begin{aligned} \left(\mathcal{E}_T^{t,n} Z\right)_{n\in\mathbb{N}} &\quad \text{is uniformly P-integrable, and} \\ \left(\left(\mathcal{E}_s^{t,n}H_s + \tfrac{I_s^{t,n}}{n!}\alpha_s X_s\right)_{s\in[t,T]}\right)_{n\in\mathbb{N}} &\quad \text{is uniformly } P \otimes dt\text{-integrable.} \end{aligned} \tag{UI}$$

Note: Of course, it suffices to require (I) and the first condition in (UI) merely for $t = 0$; then, the statements for $t > 0$ hold a fortiori. We have opted for the above formulation for the sake of transparency.

We will prove the following general Gronwall inequality:

THEOREM 3.15. *Let (α, X, H, Z) be a Gronwall parameter such that* (I) *and* (UI) *are satisfied, and suppose that*

$$X_t \leqslant E_t\left[\int_t^T (\alpha_s X_s + H_s)ds + Z\right] \quad \textit{for all } t \in [0, T]. \tag{G}$$

Then X satisfies the inequality

$$X_t \leqslant E_t\left[\int_t^T e^{\int_t^s \alpha_u du} H_s ds + e^{\int_t^T \alpha_u du} Z\right] \quad \textit{for all } t \in [0, T].$$

Since $I_s^{t,n}/n! \to 0$ and $\mathcal{E}_s^{t,n} \to e^{\int_t^s \alpha_u du}$ as $n \to \infty$ for all $0 \leqslant t \leqslant s \leqslant T$, the abstract integrability condition (UI) implies the L^1-convergence

$$\int_t^T (\mathcal{E}_s^{t,n}H_s + \tfrac{I_s^{t,n}}{n!}\alpha_s X_s)ds + \mathcal{E}_T^{t,n} Z \to \int_t^T e^{\int_t^s \alpha_u du} H_s ds + e^{\int_t^T \alpha_u du} Z. \tag{3.12}$$

Thus we prove Theorem 3.15 by iterating the linear integral inequality (G) to obtain

$$X_t \leqslant E_t\left[\int_t^T \left(\mathcal{E}_s^{t,n}H_s + \tfrac{I_s^{t,n}}{n!}\alpha_s X_s\right) ds + \mathcal{E}_T^{t,n} Z\right], \quad n \in \mathbb{N}. \tag{3.13}$$

Then the claim follows by (3.12), upon sending $n \to \infty$.

Before we address the proof of Theorem 3.15, we discuss the integrability assumptions (I) and (UI) and provide sufficient conditions. First, requiring (I) is necessary for (3.13) to be a meaningful upper bound. Second, with (I) in place, the uniform integrability assumptions in (UI) are equivalent to the $L^1(P)$-convergence in (3.12). However, both (I) and (UI) are abstract conditions that are difficult to verify directly. In the following, we give some simple sufficient conditions and record the corresponding corollaries of Theorem 3.15.

First, we note that

$$\mathcal{E}_s^{t,n}|H_s| \leqslant e^{\int_t^s \alpha_u du}|H_s| \leqslant e^{\int_0^s \alpha_u du}|H_s|, \quad s \in [t,T].$$

Hence $\{(\mathcal{E}_s^{t,n}H_s)_{s\in[t,T]}\}$ is uniformly integrable for all $t \in [0,T]$, if the process $e^{\int_0^\cdot \alpha_u du}H$ is integrable. The same argument applies to $\mathcal{E}_T^{t,n}|Z|$ and shows that $\{\mathcal{E}_T^{t,n}|Z| : n \in \mathbb{N}\}$ is uniformly integrable for all $t \in [0,T]$ provided that $e^{\int_0^T \alpha_u du}Z$ is an integrable random variable; thus, the following conditions (EI) and (UI′) imply (I) and (UI):

$$E\left[\int_0^T e^{\int_0^t \alpha_u du}|H_t|dt + e^{\int_0^T \alpha_u du}|Z|\right] < \infty, \tag{EI}$$

$$\left(\left(\int_0^\cdot \alpha_s ds\right)^n \frac{\alpha X}{n!}\right)_{n\in\mathbb{N}} \quad \text{is uniformly } P\otimes dt\text{-integrable.} \tag{UI′}$$

A sufficient condition for (EI) is the (exponential) moment condition

$$E\left[e^{p\int_0^T \alpha_u du}\right] < \infty \quad \text{and} \quad E\left[\int_0^T |H_s|^q ds + |Z|^q\right] < \infty \quad \text{for} \quad \frac{1}{p}+\frac{1}{q}=1. \tag{EM}$$

If additionally $E[\int_0^T |\alpha_s X_s|^q ds] < \infty$, then (EM) implies (UI′), since

$$\frac{1}{n!}\left(\int_0^t \alpha_s ds\right)^n \alpha_t X_t \leqslant e^{\int_0^t \alpha_s ds}|\alpha_t X_t|, \quad t \in [0,T], \; n \in \mathbb{N}.$$

COROLLARY 3.16. *Let α be a non-negative progressive process, and let X and H be progressive processes such that*

$$e^{\int_0^T \alpha_u du} \in L^p(P) \quad \textit{and} \quad \alpha X, H \in L^q(P\otimes dt)$$

for $p, q \in [1,\infty]$ with $1/p + 1/q = 1$; moreover, let $Z \in L^q(P)$. If X satisfies the linear integral inequality (G)*, i.e., if*

$$X_t \leqslant E_t\left[\int_t^T (\alpha_s X_s + H_s)ds + Z\right] \quad \textit{for all } t \in [0,T],$$

then, it follows that

$$X_t \leqslant E_t\left[\int_t^T e^{\int_t^s \alpha_u du}H_s ds + e^{\int_t^T \alpha_u du}Z\right] \quad \textit{for all } t \in [0,T].$$

Proof. Clearly, (α, X, H, Z) is a Gronwall parameter. By the discussion preceding this corollary, the Gronwall parameter (α, X, H, Z) satisfies (I) and (UI), and the result is implied by Theorem 3.15. □

Of course, (EI) is automatically fulfilled if $H = 0$ and $Z = 0$, and only (UI′) is needed.

COROLLARY 3.17. *Let $0 \leqslant \alpha$ and X be progressive processes and suppose that*

$$\left(\frac{1}{n!}\left(\int_0^\cdot \alpha_s ds\right)^n \alpha X\right)_{n\in\mathbb{N}} \quad \textit{is uniformly } P\otimes dt\textit{-integrable.} \tag{UI′}$$

If $\int_0^T \alpha_s ds < \infty$, and if X satisfies the homogeneous linear integral inequality

$$X_t \leqslant E_t\left[\int_t^T \alpha_s X_s ds\right], \qquad t \in [0,T],$$

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

As indicated above, the proof of Theorem 3.15 relies on iterating the linear integral inequality (G). This procedure yields the conclusion of Lemma 3.18. While the formal calculations are easy, some work is required to ensure that the involved quantities are sufficiently measurable and integrable.

Iterating the inequality

LEMMA 3.18. *Under the assumptions of Theorem 3.15, we have*

$$X_t \leqslant E_t \left[\int_t^T \left(\mathcal{E}_s^{t,n} H_s + \frac{I_s^{t,n}}{n!} \alpha_s X_s \right) ds + \mathcal{E}_T^{t,n} Z \right], \quad t \in [0,T],\ n \in \mathbb{N}.$$

Proof. See Lemma B.14, p. 178ff. □

Once Lemma 3.18 is in place, establishing Theorem 3.15 is a simple matter of sending $n \to \infty$. For completeness, we explicitly record the proof.

Proof of Theorem 3.15. By Lemma 3.18 we have

$$X_t \leqslant E_t \left[\int_t^T \left(\mathcal{E}_s^{t,n} H_s + I_s^{t,n} \alpha_s X_s \right) ds + \mathcal{E}_T^{t,n} Z \right], \quad t \in [0,T]$$

for all $n \in \mathbb{N}$. Since $\int_0^T \alpha_s ds < \infty$, for all $0 \leqslant t \leqslant s \leqslant T$, we get

$$I_s^{t,n} = \tfrac{1}{n!} \left(\int_t^s \alpha_u du \right)^n \to 0 \quad \text{and} \quad \mathcal{E}_s^{t,n} = \sum_{k=0}^n I_s^{t,n} \to e^{\int_t^s \alpha_u du}. \tag{3.14}$$

Together with the assumed uniform integrability (UI) of

$$\left(\left(\mathcal{E}_{t,s}^n H_s + \frac{I_{t,s}^n}{n!} \alpha_s X_s \right)_{s \in [t,T]} \right)_{n \in \mathbb{N}} \quad \text{and} \quad \left(\mathcal{E}_{t,T}^n Z \right)_{n \in \mathbb{N}},$$

the convergence statement (3.14) implies

$$\int_t^T \left(\mathcal{E}_s^{t,n} H_s + I_s^{t,n} \alpha_s X_s \right) ds + \mathcal{E}_T^{t,n} Z \to \int_t^T e^{\int_t^s \alpha_u du} H_s ds + e^{\int_t^T \alpha_u du} Z$$

in $L^1(P)$, and hence we obtain

$$X_t \leqslant E_t \left[\int_t^T e^{\int_t^s \alpha_u du} H_s ds + e^{\int_t^T \alpha_u du} Z \right], \quad \text{for all } t \in [0,T]. \qquad \square$$

3.4.2 *Stopping time extensions*

Using stopping arguments, it is possible to drop the positivity assumption on α in the general Gronwall inequality in Theorem 3.15. The price to pay is that the linear integral inequality (G) must be required to hold for all stopping times.

We begin with the homogeneous case; closely related arguments are employed by Schroder and Skiadas (1999).

THEOREM 3.19. *Let α be a progressive process with $\int_0^T \alpha_s^+ ds < \infty$, and let $(X_t)_{t \in [0,T]}$ be right-continuous and adapted. Suppose that the family*

$$\left(\tfrac{1}{n!} \left(\int_0^\cdot \alpha_s^+ ds \right)^n \alpha^+ X \right)_{n \in \mathbb{N}} \quad \textit{is uniformly } P \otimes dt\textit{-integrable.} \tag{UI$^+$}$$

If $X_T \leqslant 0$ *and for all stopping times* τ *with* $E[X_\tau^+] < \infty$ *we have*

$$1_{\{\tau>t\}}X_t \leqslant E_t\left[1_{\{\tau>t\}}\int_t^\tau \alpha_s X_s ds + 1_{\{\tau>t\}}X_\tau\right], \quad t \in [0,T], \tag{3.15}$$

then $X_t \leqslant 0$ *for all* $t \in [0,T]$ *a.s.*

Note: For a random variable ξ with $E[\xi^+] < \infty$, we define the conditional expectation as the $[-\infty,\infty)$-valued random variable $E[\xi \mid \mathcal{G}] = E[\xi^+ \mid \mathcal{G}] - \sup_{n\in\mathbb{N}} E[\xi^- \wedge n \mid \mathcal{G}]$. A version of the conditional expectation in (3.15) never takes the value $-\infty$ as X is a real-valued process.

Proof. Suppose by contradiction that the event $A \triangleq \{X_{t_0} > 0\} \in \mathcal{F}_{t_0}$ has positive probability for some $t_0 \in [0,T]$, and consider the stopping time $\tau \triangleq \inf\{t \geqslant t_0 : X_t \leqslant 0\} \leqslant T$. Now, (3.15) yields

$$1_{\{\tau>t\}}X_t \leqslant E_t\left[1_{\{\tau>t\}}\int_t^\tau \alpha_s X_s ds + 1_{\{\tau>t\}}X_\tau\right] \leqslant E_t\left[1_{\{\tau>t\}}\int_t^\tau \alpha_s^+ X_s ds\right]$$

for all $t \in [0,T]$, since $X > 0$ on $(t,\tau]$ and $X_\tau \leqslant 0$ by right-continuity; thus, the process $Y_t \triangleq 1_{\{\tau>t\}}X_t$ satisfies

$$Y_t \leqslant E_t\left[\int_t^T \alpha_s^+ Y_s ds\right] \quad \text{for all } t \in [0,T].$$

Since α^+ is a non-negative progressive process with $\int_0^T \alpha_s^+ ds < \infty$ and

$$\tfrac{1}{n!}\left(\int_0^\cdot \alpha_s^+ ds\right)^n \alpha_t^+ |Y_t| \leqslant \tfrac{1}{n!}\left(\int_0^\cdot \alpha_s^+ ds\right)^n \alpha_t^+ |X_t|, \quad t \in [0,T],$$

(UI^+) implies that the assumptions of Corollary 3.17 are satisfied. But then

$$Y_{t_0} = 1_A X_{t_0} \leqslant 0, \quad \text{where} \quad A = \{X_{t_0} > 0\} \quad \text{with} \quad P(A) > 0,$$

which is a contradiction. □

The above proofs indicate that the method of iterating the inequality does not take stochastic Gronwall inequalities farther than the abstract uniform integrability condition (UI^+). One way to establish (UI^+) is to verify an exponential moment condition; however, under this condition, the result is known in the literature.

In the remainder of this thesis, we shall only rely on the two stochastic Gronwall inequalities given below. We derive them from Theorem 3.19; of course, they could alternatively be proven directly.

Standard homogeneous Gronwall inequality

PROPOSITION 3.20. *Let* $(\alpha)_{t\in[0,T]}$ *be progressively measurable and bounded above, and let* $(X_t)_{t\in[0,T]}$ *be a right-continuous and adapted process with*

$$E\left[|X_\tau|\right] < \infty \quad \textit{for all } [0,T]\textit{-valued stopping times } \tau.$$

If $X_T \leqslant 0$ *and, for every stopping time* τ*, we have*

$$1_{\{\tau>t\}}X_t \leqslant E_t\left[1_{\{\tau>t\}}\int_t^\tau \alpha_s X_s ds + 1_{\{\tau>t\}}X_\tau\right], \quad t \in [0,T],$$

then $X_t \leqslant 0$ *for all* $t \in [0,T]$ *a.s.*

Proof. This is a simple corollary of Theorem 3.19. □

Note: The condition $E[|X_\tau|] < \infty$ is not a serious requirement; usually, we will have $\sup\{E[|X_\tau|] : \tau \text{ is a } [0,T]\text{-valued stopping time}\} < \infty$ or even $E[\sup_{t\in[0,T]}|X_t|] < \infty$.

The "stopping time-version" of the stochastic Gronwall inequality is particularly useful in the context of BSDEs. Let us say that a semimartingale Y is a *subsolution* of the BSDE

$$dX_t = -f(t, X_t)dt + dM_t, \qquad X_T = \xi, \tag{3.16}$$

if there exists a martingale M^Y and a decreasing (right-continuous) process A^Y such that

$$dY_t = -f(t, Y_t)dt + dM^Y_t - dA^Y_t, \qquad Y_T \leqslant \xi, \tag{3.17}$$

and if (Y, M^Y, A^Y) satisfies

$$E\left[\int_0^T |f(s, Y_s)|ds + |Y_T| + |A^Y_T - A^Y_0|\right] < \infty. \tag{3.18}$$

Similarly, we call Y a *supersolution* of (3.16) if $Y_T \geqslant X_T = \xi$ and A^Y is increasing; these definitions are in line with standard terminology (see, e.g., Peng (1999)). Here, we do not require that (f, ξ) is a BSDE-standard parameter, but f is understood to be measurable, of course.

LEMMA 3.21. *If Y is a subsolution of the BSDE (3.16), then,*

$$\sup\{E[|Y_\tau|] : \tau \textit{ is a } [0,T]\textit{-valued stopping time}\} < \infty, \tag{3.19}$$

and, for all stopping times τ, we have

$$1_{\{\tau>t\}}Y_t \leqslant E_t\left[1_{\{\tau>t\}}\int_t^\tau f(s, Y_s)ds + 1_{\{\tau>t\}}Y_\tau\right], \quad t \in [0,T]. \tag{3.20}$$

If Y is a supersolution, then (3.20) holds with "$\geqslant$", and if Y is a solution, then (3.20) holds with "$=$".

Proof. Let $\sigma \leqslant \tau$ be stopping times. Integrating (3.17) from time σ to time τ, we get

$$Y_\sigma + \left(M^Y_\tau - M^Y_\sigma\right) = \int_\sigma^\tau f(s, Y_s)ds + \left(A^Y_\tau - A^Y_\sigma\right) + Y_\tau \tag{3.21}$$

Choosing $\tau = T$ and taking $\mathcal{F}_\sigma$-conditional expectations, we obtain

$$Y_\sigma = E_\sigma\left[\int_\sigma^T f(s, Y_s)ds + A^Y_T - A^Y_\sigma + Y_T\right],$$

and hence (3.19) follows from (3.18). The same argument works for supersolutions. For the second part, let $t \in [0,T]$ and choose $\sigma = t \wedge \tau$. Then (3.21) yields

$$Y_{t\wedge\tau} + \left(M^Y_\tau - M^Y_{t\wedge\tau}\right) \leqslant \int_{t\wedge\tau}^\tau f(s, Y_s)ds + Y_\tau,$$

since A^Y is decreasing. Taking time-t conditional expectations, and then multiplying by the indicator of the set $\{\tau > t\}$, we obtain

$$1_{\{\tau>t\}}Y_t \leqslant E_t\left[1_{\{\tau>t\}}\int_t^\tau f(s, Y_s)ds + 1_{\{\tau>t\}}Y_\tau\right].$$

For supersolutions A^Y is increasing, and we get "$\geqslant$" in (3.20). □

Standard inhomogeneous Gronwall inequality

PROPOSITION 3.22. *Let* α *be a bounded progressively measurable process and* $\xi \in L^1(P)$. *Let* $\beta \in L^1(P \otimes dt)$ *be a progressive process and let* $(X_t)_{t\in[0,T]}$ *be right-continuous and adapted with*

$$E\big[|X_\tau|\big] < \infty \quad \textit{for all } [0,T]\textit{-valued stopping times } \tau.$$

If $X_T \leqslant \xi$ *and, for every stopping time* τ, *we have*

$$1_{\{\tau>t\}}X_t \leqslant E_t\left[1_{\{\tau>t\}}\int_t^\tau(\alpha_s X_s + \beta_s)ds + 1_{\{\tau>t\}}X_\tau\right] \quad \textit{for all } t \in [0,T],$$

then, it follows that

$$X_t \leqslant E_t\left[\int_t^T e^{\int_t^s \alpha_u du}H_s ds + e^{\int_t^T \alpha_u du}Z\right] \quad \textit{for all } t \in [0,T] \text{ a.s.}$$

Proof. By Proposition 2.77, the unique solution of the linear BSDE $dY_t = -[\alpha_t Y_t + \beta_t]dt + dM_t$, $Y_T = \xi$, is given by

$$Y_t = E_t\left[\int_t^T e^{\int_t^s \alpha_u du}\beta_s ds + e^{\int_t^T \alpha_u du}\xi\right], \qquad t \in [0,T].$$

The above Lemma 3.21 yields

$$\sup\big\{E\big[|Y_\tau|\big] : \tau \text{ is a } [0,T]\text{-valued stopping time}\big\} < \infty$$

and shows that, for all stopping times τ, we have

$$1_{\{\tau>t\}}Y_t = E_t\left[1_{\{\tau>t\}}\int_t^\tau \alpha_s Y_s ds + 1_{\{\tau>t\}}Y_\tau\right], \qquad t \in [0,T].$$

Thus $\Delta \triangleq X - Y$ is a right-continuous adapted process with

$$\Delta_T \leqslant 0, \qquad E\big[|\Delta_\tau|\big] < \infty, \text{ and}$$
$$1_{\{\tau>t\}}\Delta_t \leqslant E_t\left[1_{\{\tau>t\}}\int_t^\tau \alpha_s \Delta_s ds + 1_{\{\tau>t\}}\Delta_\tau\right], \quad t \in [0,T],$$

for all stopping times τ; Proposition 3.20 implies that $\Delta \leqslant 0$. □

As a corollary of Proposition 3.20, we give a proof of the comparison result which was used to establish the power utility bounds in Lemma 3.13 at the end of the previous Section 3.3.

Subsolutions are smaller than solutions

COROLLARY 3.23. *Let* (f, ξ) *be a BSDE*^{p}*-standard parameter, and let* X *denote the solution of the corresponding BSDE*

$$dX_t = -f(t, X_t)dt + dM_t, \qquad X_T = \xi. \tag{3.22}$$

If Y *is a subsolution (supersolution) of* (3.22), *then* $Y \leqslant X$ *(*$Y \geqslant X$*).*

Proof. We let Y be a subsolution of (3.22), set $\Delta \triangleq Y - X$ and note that $\Delta_T \leqslant 0$. Lemma 3.21 yields $E[|\Delta_\tau|] < \infty$ and

$$1_{\{\tau>t\}}\Delta_t \leqslant E_t\left[1_{\{\tau>t\}}\int_t^\tau\big(f(s,Y_s) - f(s,X_s)ds + 1_{\{\tau>t\}}\Delta_\tau\right], \quad t \in [0,T],$$

for all stopping times τ. Define a bounded progressive process α by

$$\alpha_s \triangleq \big[f(s,Y_s) - f(s,X_s)\big]/\big[Y_s - X_s\big] \quad \text{if } Y_s \neq X_s$$

and $\alpha_s \triangleq 0$ if $Y_s = X_s$. Then, for all stoppping times τ, we have

$$1_{\{\tau<t\}}\Delta_t \leqslant E_t\left[1_{\{\tau>t\}}\int_t^\tau \alpha_s \Delta_s ds + 1_{\{\tau>t\}}\Delta_\tau\right], \quad t \in [0,T],$$

and Proposition 3.20 shows that $Y - X = \Delta \leqslant 0$. If Y is a supersolution, then the proof is literally the same upon replacing Δ by $X - Y$. □

Note: Of course, Corollary 3.23 can be combined with the comparison result of Theorem 2.76 to obtain a comparison theorem for sub-/supersolutions of BSDEs; see, e.g., Peng (1999) for the explicit formulation of such a result.

Having established the stochastic Gronwall inequalites in Proposition 3.20 and Proposition 3.22 and the comparison result for sub-/ supersolutions of BSDEs, we are now in a position to return to our investigation of continuous-time Epstein-Zin utility.

3.5 PROOFS OF THE MAIN RESULTS

In this section, we give the proofs of our main results, which have been formulated in Section 3.2; thus, we restrict our considerations to the Epstein-Zin parameterization of stochastic differential utility with relative risk aversion and elasticity of intertemporal substitution

$$\gamma, \psi > 1, \quad \text{so that} \quad \phi = \tfrac{1}{\psi} < 1, \quad \theta = \tfrac{1-\gamma}{1-\phi} < 0, \quad q = \tfrac{\phi-\gamma}{1-\gamma} > 1.$$

Recall that the Epstein-Zin aggregator is given by

$$f(c,v) \triangleq \tfrac{\delta}{1-\phi} c^{1-\phi}[(1-\gamma)v]^{q} - \delta\theta v, \qquad \text{for } c, (1-\gamma)v \geqslant 0. \tag{3.2}$$

We fix an arbitrary $\beta > q$ and select $\alpha \geqslant 1$ such that ${}^{1-\phi}/_{\alpha} + {}^{q}/_{\beta} \leqslant 1$. Then, by the integrability condition in the definition of the family $\mathcal{C}$ of consumption plans (Definition 3.1), we have

$$E\big[\textstyle\int_0^T c_t^{\alpha} dt + c_T^{(1-\gamma)\beta} + c_T\big] < \infty \quad \text{for every } c \in \mathcal{C}. \tag{3.23}$$

Since it is only this condition that is required for the following proofs, the integrability condition in Definition 3.1 can be relaxed to a requirement of the form (3.23), if $\mathcal{S}^{\cap}$ is replaced by $\mathcal{S}^{\beta}$ in the definition (3.4) of $\mathcal{V}$. Finally, we set $\mathcal{V}^{\beta} \triangleq \{V \in \mathcal{S}^{\beta} : (1-\gamma)V_t \geqslant 0 \text{ for } t \in [0,T]\}$.

LEMMA 3.24. *Let $V \in \mathcal{V}^{\beta}$ satisfy*

$$V_t = E_t\big[\textstyle\int_t^T f(c_s, V_s) ds + U(c_T)\big], \qquad t \in [0,T], \tag{3.1}$$

for some $c \in \mathcal{C}$. Then $M_t = E_t[\int_0^T f(c_s,V_s)ds + U(c_T)]$, $t \in [0,T]$, defines a uniformly integrable martingale, and we have

$$dV_t = -f(c_t, V_t)dt + dM_t, \qquad V_T = U(c_T). \tag{3.24}$$

In particular, for all stopping times τ, we have

$$1_{\{\tau>t\}} V_t = E_t\big[1_{\{\tau>t\}} \textstyle\int_t^T f(c_s,V_s)ds + 1_{\{\tau>t\}} V_{\tau}\big], \quad t \in [0,T]. \tag{3.25}$$

Proof. Setting $\bar{\alpha} \triangleq \alpha/(1-\phi) > 1$ and $\bar{\beta} \triangleq \beta/q > 1$, we get $r^{-1} \triangleq \bar{\alpha}^{-1} + \bar{\beta}^{-1} \leqslant 1$. Hölder's inequality yields

$$\Big(\textstyle\int_0^T |f(c_s,V_s)|^r ds\Big)^{1/r} \leqslant \tfrac{\delta}{1-\phi}\Big(\textstyle\int_0^T c_s^{\alpha} ds\Big)^{1/\bar{\alpha}} \Big(\textstyle\int_0^T [(1-\gamma)V_s]^{\beta} ds\Big)^{1/\bar{\alpha}} + T^{1/r}|\delta\theta| \sup_{t\in[0,T]} |V_t|, \tag{3.26}$$

and hence the integrability assumption in (3.23) implies that

$$\int_0^T |f(c_s, V_s)| ds + |U(c_T)| \in L^1(P).$$

In particular, M is a uniformly integrable martingale, and thus V is a solution of the BSDE (3.24). The stopping time-representation of V in (3.25) is now an immediate consequence of Lemma 3.21. □

In Corollary 3.10 in Section 3.3 above, we have shown that for each bounded consumption plan $c \in \mathcal{C}_b$ there exists a unique $V = V(c) \in \mathcal{V}_b$ with

$$V_t = E_t\left[\int_t^T f(c_s, V_s) ds + U(c_T)\right], \qquad t \in [0, T]. \tag{3.1}$$

We recall that $c \in \mathcal{C}_b$ and $V \in \mathcal{V}_b$, if $c \in \mathcal{C}$ is a consumption plan and V is a semimartingale such that

$$k_0 \preccurlyeq c \preccurlyeq k_1 \quad \text{and} \quad k_0 \leqslant (1-\gamma)V \leqslant k_1$$

for some constants $0 < k_0 < k_1$; thus, a mapping

$$V : \mathcal{C}_b \to \mathcal{V}_b \subset \mathcal{V}^\beta, \quad c \mapsto V(c), \quad \text{where } V = V(c) \text{ satisfies (3.1)}$$

is already constructed. Our main task is to extend this mapping from $\mathcal{C}_b$ to $\mathcal{C}$ by a suitable limiting argument and to prove that the extension is monotone, concave and unique.

For this, the following general comparison result is key. A closely related result can be found in Schroder and Skiadas (1999).

Comparison result

THEOREM 3.25. *Let $c \in \mathcal{C}$ and let $V \in \mathcal{V}^\beta$ satisfy (3.1). Suppose that $Y \in \mathcal{V}^\beta$ with $Y_T \leqslant V_T$ and that for every stopping time τ*

$$1_{\{\tau>t\}} Y_t \leqslant E_t\left[1_{\{\tau>t\}} \int_t^\tau f(c_s, Y_s) ds + 1_{\{\tau>t\}} Y_\tau\right] \quad \text{for all } t \in [0, T].$$

Then it follows that $Y \leqslant V$.

Proof. Let τ be a stopping time. Using the representation (3.25) from Lemma 3.24, we obtain

$$1_{\{\tau>t\}} V_t = E_t\left[1_{\{\tau>t\}} \int_t^\tau f(c_s, V_s) ds + 1_{\{\tau>t\}} V_\tau\right] \quad \text{for all } t \in [0, T].$$

For all $c \in \mathfrak{C}$, the map $\mathfrak{U} \to \mathbb{R}$, $\nu \mapsto f(c, \nu)$ is convex by Lemma B.1, and hence we have

$$f(c_s, V_s) \geqslant f(c_s, Y_s) + f_\nu(c_s, Y_s)(V_s - Y_s) \quad \text{for dt-a.e. } s \in [0, T].$$

Thus we obtain

$$1_{\{\tau>t\}}(Y_t - V_t) \leqslant E_t\left[1_{\{\tau>t\}} \int_t^\tau f_\nu(c_s, V_s)(Y_s - V_s) ds + 1_{\{\tau>t\}}(Y_\tau - V_\tau)\right]$$

for all $t \in [0, T]$, where $f_\nu(c, Y)$ is a progressively measurable process that is bounded above by $-\delta\theta$. Since $Y_T - V_T \leqslant 0$, the stochastic Gronwall inequality in Proposition 3.20 implies that $Y - V \leqslant 0$. □

COROLLARY 3.26. *Let* $c^1 \preccurlyeq c^2$ *in* $\mathcal{C}$ *and* $V^1, V^2 \in \mathcal{V}^\beta$ *with*

Monotonicity and uniqueness

$$V^i_t = E_t\left[\int_t^T f(c^i_s, V^i_s)ds + U(c^i_T)\right], \quad t \in [0,T],\ i = 1,2.$$

Then $V^1 \leqslant V^2$*. In particular, for every* $c \in \mathcal{C}$*, there is at most one* $V(c) = V \in \mathcal{V}^\beta$ *that satisfies* (3.1)

Proof. The Epstein-Zin aggregator is increasing in c by Lemma B.1; hence, the representation (3.25) yields

$$1_{\{\tau>t\}}V^1_t \leqslant E_t\left[1_{\{\tau>t\}}\int_t^\tau f(c^2_s, V^1_s)ds + 1_{\{\tau>t\}}V^1_\tau\right], \qquad t \in [0,T].$$

By Theorem 3.25 the proof is complete. □

Having established uniqueness of stochastic differential utility, it now suffices to produce some $V \in \mathcal{V}^\beta$ that satisfies (3.1). We will obtain such a process by a monotone convergence argument, as a pointwise limit of utility processes associated to bounded consumption plans.

3.5.1 *Monotone convergence*

We say that a sequence $(c^n)_{n\in\mathbb{N}} \subset \mathcal{C}$ is *increasing* if $c^n \preccurlyeq c^{n+1}$ for all $n \in \mathbb{N}$, and we write $c^n \to c$ in $\mathcal{C}$ if and only if

$$c^n_t \to c_t \text{ for a.e. } t \in [0,T] \quad \text{and} \quad c^n_T \to c_T \quad \text{with} \quad c \in \mathcal{C}.$$

If $(c^n)_{n\in\mathbb{N}} \subset \mathcal{C}$ is increasing with $c^n \to c$ in $\mathcal{C}$, we briefly write $c^n \uparrow c$ in $\mathcal{C}$. The decreasing counterparts are defined analogously.

LEMMA 3.27. *Let* $(c^n)_{n\in\mathbb{N}} \subset \mathcal{C}$ *and* $(V^n)_{n\in\mathbb{N}} \subset \mathcal{V}^\beta$ *such that*

Monotone convergence

$$V^n_t = E_t\left[\int_t^T f(c^n_s, V^n_s)ds + U(c^n_T)\right], \quad t \in [0,T], \quad n \in \mathbb{N}. \tag{3.27}$$

If $c^n \uparrow c$ *or* $c^n \downarrow c$ *in* $\mathcal{C}$*, then there exists* $V \in \mathcal{V}^\beta$ *with*

$$V_t = E_t\left[\int_t^T f(c_s, V_s)ds + U(c_T)\right], \quad t \in [0,T],$$

and we have $V^n_t \to V_t$ *for all* $t \in [0,T]$.

Proof. If $(c^n)_{n\in\mathbb{N}}$ is increasing (decreasing), then Corollary 3.26 shows that $(V^n)_{n\in\mathbb{N}}$ is increasing (decreasing). In the increasing case, we have $V^1 \leqslant V^n \leqslant 0$ for all $n \in \mathbb{N}$, that is, $(V^n)_{n\in\mathbb{N}}$ is bounded in $\mathcal{V}^\beta$. To derive a similar bound in the decreasing case, we note that the Epstein-Zin aggregator (3.2) satisfies $f(c,v) \geqslant -\delta\theta v$, and hence, using the stopping time-representation in (3.25), we obtain

$$1_{\{\tau>t\}}V^n_t \geqslant E_t\left[1_{\{\tau>t\}}\int_t^\tau -\delta\theta V_s ds + 1_{\{\tau>t\}}V^n_\tau\right], \qquad t \in [0,T],$$

with $V^n_T = U(c^n_T) \geqslant U(c_T)$; thus the stochastic Gronwall inequality of Proposition 3.22 implies $V \geqslant \mathcal{U}$, where $\mathcal{U}_t \triangleq E_t[e^{-\delta\theta(T-t)}U(c_T)]$ and $\mathcal{U} \in \mathcal{V}^\beta$ by the integrability assumptions in (3.23).

Hence in both cases, it follows that $U \leqslant V^n \leqslant 0$ for all $n \in \mathbb{N}$ and some $U \in \mathcal{V}^\beta$; therefore, we can define the process V as the monotone pointwise limit $V_t \triangleq \lim_{n\to\infty} V_t^n$, $t \in [0,T]$. For each $n \in \mathbb{N}$ and dt-a.e. $s \in [0,T]$, we estimate

$$|f(c_s^n, V_s^n)| \leqslant \tfrac{\delta}{1-\phi}|c_s^1 + c_s|^{1-\phi}\big[(1-\gamma)U_s\big]^q + \tfrac{\delta}{1-\phi}(1-\gamma)U_s =: B_s.$$

Now, the integrability assumptions in (3.23) and Hölder's inequality imply that $B \in L^1(P \otimes dt)$; this argument appeared previously in the proof of Lemma 3.24 (see in particular inequality (3.26)). Moreover,

$$|U(c_t^n)| \leqslant \varepsilon^{1-\gamma}\big[(c_T^1)^{1-\gamma} + c_T^{1-\gamma}\big] \in L^1(P).$$

We have thus found an integrable upper bound on the relevant quantities in (3.27); therefore, sending $n \to \infty$ in (3.27), we obtain

$$V_t = \lim_{n\to\infty} E_t\left[\int_t^T f(c_s^n, V_s^n)ds + U(c_T^n)\right] = E_t\left[\int_t^T f(c_s, V_s)ds + U(c_T)\right]$$

for all $t \in [0,T]$, by dominated convergence. In particular, V is a semimartingale, and hence $V \in \mathcal{V}^\beta$ is a utility process for c as asserted. □

Remark. Using Doob's maximal inequality, one can show that in the situation of Lemma 3.27, we also have $V^n \to V$ in $\mathcal{S}^\beta$. This is, however, not needed in the following. ◇

Lemma 3.27 can be used to prove the existence of a utility process for an unbounded consumption plan c, by applying it to suitable, monotonically converging truncations c^n of c. Before, we carry this out, we first show that the map $\mathcal{C}_b \to \mathcal{V}_b, c \mapsto V(c)$ is concave, which will then easily carry over to the extension.

3.5.2 *Concavity for bounded consumption plans*

To establish concavity, it is convenient to consider the ordinally equivalent transformation $Y = u_\phi \circ u_\gamma^{-1}(V)$ as in the proof of the power utility bounds in Lemma 3.13 above.

LEMMA 3.28. *Let $c \in \mathcal{C}_b$ and let $V = V(c) \in \mathcal{V}_b$ be the associated bounded utility process. Then the process $Y = u_\phi \circ u_\gamma^{-1}(V)$ satisfies*

$$\begin{aligned} dY_t = &-\Big(\delta\big[u_\phi(c_t) - Y_t\big]dt + \tfrac{1}{2}\tfrac{\phi-\gamma}{1-\phi}\tfrac{d[Y_t]^c}{Y_{t-}} + \tfrac{1}{\theta}Y_{t-}\Delta Y_t^\theta - \Delta Y_t\Big) \\ &+\big((1-\gamma)V_{t-}\big)^{-q}dM_t, \end{aligned} \tag{3.28}$$

where M is a bounded martingale.

Proof. The process V has dynamics $dV_t = -f(c_t, V_t)dt + dM_t$, where M is a bounded martingale. Moreover $0 < k_0 \leqslant (1-\gamma)V \leqslant k_1$ for some constants k_0, k_1. A calculation using Itō's formula (see Lemma B.4, p. 173ff.) shows that Y has dynamics as in (3.28). □

Next we show that the convex combination of the transformed processes Y^1 and Y^2 associated with two bounded consumption plans $c^1, c^2 \in \mathcal{C}_b$ is, in the terminology of BSDE theory (see, e.g., Peng (1999)), a subsolution of (3.28) with $c \triangleq \lambda c^1 + (1-\lambda)c^2$. This boils down to showing that the driver of the BSDE is concave; however, the driver is integrated with respect to the quadratic variation process of the solution, which complicates the matter. To deal with these difficulties, we rely on the following convexity result.

LEMMA 3.29. *Let $\mathcal{S}_\bullet \triangleq \{X \in \mathcal{S} : X > 0\}$ denote the set of positive semimartingales and let $t, s \in [0,T]$, $t \leqslant s$. The mapping*

$$\mathcal{S}_\bullet \to L^0_+(P), \quad X \mapsto \int_t^s \frac{d[X]^c_\tau}{X_{\tau-}} \qquad \text{is convex.}$$

Proof. Corollary B.9; Section B.2 (p. 174ff.) is devoted to the proof. □

The jumps, on the other hand, are easily accommodated.

LEMMA 3.30. *The function*

$$h : (0,\infty)^2 \to \mathbb{R}, \quad h(x,y) = \tfrac{1}{\theta} x^{1-\theta}\big[y^\theta - x^\theta\big] \quad \text{is concave.}$$

Proof. Differentiating h, we get $h_{xx}(x,y) = (\theta-1)y^\theta/x^{\theta+1}$, $h_{xy}(x,y) = (1-\theta)y^{\theta-1}/x^\theta$ and $h_{yy}(x,y) = (\theta-1)x^{1-\theta}y^{\theta-2}$. Now $h_{xx} < 0$ and $h_{xx}h_{yy} - h^2_{xy} = 0$, so the Hessian of h is negative semidefinite. □

LEMMA 3.31. *Suppose that $c^i \in \mathcal{C}_b$, $i = 1,2$, let $V^i = V(c^i) \in \mathcal{V}_b$ denote the associated bounded utility process and let $\lambda \in [0,1]$. Moreover, set $Y^i \triangleq u_\phi \circ u_\gamma^{-1}(V^i)$ and $c \triangleq \lambda c^1 + (1-\lambda)c^2$ as well as $Y \triangleq \lambda Y^1 + (1-\lambda)Y^2$. Then*

$$dY_t = -\Big[\delta\big[u_\phi(c_t) - Y_t\big]dt + \tfrac{1}{2}\tfrac{\phi-\gamma}{1-\phi}\tfrac{d[Y_t]^c}{Y_{t-}} + \tfrac{1}{\theta}(Y_{t-})^{1-\theta}\Delta(Y_t)^\theta - \Delta Y_t\Big] - dA_t + dM_t, \tag{3.29}$$

where the martingale M is specified in (3.30) and A is a decreasing process.

Proof. Using the representation (3.28) of Y^1, Y^2 in Lemma 3.28, we get

$$\begin{aligned} dY_t = &-\big(\delta[u_\phi(c_t) - Y_t]dt + dA^1_t\big) - \Big(\tfrac{1}{2}\tfrac{\phi-\gamma}{1-\phi}\tfrac{d[Y_t]^c}{Y_{t-}} + dA^2_t\Big) \\ &- \big(\tfrac{1}{\theta}Y_{t-}\Delta Y^\theta_t - \Delta Y_t + dA^3_t\big) + dM_t, \end{aligned}$$

$$\begin{aligned} dA^1_t &= \delta\big[\lambda u_\phi(c^1_t) + (1-\lambda)u_\phi(c^2_t) - u_\phi(c_t)\big]dt, \\ dA^2_t &= \frac{1}{2}\frac{\phi-\gamma}{1-\phi}\left[\lambda\frac{d[Y^1_t]^c}{Y^1_{t-}} + (1-\lambda)\frac{d[Y^2_t]^c}{Y^2_{t-}} - \frac{d[Y_t]^c}{Y_{t-}}\right], \\ dA^3_t &= \lambda\left(\tfrac{1}{\theta}Y^1_{t-}\Delta(Y^1_t)^\theta\right) + (1-\lambda)\left(\tfrac{1}{\theta}Y^2_{t-}\Delta(Y^2_t)^\theta\right) - \left(\tfrac{1}{\theta}Y_{t-}\Delta Y^\theta_t\right), \\ dM_t &= \lambda\big((1-\gamma)V^1_{t-}\big)^{-q}dM^1_t + (1-\lambda)\big((1-\gamma)V^2_{t-}\big)^{-q}dM^2_t \end{aligned} \tag{3.30}$$

for some bounded martingales M^1 and M^2. Recalling that $V^1, V^2 \in \mathcal{V}_b$ take values in a compact subset of $\mathfrak{U} = (-\infty, 0)$, it follows that M is a martingale, as well. To complete the proof, it thus remains to show that $A \triangleq A_1 + A_2 + A_3$ is a decreasing process.

Concavity of power utility functions implies

$$\lambda u_\phi(c^1_t) + (1-\lambda)u_\phi(c^2_t) - u_\phi(c_t) \leqslant 0, \qquad t \in [0,T],$$

hence A^1 is decreasing. Moreover, Lemma 3.29 implies that for all $s,t \in [0,T]$ with $s \leqslant t$

$$Q(s,t) \triangleq \lambda \int_s^t \frac{d[Y^1]^c_\tau}{Y^1_{\tau-}} + (1-\lambda)\int_s^t \frac{d[Y^2]^c_\tau}{Y^2_{\tau-}} - \int_s^t \frac{d[Y]^c_\tau}{Y_{\tau-}} - \geqslant 0,$$

and thus $A^2 = \frac{1}{2}\frac{\phi-\gamma}{1-\phi}Q(0,\cdot)$ is a decreasing process. Lemma 3.30 yields $\lambda h(Y^1_{t-}, Y^1_t) + (1-\lambda)h(Y^2_{t-}, Y^2_t) - h(Y_{t-}, Y_t) \leqslant 0$ for all $t \in [0,T]$, i.e., only jumps of negative height occur in A^3, and hence A^3 is a decreasing pure jump process. □

Now, it remains to reverse the transformation and apply the comparison result from Proposition 3.25.

PROPOSITION 3.32. *The map* $\mathcal{C}_b \to \mathcal{S}_b$, $c \mapsto V(c)$ *is concave.*

Proof. Let $\lambda \in [0,1]$, $c^i \in \mathcal{C}_b$, and let $V^i = V(c^i) \in \mathcal{V}_b$ be the associated utility processes. Moreover define c, Y^1, Y^2 and Y as in Lemma 3.31 and let $V = V(c)$ be the utility process corresponding to c. The map $g \triangleq u_\gamma \circ u_\phi^{-1}$ is concave, so we have

$$\lambda V^1 + (1-\lambda)V^2 = \left(\lambda g(Y^1) + (1-\lambda)g(Y^2)\right) \leqslant g\left(\lambda Y^1 + (1-\lambda)Y^2\right) = g(Y),$$

and hence, to prove the assertion, it suffices to show that $g(Y) \leqslant V$.

By (3.29) and Itō's formula (see, e.g., Theorem I.4.57 in Jacod and Shiryaev (2003)), the process $X \triangleq g(Y)$ has dynamics

$$dX_t = -f(c_t, X_t)dt - g'(Y_{t-})dA_t + g'(Y_{t-})dM_t,$$

where M is given by (3.30); therefore, we obtain

$$\begin{aligned}1_{\{\tau>t\}}X_t = 1_{\{\tau>t\}}X_\tau + 1_{\{\tau>t\}}\textstyle\int_t^\tau f(c_s,X_s)ds + 1_{\{\tau>t\}}\int_t^\tau g'(Y_{s-})dA_s \\ - 1_{\{\tau>t\}}(N_\tau - N_t) \qquad (3.31)\end{aligned}$$

for any stopping time τ, where

$$N_t \triangleq \textstyle\int_0^t g'(Y_{s-})\Big[\lambda\big((1-\gamma)V^1_{s-}\big)^{-q}dM^1_s + (1-\lambda)\big(1-\gamma)V^2_{s-}\big)^{-q}dM^2_s\Big]$$

is martingale because the integrands are bounded processes and M^1 and M^2 are bounded martingales. Taking time-t conditional expectations in (3.31), it thus follows that

$$1_{\{\tau>t\}}X_t \leqslant E_t\left[1_{\{\tau>t\}}\textstyle\int_t^\tau f(c_s,X_s)ds + 1_{\{\tau>t\}}X_\tau\right]$$

since A is decreasing and $g' \leqslant 0$. Moreover, we have

$$Y_T = \lambda u_\phi(\varepsilon c^1_T) + (1-\lambda)u_\phi(\varepsilon c^2_T) \leqslant u_\phi(\lambda c^1_T + (1-\lambda)c^2_T) = u_\phi(\varepsilon c_T),$$

and hence $X_T = u_\gamma \circ u_\phi^{-1}(Y_T) \leqslant u_\gamma(\varepsilon c_T) = U(c_T) = V_T$; therefore, Proposition 3.25 implies that $X \leqslant V$ and concavity is established. □

3.5.3 *Extension to unbounded consumption plans*

THEOREM 3.33. *The map* $\mathcal{C}_b \subset \mathcal{C} \to \mathcal{V}^\beta, c \mapsto V(c)$ *admits a unique, concave extension to the class* $\mathcal{C}$ *of all consumption plans. This extension is uniquely characterized as follows: For every* $c \in \mathcal{C}$ *and* $V \in \mathcal{V}^\beta$ *we have*

$$V = V(c) \quad \Longleftrightarrow \quad V_t = E_t\left[\int_t^T f(c_s, V_s)ds + U(c_T)\right], \quad t \in [0,T]. \tag{3.32}$$

Proof. Let $c \in \mathcal{C}$. Corollary 3.26 implies that there is at most one process $V = V(c) \in \mathcal{V}^\beta$ which satisfies the BSDE in (3.32). To construct V, we first consider $\bar{c} \in \mathcal{C}$ with $k_0 \preccurlyeq \bar{c}$ for some constant $k_0 > 0$ and define, for each $n \in \mathbb{N}$, an associated bounded consumption plan $\bar{c}^n \in \mathcal{C}_b$ by $\bar{c}^n_t \triangleq \bar{c}_t \wedge n$, $t \in [0,T]$. Lemma 3.9 yields a bounded utility process $\bar{V}^n = V(\bar{c}^n) \in \mathcal{V}_b$ associated to $\bar{c}_n$.

Since $\bar{c}^n \uparrow \bar{c}$ in $\mathcal{C}$, the monotone convergence result of Lemma 3.27 implies that there exists $\bar{V} \in \mathcal{V}^\beta$ with

$$\bar{V}_t = E_t\left[\int_t^T f(\bar{c}_s, \bar{V}_s)ds + U(\bar{c}_T)\right], \quad t \in [0,T].$$

In the general case, we put $c^n \triangleq c + \frac{1}{n}$ and use the previous argument to construct $V^n \in \mathcal{V}^\beta$ with

$$V^n_t = E_t\left[\int_t^T f(c^n_s, V^n_s)ds + U(c^n_T)\right], \quad t \in [0,T], \quad n \in \mathbb{N}.$$

Since $c^n \downarrow c$ in $\mathcal{C}$, Lemma 3.27 shows that $V^n_t \to V_t$, where $V \in \mathcal{V}^\beta$ and

$$V_t = E_t\left[\int_t^T f(c_s, V_s)ds + U(c_T)\right], \quad t \in [0,T].$$

Thus $V(c)$ is well-defined via (3.32) for every $c \in \mathcal{C}$. Finally, concavity follows from Proposition 3.32 by a limiting argument. □

With the previous Theorem 3.33, we have also established Theorem 3.2 and Theorem 3.4.

COROLLARY 3.34. *Let* $(c^n)_{n\in\mathbb{N}} \subset \mathcal{C}$ *and suppose there are* $\underline{c}, \bar{c} \in \mathcal{C}$ *such that* $\underline{c} \preccurlyeq c^n \preccurlyeq \bar{c}$ *for all* $n \in \mathbb{N}$*. If* $c^n \to c$ *in* $\mathcal{C}$*, then* $V_t(c^n) \to V_t(c)$ *for all* $t \in [0,T]$.

Dominated convergence

Proof. The assumptions imply that $a^n \triangleq \inf_{k\geqslant n} c^n \in \mathcal{C}$ and $b^n \triangleq \sup_{k\geqslant n} c^k \in \mathcal{C}$. Clearly $a^n \uparrow c$ and $b^n \downarrow c$ in $\mathcal{C}$, so by monotone convergence (Lemma 3.27)

$$V_t(a^n), V_t(b^n) \to V_t(c) \quad \text{for all } t \in [0,T].$$

On the other hand, monotonicity (Corollary 3.26) yields $V(a^n) \geqslant V(c^n) \geqslant V(b^n)$ for every $n \in \mathbb{N}$, and hence the proof is complete. □

Using the dominated convergence result of Corollary 3.34, we can extend the power utility bound (3.9) from Corollary 3.14 to all of $\mathcal{C}$. Recall that

$$\mathbf{U}_\gamma(c) \leqslant \mathbf{V}(c) \leqslant \mathbf{U}_\phi(c) \quad \text{for all } c \in \mathcal{C}_b, \tag{3.9}$$

where $\mathbf{U}_\rho(c) = u_\gamma \circ u_\rho^{-1}(Y^\rho(c))$, and $Y^\rho(c) = Y^\rho$ is given by

$$Y_t^\rho = \delta e^{\delta t} E_t\big[\int_t^T e^{-\delta s} u_\rho(c_s) ds + u_\rho(\varepsilon c_T)\big], \quad \rho \in \{\phi, \gamma\}. \tag{3.33}$$

Power utility bounds

PROPOSITION 3.35. *Let* $c \in \mathcal{C}$ *and suppose that*

$$E\left[\int_0^T (c_t^{1-\phi} + c_t^{1-\gamma}) dt + c_T^{1-\gamma} + c_T^{1-\phi}\right] < \infty.$$

Then $\mathbf{U}_\gamma(c) \leqslant \mathbf{V}(c) \leqslant \mathbf{U}_\phi(c)$. *In particular,* $(1-\gamma)\mathbf{V}(c) > 0$.

Note: The integrability condition in the formulation of Proposition 3.35 is implied by the integrability condition in the definition of the family of consumption plans $\mathcal{C}$ (Definition 3.1). It has been added here so that the proposition remains valid if $\mathcal{C}$ is defined by the weaker integrability condition (3.23), which is all that is required in this section. Under the original integrability assumption from Definition 3.1, the above proposition also shows that $E[\sup_{t\in[0,T]} |\mathbf{V}_t(c)|^p] < \infty$ for all $p \in \mathbb{R}$.

Proof. For $n \in \mathbb{N}$, we set $c^n \triangleq (\frac{1}{n} \vee c) \wedge n$ and note that $c^n \to c$ with $c \wedge 1 \preccurlyeq c^n \preccurlyeq c \vee 1$. Corollary 3.34 implies that $\mathbf{V}_t(c^n) \to \mathbf{V}_t(c)$ for all $t \in [0,T]$. At the same time, (3.9) shows that

$$Y_n^\gamma = \mathbf{U}_\gamma(c^n) \leqslant \mathbf{V}(c^n) \leqslant \mathbf{U}_\phi(c^n) \leqslant u_\gamma \circ u_\phi^{-1}(Y_n^\phi),$$

where Y_n^γ and Y_n^ϕ are given by (3.33), with c replaced by c^n. By dominated convergence, we have $(Y_n^\gamma)_t \to Y_t^\gamma = [\mathbf{U}_\gamma(c)]_t$ and $(Y_n^\phi)_t \to Y_t^\phi = [u_\phi \circ u_\gamma^{-1}(\mathbf{U}_\phi(c))]_t$ for all $t \in [0,T]$, and hence the claim follows by sending $n \to \infty$. Note that $E[\sup_{t\in[0,T]}((1-\gamma)\mathbf{U}_\phi(c)_t)^{1/\theta}] < \infty$, so

$$E\big[\sup_{t\in[0,T]} ((1-\gamma)\mathbf{V}_t(c))^{1/\theta}\big] < \infty \quad \text{and} \quad \mathbf{V}(c) > 0. \qquad \square$$

3.5.4 *Proof of Theorem 3.5*

This subsection is devoted to the proof of

Utility gradient inequality

THEOREM 3.5. *Suppose that* $\hat{c} \in \mathcal{C}$ *satisfies*

$$E\left[\int_0^T f_c(\hat{c}_t, V_t(\hat{c}))\hat{c}_t\, dt + U'(\hat{c}_T)\hat{c}_T\right] < \infty. \tag{3.34}$$

Then, the following utility gradient inequality holds:
For every $c \in \mathcal{C}$ *and all* $t \in [0,T]$, *we have*

$$V_t(c) \leqslant V_t(\hat{c}) + \langle m^t(\hat{c}), c - \hat{c}\rangle_t \tag{3.35}$$

where $\langle m, y\rangle_t \triangleq E_t[\int_t^T m_s y_s\, ds + m_T y_T]$ *and the time-*t *utility gradient* $m^t(\hat{c}) \triangleq \hat{m}^t$ *is given by*

$$\hat{m}_s^t \triangleq e^{\int_t^s f_v(\hat{c}_\tau, V_\tau(\hat{c}))d\tau}\big(f_c(\hat{c}_s, V_s(\hat{c}))1_{[t,T)}(s) + U'(\hat{c}_T)1_{\{T\}}(s)\big), \quad s \in [t,T].$$

The idea of the proof is the following: By concavity of $c \mapsto V(c)$, we can establish (3.35) by identifying $\langle m^t(\hat{c}), c - \hat{c}\rangle_t$ as the derivative of the utility process at $\hat{c}$, in the direction of c,

$$\lim_{\lambda\to 0} \tfrac{1}{\lambda}\big(V_t(\hat{c}+\lambda(c-\hat{c})) - V_t(\hat{c})\big) = \lim_{\lambda\to 0} \tfrac{1}{\lambda}\big(V_t(c^\lambda) - V_t(\hat{c})\big),$$

where $c^\lambda \triangleq (1-\lambda)\hat{c} + \lambda c$. The scaled difference $\Delta^\lambda \triangleq (V(c^\lambda) - V(\hat{c}))/\lambda$ of the two utility processes in the above equation is given by

$$\lambda\Delta_t^\lambda = E_t\left[\int_t^T f(c_s^\lambda, V_s(c^\lambda)) - f(\hat{c}_s, V_s(\hat{c}))\,ds + U(c_T^\lambda) - U(\hat{c}_T)\right]. \quad (3.36)$$

The implicit nature of the above equation complicates the matter of taking the limit $\lambda \to 0$; thus, the first step is to derive an explicit representation. For this, we linearize the equation, using the following measurable selection from the mean value theorem.

Measurable selection: Mean value theorem

LEMMA 3.36. *Let X, Y be real-valued optional processes and $g : \Omega \times [0,T] \times \mathbb{R}$ be such that $(\omega, t) \mapsto g(\omega, t, x)$ is optional for every $x \in \mathbb{R}$ and $x \mapsto g(\omega, t, x)$ is continuously differentiable for all $(\omega, t) \in \Omega \times [0,T]$. Then, there exists an optional process Z with $X \wedge Y \leqslant Z \leqslant X \vee Y$ such that*

$$g_x(\cdot, Z)(X - Y) = g(\cdot, X) - g(\cdot, Y). \quad (3.37)$$

Proof. Endow $\Omega \times [0,T]$ with the optional σ-algebra and define a compact-valued correspondence $\varphi : \Omega \times [0,T] \to 2^{\mathbb{R}}$ by $\varphi(\omega, t) = [X_t(\omega) \wedge Y_t(\omega), X_t(\omega) \vee Y_t(\omega)]$. Theorem 18.15 in Aliprantis and Border (2006) implies that φ is weakly measurable. Consider the function

$$f(\cdot, x) \triangleq -\big|g_x(\cdot, x)(X - Y) - \big(g(\cdot, X) - g(\cdot, Y)\big)\big|.$$

For fixed (ω, t), the map $x \mapsto f(\omega, t, x)$ is continuous; moreover, the process $(\omega, t) \to f(\omega, t, x)$ is optional for all $x \in \mathbb{R}$. Theorem 19.19 in Aliprantis and Border (2006) shows that the argmax-correspondence

$$(\omega, t) \mapsto \Big\{z \in \varphi(\omega, t) \,:\, f(\omega, t, z) = \max_{x\in\varphi(\omega,t)} f(\omega, t, x)\Big\}$$

admits an optional selector Z. Since $\max_{x\in\varphi(\omega,t)} f(\omega, t, x) = 0$ by the mean-value theorem, the process Z satisfies (3.37). □

Returning to the proof of Theorem 3.5, we use Lemma 3.36 to obtain an optional process $\tilde{c}^\lambda$ with $c^\lambda \wedge \hat{c} \preccurlyeq \tilde{c}^\lambda \preccurlyeq c^\lambda \vee \hat{c}$ such that

$$f(c^\lambda, V(c^\lambda)) - f(\hat{c}, V(c^\lambda)) = f_c(\tilde{c}^\lambda, V(c^\lambda))(c^\lambda - \hat{c}) \quad (3.38)$$

and $U(c_T^\lambda) - U(\hat{c}_T) = U'(\tilde{c}_T^\lambda)(c_T^\lambda - \hat{c}_T)$. Moreover, we get an optional process V^λ with $V(\hat{c}) \wedge V(c^\lambda) \leqslant V^\lambda \leqslant V(\hat{c}) \vee V(c^\lambda)$ such that

$$f(\hat{c}, V(c^\lambda)) - f(\hat{c}, V(\hat{c})) = f_v(\hat{c}, V^\lambda)(V(c^\lambda) - V(\hat{c})). \quad (3.39)$$

Recalling $c^\lambda \triangleq (1-\lambda)\hat{c}+\lambda c$, we see that $(c^\lambda - \hat{c})/\lambda = c - \hat{c}$; thus inserting (3.38) and (3.39) into (3.36), we arrive at

$$\Delta_t^\lambda = E_t\left[\int_t^T f_c(\tilde{c}_s^\lambda, V_s(c^\lambda))(c_s-\hat{c}_s) + f_v(\hat{c}_s, V_s^\lambda)\Delta_s^\lambda ds + U'(\tilde{c}_T^\lambda)(c_T-\hat{c}_T)\right], \tag{3.40}$$

that is, Δ^λ satisfies a linear BSDE. An explicit formula for the solutions of linear BSDEs is available; however, $f_v(\hat{c}, V^\lambda)$ is not bounded, nor can we expect that the other quantities involved are integrable for any choice of $\hat{c}$ and c. Hence the standard result (e.g., our Propostion 2.77) is not applicable, and we have to rely on approximation arguments.

Proof for bounded consumption plans

STEP 1. In a first step, we prove Theorem 3.5 for consumption plans $\hat{c}, c \in \mathcal{C}$ that satisfy

$$(c-\hat{c})/\hat{c} \leqslant K \qquad \text{for some } K > 0, \tag{B1}$$
$$2k \preccurlyeq \hat{c} \preccurlyeq \ell/2 \qquad \text{for some } k, \ell > 0. \tag{B2}$$

For all $\lambda < \bar{\lambda} \triangleq \min(2^{-1}, K^{-1})$, assumptions (B1) and (B2) imply

$$k \preccurlyeq \hat{c}^2/2 \preccurlyeq c^\lambda \wedge \hat{c} \preccurlyeq \tilde{c}^\lambda \preccurlyeq c^\lambda \vee \hat{c} \preccurlyeq 2\hat{c} \preccurlyeq \ell, \tag{3.41}$$

i.e., $\tilde{c}^\lambda \in \mathcal{C}_b$ is a bounded consumption plan. Recall that

$$\bar{c} \in \mathcal{C}_b \implies k_0(\bar{c}) \leqslant V(\bar{c}) \leqslant k_1(\bar{c}) \quad \text{for } k_0(\bar{c}), k_1(\bar{c}) < 0,$$

by Corollary 3.10 and Theorem 3.2; hence, by monotonicity of $c \mapsto V(c)$ (Theorem 3.4), we can find constants $k_0, k_1 < 0$ such that

$$k_0 \leqslant V(\hat{c}/2) \leqslant V(\hat{c}) \wedge V(c^\lambda) \leqslant V^\lambda \leqslant V(\hat{c}) \vee V(c^\lambda) \leqslant V(2\hat{c}) \leqslant k_1.$$

In summary, (B1) and (B2) guarantee that $\tilde{c}_\lambda$ and $(1-\gamma)V^\lambda$ take values in a compact subset of $(0,\infty)$ for all $\lambda < \bar{\lambda}$; thus there exists some $M > 0$ such that

$$|f_c(\tilde{c}_s^\lambda, V_s(c^\lambda))| + |f_v(\hat{c}_s, V_s^\lambda)| \leqslant M \quad \text{for } dt\text{-a.e. } t \in [0,T], \tag{3.42}$$

and $|U'(\tilde{c}_T^\lambda)| \leqslant M$ for all $\lambda < \bar{\lambda}$. Moreover, $|c-\hat{c}| \leqslant K\hat{c} \leqslant K\ell/2$; hence, all quantities involved in (3.40) are bounded by a deterministic constant. In particular, Propostion 2.77 is applicable, and it shows that the unique solution of (3.40) is given by

$$\Delta_t^\lambda = E_t\left[\int_t^T G_s^{t,\lambda}(c_s-\hat{c}_s)ds + G_T^{t,\lambda}(c_T-\hat{c}_T)\right], \quad t \in [0,T], \tag{3.43}$$

where for $t, s \in [0,T]$, $t \leqslant s$,

$$G_s^{t,\lambda} \triangleq e^{\int_t^s f_v(\hat{c}_\tau, V_\tau^\lambda)d\tau}\big(f_c(\tilde{c}_s^\lambda, V_s(c^\lambda))1_{[0,T)}(s) + U'(\tilde{c}_T^\lambda)1_{\{T\}}(s)\big).$$

Sending $\lambda \to 0$ in (3.43), we can prove the utility gradient inequality under the assumptions (B1) and (B2).

LEMMA 3.37. *Suppose that* (B1) *and* (B2) *hold. Then, for all* $t \in [0,T]$, *we have the utility gradient inequality*

Utility gradient inequality for bounded plans (I)

$$V_t(c) \leqslant V_t(\hat{c}) + \langle m^t(\hat{c}), c - \hat{c}\rangle_t. \tag{3.44}$$

Proof. Let $t \in [0,T]$ and recall that $m^t(\hat{c})$ is defined as

$$m^t_s(\hat{c}) \triangleq e^{\int_t^s f_v(\hat{c}_\tau, V_\tau(\hat{c}))d\tau}\big(f_c(\hat{c}_s, V_s(\hat{c}))1_{[0,T)}(s) + U'(\hat{c}_T)1_{\{T\}}(s)\big)$$

for $s \in [t,T]$. Since $c^\lambda \to \hat{c}$ in $\mathcal{C}$ as $\lambda \to 0$, our dominated convergence result (Corollary 3.34) implies that $V_t(c^\lambda) \to V_t(\hat{c})$ as $\lambda \to 0$, where the relevant bounds are provided by (3.41). In particular, $\tilde{c}^\lambda \to \hat{c}$ and $V^\lambda_t \to V_t(\hat{c})$ as $\lambda \to 0$, and therefore

$$G^{t,\lambda}_s \to m^t_s(\hat{c}) \quad \text{for } dt\text{-a.e. } s \in [0,T] \quad \text{and} \quad G^{t,\lambda}_T \to m^t_T(\hat{c}) \quad \text{as } \lambda \to 0,$$

establishing pointwise convergence inside the conditional expectation in (3.43); thus, (3.42) and dominated convergence imply

$$\Delta^\lambda_t \to \langle m^t(\hat{c}), c - \hat{c}\rangle_t \quad \text{as } \lambda \to 0.$$

Noting that $V_t(c^\lambda) \geqslant V_t(\hat{c}) + \lambda[V_t(c) - V_t(\hat{c})]$ by concavity (Theorem 3.4), we get $V_t(c) \leqslant V_t(\hat{c}) + \Delta^\lambda_t$; hence, the utility gradient inequality (3.44) obtains in the limit $\lambda \to 0$. □

STEP 2. In a second step, we relax the assumptions on $\hat{c}, c \in \mathcal{C}$ to

Relaxing assumption (B2)

$$(c - \hat{c})/\hat{c} \leqslant K \qquad \text{for some } K > 0, \tag{B1}$$

$$k \preccurlyeq \hat{c} \qquad \text{for some } k > 0, \tag{B2'}$$

We set $h \triangleq (c - \hat{c})/\hat{c}$ and note that $h \leqslant K$ by (B1). For $n \in \mathbb{N}$, we put $\hat{c}^n \triangleq \hat{c} \wedge n$ and $c^n \triangleq (1+h)\hat{c}^n$. Then

$$\frac{c^n - \hat{c}^n}{\hat{c}^n} = \frac{(1+h)\hat{c}^n - \hat{c}^n}{\hat{c}^n} = h \leqslant K,$$

so $\hat{c}^n$ and c^n do satisfy (B1). Moreover, $\hat{c}^n \leqslant n$, and hence (B2) is satisfied, as well. Thus the first step, Lemma 3.37, implies

$$V_t(c^n) \leqslant V_t(\hat{c}^n) + \langle m^t(\hat{c}^n), c^n - \hat{c}^n\rangle_t \quad \text{for all } t \in [0,T], \tag{3.45}$$

where the relevant utility gradient is given as

$$m^t_s(\hat{c}^n) \triangleq e^{\int_t^s f_v(\hat{c}^n_\tau, V_\tau(\hat{c}^n))d\tau}\big(f_c(\hat{c}^n_s, V_s(\hat{c}^n))1_{[t,T)}(s) + U'(\hat{c}^n_T)1_{\{T\}}(s)\big),$$

for $s \in [t,T]$. Clearly, $\hat{c}^n \uparrow \hat{c}$, and hence $c^n \uparrow c$, as well. Lemma 3.27 shows that $V_t(c^n) \uparrow V_t(c)$ and $V_t(\hat{c}^n) \uparrow V_t(\hat{c})$, $t \in [0,T]$, and therefore

$$m^t_s(\hat{c}^n) \to m^t_s(\hat{c}) \quad \text{for } dt\text{-a.e. } s \in [t,T] \quad \text{and} \quad m^t_T(\hat{c}^n) \to m^t_T.$$

To prove the utility gradient inequality, it thus remains to justify that we can take the limit inside the inner product in (3.45). Since $f_v \leqslant -\delta\frac{1-\gamma}{1-\phi}$ (see, e.g., Lemma B.1), for all $n \in \mathbb{N}$, we have

$$e^{\delta\frac{1-\gamma}{1-\phi}T} m^t_s(\hat{c}^n) \leqslant f_c(\hat{c}^n_s, V_s(\hat{c}^n)) \leqslant f_c(k, V_s(k)) \quad \text{for } dt\text{-a.e. } s \in [t,T],$$

where the last inequality is due to $f_{cc} \leqslant 0$, $f_{cv} \leqslant 0$; see again Lemma B.1. Since $k_0 \leqslant V(k) \leqslant k_1$ for some constants $k_0, k_1 < 0$, we see that $0 \leqslant m^t(\hat{c}^n)$ is bounded by a positive constant. Similarly, $0 \leqslant e^{\delta\frac{1-\gamma}{1-\phi}T} m^t_T(\hat{c}^n) \leqslant U'(\hat{c}^n_T) \leqslant U'(k)$. Finally, $|c^n - \hat{c}^n| = |h|\hat{c} \leqslant K\hat{c}$; therefore, we find $M > 0$ such that

$$\left|m^t_s(\hat{c}^n)(c^n_s - \hat{c}^n_s)\right| \leqslant M\hat{c}_s \quad \text{for } dt\text{-a.e. } s \in [t, T],$$

and $|m^t_T(\hat{c}^n)(c^n_T - \hat{c}^n_T)| \leqslant M\hat{c}_T$. This uniform upper bound is integrable as $\hat{c} \in \mathcal{C}$, or, more generally, by the integrability condition in (3.23). Sending $n \to \infty$ in (3.45), the dominated convergence theorem implies (3.46), thus completing the second step. We have just proven the following lemma:

Utility gradient inequality for bounded plans (II)

LEMMA 3.38. *Suppose that* (B1) *and* (B2′) *hold. Then, for all* $t \in [0, T]$, *we have the utility gradient inequality*

$$V_t(c) \leqslant V_t(\hat{c}) + \langle m^t(\hat{c}), c - \hat{c}\rangle_t. \tag{3.46}$$

Dropping assumption (B2′)

STEP 3. In a third step, we relax the assumptions on $\hat{c}, c \in \mathcal{C}$ to

$$(c - \hat{c})/\hat{c} \leqslant K \qquad \text{for some } K > 0, \tag{B1}$$

$$E\left[\int_0^T f_c(\hat{c}_t, V_t(\hat{c}))\hat{c}_t dt + U'(\hat{c}_T)\hat{c}_T\right] < \infty. \tag{3.34}$$

Utility gradient inequality with (B1)

LEMMA 3.39. *Suppose that* (B1) *and* (3.34) *hold. Then for all* $t \in [0, T]$, *we have the utility gradient inequality*

$$V_t(c) \leqslant V_t(\hat{c}) + \langle m^t(\hat{c}), c - \hat{c}\rangle_t.$$

Proof. For each $n \in \mathbb{N}$, we set $\hat{c}^n \triangleq \hat{c} + \frac{1}{n}$. Then

$$h^n \triangleq \frac{c - \hat{c}^n}{\hat{c}^n} = \frac{c}{\hat{c}^n} - 1 \leqslant \frac{c}{\hat{c}} - 1 = h \leqslant K$$

by (B1), and $\hat{c}^n \succcurlyeq 1/n$; thus, $\hat{c}^n$ and c satisfy (B1) and (B2′). The second step, Lemma 3.38, yields

$$V_t(c) \leqslant V_t(\hat{c}) + \langle m^t(\hat{c}^n), c - \hat{c}^n\rangle_t, \quad \text{where} \tag{3.47}$$

$$m^t_s(\hat{c}^n) \triangleq e^{\int_t^s f_v(\hat{c}^n_\tau, V_\tau(\hat{c}^n))d\tau}\left(f_c(\hat{c}^n_s, V_s(\hat{c}^n))1_{[t,T)}(s) + U'(\hat{c}^n_T)1_{\{T\}}(s)\right),$$

for $s \in [t, T]$. Since $\hat{c}^n \downarrow \hat{c}$, Lemma 3.27 applies to show that $V_t(\hat{c}^n) \downarrow V_t(\hat{c})$, $t \in [0, T]$; hence, it remains to show that we can take the limit inside the inner product in (3.47). Recalling that $f_v \leqslant -\delta\frac{1-\gamma}{1-\phi}$, we have

$$0 \leqslant e^{\delta\frac{1-\gamma}{1-\phi}T} m^t_s(\hat{c}^n) \leqslant f_c(\hat{c}_s, V_s(\hat{c}_s)) \quad \text{for } dt\text{-a.e. } s \in [t, T]$$

since $f_{cc} < 0$, $f_{cv} < 0$. Similarly, because $U'' < 0$, we get

$$0 \leqslant e^{\delta\frac{1-\gamma}{1-\phi}T} m^t_T(\hat{c}^n) \leqslant U'(\hat{c}_T).$$

In view of

$$|c - \hat{c}^n| = |h^n|\hat{c} \leqslant K\hat{c} \qquad \text{and the integrability assumption in (3.34),}$$

we have found an integrable upper bound for the relevant quantities in (3.47); the claim follows by dominated convergence. □

STEP 4. In the final step, we drop all additional assumptions and only suppose that $\hat{c} \in \mathcal{C}$ satisfies

Proof of the general case

$$E\left[\int_0^T f_c\big(\hat{c}_t, V_t(\hat{c})\big)\hat{c}_t dt + U'(\hat{c}_T)\hat{c}_T\right] < \infty, \tag{3.34}$$

as in the statement of Theorem 3.5.

Proof of Theorem 3.5. For $c \in \mathcal{C}$, we define the predictable $[-1,\infty)$-valued process h by $h \triangleq (c-\hat{c})/\hat{c} = c/\hat{c} - 1$. For each $n \in \mathbb{N}$, we put $h^n \triangleq h \wedge n$ and $c^n \triangleq (1+h^n)\hat{c}$. Then $c^n \in \mathcal{C}$, and we have

$$(c^n - \hat{c})/\hat{c} = h^n \leqslant n.$$

Then $\hat{c}$ and c^n satisfy assumption (B1) and (3.34), and hence the third step, Lemma 3.39, yields

$$V_t(c^n) \leqslant V_t(\hat{c}) + E_t\left[\int_t^T m_s^t(\hat{c})(c_s^n - \hat{c}_s)ds + m_T^t(\hat{c})(c_T^n - \hat{c}_T)\right]. \tag{3.48}$$

Clearly, $h^n \uparrow h$, and hence $c^n \uparrow c$ in $\mathcal{C}$, so Lemma 3.27 shows that $V_t(c^n) \to V_t(c)$, $t \in [0,T]$. Since $\hat{m}_s^t \triangleq m^t(\hat{c}) \geqslant 0$, we have

$$\begin{aligned} -\hat{m}_s^t\hat{c}_s &\leqslant \hat{m}_s^t(c_s^n - \hat{c}_s) \leqslant \hat{m}_s^t(c_s^{n+1} - \hat{c}_s) \qquad \text{for } dt\text{-a.e } s \in [t,T], \\ -\hat{m}_T^t\hat{c}_T &\leqslant \hat{m}_T^t(c_T^n - \hat{c}_T) \leqslant \hat{m}_T^t(c_T^{n+1} - \hat{c}_T) \qquad \text{for all } n \in \mathbb{N}. \end{aligned}$$

Recalling $f_v \leqslant -\delta\frac{1-\gamma}{1-\phi}$, for the lower bound, we obtain

$$E\left[\int_t^T |\hat{m}_s^t\hat{c}_s|ds + |\hat{m}_T^t\hat{c}_T|\right] \leqslant e^{-\delta\frac{1-\gamma}{1-\phi}T}\, E\left[\int_t^T f_c\big(\hat{c}_s, V_s(\hat{c})\big)\hat{c}_s ds + U'(\hat{c}_T)\hat{c}_s\right],$$

which is integrable by (3.34); thus the assertion follows from monotone convergence upon letting $n \to \infty$ in (3.48). □

With the proof of Theorem 3.5, we have provided the proofs of all results formulated in Section 3.2. Our investigation of the Epstein-Zin parameterization of SDU with $\gamma, \psi > 1$ is completed.

4 CONSUMPTION-PORTFOLIO OPTIMIZATION WITH STOCHASTIC DIFFERENTIAL UTILITY

In this chapter, we study the incomplete-market consumption-portfolio problem of an investor with continuous-time Epstein-Zin preferences. The exposition is largely based on Kraft, Seiferling, and Seifried (2015).

First and foremost, we provide an explicit construction of bounded, positive $C^{1,2}$-solutions for a class of semilinear partial differential equations (PDEs). This construction is based on fixed point arguments for the associated system of forward-backward stochastic differential equations. We study the Feynman-Kac representation mapping Φ that is associated with the semilinear PDE and obtain a fixed point in the space of continuous functions as a limit of iterations of Φ. We are able to improve uniform convergence to convergence in $C^{0,1}$, using the probabilistic representation of the solution. This not only yields a theoretical convergence result, but also leads directly to a numerical method with superexponential speed of convergence that allows us to determine optimal strategies efficiently via iteratively solving linear PDEs. Moreover, we establish a verification theorem which characterizes the value function of the consumption-portfolio problem in terms of a bounded, positive $C^{1,2}$ solution of such a semilinear partial differential equation, which appears as a reduced version of the Hamilton-Jacobi-Bellman equation. The proof of this result is based on a combination of dynamic programming arguments and utility gradient inequalities for recursive utility.

The above-mentioned results provide a new method to solve incomplete-market consumption-portfolio problems and asset pricing models with unspanned risk and recursive preferences: In both settings the agent's value function is characterized by a semilinear partial differential equation. In the literature, solutions of this equation have only been obtained in special cases, and general existence and uniqueness results have not been available; thus, researchers have resorted to analytical approximations of unclear precision. Here, we establish both theoretical existence and uniqueness results and an efficient numerical method for that equation. Our results are neither restricted to affine asset dynamics nor do we have to impose any constraints on the agent's risk aversion or elasticity of intertemporal substitution.

The chapter is structured as follows: In Section 4.1 we formulate the consumption-portfolio problem and the associated dynamic programming equation, and we derive candidate optimal strategies. Moreover, we provide links to the literature. Section 4.2 is concerned with existence and uniqueness of classical solutions for a class of semilinear PDEs and contains the proof of our first main result – Theorem 4.10. Theorem 4.10 guarantees in particular that the dynamic programming equation has a unique solution. A verification result, which demonstrates that the associated candidate optimal strategies do indeed provide the solution of the consumption-portfolio problem, is then provided in Section 4.3. After that, Section 4.4 briefly relates our findings to the asset pricing literature. Building on the Feynman-Kac iteration method from Section 4.2, Section 4.5 sets the basis for our numerical method which is applied to several examples of consumption-portfolio and asset pricing problems in Section 4.6.

4.1 OPTIMAL CONSUMPTION-PORTFOLIO SELECTION WITH EPSTEIN-ZIN PREFERENCES

To formulate the consumption-portfolio problem, we first set up the mathematical framework, recall the definition of continuous-time Epstein-Zin preferences, and specify the financial market model.

4.1.1 *Mathematical model and Epstein-Zin preferences*

We fix a probability space $(\Omega, \mathcal{F}, P)$ with a complete right-continuous filtration $(\mathcal{F}_t)_{t\in[0,T]}$ that is generated by a Wiener process $(W, \bar{W})$. We denote the *consumption space* by $\mathfrak{C} \triangleq (0, \infty)$. In the following, we are interested in an agent's preferences on the set of dynamic consumption plans.

Consumption plans

DEFINITION 4.1. A progressively measurable process c with values in $\mathfrak{C}$ is a *consumption plan*, if

$$c \in \mathcal{C} \triangleq \left\{c \in \mathcal{D}^+ : E\left[\int_0^T c_t^p dt + c_T^p\right] < \infty \quad \text{for all } p \in \mathbb{R}\right\}.$$

Here, we denote the space of square-integrable, progressively measurable processes by

$$\mathcal{D} \triangleq \left\{X = (X_t)_{t\in[0,T]} \text{ progressive} : E\left[\int_0^T X_t^2 dt + X_T^2\right] < \infty\right\},$$

and we write $\mathcal{D}^+ \triangleq \{X \in \mathcal{D} : X_t > 0 \text{ for } dt\text{-a.e. } t \in [0,T]\}$ for its strictly positive cone. ◇

The agent's preferences on $\mathcal{C}$ are described by a utility index $\nu : \mathcal{C} \to \mathbb{R}$, that is,

$$c \in \mathcal{C} \text{ is weakly preferred to } \bar{c} \in \mathcal{C} \quad \text{if and only if} \quad \nu(c) \geqslant \nu(\bar{c}),$$

see Duffie and Epstein (1992b) and Epstein and Zin (1989). To construct the Epstein-Zin utility index, let

$$\delta > 0, \quad \gamma > 0, \quad \psi > 0 \quad \text{with } \gamma, \psi \neq 1$$

be given, and put $\phi \triangleq \frac{1}{\psi}$. If $\gamma < 1$, we set $\mathfrak{U} \triangleq (0, \infty)$, and if $\gamma > 1$, we set $\mathfrak{U} \triangleq (-\infty, 0)$. Then the continuous-time Epstein-Zin aggregator is given by $f\colon \mathfrak{C} \times \mathfrak{U} \to \mathbb{R}$,

$$f(c, v) \triangleq \delta\theta v \left[\left(\frac{c}{((1-\gamma)v)^{\frac{1}{1-\gamma}}} \right)^{1-\frac{1}{\psi}} - 1 \right], \quad \text{where } \theta = \frac{1-\gamma}{1-\phi} \in \mathbb{R},\, \theta \neq 0.$$

Here, γ represents the agent's relative risk aversion, ψ is his elasticity of intertemporal substitution (EIS) and δ is his rate of time preference. One can show that for every consumption plan $c \in \mathcal{C}$, there exists a unique semimartingale V^c satisfying

$$V^c_t = E_t \left[\int_t^T f(c_s, V^c_s) ds + U(c_T) \right] \quad \text{for all } t \in [0, T], \tag{4.1}$$

where $U : \mathfrak{C} \to \mathbb{R}$, $U(x) \triangleq \varepsilon^{1-\gamma} \frac{1}{1-\gamma} x^{1-\gamma}$ is a CRRA utility function for bequest and $\varepsilon \in (0, \infty)$ is a weight factor; see Schroder and Skiadas (1999) and Theorem 3.2 in this thesis for the case $\gamma > 1$, $\psi > 1$. This leads to the following definition:

DEFINITION 4.2. The Epstein-Zin utility index $\nu : \mathcal{C} \to \mathfrak{U}$ is given by $\nu(c) \triangleq V^c_0$ where V^c is the unique process satisfying (4.1). ◇ *Utility index*

The classical time-additive utility specification

$$\nu(c) = E\left[\int_0^T e^{-\delta s} u(c_s) ds + e^{-\delta T} U(c_T) \right],$$

where $u : \mathfrak{C} \to \mathbb{R}$, $u(x) \triangleq \frac{1}{1-\gamma} x^{1-\gamma}$, is subsumed as the special case of the Epstein-Zin parametrization where $\gamma = \phi$; hence our analysis applies in particular to consumption-portfolio optimization with additive CRRA preferences and arbitrary risk aversion parameter $\gamma \neq 1$.

Remark. The specifications $\gamma = 1$ or $\phi = 1$ correspond to unit relative risk aversion or unit EIS, respectively; $\gamma = \phi = 1$ represents time-additive logarithmic utility. The case of unit EIS, $\phi = 1$, is well-understood and has been studied extensively in the literature; see, e.g., Schroder and Skiadas (2003) and Chacko and Viceira (2005). ◇

4.1.2 *Financial market model*

Two securities are traded. The first is a locally risk-free asset (e.g., a money market account) M with dynamics

$$dM_t = r(Y_t) M_t dt,$$

while the second asset (e.g., a stock or stock index) S is risky and satisfies

$$dS_t = S_t \left[(r(Y_t) + \lambda(Y_t)) dt + \sigma(Y_t) dW_t \right].$$

The interest rate $r : \mathbb{R} \to \mathbb{R}$ and the stock's excess return and volatility $\lambda, \sigma : \mathbb{R} \to \mathbb{R}$ are assumed to be measurable functions of a state process Y with dynamics

$$dY_t = \alpha(Y_t) dt + \beta(Y_t) \left(\rho dW_t + \sqrt{1-\rho^2} d\bar{W}_t \right), \quad Y_0 = y.$$

Here $\alpha, \beta : \mathbb{R} \to \mathbb{R}$ are measurable functions and $\rho \in [-1,1]$ denotes the correlation between stock returns and the state process. Throughout this chapter, we assume:

(A1) The coefficients $r, \lambda, \sigma, \alpha$ are bounded and Lipschitz continuous; the coefficient β is bounded and has a bounded Lipschitz continuous derivative.

(A2) Ellipticity condition: $\inf_{y\in\mathbb{R}} \sigma(y) > 0$ and $\inf_{y\in\mathbb{R}} \beta(y) > 0$.

The investor's wealth dynamics are given by

$$dX_t^{\pi,c} = X_t^{\pi,c}\left[(r(Y_t) + \pi_t\lambda(Y_t))dt + \pi_t\sigma(Y_t)dW_t\right] - c_t dt, \quad X_0 = x, \quad (4.2)$$

where π_t denotes the fraction of wealth invested in the risky asset at time t, the constant $x > 0$ is the investor's initial wealth, and c his consumption plan.

Admissible strategies

DEFINITION 4.3. The pair of strategies (π, c) is *admissible* for initial wealth $x > 0$ if it belongs to the set

$$\mathcal{A}(x) \triangleq \left\{(\pi, c) \in \mathcal{D} \times \mathcal{C} : X_t^{\pi,c} > 0 \text{ for all } t \in [0,T] \text{ and } c_T = X_T^{\pi,c}\right\}. \quad \diamond$$

The investor's preferences are described by a recursive utility functional of Epstein-Zin type. Hence an admissible pair of strategies $(\pi, c) \in \mathcal{A}(x)$ yields utility

$$v(c) \triangleq V_0^c, \quad \text{where } V_t^c \triangleq E_t\left[\int_t^T f(c_s, V_s^c)ds + U(X_T^{\pi,c})\right] \text{ for } t \in [0,T].$$

4.1.3 *Consumption-portfolio optimization problem*

Consumption-portfolio problem

DEFINITION 4.4. Given initial wealth $x > 0$, the investor's *consumption-portfolio problem* is to maximize utility over the class of admissible strategies $\mathcal{A}(x)$, i.e., to

$$\text{(P)} \quad \text{find} \quad (\pi^\star, c^\star) \in \mathcal{A}(x) \quad \text{such that} \quad v(c^\star) = \sup_{(\pi,c)\in\mathcal{A}(x)} v(c). \quad \diamond$$

Remark. The consumption-porfolio problem (P) has been widely studied in the literature. Schroder and Skiadas (1999) investigate the case of complete markets, and Schroder and Skiadas (2003, 2005, 2008) provide necessary and sufficient optimality conditions for general homothetic and translation-invariant preferences. Moreover, Schroder and Skiadas (2003) solve the consumption-portfolio problem for an investor with unit EIS in closed form. Chacko and Viceira (2005) obtain closed-form solutions for an investor with unit EIS in an inverse Heston stochastic volatility model, and Kraft et al. (2013) derive explicit solutions for a non-unit EIS investor whose preference parameters satisfy the condition

$$\psi = 2 - \gamma + \tfrac{(1-\gamma)^2}{\gamma}\rho^2. \quad \text{(H)}$$

Berdjane and Pergamenshchikov (2013) study the above-described consumption-portfolio problem in the special case where the investor has additive preferences with relative risk aversion $\gamma \in (0, 1)$. Figure 1 depicts the parametrizations for which solutions are known in the literature. $\diamond$

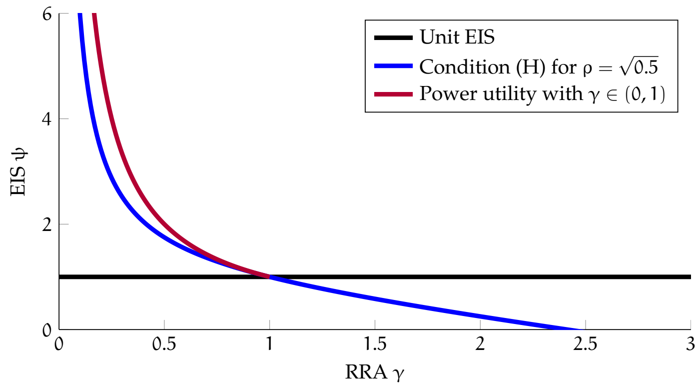

Figure 1: Combinations of RRA γ and EIS ψ for which solutions of consumption-portfolio problems with unspanned risk are known.

The HJB equation

We consider the dynamic programming equation associated with the consumption-portfolio problem (P),

$$0 = \sup_{\pi \in \mathbb{R}, c \in (0,\infty)} \Big\{ w_t + x(r + \pi\lambda)w_x - cw_x + \tfrac{1}{2}x^2\pi^2\sigma^2 w_{xx} + \alpha w_y + \tfrac{1}{2}\beta^2 w_{yy} + x\pi\sigma\beta\rho w_{xy} + f(c, w)\Big\}, \tag{4.3}$$

subject to the terminal condition $w(T, x, y) = \varepsilon^{1-\gamma}\frac{1}{1-\gamma}x^{1-\gamma}$.

By homotheticity of Epstein-Zin utility, one certainly expects solutions to take the form

$$w(t, x, y) = \tfrac{1}{1-\gamma}x^{1-\gamma}h(t, y)^k, \quad (t, x, y) \in [0, T] \times (0, \infty) \times \mathbb{R}, \tag{4.4}$$

where k is a constant and $h \in C^{1,2}([0, T] \times \mathbb{R})$ is strictly positive with $h(T, \cdot) = \hat{\varepsilon} \triangleq \varepsilon^{\frac{1-\gamma}{k}}$. Choosing $k \triangleq \frac{\gamma}{\gamma + (1-\gamma)\rho^2}$ and solving the first-order conditions leads to the following definition:

DEFINITION 4.5. The *candidate optimal strategies* are given by

$$\hat{\pi} \triangleq \frac{\lambda}{\gamma\sigma^2} + \frac{k}{\gamma}\frac{\beta\rho}{\sigma}\frac{h_y}{h} \quad \text{and} \quad \hat{c} \triangleq \delta^{\psi}h^{q-1}x, \tag{4.5}$$

where $q \in \mathbb{R}$, $q \neq 1$ is given by

$$q \triangleq 1 - \frac{\psi k}{\theta},$$

and where h is a strictly positive solution of the semilinear partial differential equation (PDE)

$$0 = h_t - \tilde{r}h + \tilde{\alpha}h_y + \tfrac{1}{2}\beta^2 h_{yy} + \tfrac{\delta^\psi}{1-q}h^q, \quad h(T,\cdot) = \hat{\varepsilon}, \tag{4.6}$$

with coefficients

$$\tilde{r} \triangleq -\tfrac{1}{k}\left[r(1-\gamma) + \tfrac{1}{2}\tfrac{1-\gamma}{\gamma}\tfrac{\lambda^2}{\sigma^2} - \delta\theta\right] \quad \text{and} \quad \tilde{\alpha} \triangleq \alpha + \tfrac{1-\gamma}{\gamma}\tfrac{\lambda\beta\rho}{\sigma}. \tag{4.7}$$

In the following, we refer to (4.6) as the *reduced HJB equation*. ◇

Note: The function h in the separation (4.4) is closely related to the candidate for the agent's optimal consumption-wealth ratio as used in, e.g., Campbell et al. (2004), Campbell and Viceira (2002), and Chacko and Viceira (2005). More precisely, by (4.5) we have $\hat{c}/x = \delta^\psi h^{-\psi k/\theta}$, so we can represent the candidate for the value function equivalently as $w(t,x,y) = x^{1-\gamma}\delta^\theta(\hat{c}/x)^{-\theta/\psi}/(1-\gamma)$.

LEMMA 4.6. *If $h \in C^{1,2}([0,T] \times \mathbb{R})$ is a strictly positive solution of (4.6), then the function given by $w(t,x,y) = \frac{1}{1-\gamma}x^{1-\gamma}h(t,y)^k$ solves the HJB equation (4.3).*

Proof. This follows by a direct calculation, see Section C.1, p. 181. □

LEMMA 4.7. *The functions $\tilde{r}$ and $\tilde{\alpha}$ are bounded and Lipschitz continuous.*

Proof. A very simple calculation, see Section C.1, p. 181. □

Remark. Note that for all $\rho \in [-1,1]$ we have

$$q = 1 - \tfrac{1-\phi}{1-\gamma}C, \qquad \text{where} \quad C \triangleq \tfrac{\psi\gamma}{\gamma(1-\rho^2)+\rho^2} > 0.$$

Thus $q < 1$ if and only if $\frac{1-\phi}{1-\gamma} > 0$, and $q > 1$ if and only if $\frac{1-\phi}{1-\gamma} < 0$; see Table 1 and Figure 2. ◇

$q < 1$	$q = 1$	$q > 1$
$\frac{1-\phi}{1-\gamma} > 0$	$\phi = 1$	$\frac{1-\phi}{1-\gamma} < 0$

Table 1: Range of the exponent q in (4.6) depending on the risk aversion γ and the reciprocal of the elasticity of intertemporal substitution ϕ.

To solve the consumption-portfolio problem (P), we proceed as follows: First, we prove the following general existence and uniqueness result for the reduced HJB equation (4.6).

Solution of HJB equation

THEOREM 4.8. *For all $\gamma, \psi, \delta > 0$ with $\gamma, \psi \neq 1$, there exists a unique solution $h \in C^{1,2}([0,T] \times \mathbb{R})$ to the reduced HJB equation (4.6) such that*

$$\underline{h} \leqslant h \leqslant \overline{h} \quad \textit{for positive constants } 0 < \underline{h} < \overline{h} \quad \textit{and} \quad \|h_y\|_\infty < \infty.$$

Proof. Lemma 4.7 shows that Theorem 4.8 is a consequence of Theorem 4.10 below. Theorem 4.10 is the first main result of this chapter, and Section 4.2 is dedicated to its proof. □

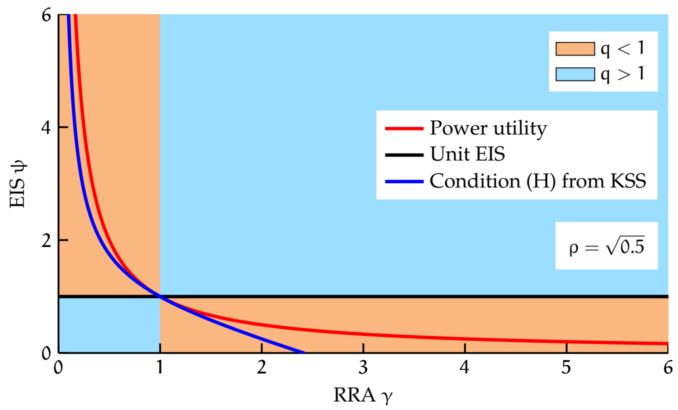

Figure 2: Range of the exponent q in (4.6) depending on the RRA γ and the EIS ψ. Condition (H) is calculated for $\rho = \sqrt{0.5}$.

Once existence and uniqueness for the reduced HJB equation (4.6) are established, we show that the associated candidate strategies are indeed optimal. This second main result is contained in Section 4.3.

THEOREM 4.9. *Let h be a solution of the reduced HJB equation (4.6) as given by Theorem 4.8. Then the corresponding candidate strategies $(\hat{\pi}, \hat{c})$,* *Solution of consumption-portfolio problem*

$$\begin{aligned} \hat{\pi}_t &= \frac{\lambda(Y_t)}{\gamma\sigma(Y_t)^2} + \frac{k}{\gamma}\frac{\beta(Y_t)\rho}{\sigma(Y_t)}\frac{h_y(t,Y_t)}{h(t,Y_t)}, && t \in [0,T), \\ \hat{c}_t &= \delta^{\psi} h(t,Y_t)^{q-1} X_t^{\hat{\pi},\hat{c}}, && t \in [0,T), \end{aligned} \tag{4.8}$$

are optimal for the consumption-portfolio problem (P).

By a slight abuse of notation, we write $\hat{\pi}_t = \hat{\pi}(t,Y_t)$ and $\hat{c}_t \triangleq \hat{c}(t,X_t^{\hat{\pi},\hat{c}},Y_t)$ for $t \in [0,T)$. This will not give rise to confusion in the following.

4.1.4 *Links to the literature*

The current chapter on optimal consumption-portfolio selection with stochastic differential utility is related to several strands of literature.

First, we contribute to the literature on dynamic incomplete-market portfolio theory. Liu (2007) considers portfolio problems with unspanned risk and time-additive utility. His framework already nests a number of popular models, including those of Kim and Omberg (1996), Campbell and Viceira (1999), Barberis (2000), and Wachter (2002), as special cases. Given the boundedness conditions (A1) and (A2), our approach can be used to generalize several of his solutions to settings where asset price dynamics are non-affine or non-quadratic and the agent has recursive utility. Recursive utility has been developed by Kreps and Porteus (1978, 1979), Epstein and Zin (1989) and Duffie and Epstein (1992b). Chacko and Viceira (2005) study a consumption-portfolio problem with affine stochastic volatility and recursive preferences. They find an explicit solution for unit

EIS and approximate the solution for non-unit EIS using the Campbell-Shiller technique. Our approach makes it possible to extend their analysis to problems with non-affine specifications of stochastic volatility, without having to rely on approximations.

Finally, our results are also related to Schroder and Skiadas (1999), who focus on complete markets, and to Schroder and Skiadas (2003), who provide necessary and sufficient optimality conditions in a general homothetic setting by duality methods and obtain explicit solutions for unit EIS.

Second, we add to the asset pricing literature by establishing a novel solution method for the agent's value function and the consumption-wealth ratio. In particular, this includes research on long-run risk and disasters (see, e.g., Bansal and Yaron (2004), Barro (2006), Benzoni et al. (2011), Gabaix (2012), Rietz (1988), Wachter (2013)).

Our mathematical approach has various ties to the literature: The verification argument used to solve the consumption-portfolio problem builds on the so-called utility gradient approach that has been developed in a series of papers by Duffie, Schroder, and Skiadas, including Duffie and Skiadas (1994) and Schroder and Skiadas (1999, 2003, 2008). We generalize the verification results in Duffie and Epstein (1992b), who derive a verification result for aggregators satisfying a Lipschitz condition, and of Kraft et al. (2013), who consider Epstein-Zin preferences under parameter restrictions. Our results are also related to the findings of Duffie and Lions (1992), who study the existence of stochastic differential utility using PDE methods, and to Marinacci and Montrucchio (2010), who establish existence and uniqueness of recursive utility in discrete time. The analysis of Berdjane and Pergamenshchikov (2013) is based on a fixed point argument related to the one we use, but is focused on the special case where the agent has time-additive utility with risk aversion below unity and the state process has constant volatility. In a recent paper that appeared after our article, Kraft, Seiferling, and Seifried (2015), had been finished, Xing (2015) addresses a closely related class of portfolio optimization problems using BSDE techniques and thus complements our analysis: He requires weaker boundedness (rsp., integrability) conditions, but does not provide information on how to determine optimal strategies. In addition, his analysis is restricted to the case of RRA and EIS both being greater than one.

Finally, our analysis also contributes to the literature on quasilinear partial differential equations (PDEs) and backward and forward-backward stochastic differential equations (BSDEs and FBSDEs, respectively). We demonstrate that the FBSDE associated with the semilinear PDE which is relevant for our applications in consumption-portfolio choice and asset pricing, admits a unique bounded solution. Importantly, the driver of this FBSDE is not Lipschitz, so standard results do not apply. We thus contribute to the growing literature on non-Lipschitz BSDEs and FBSDEs, including, among others, Kobylanski (2000), Briand and Carmona (2000), Briand and Hu (2008), and Delbaen et al. (2011). In addition, by deriving an associated Feynman-Kac representation, this paper adds to the literature that connects FBSDEs to semilinear Cauchy problems; see, e.g., Pardoux and Peng (1992), Delarue (2002) and Ma et al. (2012) and the references therein.

4.2 FEYNMAN-KAC FIXED POINT APPROACH

The goal of this section is to establish existence and uniqueness for the reduced HJB equation (4.6). Abstracting away from the financial market of Subsection 4.1.2, we present a constructive method to obtain a classical solution of the semilinear PDE

$$0 = h_t + a h_{yy} + b h_y + c h + \tfrac{d}{1-q} h^q, \quad h(T, \cdot) = \hat{\varepsilon}, \tag{4.9}$$

where the coefficients

Standing assumptions

(C1) $a, b, c : \mathbb{R} \to \mathbb{R}$ are bounded and Lipschitz continuous,

(C2) the function a has a bounded Lipschitz continuous derivative and satisfies $\inf_{y \in \mathbb{R}} a(y) > 0$,

(C3) d and $\hat{\varepsilon}$ are positive constants.

By assumptions (A1) and (A2) and Lemma 4.7, the reduced HJB equation (4.6) is the special instance of (4.9) where

$$a = \beta^2/2, \quad b = \tilde{\alpha}, \quad \text{and} \quad c = -\tilde{r}$$

are given by (4.7), in terms of the coefficients of the financial market of Subsection 4.1.2. In particular, Theorem 4.8 is a special case of

Existence and uniqueness for the semilinar PDE

THEOREM 4.10. *Under the above assumptions, the semilinear Cauchy problem* (4.9) *has a unique solution in the class*

$$\left\{ h \in C^{1,2}([0,T] \times \mathbb{R}) \;:\; \exists\, c_1, c_2 > 0 \;\text{ s.t. }\; c_1 \leqslant h \leqslant c_2 \text{ and } \|h_y\|_\infty < \infty \right\}.$$

To prove Theorem 4.10, we study the associated system of forward-backward stochastic differential equations (FBSDEs)

$$d\eta_t^{t_0,y_0} = b(\eta_t^{t_0,y_0})dt + \sqrt{2a(\eta_t^{t_0,y_0})}\,dW_t, \tag{4.10}$$

$$dX_t^{t_0,y_0} = -\left[\tfrac{d}{1-q} (X_t^{t_0,y_0})^q + c(\eta_t^{t_0,y_0}) X_t^{t_0,y_0} \right] dt + Z_t^{t_0,y_0} dW_t, \tag{4.11}$$

where $t_0 \in [0,T]$, $y_0 \in \mathbb{R}$, and $\eta_{t_0}^{t_0,y_0} = y_0$ and $X_T^{t_0,y_0} = \hat{\varepsilon}$. We will demonstrate that there exists a unique family $(X^{t,y})_{t\in[0,T]}^{y\in\mathbb{R}}$ of bounded positive solutions of this FBSDE system, and that this family yields a solution of the reduced HJB equation via the generalized Feynman-Kac formula

$$h(t,y) = X_t^{t,y} = E_t \left[\int_t^T e^{\int_t^s c(\eta_\tau^{t,y}) d\tau} \tfrac{d}{1-q} (X_s^{t,y})^q ds + \hat{\varepsilon} e^{\int_t^T c(\eta_\tau^{t,y}) d\tau} \right].$$

Remark. In this context, a natural way to think of the function h is as the fixed point of the *Feynman-Kac operator* $\Phi : C_b([0,T] \times \mathbb{R}) \to C_b([0,T] \times \mathbb{R})$,

$$(\Phi h)(t,y) \triangleq E_t\left[\int_t^T e^{\int_t^s c(\eta_\tau^{t,y})d\tau} \tfrac{d}{1-q} h(s,\eta_s^{t,y})^q ds + \hat{\varepsilon} e^{\int_t^T c(\eta_\tau^{t,y})d\tau}\right].$$

In Section 4.5, we elaborate on this perspective in detail. ⋄

The connection between semilinear PDEs and (F)BSDEs is well-established in the mathematical literature. While classical results, including Pardoux and Peng (1992) and Ma et al. (1994, 2012), impose a Lipschitz condition on the generator, recent research has focused on relaxing that assumption. Starting from Kobylanski (2000), existence and uniqueness results for BSDEs with quadratic and convex drivers have been obtained: Briand and Carmona (2000), Delarue (2002), Briand and Hu (2008) and Delbaen et al. (2011) replace the Lipschitz assumption by a so-called monotonicity condition, while retaining a polynomial growth condition. In general, however, the driver in the FBSDE system (4.10), (4.11) is neither Lipschitz, nor does it satisfy monotonicity or polynomial growth conditions; hence, results from this literature cannot be applied to that equation. By establishing suitable a priori estimates for (4.10), (4.11) and (4.9), we prove the relevant existence, uniqueness and representation results in the following.

4.2.1 *Solving the FBSDE system: A fixed point approach*

Until further notice, we fix $t_0 \in [0,T]$ and $y_0 \in \mathbb{R}$. Assumptions (C1) and (C2) guarantee that the forward equation (4.10) has a unique strong solution $\eta \triangleq \eta^{t_0,y_0}$; see, e.g., Karatzas and Shreve (1991).

For a progressively measurable process $(X)_{t\in[t_0,T]}$, we write

$$\|X\|_\infty = \operatorname{ess\,sup}_{dt\otimes P} |X_t|$$

and denote by $\mathcal{D}_\infty$ the space of all progressively measurable processes $(X_t)_{t\in[t_0,T]}$ with $\|X\|_\infty < \infty$. Clearly, $(\mathcal{D}_\infty, \|\cdot\|_\infty)$ forms a Banach space upon identifying processes that coincide $dt \otimes P$-a.e. In the following, we construct a fixed point of the operator

Fixed point operator

$$\Psi : A^q \subset \mathcal{D}_\infty \to \mathcal{D}_\infty, \quad X \mapsto \Psi X$$

defined on its domain, the closed subset

$$A^q \triangleq \left\{X \in \mathcal{D}^\infty \,:\, (1-q)X \geqslant (1-q)\hat{\varepsilon} e^{\operatorname{sign}(q-1)T\|c\|_\infty} \quad dt \otimes P\text{-a.e.}\right\},$$

via the formula

$$(\Psi X)_t \triangleq E_t\left[\int_t^T e^{\int_t^s c(\eta_\tau)d\tau} \tfrac{d}{1-q}(0 \vee X_s)^q ds + \hat{\varepsilon} e^{\int_t^T c(\eta_\tau)d\tau}\right].$$

LEMMA 4.11. *The operator $\Psi : A^q \to A^q$ is well-defined.*

Proof. For $X \in \mathcal{A}^q$ and $q < 0$, we have $X \geqslant \hat{\varepsilon}e^{-T\|c\|_\infty}$ and hence

$$(0 \vee X)^q \leqslant \hat{\varepsilon}^q e^{-qT\|c\|_\infty}, \quad \text{whereas} \quad (0 \vee X)^q \leqslant \|X\|_\infty^q \quad \text{if } q > 0.$$

Thus $(0 \vee X)^q \in \mathcal{D}^\infty$ and

$$M_t \triangleq E_t\left[\int_{t_0}^T e^{\int_{t_0}^s c(\eta_\tau)d\tau} \frac{d}{1-q}(0 \vee X_s)^q ds + \hat{\varepsilon} e^{\int_{t_0}^T c(\eta_\tau)d\tau}\right] \tag{4.12}$$

defines a bounded martingale. Therefore the process

$$(\Psi X)_t = e^{-\int_{t_0}^t c(\eta_\tau)d\tau}\left[M_t - \int_{t_0}^t e^{\int_{t_0}^s c(\eta_\tau)d\tau} \frac{d}{1-q}(0 \vee X_s)^q ds\right]$$

has a continuous modification. In particular, $\Psi X \in \mathcal{D}^\infty$. Finally,

$$\begin{aligned}(1-q)(\Psi X)_t &= E_t\left[\int_t^T d e^{\int_t^s c(\eta_\tau)d\tau}(0 \vee X_s)^q ds + (1-q)\hat{\varepsilon} e^{\int_t^T c(\eta_\tau)d\tau}\right] \\ &\geqslant (1-q)\, E_t\left[\hat{\varepsilon} e^{\int_t^T c(\eta_\tau)d\tau}\right] \geqslant (1-q)\hat{\varepsilon} e^{\operatorname{sign}(q-1)T\|c\|_\infty},\end{aligned}$$

and hence $\Psi X \in \mathcal{A}^q$. □

Fixed points of the operator Ψ yield solutions of the forward-backward system.

LEMMA 4.12. *Let $X \in \mathcal{A}^q$. Then $\Psi X = X$ if and only if X solves the BSDE*

$$dX_t = -\left[\frac{d}{1-q}(0 \vee X_t)^q + c(\eta_t)X_t\right] dt + dN_t, \quad X_T = \hat{\varepsilon}, \tag{4.13}$$

with some L^2-martingale N. In particular, if $\Psi X = X$ is positive, then X solves the backward equation (4.11).

Proof. Let $X \in \mathcal{A}^q$ with $\Psi X = X$, and let M be the bounded martingale from (4.12). Then $Y_t \triangleq e^{\int_{t_0}^t c(\eta_\tau)d\tau} X_t$ satisfies

$$Y_t = M_t - \int_{t_0}^t e^{\int_{t_0}^s c(\eta_\tau)d\tau} \frac{d}{1-q}(0 \vee X_s)^q ds.$$

Integrating by parts, it follows that X solves (4.13) with

$$dN_t = e^{-\int_{t_0}^t c(\eta_\tau)d\tau} dM_t,$$

where N is an L^2-martingale by Burkholder's inequality. If X is positive, then $X = 0 \vee X$ and thus X also solves (4.11).

Conversely, if $X \in \mathcal{A}^q$ solves (4.13), then $Y \triangleq e^{\int_{t_0}^t c(\eta_\tau)d\tau} X_t$ satisfies

$$dY_t = -e^{\int_{t_0}^t c(\eta_\tau)d\tau} \frac{d}{1-q}(0 \vee X_t)^q dt + e^{\int_{t_0}^t c(\eta_\tau)d\tau} dN_t,$$

where $e^{\int_{t_0}^t c(\eta_\tau)d\tau} dN_t$ is an L^2-martingale; thus, integrating from t to T and taking conditional expectations, we get $Y_t = e^{\int_{t_0}^t c(\eta_\tau)d\tau}(\Psi X)_t$ and hence $X = \Psi X$. □

Our construction of a fixed point of Ψ is based on the following ramification of the classical Banach fixed point argument.

Fixed point iteration in $\mathcal{D}_\infty$

PROPOSITION 4.13. *Let $S : A \to A$ be an operator on a closed, non-empty subset A of $\mathcal{D}_\infty$ and assume that there are constants $c > 0$, $\rho \geqslant 0$ such that, for all $X, Y \in A$, we have a Lipschitz condition of the form*

$$|(SX)_t - (SY)_t| \leqslant c \int_t^T E_t \left[e^{(s-t)\rho} |X_s - Y_s| \right] ds \quad \text{for all } t \in [t_0, T].$$

Then S has a unique fixed point $X \in A$. Moreover, the iterative sequence $X_{(n)} \triangleq SX_{(n-1)}$ $(n = 1, 2, \dots)$ with an arbitrarily chosen $X_{(0)} \in A$ satisfies

$$\|X_{(n)} - X\|_\infty \leqslant e^{T\rho} (\|X_{(0)}\| + \|X\|_\infty) \left(\tfrac{ecT}{n}\right)^n \quad \text{for all } n > cT.$$

Proof. The proof is provided in Appendix C.1. □

We come to the main result of Section 4.2.

Fixed point and convergence

THEOREM 4.14. *For all $t_0 \in [0, T]$ and $y_0 \in \mathbb{R}$, there exists*

- *a unique progressive process $X^{t_0, y_0} \in \mathcal{D}_\infty$ which solves (4.11).*

Moreover, there are

- *positive constants $0 < \underline{h} < \overline{h}$ such that*

$$\underline{h} \leqslant X^{t_0,y_0} \leqslant \overline{h} \quad \text{for all } (t_0, y_0) \in [0, T] \times \mathbb{R}, \text{ and} \tag{4.14}$$

- *positive constants $K, k > 0$ such that*

$$\|X^{t_0,y_0}_{(n)} - X^{t_0,y_0}\|_\infty \leqslant K \left(\tfrac{k}{n}\right)^n \quad \text{for all } n > \tfrac{k}{e} \tag{4.15}$$

and all $(t_0, y_0) \in [0, T] \times \mathbb{R}$, where

$$X^{t_0,y_0}_{(0)} \triangleq \hat{\varepsilon}, \quad \text{and} \quad X^{t_0,y_0}_{(n)} \triangleq \Psi X^{t_0,y_0}_{(n-1)}, \quad n \in \mathbb{N}.$$

The constants $K, k > 0$ are explicitly given by $K \triangleq e^{T\|c\|_\infty}(\hat{\varepsilon} + \overline{h})$ and

$$\begin{aligned} k &\triangleq eT \left|\tfrac{qd}{1-q}\right| \left(\hat{\varepsilon} e^{T\|c\|_\infty}\right)^{q-1} && \text{for } q > 1 \text{ and} \\ k &\triangleq eT \left|\tfrac{qd}{1-q}\right| \left(\hat{\varepsilon} e^{T\|c\|_\infty}\right)^{1-q} && \text{for } q < 1, \end{aligned} \tag{4.16}$$

and the constants $\overline{h} > \underline{h} > 0$ are given explicitly by

$$\begin{aligned} &\underline{h} \triangleq \hat{\varepsilon} \exp\left(-\tfrac{k}{e} - T\|c\|_\infty\right) && \text{and} && \overline{h} \triangleq \hat{\varepsilon} \exp\left(T\|c\|_\infty\right) && \text{if } q > 1 \\ &\underline{h} \triangleq \hat{\varepsilon} \exp\left(-T\|c\|_\infty\right) && \text{and} && \overline{h} \triangleq \hat{\varepsilon} \exp\left(\tfrac{k}{e} + T\|c\|_\infty\right) && \text{if } q < 1. \end{aligned}$$

With Lemma 4.11, we have already seen that $\Psi : A^q \to A^q$ is defined on a closed non-empty subset of $\mathcal{D}^\infty$. To apply Proposition 4.13 and obtain a unique fixed point and the convergence statement, we thus need to verify the uniform Lipschitz condition.

Lipschitz continuity of Ψ

LEMMA 4.15. *With k given by (4.16), for all $X, Y \in A^q$, we have*

$$\left|(\Psi X)_t - (\Psi Y)_t\right| \leqslant \tfrac{k}{eT} \int_t^T E_t \left[e^{(s-t)\|c\|_\infty} |X_s - Y_s| \right] ds, \quad t \in [t_0, T]. \tag{4.17}$$

Proof. For $X, Y \in \mathcal{A}^q$, we immediately get

$$|(\Psi X)_t - (\Psi Y)_t| \leqslant E_t \left[\int_t^T e^{\int_t^s \|c\|_\infty d\tau} \frac{d}{|1-q|} \left|(0 \vee X_s)^q - (0 \vee Y_s)^q\right| ds \right]. \quad (4.18)$$

By definition of $\mathcal{A}^q$, the processes $0 \vee X$ and $0 \vee Y$ take values in $I_1 \triangleq [0, \hat{\varepsilon} e^{T\|c\|_\infty}]$ if $q > 1$ and values in $I_2 \triangleq [\hat{\varepsilon} e^{-T\|c\|_\infty}, \infty)$ if $q < 1$. In the first case, the function $\varphi(x) = (0 \vee x)^q$ is Lipschitz continuous on I_1 with constant $L_1 \triangleq q(\hat{\varepsilon} e^{T\|c\|_\infty})^{q-1}$. In the second case, it is Lipschitz continuous on I_2 with constant $L_2 \triangleq |q|(\hat{\varepsilon} e^{T\|c\|_\infty})^{1-q}$. Now, by (4.16),

$$k/(eT) = \left|\frac{qd}{1-q}\right| \left(\hat{\varepsilon} e^{T\|c\|_\infty}\right)^{q-1} = \frac{d}{|1-q|} L_1 \qquad \text{if } q > 1, \text{ and}$$
$$k/(eT) = \left|\frac{qd}{1-q}\right| \left(\hat{\varepsilon} e^{T\|c\|_\infty}\right)^{1-q} = \frac{d}{|1-q|} L_2 \qquad \text{if } q < 1.$$

Thus we have

$$\left|(0 \vee X_s)^q - (0 \vee Y_s)^q\right| \leqslant \frac{k}{d} \frac{|1-q|}{eT} |X_s - Y_s|, \qquad s \in [t_0, T], \quad (4.19)$$

and (4.18) implies (4.17). □

Now, Proposition 4.13 applies to Ψ and yields nearly all of Theorem 4.14. Only the uniform lower bound for $q > 1$ and the uniform upper bound for $q < 1$ require an additional argument.

LEMMA 4.16. *Let $X \in \mathcal{A}^q$ with $\Psi X = X$. Then*

Additional bounds

$$X \geqslant \underline{h} \quad \text{if } q > 1, \quad \text{and} \quad X \leqslant \overline{h} \quad \text{if } q < 1.$$

Proof. If $X \in \mathcal{A}^q$ with $\Psi X = X$, then, by Lemma 4.12, X solves the BSDE

$$dX_t = -\left[\frac{d}{1-q}(0 \vee X_t)^q + c(\eta_t) X_t\right] dt + dN_t, \quad X_T = \hat{\varepsilon},$$

where N is an L^2-martingale. Thus for every stopping time τ and all $t \in [t_0, T]$, we have

$$1_{\{\tau > t\}} X_t = E_t \left[1_{\{\tau > t\}} \int_t^\tau \left(\frac{d}{1-q}(0 \vee X_s)^q + c(\eta_s) X_s \right) ds + 1_{\{\tau > t\}} X_\tau \right].$$

Recalling inequality (4.19) from the proof of Lemma 4.15, we see that

$$\frac{d}{|1-q|}(0 \vee X_s)^q \leqslant \frac{k}{eT} 1_{\{X_s > 0\}} X_s, \qquad s \in [t_0, T]. \quad (4.20)$$

Consequently, if $q > 1$, we have

$$\frac{d}{1-q}(0 \vee X_s)^q + c(\eta_s) X_s \geqslant \alpha_s X_s, \quad \text{where } \alpha_s \triangleq -\frac{k}{eT} 1_{\{X_s > 0\}} + c(\eta_s)$$

is a bounded progressively measurable process, and therefore we get

$$1_{\{\tau > t\}} X_t \geqslant E_t \left[1_{\{\tau > t\}} \int_t^\tau \alpha_s X_s ds + 1_{\{\tau > t\}} X_\tau \right], \quad t \in [t_0, T],$$

for all stopping times τ. The stochastic Gronwall inequality from Proposition 3.22 (p. 92) yields

$$X_t \geqslant E_t \left[e^{\int_t^T \alpha_u du} \hat{\varepsilon} \right] \geqslant \hat{\varepsilon} e^{-\frac{k}{e} - T\|c\|_\infty} = \underline{h} \qquad \text{for all } t \in [t_0, T].$$

On the other hand, if $q < 1$, then (4.20) yields

$$\tfrac{d}{1-q}(0 \vee X_s)^q + c(\eta_s)X_s \leqslant \alpha_s X_s, \quad \text{where } \alpha_s \triangleq \tfrac{k}{eT} + c(\eta_s),$$

and hence

$$1_{\{\tau>t\}}X_t \leqslant E_t\left[1_{\{\tau>t\}}\int_t^\tau \alpha_s X_s ds + 1_{\{\tau>t\}}X_\tau\right], \quad t \in [t_0, T],$$

for all stopping times and a bounded progressive process α. Once again the stochastic Gronwall inequality from Proposition 3.22 applies. This time, it shows that

$$X_t \leqslant E_t\left[e^{\int_t^T \alpha_u du}\hat{\varepsilon}\right] \leqslant \hat{\varepsilon}e^{\frac{k}{e}+T\|c\|_\infty} = \overline{h}, \qquad \text{for all } t \in [t_0, T]. \qquad \square$$

Now, we collect the above results and give the

Proof of Theorem 4.14. It is clear that $\underline{h} \leqslant \hat{\varepsilon} = X^{(0)} \leqslant \overline{h}$ and thus $X^{(0)} \in A^q$; hence, Lemma 4.11 implies that $X^{(n)} \in A^q$ for each member of the sequence $X^{(n)} = \Psi X^{(n-1)}$. In particular,

$$X^{(n)} \leqslant \overline{h} \quad \text{if } q > 1 \quad \text{and} \quad X^{(n)} \geqslant \underline{h} \quad \text{if } q < 1, \quad n \in \mathbb{N}. \tag{4.21}$$

By Lemma 4.15, Proposition 4.13 applies to the operator $\Psi : A^q \to A^q$. It follows that

$$\|X_{(n)} - X\|_\infty \leqslant e^{T\|c\|_\infty}(\|X_{(0)}\| + \|X\|_\infty)\left(\tfrac{k}{n}\right)^n \quad \text{for all } n > \tfrac{k}{e}, \tag{4.22}$$

where $\Psi X = X \in A^q$ is the unique fixed point of Ψ. Together with Lemma 4.12, this proves the first claim. Moreover, inequality (4.21) and Lemma 4.16 show that $\underline{h} \leqslant X \leqslant \overline{h}$ and thus establish (4.14). Finally, since $X_{(0)} = \hat{\varepsilon}$ and $\|X\|_\infty \leqslant \overline{h}$, estimate (4.22) yields (4.15). $\square$

4.2.2 *Differentiability of the fixed point*

In this subsection, we demonstrate that the solutions X^{t_0,y_0} of (4.11) provided by Theorem 4.14 yield a solution h to the semilinear Cauchy problem (4.9)

$$0 = h_t + ah_{yy} + bh_y + ch + \tfrac{d}{1-q}h^q, \quad h(T, \cdot) = \hat{\varepsilon}.$$

For that purpose, we cut off the nonlinearity using the *a priori* estimates provided by Theorem 4.14; this leads to a PDE that is known to have a classical solution $g \in C_b^{1,2}([0,T] \times \mathbb{R})$. We then conclude by proving that $g = h$. Here and in the following, $C_b^{1,2}([0,T] \times \mathbb{R})$ denotes the Banach space of all functions $u : [0,T] \times \mathbb{R}, (t,y) \to u(t,y)$ which are once continuously differentiable with respect to t and twice continuously differentiable with respect to y, and which have finite norm

$$\|u\|_{C^{1,2}} \triangleq \|u\|_\infty + \|u_t\|_\infty + \|u_y\|_\infty + \|u_{yy}\|_\infty.$$

THEOREM 4.17. *Let X^{t_0,y_0} denote the solutions to the FBSDEs (4.11) given by Theorem 4.14 and define*

Differentiability, Probabilistic Representation

$$h(t_0,y_0) \triangleq X_{t_0}^{t_0,y_0} \quad \text{for } (t_0,y_0) \in [0,T]\times\mathbb{R}.$$

Then $h \in C_b^{1,2}([0,T]\times\mathbb{R})$, and h satisfies the semilinear PDE (4.9). Moreover, h is the unique solution of (4.9) in the class

$$\left\{h \in C^{1,2}([0,T]\times\mathbb{R}) \,:\, \exists\, c_1, c_2 > 0 \text{ s.t. } c_1 \leqslant h \leqslant c_2 \text{ and } \|h_y\|_\infty < \infty\right\}.$$

In addition, h admits the probabilistic representation

$$h(t,y) = E_t\left[\int_t^T \left(c(\eta_s^{t,y})h(s,\eta_s^{t,y}) + \tfrac{d}{1-q}h(s,\eta_s^{t,y})^q\right)ds + \hat{\varepsilon}\right]. \qquad (4.23)$$

Proof. We take $\underline{h}$ and $\overline{h}$ as in Theorem 4.14 and choose a smooth cut-off function $\varphi \in C_b^1(\mathbb{R})$ with

$$\varphi(v) = \tfrac{1}{2}\underline{h} \quad \text{for } v \leqslant \tfrac{1}{2}\underline{h}, \qquad \varphi(v) = \overline{h}+1 \quad \text{for } v \geqslant \overline{h}+1, \text{and}$$
$$\varphi(v) = v \text{ for } v \in [\underline{h},\overline{h}].$$

We set $f(v) \triangleq \frac{d}{1-q}\varphi(v)^q$ and consider the semilinear Cauchy problem

$$0 = g_t + a g_{yy} + b g_y + c g + \tfrac{d}{1-q} f(g), \quad g(T,\cdot) = \hat{\varepsilon}. \qquad (4.24)$$

The function f is clearly continuously differentiable and bounded with a bounded derivative; hence, by a classical result on semilinear PDEs there exists a (not necessarily unique) classical solution $g \in C_b^{1,2}([0,T]\times\mathbb{R})$ of equation (4.24); see, e.g., Corollary C.4 (p. 189) in this thesis, or Theorem 8.1 in Ladyženskaja et al. (1968) (p. 495).

To demonstrate that $g = h$, we fix $(t_0,y_0) \in [0,T]\times\mathbb{R}$ and set $\bar{X}_t^{t_0,y_0} \triangleq \bar{X}_t \triangleq g(t,\eta_t)$, $t \in [t_0,T]$, where $\eta \triangleq \eta^{t_0,y_0}$ is given by (4.10). By Itō's formula and (4.24) we have

$$d\bar{X}_t = -\left[f(\bar{X}_t) + c(\eta_t)\bar{X}_t\right]dt + \bar{Z}_t dW_t, \quad \bar{X}_T = \hat{\varepsilon}, \qquad (4.25)$$

where $\bar{Z}_t \triangleq g_y(t,\eta_t)\sqrt{2a(\eta_t)}$ is bounded. On the other hand, Theorem 4.14 yields a unique solution $X \triangleq X^{t_0,y_0}$ of (4.11), i.e.,

$$dX_t = -\left[\tfrac{d}{1-q}X_t^q + c(\eta_t)X_t\right]dt + Z_t dW_t, \quad X_T = \hat{\varepsilon}.$$

Since $\underline{h} \leqslant X \leqslant \overline{h}$, we have $f(X_t) = \frac{d}{1-q}X_t^q$ and therefore X also satisfies

$$dX_t = -\left[f(X_t) + c(\eta_t)X_t\right]dt + Z_t dW_t, \qquad X_T = \hat{\varepsilon}. \qquad (4.26)$$

Thus we conclude that X solves (4.25), too. Since (4.25) is a BSDE with a Lipschitz driver, the standard uniqueness result for BSDEs implies that $X = \bar{X}$; see, e.g., Theorem 2.74 (p. 68). In particular, we have $h(t_0,y_0) = X_{t_0}^{t_0,y_0} = \bar{X}_{t_0}^{t_0,y_0} = g(t_0,y_0)$.

To show uniqueness, let $u \in C^{1,2}$ be another solution of (4.9) in the class under consideration, i.e., $\|u_y\|_\infty < \infty$ and there exist positive constants $\underline{u}, \overline{u}$ such that $0 < \underline{u} \leqslant u \leqslant \overline{u}$.

Replacing $\underline{h}$ and $\overline{h}$ by $\underline{h} \wedge \underline{u}$ and $\overline{h} \vee \overline{u}$ in the first part of the proof, Itō's formula and (4.24) show that $Y_t \triangleq u(t, \eta_t)$ satisfies the BSDE

$$dY_t = -\big[f(Y_s) + c(\eta_t)Y_t\big]dt + \tilde{Z}_t dW_t, \quad Y_T = \hat{\varepsilon},$$

where $\tilde{Z}_t \triangleq u_y(t, \eta_t)\sqrt{2a(\eta_t)}$ is bounded. Recall from (4.26) that X also solves that BSDE; hence, $\bar{X} = X = Y$ by uniqueness for BSDEs, and thus $g = h = u$. □

In particular, we have completed the

Proof of Theorem 4.10. Combine Theorem 4.14 and Theorem 4.17. □

4.3 VERIFICATION

Let h be the unique solution of the reduced HJB equation (4.6), as in Theorem 4.8, and consider the associated candidate optimal strategies

$$\begin{aligned} \hat{\pi}_t &= \tfrac{\lambda(Y_t)}{\gamma\sigma(Y_t)^2} + \tfrac{k}{\gamma}\tfrac{\beta(Y_t)\rho}{\sigma(Y_t)}\tfrac{h_y(t,Y_t)}{h(t,Y_t)}, && t \in [0,T), \\ \hat{c}_t &= \delta^{\psi} h(t, Y_t)^{q-1} X_t^{\hat{\pi},\hat{c}}, && t \in [0,T). \end{aligned} \tag{4.8}$$

In this section, we verify that these candidate optimal strategies are indeed optimal for the consumption-portfolio problem (P), i.e, we prove Theorem 4.9 relying on the utility gradient approach.

4.3.1 *Abstract utility gradient approach*

Let $(\bar{\pi}, \bar{c}) \in \mathcal{A}(x)$ be a given fixed consumption-portfolio strategy (below we take the candidate solution in (4.8), but the abstract argument here does not rely on that specific choice). We put

$$\nabla_t(\bar{c}) \triangleq \begin{cases} f_c(\bar{c}_t, V_t^{\bar{c}}) & \text{if } t < T, \\ U'(\bar{c}_T) & \text{if } t = T, \end{cases}$$

and define the corresponding *utility gradient* by

$$m_t(\bar{c}) \triangleq \exp\left(\int_0^t f_v(\bar{c}_s, V_s^{\bar{c}})ds\right) \nabla_t(\bar{c}). \tag{4.27}$$

If $\bar{c}$ satisfies the integrability condition

$$E\left[\int_0^T f_c(\bar{c}_s, V_s^{\bar{c}})^p ds + \exp\left(p\int_0^T f_v(\bar{c}_s, V_s^{\bar{c}})ds\right)\right] < \infty \quad \text{for all } p > 0,$$

then we have the utility gradient inequality

$$V_0^c \leqslant V_0^{\bar{c}} + \langle m(\bar{c}), c - \bar{c}\rangle \quad \text{for all } c \in \mathcal{C}, \tag{UGI}$$

where the inner product on $\mathcal{D}$ is given by

$$\langle X, Y\rangle = E\big[\int_0^T X_t Y_t dt + X_T Y_T\big].$$

For $\gamma > 1$ and $\psi > 1$, the utility gradient inequality (UGI) is a consequence of Theorem 3.5 (on p. 79 above). For the remaining parameterizations, we refer to Lemma 2 in Schroder and Skiadas (1999) and its ramifications in Section 7 therein.

For every strategy $(\pi, c) \in \mathcal{A}(x)$, we now introduce the *deflated wealth processes*

$$Z_t^{\pi,c} \triangleq \bar{m}_t X_t^{\pi,c} + \int_0^t \bar{m}_s c_s ds, \quad \text{where } \bar{m} \triangleq m(\bar{c}).$$

With this, we can state the following general verification theorem:

THEOREM 4.18. *Suppose that the deflated wealth process $Z^{\pi,c}$ is a local martingale for every admissible strategy $(\pi, c) \in \mathcal{A}(x)$, and that $Z^{\bar{\pi},\bar{c}}$ is a true martingale. Moreover, assume that*

Abstract verification

$$E\left[\int_0^T f_c(\bar{c}_s, V_s^{\bar{c}})^p ds + \exp\left(p\int_0^T f_v(\bar{c}_s, V_s^{\bar{c}})ds\right)\right] < \infty \quad \textit{for all } p > 0.$$

Then $(\bar{\pi}, \bar{c})$ is optimal.

Proof. The utility gradient inequality (UGI) evaluated at $\bar{c}$ implies

$$V_0^c \leqslant V_0^{\bar{c}} + \langle \bar{m}, c - \bar{c} \rangle = V_0^{\bar{c}} + E\left[\int_0^T \bar{m}_s(c_s - \bar{c}_s)ds + \bar{m}_T(X_T^{\pi,c} - X_T^{\bar{\pi},\bar{c}})\right],$$

where

$$\int_0^T \bar{m}_s(c_s - \bar{c}_s)ds + \bar{m}_T(X_T^{\pi,c} - X_T^{\bar{\pi},\bar{c}}) = Z_T^{\pi,c} - Z_T^{\bar{\pi},\bar{c}}.$$

Here the process $Z^{\pi,c}$ is a positive local martingale, hence a supermartingale, while $Z^{\bar{\pi},\bar{c}}$ is a martingale by assumption. Since $X_0^{\pi,c} = X_0^{\bar{\pi},\bar{c}} = x$, we obtain

$$E[Z_T^{\pi,c} - Z_T^{\bar{\pi},\bar{c}}] \leqslant E[Z_0^{\pi,c} - Z_0^{\bar{\pi},\bar{c}}] = f_c(\bar{c}_0, V_0^{\bar{c}})(X_0^{\pi,c} - X_0^{\bar{\pi},\bar{c}}) = 0. \qquad \square$$

4.3.2 *Admissibility of the candidate solution*

In the proof of Theorem 4.9 below, we apply the abstract verification result in Theorem 4.18 to the candidate $(\hat{\pi}, \hat{c})$ given by (4.8). In the following, we thus verify that the conditions of Theorem 4.18 are satisfied for that strategy.

We first establish admissibility of $(\hat{\pi}, \hat{c})$. Suppose that h is the unique solution of the reduced HJB equation (4.6), as provided by Theorem 4.8, and let $(\hat{\pi}, \hat{c})$ be given by (4.8). For simplicity of notation we write

$$\hat{V} \triangleq V^{\hat{c}}, \qquad \hat{X} \triangleq X^{\hat{\pi},\hat{c}}, \qquad \hat{m} \triangleq m(\hat{c}),$$

for the utility process, the wealth process and the utility gradient associated with $(\hat{\pi}, \hat{c})$. The proofs of the following two results are provided in Appendix C.1.

Moments of $\hat{X}$ **LEMMA 4.19.** *The candidate optimal wealth process has all moments, i.e.,*

$$E\left[\sup_{t\in[0,T]}\hat{X}_t^p\right] < \infty \quad \textit{for all } p \in \mathbb{R}.$$

In particular, $\hat{X}_t > 0$ *for all* $t \in [0,T]$ *a.s.*

As a consequence, we can show that $\hat{c} \in \mathcal{C}$ and $\hat{V}_t = w(t, \hat{X}_t, Y_t)$, where the function $w(t,x,y) \triangleq \frac{1}{1-\gamma}x^{1-\gamma}h(t,y)^k$ solves the HJB equation (4.3) by Lemma 4.6:

Admissibility of $\hat{c}$ **LEMMA 4.20.** *Let* $V_t \triangleq w(t, \hat{X}_t, Y_t)$, $t \in [0,T]$. *Then*

$$V = \hat{V} \quad \textit{and} \quad w_x(t, \hat{X}_t, Y_t) = f_c(\hat{c}_t, \hat{V}_t).$$

Moreover, we have

$$E\left[\sup_{t\in[0,T]}|\hat{c}_t|^p + \sup_{t\in[0,T]}|\hat{V}_t|^p\right] < \infty \quad \textit{for all } p \in \mathbb{R},$$

and, in particular, $\hat{c} \in \mathcal{C}$.

COROLLARY 4.21. *The candidate* $(\hat{\pi}, \hat{c}) \in \mathcal{A}(x)$ *is admissible.*

Proof. Combine Lemmas 4.19 and 4.20. □

4.3.3 *Optimality of the candidate solution*

Next we show that the deflated wealth process $Z^{\pi,c}$ is a local martingale for every admissible consumption-portfolio strategy $(\pi, c) \in \mathcal{A}(x)$. The proofs can again be found in Appendix C.1.

Dynamics of $Z^{\pi,c}$ **LEMMA 4.22.** *For all* $(\pi, c) \in \mathcal{A}(x)$, *the deflated wealth process* $Z^{\pi,c}$ *is a local martingale with dynamics*

$$dZ_t^{\pi,c} = \hat{m}_t X_t^{\pi,c}\left[\left(\pi_t\sigma(Y_t) - \tfrac{\lambda(Y_t)}{\sigma(Y_t)}\right)dW_t + k\sqrt{1-\rho^2}\beta(Y_t)\tfrac{h_y(t,Y_t)}{h(t,Y_t)}d\bar{W}_t\right].$$

For the candidate optimal process $(\hat{\pi}, \hat{c})$, *this implies*

$$dZ_t^{\hat{\pi},\hat{c}} = \hat{m}_t\hat{X}_t\left[\left(\tfrac{1-\gamma}{\gamma}\tfrac{\lambda(Y_t)}{\sigma(Y_t)} + \tfrac{k}{\gamma}\beta(Y_t)\rho\tfrac{h_y(t,Y_t)}{h(t,Y_t)}\right)dW_t \ldots\right.$$
$$\left.\ldots + k\sqrt{1-\rho^2}\beta(Y_t)\tfrac{h_y(t,Y_t)}{h(t,Y_t)}d\bar{W}_t\right].$$

To prove Theorem 4.9, it remains to verify that $Z^{\hat{\pi},\hat{c}}$ is in fact a true martingale, and that the utility gradient inequality holds at $\hat{c}$. This is guaranteed by the following result:

LEMMA 4.23. *For any* $p > 0$ *we have*

$$E\left[\int_0^T f_c(\hat{c}_s, \hat{V}_s)^p ds + \exp\left(p\int_0^T f_v(\hat{c}_s, \hat{V}_s)ds\right)\right],\ E\left[\sup_{t\in[0,T]}|\hat{m}_t|^p\right] < \infty.$$

Moreover, the process $Z^{\hat{\pi},\hat{c}}$ *is a martingale.*

Combining the preceding results, we can complete the

Proof of Theorem 4.9. By Lemmas 4.22 and 4.23, the conditions of Theorem 4.18 are fulfilled; thus, Theorem 4.18 implies that $(\hat{\pi}, \hat{c})$ is optimal for the consumption-portfolio problem (P). □

4.4 ASSET PRICING WITH EPSTEIN-ZIN PREFERENCES

The purpose of this section is to demonstrate the significance of our PDE results from Section 4.2 for asset pricing applications. For that purpose, we introduce a model that nests a continuous-time version of the disaster model of Barro (2006) as well as (a suitably truncated version of) the model by Wachter (2013) as special cases.

Endowment process

We assume an endowment economy, populated by a representative agent. His endowment (aggregate consumption) satisfies

$$dC_t = C_{t-}[\mu(Y_t)dt + \sigma(Y_t)dW_t + (e^{Z_t} - 1)dN_t],$$

where $dY_t = \alpha(Y_t)dt + \beta(Y_t)\big(\rho dW_t + \sqrt{1-\rho^2}d\bar{W}_t\big)$, $Y_0 = y$. Here N is a counting process with intensity $\Lambda_t = \Lambda(Y_t)$. We assume that all coefficients satisfy conditions (A1) and (A2) from Section 4.1.3. The random variables Z_t are taken to be independent of W, $\bar{W}$, and N with time-invariant distribution ν. We also assume that $E^\nu[e^{(1-\gamma)Z_t}] < \infty$, where $E^\nu[\cdot]$ denotes the expectation with respect to ν (i.e., $\int e^{(1-\gamma)z}\nu(dz) < \infty$).

Value function and state-price deflator

The representative agent's utility functional is given by

$$V_t^C = E_t\left[\int_t^T f(C_s, V_s^C)ds + U(C_T)\right] \quad \text{for all } t \in [0, T],$$

where f is the continuous-time Epstein-Zin aggregator and $U(c) \triangleq \frac{\varepsilon^{1-\gamma}}{1-\gamma}c^{1-\gamma}$. Similarly as in Section 4.1.3, the agent's value function $V_t^C = w(t, C_t, Y_t)$ satisfies a PDE of the form

$$\begin{aligned} 0 = w_t + \mu c w_c + \tfrac{1}{2}c^2\sigma^2 w_{cc} + \alpha w_y + \tfrac{1}{2}\beta^2 w_{yy} + c\sigma\beta\rho w_{cy} \\ + f(c, w) + \Lambda\, E^\nu[\Delta w], \qquad w(T, c, y) = \varepsilon^{1-\gamma}\tfrac{1}{1-\gamma}c^{1-\gamma}. \end{aligned}$$

Here,

$$E^\nu[\Delta w](t, c, y) = E^\nu[w(t, ce^{Z_t}, y)] - w(t, c, y)$$

is the expected change of the value function upon a jump of the endowment process.

Note: Here, we use a finite time horizon. By choosing a large T and a suitable weight on bequest, this can be used to approximate the infinite horizon case; see Algorithm 4.2 and Section 4.6.

As in Subsection 4.1.3, the solution takes the form

$$w(t, c, y) = \tfrac{1}{1-\gamma}c^{1-\gamma}h(t, y), \quad (t, c, y) \in [0, T] \times (0, \infty) \times \mathbb{R}.$$

This leads to the following semilinear PDE for h:

$$0 = h_t + \tilde{r}h + \tilde{\alpha}h_y + \tfrac{1}{2}\beta^2 h_{yy} + \tfrac{\delta}{1-q}h^q, \quad h(T,\cdot) = \hat{\varepsilon} = \varepsilon^{1-\gamma}, \tag{4.28}$$

where $q \triangleq 1 - 1/\theta$, $\tilde{\alpha} \triangleq \alpha + (1-\gamma)\sigma\beta\rho$, and

$$\tilde{r} \triangleq (1-\gamma)\mu - \tfrac{1}{2}\gamma(1-\gamma)\sigma^2 - \delta\theta + \Lambda(E^{\nu}[e^{(1-\gamma)Z_t}] - 1).$$

Since $E^{\nu}[e^{(1-\gamma)Z_t}]$ is a time-independent constant, the PDE (4.28) is of the form (4.9); hence, it can be solved with the methods from Section 4.2. Given the solution h of (4.28), the state-price deflator m in this economy (i.e., the representative agent's utility gradient) can be expressed in closed form via

$$m_t = \delta \exp\Big(\delta \tfrac{\phi-\gamma}{1-\phi} \int_0^t h(s, Y_s)^{-\frac{1}{\theta}} ds - \delta\theta t\Big) C_t^{-\gamma} h(t, Y_t)^{1-\frac{1}{\theta}}; \tag{4.29}$$

see, e.g., Duffie and Epstein (1992a,b), Duffie and Skiadas (1994), and the utility gradient in (4.27).

Using the state-price deflator (4.29), equilibrium asset prices can be calculated in a straightforward manner. For instance, the value of the claim to aggregate consumption, i.e., the present value of all future consumption, is given by

$$P_t^C = \int_t^T E_t[\tfrac{m_s}{m_t} C_s] ds + E_t[\tfrac{m_T}{m_t} C_T].$$

In particular, see, e.g., Benzoni et al. (2011), we obtain the consumption-wealth ratio as

$$\frac{C_t}{P_t^C} = \delta h(t, Y_t)^{-\frac{1}{\theta}}.$$

4.5 PDE ITERATION APPROACH

In this section, we develop an explicit constructive method to obtain the solution of the semilinear PDE (4.9) and, in particular, the reduced HJB equation. Existence and uniqueness of the solution are guaranteed by Theorem 4.10 above. More precisely, we will show that $h_n \triangleq \Phi^n \hat{\varepsilon} \overset{n\to\infty}{\longrightarrow} h$ in $C^{0,1}$, where the operator Φ is given by

$$\Phi : D(\Phi) \subset C_b^{1,2}([0,T] \times \mathbb{R}) \to C_b^{1,2}([0,T] \times \mathbb{R}), \quad f \mapsto \Phi f$$

and $\Phi f \triangleq g$ is the unique classical solution of the linear PDE

$$0 = g_t + a g_{yy} + b g_y + c g + \tfrac{d}{1-q}(0 \vee f)^q \quad \text{with} \quad g(T,\cdot) = \hat{\varepsilon}.$$

Thus h can be determined by iteratively solving linear PDEs.

4.5.1 *PDE iteration*

Our first step is to show that the iteration of PDEs as above is feasible. Thus we verify that the operator Φ is well-defined on its domain $D(\Phi)$, where

$$D(\Phi) \triangleq \{f \in C_b^{1,2}([0,T] \times \mathbb{R}) \, : \, f \geqslant \underline{h}\} \quad \text{for } q < 1, \quad \text{and}$$
$$D(\Phi) \triangleq \{f \in C_b^{1,2}([0,T] \times \mathbb{R}) \, : \, f \leqslant \overline{h}\} \quad \text{for } q > 1,$$

and $\underline{h}, \overline{h}$ are the constants specified in Theorem 4.14.

LEMMA 4.24. *If $u \in D(\Phi)$, then there exists a unique $g \in C_b^{1,2}([0,T] \times \mathbb{R})$ that satisfies*

$$0 = g_t + a g_{yy} + b g_y + c g + \tfrac{d}{1-q}(0 \vee u)^q, \qquad g(T, \cdot) = \hat{\varepsilon}. \quad (4.30)$$

Proof. If $q < 1$ and $u \geqslant \underline{h} > 0$, then $f \triangleq \frac{d}{1-q}(0 \vee u)^q \in C_b^{1,2}([0,T] \times \mathbb{R})$. If $q > 1$ with $u \leqslant \overline{h} < \infty$, then f is Lipschitz continuous since

$$|f(t,y) - f(t',y')| \leqslant |\tfrac{d}{1-q}| q \overline{h}^{q-1} |u(t,y) - u(t',y')|.$$

In either case, by classical results, there exists a unique $g \in C_b^{1,2}([0,T] \times \mathbb{R})$ satisfying (4.30); see, e.g., Corollary C.2 (p. 188) in this thesis, or Theorem 5.1 in Ladyženskaja et al. (1968), p. 320. □

To establish the link between the iterated solutions h_n of the Cauchy problem and the stochastic processes $X_{(n)}^{t_0,y_0}$ of Section 4.2, we first record a simple uniqueness result:

LEMMA 4.25. *For every $n \in \mathbb{N}$, the process $X^{(n)} \triangleq X_{(n)}^{t_0,y_0}$ defined in Theorem 4.14 is the unique solution of the linear BSDE*

$$dX_t^{(n)} = -\left[\tfrac{d}{1-q}\left(0 \vee X_t^{(n-1)}\right)^q + c(\eta_t^{t_0,y_0}) X_t^{(n)}\right] dt + Z_t^{(n)} dW_t, \quad (4.31)$$

with terminal condition $X_T^{(n)} = \hat{\varepsilon}$.

Proof. With $\varphi \triangleq \frac{d}{1-q}(0 \vee X^{(n-1)})^q$, by definition of $X^{(n)}$, we have

$$X_t^{(n)} = E_t\left[\int_t^T e^{\int_t^s c(\eta_\tau^{t_0,y_0}) d\tau} \varphi_s ds + \hat{\varepsilon} e^{\int_t^T c(\eta_\tau^{t_0,y_0}) d\tau}\right].$$

Since φ and $c(\eta^{t_0,y_0})$ are bounded processes, Propostion 2.77 shows that $X^{(n)}$ is the unique solution of the linear backward equation $dX_t^{(n)} = -[\varphi_t + c(\eta_t^{t_0,y_0}) X_t^{(n)}] dt + Z_t^{(n)} dW_t$. □

The connection between h_n and $X_{(n)}^{t_0,y_0}$ is now given as follows:

THEOREM 4.26. *For each $n \in \mathbb{N}$, we have $h_n = \Phi^n \hat{\varepsilon} \in D(\Phi)$ and*

$$h_n(t, \eta_t^{t_0,y_0}) = \left(X_{(n)}^{t_0,y_0}\right)_t \quad \textit{for all } t \in [t_0, T], \; (t_0, y_0) \in [0,T] \times \mathbb{R}.$$

Proof. The assertion is clearly true for $n = 0$, since $h_0 = \Phi^0 \hat{\varepsilon} = \hat{\varepsilon}$ and $X^{t_0,y_0}_{(0)} = \hat{\varepsilon}$. Assume by induction that $h_{n-1} = \Phi^{n-1}\hat{\varepsilon} \in D(\Phi)$ with

$$h_{n-1}(t, \eta^{t_0,y_0}_t) = \big(X^{t_0,y_0}_{(n-1)}\big)_t \quad \text{for all } t \in [t_0, T] \tag{4.32}$$

and all $(t_0, y_0) \in [0,T] \times \mathbb{R}$. By Lemma 4.24, the function $g \triangleq h_n = \Phi h_{n-1} \in C^{1,2}_b([0,T] \times \mathbb{R})$ is well-defined and satisfies

$$0 = g_t + a g_{yy} + b g_y + c g + \tfrac{d}{1-q}(0 \vee h_{n-1})^q, \qquad g(T,\cdot) = \hat{\varepsilon}. \tag{4.33}$$

Let $(t_0, y_0) \in [0,T] \times \mathbb{R}$, let $\eta \triangleq \eta^{t_0,y_0}$ be given by (4.10), and set $X_t \triangleq g(t, \eta_t)$. By (4.32), (4.33) and Itō's formula we have

$$dX_t = -\Big[\tfrac{d}{1-q}\Big(0 \vee \big(X^{t_0,y_0}_{(n-1)}\big)_t\Big)^q + c(\eta_t)X_t\Big]\,dt + Z_t dW_t,$$

where $Z_t \triangleq g_y(t, \eta_t)\sqrt{2a(\eta_t)}$ is bounded. Consequently, X is a solution of (4.31), so, by Lemma 4.25, we must have $X = X^{t_0,y_0}_{(n)}$. Thus

$$h_n(t, \eta^{t_0,y_0}_t) = \big(X^{t_0,y_0}_{(n)}\big)_t \quad \text{for all } t \in [t_0, T],\ (t_0, y_0) \in [0,T] \times \mathbb{R}.$$

Theorem 4.14 implies $\underline{h} \leqslant X^{t_0,y_0}_{(n)}$ for $q < 1$ and $X^{t_0,y_0}_{(n)} \leqslant \overline{h}$ for $q > 1$. Hence $h_n \in D(\Phi)$, and the induction is complete. □

The convergence $h_n \to h$ is now a corollary of the analysis in Section 4.2.

COROLLARY 4.27. *Let $h \in C^{1,2}_b([0,T] \times \mathbb{R})$ be the unique solution of the semilinear Cauchy problem equation (4.9) as given by Theorem 4.10. Moreover, let $h_n \triangleq \Phi^n \hat{\varepsilon} \in C^{1,2}_b([0,T] \times \mathbb{R})$ be defined recursively as the unique bounded solution of the Cauchy problem*

$$0 = (h_n)_t + a(h_n)_{yy} + b(h_n)_y + c h_n + \tfrac{d}{1-q}(0 \vee h_{n-1})^q, \quad h_n(T,\cdot) = \hat{\varepsilon}.$$

Then, with the constants $K, k > 0$ given in (4.16), we have

$$\|h_n - h\|_\infty \leqslant K\left(\tfrac{k}{n}\right)^n \quad \textit{for all } n > \tfrac{k}{e}.$$

Proof. By Theorem 4.26, we have $h_n(t, \eta^{t_0,y_0}_t) = \big(X^{t_0,y_0}_{(n)}\big)_t$ for all $t \in [t_0, T]$ and all $(t_0, y_0) \in [0,T] \times \mathbb{R}$. Thus Theorem 4.14 yields

$$\begin{aligned} |h_n(t_0,y_0) - h(t_0,y_0)| &= \big|\big(X^{t_0,y_0}_{(n)}\big)_{t_0} - X^{t_0,y_0}_{t_0}\big| \\ &\leqslant \|X^{t_0,y_0}_{(n)} - X^{t_0,y_0}\|_\infty \leqslant K\left(\tfrac{k}{n}\right)^n \end{aligned}$$

for all $n > \frac{k}{e}$ uniformly in $(t_0, y_0) \in [0,T] \times \mathbb{R}$. □

4.5.2 *Convergence rate of the PDE iteration in $C^{0,1}$*

In this section, we use the probabilistic representation (4.23) of h to demonstrate that both h_n and $(h_n)_y$ converge uniformly to h and h_y. We also identify the relevant convergence rate.

For this, we require slightly stronger regularity conditions on the coefficients in

$$0 = h_t + ah_{yy} + bh_y + ch + \tfrac{d}{1-q}h^q, \quad h(T,\cdot) = \hat{\varepsilon}. \tag{4.9}$$

In addition, we shall assume that b has a bounded Lipschitz continuous derivative. For easy reference, we explicitly list all assumptions on the coefficients:

(C1) $a, b, c : \mathbb{R} \to \mathbb{R}$ are bounded and Lipschitz continuous.

(C2′) a and b have bounded Lipschitz continuous derivatives and satisfy $\inf_{y\in\mathbb{R}} a(y) > 0$.

(C3) d and $\hat{\varepsilon}$ are positive constants.

For the reduced HJB equation (4.6), these conditions are met if we replace (A1) by the following slightly stronger regularity condition:

(A1′) The coefficients $r, \lambda, \sigma, \alpha, \beta$ are bounded with bounded and Lipschitz continuous derivatives.

Similarly as in Lemma 4.7, this assumption guarantees that

$$b = \tilde{\alpha} \quad \text{and} \quad a = \beta^2/2$$

have a bounded Lipschitz continuous derivative.

Our assumptions on a and b imply the following estimate for the derivative of the semigroup $(P_s)_{s\in[0,T]}$ generated by $\eta^{0,\cdot}$:

Derivative of the semigroup

PROPOSITION 4.28. *Assume that* (C1) *and* (C2′) *are satisfied, and let* $(P_s)_{s\in[0,T]}$ *be the semigroup associated with the process* $\eta^{0,\cdot}$ *which is given by* (4.10)*. Then, there exists a constant* $M > 0$ *such that*

$$\|D(P_t f)\|_\infty \leqslant M t^{-\frac{1}{2}} \|f\|_\infty \quad \textit{for all } t \in [0,T] \textit{ and all } f \in C_b(\mathbb{R}).$$

Proof. See Theorem 1.5.2 in Cerrai (2001) or Theorem 3.3 in Bertoldi and Lorenzi (2005). □

Remark. We refer to Elworthy and Li (1994) and Cerrai (1996) for related results. For Hölder-continuous $f \in C_b(\mathbb{R})$, results like Proposition 4.28 are well-known in the literature on parabolic PDEs; see, e.g., Ladyženskaja et al. (1968). ◇

We are now in a position to establish the convergence of our fixed point iteration in $C_b^{0,1}([0,T]\times\mathbb{R})$, endowed with the norm $\|h\|_{C^{0,1}} \triangleq \|h\|_\infty + \|\frac{\partial}{\partial y}h\|_\infty$. This provides the rigorous basis for the numerical method in Section 4.6 below.

Convergence of derivative

THEOREM 4.29. *The functions* h_n $(n = 1, 2, \dots)$ *are uniformly bounded in* $C_b^{0,1}([0,T]\times\mathbb{R})$*, and we have*

$$\|h_n - h\|_{C^{0,1}} \leqslant 2kM\sqrt{T}\left(\|c\|_\infty \tfrac{K}{n} + \tfrac{1}{eT}\right)\left(\tfrac{k}{n-1}\right)^{n-1} \quad \textit{for all } n > \tfrac{k}{e} + 1,$$

where $K, k > 0$ *are given by* (4.16) *and* $M > 0$ *is given in Proposition* 4.28*.*

Proof. Lemma 4.25 shows that

$$(X^{t_0,y_0}_{(n)})_t = E_t\Big[\int_t^T \Big(\tfrac{d}{1-q}(0 \vee (X^{t_0,y_0}_{(n-1)})_s)^q + c(\eta_s)(X^{t_0,y_0}_{(n)})_s\Big)ds + \hat{\varepsilon}\Big]$$

for all $t \in [t_0, T]$ and all $n \in \mathbb{N}$. Moreover, Theorem 4.26 implies that $h_n(t, \eta_t^{t_0,y_0}) = (X^{t_0,y_0}_{(n)})_t$ for each $n \in \mathbb{N}$. With $f_n \triangleq \frac{d}{1-q}(0 \vee h_n)^q$, we can thus represent h_n via

$$h_n(t_0, y_0) = \int_0^{T-t_0} (P_s \tilde{h}_n(t_0, s, \cdot))(y_0) ds + \hat{\varepsilon},$$

where $(P_s)_{s\in[0,T]}$ denotes the semigroup corresponding to $\eta^{0,\cdot}$, and where $\tilde{h}_n(t, s, y) \triangleq f_{n-1}(s+t, y) + c(y)h_n(s+t, y)$. Analogously, by Theorem 4.17, we obtain

$$h(t_0, y_0) = \int_0^{T-t} (P_s \tilde{h}(t_0, s, \cdot))(y) ds + \hat{\varepsilon},$$

where $\tilde{h}(t, s, \cdot) \triangleq \frac{d}{1-q} h(s+t, \cdot)^q + ch(s+t, \cdot)$. Setting $v_n \triangleq \tilde{h}_n - \tilde{h}$, we thus have

$$h_n(t_0, \cdot) - h(t_0, \cdot) = \int_0^{T-t_0} P_s v_n(t_0, s, \cdot) ds.$$

With $K, k > 0$ given by (4.16), inequality (4.19) from the proof of Lemma 4.15 (p. 118) yields

$$\begin{aligned}\|v_n\|_\infty &\leqslant \|c\|_\infty \|h_n - h\|_\infty + \tfrac{d}{|1-q|}\tfrac{k}{d}\tfrac{|1-q|}{eT}\|h_{n-1} - h\|_\infty \\ &\leqslant \|c\|_\infty K\big(\tfrac{k}{n}\big)^n + \tfrac{k}{eT}\big(\tfrac{k}{n-1}\big)^{n-1},\end{aligned}$$

where the last inequality follows from Corollary 4.27.

Now, Proposition 4.28 implies

$$\|\tfrac{\partial}{\partial y} h_n(t_0, \cdot) - \tfrac{\partial}{\partial y} h(t_0, \cdot)\|_\infty \leqslant M\|v_n\|_\infty \int_0^{T-t_0} \tfrac{1}{\sqrt{s}} ds \leqslant 2\sqrt{T} M \|v_n\|_\infty,$$

and the proof is complete. □

4.6 NUMERICAL RESULTS

4.6.1 *User's guide*

Before we study specific applications, we provide a general outline that explains how to apply our theoretical results to concrete consumption-portfolio problems and asset pricing models. By Theorem 4.9, the solution to the consumption-portfolio problem (P) is given by the optimal policies $(\hat{\pi}, \hat{c})$ in (4.8). These depend on the solution of the reduced HJB equation

$$0 = h_t - \tilde{r}h + \tilde{\alpha} h_y + \tfrac{1}{2}\beta^2 h_{yy} + \tfrac{\delta^\psi}{1-q} h^q, \quad h(T, \cdot) = \hat{\varepsilon}, \tag{4.6}$$

see also Definition 4.5. Analogously, in the asset pricing framework of Section 4.4, the state-price deflator is given by

$$m_t = \delta \exp\Big(\delta \tfrac{\phi-\gamma}{1-\phi} \int_0^t h(s, Y_s)^{-\frac{1}{\theta}} ds - \delta\theta t\Big) C_t^{-\gamma} h(t, Y_t)^{1-\frac{1}{\theta}}, \tag{4.29}$$

where h satisfies the semilinear partial differential equation

$$0 = h_t + \tilde{r}h + \tilde{\alpha}h_y + \tfrac{1}{2}\beta^2 h_{yy} + \tfrac{\delta}{1-q}h^q, \quad h(T,\cdot) = \hat{\varepsilon} = \varepsilon^{1-\gamma}. \qquad (4.28)$$

Both equation (4.6) and equation (4.28) are of the form

$$0 = h_t + ah_{yy} + bh_y + ch + \tfrac{d}{1-q}h^q, \quad h(T,\cdot) = \hat{\varepsilon}, \qquad (4.9)$$

and hence Theorem 4.10 implies that both PDEs have a unique bounded classical solution.

Algorithm 4.1 below provides a step-by-step method for the construction of solutions to PDEs of the form (4.9). This algorithm is easy to implement and relies solely on an efficient method for solving linear PDEs as a prerequisite. Consistency of this approach is guaranteed by Theorem 4.29, which demonstrates that the sequence of solutions provided by Algorithm 4.1 converges to the solution of (4.9). Theorem 4.29 also implies that the same is true for the associated derivatives. Additionally, Theorem 4.29 ensures a superexponential speed of convergence.

ALGORITHM 4.1.

(1) Set $h_0 \triangleq \hat{\varepsilon}$ and $n \triangleq 1$.

(2) Compute h_n as the solution g of the linear inhomogeneous PDE

$$0 = g_t + ag_{yy} + bg_y + cg + \tfrac{d}{1-q}(0 \vee h_{n-1})^q, \quad g(T,\cdot) = \hat{\varepsilon}. \qquad (*)$$

(3) If h_n is not sufficiently close to h_{n-1}, increase n by 1 and return to step (2).

To solve the linear PDE (*) in Step (2), we use a semi-implicit Crank-Nicolson scheme. Notice that the relevant finite-difference matrices depend on the linear part of the PDE (*) only. Therefore, the construction and LU decomposition of these matrices need to be carried out only once, in a precomputation step. This is one important feature that contributes to the excellent numerical performance of our method.

Remark. Our results require the coefficients of the state process to satisfy assumptions (A1′) and (A2). These are standard regularity conditions, but they may not be satisfied for specific models such as the Heston (1993) model below. In this case, *we implicitly understand that the relevant model has been suitably truncated* (say at stochastic volatility levels 0.0001% and 10000%), so that these conditions are satisfied. Notice that truncations of this kind are implicit in *any* numerical implementation of a possibly unbounded model on a finite grid. ◇

In some applications (e.g., asset pricing), the solution to an infinite-horizon problem is needed. In this case, the following extension of Algorithm 4.1 can be used:

ALGORITHM 4.2.

(1) Fix $\varepsilon > 0$ and a *moderate* time horizon T and set $h^0 \triangleq \hat{\varepsilon}$ and $n \triangleq 1$.

(2) Use Algorithm 4.1 to compute h^n as the solution h of the finite-horizon semilinear PDE

$$0 = h_t + ah_{yy} + bh_y + ch + \tfrac{d}{1-q}h^q, \quad h(T,\cdot) = h^{n-1}(0,\cdot).$$

(3) If $h^n(0,\cdot)$ is not sufficiently close to $h^{n-1}(0,\cdot)$, increase n by 1 and return to (2); otherwise return $h \triangleq h^n(0,\cdot)$.

In Step (1) one may take, e.g., $\varepsilon = 1$ and $T = 1$. By construction, it is clear that $g^n : [0, nT] \times \mathbb{R} \to \mathbb{R}$ with $g^n(t,y) \triangleq h^{n-k}(t-kT, y)$ for $t \in [kT, (k+1)T]$ solves

$$0 = g_t + ag_{yy} + bg_y + cg + \tfrac{d}{1-q}g^q, \quad g(nT,\cdot) = \hat{\varepsilon}.$$

Under a suitable transversality condition, the limit

$$h \triangleq \lim_{n\to\infty} g^n(0,\cdot)$$

is a solution of the infinite-horizon equation

$$0 = ah_{yy} + bh_y + ch + \tfrac{d}{1-q}h^q;$$

see Duffie and Lions (1992) or Appendix C of Duffie and Epstein (1992b) with C. Skiadas. The specific choice of ε and T becomes irrelevant in the limit $n \to \infty$.

4.6.2 *Consumption-portfolio optimization with stochastic volatility*

Generalized square-root and GARCH diffusions

We first illustrate our approach for the model specification

$$dS_t = S_t[(r + \bar{\lambda}Y_t)dt + \sqrt{Y_t}dW_t] \tag{4.34}$$

with constant interest rate r and constant $\bar{\lambda}$, i.e., we consider a stochastic volatility model with stochastic excess return. The state process satisfies

$$dY_t = (\vartheta - \kappa Y_t)dt + \bar{\beta}Y_t^p(\rho dW_t + \sqrt{1-\rho^2}d\bar{W}_t) \tag{4.35}$$

with mean reversion level ϑ/κ, mean reversion speed κ, and $p \in [0.5, 1]$. For $p = 0.5$ we obtain the Heston (1993) model and for $p = 1$ a GARCH diffusion model. Note that closed-form solutions for consumption-portfolio problems with such dynamics are only available in the special case $p = 0.5$, but solely with specific parameter choices. Further note that for $p > 0.5$ the model is not affine, and explicit solutions cannot be expected. The model coefficients are chosen as follows:

$$r = 0.02,\ \kappa = 5,\ \frac{\vartheta}{\kappa} = 0.15^2,\ \bar{\lambda} = 3.11,\ \rho = -0.5,\ \text{and } \bar{\beta} = 0.25, \quad (4.36)$$

so that for $p = 0.5$ the calibration is similar to that of Liu and Pan (2003). Furthermore, we assume that the agent's rate of time preference is $\delta = 0.05$ and that his bequest motive is $\varepsilon = 1$. The time horizon is set to $T = 10$ years. We begin with numerical examples for the Heston model (i.e., $p = 0.5$ in (4.35)); unless explicitly stated otherwise, all of the following figures are based on a Heston model with parameters (4.36).

COMPUTATIONAL EFFICIENCY The theoretical convergence rate identified in Theorem 4.29 materializes quickly in practice. Typical running times for the solutions reported below are well under 5 seconds.[1] To quantify the convergence speed, Figure 3 depicts the logarithmic relative deviations

$$\log_{10}\left(\frac{\|h_n - h_{n-1}\|_\infty}{\|h_{n-1}\|_\infty}\right) \quad \text{and} \quad \log_{10}\left(\frac{\|\frac{\partial}{\partial y}h_n - \frac{\partial}{\partial y}h_{n-1}\|_\infty}{1 + \|\frac{\partial}{\partial y}h_{n-1}\|_\infty}\right) \quad (4.37)$$

as a function of the number of iterations n. Figure 3 clearly illustrates the superlinear convergence of our method. Figure 4 shows the convergence of Algorithm 4.1. We plot the intermediate solutions after $n = 1, 2, \ldots, 5, 10, 15$ steps of the iteration. It is apparent that the algorithm converges quickly: After $n = 5$ steps the solution is visually indiscernible from subsequent iterations; the solutions for $n \geqslant 15$ are even numerically indistinguishable.

OPTIMAL STRATEGIES Figure 5 illustrates the optimal consumption-wealth ratio $(c/x)^\star$ at time $t = 0$, as a function of initial volatility

1 Machine: Intel® Core™ i3-540 Processor (4M Cache, 3.06 GHz), 4 GB RAM.

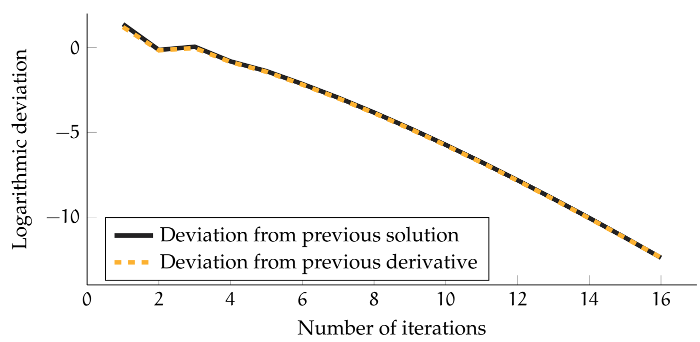

Figure 3: *Logarithmic deviation from previous solution.* This figure depicts the convergence speed (4.37) of the value function.

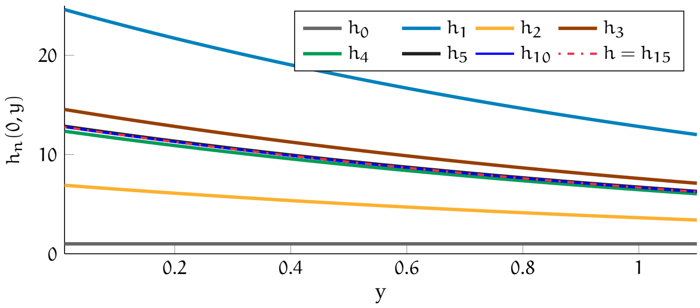

Figure 4: *Approximation after n iteration steps.* The functions h_n described in Algorithm 4.1 converge to the solution h of (4.6).

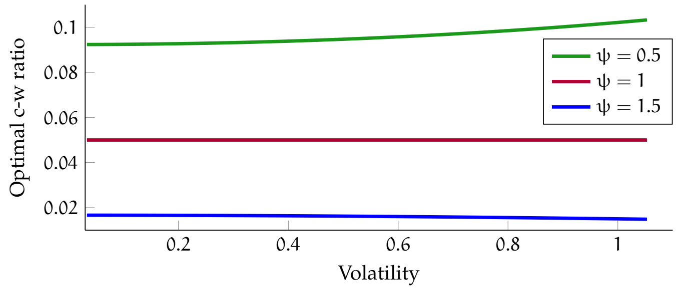

Figure 5: *Optimal consumption-wealth ratio* $(c/x)^*$ at time $t = 0$ as a function of initial volatility σ_0 for a risk aversion of $\gamma = 5$.

σ_0 for a risk aversion of $\gamma = 5$ and an EIS of $\psi \in \{0.5, 1, 1.5\}$. For reasonable risk aversions, the optimal stock allocations as a function of σ_0 are almost flat. For instance, for $\gamma \in \{3, 4, 5, 6, 10\}$ and $\psi = 0.5$ the demands vary between about 110% and 30%.

COMPARISON WITH KNOWN SOLUTIONS Figure 6 shows a range of solutions of (4.6) as the EIS ψ varies. Here we have chosen $\gamma = 2$ so that for $\psi = 0.125$ (the lowest graph in Figure 6) an explicit solution is available (see Kraft et al. (2013)). For $\psi = 1$, we use the finite-horizon analog of the explicit solution in Chacko and Viceira (2005). The solutions for the other values of EIS are computed by applying Algorithm 4.1. Note that Figure 6 depicts $g \triangleq h^{\frac{k}{1-\gamma}}$ so that the value function can be represented as

$$w(t, x, y) = \frac{1}{1-\gamma} x^{1-\gamma} h(t, y)^k = \frac{1}{1-\gamma} \left(g(t, y) x\right)^{1-\gamma}.$$

In this context, g can be interpreted as a cash multiplier.

Finally, we present comparative statics for the model (4.34) where we vary the power p. Figure 7 shows the value of the stock demand π^* at time $t = 0$ as a function of the initial volatility σ_0 and the power p; here, $\gamma = 5$ and $\psi = 1.5$.

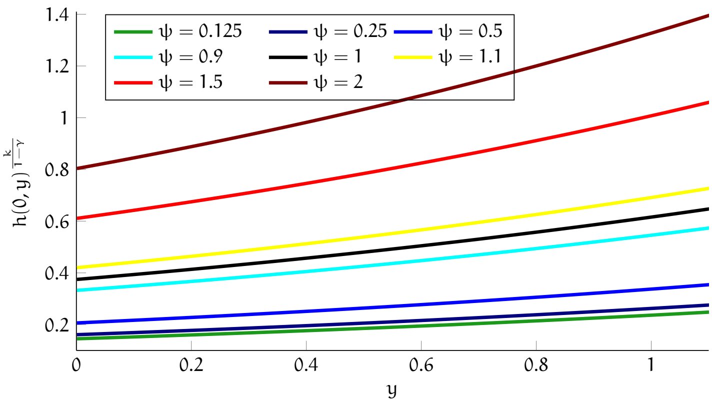

Figure 6: *Value function for different EIS.* This figure compares the function $h^{\frac{k}{1-\gamma}}$ at time $t = 0$ for a RRA of $\gamma = 2$.

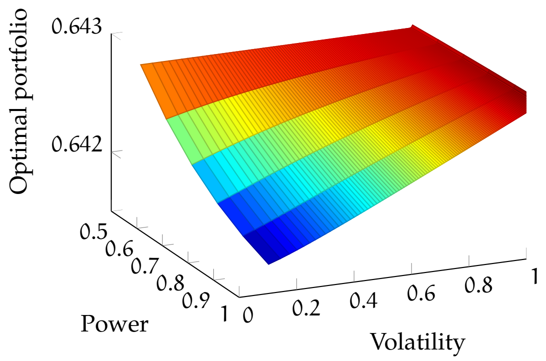

Figure 7: *Optimal stock demand and power.* This figure depicts the optimal stock demand $\pi^\star$ at time $t = 0$ as a function of initial volatility σ_0 and the power p. The model is (4.34) so $p = 0.5$ corresponds to the Heston model. The calibration is given by parameters (4.36), the agent's RRA is $\gamma = 5$ and his EIS is $\psi = 1.5$.

Exponential Vašíček model

As another application, we consider a stochastic volatility model where the volatility is lognormally distributed. The asset price dynamics are

$$dS_t = S_t[(r + \bar{\lambda} e^{2Y_t})dt + e^{Y_t} dW_t]$$

with interest rate $r = 0.05$ and $\bar{\lambda} = 3.11$. The state process

$$dY_t = (\vartheta - \kappa Y_t)dt + \bar{\beta}(\rho dW_t + \sqrt{1-\rho^2} d\bar{W}_t)$$

is an Ornstein-Uhlenbeck process with mean reversion speed $\kappa = 5$ and mean reversion level $\vartheta/\kappa = -1.933$. The correlation is set to $\rho = -0.5$ and we put $\bar{\beta} = 0.57$. These parameters are chosen in such a way that the long-term mean and variance of the squared-volatility process $\sigma_t = e^{2Y_t}$ coincide with those of the squared volatility process in the Heston model (4.35) calibrated according to (4.36). We continue to use the time preference rate $\delta = 0.05$ and bequest weight $\varepsilon = 1$. Unless stated otherwise, from now on, all figures are based on the above exponential Vašíček model with parameters

$$\kappa = 5, \quad \vartheta/\kappa = -1.933, \quad \rho = -0.5 \quad \text{and} \quad \bar{\beta} = 0.57.$$

OPTIMAL STRATEGIES Figure 8 depicts the optimal consumption-wealth ratio at time $t = 0$ as a function of initial volatility for a risk aversion of $\gamma = 5$ and an EIS of $\psi \in \{0.5, 1, 1.5\}$.

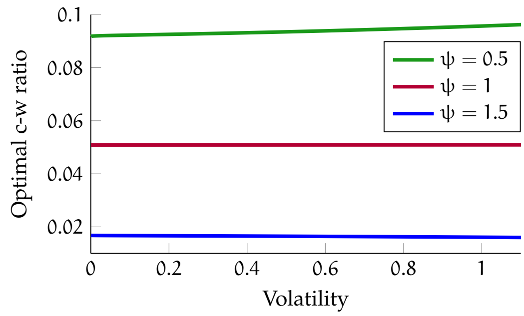

Figure 8: *Optimal consumption-wealth ratio* $(c/x)^\star$ at time $t = 0$ as a function of initial volatility σ_0 for $\gamma = 5$.

Figure 9 shows optimal stock allocations as a function of initial volatility for $\gamma \in \{3, 4, 5, 6, 10\}$ and $\psi = 0.5$.

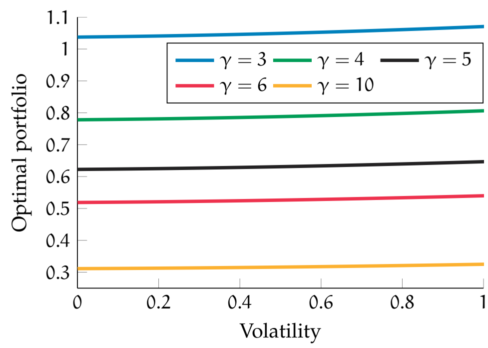

Figure 9: *Optimal stock demand and risk aversion.* The optimal stock allocation $\pi^\star$ at time $t = 0$ is shown as a function of initial volatility σ_0 for different values of the RRA γ and an EIS of $\psi = 0.5$.

4.6.3 *Asset pricing in disaster models*

Generalized square-root and GARCH diffusions

In this subsection, we illustrate our general approach for disaster models, which play an important role in asset pricing (see, e.g., Barro (2006)). The endowment process is given by

$$dC_t = C_{t-}\left[\mu dt + \sigma dW_t + (e^{Z_t} - 1)dN_t\right], \tag{4.38}$$

where N is a counting process with intensity $\lambda_t = Y_t$. For $p \in [0.5, 1]$ the state process Y satisfies

$$dY = \kappa(\bar{\lambda} - Y_t)dt + \bar{\beta} Y_t^p (\rho dW_t + \sqrt{1-\rho^2} d\bar{W}_t) \qquad (4.39)$$

with mean reversion speed $\kappa = 0.080$ and mean reversion level $\bar{\lambda} = 0.0355$. Moreover, we set $\mu = 0.0252$, $\sigma = 0.02$ and $\bar{\beta} = 0.067$. The time preference rate is $\delta = 0.012$. The random variables Z_t that model the sizes of disaster events are independent of W, $\bar{W}$ and N and satisfy $E^{\nu}[e^{(1-\gamma)Z_t}] = e^{(1-\gamma)0.15}$. The parameters are calibrated such that for $p = 0.5$ the model of Wachter (2013) obtains. Until stated otherwise, all figures that follow are based on (4.38), (4.39) with the above parameterization. Moreover, we fix $p = 0.5$ unless stated otherwise.

In the following, we present results for an infinite-horizon economy by applying Algorithm 4.2. Depending on the choice of the model parameters, typical computation times until a steady state is reached vary between 30 and 90 seconds.[2] To demonstrate the efficiency of the algorithm, we first study the convergence to the steady state for bequest motives $\varepsilon \in \{0.1, 1, 10\}$. Figure 10 shows the maximal distance of the corresponding finite time-horizon PDE solution to the infinite-horizon stationary solution if $\gamma = 3$, $\psi = 1.5$, and $\rho = 0$. As expected, the steady-state solutions are independent of the weight on the bequest motive.

Figure 11 depicts the consumption-wealth ratio as a function of the agent's risk aversion for an initial intensity of $\lambda_0 = \bar{\lambda}$, a correlation of $\rho = 0$, and an EIS of $\psi \in \{0.5, 1, 1.5\}$.

Figure 12 shows the consumption-wealth ratio as a function of ρ and λ_0. Here the representative agent's EIS is set to $\psi = 0.5$ and his risk aversion is $\gamma = 3$.

Finally, we analyze the influence of the power p in (4.39). Figure 13 shows the consumption-wealth ratio as a function of the power p and the initial intensity λ_0. Here we set $\gamma = 3$, $\psi = 1.5$ and $\rho = 0$. Note that for $p > 0.5$ the model fails to be affine, and closed-form solutions are not available even for unit EIS.

2 Machine: Intel® Core™ i3-540 Processor (4M Cache, 3.06 GHz), 4 GB RAM

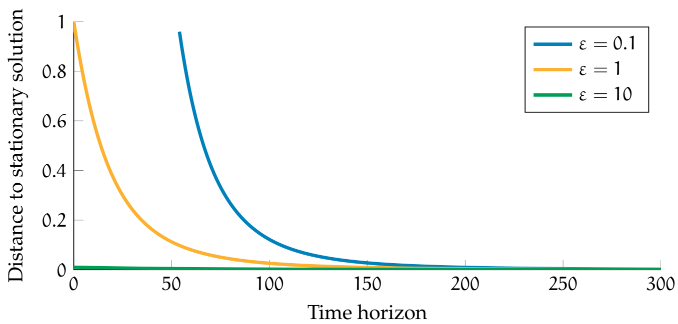

Figure 10: *Maximal distance to the stationary solution.* This figure shows the convergence speed for alternative values of the bequest motive ε.

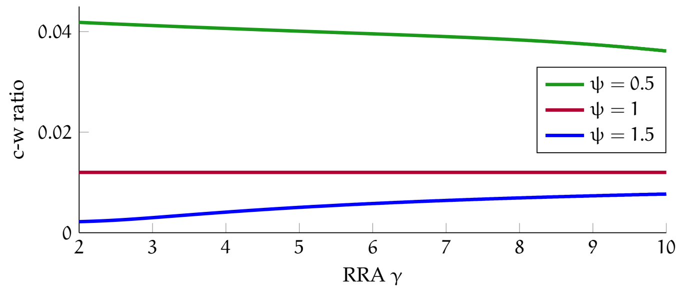

Figure 11: *Consumption-wealth ratio in Wachter's model* as a function of the RRA γ for alternative levels of EIS ψ and correlation $\rho = 0$.

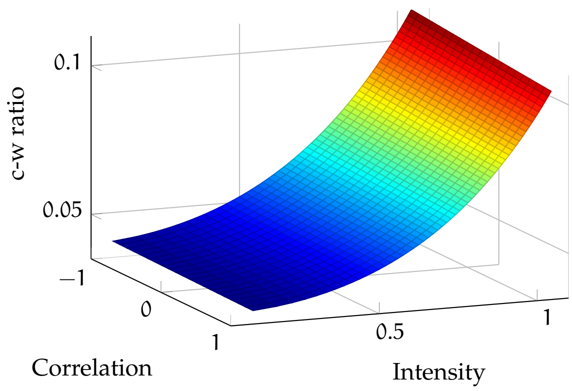

Figure 12: *Consumption-wealth ratio in Wachter's model* as a function of correlation ρ and initial intensity λ_0. The representative agent's RRA is $\gamma = 3$ and his EIS is $\psi = 0.5$.

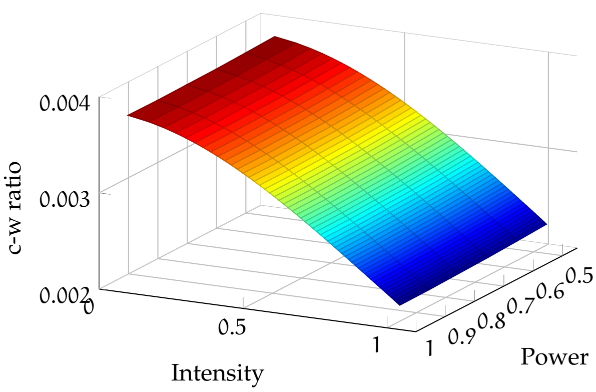

Figure 13: *Consumption-wealth ratio in the generalized square-root and GARCH models* as a function of the power p and the initial intensity λ_0. Correlation, RRA and EIS are $\rho = 0$, $\gamma = 3$ and $\psi = 1.5$.

Exponential Vašíček model

As our last application, we consider a variant of Wachter's model where the intensity process follows an exponential Vašíček process. Aggregate consumption follows the dynamics (4.38) where the counting process N has intensity $\lambda_t = e^{Y_t}$ and the state process Y satisfies $dY = (\vartheta - \kappa Y_t)dt + \bar{\beta}(\rho dW_t + \sqrt{1-\rho^2}d\bar{W}_t)$. The mean reversion speed is $\kappa = 0.080$ and the mean reversion level $\bar{y} \triangleq \vartheta/\kappa = -0.058$. Moreover, we set $\mu = 0.0252$, $\sigma = 0.02$ and $\bar{\beta} = 0.305$. These parameters are chosen such that the long-term mean and variance of the intensity process λ match those of the previous disaster model (4.39) for $p = 0.5$. The time preference rate is set to $\delta = 0.012$ and we assume $E^\nu[e^{(1-\gamma)Z_t}] = e^{(1-\gamma)0.15}$. Figure 14 depicts the consumption-wealth ratio as a function of γ for $\psi \in \{0.5, 1, 1.5\}$, $\lambda_0 = e^{\bar{y}}$, and $\rho = 0$.

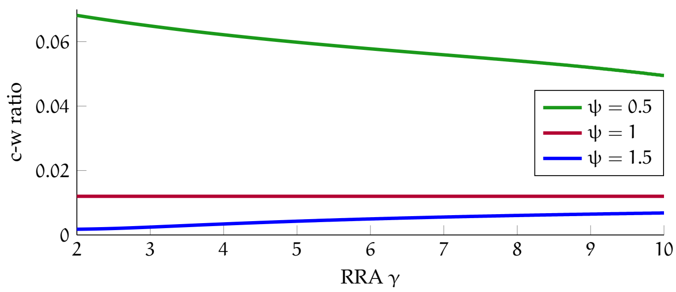

Figure 14: *Consumption-wealth ratio in exponential Vašíček model* as a function of the RRA γ for different values ψ of the EIS and $\rho = 0$.

A

APPENDIX: SUBLINEAR EXPECTATIONS

This chapter complements Chapter 2. While it provides proofs for some of the results given in Section 2.4, it aims at being essentially self-contained.

In Section A.1, an integral for measurable functions with respect to finitely additive measures is constructed. Section A.2 is concerned with sublinear expectation operators: First, sublinear expectation operators given by families of finitely additive probability measures are studied. Then, attention is restricted to sublinear expectation operators satisfying a Fatou property; completeness of associated function and process spaces is proven. Finally, we show that every sublinear expectation operator on a space of bounded random variables is given by finitely additive probabilities.

Notation

In all of Appendix A, we fix a measurable space $(\Omega, \mathcal{A})$. Unless explicitly stated otherwise, all notions requiring a measurable space are to be understood with respect to $(\Omega, \mathcal{A})$. For instance, an $\mathcal{A}$-*measurable function* $g : \Omega \to \mathbb{R}$ will simply be referred to as a real-valued *measurable* function, and the term *finitely additive probability* will signify a finitely additive probability measure on $(\Omega, \mathcal{A})$. If S is a topological space, $\mathcal{B}(S)$ denotes the Borel σ-algebra on S. We write $\mathcal{L}^0(\mathcal{A}; S)$ for the collection of all measurable functions $(\Omega, \mathcal{A}) \to (S, \mathcal{B}(S))$. The linear space $\mathcal{L}^0(\mathcal{A}; \mathbb{R})$ of all real-valued measurable functions is simply designated $\mathcal{L}^0(\mathcal{A})$. Moreover, $\mathcal{L}^\infty(\mathcal{A}) \subset \mathcal{L}^0(\mathcal{A})$ denotes the subspace of all bounded measurable functions and $\|\cdot\|_\infty$ the corresponding uniform norm.

A.1 INTEGRATION WITH RESPECT TO FINITELY ADDITIVE MEASURES

Let μ be a finite, finitely additive measure on $\mathcal{A}$. The purpose of this section is to define an integral with respect to μ for all positive measurable functions. We adapt the usual development of the Lebesgue integral (in the presentation of Rudin (1974)) to the finitely additive situation. The starting point is the unambiguously defined integral for simple functions.

Integration of simple functions

Simple function

A measurable function s is called *simple* if it takes only finitely many different values $\alpha_1, \dots \alpha_n \in [0, \infty)$. In that case, the *canonical representation* of s is

$$s = \sum_{i=1}^n \alpha_i 1_{A_i}, \qquad \text{where} \quad A_i \triangleq \{s = \alpha_i\},\ i = 1, \dots, n,$$

Integral of simple function

and the *integral of* s *(with respect to* μ*)* is defined as

$$\mu[s] \triangleq \int s d\mu \triangleq \sum_{i=1}^n \alpha_i \mu(A_i) \in \mathbb{R}.$$

The integral of s over a subset $A \in \mathcal{A}$ is denoted by

$$\int_A s d\mu \triangleq \mu[1_A s] = \sum_{i=1}^n \alpha_i \mu(A_i \cap A).$$

LEMMA A.1. *Let* s *be a simple function. Then*

$$s \cdot \mu : \mathcal{A} \to [0, \infty),\ A \mapsto \int_A s d\mu$$

is a finite, finitely additive measure on $\mathcal{A}$*.*

Proof. Suppose that s has the canonical representation $s = \sum_{i=1}^n \alpha_i 1_{A_i}$ and let $A, B \in \mathcal{A}$ be disjoint. Since μ is additive, we get

$$\begin{aligned} \int_{A \cup B} s d\mu &= \sum_{i=1}^n \alpha_i \mu(A_i \cap (A \cup B)) \\ &= \sum_{i=1}^n \alpha_i \mu(A_i \cap A) + \sum_{i=1}^n \alpha_i \mu(A_i \cap B) = \int_A s d\mu + \int_B s d\mu, \end{aligned}$$

which shows that $s \cdot \mu$ is also additive. □

Additivity of integral for simple functions

LEMMA A.2. *Let* s, t *be simple functions and* $\alpha \in [0, \infty)$*. Then*

$$\mu[s + t] = \mu[s] + \mu[t] \quad \text{and} \quad \mu[\alpha s] = \alpha \mu[s].$$

Proof. Let s and t have canonical representations $s = \sum_{i=1}^n \alpha_i 1_{A_i}$ and $t = \sum_{j=1}^m \beta_j 1_{B_j}$. For $i = 1, \dots, n$ and $j = 1, \dots, m$, we put $C_{ij} \triangleq A_i \cap B_j$ and note that Ω is the disjoint union of the sets C_{ij}. We have

$$(s + t) \cdot \mu(C_{ij}) = \int_{C_{ij}} (s + t) d\mu = \alpha_i \mu(C_{ij}) + \beta_j \mu(C_{ij})$$

as well as

$$s \cdot \mu(C_{ij}) = \alpha_i \mu(C_{ij}) \quad \text{and} \quad t \cdot \mu(C_{ij}) = \beta_j \mu(C_{ij}).$$

Now, Lemma A.1 implies that $\mu[s + t] = \mu[s] + \mu[t]$. Moreover, since αs has the canonical representation $\alpha s = \sum_{i=1}^n (\alpha \alpha_i) 1_{A_i}$, we immediately get $\mu[\alpha s] = \alpha \mu[s]$. □

Integration of non-negative functions

Integral of non-negative functions

DEFINITION A.3. The *integral* of a measurable function $f : \Omega \to [0, \infty]$ (with respect to μ) is given by

$$\mu[f] \triangleq \int f d\mu \triangleq \sup_{g \in S(f)} \mu[g],$$

where $S(f)$ denotes the collection of all simple functions $g \leqslant f$. ◇

Some immediate consequences of the above definition are collected in the following lemma:

LEMMA A.4. *For all measurable functions* $f, g : \Omega \to [0, \infty]$ *and* $A, B \in \mathcal{A}$, *the following statements hold:*

Elementary properties of the integral

(i) *If* $\mu(\{f = \infty\}) > 0$, *then* $\mu[f] = \infty$.

(ii) *If* $\mu(\{f \neq 0\}) = 0$, *then* $\mu[f] = 0$.

(iii) *If* $\mu(A) = 0$, *then* $\int_A f \, d\mu = 0$.

(iv) *If* $f \leqslant g$, *then* $\mu[f] \leqslant \mu[g]$.

(v) *If* $A \subset B$, *then* $\int_A f \, d\mu \leqslant \int_B f \, d\mu$.

(vi) *For all constants* $0 \leqslant c < \infty$, *we have* $\mu[cf] = c\mu[f]$.

Proof. If $\delta \triangleq \mu(\{f = \infty\}) > 0$, we have $h_n \triangleq 1_{\{f=\infty\}} n \in S(f)$ for all $n \in \mathbb{N}$ and thus $\mu[f] \geqslant \mu[h_n] = \delta n$ for all $n \in \mathbb{N}$, establishing (i). If $\mu(\{f > 0\}) = 0$ and $h \in S(f)$, then $\mu(\{h > 0\}) = 0$, and we see that $\mu[h] = 0$. This yields (ii), which immediately implies (iii). If $f \leqslant g$, then $S(f) \subset S(g)$ and (iv) follows; (v) is immediate from (iv).

To prove (vi), we note that $cS(f) = S(cf)$. Hence linearity of $\mu[\cdot]$ on step functions (Lemma A.2) implies

$$\mu[cf] = \sup_{h \in S(f)} \mu[cf] = \sup_{h \in S(f)} c\mu[f] = c\mu[f],$$

and positive homogeneity is established. □

Remark. The monotone convergence theorem is valid for the integral from Definition A.3 if and only if μ is countably additive: Clearly, validity of the monotone convergence theorem implies that μ is σ-continuous from below, which in turn implies countable additivity. On the other hand, if μ is countably additive, then monotone convergence is in force (see, e.g., Theorem 1.26 in Rudin (1974)) and Definition A.3 yields the classical Lebesgue integral; in particular, Fatou's Lemma holds, i.e.,

Monotone convergence, Fatou's Lemma and countable additivity

$$\int \liminf_{n\to\infty} f_n d\mu \leqslant \liminf_{n\to\infty} \int f_n d\mu \quad \text{for all mbl. fcts. } f_n : \Omega \to [0, \infty]. \quad (\odot)$$

Conversely, if $(\odot)$ holds and $A_n \in \mathcal{A}$ for each $n \in \mathbb{N}$ with $A_n \downarrow A$, then

$$\lim_{n\to\infty} \mu(A_n) = \liminf_{n\to\infty} \int 1_{A_n} d\mu \leqslant \int \liminf_{n\to\infty} 1_{A_n} d\mu = \mu(A).$$

Thus μ is continuous from above; since μ is finite, this implies that μ is σ-additive. ◇

Additivity of the integral can be extended from simple functions to arbitrary measurable functions without appealing to the monotone convergence theorem. We will make use of the following elementary approximation result for bounded measurable functions:

LEMMA A.5. *For every* $0 \leqslant f \in \mathcal{L}^\infty(\mathcal{A})$, *there exists an increasing sequence* $(f_n)_{n\in\mathbb{N}} \subset S(f)$ *converging uniformly to* f. *Moreover, for any such sequence, we have* $\mu[f] = \lim_{n\to\infty} \mu[f_n]$.

Proof. To obtain such a sequence, simply choose $f^n = 2^{-n}\lfloor 2^n f\rfloor$. Since $(f_n)_{n\in\mathbb{N}} \subset S(f)$ is increasing, monotonicity (Lemma A.4 (iv)) and the definition of $\mu[f]$ immediately yield

$$\lim_{n\to\infty} \mu[f_n] = \sup_{n\in\mathbb{N}} \mu[f_n] \leqslant \sup_{h\in S(f)} \mu[h] = \mu[f].$$

On the other hand, for every $h \in S(f)$ and every $\varepsilon > 0$, we have $h \leqslant f_n + \varepsilon$ for all but finitely many $n \in \mathbb{N}$, because $f_n \to f$ uniformly; thus, Lemma A.4 (iv) and Lemma A.2 imply

$$\mu[h] \leqslant \mu[f_n + \varepsilon] = \mu[f_n] + \varepsilon \leqslant \sup_{n\in\mathbb{N}} \mu[f_n] + \varepsilon. \qquad \square$$

Additivity of the integral

LEMMA A.6. *Let* $f, g : \Omega \to [0,\infty]$ *be measurable functions. Then*

$$\mu[f+g] = \mu[f] + \mu[g].$$

Proof. Let $r \in S(f)$ and $s \in S(g)$. Then $r + s \in S(f+g)$ and Lemma A.2 implies

$$\mu[r] + \mu[s] = \mu[r+s] \leqslant \sup_{h\in S(f+g)} \mu[h] = \mu[f+g],$$

so that $\mu[f] + \mu[g] \leqslant \mu[f+g]$.

To prove the converse inequality, let $h \in S(f+g)$ and consider the bounded functions $\tilde{f} \triangleq f \wedge \|h\|_\infty$ and $\tilde{g} \triangleq g \wedge \|h\|_\infty$. We note that $h \leqslant \tilde{f} + \tilde{g}$. Lemma A.5 yields a sequence $(f_n)_{n\in\mathbb{N}} \subset S(\tilde{f})$ with

$$f_n \to \tilde{f} \text{ uniformly}, \quad \mu[f_n] \to \mu[\tilde{f}] \quad \text{and } f_n \leqslant \tilde{f},$$

and a sequence $(g_n)_{n\in\mathbb{N}} \subset S(\tilde{g})$ such that

$$g_n \to \tilde{g} \text{ uniformly}, \quad \mu[g_n] \to \mu[\tilde{g}] \quad \text{and } g_n \leqslant \tilde{g}.$$

Hence, for every $k \in \mathbb{N}$, we find some $k \leqslant n_k \in \mathbb{N}$ such that

$$h \leqslant \tilde{f} + \tilde{g} \leqslant f_{n_k} + g_{n_k} + \tfrac{1}{k}$$

Now, Lemma A.4 (iv) and Lemma A.2 imply

$$\mu[h] \leqslant \mu[f_{n_k} + g_{n_k} + \tfrac{1}{k}] = \mu[f_{n_k}] + \mu[g_{n_k}] + \tfrac{1}{k} \to \mu[\tilde{f}] + \mu[\tilde{g}].$$

Since $\tilde{f} \leqslant f$ and $\tilde{g} \leqslant g$, another application of Lemma A.4 (iv) gives

$$\mu[h] \leqslant \mu[\tilde{f}] + \mu[\tilde{g}] \leqslant \mu[f] + \mu[g] \quad \text{for all } h \in S(f+g).$$

Therefore $\mu[f+g] = \sup_{h\in S(f+g)} \mu[h] \leqslant \mu[f] + \mu[g]$, and the proof is complete. $\square$

Combining Lemma A.4 with Lemma A.6, we obtain two corollaries.

COROLLARY A.7. *The integral operator*

$$\mu[\cdot] : \mathcal{L}^0(\mathcal{A}; [0,\infty]) \to [0,\infty], \quad f \mapsto \mu[f]$$

is increasing, positively homogeneous and additive.

COROLLARY A.8. *Let* $f, g : \Omega \to [0,\infty]$ *be measurable functions and suppose that* $\mu(f \neq g) = 0$. *Then* $\mu[f] = \mu[g]$.

Integrable functions

A measurable function $f : \Omega \to \mathbb{R}$ is μ-*integrable* if $\int |f| \, d\mu < \infty$; the collection of all μ-integrable functions is denoted by $\mathcal{L}^1(\mu) \subset \mathcal{L}^0(\mathcal{A})$. For all $f, g \in \mathcal{L}^1(\mu)$ and $\alpha, \beta \in \mathbb{R}$, Corollary A.7 implies

$$\int |\alpha f + \beta g| \, d\mu \leqslant \int \big(|\alpha||f| + |\beta||g|\big) \, d\mu = |\alpha| \int |f| \, d\mu + |\beta| \int |g| \, d\mu < \infty$$

so that $\mathcal{L}^1(\mu) \subset \mathcal{L}^0(\mathcal{A})$ forms a linear space. For $f \in \mathcal{L}^1(\mu)$, both $\int f^+ d\mu$ and $\int f^- d\mu$ exist as real numbers.

DEFINITION A.9. The integral of a μ-integrable function f is given by *Integral*

$$\mu[f] \triangleq \int f \, d\mu \triangleq \int f^+ d\mu - \int f^- d\mu \in \mathbb{R}, \quad f \in \mathcal{L}^1(\mu). \qquad \diamond$$

LEMMA A.10. *The integral*

$$\mu[\cdot] : \mathcal{L}^1(\mu) \to \mathbb{R}, \quad f \mapsto \mu[f] = \int f \, d\mu$$

is a positive linear operator.

Proof. Positivity is a consequence of Lemma A.4, and it remains to prove linearity. Note that $(-f)^+ = f^-$ and $(-f)^- = f^+$ so that

$$\mu[-f] = \mu[f^-] - \mu[f^+] = -\mu[f]. \tag{A.1}$$

For $\alpha \geqslant 0$, Corollary A.7 yields

$$\mu[\alpha f] = \alpha\mu[f^+] - \alpha\mu[f^-] = \alpha\mu[f], \tag{A.2}$$

since $(\alpha f)^+ = \alpha f^+$ and $(\alpha f)^- = \alpha f^-$. Combining (A.2) and (A.1), we get (A.2) for $\alpha < 0$; hence, it only remains to prove $\int (f + g) \, d\mu = \int f \, d\mu + \int g \, d\mu$. Write $h = f + g$ and note that

$$h^+ - h^- = h = f + g = (f^+ - f^-) + (g^+ - g^-),$$

so that

$$h^+ + f^- + g^- = f^+ + g^+ + h^-.$$

Integrating both sides with respect to μ, the claim follows from additivity for non-negative functions (Lemma A.6) upon rearranging. □

Sometimes it will be convenient to regard the integral as an operator on the convex cone

$$\mathcal{K}(\mu) = \{f \in \mathcal{L}^0(-\infty, \infty] \, : \, \mu[f^-] < \infty\}.$$

Note: As in the main text, a convex cone $K \subset \mathcal{L}^0[-\infty, \infty]$ is, by definition, a set of functions $K \subset \mathcal{L}^0[-\infty, \infty]$ such that $\alpha f + \beta g \in K$ whenever $0 \leqslant \alpha, \beta < \infty$ and $f, g \in K$.

LEMMA A.11. *The integral*

$$\mu[\cdot] : \mathcal{K}(\mu) \to (-\infty, \infty], \quad f \mapsto \mu[f] \triangleq \mu[f^+] - \mu[f^-],$$

is an increasing, additive and positively homogeneous operator.

Proof. Let $f, g \in \mathcal{K}(\mu)$ and $\alpha \geqslant 0$. If $f \leqslant g$, then $f^+ \leqslant g^+$ and $f^- \geqslant g^-$ so that $\mu[f] = \mu[f^+] - \mu[f^-] \leqslant \mu[g^+] - \mu[g^-] = \mu[g]$ by Corollary A.7. Corollary A.7 also yields

$$\mu[\alpha f] = \mu[\alpha f^+] - \mu[\alpha f^-] = \alpha\mu[f^+] - \alpha\mu[f^-] = \alpha\mu[f] \in (-\infty, \infty].$$

If $\mu[f^+] + \mu[g^+] < \infty$, then $\mu(\{f = \infty\}) = \mu(\{g = \infty\}) = 0$ by Lemma A.4; thus, $\bar{f} \triangleq f1_{\{f<\infty\}}$, $\bar{g} \triangleq g1_{\{g<\infty\}} \in \mathcal{L}^1(\mu)$, and Corollary A.8 and Lemma A.10 imply

$$\mu[f] + \mu[g] = \mu[\bar{f}] + \mu[\bar{g}] = \mu[\bar{f} + \bar{g}] = \mu[f + g].$$

It remains to prove additivity if $\mu[f^+] = \infty$; but then, $\mu[(f+g)^+] = \infty$, as well, and we have $\mu[f+g] = \infty = \mu[f] + \mu[g]$. □

On the space of bounded functions, the integral is particularly well-behaved.

LEMMA A.12. *The restriction of the integral to bounded functions,*

$$\mu[\cdot] : \mathcal{L}^\infty(\mathcal{A}) \to \mathbb{R}, \quad f \mapsto \mu[f]$$

is a positive continuous linear operator with operator norm $\|\mu[\cdot]\| = \mu(\Omega)$.

Proof. Clearly $\mathcal{L}^\infty(\mathcal{A}) \subset \mathcal{L}^1(\mu)$, and hence $\mu[\cdot]$ is positive linear operator by Lemma A.10. Moreover, for every $f \in \mathcal{L}^\infty(\mathcal{A})$, we have $f + \|f\|_\infty \geqslant 0$, and hence $0 \leqslant \mu[f + \|f\|_\infty] = \mu[f] + \mu(\Omega)\|f\|_\infty$. Therefore $-\mu(\Omega)\|f\|_\infty \leqslant \mu[f]$. The same reasoning applies to $\|f\|_\infty - f \geqslant 0$ and shows that $\mu[f] \leqslant \mu(\Omega)\|f\|_\infty$. Thus $|\mu[f]| \leqslant \mu(\Omega)\|f\|$ and $\mu[\cdot]$ is continuous with $\|\mu[\cdot]\| \leqslant \mu(\Omega)$ and $\mu[1] = \mu(\Omega)$. □

We stress one important point: The integral defined in this section, has the advantage of being defined for all non-negative measurable functions; however, it does not coincide with the Dunford (1935) integral more commonly used in functional analysis; see, e.g., Dunford and Schwartz (1958), Bhaskara Rao and Bhaskara Rao (1983), Luxemburg (1991). There, one also starts with the usual integral for simple functions

$$^D\!\!\int f\,d\mu \triangleq \int f\,d\mu = \textstyle\sum_{i=1}^n \alpha_i \mu(A_i),$$

but this integral is then extended to functions f which can be approximated by simple functions f_n in μ-measure in such a fashion that there exists a unique limit

$$^D\!\!\int f\,d\mu \triangleq \lim_{n\to\infty} {}^D\!\!\int f_n\,d\mu = \lim_{n\to\infty} \int f_n d\mu,$$

see, e.g., Dunford and Schwartz (1958), p. 101ff. Since uniform convergence implies convergence in μ-measure, Lemmas A.5 and A.12 show that Dunford's integral coincides with $\mu[\cdot]$ on $\mathcal{L}^\infty(\mathcal{A})$, where it is just the unique continuous extension of the integral for simple functions. This modest version of Dunford's integral is all that is needed to represent continuous linear functionals on $\mathcal{L}^\infty(\mathcal{A})$:

LEMMA A.13. *Let $T : \mathcal{L}^\infty(\mathcal{A}) \to \mathbb{R}$ be a positive continuous linear operator. Then there exists a unique finitely additive measure m with $m(\Omega) = T[1]$ such that T is the restriction of $m[\cdot]$ to $\mathcal{L}^\infty(\mathcal{A})$, i.e.,*

Representation of functionals on $\mathcal{L}^\infty(\mathcal{A})$

$$Tg = m[g] = \int g \, dm \quad \textit{for all } g \in \mathcal{L}^\infty(\mathcal{A}). \tag{$\star$}$$

Proof. By ($\star$), the only way to define $m : \mathcal{A} \to \mathbb{R}$ is via the formula $m(A) = T[1_A]$. Since T is linear, this gives rise to an additive set function. Positivity of T implies

$$0 \leqslant m(A) \leqslant m(B) \leqslant m(\Omega) = T[1], \qquad A \subset B, \; A, B \in \mathcal{A},$$

and hence m is in fact a finitely additive measure on $\mathcal{A}$. Lemma A.12 guarantees that $m[\cdot]$ is a positive continuous linear operator on $\mathcal{L}^\infty(\mathcal{A})$. Hence, for $f = \sum_{i=1}^n \alpha_i 1_{A_i} \in \operatorname{span}\{1_A : A \in \mathcal{A}\} \subset \mathcal{L}^\infty(\mathcal{A})$, we have

$$m[f] = \sum_{i=1}^n \alpha_i m(A_i) = \sum_{i=1}^n \alpha_i T[1_{A_i}] = T[f].$$

Since $\operatorname{span}\{1_A : A \in \mathcal{A}\}$ is dense in $\mathcal{L}^\infty(\mathcal{A})$ and both $m[\cdot]$ and T are continuous operators on $\mathcal{L}^\infty(\mathcal{A})$, representation ($\star$) is established. □

A.2 SUBLINEAR EXPECTATION OPERATORS

According to Definition 2.36, a (static) *sublinear expectation operator* $\mathcal{E}$ is an extended real-valued function which is defined on a convex cone $\mathcal{K} \subset \mathcal{L}^0(\mathcal{A}; [-\infty, \infty])$ containing all real constants and which is

- *sublinear*, i.e., $\mathcal{E}[X + Y] \leqslant \mathcal{E}[X] + \mathcal{E}[Y]$ and $\mathcal{E}[\alpha X] = \alpha \mathcal{E}[X]$ for all $X, Y \in \mathcal{K}$ and $\alpha \geqslant 0$,
- *monotone*, i.e., $\mathcal{E}[X] \leqslant \mathcal{E}[Y]$, for all $X, Y \in \mathcal{K}$ with $X \leqslant Y$, and
- *constant-preserving*, i.e., $\mathcal{E}[\alpha] = \alpha$ for all $\alpha \in \mathbb{R}$.

A.2.1 *Sublinear expectation operators via additive probabilities*

We associate a set $\mathcal{K}(M)$ and an operator $M[\cdot] : \mathcal{K}(M) \to (-\infty, \infty]$ with every (non-empty) family of finitely additive probabilities M as follows: The set $\mathcal{K}(M)$ is given by

$$\mathcal{K}(M) \triangleq \left\{ g \in \mathcal{L}^0(\mathcal{A}; (-\infty, \infty]) : \sup_{m \in M} \int g^- dm < \infty \right\} \subset \mathcal{L}^0(\mathcal{A}; (-\infty, \infty]),$$

and the operator $M[\cdot]$ is given by

$$M[\cdot] : \mathcal{K}(M) \to (-\infty, \infty], \quad f \mapsto M[f] \triangleq \sup_{m \in M} m[f].$$

In the following, when we refer to a family of finitely additive probabilities, we always mean a *non-empty* family.

LEMMA A.14. *Let M be a family of finitely additive probabilities. Then $M[\cdot]$ is a sublinear expectation operator on the convex cone $\mathcal{K}(M)$.*

Proof. Let $f, g \in \mathcal{K}(M) \subset \mathcal{L}^0(\mathcal{A}; (-\infty, \infty])$ and $\alpha, \beta \geqslant 0$. Then $\alpha f + \beta g \in \mathcal{L}^0(\mathcal{A}; (-\infty, \infty])$. Since $(\alpha f + \beta g)^- \leqslant \alpha f^- + \beta g^-$, we have

$$\int (\alpha f + \beta g)^- dq \leqslant \alpha \sup_{m \in M} \int f^- dm + \beta \sup_{m \in M} \int g^- dm < \infty$$

for all $q \in M$ so that $\alpha f + \beta g \in \mathcal{K}(M)$; hence, $\mathcal{K}(M)$ is a convex cone. Evidently, every constant function is contained in $\mathcal{K}(M)$ and $M[\cdot] : \mathcal{K}(M) \to (-\infty, \infty]$.

It remains to prove that $M[\cdot]$ is a sublinear expectation operator. Note that, for each $m \in M$, we have $\mathcal{K}(M) \subset \mathcal{K}(m)$, so that

$$m : \mathcal{K}(M) \to (-\infty, \infty], \quad f \mapsto m[f]$$

is an increasing, additive and positively homogeneous operator by Lemma A.11. Moreover, $m[\alpha] = \alpha$ for all $\alpha \in \mathbb{R}$. Therefore, the upper envelope $M[\cdot] = \sup_{m \in M} m[\cdot]$ also preserves constants and is increasing, positively homogeneous and subadditive; hence, M is a sublinear expectation operator. □

Associated function spaces

Let M be a (non-empty) family of finitely additive probability measures. We consider the corresponding function spaces

$$\mathcal{L}^p(M) \triangleq \left\{ f \in \mathcal{L}^0(A) : M\left[|f|^p\right] < \infty \right\} \subset \mathcal{L}^0(\mathcal{A}), \qquad 1 \leqslant p < \infty.$$

If $f, g \in \mathcal{L}^p(M)$, then $f + g \in \mathcal{L}^0(\mathcal{A})$, and monotonicity and sublinearity of $M[\cdot]$ imply

$$M\left[|\alpha f + g|^p\right] \leqslant 2^p \left(|\alpha|^p M\left[|f|^p\right] + M\left[|g|^p\right]\right) < \infty$$

so that $\mathcal{L}^p(M) \subset \mathcal{L}^0(\mathcal{A})$ is a linear space contained in $\mathcal{K}(M)$.

For each $p \geqslant 1$, we consider the function

$$\| \cdot \|_{M,p} : \mathcal{L}^0(\mathcal{A}; [-\infty, \infty]) \to [0, \infty], \quad f \mapsto M\left[|f|^p\right]^{\frac{1}{p}}.$$

By definition, $\| \cdot \|_{M,p}$ is finite on $\mathcal{L}^p(M)$. Below, we establish that it is a seminorm. As in Subsection 2.2.2 at the beginning of the present thesis, this is achieved via Hölder's inequality. The proof is a simpler version of the one of Lemma 2.7.

Hölder inequality

LEMMA A.15. *Let $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$. For all $X, Y \in \mathcal{L}^0(\mathcal{A}; [0, \infty])$, we have*

$$\|XY\|_{M,1} \leqslant \|X\|_{M,p} \|Y\|_{M,q}.$$

In particular, $\mathcal{L}^p(M) \cdot \mathcal{L}^q(M) \subset \mathcal{L}^1(M)$.

Proof. Clearly $XY \in \mathcal{L}^0(\mathcal{A}; [0, \infty])$. For $\varepsilon > 0$ we put

$$\bar{X} \triangleq \frac{X}{(\varepsilon + M[|X|^p])^{\frac{1}{p}}} \quad \text{and} \quad \bar{Y} \triangleq \frac{Y}{(\varepsilon + M[|Y|^q])^{\frac{1}{q}}},$$

(where $\infty/\infty \triangleq 0$) and note that $|\bar{X}\bar{Y}| \leqslant \frac{1}{p}|\bar{X}|^p + \frac{1}{q}|\bar{Y}|^q$ by Young's inequality; thus, by monotonicity and sublinearity of $M[\cdot]$, we obtain

$$M\big[|\bar{X}\bar{Y}|\big] \leqslant M\big[\tfrac{1}{p}|\bar{X}|^p + \tfrac{1}{q}|\bar{Y}|^q\big] \leqslant \tfrac{1}{p}M\big[|\bar{X}|^p\big] + \tfrac{1}{q}M\big[|\bar{Y}|^q\big].$$

Either $\bar{X} = 0$ or $M[|X|^p] < \infty$, and hence we have

$$M\big[|\bar{X}|^p\big] = M[|X|^p]/\big(\varepsilon + M\big[|X|^p\big]\big) \leqslant 1 \quad \text{and} \quad M\big[|\bar{Y}|^p\big] \leqslant 1$$

by positive homogeneity. Therefore $M[|\bar{X}\bar{Y}|] \leqslant \frac{1}{p} + \frac{1}{q} = 1$, and

$$M\big[|XY|\big] \leqslant \big(\varepsilon + M[|X|^p]\big)^{\frac{1}{p}}\big(\varepsilon + M[|Y|^q]\big)^{\frac{1}{q}}.$$

Letting $\varepsilon \to 0$ yields the claim. □

COROLLARY A.16. *Let* $1 \leqslant p \leqslant q$. *For all* $X \in \mathcal{L}^0(\mathcal{A}; [-\infty, \infty])$, *we have* $\|X\|_{L,p} \leqslant \|X\|_{L,q}$ *and, in particular,* $\mathcal{L}^q(M) \subset \mathcal{L}^p(M)$.

With Hölder's inequality, using a classical argument, one can prove that $\|\cdot\|_{M,p}$ is a seminorm on $\mathcal{L}^p(M)$. We prove a slightly more general statement which will be useful later on.

LEMMA A.17. *Let* $\mathcal{X}$ *be a convex cone*[1] *and suppose that* $\rho : \mathcal{X} \to \mathcal{L}^0(\mathcal{A}; [0, \infty])$ *is sublinear. Then, the function* *Norms*

$$\|\rho\|_{M,p} : \mathcal{X} \to [0, \infty], \quad X \mapsto \|\rho(X)\|_{M,p}$$

is sublinear. If $\mathcal{X}$ *is a linear space, and if* ρ *is homogeneous and maps into* $\mathcal{L}^p(M)$, *then* $\|\rho\|_{M,p}$ *is a seminorm on* $\mathcal{X}$.

Proof. Positive homogeneity is obvious. To prove subadditivity of $\|\rho\|_{M,p}$, let $X, Y \in \mathcal{X}$ and set $\bar{X} = \rho(X)$, $\bar{Y} = \rho(Y)$ and $\bar{Z} = \rho(X+Y)$. Then $\bar{X}, \bar{Y}, \bar{Z} \in \mathcal{L}^0(\mathcal{A}; [0, \infty])$ and

$$|\bar{Z}|^p = \bar{Z} \cdot |\bar{Z}|^{p-1} \leqslant \bar{X} \cdot |\bar{Z}|^{p-1} + \bar{Y} \cdot |\bar{Z}|^{p-1},$$

since $\rho : \mathcal{X} \to \mathcal{L}^0(\mathcal{A}; [0, \infty])$ is subadditive. Monotonicity and sublinearity of $M[\cdot]$ imply

$$M\left[|\bar{Z}|^p\right] \leqslant M\left[\bar{X} \cdot |\bar{Z}|^{p-1}\right] + M\left[\bar{Y} \cdot |\bar{Z}|^{p-1}\right].$$

For $p = 1$, this gives $\|\bar{Z}\|_{M,1} \leqslant \|\bar{X}\|_{M,1} + \|\bar{Y}\|_{M,1}$, thus establishing subadditivity. To prove the claim for $p > 1$, we set $q \triangleq \frac{p}{p-1}$ and note that

$$\big\||\bar{Z}|^{p-1}\big\|_{M,q}^q = M\left[\big||\bar{Z}|^{p-1}\big|^q\right] = M\left[|\bar{Z}|^p\right] = \|\bar{Z}\|_{M,p}^p.$$

1 Consistent with the meaning of convex cone above, here, a convex cone $(K, +, \cdot)$ means any commutative cancellative monoid $(K, +)$ together with a scalar multiplication $\cdot : [0, \infty) \times K \to K$ such that, for all $\alpha, \beta \in [0, \infty)$ and $x, y \in K$, we have

$$1x = x, \quad \alpha(\beta x) = (\alpha\beta)x, \quad \alpha(x+y) = \alpha x + \alpha y, \quad (\alpha+\beta)x = \alpha x + \beta y, \quad \alpha, \beta \in [0, \infty).$$

Note that any such convex cone can be embedded into a vector space in which it is a convex cone in the usual sense.

Since $\frac{1}{p} + \frac{1}{q} = 1$, Lemma A.15 yields

$$\left\|\bar{X} \cdot |\bar{Z}|^{p-1}\right\|_{M,1} \leqslant \|\bar{X}\|_{M,p} \||\bar{Z}|^{p-1}\|_{M,q} = \|\bar{X}\|_{M,p} \|\bar{Z}\|_{M,p}^{\frac{p}{q}}$$

as well as

$$\left\|\bar{Y} \cdot |\bar{Z}|^{p-1}\right\|_{M,1} \leqslant \|\bar{Y}\|_{M,p} \||\bar{Z}|^{p-1}\|_{M,q} = \|\bar{Y}\|_{M,p} \|\bar{Z}\|_{M,p}^{\frac{p}{q}},$$

and we arrive at

$$\|\bar{Z}\|_{M,p}^{p} \leqslant \left(\|\bar{X}\|_{M,p} + \|\bar{Y}\|_{M,p}\right) \|\bar{Z}\|_{M,p}^{\frac{p}{q}}.$$

If $\|\bar{Z}\|_{M,p} = 0$, there is nothing to prove; otherwise, rearrange to obtain

$$\|\bar{Z}\|_{M,p} = \|\bar{Z}\|_{M,p}^{p-\frac{p}{q}} \leqslant \|\bar{X}\|_{M,p} + \|\bar{Y}\|_{M,p}.$$

Thus, the function $\|\rho\|_{M,p}$ is subadditive.

Finally, if ρ is homogeneous and maps to $\mathcal{L}^p(M)$, and if $\mathfrak{X}$ is a linear space, then we have $\|\rho\|_{M,p} : \mathfrak{X} \to [0, \infty)$ as well as

$$\|\rho(\alpha X)\|_{M,p} = M\left[\left||\alpha|\rho(X)\right|^p\right]^{\frac{1}{p}} = |\alpha| \|\rho(X)\|_{M,p}, \quad X \in \mathfrak{X}, \alpha \in \mathbb{R},$$

by sublinearity of $M[\cdot]$; thus, $\|\rho\|_{M,p}$ is a seminorm on $\mathfrak{X}$. □

$\|\cdot\|_{M,p}$-seminorm

COROLLARY A.18. *For each $p \geqslant 1$, the map $\|\cdot\|_{M,p}$ is a seminorm on $\mathcal{L}^p(M)$.*

Proof. Apply Lemma A.17 to $\mathfrak{X} = \mathcal{L}^p(M)$ and the homogeneous and subadditive map $\rho : \mathcal{L}^p(M) \to \mathcal{L}^p(M) \cap \mathcal{L}^0(\mathcal{A}; [0, \infty])$, $X \mapsto |X|$. □

Associated process spaces

Let $(T, \mathcal{T}, \mu)$ be a finite measure space and $\mathcal{G} \subset \mathcal{A} \otimes \mathcal{T}$ a σ-algebra. We write $\mathcal{P}^0(\mathcal{G}) = \mathcal{L}^0(\mathcal{G}; \mathbb{R})$ for the linear space of all $\mathcal{G}$-measurable processes

$$X : \Omega \times T \to \mathbb{R}, \quad (\omega, t) \mapsto X_t(\omega).$$

For all numbers $p, q \geqslant 1$, we consider the p-q-*seminorm*

$$\|X\|_{p,q} \triangleq \left\||X|_q\right\|_{M,p} = M\left[\left(\int |X_t|^q \mu(dt)\right)^{\frac{p}{q}}\right]^{\frac{1}{p}}. \tag{A.3}$$

Here, $|f|_q = \left(\int |f(t)|^q \mu(dt)\right)^{\frac{1}{q}}$ denotes the seminorm of the Lebesgue space $\mathcal{L}^q(\mu)$. Lemma A.17 shows that $\|\cdot\|_{p,q}$ is a sublinear map on the convex cone $\mathcal{L}^0(\mathcal{A} \otimes \mathcal{T}; (-\infty, \infty])$. Thus, for all $p, q \geqslant 1$,

$$\mathcal{P}^{p,q}(M, \mu; \mathcal{G}) \triangleq \left\{X \in \mathcal{P}^0(\mathcal{G}) \,:\, \|X\|_{p,q} < \infty\right\} \tag{A.4}$$

is a linear spaces of $\mathcal{G}$-measurable real-valued processes, and $\|\cdot\|_{p,q}$ is a seminorm on $\mathcal{P}^{p,q}(M, \mu; \mathcal{G})$.

To guarantee completeness of these process spaces and, in particular, of $(\mathcal{L}^p(M), \|\cdot\|_{M,p})$, we need an additional assumption – the Fatou property. The following subsections are devoted to the Fatou property and its ramifications.

A.2.2 *The Fatou property and upper probabilities*

DEFINITION A.19. A family Q of finitely additive probabilities has the *Fatou property* if

Fatou property

$$Q\big[\liminf_{n\to\infty} g_n\big] \leqslant \liminf_{n\to\infty} Q[g_n] \quad \text{for all } (g_n)_{n\in\mathbb{N}} \subset \mathcal{L}^0(\mathcal{A};[0,\infty]).$$

In that case, the associated sublinear expectation operator $Q[\cdot]$ is called an *upper expectation*, and the corresponding set function

Upper expectation

$$Q(\cdot):\mathcal{A}\to[0,1], \quad A\mapsto Q(A) \triangleq Q[1_A] = \sup_{q\in Q} q(A)$$

is said to be an *upper probability*. An event $A\in\mathcal{A}$ has upper probability $u\in[0,1]$ (w.r.t. Q) if $Q(A)=u$. An event $A\in\mathcal{A}$ is said to be Q-*negligible* (or simply negligible if Q is clear from the context) if it has upper probability zero. If $S[\omega]$ is a statement depending on ω and the event $\{\omega\in\Omega : S[\omega] \text{ is false}\}$ is negligible, then we say that S holds Q-essentially. ◇

Upper probability
Negligible event

It is easy to show that upper envelopes of *countably* additive probability measures have the Fatou property, see Example 2.39 (p. 38) in the main part of this thesis. Indeed, most prominent examples of upper probabilities are of that form.

Upper probabilities are countably subadditive:

LEMMA A.20. *All upper probabilities Q are countably subadditive; i.e., for all $A_n\in\mathcal{A}$, $n\in\mathbb{N}$, we have*

Countable subadditivity of upper probability

$$Q\left(\textstyle\bigcup_{n=1}^{\infty} A_n\right) \leqslant \textstyle\sum_{n=1}^{\infty} Q(A_n).$$

In particular, countable unions of negligible events are negligible.

Proof. We set $B_1 \triangleq A_1$ and $B_n \triangleq A_n\setminus B_{n-1}$. Then $B_n\in\mathcal{A}$, $n\in\mathbb{N}$, are pairwise disjoint with $A \triangleq \bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$. Now, we put $f_n \triangleq \sum_{k=1}^{n} 1_{B_k}$. Then $0\leqslant f_n\uparrow f = 1_A$ and, by the Fatou property of the upper expectation Q, it follows that

$$Q(A) = Q[f] \leqslant \liminf_{n\to\infty} Q[f_n] \leqslant \liminf_{n\to\infty} \textstyle\sum_{k=1}^{n} Q[1_{B_k}],$$

as $Q[\cdot]$ is sublinear. Since $B_k\subset A_k$, monotonicity of $Q[\cdot]$ implies that $Q[1_{B_k}]\leqslant Q[1_{A_k}]$, and thus

$$Q(A) \leqslant \liminf_{n\to\infty} \textstyle\sum_{k=1}^{n} Q[1_{A_k}] = \sum_{n=1}^{\infty} Q(A_n). \qquad \square$$

LEMMA A.21. *Let Q be an upper expectation and $X:\Omega\to[-\infty,\infty]$ measurable. If $Q[|X|]=0$, then X vanishes Q-essentially.*

Proof. By monotonicity and positive homogeneity of $Q[\cdot]$, we have

$$Q\big[1_{\{|X|>\frac{1}{n}\}}\big] \leqslant nQ\big[1_{\{|X|>\frac{1}{n}\}}|X|\big] \leqslant n\,Q\big[|X|\big] = 0.$$

Since Q has the Fatou property, we get

$$Q(X\neq 0) = Q\left[1_{\{|X|>0\}}\right] \leqslant \liminf_{n\to\infty} Q\big[1_{\{|X|>\frac{1}{n}\}}\big] = 0. \qquad \square$$

Integrable functions are essentially finite

LEMMA A.22. *Let* Q *be an upper expectation and* $X : \Omega \to [0, \infty]$ *measurable. If* $Q[X] < \infty$*, then* X *is* Q*-essentially everywhere finite.*

Proof. Assume by contradiction that $\{X = \infty\}$ has positive upper probability; then $q(\{X = \infty\}) > 0$ for some $q \in Q$, and hence $q[X] = \infty$ by Lemma A.4. This implies that $Q[X] = \infty$, a contradiction. □

LEMMA A.23. *Let* Q *be an upper expectation and* $X, Y \in \mathcal{L}^1(Q)$*. If* X *and* Y *coincide* Q*-essentially, then* $Q[X] = Q[Y]$*.*

Proof. By assumption, the sets $A^+ \triangleq \{X^+ \neq Y^+\}$ and $A^- \triangleq \{X^- \neq Y^-\}$ are Q-negligible. In particular, $q[X^+] = q[Y^+]$ and $q[X^-] = q[Y^-]$ for every $q \in Q$ by Corollary A.8. Therefore

$$Q[X] = \sup_{q \in Q} \big(q[X^+] - q[X^-] \big) = \sup_{q \in Q} \big(q[Y^+] - q[Y^-] \big) = Q[Y]. \qquad \square$$

A.2.3 *Completeness results*

This section is concerned with proving completeness of the process spaces $\mathcal{P}^{p,q}(Q, \mu; \mathcal{G})$, as introduced at the end of Subsection A.2.1 on page 150.

STANDING ASSUMPTIONS In the following,

- Q is an upper probability,
- $(T, \mathcal{T}, \mu)$ is a finite measure space, and
- $\mathcal{G} \subset \mathcal{A} \otimes \mathcal{T}$ is a σ-algebra.

The proof of completeness hinges on the Fatou property of Q and is very close to that of the classical Riesz-Fischer completeness theorem for L^p-spaces (see, e.g., Rudin (1974), Theorem 3.11, p. 69ff.). The Fatou property of Q yields a Fatou property for the p-q-seminorms.

Fatou property of p-q-seminorm

LEMMA A.24. *For all* $p, q \geqslant 1$*, the* p-q*-seminorm* $\|\cdot\|_{p,q}$ *(as given in* (A.3) *on p. 150 with* $M = Q$*) satisfies the following Fatou property:*

If $(X^n)_{n \in \mathbb{N}}$ *is a sequence in* $\mathcal{L}^0(\mathcal{A} \otimes \mathcal{T}; [0, \infty])$*, then*

$$\| \liminf_{n \to \infty} X_n \|_{p,q} \leqslant \liminf_{n \to \infty} \|X^n\|_{p,q}.$$

Proof. First, we note that the section $X^n(\omega)$ is $\mathcal{T}$-measurable for all $\omega \in \Omega$ and every $n \in \mathbb{N}$, since $\mathcal{G} \subset \mathcal{A} \otimes \mathcal{T}$. Setting $X \triangleq \liminf_{n \to \infty} X^n$, Fatou's lemma implies

$$\int X_t(\omega)^q \, \mu(dt) \leqslant \liminf_{n \to \infty} \int X^n_t(\omega)^q \, \mu(dt) \quad \text{for all } \omega \in \Omega.$$

By monotonicity of $Q[\cdot]$, we get

$$\| \liminf_{n \to \infty} X^n \|_{p,q}^p \leqslant Q\Big[\liminf_{n \to \infty} \Big(\int |X^n_t|^q \, \mu(dt) \Big)^{\frac{p}{q}} \Big].$$

Now, the Fatou property of $Q[\cdot]$ implies

$$\|X\|_{p,q}^{p} \leqslant \liminf_{n\to\infty} Q\Big[\Big(\int |X_t^n|^q \,\mu(dt)\Big)^{\frac{p}{q}}\Big] = \liminf_{n\to\infty} \|X^n\|_{p,q}^{p}. \qquad \square$$

As a first consequence of the Fatou property, we see that Cauchy sequences have subsequences which converge pointwise:

LEMMA A.25. *Every $\|\cdot\|_{p,q}$-Cauchy sequence $(X^n)_{n\in\mathbb{N}} \subset \mathcal{P}^{p,q}(Q,\mu;\mathcal{G})$ has a subsequence $(X^{n_k})_{k\in\mathbb{N}}$ depending only on $(\|X^n - X^m\|_{p,q})_{n,m\in\mathbb{N}}$ such that the following holds: There exists a Q-negligible set $N \in \mathcal{A}$ and a $\mathcal{G}$-measurable real-valued process such that*

$$\lim_{k\to\infty} X_t^{n_k}(\omega) = X_t(\omega) \quad \textit{for } \mu\textit{-a.e. } t \in T \textit{ and all } \omega \in N^c. \tag{A.5}$$

Proof. We choose a subsequence (again denoted by $(X^n)_{n\in\mathbb{N}}$) such that

$$\|X^{n+1} - X^n\|_{p,q} \leqslant 2^{-n} \quad \text{for all } n \in \mathbb{N}. \tag{A.6}$$

Clearly, this subsequence depends only on $(\|X^n - X^m\|_{p,q})_{n,m\in\mathbb{N}}$.

Now, for each $n \in \mathbb{N}$, we define a $\mathcal{G}$-measurable process Y^n via

$$Y^n \triangleq \textstyle\sum_{k=1}^{n} |X^{k+1} - X^k| \geqslant 0. \tag{A.7}$$

Clearly, $Y_t^n(\omega) \uparrow Y_t(\omega)$ for all $t \in [0,T]$ and $\omega \in \Omega$, where $(t,\omega) \mapsto Y_t(\omega)$ is a non-negative $\mathcal{G}$-measurable process. Lemma A.24 implies

$$\|Y\|_{p,q} \leqslant \liminf_{n\to\infty} \|Y^n\|_{p,q}, \tag{A.8}$$

where we can estimate

$$\|Y^n\|_{p,q} \leqslant \textstyle\sum_{k=1}^{n} \big\|X^{k+1} - X^k\big\|_{p,q} \leqslant \sum_{k=1}^{\infty} 2^{-k} = 1, \tag{A.9}$$

by definition of Y^n (A.7), subadditivity of $\|\cdot\|_{p,q}$ and (A.6). In conjunction with (A.8), estimate (A.9) shows that

$$Q\left[\left(\int |Y_t|^q \,\mu(dt)\right)^{\frac{p}{q}}\right] = \|Y\|_{p,q}^{p} \leqslant 1.$$

Hence, by Lemma A.22, the set $N \in \mathcal{A}$ of all ω with $\int |Y_t(\omega)|^q \mu(dt) = \infty$ is a Q-negligible event. Therefore, for all $\omega \in N^c$, there is a μ-null set $A_\omega \in \mathcal{T}$ such that

$$\textstyle\sum_{k=1}^{\infty} |X_t^{k+1}(\omega) - X_t^k(\omega)| = Y_t(\omega) < \infty \quad \text{for all } t \in A_\omega^c. \tag{A.10}$$

From (A.10), we see that for every $\varepsilon > 0$, each $\omega \in N^c$ and every $t \in A_\omega^c$ there is some $n_\varepsilon(\omega,t)$ such that

$$\textstyle\sum_{k=n_\varepsilon(\omega,t)}^{\infty} |X_t^{k+1}(\omega) - X_t^k(\omega)| < \varepsilon. \tag{A.11}$$

By a telescoping sum argument, (A.11) implies that

$$(X_t^n(\omega))_{n\in\mathbb{N}} \text{ is Cauchy in } \mathbb{R} \text{ for each } \omega \in N^c \text{ and all } t \in A_\omega^c. \tag{A.12}$$

Now, for all $(\omega, t) \in \Omega \times [0, T]$, we set

$$X_t(\omega) \triangleq 1_{\{\limsup_{n\to\infty} |X_t^n| < \infty\}}(\omega) \cdot \limsup_{n\to\infty} X_t^n(\omega).$$

The process $X : \Omega \times [0, T] \to \mathbb{R}$, $(\omega, t) \to X_t(\omega)$ thus defined is $\mathcal{G}$-measurable and, by (A.12), we have

$$X_t(\omega) = \lim_{n\to\infty} X_t^n(\omega) \quad \text{for all } t \in A_\omega^c \text{ and every } \omega \in N^c.$$

This establishes (A.5) and concludes the proof. □

Completeness of the process spaces is a consequence of pointwise convergence and the Fatou property from Lemma A.24:

THEOREM A.26. *The spaces* $\mathcal{P}^{p,q}(Q, \mu; \mathcal{G})$ *are complete.*

Proof. Let $(X^n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\mathcal{P}^{p,q}(Q, \mu; \mathcal{G})$. For brevity, we write $\|\cdot\| \triangleq \|\cdot\|_{p,q}$. Lemma A.25 yields a $\mathcal{G}$-measurable real-valued process X such that

$$\lim_{k\to\infty} X_t^{n_k} = X_t \quad \text{for } \mu\text{-a.e. } t \in T, \quad Q\text{-essentially.}$$

Therefore

$$\|X\| = \|\liminf_{k\to\infty} X^{n_k}\| \quad \text{and} \quad \|X^{n_k} - X\| = \|\liminf_{\ell\to\infty} (X^{n_k} - X^{n_\ell})\|,$$

and hence Lemma A.24 implies

$$\|X\| \leqslant \liminf_{k\to\infty} \|X^{n_k}\| < \infty \;\; \text{and} \;\; \|X^{n_k} - X\| \leqslant \liminf_{\ell\to\infty} \|X^{n_k} - X^{n_\ell}\| \overset{k\to\infty}{\longrightarrow} 0.$$

Thus $X \in \mathcal{P}^{p,q}$ and $X^n \to X$ in $\mathcal{P}^{p,q}$. □

We now apply Theorem A.26 to the seminormed spaces

$$\mathcal{L}^p(Q) \triangleq \left\{ f \in \mathcal{L}^0(\mathcal{A}) \, : \, \|f\|_{L,p} \triangleq M\big[|f|^p\big]^{\frac{1}{p}} < \infty \right\}, \qquad 1 \leqslant p < \infty,$$

which have been introduced in Subsection A.2.1, p. 148. Clearly,

$$(\mathcal{L}^p(Q), \|\cdot\|_{L,p}) \quad \text{can be identified with} \quad (\mathcal{P}^{p,p}(Q, \mu; \mathcal{G}), \|\cdot\|_{p,p}),$$

upon using the trivial specifications

$$T = \{0\}, \quad \mathcal{T} = 2^T, \quad \mathcal{G} = \mathcal{A} \otimes \mathcal{T} = \{A \times \{0\} \, : \, A \in \mathcal{A}\}$$

and μ as the measure assigning mass 1 to $\{0\}$. Thus by Theorem A.26 and Lemma A.25 we obtain

Completeness of $\mathcal{L}^p(Q)$

COROLLARY A.27. *The spaces* $\mathcal{L}^p(Q)$, $p \geqslant 1$, *are complete. Every sequence* $X_n \to X$ *in* $\mathcal{L}^p(Q)$ *has a subsequence* $(X_{n_k})_{k\in\mathbb{N}}$ *depending only on the seminorms* $(\|X_n - X_m\|_{L,p})_{n,m\in\mathbb{N}}$ *such that*

$$\lim_{k\to\infty} X_{n_k}(\omega) = X(\omega) \quad \textit{for } Q\textit{-essentially every } \omega \in \Omega.$$

A.2.4 *The Banach space* $L^p(Q)$

We write $\mathcal{N}_Q$ for the collection of all Q-negligible sets and consider the equivalence relation $\sim_Q$ on $\Omega^{\mathbb{R}}$, induced by Q via

$$f \sim_Q g \iff f(\omega) = g(\omega) \quad \text{for all } \omega \in \Omega \setminus N \text{ and some } N \in \mathcal{N}_Q.$$

The $\sim_Q$-equivalence class of a function $f \in \Omega^{\mathbb{R}}$ will be denoted by $[f]_Q$, and we define

$$L^p(Q) \triangleq \left\{[f]_Q \,:\, f \in \mathcal{L}^p(Q)\right\} \subset {}^{\Omega^{\mathbb{R}}}\!/\!_{\mathcal{N}_Q} \triangleq {}^{\Omega^{\mathbb{R}}}\!/\!_{\sim_Q}$$

as the collection of all such equivalence classes containing a member of $\mathcal{L}^p(Q)$. We have the following (trivial) result:

LEMMA A.28. *For all* $p \geqslant 1$, *the operator*

$$Q : L^p(Q) \to \mathbb{R}, \quad X \mapsto Q[f], \quad \textit{where } f \in X \cap \mathcal{L}^0(\mathcal{A}),$$

is well-defined, and $(L^p, \|\cdot\|_{L,p})$ *is a normed space.*

Proof. Let $X \in L^p(Q)$. By definition of $L^p(Q)$, there is some $f \in \mathcal{L}^p(Q)$ such that $X = [f]_Q$. In particular, $f \in X \cap \mathcal{L}^0(\mathcal{A})$. Let $g \in X \cap \mathcal{L}^0(\mathcal{A})$ be arbitrary. Then f and g coincide Q-essentially, and hence $Q[f] = Q[g]$ (see Lemma A.23). Thus $Q[\cdot]$ is well-defined.

By definition, $L^p(Q)$ is contained in the linear space ${}^{\Omega^{\mathbb{R}}}\!/\!_{\mathcal{N}_Q}$. Let $X, Y \in L^p(Q)$ and $\alpha \in \mathbb{R}$, and choose $f \in X \cap \mathcal{L}^p(Q)$ and $g \in Y \cap \mathcal{L}^p(Q)$. Then $\alpha f + g \in \mathcal{L}^p(Q)$, since $\mathcal{L}^p(Q)$ is a linear space, and thus

$$\alpha X + Y = \alpha[f]_Q + [g]_Q = [\alpha f + g]_Q \in L^p(Q).$$

Since $\|\cdot\|_{L,p}$ is a seminorm on $\mathcal{L}^p(Q)$ (Corollary A.18) and $Q[\cdot]$ is well-defined on $L^p(Q)$, we obtain

$$\|X + Y\|_{L,p} = \|f + g\|_{L,p} \leqslant \|f\|_{L,p} + \|g\|_{L,p} = \|X\|_{L,p} + \|Y\|_{L,p}$$

and, similarly, $\|\alpha X\|_{L,p} = |\alpha|\|X\|_{L,p}$. It remains to check that

$$f \sim_Q 0 \iff \|f\|_{L,p} = 0$$

for all $f \in \Omega^{\mathbb{R}}$. Since $Q[\cdot]$ is well-defined on $L^p(Q)$, the validity of " $\Longrightarrow$ " is obvious. For the other direction, we note that $\|f\|_{L,p} = 0$ means that $Q[|g|^p] = 0$ for some $g \in \mathcal{L}^p(Q)$ with $f \sim_Q g$. But then g vanishes Q-essentially by Lemma A.21, and hence $f \sim_Q g \sim_Q 0$. □

COROLLARY A.29. *If* Q *has the Fatou property, then* $L^p(Q) \subset {}^{\Omega^{\mathbb{R}}}\!/\!_{\mathcal{N}_Q}$ *is a Banach space for all* $p \geqslant 1$. *If* $X_n \to X$ *in* $L^p(Q)$, *then there exists a subsequence* $(X_{n_k})_{k \in \mathbb{N}} \subset L^p(Q)$ *such that* $X_{n_k} \to X$ *converges* Q*-essentially.*

Proof. In view of Lemma 2.41, completeness of $L^p(Q)$ and the convergence statement are guaranteed by Corollary A.27. □

We continue this subsection with a brief investigation of convergence in upper probability.

DEFINITION A.30. We say that a sequence $(X_n)_{n\in\mathbb{N}} \subset \mathcal{L}^0(\mathcal{A})$ *converges to* X *in upper probability* (in symbols: $X_n \xrightarrow{Q} X$) if

$$\lim_{n\to\infty} Q\big(|X_n - X| > \varepsilon\big) = 0 \quad \text{for all } \varepsilon > 0.$$

We have the following simple relations between convergence in upper probability and L^p-convergence:

LEMMA A.31. *Let* $(X_n)_{n\in\mathbb{N}} \subset \mathcal{L}^0(\mathcal{A})$.

(i) If $X_n \xrightarrow{Q} X$*, then,* Q*-essentially,* $X_{n_k} \to X$ *for a subsequence.*

(ii) If $X_n \to X$ *in* $L^p(Q)$*, then* $X_n \xrightarrow{Q} X$.

(iii) If $(X_n)_{n\in\mathbb{N}}$ *is uniformly bounded by a constant and* $X_n \xrightarrow{Q} X$*, then* $X_n \to X$ *in* L^p.

(iv) Let $g : \mathbb{R} \to \mathbb{R}$ *be continuous and suppose that* $X_n \xrightarrow{Q} X$*. If* g *is uniformly continuous, or if* $X \in L^1(Q)$*, then* $g(X_n) \xrightarrow{Q} g(X)$.

Proof. (i) If $X_n \xrightarrow{Q} X$, we can choose a subsequence such that

$$Q\big(|X_{n_k} - X| > 2^{-k}\big) \leqslant 2^{-k}.$$

Upper probabilities are countably subadditive by Lemma A.20; hence, the event $A_\ell \triangleq \bigcup_{k\geqslant \ell}\{|X_{n_k} - X| > 2^{-k}\}$ satisfies $Q(A_\ell) \leqslant 2^{-\ell+1}$. The event on which X_{n_k} fails to converge to X is contained in A_ℓ for all $\ell \in \mathbb{N}$, and is thus Q-negligible.

(ii) By the properties of sublinear expection operators, we have

$$Q\big(|X_n - X| > \varepsilon\big) \leqslant \varepsilon^{-p} Q\big[1_{\{|X_n - X| > \varepsilon\}}|X_n - X|^p\big] \leqslant \varepsilon^{-p}\|X_n - X\|^p_{L,p},$$

and hence $X_n \xrightarrow{Q} X$ whenever $X_n \to X$ in $L^p(Q)$.

(iii) Suppose that $|X_n| \leqslant K$ for all $n \in \mathbb{N}$ and some $K > 0$. Then $|X| \leqslant K$ as well, by (i). If $X_n \xrightarrow{Q} X$, then sublinearity and monotonicity of $Q[\cdot]$ imply

$$Q\big[|X_n - X|^p\big] \leqslant \varepsilon^p + 2^p K^p Q\big(|X_n - X| > \varepsilon\big) \to \varepsilon^p$$

for all $\varepsilon > 0$; thus, $X_n \to X$ in $L^p(Q)$.

(iv) If g is uniformly continuous, then for each $\varepsilon > 0$, there is some $\delta > 0$ such that $|g(x) - g(y)| \leqslant \varepsilon$ whenever $|x - y| \leqslant \delta$. Therefore

$$Q\big(|g(X_n) - g(X)| > \varepsilon\big) \leqslant Q\big(|X_n - X| > \delta\big) \to 0,$$

that is, $g(X_n) \xrightarrow{Q} g(X)$.

If g is merely continuous, we apply the previous step to the uniformly continuous function $g^m(x) \triangleq g((-m \vee x) \wedge m)$, and obtain

$$Q\big(|g(X_n) - g(X)| > \varepsilon\big) \leqslant Q\big(|g^m(X_n) - g^m(X)| > \varepsilon\big) + Q\big(|X_n - X| > 1\big) \\ + Q(|X| > m-1) \overset{n\to\infty}{\longrightarrow} Q(|X| > m-1).$$

If $X \in L^1(Q)$, then we have $Q(|X| > m) \leqslant Q[|X|]/m \to 0$, and thus $g(X_n) \overset{Q}{\longrightarrow} g(X)$. □

In the next subsection, the following complete subspace of $L^p(Q)$ and its properties will be a useful tool. We define

$$L^p_b(Q) \quad \text{as the closure of} \quad \big\{[f]_Q \in L^p(Q) : f \in \mathcal{L}^\infty(\mathcal{A})\big\} \quad \text{in } L^p(Q).$$

Similar to Denis et al. (2011), we have

LEMMA A.32. *For $X \in \mathcal{L}^0(\mathcal{A})$, we have*

$$X \in L^p_b(Q) \iff Q\left[1_{\{|X|>n\}}|X|^p\right] \to 0.$$

In this case, $(-n \vee X) \wedge n \to X$ in $L^p(Q)$. In particular, we have

$$|X-Y|,\ X \wedge Y,\ X \vee Y \in L^p_b(Q) \quad \textit{whenever } X, Y \in L^p_b(Q).$$

Proof. "⇒" Let $X \in L^p_b(Q)$ and $\varepsilon > 0$. Choose $Y \in \mathcal{L}^\infty(\mathcal{A})$ with $\|X - Y\|_{L,p} < \varepsilon$, and note that

$$\|1_{\{|X|>n\}}|X|\|_{L,p} \leqslant \|X-Y\|_{L,p} + \|Y\|_\infty Q\big(|X|>n\big)^{\frac{1}{p}} \leqslant \varepsilon + \tfrac{\|Y\|_\infty \|X\|_{L,p}}{n}.$$

"⇐" We set $X_n \triangleq (-n \vee X) \wedge n \in \mathcal{L}^\infty(\mathcal{A})$, and note that $|X_n - X| \leqslant 1_{\{|X|>n\}}|X| \to 0$ in $L^p(Q)$; thus, $X \in L^p_b(Q)$. This proves the characterization and the $L^p(Q)$-convergence statement. The characterization of $L^p_b(Q)$ immediately implies that $|Z| \in L^p_b(Q)$ whenever $Z \in L^p_b(Q)$, which immediately yields the "in particular"-statement. □

We note that Lemma A.32 shows in particular that $X \in L^p_b(Q)$ whenever $|X| \leqslant |Y|$ for some $Y \in L^p_b(Q)$.

COROLLARY A.33. *For $X \in \mathcal{L}^0(\mathcal{A})$, we have*

$$|X|^p \in L^1_b(Q) \iff X \in L^p_b(Q).$$

Proof. If $|X|^p \in L^1_b(Q)$, then

$$Q\left[1_{\{|X|>n\}}|X|^p\right] \leqslant Q\left[1_{\{|X|^p>n\}}|X|^p\right] \to 0,$$

by Lemma A.32; thus, $X \in L^p_b(Q)$. On the other hand, if $X \in L^p_b(Q)$, we put $\xi^n \triangleq (|X| \wedge n)^p \in \mathcal{L}^\infty(\mathcal{A})$ and note that

$$\left\|\xi^n - |X|^p\right\|_{L,1} \leqslant Q\big[1_{\{|X|>n\}}|X|^p\big] \to 0$$

by Lemma A.32; hence, $|X|^p \in L^1_b(Q)$. □

On $L_b^p(Q)$, we have a dominated convergence theorem:

Lemma A.34. *Let $X_n \xrightarrow{Q} X$ and suppose that $|X_n| \leqslant Y$ for all $n \in \mathbb{N}$ and some $Y \in L_b^p(Q)$. Then $X_n \to X$ in $L^p(Q)$.*

Proof. First note that $|X| \leqslant Y$ by Lemma A.31 (i). We let $\varepsilon > 0$ and set $X_n^k \triangleq (-k \vee X_n) \wedge k$ and $X^k \triangleq (-k \vee X) \wedge k$. Since $|X_n| \leqslant Y$ for all $n \in \mathbb{N}$, we have

$$\|X_n^k - X_n\|_{L^p}^p \leqslant Q\big[1_{\{|X_n|>k\}}|X_n|^p\big] \leqslant Q\big[1_{\{|Y|>k\}}|Y|^p\big] \to 0$$

by Lemma A.32. Hence, we find $k \in \mathbb{N}$ such that $\|X_n^k - X_n\|_{L^p} < \varepsilon/2$ for all $n \in \mathbb{N}$. The same argument shows that $\|X^k - X\|_{L^p} < \varepsilon/2$, as well. Therefore

$$\|X_n - X\|_{L,p} \leqslant \varepsilon + \|X_n^k - X^k\|_{L,p}.$$

Lemma A.31 (iv) implies that $X_n^k \xrightarrow{Q} X^k$ as $n \to \infty$. But $(X_n^k)_{n\in\mathbb{N}}$ is uniformly bounded by k, and thus Lemma A.31 (iii) yields the $L^p(Q)$-convergence $\|X_n^k - X^k\|_{L,p} \to 0$. □

A.2.5 *Lebesgue families*

Throughout this section, Q is an upper probability. We use our results on the spaces $L^p(Q)$ to build Lebesgue families (as in Definition 2.6). To this end, let $(\mathcal{F}_t)_{t\in[0,T]}$ be a filtration with $\mathcal{F}_T = \mathcal{A}$ and suppose that $\mathcal{F}_0$ is Q-trivial, i.e.,

$$A \in \mathcal{F}_0 \implies Q(A) = 0.$$

Defining

$$L_t^p(Q) \triangleq \big\{X \in L^p(Q) \,:\, X \cap \mathcal{L}^0(\mathcal{F}_t) \neq \emptyset\big\}$$

as the collection of all equivalence classes $X \in L^p(Q)$ which have an $\mathcal{F}_t$-measurable representative, it is straightforward to see that

$$\big\{(L_t^p(Q), \|\cdot\|_{L,p}) \,:\, t \in [0,T],\ p \geqslant 1\big\}$$

is a Lebesgue family, that is, for all $t \in [0,T]$, we have

(L1) $L_t^p(Q) \leqslant \Omega^{\mathbb{R}}/\mathcal{N}_Q$ for all $p \geqslant 1$,

(L2) $L_s^p(Q) \subset L_t^q(Q)$ for all $p \geqslant q \geqslant 1$ and $0 \leqslant s \leqslant t$,

(L3) $L_t^p(Q) = \{X \in L_t^1(Q) \,:\, |X|^p \in L_t^1(Q)\}$ for all $1 \leqslant p \leqslant \infty$,

(L4) $XY \in L_t^1(Q)$, if $X \in L_t^p(Q)$, $Y \in L_t^q(Q)$ and $\frac{1}{p} + \frac{1}{q} = 1$,

(L5) $L_0^p(Q) = i(\mathbb{R})$ for all $p \geqslant 1$, where $i : \mathbb{R} \to L_0^1(Q)$, $x \mapsto [\omega \to x]_Q$ is an order-preserving isometric isomorphism,

(L6) $L_+^p(Q) \triangleq \{X \in L^p(Q) \,:\, X \geqslant 0\} \subset L^p$ is closed.

Here, (L1) is satisfied by construction and (L2)-(L4) follow from Hölder's inequality (see Lemma A.15 and Corollary A.16). Since $\mathcal{F}_0$ is Q-trivial, every $\mathcal{F}_0$-measurable function is Q-essentially constant, and we get (L5). Pointwise convergence for subsequences (Corollary A.29) yields (L6).

To serve as an appropriate domain for a sublinear expectation, the above family is usually to large. Instead, one works with a smaller family that is obtained by taking $L^p(Q)$-closures of a $|\cdot|^p$-stable algebra of (regular) functions which contains all constants:

From here on out, $\mathcal{H} \subset \mathcal{L}^\infty(\mathcal{A})$ is a $|\cdot|^p$-stable function algebra containing all constants, i.e.,

(H1) $\mathcal{H} \subset \mathcal{L}^\infty(\mathcal{A})$,

(H2) $\alpha g + \beta h \in \mathcal{H}$ if $\alpha, \beta \in \mathbb{R}$, $g, h \in \mathcal{H}$,

(H3) $fg \in \mathcal{H}$ if $f, g \in \mathcal{H}$,

(H4) $|h|^p \in \mathcal{H}$ for all $p \geqslant 1$ if $h \in \mathcal{H}$,

(H5) $1 \in \mathcal{H}$.

We note that $|\cdot|^p$-stability (H4) entails in particular that $\mathcal{H}$ is stable under taking the maximum and minimum of finitely many elements. We have the following result:

THEOREM A.35. *Let $\mathcal{H} \subset \mathcal{L}^\infty(\mathcal{A})$ be a $|\cdot|^p$-stable algebra of functions which contains all constants, and set*

$$L_t^p \triangleq \operatorname{clos}\Big\{[h]_Q \in \Omega^{\mathbb{R}}/\mathcal{N}_Q \,:\, h \in \mathcal{H} \cap \mathcal{L}^0(\mathcal{F}_t)\Big\} \subset L_t^p(Q), \quad t \in [0,T],\ p \geqslant 1.$$

Then $\{(L_t^p, \|\cdot\|_{L,p}) : t \in [0,T],\ p \geqslant 1\}$ is a Lebesgue family.

We note that $L_t^p \subset L_b^p$ for all $t \in [0,T]$ and $p \geqslant 1$ since $\mathcal{H} \subset \mathcal{L}^\infty(\mathcal{A})$.

Proof. Let $t \in [0,T]$. Since $L_t^p(Q)$ is a Banach space, the closure L_t^p of its subspace of Q-equivalence classes of functions from $\mathcal{H} \cap \mathcal{L}^0(\mathcal{F}_t)$ is itself a Banach space. We now verify the properties (L1)-(L6) of a Lebesgue family one by one:

(L1) By definition, $L_t^p \leqslant L_t^p(Q) \leqslant L^p(Q) \leqslant \Omega^{\mathbb{R}}/\mathcal{N}_Q$.

(L2) Let $1 \leqslant q \leqslant p$ and $0 \leqslant s \leqslant t \leqslant T$. Since $\mathcal{H} \cap \mathcal{L}^0(\mathcal{F}_s) \subset \mathcal{H} \cap \mathcal{L}^0(\mathcal{F}_t)$, the same is true for their $L^q(Q)$-closures L_s^q and L_t^q. If $X_n \to X$ in $L^p(Q)$, then also $X_n \to X$ in $L^q(Q)$, as $\|X\|_{L,p} \leqslant \|X\|_{L,q}$ for all $X \in L^1(Q)$ by Corollary A.16. Thus the $L^q(Q)$-closure is bigger and $L_s^p \subset L_s^q$. All in all, we have $L_s^p \subset L_s^q \subset L_t^q$.

(L3) We have to show that $L_t^p = \{X \in L_t^1 : |X|^p \in L_t^1\}$.

"$\subset$" Let $X \in L_t^p$ and take $h_n \in \mathcal{H} \cap \mathcal{L}^0(\mathcal{F}_t)$ such that $\|h_n - X\|_{L,p} \to 0$. By Corollary A.16, this implies $\|h_n - X\|_{L,1} \to 0$, and we get $X \in L_t^1$. Moreover, we have $X \in L_b^p(Q)$ since $\mathcal{H} \subset \mathcal{L}^\infty(\mathcal{A})$. Hence

$|X|^p \in L^1_b(Q)$ by Corollary A.33, and $\xi_m \triangleq |X|^p \wedge m \to |X|^p$ in $L^1(Q)$ by Lemma A.32. To prove that $|X|^p \in L^1_t$, it thus suffices to check that $\xi_m \in L^1_t$ for all $m \in \mathbb{N}$. For this, set $\eta_n \triangleq |h_n|^p \wedge m$ and note that $\eta_n \in \mathcal{H}$ since $\mathcal{H}$ is $|\cdot|^p$-stable and contains all real constants. Now, we have $\eta_n \xrightarrow{Q} \xi_m$ as $n \to \infty$ by Lemma A.31 (iv). But then, Lemma A.31 (iii) immediately yields $\eta_n \to \xi_m$ in $L^p(Q)$; hence $\xi_m \in L^1_t$.

"$\supset$" Now, let $X \in L^1_t$ with $|X|^p \in L^1_t$, and take $h_n \in \mathcal{H} \cap \mathcal{L}^0(\mathcal{F}_t)$ such that $h_n \to X$ in $L^1(Q)$. In particular, $|X|^p \in L^1_b(Q)$, and thus $X \in L^p_b(Q)$ by Corollary A.33. Lemma A.32 shows that $\xi_m \triangleq (-m \vee X) \wedge m \to X$ in $L^p(Q)$. To prove that $X \in L^p_t$, it thus suffices to check that $\xi_m \in L^p_t$ for all $m \in \mathbb{N}$. This is again a consequence of Lemma A.31, since $\eta_n \triangleq (-m \vee h_n) \wedge m \in \mathcal{H} \cap \mathcal{L}^0(\mathcal{F}_t)$.

(L4) Let $\frac{1}{p} + \frac{1}{q} = 1$, $X \in L^p_t$, and $Y \in L^q_t$. We take $X_n, Y_n \in \mathcal{H} \cap \mathcal{L}^0(\mathcal{F}_t)$ such that $\|X_n - X\|_{L,p} + \|Y_n - Y\|_{L,q} \to 0$. Then $P_n \triangleq X_n Y_n \in \mathcal{H} \cap \mathcal{L}^0(\mathcal{F}_t)$ by (H4), and we have

$$\|P_n - XY\|_{L,1} \leqslant \|X_n(Y_n - Y)\|_{L,1} + \|Y(X_n - X)\|_{L,1}.$$

Now, Hölder's inequality (Lemma A.15) implies

$$\|P_n - XY\|_{L,1} \leqslant \|X_n\|_{L,p} \|Y_n - Y\|_{L,q} + \|Y\|_{L,q} \|X_n - X\|_{L,p} \to 0,$$

and we have shown that $XY \in L^1_t$.

(L5) Since $\mathcal{F}_0$ is Q-trivial, the intersection $\mathcal{H} \cap \mathcal{L}^0(\mathcal{F}_0)$ consists solely of Q-essentially constant functions. Hence $L^p_0 = \mathrm{clos}(i_Q(\mathbb{R}))$, where $i_Q : \mathbb{R} \to \Omega^{\mathbb{R}}/\mathcal{N}_Q$ maps a constant $x \in \mathbb{R}$ to the $\sim_Q$-equivalence class of the constant function $\omega \mapsto x$. Since i_Q is isometric, $i_Q(\mathbb{R})$ is complete, and thus $L^p_0 = \mathrm{clos}(i_Q(\mathbb{R})) = i_Q(\mathbb{R})$.

(L6) Let $0 \leqslant X_n \in L^p_t$ and suppose that $X_n \to X$ in $L^p(Q)$. By Corollary A.29, $X_{n_k} \to X$ outside a Q-negligible set for some subsequence, and thus $X \geqslant 0$ in L^p. □

A.2.6 *Sublinear expectation operators on $\mathcal{L}^\infty(\mathcal{A})$*

In Subsection A.2.1, we have seen how a family Q of finitely additive probability measures gives rise to a sublinear expectation operator

$$Q[X] = \sup_{q \in Q} q[X].$$

In this subsection, we give a partial converse: We show that every sublinear expectation operator on $\mathcal{L}^\infty(A)$ is given by a family of finitely additive probability measures. This representation result is by now classic, see, e.g., Föllmer and Schied (2004). Our presentation is close to the one in Peng (2010).

THEOREM A.36. *Let $\mathcal{H} \leqslant \mathcal{L}^\infty(\mathcal{A})$ be a linear space of bounded measurable functions containing all constants and let $\mathcal{E}$ be a real-valued function on $\mathcal{H}$. Then $\mathcal{E}$ is a sublinear expectation operator if and only if there exists a (non-empty) family Q of finitely additive probability measures such that $\mathcal{E} = Q[\cdot]|_{\mathcal{H}}$, i.e.,*

$$\mathcal{E}[h] = \sup_{q \in Q} \int h \, dq \qquad \textit{for all } h \in \mathcal{H}. \tag{A.13}$$

Since every family Q of finitely additive probability measures induces a sublinear expectation operator $Q[\cdot]$ on $\mathcal{K}(Q)$ (and hence on $\mathcal{L}^\infty(\mathcal{A})$) via (A.13), it remains to show that every sublinear expectation operator admits a representation of the form (A.13). The basic idea is simple: We prove that $\mathcal{E}$ can be represented as a supremum of continuous linear functionals on $\mathcal{L}^\infty(A)$. Then (A.13) follows, since every continuous linear functional on $\mathcal{L}^\infty(\mathcal{A})$ is given by a finitely additive measure. We begin by extending $\mathcal{E}$ to all of $\mathcal{L}^\infty(\mathcal{A})$.

LEMMA A.37. *Let $\mathcal{E}$ be a sublinear expectation operator defined on some subspace $\mathcal{H} \leqslant \mathcal{L}^\infty(\mathcal{A})$ which contains all constants. Then the envelope*

$$\bar{\mathcal{E}}[g] \triangleq \inf\{\mathcal{E}[h] : h \in \mathcal{H} \textit{ with } h \geqslant g\}, \qquad g \in \mathcal{L}^\infty(\mathcal{A}),$$

defines a sublinear expectation operator $\bar{\mathcal{E}}$ on $\mathcal{L}^\infty(\mathcal{A})$ with $\bar{\mathcal{E}}|_{\mathcal{H}} = \mathcal{E}$.

Proof. Given $g \in \mathcal{L}^\infty(\mathcal{A})$, we write $U(g) \subset \mathcal{H}$ for the collection of all $h \in \mathcal{H}$ with $h \geqslant g$. Note that $\|g\|_\infty \in U(g)$ and that $\inf_{\omega \in \Omega} g(\omega) \leqslant h$ for every $h \in U(g)$; hence,

$$-\|g\|_\infty \leqslant \inf_{\omega \in \Omega} g(\omega) \leqslant \bar{\mathcal{E}}[g] \leqslant \|g\|_\infty \quad \text{for every } g \in \mathcal{L}^\infty(\mathcal{A}).$$

In particular, $\bar{\mathcal{E}}$ maps $\mathcal{L}^\infty(\mathcal{A})$ onto $\mathbb{R}$. Moreover, since $\mathcal{E}[g] \leqslant \mathcal{E}[h]$ for all $g, h \in \mathcal{H}$ with $h \geqslant g$, it is clear that $\bar{\mathcal{E}}[g] = \mathcal{E}[g]$ for all $g \in \mathcal{H}$. Notably, $\bar{\mathcal{E}}$ preserves constants. It remains to check that $\bar{\mathcal{E}}$ is monotone and sublinear. To prove monotonicity, let $f, g \in \mathcal{L}^\infty(\mathcal{A})$ with $f \leqslant g$. Then $U(g) \subset U(f)$, and hence

$$\bar{\mathcal{E}}[f] = \inf_{h \in U(f)} \mathcal{E}[h] \leqslant \inf_{h \in U(g)} \mathcal{E}[h] = \bar{\mathcal{E}}[g].$$

To establish subadditivity, let $f, g \in \mathcal{L}^\infty(A)$ and choose $(f_n)_{n \in \mathbb{N}} \subset U(f)$ and $(g_n)_{n \in \mathbb{N}} \subset U(g)$ such that $\mathcal{E}[f_n] \to \bar{\mathcal{E}}[f]$ as well as $\mathcal{E}[g_n] \to \bar{\mathcal{E}}[g]$. Then, we have $f_n + g_n \in U(f+g)$, and we obtain

$$\bar{\mathcal{E}}[f+g] \leqslant \mathcal{E}[f_n + g_n] \leqslant \mathcal{E}[f_n] + \mathcal{E}[g_n] \to \mathcal{E}[f] + \mathcal{E}[g].$$

Positive homogeneity follows upon observing that $U(\alpha g) = \alpha U(g)$ for all $\alpha > 0$ and $g \in \mathcal{L}^\infty(A)$. □

Now, we show that $\mathcal{E}$ can be written as a supremum of continuous linear functionals on $\mathcal{L}^\infty(\mathcal{A})$.

LEMMA A.38. *Under the assumptions of Theorem A.36, there exists a family* M *of positive linear contractions* $T : \mathcal{H} \to \mathbb{R}$ *with* $T1 = 1$ *such that*

$$\mathcal{E}[h] = \sup_{T \in M} Th \quad \textit{for all } h \in \mathcal{H}. \tag{A.14}$$

Proof. In view of Lemma A.37, we may suppose without loss of generality that $\mathcal{E}$ is defined on all of $\mathcal{L}^\infty(\mathcal{A})$.

We say that a linear functional $T : \mathcal{Y} \subset \mathcal{L}^\infty(\mathcal{A}) \to \mathbb{R}$ is dominated by $\mathcal{E}$ if $Th \leqslant \mathcal{E}[h]$ for all $h \in \mathcal{Y}$, and we denote by M the set of all linear functionals on $\mathcal{L}^\infty(\mathcal{A})$ which are dominated by $\mathcal{E}$. We proceed to show that for each $h \in \mathcal{L}^\infty(\mathcal{A})$, there is some $T_h \in M$ with $T_h h = \mathcal{E}[h]$. Then M is non-empty and it immediately follows that $\mathcal{E}[h] = \sup_{T \in M} Th$.

Thus let $h \in \mathcal{L}^\infty(\mathcal{A})$ and consider the linear functional $\tau_h[\alpha \cdot h] \triangleq \alpha \cdot \mathcal{E}[h]$ on $\mathcal{Y} \triangleq \operatorname{span}\{h\} \subset \mathcal{H}$.

For all $\alpha \geqslant 0$, sublinearity of $\mathcal{E}$ implies $\tau_h[\alpha \cdot h] = \mathcal{E}[\alpha \cdot h]$ as well as

$$0 = \mathcal{E}[0] = \mathcal{E}[\alpha \cdot h + (-\alpha) \cdot h] \leqslant \alpha \cdot \mathcal{E}[h] + \mathcal{E}[(-\alpha) \cdot h],$$

which rearranges to $\tau_h[(-\alpha) \cdot h] = -\alpha \cdot \mathcal{E}[h] \leqslant \mathcal{E}[(-\alpha) \cdot h]$. Thus τ_h is dominated by $\mathcal{E}$, and the Hahn-Banach Theorem (see, e.g., Theorem 5.53 in Aliprantis and Border (2006), p. 195) yields a linear functional $T_h : \mathcal{H} \to \mathbb{R}$ which is dominated by $\mathcal{E}$ and whose restriction to $\mathcal{Y}$ coincides with τ_h. Hence $T_h \in M$ with $T_h h = \mathcal{E}[h]$ and

$$\mathcal{E}[h] = \sup_{T \in M} Th \quad \text{for all } h \in \mathcal{H},$$

which proves (A.14). It remains to show that M consists solely of positive linear contractions with $T1 = 1$:

If $T \in M$ and $0 \leqslant h \in \mathcal{L}^\infty(\mathcal{A})$, then $T[-h] \leqslant \mathcal{E}[-h] \leqslant 0$ by monotonicity of $\mathcal{E}$, and hence $Th = -T[-h] \geqslant 0$. Moreover, if $T \in M$ and $\alpha \in \mathbb{R}$, then

$$T\alpha \leqslant \mathcal{E}[\alpha] = \alpha \quad \text{and} \quad -T\alpha = T[-\alpha] \leqslant \mathcal{E}[-\alpha] = -\alpha, \quad \text{i.e.,} \quad T\alpha = \alpha,$$

because $\mathcal{E}$ dominates T and preserves constants. For the same reasons, and because $\mathcal{E}[\cdot]$ is monotone, for $T \in M$ and $h \in \mathcal{L}^\infty(A)$, we have

$$Th \leqslant \mathcal{E}[h] \leqslant \mathcal{E}[\|h\|_\infty] = \|h\|_\infty \qquad \text{and}$$
$$-(Th) = T(-h) \leqslant \mathcal{E}[-h] \leqslant \mathcal{E}[-\|h\|_\infty] = -\|h\|_\infty,$$

and hence $|Th| \leqslant \|h\|_\infty$; thus, T is a contraction. □

We now give the proof of the probabilistic representation result.

Proof of Theorem A.36. Suppose that $\mathcal{E} : \mathcal{H} \subset \mathcal{L}^\infty(\mathcal{A})$ is a sublinear expectation operator. By Lemma A.38, there exists a family M of positive linear contractions $T : \mathcal{L}^\infty(\mathcal{A}) \to \mathbb{R}$ with $T1 = 1$ such that $\mathcal{E}[h] = \sup_{T \in M} Th$ for all $h \in \mathcal{H}$. By Lemma A.13, each $T \in M$ correspond to a unique finitely additive probability m_T with $Tg = \int g \, dm_T$ for all $g \in \mathcal{L}^\infty(\mathcal{A})$. Setting $Q \triangleq \{m_T : T \in M\}$, we have

$$\mathcal{E}[h] = \sup_{q \in Q} \int h \, dq \quad \text{for all } h \in \mathcal{H}.$$ □

A.3 INTEGRATION OF VECTOR-VALUED FUNCTIONS

Let $(S, \mathcal{S}, \mu)$ be a finite measure space and $(X, |\cdot|_X)$ a Banach space. We briefly review the theory of integration for functions $f : S \to X$ with respect to μ. We focus on the aspects relevant for Chapter 2 of this thesis, complementing the short outline of Subsection 2.2.4, and confine ourself to presenting vector-valued integration as a consequence of the real-valued theory. For additional background, we refer to Dunford and Schwartz (1958) and Diestel and Uhl (1977), which are the sources of the following exposition.

The (Lebesgue-Bochner) integration theory is developed starting from simple functions. A *measurable simple function* $f : S \to X$ takes finitely many values $x_1, \dots, x_n \in X$ and satisfies $A_i \triangleq f^{-1}(\{x_i\}) \in \mathcal{S}$, $i = 1, \dots, n$; hence, $f = \sum_{i=1}^n 1_{A_i} x_i$. The collection of all measurable simple functions forms a vector space and is denoted by $E(\mu; X)$.

Integral for simple functions

The *integral of a simple function* $f \in E(\mu; X)$ with respect to μ is

$$\int_A f \, d\mu \triangleq \sum_{i=1}^n \mu(A_i \cap A) x_i \in X, \quad A \in \mathcal{S}. \tag{A.15}$$

Clearly, we have

$$\left| \int_A f \, d\mu \right|_X \leqslant \sum_{i=1}^n \mu(A_i \cap A) |x_i|_X = \int_A |f(\cdot)|_X \, d\mu, \tag{A.16}$$

where the integral on the right-hand side is the classical Lebesgue integral of the measurable function $|f(\cdot)|_X : S \to [0, \infty)$. The expression

$$\|f\|_p \triangleq \big| |f(\cdot)|_X \big|_p \triangleq \left(\int |f(\cdot)|_X^p \, d\mu \right)^{1/p}, \qquad p \geqslant 1,$$

makes sense for all measurable functions $f : S \to X$. On the space $\mathbb{L}^p(\mu; X)$ of all measurable functions $f : S \to X$ for which it is finite, the mapping $\|\cdot\|_p$ is clearly a seminorm.

In particular, $\|\cdot\|_p$ is a seminorm on the space of simple functions $E(\mu; X)$, and hence (A.16) shows that the integral $\int_A : E(\mu; X) \to X$ is a continuous linear operator with respect to that seminorm. If $(f_n)_{n \in \mathbb{N}} \subset E(\mu; X)$ is a $\|\cdot\|_p$-Cauchy sequence, then $(f_{n_k}(s))_{k \in \mathbb{N}}$ is a Cauchy sequence in X for μ-a.e. $s \in S$ and a suitable subsequence $(n_k)_{k \in \mathbb{N}}$. Thus f_{n_k} converges μ almost everywhere to some measurable $f \in \mathbb{L}^p(\mu; X)$, and we have $\|f_n - f\|_p \to 0$.

Extension of the integral

The completion $\mathcal{L}^p(\mu; X)$ of $E(\mu; X)$ in $\mathbb{L}^p(\mu; X)$ is a space of measurable functions $S \to X$ on which the (Lebesgue-Bochner) *integral* $\int_A$ is uniquely defined by continuous extension. Identifying functions that coincide μ-a.e., we obtain a Banach space, which we denote by $L^p(\mu; X)$. The following characterization of $\mathcal{L}^p(\mu; X)$ is very useful:

Characterization of integrable functions

LEMMA A.39. *For an X-valued measurable function f on S, the following statements imply each other:*

(a) $f \in \mathcal{L}^p(\mu; X)$.

(b) *f is μ-a.e. separably valued (i.e., $\{f(t) : t \in N^c\} \subset X$ is separable for some μ-null set N) and $\|f\|_p < \infty$.*

In particular, $\mathcal{L}^p(\mu;X) = \mathbb{L}^p(\mu;X)$ *if* X *is a separable Banach space.*

Proof. Every a.e. limit of a sequence of simple functions is clearly a.e. separably valued. On the other hand, if f is a.e. separably valued it is straightforward to produce a sequence of simple functions with $f_n \to f$ a.e. and $|f_n(\cdot)|_X \leqslant |f(\cdot)|_X + 1$ a.e. Then $f_n \to f \in \mathbb{L}^p(\mu;X)$ by dominated convergence, and hence $f \in \mathcal{L}^p(\mu;X)$; see also Lemma III.6.7.9 in Dunford and Schwartz (1958), p. 147. □

A.3.1 *Integration on Lebesgue families*

Let $\mathcal{L} = \{(L^p_t, \|\cdot\|_{L,p}) : t \in [0,T],\ p \geqslant 1\}$ be a *Lebesgue family* (see Definition 2.6, p. 16) and let μ be a finite Borel measure on $[0,T]$. Every one of the spaces L^p_t is a Banach space; the space $\mathcal{L}^1(\mu;L^p_t)$ of L^p_t-valued (Bochner) integrable functions was introduced above. Recall that the (Bochner) integral is defined on $\mathcal{L}^1(\mu;L^p_t)$ as the unique continuous linear operator

$$\textstyle\int_A : \mathcal{L}^1(\mu;L^p_t) \to L^p_t \quad \text{with} \quad \big\|\int_A f\,d\mu\big\|_{L,p} \leqslant \int_A \|f(\cdot)\|_{L,p}\,d\mu$$

that satisfies (A.15) for all measurable simple functions $[0,T] \to L^p_t$.

Recall Definition 2.16 (p. 22): An L^p*-process* is function $X : [0,T] \to L^p$, $X \mapsto X_t$. If $X_t \in L^p_t$ for all $t \in [0,T]$, then X is *adapted*; if X is $\mathcal{B}([0,T])$-$\mathcal{B}(L^p)$-measurable, then X is *measurable*. The space of all measurable and adapted L^p-processes is denoted by $\mathcal{X}^p(\mathcal{L})$.

The space of *adapted* μ*-integrable* L^p*-processes* is defined by

$$\mathcal{P}^p \triangleq \mathcal{P}^p(\mathcal{L},\mu) \triangleq \big\{X \in \mathcal{X}^p(\mathcal{L}) : X1_{[0,t]} \in \mathcal{L}^1(\mu;L^p_t) \text{ for all } t \in [0,T]\big\}$$

and is equipped with the seminorm

$$\|X\|_{P,p} \triangleq \int_{[0,T]} \|X_t\|_{L,p}\,\mu(dt), \qquad X \in \mathcal{P}^p(\mathcal{L},\mu).$$

Remark. If L^p_T is separable, then Lemma A.39 above shows that

$$\mathcal{P}^p(\mathcal{L},\mu) = \big\{X \in \mathcal{X}^p(\mathcal{L}) : \|X\|_{P,p} < \infty\big\}. \qquad \diamond$$

LEMMA A.40. *The seminormed space* $(\mathcal{P}^p, \|\cdot\|_{P,p})$ *is complete.*

Proof. Since $\mathcal{X}^p(\mathcal{L})$ and $\mathcal{L}^1(\mu;L^p_t)$ are linear spaces for all $t \in [0,T]$, it is clear that $\mathcal{P}^p = \mathcal{P}^p(\mathcal{L},\mu)$ is a linear space as well. It is immediate that $\|\cdot\|_{P,p}$ is a homogeneous and subadditive function on $\mathcal{X}^p(\mathcal{L})$; moreover $\|\cdot\|_{P,p}$ is finite on $\mathcal{L}^1(\mu;L^p) \supset \mathcal{P}^p$. If $(X^n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{P}^p$, then $(X^n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{L}^1(\mu;L^p)$. Hence there is some $X \in \mathcal{L}^1(\mu;L^p)$ with $\|X^n - X\|_{P,p} \to 0$, and there is a subsequence $(n_k)_{k\in\mathbb{N}}$ and some μ-null set $N \in \mathcal{B}([0,T])$ such that $\lim_{k\to\infty} X^{n_k}_t = X_t$ in L^p_t for all $t \in N^c$. Hence the process $Y \triangleq X1_{N^c}$ is both measurable and adapted. Moreover, $Y = X1_{N^c} \in \mathcal{L}^1(\mu;L^p)$ and $\|X^n - Y\|_{P,p} \to 0$. In particular, $(X^n 1_{[0,t]})_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{L}^1(\mu,L^p_t)$ and $\|X^n 1_{[0,t]} - Y1_{[0,t]}\|_{P,p} \to 0$. This shows that $Y1_{[0,t]} \in \mathcal{L}^1(\mu;L^p_t)$; hence, $Y \in \mathcal{P}^p$ and $\mathcal{P}^p$ is complete. □

By Lemma A.40, we obtain a Banach space P^p as the quotient space

$$P^p \triangleq \mathcal{P}^p/\mathcal{N}^p, \qquad \text{where } \mathcal{N}^p \triangleq \{X \in \mathcal{P}^p : \|X\|_{P,p} = 0\},$$

that is, by identifying $X, Y \in \mathcal{P}^p$ if $X_t = Y_t$ for μ-a.e. $t \in [0,T]$. We have already seen that the integral is a continuous linear operator on $\mathcal{L}^1(\mu; L^p)$, and hence we have

LEMMA A.41. *The integral*

$$\textstyle\int_A : P^p \to L^p, \quad X \mapsto \int_A X \, d\mu$$

is a continuous linear operator satisfying

$$\left\| \textstyle\int_A X \, d\mu \right\|_{L,p} \leqslant \textstyle\int_A \|X_t\|_{L,p} \, \mu(dt) \leqslant \|X\|_{P,p}, \qquad X \in P^p,$$

for all $A \in \mathcal{B}([0,T])$.

A.3.2 *Integration and classical probabilities*

We now consider the special case where the Lebesgue family

$$\mathcal{L} \triangleq \{(L^p_t, \|\cdot\|_{L,p}) : t \in [0,T], \ p \geqslant 1\}$$

consists of the classical Lebesgue spaces

$$L^p_t = L^p(\Omega, \mathcal{F}_t, P), \quad t \in [0,T], \ p \geqslant 1,$$

on a complete filtered probability space $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{t \in [0,T]}, P)$. Since all of the spaces L^p_t are separable, the space

$$\mathcal{P}^p = \mathcal{P}^p(\mathcal{L}, dt) \quad \text{of adapted } dt\text{-integrable } L^p\text{-processes}$$

consists precisely of all measurable functions $f : ([0,T], \mathcal{B}([0,T])) \to (L^p, \mathcal{B}(L^p))$ with $f(t) \in L^p_t$ for all $t \in [0,T]$ and

$$\|X\|_{P,p} = \textstyle\int_{[0,T]} \|X(t)\|_{L,p} \, dt < \infty.$$

In the present setting, one is more used to working with (progressively) measurable stochastic processes $X : \Omega \times [0,T] \to \mathbb{R}$. Let us consider the space

$$\tilde{\mathbb{P}}^p \triangleq \Big\{ X \text{ progressive} : \textstyle\int_{[0,T]} E\big[|X_t|^p\big]^{\frac{1}{p}} \, dt < \infty \Big\}.$$

Here, $t \mapsto E^P\left[|X|^p\right]^{\frac{1}{p}} = \|X_t\|_{L,p}$ is measurable by Tonelli's theorem; for $t \mapsto \sup_{q \in Q} E^q\left[|X|^p\right]^{\frac{1}{p}}$ this does not need to be the case if Q is uncountable. Clearly, the expression $\int_{[0,T]} \|X_t\|_{L,p} dt$ defines a seminorm on $\tilde{\mathbb{P}}^p$; since it is given by the same formula as the the seminorm $\|\cdot\|_{P,p}$ on $\mathcal{P}^p$, it will be denoted by $\|\cdot\|_{P,p}$ as well, with a slight abuse of notation. For the weakest norm $\|X\|_{P,1}$, the space $\tilde{\mathbb{P}}^1$ is the well-known complete seminormed space of product-integrable progressive processes.

Identifying processes in $\tilde{\mathbb{P}}^p$ that coincide $P \otimes dt$-a.e., we obtain the corresponding normed space, which we denote by $\mathbb{P}^p$.

LEMMA A.42. *For each* $p \geqslant 1$, $(\mathbb{P}^p, \|\cdot\|_{P,p})$ *is a Banach space.*

Proof. Let $\mathcal{G}$ denote the progressive σ-algebra on $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{t\in[0,T]}, P)$. Then $\mathbb{P}^1 = L^1(\Omega \times [0,T], \mathcal{G}, P \otimes dt)$ is a Banach space. If $X^n \to X$ in $\mathbb{P}^1$, then $X^{n_k} \to X$ converges $P \otimes dt$-a.e. for a subsequence.

Now, let $(X^n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\mathbb{P}^p$. Then

$$\|X^n - X^m\|_{P,1} = \int_{[0,T]} \|X(t)\|_{L,1}\, dt \leqslant \int_{[0,T]} \|X(t)\|_{L,p}\, dt = \|X^n - X^m\|_{P,p}$$

by Hölder's inequality. Thus $X^{n_k} \to X$ converges $P \otimes dt$-a.e. for a subsequence and some $X \in \mathbb{P}^1$. Fatou's lemma implies

$$\|X_t\|_{L,p} = \|\liminf_{k\to\infty} X_t^{n_k}\|_{L,p} \leqslant \sup_{n\in\mathbb{N}} \|X^n\|_{L,p} < \infty \quad \text{for } dt\text{-a.e } t \in [0,T].$$

Similarly, we obtain $\|X_t^{n_k} - X_t\|_{L,p} \leqslant \liminf_{\ell\to\infty} \|X_t^{n_k} - X_t^{n_\ell}\|_{L,p}$ for dt-a.e. $t \in [0,T]$. Integrating these inequalities from 0 to T and using Fatou's lemma again, we get

$$\|X\|_{P,p} < \infty \quad \text{and} \quad \|X^{n_k} - X\|_{P,p} \leqslant \liminf_{\ell\to\infty} \|X_t^{n_k} - X_t^{n_\ell}\|_{P,p} \overset{k\to\infty}{\longrightarrow} 0.$$

Thus $X \in \mathbb{P}^p$ and $X^n \to X$ in $\mathbb{P}^p$. □

In the following, we prove that P^p and $\mathbb{P}^p$ are isometrically isomorphic via the mapping

$$\alpha : (\mathbb{P}^p, \|\cdot\|_{P,p}) \to (P^p, \|\cdot\|_{P,p}), \quad X \mapsto \alpha(X) \triangleq X 1_{[0,\infty)}\big(E[|X|^p]\big).$$

Since $\|X\|_{P,p} = \int_0^T E[|X_t|^p]^{\frac{1}{p}} dt < \infty$, we have $E[|X_t|^p] < \infty$ for dt-a.e. $t \in [0,T]$, and hence

$$X_t = X_t 1_{[0,\infty)}\big(E[|X_t|^p]\big) = [\alpha(X)]_t \quad \text{in } L_t^p \text{ for dt-a.e. } t \in [0,T].$$

Thus $\alpha(X)$ is an adapted L^p-process which is a.e. equal to X; hence, if α is well-defined, it clearly is an isometric homomorphism. To show that α is well-defined, it remains to prove that $\alpha(X)$ is $\mathcal{B}([0,T])$-$\mathcal{B}(L^p)$-measurable. This, in turn, is a consequence of the following lemma:

LEMMA A.43. *Let* $X : \Omega \times [0,T] \to \mathbb{R}$ *be* $\mathcal{A} \otimes \mathcal{B}([0,T])$*-measurable. Then* $\alpha(X) = X 1_{[0,\infty)}(E[|X|^p])$ *is* $\mathcal{B}([0,T]) \otimes \mathcal{B}(L^p)$*-measurable.*

Proof. Since L^p is a separable Banach space, it suffices to show that the distance function

$$d_\xi : [0,T] \to \mathbb{R}, \quad t \mapsto \|\alpha(X_t) - \xi\|_{L^p} \quad \text{is } \mathcal{B}([0,T])\text{-measurable}$$

for every $\xi \in L^p$. For every $\xi \in \mathcal{L}^p(P) \subset \mathcal{L}^0(\mathcal{A})$, the process $0 \leqslant D_\xi \triangleq |X - \xi|^p$ is $\mathcal{A} \otimes \mathcal{B}([0,T])$-measurable. Hence, by Tonelli's theorem,

$$\text{the map} \quad \tilde{d}_\xi : [0,T] \to [0,\infty], \ t \mapsto E[D_t] = E\big[|X_t - \xi|^p\big] \quad \text{is measurable}$$

for every $\xi \in L^p$. Noting that

$$d_\xi = \big(\tilde{d}_\xi\big)^{\frac{1}{p}} \cdot 1_{[0,\infty)}\big(\tilde{d}_0\big) + \|\xi\|_{L,p} \cdot 1_{\{\infty\}}\big(\tilde{d}_0\big),$$

it follows that d_ξ is measurable for every $\xi \in L^p$. □

To prove that α is onto, for each $f \in P^p$, we need to produce an a.e.-modification $X \in \mathbb{P}^p$. This is achieved by an approximation argument.

LEMMA A.44. *For every* $f \in \mathcal{P}^p$, *there exists some* $X \in \tilde{\mathbb{P}}^p$ *such that*

$$X_t = f(t) \quad \textit{in } L^p \quad \textit{for } dt\textit{-a.e. } t \in [0,T].$$

Proof. For each $n \in \mathbb{N}$ and $k \in \{0, \dots, n\}$, we put $t_k^n \triangleq \lceil T \rceil k/n \wedge T$ and choose a simple function

$$f_k^n \in E\big(\mu; L^p_{t_k^n}\big) \quad \text{with} \quad \|f_k^n 1_{[0,t_k^n]} - f 1_{[0,t_k^n]}\|_{P,p} < 1/n^2.$$

Now, for each $n \in \mathbb{N}$, we set

$$f^n(t) \triangleq f(0) 1_{\{0\}}(t) + \textstyle\sum_{k=1}^n f_k^n 1_{(t_{k-1}^n, t_k^n]}(t) \in E(\mu; L^p), \qquad t \in [0,T],$$

and note that

$$\|f^n - f\|_{P,p} = \textstyle\sum_{k=1}^n \int_{t_{k-1}^n}^{t_k^n} \|f_k^n(t) - f(t)\|_{L,p}\, \mu(dt) \leqslant 1/n. \qquad \text{(A.17)}$$

Fixing some representation $\xi_k \in \mathcal{L}^p(P)$ for each one of the values $\eta_1, \dots \eta_{N_n} \in L^p$ of f^n, $n \in \mathbb{N}$, we obtain a stochastic process

$$X^n : \Omega \times [0,T] \to \mathbb{R} \quad \text{with} \quad X_t^n = f(t) \text{ in } L^p \text{ for all } t \in [0,T]. \qquad \text{(A.18)}$$

By construction, for each $t \in [0,T]$, the restriction

$$X^n : \Omega \times [0,t] \to \mathbb{R} \quad \text{is} \quad \mathcal{F}_{t+1/n} \otimes \mathcal{B}([0,t])\text{-measurable}, \qquad \text{(A.19)}$$

where $\mathcal{F}_s = \mathcal{F}_T$ for $s > T$. In particular, X^n is $\mathcal{F}_T \otimes \mathcal{B}([0,T])$-measurable.

For $n, m \in \mathbb{N}$, by Hölder's inequality and (A.18), we have

$$\int_0^T E^P\left[|X_t^n - X_t^m|\right] dt \leqslant \int_0^T E^P\left[|X_t^n - X_t^m|^p\right]^{\frac{1}{p}} dt = \|f^n - f^m\|_{P,p},$$

and hence (A.17) implies that $(X^n)_{n \in \mathbb{N}}$ is a Cauchy sequence in $L^1(\Omega \times [0,T], \mathcal{F}_T \otimes \mathcal{B}([0,T]), P \otimes dt)$. Passing to a subsequence, we may thus assume that

$$X_t^n \text{ converges in } \mathbb{R}, \quad P\text{-a.s. for } dt\text{-a.e. } t \in [0,T]. \qquad \text{(A.20)}$$

Selecting a further subsequence if necessary, we may additionally assume that

$$f^n(t) \to f(t) \quad \text{in} \quad L^p \quad \text{for } dt\text{-a.e. } t \in [0,T]. \qquad \text{(A.21)}$$

We now define $X : \Omega \times [0,T] \to \mathbb{R}$ as the measurable limit

$$X_t \triangleq 1_{\{\limsup_{n\to\infty} |X_t^n| < \infty\}} \limsup_{n\to\infty} X_t^n, \quad t \in [0,T].$$

Then (A.18), (A.20) and (A.21) show that $X_t = f(t)$ in L_t^p for dt-a.e. $t \in [0,T]$. Moreover, it follows from (A.19) that the restriction

$$X : \Omega \times [0,t] \to \mathbb{R} \quad \text{is } \mathcal{F}_{t+\varepsilon} \otimes \mathcal{B}([0,t])\text{-measurable}$$

for all $\varepsilon > 0$. Thus X is a progressively measurable a.e.-modification of f. In particular, $\|X\|_{P,p} = \|f\|_{P,p} < \infty$, and $X \in \tilde{\mathbb{P}}^p$ is as desired. □

Note: Every $f \in \mathcal{P}^p$ is a limit of measurable simple functions $[0,T] \to L^p$ and hence it has some $\mathcal{F}_T \otimes \mathcal{B}([0,T])$-measurable and adapted a.e.-modification X. Now, we can take the optional projection ${}^oX$ (see e.g., Dellacherie and Meyer (1982), Theorem 43, p. 103) as the optional (and, in particular, progressive) modification of f in Lemma A.44. As much more is known about the structure of f, we prefer the direct and elementary argument of Lemma A.44.

In view of Lemma A.43 and Lemma A.44, we have proven the following proposition.

One-to-one correspondence of $\mathbb{P}^p$ and P^p

PROPOSITION A.45. *The mapping*

$$\alpha : (\mathbb{P}^p, \|\cdot\|_{P,p}) \to (P^p, \|\cdot\|_{P,p}), \quad X \mapsto \alpha(X) \triangleq X1_{[0,\infty)}\big(E[|X|^p]\big).$$

is an isometric isomorphism. In particular, for every $X \in \mathbb{P}^p$, the image $f = \alpha(X) \in P^p$ is uniquely characterized by the property

$$f(t) = X_t \quad \textit{in } L^p \textit{ for } dt\textit{-a.e. } t \in [0,T].$$

On $(\mathbb{P}^p, \|\cdot\|_{P,p})$, the pathwise Lebesgue integral

$$\oint_A : \mathbb{P}^p \to L^1, \quad X \mapsto \oint_A X_t\, dt \triangleq \big[\omega \mapsto \int_A X_t(\omega)\, dt\big]$$

is well-defined: Indeed, $\oint_A X_t\, dt$ is measurable, and we have

$$E^P\left[\left|\oint_A X_t\, dt\right|\right] \leqslant E^P\left[\oint_A |X_t|\, dt\right] = \int_A E^P\left[|X_t|\right] dt \leqslant \|X\|_{P,p}$$

by Fubini's theorem. Thus, $\oint_A : \mathbb{P}^p \to L^1$ is a continuous linear operator. Clearly, $\oint_A$ is the canonical way to define an integral on $\mathbb{P}^p$. It is now natural to ask about the relationship between this pathwise Lebesgue integral on $\mathbb{P}^p$ and the Bochner integral $\int_A X\, dt$ on P^p. An answer is readily given:

Consitency of integration

PROPOSITION A.46. *Let $X \in \mathbb{P}^p$ and $Y \in P^p$ such that $X_t = Y_t$ in L^p for dt-a.e. $t \in [0,T]$. Then, the* pathwise *Lebesgue integral*

$$\oint_A X_t\, dt \triangleq \big[\omega \mapsto \int_A X_s(\omega)\, dt\big], \quad A \in \mathcal{B}([0,T]),$$

defines a member of L^p which coincides with the Bochner integral $\int_A Y_t\, dt$. In particular, we have

$$\oint_A X_t\, dt = \int_A [\alpha(X)]_t\, dt \quad \textit{for all } X \in \mathbb{P}^p.$$

Proof. Let X be an elementary predictable process, i.e.,

$$X_t(\omega) = \xi_0(\omega)1_{\{0\}}(t) + \sum_{k=1}^n \xi_k(\omega)1_{(t_k,t_{k+1}]}(t), \quad (\omega,t) \in \Omega \times [0,T],$$

where $0 \leqslant t_1 < \cdots < t_n \leqslant T$ and $\xi_k \in \mathcal{L}^\infty(\mathcal{F}_{t_k})$. Then $X \in \mathbb{P}^p$ and X induces a simple function $f \in S(dt; L^p)$ via $t \mapsto f(t) \triangleq X_t \in L^p$. In particular, $\alpha(X) = f \in P^p$. For each $\omega \in \Omega$, the Lebesgue integral $I(\omega)$ of $X(\omega)$ is given by

$$I(\omega) \triangleq \int_{[0,T]} X_t(\omega) dt = \sum_{k=1}^n (t_{k+1} - t_k)\xi_k(\omega), \quad \omega \in \Omega,$$

and hence

$$\oint_{[0,T]} X_t dt = \sum_{k=1}^{n}(t_{k+1}-t_k)\xi_k = \sum_{k=1}^{n} f(t_{t_k+1})(t_{k+1}-t_k)$$
$$= \int_{[0,T]} f\,dt = \int_{[0,T]} \alpha(X)\,dt \quad \text{in } L^p.$$

Thus the pathwise Lebesgue integral satisfies

$$\oint_A X_t\,dt = \int_A \alpha(X)\,dt \quad \text{in } L^p \tag{A.22}$$

for all elementary predictable processes X. In particular, by continuity of the Bochner integral (Lemma A.41), we have

$$\left\|\oint_A X_t\,dt\right\|_{L,p} = \left\|\int_A \alpha(X)\,dt\right\|_{L,p} \leqslant \|\alpha(X)\|_{P,p} = \|X\|_{P,p}. \tag{A.23}$$

By Lemma A.47 below, the collection of all elementary predictable processes is dense in $\mathbb{P}^p$; hence, we obtain

$$\oint_A X_t\,dt = \int_A \alpha(X)\,dt \quad \text{in } L^p \qquad \text{for all } X \in \mathbb{P}^p,$$

from (A.22), by continuity of the integrals (A.23) and since α is an isometric isomorphism. □

LEMMA A.47. *Let $X \in \tilde{\mathbb{P}}^p$. Then there exists a sequence $(X^n)_{n\in\mathbb{N}}$ of elementary predictable process such that $\|X^n - X\|_{P,p} \to 0$.*

Proof. We follow the standard argument, see, e.g., Ethier and Kurtz (1986), Lemma 2.2 (p. 281).

By dominated convergence, we may assume that X is bounded and it suffices to construct approximations that converge $P \otimes dt$-a.e. It is straightforward, to approximate continuous progressive processes by elementary predictable processes, and hence we only need to show that X can be approximated by continuous processes. We define a bounded, progressively measurable and continuous processes I by $I_t \triangleq \int_0^t X_s ds$. For $n \in \mathbb{N}$ and $t \in [0,T]$, we set

$$X_t^n = n(I_t - I_{(t-1/n)\vee 0}).$$

Then the X^n are uniformly bounded, progressively measurable and continuous. Moreover, for all $t > 0$ and $n > 1/t$, we have

$$X_t^n = n\big(\int_0^t X_s ds - \int_0^{t-1/n} X_s ds\big) \to X_t \quad P\text{-a.s.}$$

by Lebesgue's differentiation theorem, see e.g., Rudin (1974), Theorem 8.17, p. 176. The approximation is constructed. □

BNEE parameter via BSDE parameter

LEMMA A.48. *Let (f,ξ) be a BSDE^p-standard parameter. Then, there exists a BNEE^p-standard parameter $g : [0,T] \times L^p \to L^p$ such that*

$$\alpha\big(f(\cdot,X)\big) = g\big(\cdot,\alpha(X)\big) \qquad \textit{for all } X \in \mathbb{P}^p.$$

Proof. By (B2), $\|f(t,0)\|_{L,p} < \infty$ for all $t \in N^c$ and some dt-null set $N \in \mathcal{B}([0,T])$. We put $\bar{f} \triangleq f1_{N^c}$ and note that $(\bar{f},\xi)$ is still a BSDE^p-standard parameter. For each $\eta \in L^p$, it is possible to define

$$g(t,\eta) \triangleq \bar{f}(\cdot,t,\eta) \in L^p.$$

We claim that g maps $[0,t] \times L^p_t$ into L^p_t:

For each $s \in [0,t]$ and all $\eta \in L^p_t$, the function $\omega \mapsto \bar{f}(\omega,s,\eta(\omega))$ is easily seen to be $\mathcal{F}_t$-measurable. Moreover, by (B2), we have

$$\|g(s,\eta)\|_{L,p} = \|\bar{f}(\cdot,s,\eta)\|_{L,p} \leqslant L\|\eta\|_{L,p} + 1_{N^c}(t)\|f(s,0)\|_{L,p} < \infty. \quad \text{(A.24)}$$

This shows that $g(s,\eta)$ is in L^p_t whenever $\eta \in L^p_t$ and $s \in [0,t]$. In particular, g maps $[0,T] \times L^p$ into L^p.

For all $\eta,\zeta \in L^p$, (B2) implies

$$\|g(t,\eta) - g(t,\zeta)\|_{L^p} \leqslant E\big[L^p|\eta-\zeta|^p\big]^{\frac{1}{p}} = L\|\eta-\zeta\|_{L,p},$$

which establishes the Lipschitz property (2.14) of a BNEE^p-standard parameter. To prove that g is a BNEE^p-standard parameter, it thus remains to show that $g(\cdot,X) \in P^p$ whenever $X \in S^p$.

Let $Y \in P^p$. Proposition 2.66 yields $X \in \mathbb{P}^p$ with $\alpha(X) = Y$ in P^p. Replacing X by $X1_{[0,\infty)}(\|X\|_{L,p})$ if necessary, we may assume that

$$\|X_t\|_{L,p} < \infty \quad \text{and} \quad X_t = Y_t \quad \text{in } L^p \quad \text{for all } t \in [0,T].$$

Now, $\bar{X} \triangleq \bar{f}(\cdot,X) \in \mathbb{P}^p$ by (2.54) above, and $\|\bar{X}_t\|_{L,p} < \infty$ for all $t \in [0,T]$ by (A.24). Therefore, as a consequence of the definition of α in Proposition 2.66, the L^p-process $\bar{Y} \triangleq \alpha(\bar{X}) \in P^p$ satisfies

$$\bar{Y}_t = \bar{X}_t = \bar{f}(t,X_t) = \bar{f}(t,Y_t) = g(t,Y_t) \quad \text{in } L^p \text{ for all } t \in [0,T].$$

Thus $g(\cdot,Y) = \bar{Y} \in P^p$ for all $X \in P^p \supset S^p$. Finally, by the definitions of g and α, for any $X \in \mathbb{P}^p$, we have

$$\big[\alpha\big(f(\cdot,X)\big)\big]_t = f(t,X_t) = f(t,X_t)1_{N^c}(t) = g\big(t,[\alpha(X)]_t\big) \quad \text{in } L^p$$

for dt-a.e. $t \in [0,T]$, that is, $\alpha(f(\cdot,X)) = g(\cdot,\alpha(X))$ in P^p. □

B

APPENDIX: STOCHASTIC DIFFERENTIAL UTILITY

This appendix supplements Chapter 3. In Section B.1, we collect elementary properties of the Epstein-Zin aggregator and provide some proofs that have been omitted from the main text. Section B.2 is concerned with showing that

$$X \mapsto \int_t^s \frac{d[X]^c_\tau}{X_{\tau-}}$$

is a convex mapping on the set of all positive semimartingales. This fact is used in Section 3.5 to prove concavity of stochastic differential utility. In Section B.3, we give several results required in the proof of the general stochastic Gronwall inequality in Section 3.4.

Throughout this appendix, we adhere to the notation of Chapter 3, and consequently

$$\gamma \in (0,\infty),\ \gamma \neq 1, \qquad \psi \in (0,\infty),\ \psi \neq 1, \quad \text{and} \quad \delta > 0$$

are fixed constants. Moreover,

$$\mathfrak{C} \triangleq (0,\infty) \quad \text{and} \quad \mathfrak{U} \triangleq (1-\gamma)\mathfrak{C}$$

as well as

$$\phi \triangleq 1/\psi, \quad \theta \triangleq \tfrac{1-\gamma}{1-\phi}, \quad \text{and} \quad q \triangleq \tfrac{\theta-1}{\theta} = \tfrac{\phi-\gamma}{1-\gamma}.$$

By $f : \mathfrak{C} \times \mathfrak{U} \to \mathbb{R}$, we denote the Epstein-Zin aggregator

$$f(c,v) \triangleq \tfrac{\delta}{1-\phi} c^{1-\phi}[(1-\gamma)v]^q - \delta\theta v.$$

B.1 PROPERTIES OF THE EPSTEIN-ZIN AGGREGATOR

In the following, we collect elementary properties of the Epstein-Zin aggregator. The next lemma lists its derivatives.

Derivatives of Epstein-Zin aggregator

LEMMA B.1. *For all* $c \in \mathfrak{C}$, $v \in \mathfrak{U}$, *we have*

$$\begin{aligned}
f_c(c,v) &= \delta c^{-\phi}[(1-\gamma)v]^{\frac{\phi-\gamma}{1-\gamma}} \\
f_{cc}(c,v) &= -\phi\delta c^{-\phi-1}[(1-\gamma)v]^{\frac{\phi-\gamma}{1-\gamma}} \\
f_{cv}(c,v) &= \delta(\phi-\gamma)c^{-\phi}[(1-\gamma)v]^{\frac{\phi-1}{1-\gamma}} \\
f_v(c,v) &= \delta\tfrac{\phi-\gamma}{1-\phi}c^{1-\phi}[(1-\gamma)v]^{\frac{\phi-1}{1-\gamma}} - \delta\tfrac{1-\gamma}{1-\phi} \\
f_{vv}(c,v) &= \delta(\gamma-\phi)c^{1-\phi}[(1-\gamma)v]^{\frac{\phi-1}{1-\gamma}-1},
\end{aligned}$$

and, in particular,

$$f_c > 0, \quad f_{cc} < 0, \quad \text{sign}(f_{cv}) = \text{sign}(\phi - \gamma), \quad \text{sign}(f_{vv}) = \text{sign}(\gamma - \phi).$$

Properties of Epstein-Zin aggregator

Thus f is always increasing and concave in c. Moreover, f is convex in v if $\gamma > \phi$ and concave in v if $\gamma < \phi$. Finally, $f : \mathfrak{C} \times \mathfrak{U} \to \mathbb{R}$ is concave if and only if $\gamma < \phi$.

Proof. To prove the statement concerning concavity of f (jointly in c, v), we note that

$$(f_{cc} \cdot f_{vv})(c, v) = -\phi(\gamma - \phi)\delta^2 c^{-2\phi}[(1-\gamma)v]^{2\frac{\phi-1}{1-\gamma}}, \quad \text{and}$$
$$f_{cv}(c, v)^2 = (\phi - \gamma)^2 \delta^2 c^{-2\phi}[(1-\gamma)v]^{2\frac{\phi-1}{1-\gamma}}$$

so that

$$(f_{cc} \cdot f_{vv})(c, v) - f_{cv}(c, v)^2 = (\gamma(\phi - \gamma))\, \delta^2 c^{-2\phi}[(1-\gamma)v]^{2\frac{\phi-1}{1-\gamma}}.$$

Since $f_{cc} < 0$, this shows that the Hessian of f is negative definite if and only if $\phi > \gamma$ and otherwise indefinite. □

COROLLARY B.2. *The derivative f_v of the Epstein-Zin aggregator enjoys the following boundedness properties:*

f_v	$\phi < 1$	$\phi > 1$
$\gamma > \phi$	bounded above	bounded below
$\gamma < \phi$	bounded below	bounded above

Table 2: Boundedness of f_v

In the following, given $\rho \in (0, \infty) \setminus \{1\}$, we write $u_\rho(c) = c^{1-\rho}/(1-\rho)$ for the corresponding ρ-power utility function.

LEMMA B.3. *For all $c \in \mathfrak{C}$, we have*

$$f(c, u_\gamma(c)) = 0 \quad \textit{and} \quad f_v(c, u_\gamma(c)) = -\delta.$$

Moreover, we have

$$f(c, v) \geqslant \delta[u_\gamma(c) - v] \quad \textit{if} \quad \gamma \geqslant \phi, \qquad \textit{and}$$
$$f(c, v) \leqslant \delta[u_\gamma(c) - v] \quad \textit{if} \quad \gamma \leqslant \phi.$$

Proof. The first claim follows immediately by direct calculation. For the second, recall that $v \mapsto f(c, v)$ is convex for $\gamma > \phi$ by Lemma B.1. Therefore, we have

$$f(c, v) \geqslant f(c, u_\gamma(c)) + f_v(c, u_\gamma(c))\big[v - u_\gamma(c)\big] = \delta[u_\gamma(c) - v].$$

If $\gamma < \phi$, then $v \mapsto f(c, v)$ is concave for $\gamma > \phi$ by Lemma B.1 and all of the above inequalities are reversed. Finally, if $\gamma = \phi$ then the Epstein-Zin aggregator degenerates to $f(c, v) = \delta[u_\gamma(c) - v]$. □

LEMMA B.4. *Let V be a semimartingale with dynamics*

$$dV_t = -f(c_t, V_t)dt + dM_t,$$

where M is a local martingale, and suppose that

$$(1-\gamma)V > 0 \quad \text{and} \quad (1-\gamma)V_- > 0. \tag{B.1}$$

Then, the process $Y = u_\phi \circ u_\gamma^{-1}(V)$ satisfies

$$dY_t = -\left(\delta\left[u_\phi(c_t) - Y_t\right]dt + dA_t\right) + \left((1-\gamma)V_{t-}\right)^{-q} dM_t, \tag{B.2}$$

where the process A is increasing if $\phi \geqslant \gamma$ and decreasing if $\phi \leqslant \gamma$; the process A is given by

$$A_t = \int_0^t \frac{\phi-\gamma}{2} \frac{d[Y]_s^c}{(1-\phi)Y_{s-}} + \frac{1}{1-\gamma} \sum_{0<s\leqslant t} J\left((1-\phi)Y_{s-}, (1-\phi)Y_s\right),$$

where $J(x,y) = x^{1-\theta}\left(y^\theta - x^\theta\right) - \theta(y-x)$.

Proof. We set $g(v) \triangleq u_\phi \circ u_\gamma^{-1}(v) = \frac{1}{1-\phi}((1-\gamma)v)^{1-q}$, and note that $g'(v) = ((1-\gamma)v)^{-q}$ and $g''(v) = (\gamma-\phi)((1-\gamma)v)^{-q-1}$. Now, $Y = g(V)$ and, by (B.1), we have

$$(1-\phi)Y > 0 \quad \text{and} \quad (1-\phi)Y_- > 0. \tag{B.3}$$

A direct calculation using Itō's formula (see e.g., Theorem I.4.57 in Jacod and Shiryaev (2003)) yields

$$dY_t = -\delta\left[u_\phi(c_t) - Y_t\right]dt + g'(V_{t-})dM_t + \tfrac{1}{2}g''(V_{t-})d[M^c]_t - dR_t, \tag{B.4}$$

where R is of finite variation and satisfies

$$dR_t = g'(V_{t-})\Delta V_t - \Delta Y_t = \tfrac{1}{1-\gamma}\left((1-\phi)Y_{t-}\right)^{1-\theta}\Delta\left((1-\phi)Y_t\right)^\theta - \Delta Y_t.$$

In particular, we have

$$dY_t^c = g'(V_{t-})dM_t^c, \quad \text{and hence} \quad d[Y]_t^c = g'(V_{t-})^2 d[M^c]_t.$$

For the continuous quadratic variation part in (B.4), we thus obtain

$$g''(V_{t-})d[M^c]_t = \frac{g''(V_{t-})}{g'(V_{t-})^2} d[Y]_t^c = (\gamma-\phi)\frac{d[Y]_t^c}{(1-\phi)Y_{t-}}.$$

Inserting this into (B.4), we arrive at

$$dY_t = -\left(\delta\left[u_\phi(c_t) - Y_t\right]dt + \frac{\phi-\gamma}{2}\frac{d[Y]_t^c}{(1-\phi)Y_{t-}} + dR_t\right) + g'(V_{t-})dM_t,$$

which establishes (B.2) with $dA_t = \frac{\phi-\gamma}{2}\frac{d[Y]_t^c}{(1-\phi)Y_{t-}} + dR_t$.

If $\gamma \geqslant \phi$, then $\frac{\phi-\gamma}{2}\frac{d[Y]_t^c}{(1-\phi)Y_{t-}}$ is decreasing by (B.3). Moreover, g is convex, and hence we have

$$g(V_t) \geqslant g(V_{t-}) + g'(V_{t-})(V_t - V_{t-}).$$

Therefore only jumps of non-positive height occur in R and thus R is decreasing. Consequently,

$$dA_t = \tfrac{\phi-\gamma}{2}\tfrac{d[Y]_t^c}{(1-\phi)Y_{t-}} + dR_t$$

is a decreasing process. If $\gamma \leqslant \phi$, then g is concave and $\frac{\phi-\gamma}{(1-\phi)Y_{t-}} \geqslant 0$, and hence A is increasing. □

B.2 A CONVEXITY RESULT

We now prove that

$$X \mapsto \int_t^s \frac{d[X]^c_\tau}{X_{\tau-}} \tag{B.5}$$

is a convex mapping on the set of all positive semimartingales. We start with an auxiliary result.

LEMMA B.5. *Consider the convex subset* $E \triangleq \{(x,y) \in \mathbb{R}^2 : x+y > 0\} \subset \mathbb{R}^2$ *and let*

$$h : E \times (0,\infty) \to (0,\infty), \quad ((x,y),\bar{x}) \mapsto \frac{(\bar{x}-x)^2}{x+y}.$$

Then h *is convex.*

Proof. The Hessian H of h is given by

$$H((x,y),\bar{x}) = \begin{pmatrix} \frac{2\,(\bar{x}+y)^2}{(x+y)^3} & -\frac{2\,(\bar{x}+y)}{(x+y)^2} & \frac{2\,(\bar{x}+y)\,(\bar{x}-x)}{(x+y)^3} \\ -\frac{2\,(\bar{x}+y)}{(x+y)^2} & \frac{2}{x+y} & \frac{2\,x-2\,\bar{x}}{(x+y)^2} \\ \frac{2\,(\bar{x}+y)\,(\bar{x}-x)}{(x+y)^3} & \frac{2\,x-2\,\bar{x}}{(x+y)^2} & \frac{2\,(\bar{x}-x)^2}{(x+y)^3} \end{pmatrix}.$$

Since $x+y > 0$, we get $\frac{2\,(\bar{x}+y)^2}{(x+y)^3} \geqslant 0$. The remaining two principal minors vanish. Hence H is positive semidefinite on $E \times (0,\infty)$ by Sylvester's criterion, and h is convex. □

COROLLARY B.6. *For each* $n \in \mathbb{N}$ *the function* $F^n : E^n \times (0,\infty) \to \mathbb{R}$,

$$((x_0,y_0),(x_1,y_1),\dots,(x_{n-1},y_{n-1}),x_n) \mapsto \sum_{k=0}^{n-1} \frac{(x_{k+1}-x_k)^2}{x_k+y_k}$$

is convex.

Proof. Clearly $F^n = \sum_{k=0}^{n-1} h^n_k$ with

$$h^n_k\,((x_0,y_0),(x_1,y_1),\dots,(x_{n-1},y_{n-1}),x_n) \triangleq h((x_k,y_k),x_{k+1}),$$

and so F^n is convex by Lemma B.5. □

To exploit the convexity result of Corollary B.6, we rely on a Riemann sum approximation of the integral in (B.5).

LEMMA B.7. *Let* X *and* Y *be semimartingales such that* $X+Y$ *is bounded away from zero and let* $t,s \in [0,T]$, $t \leqslant s$. *For each* $n \in \mathbb{N}$, *let* $t = t^n_0 < t^n_1 < \dots < t^n_n = s$ *be a partition of* $[t,s]$ *and set*

$$I^n_{t,s}[X,Y] \triangleq F^n\left(X_{t^n_0}, Y_{t^n_0}, \dots, X_{t^n_{n-1}}, Y_{t^n_{n-1}}, X_{t^n_n}\right).$$

If $\max_{k=1,\dots,n} |t^n_k - t^n_{k-1}| \to 0$ *as* $n \to \infty$, *we have*

$$I^n_{t,s}[X,Y] \to \int_t^s \frac{d[X]_\tau}{X_{\tau-}+Y_{\tau-}} \quad \textit{in probability.}$$

Proof. With $Q^n_\tau[X] \triangleq \sum_{k=1}^n (X_{t^n_k\wedge\tau} - X_{t^n_{k-1}\wedge\tau})^2$, $\tau \in [0,T]$, Theorem 4.47 in Jacod and Shiryaev (2003) implies that $\sup_{\tau\in[0,T]} |Q^n_\tau[X] - [X]_\tau| \to 0$ in probability. The Riemann sums

$$J^n_{t,s}[X,Y] \triangleq \sum_{k=1}^n (X_{t^n_{k-1}} + Y_{t^n_{k-1}})^{-1}([X]_{t^n_k} - [X]_{t^n_{k-1}})$$

satisfy $J^n_{t,s}[X,Y] \to \int_t^s \frac{d[X]_\tau}{X_{\tau-}+Y_{\tau-}}$ as $n \to \infty$. Taking $\varepsilon > 0$ with $X + Y \geqslant \varepsilon$ we obtain

$$|I^n_{t,s}[X,Y] - J^n_{t,s}[X,Y]| \leqslant \frac{2}{\varepsilon} \sup_{\tau\in[0,T]} |Q^n_\tau[X] - [X]_\tau| \to 0 \quad \text{in probability,}$$

and hence the claim follows. □

Recall that $\mathcal{S}$ denotes the space of all semimartingales. Combining Lemma B.7 with Corollary B.6, we get the essence of our convexity result.

LEMMA B.8. *Let* $\mathcal{D} \triangleq \{(X,Y) \in \mathcal{S}\times\mathcal{S} : X+Y > 0\}$. *For* $t,s \in [0,T]$ *with* $t \leqslant s$, *we define*

$$I_{t,s} : \mathcal{D} \to L^0_+(P), \quad (X,Y) \mapsto I_{t,s}[X,Y] \triangleq \int_s^t \frac{d[X]_\tau}{X_{\tau-}+Y_{\tau-}}.$$

Then $I_{t,s}$ *is convex.*

Proof. Let $(X^i,Y^i) \in \mathcal{D}$ for $i = 1,2$ and $\lambda \in (0,1)$. For each $\varepsilon > 0$ we put $Y^{i,\varepsilon} \triangleq Y^i + \varepsilon$, $i = 1,2$. Then, by Lemma B.7, we have

$$I^n_{t,s}\left[X^i, Y^{i,\varepsilon}\right] \to \int_t^s \frac{d[X^i]_\tau}{X^i_{\tau-}+Y^{i,\varepsilon}_{\tau-}} = I_{t,s}\left[X^i,Y^{i,\varepsilon}\right] \quad \text{in probability, } i = 1,2.$$

With $(X,Y^\varepsilon) \triangleq (\lambda X^1 + (1-\lambda)X^2, \lambda Y^{1,\varepsilon} + (1-\lambda)Y^{2,\varepsilon})$, the same result yields

$$I^n_{t,s}[X,Y^\varepsilon] \to \int_t^s \frac{d[X]_\tau}{X_{\tau-}+Y^\varepsilon_{\tau-}} = I_{t,s}[X,Y^\varepsilon] \quad \text{in probability.}$$

Now, Corollary B.6 implies that

$$I^n_{t,s}[X,Y^\varepsilon] \leqslant \lambda I^n_{t,s}\left[X^1,Y^{1,\varepsilon}\right] + (1-\lambda) I^n_{t,s}\left[X^2,Y^{2,\varepsilon}\right],$$

so letting $n \to \infty$, we obtain

$$\int_t^s \frac{d[X]_\tau}{X_{\tau-}+Y^\varepsilon_{\tau-}} \leqslant \lambda \int_t^s \frac{d[X^1]_\tau}{X^1_{\tau-}+Y^{1,\varepsilon}_{\tau-}} + (1-\lambda)\int_t^s \frac{d[X^2]_\tau}{X^2_{\tau-}+Y^{2,\varepsilon}_{\tau-}}.$$

Sending $\varepsilon \downarrow 0$, we conclude by monotone convergence. □

COROLLARY B.9. *Let* $\mathcal{S}_\bullet \triangleq \{X \in \mathcal{S} : X > 0\}$ *denote the set of positive semimartingales, and let* $t,s \in [0,T]$, $t \leqslant s$. *The mapping*

$$\mathcal{S}_\bullet \to L^0_+(P), \quad X \mapsto \int_t^s \frac{d[X]^c_\tau}{X_{\tau-}} \qquad \text{is convex.}$$

Proof. For every semimartingale X there exists a unique continuous local martingale X^c such that $[X^c] = [X]^c$; see, e.g., Proposition I.4.27 and Theorem I.4.52 in Jacod and Shiryaev (2003). Uniqueness implies in particular that the mapping $D : \mathcal{S} \to \mathcal{S}\times\mathcal{S}$, $X \mapsto (X^c, X - X^c)$ is linear. Since $D(\mathcal{S}_\bullet) \subset \mathcal{D}$, Lemma B.8 implies that $I_{t,s} \circ D$ is convex. □

B.3 PROOF OF THE STOCHASTIC GRONWALL INEQUALITY

This section supplements Section 3.4 on Gronwall inequalities. After giving several preliminaries, we shall prove that iterating the linear integral inequality

$$X_t \leqslant E_t\left[\int_t^T (\alpha_s X_s + \beta_s) d + Z\right], \quad t \in [0,T], \tag{G}$$

is feasible under the assumptions of Theorem 3.15 (p. 87). More precisely, we will show that

$$X_t \leqslant E_t\left[\int_t^T \left(\mathcal{E}_s^{t,n} H_s + \tfrac{I_s^{t,n}}{n!}\alpha_s X_s\right) ds + \mathcal{E}_T^{t,n} Z\right], \quad t \in [0,T],\ n \in \mathbb{N}.$$

To carry out the proof, we first recall some well-known results on the measure of the n-dimensional simplex.

Measure of the simplex

Let $0 \leqslant s \leqslant t \leqslant T$. For each $n \in \mathbb{N}$, we let $\mathcal{S}^n_{(s,t)}$ denote the n-dimensional simplex within the open cube $(s,t)^n$, i.e.,

$$\mathcal{S}^n_{(s,t)} = \left\{(s_1,\dots,s_n) \in (s,t)^n \,:\, t < s_1 < s_2 < \dots < s_n < s\right\}, \quad n \in \mathbb{N}.$$

Given a finite atomless Borel measure μ on $[0,T]$, we denote its n-fold product measure by $\mu^{\otimes n}$. In the following, we are interested in the volume $\mu^{\otimes n}(\mathcal{S}^n_{(s,t)})$; we will use the convention $\mu^{\otimes 0}(\mathcal{S}^0_{(s,t)}) = 1$.

LEMMA B.10. *For all $n \in \mathbb{N}$, we have*

$$\mu^{\otimes n}(\mathcal{S}^n_{(s,t)}) = \int_{(s,t)} \mu^{\otimes(n-1)}(\mathcal{S}^{n-1}_{(s,s_n)})\mu(ds_n)$$

Proof. Fubini's theorem implies

$$\mu^{\otimes n}(\mathcal{S}^n_{(s,t)}) = \int_{(s,t)} \mu^{\otimes(n-1)}(\mathcal{S}^n_{(s,t)}|_{s_n})\mu(ds_n),$$

where the s_n-section of $\mathcal{S}^n_{(s,t)}$,

$$\mathcal{S}^n_{(s,t)}|_{s_n} = \{(t_1,\dots t_{n-1}) \in (t,s)^{n-1} \,:\, (t_1,\dots t_{n-1},t_n) \in \mathcal{S}^{t,n}_s,\ t_n = s_n\},$$

is simply $\mathcal{S}^{n-1}_{(s,s_n)}$. Thus

$$\mu^{\otimes n}(\mathcal{S}^n_{(s,t)}) = \int_t^s \mu^{\otimes(n-1)}(\mathcal{S}^{n-1}_{(s,s_n)})\mu(ds_n),$$

establishing the claimed identity. □

LEMMA B.11. *For all $n \in \mathbb{N}$, we have*

$$\mu^{\otimes n}(\mathcal{S}^n_{(s,t)}) = \tfrac{1}{n!}\mu\big((s,t)\big)^n.$$

Proof. To compute the volume, let $\sigma \in \mathcal{S}_n$ be a permutation and put

$$\mathcal{S}^{n,\sigma}_{(s,t)} \triangleq \{(s_1,\dots,s_n) \in (s,t)^n : t < s_{\sigma_1} < s_{\sigma_2} < \cdots < s_{\sigma_n} < s\}.$$

Then $\mu^{\otimes n}(\mathcal{S}^{n,\sigma}_{(s,t)}) = \mu^{\otimes n}(\mathcal{S}^{n}_{(s,t)})$ since product measures are unaffected by coordinate permutations. Moreover, we have

$$\bigcup_{\sigma\in\mathcal{S}_n} \mathcal{S}^{n,\sigma}_{(s,t)} = (s,t)^n \setminus \bigcup_{i,j=1}^n N^{s,t,n}_{i,j},$$

where $N^{s,t,n}_{i,j} \triangleq \{(s_1,\dots,s_n) \in (s,t)^n : s_i = s_j\}$. Since μ is atomless, it follows that $\mu^{\otimes n}(N^{s,t,n}_{i,j}) = 0$ for all $i,j,=1,\dots,n$. Hence, we obtain

$$\mu^{\otimes n}(\mathcal{S}^{n,\sigma}_{(s,t)}) \cdot n! = \sum_{\sigma\in\mathcal{S}_n} \mu^{\otimes n}(\mathcal{S}^{n,\sigma}_{(s,t)}) = \mu((t,s))^n. \qquad \square$$

COROLLARY B.12. *Let α be a non-negative progressively measurable process and define the random measure μ via $\mu((a,b]) = \int_a^b \alpha_s ds$. Then*

$$\mu^{\otimes n}(\mathcal{S}^n_{(s,t)}) = \int_t^s \mu^{\otimes(n-1)}(\mathcal{S}^n_{(s,u)})\alpha_u du, \qquad n \in \mathbb{N}, \tag{B.6}$$

where $\mu^{\otimes 0}(\mathcal{S}^0_{(t,s)}) = 1$; in particular, $(\mu^{\otimes n}(\mathcal{S}^n_{(t,s)}))_{s\in[t,T]}$ is progressively measurable for all $n \in \mathbb{N}$ and each $t \in [0,T]$. Moreover,

$$\mu^{\otimes n}(\mathcal{S}^n_{(t,s)}) = \tfrac{1}{n!}\left(\int_t^s \alpha_u du\right)^n, \qquad n \in \mathbb{N}.$$

Proof. Apply Lemmas B.10 and B.11 path-by-path. Since $\mu^{\otimes 0}(\mathcal{S}^0_{(t,s)}) = 1$ is trivially progressive, progressiveness of $(\mu^{\otimes n}(\mathcal{S}^n_{(t,s)}))_{s\in[t,T]}$ follows from (B.6) by induction. $\square$

Conditional Fubini theorem

To iterate the inequality, we need to interchange conditional expectations and integrals with respect to time. The following Fubini type theorem for conditional expectations shows that this is possible.

Conditional Fubini

PROPOSITION B.13. *Let $\mathcal{G} \subset \mathcal{A}$ be a sub-σ-field and $Y = (Y_s)_{s\in[t,T]}$ a measurable process with $\int_t^T E[|Y_s|]ds < \infty$. There exists a measurable process $H = (H_s)_{s\in[t,T]}$ with*

$$H_s = E\left[Y_s \mid \mathcal{G}\right] \quad \textit{for all } s \in [t,T]. \tag{B.7}$$

Moreover, any such process satisfies

$$E\left[\int_t^s Y_u du \mid \mathcal{G}\right] = \int_t^s H_u du \quad \textit{for all } s \in [t,T]. \tag{B.8}$$

Proof. By Proposition 4.6 in Ethier and Kurtz (1986) (p. 74), there exists a measurable process H which satisfies (B.7) and (B.8). Suppose now that K is a measurable process with $K_s = E[Y_s \mid \mathcal{G}]$ for all $s \in [t,T]$. Then

$$E\left[\left|\int_t^s H_u du - \int_t^s K_u du\right|\right] \leqslant E\left[\int_t^s |H_u - K_u| du\right] = \int_t^s E\left[|H_u - K_u|\right] du = 0$$

by Fubini's theorem, and hence

$$\int_t^s K_u du = \int_t^s H_u du = E\left[\int_t^s Y_u du \mid \mathcal{G}\right] \quad \text{a.s. for all } s \in [t,T]. \qquad \square$$

Proof of the stochastic Gronwall inequality

We are now in a position to iterate the linear integral inequality

$$X_t \leqslant E_t\left[\int_t^T(\alpha_s X_s + \beta_s)ds + Z\right], \quad t \in [0,T], \tag{G}$$

where $0 \leqslant \alpha$, X and H are progressive processes and Z is a random variable. Recall that the assumptions of Theorem 3.15 include the following integrability conditions: For all $t \in [0,T]$, we have

$$E\left[\int_t^T(\mathcal{E}_s^{t,n}|H_s| + I_s^{t,n}\alpha_s|X_s|)ds + \mathcal{E}_T^{t,n}|Z|\right] < \infty \quad \text{for all } n \in \mathbb{N}, \tag{I}$$

where

$$I_s^{t,n} \triangleq \tfrac{1}{n!}\left(\int_t^s \alpha_u du\right)^n \quad \text{and} \quad \mathcal{E}_s^{t,n} \triangleq \sum_{k=0}^n I_s^{t,k}, \qquad 0 \leqslant t \leqslant s \leqslant T.$$

Iterating the inequality

LEMMA B.14. *Under the assumptions of Theorem 3.15, we have*

$$X_t \leqslant E_t\left[\int_t^T\left(\mathcal{E}_s^{t,n}H_s + \tfrac{I_s^{t,n}}{n!}\alpha_s X_s\right)ds + \mathcal{E}_T^{t,n}Z\right], \quad t \in [0,T],\ n \in \mathbb{N}.$$

Proof. With Corollary B.12, we have established the identity

$$\tfrac{1}{n!}I_s^{t,n} = \mu^{\otimes n}(\mathcal{S}_s^{t,n}) = \int_t^s \mu^{\otimes(n-1)}(\mathcal{S}_u^{t,n-1})\alpha_u\, du, \quad n \in \mathbb{N}, \tag{B.9}$$

where μ denotes the random measure with $\mu((a,b]) = \int_a^b \alpha_u du$. Thus, we have to show that

$$X_t \leqslant E_t\left[\int_t^T\left(\mathcal{E}_s^{t,n}H_s + \mu^{\otimes n}(\mathcal{S}_s^{t,n})\alpha_s X_s\right)ds + \mathcal{E}_T^{t,n}Z\right], \quad t \in [0,T],\ n \in \mathbb{N}. \tag{B.10}$$

We proceed by induction; since $\mu^{\otimes 0}(\mathcal{S}_s^{t,0}) = 1$, for $n = 0$, inequality (B.10) is just the linear integral inequality (G),

$$X_s \leqslant E_s\left[\int_s^T(\alpha_u X_u + H_u)du + Z\right] =: \bar{X}_s, \quad s \in [0,T], \tag{B.11}$$

which X satisfies by assumption. Assume by induction that (B.10) holds for $n \in \mathbb{N}$. To perform the induction step, we note that

$$\bar{X}_s = E_s\left[\int_0^T(\alpha_u X_u + H_u)du + Z\right] - \int_0^s(\alpha_u X_u + H_u)du,$$

and so we can choose a progressive modification of $\bar{X}$. Since $\alpha \geqslant 0$, by (B.11), we thus obtain

$$\mathcal{E}_s^{t,n}H_s + \mu^{\otimes n}(\mathcal{S}_s^{t,n})\alpha_s X_s \leqslant \mathcal{E}_s^{t,n}H_s + \mu^{\otimes n}(\mathcal{S}_s^{t,n})\alpha_s \bar{X}_s =: J_s^{t,n} \tag{B.12}$$

for all $s \in [t,T]$, where $J^{t,n}$ is a progressively measurable process. Hence, the induction hypothesis (B.10) implies

$$X_t \leqslant E_t\left[\int_t^T J_s^{t,n}ds + \mathcal{E}_T^{t,n}Z\right], \qquad s \in [t,T]. \tag{B.13}$$

To proceed, we have to interchange the order of integration. By Proposition B.13, this is possible if $J^{t,n}$ is integrable.

To establish the necessary integrability, we first note that

$$E\left[\int_t^T \left(\mathcal{E}_s^{t,n+1}|H_s| + \mu^{\otimes(n+1)}(\mathcal{S}_s^{t,n+1})\alpha_s|X_s|\right) ds + \mathcal{E}_T^{t,n+1}|Z|\right] < \infty, \quad \text{(B.14)}$$

by the integrability assumptions (I). Recalling (B.9), we have

$$\int_t^T \mu^{\otimes(n+1)}(\mathcal{S}_s^{t,n+1})\alpha_s X_s ds = \int_t^T \left(\int_t^s \mu^{\otimes n}(\mathcal{S}_u^{t,n})\alpha_u\, du\right)\alpha_s X_s\, ds.$$

By (B.14), the use of Fubini's theorem is justified, and we get

$$\int_t^T \mu^{\otimes(n+1)}(\mathcal{S}_s^{t,n+1})\alpha_s X_s ds = \int_t^T \mu^{\otimes n}(\mathcal{S}_s^{t,n})\alpha_s \left(\int_s^T \alpha_u X_u\, du\right) ds. \quad \text{(B.15)}$$

Since $\mathcal{E}_s^{t,n} = \sum_{k=0}^{n} \mu^{\otimes k}(\mathcal{S}_s^{t,k})$, $n \in \mathbb{N}$, we obtain the identity

$$\int_t^T \mathcal{E}_s^{t,n+1} H_s\, ds = \int_t^T \mathcal{E}_s^{t,n} H_s\, ds + \int_t^T \mu^{\otimes(n+1)}(\mathcal{S}_s^{t,n+1}) H_s\, ds. \quad \text{(B.16)}$$

Moreover, we see that

$$\mu^{\otimes(n+1)}(\mathcal{S}_s^{t,n+1})|H_s| \leqslant \mathcal{E}_s^{t,n+1}|H_s|, \qquad \mu^{\otimes(n+1)}(S_T^{t,n+1})|Z| \leqslant \mathcal{E}_T^{t,n+1}|Z|. \quad \text{(B.17)}$$

Again, by (B.14), Fubini's theorem applies and shows that

$$\int_t^T \mu^{\otimes(n+1)}(\mathcal{S}_s^{t,n+1}) H_s\, ds = \int_t^T \mu^{\otimes n}(\mathcal{S}_s^{t,n})\alpha_s \left(\int_s^T H_u\, du\right) ds. \quad \text{(B.18)}$$

Additionally, we have

$$\mu^{\otimes(n+1)}(S_T^{t,n+1}) Z = \int_t^T \mu^{\otimes n}(\mathcal{S}_s^{t,n})\alpha_s Z\, ds. \quad \text{(B.19)}$$

We now establish integrability of $J^{t,n}$. From (B.12), we obtain

$$0 \leqslant |J_s^{t,n}| \leqslant E_s\left[\mathcal{E}_s^{t,n}|H_s| + \mu^{\otimes n}(\mathcal{S}_s^{t,n})\alpha_s\left(\int_s^T |\alpha_u X_u + H_u| du + |Z|\right)\right].$$

Hence, Tonelli's theorem yields

$$E\left[\int_t^T |J_s^{t,n}|\, ds\right] \leqslant E\left[\int_t^T \mu^{\otimes n}(\mathcal{S}_s^{t,n})\alpha_s\left(\int_s^T |\alpha_u X_u + H_u| du + |Z|\right) ds\right] + E\left[\int_t^T \mathcal{E}_s^{t,n}|H_s|\, ds\right]. \quad \text{(B.20)}$$

Now, (B.15), (B.18) and (B.19) imply

$$E\left[\int_t^T |J_s^{t,n}|\, ds\right] \leqslant E\left[\int_t^T \frac{1}{(n+1)!} I_s^{t,n+1}\left(\alpha_s|X_s| + |H_s|\right) ds + I_T^{t,n+1} Z\right], \quad \text{(B.21)}$$

where $I_s^{t,n+1}/(n+1)! = \mu^{\otimes(n+1)}(\mathcal{S}_s^{t,n+1})$. Thus the simple estimate (B.17) and the integrability statement (B.14) show that

$$E\left[\int_t^T |J_s^{t,n}|\, ds\right] < \infty. \quad \text{(B.22)}$$

Therefore, we are now in a position to interchange integrals in inequality (B.13): By Proposition B.13, the process $s \mapsto E_t[J_s^{t,n}]$ has a measurable modification K with $\int_t^T K_s ds = E_t[\int_t^T J_s^{t,n} ds]$ and we get

$$X_t \leqslant E_t\left[\int_t^T J_s^{t,n} ds + \mathcal{E}_T^{t,n} Z\right] = \int_t^T K_s ds + E_t\left[\mathcal{E}_T^{t,n} Z\right], \quad s \in [t, T]. \quad \text{(B.23)}$$

For all $s \in [t, T]$, the tower property of conditional expectations yields $K_s = E_t[J_s^{t,n}] = E_t[Y_s^{t,n}]$, where

$$Y_s^{t,n} \triangleq \mathcal{E}_s^{t,n} H_s + \mu^{\otimes n}(\mathcal{S}_s^{t,n})\alpha_s \left(\int_s^T (\alpha_u X_u + H_u) du + Z\right)$$

is a measurable process. Clearly, the bound (B.20) is also valid for $Y^{t,n}$. Thus, as in (B.21) and (B.22), we obtain $E[\int_t^T |Y_s^{t,n}| ds] < \infty$. The Fubini identities (B.15), (B.18), (B.19) and the identity (B.16) show that the integral of $Y^{t,n}$ is given by

$$\int_t^T Y_s^{t,n} ds = \int_t^T \mathcal{E}_s^{t,n+1} H_s + \mu^{\otimes(n+1)}(\mathcal{S}_s^{t,n+1})\alpha_s X_s ds + \mu^{\otimes(n+1)}(\mathcal{S}_T^{t,n+1}) Z.$$

Since $Y^{t,n}$ is integrable, Proposition B.13 yields a measurable modification $\bar{K}$ of $s \mapsto E_t[Y_s^{t,n}]$ such that $\int_t^T \bar{K}_s ds = E_t[\int_t^T Y_s^{t,n} ds]$. But then, $\bar{K}$ is also a measurable modification of $s \mapsto E_t[J_s^{t,n}]$, and the second part of Proposition B.13 implies

$$E_t[\int_t^T J_s^{t,n} ds] = \int_t^T \bar{K}_s ds = E_t[\int_t^T Y_s^{t,n} ds].$$

We insert this into (B.23) and obtain

$$\begin{aligned} X_t &\leqslant E_t\left[\int_t^T Y_s^{t,n} ds\right] + E_t\left[\mathcal{E}_T^{t,n} Z\right] \\ &= E_t\left[\int_t^T \mathcal{E}_s^{t,n+1} H_s + \mu^{\otimes(n+1)}(\mathcal{S}_s^{t,n+1})\alpha_s X_s ds + \mathcal{E}_T^{t,n+1} Z\right] \end{aligned}$$

since $\mu^{\otimes(n+1)}(\mathcal{S}_T^{t,n+1}) Z + \mathcal{E}_T^{t,n} Z = \mathcal{E}_T^{t,n+1} Z$. We have thus established (B.10) for $n+1$, and the proof if complete. □

C

APPENDIX: CONSUMPTION-PORTFOLIO OPTIMIZATION

C.1 PROOFS OMITTED FROM THE MAIN TEXT

PROOFS FOR SECTION 4.1.3

LEMMA 4.6. *If* $h \in C^{1,2}([0,T] \times \mathbb{R})$ *is a strictly positive solution of*

$$0 = h_t - \tilde{r}h + \tilde{\alpha}h_y + \tfrac{1}{2}\beta^2 h_{yy} + \tfrac{\delta\psi}{1-q}h^q, \quad h(T,\cdot) = \hat{\varepsilon}, \tag{4.6}$$

then $w(t,x,y) = \frac{1}{1-\gamma}x^{1-\gamma}h(t,y)^k$ *solves the HJB equation*

$$\begin{aligned} 0 = \sup_{\pi\in\mathbb{R},\, c\in(0,\infty)} \Big\{ & w_t + x(r+\pi\lambda)w_x - cw_x + \tfrac{1}{2}x^2\pi^2\sigma^2 w_{xx} \\ & + \alpha w_y + \tfrac{1}{2}\beta^2 w_{yy} + x\pi\sigma\beta\rho w_{xy} + f(c,w)\Big\}. \end{aligned} \tag{4.3}$$

Proof of Lemma 4.6. We set

$$\begin{aligned} H(z,\pi,c) \triangleq w_t + x(r+\pi\lambda)w_x - cw_x + \tfrac{1}{2}x^2\pi^2\sigma^2 w_{xx} + \alpha w_y \\ + \tfrac{1}{2}\beta^2 w_{yy} + x\pi\sigma\beta\rho w_{xy} + f(c,w), \end{aligned}$$

where $z \triangleq (t,x,y,w_x,w_y,w_{x_y}w_{xx},w_{yy})$. Separating

$$H(z,\pi,c) =: u(z,\pi) + s(z,c) + q(z),$$

it is easy to see that the candidate solutions $\hat{\pi}$ and $\hat{c}$ defined in (4.5) are the unique solutions of the associated first-order conditions

$$\begin{aligned} 0 = s_c(z,c) = -w_x + f_c(c,w), \\ 0 = u_\pi(z,\pi) = x\lambda w_x + \pi x^2\sigma^2 w_{xx} + x\sigma\beta\rho w_{xy}. \end{aligned} \tag{C.1}$$

Hence, concavity of u and s implies that

$$H(z,\hat{\pi},\hat{c}) = \sup_{\pi\in\mathbb{R},\, c\in(0,\infty)} H(z,\pi,c).$$

A direct computation shows that $H(z,\hat{\pi},\hat{c}) = 0$ since h solves the reduced HJB equation (4.6). Thus w solves the HJB equation (4.3). □

LEMMA 4.7. *The functions* $\tilde{r}$ *and* $\tilde{\alpha}$ *are bounded and Lipschitz continuous.*

Proof of Lemma 4.7. By (A1) and (A2), $\tilde{\alpha}$ and $\tilde{r}$ are bounded. Moreover

$$\begin{aligned} |\tilde{\alpha}(y) - \tilde{\alpha}(\bar{y})| \leqslant |\tfrac{1-\gamma}{\gamma}|\rho\left(|\tfrac{\lambda(y)}{\sigma(y)}||\beta(y)-\beta(\bar{y})| + |\tfrac{\beta(\bar{y})}{\sigma(y)}||\lambda(y)-\lambda(\bar{y})|\right) \\ + |\beta(\bar{y})\lambda(\bar{y})||\tfrac{\sigma(\bar{y})-\sigma(y)}{\sigma(y)\sigma(\bar{y})}| + |\alpha(y)-\alpha(\bar{y})| \end{aligned}$$

so $\tilde{\alpha}$ is Lipschitz continuous. Finally,

$$k|\tilde{r}(y)-\tilde{r}(\bar{y})| \leqslant |1-\gamma||r(y)-r(\bar{y})| + |\tfrac{1-\gamma}{\gamma}|\|\lambda\|_\infty(\inf_{x\in\mathbb{R}}\sigma(x))^{-2}|\lambda(y)-\lambda(\bar{y})|$$

$$+|\tfrac{1-\gamma}{\gamma}|\|\lambda\|^2_\infty\|\sigma\|_\infty(\inf_{x\in\mathbb{R}}\sigma(x))^{-4}|\sigma(\bar{y})-\sigma(y)|. \qquad \square$$

PROOFS FOR SECTION 4.2

PROPOSITION 4.13. *Let $S : A \to A$ be an operator on a closed, non-empty subset A of $\mathcal{D}_\infty$ and assume that there are constants $c > 0$, $\rho \geqslant 0$ such that for all $X, Y \in A$ we have a Lipschitz condition of the form*

$$|(SX)_t-(SY)_t| \leqslant c\int_t^T E_t\left[e^{(s-t)\rho}|X_s-Y_s|\right]ds \quad \textit{for all } t\in[t_0,T]. \qquad (C.2)$$

Then S has a unique fixed point $X \in A$. Moreover, the iterative sequence $X_{(n)} \triangleq SX_{(n-1)}$ $(n = 1, 2, \dots)$ with an arbitrarily chosen $X_{(0)} \in A$ satisfies

$$\|X_{(n)}-X\|_\infty \leqslant e^{T\rho}(\|X_{(0)}\|+\|X\|_\infty)\left(\tfrac{ecT}{n}\right)^n \quad \text{for all } n > cT.$$

Proof of Proposition 4.13. For any fixed $\kappa > c+\rho$, define a metric d equivalent to $\|\cdot\|_\infty$ by $d(X,Y) \triangleq \operatorname{ess\,sup}_{dt\otimes P} e^{-\kappa(T-t)}|X_t-Y_t|$. Then (A, d) is a complete metric space. By definition of d, we have

$$|X_s-Y_s| \leqslant e^{\kappa(T-s)}d(X,Y) \qquad \text{for } dt\text{-a.e. } s\in[t_0,T],$$

and hence (C.2) implies

$$e^{-\kappa(T-t)}|(SX)_t-(SY)_t| \leqslant e^{-\kappa(T-t)}c\int_t^T e^{(s-t)\rho}e^{\kappa(T-s)}d(X,Y)ds.$$

Calculating the integral, we obtain

$$e^{-\kappa(T-t)}|(SX)_t-(SY)_t| \leqslant \tfrac{c}{\kappa-\rho}d(X,Y) \quad \text{for } dt\text{-a.e. } t\in[t_0,T],$$

and we conclude that $d(SX,SY) \leqslant \frac{c}{\kappa-\rho}d(X,Y)$, where $\frac{c}{\kappa-\rho} < 1$. Hence S is a contraction on (A, d). Banach's fixed point theorem yields a unique $X \in A$ with $SX = X$; for all $n \in \mathbb{N}$, we have $d(X_{(n)}, X) \leqslant (\frac{c}{\kappa-\rho})^n d(X_{(0)}, X)$. Hence it follows that

$$\begin{aligned}|(X_{(n)})_t-X_t| &\leqslant e^{\kappa T}d(X_{(n)},X) \leqslant (\tfrac{c}{\kappa-\rho})^n e^{\kappa T}d(X_{(0)},X)\\ &\leqslant e^{\kappa T}(\|X_{(0)}\|_\infty+\|X\|_\infty)(\tfrac{c}{\kappa-\rho})^n,\end{aligned}$$

and thus $\|X_{(n)}-X\|_\infty \leqslant e^{\kappa T}(\|X_{(0)}\|_\infty+\|X\|_\infty)(\frac{c}{\kappa-\rho})^n$ for every $n\in\mathbb{N}$ and every choice of $\kappa > c+\rho$. Setting $\kappa = \frac{n+T\rho}{T}$ for $n > cT$, we obtain the asserted error bound. $\square$

PROOFS FOR SECTION 4.3

LEMMA 4.19. *The candidate optimal wealth process has all moments, i.e.,*

$$E\big[\sup_{t\in[0,T]}\hat{X}_t^p\big] < \infty \quad \textit{for all } p \in \mathbb{R}.$$

Proof Lemma 4.19. According to (4.2), the candidate optimal wealth process $\hat{X} = X^{\hat{\pi},\hat{c}}$ has dynamics

$$d\hat{X}_t = \hat{X}_t\big[a_t dt + b_t dW_t\big], \quad \hat{X}_0 = x,$$

where $a_t \triangleq r_t + \frac{1}{\gamma}\frac{\lambda_t^2}{\sigma_t^2} + \frac{k}{\gamma}\frac{\lambda_t\beta_t\rho}{\sigma_t}\frac{h_y}{h} - \delta^\psi h^{q-1}$ and $b_t \triangleq \frac{1}{\gamma}\frac{\lambda_t}{\sigma_t} + \frac{k}{\gamma}\beta_t\rho\frac{h_y}{h}$ are bounded by (A1), (A2) and Theorem 4.8. By Itō's formula

$$\hat{X}_t^p = x^p \exp\left(p\int_0^t \left(a_s + \tfrac{1}{2}(p-1)b_s^2\right) ds\right) \mathcal{E}_t\left(p\int_0^\cdot b_s dW_s\right)$$

where $\mathcal{E}_t(\cdot)$ denotes the stochastic exponential. Choose $M > 0$ such that $|pa_t| + |p(p-1)b_t^2|, |pb_t| < M$ for all $t \in [0,T]$. By Novikov's condition $\mathcal{E}_t\left(p\int_0^\cdot b_s dW_s\right)$ is an L^2-martingale, so using Doob's L^2-inequality we get

$$E\left[\sup_{t\in[0,T]}\hat{X}_t^p\right] \leqslant 2x^p e^{MT}\, E\left[\mathcal{E}_T\left(p\int_0^\cdot b_s dW_s\right)^2\right]^{\frac{1}{2}} < \infty. \qquad \square$$

LEMMA 4.20. *Let* $V_t \triangleq w(t,\hat{X}_t,Y_t)$, $t \in [0,T]$. *Then*

$$dV_t = -f(\hat{c}_t, V_t)dt + d\tilde{M}_t,$$

where

$$d\tilde{M}_t = V_t\left[\frac{1-\gamma}{\gamma}\frac{\lambda_t}{\sigma_t} + \frac{\rho k}{\gamma}\beta_t\frac{h_y}{h}\right]dW_t + V_t k\sqrt{1-\rho^2}\beta_t\frac{h_y}{h}d\bar{W}_t.$$

In particular, $V = \hat{V}$ *and* $w_x(t,\hat{X}_t,Y_t) = f_c(\hat{c}_t,\hat{V}_t)$. *Moreover, we have*

$$E\left[\sup_{t\in[0,T]}|\hat{c}_t|^p + \sup_{t\in[0,T]}|\hat{V}_t|^p\right] < \infty \quad \text{for all } p \in \mathbb{R}.$$

Proof of Lemma 4.20. By Itō's formula we have

$$dV_t = \big[w_t + \hat{X}_t(r_t + \hat{\pi}_t\lambda_t)w_x - \hat{c}_t w_x + \tfrac{1}{2}\hat{X}_t^2\hat{\pi}_t^2\sigma_t^2 w_{xx} + \alpha_t w_y + \tfrac{1}{2}\beta_t^2 w_{yy}$$
$$+\hat{X}_t\hat{\pi}_t\sigma_t\beta_t\rho w_{xy}\big]dt + d\tilde{M}_t$$

where $\tilde{M}$ is a local martingale. Hence, $dV_t = -f(\hat{c}_t,V_t)dt + d\tilde{M}_t$ by Lemma 4.6. Moreover, exploiting the special form of w we get

$$d\tilde{M}_t = V_t\left[\frac{1-\gamma}{\gamma}\frac{\lambda_t}{\sigma_t} + \frac{\rho k}{\gamma}\beta_t\frac{h_y}{h}\right]dW_t + V_t k\sqrt{1-\rho^2}\beta_t\frac{h_y}{h}d\bar{W}_t.$$

Here V_t can be rewritten as $V_t = w(t,\hat{X}_t,Y_t) = \frac{1}{1-\gamma}\hat{X}_t^{1-\gamma}h(t,Y_t)^k$. By Theorem 4.8, the function h is bounded and bounded away from zero. Thus $E[\sup_{t\in[0,T]}|V_t|^p] < \infty$ for all $p \in \mathbb{R}$ by Lemma 4.19. As h_y, λ, β

and σ^{-1} are bounded and h is bounded away from zero, the local martingale part in the Itō decomposition of V is an L^2-martingale. By uniqueness of stochastic differential utility, V is the unique utility process $\hat{V} = V^{\hat{c}}$ associated with $\hat{c}$. The first-order condition (C.1) for optimal consumption implies $w_x(t, \hat{X}_t, Y_t) = f_c(t, w(t, \hat{X}_t, Y_t)) = f_c(\hat{c}_t, \hat{V}_t)$. Finally, Lemma 4.19 and the boundedness of $\delta^\psi h(t, Y_t)^{q-1}$ imply that $E[\sup_{t\in[0,T]} |\hat{c}_t|^p] < \infty$ for all $p \in \mathbb{R}$. In particular, $\hat{c} \in \mathcal{C}$. □

LEMMA 4.22. *For all $(\pi, c) \in \mathcal{A}(x)$, the deflated wealth process $Z^{\pi,c}$ is a local martingale with dynamics*

$$dZ_t^{\pi,c} = \hat{m}_t X_t^{\pi,c} \Big[\Big(\pi_t \sigma(Y_t) - \tfrac{\lambda(Y_t)}{\sigma(Y_t)} \Big) dW_t + k\sqrt{1-\rho^2} \beta(Y_t) \tfrac{h_y(t,Y_t)}{h(t,Y_t)} d\bar{W}_t \Big].$$

For the candidate optimal process $(\hat{\pi}, \hat{c})$ this implies

$$\begin{aligned} dZ_t^{\hat{\pi},\hat{c}} = \hat{m}_t \hat{X}_t \Big[\Big(\tfrac{1-\gamma}{\gamma} \tfrac{\lambda(Y_t)}{\sigma(Y_t)} + \tfrac{k}{\gamma} \beta(Y_t) \rho \tfrac{h_y(t,Y_t)}{h(t,Y_t)} \Big) dW_t \ldots \\ \ldots + k\sqrt{1-\rho^2} \beta(Y_t) \tfrac{h_y(t,Y_t)}{h(t,Y_t)} d\bar{W}_t \Big]. \end{aligned} \tag{C.3}$$

For the proof of Lemma 4.22, recall the candidate optimal strategies

$$\begin{aligned} \hat{\pi}_t &= \tfrac{\lambda(Y_t)}{\gamma \sigma(Y_t)^2} + \tfrac{k}{\gamma} \tfrac{\beta(Y_t)\rho}{\sigma(Y_t)} \tfrac{h_y(t,Y_t)}{h(t,Y_t)}, \quad && t \in [0,T), \\ \hat{c}_t &= \delta^\psi h(t, Y_t)^{q-1} X_t^{\hat{\pi},\hat{c}}. && t \in [0,T), \end{aligned} \tag{4.8}$$

Proof of Lemma 4.22. For simplicity of notation, we set

$$r_t \triangleq r(Y_t), \quad \lambda_t \triangleq \lambda(Y_t) \quad \text{and} \quad \sigma_t \triangleq \sigma(Y_t).$$

By the product rule, we have

$$dZ_t^{\pi,c} = \hat{m}_t c_t dt + \hat{m}_t dX_t^{\pi,c} + X_t^{\pi,c} d\hat{m}_t + d[\hat{m}, X^{\pi,c}]_t.$$

Inserting the dynamics of $X^{\pi,c}$ from (4.2), we get

$$dZ_t^{\pi,c} = \hat{m}_t X_t^{\pi,c} [(r_t + \pi_t \lambda_t) dt + \pi_t \sigma_t dW_t] + X_t^{\pi,c} d\hat{m}_t + d[\hat{m}, X^{\pi,c}]_t. \tag{C.4}$$

By Lemma 4.20, we have

$$\hat{V}_t = w(t, \hat{X}_t, Y_t) \quad \text{and} \quad \hat{m}_t = e^{\int_0^t f_v(\hat{c}_s, \hat{V}_s) ds} w_x(t, \hat{X}_t, Y_t). \tag{C.5}$$

From here on, we abbreviate $f_v = f_v(\hat{c}_t, \hat{V}_t)$, $w_x = w_x(t, \hat{X}_t, Y_t)$ etc. From (C.5), we see that

$$d\hat{m}_t = \hat{m}_t \big[f_v dt + \tfrac{dw_x}{w_x} \big], \tag{C.6}$$

Note that $f_v(c, v) = \delta \frac{\phi-\gamma}{1-\phi} c^{1-\phi} [(1-\gamma)v]^{\frac{\phi-1}{1-\gamma}} - \delta\theta$. Plugging in the candidate $\hat{c}$ from (4.8) and $\hat{V}_t = w(t, \hat{X}_t, Y_t)$, we obtain

$$f_v = \tfrac{\phi-\gamma}{1-\phi} \delta^\psi h^{q-1} - \delta\theta. \tag{C.7}$$

Moreover, Itō's formula yields

$$dw_x = w_x\Big[\tfrac{w_{xt}}{w_x}dt + \tfrac{w_{xx}}{w_x}d\hat{X}_t + \tfrac{w_{xy}}{w_x}dY_t + \tfrac{1}{2}\tfrac{w_{xxx}}{w_x}d[\hat{X}_t] + \tfrac{1}{2}\tfrac{w_{xyy}}{w_x}d[Y_t] + \tfrac{w_{xxy}}{w_x}d[\hat{X}_t, Y_t]\Big].$$

Substituting for w, we find

$$\frac{dw_x}{kw_x} = \frac{h_t}{h}dt - \frac{\gamma}{k}\frac{d\hat{X}_t}{\hat{X}_t} + \frac{h_y}{h}dY_t + \frac{1}{2}\frac{\gamma(1+\gamma)}{k}\frac{d[\hat{X}_t]}{\hat{X}_t^2} + \frac{1}{2}\left((k-1)\frac{h_y^2}{h^2} + \frac{h_{yy}}{h}\right)d[Y_t] - \frac{\gamma}{\hat{X}_t}\frac{h_y}{h}d[\hat{X}_t, Y_t].$$

Plugging in the candidate $\hat{\pi}$ from (4.8) and the dynamics of $\hat{X}$ and Y, we get

$$\frac{dw_x}{kw_x} = A_t^1 dt + A_t^2 dt - \frac{1}{k}\frac{\lambda_t}{\sigma_t}dW_t + \sqrt{1-\rho^2}\beta_t\frac{h_y}{h}d\bar{W}_t, \quad \text{where} \tag{C.8}$$

$$A_t^1 \triangleq \frac{h_t}{h} - \frac{\gamma}{k}r_t + \frac{1}{2}\frac{1}{k}\frac{1-\gamma}{\gamma}\frac{\lambda_t^2}{\sigma_t^2} + \frac{1}{\gamma}\frac{\lambda_t\beta_t\rho}{\sigma}\frac{h_y}{h} + \frac{\gamma}{k}\delta^\psi h^{q-1} + \frac{k}{2}\frac{1+\gamma}{\gamma}\beta_t^2\rho^2\frac{h_y^2}{h^2}$$

$$A_t^2 \triangleq \frac{h_y}{h}\left(\alpha_t - \frac{\rho\beta_t\lambda_t}{\sigma_t}\right) + \frac{h_y^2}{h^2}\left(\frac{k-1}{2}\beta_t^2 - k\beta_t^2\rho^2\right) + \frac{\beta_t^2}{2}\frac{h_{yy}}{h}.$$

For the sum of the $\frac{h_y^2}{h^2}$-terms, we have

$$\frac{k}{2}\frac{1+\gamma}{\gamma}\beta_t^2\rho^2\frac{h_y^2}{h^2} + \frac{h_y^2}{h^2}\left(\frac{k-1}{2}\beta_t^2 - k\beta_t^2\rho^2\right) = \beta_t^2\frac{h_y^2}{h^2}\left(\frac{k}{2}\rho^2\frac{1+\gamma}{\gamma} + \frac{k-1}{2} - \rho^2 k\right) = 0$$

by our choice of k. Combining (C.6), (C.7) and (C.8), and recalling

$$\tilde{r} = -\frac{1}{k}\left[r(1-\gamma) + \frac{1}{2}\frac{1-\gamma}{\gamma}\frac{\lambda^2}{\sigma^2} - \delta\theta\right] \quad \text{and} \quad \tilde{\alpha} = \alpha + \frac{1-\gamma}{\gamma}\frac{\lambda\beta\rho}{\sigma}, \tag{4.7}$$

we thus obtain

$$\frac{d\hat{m}_t}{k\hat{m}_t} = -\frac{1}{k}\frac{\lambda_t}{\sigma_t}dW_t + \sqrt{1-\rho^2}\beta_t\frac{h_y}{h}d\bar{W}_t \dots \tag{C.9}$$

$$\dots + \left[\frac{h_t}{h} + \frac{1}{k}\left(-\gamma r_t + \frac{1}{2}\frac{1-\gamma}{\gamma}\frac{\lambda_t^2}{\sigma_t^2} - \delta\theta\right) + \tilde{\alpha}_t\frac{h_y}{h} + \frac{\beta_t^2}{2}\frac{h_{yy}}{h} + \frac{\delta^\psi}{1-q}\frac{h^q}{h}\right]dt.$$

Moreover, (C.9) yields

$$d[\hat{m}_t, X_t^{\pi,c}] = -\lambda_t\pi_t\hat{m}_t X_t^{\pi,c}dt.$$

Coming back to (C.4), we thus get

$$dZ_t^{\pi,c} = \hat{m}_t X_t^{\pi,c}[(r_t + \pi_t\lambda_t)dt + \pi_t\sigma_t dW_t] + X_t^{\pi,c}d\hat{m}_t + d[\hat{m}, X^{\pi,c}]_t$$
$$= k\hat{m}_t X_t^{\pi,c}\frac{1}{h}\left[h_t - \tilde{r}_t h + \tilde{\alpha}_t h_y + \frac{1}{2}\beta_t^2 h_{yy} + \frac{\delta^\psi}{1-q}h^q\right]dt + d\overline{M}_t,$$

where $d\overline{M}_t \triangleq \hat{m}_t X_t^{\pi,c}[(\pi_t\sigma_t - \frac{\lambda_t}{\sigma_t})dW_t + k\sqrt{1-\rho^2}\beta_t\frac{h_y}{h}d\bar{W}_t]$ defines a local martingale $\overline{M}$. Since h solves (4.6), we get

$$dZ_t^{\pi,c} = d\overline{M}_t = \hat{m}_t X_t^{\pi,c}[(\pi_t\sigma_t - \frac{\lambda_t}{\sigma_t})dW_t + k\sqrt{1-\rho^2}\beta_t\frac{h_y}{h}d\bar{W}_t],$$

Plugging in the definition of $\hat{\pi}$ immediately yields (C.3). □

LEMMA 4.23. *For any* $p > 0$, *we have*

$$E\left[\int_0^T f_c(\hat{c}_s, \hat{V}_s)^p ds + \exp\left(p\int_0^T f_v(\hat{c}_s, \hat{V}_s)ds\right)\right], \; E\left[\sup_{t\in[0,T]}|\hat{m}_t|^p\right] < \infty.$$

Moreover, the process $Z^{\hat{\pi},\hat{c}}$ *is a martingale.*

Proof of Lemma 4.23. Recalling $\underline{h} \leqslant h \leqslant \overline{h}$ and (C.7), we have

$$\begin{aligned} f_v(\hat{c}_s, \hat{V}_s) &= \tfrac{\phi-\gamma}{1-\phi}\delta^{\psi} h(s, Y_s)^{q-1} - \delta\theta \\ &\leqslant |\tfrac{\phi-\gamma}{1-\phi}|\delta^{\psi}\left(\underline{h}^{q-1} + \overline{h}^{q-1}\right) + |\delta\theta| \triangleq m_1, \end{aligned}$$

and we get $0 \leqslant \exp(p\int_0^T f_v(\hat{c}_s, \hat{V}_s)ds) \leqslant e^{Tpm_1}$. On the other hand, from Lemma 4.20, it follows that $E[\sup_{t\in[0,T]} f_c(\hat{c}_t, \hat{V}_t)^p] < \infty$ for all $p \in \mathbb{R}$. This proves the first part of the claim and implies that $\hat{m}_t = \exp(\int_0^t f_v(\hat{c}_s, \hat{V}_s)ds) f_c(\hat{c}_s, \hat{V}_s)$ has all moments, as well. To show that $Z^{\hat{\pi},\hat{c}}$ is a martingale, note that $\frac{1-\gamma}{\gamma}\frac{\lambda_t}{\sigma_t} + \frac{k}{\gamma}\beta_t\rho\frac{h_y}{h}$ is uniformly bounded by some $c > 0$. By Lemma 4.23 and Lemma 4.19, we have

$$\int_0^T E\left[\hat{m}_t^2\hat{X}_t^2(\tfrac{1-\gamma}{\gamma}\tfrac{\lambda_t}{\sigma_t} + \tfrac{k}{\gamma}\beta_t\rho\tfrac{h_y}{h})^2\right]dt \leqslant c^2\int_0^T \sqrt{E[\hat{m}_t^4]\,E[\hat{X}_t^4]}dt < \infty.$$

Analogously, we obtain $\int_0^T E[\hat{m}_t^2\hat{X}_t^2(k\sqrt{1-\rho^2}\beta_t\frac{h_y}{h})^2]dt < \infty$. From this and Lemma 4.22, we conclude that $Z^{\hat{\pi},\hat{c}}$ is an L^2-martingale. ☐

C.2 SOME FACTS ON PARABOLIC PARTIAL DIFFERENTIAL EQUATIONS

This appendix collects the relevant results on linear and semilinear parabolic partial differential equations that are used in this article. Following Ladyženskaja et al. (1968), we first introduce the Hölder spaces $H^{r/2,r}([0,T]\times\mathbb{R}^d)$ for $r \in \mathbb{R}^+$. For a continuous function

$$u : [0,T]\times\mathbb{R}^d \to \mathbb{R}, \quad (t,x) \mapsto u(t,x)$$

Hölder coefficients

and $q \in (0,1)$, we define the Hölder coefficient $\langle u\rangle_x^q$ in space via

$$\langle u\rangle_x^q \triangleq \sup_{t\in[0,T],\; x,x'\in\mathbb{R}^d,\; |x-x'|\leqslant 1} \frac{|u(t,x)-u(t,x')|}{|x-x'|^q}$$

and the Hölder coefficient $\langle u\rangle_t^q$ in time via

$$\langle u\rangle_t^q \triangleq \sup_{t,t'\in[0,T],\; x\in\mathbb{R}^d,\; |t-t'|\leqslant 1} \frac{|u(t,x)-u(t,x')|}{|t-t'|^q}.$$

Hölder spaces

The space $H^{r/2,r}([0,T]\times\mathbb{R}^d)$ consists of all functions $u : [0,T]\times\mathbb{R}^d \to \mathbb{R}$ that are continuous along with all derivatives $D_t^{\alpha}D_x^{\beta}u$ of "order" $2|\alpha|+|\beta| \leqslant r$ and that satisfy $\|u\|_H^{r/2,r} < \infty$. Here, the norm $\|u\|_H^{r/2,r}$ of u is given by

$$\|u\|_H^{r/2,r} \triangleq \langle u\rangle_{\bullet}^{r/2,r} + \sum_{2|\alpha|+|\beta|\leqslant\lfloor r\rfloor} \|D_t^{\alpha}D_x^{\beta}u\|_{\infty},$$

where the space-time Hölder coefficient $\langle u\rangle_{\bullet}^{r/2,r}$ of u is given by

$$\langle u\rangle_{\bullet}^{r/2,r} \triangleq \sum_{2|\alpha|+|\beta|=\lfloor r\rfloor} \langle D_t^{\alpha} D_x^{\beta} u\rangle_x^{r-\lfloor r\rfloor} + \sum_{r-2<2|\alpha|+|\beta|<r} \langle D_t^{\alpha} D_x^{\beta} u\rangle_t^{\frac{r-2|\alpha|-|\beta|}{2}}.$$

Thus $\|u\|_H^{r/2,r}$ sums up the L^∞-norms of all relevant derivatives, plus the Hölder coefficients of the highest-order derivatives. Analogously, for $r \in \mathbb{R}^+$, the space $H^r(\mathbb{R}^d)$ is defined as the collection of all $\lfloor r\rfloor$-times continuously differentiable functions $u : \mathbb{R}^d \to \mathbb{R}$ with $\|u\|_H^r < \infty$, where the norm is given by

$$\|u\|_H^r \triangleq \langle u\rangle_{\bullet}^r + \sum_{|\beta|\leqslant\lfloor r\rfloor} \|D^\beta u\|_\infty \quad \text{and} \quad \langle u\rangle_{\bullet}^r \triangleq \sum_{|\beta|=\lfloor r\rfloor} \langle D^\beta u\rangle^{r-\lfloor r\rfloor}.$$

Note: Here, we slightly abuse notation since $\langle u\rangle_x^q$ has only been defined for functions on $[0,T] \times \mathbb{R}^d$. Of course, for $u : \mathbb{R}^d \to \mathbb{R}$ and $q \in (0,1)$ we understand that $\langle u\rangle_x^q \triangleq \sup_{x,x'\in\mathbb{R}^d,\,|x-x'|\leqslant 1} \frac{|u(x)-u(x')|}{|x-x'|^q}$.

C.2.1 *Linear Cauchy problem*

We now consider a linear second-order differential operator

$$Lu \triangleq \frac{\partial u}{\partial t} - \sum_{i,j=1}^d a_{ij}(t,x)\frac{\partial^2 u}{\partial x_i \partial x_j} - \sum_{i=1}^d b_i(t,x)\frac{\partial u}{\partial x_i} - c(t,x)u,$$

where the coefficients a, b, c are defined on $[0,T] \times \mathbb{R}^d$ and $(a_{ij}(t,x))_{i,j}$ is a symmetric matrix for all $(t,x) \in [0,T] \times \mathbb{R}^d$. Theorem C.1 relies on the following conditions:

(P1) The operator L is uniformly parabolic, i.e., there exist $0 < c_1 < c_2 < \infty$ such that for every $(t,x) \in [0,T] \times \mathbb{R}^d$ we have

Assumptions on linear differential operator

$$c_1|y|^2 \leqslant \sum_{i,j=1}^d a_{ij}(t,x) y_i y_j \leqslant c_2|y|^2 \quad \text{for all } y \in \mathbb{R}^d.$$

$(P2)^r$ The coefficients satisfy $a_{ij}, b_i, c \in H^{r/2,r}([0,T] \times \mathbb{R}^d)$ for all $i,j = 1,\dots,d$.

We can now state the main existence and uniqueness result on the Cauchy problem in $\mathbb{R}^d$:

Existence and uniqueness for linear Cauchy problem

THEOREM C.1. *Suppose that* (P1) *and* $(P2)^r$ *are satisfied with* $r \in \mathbb{R}^+$, $r \notin \mathbb{N}$. *Further assume that* $\varphi \in H^{r+2}(\mathbb{R}^d)$ *and* $f \in H^{r/2,r}([0,T] \times \mathbb{R}^d)$. *Then, there exists a unique* $u \in H^{(r+2)/2,r+2}([0,T] \times \mathbb{R}^d)$ *such that*

$$Lu = f, \quad u(0,\cdot) = \varphi.$$

Moreover, u *satisfies*

$$\|u\|_H^{r/2+1,r+2} \leqslant c\left(\|\varphi\|_H^{r+2} + \|f\|_H^{r/2,r}\right),$$

where $c > 0$ *is a global constant that is independent of* φ *and* f.

Proof. See Theorem 5.1 in Ladyženskaja et al. (1968), p. 320. □

As a special case, we obtain the result we have used in the proof of Lemma 4.24.

COROLLARY C.2. *Suppose that*

(C1) $a, b, c : \mathbb{R} \to \mathbb{R}$ *are bounded and Lipschitz continuous,*

(C2) *the function* a *has a bounded Lipschitz continuous derivative and satisfies* $\inf_{y \in \mathbb{R}} a(y) > 0$,

(C3′) $\hat{\varepsilon} \in H^{r+2}(\mathbb{R})$ *for some* $r \in (0, 1)$.

Then, for each bounded and Lipschitz continuous $f : [0, T] \times \mathbb{R} \to \mathbb{R}$, *there exists a unique* $g \in C_b^{1,2}([0, T] \times \mathbb{R})$ *that solves*

$$0 = g_t + a g_{yy} + b g_y + c g + f, \quad g(T, \cdot) = \hat{\varepsilon}.$$

Proof. Consider the second-order differential operator

$$Lu = \tfrac{\partial u}{\partial t} - a \tfrac{\partial^2 u}{\partial y \partial y} - b \tfrac{\partial u}{\partial y} - cu.$$

By assumption (C1) and (C2), the differential operator L satisfies (P1) and $(P2)^r$ for $r \in (0, 1)$. Moreover, $f \in H^{r/2,r}([0, T] \times \mathbb{R})$ since f is Lipschitz continuous. Hence, Theorem C.1 yields $u \in H^{(r+2)/2,r+2}([0, T] \times \mathbb{R})$ such that

$$Lu = f(T - t, \cdot), \quad u(0, \cdot) = \hat{\varepsilon} \quad \text{and} \quad \|u\|_{C^{1,2}} \leqslant \|u\|_H^{(r+2)/2,r+2} < \infty.$$

Thus defining $g \in C_b^{1,2}([0, T] \times \mathbb{R})$ by $g(t, y) \triangleq u(T - t, y)$, we obtain

$$0 = g_t + a g_{yy} + b g_y + c g + f, \quad g(T, \cdot) = \hat{\varepsilon}. \qquad \square$$

C.2.2 *Quasilinear Cauchy problem*

We consider the nonlinear differential operator

$$Lu \triangleq u_t - \sum_{i=1}^d \left(\tfrac{d}{dx_i} a_i(t, x, u, u_x) \right) + a(t, x, u, u_x)$$

with principal part in divergence form. In order to formulate conditions on the coefficients of L, we set

$$a_{ij}(t, x, u, p) \triangleq \frac{\partial a_i(x, t, u, p)}{\partial p_j}, \qquad \text{and} \tag{C.10}$$
$$A(t, x, u, p) \triangleq a(t, x, u, p) - \sum_{i=1}^d \left(\tfrac{\partial a_i}{\partial u} p_i + \tfrac{\partial a_i}{\partial x_i} \right).$$

We formulate the following conditions:

Assumptions on nonlinear differential operator

(Q1) For all $t \in (0, T]$, $x, p \in \mathbb{R}^d$ and $u \in \mathbb{R}$, we have

$$\sum_{i,j=1}^d a_{ij}(t, x, u, p) y_i y_j \geqslant 0 \quad \text{for all } y \in \mathbb{R}^d.$$

(Q2) There exist $b_1, b_2 \geqslant 0$ such that for all $t \in (0,T]$, $x \in \mathbb{R}^d$ and $u \in \mathbb{R}$ we have

$$A(t,x,u,0) \geqslant -b_1 u^2 - b_2.$$

(Q3) The functions a and a_i are continuous and a_i is differentiable with respect to x, u, and p. Moreover, there exist $c_1, c_2 > 0$ such that for all $v = (t,x,u,p) \in [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$ we have

$$c_1|y|^2 \leqslant \textstyle\sum_{i,j=1}^n a_{ij}(v) y_i y_j \leqslant c_2 |y|^2 \quad \text{for all } y \in \mathbb{R}^d$$

and, with a_{ij} given by (C.10),

$$|a(v)| + \textstyle\sum_{i=1}^d \left(|a_i(v)| + |\frac{\partial a_i(v)}{\partial u}\right)(1+|p|) + \sum_{i,j=1}^d |a_{ij}(v)| \leqslant c_2(1+|u|)(1+|p|)^2.$$

$(Q4)^\beta$ There exists $\beta \in (0,1)$ such that for all compact sets $K \subset \mathbb{R}$, $\bar{K} \subset \mathbb{R}^d$ the functions

$$a_i, a, a_{i,j}, \frac{\partial a_i}{\partial u}, \frac{\partial a_i}{\partial x_i} : [0,T] \times \mathbb{R}^d \times K \times \bar{K} \to \mathbb{R}$$

are Hölder continuous in t, x, u and p with exponents $\frac{\beta}{2}$, β, β and β, respectively.

Here, we say that $f : [0,T] \times \mathbb{R}^d \times K \times \bar{K} \to \mathbb{R}, z = (z^1, z^2, z^3, z^4) \mapsto f(z)$ is β-Hölder continuous in z^i if

$$\langle u \rangle_i^\beta \triangleq \sup_{z,\bar{z} \in \mathrm{Dom}(f),\ z^j = \bar{z}^j,\ j \neq i,\ |z^i - \bar{z}^i| \leqslant 1} \frac{|f(z) - f(\bar{z})|}{|z^i - \bar{z}^i|^\beta} < \infty.$$

THEOREM C.3. *Suppose that* $\psi_0 \in H^{2+\beta}(\mathbb{R}^d)$ *and that* (Q1), (Q2), (Q3) *and* $(Q4)^\beta$ *are satisfied for some* $\beta \in (0,1)$. *Then, there exists a solution* $u \in H^{(2+\beta)/2, 2+\beta}([0,T] \times \mathbb{R}^d)$ *of the Cauchy problem*

Solution of quasilinear Cauchy problem

$$Lu = 0, \qquad u(0,\cdot) = \psi_0.$$

Proof. See Theorem 8.1 in Ladyženskaja et al. (1968), p. 495. □

For our purposes in Section 4.2, we require the following result:

COROLLARY C.4. *Suppose that*

(C1) $a, b, c : \mathbb{R} \to \mathbb{R}$ *are bounded and Lipschitz continuous,*

(C2) *the function* a *has a bounded Lipschitz continuous derivative and satisfies* $\inf_{y \in \mathbb{R}} a(y) > 0$,

(C3') $\hat{\varepsilon} \in H^{r+2}(\mathbb{R})$ *for some* $r \in (0,1)$,

and let $f \in C_b^1(\mathbb{R})$. *Then, the semilinear PDE*

$$0 = g_t + a g_{yy} + b g_y + c g + f(g), \qquad g(T,\cdot) = \hat{\varepsilon}.$$

has a solution $g \in C_b^{1,2}([0,T] \times \mathbb{R})$.

Proof. After setting $a_1(t,x,u,p) \triangleq pa(x)$ and

$$\bar{a}(t,x,u,p) \triangleq -b(x)p - c(x)u - f(u) + pa'(x),$$

we can represent the relevant differential operator as

$$\begin{aligned} Lu &\triangleq u_t - \tfrac{d}{dx} a_1(t,x,u,u_x) + \bar{a}(t,x,u,u_x) \\ &= u_t - \tfrac{d}{dx}(u_x a(x)) - b(x)u_x - c(x)u - f(u) + u_x a'(x) \\ &= u_t - a(x)u_{xx} - b(x)u_x - c(x)u - f(u). \end{aligned}$$

Hence if $u \in C_b^{1,2}([0,T]\times\mathbb{R})$ solves $Lu = 0$, $u(0,\cdot) = \hat{\varepsilon}$, then $g(t,x) \triangleq u(T-t,x)$ defines member of $C_b^{1,2}([0,T]\times\mathbb{R})$ that satisfies

$$0 = g_t + ag_{yy} + bg_y + cg + f(g), \quad g(T,\cdot) = \hat{\varepsilon}.$$

We now verify the assumptions of Theorem C.3 for L. Note that

$$a_{11}(t,x,u,p) = \frac{\partial a_1(x,t,u,p)}{\partial p_1} = a(x)$$

so (Q1) holds since

$$a_{11}(t,x,u,p)y^2 = a(x)y^2 \geqslant 0 \quad \text{by (C2).}$$

Next observe that

$$\begin{aligned} A(t,x,u,p) &= \bar{a}(t,x,u,p) - \frac{\partial a_1(t,x,u,p)}{\partial u}p - \frac{\partial a_1(t,x,u,p)}{\partial x} \\ &= -b(x)p - c(x)u - f(u). \end{aligned}$$

Thus (Q2) is satisfied since

$$A(t,x,u,0) = -c(x)u - f(u) \geqslant -\|c\|_\infty|u| - \|f\|_\infty \geqslant -b_1u^2 - b_2$$

with $b_1 \triangleq \|c\|_\infty$ and $b_2 \triangleq \|c\|_\infty + \|f\|_\infty$. To check (Q3), note that by (C1) and (C2) the functions a_1 and $\bar{a}$ are continuous and that a_1 is differentiable. Moreover

$$\inf_{x\in\mathbb{R}} a(x)|y|^2 \leqslant a_{11}(t,x,u,p)y^2 \leqslant \|\beta\|_\infty|y|^2$$

for all $t \in [0,T]$ and $x,u,p,y \in \mathbb{R}$. In addition, for $\nu = (t,x,u,p) \in [0,T]\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}$, we have

$$\begin{aligned} &|\bar{a}(\nu)| + \left(|a_1(\nu)| + \left|\frac{\partial a_1(\nu)}{\partial u}\right|\right)(1+|p|) + |a_{11}(\nu)| \\ &\leqslant \|b\|_\infty|p| + \|c\|_\infty|u| + \|f\|_\infty + \|a'\|_\infty|p| + |p|\|a\|_\infty(1+|p|) + \|a\|_\infty \\ &\leqslant (\|a\|_\infty + \|b\|_\infty + \|c\|_\infty + \|f\|_\infty + \|a'\|_\infty)(1+|u|)(1+|p|)^2 \end{aligned}$$

since a, b, c, f, and a' are bounded. Thus (Q3) holds with

$$c_2 \triangleq \|a\|_\infty + \|b\|_\infty + \|c\|_\infty + \|f\|_\infty + \|a'\|_\infty$$

and $c_1 \triangleq \inf_{x \in \mathbb{R}} a(x) > 0$. Finally, by assumptions (C1), (C2) for any compact set $K \subset \mathbb{R}$ the functions

$$a_1(v) = pa(x), \quad a(v) = -b(x)p - c(x)u - f(u) + pa'(x),$$
$$a_{11}(v) = a(x), \quad \frac{\partial a_1}{\partial u}(v) = 0, \quad \frac{\partial a_1}{\partial p}(v) = a'(x)$$

restricted to $[0, T] \times \mathbb{R} \times K \times K$ are Lipschitz continuous in x, u and p: Indeed, a, a' b and c are bounded and Lipschitz continuous by (C1) and (C2), and $f \in C_b^1(\mathbb{R})$. Hence $(Q4)^{\frac{1}{2}}$ holds as well, and, by Theorem C.3, the Cauchy problem

$$Lu = 0, \quad u(0, \cdot) = \hat{\varepsilon}$$

has a solution $u \in H^{5/4,5/2}([0, T] \times \mathbb{R}^d) \subset C_b^{1,2}([0, T] \times \mathbb{R}^d)$. □

BIBLIOGRAPHY

Acciaio, B., H. Föllmer, and I. Penner (2012): "Risk assessment for uncertain cash flows: model ambiguity, discounting ambiguity, and the role of bubbles," *Finance and Stochastics*, 16, 669–709.

Aliprantis, C. D. and K. C. Border (2006): *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer, Berlin, third ed.

Aliprantis, C. D. and O. Burkinshaw (1998): *Principles of Real Analysis*, Academic Press, Inc., San Diego, CA, third ed.

Antonelli, F. (1993): "Backward-forward stochastic differential equations," *The Annals of Applied Probability*, 3, 777–793.

Antonelli, F. (1996): "Stability of backward stochastic differential equations," *Stochastic Processes and their Applications*, 62, 103–114.

Antonelli, F. and A. Kohatsu-Higa (2000): "Filtration stability of backward SDE's," *Stochastic Analysis and Applications*, 18, 11–37.

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1999): "Coherent measures of risk," *Mathematical Finance*, 9, 203–228.

Artzner, P., F. Delbaen, J.-M. Eber, D. Heath, and H. Ku (2007): "Coherent multiperiod risk adjusted values and Bellman's principle," *Annals of Operations Research*, 152, 5–22.

Bank, P. and F. Riedel (2001a): "Existence and structure of stochastic equilibria with intertemporal substitution," *Finance and Stochastics*, 5, 487–509.

Bank, P. and F. Riedel (2001b): "Optimal consumption choice with intertemporal substitution," *The Annals of Applied Probability*, 11, 750–788.

Bansal, R. and A. Yaron (2004): "Risks for the long run: A potential resolution of asset pricing puzzles," *The Journal of Finance*, 59, 1481–1509.

Barberis, N. C. (2000): "Investing for the long run when returns are predictable," *The Journal of Finance*, 55, 225–264.

Barles, G., R. Buckdahn, and E. Pardoux (1997): "Backward stochastic differential equations and integral-partial differential equations," *Stochastics and Stochastics Reports*, 60, 57–83.

Barro, R. J. (2006): "Rare disasters and asset markets in the twentieth century," *The Quarterly Journal of Economics*, 121, 823–866.

Becherer, D. (2006): "Bounded solutions to backward SDEs with jumps for utility optimization and indifference hedging," *The Annals of Applied Probability*, 16, 2027–2054.

Belak, C., T. Seiferling, and F. T. Seifried (2015): "Backward nonlinear expectation equations," available at SSRN: http://ssrn.com/abstract=2547940.

Benzoni, L., P. Collin-Dufresne, and R. S. Goldstein (2011): "Explaining asset pricing puzzles associated with the 1987 market crash," *Journal of Financial Economics*, 101, 552–573.

Berdjane, B. and S. Pergamenshchikov (2013): "Optimal consumption and investment for markets with random coefficients," *Finance and Stochastics*, 17, 419–446.

Bertoldi, M. and L. Lorenzi (2005): "Estimates of the derivatives for parabolic operators with unbounded coefficients," *Transactions of the American Mathematical Society*, 357, 2627–2664.

Bhaskara Rao, K. P. S. and M. Bhaskara Rao (1983): *Theory of Charges: A Study of Finitely Additive Measures*, vol. 109 of *Pure and Applied Mathematics*, Academic Press, Inc., New York.

Bismut, J.-M. (1973): "Conjugate convex functions in optimal stochastic control," *Journal of Mathematical Analysis and Applications*, 44, 384–404.

Borovička, J., L. P. Hansen, M. Hendricks, and J. A. Scheinkman (2011): "Risk-price dynamics," *Journal of Financial Econometrics*, 9, 3–65.

Bouchard, B. and R. Elie (2008): "Discrete-time approximation of decoupled forward-backward SDE with jumps," *Stochastic Processes and their Applications*, 118, 53–75.

Bouchard, B. and N. Touzi (2004): "Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations," *Stochastic Processes and their Applications*, 111, 175–206.

Briand, P. and R. Carmona (2000): "BSDEs with polynomial growth generators," *Journal of Applied Mathematics and Stochastic Analysis*, 13, 207–238.

Briand, P., F. Coquet, Y. Hu, J. Mémin, and S. Peng (2000): "A converse comparison theorem for BSDEs and related properties of g-expectation," *Electronic Communications in Probability*, 5, 101–117.

Briand, P. and Y. Hu (2008): "Quadratic BSDEs with convex generators and unbounded terminal conditions," *Probability Theory and Related Fields*, 141, 543–567.

Campbell, J. Y. (2003): "Consumption-Based Asset Pricing," in *Financial Markets and Asset Pricing*, ed. by G. M. Constantinides, M. Harris, and R. Stulz, Elsevier, vol. 1 of *Handbook of the Economics of Finance*, chap. 13, 803–887.

Campbell, J. Y., G. Chacko, J. Rodriguez, and L. M. Viceira (2004): "Strategic asset allocation in a continuous-time VAR model," *Journal of Economic Dynamics & Control*, 28, 2195–2214.

Campbell, J. Y. and L. M. Viceira (1999): "Consumption and portfolio decisions when expected returns are time varying," *The Quarterly Journal of Economics*, 114, 433–495.

Campbell, J. Y. and L. M. Viceira (2002): *Strategic Asset Allocation: Portfolio Ahoice for Aong-Term Investors*, Clarendon Lectures in Economics, Oxford University Press.

Cerrai, S. (1996): "Elliptic and parabolic equations in $\mathbb{R}^n$ with coefficients having polynomial growth," *Communications in Partial Differential Equations*, 21, 281–317.

Cerrai, S. (2001): *Second Order PDE's in Finite and Infinite Dimension: A Probabilistic Approach*, vol. 1762 of *Lecture Notes in Mathematics*, Springer-Verlag, Berlin.

Chacko, G. and L. M. Viceira (2005): "Dynamic consumption and portfolio choice with stochastic volatility in incomplete markets," *The Review of Financial Studies*, 18, 1369–1402.

Chen, Z. and L. Epstein (2002): "Ambiguity, risk, and asset returns in continuous time," *Econometrica*, 70, 1403–1443.

Chen, Z. and S. Peng (2000): "A general downcrossing inequality for g-martingales," *Statistics & Probability Letters*, 46, 169–175.

Cheridito, P., F. Delbaen, and M. Kupper (2004): "Coherent and convex monetary risk measures for bounded càdlàg processes," *Stochastic Processes and their Applications*, 112, 1–22.

Cheridito, P., F. Delbaen, and M. Kupper (2006): "Coherent and convex monetary risk measures for unbounded càdlàg processes," *Finance and Stochastics*, 10, 427–448.

Cheridito, P., H. M. Soner, N. Touzi, and N. Victoir (2007): "Second-order backward stochastic differential equations and fully nonlinear parabolic PDEs," *Communications on Pure and Applied Mathematics*, 60, 1081–1110.

Cheridito, P. and M. Stadje (2013): "BSΔEs and BSDEs with non-Lipschitz drivers: Comparison, convergence and robustness," *Bernoulli*, 19, 1047–1085.

Choquet, G. (1954): "Theory of capacities," *Annales de l'institut Fourier*, 5, 131–295.

Choquet, G. (1959): "Forme abstraite du théorème de capacitabilité," *Annales de l'institut Fourier*, 9, 83–89.

Cohen, S., S. Ji, and S. Peng (2011): "Sublinear expectations and martingales in discrete time," available at arXiv: http://arxiv.org/abs/1104.5390.

Coolen, F. P., M. C. Troffaes, and T. Augustin (2011): "Imprecise Probability," in *International Encyclopedia of Statistical Science*, ed. by M. Lovric, Springer, Berlin, 645–648.

Coquet, F., Y. Hu, J. Mémin, and S. Peng (2002): "Filtration-consistent nonlinear expectations and related g-expectations," *Probability Theory and Related Fields*, 123, 1–27.

Cox, J. C. and C.-f. Huang (1989): "Optimal consumption and portfolio policies when asset prices follow a diffusion process," *Journal of Economic Theory*, 49, 33–83.

Delarue, F. (2002): "On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case," *Stochastic Processes and their Applications*, 99, 209–286.

Delbaen, F. (2002): "Coherent Risk Measures on General Probability Spaces," in *Advances in Finance and Stochastics*, ed. by K. Sandmann and P. J. Schönbucher, Springer, Berlin, 1–37.

Delbaen, F., Y. Hu, and A. Richou (2011): "On the uniqueness of solutions to quadratic BSDEs with convex generators and unbounded terminal conditions," *Annales de l'Institut Henri Poincaré Probabilités et Statistiques*, 47, 559–574.

Dellacherie, C. (1972): *Capacités et processus stochastiques*, Springer-Verlag, Berlin-New York.

Dellacherie, C. and P.-A. Meyer (1982): *Probabilities and Potential B: Theory of Martingales*, vol. 72 of *North-Holland Mathematics Studies*, North-Holland Publishing Co., Amsterdam.

Delong, Ł. (2013): *Backward Stochastic Differential Equations with Jumps and Their Actuarial and Financial Applications*, European Actuarial Academy (EAA) Series, Springer, London.

Denis, L., M. Hu, and S. Peng (2011): "Function spaces and capacity related to a sublinear expectation: Application to G-Brownian motion paths," *Potential Analysis*, 34, 139–161.

Denis, L. and C. Martini (2006): "A theoretical framework for the pricing of contingent claims in the presence of model uncertainty," *The Annals of Applied Probability*, 16, 827–852.

Di Tella, P. and H.-J. Engelbert (2015): "On the predictable representation property of martingales associated with Lévy processes," *Stochastics*, 87, 170–184.

Diestel, J. and J. J. Uhl, Jr. (1977): *Vector Measures*, vol. 15 of *Mathematical Surveys and Monographs*, American Mathematical Society, Providence, Rhode Island.

Duffie, D. and L. G. Epstein (1992a): "Asset pricing with stochastic differential utility," *The Review of Financial Studies*, 5, 411–436.

Duffie, D. and L. G. Epstein (1992b): "Stochastic differential utility," *Econometrica*, 60, 353–394, with an appendix by the authors and C. Skiadas.

Duffie, D., P.-Y. Geoffard, and C. Skiadas (1994): "Efficient and equilibrium allocations with stochastic differential utility," *Journal of Mathematical Economics*, 23, 133–146.

Duffie, D. and P.-L. Lions (1992): "PDE solutions of stochastic differential utility," *Journal of Mathematical Economics*, 21, 577–606.

Duffie, D., M. Schroder, and C. Skiadas (1997): "A term structure model with preferences for the timing of resolution of uncertainty," *Economic Theory*, 9, 3–22.

Duffie, D. and C. Skiadas (1994): "Continuous-time security pricing: A utility gradient approach," *Journal of Mathematical Economics*, 23, 107–131.

Dunford, N. (1935): "Integration in general analysis," *Transactions of the American Mathematical Society*, 37, 441–453.

Dunford, N. and J. T. Schwartz (1958): *Linear Operators I: General Theory*, vol. 7 of *Pure and Applied Mathematics*, Interscience Publishers, Inc., New York, with the assistance of W. G. Bade and R. G. Bartle.

El Karoui, N., S. Peng, and M. C. Quenez (1997): "Backward stochastic differential equations in finance," *Mathematical Finance*, 7, 1–71.

El Karoui, N., S. Peng, and M. C. Quenez (2001): "A dynamic maximum principle for the optimization of recursive utilities under constraints," *The Annals of Applied Probability*, 11, 664–693.

Elworthy, K. D. and X.-M. Li (1994): "Formulae for the derivatives of heat semigroups," *Journal of Functional Analysis*, 125, 252–286.

Epstein, L. G. (1987): "The global stability of efficient intertemporal allocations," *Econometrica*, 55, 329–355.

Epstein, L. G. and S. Ji (2013): "Ambiguous volatility and asset pricing in continuous time," *The Review of Financial Studies*, 26, 1740–1786.

Epstein, L. G. and S. Ji (2014): "Ambiguous volatility, possibility and utility in continuous time," *Journal of Mathematical Economics*, 50, 269–282.

Epstein, L. G. and M. Schneider (2003): "Recursive multiple-priors," *Journal of Economic Theory*, 113, 1–31.

Epstein, L. G. and T. Wang (1994): "Intertemporal asset pricing under Knightian uncertainty," *Econometrica*, 62, 283–322.

Epstein, L. G. and S. E. Zin (1989): "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework," *Econometrica*, 57, 937–969.

Ethier, S. N. and T. G. Kurtz (1986): *Markov Processes: Characterization and Convergence*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York.

Feldman, J. and M. Smorodinsky (1997): "Simple examples of non-generating Girsanov processes," in *Séminaire de Probabilités XXXI*, Springer, Berlin, vol. 1655 of *Lecture Notes in Mathematics*, 247–251.

Föllmer, H. and A. Schied (2004): *Stochastic Finance: An Introduction in Discrete Time*, vol. 27 of *de Gruyter Studies in Mathematics*, Walter de Gruyter & Co., Berlin, second revised and extended ed.

Gabaix, X. (2012): "Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance," *The Quarterly Journal of Economics*, 127, 645–700.

Gilboa, I. and D. Schmeidler (1989): "Maxmin expected utility with non-unique prior," *Journal of Mathematical Economics*, 18, 141–153.

Guvenen, F. (2009): "A parsimonious macroeconomic model for asset pricing," *Econometrica*, 77, 1711–1750.

Hampel, P. (2009): "Nonadditive probabilities in statistics," *Journal of Statistical Theory and Practice*, 3, 11–23.

Hansen, L., J. Heaton, and N. Li (2008): "Consumption strikes back? Measuring long run risk," *Journal of Political Economy*, 116, 260–302.

Harrison, J. M. and D. M. Kreps (1979): "Martingales and arbitrage in multiperiod securities markets," *Journal of Economic Theory*, 20, 381–408.

Hayashi, T. (2005): "Intertemporal substitution, risk aversion and ambiguity aversion," *Economic Theory*, 25, 933–956.

Heston, S. L. (1993): "A closed-form solution for options with stochastic volatility with applications to bond and currency options," *The Review of Financial Studies*, 6, 327–343.

Hu, M., S. Ji, S. Peng, and Y. Song (2014a): "Backward stochastic differential equations driven by G-Brownian motion," *Stochastic Processes and their Applications*, 124, 759–784.

Hu, M., S. Ji, S. Peng, and Y. Song (2014b): "Comparison theorem, Feynman-Kac formula and Girsanov transformation for BSDEs driven by G-Brownian motion," *Stochastic Processes and their Applications*, 124, 1170–1195.

Hu, M. and S. Peng (2009): "On representation theorem of G-expectations and paths of G-Brownian motion," *Acta Mathematicae Applicatae Sinica*, 25, 539–546.

Huber, P. J. (1981): *Robust Statistics*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York.

Huber, P. J. and V. Strassen (1973): "Minimax tests and the Neyman-Pearson lemma for capacities," *The Annals of Statistics*, 1, 251–263.

Jacod, J. and A. N. Shiryaev (2003): *Limit Theorems for Stochastic Processes*, vol. 288 of *Grundlehren der Mathematischen Wissenschaften*, Springer-Verlag, Berlin, second ed.

Jobert, A. and L. C. G. Rogers (2008): "Valuations and dynamic convex risk measures," *Mathematical Finance*, 18, 1–22.

Kahneman, D. (2011): *Thinking, Fast and Slow*, Farrar, Straus and Giroux, New York.

Kallenberg, O. (2002): *Foundations of Modern Probability*, Probability and Its Applications, Springer-Verlag, New York, second ed.

Kallsen, J. and J. Muhle-Karbe (2010): "Utility maximization in models with conditionally independent increments," *The Annals of Applied Probability*, 20, 2162–2177.

Kaltenbrunner, G. and L. A. Lochstoer (2010): "Long-run risk through consumption smoothing," *The Review of Financial Studies*, 23, 3190–3224.

Karatzas, I., J. P. Lehoczky, S. E. Shreve, and G.-L. Xu (1991): "Martingale and duality methods for utility maximization in an incomplete market," *SIAM Journal on Control and Optimization*, 29, 702–730.

Karatzas, I. and S. E. Shreve (1991): *Brownian Motion and Stochastic Calculus*, vol. 113 of *Graduate Texts in Mathematics*, Springer-Verlag, New York, second ed.

Kharroubi, I. and H. Pham (2015): "Feynman–Kac representation for Hamilton–Jacobi–Bellman IPDE," *The Annals of Probability*, 43, 1823–1865.

Kim, T. S. and E. Omberg (1996): "Dynamic nonmyopic portfolio behavior," *The Review of Financial Studies*, 9, 141–161.

Kobylanski, M. (2000): "Backward stochastic differential equations and partial differential equations with quadratic growth," *The Annals of Probability*, 28, 558–602.

Kolmogorov, A. (1933): *Grundbegriffe der Wahrscheinlichkeitsrechnung*, vol. 2 of *Ergebnisse der Mathematik und Ihrer Grenzgebiete*, Springer-Verlag, Berlin.

Kraft, H., T. Seiferling, and F. T. Seifried (2015): "Asset pricing and consumption-portfolio choice with recursive utility and unspanned risk," available at at SSRN: http://ssrn.com/abstract=2424706.

Kraft, H. and F. T. Seifried (2010): "Foundations of continuous-time recursive utility: Differentiability and normalization of certainty equivalents," *Mathematics and Financial Economics*, 3, 115–138.

Kraft, H. and F. T. Seifried (2014): "Stochastic differential utility as the continuous-time limit of recursive utility," *Journal of Economic Theory*, 151, 528–550.

Kraft, H., F. T. Seifried, and M. Steffensen (2013): "Consumption-portfolio optimization with recursive utility in incomplete markets," *Finance and Stochastics*, 17, 161–196.

Kramkov, D. and W. Schachermayer (1999): "The asymptotic elasticity of utility functions and optimal investment in incomplete markets," *The Annals of Applied Probability*, 9, 904–950.

Kreps, D. M. and E. L. Porteus (1978): "Temporal resolution of uncertainty and dynamic choice theory," *Econometrica*, 46, 185–200.

Kreps, D. M. and E. L. Porteus (1979): "Temporal von Neumann-Morgenstern and induced preferences," *Journal of Economic Theory*, 20, 81–109.

Ladyženskaja, O. A., V. A. Solonnikov, and N. N. Ural′ceva (1968): *Linear and Quasi-linear Equations of Parabolic Type*, vol. 23 of *Translations of Mathematical Monographs*, American Mathematical Society, Providence, R.I.

Lazrak, A. and M. C. Quenez (2003): "A generalized stochastic differential utility," *Mathematics of Operations Research*, 28, 154–180.

Levental, S., S. Sinha, and M. Schroder (2013): "Linked recursive preferences and optimality," *Mathematical Finance*, to appear.

Liu, J. (2007): "Portfolio selection in stochastic environments," *The Review of Financial Studies*, 20, 1–39.

Liu, J. and J. Pan (2003): "Dynamic derivative strategies," *Journal of Financial Economics*, 69, 401–430.

Luxemburg, W. A. J. (1991): "Integration with Respect to Finitely Additive Measures," in *Positive Operators, Riesz Spaces, and Economics*, Springer, Berlin, vol. 2 of *Studies in Economic Theory*, 109–150.

Ma, C. (2000): "An existence theorem of intertemporal recursive utility in the presence of Lévy jumps," *Journal of Mathematical Economics*, 34, 509–526.

Ma, J., P. Protter, and J. M. Yong (1994): "Solving forward-backward stochastic differential equations explicitly – a four step scheme," *Probability Theory and Related Fields*, 98, 339–359.

Ma, J., H. Yin, and J. Zhang (2012): "On non-Markovian forward-backward SDEs and backward stochastic PDEs," *Stochastic Processes and their Applications*, 122, 3980–4004.

Mandelkern, M. (1990): "On the uniform continuity of Tietze extensions," *Archiv der Mathematik*, 55, 387–388.

Marinacci, M. and L. Montrucchio (2010): "Unique solutions for stochastic recursive utilities," *Journal of Economic Theory*, 145, 1776–1804.

Mehra, R. and E. C. Prescott (1985): "The equity premium: A puzzle," *Journal of Monetary Economics*, 15, 145–161.

Miranda, E. (2008): "A survey of the theory of coherent lower previsions," *International Journal of Approximate Reasoning*, 48, 628–658.

Nualart, D. and W. Schoutens (2000): "Chaotic and predictable representations for Lévy processes," *Stochastic Processes and their Applications*, 90, 109–122.

Nualart, D. and W. Schoutens (2001): "Backward stochastic differential equations and Feynman-Kac formula for Lévy processes, with applications in finance," *Bernoulli*, 7, 761–776.

Nutz, M. (2012): "A quasi-sure approach to the control of non-Markovian stochastic differential equations," *Electronic Journal of Probability*, 17, 1–23.

Nutz, M. (2013): "Random G-expectations," *The Annals of Applied Probability*, 23, 1755–1777.

Nutz, M. and H. M. Soner (2012): "Superhedging and dynamic risk measures under volatility uncertainty," *SIAM Journal on Control and Optimization*, 50, 2065–2089.

Nutz, M. and R. van Handel (2013): "Constructing sublinear expectations on path space," *Stochastic Processes and their Applications*, 123, 3100–3121.

Obstfeld, M. (1994): "Evaluating risky consumption paths: The role of intertemporal substitutability," *European Economic Review*, 38, 1471–1486.

Pardoux, É. and S. Peng (1990): "Adapted solution of a backward stochastic differential equation," *Systems & Control Letters*, 14, 55–61.

Pardoux, É. and S. Peng (1992): "Backward stochastic differential equations and quasilinear parabolic partial differential equations," in *Stochastic Partial Differential Equations and Their Applications*, Springer, Berlin, vol. 176 of *Lecture Notes in Control and Information Sciences*, 200–217.

Pelsser, A. and M. Stadje (2014): "Time-consistent and market-consistent evaluations," *Mathematical Finance*, 24, 25–65.

Peng, S. (1997): "Backward SDE and related g-expectation," in *Backward Stochastic Differential Equations*, Longman, Harlow, vol. 364 of *Pitman Research Notes Mathematical Series*, 141–159.

Peng, S. (1999): "Monotonic limit theorem of BSDE and nonlinear decomposition theorem of Doob-Meyer's type," *Probability Theory and Related Fields*, 113, 473–499.

Peng, S. (2004a): "Filtration consistent nonlinear expectations and evaluations of contingent claims," *Acta Mathematicae Applicatae Sinica*, 20, 191–214.

Peng, S. (2004b): "Nonlinear Expectations, Nonlinear Evaluations and Risk Measures," in *Stochastic Methods in Finance*, Springer, Berlin, vol. 1856 of *Lecture Notes in Mathematics*, 165–253.

Peng, S. (2005a): "Dynamically consistent nonlinear evaluations and expectations," available at arXiv: http://arxiv.org/abs/math/0501415.

Peng, S. (2005b): "Nonlinear expectations and nonlinear Markov chains," *Chinese Annals of Mathematics*, 26, 159–184.

Peng, S. (2007): "G-Expectation, G-Brownian Motion and Related Stochastic Calculus of Itô Type," in *Stochastic Analysis and Applications*, Springer, Berlin, vol. 2 of *Abel Symposia*, 541–567.

Peng, S. (2008): "Multi-dimensional G-Brownian motion and related stochastic calculus under G-expectation," *Stochastic Processes and their Applications*, 118, 2223–2253.

Peng, S. (2010): *Nonlinear Expectations and Stochastic Calculus under Uncertainty*, available at arXiv: http://arxiv.org/abs/1002.4546.

Pham, H. (2009): *Continuous-time Stochastic Control and Optimization with Financial Applications*, vol. 61 of *Stochastic Modelling and Applied Probability*, Springer-Verlag, Berlin.

Quenez, M.-C. and A. Sulem (2013): "BSDEs with jumps, optimization and applications to dynamic risk measures," *Stochastic Processes and their Applications*, 123, 3328–3357.

Rietz, T. A. (1988): "The equity premium: A solution," *Journal of Monetary Economics*, 22, 117–131.

Royer, M. (2006): "Backward stochastic differential equations with jumps and related non-linear expectations," *Stochastic Processes and their Applications*, 116, 1358–1376.

Rudin, W. (1974): *Real and Complex Analysis*, McGraw-Hill Series in Higher Mathematics, McGraw-Hill Book Co., New York, second ed.

Schroder, M. and C. Skiadas (1999): "Optimal consumption and portfolio selection with stochastic differential utility," *Journal of Economic Theory*, 89, 68–126.

Schroder, M. and C. Skiadas (2003): "Optimal lifetime consumption-portfolio strategies under trading constraints and generalized recursive preferences," *Stochastic Processes and their Applications*, 108, 155–202.

Schroder, M. and C. Skiadas (2005): "Lifetime consumption-portfolio choice under trading constraints, recursive preferences, and nontradeable income," *Stochastic Processes and their Applications*, 115, 1–30.

Schroder, M. and C. Skiadas (2008): "Optimality and state pricing in constrained financial markets with recursive utility under continuous and discontinuous information," *Mathematical Finance*, 18, 199–238.

Seiferling, T. and F. T. Seifried (2015): "Stochastic differential utility with preference for information: Existence, uniqueness, concavity, and utility gradients," available at SSRN: http://ssrn.com/abstract=2625800.

Skiadas, C. (1998): "Recursive utility and preferences for information," *Economic Theory*, 12, 293–312.

Skiadas, C. (2008): "Dynamic Portfolio Choice and Risk Aversion," in *Financial Engineering*, ed. by J. R. Birge and V. Linetsky, Elsevier, vol. 15 of *Handbooks in Operations Research and Management Science*, chap. 19, 789–843.

Skiadas, C. (2013): "Scale-invariant asset pricing and consumption/portfolio choice with general attitudes toward risk and uncertainty," *Mathematics and Financial Economics*, 7, 431–456.

Soner, H. M., N. Touzi, and J. Zhang (2011a): "Martingale representation theorem for the G-expectation," *Stochastic Processes and their Applications*, 121, 265–287.

Soner, H. M., N. Touzi, and J. Zhang (2011b): "Quasi-sure stochastic analysis through aggregation," *Electronic Journal of Probability*, 16, 1844–1879.

Soner, H. M., N. Touzi, and J. Zhang (2012): "Wellposedness of second order backward SDEs," *Probability Theory and Related Fields*, 153, 149–190.

Soner, H. M., N. Touzi, and J. Zhang (2013): "Dual formulation of second order target problems," *The Annals of Applied Probability*, 23, 308–347.

Stadje, M. (2010): "Extending dynamic convex risk measures from discrete time to continuous time: A convergence approach," *Insurance: Mathematics & Economics*, 47, 391–404.

Svensson, L. E. O. (1989): "Portfolio choice with non-expected utility in continuous time," *Economics Letters*, 30, 313–317.

Tallarini, Jr, T. D. (2000): "Risk-sensitive real business cycles," *Journal of Monetary Economics*, 45, 507–532.

Tang, S. J. and X. J. Li (1994): "Necessary conditions for optimal control of stochastic systems with random jumps," *SIAM Journal on Control and Optimization*, 32, 1447–1475.

von Neumann, J. and O. Morgenstern (1944): *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, New Jersey.

Wachter, J. A. (2002): "Portfolio and consumption decisions under mean-reverting returns: An exact solution for complete markets," *Journal of Financial and Quantitative Analysis*, 37, 63–91.

Wachter, J. A. (2013): "Can time-varying risk of rare disasters explain aggregate stock market volatility?" *The Journal of Finance*, 68, 987–1035.

Walley, P. (1991): *Statistical Reasoning with Imprecise Probabilities*, vol. 42 of *Monographs on Statistics and Applied Probability*, Chapman and Hall, Ltd., London.

Weil, P. (1989): "The equity premium puzzle and the risk-free rate puzzle," *Journal of Monetary Economics*, 24, 401–421.

Weil, P. (1990): "Nonexpected utility in macroeconomics," *The Quarterly Journal of Economics*, 105, 29–42.

Xing, H. (2015): "Consumption investment optimization with Epstein-Zin utility in incomplete markets," to appear in Finance and Stochastics, available at arXiv: http://arxiv.org/abs/1501.04747.

Zhang, J. (2004): "A numerical scheme for BSDEs," *The Annals of Applied Probability*, 14, 459–488.

SCIENTIFIC CAREER

06/2007	*Allgemeine Hochschulreife*, Hanns-Seidel-Gymnasium Hösbach
10/2007 – 01/2012	University of Kaiserslautern
07/2009	*Vordiplom* in Mathematics (Minor: Computer Science)
07/2009 – 12/2009	Exchange student at the National University of Singapore
01/2012	*Diplom* (MSc) in Mathematics (specialization: *Analysis*), Program: *Mathematics International*
04/2012 – 03/2013	Scientific assistant at the Department of Mathematics of the University of Kaiserslautern
09/2012-04/2016	PhD student of Prof. Dr. Frank T. Seifried, University of Kaiserslautern
04/2013-03/2016	PhD grant from the *German Academic Scholarship Foundation*

WISSENSCHAFTLICHER WERDEGANG

06/2007	Allgemeine Hochschulreife, Hanns-Seidel-Gymnasium Hösbach
10/2007 – 01/2012	Technische Universität Kaiserslautern
07/2009	Vordiplom in Mathematik (Anwendungsfach: Informatik)
07/2009 – 12/2009	Auslandsstudium an der National University of Singapore
01/2012	Diplom in Mathematik, Schwerpunkt: *Angewandte Analysis*, Studienprogramm: *Mathematics International*
04/2012 – 03/2013	Wissenschaftlicher Mitarbeiter am Fachbereich Mathematik der Technischen Universität Kaiserslautern
09/2012-04/2016	Promotionsstudent bei Prof. Dr. Frank T. Seifried, Technische Universität Kaiserslautern
04/2013-03/2016	Promotionsförderung durch die *Studienstiftung des deutschen Volkes*